Différence d'échelle spatiale entre les mesures satellitaires et in situ d'humidité du sol : analyse par des approches spatio-temporelles

Beatriz Molero

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *18/12/2017* **par :**

Beatriz MOLERO RODENAS

Différence d'échelle spatiale entre les mesures satellitaires et *in situ* d'humidité du sol : analyse par des approches spatio-temporelles

---

JURY

Jean-Philippe GASTELLU-ETCHEGORRY Professeur d'Université UT3	Président du Jury
Sylvie THIRIA Professeur d'Université UPMC-LOCEAN	Rapporteur
Thierry PELLARIN Chargé de Recherche IGE	Rapporteur
Jean-Pierre WIGNERON Directeur de Recherche INRA	Examinateur
Jean-Luc VERGELY Ingénieur de Recherche ACRI-ST	Examinateur
María José ESCORIHUELA Ingénieur de Recherche IsardSAT	Examinateur
Yann H. KERR Directeur de Recherche CNES	Directeur de thèse
Delphine J. LEROUX Chercheur CNRM	Co-directeur de thèse
Philippe RICHAUME Ingénieur de Recherche UT3	Co-encadrant de thèse

---

École doctorale et spécialité :

SDU2E : Surfaces et interfaces continentales, Hydrologie

Unité de Recherche :

Centre d'Études Spatiales de la BIOsphère - CESBIO (UMR 5126)

Directeur(s) de Thèse :

Yann H. KERR, *Delphine J. LEROUX* et *Philippe RICHAUME*

"If you want to know the way nature works, we looked at it, carefully, (look at it, see) that's the way it looks! You don't like it? Go somewhere else – to another universe! Where the rules are simpler, philosophically more pleasing, more psychologically easy. I can't help it! OK! If I'm going to tell you honestly what the world looks like to the human beings who have struggled as hard as they can to understand it, I can only tell you what it looks like."

Richard Feynman
QED: Photons, Corpuscles of Light
Sir Douglas Robb Lectures, University of Auckland, June 1979

Abstract

Soil moisture (SM) is an essential climate variable and its monitoring is required for many applications, such as meteorological forecasting and hydrological modelling. Satellite borne microwave instruments provide surface SM estimates continuously in time and globally in space, with spatial resolutions (footprint) ranging from 30 to 100 km.

The accuracy of satellite SM products needs to be assessed through validation, which usually consists in comparing satellite-derived and *in situ* SM time series. *In situ* sensors typically measure over an area of a few centimetres (point resolution). Their representativeness area can be larger as SM exhibits spatial connection, but it never reaches the satellite footprint size. This spatial scale mismatch between satellite and *in situ* SM estimates impairs the validation process and the respective summary statistics to an extent that is not currently known.

This thesis aims at improving the knowledge of the spatial scale mismatch, as well as providing methods for its assessment applicable to any validation area. To this end, I first explored the spatial redistribution of SM with a disaggregation model, DISPATCH. This showed that the principal drivers of SM spatial variability do vary as functions of the landscape and the climatic conditions. Regarding the time domain, SM shown time scale dependence, which I studied using wavelet transforms.

From this basis, the connection between the SM spatial and time scales was investigated. Modelled and measured point series at Yanco and Little Washita *in situ* networks were decomposed into time scales ranging from 0.5 to 128 days, using wavelet transforms. The spatial representativeness of the point measurements was then assessed, on a per time scale basis, with 4 different approaches: temporal stability analysis, triple collocation, the percentage of correlated areas (CArea) and a new approach that uses wavelet-based correlations (WCor). I found that the average of the spatial representativeness values tends to increase with the time scales but so does their dispersion. This implies that some stations had large representativeness areas at seasonal scales, while others did not. At sub-weekly scales, all stations exhibited very small representativeness areas. Regarding the methods, WCor and CArea gave consistent results at all timescales. WCor is the most robust as it is the least sensitive to the number of stations.

This latter experiment revealed that isolating the spatial scale mismatch is not straightforward as its magnitude varies with the *in situ* location and the time scale. Nevertheless, one of the components of the mismatch, the sampling uncertainty, has been assessed separately with bootstrap and Monte Carlo simulations of point-support series. The series followed a temporal model describing the typical SM transitions in the region, and a spatial model describing the statistical distribution of SM.

This thesis has shown that SM time and spatial scales are connected and it has provided some generic tools to study this connection. This has important implications for validation and modelling, which could be adjusted depending on the time scale. Non-parametric statistical approaches, like Monte Carlo and bootstrap, could also be already operationally used in satellite validation campaigns.

Résumé

L'humidité du sol (soil moisture, SM, en anglais) est une variable climatique essentielle dont le suivi est requis dans de nombreuses applications, telles que la prévision météorologique et la modélisation hydrologique. Les instruments microondes à bord des satellites fournissent des estimations SM de surface en continu dans le temps et globalement dans l'espace, avec des résolutions spatiales (empreintes) allant de 30 à 100 km.

La précision des produits satellitaires de SM doit être validée, ce qui consiste généralement à comparer des séries temporelles de SM obtenues par satellite avec celles mesurées au sol. Les capteurs *in situ* sont généralement représentatifs d'une zone de quelques centimètres (résolution ponctuelle). Celle-ci peut éventuellement être plus grande compte tenu du fait que la SM est reliée spatialement, mais elle n'atteint jamais la taille de l'empreinte satellitaire. Cette différence entre l'échelle spatiale des estimations satellitaires et *in situ* altère le processus de validation et les statistiques obtenues à un niveau qui n'est pas connu actuellement.

Cette thèse vise à améliorer la connaissance de l'impact de la différence de l'échelle spatiale, ainsi qu'à fournir des méthodes d'évaluation de celle-ci applicables à toute zone de validation. Pour ce faire, j'ai d'abord exploré la redistribution spatiale de SM avec un modèle de désagrégation, DISPATCH. Cela a montré que les principaux facteurs de variabilité spatiale de SM varient en fonction de l'occupation des sols et du climat. En ce qui concerne le domaine temporel, la SM a montré une dépendance avec l'échelle de temps, que j'ai étudiée en utilisant des transformées en ondelettes.

Sur cette base, la relation entre les échelles spatiales et temporelles a été étudiée. Les séries modélisées à résolution ponctuelle et mesurées sur les régions de Yanco et Little Washita ont été décomposées en échelles de temps comprises entre 0,5 et 128 jours, en utilisant des transformées en ondelettes. La représentativité spatiale des mesures à résolution ponctuelle a ensuite été évaluée, par échelle de temps, avec 4 approches différentes : l'analyse de la stabilité temporelle, la triple colocation, le pourcentage de zones corrélées (CArea) et une nouvelle approche utilisant des corrélations basées sur des ondelettes (WCor). J'ai constaté que la moyenne des valeurs de représentativité spatiale obtenues tend à augmenter avec l'échelle de temps, mais aussi leur dispersion. Cela implique que certaines stations ont de vastes zones de représentativité à des échelles saisonnières, tandis que d'autres ne l'ont pas. Aux échelles sous-hebdomadaires, toutes les stations présentaient de très petites zones de représentativité. En ce qui concerne les méthodes de représentativité, WCor et CArea ont donné des résultats cohérents à toutes les échelles de temps. WCor est le plus robuste car elle est la moins sensible au nombre de stations.

Cette dernière expérience a révélé qu'isoler la différence d'échelle spatiale n'est pas direct car sa magnitude varie avec le point de mesure et l'échelle de temps. Néanmoins, l'une de ses composantes, l'incertitude d'échantillonnage, a été évaluée séparément avec des approches bootstrap et des simulations de Monte Carlo de séries à résolution ponctuelle. Les séries ont été contraintes par des modèles temporels décrivant la dynamique typique de la SM dans la région et un modèle spatial décrivant la distribution statistique de la SM. Cette thèse a montré que les échelles temporelles et spatiales de l'humidité du sol sont connectées et elle a fourni des outils génériques pour étudier cette connection. Ceci a des implications importantes pour la validation et la modélisation, qui pourraient être ajustées en fonction de l'échelle de temps. Les approches statistiques non paramétriques, comme Monte Carlo et bootstrap, pourraient également être utilisées de manière opérationnelle dans les campagnes de validation satellitaires.

Acknowledgements

Foremost, I am particularly indebted to my three supervisors, Yann Kerr, Delphine Leroux and Philippe Richaume, for their constant faith in my work. I am deeply grateful to Yann for giving me so many wonderful opportunities (including this thesis) and for his guidance. I am equally grateful to Delphine for being supportive and her help on organizing ideas. I am equally grateful to Philippe for proposing this thesis topic -quite an adventure-, for our long discussions and his understanding of the highs and lows of a thesis journey.

Further, I would like to acknowledge the jury members of this thesis and the members of my thesis committee. They have provided very useful hints for improving this thesis manuscript and continuing research.

I am also grateful to many colleagues I have worked with. Special mention goes to Olivier Merlin, Yoann Malbéteau and Ahmad Al Bitar, whose research enthusiasm was inspirational for me. I thank Sylvain Mangiarotti for his motivation and his enlightening discussions about soil moisture and chaos. I am also very grateful to Marc Thibeault and his team at the CONAE for their warm welcome and their interest and exchanges on my research. Special thanks go to Martin Guerrieri and Alvaro Soldano for the long but very fun hours of driving through the Pampas. Finally, I am grateful to the whole SMOS team at the CESBIO, the group has been a source of good advice and collaboration. I am specially thankful to Ali Mahmoodi and Simone Bircher who have been particularly attentive.

This thesis has been possible thanks to a number of *in situ* soil moisture datasets of high quality. I thank the people working in the USDA-ARS, the CONAE and the Murrumbidgee monitoring network for their good work on gathering these data. I also thank the CNES for funding my research through the TOSCA and CATDS programs and the CESBIO and its staff for providing a pleasant workplace.

Last but not least, I would like to thank my family and so many friends, from inside and outside the CESBIO, for their love and encouragement. And most of all, to my partner for its good sense of humour and its faithful support.

Contents

Abstract iii

Résumé v

Acknowledgements vii

List of Tables xi

List of Figures xvii

Acronyms xix

Symbols xxiii

1 Introduction 1
1.1 The surface soil moisture (SM) variable . . . 2
1.2 Measurement of SM . . . 5
1.3 Error budget and spatial scale mismatch . . . 8
1.4 Questions and objectives of this thesis . . . 11

1 Introduction (français) 15
1.1 La variable humidité de surface du sol . . . 16
1.2 Mesurer l'humidité du sol . . . 19
1.3 Bilan d'erreur et d'écart d'échelle spatiale . . . 22
1.4 Questions et objectifs de cette thèse . . . 26

2 Spatial variability of surface soil moisture (SM) 29
2.1 Sources of soil moisture variability . . . 30
2.2 Spatial variability metrics . . . 31
2.3 Observed spatial variability . . . 33
2.4 DISPATCH: a spatial disaggregation approach . . . 37
2.5 Conclusions . . . 49

3 Multi-time scale decomposition of soil moisture series 51
3.1 Introduction . . . 52
3.2 Time scales: wavelet-based approaches . . . 53
3.3 Datasets . . . 59
3.4 Decomposition of soil moisture series . . . 66
3.5 Conclusions . . . 76

4 Spatio-temporal dynamics of soil moisture within the footprint 79
4.1 Connections between spatial and temporal scales . . . 80
4.2 Methods for spatial representativeness assessment . . . 84
4.3 Assessment of the connection between spatial and time scales . . . 88
4.4 Assessment of the spatial representativeness: comparison of methods . . . 92
4.5 Assessment of the spatial representativeness of decomposed *in situ* series . . . 97
4.6 Discussion: links with chapter 3 . . . 103

4.7 Conclusions . . . 104

5 Uncertainty in temporal validation statistics due to spatial sampling 107
5.1 Assessment of the sampling uncertainty of individual locations (SUL) . . . 108
5.2 Assessment of the sampling uncertainty of the mean (SUM) . . . 120
5.3 Conclusions . . . 129

6 Conclusions and perspectives 131
6.1 Summary . . . 132
6.2 Perspectives . . . 137
6.3 Open questions . . . 138
6.4 Final note . . . 139

6 Conclusions et perspectives (français) 141
6.1 Résumé . . . 142
6.2 Perspectives . . . 148
6.3 Questions ouvertes . . . 150
6.4 Note finale . . . 150

A The DISPATCH processor 153
A.1 DISPATCH algorithm . . . 153
A.2 The C4DIS processor . . . 155

B Confidence intervals (CI) for the wavelet variance 157

C Confidence intervals (CI) for the wavelet correlation 159

D *In situ* networks: land cover maps 163

E Temporal empirical models of 5 *in situ* series 167

F Bootstrap distributions and sample size 171

Bibliography 175

List of Tables

1.1 Main characteristics of *in situ* and satellite SM measurements. 7
1.2 Principal spatial supports 11
1.3 Principal spatial extents 11

1.1 Principales caractéristiques des mesures *in situ* et par satellite de la SM. 22
1.2 Supports spatiaux principaux. 26
1.3 Étendues spatiales principales. 26

2.1 Summary of previous studies of soil moisture spatial variability classified according to the relationship between variance and mean (left column). "+" denotes increasing variance with increasing mean and "-" denotes decreasing variance with increasing mean. All are based on point-support (*in situ*) data, except for $*_a$ and $*_b$ 34
2.2 Dependence of spatial variance and correlation length with spacing, support and extent according to (Western and Blöschl, 1999). "+" describes an increasing trend (if one increases, the other increases), "−" describes a decreasing trend (if one increases, the other decreases) 36
2.3 Main characteristics of validation areas for DISPATCH 41
2.4 Number of samples used to evaluate DISPATCH SM products, classified per *in situ* network 42

3.1 Own calibration values for the parameters of the model of Pan et al. (2003) . . . 60
3.2 Characteristics of the modelled series of Little Washita. Only the differences with respect to the respective reference series (*ref-a* or *ref-b*) are indicated 61
3.3 Main characteristics of the *in situ* neworks used in this chapter. 63
3.4 Wavelet scales for time series sampled at $\boldsymbol{\Delta t = 3}$ **h** (*in situ* and modelled series) 65
3.5 Wavelet scales for time series sampled at $\boldsymbol{\Delta t = 12}$ **h** (*in situ* and SMOS series) . 65
3.6 Pre-processing and wavelet parameters of the modelled point SM series 66
3.7 Pre-processing and wavelet parameters of the *in situ* and SMOS soil moisture series. 66

4.1 Pre-processing of the data used to assess the connections between spatial and time scales in the Yanco region. 88
4.2 Pre-processing of the data used to assess the spatial representativeness in the Yanco region 94
4.3 Pre-processing of the data used to assess the spatial representativeness in Little Washita. 98

5.1 Characteristics of the distributions of the bootstrapped statistics obtained with different approaches. The approaches (A, B, C) assess the sampling uncertainty in groups of 5 stations. 128

List of Figures

1.1 Example of *in situ* SM and precipitation time series. The series were measured at station #9 of the Little Washita network (Oklahoma, U.S.). 3
1.2 Schematic of the water and energy fluxes at the land-atmosphere boundary . . . 3
1.3 Schematic of the scaling triplet: extent, support and spacing 5
1.4 Schematic of the transformation of brightness temperature (TB) into SM in radiometer-based systems. Note that the arrows indicate the direction of the modelling. 6
1.5 The 3 components of the spatial scale mismatch 10
1.6 Relationship between dielectric constant and SM according to the Mironov's model (Mironov and Fomin, 2009; Mironov et al., 2009) used in the SMOS SM retrieval processor (Kerr et al., 2014) for different clay contents. 11
1.7 Conceptual representation of the geophysical mismatch 11
1.8 Comparison of one SMOS soil moisture time series with 20 *in situ* time series corresponding to the same SMOS footprint 12

1.1 Exemple des séries temporelles de mesures *in situ* d'humidité du sol (SM) et de précipitation. Les séries ont été mesurées à la station n°9 du réseau Little Washita (Oklahoma, États-Unis). 17
1.2 Schéma des flux d'eau et d'énergie à l'interface terre-atmosphère. 17
1.3 Schéma du triplet d'échelle : étendue, support et espacement 19
1.4 Schéma de la transformation de la température de brillance (TB) en humidité du sol dans les systèmes utilisant des radiomètres satellitaires. Les flèches indiquent la direction de la modélisation. 21
1.5 Les 3 composants de la différence d'échelle spatiale. 25
1.6 Relation entre la constante diélectrique du sol et l'humidité du sol selon le modèle de Mironov (Mironov et Fomin, 2009 ; Mironov et al., 2009) utilisé dans le processeur SMOS d'inversion de l'humidité du sol (Kerr et al., 2014) pour différentes fractions d'argile. 26
1.7 Représentation conceptuelle de l'écart géophysique. 26
1.8 Comparaison d'une série temporelle d'humidité du sol du satellite SMOS avec 20 séries temporelles d'humidité du sol *in situ* mesurées au sein de la même empreinte satellitaire. 27

2.1 Standard deviation as a function of mean SM content from the for the 154 footprint-support samples of SGP97. Source: Ryu and Famiglietti (2005) 35
2.2 Evolution of the classical and the bounded CV (top row) and the standard deviation (bottom row) for a typical SM spatial distributions 36
2.3 Th spatial standard deviation *vs* the spatial mean of the *in situ* samples of Yanco, MB, LW and WG, at the SMOS overpass times. 44
2.4 Maps of L3 soil moisture and L4/DISPATCH soil moisture for the Yanco area on 2010/11/22. The black contours define the Coleambally irrigated farms. Source: Molero et al. (2016) 45
2.5 Maps of L4 disaggregated soil moisture for MB showing the progression of floods on the first days of January 2011. The day and the orbit (A/ascending, D/descending) is indicated in the title of each image. The black squares correspond from left to right to Yanco, Kyemba and Adelong areas. Source: Molero et al. (2016) 45

2.6 Maps of L3 soil moisture and L4 disaggregated soil moisture for the LW watershed (left) and the WG watershed (right) on 2011/05/02 and 2012/01/03, respectively. Solid contours indicate the watershed boundaries. In the LW maps, the river and lakes are represented by the bold dotted lines. Source: Molero et al. (2016) . . . 46
2.7 Distributions of the spatial statistics from the comparison of *in situ* to L3 and L4/DISPATCH soil moisture products. L3 statistics are in blue and L4/DISPATCH statistics are in red . . . 47
2.8 Distributions of the spatial statistics from the comparison of *in situ* to L3 and L4/DISPATCH soil moisture products. L3 statistics are in blue and L4/DISPATCH statistics are in red . . . 48
2.9 Distribution of spatial correlation values from the comparison of *in situ* to L3 and L4/DISPATCH soil moisture products. The statistics are calculated for (a) Yanco during summer (2010/12/01-2011/03/01) and for (b) WG during winters (December, January and February months of years 2010 to 2014) . . . 48
2.10 Scatter plots of *in situ* samples compared to L3 (upper row) and L4 disaggregated samples (bottom row), classified per *in situ* network. Source: Molero et al. (2016) 49

3.1 Continuous Haar wavelet at scale $\tau = 1$ and $t = 0$ . . . 54
3.2 Different types of MODWT mother wavelets $h(t)$ . . . 55
3.3 Daubechies-4 wavelet and scale filters at scales 1 to 5 . . . 56
3.4 Flow diagram illustrating analysis of X into W_1, W_2, W_3 and V_3 using the pyramid algorithm . . . 56
3.5 Flow diagram illustrating the wavelet decomposition of a signal X for a maximum decomposition level of 3. . . . 57
3.6 Modelled and measured SM series at station #1 (top), with the adjusted γ_c parameter (middle) and the LAI series (bottom) at the same location. . . . 61
3.7 Little Washita modelled series for the 2012/07/03 - 2014/07/02 period . . . 62
3.8 Pictures taken at Monte Buey during the Cal/Val campaing in November 2016 that illustrate the water table level (left) and the trenches built for water drainage (right) . . . 63
3.9 Haar-based multi-resolution analysis (MRA) of the reference *ref-a* series of the Little Washita modelled dataset for the 2012/07/03 - 2014/07/02 period . . . 67
3.10 Wavelet variance of the modelled series. . . . 68
3.11 Wavelet-based correlation between the modelled sample and reference series. . . . 68
3.12 Variance decomposition of the point *in situ* series (color lines) and the *in situ* spatial average series (bold black line) . . . 71
3.13 Correlation between detail series of the *in situ* spatial average and the point *in situ* series . . . 72
3.14 Pictures taken during the Monte Buey Cal/Val campaign on november 2016. The white notebook indicates the position of the station . . . 73
3.15 Variance decomposition of the average *in situ* series and the SMOS series . . . 75
3.16 Variance decomposition of the average *in situ* series and the SMOS series after setting the same gaps in *in situ* series as in SMOS series . . . 75
3.17 Correlation between the detail series of the average *in situ* series and the SMOS series . . . 76
3.18 Correlation between the detail series of the average *in situ* series and the SMOS series after setting the same gaps in *in situ* series as in SMOS series . . . 76

4.1 Hydrological processes at a range of characteristic spatial and temporal scales. Source: (Blöschl and Sivapalan, 1995) . . . 82
4.2 Maps of wavelet correlation per time scale (R_j) between the detail series of each pixel and the detail series of the average of all the pixels (FAvg, field average) of the gridded SM dataset for the 2014/09-2015/03 6-month period. Source: (Molero et al., 2018) . . . 89

4.3 CArea scores as a function of the correlation threshold for some specific pixels of the gridded SM dataset for the 2014/09-2015/03 6-month period. Each plot corresponds ot a different time scale. Source: (Molero et al., 2018) 90
4.4 Wavelet variance for 3 different disaggregated series: the average (black), and those of pixels #6 (blue) and #13 (cyan). Dotted lines depict 95 % confidence intervals (CIs) (Appendix B). 91
4.5 Spatial representativeness scores (vertical axis) from different methods (horizontal axis) for different pixels of the gridded SM dataset. The representativeness is evaluated on a) full series and on b) detrended series (components $>$ 32 days are removed). The CArea correlation threshold is 0.55. Source: (Molero et al., 2018) 95
4.6 Spatial representativeness scores from different methods for different pixels of the gridded SM dataset. The representativeness is evaluated on wavelet-decomposed series (detail series). Source: (Molero et al., 2018) 96
4.7 Correlation between the detail series of the TC datasets. The left column represents the TC-ECMWF triplet and the right column, the TC-AMSR2 triplet. In the top row, the gridded SM dataset is the network average, in the bottom row, the gridded SM dataset is the pixel #4 time series 97
4.8 Spatial representativeness scores for different *in situ* locations of the LW network. The representativeness is evaluated on wavelet-decomposed series (detail series). Source: (Molero et al., 2018) 100
4.9 Correlation between the detail series of the TC datasets. The left column represents the TC-ECMWF triplet and the right column, the TC-AMSR2 triplet. In the top row, the *in situ* dataset is the network average, in the bottom row, the *in situ* dataset is the station #4 time series 101
4.10 Wavelet variance of the TC datasets. The left column represents the TC-ECMWF triplet and the right column, the TC-AMSR2 triplet. In the top row, the *in situ* dataset is the network average, in the bottom row, the *in situ* dataset is the station #8 time series 103

5.1 Necessary elements to obtain the distribution of validation statistics within a satellite footprint of size A. 109
5.2 Procedure to estimate the effect of the sampling uncertainty in validation statistics. 110
5.3 Final model fit for station #9 112
5.4 Spatial model: spatial PDFs produced with a standard deviation of 0.05 m^3/m^3 and modes 0.1, 0.25 and 0.4, from left to right 113
5.5 Distribution of samples before and after CDF-matching 115
5.6 Simulation of M=10,000 point series for the Little Washita region. In the top row each coloured line is one different simulated series and in the 2^{nd}-top row, the density of series is represented in gray. The 3^{rd} row shows the spatial standard deviations of the 5 measured series and of the 10,000 simulated series. The bottom row shows the 5 measured *in situ* series. 115
5.7 The network average (NAvg), SMOS and AMSR2 series on top of the simulated point series for the Little Washita region 116
5.8 Comparison of the network *in situ* average (NAvg), SMOS and AMSR2 time series with the simulated point series 117
5.9 Validation of SMOS and AMSR2 datasets with 5 *in situ* series with simoultaneous assessment of the statistical sampling uncertainty (box plots). In the box plots, the middle red line is the median, the limits of the box are the 25 % and the 75 % quartiles and the whiskers are the 5 % and the 95 % percentiles of the uncertainty distribution. Statistics falling outside the 90 % percentile are represented as blue circles. 118

5.10 Bootstrap distribution of the correlation between SMOS and the *in situ* NAvg, as a function of the number of *in situ* time series taken to compute the NAvg. The dashed line correspond to the median and the dotted lines to the 5 % and 95 % confidence limits 125
5.11 Width of the 90 % CI of the bootstrap distributions of the validation statistics, as a function of the number of *in situ* time series taken to compute the NAvg in Little Washita 125
5.12 Temporal statistics and their respective uncertainty distributions obtained when validating SMOS soil moisture series with only 5 ground series (K'=5). The uncertainty distributions are obtained with 3 different approaches: A) Bootstrap 1,000 subsets of K'=5 series out of 20 *in situ* series; ; B) Bootstrap 1,000 subsets of K'=5 series out of 10,000 simulated series; C) Bootstrap one group of K'=K=5 *in situ* series. 127
5.13 Validation of SMOS and AMSR2 datasets with the NAvg of 5 *in situ* series with simultaneous assessment of the sampling uncertainty of the NAvg (box plots). The triangle represents the comparison with the measured NAvg 128
5.14 Validation of SMOS and AMSR2 datasets with the NAvg of 20 *in situ* series with simultaneous assessment of the statistical sampling uncertainty (box plots). The triangle represents the comparison with the measured NAvg 129

A.1 Example of the triangle method of Moran et al. (1994), which is based on the position of the pixel in the LST-f_v space, applied to the ASTER data acquired on the 17/08/2011 in Urgell, Spain. Source: Merlin (2016) 154

B.1 Variance decomposition of the *in situ* time series (color lines) and the average *in situ* series (black line) of the Little Washita network, with the respective confidence intervals 158

C.1 Correlation between the detail series of the point *in situ* series and the average *in situ* series of Little Washita, with the respective 95 % CIs calculated with the overlapping block bootstrap method 161
C.2 Bootstrap distributions of the correlation between **station #8** series and the average *in situ* series of the Little Washita network 162
C.3 Bootstrap distributions of the correlation between **station #12** series and the average *in situ* series of the Little Washita network 162

D.1 Ecoclimap land cover maps for the four dense *in situ* networks used in this thesis. The circle represents a mean SMOS footprint size of 40 km. 164
D.2 Aggregated Ecoclimap land cover maps for the four dense *in situ* networks used in this thesis. The aggregated classes are those considered in the SMOS retrieval algorithms. The circle represents a mean SMOS footprint size of 40 km. 164
D.3 NDVI time series observed at each of the station locations (the station ID is indicated in the axis labels at the left). The values are extracted from the NDVI raster, input to the disaggregation algorithm in the C4DIS processor (Molero et al., 2016). 165

E.1 Pre-selection of peaks and decays for the fit of the peak model for station #9 . . 168
E.2 Fit of the peaks model to the observed data at station #9 168
E.3 Fit of the decay model to the observed data at station #9 169
E.4 Observed and modelled time series at each of the 5 selected stations locations: #4, #6, #9, #10 and #13 169

F.1 Bootstrap distribution of the correlation between SMOS and the *in situ* NAvg, as a function of the number of *in situ* time series taken to compute the NAvg. The dashed line correspond to the median and the dotted lines to the 5 % and 95 % confidence limits 171
F.2 Bootstrap distribution of the bias between SMOS and the *in situ* NAvg, as a function of the number of *in situ* time series taken to compute the NAvg. The dashed line correspond to the median and the dotted lines to the 5 % and 95 % confidence limits 172
F.3 Bootstrap distribution of the STDD between SMOS and the *in situ* NAvg, as a function of the number of *in situ* time series taken to compute the NAvg. The dashed line correspond to the median and the dotted lines to the 5 % and 95 % confidence limits 173

Acronyms

AMSR2 Advanced Microwave Scanning Radiometer-2.

bCV bounded coefficient of variation.

C4DIS CATDS Level-4 Disaggregation processor.

Cal/Val calibration and validation.

CArea correlated area.

CATDS Centre Aval de Traitement des Données SMOS.

CDF cumulative density function.

CDTI Centro Para el Desarrollo Tecnologico Industrial.

CESBIO Centre d'Études Spatiales de la Biosphère.

CI confidence interval.

CNES Centre National d'Etudes Spatiales.

CV coefficient of variation.

CWT continuous wavelet transform.

D Daubechies.

DCT discrete cosine transform.

DEM digital elevation model.

DISPATCH DISaggregation based on Physical And Theoretical scale Change.

DQX data quality index.

DWT discrete wavelet transform.

EASE Equal-Area Scalable Earth.

eCDF empirical cumulative density function.

ECMWF European Centre for Medium-range Weather Forecasting.

ESA European Space Agency.

ESTAR Electronically Scanned Thinned Array Radiometer.

FAvg field average.

HR higher resolution.

JAXA Japan Aerospace Exploration Agency.

L-MEB L-band Microwave Emission of the Biosphere.

L2 Level-2.

L3 Level-3.

L4 Level-4.

LA least-asymmetric.

LAI leaf area index.

LP DAAC NASA Land Processes Distributed Active Archive Center.

LPRM Land Parameter Retrieval Model.

LR low resolution.

LST land surface temperature.

LW Little Washita.

MB Murrumbidgee.

MBy Monte Buey.

MIR CLF31A/D SMOS Level-3 1-day global SM product.

MOD11A1 MODIS/Terra daily L3 global LST and emissivity 1-km grid.

MOD13A2 MODIS/Terra 16-day L3 global vegetation indices 1-km grid.

MODIS MODerate resolution Imaging Spectroradiometer.

MODWT maximal overlap discrete wavelet transform.

MRA multi-resolution analysis.

MRD mean relative difference.

MYD11A1 MODIS/Aqua daily L3 global LST and emissivity 1-km grid.

NASA National Aeronautics and Space Administration.

NAvg network average.

NDVI normalized difference vegetation index.

NWP numerical weather prediction.

PDF probability density function.

RD relative differences.

RFI radio-frequency interference.

RMSD root mean squared difference.

RMSE root mean square error.

RZSM root zone soil moisture.

SAR synthetic aperture radar.

SDRD standard deviation of the relative differences.

SEE soil evaporative efficiency.

SGP97 Southern Great Plains 1997.

SM soil moisture.

SMAP Soil Moisture Active Passive.

SMOS Soil Moisture Ocean Salinity.

SR spatial representativeness.

STDD standard deviation of the differences.

SU sampling uncertainty.

SUL sampling uncertainty of the locations.

SUM sampling uncertainty of the mean.

TB brightness temperature.

TC triple collocation.

TDR time-domain reflectometry.

TStab temporal stability analysis.

ubRMSD unbiased root mean squared difference.

USGS U.S. Geological Survey.

VUA Vrije Universiteit Amsterdam.

WCor wavelet-based correlation.

WG Walnut Gulch.

Symbols

α confidence level.

B bias.

df degrees of freedom.

D_j detail series at scale j.

ϵ Noise variable.

ε relative permitivity.

f_v fractional vegetation cover.

γ mean lapse rate ($^\circ C/m$).

γ_c canopy interception coefficient.

$h_j^\circ(t)$ wavelet filter periodized to N (transfer function $H_j(f/N)$).

$h_j(t)$ wavelet filter at scale j (transfer function $H_j(f)$).

j time scale (unit less).

$J0$ last time scale or decomposition level.

K number of samples in space (locations).

Ks soil drainage coefficient.

L_j width of the wavelet filter at scale j.

η_l loss coefficient.

M number of samples in Monte Carlo simulations.

μ population mean or mean of a random variable.

$\hat{\mu}^*$ mean of the bootstrapped samples.

$\hat{\mu}$ sample mean.

N number of samples or number of samples in time.

N_B number of bootstrap samples.

$N(\mu, \sigma)$ normal distribution with mean μ and standard deviation σ.

p precipitation.

R sample correlation.

ρ population correlation.

R_j wavelet-based correlation.

$\hat{r}_{k,\theta}$ triple collocation correlation.

R_{th} correlation threshold of the CArea method.

S slope of the regression line.

σ population standard deviation.

σ^2 population variance or variance of a random variable.

$\hat{\sigma}^2$ sample variance.

$\hat{\sigma}_j^2$ wavelet variance at scale j.

$\sigma_{X,Y}$ population covariance between X and Y.

$\hat{\sigma}_{X,Y}$ sample covariance between X and Y.

$\hat{\sigma}^*$ standard deviation of the bootstrapped samples.

$\hat{\sigma}$ sample standard deviation.

S_j smooth series at scale j.

τ (τ_j) time scale (in time units).

θ true signal (true soil moisture).

Δt sampling period.

T_s soil skin temperature.

T_v vegetation temperature.

V_j scaling coefficients at scale j.

$v_j(t)$ scaling filter at scale j (transfer function $V_j(f)$).

$v_j^{\circ}(t)$ scaling filter periodized to N (transfer function $V_j(f/N)$).

ψ wavelet function.

W_j wavelet coefficients at scale j.

X^* bootstrapped sample.

Z soil depth.

Chapter 1

Introduction

Contents

1.1 The surface soil moisture (SM) variable 2
1.1.1 Definition 2
1.1.2 SM related processes 2
1.1.3 Why is SM information relevant 3
1.2 Measurement of SM 5
1.2.1 Large-support observations: satellite sensors 5
1.2.2 Point-support measurements: *in situ* sensors 6
1.3 Error budget and spatial scale mismatch 8
1.3.1 What is soil moisture at 40 km resolution? 8
1.3.2 Observation systems errors and uncertainties 8
1.3.3 *In situ* and satellite: the spatial scale mismatch 9
1.3.4 Terminology 11
1.4 Questions and objectives of this thesis 11
1.4.1 Manuscript structure 13

Soil moisture (SM) represents about 0.009 % of the total Earth's water storage and about 0.8 % of the Earth's liquid fresh water resources (Trenberth et al., 2007), yet it is a fundamental physical variable that participates in the water and energy exchanges between the land and the atmosphere. Since 2010, soil moisture belongs to the 50 Essential Climate Variables (ECVs) defined by the Global Climate Observing System (GCOS).

One of the future trends for climate change is the intensification of the hydrological cycle (Trenberth, 2009; Seneviratne et al., 2010). The raise in temperatures will increase the evaporation of soil moisture from soil and the transpiration from vegetation (evapotranspiration), easing the probability of droughts and wildfires in some regions of the Earth. The increased water vapor will lead to more intense rainfalls and floods in other parts of the globe. In less than 100 years, soil moisture will suffer a global redistribution with maximum changes of +25 % and -25% during boreal summer (Seneviratne et al., 2010). Its monitoring is thus essential.

Global monitoring requires global observations. Passive microwave sensors onboard satellites, like AMSR-E/2, Aquarius, SMOS and SMAP, provide global observations of *surface* SM. Surface SM can also be quantified from the ground via *in situ* measurement networks. There are important differences in the spatial sampling of satellite and *in situ* sensors: Satellites provide SM estimations with a resolution of ∼40 km that fully cover the whole globe surface. In contrast,

ground sensors have a resolution of some centimetres, are distributed over smaller areas (catchment or regionwide extents) and do not fully cover the surface since the stations are separated by some kilometres. In consequence, there exist a *spatial scale mismatch* between satellite and *in situ* sensors that capture different regimes of SM variability, both in space and time.

This chapter provides some introductory material on surface SM and its importance in the climatic system. It also presents the research problem of this thesis: the spatial scale mismatch between measurement techniques. It is described in detail along with the terminology that will be used in this thesis manuscript.

1.1 The surface soil moisture (SM) variable

The soil moisture (SM) variable participates in the exchanges of water and energy between the land surface and the atmosphere. As a consequence, it plays an important role in climatic, atmospheric, hydrologic and ecological processes (Rodriguez-Iturbe, 2000; Daly and Porporato, 2005; Legates et al., 2011). The following sections describe the principal contributions of SM to the climatic system and diverse applications.

1.1.1 Definition

Soil moisture is the amount of water stored in the unsaturated soil zone. The term *surface* SM is used when it only concerns the top 10 cm of soil. It can be expressed using different conventions, like the gravimetric (mass of water divided by mass of dry soil) or the volumetric (volume of water divided by volume of total soil sample). This thesis is dedicated to the study of surface SM in the top 5 cm of soil, which is the average sensing depth of the two SM-dedicated satellite missions (SMOS, SMAP) and of most of the operational *in situ* networks. The volumetric unit convention for SM is adopted and expressed as m^3/m^3.

L'humidité du sol est une variable positive bornée. Le minimum théorique est 0 m^3/m^3 (sol totalement sec) mais en pratique de l'eau résiduelle reste dans le sol. Par exemple, dans la zone racinaire, les plantes ne sont pas capables d'extraire de l'eau au-dessous d'un seuil spécifique (*point de flétrissement*). Concernant la valeur maximale d'humidité du sol, deux seuils sont possibles. À *saturation*, tous les pores sont remplis d'eau. La saturation arrive après de fortes précipitations et induit une forte probabilité de ruissellement.

SM is a bounded positive variable. The theoretical minimum is zero (totally dry soil), but in practice, residual SM remains. For ex ample, in the root zone, plants are not able to extract water below a specific threshold (*wilting point*). Concerning the upper limit, two thresholds are possible. At *saturation*, all pores are filled with water. Saturation takes place after strong precipitation and makes lateral runoff highly probable. After some time, part of the water flows in the form of *drainage*, plant water intake or soil evaporation and plant transpiration (*evapotranspiration*). The *field capacity* threshold is reached as soon as some air enters in the soil pores. A more detailed review of these processes is given by the Food and Agriculture Organization (FAO) of the United Nations (Brouwer et al., 1985, ch. 2) and by Western et al. (2002).

1.1.2 SM related processes

SM partitions the quantity of water and energy transferred between the land and the atmosphere. Figure 1.2 illustrates the main energy and water processes where SM is involved. Regarding the water fluxes, the principal forcing of SM is precipitation. This can be easily seen in Figure 1.1 which presents the SM and precipitation time series measured at the same location.

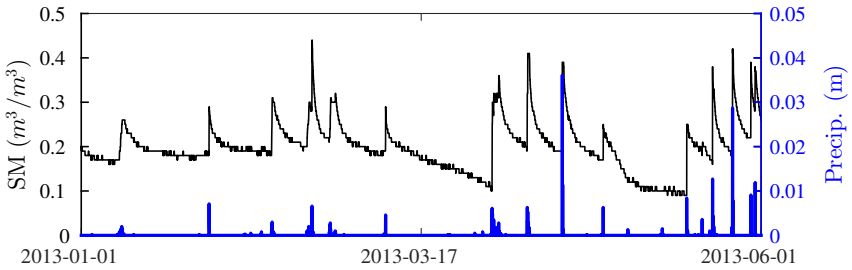


Figure 1.1 – Example of *in situ* SM and precipitation time series. The series were measured at station #9 of the Little Washita network (Oklahoma, U.S.).

Surface SM is depleted by downwards drainage and evapotranspiration. Lateral runoff can either contribute or reduce the SM levels at a location, depending on the direction of the lateral flow. Other processes like capillary rise or the increase of the water table can induce an increment of surface SM levels.

The energy and moisture fluxes are related via *evapotranspiration*. It is estimated that more than a half of the total solar energy absorbed by the land is spent in evapotranspiration (Oki and Kanae, 2006). In short, the change of energy within the considered surface layer equals the *net radiation* minus the *latent heat flux*, which is proportional to evapotranspiration. Other terms also contribute to the energy exchange: the *transmission of heat to deeper layers* and the *sensible heat flux*, associated with changes in temperature. More thorough descriptions of the processes are available in Western et al. (2002) and Seneviratne et al. (2010).

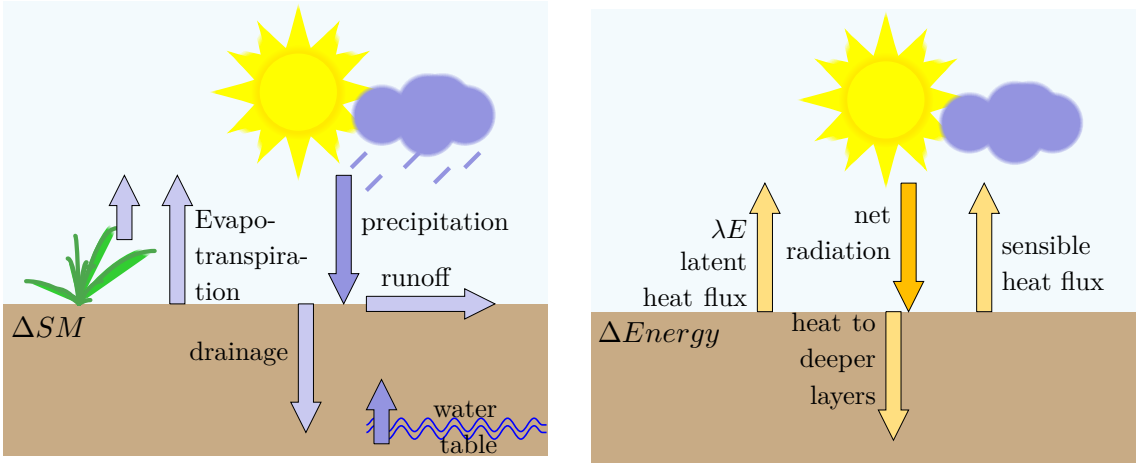


Figure 1.2 – Schematic of the water and energy fluxes at the land-atmosphere boundary

1.1.3 Why is SM information relevant

Surface SM plays a major role in different disciplines. Surface SM is a key variable in land-atmosphere interactions through evapotranspiration and energy fluxes, so it is of high value in numerical weather predictions (NWPs) and climatic general circulation models (GCMs). Because it participates in the water cycle, surface SM also helps to improve hydrological forecasts. In agriculture, surface SM is linked to root zone soil moisture (RZSM) and thus, to yield production. Finally, ecological processes can be influenced by SM since it contributes to the available carbon in soil, plant growth, plant species distribution and the presence of other living organisms such as bacteria and insects.

Climate and meteorological forecasting. The Earth's climate can be seen as the result of a dynamic feedback equilibrium where oceans and land affect the atmosphere and the other way

round (Douville, 2004). The land-atmosphere boundary is specially heterogeneous in space so information on surface SM can help for a better modelling of the boundary. The importance of SM is strongest in transitional regions between wet and dry climates (Seneviratne et al., 2010). Given that SM is both a water and energy storage, it induces persistence in the climate system (e.g. Lorenz et al., 2010; Nicolai-Shaw et al., 2016) and SM initialization is useful in seasonal forecasting (e.g. Douville, 2004; Koster et al., 2010).

Similarly, surface SM may play a important role in weather forecasting as it has been shown that some regions exhibit strong coupling between the land surface and the lower atmosphere (Koster et al., 2004; Seneviratne et al., 2010). In these regions, it influences key low atmosphere variables like relative humidity and temperature of the air (Brocca et al., 2017). However, the assimilation of SM data from the ASCAT satellite in the meteorological models at the European Centre for Medium-range Weather Forecasting (ECMWF) (De Rosnay et al., 2013) and the UK Met Office (Dharssi et al., 2011), has not shown considerable improvement of the SM analysis when compared to *in situ* measurements. In contrast, the assimilation of SM data from the SMAP satellite (Kolassa et al., 2017) and from the SMOS satellite (Rodríguez-Fernández et al., 2017a) has provided promising results. Finally, the improvements reported in the analysis of SM seem to be concentrated in some regions of the Earth. The regions are different depending on the satellite sensor (ASCAT, SMOS) and the assimilation scheme (Rodríguez-Fernández et al., 2017a).

Hydrological modelling and land applications. The assimilation of surface SM in hydrological models is useful for various applications. In the case of catchment-scale moddels, SM assimilation can improve the accuracy of different output variables like water table depth and streamflow (Leroux et al., 2016; Crow et al., 2017) and discharge (Laiolo et al., 2016). By the same means, satellite rainfall estimates, which have been shown to be quite inaccurate in some regions of the globe, can be corrected (Brocca et al., 2014; Pellarin et al., 2017).

Land applications like flood and drought forecasting also benefit from improved predictions thanks to surface SM assimilation. SM serves as a proxy for the amount of water excess and deficit because it partitions the incoming precipitation into infiltration, runoff and evapotranspiration (Seneviratne et al., 2010). Flood modelling can be direclty improved by correcting the rainfall estimations assimilated in rainfall-runoff models (Massari et al., 2014). This has been also observed in operational systems like the European Flood Awareness System (EFAS), which improved flood predictions after assimilation of SM in its internal hydrological model (Wanders et al., 2014b).

Regarding droughts, surface SM can help predict agricultural droughts (Velpuri et al., 2016; Bolten et al., 2016; Martínez-Fernández et al., 2016). While the meteorological drought responds to a simple lack of precipitation, the agricultural drought appears when the crop production is impaired as a result of not enough water resources for the plants (Mishra and Singh, 2010). Different indices have been used to predict agricultural droughts. Torres et al. (2013) based the drought predictions on historical records of the soil water deficit (SWD), an index only dependent on SM and soil depth. The U.S. Drought Monitor (Svoboda et al., 2002) provides weekly maps based on 5 severity levels of drought calculated from a number of indicators and variables blended together, SM included. Al Bitar et al. (2013) proposed to predict drought from the root zone soil moisture (RZSM), calculated with a double bucket model that assimilates surface SM and leaf area index (LAI) variables.

Agricultural applications like forecasting of crop production directly derive from the research on the relation between SM and agricultural droughts. For example, the RZSM derived from surface SM observations has allowed Gibon et al. (2017) to explain 89 % of the crop yield variation in Niger, 72 % in Burkina Faso, 82 % in Mali and 84 % in Senegal, after assimilation in a statistical model. Similarly, surface SM observations allowed Ines et al. (2013) to improve maize crop predictions. Other agricultural applications include detection of irrigation for water

resources management. For example, Singh et al. (2016) used SM data from the AMSR-E satellite to reveal a temporal shift in irrigation practices in India between 2002 and 2011.

1.2 Measurement of SM

There are several technologies providing SM estimates. **Direct measurements** of SM are only possible with destructive methods like gravimetric measurements. They consist in evaluating in laboratory the quantity of water evaporated from a volume of soil that was previously extracted from the area of study. Direct methods are manual and long SM time series can only be provided by **indirect methods**. Indirect methods measure a physical variable that is strongly linked to SM (e.g. surface emitted energy, capacitance of soil, travel time of a signal in the soil) that is then translated into SM. These sensors operate at various spatial scales.

The term *scale* refers to a triplet of entities, which is represented in Figure 1.3: the **extent**, the area enclosing all the measurements; the **spacing**, the distance between measurements; and the **support**, the area actually sensed by the instrument (Blöschl and Sivapalan, 1995). SM measurement systems can be classified according to this scale triplet. In this thesis, two types of observing systems will be distinguished: those whose support is *large* ($\sim 40^2$ km^2) and those whose support is *small* ($<1^2$ m^2).

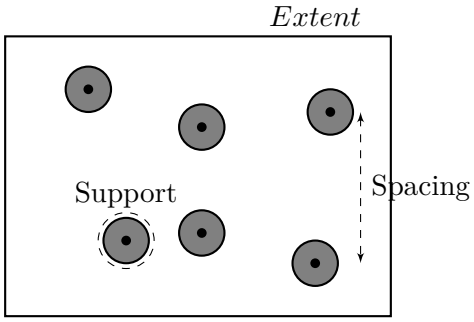


Figure 1.3 – Schematic of the scaling triplet: extent, support and spacing

1.2.1 Large-support observations: satellite sensors

Remote sensing instruments can be operated from the ground (placed on towers or elevated places), from airborne platforms or from satellites. In the first two cases, the support varies between some meters to $\sim$1 km, but the observed extents are limited to the watershed scale (10^2 km^2) at best. In the case of satellite sensors, global coverage is often achieved in some few days and the surface is fully and regularly sampled.

The support of satellite observations depends on the type of instrument in a first place. Surface SM can be derived from optical sensors, in which case the support is defined by the optical pixel and ranges between some meters to some few kilometres. However, no readily global operational optical-SM products are available since they suffer from diverse major drawbacks: They are sensitive to the skin soil layer where SM is controlled by top surface phenomena such as wind and shadowing, to cite some. In addition, optical observations cannot be acquired under cloudy and night-time conditions and they are highly sensitive to the Earth's atmosphere and to different surface elements (vegetation, roughness, etc.) (Petropoulos et al., 2015).

In the case of **microwave sensors**, the support is equivalent to the resolution of the instrument and it is defined by the antenna footprint as the area containing half of the total signal power. In this thesis, the words **support**, **resolution** and **footprint** will be used indistinctly to refer to the size of the support of microwave observations. This one varies from some meters to

hundreds of kilometres, depending on the frequency measured by the sensor and other factors (distance to the Earth surface, sensor characteristics, inclination, etc.). Although they do not suffer from the same drawbacks as optical sensors, microwave sensors can still be sensitive to surface elements like vegetation and surface roughness: a careful choice of the sensor regarding the antenna configuration (active or passive) and the microwave frequency can mitigate these influences. Active sensors transmit a signal to the Earth that they analyse after reflection, refraction and scattering by the Earth's constituents while passive sensors simply measure their respective natural emissions. Active sensors are more sensitive than passive sensors to the scattering produced by vegetation structure and surface roughness, among other factors. In contrast, they can provide observations with resolutions up to some tens of meters. Active sensors include synthetic aperture radars (SARs) (ERS, ALOS, Sentinel 1) and scatterometers (ASCAT). **Passive instruments (radiometers)** have coarser resolutions of some tens of kilometres but are more immune to surface elements. C- and X-band radiometers like AMSR-E/2 and WindSat (Wagner et al., 2007; Mladenova et al., 2011; Parinussa et al., 2012) as well as L-band radiometers like SMOS and SMAP (Al Bitar et al., 2012; Kerr et al., 2016; Colliander et al., 2017) have shown good skills in estimating surface SM.

This thesis is dedicated to the study of SM observations from satellite radiometers, like AMSR-E, AMSR2, SMOS or SMAP, when compared to point-support measurements. Although they present important differences in terms of sensor technology and retrieval models, they all coincide in having similar resolutions (27^2-55^2 km^2) and similar scheme for inferring SM. Their main characteristics are summarised in Table 1.1. The energy measured by radiometers at a specific microwave frequency or brightness temperature (TB) is transformed into SM through the inversion of a series of models that depend on the surface characteristics (land cover, roughness, soil texture, temperature). A general diagram of a typical retrieval system is presented in Figure 1.4. Hereafter, any reference to satellite observations and supports will concern this kind of sensors.

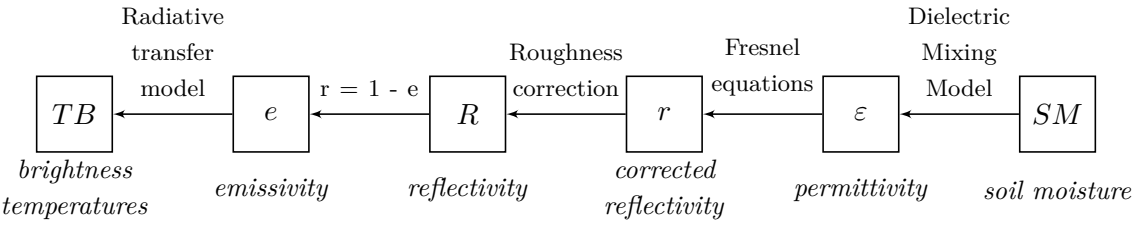


Figure 1.4 – Schematic of the transformation of brightness temperature (TB) into SM in radiometer-based systems. Note that the arrows indicate the direction of the modelling.

1.2.2 Point-support measurements: *in situ* sensors

Typical *in situ* measurements have a support of just some few centimetres. One special type, cosmic-ray neutron measurements, can reach resolutions of 200-300 m (Zreda et al., 2012, COSMOS probes). The resolution mismatch with respect to satellite measurements is similar, so all *in situ* measurements are considered as *point* observations. Hereafter, the terms **point-**, **small-** and **local-support** will be used indistinctly to refer to this kind of measurements. In practice, *in situ* measurements can represent a larger area (**representativeness area**) because the factors driving SM variability (vegetation, soil texture, topography, rainfall) are spatially connected. The representativeness area can be considered as the *effective* support of the *in situ* measurement. The estimation of the representativeness area is not straightforward and highly dependent on the method and criteria followed to estimate it. In any case, it cannot cover the full satellite footprint because SM-related factors are never perfectly homogeneous within the footprint.

Permanent *in situ* stations consist in probes installed permanently in the soil. They usually

belong to a network, which is a group of spatially distributed *in situ* stations managed by the same organisation and designed to monitor SM over a specific area. Networks vary in extent, spacing and number of stations. They can be divided into two types depending on the number of stations available per satellite footprint. Networks with only one station per footprint are called **sparse networks**. This implies that their minimum spacing is equal or larger than the satellite resolution. They are also referred as *large-scale* networks because they usually cover extents larger than 10,000 km^2 (Crow et al., 2012). Networks with multiple stations per footprint are called **dense networks** or *small-scale* networks because their extents vary between 100 km^2 to 10,000 km^2 (Crow et al., 2012). As a consequence, their spacing is smaller than the satellite footprint. Examples of sparse networks include the SMOSMANIA network in France (Calvet et al., 2007) and the SCAN network in USA (http://www.wcc.nrcs.usda.gov). Examples of dense networks are the watersheds networks in USA, such as Little Washita and Walnut Gulch (Cosh et al., 2006, 2008), the Yanco network in Australia (Smith et al., 2012) and the Monte Buey network in Argentina (Thibeault et al., 2015). For a detailed description of available *in situ* networks and their respective spatial characteristics the reader is referred to Crow et al. (2012). The main characteristics of *in situ* measurements are summarised in Table 1.1.

There is a variety of *in situ* sensor types: capacitance sensors, impedance sensors, time- and frequency domain reflectometry (TDR, FDR) sensors and time- and frequency-domain transmissometry (TDT, FDT) sensors (Robinson et al., 2008). Although the variable they measure is different (for example, the charge time of a capacitor in the case of capacitance sensors and the travel time of a a voltage pulse in the case of TDR sensors), they all take advantage of the difference between the relative permittivity (ε) of dry soil and water (Topp, 2003). Basically, the raw measured variable is translated into ε, then ε is translated into SM using dielectric mixing models or empirical calibrations.

TDR/FDR and TDT/FDT are considered more accurate than impedance and capacitance sensors. The latter are on the contrary cheaper and less power consuming (Bircher et al., 2016), which makes possible to sample more densely in time and space. For this reason, they have been widely used in satellite validation campaigns (Albergel et al., 2012), design of *in situ* networks (Bircher et al., 2012) and SM spatio-temporal studies (Famiglietti et al., 1999).

Table 1.1 – Main characteristics of *in situ* and satellite SM measurements.

	Ground measurements (permanent networks)	**Microwave remote sensors (radiometers)**
Support	Point	27^2-55^2 km^2
Extent	Dense networks: $< 50^2$ km^2 Sparse networks: $>> 50^2$ km^2	Global
Spacing	Dense networks: 1-10 km Sparse networks: $>$ 50 km	Equals the support
Spatial sampling	Irregular surface is not fully sampled	Regular surface is fully sampled
Temporal sampling	Sub-daily, regular	1-4 days, regular or irregular
Temporal coverage	$>$ 1 year (depending on the network)	Since the 2000s: AMSR-E (2002-), WindSat (2013-), SMOS (2009-), AMSR2 (2012-), SMAP (2015-)
Advantages	High temporal resolution	Global coverage
Disadvantages	Geographically limited, expensive maintenance, large periods of sensor fault are usual	Lower temporal resolution Reduced quality under some conditions

1.3 Error budget and spatial scale mismatch

Ground and satellite sensors provide SM estimations with very different spatial support sizes an sampling densities. As a consequence, ground sensors give access to the state of the SM variable and related processes at particular points in space while satellite sensors give a spatially integrated view. This observational mismatch is referred as the **spatial scale mismatch** between ground and satellite SM measurements and is the problem investigated in this thesis.

Despite the spatial scale mismatch, the accuracy of satellite observations is often assessed by direct comparison with *in situ* measurements. Therefore, the comparison does not only contain satellite errors but also *in situ* errors and the spatial scale mismatch. This section describes the error budget of satellite-insitu comparisons, with special attention to the spatial scale mismatch.

1.3.1 What is soil moisture at 40 km resolution?

In order to define the spatial scale mismatch, the definitions for the SM observed at the local and the footprint scales need to be settled first. The definition for the point SM x_i at a location i seems straightforward and there is unanimity in accepting that the SM measured with indirect and direct methods should be the same, regardless instrumental and modelling errors. The meaning of the footprint-support SM y is, however, not clear. There exist two main interpretations for the footprint-support SM:

Definition I: y is the spatial mean of X, i.e. the x_i measured at all the infinitesimal points i within the footprint support A. If $f(X)$ is the probability density function (PDF) of X, then the footprint SM is expressed as

$$y_I = \mu_X = \int_A x_i f(x_i) dx_i \tag{1.1}$$

In the case of *in situ* dense networks, this footprint soil moisture y_I is approximated by linear averaging the values recorded by the stations at each time step (also called *network average*, NAvg).

Definition II: y is the SM value that corresponds to the measured energy level TB integrated over the footprint. The retrieval system (Figure 1.4) can be summarised as a global function $G(\cdot)$ that obtains SM from the ensemble of TB_i and a set of parameters $\{P_1, P_2, \ldots, P_K\}$. Let $f(TB, P_1, \ldots, P_K)$ be the joint density function of the TB_i variable and the parameter set, then the footprint SM is expressed as

$$y_{II} = \int_A \int_{P_1} \cdots \int_{P_K} G(TB_i, P_1, \ldots, P_K) f(TB_i, P_1, \ldots, P_K) dTB_i dP_1 \ldots dP_K \tag{1.2}$$

This is basically how radiometers perceive the footprint SM. The footprint soil moisture y_{II} is not the average of the infinitesimal punctual SM but the translation of the average of the infinitesimal punctual TBs into SM. As it will be explained later, even if the systems, parameters and models were free of errors, y_I is not equal to y_{II} in general.

1.3.2 Observation systems errors and uncertainties

In the SM literature, when the accuracy of observational systems is addressed, the terms *error* and *uncertainty* are usually considered as synonyms. This does not have significant consequences

in practice, but it is important to distinguish them when defining the contributions to the spatial scale mismatch.

The level of **uncertainty** is the level of lack of knowledge about how the observed state or variable should behave. The lack of knowledge associated to the number of observations or samples is called the *sampling uncertainty*. Generally speaking,our knowledge about the true state of our variable increases when the number of samples increases. For example, the saturation point estimated from a 1-week SM series is much more uncertain that the saturation point estimated from a 1-year series. Similarly, the uncertainty of the average SM for a particular area increases with decreasing number of sampled locations.

Another type of uncertainty is the *epistemic uncertainty* that refers to a fundamental ignorance about the true state or process. The footprint SM is subject to epistemic uncertainty because there is no unequivocal definition for it (section 1.3.1). Other parameters and variables can also be concerned by epistemic uncertainty. For example, the definition of the "land cover at 40 km resolution" is not clear if half of the pixel is urban and half is forest. This epistemic uncertainty about footprint-scale variables entails *scaling uncertainty* because it is not known how the variable should be scaled from the small to the large support.

The term **error** does not allude to the lack of knowledge but to the level of deviation of the observed state from the true state. For example, supposing that gravimetric measurements are free of errors, if the capacitance SM measurements at the same location differ, the difference is the *measurement error* of the capacitance probe. In the case of errors, usually the true state cannot be known but the definition of the true state remains known. For example, the definition of the energy emitted by the Earth within a 40-km area is known and it can be analytically expressed, but it can only be accessed via the observation systems (antennas) that are subject to errors (not uncertainty). The variables and parameters used in SM retrieval models are also affected by classification errors. These arise when the class associated to a location or area does not coincide with the reality. Finally, the term *error* can also be synonym of *inaccuracy*: a model is a simplification of the reality that could never reproduce its whole complexity. Inaccuracies in the representation of the reality are unavoidable and lead to errors in the model outputs.

The list of uncertainties and errors in observed SM datasets can be summarised as follows:

- Measurement and instrumental errors: errors in the measured variable (TB, capacitance of soil, travel time of a signal, etc.) due to the instrument or human actions.
- Errors in input parameters describing soil and land cover (soil texture, vegetation type and state, etc.).
- Inaccuracies of the models (radiative transfer, dielectric mixing model, etc.).
- Scaling uncertainty (in large-support systems): level of ignorance about how the footprint-support variable should be at the large-support scale (e.g. SM, land cover).
- Sampling uncertainty: level of ignorance about the variable given the available number and the location of samples.

1.3.3 *In situ* and satellite: the spatial scale mismatch

The spatial scale mismatch encompasses all the factors contributing to a mismatch in the spatial scales of the observation systems, the satellite and the *in situ*. Three elements contribute to the spatial scale mismatch: the *sampling* and the *scaling* uncertainties, introduced in the previous section, and the *geophysical mismatch*. The geophysical mismatch concerns the different processes and scales of processes captured by both observation systems as a result of their different

supports (footprint *vs* point). The elements of the spatial scale mismatch are summarised in Figure 1.5 and described in the following paragraphs.

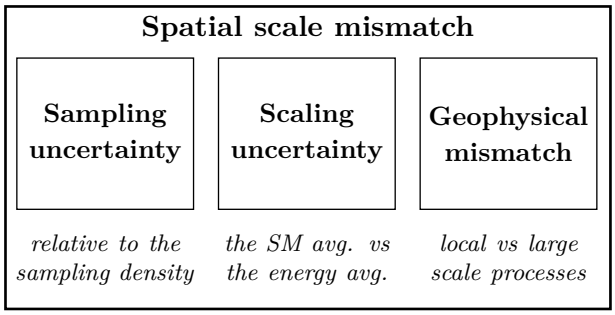


Figure 1.5 – The 3 components of the spatial scale mismatch

***In situ* sampling uncertainty:** Point *in situ* measurements are a proxy to the footprint SM as defined in Definition I (section 1.3.1): they can be averaged together into the network average (NAvg), which is an estimate of the true footprint soil moisture y_I. The uncertainty of the NAvg decreases with the number of spatial samples and depends on the location of the stations. Multiple sampling techniques have been used to estimate this uncertainty, with more or less success: statistical, geostatistical, stratified and bootstrap sampling. A descriptive review can be found in (Wang et al., 2008, 2015a).

Scaling uncertainty: When *in situ* and satellite observations are compared, the scaling uncertainty is made of two components:

(i) the epistemic uncertainty about about how the footprint-support SM should be (y_I or y_{II})
(ii) the differences between the (true) footprint SM derived from *in situ* measurements (y_I) and that derived from satellite observations (y_{II}).

The first component has been already explained in the previous section. The second makes reference to the fact that y_I and y_{II} are fundamentally different because of the non-linearity of the SM retrieval process. This difference is typical from non-linear systems where the average of the processed inputs is not equal to the process of the averaged inputs. In the case of satellite systems, there are two main sources of non-linearities. On the one hand, the function linking TB and SM is non linear (see Figure 1.6 for the non-linear link between dielectric constant ε and SM). On the other hand, model parameters, like soil and land cover, are fundamentally heterogeneous within the footprint and each soil/land cover type exhibit different functional relationships between ε and SM. Numerous studies have confirmed that this scaling uncertainty is non-negligible in the case of remote sensing radiometers (Crosson et al., 2010), such as AMSR-E (Crow et al., 2001), SMOS (Pellarin et al., 2003), Hydros (Crow et al., 2005) and SMAP (Zhan et al., 2008).

Geophysical mismatch: The geophysical mismatch is related to the observation support: while the satellite integrates the processes within a surface, the *in situ* probe observes a point inside it. In other words, they observe the same variable from different perspectives. For this reason, they capture different processes or different scales of the same processes. For example, when precipitation takes place within the footprint, the SM observed at a specific location is probably different from the averaged SM response of the footprint. This is easily understood by analogy to zooming in and out a picture. Figure 1.7 illustrates it: if the observational reference is set at the local scale, the observed value can be "blue", "yellow" or "green"; when the observational reference is set at the large scale, the observed value is an integral of the small scale values, "green" in this case.

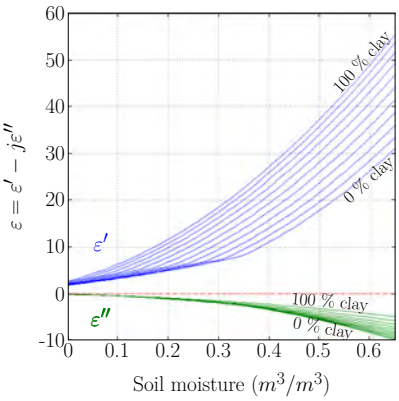


Figure 1.6 – Relationship between dielectric constant and SM according to the Mironov's model (Mironov and Fomin, 2009; Mironov et al., 2009) used in the SMOS SM retrieval processor (Kerr et al., 2014) for different clay contents.

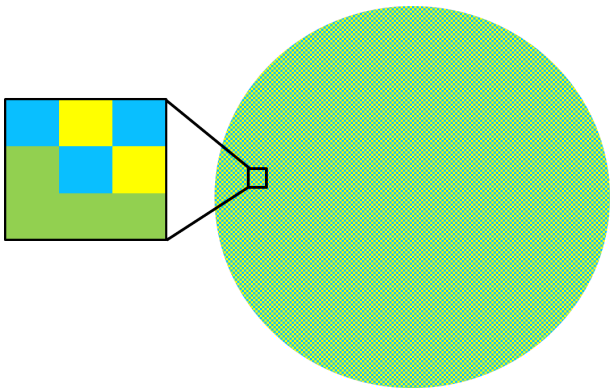


Figure 1.7 – Conceptual representation of the geophysical mismatch

1.3.4 Terminology

Herein, I summarise the terminology used in this manuscript concerning spatial scales. In terms of support, two main scales are distinguished (Table 1.2): the point support (*in situ* SM measurements) and the large support (satellite SM estimates). A third support, the small support, is added to refer to high-resolution SM model datasets. Note that this nomenclature applies to the physical supports, not to the *effective* supports or representativeness areas of *in situ* stations. Regarding the extent, the most commonly used terms are listed in Table 1.3.

Table 1.2 – Principal spatial supports

Spatial supports	
local/point	$\lesssim 1^2$ m^2
small/high-resolution	$\lesssim 1^2$ km^2
large/footprint/low-resolution	$\sim 40^2$-50^2 km^2

Table 1.3 – Principal spatial extents

Spatial extents	
field	$\lesssim 100^2$ m^2
hill-slope	$\sim 100^2$ m^2
catchment	$\sim 10^2$ km^2
footprint	$\sim 40^2$-50^2 km^2
regional	$\sim 1{,}000^2$ km^2

1.4 Questions and objectives of this thesis

The accuracy of satellite SM observations is usually evaluated by direct comparison to *in situ* measurements. Given the limited spatial coverage of *in situ* networks, most validation campaigns consist in confronting the observations in time, i.e. confronting a satellite and an *in situ* time series taken at a particular location or region. The statistics derived (correlation, bias, RMSE) are supposed to represent satellite skills. Nevertheless, they are also affected by ground probe errors and the spatial scale mismatch between the satellite and the *in situ* observations. While *in situ* errors are expected small (<0.05 m^3/m^3) if the *in situ* network is properly maintained, the contribution of **the spatial scale mismatch is supposed important although it remains undetermined**. In consequence, the inherent errors of the observation systems remain also undetermined.

Figure 1.8 provides a preliminary view of the spatial scale mismatch. It shows the comparison between one SMOS time series and the *in situ* time series measured at 20 different locations within the same footprint. The statistics are presented under the form of a Taylor plot (Taylor, 2001), which allows to summarise all the validation results of one network in just one plot. The

bias is in colour scheme, the correlation in the outer circle and the standard deviation of the differences (STDD, or *unbiased* RMSE) in the inner green circle. The scatter in the markers suggests that the magnitude of the spatial scale mismatch in this area is non-negligible. As an example, the STDD ranges from 0.045 to 0.08 m^3/m^3, depending on the location chosen for the comparison. However, it is not known what percentage of the STDD is attributed to the spatial scale mismatch and what to the satellite and *in situ* errors.

The spatial scale mismatch is not only a spatial problem but also a temporal problem. SM fields are not stationary in time and the spatial redistribution of SM changes in time. This means that the distance in terms of SM between one particular location and the (unknown) large-support observation varies in time. Therefore, the spatial scale mismatch will more likely require to be assessed with spatio-temporal approaches.

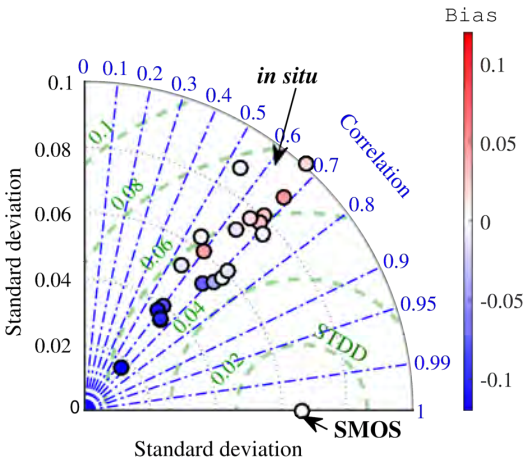


Figure 1.8 – Comparison of one SMOS soil moisture time series with 20 *in situ* time series corresponding to the same SMOS footprint

The spatial scale mismatch is a recurrent problem in the validation of satellite SM products. It is present since the launch in the late 70s of the Scanning Multichannel Microwave Radiometer (SMMR), which was first used to retrieve SM from space. Since then, various approaches have been developed to either overcome the spatial scale mismatch or to evaluate it. For example, *downscaling methods* have been conceived to improve the resolution of satellite SM observations with the aid of auxiliary higher-resolution data. Statistical analysis approaches like *triple collocation* (TC) have been also adapted to assess the spatial scale mismatch between a location and the footprint SM. The drawbacks of the methods proposed to date is that they are specific to a region, to a series of landscape and climatic conditions and/or they depend on a set of statistical hypotheses that are not always met and difficult to verify. As a consequence, **there is no standard approach to handle the spatial scale mismatch** and, as a proof, satellite observations are still validated operationally without taking it into account.

Different **questions** arise when addressing the problem of the spatial scale mismatch:

- what is a good definition for the SM at the satellite resolution?
- what is exactly behind the spatial scale mismatch?
- is the spatial scale mismatch measurable?
- is it possible to quantify its impact in the classical validation metrics?

This thesis seeks to **contribute to the resolution of the spatial scale mismatch** problem and related questions. The scope is to **improve the current knowledge on the spatial scale mismatch** and to **provide novel or improved methods** for its assessment. The proposed approaches need to signify a breakthrough with respect to existing ones in terms of standardisation: they should be **easily applicable to all validation regions**. To to this, the number of geophysical and statistical assumptions should be kept as low as possible. This should prepare the path for an integrated assessment of the spatial scale mismatch in satellite validation campaigns.

1.4.1 Manuscript structure

This manuscript is structured in 6 chapters, including the introduction and the conclusion. In this first introductory chapter, I have described the research problem and questions. The spatial scale mismatch problem and related concepts applied to the surface SM variable will be the guiding thread for the rest of the text.

Chapters 2 and 3 provide a solid background in SM spatial and temporal variability as well as they introduce most of the datasets and techniques that will be used in the following chapters. **Chapter 2** addresses the study of the spatial variability of SM within a typical satellite footprint. It describes the state of the art concerning the physical drivers and the statistical behaviour of SM fields. Then, the DISPATCH downscaling method is evaluated, which illustrates the potentials and drawbacks of downscaling approaches for overcoming the spatial scale mismatch. **Chapter 3** is dedicated to the temporal variability of SM. The time scales of several SM datasets are analysed with the help of wavelet transforms, a signal processing tool that has been marginally used to date in the SM literature.

This thesis develops the hypothesis that the spatial scale mismatch can be assessed from the temporal domain because SM spatial and time scales are connected. **Chapter 4** is dedicated to the description and evaluation of this core hypothesis. To this end, it provides novel methods based on wavelet analysis to assess the spatial scale mismatch on a per time-scale basis. The experiments are applied to modelled and measured SM time series from different typical satellite validation regions.

While chapter 4 addresses the spatial scale mismatch as a whole, **chapter 5** is focused on the sampling uncertainty. It is dedicated to the transformation of the *in situ* sampling uncertainty (spatial) into uncertainty in the usual satellite validation metrics (temporal). In other words, the scope of this chapter is to provide confidence intervals for the temporal metrics used in satellite validation campaigns (correlation, bias, RMSE) based on the *a priori* known sampling uncertainty. The concepts of chapter 2 will be of use here.

Finally, the last chapter is dedicated to the general conclusions of the thesis as well as the research perspectives that stem from the works presented here.

Chapitre 1

Introduction (français)

Contents

1.1 La variable humidité de surface du sol 16
1.1.1 Définition 16
1.1.2 Processus liés à l'humidité du sol 16
1.1.3 Pourquoi connaitre l'humidité du sol est-il pertinent ? 18
1.2 Mesurer l'humidité du sol 19
1.2.1 Observations ayant un grand support : les capteurs satellitaires 20
1.2.2 Mesures ponctuelles : les capteurs *in situ* 21
1.3 Bilan d'erreur et d'écart d'échelle spatiale 22
1.3.1 Qu'est-ce que l'humidité du sol à une résolution de 40 km ? 22
1.3.2 Erreurs et incertitudes des systèmes d'observation 23
1.3.3 Mesures *in situ* et satellite : la différence d'échelle spatiale 24
1.3.4 Terminologie 26
1.4 Questions et objectifs de cette thèse 26
1.4.1 Structure du manuscrit 28

L'humidité du sol représente environ 0,009 % du stockage total de l'eau terrestre et environ 0,8 % des ressources liquides en eau douce de la Terre (Trenberth et al., 2007), il s'agit pourtant d'une variable géophysique fondamentale qui participe aux échanges d'eau et d'énergie entre la terre et l'atmosphère. Depuis 2010, l'humidité du sol fait partie des 50 variables climatiques essentielles (ECV) définies par le Global Climate Observing System (GCOS).

L'une des tendances futures du changement climatique est l'intensification du cycle hydrologique (Trenberth, 2009 ; Seneviratne et al., 2010). La hausse des températures augmentera l'évaporation de l'eau du sol et la transpiration de la végétation (évapotranspiration), augmentant de ce fait la probabilité de sécheresses et de feux de forêt dans certaines régions du globe. L'augmentation de la vapeur d'eau entraînera des pluies et donc des inondations plus intenses dans d'autres parties du globe. En moins de 100 ans, l'humidité du sol devrait provoquer une redistribution globale avec des changements maximaux allant de +25 % à -25 % pendant l'été boréal (Seneviratne et al., 2010). Son suivi est donc essentiel.

Un suivi global de l'humidité du sol nécessite des observations globales. Les capteurs à micro-ondes passives à bord des satellites, comme AMSR-E / 2, SMOS, Aquarius et SMAP, fournissent des observations globales de l'humidité *de surface* du sol. Cette humidité peut également être mesurée au sol via des réseaux de capteurs *in situ*. Il y a des différences importantes entre l'échantillonnage spatial des capteurs satellitaires et celui des capteurs *in situ* : les satellites

fournissent des estimations d'humidité du sol avec une résolution de ∼40 km et qui couvrent l'ensemble du globe. En revanche, les capteurs terrain ont une résolution de quelques centimètres, sont répartis sur des zones plus petites (à l'échelle de la parcelle ou de la région) et ne couvrent pas complètement la surface puisque les stations sont souvent séparées par plusieurs kilomètres. Par conséquent, il existe une *différence d'échelle spatiale* entre la mesure satellite et celle des capteurs *in situ* qui sont représentatifs de différents régimes de variabilité de l'humidité du sol, tant dans l'espace que dans le temps.

Ce chapitre fournit des éléments introductifs à propos de l'humidité de surface du sol et de son importance dans les systèmes climatiques. Il présente également la problématique de recherche de cette thèse : la différence d'échelle spatiale entre les différentes techniques de mesure. La problématique ainsi que la terminologie employée tout au long de ce manuscrit de thèse seront décrites en détail.

1.1 La variable humidité de surface du sol

La variable humidité du sol participe aux échanges d'eau et d'énergie entre la surface terrestre et l'atmosphère. Par conséquent, elle joue un rôle clef dans le climat et les processus atmosphériques, hydrologiques et écologiques (Rodriguez-Iturbe, 2000 ; Daly et Porporato, 2005 ; Legates et al., 2011). Les sections suivantes décrivent les principales contributions de l'humidité du sol au système climatique ainsi qu'à diverses applications.

1.1.1 Définition

L'humidité du sol est la quantité d'eau stockée dans la zone non saturée du sol. Le terme humidité *de surface* est utilisé lorsqu'il ne concerne que les 10 premiers centimètres de sol. Il peut être exprimé dans plusieurs types d'unités, comme par exemple l'unité gravimétrique (masse d'eau divisée par masse de sol sec) ou volumique (volume d'eau divisée par volume d'échantillon total de sol). Cette thèse est consacrée à l'étude de l'humidité de surface du sol contenue dans les 5 premiers centimètres de sol, soit la profondeur de détection moyenne des deux missions satellitaires dédiées à l'observation de l'humidité du sol (SMOS, SMAP) et de la plupart des réseaux opérationnels de mesures *in situ*. La convention volumique est adoptée et donc le niveau d'humidité du sol est exprimée en m^3/m^3.

L'humidité du sol est une variable positive bornée. Le minimum théorique est 0 m^3/m^3 (sol totalement sec) mais en pratique de l'eau résiduelle reste dans le sol. Par exemple, dans la zone racinaire, les plantes ne sont pas capables d'extraire de l'eau au-dessous d'un seuil spécifique (*point de flétrissement*). Concernant la valeur maximale d'humidité du sol, deux seuils sont possibles. À *saturation*, tous les pores sont remplis d'eau. La saturation arrive après de fortes précipitations et induit une forte probabilité de ruissellement. Après un certain temps, une partie de l'eau s'écoule sous forme de *drainage*, de consommation par les plantes et/ou d'évaporation du sol et transpiration des plantes (*évapotranspiration*). La *capacité au champ* (field capacity, en anglais) est atteinte à ce moment-là, dès que l'air rentre dans les pores du sol. Un examen plus détaillé de ces processus est donné par l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO, pour son sigle en anglais) (Brouwer et al., 1985, ch. 2) et Western et al. (2002).

1.1.2 Processus liés à l'humidité du sol

L'humidité du sol pilote les échanges d'eau et d'énergie à l'interface sol atmosphère. La Figure 1.2 illustre les principaux processus d'échange d'eau et d'énergie pour lesquels l'humidité du sol joue un rôle. En ce qui concerne les flux d'eau, le forçage principal de l'humidité du sol est

la précipitation. Cela se voit facilement sur la Figure 1.1, qui présente des séries temporelles d'humidité du sol et de précipitation mesurées au même endroit.

L'humidité de surface du sol diminue sous l'effet du drainage et de l'évapotranspiration. Le ruissellement latéral peut soit augmenter ou réduire les niveaux de l'humidité du sol par endroits, en fonction de la direction du flux latéral.

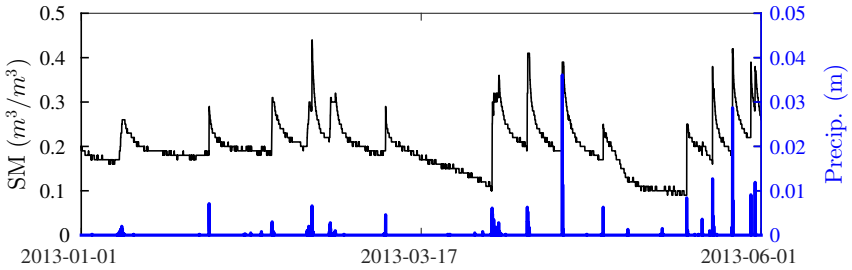


Figure 1.1 – Exemple des séries temporelles de mesures *in situ* d'humidité du sol (SM) et de précipitation. Les séries ont été mesurées à la station n°9 du réseau Little Washita (Oklahoma, États-Unis).

D'autres processus comme la remontée d'eau par capillarité ou la remontée de la nappe phréatique peuvent induire un accroissement des niveaux d'humidité de la surface de la terre.

Les flux d'énergie et d'humidité sont liés par l'*évapotranspiration*. On estime que plus de la moitié de l'énergie solaire totale absorbée par la terre est consommée par évapotranspiration (Oki et Kanae, 2006). En bref, les variations d'énergie dans la couche de surface sont égales au *rayonnement net* moins le flux de *chaleur latente*, lui-même proportionnel à l'évapotranspiration. D'autres termes contribuent également à l'échange d'énergie : la transmission de la chaleur aux couches plus profondes et le flux de *chaleur sensible*, associé à des changements de température. Des descriptions plus détaillées de ces processus sont disponibles dans Western et al. (2002) et Seneviratne et al. (2010).

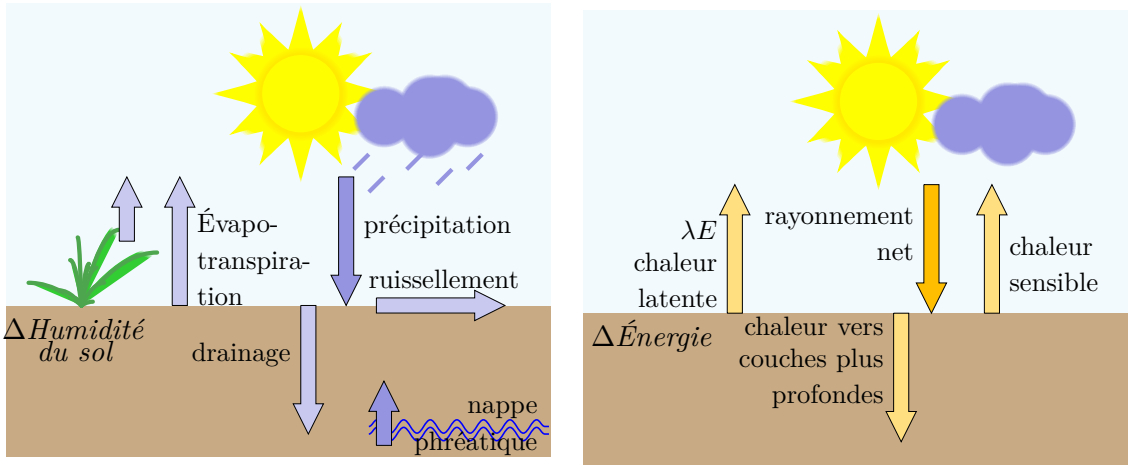


Figure 1.2 – Schéma des flux d'eau et d'énergie à l'interface terre-atmosphère.

1.1.3 Pourquoi connaitre l'humidité du sol est-il pertinent ?

L'humidité de surface du sol joue un rôle majeur dans différentes disciplines. C'est une variable clé dans les interactions terre-atmosphère par le biais de l'évapotranspiration et des flux d'énergie. Par conséquent, elle joue un rôle important dans les modèles de prévisions météorologiques numériques (NWP, pour son sigle en anglais) et dans les modèles climatiques de circulation générale. Puisqu'elle participe au cycle de l'eau, l'humidité de surface du sol contribue également à l'amélioration des prévisions hydrologiques. Dans l'agriculture, l'humidité de surface du sol est liée à l'humidité dans la zone racinaire et donc à la production agricole. Enfin, les processus écologiques peuvent être influencés par l'humidité du sol car elle contribue à la disponibilité de carbone dans le sol, la croissance des plantes, la répartition des espèces végétales et la présence d'autres organismes vivants comme les bactéries et les insectes ou les processus de nitrification.

Prévisions climatiques et météorologiques. Le climat de la Terre peut être considéré comme le résultat d'un équilibre dynamique où les océans et les terres affectent l'atmosphère e et réciproquement (Douville, 2004). L'interface terre-atmosphère est spécialement hétérogène dans l'espace, donc l'information sur l'humidité de surface du sol peut aider à une meilleure modélisation de l'interface. L'importance de l'humidité du sol est la plus forte dans les régions de transition entre les climats humides et secs (Seneviratne et al., 2010). Étant donné que l'humidité du sol est à la fois un stockage d'eau et d'énergie, elle induit une persistance dans le système climatique(e.g. Lorenz et al., 2010 ; Nicolai-Shaw et al., 2016) et l'initialisation par l'humidité du sol est utile dans les prévisions saisonnières (e.g. Douville, 2004 ; Koster et al., 2010).

De même, l'humidité de surface du sol peut jouer un rôle important dans les prévisions météorologiques puisqu'il a été montré que certaines régions présentent un fort couplage entre la surface et les basses couches de l'atmosphère (Koster et al., 2004 ; Seneviratne et al., 2010). Dans ces régions, elle influence des variables géophysiques de la basse atmosphère telles que l'humidité relative et la température de l'air (Brocca et al., 2017). Cependant, l'assimilation des données d'humidité du sol issues du satellite ASCAT dans les modèles météorologiques au Centre Européen de Prévision Météorologique à Moyen Terme (CEPMMT) (De Rosnay et al., 2013) et au UK Met Office (Dharssi et al., 2011), n'a pas montré d'amélioration considérable lors de l'analyse des humidités prédites par rapport aux mesures *in situ*. En revanche, les exercices d'assimilation des données d'humidité du sol provenant des satellites SMAP (Kolassa et al., 2017) et SMOS (Rodríguez-Fernández et al., 2017a) ont fourni des résultats prometteurs. Enfin, les améliorations rapportées dans l'analyse de l'humidité du sol semblent concentrées dans certaines régions de la Terre. Les régions sont différentes en fonction du capteur satellitaire (ASCAT, SMOS) et du schéma d'assimilation utilisé (Rodríguez-Fernández et al., 2017a).

Modélisation hydrologique et applications terrestres. L'assimilation de l'humidité de surface du sol dans des modèles hydrologiques est utile pour diverses applications. Dans le cas des modèles à l'échelle du bassin versant, l'assimilation de l'humidité du sol peut améliorer la précision de différentes variables de sortie comme la profondeur de la nappe phréatique et l'écoulement fluvial (Leroux et al., 2016 ; Crow et al., 2017) et le débit d'eau (Laiolo et al., 2016). Également, les estimations des pluies par satellite, qui ont été jugées assez inexactes dans certaines régions du globe, peuvent être corrigées (Brocca et al., 2014 ; Pellarin et al., 2017).

D'autres applications terrestres comme la prévision des inondations et des sécheresses bénéficient aussi de prévisions améliorées grâce à l'assimilation l'humidité de surface du sol. L'humidité du sol sert de substitut à la quantité d'eau en excès et en déficit car elle partitionne la précipitation entre infiltration, ruissellement et évapotranspiration (Seneviratne et al., 2010). La prévision des inondations peut être nettement améliorée en corrigeant les précipitations assimilées dans les modèles de type précipitation-ruissellement (Massari et al., 2014). Cela a été aussi observé dans des systèmes opérationnels comme le European Flood Awareness System (EFAS), qui a amélioré ses prévisions d'inondation après l'assimilation de l'humidité du sol dans son modèle hydrologique interne (Wanders et al., 2014b).

En ce qui concerne les sécheresses, l'humidité de surface du sol peut aider à anticiper les sécheresses agricoles (Velpuri et al., 2016 ; Bolten et al., 2016 ; Martínez-Fernández et al., 2016). Alors que la sécheresse météorologique répond à un simple manque de précipitations, la sécheresse agricole apparaît lorsque la production végétale est touchée en raison de l'insuffisance des ressources en eau pour les plantes (Mishra et Singh, 2010). Différents indices ont été utilisés pour prédire les sécheresses agricoles. Torres et al. (2013) a basé ses prévisions de la sécheresse sur des séries historiques du "soil water deficit" (SWD), un indice uniquement dépendant de l'humidité du sol et de la profondeur du sol. Le U.S. Drought Monitor (Svoboda et al., 2002) fournit des cartes hebdomadaires sur la base de 5 niveaux de gravité de sécheresse calculés à partir d'un certain nombre d'indicateurs et de variables combinés ensemble, dont l'humidité du sol. (Al Bitar et al., 2013) a proposé de prédire la sécheresse à partir de l'humidité du sol en zone racinaire, calculée avec un modèle de type "double bucket" qui assimile l'humidité de surface du sol et l'indice de surface foliaire.

Diverses applications agricoles, comme par exemple la prévision de la production végétale, découlent directement de la recherche sur la relation entre l'humidité du sol et les sécheresses agricoles. Par exemple, l'humidité du sol en zone racinaire dérivée des données d'humidité de surface du sol ont permis à Gibon et al. (2017) d'expliquer 89 % du rendement des cultures au Niger, 72 % au Burkina Faso, 82 % au Mali et 84 % au Sénégal, après assimilation dans un modèle statistique. De même, les observations de l'humidité de surface du sol ont permis à Ines et al. (2013) de mieux prévoir la production du maïs. D'autres applications agricoles incluent la détection de l'irrigation pour la gestion des ressources en eau. Par exemple, Singh et al. (2016) ont utilisé des données d'humidité du sol du satellite AMSR-E pour détecter un changement dans les pratiques d'irrigation en Inde entre 2002 et 2011.

1.2 Mesurer l'humidité du sol

Plusieurs techniques fournissent des estimations d'humidité de surface du sol. Les **mesures directes** de l'humidité du sol ne sont possibles qu'avec des méthodes destructrices comme les mesures gravimétriques. Elles consistent à évaluer en laboratoire la quantité d'eau évaporée d'un volume de sol préalablement extrait de la zone d'étude. Les **méthodes directes** sont manuelles et donc peu compatibles avec la génération de longues séries temporelles. Les méthodes indirectes mesurent une variable physique fortement liée à l'humidité du sol (par exemple l'énergie émise par la surface, la capacitance du sol, le temps du parcours d'un signal dans le sol) qui est traduit en humidité du sol. Ces capteurs fonctionnent à différentes échelles spatiales.

Le terme *échelle spatiale* réfère à un triplet d'entités, qui est représenté à la Figure 1.3 : **l'étendue**, la zone entourant toutes les mesures ; **l'espacement**, la distance entre les mesures ; et **le support**, la zone réellement détectée par l'instrument (Blöschl et Sivapalan, 1995). Les systèmes de mesure de l'humidité du sol peuvent être classés en fonction de ce triplet d'échelle. Dans cette thèse, deux types de systèmes d'observation seront distingués : ceux dont le support est *grand* ($\sim 40^2$ km^2) et ceux dont le support est *petit* ($<1^2$ m^2).

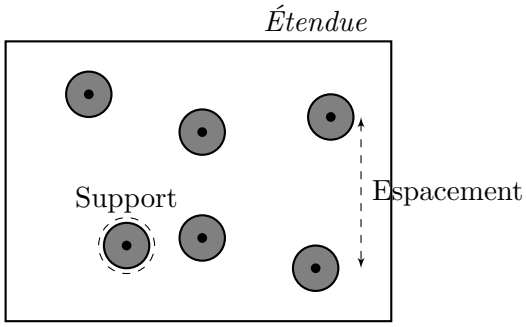


Figure 1.3 – Schéma du triplet d'échelle : étendue, support et espacement

1.2.1 Observations ayant un grand support : les capteurs satellitaires

Les instruments de télédétection peuvent être placés à proximité du sol (sur des tours ou des endroits surélevés), sur des plates-formes aéroportées ou à bord de satellites. Dans les deux premiers cas, le support varie de quelques mètres à ∼1 km, mais les étendues observées sont limitées à l'échelle du bassin versant (10^2 km^2) au mieux. Dans le cas des capteurs satellitaires, la couverture mondiale est souvent réalisée en quelques jours et la surface est échantillonnée de façon complète, homogène et régulière.

En premier lieu, le support des observations satellitaires dépend du type d'instrument. L'humidité de surface du sol peut être obtenue au moyen de capteurs optiques, auquel cas le support est défini par le pixel optique et varie entre quelques mètres à quelques kilomètres. Cependant, il n'existe pas de produits optiques opérationnels car ils présentent des inconvénients majeurs : ils sont sensibles à la surface du sol où l'humidité du sol est contrôlé par des phénomènes surfaciques tels que le vent et l'ombre, pour citer certains. En outre, les observations optiques ne peuvent être acquises en conditions nuageuses et nocturnes et sont très sensibles à l'atmosphère terrestre et à différents éléments de surface (végétation, rugosité, etc.) (Petropoulos et al., 2015).

Dans le cas des **capteurs à micro-ondes**, le support équivaut à la résolution de l'instrument qui est défini par l'empreinte de l'antenne au sol comme la zone représentant la moitié de la puissance totale du signal reçu. Dans cette thèse, les mots **support**, **résolution** et **empreinte satellitaire** seront utilisés indistinctement pour se référer à la taille du support des observations micro-ondes. Celui-ci varie de quelques mètres à des centaines de kilomètres, en fonction de la longueur d'onde d'observation ainsi que d'autres facteurs (distance à la surface de la terre, les caractéristiques du capteur, l'inclinaison, etc.). Bien qu'ils ne souffrent pas des mêmes inconvénients que les capteurs optiques, les capteurs micro-ondes peuvent encore être sensibles aux éléments de surface comme la végétation et la rugosité de la surface : en fonction des caractéristiques du capteur (actif ou passif, configuration de l'antenne, longueur d'onde, etc.) ces effets perturbateurs peuvent être plus ou moins importants. Les capteurs actifs transmettent un signal à la Terre qu'ils analysent après réflexion, réfraction et dispersion par les milieux traversés ou rencontrés alors que les capteurs passifs mesurent simplement leurs émissions naturelles. Les capteurs actifs sont plus sensibles à la diffusion de surface ou de volume que les capteurs passifs. En revanche, ils peuvent fournir des observations jusqu'à des résolutions décamétriques (au dépend de la couverture temporelle le plus souvent). Les capteurs actifs comprennent les radars à synthèse d'ouverture (ROS ou SAR pour son sigle en anglais) (ERS, ALOS, Sentinel 1) et des diffusiomètres (ASCAT). Les **instruments passifs** ou **radiomètres** possèdent des résolutions plus grossières à fréquence égale (de quelques dizaines de kilomètres) mais ils sont moins sensibles aux effets de structure. Les radiomètres en bande C et X comme AMSR-E / 2 et WindSat (Wagner et al., 2007 ; Mladenova et al., 2011 ; Parinussa et al., 2012) ainsi que les radiomètres en bande L comme SMOS et SMAP (Al Bitar et al., 2012 ; Kerr et al., 2016 ; Colliander et al., 2017) fournissent des estimations de qualité de l'humidité de surface du sol.

Cette thèse est consacrée à l'étude de la comparaison entre l'humidité du sol obtenue par des radiomètres satellitaires, comme AMSR-E, AMSR2, SMOS ou SMAP, par rapport aux mesures *in situ* qui possèdent un support considéré ponctuel. Bien que les radiomètres satellitaire présentent des différences importantes en termes de technologie des capteurs et de modèles d'inversion, ils ont tous des résolutions similaires (27^2-55^2 km^2) et un schéma similaire pour estimer l'humidité du sol à partir du signal observé. Leurs caractéristiques principales sont résumées dans le Tableau 1.1. L'énergie mesurée par les radiomètres à une fréquence spécifique ou température de brillance (TB) est transformée en humidité du sol par l'inversion d'une série de modèles qui dépendent des caractéristiques de la surface (couverture terrestre, rugosité, texture du sol, température). Un schéma général d'un système d'inversion typique est présenté par la Figure 1.4. Par la suite, toute référence aux observations et aux supports satellitaires concernera ce type de capteurs.

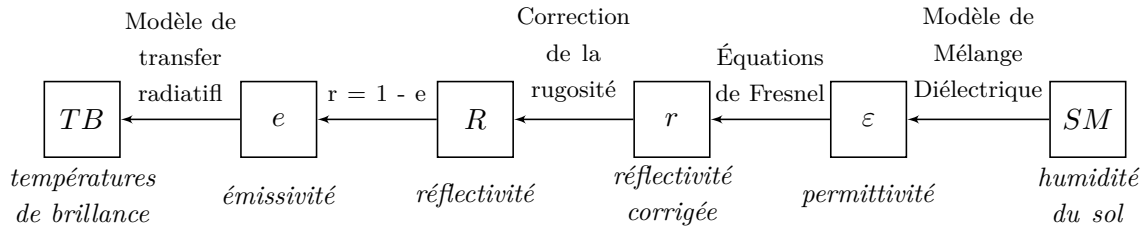


Figure 1.4 – Schéma de la transformation de la température de brillance (TB) en humidité du sol dans les systèmes utilisant des radiomètres satellitaires. Les flèches indiquent la direction de la modélisation.

1.2.2 Mesures ponctuelles : les capteurs *in situ*

Les mesures *in situ* typiques ont un support de quelques centimètres seulement. Un certain type, les mesures des neutrons du rayonnement cosmique (Zreda et al., 2012, capteurs COSMOS) permettent d'obtenir des résolutions de l'ordre de 200-300 m. Dans les deux cas, l'écart de résolution par rapport aux mesures satellitaires est similaire, donc toutes les mesures *in situ* sont assimilées à des observations à *support ponctuel*. Par la suite, les termes **support ponctuel**, **petit** et **local** seront utilisés indistinctement pour se référer à ce type de mesures. En pratique, les mesures *in situ* peuvent représenter une zone plus grande (zone de représentativité) parce que les facteurs générant la variabilité de l'humidité du sol (végétation, texture du sol, topographie, précipitations) sont spatialement liés. La zone de représentativité peut être considérée comme le support *effectif* de la mesure *in situ*. L'estimation de la zone de représentativité n'est pas simple et dépend fortement de la méthode suivis et des critères utilisés pour l'estimer. Dans tous les cas, ce support efficace ne peut pas représenter l'empreinte satellitaire complète car les facteurs liés à l'humidité du sol ne sont jamais parfaitement homogènes dans l'empreinte.

Les stations de mesures *in situ* de type permanent consistent en des sondes installées en permanence dans le sol. D'ordinaire elles font partie d'un réseau, qui est un groupe de stations contrôlées par la même organisation et conçu pour surveiller l'humidité du sol sur une zone donnée. Les réseaux varient en termes d'étendue, d'espacement et de nombre de stations. Ils peuvent être divisés en deux types en fonction du nombre de stations disponibles par empreinte satellitaire. Les réseaux avec une seule station par empreinte sont appelés **réseaux de type éparse**. Cela implique que leur espacement de l'ordre ou supérieur à la résolution du satellite. Ils sont aussi désignés comme des réseaux *à grande échelle* car ils couvrent généralement des étendues de plus de 10.000 km^2 (Crow et al., 2012). Les réseaux avec plusieurs stations par empreinte sont appelés des **réseaux de type dense** ou des réseaux *à petite échelle* parce que leurs étendues varient entre 100 km^2 et 10.000 km^2 (Crow et al., 2012). En conséquence, leur espacement est plus petit que l'empreinte satellitaire. Des exemples de réseaux épars incluent le réseau SMOSMANIA en France (Calvet et al., 2007) et le réseau SCAN aux États-Unis (`http://www.wcc.nrcs.usda.gov`). Exemples des réseaux denses sont les bassins versants de l'USDA (United States Department of Agriculture) comme par exemple Little Washita et Walnut Gulch (Cosh et al., 2006, 2008), le réseau Yanco en Australie (Smith et al., 2012) et le réseau Monte Buey en Argentine (Thibeault et al., 2015). Pour une description détaillée des réseaux de mesures *in situ* disponibles et leurs caractéristiques spatiales respectives, le lecteur peut se référer à Crow et al. (2012). Les caractéristiques principales des mesures *in situ* sont résumées dans le Tableau 1.1.

Il existe une grande variété de types de capteurs *in situ* : capteurs de mesures de capacitance, de mesures d'impédance, capteurs de mesure de réflectométrie de domaine temporel (TDR) et fréquentiel (FDR) et les capteurs basés sur la transmissométrie (TDT, FDT) (Robinson et al., 2008). Bien que l'information mesurée soit différente (par exemple, le temps de charge d'un condensateur dans le cas de capteurs capacitifs ou le temps de déplacement d'une impulsion de tension dans le cas des capteurs TDR), ils exploitent tous la différence entre la permittivité relative (ε) du sol sec et de l'eau (Topp, 2003). Fondamentalement, l'information bute mesurée

est traduite en ε, puis ε est traduite en humidité du sol en utilisant un modèle diélectrique ou un étalonnage empirique.

Les TDR / FDR et les TDT / FDT sont considérés comme plus précis que les capteurs d'impédance et de capacitance. Ces derniers sont par contre moins coûteux et moins consommateurs d'énergie (Bircher et al., 2016), ce qui permet un échantillonnage plus dense dans le temps et l'espace. Pour cette raison, ils ont été largement utilisés dans les campagnes de validation par satellite (Albergel et al., 2012), dans la conception de réseaux de mesures *in situ* (Bircher et al., 2012) et dans des études spatio-temporelles l'humidité du sol (Famiglietti et al., 1999).

Tableau 1.1 – Principales caractéristiques des mesures *in situ* et par satellite de la SM.

	Measures terrain (réseaux permanents)	**Capteurs micro-ondes satellitaires (radiomètres)**
Support	Ponctuel	27^2-55^2 km^2
Étendue	Réseaux denses : $< 50^2$ km^2	Global
	Réseaux éparses : $>> 50^2$ km^2	
Espacement	Réseaux denses : 1-10 km	Égal au support
	Réseaux éparses : $>$ 50 km	
Échantillonnage spatial	Irrégulier	Régulier
	Surface pas complétement échantillonnée	Surface complétement échantillonnée
Échantillonnage temporel	Plusieurs fois par jour, régulier	1-4 jours, régulier ou pas
Couverture temporelle	$>$ 1 an (en fonction du réseau)	Depuis l'année 2000 : AMSR-E (2002-), WindSat (2013-), SMOS (2009-), AMSR2 (2012-), SMAP (2015-)
Avantages	Haute résolution temporelle	Couverture globale
Désavantages	Limitées géographiquement	Plus basse résolution temporelle
	Entretien cher	Plus basse qualité dans certaines conditions
	Défaillances des capteurs courantes	

1.3 Bilan d'erreur et d'écart d'échelle spatiale

Les capteurs terrain et satellitaires fournissent des estimations de l'humidité du sol avec des dimensions de support spatial et des densités d'échantillonnage très différentes. Les capteurs terrain donnent accès à l'état de la variable humidité du sol et les processus associés à des points particuliers dans l'espace, tandis que les capteurs satellitaires donnent accès à une vue intégrée. Cette différence de perspectives d'observation est appelée la différence d'échelle spatiale entre les mesures terrain et par satellite d'humidité du sol. C'est le problème central étudié dans cette thèse.

Malgré l'incompatibilité de leurs échelles spatiales respectives, la précision des observations satellitaires est souvent évaluée par comparaison directe avec les mesures *in situ*. Par conséquent, la comparaison ne contient pas seulement les erreurs de satellites, mais aussi les erreurs des mesures *in situ* et la différence d'échelle spatiale. Cette section décrit le bilan d'erreur des comparaisons satellite-insitu, avec une attention particulière portée à la différence de l'échelle spatiale.

1.3.1 Qu'est-ce que l'humidité du sol à une résolution de 40 km ?

Afin de définir la disparité entre les échelles spatiales, les définitions de l'humidité du sol observée à l'échelle locale et à l'échelle de l'empreinte satellitaire doivent d'abord être posées. La définition

de l'humidité du sol mesurée à un point dans l'espace semble simple et il y a un consensus pour considérer que les mesures ponctuelles effectuées au moyen de méthodes indirectes et directes devraient être identiques, indépendamment des erreurs d'instrumentation et de modélisation. La définition de l'humidité du sol à l'échelle du support satellitaire (grand support) n'est cependant pas claire. Il existe deux interprétations principales pour l'humidité du sol à grand support :

Définition I : y est la moyenne spatiale de X, c'est-à-dire les x_i mesurés à tous les points infinitésimaux i dans le support A. Si $f(X)$ est la fonction de densité de probabilité (PDF) de X, alors l'humidité du sol à grand support est exprimée comme

$$y_I = \mu_X = \int_A x_i f(x_i) dx_i \tag{1.1}$$

Dans le cas de réseaux de mesures *in situ* de type dense, cette humidité du sol est approximée en moyennent linéairement les valeurs enregistrées par les stations de mesures *in situ* à chaque pas de temps (également appelée "network average", NAvg).

Définition II : y est la valeur l'humidité du sol qui correspond au niveau d'énergie mesurée (TB) intégrée sur l'empreinte satellitaire. Le système d'inversion (Figure 1.4) peut être résumé comme une fonction globale $G(\mathring{u})$ qui dérive l'humidité du sol des TB_i et un ensemble de paramètres $P_1, P_2, ..., P_K$. Soit $f(TB_i, P_1, ..., P_K)$ la fonction de densité de la variable TB_i et le jeu de paramètres, l'humidité du sol à grand support peut être exprimée comme :

$$y_{II} = \int_A \int_{P_1} \cdots \int_{P_K} G(TB_i, P_1, \ldots, P_K) f(TB_i, P_1, \ldots, P_K) dTB_i dP_1 \ldots dP_K \tag{1.2}$$

Cette équation représente essentiellement la façon dont les radiomètres perçoivent l'humidité du sol. L'humidité du sol à grand support n'est pas l'humidité du sol moyenne mais la traduction de la TB moyenne en humidité du sol. Comme expliqué plus loin, même si les systèmes, les paramètres et les modèles étaient exempts d'erreurs, y_I ne serait pas égal à y_{II} en général.

1.3.2 Erreurs et incertitudes des systèmes d'observation

Dans la littérature sur l'humidité du sol, lorsque l'exactitude des systèmes d'observation est traitée, les termes *erreur* et *incertitude* sont habituellement considérés comme synonymes. Cela n'a pas de conséquences importantes en pratique, mais il est important de les distinguer lors de la définition de leurs contributions à la différence d'échelle spatiale.

Le niveau d'**incertitude** est le niveau de méconnaissance sur la manière dont l'état ou la variable observée devrait se comporter. Le manque de connaissance associé au nombre d'observations ou échantillons est appelé *incertitude d'échantillonnage*. D'une manière générale, notre connaissance de l'état vrai de la variable étudiée augmente lorsque le nombre d'échantillons augmente. Par exemple, le point de saturation du sol estimé à partir d'une série d'humidité du sol mesuré pendant une semaine est beaucoup plus incertain que le point de saturation estimé à partir d'une série mesurée pendant une année (à échantillonnage temporel identique). De même, l'incertitude de l'humidité du sol moyenne sur une zone donnée augmente avec la diminution du nombre de points échantillonnés.

Un autre type d'incertitude est l'*incertitude épistémique* qui se réfère à une ignorance fondamentale sur l'état ou le processus réel. L'humidité du sol à l'échelle de l'empreinte satellitaire est soumise à une incertitude épistémique parce qu'il n'y a pas de définition univoque (section 1.3.1). D'autres paramètres et variables peuvent être également concernés par l'incertitude épistémique. Par exemple, la définition de "l'occupation du sol à la résolution de 40 km" n'est pas

claire si la moitié du pixel est urbain et l'autre moitié est une forêt. Cette incertitude épistémique sur les variables à l'échelle de l'empreinte satellitaire implique une *incertitude d'échelle* (scaling uncertainty) car le processus de changement d'échelle du petit au grand support n'est pas connu avec certitude.

Le terme **erreur** ne fait pas allusion au manque de connaissances, mais à la différence entre l'état observé et l'état réel. Par exemple, en supposant que les mesures gravimétriques sont sans erreurs, si les mesures de l'humidité du sol par des sondes de capacitance aux mêmes endroits diffèrent, la différence est l'*erreur de mesure* de la sonde de capacitance. Dans le cas des erreurs, habituellement l'état vrai de la variable ne peut pas être connu, mais la définition de l'état vrai reste connue. Par exemple, la définition de l'énergie émise par la Terre dans une zone de 40 km est connue et peut être analytiquement exprimée, mais elle ne peut être relevée qu'imparfaitement par les systèmes d'observation (antennes) qui ont leurs sources inhérentes d'erreurs (pas d'incertitude). Les variables et les paramètres utilisés dans les modèles d'inversion de l'humidité du sol sont également affectés par des erreurs liés aux méconnaissances exactes des caractéristiques de la surface (classification de l'occupation des sols par exemple). Ceux-ci surviennent lorsque la classe associée à un emplacement ou une zone ne coïncide pas avec la réalité. Enfin, le terme *erreur* peut également être synonyme d'imprécision ou *inexactitude* : un modèle est une simplification de la réalité qui ne pourrait jamais reproduire toute sa complexité. Les inexactitudes dans la représentation de la réalité sont inévitables et entraînent des erreurs dans les sorties de modèle.

La liste des incertitudes et des erreurs des données d'humidité du sol peut être résumée comme suit :

- Erreurs instrumentales et de la mesure : erreurs dans la variable mesurée (TB, capacitance du sol, temps du parcours d'un signal, etc.) à cause de l'action de l'instrument ou de l'action humaine.
- Erreurs dans les paramètres d'entrée décrivant le sol et le couvert terrestre (texture du sol, type de végétation et état, etc.).
- Imprécisions des modèles (approximations, transfert radiatif, modèle diélectrique, etc.).
- L'incertitude d'échelle (dans les systèmes à grand support) : niveau d'ignorance sur la façon dont la variable doit être exprimé à l'échelle du grand support (e.g. l'humidité du sol, le couvert terrestre).
- L'incertitude d'échantillonnage : niveau d'ignorance sur la variable compte tenu du nombre disponible et l'emplacement des échantillons.

1.3.3 Mesures *in situ* et satellite : la différence d'échelle spatiale

La **différence d'échelle spatiale** englobe tous les facteurs contribuant à une inadéquation des échelles spatiales des systèmes d'observation, satellite et *in situ*. Trois éléments contribuent à la différence d'échelle spatiale : l'*incertitude d'échantillonnage* et l'*incertitude d'échelle*, introduites dans la précédente section, et l'*écart géophysique*. L'écart géophysique concerne les différents processus et échelles de processus captés par les deux systèmes d'observation en fonction de leurs supports (empreinte *vs* point). Les éléments de la différence d'échelle spatiale sont résumés dans la Figure 1.5 et décrits dans les paragraphes suivants.

Incertitude d'échantillonnage (*in situ*) : Les mesures ponctuelles *in situ* sont une approximation de l'humidité du sol à grand support tel que défini dans la définition I (section 1.3.1) : elles peuvent être moyennées produisant ce qu'on appelle le network average (NAvg), qui est une estimation de l'humidité du sol à grand support y_I. L'incertitude de la NAvg diminue

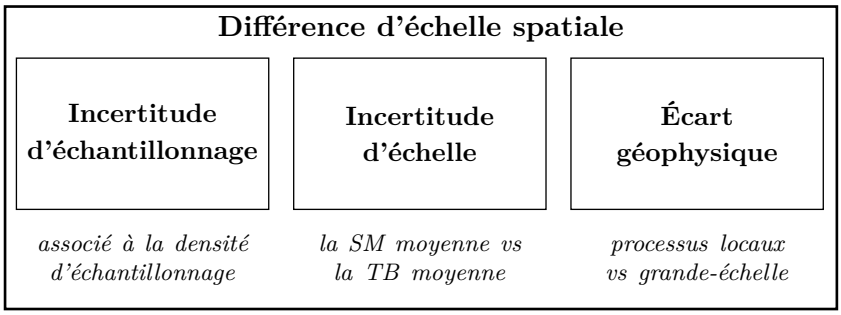


Figure 1.5 – Les 3 composants de la différence d'échelle spatiale.

avec le nombre d'échantillons spatiaux et dépend de l'emplacement des stations. Plusieurs techniques d'échantillonnage ont été utilisées pour estimer cette incertitude, avec plus ou moins de succès : échantillonnage statistique, géostatistique, stratifié et bootstrap. Une revue descriptive peut être trouvée dans (Wang et al., 2008, 2015a).

Incertitude d'échelle : Lorsque les mesures *in situ* et satellite sont comparées, l'incertitude d'échelle est constituée de deux composantes :

(i) l'incertitude épistémique qui affecte la définition de l'humidité du sol à grand support (y_I ou y_{II})
(ii) les différences entre l'humidité du sol (vraie) à grand support obtenue en moyennant les humidités ponctuelles (y_I) et celle dérivée des observations par satellite (y_{II}).

Le premier composant a déjà été expliqué dans la section précédente. Le second fait référence au fait que y_I et y_{II} sont fondamentalement différents en raison de la non-linéarité du processus d'obtention de l'humidité du sol à partir de l'énergie observée. Cette différence est typique des systèmes non linéaires où la moyenne des sorties n'est pas égal à la sortie système des entrées moyennées. Dans le cas des systèmes satellitaires, il existe deux sources principales de non-linéarités. D'une part, la fonction reliant TB et humidité du sol n'est pas linéaire (voir la Figure 1.6 pour le lien non linéaire entre la constante diélectrique ε et l'humidité du sol, par exemple). Par ailleurs, les paramètres du modèle, comme les caractéristiques du sol et la couverture du sol, sont fondamentalement hétérogènes au sein de l'empreinte satellitaire et chaque type de sol et de couverture donne lieu à des relations différentes entre ε et l'humidité du sol. De nombreuses études ont confirmé que cette incertitude d'échelle est non négligeable dans le cas des radiomètres satellitaires (Crosson et al., 2010), tels que AMSR-E (Crow et al., 2001), SMOS (Pellarin et al., 2003), Hydros (Crow et al., 2005) et SMAP (Zhan et al., 2008).

Écart géophysique : L'écart géophysique est lié au support d'observation : tandis que le satellite intègre spatialement les processus dans son empreinte, la sonde *in situ* observe un point à l'intérieur. En d'autres termes, ils observent la même variable sous différentes perspectives. Pour cette raison, ils captent des processus différents ou des échelles différentes des mêmes processus. Par exemple, lorsque des précipitations ont lieu au sein de l'empreinte, l'humidité du sol observée à un endroit spécifique est probablement différente de la réponse en humidité moyenne de l'empreinte. C'est facile à comprendre par analogie au zoom avant et arrière d'une image. La Figure 1.7 en est une illustration : si la référence d'observation est définie à l'échelle locale, la valeur observée peut être “bleu”, “jaune” ou “vert” ; quand la référence d'observation est fixée à grande échelle, la valeur observée est une intégrale de ces petites valeurs d'échelle, "vert" dans ce cas.

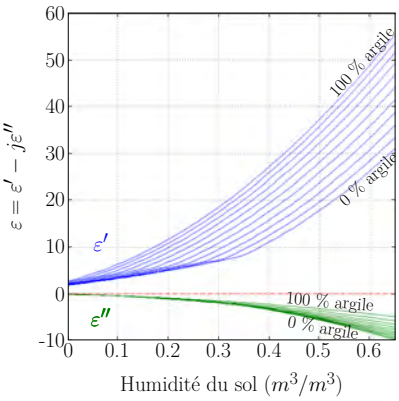


Figure 1.6 – Relation entre la constante diélectrique du sol et l'humidité du sol selon le modèle de Mironov (Mironov et Fomin, 2009 ; Mironov et al., 2009) utilisé dans le processeur SMOS d'inversion de l'humidité du sol (Kerr et al., 2014) pour différentes fractions d'argile.

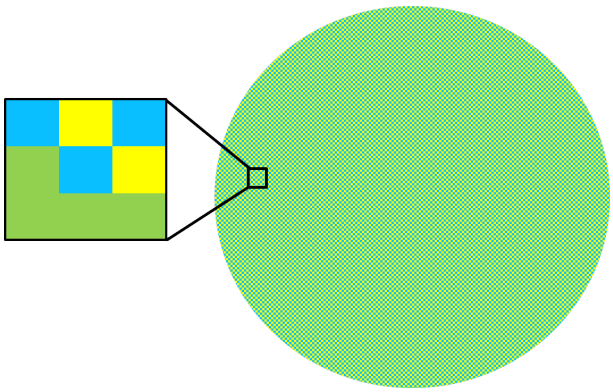


Figure 1.7 – Représentation conceptuelle de l'écart géophysique.

1.3.4 Terminologie

Cette section résume la terminologie utilisée dans ce manuscrit pour les différentes échelles spatiales. Concernant le support, deux grandes échelles sont distinguées (Tableau 1.2) : le support ponctuel (mesures *in situ*) et le grand support (estimations par satellite). Un troisième support, le petit support, est ajouté au jeu de données pour nommer l'humidité du sol à haute résolution provenant des modèles. Notez que cette nomenclature s'applique aux supports physiques, pas au support *effectif* ou zone de représentativité des stations de mesures *in situ*. En ce qui concerne l'étendue, les termes les plus couramment utilisés figurent dans le Tableau 1.3.

Tableau 1.2 – Supports spatiaux principaux.

Supports spatiaux	
local/point	$\lesssim 1^2$ m^2
petite/haute-résolution	$\lesssim 1^2$ km^2
grand/empreinte/petite-résolution	$\sim 40^2$-50^2 km^2

Tableau 1.3 – Étendues spatiales principales.

Étendues spatiales	
parcelle	$\lesssim 100^2$ m^2
flanc d'une colline	$\sim 100^2$ m^2
bassin versant	$\sim 10^2$ km^2
empreinte satellitaire	$\sim 40^2$-50^2 km^2
régional	$\sim 1.000^2$ km^2

1.4 Questions et objectifs de cette thèse

La précision des estimations d'humidité du sol par satellite est généralement évaluée par comparaison directe à des mesures *in situ*. Compte tenu de la couverture spatiale limitée des réseaux de mesures *in situ*, la plupart des campagnes de validation consistent à confronter les deux types de mesures dans le temps, c'est-à-dire en confrontant une série temporelle du satellite avec une série temporelle de mesures *in situ*, les deux prises à un endroit ou une région particulière. Les statistiques dérivées (corrélation, biais, RMSE) sont censés représenter les performances du capteur satellitaire. Néanmoins, elles sont également affectés par les erreurs de la sonde terrain et la différence spatiale entre le satellite et les observations *in situ*. Si l'on s'attend à ce que les erreurs *in situ* soient faibles (<0,05 m^3/m^3) sous réserve que le réseau *in situ* soit correctement entretenu, **la contribution de la différence de l'échelle spatiale est supposée importante bien qu'elle reste indéterminée**. En conséquence, les erreurs inhérentes aux systèmes d'observation restent également indéterminées.

La Figure 1.8 fournit une illustration de la différence d'échelle spatiale. Elle montre la comparaison entre une série temporelle SMOS et les séries temporelles de mesures *in situ* prises à 20 endroits différents dans la même empreinte satellitaire. Les statistiques sont présentées sous la forme d'un graphique de Taylor (Taylor, 2001) qui permet de résumer tous les résultats de validation pour une même référence. Le biais est représenté en code couleur, la corrélation sous forme angulaire sur le cercle extérieur et l'écart-type des différences (STDD) dans le cercle vert intérieur. La dispersion dans les marqueurs suggère que l'ampleur de la différence d'échelle spatiale dans cette zone est non négligeable. Par exemple, la STDD varie de 0,045 à 0,08 m^3/m^3, selon l'emplacement choisi pour la comparaison. Cependant, on ne sait pas quel est le pourcentage de la STDD est attribué à la différence d'échelle spatiale et ceux provenant erreurs des mesures *in situ* et satellitaires.

La différence d'échelle spatiale n'est pas seulement un problème spatial mais aussi un problème temporel. Le champ spatial d'humidité du sol n'est pas stationnaire dans le temps et sa redistribution spatiale varie dans le temps. Ceci veut dire que la distance en termes d'humidité du sol entre un point particulier et le grand support (inconnu) varie dans le temps. Par conséquent, la différence d'échelle spatiale nécessitera des approches spatio-temporelles pour être étudiée.

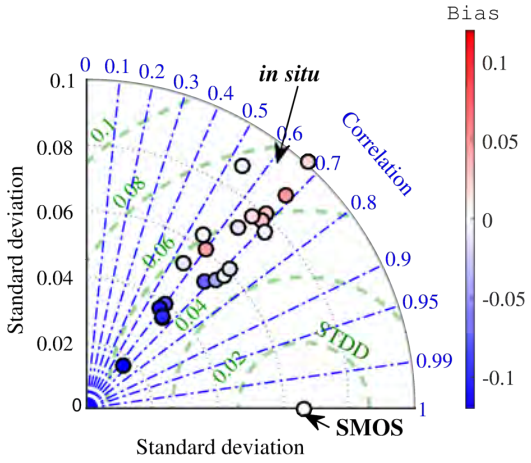


Figure 1.8 – Comparaison d'une série temporelle d'humidité du sol du satellite SMOS avec 20 séries temporelles d'humidité du sol *in situ* mesurées au sein de la même empreinte satellitaire.

La différence d'échelle spatiale est un problème récurrent dans la validation des produits satellitaire d'humidité du sol. Il est présent depuis le lancement à la fin des années 70 du Scanning Multichannel Microwave Radiomete (SMMR), le premier qui ait été utilisé pour estimer l'humidité du sol depuis l'espace. Depuis lors, différentes approches ont été utilisées pour réduire ou pour évaluer la différence d'échelle spatiale. Par exemple, des méthodes de réduction d'échelle (*downscaling*) ont été conçues pour améliorer la résolution des observations d'humidité du sol par satellite à l'aide de données auxiliaires à plus haute résolution. Des approches d'analyse statistique comme la *triple collocation* (TC) ont également été adaptées pour évaluer la différence d'échelle spatiale entre un emplacement et l'empreinte. Les inconvénients des méthodes proposées jusqu'à maintenant est qu'elles sont spécifiques à une région, à un ensemble de conditions paysagères et climatiques et / ou qu'elles dépendent d'un ensemble d'hypothèses statistiques qui ne sont pas toujours respectées et difficiles à vérifier. En conséquence, **il n'y a pas d'approche standardisée pour aborder la différence d'échelle spatiale** et, à titre de preuve, les observations par satellite sont toujours validées opérationnellement sans en prendre compte.

Des **questions** différentes se posent lorsque l'on aborde le problème de l'inadéquation spatiale :

- quelle est la bonne définition de l'humidité du sol à la résolution du satellite ?
- qu'est-ce qu'il y a exactement derrière la différence d'échelle spatiale ?
- la différence d'échelle spatiale est-elle mesurable ?
- est-il possible de quantifier son impact dans les métriques de validation classiques ?

Cette thèse vise à **contribuer à la résolution du problème de différence d'échelle spatiale** et les questions associées. Le but est d'**améliorer les connaissances actuelles** sur la différence d'échelle spatiale et de fournir **de nouvelles méthodes ou d'améliorer celles existantes** pour son évaluation. Les approches proposées doivent représenter une amélioration par rapport aux approches existantes en termes d'application générale : elles devraient être **faci-**

lement applicables à toutes les régions de validation. Pour cela, le nombre d'hypothèses géophysiques et statistiques devrait être maintenu aussi bas que possible. Cela permettrait de préparer la voie à une évaluation intégrée de la différence d'échelle spatiale dans les campagnes de validation des produits satellitaires d'humidité du sol.

1.4.1 Structure du manuscrit

Ce manuscrit est structuré en 6 chapitres, y compris l'introduction et la conclusion. Dans ce premier chapitre d'introduction, j'ai décrit la problématique de recherche et les questions associées. Le problème de la différence d'échelle spatiale dans l'observation de la variable humidité de surface du sol sera le fil conducteur pour le reste du texte.

Les chapitres 2 et 3 fournissent une base solide dans la variabilité spatio-temporelle de l'humidité du sol et présentent la plupart des ensembles de données et des techniques qui seront utilisés dans les chapitres suivants. Le **chapitre 2** aborde l'étude de la variabilité spatiale de l'humidité du sol au sein d'une empreinte satellitaire typique. Il décrit l'état de l'art concernant le comportement statistique et les éléments géophysiques contrôlant les champs d'humidité du sol. La méthode de désagrégation DISPATCH est évaluée, permettant d'illustrer les potentiels et les inconvénients de ce type d'approches pour réduire la différence d'échelle spatiale. Le **chapitre 3** est dédié à la variabilité temporelle de l'humidité du sol. Les échelles de temps de plusieurs jeux de données d'humidité du sol sont analysés à l'aide de transformées en ondelettes, un outil de traitement de signal marginalement utilisé à ce jour dans la littérature sur l'humidité du sol.

Cette thèse postule que la différence d'échelle spatiale peut être évaluée à partir du domaine temporel car les échelles spatiales et temporelles de l'humidité du sol sont connectées. Le **chapitre 4** est dédié à la description et l'évaluation de cette hypothèse fondamentale. À cette fin, il fournit de nouvelles méthodes basées sur l'analyse par ondelettes pour évaluer la différence d'échelle spatiale sur une base d'échelle de temps. Les expériences sont développées sur des séries temporelles d'humidité du sol modélisées et mesurées sur des régions de validation de produits satellitaires.

Bien que le chapitre 4 traite de la différence d'échelle spatiale dans son ensemble, le **chapitre 5** se concentre sur l'incertitude d'échantillonnage. Il est dédié à la transformation de l'incertitude d'échantillonnage *in situ* (spatial) en incertitude dans les statistiques de validation des données satellitaires (temporelles). En d'autres termes, l'objectif de ce chapitre est de fournir des intervalles de confiance pour les métriques temporelles utilisées dans les campagnes de validation des produits satellitaires (corrélation, biais, RMSE), en se basant sur l'incertitude d'échantillonnage qui devra être estimée au préalable. Les concepts du chapitre 2 seront utilisés à ce niveau.

Enfin, le dernier chapitre est consacré aux conclusions générales de la thèse ainsi qu'aux perspectives de recherche qui en découlent.

Chapter 2

Spatial variability of surface soil moisture (SM)

Contents

2.1 Sources of soil moisture variability 30
2.2 Spatial variability metrics 31
2.2.1 Variance 31
2.2.2 Coefficient of Variation (CV) 32
2.3 Observed spatial variability 33
2.3.1 Probability distribution 33
2.3.2 Dependence with the mean 33
2.3.3 The coefficient of variation (CV) and the mean 35
2.3.4 Dependence with the support and extent 36
2.4 DISPATCH: a spatial disaggregation approach 37
2.4.1 Background on optical-based downscaling 37
2.4.2 Datasets required by DISPATCH 39
2.4.2.1 Large-support datasets (SMOS) 39
2.4.2.2 Higher-resolution (HR) datasets 39
2.4.3 DISPATCH algorithm 40
2.4.4 Validating DISPATCH products 41
2.4.4.1 Validation datasets: dense *in situ* networks 41
2.4.4.2 Data preparation 42
2.4.4.3 Evaluation metrics 42
2.4.5 Results and discussion 43
2.4.5.1 Preliminary analysis 43
2.4.5.2 Visual inspection of disaggregated maps 44
2.4.5.3 Spatial evaluation 46
2.4.5.4 Temporal evaluation 48
2.4.5.5 Discussion 49
2.5 Conclusions 49

The spatial scale mismatch problem is intrinsically related to the degree of spatial variability. Intuitively, the point-support measurements should be equal to the footprint measurement if the surface is uniform and the observation systems are free of non-linearities, errors and uncertainties (instrument, paratemers, models, etc.). Heterogeneous surfaces would however induce differences between the point measurements and the footprint measurement. Based on former studies, Western et al. (2002) reached to a similar conclusion and stated that significant scale effects were associated with the spatial variability in soil moisture (SM) and related variables.

Knowing the statistical and physical distribution of SM as well as its geophysical drivers within the satellite footprint is also important for many other applications: determining the number and location of *in situ* stations, building downscaling models for improving the resolution of satellite SM products, designing land surface models, etc..

This chapter provides the essential notions of SM spatial variability, from both statistical and geophysical perspectives. These notions are fundamental for the best understanding of chapters 4 and 5, where the spatial scale mismatch between SM measurements will be studied from a spatio-temporal perspective. This chapter is structured as follows: Sections 2.1 to 2.3 describe the main aspects of the SM spatial variability, its geophysical drivers, its evaluation metrics and its statistical distribution. All this constitute the basic background on the SM spatial distribution. Section 2.4 will be dedicated to the analysis and validation of a semi-physical disaggregation approach, DISPATCH.

The validation of DISPATCH is important for three reasons: First, it illustrates that the knowledge about the sub-footprint variability can be used to build downscaling/disaggregation approaches that reduce the spatial scale mismatch. However, disaggregation methods are usually limited by specific climatic and landscape conditions. Secondly, DISPATCH SM products are proxies for understanding the spatial redistribution of SM. For this reason, DISPATCH will be used again in chapters 3 and 4. Finally, some of the *in situ* networks used for validating DISPATCH (Yanco, Little Washita and Walnut Gulch) will be also used in the temporal and spatio-temporal analyses of the subsequent chapters.

2.1 Sources of soil moisture variability

Surface SM is a function of the precipitation frequency and intensity, the infiltration rate and the run-off rate (Hawley et al., 1983). Infiltration and run-off rates are directly controlled by the antecedent SM conditions and three main landscape properties: soil, vegetation and topography (Crow et al., 2012). The dependence between precipitation and the surface layer is complex so, for simplicity, it is considered as an external independent forcing of SM when analysing SM drivers. Irrigation is also an external independent forcing.

The relative influence of these four components, soil, vegetation, topography and precipitation, on the total SM variability is different depending on the area under study and its climatic conditions. Knowing the most predominant ones can help in the design of *in situ* networks and the validation of satellite SM estimates. For example, Bircher et al. (2012) designed a SM *in situ* network in Western Denmark based on land cover and texture, which they find to be the principal variability drivers in the region. Additionally, a good understanding of the variability drivers and their related water and energy fluxes can help in the development of downscaling approaches. This will be illustrated in section 2.4 with the DISPATCH downscaling method.

Soil heterogeneity concerns a list of elements: texture, porosity, structure, rock bed depth and organic content. Porous soils, granular structures and coarse textures like sand drain faster, while less porous soils, massive structures and finer textures like silt and organic matter resist more to drainage (Schoonover and Crim, 2015). Differences in soil properties are typical at small and medium spatial scales ($\leq$ 40 km) (Hawley et al., 1983; Famiglietti et al., 1998). Soil properties can be one of the most important drivers of SM spatial variability in some areas.

For example, in the sites analysed by Choi et al. (2007), whose extents ranged from the field to the basin scales, they showed that most of the SM spatial variance could be explained by soil heterogeneity (porosity, wilting point and field capacity).

Similarly to soil characteristics, **vegetation** also plays a important role in SM heterogeneity within the ∼40 km footprint (Mohanty and Skaggs, 2001; Qiu et al., 2001). This is because vegetation directly affects water and energy balance processes that control SM levels. It intercepts a fraction of the precipitation (canopy), limits soil evaporation and consumes a part of the soil available water (vegetation intake) (Seyfried and Wilcox, 1995).

Concerning **topography**, it affects both the vertical and the lateral redistribution of SM through three elements: slope, aspect and location on the slope (relative elevation) (Hawley et al., 1983). Steep slopes redistribute the water to the bottom of the slope, reducing infiltration at the top. Slope orientation and aspect (Malbéteau et al., 2017) as well as elevation (Merlin et al., 2013) influences the evapotranspiration and thus, soil moisture. Kornelsen and Coulibaly (2013) found that SM variability was mainly controlled by topography at two meadow hillslope scale sites during winter and spring. Topography can also explain SM variability specially when soils are drying (Choi et al., 2007). The structure of soil layers can also influence the redistribution of SM.

Finally, **precipitation** is the main meteorological forcing of SM (Crow et al., 2012). Variations in rainfall occur at different spatial scales. Storm cells can affect just some hundreds of meters (Goodrich et al., 1995), while precipitation fronts can sweep some hundreds of kilometres (Entin et al., 2000; Skøien et al., 2003). Rosenbaum et al. (2012) identified rainfall as one of the main drivers of SM variability at the small catchment scale (< 1 km^2). At the regional scale ($> 10{,}000$ km^2) Jackson et al. (1999) explicitly showed the pattern produced by rainfall on a SM field during the Southern Great Plains 1997 (SGP97) campaign. At similar spatial scales, Choi et al. (2007) identified rainfall as the second most important source of SM variability after soil heterogeneity.

2.2 Spatial variability metrics

Different metrics are used to evaluate SM spatial variability. The most used ones in the SM literature are related to the dispersion of the spatial PDF of SM. The dispersion refers to the scatter of the data around the most common or central values of the distribution. Some examples of metrics include the standard deviation (σ) (and its squared form, the variance, σ^2), the distribution percentiles and the coefficient of variation (CV). This section provides an outlook of these metrics, their advantages and limitations.

2.2.1 Variance

The sample variance is defined as

$$\hat{\sigma}^2 \triangleq \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \hat{\mu})^2 \tag{2.1}$$

where N is the sample size and $\hat{\mu}$ is the sample mean value:

$$\hat{\mu} \triangleq \frac{1}{N}\sum_{i=1}^{N} x_i^2 \tag{2.2}$$

The $N-1$ denominator in Equation 2.1 ensures that the sample variance is an unbiased estimator of the population variance σ^2, given that the samples x_i are independent (Manolakis et al., 2005, p. 139-140).

The sample variance (or the standard deviation) is usually used as measure of dispersion of SM datasets, for both time series and spatial fields. In the case of normal distributions, the intervals $[\mu - \sigma, \mu + \sigma]$ and $[\mu - 2\sigma, \mu + 2\sigma]$ contain approximately 68 % and 95 % of the samples. As section 2.3.4 will describe, SM spatial distributions are not always normal, so the comparison of σ^2 values of different distributions can lead to erroneous interpretations. Instead, distribution percentiles could be used to provide unambiguous estimates of the dispersion. While the $[\mu - 2\sigma, \mu + 2\sigma]$ interval does not necessarily contain 95 % of the samples, the $[2.5th, 97.5th]$ interval of percentiles always contain 95 % of the samples.

2.2.2 Coefficient of Variation (CV)

The coefficient of variation (CV) is defined as

$$CV \triangleq \frac{\sigma}{\mu} \tag{2.3}$$

The advantage of the CV is that it allows to compare the dispersion of variables with different scales and units. It is mainly used for distributions that have only positive values and positive skewness. It can be applied to other types of distributions, but in this case its usefulness as a measure of dispersion is questionable (Isaaks and Srivastava, 1989, p. 21): Negative values move the mean μ towards values equal or smaller than zero where the CV is not defined. In the case of left-skewed distributions, the left tail artificially inflates the CV.

The CV has been extensively applied to study the spatial distribution of SM (Brocca et al., 2007; Famiglietti et al., 2008; Brocca et al., 2012; Mittelbach and Seneviratne, 2012), where it has been used as a normalized measure of spatial variability. The objective was to balance the SM range and give more weight to the dry-end where SM variations were much smaller than in the middle SM range (see section 2.3).

From my point of view, the use of the CV for SM spatial studies is equivocal: First, recent experiments in the spatial domain suggest that the dependence of the variance with the mean follows a convex law (the peak of variance is reached in the midrange, see section 2.3.3). This implies that the CV would penalize the variations ΔSM close to the wet-end with respect to the same variations ΔSM close to the the dry-end. Secondly, as Famiglietti et al. (1999) highlighted, the range of the observed mean is much larger than that of the standard deviation, which implicitly leads the CV to monotonically decrease with the mean. This will be further discussed in section 2.3.3.

I propose to adapt the CV equation to overcome the aforementioned limitations. The coefficient first proposed by Muilwijk (1966) and Moors and Muilwijk (1971) and later formalized by Sharma et al. (2011) was specifically adapted for the case of variables that, like SM, were bounded between a minimum and a maximum value. This **"bounded" coefficient of variation (bCV)** is expressed as

$$bCV = \frac{\sigma}{\sqrt{(x_{max} - \mu)(\mu - x_{min})}} \tag{2.4}$$

where in the case of SM, x_{max} should be replaced by the saturation point and x_{min} by the wilting point. In section 2.3.3, the performances of the CV and the bCV are compared on an example.

2.3 Observed spatial variability

Since the late 70s, several authors have analysed the spatial distribution of SM based on the metrics discussed in the previous section. They have been usually interested in finding the relationship of spatial variability with mean SM content and climatic and environmental conditions. Herein, I review the state of the art and derive conclusions about the observed spatial distribution and variability of SM.

2.3.1 Probability distribution

The PDF of spatial surface soil moisture approaches a normal function when the mean SM is in the mid-range. Bi-modal distributions have also been observed (Rosenbaum et al., 2012) specially when the spatial extent is large (Ryu and Famiglietti, 2005). Given the bounded nature of SM between zero water content and porosity, the PDF becomes skewed and less variable when the mean soil moisture content approaches the dry and wet ends (Famiglietti et al., 1999; Western et al., 2002; Ryu and Famiglietti, 2005; Brocca et al., 2007, 2012; Rosenbaum et al., 2012).

The work of Ryu and Famiglietti (2005) was particularly comprehensive and is one of the few that analyses the SM spatial distribution within the 50 km footprint. They used the 800 m resolution soil moisture data acquired during the SGP97 hydrology experiment (Jackson et al., 1999). The soil moisture images were acquired by the Electronically Scanned Thinned Array Radiometer (ESTAR) instrument, over a 50-km by 250-km area in Oklahoma, U.S.. From this experiment, a total of 154 50-km samples were analysed. They observed bimodal Gaussian PDFs that they explained by heterogeneous precipitation patterns in the area. They confirmed the positive skewness under dry conditions reported before by Famiglietti et al. (1999). However, they could not observe negatively skewed distributions when the mean SM reached the wettest values: the distributions appeared truncated, which could be explained by an insufficient sampling of the wet-edge zone.

2.3.2 Dependence with the mean

Several studies have investigated the relationship between the variance of SM fields and the mean SM content. Three different behaviours have been observed: variance increases with SM content (**increasing trend**), variance decreases with SM content (**decreasing trend**) and variance exhibits a **convex law** with respect to SM content whose maximum is reached in the SM midrange. Table 2.1 summarises what I consider the most representative studies, classified per type of variance-mean relationships.

Most of the former studies focused on small spatial extents (< 1 km^2) and only some few investigated the variability within footprint extents. These are highlighted in gray in Table 2.1. The literature on field and small watershed scales is much more extensive than what is presented here. Supplementary sources of information are the publications of Western et al. (1998), Brocca et al. (2007) and Famiglietti et al. (2008), who reviewed the studies on SM spatial variability until 1996, 2005 and 2004, respectively.

Why different variance-mean laws have been observed? They seem to be dependent on the climatology, the period and instant of analysis and the effective range of SM values observed. Western et al. (2004) observed the decreasing trend in the wetter New Zealand catchments while the increasing trend in the drier Australian catchments. They associated this with more important lateral flow processes in extremely humid conditions. Similarly, based on previous studies, Brocca et al. (2007) found decreasing trends of variance with mean SM in humid climates while increasing trends in semi-arid environments. However, a more comprehensive review, as

Table 2.1 – Summary of previous studies of soil moisture spatial variability classified according to the relationship between variance and mean (left column). "+" denotes increasing variance with increasing mean and "-" denotes decreasing variance with increasing mean. All are based on point-support (*in situ*) data, except for $*_a$ and $*_b$

Variance - mean link	Sensing depth (cm)	Extent (km^2)	Nb sites	Nb samples per site and date	Spacing	Temporal sampling per site	Period	Climate	Reference
+	30	1.05	1	500	10-20 m	13 dates	over 1 year	Sub-humid temperate	(Western et al., 1998)
+ *1	30	< 0.6	3	280-485	10-40 m	6-8 dates	over 1 year	Sub-humid temperate	(Western et al., 2004)
+	0-5	3 x 3	2	∼140	250 m	6 & 5 dates (consecutive)	winter & spring	Semi-arid	(Yee et al., 2016)
+ *a	∼5	36 x 36	1	1296	1 km	9 dates (∼consecutive)	spring	Semi-arid	(Yee et al., 2016)
−	0-6	0.64	6	49	100 m	28 dates (consecutive)	summer	Sub-humid	(Famiglietti et al., 1999)
− *1	30	< 0.6	3	250-570	20-40 m	6-8 dates	over 1 year	Humid	(Western et al., 2004)
−	15	< 0.01	3	45-120	3-10 m	14 dates	spring	Semi-humid	(Brocca et al., 2007)
Convex *b	∼5	50 x 50	10	∼4000	0.8 km	16 dates	1 month in summer	Sub-humid	(Ryu and Famiglietti, 2005)
Convex *2	10	0.04	4	12	∼10 m	Each hour	1 year	-	(De Lannoy et al., 2006)
Convex	0-6	0.8 x 0.8	52	14-27	100-200 m	11-28 dates	summer	Varied	(Famiglietti et al., 2008)
Convex	0-6	50 x 100	3	20-26	∼15 km	11-16 dates	summer	Varied	(Famiglietti et al., 2008)
Convex	0-15	178, 242	2	46	∼2-3 km	35 dates	over 1 year	Semi-humid	(Brocca et al., 2012)
Convex	5	0.27	1	74	60 m	Each hour	1 month of summer	Humid temperate	(Rosenbaum et al., 2012)
Convex	10	Hill-slope	4	9	-	Each day	5.5 years	Humid continental	(Kornelsen and Coulibaly, 2013)

*a The support size is 1 km (airborne PLMR sensor) and the surface is fully sampled
*b The support size is 0.8 km (airborne ESTAR sensor) and the surface is fully sampled
*1 Different variance-mean relationships are observed depending on the sites
*2 Not exactly convex: variance increases with mean but it saturates at high SM values

the one in Table 2.1, does not reveal any clear correlation between climate and variance, so the classification according to climate does not seem conclusive.

Until the 2000s, the variance-mean dependence was classified to describe either an increasing or a decreasing trend. To my knowledge, the first to draw the convex relationship were Ryu and Famiglietti (2005). The convex relationship can be seen in Figure 2.1, which has been extracted from their article for illustrative purposes only. The convex relationship makes sense from a statistical point of view: the SM is a bounded variable so the variability should tend to be smaller at the dry and wet edges. The relationship can be also explained from a physical point of view: the wet-end can only be reached after a rather spatially homogeneous precipitation event. Just after it, the surface is homogeneously wet so the spatial variance is low and the only differences come from heterogeneity in soil porosity point values. Different dry-down decays appear afterwards as a result of local heterogeneities in texture, vegetation and topography. Finally, the dry-end is reached when most of the surface has dried out completely so the variability becomes again small. Considering this, it can be concluded that the most probable reason for the observation of only increasing trends is that samples were not available just after rain events, that the saturation point was not reached or that precipitation heterogeneity was high. On the other hand, negative relationships can appear when some areas cannot completely dry out. This can be produced by a number of elements like topography or too high water table levels.

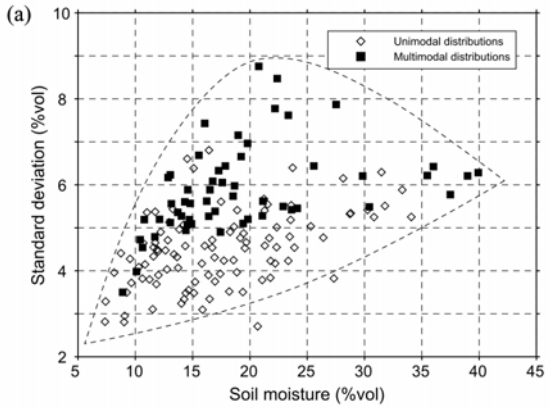


Figure 2.1 – Standard deviation as a function of mean SM content from the for the 154 footprint-support samples of SGP97. Source: Ryu and Famiglietti (2005)

Similar conclusions were derived by Crow et al. (2012), who pointed out that the spatial variability was dependent on the range of SM values observed within the period of analysis. From SGP97, SGP99, SMEX02 and SMEX03 campaign data, they derived a convex law for the variance. The law was applicable at different spatial extents, from the smallest 2.5 x 2.5 m^2 to the largest footprint 50 x 50 km^2 extent. Famiglietti et al. (2008) fitted the same variance model (with different parameter values) for both the field (< 1 km^2) and the footprint (50 x 100 km) scales. These findings are of great importance for this thesis, because they imply that the variability-mean relationship observed at the field scale can be transposed to the footprint scale. The convex relationship appears in most of the (few) studies that concerned footprint extents (in grey in Table 2.1). The only exception is the study of Yee et al. (2016), who found that the variance only increased with the mean SM content.

2.3.3 The coefficient of variation (CV) and the mean

While there is no consensus on the evolution of SM variance with SM content in space, all studies agree on the coefficient of variation (CV) to systematically decrease with increasing SM content. This has been observed at both the field and small watershed scales (Famiglietti et al., 1999; Brocca et al., 2007; Korres et al., 2015) and the footprint scale (Choi et al., 2007; Famiglietti et al., 2008; Brocca et al., 2012; Korres et al., 2015). The relationship follows an exponential law (Brocca et al., 2007; Choi et al., 2007; Korres et al., 2015), whose parameters vary with the extent, support and spacing of the measured samples (Choi et al., 2007).

As already introduced in section 2.2.2, the monotonic decrease of CV with the SM mean seems more likely due to purely statistical reasons: the decrease is more controlled by the increase in SM content than the decrease in SM standard deviation (Famiglietti et al., 1999). In the same section, I have proposed the bounded coefficient of variation (bCV) as an alternative for bounded variables like SM. Herein, I present an example that illustrates the performances of the classical and the bounded CV for SM analyses.

Example: classical *vs* bounded CV. The CV and the bCV are obtained for a range of possible mean SM values $\{\mu_k = 0,\ 0.001,\ 0.002 \ldots 0.449, 0.45\}$, $k = 1, 2 \ldots 1000$. The minimum and maximum SM values are set to 0 and 0.5 m^3/m^3, respectively. The standard deviation σ_k associated to each mean value is derived from μ_k following the exponential formula of Famiglietti et al. (1998), which they found suitable for spatial scales from some meters to 50 km:

$$\sigma = \mu \cdot A \cdot e^{-\mu \cdot B \cdot 100} \qquad [m^3/m^3] \qquad (2.5)$$

with $A = 0.564$ and $B = 0.044$ as in Brocca et al. (2012).

The first row of Figure 2.2 shows the CV and the bCV values (vertical axis) obtained for the same 1,000 synthetic SM spatial mean values μ_k (horizontal axis) and the respective standard deviation values σ_k (top row). It can be observed that the CV-mean relationship greatly differs between the 2 types of CV. While the CV decreases monotonically, the bCV is more balanced in the dry and wet edges.

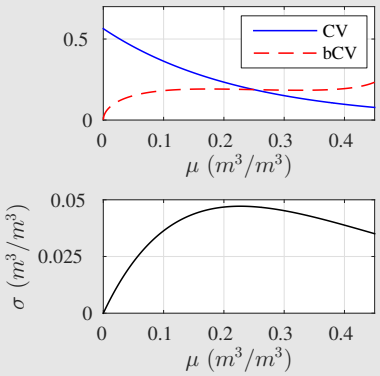


Figure 2.2 – Evolution of the classical and the bounded CV (top row) and the standard deviation (bottom row) for a typical SM spatial distributions

2.3.4 Dependence with the support and extent

From a probabilistic point of view, SM variability should be related to the measurement extent and support sizes: the larger the observed surface, the more probable it is to find contrasting environmental and landscape conditions that would increase the SM spatial variability. In the case of the aggregated support, the aggregation reduces the probability of finding extreme SM values within the same extent. This is true only if the number of supports contained within the extent is "large enough": very large aggregation supports reduce the number of SM samples within the extent, making the SM variance extremely sensitive to outliers.

Western and Blöschl (1999) provided a detailed description of the evolution of SM spatial variance and correlation length with spatial scale (support, extent and spacing). From the variogram analysis of a field of 1.05 km^2, intensively sampled at 30 cm depth, at 500-2000 locations and at 13 different days, they concluded that variance increases with extent and decreases with support, while the correlation range tends to increase with increasing spatial scale. These conclusions are summarised in Table 2.2.

Table 2.2 – Dependence of spatial variance and correlation length with spacing, support and extent according to (Western and Blöschl, 1999). "+" describes an increasing trend (if one increases, the other increases), "−" describes a decreasing trend (if one increases, the other decreases)

	Variance	Corr. length
Spacing	No conclusive	+
Extent	+	+
Support	−	+

Although the study of Western and Blöschl (1999) concerned small extents and did not use top surface measurements, different studies with 5-6 cm depth data on larger areas, have described the same type of behaviour so Table 2.2 is also valid for such measurements. Nykanen and Foufoula-Georgiou (2001) observed that variance increased with extent and decreased with the

support using the airborne 0.8 km ESTAR images acquired during the SGP97 campaign over an area of 51.2 km x 51.2 km. Later on, SMEX02 *in situ* TDR data (0-6 cm depth) allowed Choi et al. (2007) observing that variance decreased with aggregation within the basin scale (5000 km^2). At the same time, Famiglietti et al. (2008) found increasing standard deviation with extent in the SM *in situ* data gathered at five different extents, 2.5 x 2.5 m^2, 16 x 16 m^2, 100 x 100 m^2, 800 x 800 m^2 and 50 x 50 km^2, during 1 month. Most recently, SM *in situ* measurements and satellite observations (ALOS, ENVISAT) in the Rur catchment, at extents varying from 14 to 70 km, have also described the same dependence of variance with extent and support (Korres et al., 2015).

2.4 DISPATCH: a spatial disaggregation approach

The previous sections have described the main geophysical drivers of SM spatial variability, as well as its statistical behaviour within the satellite footprint. Knowing the principal geophysical drivers and their interactions is a necessary step for building downscaling models. Downscaling models improve the resolution of satellite SM estimates and consequently, they reduce the spatial scale mismatch between satellite and *in situ* measurements. It has been even suggested that downscaled SM products could be used for validation of the respective original satellite products (Crow et al., 2012; Malbéteau et al., 2016).

Direct validation with downscaled SM products suffers from an important drawback: downscaling models and inputs also contain errors and uncertainties. However, the downscaled SM maps could be used to provide a first estimation of the spatial scale mismatch. As mentioned at the beginning of this chapter, the spatial scale mismatch depends on the level of sub-footprint heterogeneity.

One of the premises of this thesis is that the approaches developed need to be globally applicable. There are no global disaggregation methods, but recently, the DISaggregation based on Physical And Theoretical scale Change (DISPATCH) algorithm (Merlin et al., 2012, 2013) was integrated as a Level-4 (L4) processor in the SMOS French ground segment to be run globally. The scope of this section is to investigate if this method can be effectively applied globally. If so, it could be used as a generic approach to estimate the spatial scale mismatch.

I personally participated in the development of the L4 DISPATCH processor, an activity that took place before the start of this thesis. Its validation was one of the first experiments of this thesis and both, the development and the validation, have been described in detail in a scientific article (Molero et al., 2016). The following sections present part of the validation results together with additional analyses, all placed in the perspective of the contents of this chapter: the SM spatial variability and the spatial scale mismatch. It is important to add that DISPATCH and the datasets presented in this section will be used again later in this manuscript.

Finally, in the downscaling literature, the terms *large-support* and *small-support* are usually replaced by *low-resolution* and *high(er)-resolution*, respectively. These concepts appeared summarised in Table 1.2, in chapter 1. To keep consistency, I will adopt this nomenclature in this section.

2.4.1 Background on optical-based downscaling

Passive microwave sensors on-board satellites provide surface SM observations at global extents with a support of some tens of km. The Soil Moisture Ocean Salinity (SMOS) mission is one of the most outstanding examples of this technology. The Soil Moisture Ocean Salinity (SMOS) sensor provides surface SM with a resolution of 27 to 55 km and a sensing depth of ∼5 cm

(Kerr et al., 2001, 2010). The SMOS SM products have been considered as suitable for hydro-climate applications (Wanders et al., 2014a; Lievens et al., 2015). However, hydro-agricultural applications need smaller-support observations (sub-kilometric) (Walker and Houser, 2004).

Large-support observations can be translated to a scale with smaller support in an operation called *downscaling* or *disaggregation*. It consists in using data at the target higher resolution to describe the variability within the large support. The ancillary data usually provide information on SM-related variables (brightness temperature (TB), surface temperature, evaporation, vegetation, precipitation) and static factors (texture, topography). As pointed out in section 2.1, the influence of each component in the spatial redistribution of SM should vary depending on the region of study.

Observations from active microwave sensors have been used to downscale passive SM observations (Das et al., 2011; Narayan et al., 2006; Zhan et al., 2006; Tomer et al., 2016). The Soil Moisture Active Passive (SMAP) mission, launched in 2015, was specifically designed for the delivering of SM at 3 km and 9 km resolutions by merging the L-band TBs with the higher resolution L-band backscatter data, both sensors on board.

Higher resolutions can be achieved by downscaling with optical data (visible, near-infrared, thermal-infrared). However, the temporal availability of the optical products and thus, of the disaggregated SM product, is degraded when clouds are present. The most common optical data used for downscaling include soil temperature and vegetation data, such as the LAI or the normalized difference vegetation index (NDVI) (Fang et al., 2013). In the 90s, it was observed that the land surface temperature (LST) and the NDVI described a dependence law with soil surface wetness and evapotranspiration. This was formalized under the *triangle* (Carlson et al., 1994; Carlson, 2007) and the *trapezoid* (Moran et al., 1994) approaches.

Chauhan et al. (2003) proposed a regression formula based on the triangle and trapezoid concepts that related SM, NDVI and LST. Later on, Piles et al. (2011) added SMOS TBs to the equation. The major drawback of these empirical approaches is that the regression coefficients need to be calibrated according to each specific area of study. In contrast, semi-empirical approaches replace the regression formula by evapotranspiration-related models that do not require local calibration. An example is that of Merlin et al. (2008b), who used the soil evaporative efficiency (SEE) as finer scale information to downscale SMOS observations. The SEE, which is the ratio of the actual to potential soil evaporation, was estimated as in Nishida et al. (2003), where the soil skin temperature was derived from MODIS data following the triangle approach. Later on, Merlin et al. (2012, 2013) improved the estimation of the soil skin temperature by taking into account the vegetation water status (before, it was considered as uniformly unstresses within the SMOS footprint). The whole disaggregation approach, called **DISaggregation based on Physical And Theoretical scale Change (DISPATCH)**, also included corrections for the SMOS antenna pattern, SMOS grid oversampling and estimations of the uncertainty in the output disaggregation data. The linear version of the SEE-SM model was validated in semi-arid regions, concretely in a 500 x 100 km area within the Murrumbidgee catchment in Australia (Merlin et al., 2012; Malbéteau et al., 2016) and in a 60 x 60 km area in Catalunya, Spain (Merlin et al., 2013). These studies showed that DISPATCH soil moisture had better spatio-temporal correlation with *in situ* measurements than the original SMOS products in semi-arid regions.

Based on the conclusions drawn from Merlin et al. (2012) and Merlin et al. (2013), a new version of DISPATCH was designed for operational and global purposes: some features were selected to prioritize robustness over accuracy, i.e. to facilitate its application over larger regions of the globe. The new version was integrated in a processor in the Level-4 (L4) of the Centre Aval de Traitement des Données SMOS (CATDS), the SMOS French ground segment. The L4 processor disaggregates SM products from the SMOS Level-3 (L3) segment into higher resolution (0.01°) SM products (Molero et al., 2016).

The following sections describe and evaluate this operational DISPATCH version. Note that in

the following, the term "L3 SM" refers to the SMOS L3 soil moisture product and the temrs "L4 SM" or "L4/DISPATCH SM" refer to the "disaggregated products". The aim of the evaluations is to investigate if DISPATCH succeeds to reduce the spatial scale mismatch under different climatic conditions (semi-arid, arid and sub-humid). DISPATCH relies on the assumption that soil temperature and vegetation green status are driving within-footprint SM spatial variability. Therefore, the results presented will serve also to assess whether these are global drivers of SM variability.

2.4.2 Datasets required by DISPATCH

2.4.2.1 Large-support datasets (SMOS)

The CATDS Level-4 Disaggregation processor (C4DIS) disaggregates the SM provided by the SMOS Level-3 1-day global SM product (MIR CLF31A/D). In this thesis, the product version 2.72 is used. L3 products are presented on the Equal-Area Scalable Earth (EASE) grid, with a grid spacing of ∼25 x 25 km.

Launched in November 2009, the SMOS (Kerr et al., 2001) is led by the European Space Agency (ESA) with collaboration the Centre National d'Etudes Spatiales (CNES) and the Centro Para el Desarrollo Tecnologico Industrial (CDTI). The sensor consists in a passive interferometer at L-band (1.4 GHz), whose approximate sensing depth is 5 cm. SMOS acquires TBs with different angles within its field of view, with resolutions that vary from 27 to 55 km, depending on the angle. In practice, the resolution of the SM product is considered 40 km on average. The maximum revisit time of the satellite is 3 days with crossing nodes at 6 a.m. and 6 p.m. local solar time for ascending and descending orbits, respectively.

The L3 SM products are directly computed from custom SMOS Level-1 brightness temperatures at the CATDS. The SM retrieval algorithm is similar to that of Level-2 (L2) (Wigneron et al., 2007; Kerr et al., 2012). The main difference is that the L2 processor considers the multi-angular observations of the same day and orbit, while the L3 processor uses several overpasses, 3 at most, within a 7-day window. Also, L3 products are presented in a 25-km grid (the EASE-Grid 2.0) while the L2 products are in a 15-km grid (the ISEA-4H9). The grid resolution must not be confused with the resolution of the sensor and the SM products, which ranges between 27 and 55 km as mentioned before. Details on the L3 algorithms and processor can be found in (Kerr et al., 2013a,b).

2.4.2.2 Higher-resolution (HR) datasets

The C4DIS processor uses the LST and NDVI datasets from the MODerate resolution Imaging Spectroradiometer (MODIS) mission to describe the within-SMOS pixel spatial variability. The MODIS sensor is a visible/near-infrared/thermal-infrared sensor onboard the NASA Terra and Aqua satellites since 1999 and 2002, respectively. Their revisit time is 1 day, with crossing nodes at 10:30 a.m. (Terra, descending orbit) and 1:30 p.m. (Aqua, ascending node). Data are acquired in 36 spectral bands with wavelengths ranging from 0.6 to 14.4 μm. The spatial resolution of the products varies between 250 m and 1 km, depending on the band.

The **LST datasets** are obtained from the version 5 of the MODIS/Terra daily L3 global LST and emissivity 1-km grid (MOD11A1) and MODIS/Aqua daily L3 global LST and emissivity 1-km grid (MYD11A1) products. The **NDVI dataset** is obtained from the version 5 of MODIS/Terra 16-day L3 global vegetation indices 1-km grid (MOD13A2). The products are retrieved from the NASA Land Processes Distributed Active Archive Center (LP DAAC). As their names indicate, the LST products have a daily frequency while the NDVI products have a 16-days frequency. Both are presented on a 1-km sinusoidal grid (Wan, 1999, 2006; Solano et al., 2010).

Finally, the C4DIS processor requires **elevation information**, which is extracted from the GTOPO30 digital elevation model (DEM) product, available in WGS84 coordinates at a 30-arc sec resolution and distributed by the U.S. Geological Survey (USGS) EROS Data Center.

2.4.3 DISPATCH algorithm

The main principles of the DISPATCH version used in the C4DIS processor is shortly described here. A detailed description is included in A. DISPATCH distributes the SM within the large-support pixel according to the spatial distribution of the soil evaporative efficiency (SEE). The SEE is the ratio of the actual to potential soil evaporation. In regimes where evaporation is moisture-controlled, i.e. where the evaporation rate depends on the amount of SM, the total potential soil evaporation is proportional to the difference in temperature between a wet soil and a dry soil. This is typical from semi-arid climates. Note that in energy-controlled scenarios, the presence of SM is not sufficient for evaporation. Humid climates are often energy-controlled.

The SEE is estimated at the target higher resolution (HR) following the model

$$SEE_{HR} = \frac{T_{s,max} - T_s}{T_{s,max} - T_{s,min}} \tag{2.6}$$

with T_s being the soil temperature and $T_{s,max}$ and $T_{s,min}$ being the soil temperatures at minimum and maximum soil moisture, respectively. They are derived from the optical observations (MODIS). Using the triangle approach (Carlson et al., 1994), T_s can be expressed as:

$$T_s = \frac{T_{surf} - f_v T_v}{1 - f_v} \tag{2.7}$$

with T_{surf} being the LST from the higher-resolution sensor (MODIS), f_v the fractional vegetation cover and T_v the vegetation temperature, both to be estimated.

The fractional vegetation cover is calculated from the NDVI obtained from the higher-resolution sensor (MODIS). Vegetation temperature T_v in Equation 2.7 is estimated with the "hourglass" approach in Moran et al. (1994). Temperature end-members in Equation 2.6 ($T_{s,min}$, $T_{s,max}$) are obtained from the minimum and the maximum surface temperatures in the scene ($T_{surf,min}$, $T_{surf,max}$), respectively. This is done through approximations the LST-f_v space (Merlin et al., 2013).

Once SEE_{HR} has been obtained for all the higher resolution (HR) pixels within the large-support pixel, the **downscaling relationship** can be applied to obtain soil moisture at the target resolution (SM_{HR}):

$$SM_{HR} = SM_{LR} + \frac{SM_p}{\partial SEE_{LR}} \cdot (SEE_{HR} - SEE_{LR}) \tag{2.8}$$

with SM_{LR} being the soil moisture observed by the low resolution sensor (SMOS), SEE_{LR} the spatial average of the SEE_{HR} values within the low resolution (LR) pixel and SM_p a parameter (its calculation is included in A).

Finally, it is worth highlighting that the downscaling equation (2.8) ensures that **the mean SM is preserved along the disaggregation**: the low-resolution SM equals the spatial mean of the high-resolution SM values.

2.4.4 Validating DISPATCH products

2.4.4.1 Validation datasets: dense *in situ* networks

Four *in situ* networks have been selected to evaluate the disaggregated products: the Murrumbidgee (MB) network and its Yanco area in Australia (Smith et al., 2012) and the Little Washita (LW) (Cosh et al., 2006) and Walnut Gulch (WG) (Cosh et al., 2008) networks in the U.S.. They provide multiple measurement points within the SMOS footprint and they exhibit contrasting climate and landscape characteristics.

The MB network is located in southern New South Wales and is considered as a sparse network (82,000 km^2). Its climate changes from semi-arid in the west (the average annual precipitation is 300 mm), to humid in the east (1900 mm in the Snowy Mountains). The MB network consists of 38 stations: 18 provide SM integrated over the top 8 cm of soil (TDR Campbell CS615 sensors) and the rest provide SM for the top 5 cm of soil (capacitance Stevens Hydra Probes). Seven stations are located in the limits of the catchment, 5 in Adelong Creek (a grazing 145-km^2 area with steep slopes), 13 in Kyemba Creek (a grazing and dairy 600-km^2 area with gentle slopes), and 13 in the Yanco area.

The Yanco area (3,000 km^2) belongs to the MB network. In this study, it is analysed separately because it constitutes itself a dense *in situ* network and because the region meets DISPATCH nominal environmental conditions: Yanco is a flat low-vegetated semi arid area. Yanco is dedicated to irrigated and dry farming, with some pastures. The climate is semi-arid (average annual precipitation of 400 mm), with most of the precipitation occurring in winter and spring. Soils are mainly silty-loam. The area has been extensively monitored since 2001 (Merlin et al., 2008a; Peischl et al., 2012; Smith et al., 2012; Panciera et al., 2014).

The Little Washita (LW) and Walnut Gulch (WG) networks consist of capacitance probes (Stevens Hydra Probe) installed at 5 cm depth. They started operating in 2002 and they have been also used in numerous validation campaigns (Jackson et al., 2010, 2012; Leroux et al., 2013a; Bindlish et al., 2015).

The LW network (610 km^2) is located in southwest Oklahoma. The climate is sub-humid (average annual precipitation of 750 mm). Winters are short and temperate, summers are hot and relatively dry and spring and autumn concentrate most of the annual precipitation (Allen and Naney, 1991). The land use is mainly rangeland and crops like winter wheat, corn and grasses. Soil texture is diverse, with large regions of sands, loams and clays. The surface is moderately rolling.

The WG network (148 km^2) is located in southeastern Arizona. The climate is semi-arid to arid (324 mm of precipitation per year), dryer than the Yanco area. Most of the rains occur in the form of small scale high-intensity thunderstorms in summer (Cosh et al., 2008). The land cover consists of desert shrubs and short grasses. Soils are mainly sands and gravels. The topography is rolling with significant rock cover.

Table 2.3 – Main characteristics of validation areas for DISPATCH

	Yanco	Murrumbidgee (MB)	Little Washita (LW)	Walnut Gulch (WG)
Extent	3,000 km^2	82,000 km^2	610 km^2	148 km^2
SMOS pixels covered	∼1	> 1	∼1/4	∼1/16
Climate	Semi-arid	Semi-arid to wet	Sub-humid	Semi-arid / arid
Annual precipitation	400 mm	300-1900 mm	750 mm	324 mm
Main precip. periods	Winter, spring	∼constant at the basin scale	Autumn, spring	Summer
Soils	Silty-loam	Clays (west) to sands (east)	Diverse	Sands and gravel
Topography	Flat	Diverse	Moderately rolling	Rolling

2.4.4.2 Data preparation

The periods selected for the Australian and the U.S. networks are different: 1 year (2010/06-2011/05) in the case of Yanco and MB, and 5 years (2010/06-2014/05) in the case of LW and WG. This was motivated by the availability of *in situ* data.

Data has been filtered out when the radio-frequency interference (RFI) probability of the original SMOS product exceeded 10 %. The RFI dataset accounts for the percentage of TBs affected by RFI presence (Oliva et al., 2012; Kerr et al., 2013a). Only the *in situ* samples taken at SMOS overpass times (± 1 h) are selected for evaluation. The three time series (*in situ*, L3 and L4) are evaluated on common dates with valid SM values (DISPATCH can produce negative SM values that were filtered out). The final number of samples used is indicated in Table 2.4.

Table 2.4 – Number of samples used to evaluate DISPATCH SM products, classified per *in situ* network

	Yanco	Murrumbidgee (MB)	Little Washita (LW)	Walnut Gulch (WG)
Nb of days (with 5 or more spatial samples)	140	195	1130	1097
Total nb of samples (both in time and space)	1477	2838	19,084	18,364

2.4.4.3 Evaluation metrics

We compare the L3 and L4/DISPATCH SM datasets with the *in situ* datasets using conventional validation metrics: correlation (R), bias (B) and standard deviation of the differences (STDD). The STDD is expressed as

$$STDD = \sqrt{E\left\{\left[(SM_{satellite} - E\{SM_{satellite}\}) - (SM_{insitu} - E\{SM_{insitu}\})\right]^2\right\}} \tag{2.9}$$

where $E\{\cdot\}$ is the mathematical expectation operator. I prefer the standard deviation of the differences (STDD) to the RMSD (Equation 2.10) because it is independent from biases in the mean and amplitude values:

$$RMSD = \sqrt{E\left\{\left[SM_{satellite} - SM_{insitu}\right]^2\right\}} = \sqrt{STDD^2 + B^2} \tag{2.10}$$

where B is the bias and "satellite" stands for either the L3 or the L4 dataset. The STDD is also referred as the unbiased root mean squared difference (ubRMSD) in the SM literature. The term STDD is used here to be consistent with the original definition of the metric (Mood et al., 1974; Salkind, 2010).

One more metric is used, the slope of the regression line (S), to evaluate the gain in spatial representativeness:

$$S = R \cdot \frac{\sigma_{satellite}}{\sigma_{insitu}} \tag{2.11}$$

While R is the slope of the *standardized* regression line, S takes into account the mismatch in standard deviations (σ) between the satellite and the *in situ* datasets. Generally speaking, the ultimate objective of a disaggregation approach like DISPATCH is that disaggregated and *in situ* values are as close as possible to the 1:1 line with minimum dispersion (Merlin et al., 2015).

2.4.5 Results and discussion

The scope is to evaluate whether DISPATCH is valid as a generic disaggregation approach. The analyses here follow 3 complementary strategies: visual inspection, spatial and temporal evaluation. It is important to remind here that the spatial scale mismatch is not only of spatial nature but also of temporal nature. As a proof, SM measurements acquired with different support sizes or at different locations exhibit different temporal dynamics. For this reason, the evaluation of DISPATCH not only includes visual and quantitative spatial evaluations but also temporal ones.

Initially, disaggregated products derived from different SMOS orbits (ascending and descending) were evaluated separately, but the statistics obtained were similar (not shown here). For this reason, in the subsequent analyses, ascending and descending orbits are combined together.

Finally, prior to the evaluations of the L4/DISPATCH dataset, I conduct a preliminary analysis of the *in situ* spatial variability. This will be important in the performance of DISPATCH as it will be shown later. This is presented in the following section.

2.4.5.1 Preliminary analysis

DISPATCH is contextual because the SEE depends upon the scene LST maximum and minimum values. The accuracy of their estimation depends on the dry-wet contrast in the area: the higher the contrast, the better their accuracy (Merlin et al., 2010). For this reason, a preliminary study on the *in situ* SM spatial variability is needed.

The spatial variance can be estimated with the sample variance of the SM values measured by the stations located within the same satellite footprint. This is only feasible in dense *in situ* networks, like those used in this chapter. In the case of sparse networks, the level of spatial variability could be guessed from the climatic conditions of the area. The literature review in section 2.3.2 indicated that the spatial variance describes a convex relationship with the mean SM content. Therefore, we could foresee that in regions where the mean SM content is either very low or very high, the SM spatial variability will be low and DISPATCH might not be an adequate proxy to estimate the spatial scale mismatch.

This idea will be verified with the *in situ* data of the networks used in this chapter (Yanco, MB, LW, WG). Since the objective is to predict the performance of the DISPATCH-SMOS products, I only select the measurements closer to the SMOS ascending and descending overpass times. Figure 2.3 shows the spatial standard deviation *vs* the spatial mean of the *in situ* measurements at each time step, classified per network. The standard deviation describes an increasing trend with the mean SM content: when the mean SM level increases, the standard deviation also increases. This seems in opposition with section 2.3.2, which concluded that the convex relationship was the actual law in footprint extents. However, it is not: the convex relationship was only observed when mean SM values higher than 0.3 m^3/m^3 were present (e.g. Figure 2.1). Otherwise, the convex function appeared truncated, resulting in an only-increasing trend. Such high SM values can only be registered just after spatially homogeneous rain events. They are not observed here, either because the temporal sampling is not concurrent in time with the rainfall (and the soil succeeds to drain the income water before the sampling), either because precipitation is heterogeneous in the regions studied.

Yanco and the MB show a wide standard deviation range: the spatial standard deviation ranges between 0 and 0.18 m^3/m^3. This implies that DISPATCH should help to reduce the spatial scale mismatch on a big number of dates. There are more high standard deviation values in Yanco than in MB, so probably the disaggregation will be gainful for more dates in Yanco than in the MB.

The wide range of variability values in Yanco is probably due to the irrigation and vegetation (crops) heterogeneity of the area. In the case of the MB, its wide extension ensures a wide range of observed SM values (the MB covers multiple SMOS pixels).

Little Washita (LW) mean SM values reach moderately wet values (~0.3 m^3/m^3). However, when compared to Yanco and MB, its standard deviation saturates at 0.08-0.09 m^3/m^3 (3rd column of Figure 2.3). This can be due to different reasons. First, the extent of the LW network is smaller than Yanco and MB. The same is also true for Walnut Gulch (WG). For example, LW is 4 times smaller than the extent of Yanco. Section 2.3.4 indicated that smaller extents tend to show lower spatial variability. This seems geophysically plausible since smaller extents tend to concentrate less heterogeneity in precipitation, soils and vegetation. The second reason is that the sub-humid climate of LW contributes to the spatial persistence of SM and probably makes soil evaporation to be energy-driven (the presence of humidity in the soil does not forcedly leads to evaporation). This together with the low spatial variability are important disadvantages for the application of DISPATCH in this area.

Finally, mean SM values are extremely low (< 0.2 m^3/m^3) in Walnut Gulch, which is typical from an arid climate (4th column of Figure 2.3). Since SM content is low, so it is the spatial variability: most of the values are concentrated below 0.08 m^3/m^3 of standard deviation, which contrasts with the Yanco values that can reach 0.18 m^3/m^3. Although the evapotranspiration in WG should be moisture-driven as in Yanco, the low spatial variability could prevent DISPATCH from producing meaningful disaggregated SM maps.

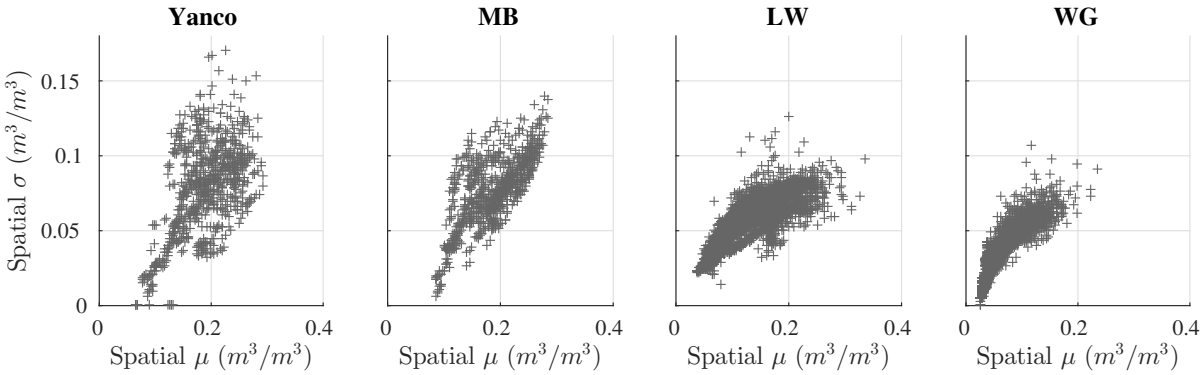


Figure 2.3 – Th spatial standard deviation *vs* the spatial mean of the *in situ* samples of Yanco, MB, LW and WG, at the SMOS overpass times.

2.4.5.2 Visual inspection of disaggregated maps

Visual inspection of disaggregated maps is a good first step in the evaluation of the method. Regarding Yanco and the MB, Figures 2.4 and 2.5 give evidence on how DISPATCH succeeds to reveal important SM spatial features. In Figure 2.4, we observe the irrigated farms of the Coleambally cooperative. Figure 2.5 clearly shows the evolution of the floods that affected New South Wales and Victoria states on the first days of January 2011. All this is consistent with the analysis of the previous section, where the climatic conditions of the areas together with high spatial variability values suggested that DISPATCH could effectively improve the resolution of SMOS products.

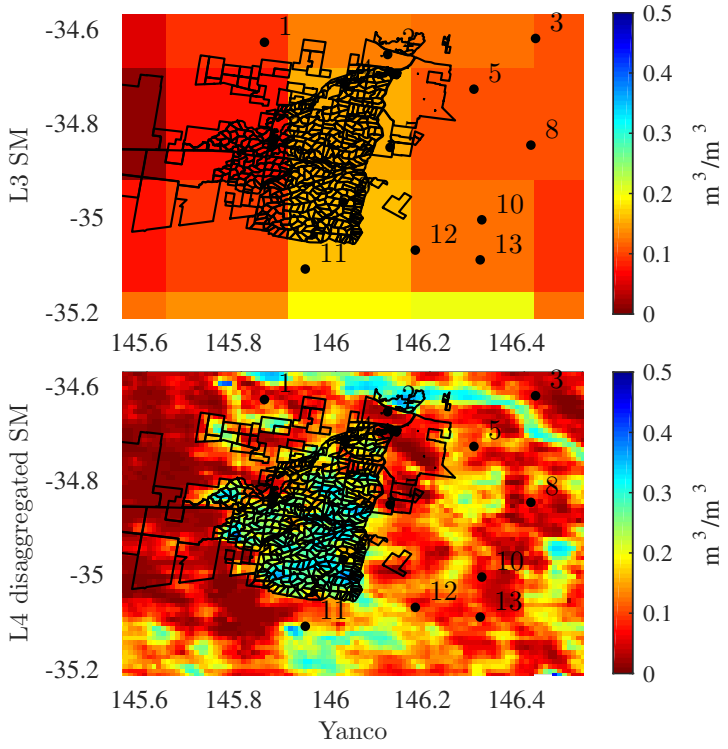


Figure 2.4 – Maps of L3 soil moisture and L4/DISPATCH soil moisture for the Yanco area on 2010/11/22. The black contours define the Coleambally irrigated farms. Source: Molero et al. (2016)

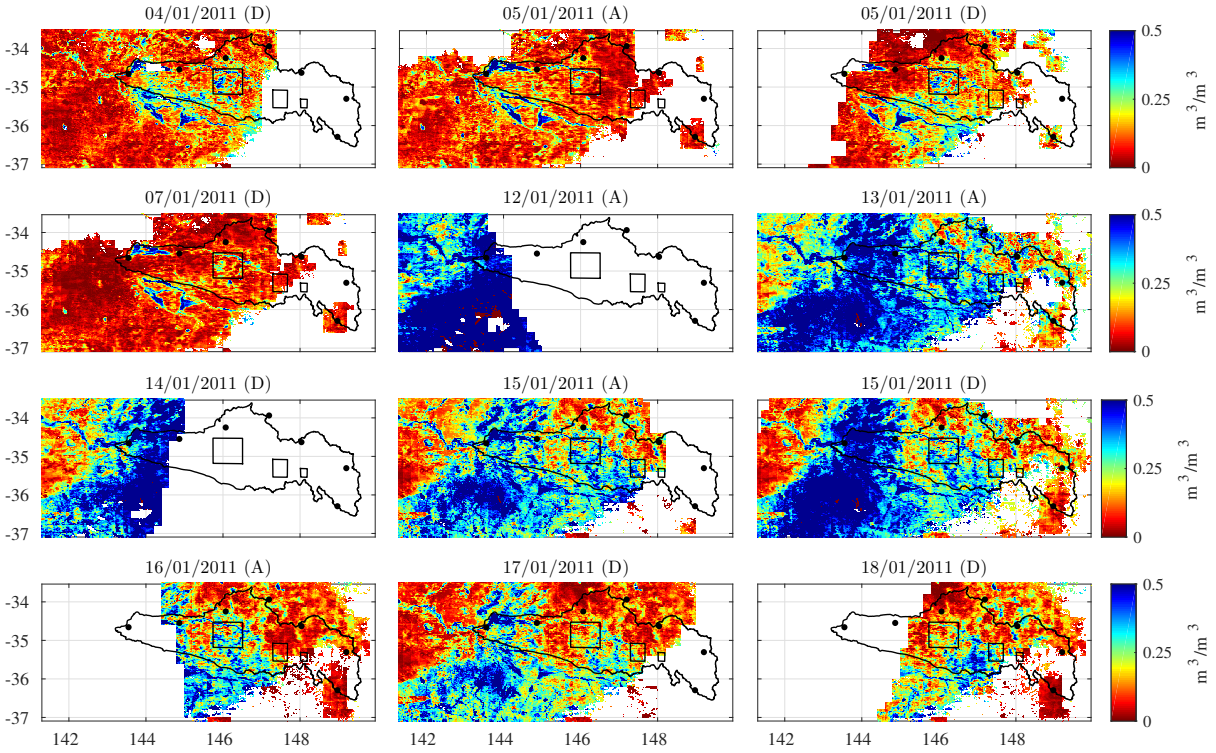


Figure 2.5 – Maps of L4 disaggregated soil moisture for MB showing the progression of floods on the first days of January 2011. The day and the orbit (A/ascending, D/descending) is indicated in the title of each image. The black squares correspond from left to right to Yanco, Kyemba and Adelong areas. Source: Molero et al. (2016)

Figure 2.6 shows two sample maps for the LW and the WG watersheds. In the LW case (left), disaggregation succeeds to reveal the surrounding lakes but not the LW river. A poor performance of DISPATCH is expected as result of the climatic and statistical conditions of the area, which were described in the previous section. Moreover, the limits between the original SMOS pixels are apparent. This can be explained because DISPATCH disaggregates each SMOS pixel independently without taking into account the information of surrounding pixels.

The WG images (right column in 2.6) were selected at one day when the spatial σ was high for that region (0.05 m^3/m^3). Despite this, the WG disaggregated field still exhibits little SM heterogeneity. This can be due to the poorer performance of DISPATCH under spatially homogeneous conditions of moisture, but also to the natural variability of SM: as reminded in the previous section, low mean SM is linked to low spatial variability. Moreover, WG is a fairly homogeneous area in terms of soil, vegetation and climate, and SM heterogeneity is mainly driven by precipitation.

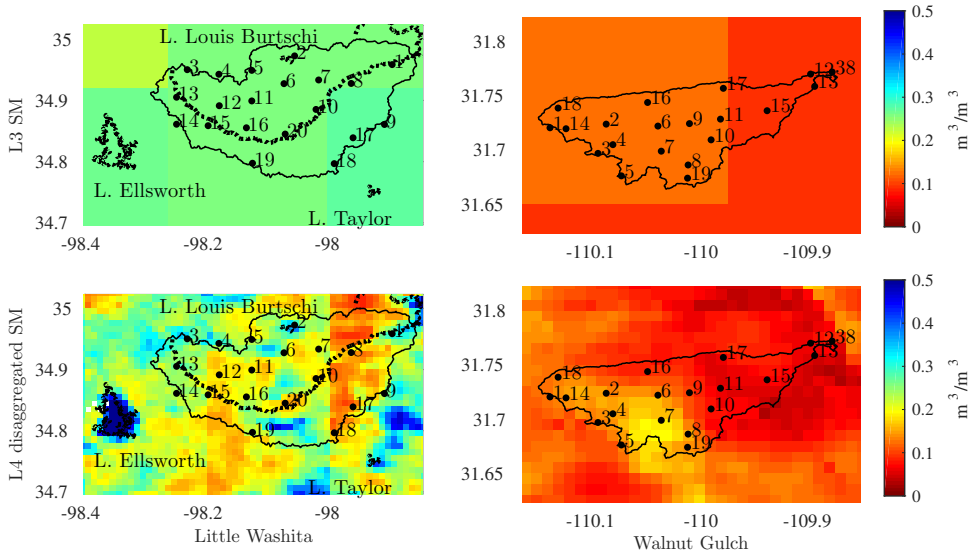


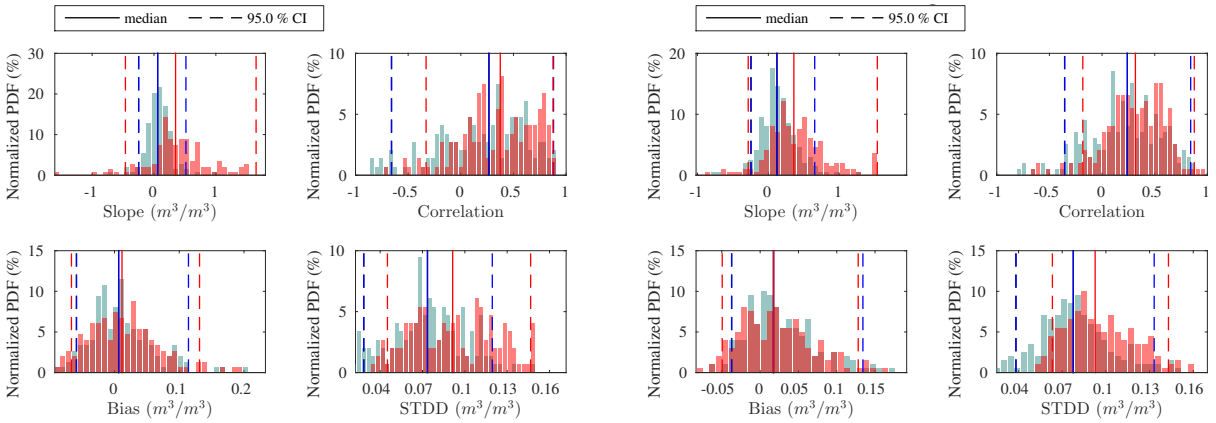
Figure 2.6 – Maps of L3 soil moisture and L4 disaggregated soil moisture for the LW watershed (left) and the WG watershed (right) on 2011/05/02 and 2012/01/03, respectively. Solid contours indicate the watershed boundaries. In the LW maps, the river and lakes are represented by the bold dotted lines. Source: Molero et al. (2016)

2.4.5.3 Spatial evaluation

The scope is to verify if DISPATCH can be applied globally to estimate the within-footprint spatial variability that should give access to the spatial scale mismatch. The preliminary and visual assessments have already given some hints: DISPATCH should provide a good estimation of the spatial redistribution of SM for those scenarios where the redistribution is controlled by evapotranspiration and where the spatial variability of SM is high (probably Yanco and the MB, possibly WG, unlikely LW). This section evaluates if DISPATCH from a spatial perspective. The following section will be dedicated to the temporal analysis.

The spatial evaluation consists in comparing the L4/DISPATCH and the L3 SM values at the *in situ* station locations with the respective *in situ* measurements at each timestamp. Usual statistics are calculated at each timestamp. Note that the number of values to compute the spatial statistics is quite low: the maximum dataset sizes are 13, 38, 20 and 20 for Yanco, MB, LW and WG, respectively. In practice, the number is lower if clouds are present. I arbitrarily set 5 as the minimum dataset size to calculate spatial statistics. Due to the small sample size, the uncertainty in the statistics is non negligible and should be considered when analysing the results.

The distribution of spatial statistics with their median and 95 % percentile values for Yanco and MB are shown in Figures 2.7a and 2.7b. The distributions here are consistent with the mean statistics presented in Merlin et al. (2012) and Malbéteau et al. (2016) for the same areas. In Figures 2.7a and 2.7b the distributions of slope and correlation move rightwards after disaggregation, which indicates that **the spatial correlation is improved**. This is more visible in the case of Yanco. The bias remains similar after the disaggregation but the STDD is clearly degraded, although not too much (∼0.03 m^3/m^3). This is expected since the disaggregation step can introduce noise from the ancillary data and the method itself. These results confirm that a) **DISPATCH reduces the spatial scale mismatch in semi-arid areas with high spatial variability**, and b) **soil evaporative efficiency (SEE) is a good proxy for SM variability in those areas**(the Murrumbidgee and Yanco).



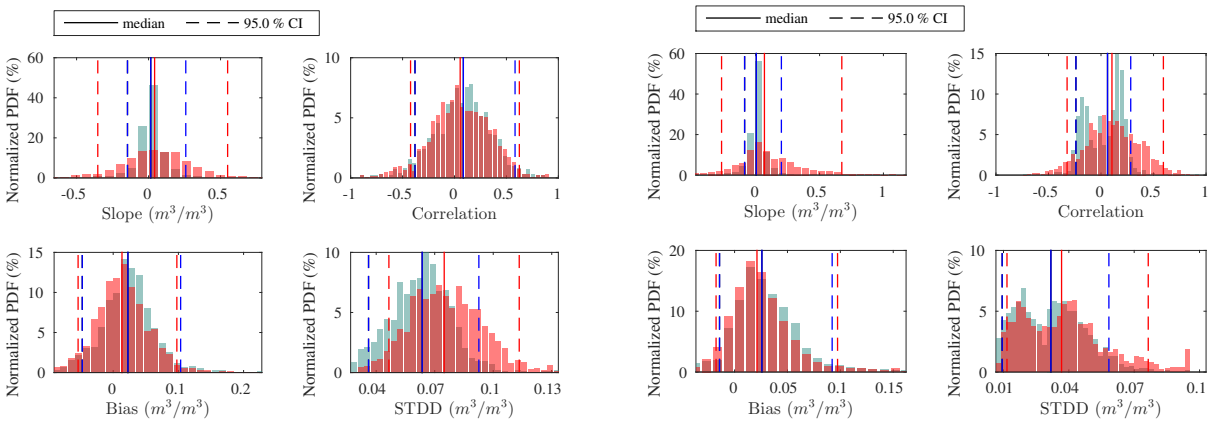
(a) Yanco (b) Murrumbidgee

Figure 2.7 – Distributions of the spatial statistics from the comparison of *in situ* to L3 and L4/DISPATCH soil moisture products. L3 statistics are in blue and L4/DISPATCH statistics are in red

The results for Little Washita and Walnut Gulch, shown in Figure 2.8, are less convincing. In the case of LW, the only improvement appears in the S metric: more values approach 1 after disaggregation, although the scatter increases. This reflects that **the SEE is not a good proxy for SM spatial variability in the LW watershed**. This seems plausible since the climate of LW is sub-humid, which implies that the evaporation rate is energy-controlled and not driven by the available SM. In Molero et al. (2016), it is shown that this situation is partially reverted during drier conditions in summer. Finally, the LW watershed exhibits large variability in vegetation and texture types, which are not taken into account by DISPATCH. As shown in Pan et al. (2003), they are important sources of SM variability in LW. New SEE models that rely on texture information are promising (Merlin et al., 2016).

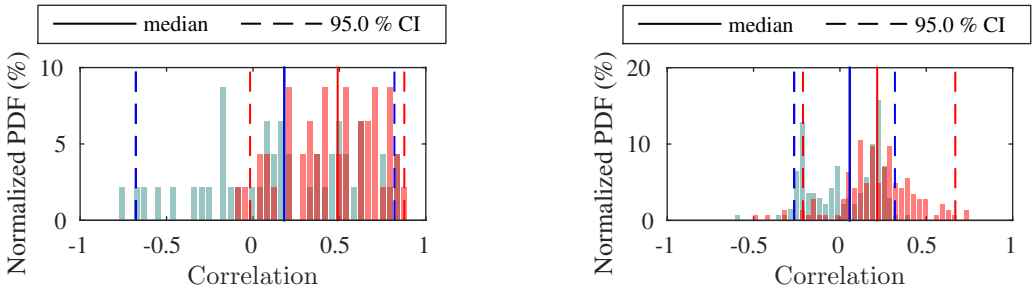
Regarding WG (Figure 2.8b), **the main improvement concerns the spatial correlation**: the proportion of negative correlations decreases while the proportion of positive correlations increases after disaggregation. However correlation values are lower than in the Australian cases, which could be explained by the lower SM spatial variability of the WG region. The sandy and gravel soils of WG make infiltration much faster, so even if precipitation heterogeneity was present, the measured sample at the SMOS overpass times would miss the SM heterogeneity induced by precipitation.

Finally, Figure 2.9 illustrates the gain of applying DISPATCH in Yanco and WG in periods of the year where the spatial variability of SM is considerable. For simplicity, only the spatial correlation is shown. In the case of Yanco (left), the optimal period is summer. When the summer distribution (Figure 2.9a) is compared against the year distribution (Figure 2.7a), it can be seen that the number of negative correlations has been drastically reduced and that they have moved towards higher values of correlation. In the case of WG (Figure 2.9b), the performance is better during winter than during the rest of the year (Figure 2.8b).



(a) Little Washita (b) Walnut Gulch

Figure 2.8 – Distributions of the spatial statistics from the comparison of *in situ* to L3 and L4/DISPATCH soil moisture products. L3 statistics are in blue and L4/DISPATCH statistics are in red



(a) Yanco summers (b) Walnut Gulch winters

Figure 2.9 – Distribution of spatial correlation values from the comparison of *in situ* to L3 and L4/DISPATCH soil moisture products. The statistics are calculated for (a) Yanco during summer (2010/12/01-2011/03/01) and for (b) WG during winters (December, January and February months of years 2010 to 2014)

2.4.5.4 Temporal evaluation

The principal objective of disaggregation is to improve the spatial resolution but the temporal dynamics might be also influenced. Figure 2.10 presents all the *in situ* SM samples compared to their respective L3 and L4 disaggretated SM values under the form of scatter plots. It should be reminded that here the values displayed are not spatial means: they correspond to each of the samples taken in space and time. Therefore, there is not necessarily any link between the SM values and the spatial variability of the area.

In the case of Yanco and MB (two left-most columns in Figure 2.10), the central part of the distribution is closer to the 1:1 line although it becomes more scattered after disaggregation. This can be explained by the noise introduced by the disaggregation method and the ancillary data.

In the case of LW, a small improvement is seen but only for very dry conditions (< 0.15 m^3/m^3). In this part of the SM range, the evaporation rate may be moisture-driven instead of energy-driven. Similar conclusions were reached when LW summers were analysed (this is not shown here but can be found in Molero et al. (2016)).

Finally, in the case of WG, no temporal gain can be observed after disaggregation: the range of both satellite and *in situ* SM values is very small and concentrated in the driest region. This is expected since the spatial variability in the WG region is low most of the time excepting for particular dates.

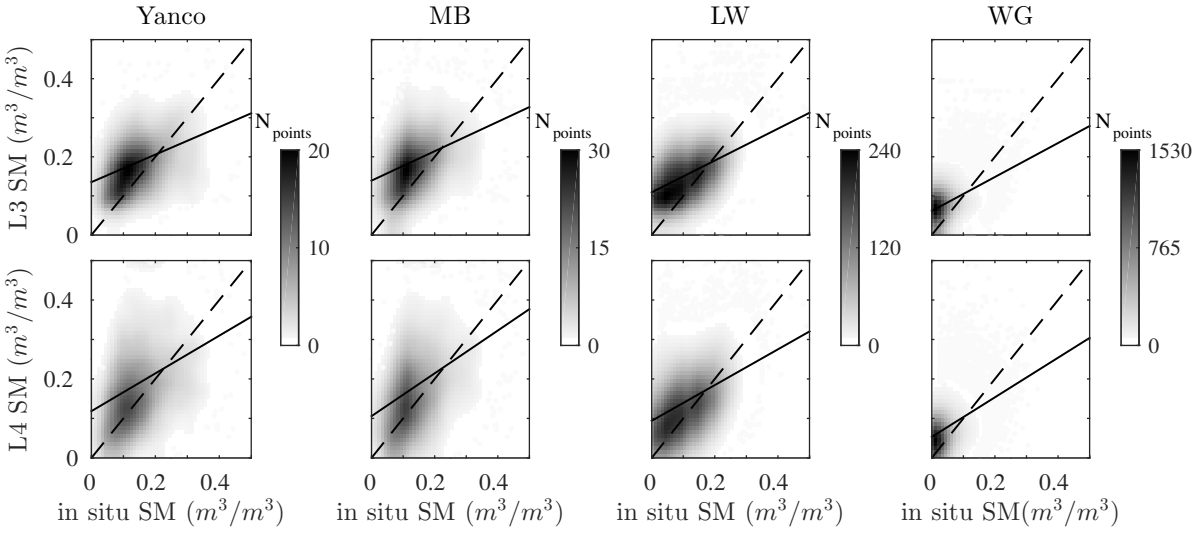


Figure 2.10 – Scatter plots of *in situ* samples compared to L3 (upper row) and L4 disaggregated samples (bottom row), classified per *in situ* network. Source: Molero et al. (2016)

2.4.5.5 Discussion

It has been showed that the L4/DISPATCH algorithm reduces the spatial scale mismatch between SMOS soil moisture products and *in situ* measurements under two simultaneous conditions:

1. Moisture-driven scenarios in regions whose spatial redistribution of SM is principally controlled by precipitation and evapotranspiration
2. Regions with a relatively high spatial variability of SM

These conditions were met by Yanco, the Murrumbidgee and Walnut Gulch. The degree of improvement was controlled by the number of dates with good spatial contrast. The improvement was far less important in Walnut Gulch because the SM spatial heterogeneity did not last in time and could not be registered by the temporal sampling. In contrast, the analysis of the disaggregated products in the Little Washita (LW) indicated that DISPATCH is not suited for sub-humid climates with low SM spatial variability.

All this suggests that unfortunately, DISPATCH products cannot be used globally to estimate/overcome the spatial scale mismatch. A preliminary assessment on the adequacy of the region of study should be conducted. Nevertheless, the contribution of DISPATCH is highly valuable: their disaggregated products are a good proxy for point SM measurements under the conditions stated before. DISPATCH products improve the spatio-temporal correlation with *in situ* measurements. DISPATCH may not be used to estimate the spatial scale mismatch at a global scale but can help to study it in some regions of the globe.

2.5 Conclusions

A good understanding of the spatial variability of soil moisture (SM) is necessary for better addressing the study and evaluation of the spatial scale mismatch: the spatial scale mismatch is intrinsically related to the sub-footprint SM redistribution. From a physical perspective, 4 landscape and environmental factors that drive SM variability were identified: precipitation, soil, vegetation and topography. They influence a variety of SM-related processes and variables, such as evapotranspiration, temperature and drainage. If their respective roles in the SM spatial redistribution are identified, downscaling approaches can be built to improve the resolution of

large-support SM measurements. I initially suggested that disaggregated SM fields could be used as a proxy to better estimate the sub-footprint SM variability.

However, it has been shown that downscaling approaches have a main drawback: they are not applicable globally since the interactions between soil moisture and SM-related variables varies between regions. This was illustrated in this chapter with the evaluation of DISPATCH, a semi-physical disaggregation algorithm. DISPATCH was able to reduce the spatial scale mismatch in areas where the soil evaporative efficiency (SEE) was a good proxy for soil moisture variability, like the Yanco region in the Murrumbidgee catchment (Australia) and the Walnut Gulch watershed in Arizona (Oklahoma, U.S.). It was not the case in other sites like the Little Washita watershed (U.S.), where the spatial variability was more controlled by vegetation and texture than by temperature and evaporative demand.

An important part of this chapter was dedicated to the study of former literature on the statistical spatial distribution of SM. From this I concluded that the SM distribution within the satellite footprint is right-skewed when the mean SM is close to the dry-end, left-skewed when the mean SM approaches the saturation point and bell-shaped in the SM midrange. The analysis of *in situ* data from Yanco, the Murrumbidgee, Little Washita and Walnut Gulch, confirmed the conclusions derived from literature. The mean-variability dependence is an important notion that will be used for modelling the sampling uncertainty in chapter 5.

Knowing the mean-variability dependence for the region of study has resulted also useful to improve the evaluation of disaggregated SM products. Section 2.4.5.1 showed that a preliminary analysis of the spatial variability of SM allowed to predict the downscaling performance: the areas with limited spatial variability at the satellite overpass times were those where DISPATCH performed the worst. In this case, a first estimate of the spatial variability was derived from the *in situ* measurements. In the case of sparse networks, where only one station is available per satellite footprint, the level of spatial variability could be guessed if the climatic characteristics of the area are known.

Finally, the disaggregation of SMOS SM in the semi-arid Yanco region reduced the distance, in terms of correlation and bias, between the satellite and the *in situ* measurements. It also succeeded in revealing strong SM transitions within the footprint due to irrigation. For this reason, the disaggregated products will be used in chapters 3 and 4 as proxies for the SM spatial variability in Yanco.

Chapter 3

Multi-time scale decomposition of soil moisture series

Contents

3.1 Introduction . . . 52
3.2 Time scales: wavelet-based approaches . . . 53
3.2.1 General definition . . . 53
3.2.2 The maximal overlap discrete wavelet transform (MODWT) . . . 54
3.2.3 Practical considerations . . . 57
3.2.3.1 Boundary conditions . . . 57
3.2.3.2 Choice of the base wavelet filter . . . 57
3.2.3.3 Choice of the maximum level of decomposition J0 . . . 58
3.2.4 Wavelet-based statistics . . . 58
3.3 Datasets . . . 59
3.3.1 Modelled series (Little Washita) . . . 60
3.3.2 *In situ* measurements . . . 62
3.3.3 SMOS data . . . 63
3.3.4 Data preparation and wavelet parameters . . . 64
3.3.4.1 Periods of analysis, sampling periods and wavelet scales . . . 64
3.3.4.2 Wavelet filters and maximum level of decomposition J0 . . . 64
3.3.4.3 Pre-processing of *in situ* series . . . 65
3.3.4.4 Pre-processing of SMOS time series . . . 65
3.4 Decomposition of soil moisture series . . . 66
3.4.1 Decomposition of point modelled series . . . 66
3.4.2 Time-scale decomposition of *in situ* series . . . 69
3.4.3 Time-scale decomposition of SMOS series . . . 74
3.5 Conclusions . . . 76

The previous chapter has been focused on the spatial dimension of the surface soil moisture (SM) variable. However, surface SM is highly variable not only in space but also in time. The drivers that control SM variability in the spatial domain also control the SM variability in the time domain: for example, precipitation forms the characteristics peaks of SM time series, soil characteristics shape the dry-down decay after a precipitation event, topography can create SM trends and vegetation can be responsible of both seasonal and dry-down responses. It can be inferred that the responses linked to the SM drivers are associated to characteristic time scales, implicitly implying that SM is a multi-time scale geophysical variable.

This chapter is dedicated to the analysis of the SM variable from a multi-time scale perspective. After reviewing the past research on SM time scales, the analyses of SM time series in the time domain and on a per-time scale basis will be presented. The method used for time-scale decomposition is based on wavelet transforms and is also described.

This chapter, together with chapter 2, are preparatory and necessary steps for chapter 4, which will be dedicated to the connections between SM spatial and temporal scales.

3.1 Introduction

Several studies have highlighted the multi-time scale nature of the SM process. They have usually divided the SM temporal dynamics in slow variations and fast variations. The publication of Entin et al. (2000) is the most relevant study on SM time scales until now. Based on a vast *in situ* SM dataset (measurements taken at 0.01-1 m depth for 11-25 years in the U.S., Russia, Mongolia and China), Entin et al. (2000) built a temporal autocorrelation model made of a short-time scale component and a large-time scale component. The short one was of the order of a few days and was related to land local characteristics (soils, topography, vegetation and root structure). The large-scale component was of the order of 1-2 months and was associated to atmospheric conditions (precipitation and evaporation patterns). Similarly, Nicolai-Shaw et al. (2016), based on satellite and model-based SM, showed that the seasonal scale of SM (4-5 months) was more driven by large-scale atmospheric conditions than by the natural tendency of SM to persist in time. In contrast, the sub-seasonal scale (1-3 months) was mainly driven by temporal SM persistence, i.e. the integration of short-scale effects over time.

Several other studies have also distinguished different behaviours between the SM fast-varying and the slow-varying components. Soil moisture "anomalies" (the fast-varying component) are usually obtained by removing the "trend" (slow-varying component), which is calculated as a temporal mean or with a ∼30-day moving average. For example, Gruber et al. (2013a) cited several publications (e.g. Cosh et al., 2006; Brocca et al., 2007; Famiglietti et al., 2008) that highlighted that, for the same location, the anomalies of SM datasets from different sensors and models tend to be better synchronised than the absolute values. Also, some validation approaches like triple collocation (TC) are systematically applied to detrended SM series because it is assumed that slow-varying components (> 30 days) mainly contain systematic differences between the datasets (Gruber et al., 2016). Similarly, Mittelbach and Seneviratne (2012) found that the differences between the trends of different point SM datasets are larger than those between their anomalies.

The geophysical and statistical analyses mentioned above seem to indicate that SM time scales are induced by different geophysical factors and that SM time scales have different properties. This has important implications on SM validation studies (e.g. triple collocation). However, SM time scales have not been yet addressed exhaustively and only the division in fast/slow variations has been studied. The following sections propose a comprehensive study of the time scales of both point-support and footprint-support SM datasets acquired at different geographical areas with contrasting climatic and environmental conditions. Soil moisture time scales will be obtained with wavelet transforms, a well known technique for decomposition of non-stationary

signals such as SM. SM is non-stationary because its variance and mean vary in time. For example, mean SM level and variability increases in seasons with more intense and frequent rains. Section 3.2 describes wavelet transforms and section 3.3 describes the datasets used in this chapter. Section 3.4 presents the experiments and is structured in 3 parts: The first one aims to understand the geophysical controls of each time scale, so it is based on a SM dataset generated with a simple temporal model. The second and third experiments analyse measured *in situ* series and satellite series (SMOS). It is worth highlighting that the experiments include small-support datasets (modelled and *in situ*) and large-support datasets (the *in situ* spatial average and SMOS).

3.2 Time scales: wavelet-based approaches

Wavelets are mathematical functions that allow decomposing time series in time scales (Foufoula-Georgiou and Kumar, 1994; Percival and Walden, 2000). In geophysics they have been applied, for example, to decompose vegetation (Blinowska and Durka, 1997) and temperature time series (Craigmile and Guttorp, 2011), to reveal causality in rain (Molini et al., 2010) and to identify climatic connections between different variables, like between SM and temperature (Casagrande et al., 2015) or between the winter Artic Oscillation and the Baltic Sea ice extent (Grinsted et al., 2004), to quote some of them. A review of wavelets in geophysics can be found in Foufoula-Georgiou and Kumar (1994).

The main advantage of wavelet transforms is that they are time-frequency transformations: not only they identify the frequency components of the signal but also when they occur in time (*time localisation* property). This is very useful when analysing non-stationary series like SM. Although wavelets are not the only techniques that provide time-frequency transformations, they are the most adequate. Moving average filters are easy to use but they unsuccessfully separate one band of frequencies from another (Smith, 1997), i.e. they have poor *frequency accuracy*. The short-time Fourier transform also separates time scales of non-stationary signals by dividing the signal in chunks and then applying the Fourier transform to each of them. However it exhibits some limitations related to the fixed width of the chunks: at high frequencies, several cycles are included within the window, providing good accuracy in frequency but bad localisation in time; at low frequencies, there is exactly the opposite effect (Barford et al., 1992). Wavelet transforms solve this issue by allocating larger windows to lower frequencies and shorter windows to higher frequencies. They rely of a set of filters (wavelets) that are dilated with the time scale.

3.2.1 General definition

Wavelets are, as their name indicates, "small waves". They can be described as an oscillation that starts at zero, grows and then decays and comes back to zero. More specifically, a wavelet is a real-value function $\psi(\cdot)$ that satisfies the following properties:

1. The integral of $\psi^2(\cdot)$ is unity: $\int_{-\infty}^{\infty} \psi^2(u)du = 1$
2. The integral of $\psi(\cdot)$ is zero: $\int_{-\infty}^{\infty} \psi(u)du = 0$

By property 1, the wavelet is defined over a finite interval, and by property 2, the wavelet is balanced above and below the horizontal axis.

Although the first references to wavelets appeared in the first half of last century with the works of Haar and Levy, it was not until 1984 that the theoretical expression of the wavelet transform appeared (Grossmann and Morlet, 1984). The continuous wavelet transform (CWT) is expressed as a collection of variables $\{W(\tau, t) : \tau > 0, -\infty < t < \infty\}$, where $\boldsymbol{\tau}$ designates the **time scale** (Equation 3.1). It consists in the convolution of the original signal $x(u)$ with a set of translated and stretched/shrinked versions of the wavelet function $\psi(u)$.

$$W(\tau,t) = \int_{-\infty}^{\infty} x(u)\psi_{\tau,t}(u)du = \int_{-\infty}^{\infty} x(u)\psi\left(\frac{u-t}{\tau}\right)du \tag{3.1}$$

The notion of *time scale* is easily understood with the aid of the Haar wavelet (Haar, 1910). For a given time scale τ, the Haar wavelet is simply a negative pulse followed by a positive pulse along the interval $[-\tau, \tau]$ (Figure 3.1).

For scale $\tau = 1$ and time $t = 0$, Equation 3.1 looks like:

$$W(1,0) = \frac{1}{\sqrt{2}}\int_0^1 x(t)dt - \frac{1}{\sqrt{2}}\int_{-1}^0 x(t)dt \tag{3.2}$$

which is proportional to the difference between the averages over two different periods of length 1: the [-1,0] period and the [0, 1] period. Thus, the Haar-based wavelet transform of a time series at a specific scale τ, $W(\tau,t)$, is a time series representing **differences between averages of periods of length τ**. For other types of wavelets that exhibit different shapes than Haar, a similar interpretation applies: the transformed series represent *weighted* averages computed along periods of length τ or *slightly longer* (Percival and Walden, 2000, pp. 11, 59).

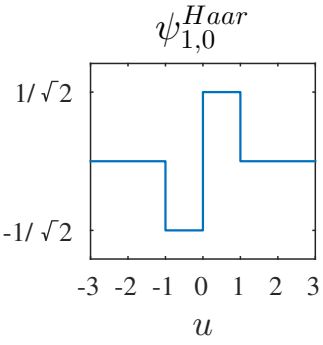


Figure 3.1 – Continuous Haar wavelet at scale $\tau = 1$ and $t = 0$

It should be noted that there exist a difference between the notion of *scale* in the wavelet field and the notion of *period* in the Fourier field. The wavelet transform at scale τ represents differences between periods of length τ. However, the wavelet filter itself expands over a 2τ (or longer) interval so it selects components of the signal that have a characteristic period of length 2τ (or longer). For example, if the scale of interest is 0.5 days, then the output of the wavelet transform will only contain diurnal variations of the signal. This is an approximation to understand how wavelets work. The nominal width of the discrete wavelet filters is larger: $[1/4\tau, 1/2\tau]$ in the frequency domain. This will be described in the next section.

3.2.2 The maximal overlap discrete wavelet transform (MODWT)

The CWT is a continuous time-scale representation, useful for exploratory analyses and to evaluate coherence between processes (e.g. Grinsted et al., 2004). However, it is fundamentally redundant, particularly at large scales were nearby scales show very similar behaviour (Percival and Walden, 2000, p. 13). The discrete wavelet transform (DWT) can be seen as a sub-sampling of the CWT that increases the separation between scales as the time scale increases. **Scales are dyadic**:

$$\tau_j = 2^{j-1} \cdot \Delta t, \qquad j = 1, 2, \ldots, J0 \qquad [s] \tag{3.3}$$

where J0 is the last decomposition level and Δt is the sampling period of the original signal. The DWT has two disadvantages: it needs the length of the original signal to be a power of two ($N = 2^{J0}$) and it is not shift invariant (Percival and Walden, 2000, pp. 159, 160). The shift dependence implies that if $x(t)$ is the original signal and $x(t - t_1)$ a time shifted version, there may exist a significant difference in the energy of their transforms. In practice, this means that the DWT is sensitive to the starting point of the input signal. The maximal overlap discrete wavelet transform (MODWT) is a non-orthogonal DWT that does not exhibit these problems. Moreover, the MODWT produces one transformed coefficient per original sample $x_t \in X(t)$ at all scales. As a consequence, it moderately increases the number of degrees of freedom per scale, thus decreasing the uncertainty of wavelet-based statistics (Cornish et al., 2006).

The MODWT, as the DWT, decomposes a finite sequence $X(t)$ into the J0 levels or scales of Equation 3.3. At scale j, the transformed coefficients consists in a time series of **wavelet coefficients $W_j(t)$** that represent variations of the signal at scale j, and a time series of **scaling coefficients $V_j(t)$** that represent variations at scales larger than j. More precisely, the wavelet coefficients represent weighted differences in averages of consecutive periods of length τ_j while scaling coefficients represent averaged fluctuations at scales larger than τ_j. They are obtained by convolution with the *wavelet* $h_j(t)$ and *scaling* $v_j(t)$ filters as follows:

$$W_j(t) = \sum_{l=0}^{N-1} h_j^{\circ}(l) X(t - l \ mod \ N), \qquad t = 0, 1, \ldots, N-1 \tag{3.4}$$

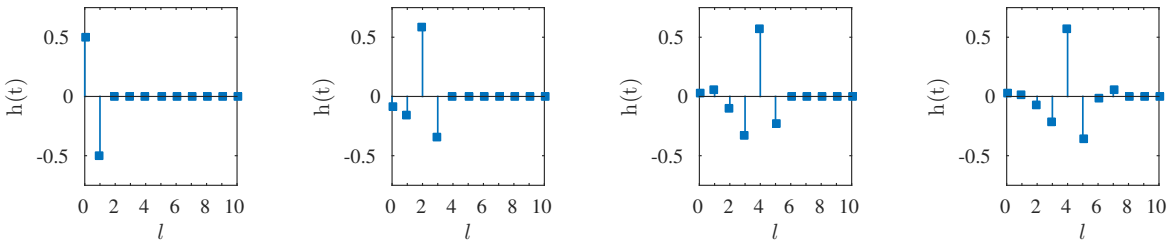
$$V_j(t) = \sum_{l=0}^{N-1} g_j^{\circ}(l) X(t - l \ mod \ N), \qquad t = 0, 1, \ldots, N-1 \tag{3.5}$$

where "$t - l \ mod \ N$" denotes periodic extension of the time series and $h_j^{\circ}(t)$ and ${v_j}^{\circ}(t)$ are the filters obtained by circularising $h_j(t)$ and $v_j(t)$ to length N. Note that, for simplicity of notation, in this thesis I suppress the " $\sim$ " symbol that usually distinguishes MODWT coefficients and filters ($\tilde{W}_j$, $\tilde{h}_j$, ...) from their DWT counterparts (W_j, h_j, ...).

For a maximum level of decomposition J0, the set of wavelet coefficients for scales $j = 1, 2, \ldots, J0$ and the scale coefficients at scale J0 yields an **energy decomposition** of the signal X:

$$\|X\| = \sum_{j=1}^{J0} \|W_j\|^2 + \|V_{J0}\|^2 \tag{3.6}$$

Different types of wavelets can be used to obtain the wavelet and scale coefficients. The most common MODWT wavelet basis functions are shown in Figure 3.2.



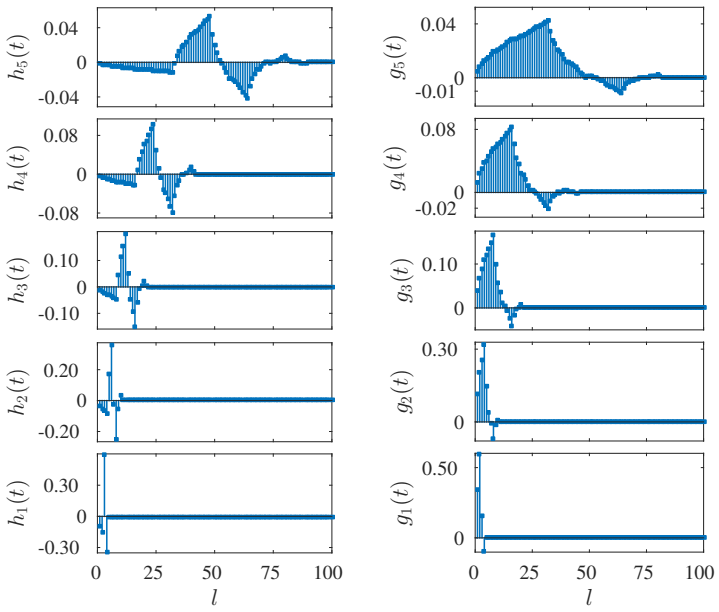
(a) Haar (L=2) (b) Daubechies-4 (L=4) (c) Daubechies-6 (L=6) (d) Least asymmetric-8 (L=8)

Figure 3.2 – Different types of MODWT mother wavelets $h(t)$

The wavelet basis functions need to be dilated at each time scale j to produce the corresponding wavelet and scale coefficients. If the width of the basis function is L, the width of the jth-scale filter is given by

$$L_j \equiv (2^j - 1)(L - 1) + 1 \tag{3.7}$$

Note that L changes with the type of filter (e.g. Figure 3.2) and that, later on, it will be a decisive characteristic when choosing the number of scales $J0$ and the accuracy of the decomposed series. To illustrate how dilation works, Figure 3.3 presents the Daubechies-4 wavelet and scaling filters at different time scales.



(a) Wavelet filters (b) Scale filters

Figure 3.3 – Daubechies-4 wavelet and scale filters at scales 1 to 5

In the frequency domain, the transfer functions for $h_j(t)$ and $v_j(t)$ are given by $H_j(f)$ and $V_j(f)$ and they correspond to band-pass filters and low-pass filters, respectively. The nominal band-pass of $H_j(f)$ is given by $1/2^{j+1} \leq |f| \leq 1/2^j$ and that of $V_j(f)$ is given by $0 \leq |f| \leq 1/2^j$. This allows the full set of coefficients for scales 1 to J0 to be computed using a pyramid algorithm (Percival and Walden, 2000, p. 307). This is illustrated in Figure 3.4. Note that in this figure the transfer functions $H_j(k/N)$ and $V_j(k/N)$ correspond to the circularised $h_j^\circ(l)$ and $v_j{}^\circ(l)$ filters.

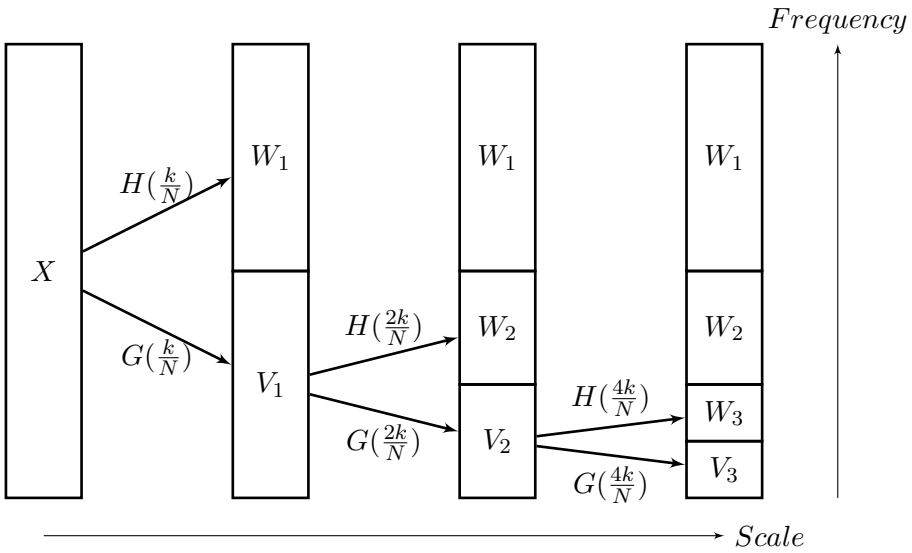


Figure 3.4 – Flow diagram illustrating analysis of X into W_1, W_2, W_3 and V_3 using the pyramid algorithm

The inverse transform of the wavelet and scaling coefficients produce the **detail series $D_j(t)$** and the **smooth series $S_j(t)$**, respectively. In practice, D_j is obtained by filtering W_j with a filter whose transfer function is $H^*(\cdot)$, the complex conjugate of $H(\cdot)$, and S_{J0} is obtained by filtering V_{J0} with a filter whose transfer function is $G^*(\cdot)$, the complex conjugate of $G(\cdot)$. The overall wavelet decomposition process is illustrated in Figure 3.5.

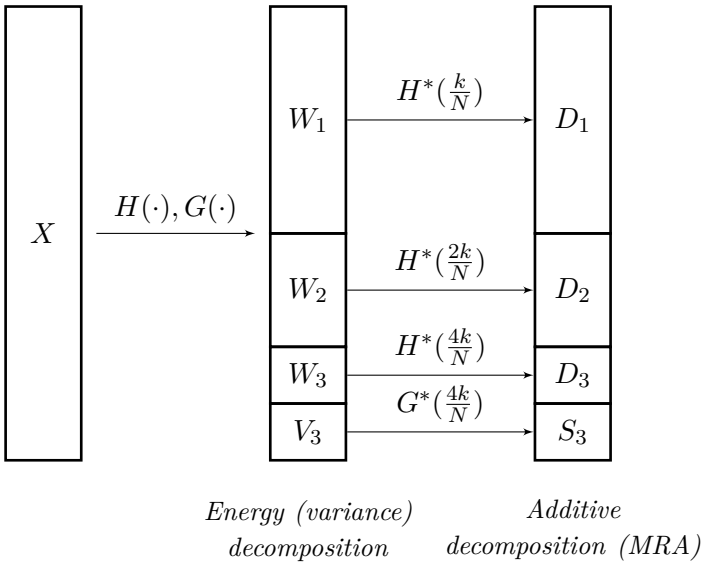


Figure 3.5 – Flow diagram illustrating the wavelet decomposition of a signal X for a maximum decomposition level of 3.

Similarly to W_j, D_j represents variations of the signal at scale j and similarly to V_j, S_j represents variations at scales larger than j. Unlikely the transformed coefficients, the detail and smooth series cannot yield an energy decomposition of the original signal. Instead, they yield an **additive decomposition** or **multi-resolution analysis (MRA)**:

$$X = \sum_{j=1}^{J0} D_j + S_{J0} \tag{3.8}$$

3.2.3 Practical considerations

3.2.3.1 Boundary conditions

When the filtering operation gets closer to the start and the end of the series $X(t)$, a smaller number of samples are available for the convolution. A typical solution for this is to circularise or reflect the series. Circularising is a more appropriate approach when the signal X contains periodic features. This is the case of SM, which can show diurnal cycles due to evapotranspiration cycles, or seasonal cycles, due to seasonal changes in temperature, rain or land cover. The MODWT coefficients that have been computed using part of the circularised signal are referred as the coefficients affected by boundary conditions of simply the *boundary coefficients*.

3.2.3.2 Choice of the base wavelet filter

The choice of the base wavelet filter is a complex task that should take into account three aspects: width, shape and phase. The most common ones are those that approximately represent differences of contiguous weighted averages (Percival and Walden, 2000, p. 105): Haar, Daubechies, Least-asymmetric or "symlets" and Coiflets.

Width. The wider the $h_j(t)$ and $v_j(t)$ filters are in time, the sharper the edges of $H(\cdot)$ and $V(\cdot)$ are in the frequency domain. This implies that the separation between frequency bands is more accurate. However, longer filters produce decompositions with more coefficients affected by boundary conditions and with poorer time localisation. Haar is the shortest filter possible

($L = 2$), followed by Daubechies ($L = 4, 6, \dots$), Coiflets ($L = 6, 12, \dots$) and least-asymmetric (LA) filters ($L = 8, 12, \dots$).

Shape. The shape of the filter affects how the generalized differences between adjacent intervals are computed. When filters with short widths are used, the shape can induce artifacts in the decomposed detail and smooth series at large time scales (for examples, see Percival and Walden, 2000, pp. 130-133). This is critical in DWT decompositions, but the impact is much lower in MODWT decompositions (Percival and Walden, 2000, p. 197).

Phase. Unlike the DWT, the MODWT detail series are not affected by the filter phase characteristics. The MODWT is implemented as a two-stage filter cascade (first $H_j(.)$, then $H_j^*(.)$, see Figure 3.5), so the equivalent filter is $|H_j(.)|^2$ that has a phase function that is always equal to zero, independently from the filter selected (Percival and Walden, 2000, p. 197). For the same reason, MODWT D_j and S_j series are less affected by filter artefacts and ripples than the transformed coefficients series Wj and Vj.

3.2.3.3 Choice of the maximum level of decomposition J0

The largest time scale should be smaller than the length of the series:

$$2^{J0-1} \leq N \tag{3.9}$$

This is natural since, for example, it would not make sense to look for seasonal signatures in 1-month series. A more restrictive condition can be applied to ensure that the filter does not extends beyond the series limits:

$$L_{J0} \leq N \tag{3.10}$$

where L_{J0} is computed using Equation 3.7. Shorter filters, like Haar or Daubechies, allow larger $J0$ values.

3.2.4 Wavelet-based statistics

Variance

There are two types of estimators of the wavelet variance: while the *unbiased* estimator does not consider wavelet coefficients affected by boundary conditions, the *biased* estimator takes all the coefficients into account. The *biased* estimator will be chosen in this thesis because it leads to an exact decomposition of the signal variance (the sum of the wavelet variances at each scale gives the total signal variance) (Percival and Walden, 2000, pp. 306, 307).

The **decomposition of the signal variance** is derived from the energy decomposition of Equation 3.6:

$$\hat{\sigma}_X^2 = \frac{1}{N} \sum_{j=1}^{J0} \|W_j\|^2 + \left(\frac{1}{N} \|V_{J0}\|^2 - \hat{\mu}^2 \right) \tag{3.11}$$

where $\hat{\mu}$ is the sample mean and the terms between parentheses are equal to the sample variance of the scale coefficients, $\hat{\sigma}_{V_{J0}}^2$, i.e. the variance of scales larger than J0. The sample variance is thus the sum of the variances at each time scale:

$$\hat{\sigma}_X^2 = \sum_{j=1}^{J0} \hat{\sigma}_{W_j}^2 + \hat{\sigma}_{V_{J0}}^2 \tag{3.12}$$

For simplicity, hereafter $\hat{\sigma}_{W_j}^2$ will be referred as $\hat{\sigma}_j^2$. The calculation of the confidence intervals (CIs) for the sample variance is described in Appendix B.

Correlation

A similar decomposition to that of the variance applies to the covariance of two time series X and Y.

$$\hat{\sigma}_{X,Y} = \sum_{j=1}^{J0} cov\{W_{X,j}(t), W_{Y,j}(t)\} + cov\{V_{X,J0}(t), V_{Y,J0}(t)\} \tag{3.13}$$

However, the correlation between X and Y cannot be expressed as the sum of the correlations at each time scale:

$$R_{X,Y} = \frac{\hat{\sigma}_{X,Y}}{\sigma_X \sigma_Y} \neq \sum_{j=1}^{J0} \frac{cov\{W_{X,j}(t), W_{Y,j}(t)\}}{\sigma_{X,j}\sigma_{Y,j}} + \frac{cov\{V_{X,J0}(t), V_{Y,J0}(t)\}}{\sigma_{V_{X,J0}}\sigma_{V_{Y,J0}}} \tag{3.14}$$

Given that the correlation cannot be decomposed in time scales, we consider that there is no gain in using the transformed coefficients W_j and V_{J0} for its calculation. Moreover, as explained in section 3.2.3.2, the D_j and S_j series are less affected by filter artefacts than the W_j and V_j coefficients, while being similar in terms of interpretation. For these reasons, hereafter the *wavelet-based correlation* will refer to the correlation between detail series and not between wavelet coefficients:

$$R_j = R\{D_{X,j}(t), D_{Y,j}(t)\} \tag{3.15a}$$

$$R_{S_{J0}} = R\{S_{X,J0}(t), S_{Y,J0}(t)\} \tag{3.15b}$$

$$R_{X,Y} \neq \sum_{j=1}^{J0} R_j + R_{S_{J0}} \tag{3.15c}$$

The calculation of the CIs for the wavelet-based correlation is described in Appendix C.

3.3 Datasets

The surface SM time series analysed in this chapter belong to 3 different groups of datasets. The first consists of time series generated with my own adaptation of the model designed by Pan et al. (2003). This first dataset will help to understand the relationship between SM drivers (precipitation, surface, climatology) and time scales. The second set includes time series from 4 dense *in situ* networks (Little Washita, Walnut Gulch, Yanco and Monte Buey). The third dataset consists of SMOS observations at the same locations.

3.3.1 Modelled series (Little Washita)

SM series generated with the model of Pan et al. (2003) will serve to explore the possible links between SM time scales and related geophysical variables. The model was specifically designed for the Little Washita (LW) network and has been used with improved calibrations in other studies (e.g. Coopersmith et al., 2016; Coopersmith et al., 2017). It is of interest here for the following reasons:

(i) it produces realistic point-support SM series
(ii) it is simple and is only based on 3 input variables: precipitation (p), soil texture (Ks) and vegetation (LAI)
(iii) the input variables are associated with specific SM temporal dynamics: precipitation with short time scales (impulse responses), texture with surface memory responses (dry-down decays) and LAI controls the seasonal behaviour of the SM series
(iv) its simplicity should allow detecting correspondences between the time scales of the modelled SM and the input variables

For a given location, the model of Pan et al. (2003) can be expressed as

$$SM(t) = min\left\{max\left\{SM(t-1)\cdot e^{\frac{-\eta_l(t)\cdot\Delta t}{Z}} + \frac{\gamma_c\cdot p(t)}{\eta_l(t)\cdot\Delta t}\cdot\left[1-e^{\frac{-\eta_l(t)\cdot\Delta t}{Z}}\right], SM_{min}\right\}, SM_{max}\right\} \quad (3.16)$$

where Δt is the sampling period in hours, $p(t)$ is the cumulative precipitation in metres (m) between $t-1$ and t, γ_c is the interception by vegetation, η_l is the loss coefficient (m/h) and Z is the penetration depth (m). The loss coefficient depends on the drainage coefficient Ks (cm/h) and the LAI (m^2/m^2).

The model was specifically designed for summer, so specific calibration is needed when applied to different periods (Pan, 2012; Coopersmith et al., 2014). Here, the calibration sought to minimise the square of the differences with station #1 of the LW network for a 2-year period during which the station was fully functional (2012/7 - 2014/7). The γ_c parameter underwent a special temporal calibration that allowed to improve the fit with the measured SM series. The resultant γ_c is not physically plausible, but it ensures that the modelled SM series is close enough to the reality. Table 3.1 describes the calibration of the parameters and Figure 3.6 shows the model time series, the adjusted γ_c parameter and the LAI time series at station #1.

Table 3.1 – Own calibration values for the parameters of the model of Pan et al. (2003)

Parameter	Units	Value
γ_c	-	See figure 3.6
x_{min}	(m^3/m^3)	0.05
x_{max}	(m^3/m^3)	0.40
η_l	$(m/year)$	$= max\{0.1\cdot Ks + 0.5\cdot LAI), 0.5\}$ [*1]
Z	(m)	$= Z_{max} - r_{mod}\cdot(Ks - Ks_{,min})$ [*2]
Ks	(cm/h)	1.3[*3] for loam, 4 for sandy loam, 5[*3] for sand

[*1] Equation different from Pan et al. (2003).
[*2] Equation empirically derived to allow a wider range of decay rates: $r_{mod} = \frac{Z_{max}-Z_{min}}{K_{s,max}-K_{s,min}} = \frac{0.09-0.025}{5-0.05}$
[*3] Source: FAO, ftp://ftp.fao.org/fi/cdrom/fao_training/FAO_Training/General/x6706e/x6706e09.htm

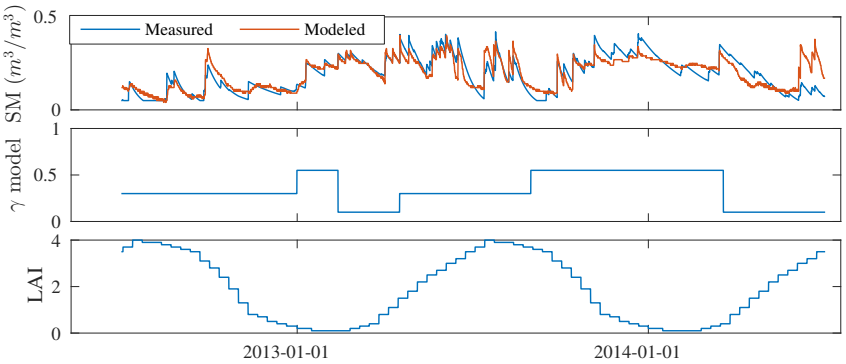


Figure 3.6 – Modelled and measured SM series at station #1 (top), with the adjusted γ_c parameter (middle) and the LAI series (bottom) at the same location.

A set of 2x9 modelled SM series is generated with the same calibration: two reference series, *ref-a* and *ref-b*, and two sets of 8 sample series. The *ref-a* series is generated for a loam texture using the *in situ* precipitation (p_{ref}) and the MODIS LAI series observed at station #1 (LAI_{ref-a}). The *ref-b* series is produced identically except for the seasonal component whose amplitude is reduced ($LAI_{ref-b} = LAI_{ref-a}/4$). In the experiments that will be presented later, each reference series is compared against a set of 8 sample series that differ from the reference in the configuration of one of the 3 variables. The goal is to generate different controlled dynamics in plausible SM series so that I can analyse how these controls are detected with wavelet decompositions.

The detailed setup is provided in Table 3.2 and the modelled SM series are shown in Figure 3.7. In short, sample series 1.x and 2.x differ in precipitation, sample series 3.x differ in texture and sample series 4.x differ in the synchronisation of the LAI signal.

Table 3.2 – Characteristics of the modelled series of Little Washita. Only the differences with respect to the respective reference series (*ref-a* or *ref-b*) are indicated

Series	Variable changed	Description	
1.1-a/b	Precip.	**Moderate precip. noise**: $p_{1.1}(t) = p_{ref}(t) + \epsilon_1(t)$ where $\epsilon_{mean}(t) \sim N(0, \sigma_{mean})$	[*1,*2]
1.2-a/b		**Maximum precip. noise**: $p_{1.2}(t) = p_{ref}(t) + \epsilon_1(t)$ where $\epsilon_{max}(t) \sim N(0, \sigma_{max})$	[*1,*2]
2.1-a/b	Precip.	**Precip. from station #6**: $p_{2.1}(t) = p_{i=6}(t)$	[*3]
2.2-a/b		**Precip. from station #16**: $p_{2.2}(t) = p_{i=16}(t)$	[*3]
3.1-a/b	Texture	**Sand**: $Ks_{3.1} = 5$	
3.2-a/b		**Sandy-loam**: $Ks_{3.2} = 4$	
4.1-a/b	LAI	**1-month shift**: $LAI_{4.1}(t) = LAI_{ref}(t - 30 \cdot 24)$	
4.2-a/b		**3-month shift**: $LAI_{4.2}(t) = LAI_{ref}(t - 3 \cdot 30 \cdot 24)$	

[*1] σ^2_{mean} and σ^2_{max} are the mean and maximum of the differences between the precipitation variance $\sigma^2_{p_1}$ of station #1 and all the other stations $\sigma^2_{p_i}$

[*2] The noise contribution $\epsilon_{mean}(t)$ and $\epsilon_{max}(t)$ is zero when $p_{ref}(t)$ is zero

[*3] There are non-synchronised precipitation events in stations #6 and #16 with respect to station #1

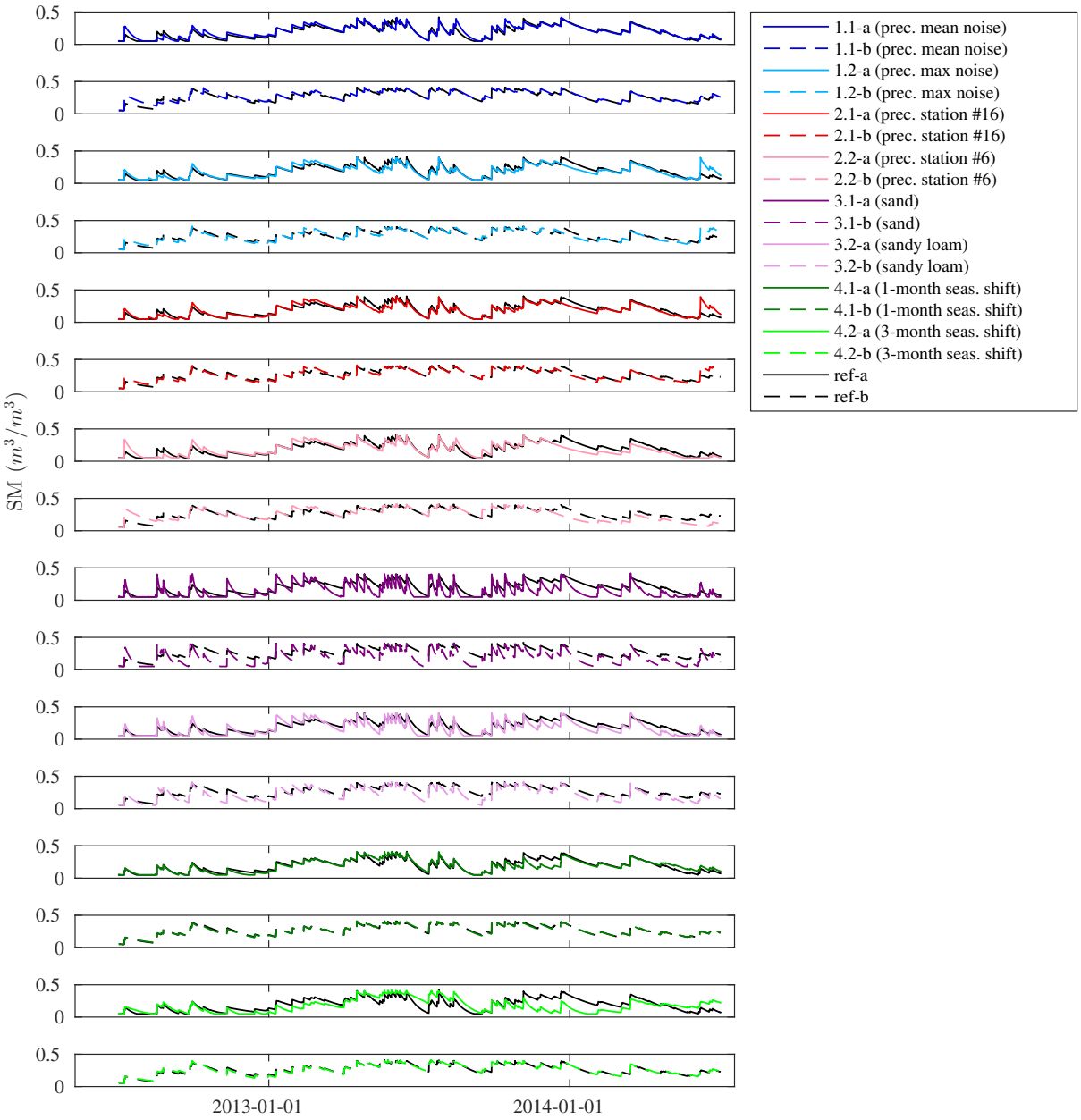


Figure 3.7 – Little Washita modelled series for the 2012/07/03 - 2014/07/02 period

3.3.2 *In situ* measurements

The *in situ* SM measurements from 4 dense networks are used in this chapter: Little Washita (LW), Walnut Gulch (WG), Yanco and Monte Buey (MBy). The first 3 networks have been already used in chapter 2 (section 2.4.4.1) so only the MBy network will be described below. The main characteristics of the networks are reminded in Table 3.3. The position of the stations and the land cover maps of the four networks are showed in Appendix D.

Monte Buey (MBy) The Monte Buey (MBy) network covers an area of 36 x 36 km^2 in the Humid Pampas region, in Argentina. The climate is temperate with average annual precipitation of 1000 mm. The area is totally flat and dedicated to non-irrigated farming and cattle. Fields are covered by wheat in winter and soya or maize in summer. Some areas can remain fallow and some plots contain alfalfa. Soil texture is homogeneous (loam to silty loam). The region is characterized by a water excess: changes in land use and the increase in precipitations in the last 10 years have raised the water table to 1-2 m below the surface. As part of my thesis, I

Table 3.3 – Main characteristics of the *in situ* neworks used in this chapter.

	Little Washita (LW)	Walnut Gulch (WG)	Yanco	Monte Buey (MBy)
Extent	610 km^2	148 km^2	3,000 km^2	1296 km^2
SMOS pixels covered	∼1/4	∼1/16	∼1	∼1
Climate	Sub-humid	Semi-arid / arid	Semi-arid	Humid
Annual precipitation	750 mm	324 mm	400 mm	1,000 mm
Main precip. periods	Autumn, spring	Summer	Winter, spring	All year
Soils	Diverse	Sands and gravel	Silty-loam	Loam, silty-loam
Topography	Moderately rolling	Rolling	Flat	Flat

participated in one calibration and validation (Cal/Val) campaign in Monte Buey in November 2016 (spring). I could witness the water excess status of the area, which is illustrated here in two pictures taken during the campaign (Figure 3.8).

The network is made of 16 Stevens Hydraprobe stations, installed horizontally at 5 cm depth. All stations, except for two, are located next to a plot, separated by 1 to 3 meters from the cropping area. Stations #14 (Los Pinos) and #8 (AAXOD) are located in the middle of the plot but protected by a fence. MBy stations present long periods of unavailability as a result of sensor faults. Despite that the network started operating in 2012, only the 2-year period from 2015/01 to 2016/12 has been selected so that the number of available stations is maximised (13 out of 16 stations).

Figure 3.8 – Pictures taken at Monte Buey during the Cal/Val campaing in November 2016 that illustrate the water table level (left) and the trenches built for water drainage (right)

3.3.3 SMOS data

SMOS soil moisture time series retrieved at the *in situ* networks locations are also analysed. For the experiments in this chapter, I use the ESA Level-2 (L2) soil moisture product (version 620), which exhibited better temporal stability in time than the L3 SM product. These temporal stability issues are currently under investigation at the SMOS SM expertise centre, the Centre d'Études Spatiales de la Biosphère (CESBIO).

The characteristics of the SMOS mission have been already described in section 2.4.2.1. In short, the resolution of the SM product is considered 40 km on average and the maximum revisit time of the satellite is 3 days with crossing nodes at 6 a.m. and 6 p.m. local solar time for ascending and descending orbits, respectively.

The L2 SM retrieval algorithm is described in detail in (Kerr et al., 2012, 2014). One of the key aspects of the retrieval algorithm is that it takes into account the landscape heterogeneity of the observed surface. The surface is divided in fractions according to the land cover (bare soil, low vegetation, forest, water, etc.). Each type of land cover is characterised by a specific model

linking brightness temperatures (TBs) and SM. SM is only retrieved for the largest or *dominant* fraction while the other fractions contribute with fixed parameters. In the *in situ* networks presented here, the dominant fraction is low-vegetated soil and SM is obtained by inversion of the L-band Microwave Emission of the Biosphere (L-MEB) model (Wigneron et al., 2007). SM is iteratively retrieved by minimising the distance between the modelled and the acquired TB. The iteration requires a initial guess for SM, among other variables, which is obtained from the European Centre for Medium-range Weather Forecasting (ECMWF) operational forecast system.

3.3.4 Data preparation and wavelet parameters

The wavelet analysis of SM series requires some conditions to be fulfilled. First, the series should not present any gap. Secondly, the length of the series should ensure that its circularisation (see section 3.2.3.1) respects the signal periodic signatures, if they exist. Additionally, the sampling period must ensure that the derived wavelet scales ($\tau_j = \Delta t \cdot 2^{j-1}$) make sense from a geophysical perspective (e.g. they include daily, seasonal and yearly cycles). Finally, all the series must be sampled with the same sampling period so they are comparable. The following paragraphs describe the wavelet-related choices and pre-processing applied to the SM series used in this chapter.

3.3.4.1 Periods of analysis, sampling periods and wavelet scales

The LW modelled series cover a 2-year period (2012/7 - 2014/7) that has been selected so that the SM series measured at the reference station #1 does not present any important gap. The selected period contains complete years so that circularisation can be used in wavelet analyses (seasonal variations, if they existed, would be respected).

Unlike the modelled series, the *in situ* time series usually present gaps. The periods of analysis are different for each network and they are selected trying to minimise the number of gaps and to maximise the number of stations in operation. Periods containing complete years are again preferred . Based on this, LW series spread 3 years (2011/5 - 2014/4), WG 5 years (2010/6 - 2014/5), Yanco 1 year (2010/6 - 2011/5) and MBy 2 years (2015/1 - 2016/12). The same series will be selected for the SMOS series. The list of periods classified per dataset can be found in Table 3.6 and Table 3.7.

The *in situ* networks measure surface SM at different sampling intervals: LW and WG measurements were taken every 30 minutes and Yanco and MBy every hour. The first two experiments (sections 3.4.1 and 3.4.2) concern point modelled and *in situ* series and the common $\Delta t = 3$ h can be set for all of them (the *in situ* series must be downsampled). The 3rd experiment (section 3.4.3) compares SMOS and *in situ* series and Δt is set to 12 h so that the minimum time interval between SMOS ascending and descending overpasses is respected. The advantage of these sampling periods is that they allow a set of wavelet scales that are coincident for all the experiments and that approximate weekly, monthly and seasonal cycles. The lists of scales for both sampling periods are provided in Table 3.4 and Table 3.5.

3.3.4.2 Wavelet filters and maximum level of decomposition J0

Given that the selected periods are longer or equal to 1 year, the maximum level of decomposition is set to $2^7 = 128$ days. This is equivalent to J0 = 11 when Δt is 3 h and to J0 = 9 when Δt is 12 h. Regarding the wavelet filter, it should allow to locate events in time with precision so that correlation analyses can be performed. Therefore, the smallest length filters, Haar (Haar, 1910) and Daubechies-4 (Daubechies, 1992), should be selected. Also because of the length, Daubechies-4 has a sharper frequency response than Haar so it better isolates time scales. It

Table 3.4 – Wavelet scales for time series sampled at $\boldsymbol{\Delta t = 3}$ **h** (*in situ* and modelled series)

Time scale j	Time scale (days) $\tau_j = 2^{j-1}\Delta t$	Time scale j	Time scale (days) $\tau_j = 2^{j-1}\Delta t$
1	0.125	7	8
2	0.25	8	16
3	0.5	9	32
4	1	10	64
5	2	11	128
6	4		

Table 3.5 – Wavelet scales for time series sampled at $\boldsymbol{\Delta t = 12}$ **h** (*in situ* and SMOS series)

Time scale j	Time scale (days) $\tau_j = 2^{j-1}\Delta t$	Time scale j	Time scale (days) $\tau_j = 2^{j-1}\Delta t$
1	0.5	7	32
2	1	8	64
3	2	9	128
4	4		
5	8		
6	16		

will be preferred to Haar when its length at the last scale J0 is shorter than the length of the series N (Equation 3.10). The length of the Daubechies-4 filter at the last scale is 767 days and that of the Haar filter is 256 days. For this reason, series longer than 2 years will be analysed with the Daubechies-4 wavelet and shorter series will be analysed with the Haar wavelet. These choices are summarised in Table 3.6 for the modelled series and in Table 3.7 for the measured *in situ* and SMOS series.

3.3.4.3 Pre-processing of *in situ* series

Despite the careful choice of the periods of analysis, the *in situ* series still showed gaps. Those exhibiting long periods of unavailability are removed from the set. The remaining gaps are filled in a two step process: 1) Big gaps are filled by linear regression with the station series that was closer in terms of correlation. The minimum length of a "big gap" is set arbitrarily to 40 days for LW and WG series and to 20 days for Yanco and MBy series. 2) The remaining "small" gaps are filled with a discrete cosine transform (DCT) approach (Wang et al., 2012). The advantage of the DCT gapfilling is that it uses the full series (its signal spectrum) and not only local information to estimate the missing data. Table 3.7 describes the selected periods of analysis, stations and gap-filling percentages for the series of each *in situ* network.

3.3.4.4 Pre-processing of SMOS time series

The SMOS L2 soil moisture product is provided on the 15-km ISEA-4H9 (Icosahedral Snyder Equal Area Earth fixed) grid. For each *in situ* network, the L2 node that is in the centre of the network is selected. The identifiers of the selected nodes are indicated in Table 3.7. Ascending and descending orbits are merged in one single time series, so the final sampling period is $\Delta t = 0.5$ days. Retrievals with probability of radio-frequency interference (RFI) higher than 10 %, and data quality index (DQX) higher than 20 % are removed. Gaps are filled with the DCT method (Wang et al., 2012). The Monte Buey network was omitted from the analysis because an unknown RFI source impaired the retrievals from June 2015 onwards so it covered a 75 % of the selected period. The pre-processing of SMOS series at each network is summarised in Table 3.7.

Table 3.6 – Pre-processing and wavelet parameters of the modelled point SM series

	Little Washita (LW) modelled series
Period	2012/7 - 2014/7 (2 years)
Wavelet type	Haar
J0	11 (128 days)
Δt	3 h

Table 3.7 – Pre-processing and wavelet parameters of the *in situ* and SMOS soil moisture series.

	Little Washita (LW)	**Walnut Gulch (WG)**	**Yanco**	**Monte Buey (MBy)**
Period	2011/5 - 2014/4 (3 years)	2010/6 - 2014/5 (4 years)	2010/6 - 2011/5 (1 year)	2015/1 - 2016/12 (2 years)
Wavelet type	Daubechies-4	Daubechies-4	Haar	Haar
J0	9 or 11 (128 days)	9 or 11 (128 days)	9 or 11 (128 days)	9 or 11 (128 days)
***In situ* data**				
Original Δt	0.5 h	0.5 h	1 h	1 h
Final Δt	3 h or 12 h	3 h or 12 h	3 h or 12 h	3 h or 12 h
Stations (selected/total)	20 / 20	17 / 20	11 / 16	13 / 16
% gap-filling (regression)	4.5 %	5.7 %	16 %	10.5 %
% gap-filling (DCT)	1.4 %	3.5 %	14.5 %[*1]	30 %[*2]
SMOS data				
Node	226157	215311	8174767	-
Δt	12 h	12 h	12 h	-
% gap-filling (DCT)	58.1 %	55.6 %	54.7 %	-
Nb. consecutive gaps (on avg.)	2.8	2.3	2.3	-
Nb. consecutive samples (on avg.)	2	2.9	2.7	-

[*] Although the percentage of gaps is high, they are distributed evenly so the quality of the DCT gapfilling is not hampered. The distributions of gaps and samples are:
[*1] 1.9 consecutive gaps every 13.8 samples (on average)
[*2] 2.1 consecutive gaps every 5.4 samples (on average)

3.4 Decomposition of soil moisture series

3.4.1 Decomposition of point modelled series

In this section, I investigate how SM time scales are influenced by different drivers of SM variability: forcing events, surface memory and seasonal signals. This will be done through the wavelet decomposition of the Little Washita (LW) modelled series, which have been described in section 3.3.1 and whose wavelet parameters have been summarised in Table 3.6. The experiment will help interpret real SM measured series in the following sections.

The MRA decomposition of the reference series *ref-a* is shown in Figure 3.9. The unitless scale of the detail D_j and smooth S_{J0} series is indicated on the left and the respective time scale is indicated on the right. Vertical red lines indicate the samples affected by boundary conditions. The reduction on the number of non-boundary samples at scale 11 is striking. When interpreting this scale, it should be kept in mind that the boundary samples are computed with replications of the available input time series: they are representative of the available SM series and not of the infinite unknown SM signal.

Despite the large number of samples affected by boundary conditions, the last scales in Figure 3.9 reflect well the typical seasonal changes in the region. For example, both the 64- and the 128-day scales reveal that summers are the driest seasons. The most humid periods are not exactly synchronised every year, but in general they appear during autumn and spring, which matches with the climatological description of the area. In addition, the first scales (0.125 to 2 days), which are almost not affected by boundary conditions, reveal well the precipitation events that drive SM spikes. D_1 to D_5 series are mostly zero except when a spike in the original SM series is detected.

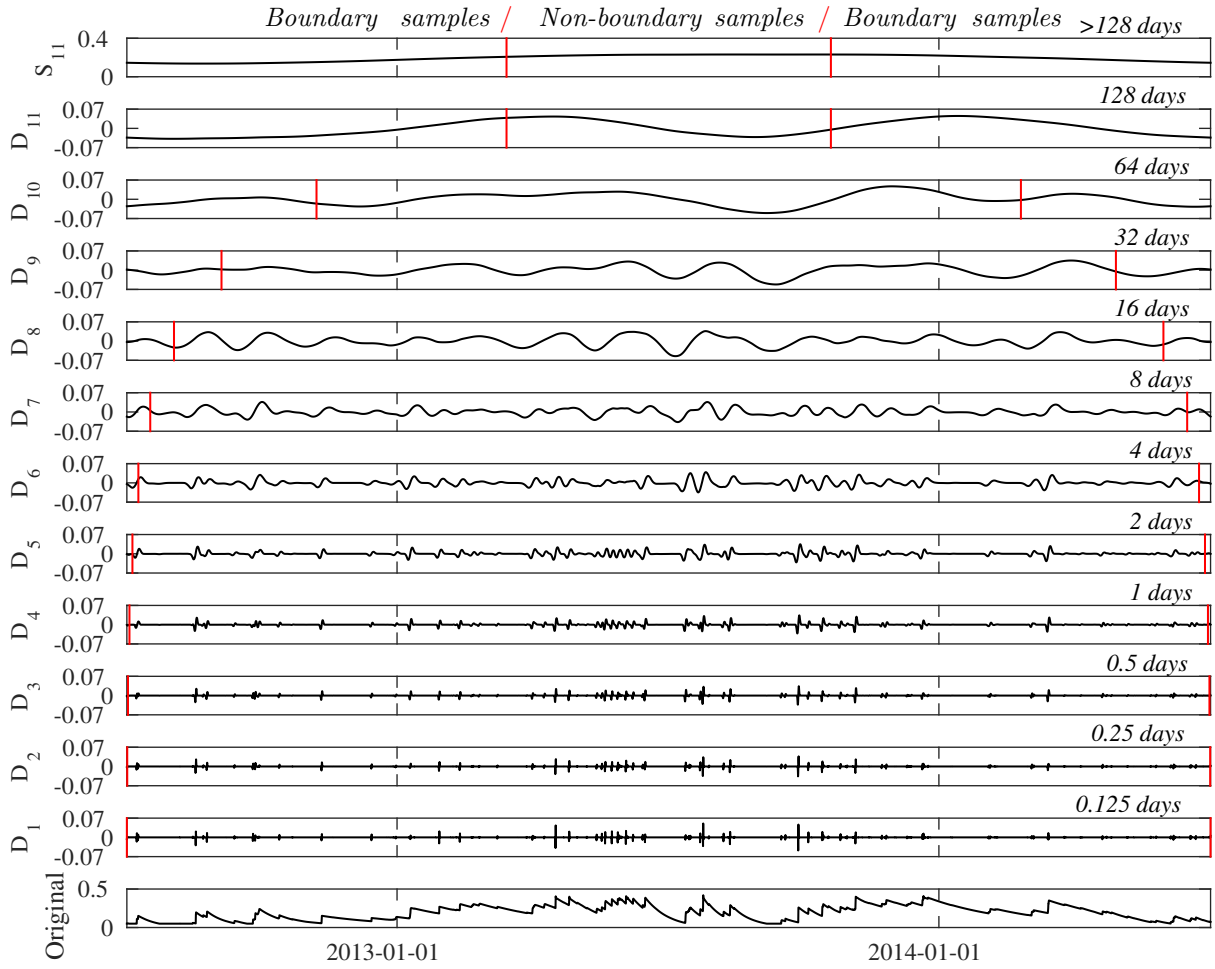


Figure 3.9 – Haar-based multi-resolution analysis (MRA) of the reference *ref-a* series of the Little Washita modelled dataset for the 2012/07/03 - 2014/07/02 period

Figure 3.10 shows the **wavelet variance** of the sample series (dot markers) with respect to the reference series (cross markers). The only difference between the series of the *a* and the *b* groups was that the LAI (the seasonal component) of the latter had a smaller amplitude. This is reflected in the smaller variance of *b*-type series at the 64- and 128-day scales with respect to the *a*-type series. The difference is reduced when the texture changes (Figure 3.10-III), which could signify that different decay rates influence the seasonal persistence. Regarding the precipitation differences generated with the model, they do not change much the variance levels (Figures 3.10-I and -II). Finally, surface memory and seasonal drivers are clearly the components that alter the shape of the variance pattern (Figures 3.10-III and -IV).

The **correlations** between *ref-a* series and series *1.x-* to *4.x-a* and between *ref-b* series and series *1.x-* to *4.x-b* are shown in Figure 3.11. Correlation patterns are far more sensitive to differences in **precipitation**. Two distinct behaviours can be noticed. When precipitation events are synchronised but vary in power (Figure 3.11-I), correlation increases with the time scale. However, when some of the precipitation events are not synchronised (Figure 3.11-II), the positive trend can be broken. For example, although station #6 apparently only misses a couple of precipitation events (e.g. the date 2014/01/01 of the series 2.2 in Figure 3.7), the repercussion in the wavelet correlation diagram is very important: correlation drops at the 16- and 32-day scales. The analysis of precipitation recorded at other locations also showed similar patterns (not shown here), implying that non-synchronised precipitation events can produce "irregular" correlation patterns whose drops cannot be easily predicted. The last scale (128 days) is rarely degraded by this kind of situations.

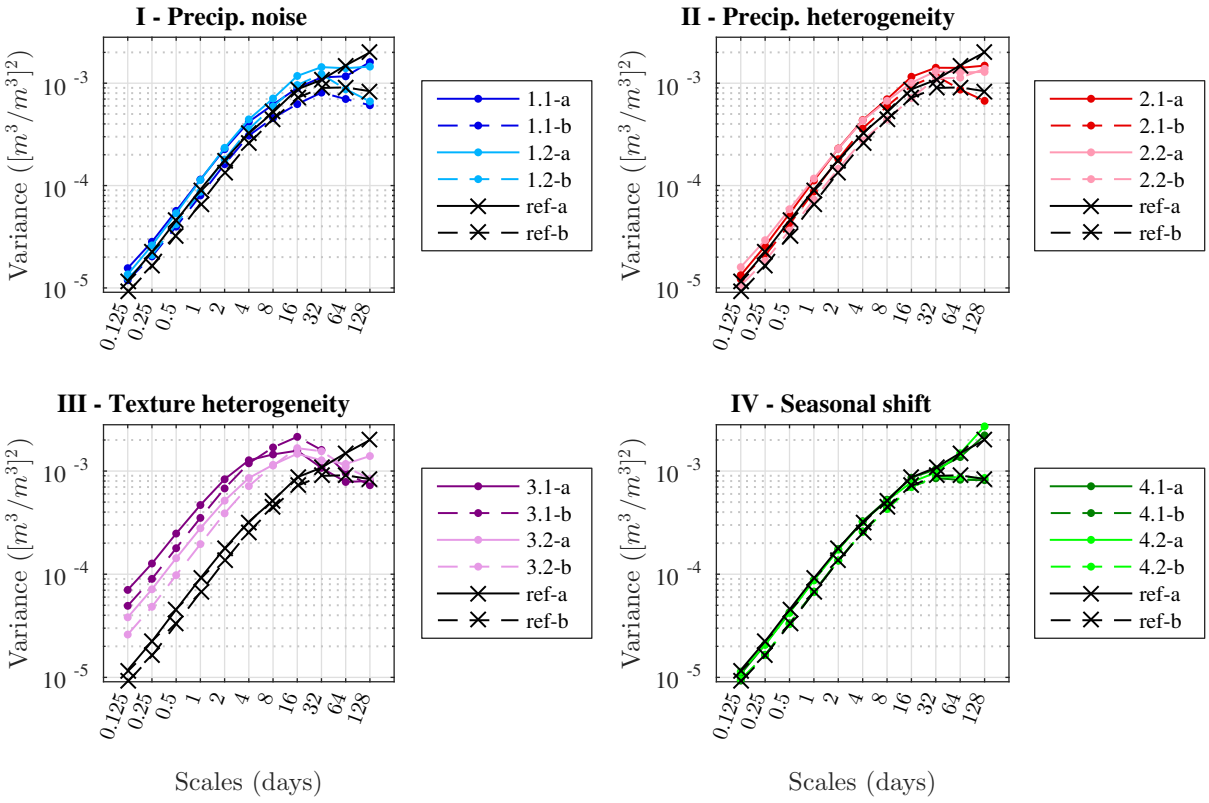


Figure 3.10 – Wavelet variance of the modelled series.

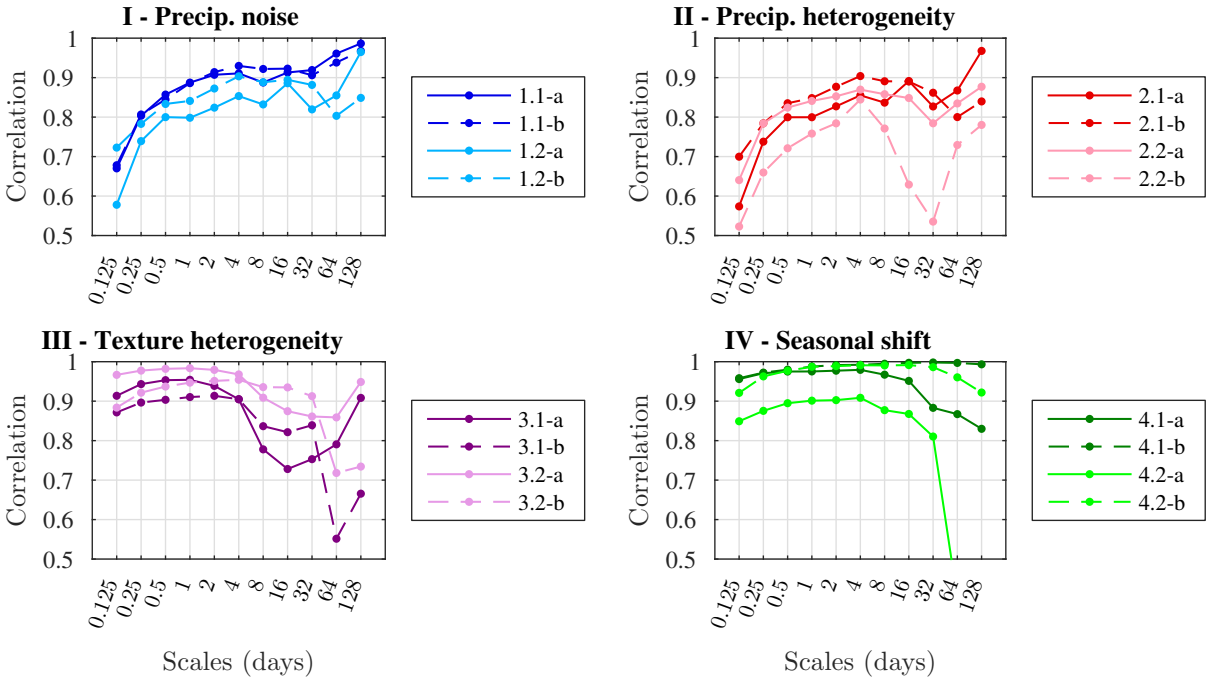


Figure 3.11 – Wavelet-based correlation between the modelled sample and reference series.

Differences in texture seem to affect larger time scales than precipitation. In Figure 3.11-III, the reference series (loam) is compared against sand- and sandy loam-based series (series 3.1 and 3.2, respectively). Since sands produce faster dry-down decays than sandy loam textures, decorrelation appears at shorter time scales for sandy and for sandy-loam textures. In any case, it can be considered that the decrease in correlation due to different decay rates starts at weekly scales (8 days) or longer. Whether it affects or not the seasonal 128-day scale depends on the texture and the power of the seasonal component: when the contribution of the seasonal component is not important (series 3.1-b and 3.2-b), the correlation between decomposed series at that scale is lost. Otherwise, the seasonality is recovered (cases 3.1-a and 3.2-a).

Thanks to further tests (not shown here), I noticed that the combination of both texture and precipitation differences can impair the 128-day scale even if the seasonal driver is important. This confirms the theories of Hasselmann (1976) and Delworth and Manabe (1988) on climatic and SM variations. Hasselmann (1976) suggested that climatic long-term variations are induced by high-frequency variations. Similarly, Delworth and Manabe (1988) proposed that the long-term anomaly of soil moisture is the response of the land surface layer to precipitation high-frequency forcings. However, the experiment here tends to indicate that there is a mixture of low and high-frequency effects at large time scales: the seasonal time scales (64-128 days) are made by two components, pure seasonal signals (vegetation growth, temperature trends, etc.) and the temporal integration of shorter responses (precipitation and surface memory). Most likely, there must exist a third component related to static fields (e.g. topography), but it has not been included in this study.

Finally, Figure 3.11-IV compares SM series with **different seasonal drivers**. There is no implication when the seasonal factor is weak (cases 4.1-b and 4.2-b). However, when there is a important seasonal driver (cases 4.1-a and 4.2-a), correlation at the monthly and seasonal scales decreases, and the magnitude of the decrease highly depends on the degree of lack of synchronism of the seasonal drivers.

The conclusions of the analysis of modelled SM series are:

1. The factors driving SM variability control variations at different time scales: precipitation at small time scales ($\leqslant$ 2 days), surface memory at middle time scales (4-64 days) and seasonal drivers at large time scales (64-128 days).
2. The seasonal scales (64-128 days) are made of two main components: standalone seasonal signals (vegetation growth, temperature trends) and the temporal integration of shorter signals (precipitation, surface memory). Static components like topography could be a third component.
3. Correlation increases with the time scale in the absence of strong heterogeneities of precipitation, texture and seasonal drivers.
4. Non-synchronised precipitation events can produce irregular correlation patterns.
5. Important differences in texture and seasonal drivers produce different variance signatures.

3.4.2 Time-scale decomposition of *in situ* series

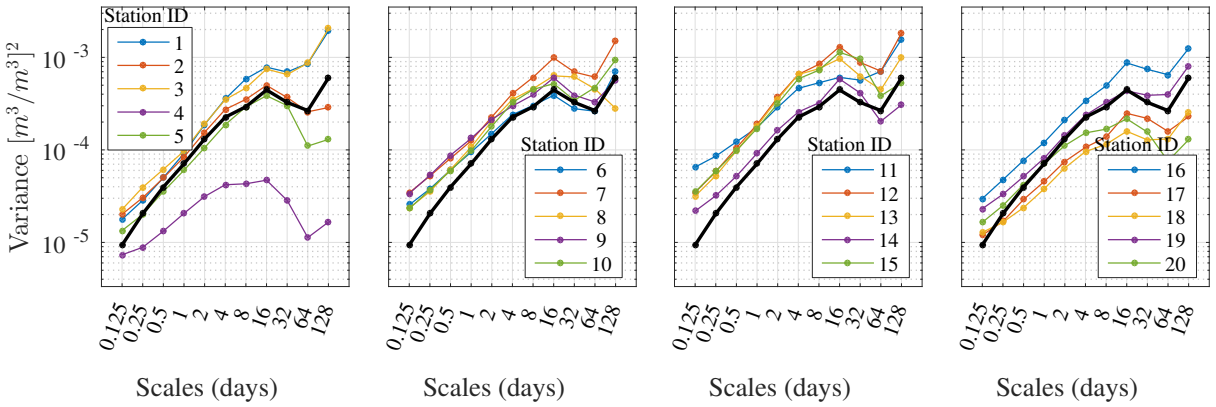
In this second experiment, real *in situ* SM data are analysed with wavelet decompositions. The conclusions from section 3.4.1 will help interpreting the results obtained. The *in situ* data is obtained from the 4 dense networks presented in section 3.3.2. The main characteristics of the networks and of the *in situ* data can be found in Table 3.3 and Table 3.7, respectively. The analysis is also based on the wavelet variance and the wavelet-based correlation as in the previous section. For readability purposes, CIs are not shown here. For further details, Appendix B and C include the statistics with their respective CIs for the *in situ* series of the Little Washita network.

Figure 3.12 presents the **variance analysis** of the *in situ* data, where each row corresponds to one of the 4 dense networks and each plotted line is the variance decomposition of a particular station series of the corresponding network. The 4 networks follow similar general variance laws: the temporal variance increases with the time scale until it reaches a maximum. It is relevant that each network has a distinct variance pattern: the variance increasing rate, the position of the maximum and the evolution after the maximum are different depending on the network. The variance pattern seems to operate as a characteristic signature of the SM process of the respective area and seems related to their climatic and environmental conditions. WG variance (Figure 3.12b) peaks at the 64-days scale, implying that the strongest differences appear when contiguous periods of 2 to 4 months are compared. This peak is clearly driven by the contrast between the summer and spring/autumn months: precipitation in the area consists in high-intensity thunderstorms concentrated in summer. On the contrary, LW (Figure 3.12a) exhibits less variation between seasons and more pronounced variation between half-year periods (128-day scale). The peak at the 16-day scale is a consequence of precipitation being usually absent for 2-week periods (see scale 1 in Figure 3.9) and a good persistence of SM levels after precipitation events. Irrigation practices predetermine the variance pattern of Yanco (Figure 3.12c): the seasonal components (64- and 128-day scales) have usually less power than smaller scales and the variance peak around the 4- to 16-day scales may be related to irrigation periods. Finally, the variance decomposition of MBy series (Figure 3.12d) should be representative of a typical humid area with no important seasonal changes where the surface remains mostly wet all the year. This variance pattern is consistent with that obtained for the modelled series with weak seasonal component (Figure 3.10): variance increases following an almost perfect power law until it reaches a plateau at the largest scales.

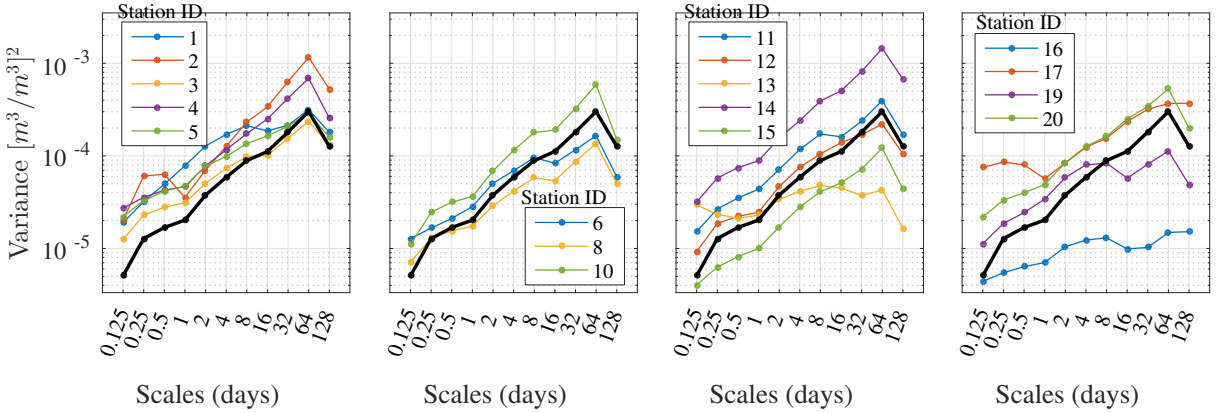
Figure 3.13 shows the **wavelet-based correlation** between the point *in situ* series and the *in situ* average series for each of the 4 networks. From a general perspective, the most striking feature is that the correlation signatures of the LW and WG stations seem less dispersed and with shapes that are more similar than those of the Yanco and MBy networks. Different aspects can contribute to that. Firstly, Yanco and MBy series are shorter than LW and WG series (1-2 years instead of 3-4). The obtained correlation values, specially at large time scales, are more uncertain (see Appendix C) and so the patterns that we observe could be questioned. However, the differences between the Yanco diagrams are so important that there should also exist relevant geophysical aspects behind. From a geophysical perspective, the main causes of the differences between LW-WG and Yanco-MBy networks may be:

(a) The LW and WG stations are much closer to each other than the Yanco and MBy stations: the extent of the later cover one satellite footprint while the former only part of it (1/4 and 1/16, Table 2.3). This can contribute to a higher degree of similarity between the measured time series.
(b) Yanco and MBy networks may contain higher heterogeneity of precipitation than the other networks because on the one side, Yanco contains irrigated and non-irrigated farms, and on the other side, their larger extents contribute to larger differences in precipitation between the stations.
(c) Yanco and MBy are agricultural sites while LW and WG are not. This implies that vegetation cover changes drastically at least once or twice per year and in a non-uniform manner in space, which should contribute to different surface responses and so different SM patterns.

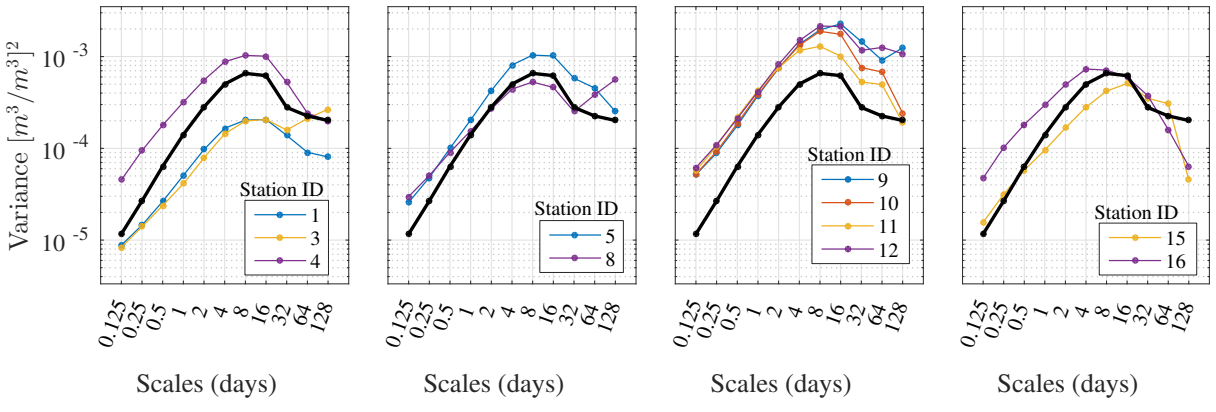
It is not possible to disentangle which of these factors is the main cause of the differences between LW-WG and Yanco-MBy networks and most probably they all play a significant role. Nevertheless, it must be highlighted that the origin of the differences is geophysical (and not noise) and they illustrate the ability of wavelet decompositions to capture different dynamics and processes.



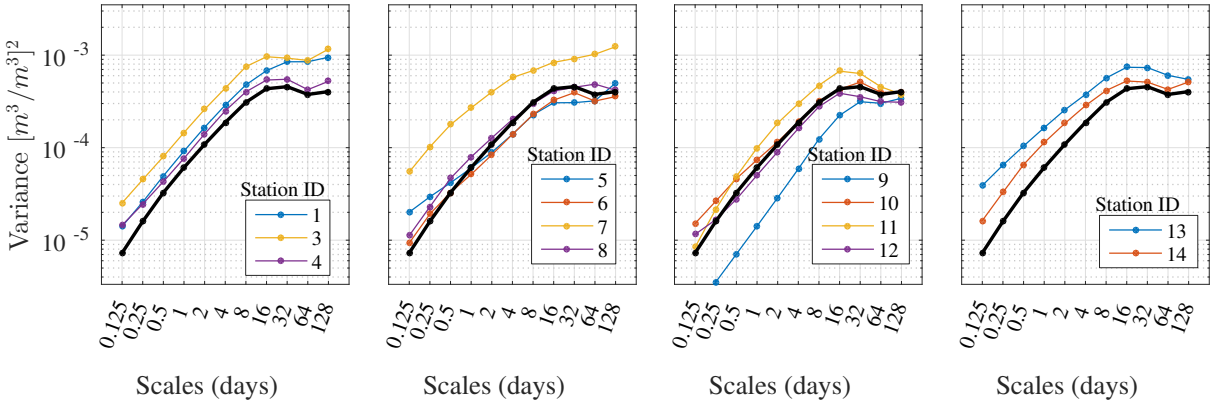
(a) Little Washita



(b) Walnut Gulch

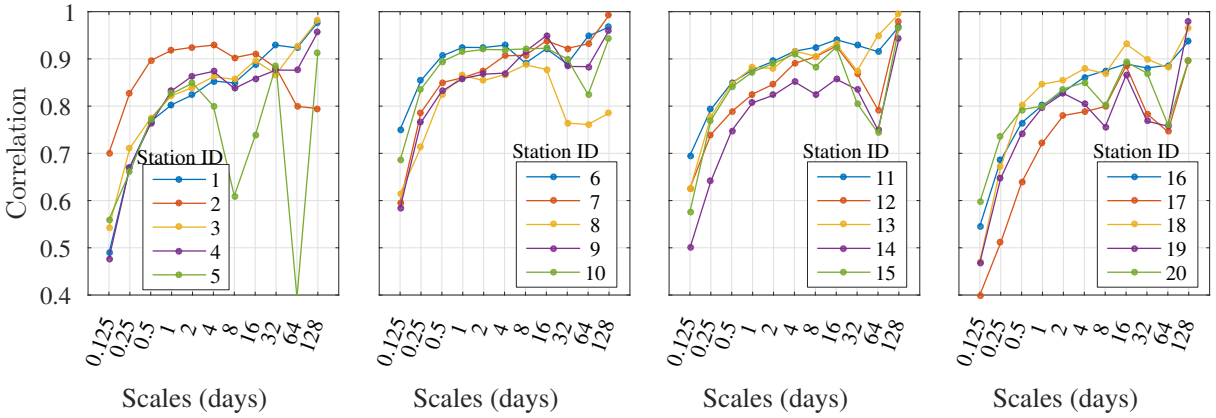


(c) Yanco

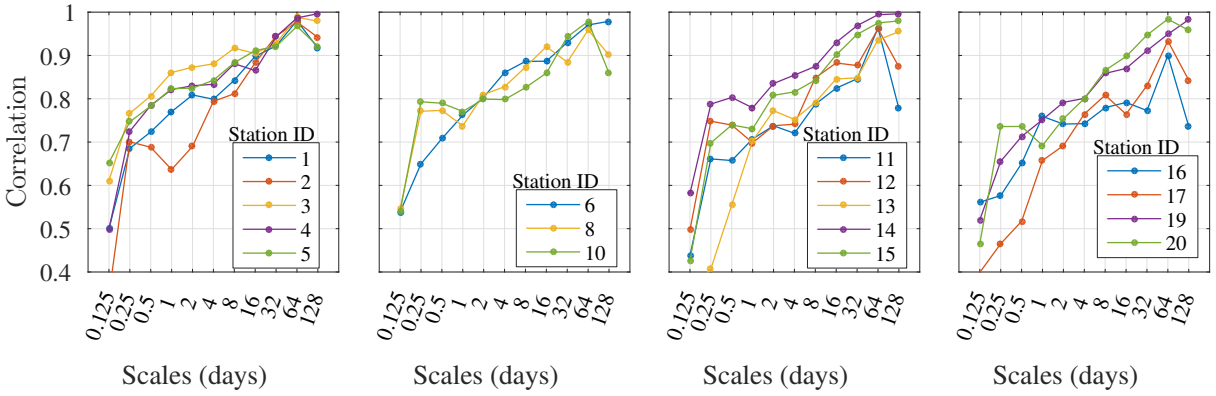


(d) Monte Buey

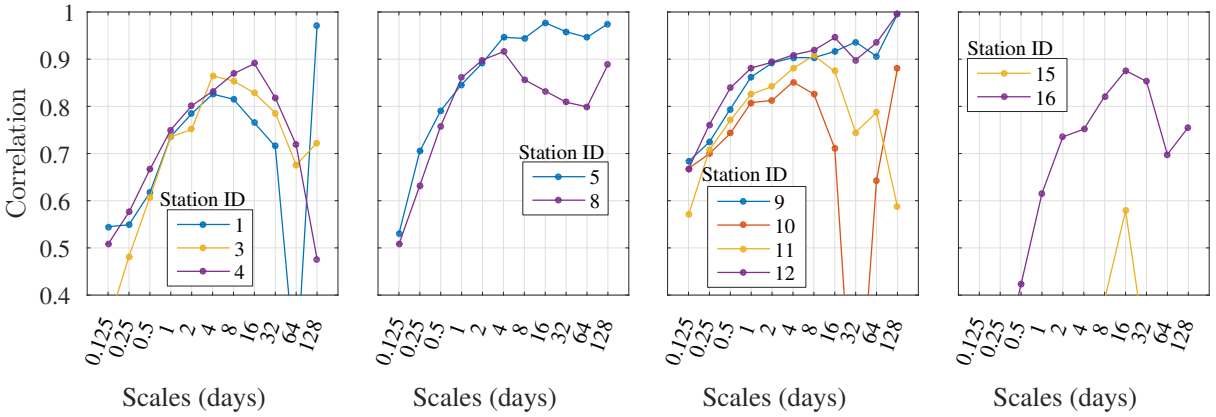
Figure 3.12 – Variance decomposition of the point *in situ* series (color lines) and the *in situ* spatial average series (bold black line)



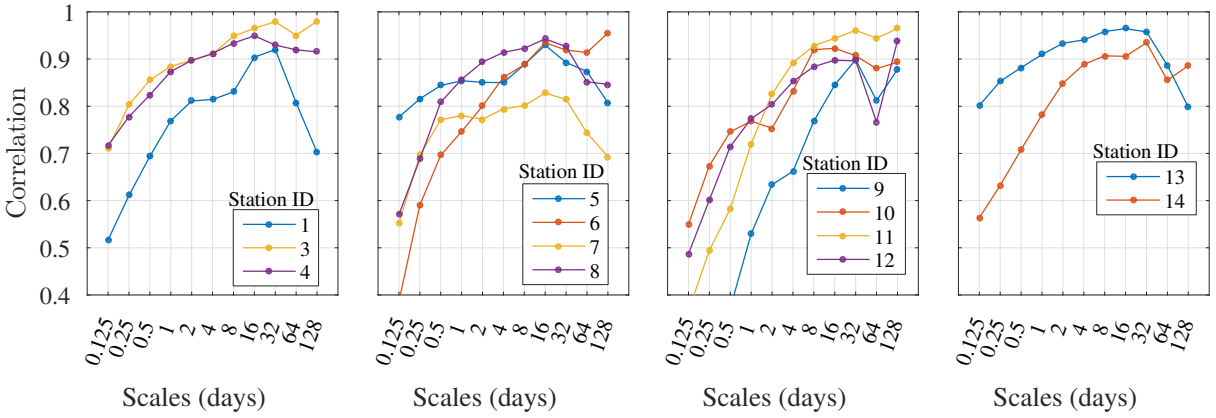
(a) Little Washita



(b) Walnut Gulch



(c) Yanco



(d) Monte Buey

Figure 3.13 – Correlation between detail series of the *in situ* spatial average and the point *in situ* series

Visual inspection of the area

Linking the variance and correlation patterns in Figures 3.12 and 3.13 to the local conditions of the stations (precipitation, soil characteristics and seasonal drivers) would require

(a) a very detailed knowledge of the local conditions at those locations
(b) a good estimation of what we understand as "average" conditions of the area

The visual inspection of the validation sites can only partially answer the aforementioned questions. I participated in a Cal/Val campaign in Monte Buey in Spring 2016. During this campaign I could realize the difficulty of predicting the level of similarity between a point and the average SM response by visual inspection. As an example, I was not able to identify the reason of the correlation decrease of stations #8 and #13 at seasonal scales (Figure 3.13d). Their local landscape conditions (illustrated in Figures 3.14a and 3.14b) were similar to those of other stations that showed instead good agreement with the average at the same scales (e.g. station #11 in Figures 3.13d and 3.14c). However, the visual inspection was proven useful in some occasions. For example, station #1 was permanently covered by 1-m height grass, which is a rare condition in the MBy area and so could explain the lack of synchronism at seasonal scales between this station and the average series (Figure 3.13d). Station #2 also exhibited decorrelation at seasonal scales and this was explained by the presence of a leaking pipe nearby. Due to this, the station was removed from the analysis. Finally, station #7 was the only station that in the close past was covered by alfalfa and was installed in the limits of an alfalfa plot (Figure 3.14d), which could explain the very different variance and correlation patterns of this station.

(a) Station #13 (La Lucia)
(b) Station #13 (La Lucia, another perspective)
(c) Station #11 (La Maya)
(d) Station #7 (El Eclipse)

Figure 3.14 – Pictures taken during the Monte Buey Cal/Val campaign on november 2016. The white notebook indicates the position of the station

Comparison with modelled series

Comparing now the correlations of modelled data (Figure 3.11) with those of *in situ* data (Figure 3.13), there are two signatures that appear in both experiments: correlation increases with the time scale and correlation increases but presents drops in the middle scales (irregular pattern). According to the conclusions derived from modelled series, the increasing pattern appears when the stations only exhibit differences in precipitation amplitudes. This is the case of many stations in the *in situ* experiment (e.g. stations #1 in LW, #1 in WG, #5 in Yanco and #3 in Mby, Figure 3.13). Regarding the irregular correlation pattern, from the modelled series it was concluded that this pattern appears when there are differences in precipitation between series that have weak seasonal components. More precisely, the variance reaches a plateau starting at the 32-day scale (Figure 3.10) and correlation increases with scale with sudden drops in the middle (Figure 3.11-I and II). It can be observed a similar behaviour in the *in situ* experiment for some stations in Figures 3.12 and 3.13: stations #5, #14, #15, #20 of LW, stations #1, #10, #16 of WG and station #12 of MBy.

Regarding texture, the model showed that series with contrasting textures exhibited a correlation decrease at middle scales (16-, 32-days) or towards the largest scales ($\geqslant$ 32 days) (Figure 3.11)

together with a different variance pattern (Figure 3.10). Stations that reproduce similar patterns in Figures 3.12 and 3.13 are stations #8, #17, #19 and #20 of LW, station #16 of WG and station #8 of Yanco. It is noticeable the case of some stations, like stations #4 of LW and #13 of WG, that exhibit a very different variance pattern but whose correlation still monotonically increases with the time scale. A difference in the sensor installation depth of some few centimetres could produce a different redistribution of the variance between time scales while keeping a fairly good synchronism with the reference time series.

Finally, the seasonal time scale of some stations is not well synchronised with that of the average series. In the conclusions of section 3.4.1, I proposed that the SM seasonal scales are made of both the contribution of seasonal drivers and the integration over time of short-time scale variables. For this reason, it is difficult to infer from the variance and correlation diagrams the reason of the differences at seasonal scales. Some examples are stations #4 and #11 of Yanco, where the decorrelation at seasonal scales can potentially be produced by different seasonal land use and temperatures and/or the integration of the combined effects of irrigation, precipitation and soil texture differences.

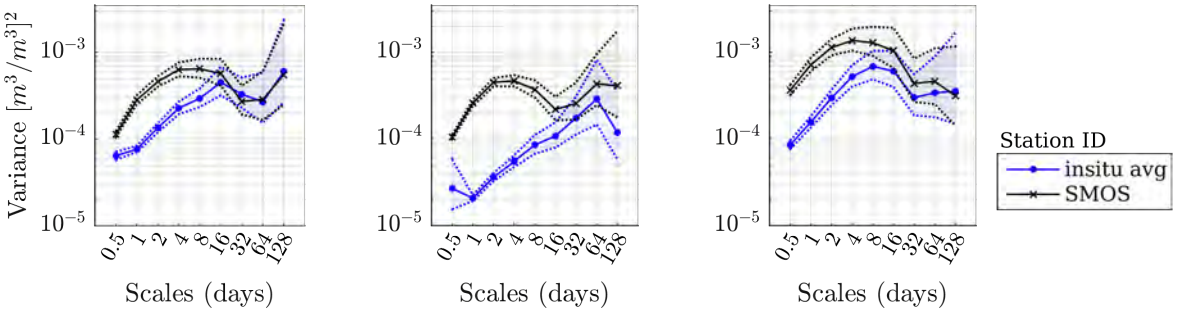
3.4.3 Time-scale decomposition of SMOS series

The previous section showed that there is a relationship between the wavelet variance signature of the stations belonging to the same network and the climatological conditions of the region. Moreover, the *in situ* average series also followed the same signature. The fact that the climate can explain the variance pattern does not mean that it cannot be influenced by other factors. Soil moisture observations are conditioned by the observing system characteristics like the instrument configuration, the models to obtain SM, the different errors and the spatial observing support. Do they also play a role in the decomposed variance and correlation patterns? In order to answer this question I propose to compare the wavelet decompositions of SMOS time series and the average of the *in situ* time series (the network average, NAvg). The datasets used in this experiment were described in section 3.3 and their respective periods and processing parameters were summarised in Table 3.7. Given that the minimum sampling period possible of SMOS is 12 h, all the time series here are sampled at a $\Delta t = 12$ h interval, the NAvg *in situ* series included. The MBy network is not included here because of RFI presence in SMOS observations.

Figure 3.15 shows the **wavelet variance** of the SMOS series and the *in situ* average series together with their respective CIs, computed following the method described in Appendix B. The difference between the SMOS and the *in situ* variance patterns is considerable. While the *in situ* patterns are characterized by a first stage of monotonically increasing variance between the 0.5-day scale and the 8- to 64-day scale, SMOS variances always reach their maximum at the 4- and 8-day scales. This "inflation" of the variance in the first scales can be due to a combination of factors. The first of them is that SMOS series have undergone a more significant gap-filling than *in situ* series (see Table 3.7). In order to clarify the effect of gap-filling, I replicated the SMOS gaps in the *in situ* series and computed wavelet variances again. The resultant variance decomposition for the 3 networks is shown in Figure 3.16. Surprisingly, the DCT gap-filling does not increase the variance but decreases it at the first 3 scales (0.5-2 days). This can be due to gap-filling working as a low-pass filter and producing smoother transitions than the real ones. The most relevant feature is that the rest of the scales (>2 days) remain unaffected by the gap-filling stage. This implies that phenomena occurring at time scales larger than 2 days can be perfectly tracked by SMOS with no degradation due to the overpass frequency.

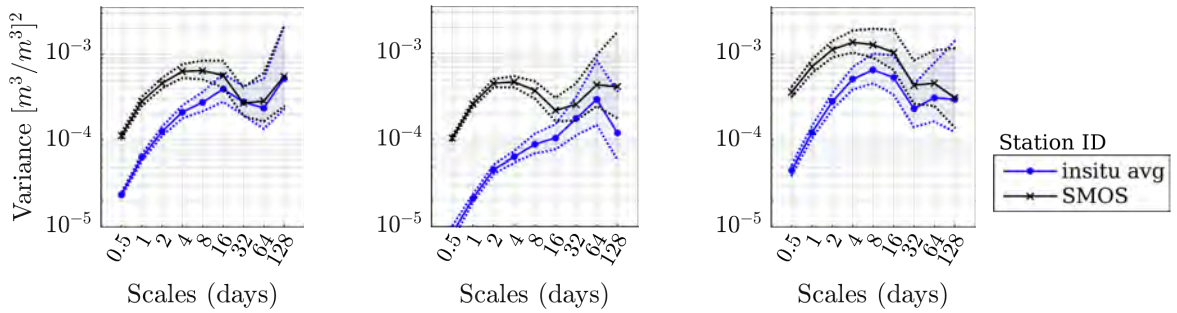
Another possible cause for the "inflated" variance of SMOS observations at sub-weekly scales is that SMOS observations may suffer from higher levels of instrumental noise. However, this argument does not explain by itself the variance levels shown in Figure 3.15. This will be further justified in the next chapter, where I will include a description of the sources of noise in SMOS soil moisture retrievals and show the variance diagrams of other large-support satellite observations

(AMSR2) and models (ECMWF). Finally, the most likely explanation is that there co-exist both noise and geophysical contributions associated to the spatial scale mismatch between SMOS and *in situ* acquisitions. While SMOS observes a continuous integrated surface through its antenna pattern, the *in situ* average is constructed with only 11 to 20 points. In plain words, SMOS observes much more local responses through its average TB than the *in situ* average. It is not surprising that the larger variance mismatch appears in the WG network: WG is the network with the smallest extent of the 3 analysed, which suggests that the variance mismatch may have a geophysical reason. This will be discussed in the next chapter.



(a) Little Washita (LW) (b) Walnut Gulch (WG) (c) Yanco

Figure 3.15 – Variance decomposition of the average *in situ* series and the SMOS series

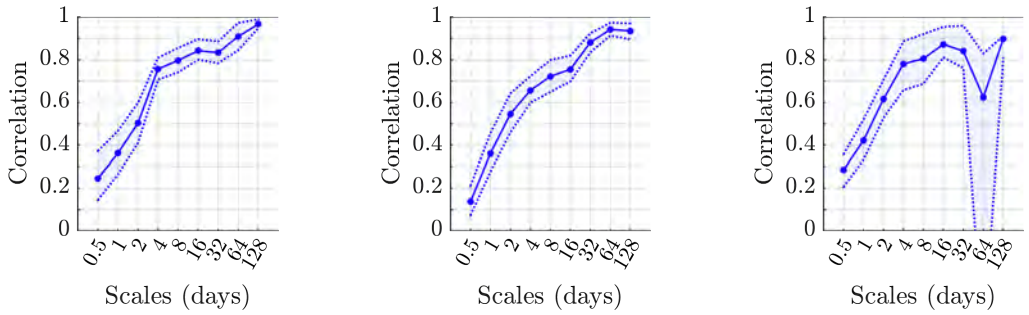


(a) Little Washita (LW) (b) Walnut Gulch (WG) (c) Yanco

Figure 3.16 – Variance decomposition of the average *in situ* series and the SMOS series after setting the same gaps in *in situ* series as in SMOS series

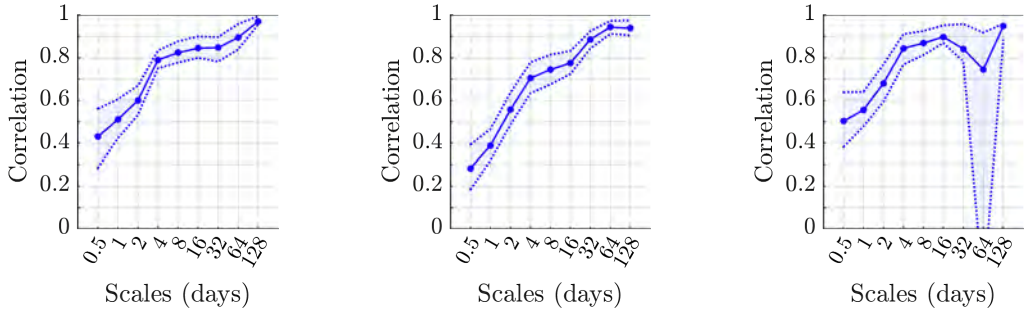
The **correlation** between the detail series of the SMOS and the *in situ* average series is shown in Figure 3.17. The CIs have been calculated following the procedure in Appendix C. There are common features between the diagrams of the 3 networks. For example, they all exhibit very low correlation at small time scales. As in the case of the variance diagrams, this seems to point out that the observation systems are affected by considerable levels of instrumental noise, and/or that the *in situ* spatial average does not succeed to represent the full surface at those scales. In addition to this, the correlation patterns in Figure 3.17 increase monotonically with the time scale, except for Yanco , which shows a decrease at the 64-day scale. The monotonic increase implies that the *in situ* spatial average succeeds to capture the middle-scale and large-scale SM responses (surface memory and climatology) the same way as SMOS. According to the previous sections, the decorrelation at the smallest scales can be due to precipitation: the *in situ* average does not capture precipitation the same way as SMOS does, which is not surprising given that the first is the average of some SM samples and the second is the SM associated with the average energy of the entire area. According to the conclusions of the two precedent sections, the irregular pattern of Yanco (Figure 3.17c) can be due to a higher heterogeneity in forcing events combined with a moderately low weight of the seasonality (Figure 3.15c). This is consistent with the fact that Yanco contains irrigated areas.

Which is the effect of gap-filling in correlation diagrams? By introducing the gaps of SMOS series in the *in situ* series and computing correlation again, I can confirm that, like in the case of the wavelet variance, the first 3 scales are the most affected by the gap-filling stage (Figure 3.18). This seems coherent with the frequency of gaps in SMOS series that is 2 gaps every 2 samples on average. In the cases presented here, the 0.5- to 2-day scales exhibit an increase of the correlation after the gap-filling experiment, most likely because the gap-fill procedure (DCT) created similar smoother temporal patterns in both *in situ* and SMOS series.



(a) Little Washita (LW) (b) Walnut Gulch (WG) (c) Yanco

Figure 3.17 – Correlation between the detail series of the average *in situ* series and the SMOS series



(a) Little Washita (LW) (b) Walnut Gulch (WG) (c) Yanco

Figure 3.18 – Correlation between the detail series of the average *in situ* series and the SMOS series after setting the same gaps in *in situ* series as in SMOS series

3.5 Conclusions

The temporal variability of surface SM is driven by the same elements that control its spatial variability: precipitation, surface characteristics, land cover and topography. These elements shape the temporal response at different time scales. For example, precipitation is responsible of the intermittent peaks in SM time series, surface properties control the velocity of the dry-down after a precipitation event, and seasonal-varying factors (e.g. land cover, temperature) can create SM seasonal-varying trends.

Although their effects in original SM series have been usually addressed in the SM literature, comprehensive studies on the SM time scales and their relation to differences in drivers (heterogeneity) was still missing. This chapter has provided the elements to better understand the relations between the exhaustive set of SM time scales (from 3 h to 128 days) and the SM drivers. The experiments concerned variance and correlation analyses of wavelet-decomposed SM series. The SM series were obtained from a modelled dataset and the *in situ* and SMOS datasets acquired at 4 *in situ* networks with constrasting climatological and environmental conditions (Little Washita, Walnut Gulch, Yanco, Monte Buey). The results of these experiments have led to the following conclusions:

1. The variances of the wavelet-decomposed series at each time scale (wavelet variance) describe a shape that is characteristic of the type of sensor (SMOS, *in situ*) and of the climatological and environmental conditions of the site. Important differences in precipitation frequency, surface characteristics and seasonal drivers can change significantly the variance pattern.
2. As inferred from the correlation diagrams (correlation between two wavelet-decomposed series at each time scale), SM drivers are associated, as expected, to characteristic time scales in SM: precipitation is associated to sub-weekly scales, surface characteristics to sub-weekly to sub-seasonal scales and seasonal drivers mainly concern monthly to seasonal scales.
3. If the only difference between two different locations concerns the amplitudes of the precipitation events (no lack of synchronism), the correlation between the respective SM series increases with the time scale.
4. Although SM drivers are associated with characteristic time scales, there also exists a propagation of the effects of precipitation and surface characteristics to larger scales (monthly and seasonal): the lack of synchronism between the precipitation events at two locations and the combination of precipitation differences with texture differences produce irregular patterns of correlation (the increasing trend is not followed).
5. The correlation diagrams of the *in situ* series belonging to larger-extent networks ($\sim$1 footprint) are much more different that those belonging to smaller-extent networks ($< 1/2$ footprint).
6. The differences between the patterns of the wavelet variance of SMOS and *in situ* series and their wavelet-based correlation are mainly associated to the spatial scale mismatch and noise. This will be further investigated in the next chapter.

This chapter has presented a series of experiments and analyses in the temporal domain but that are also related to the spatial scale problem: the last analyses concerned the comparison of point *in situ* series and footprint-like series (the *in situ* average) and the comparison of the *in situ* average and large-support series (SMOS). Moreover, the first part of the chapter concerned the comparison of series with differences in precipitation/texture/seasonal components. The term "differences" is an implicit synonym of "spatial heterogeneity". The transition to the next chapter, which will address the connection between time and spatial scales, is thus straightforward: if SM drivers affect differently each time scale, the degree of spatial heterogeneity at each time scale will be also different.

Chapter 4

Spatio-temporal dynamics of soil moisture within the footprint

Contents

4.1 Connections between spatial and temporal scales 80
4.1.1 Scales of SM-related processes 80
4.1.2 Connections between spatial and time scales in SM 81
4.1.3 Spatial representativeness of point-support SM series 83
4.1.4 Summary of hypotheses 83
4.1.5 Summary of the experiments proposed 83
4.2 Methods for spatial representativeness assessment 84
4.2.1 Existing methods 84
4.2.1.1 Temporal stability TStab 84
4.2.1.2 Triple collocation (TC) 85
4.2.1.3 Inverse footprint 86
4.2.2 Novel methods 86
4.2.2.1 Correlated areas (CArea) 86
4.2.2.2 Wavelet correlation (WCor) 87
4.3 Assessment of the connection between spatial and time scales . . . 88
4.3.1 Datasets 88
4.3.2 Spatial and time scales in SM gridded data 89
4.4 Assessment of the spatial representativeness: comparison of methods 92
4.4.1 Datasets 92
4.4.1.1 AMSR2 data 93
4.4.1.2 ECMWF data 93
4.4.2 Methods applied to full time series 94
4.4.3 Methods applied to decomposed series 95
4.5 Assessment of the spatial representativeness of decomposed *in situ* series 97
4.5.1 Datasets 98
4.5.2 Representativeness in dense networks 98
4.5.3 Representativeness in sparse networks 102
4.6 Discussion: links with chapter 3 103
4.7 Conclusions 104

Chapters 2 and 3 were dedicated to the study of the spatial and the temporal variability of surface soil moisture (SM), respectively. The former introduced the concept of *spatial scale* and showed that it is possible to go from the large scale to the small scale with *downscaling approaches*: if the adequate geophysical proxies for the region of study are chosen (temperature, vegetation, etc.), the *large-support* SM observations can be transformed into *small-support* SM observations. After that, chapter 3 showed that the same geophysical drivers of SM spatial variability cause SM to behave differently depending on the observed *time scale*.

This thesis is focused on the study of the spatial scale mismatch between point-support and footprint-support observations. Studying the spatial scale mismatch is challenging because, at the point scale, only few locations can be sampled simultaneously to be compared with the footprint-scale SM. However, in the temporal domain, the availability of SM observations is large. Would it be possible to assess the spatial mismatch from the temporal domain? This could only be possible if spatial and temporal scales are connected, so that the behaviour of the ones could be inferred from the others. This chapter is dedicated to the study of the interconnections of spatial and time scales in SM within typical satellite footprints (∼50 km). The ultimate objective is to understand how the spatial scale mismatch is distributed through the different time scales.

The chapter is structured as follows. First, the state of the art concerning the connection between spatial and temporal scales in hydrology and SM is presented (section 4.1). Then, the methodology is introduced. The methods concerning time-scale decomposition were already described in the previous chapter (wavelet transforms). The methods concerning the spatial domain are introduced here: section 4.2 is dedicated to the description of available and novel approaches for the assessment of the spatial representativeness of SM datasets. Sections 4.3, 4.4 and 4.5 present the experiments conducted on modelled and measured SM datasets. The conclusions are summarised in section 4.7. Some of the analyses and results here presented have been included in a recent publication (Molero et al., 2018). The current document is, nevertheless, a more broad and comprehensive view of the spatio-temporal scales of SM within the footprint.

4.1 Connections between spatial and temporal scales

Geophysical processes take place and are observed at a wide range of spatial and temporal scales. Some processes exhibit one or more preferred scales, i.e. certain scales contain most of the process signal energy (variations). This has diverse implications. First, processes can transfer their characteristic scales to other coupled geophysical variables. For example, daily solar cycle produces daily patterns in evapotranspiration that are reflected in SM daily patterns. Secondly, the process characteristic scale influences its observation. When the observation scale is smaller than the process scale, this one is perceived as a trend in the data; when the observation scale is larger, the same process can be perceived as noise. This is applicable to both spatial and temporal representations.

The following sections present a summary of past research on spatial and temporal scales in hydrology and more specifically in surface SM, which will lead to a general hypothesis on the interaction between SM spatial and temporal scales.

4.1.1 Scales of SM-related processes

The classification of processes by characteristic spatial and time scales has received the interest of disciplines like oceanography (Stommel, 1963) and meteorology (Fortak, 1982). To my knowledge, the first to draw a time-spatial diagram of hydrological processes were Blöschl and Sivapalan (1995), based on past literature and their own research. The diagram is presented in

Figure 4.1 and some reference scales of SM have been added to it. The minimum and maximum spatial scales, local and footprint, are highlighted on the horizontal axis. In the vertical axis, 3 time scales (daily, seasonal and yearly) are highlighted for informative purposes, although, as shown in the previous chapter, more time scales could be of interest.

The observation of Figure 4.1 already provides some useful elements for the understanding of SM temporal and spatial scales. For example, it can be seen that precipitation, which is the principal forcing of SM, can be found in various forms and scales in nature. Precipitation fronts are expected rather homogeneous within the satellite footprint. They can last for some hours (fast cold fronts) to some weeks (slower warm fronts) (Li et al., 2011; Diem, 2013). In contrast, storm cells and thunder-storms are mainly heterogeneous within the footprint and are usually restricted to sub-daily time scales.

As illustrated in Figure 4.1, phenomena related to the decrease of SM levels also apply to a range of spatial and time scales. For example, there can exist daily evapotranspiration cycles at scales smaller than the footprint. When the infiltration capacity is exceeded as a result of high rainfall intensity, water drains to deeper soil layers, a process that can last for short time periods close to the daily scale. Horizontal runoff can appear as a result of soil saturation or a raise of the groundwater table, in which case it affects field and catchment scales during longer periods (weeks, months, seasons). Finally, floods can even affect the footprint scale. In their absence, runoff is considered heterogeneous within the footprint scale.

SM is still coupled to many other processes and variables missing from Figure 4.1 that operate at distinct spatial and temporal scales. For example, SM can be coupled with evapotranspiration and temperature (Seneviratne et al., 2010) that usually exhibit seasonal patterns that cover regional extents (Craigmile and Guttorp, 2011).

As previously introduced, the observation is affected by the relative gap between the observational scale and the process scale. In consequence, the SM variable is perceived differently at the *in situ* and at the satellite spatial scales. Taking the footprint scale as the reference spatial scale of observation, precipitation fronts should be perceived as spatial trends while convective cells are probably translated in "noise" because of their within-footprint heterogeneity. A similar situation to that of convective weather patterns may happen with runoff, drainage and evapotranspiration. In contrast, seasonal temperature cycles would most probably be translated into spatial SM trends given their regional extent. This leads to the conclusion that when comparing *in situ* and satellite observations, they should contain similar spatial trends (associated to factors occurring at equal-to- or larger-than-footprint scales), and different small spatial scale variations associated with within-footprint heterogeneity. Could this remaining "noise" due to the spatial scale mismatch between the *in situ* and footprint observations be corrected or, at least, identified?

The answer could reside in one remarkable feature present in the aforementioned processes and observed in Figure 4.1: regarding SM-related processes, **small spatial scales tend to be associated with small time scales, and large spatial scales with large time scales**. However, it remains to unveil whether the same is directly applicable to the SM variable itself. Should this be true, when comparing *in situ* and satellite SM observations, their small time scale components should be different while their large time scale components should be similar.

4.1.2 Connections between spatial and time scales in SM

The association between SM spatial and temporal scales has already been addressed in the past. Although the literature is short, the conclusions are encouraging. At regional and continental scales, the studies of Cayan and Georgakakos (1995), Vinnikov et al. (1996), and Entin et al. (2000) found that small spatial scales were connected with small time scales and that large spatial scales were connected with large time scales. For example, Entin et al. (2000) identified

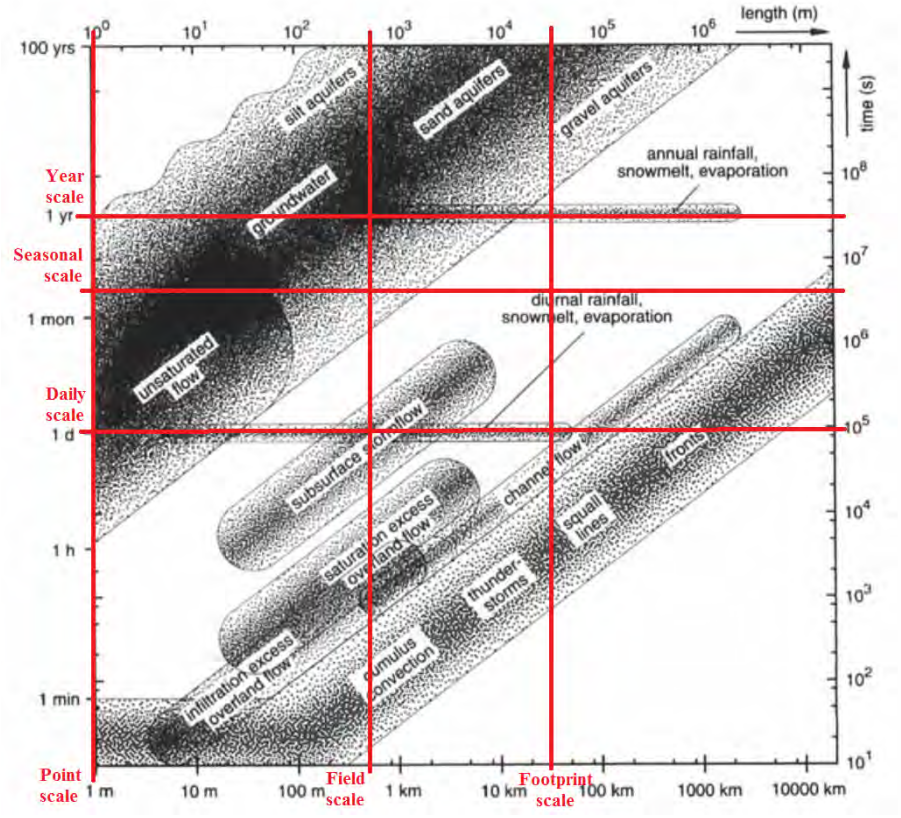


Figure 4.1 – Hydrological processes at a range of characteristic spatial and temporal scales. Source: (Blöschl and Sivapalan, 1995)

two main scales: the small scale, that was of the order of some tens of meters and of a few days, was due to local processes such as infiltration, precipitation and drainage; the large scale, that was mainly of the order of some hundreds of km and 2-3 months, was due to climatic atmospheric forcing.

At the field and small catchment scales, Skøien et al. (2003) succeeded to relate characteristic spatial and time scales of 4 variables (precipitation, runoff, water table and SM) stem from the analysis of an Austrian hydrographic dataset and five small catchments in Australia and New Zealand. The SM dataset covered spatial extents of up to 1.4 km^2. Concerning extents closer to the footprint scale, Chaney et al. (2014) demonstrated that SM seasonal variations in the Little River catchment were controlled by large spatial scale factors (land cover and evapotranspiration) and not by small spatial scale factors (soil texture). Recently, Su and Ryu (2015) have shown that the correlation between point SM time series (*in situ*) and footprint SM time series (satellite) increases with the time scale. This is expected if SM-related processes at spatial scales equal to or are larger than the footprint exhibit also large time scales, while sub-footprint processes are related to short time scales. The studies of Su and Ryu (2015) and Chaney et al. (2014) are the only ones that evaluate SM time scales at the field and footprint (or close to footprint) scales.

The spatial scale of seasonal trends. In the domain of satellite SM validation, there is a number of publications that state that temporal trends of *in situ* and satellite datasets are far from being similar. In particular, most of the publications that compare SM datasets through a method called triple collocation (TC), which I will describe later in this chapter, express that there exist systematic differences in the climatologies of *in situ* and satellite time series (Dorigo et al., 2010; Gruber et al., 2013a). For this reason, prior to the application of the method, they removed the climatology obtained with a moving average window filter of 35-days long (Crow et al., 2010; Draper et al., 2013) or 31-days long (Gruber et al., 2013a). As an example,

Gruber et al. (2013a) states that "the seasonal vegetation growing cycles strongly affect satellite observations, but not *in situ* measurements". This could be due to *in situ* stations not being installed under vegetated covers, which is usually the case of agricultural sites.

This interpretation of seasonal scales is in conflict with the aforementioned studies on SM spatial and time scales (e.g. Ryu and Famiglietti, 2005; Chaney et al., 2014) that highlight a fundamental consistency between the temporal trends of point and footprint datasets. What happens then with SM seasonal scales? Are they governed by sub-footprint or by "super-"footprint processes? I will provide the elements to elucidate this apparent divergence in interpretations that has not been yet addressed in the literature.

4.1.3 Spatial representativeness of point-support SM series

Although a typical *in situ* station has a spatial support of just some few centimetres (point/local support), it can represent a larger extension because the drivers of SM variability (precipitation, soil characteristics, vegetation and topography) show some degree of spatial connectivity and organisation. The same is usually expressed by saying that the **spatial representativeness**, the **representativeness area** or the **effective support** of the station is larger than its actual support. In soil moisture calibration and validation (Cal/Val) studies, the representativeness area of a location is measured as the surrounding extension where the observed SM is "sufficiently similar" to the SM at that location. The degree of similarity is usually expressed in terms of correlation-like metrics, although sometimes variance-based analysis like variograms can be used.

If spatial and temporal scales exhibit some sort of connection, then the spatial representativeness of a station should be also connected to the time scales. This implies that if small time scales and small spatial scales are associated, then locations have small representativeness areas at small time scales. On the contrary, if large time scales and large spatial scales are associated, locations should have large representativeness areas at large time scales. In other words, footprint-scale observations and point-scale observations are expected more similar at large than at small time scales.

4.1.4 Summary of hypotheses

Along this section 4.1, I have introduced some hypotheses concerning the connection between spatial and time scales of surface SM that I summarise here:

(i) Spatial and temporal scales are connected.
(ii) Small/large spatial scales are connected to small/large time scales, respectively.
(iii) The spatial representativeness of a location increases with the time scale.
(iv) The connection is induced by SM-related factors: those that are heterogeneous within the footprint occur at shorter time scales than those mainly homogeneous within the footprint.
(v) Processes occurring at a scale smaller than the observation scale are perceived as noise while processes occurring at larger scales are perceived as trends.

In addition, there is a remaining question concerning the seasonal time scales of SM: are they associated with large spatial scales or not? Should we expect similarity between observations of different support sizes at seasonal time scales?

4.1.5 Summary of the experiments proposed

In order to verify the aforementioned hypotheses and shed light on the related questions, a set of experiments are conducted. In a first place, modelled SM data is analysed to derive preliminary conclusions. Then, *in situ* and satellite SM data are analysed. In all experiments, time scales are obtained through wavelet transforms as in chapter 3. The spatial scales are assessed through

methods that evaluate the spatial representativeness of some selected locations. These methods are presented in section 4.2.

The **first experiment** is described in section 4.3. It takes advantage of the full regular spatial sampling of a modelled SM dataset to investigate whether spatial and time scales are potentially connected. The full sampling feature allows also to observe the spatial evolution of the representativeness area: its shapes, exact extent and spatial evolution with the time scale.

The **second experiment** is dedicated to the comparison of different methods for spatial representativeness assessment when the surface is not fully sampled (section 4.4). This is the nominal case since dense *in situ* networks provide around 20 sampling locations within a satellite footprint at most. The study is also based on the gridded modelled dataset and is a necessary in order to select the best methods for the analysis of real measured data.

The **third experiment** evaluates the spatial representativeness of multiple *in situ* stations belonging to different networks. As in the previous experiments, the evaluation is done at a range of time scales. It allows to verify the conclusions from the first and the second experiment on modelled data.

Finally, it should be reminded that in the previous chapter, it was presented how differences in SM drivers (precipitation, soil texture and vegetation growth) affect SM time scales (section 3.4.1). The spatial aspect was actually implicitly addressed: "differences" can be replaced by "spatial heterogeneity". The conclusions of section 3.4.1 will be of help interpret the results of the experiments in this chapter.

4.2 Methods for spatial representativeness assessment

The evaluation of the spatial scales of SM observations will be addressed through the evaluation of the spatial representativeness (or simply, representativeness) of point datasets with respect to footprint datasets. Herein, I describe the methods used for this purpose, the existing ones (section 4.2.1) and two novel approaches I propose (section 4.2.2).

4.2.1 Existing methods

4.2.1.1 Temporal stability TStab

Introduced in Vachaud et al. (1985), the temporal stability analysis (TStab) method has been described and used in a number of publications (Martínez-Fernández and Ceballos, 2005; Cosh et al., 2006; Mittelbach and Seneviratne, 2012). Given a point-support series $x_i(t)$ measured at location i, and a large-support series $y(t)$, the temporal stability analysis (TStab) evaluates their relative differences (RD):

$$RD_i(t) = \frac{x_i(t) - y(t)}{y(t)} \tag{4.1}$$

The most representative point time series is the one that exhibits the smallest (relative) bias and dispersion with respect to the large-support series, i.e. the one with smaller mean RD or MRD (Equation 4.2) and with the smallest mean standard deviation of the RD or SDRD (Equation 4.3).

$$MRD_i = \frac{1}{N}\sum_{t=1}^{N} RD_i(t) \tag{4.2}$$

$$SDRD_i = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(RD_i(t) - MRD_i)^2} \tag{4.3}$$

Both metrics can be summarised in one single metric like in Jacobs et al. (2004):

$$RMSE_i = \sqrt{MRD_i^2 + SDRD_i^2} \tag{4.4}$$

4.2.1.2 Triple collocation (TC)

Triple collocation (TC) serves to estimate the random errors of three collocated datasets that represent the same variable (Stoffelen, 1998). It relies on a linear error model:

$$k(t) = a_k + b_k\theta(t) + \epsilon_k(t) \tag{4.5}$$

where k denotes one of the three spatially and temporally collocated data sets (X,Y,Z), θ is the unknown true soil moisture state, a_k and b_k are systematic additive and multiplicative biases with respect to the true state, and ϵ_k is a zero-mean random noise variable. The variance of the errors $\hat{\sigma}^2_{\epsilon_k}$ can only be calculated if the following assumptions are met:

(i) zero error-signal cross-covariances: $\sigma_{\theta,\epsilon_k} = 0$
(ii) zero error cross-covariances: $\sigma_{\epsilon_X,\epsilon_Y} = 0$, $\sigma_{\epsilon_X,\epsilon_Z} = 0$, $\sigma_{\epsilon_Y,\epsilon_Z} = 0$

A detailed discussion about these assumptions is provided in Gruber et al. (2016). Under these conditions, the following equalities hold:

$$\sigma_k^2 = b_k^2\sigma_\theta^2 + \sigma_{\epsilon_k}^2 \tag{4.6a}$$
$$\sigma_{k,l} = b_k b_l \sigma_\theta^2 \tag{4.6b}$$

TC has been used for intercomparing soil moisture datasets (Draper et al., 2013; Leroux et al., 2013b) and for evaluating the spatial representativeness of point-support SM datasets. The latter can be based on two different metrics, the error variance $\hat{\sigma}^2_{\epsilon_k}$ (Gruber et al., 2013b; Miralles et al., 2010) and the correlation between the dataset and the true soil moisture $\hat{r}_{k,\theta}$ (McColl et al., 2014; Chen et al., 2016).

The error variance can be derived from Equations 4.6 as:

$$\hat{\sigma}^2_{\epsilon_X} = \hat{\sigma}^2_X - b_X^2\sigma_\theta^2 = \hat{\sigma}^2_X - \frac{\hat{\sigma}_{X,Y}\hat{\sigma}_{X,Z}}{\hat{\sigma}_{Y,Z}} \tag{4.7}$$

Similar expressions hold for the Y and Z datasets. These error variances are referred as *unscaled* because they are relative to their respective X, Y, or Z data space. In order to express the error variance estimates into a common reference data space, for example that of X, $\hat{\sigma}^2_{\epsilon_Y}$ and $\hat{\sigma}^2_{\epsilon_Z}$ should be rescaled by b_X/b_Y and b_X/b_Z respectively. The b_k parameters can only be calculated if one of them is known. Usually, X represents the *in situ* dataset and is assumed perfectly calibrated with respect to the truth (i.e. $b_X = 1$). However, this is not necessarily true, as Chen et al. (2016) demonstrated in different U.S. *in situ* networks. Naive rescaling can produce biased error estimates. For a more detailed description of TC equations the reader is referred to Gruber et al. (2013b) and McColl et al. (2014).

The correlation $\hat{r}_{k,\theta}$ is defined as:

$$\hat{r}_{k,\theta} = \frac{\hat{\sigma}_{\theta,k}}{\sigma_\theta \hat{\sigma}_k} \tag{4.8}$$

If the zero error-signal and signal cross-covariance conditions are met then:

$$\hat{r}_{k,\theta} = \frac{b_X \sigma_\theta}{\hat{\sigma}_X} \tag{4.9}$$

By combination of Equations 4.7 and 4.8, the TC correlation can be finally computed as:

$$\hat{r}_{X,\theta} = \pm \sqrt{\frac{\hat{\sigma}_{X,Y}\hat{\sigma}_{X,Z}}{\hat{\sigma}^2_X \hat{\sigma}_{Y,Z}}} \tag{4.10}$$

Similar expressions hold for the Y and Z datasets. The 3 following conditions are necessary for the computation of $\hat{r}_{k,\theta}$:

(i) Non-negative cross-correlation between all datasets.
(ii) Non-negative $\hat{\sigma}^2_{\epsilon_k}$
(iii) Non-negative $\hat{r}^2_{k,\theta}$

The TC correlation $\hat{r}_{k,\theta}$ (and $\hat{\sigma}^2_{\epsilon_k}$) can be used to evaluate the spatial representativeness of a point-support dataset (X) if Y and Z are large-support datasets. If the measurement errors are relatively small, Y and Z should show stronger similarity due to their similar spatial support. Moreover, if the main differences between the X and Y/Z are due to the geophysical scale mismatch (and not to measurement errors), the X metrics mainly characterize how distant the X dataset is from the Y/Z pair. In other words, they characterize the degree of spatial representativeness of the X dataset. This is analytically described in Gruber et al. (2016) and Vogelzang and Stoffelen (2012).

The advantage of $\hat{r}_{k,\theta}$ with respect to the error variance is that it does not need rescaling and is already normalized by the total signal power (Equation 4.9) between 0 and 1, and so it can be used to compare the scores of different triplets. Unlikely $\hat{r}_{k,\theta}$, $\hat{\sigma}^2_{\epsilon_k}$ increases with the signal variance. In this chapter, I will run TC analyses on wavelet decomposed series, which have different variances, so the TC correlation metric $\hat{r}_{k,\theta}$ is selected. Note that, for simplicity, in many of the figures in this chapter $\hat{r}_{k,\theta}$ is simple denoted as $r_{k,\theta}$.

4.2.1.3 Inverse footprint

The *inverse footprint* approach was proposed and used by Orlowsky and Seneviratne (2014) and Nicolai-Shaw et al. (2015). The representativeness area of a location $i0$ is measured as the surrounding area where other locations i exhibit temporal similarity (Spearman correlation) above an arbitrarily selected threshold.

4.2.2 Novel methods

4.2.2.1 Correlated areas (CArea)

The original *inverse footprint* (Orlowsky and Seneviratne, 2014; Nicolai-Shaw et al., 2015) relies on two assumptions: First, the representativeness area of a location $i0$ is a continuous surface centred on $i0$. Secondly, the spatial sampling surrounding $i0$ is sufficient to determine the representativeness area. As it will be shown later, the spatial representativeness of SM is not

only dependent on the distance as there are "islands" of representativeness: a location can be similar to the reference location even if the locations that are closer to the reference are not. This suggests that the spatial sampling of typical dense *in situ* networks (10-20 points per satellite footprint) is not sufficient for the inverse footprint method. For these reasons, I have modified the original method into an approach that I refer as the *correlated area (CArea)* method. The modifications are:

1. CArea is only applied to SM gridded data (surface or disaggregation model at high spatial resolution).
2. Pearson correlation replaces Spearman correlation. This respects the consistency with the other approaches described in this thesis that are based on the Pearson statistic.
3. The CArea metric is not the absolute surface of the surrounding area but the percentage of pixels that exceed a specified correlation threshold, no matter if they are contiguous or not.

The CArea method evaluates the percentage of pixel time series $x_i(t)$ that are correlated with the reference series $x_{i0}(t)$ above a specific threshold R_{th} within an extent A:

$$CArea(i_0) = \frac{1}{M} \sum_{\forall i \neq i_0 \subset A} \mathcal{H}(R_{x_i,x_{i0}} - R_{th}) \times 100 \qquad [\%] \tag{4.11}$$

where M is the number of locations (pixels) i within the area of study A, and $\mathcal{H}$ is the Heaviside function that is either 0 or 1 for negative and positive values, respectively. The greater $CArea(i_0)$ and R_{th} values are, the more representative the location i_0 is of the observed surface in terms of correlation. This method can be applied also to time scale decomposed series.

4.2.2.2 Wavelet correlation (WCor)

The *wavelet-based correlation (WCor)* or just *wavelet correlation* can be applied to evaluate the similarity, in terms of correlation, between a point-support series $x_i(t)$ and a large-support series $y(t)$ on a per-time scale basis. This is mathematically expressed as

$$R_j = R\{D_{x_0,j}(t), D_{y,j}(t)\}, \qquad j = 1, 2, \ldots, J0 \tag{4.12}$$

where D_j denotes the wavelet-decomposed detail series of the respective time series, x_i or y. The R_j values are simply a measure of linear matching: they cannot directly distinguish the proportion of errors and the spatial scale mismatch in the comparison. However, as I will show later, the analysis of multiple *in situ* and modelled series, will serve to understand how the spatial scale mismatch is distributed along the time scales and how spatial and temporal scales are connected.

4.3 Assessment of the connection between spatial and time scales

The aim of this section is to verify whether it exists a connection between SM spatial and time scales based on modelled SM datasets. In order to do this, the spatial representativeness of a set of locations (pixels) will be assessed on a per-time scale basis.

4.3.1 Datasets

Table 4.1 – Pre-processing of the data used to assess the connections between spatial and time scales in the Yanco region.

Yanco - gridded SM dataset	
Period	2014/09 - 2015/03 (6 months)
Wavelet type	Haar
Δt	0.5 days
J0	8 (64 days)
Locations (selected/total)	13 / 13
% gap-filling (regression)	0 %
% gap-filling (DCT)	23.3 %
Nb. consecutive gaps (on avg.)	5
Nb. consecutive samples (on avg.)	17.4

The dataset used in this section is a set (time series) of DISPATCH maps at 1-km resolution for the Yanco region. The DISPATCH method as well as the Yanco area were described in detail in section 2.4. In this experiment, a important difference is introduced: **the large-support SM product (SMOS L3) is replaced by the spatial average of the Yanco *in situ* time series** sampled at the SMOS overpass times (approximately 6 a.m. and 6 p.m.). In other words, it is the *in situ* spatial average that is disaggregated and not the SMOS observations. This allows to produce 1-km support time series that are closer to the *in situ* series and that are independent from the SMOS satellite errors.

The period of analysis is selected so that:

1. it has the most robust NAvg, i.e. the NAvg calculated with the largest number of *in situ* stations
2. it minimises the length and the number of data gaps

Long periods of clouds reduced dramatically the availability of MODIS data and thus, DISPATCH data, specially during the Austral winter. For this reason, only 6-month periods were initially selected. Among these, it was the 2014/09/01 - 2015/03/01 period that had the most robust NAvg (calculated with 11 stations) and the lowest percentage of gaps. These constitute $\sim$23 % of the selected series on average and are filled with the DCT approach (Wang et al., 2012). The data selection and pre-processing for this experiment is summarised in Table 4.1.

The final SM dataset is fully sampled in space and time. This allows an exhaustive evaluation of the spatial representativeness when using the CArea method. Interpreting the CArea scores of all the disaggregated series within the area would be intractable, so only the pixels that coincide with the locations of the Yanco stations are selected. The spatial representativeness of these pixels is evaluated on a per time-scale basis by applying the CArea method to their wavelet detail series. In addition, the representativeness of two other series that are expected to better represent the satellite footprint is also evaluated. These are the **field average (FAvg)** series, which is the average of the series of the pixels contained within the area of study, and the **network average (NAvg)** series, which is the average of the series of the pixels matching Yanco station locations.

4.3.2 Spatial and time scales in SM gridded data

The CArea method is performed in 2 steps:

1. The correlation between the time series of the location under analysis $i0$ and all the other locations i, R_{x_{i0},x_i}, is computed.
2. The percentage of correlation values obtained that exceeds the pre-determined correlation threshold R_{th} is calculated.

Figure 4.2 shows the outcome of the first step when the series analysed $i0$ is the **FAvg**. The black circles denote the location of Yanco stations. The analysis is applied to each time scale separately. Note that the application of the first step to the FAvg series is equivalent to applying the wavelet-correlation method to assess the representativeness of each pixel series x_i.

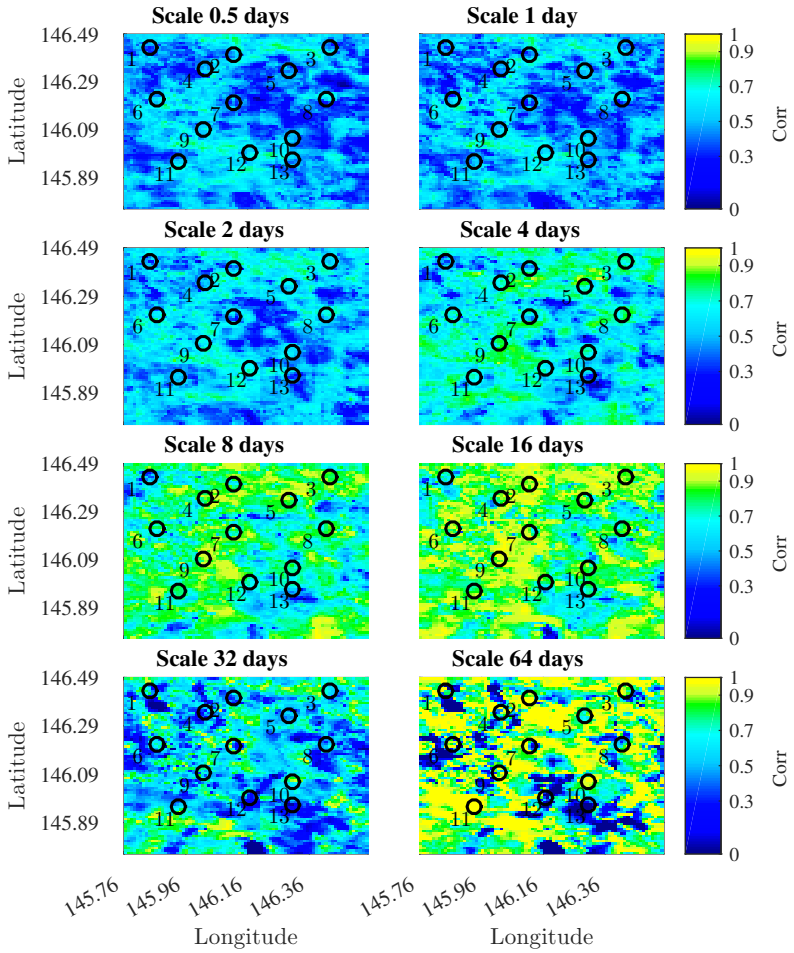


Figure 4.2 – Maps of wavelet correlation per time scale (R_j) between the detail series of each pixel and the detail series of the average of all the pixels (FAvg, field average) of the gridded SM dataset for the 2014/09-2015/03 6-month period. Source: (Molero et al., 2018)

There are 3 relevant features in Figure 4.2. The first one is that, regardless the 32-day scale, in overall, **the correlation between the pixels and the average series increases with the time scale**. Secondly, the first 3 scales (**0.5-2 days**) show very low correlation values that contrast with the other scales. Finally, **the seasonal scale (64 days)** is the one with the largest percentage of pixels with very high correlation values (> 0.9). In contrast, the percentage of pixels with very low correlation values (≤ 0.3) is non-negligible. In other words, the 64-day scale seem to be the scale that has the largest number of pixels with extreme correlation values.

To complete the analysis, the same first step is repeated for 14 other reference series $i0$: the NAvg and the 13 selected pixels. Then, the percentages are computed for a range of correlation thresholds from 0 to 1 (2^{nd} step). The final CArea scores are presented in Figure 4.3, where each

line represents one different pixel, the correlation thresholds are indicated on the horizontal axis and the spatial representativeness values are indicated on the vertical axis. The features that were observed in Figure 4.2 are also present here. This implies that **the wavelet-correlation method and the CArea method can be considered equivalent** when the surface is fully sampled. The features will be further described in the following paragraphs.

In Figure 4.3, the lines tend to move to the right as the time scale increases, meaning that the general representativeness of the pixels increases with the time scale, as it was noticed before. The FAvg and the NAvg series are the most representative series at all time scales, which is expected. Consistently with the analysis of Figure 4.2, the smallest time scales (0.5-2 days) exhibit the smallest CArea scores, with more than 75 % of the area correlated below 0.5. As introduced in section 3.4.3, the low correlation values at small time scales can be induced by gap-filling and/or by the difference in the support sizes of the local observation (the pixel, in this case) and the footprint observation (the FAvg here).

I evaluated the **impact of gap-filling** on original *in situ* series by comparing the wavelet correlation values obtained before and after setting the same gaps as those in the DISPATCH series. I found that, at the 0.5-2 day scales, correlation decreased by 0.08 on average. This means that gap-filling does not change considerably the first 3 plots in Figure 4.3 with respect to the other time scales and that, like in the experiments of section 3.4.3, the low correlation values at short time scales are probably mainly due to the geophysical mismatch between the local and the footprint observations. It should not be forgotten either that the DISPATCH model and inputs have intrinsic errors and uncertainties that can have an impact too.

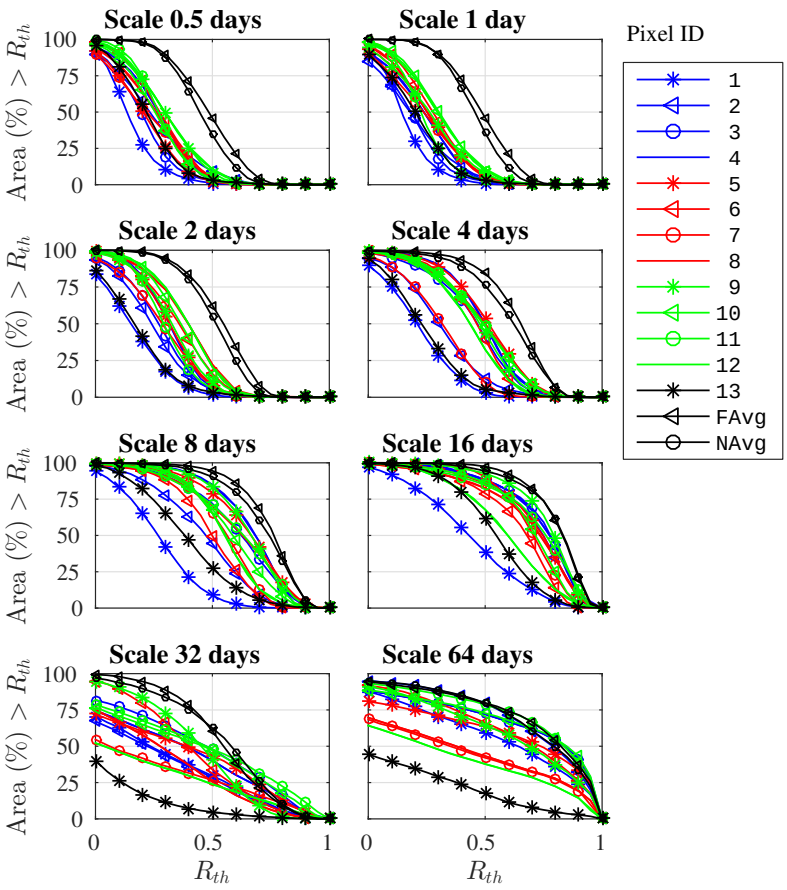


Figure 4.3 – CArea scores as a function of the correlation threshold for some specific pixels of the gridded SM dataset for the 2014/09-2015/03 6-month period. Each plot corresponds ot a different time scale. Source: (Molero et al., 2018)

Although in general, the representativeness values increase with the time scale, the lines become more distant from each other (Figure 4.3). It can be inferred that **the evolution of the spatial**

representativeness with the time scale is not the same for all locations. In this dataset, small and middle time-scale factors are represented by the surface temperature data used by DISPATCH, while the NDVI controls the last scales (32 and 64 days). The increase in the scatter of the lines with the time scale can be due to some locations being very different in terms of surface temperature and vegetation status with respect to the field average, which is not surprising since surface skin temperature observations are quite variable in space and time

At weekly scales (8-16 days), most of the series have more than 50 % of the surface with a correlation above 0.5 and 0.6, respectively (Figure 4.3). This signifies that the weekly temperature patterns are quite homogeneous in the region. However, **the 32-day scale** breaks the tendency of increasing correlation with the time scale. This is probably due to a combination of small time scale heterogeneity (precipitation/irrigation) and low temporal variance. As shown in the previous chapter, in Figure 3.11, precipitation heterogeneity can produce irregular patterns in the correlation diagram at monthly scales. This has been also shown in (Molero et al., 2018), where the comparison of time series with 20 % of the precipitation events randomly shifted by ±0.5 days led to correlation decrease at the same time scales. The hypothesis of the presence of precipitation heterogeneity in Yanco is highly plausible since it contains irrigated (and non-irrigated) farms. In addition, from a statistical point of view, low-variance signals tend to exhibit lower correlation values (Berger and Sweney, 1965; Goodwin and Leech, 2006) and the 32-day scale has a small temporal variance similar to that of the 2-day scale for this dataset. This can be verified in Figure 4.4, which presents the wavelet variance of the time series of the NAvg and the stations #6 and #13. Finally, as result of the noise propagation and the low variance, the 32-day scale may have a low signal power with respect to noise (signal-to-noise ratio or SNR).

Like in Figure 4.2, in Figure 4.3 the **seasonal scale (64 days)** is the time scale that exhibits the largest areas with very high correlation: the most representative series exhibits ∼40 % of the area correlated above 0.9. However, the dispersion in representativeness values is maximum: locations can be either highly representative of the footprint SM (series #10 in Figure 4.3, yellow areas in Figure 4.2) or not at all (series #13 in in Figure 4.3, dark blue areas in Figure 4.2). As described in section 3.4.1, this situation can be explained by two factors that come in to play at this scale: seasonal drivers and the propagation of smaller time scale variations. Regarding seasonal drivers, in the dataset used here, the main seasonal factor is vegetation. In Appendix D, I provide additional land cover information for the Yanco network, which shows that the area is mostly low vegetated (Figure D.2) but that the observed NDVI time series differ between locations (Figure D.3). However, I cannot find a clear relation between the individual NDVI series and the representativeness scores in Figure 4.3, which seems to indicate that the propagation of smaller time scale heterogeneities is more important than the seasonal driver in this experiment. The propagation of short time scale differences produces decorrelation at seasonal scales that is specially severe when the seasonal variance is weak, as indicated in Section 3.4.1. This is a highly plausible explanation in the Yanco case, because the region is rich in short time scale heterogeneities due to irrigation and the 64-day scale has a weak variance compared to the other scales (Figure 4.4).

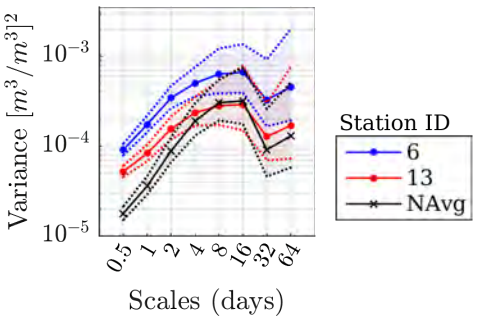


Figure 4.4 – Wavelet variance for 3 different disaggregated series: the average (black), and those of pixels #6 (blue) and #13 (cyan). Dotted lines depict 95 % CIs (Appendix B).

The results obtained for the 64-day scale shed light on the apparent opposition about seasonal scales that was introduced in section 4.1: depending on the location, the respective seasonal SM component is similar or not to the seasonal component of the footprint SM. This supports at the same time triple collocation (TC) studies, which suggested that seasonal components were different, and studies on SM scales like that of Su and Ryu (2015), which suggested that

similarity increases with the time scale up to seasonal scales.

To conclude, this section has shown that SM time and spatial scales are connected. I stress the fact that the conclusions here presented are based on modelled SM data, so they should be verified with measured SM data. This will be addressed in 4.5. The difficulties for assessing the representativeness of observed data is that the surface is not fully sampled and that the notion of representativeness depends on the method used to measure it. For this reason the following section is dedicated to the inter-comparison of representativeness methods.

Conclusions

1. Spatial representativeness scores tend to increase with the time scale.
2. The dispersion in representativeness scores also increases with the time scale: at seasonal time scales (64 days) the locations can be either highly or poorly representative of the network average.
3. Small scales exhibit uniform low representativeness scores.
4. At middle scales (32 days), representativeness scores can drop because of: a) propagation of small scale geophysical differences, and b) low SNR.

4.4 Assessment of the spatial representativeness: comparison of methods

This section is dedicated to the inter-comparison of the methods for assessing spatial representativeness that have been described in section 4.2: temporal stability analysis (TStab), triple collocation (TC), correlated area (CArea) and wavelet-based correlation (WCor). Herein, the representativeness is evaluated at the same locations of the gridded SM dataset used in the previous section. The 1-km pixels are considered as the local-support datasets. The footprint-support dataset is represented by the network average (NAvg) in the case of TStab and WCor approaches. In the case of TC, the datasets of the triplet must be independent so neither the NAvg nor the FAvg can be used. Instead, SM products from SMOS and AMSR2 satellite sensors and an ECMWF model are selected and two types of triplets are configured: the AMSR2-triplet formed by one pixel time series, SMOS and AMSR2 time series, and the ECMWF-triplet formed by one pixel time series, SMOS and ECMWF time series.

4.4.1 Datasets

The area of study is Yanco and the selected period is 2014/09 - 2015/03 (6 months), which ensures that the number of gaps in gridded SM data is minimum. All datasets are sampled and temporally collocated to match SMOS overpass times over the study area ($\Delta t = 0.5$ days). The area of study includes all the locations of the Yanco stations plus a 0.05° extension to avoid borders effects in the CArea method.

Regarding wavelet decompositions, the largest possible level of decomposition J0 is 8 (64 days) and the wavelet type chosen is Haar, since it is the only whose length at scale J0 (256 samples) is smaller than or equal to the series length (364 samples).

The gridded SM dataset was described in the previous section (4.3). The SMOS series used in this experiment are extracted from the ESA Level-2 soil moisture product (version 620) for the analysis period, which was described in section 3.3.3. The other two large-support datasets, AMSR2 and ECMWF, are described below. Finally, the pre-processing of all the datasets is summarised in Table 4.2.

4.4.1.1 AMSR2 data

In May 2012, the Japan Aerospace Exploration Agency (JAXA) launched the GCOM-W1 mission, with the Advanced Microwave Scanning Radiometer-2 (AMSR2) sensor onboard. AMSR2 is a passive multi-band scanning radiometer that scans the Earth surface with a fixed incidence angle of 55°. Its revisit time is 1-2 days with crossing nodes at 1:30 p.m. and 1:30 a.m. local solar time for ascending and descending orbits. In this thesis, I have selected the SM products derived from the lowest AMSR2 frequency band (6.9 GHz, C-band). Although more sensitive to vegetation than the lower SMOS frequency (1.4 GHz, L-band), the band selected is the least sensitive among the available AMSR2 bands. At this frequency, the resolution is ∼35 x 61 km^2 (along scan x along track) (JAXA, 2013) and the approximate sensing depth is 1-2 cm, so shallower than the SMOS one.

Different models are available to translate the AMSR2 observed brightness temperatures (TBs) into SM. The **Land Parameter Retrieval Model (LPRM)** is selected since it performs better than others, specially in the dry part of the SM range and so, in the Yanco region (Draper et al., 2009; Rüdiger et al., 2009). The LPRM is developed by the Vrije Universiteit Amsterdam (VUA) in collaboration with the National Aeronautics and Space Administration (NASA). LPRM is based on a simple radiative transfer equation for polarized C and X signals (Owe et al., 2008; Parinussa et al., 2012). The surface characteristics are considered homogeneous within the pixel (vegetation scattering albedo, surface roughness, etc.). Since the product distributed by the NASA services showed unusual temporal patterns and biases (Cho et al., 2017), in this thesis I use a LPRM-AMSR2 soil moisture dataset generated by Dr Parinussa from the re-calibrated TBs provided by the JAXA.

AMSR2 observations are nearly top surface so the accuracy of the derived SM products is highly sensitive to the surface temperature, which is largely more uniform and stable at the descending overpass (1:30 a.m.) than at the ascending overpass (1:30 p.m.). The descending overpass products have been demonstrated largely more accurate than their ascending counterparts in relatively sparsely-vegetated regions (e.g. Draper et al., 2009; Lei et al., 2015). For this reason, the descending products are the only ones automatically produced by the VUA's team of Dr Parinussa and will be the ones used here. The AMSR2 pixel that is closer to the selected SMOS grid node is chosen. The AMSR2 time series are resampled in time to match the SMOS overpass times: each temporal sample is associated to the closest SMOS overpass time in local standard time. The *in situ* samples were selected at the SMOS overpass times, so the temporal resampling of AMSR2 may influence its respective scores. Finally, I discard the SM observations equal to 0.0 m^3/m^3 and those with quality values higher than 68, which were associated with anomalous SM retrievals. The AMSR2 data is summarised in Table 4.2.

4.4.1.2 ECMWF data

The European Centre for Medium-range Weather Forecasting (ECMWF) data is obtained from the top 0-7 cm SM layer of the ECMWF forecast system that is used by the SMOS L2 processor as initial guess in the retrieval loop. It is automatically interpolated in space and time to match the SMOS L2 grid and overpass times by the SMOS processor. More information on this ECMWF auxiliary product can be found in (Kerr et al., 2012, 2014, 2016). A summary of the pre-processing of ECMWF data can be found in Table 4.2.

It should be noted that the characteristics of this custom ECMWF product may foster similarity with the SMOS soil moisture product. The SMOS retrieval algorithm consists, basically, on trying several SM values and only keeping the one whose modelled TB better fits with the observed TB. The first value tried or *initial guess* is the ECMWF forecast, so the SMOS retrieved soil moisture depends on the custom ECMWF product and the optimisation strategy. Moreover, the SMOS radiative transfer models need the effective soil temperature to translate TB into

emissivity (then into reflectivity, dielectric constant and SM). The effective soil temperature is estimated from the soil temperatures of the ECMWF forecast model, which may also contribute to a better connection between ECMWF and SMOS soil moisture.

Table 4.2 – Pre-processing of the data used to assess the spatial representativeness in the Yanco region

Yanco	
Period	2014/09 - 2015/03 (6 months)
Wavelet type	Haar
Δt	0.5 days
J0	8 (64 days)
gridded SM data	
Locations (selected/total)	13 / 13
% gap-filling (regression)	0 %
% gap-filling (DCT)	23.3 %
Nb. consecutive gaps (on avg.)	5
Nb. consecutive samples (on avg.)	17.4
SMOS data	
Node	8174767
Orbits	asc. + desc.
% gap-filling (DCT)	56 %
Nb. consecutive gaps (on avg.)	2.8
Nb. consecutive samples (on avg.)	2.2
AMSR2 data	
Pixel	Closest to SMOS node
Orbits	desc.
% gap-filling (DCT)	65.6 %
Nb. consecutive gaps (on avg.)	1.9
Nb. consecutive samples (on avg.)	1
ECMWF data	
% gap-filling (DCT)	48.3 %
Nb. consecutive gaps (on avg.)	2.5
Nb. consecutive samples (on avg.)	2.8

4.4.2 Methods applied to full time series

Initially, TC, TStab and CArea were not designed for time-scale decomposed series, so in this section I analyse their relative performances on full and de-trended time series, which are the nominal cases. Figure 4.5-a shows the spatial representativeness scores of the set of selected locations of the gridded SM dataset. The CArea scores, obtained for a $R_{th} = 0.55$, appear at the left, the TC scores in the middle and the TStab scores at the right of the plot. Note that each of the methods have a different vertical axis which is orientated so that the maximum of representativeness is in the top. Some locations are not present in the TC groups because the TC estimate of the error variance was negative. This happens when the variance of the gridded SM dataset is smaller than the product of the cross-covariances in Equation 4.7. This could be due to temporal biases.

In Figure 4.5, the ranking of the locations obtained with the different methods is not the same. TStab exhibits the largest disagreements compared to the other methods, probably because Yanco is an agricultural site so human decisions, such as cropping and irrigation, might impair the temporal stability of the spatial distribution of SM (Yee et al., 2016). In any case, differences between the rankings provided by the other methods are still important, even between the two TC variants. Given that seasonal biases can play a significant role as suggested in TC publications (e.g. Miralles et al., 2010; Chen et al., 2016), I have also conducted the same analysis after detrending the datasets for variations larger than 32 days (the 32-days wavelet smooth series S_8 is subtracted from the original time series).

Figure 4.5-b shows the representativeness scores after detrending the datasets. TStab could not

be applied because the wavelet detrending produced multiple zero-crossings of the series (detail series are zero-mean by construction), which would imply divisions by zero. Therefore, TStab will not be used in the subsequent experiments in this chapter. Concerning the two TC variants, now the ordering of the locations is more similar. This improvement is probably because the wavelet detrending has successfully removed the positive bias that AMSR2 series showed during the first half of the analysed period (not shown here). Seasonal/month-scale biases in sensors operating at frequencies higher than L-band can appear because of their increased sensibility to vegetation and atmospheric factors. This was observed, for example, in the LPRM product from the SSM/I sensor (Ku-band) (Dorigo et al., 2010).

Despite detrending, the TC and the CArea approaches still provide different results (Figure 4.5-b). They both attribute more spatial representativeness to the network average and pixels #4, #5, #8, #9, #10, and smaller representativeness to locations #1, #2, #6, #7. However, some locations (#3, #12 and #13) show large discrepancies. As a conclusion, detrending helps to improve the match between methods.

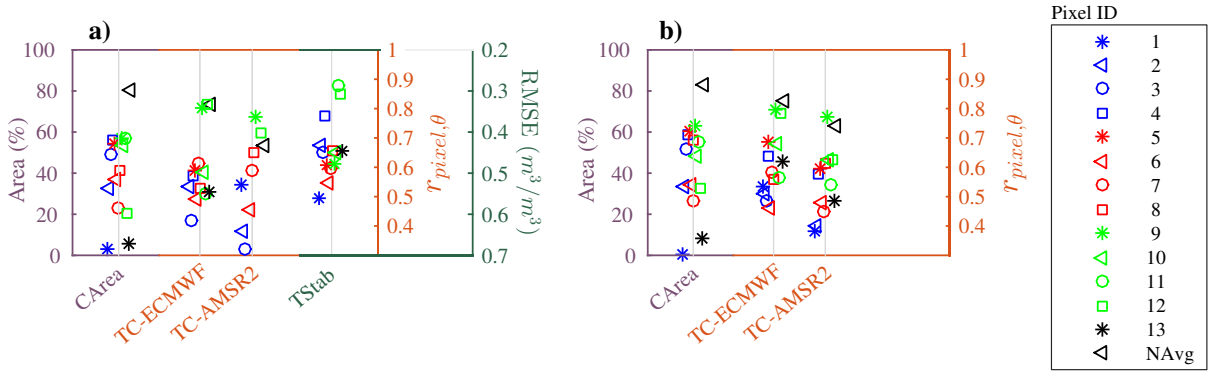


Figure 4.5 – Spatial representativeness scores (vertical axis) from different methods (horizontal axis) for different pixels of the gridded SM dataset. The representativeness is evaluated on a) full series and on b) detrended series (components > 32 days are removed). The CArea correlation threshold is 0.55. Source: (Molero et al., 2018)

4.4.3 Methods applied to decomposed series

Since detrending improves the match between methods, it can be inferred that some time scales should allow a better match than others. In order to link time scales and levels of representativeness, it is needed first to select the method that is able to better assess the spatial representativeness at all time scales. This section compares the performances of different methods on a per time scale basis. The WCor method is added to the analyses.

The results are presented in Figure 4.6. Each plot corresponds to a different time scale and it can be interpreted similarly to those in Figure 4.5. Multiple locations are missing from the TC groups at half-day, 1-day, 32- and 64-days scales: some of them were off vertical axis limits and some of them failed the TC preliminary tests (section 4.2.1.2). The principal reason of the absence of TC scores here is that some datasets exhibited very low cross-correlations (< 0.4) at those scales. Very low cross-correlations between some of the datasets is an indicator of the violation of the linear model assumption and can also produce very low TC correlations $\hat{r}_{k,\theta}$ and estimated negative error variances $\hat{\sigma}^2_{\epsilon_k}$. The correspondence between the low cross-correlations and the TC scores is illustrated with the help of Figure 4.7, that includes the datasets cross-correlations per time scale for two sample cases, the NAvg and station #4. It can be observed that when their correlations are below 0.3 points (black, red and magenta lines), their TC scores are missing from Figure 4.6

In Figure 4.6, **TC rankings** largely differ from WCor and CArea rankings. The best match between methods appears at the 4- and 8-day scales (regardless TC-AMSR2). The match

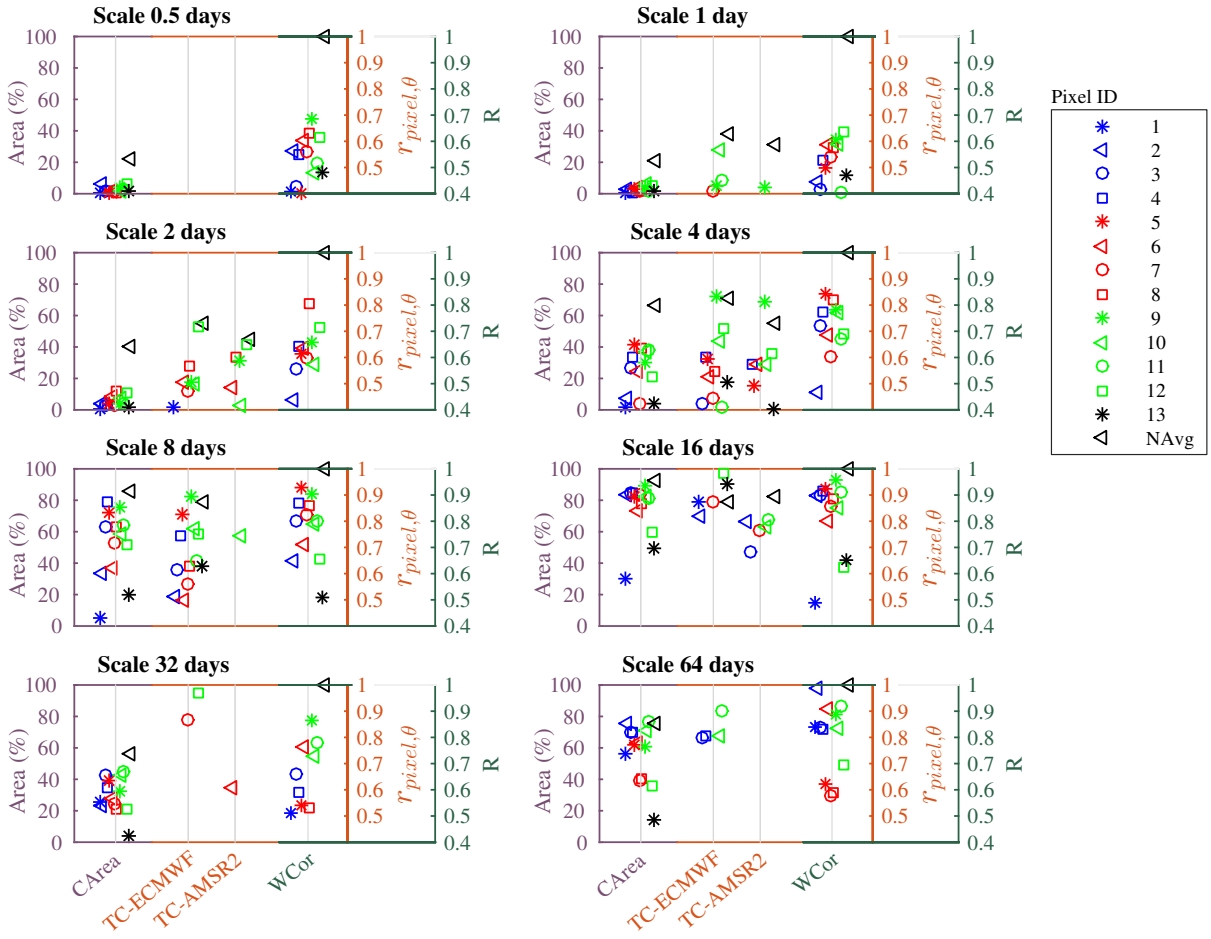
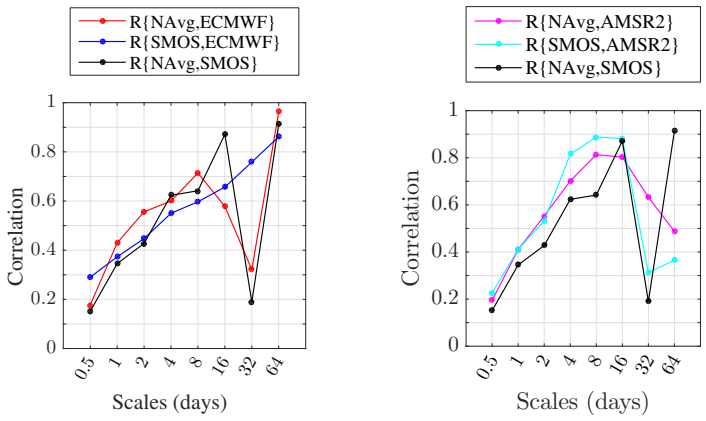


Figure 4.6 – Spatial representativeness scores from different methods for different pixels of the gridded SM dataset. The representativeness is evaluated on wavelet-decomposed series (detail series). Source: (Molero et al., 2018)

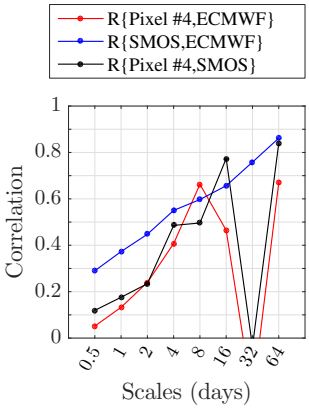
between CArea and WCor is particularly good at the 16-day scale but not with TC scores. In fact, at scales equal or larger than 16 days, the number of independent samples is drastically reduced as a result of wavelet filtering, so the performance of TC approaches is impaired. The usual recommendation for conducting TC analyses is to have a sample size of at least N=100 (Dorigo et al., 2010) or N=500 (Zwieback et al., 2012). The sample size here is 366 so the number of independent samples likely falls below the 100 threshold at large scales. Note that the exact estimation of the number of independent samples in this type of wavelet decompositions (MODWT) is not straightforward as explained in Appendix C. At small scales ($\leq$ 2 days), the mismatch between TC and the other methods is probably due to the very low correlation between the datasets (Figure 4.7). In conclusion, TC is not convenient in the case of too short series or in the case of small time scales.

In contrast to TC, the results of **WCor and CArea** approaches are consistent: the rankings of the locations are similar for all time scales. This is a significant result because WCor assesses the representativeness of a location by comparison to the network average (NAvg), while CArea assesses the representativeness of a location by comparison to all the other individual pixels: the correlation between a point series and the average is not simply the average of point-to-point correlation values. In conclusion, **WCor is a robust method for the evaluation of spatial representativeness on a per-time scale basis** when the surface is not fully sampled and the CArea method cannot be applied. This is the typical situation in validation of satellite SM products.

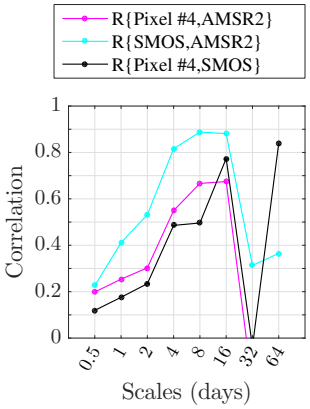


(a) NAvg-SMOS-ECMWF

(b) NAvg-SMOS-AMSR2



(c) Pixel #4-SMOS-ECMWF



(d) Pixel #4-SMOS-AMSR2

Figure 4.7 – Correlation between the detail series of the TC datasets. The left column represents the TC-ECMWF triplet and the right column, the TC-AMSR2 triplet. In the top row, the gridded SM dataset is the network average, in the bottom row, the gridded SM dataset is the pixel #4 time series

Conclusions

1. TStab cannot be applied to wavelet decomposed series because of their multiple zero-crossings
2. WCor and CArea give consistent results
3. TC scores are similar to WCor and CArea scores except when cross-correlation between the datasets is low (sub-weekly scales) and when the number of independent samples is low (seasonal scales)

4.5 Assessment of the spatial representativeness of decomposed *in situ* series

I have shown the geophysical drivers of the different spatial/temporal scales of some SM modelled datasets (sections 3.4.1 and 4.3). When different methods for assessing spatial representativeness were compared, WCor performed better than TC in modelled datasets (section 4.4). All this has prepared the analysis of actual observed SM data that will be addressed in this section. The spatial representativeness of *in situ* SM series from the Little Washita (LW) network will be assessed, on a time-scale basis, with the TC and WCor methods using different footprint-support datasets: the network *in situ* average (NAvg), SMOS, AMSR2 and ECMWF. The CArea approach will not be used because the spatial sampling is not sufficient.

4.5.1 Datasets

The selected period is 2012/07 - 2014/07 (2 years) and ensures that there is a minimum number of gaps in the *in situ* series, that a number of whole years is selected and that all datasets are available (AMSR2 observations started the 2012/05/18). As in the previous sections, all datasets are sampled and temporally collocated at SMOS overpass times over the study area. LW *in situ* series are gapfilled following the same procedure as in chapter 3 (section 3.3.4). The pre-processing of all the datasets is summarised in Table 4.3.

Regarding wavelet decompositions, the last level of decomposition is J0=9 (128 days) and the wavelet type chosen is Daubechies-4. Daubechies-4 better isolates time-scales than Haar and, although at scale 9 the length of this wavelet filter is longer than the series length ($L_j = 1534 > N = 1460$), the wrapping due to circularisation is minimum and only concerns very low energy coefficients.

Table 4.3 – Pre-processing of the data used to assess the spatial representativeness in Little Washita.

Little Washita	
Period	2012/7 - 2014/7 (2 years)
Wavelet type	Daubechies-4
Δt	0.5 days
J0	9 (128 days)
***In situ* data**	
Stations (selected/total)	20 / 20
% gap-filling (regression)	6 %
% gap-filling (DCT)	1 %
Nb. consecutive gaps (on avg.)	3.1
Nb. consecutive samples (on avg.)	714.8
SMOS data	
Node	226157
Orbits	asc. + desc.
% gap-filling (DCT)	58.1 %
Nb. consecutive gaps (on avg.)	2.8
Nb. consecutive samples (on avg.)	2
AMSR2 data	
Pixel	Closest to SMOS node
Orbits	desc.
% gap-filling (DCT)	79.3 %
Nb. consecutive gaps (on avg.)	3.8
Nb. consecutive samples (on avg.)	1
ECMWF data	
% gap-filling (DCT)	48 %
Nb. consecutive gaps (on avg.)	2.5
Nb. consecutive samples (on avg.)	2.7

4.5.2 Representativeness in dense networks

The representativeness scores obtained for the Little Washita *in situ* series with the TC and WCor methods are presented in Figure 4.8. Different configurations of the methods are possible, depending on the large-support dataset used. This section analyses their respective scores in the context of *in situ* dense networks, where the NAvg is available. In the following section, the same Figure 4.8 will be again studied but in the context of sparse networks, where no NAvg is available.

Comparison of TC with WCor-NAvg

Looking at the TC-ECMWF, TC-AMSR2 and WCor-NAvg scores in Figure 4.8, the patterns described by the spatial representativeness and the time scales are consistent with what was observed in modelled datasets in sections 4.3 and 4.4. The **overall spatial representativeness** increases with the time scale and the seasonal scales (64 and 128 days) exhibit the largest dispersion in representativeness scores among the stations. Unlike in the experiment of the previous section (Figure 4.3) where a drop in representativeness scores appeared at the 32-day scale, here the drops appears at the **64-day** and does not concern WCor-NAvg scores. Instead, it concerns all other methods and configurations: TC scores and the WCor values calculated with SMOS, AMSR2 and ECMWF datasets. A possible explanation for this is that the LW network only covers ∼1/4 of a footprint, so the locations should have more significant differences in small spatial scale factors (precipitation, texture, vegetation) with respect to the satellite observations than to the *in situ* NAvg. As explained in sections 4.3 and 4.4, such small scale heterogeneities can be propagated to larger time scales, producing drops in the correlation diagram at monthly and seasonal scales.

Regarding the different TC configurations, much more scores are present than in the gridded SM experiment, probably because the number of samples is four times bigger (2 years of data instead of 6 months). This helps also improving the matching between the representativeness rankings provided by the TC-ECMWF, TC-AMSR2 and WCor-NAvg methods. While in the experiment based on 6-months gridded SM data, only the 4- and 8-day scales showed a good match between methods, here the good match applies to 4 scales, the 2-day to 16-day scales.

Regarding the **largest scales (32- to 128-days)**, it can be observed again a reduction in the number of TC scores, due to the reduction in the number of independent samples after wavelet filtering. Despite this, the consistency between WCor and TC results still hold good at the last 128-day scale, because the correlation between datasets is extremely good at that scale. For example, both methods designate stations #3, #11, #14, #15, #17, #19 as the most representative and stations #2, #4, #5,#8, #16 and #20 as the least representative ones. Figure 4.9 shows the wavelet correlation between the TC datasets when the representativeness of the NAvg is evaluated (a,b) and when the representativeness of station #8 is evaluated (c,d). There is a clear correspondence between the correlation values in Figure 4.9 and TC representativeness scores in Figure 4.8 at the 128-day scale: station #8 triplets show very low correlation so station #8 is not present among the TC scores. Similarly, the NAvg triplets show very high correlation when ECMWF and not AMSR2 is used as 3rd dataset, so the NAvg is present among the TC-ECMWF scores but not among the TC-AMSR2 scores.

Comparison between WCor configurations

The WCor-SMOS, WCor-AMSR2 and WCor-ECMWF scores follow the progression of increasing representativeness with increasing time scale as the WCor-NAvg. However, their representativeness values at **scales smaller than 4 days** are much smaller than the WCor-NAvg values. They are actually off axis limits. There are two causes that can explain this:

(i) Geophysical spatial scale mismatch: The geophysical station-NAvg distance is much smaller than the station-satellite distance, since satellite sensors provide a SM observation that is the result of the full integration of the observed surface.

(ii) Gap-filling: The gap-filling proportion of SMOS, AMSR2 and ECMWF is much higher (> 48 %) than that of the *in situ* series (7 %) and gaps are not distributed similarly.

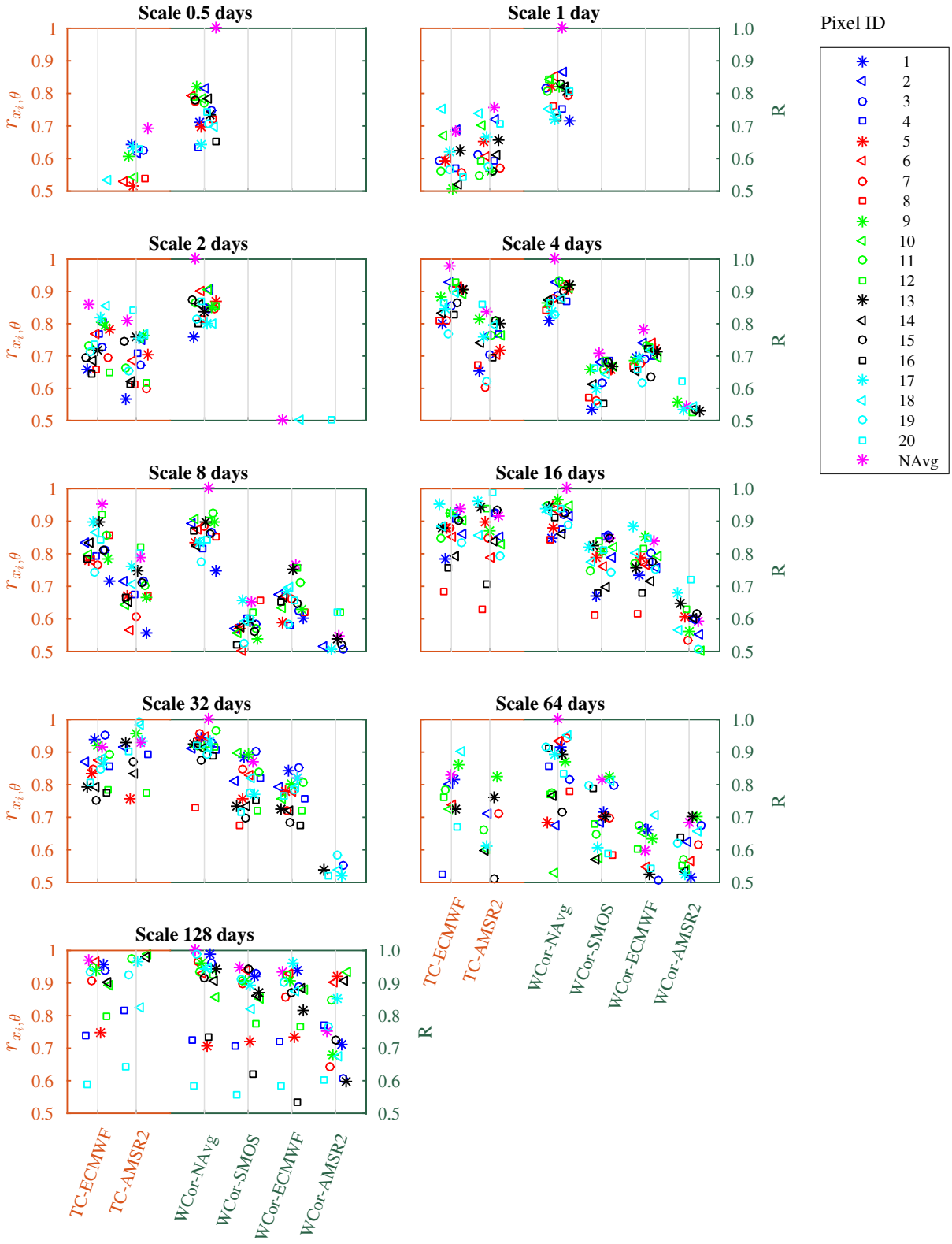
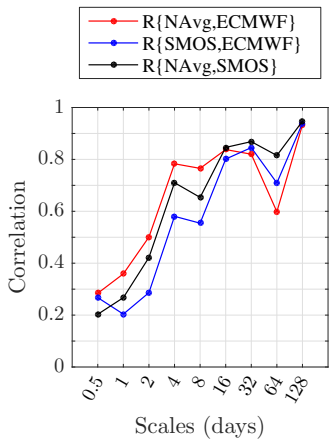
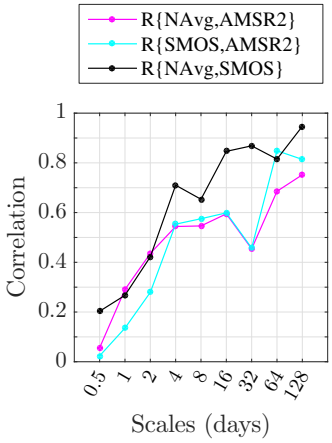


Figure 4.8 – Spatial representativeness scores for different *in situ* locations of the LW network. The representativeness is evaluated on wavelet-decomposed series (detail series). Source: (Molero et al., 2018)

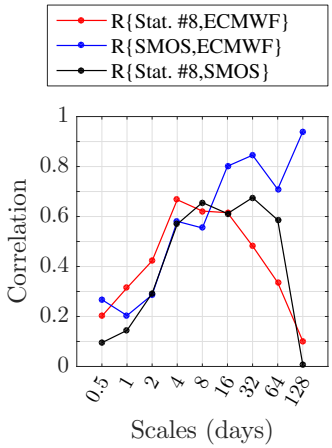
In addition, Figure 4.9 reveals a surprising feature: at sub-weekly scales, the correlation between a location (station #8) and the large-support datasets is is the same range (0 - 0.3) as the cross-correlations between the large-support datasets. Su and Ryu (2015) showed similar wavelet correlation values at the same scales for the MERRA model and AMSR-E, both large-support datasets too. Can it be said then that the low scores at sub-weekly scales are due to the spatial scale mismatch or that it is simply noise? The main contribution should still be the spatial scale mismatch because the representativeness scores obtained when the NAvg is used as large-support dataset are also low. However, when SMOS, AMSR2 and ECMWF are compared, more elements come into play. For example, there is a mismatch in observing times and in sensing depths between AMSR2 and the other datasets. Moreover, the physics of the instruments are different and they are more or less sensitive to different changes in the surface. AMSR2 is more sensitive than SMOS to top-surface elements and phenomena occurring at small time scales (wind, solar exposition, etc.) may change the perceived signal.



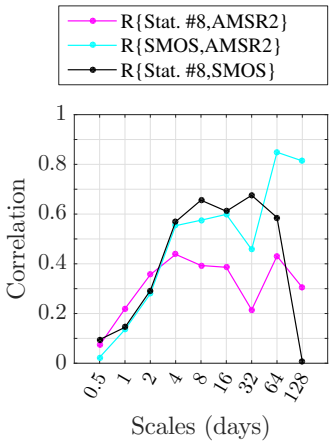
(a) NAvg-SMOS-ECMWF



(b) NAvg-SMOS-AMSR2



(c) Station #8-SMOS-ECMWF



(d) Station #8-SMOS-AMSR2

Figure 4.9 – Correlation between the detail series of the TC datasets. The left column represents the TC-ECMWF triplet and the right column, the TC-AMSR2 triplet. In the top row, the *in situ* dataset is the network average, in the bottom row, the *in situ* dataset is the station #4 time series

4.5.3 Representativeness in sparse networks

Sparse networks (only one station is available per satellite footprint) are even more common in satellite validation than dense networks. In the case of sparse networks, the large-support observation can only be provided by satellites or models. In this section, I compare the performances of WCor and TC approaches when based on large-support satellite and model data. The respective scores are also those in Figure 4.8.

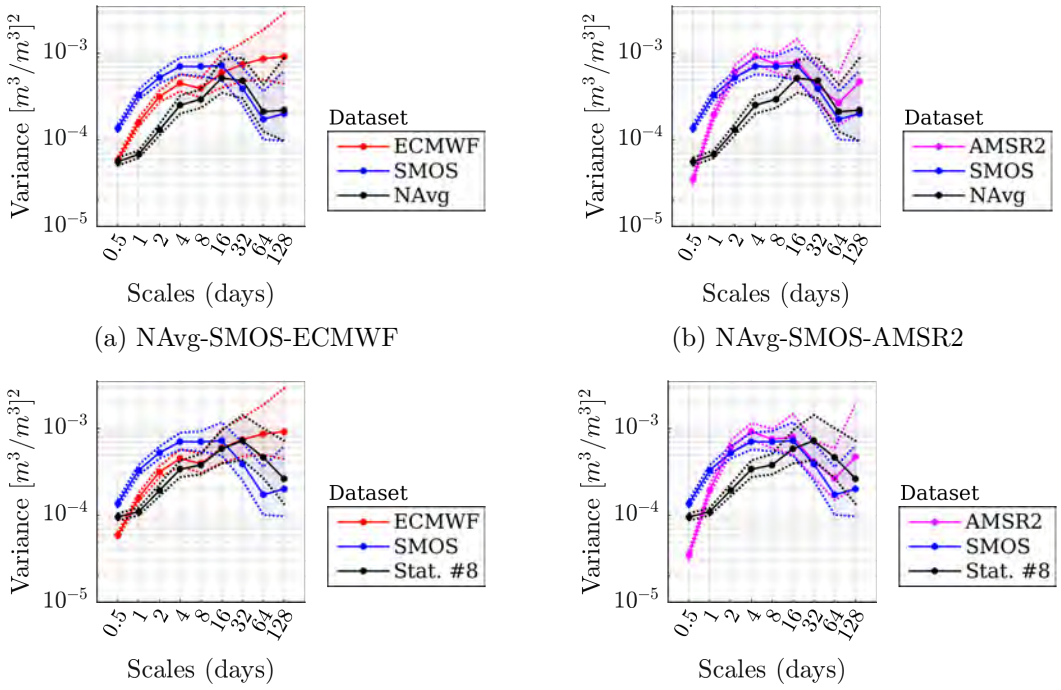
Comparison of TC and WCor configurations without the NAvg

When comparing the different TC and WCor configurations, similarities are found by groups (Figure 4.8): TC-ECMWF rankings match well with the WCor-SMOS/ECMWF ones (1st group), while TC-AMSR2 rankings match well with the WCor-AMSR2 results (2nd group). This highlights that both TC and WCor methods have a high sensitivity to the choice of the large support dataset. The consistency between methods is lost at the **64-day** scale for the 2nd group and at the last **128-day** scale for the 1st group, which is probably due to the reduction in the number of independent samples that impairs TC performance. ECMWF results are closer to SMOS than to AMSR2 results probably because their temporal gaps were similarly distributed and because SMOS is not totally independent from ECMWF soil moisture and temperatures as explained in section 4.4.1.2.

There are 3 possible large-support observations, different from the NAvg, available for the LW region. Which **large-support dataset is the most appropriate** to track the evolution of spatial and temporal scales? It is easy to see that, in this study, the WCor-NAvg results are more similar to the results of the 1st group (WCor-SMOS/ECMWF) than to the 2nd group (WCor-AMSR2). The much larger number of gaps and the sensing time mismatch in AMSR2 series impairs the performance of its group. At middle scales (4-16 days), the ranking of WCor-ECMWF is the most similar to the WCor-NAvg ranking (Figure 4.8). I attribute this to SMOS observational noise. At last scales (32-128 days), the situation is reversed and WCor-SMOS is closer to WCor-NAvg. Therefore, SMOS can be considered as a good large-support dataset to be used for spatial representativeness assessment in the Little Washita region, especially at the month and seasonal scales. Moreover, gap-filling does not seem to influence the performance of SMOS.

Finally, although I have based the evaluation of the spatial representativeness on correlation metrics, it is worth showing the **variance decomposition** of the different datasets analysed. Figure 4.10 includes the wavelet variance diagrams of the TC triplets built for evaluating the representativeness of the NAvg (top) and of station #8 (bottom). This completes the analysis started in section 3.4.3 (Figure 3.15). First, as suggested in section 3.4.3, the variance diagrams are distinctive for each type of sensor. Also, it is confirmed that the higher SMOS variance at 0.5-8 day scales is partially due to noise. AMSR2 variance keeps lower in the first two scales probably because a larger number of gaps were DCT filled, which smooths the temporal transitions so it acts as a low-pass filter.

ECMWF variance keeps lower also for scales larger than 1 day. The dynamics of ECMWF series are governed by its model structure so it does not show the same levels of high-frequency noise as sensors can show. Finally, SMOS and AMSR2 are the observation systems with more similar variance patterns. SMOS is the system that better reflects the NAvg variance pattern at large time scales, suggesting again that it is specially suitable for the Little Washita region at month and seasonal scales.



(a) NAvg-SMOS-ECMWF

(b) NAvg-SMOS-AMSR2

(c) Station #8-SMOS-ECMWF

(d) Station #8-SMOS-AMSR2

Figure 4.10 – Wavelet variance of the TC datasets. The left column represents the TC-ECMWF triplet and the right column, the TC-AMSR2 triplet. In the top row, the *in situ* dataset is the network average, in the bottom row, the *in situ* dataset is the station #8 time series

Conclusions

1. The results confirm the conclusions of previous sections on modelled data:
   (a) Spatial representativeness tends to increase with the time scale.
   (b) Dispersion of spatial representativeness scores also increases with the time scale.
   (c) Stations show uniformly low spatial representativeness at sub-weekly scales.
2. Sub-weekly time scales may be also affected by mismatch in sensing depths, in sensing times or in unavailability of data.
3. The representativeness scores obtained with satellite datasets show tendencies that cannot be revealed with the *in situ* average: lower scores at sub-weekly scales and deterioration of the scores at some specific scales. This may be due to their different approaches of obtaining the footprint SM (averaged energy *vs* averaged SM).
4. In sparse networks the representativeness scores are highly sensitive to the satellite and model selected as large-support dataset.
5. DCT gap-filling of SMOS and AMSR2 partially compensates the noise of the sensors at the two first scales. Gap-filling benefits the correlation at the first two scales when the series have similar gap schemes (ECMWF, SMOS) and slightly reduces the correlation when the series have different gap schemes (*in situ*, satellite).

4.6 Discussion: links with chapter 3

The correlation patterns showed in sections 4.3, 4.4 and 4.5 are similar to those showed in chapter 3 (section 3.4). The modelled Little Washita dataset in section 3.4.1 served to illustrate how the differences in specific temporal components of the SM signal influence the wavelet correlation diagram. Three components were distinguished: impulse responses, dry-down decays and

seasonal dynamics. They were represented by distinct parameters in the model of section 3.4.1: precipitation, soil texture and vegetation leaf area index (LAI).

The differences in one or more of these components may be due to spatial heterogeneity: two different locations where SM is measured may differ in precipitation, soil and vegetation characteristics; similarly, precipitation, soil and vegetation are perceived differently by point-support and by large-support sensors. For these reasons, the conclusions of chapter 3 can be extrapolated to this chapter.

The experiments of this chapter have revealed that locations have very small spatial representativeness at sub-weekly time scales. The model-based experiment in the previous chapter (section 3.4.1) indicated that, at sub-weekly scales, low correlation between different SM series was due to precipitation. Therefore, it can be derived that the main reason of the small representativeness at sub-weekly scales is the heterogeneity in precipitation, and maybe in other short time scale elements, within the satellite footprint. The precipitation at a specific location would be fundamentally different from the "footprint precipitation".

Another aspect highlighted in section 3.4.1 was that the propagation of short time scale responses, such as precipitation and dry-down decays, could induce irregular correlation patterns. This has been observed in the gridded SM dataset (section 4.3 and 4.4) and the *in situ* SM dataset (section 4.5). In the latter, the representativeness scores calculated by using satellite and model SM as large-support datasets revealed a drop in the correlation level at the 64-day scale. The satellite sensor does not perceive small time-scale components in the same way the point *in situ* sensor does, so it seems plausible that such differences integrate over time and reduce correlation scores at scales with low SNR.

Finally, one of the conclusions of chapter 3 was that the seasonal scale was made of 3 components: the influence of seasonal-varying geophysical factors (e.g. vegetation), the integration of small time-scale signatures and static fields (e.g. topography). As a consequence, the spatial scale mismatch might be highly variable at seasonal scales depending on the location analysed. This justifies that, in the figures presented in this chapter, some locations are highly representative while others are very poorly representative of the footprint SM estimates. It is also true that the last time scales have higher statistical uncertainty due to a reduction in the number of independent samples and a large number of boundary samples after wavelet filtering. This makes difficult to compare locations with similar moderate representativeness values at seasonal scales (e.g. correlation values between 0.5 and 0.8). Nevertheless, extremely high and extremely low representativeness scores (e.g. correlation >0.9, <0.4) respond to clearly similar and opposed temporal dynamics, respectively, that can only be due to geophysical reasons. Examples of WCor confidence intervals at seasonal scales can be seen in Appendix C.

4.7 Conclusions

This chapter has been dedicated to the study of the interactions of spatial and time scales within typical coarse scale satellite footprint-size areas. For this purpose, I have evaluated the spatial representativeness of different locations for a comprehensive range of time scales. The decomposition in time scales has been achieved with wavelet transforms and the representativeness has been assessed with various methods: triple collocation (TC), temporal stability analysis (TStab), correlated area (CArea), an approach adapted from the inverse footprint method and wavelet-based correlation (WCor), a new approach proposed here.

The analyses have been conducted on both model-based and measured SM data from Yanco and Little Washita networks. These experiments have confirmed that there exists a connection between spatial and temporal scales: the spatial scale increases with the time scale. However, this increasing relationship is not always observed. At **sub-weekly scales**, all locations are equally not representative of the footprint observations. **From sub-weekly to seasonal scales, the**

spatial representativeness tends to increase in overall. However, the representativeness scores obtained for the set of selected locations **also become more scattered**. At the end, at seasonal scales, locations are either highly or very poorly representative of the footprint observation.

The justification for the observed connections between spatial and temporal scales can be found in the factors driving SM variability. As described in section 4.1, SM-related processes usually exhibit characteristic spatial and time scales that are connected, small spatial scales with small time scales and large spatial scales with large time scales. For example, convective precipitation events are heterogeneous within the footprint and do not last in time for more than one day. Texture is fundamentally heterogeneous within the footprint and contributes to the temporal dry-down pattern whose maximum duration is around 3 days. The results of this chapter are consistent with the analysis of chapter 3: heterogeneity in precipitation and soil characteristics may induce low spatial representativeness values at sub-monthly scales, irregular correlation patterns may be related to the integration of the precipitation and soil temporal responses over time and seasonal scales may be influenced by large scale factors like global temperatures but also by the integration of within-footprint heterogeneities.

The experiments proposed have also helped to solve the question stated in section 4.1.4 about the spatial scale of seasonal trends. This was motivated by the apparent divergent interpretations concerning seasonal scales I noticed in former studies. In my experiments, a large number of locations exhibited good synchronism with the large-support dataset at those scales, which confirms the results of Su and Ryu (2015). On the other hand, I also found that many locations did not follow the footprint seasonal patterns, corroborating that in TC studies the seasonal trend can be a source of non-linearities (Gruber et al., 2016). In conclusion, both phenomena, similar and dissimilar seasonal scales co-exist within the same footprint.

Finally, a valuable outcome of this chapter is the inter-comparison of methods for the assessment of spatial representativeness. It revealed that TStab could not be applied to wavelet decomposed series because their multiple zero-crossings prevented the TStab equation to be solved. TC provided representativeness scores that were consistent with CArea and WCor scores, principally when the cross-correlation values between the datasets was above 0.4 and when the number of samples was large enough (> 100). Very low cross-correlation values were systematically observed at sub-weekly scales and, as a consequence, TC scores could not be computed. The reduction in the number of independent samples was specially harmful at large time scales (≥ 64 days in this case), so many TC scores were missing or highly uncertain. **In contrast to TC, CArea and WCor results were consistent in general at all time scales**. Since WCor is less sensitive to the spatial sampling size than CArea, I would recommend to use WCor as a reference approach when working with real measured data (dense and sparse *in situ* networks) and when time scale decompositions are desired.

This is, to my knowledge, the first comprehensive investigation on the connection between SM spatial and time scales within the satellite footprint ($\sim$50 km). The understanding of these interactions can help to improve validation: the spatial scale mismatch could be estimated from the temporal observations and then included in the validation error budget. Meanwhile, small time scales could be removed from the validation of satellite products: it has been shown that locations are not representative of the footprint at such time scales. Finally, time decompositions along with the WCor method are promising tools for improving not only the satellite validation but also the modelling of surface SM. Multi-scale algorithms could be built based on the specific interactions at each time and spatial scales. Given this time-scale dependence, spatial variability should be addressed differently depending on the time scale.

Chapter 5

Uncertainty in temporal validation statistics due to spatial sampling

Contents

5.1 Assessment of the sampling uncertainty of individual locations (SUL) 108
5.1.1 How to evaluate the contribution of the SUL to validation metrics? . . . 109
5.1.1.1 Proposed approach . . . 109
5.1.1.2 Datasets . . . 110
5.1.2 Temporal model . . . 111
5.1.3 Spatial model . . . 112
5.1.4 Spatio-temporal simulations . . . 113
5.1.5 Calculation of statistics and final results . . . 116
5.1.6 Inter-comparison of validation statistics affected by the SUL . . . 118
5.1.7 Discussion . . . 119
5.2 Assessment of the sampling uncertainty of the mean (SUM) . . . 120
5.2.1 How to assess the uncertainty of the mean? . . . 120
5.2.2 Bootstrap approaches . . . 122
5.2.3 Dependency between the SUM and the sample size . . . 123
5.2.3.1 Datasets . . . 123
5.2.3.2 Methodology . . . 123
5.2.3.3 Experiment and results . . . 124
5.2.4 Comparison of different SUM estimation approaches . . . 126
5.2.4.1 Datasets and methods . . . 126
5.2.4.2 Results . . . 127
5.2.5 Inter-comparison of validation statistics affected by the SUM . . . 128
5.3 Conclusions . . . 129

The previous chapters have addressed the spatial scale mismatch as a whole, without differentiating its 3 components: the sampling uncertainty, the scaling uncertainty and the geophysical mismatch. The **sampling uncertainty (SU)** refers to the lack of knowledge about the required sampling scheme of the study area: what would have happened if the area had been sampled differently? It is the only component of the spatial scale mismatch that is independent from satellite observations.

When a satellite time series is compared against an *in situ* time series a set of summary statistics is obtained. They are uncertain to some level due to the SU, since they would have been different if we had sampled at another location. However, we cannot know to what extent. In practice, the objective is not only to estimate the SU but to evaluate its influence in satellite validation statistics. It should be kept in mind that the SU varies in time but it has a spatial origin. In contrast, satellite validation statistics are temporal.

Two different SU cases can be distinguished: the **sampling uncertainty of the locations (SUL)**, when the satellite observations are compared to time series from individual ground stations x_i, and the **sampling uncertainty of the mean (SUM)**, when the satellite observations are compared to the sample spatial average of multiple ground series (the network average or NAvg). The scope of this chapter is to provide procedures to evaluate the uncertainty in the satellite (temporal) validation statistics that is due to the SUL and the SUM, in the form of confidence intervals. The premise is that the approaches developed should be easily applicable to any region of study, so that a standard protocol could be established in the future in satellite validation campaigns.

5.1 Assessment of the sampling uncertainty of individual locations (SUL)

One of the most frequent cases of satellite SM validation approaches consists in comparing the satellite observations with the measurements of only one *in situ* station. Such is the situation when validating using sparse networks like SCAN or SNOTEL. The *in situ* series and thus, the statistics obtained, would be different if the *in situ* samples were taken at a different location: they are affected by the sampling uncertainty of the locations (SUL).

Intuitively, if SM-related factors such as soil composition, vegetation, topography and precipitation, affect homogeneously the study area, then the sampling uncertainty is expected to be small. For a given surface size and number of sampled locations, the SUL is more important as the spatial variability of the area increases. In addition, the SUL is not constant in time because the spatial distribution of SM is not either. Chapter 2 showed that the spatial distribution of SM and its respective mean and variance, vary in time.

Satellite SM products are usually validated by comparing them to point *in situ* SM measurements under the form of time series. Their similarity is evaluated through a set of metrics, namely correlation (R), bias (B) and standard deviation of the differences (STDD). The statistics obtained are temporal but they are affected by the SUL: they would have been different if the ground station had been placed at another location within the satellite footprint. The approximate influence of the SUL in validation statistics can be observed in one of the figures of the introduction, Figure 1.8. There, one SMOS SM time series was compared with the *in situ* series measured at different stations within the footprint and a wide range of possible statistics was obtained.

The SUL is critical when comparing the performances of different large-support SM products. This is easily explained with an example. We compare SMOS and AMSR2 soil moisture time series with the *in situ* time series measured at station i_1. The respective correlations obtained are 0.7 and 0.8. Then, we compare the SMOS and AMSR2 series to a different station, i_2, and

we obtain 0.75 and 0.61 instead. Which of the two satellite SM products is closer to the *in situ* measurements then?

The comparison between the SM products of the two satellites would be easier if we could draw the confidence intervals (CIs) of the temporal correlation considering the SUL of the area of study. If we knew that the 90 % CIs are [0.6, 0.77] for SMOS and [0.61, 0.81] for AMSR2, then we could affirm that, at the view of the SUL and the statistics at the sampled locations, SMOS and AMSR2 seem equivalent in terms of correlation for the region of study.

This section aims to provide an approach for translating the SUL into uncertainty of the temporal validation metrics. As stated in the introduction of this chapter, the approach should be as generic as possible. Section 5.1.1 describes the approach in overall and the sub-sequent sections (5.1.2 - 5.1.7) describe each of the steps.

5.1.1 How to evaluate the contribution of the SUL to validation metrics?

As explained before, the SUL is intrinsically dependent on the spatial variability of SM. Therefore, it can only be estimated if the spatial variability is estimated first. The typical spatial sample size of dense networks allows to obtain multiple statistics for the same satellite footprint (10-20 stations so 10-20 values). The dispersion of these statistics gives an idea of the impact that the SUL has in them, but is not sufficient to directly deduce confidence intervals. Some approach is needed to "fill" the distributions of statistics.

5.1.1.1 Proposed approach

For a given footprint area A, the computation of the CIs of the validation statistics could be derived if we had access to the distributions of all the possible statistics within that area. Let us suppose that at each time step t, SM(t) is a continuous spatial field, the distributions of statistics could be obtained if we had access to all the point SM series in the area, $x_i(t)\ \forall i \in A$. The statistics could be simply obtained by comparing the satellite time series $y(t)$ and each of the point series $x_i(t)$ (Figure 5.1).

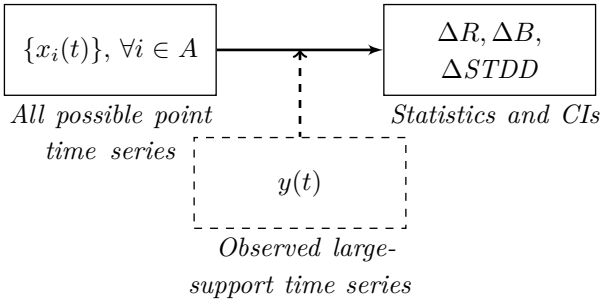


Figure 5.1 – Necessary elements to obtain the distribution of validation statistics within a satellite footprint of size A.

In reality, we do not have access to all the possible point time series of the region, but models can represent them and the continuous A area can be approximated by high-resolution grids. There are two main types of models: physical and empirical. **Physical models** describe the underlying physics of the process or target variable. Most of the hydrological and land surface models that provide spatialized SM as an output are physically-based. This implies that they are designed for specific environmental conditions or for a specific region of study. Moreover, the statistical spatial distribution of the *true* SM field is not necessarily reproduced by the model. Due to this, physical models do not seem appropriate for building a standard procedure applicable to all validation regions.

Empirical models describe the relationships between a set of variables (explanatory variables) and the response variable (SM in this case), relationships that cannot necessarily be explained from a physical point of view. The structure and coefficients of empirical models can be specific to the area of study, but empirical models are usually simpler and much easier to build than physical models. In some cases, the same model structure can be re-used for different regions. For these reasons, empirical (or semi-empirical) modelling is selected in this chapter to simulate SM time series.

With empirical modelling, if we have K measured *in situ* series in the region of study, then K models can be built to reproduce their dynamics. However, the question on how to generate *all the possible* time series within the region remains to be answered. Moreover, some kind of spatial information should be necessary to spatialize them. We usually have access to spatial information of the region of study because intensive ground campaigns are not rare in validation regions. However, it is not clear how to harmonize both temporal and spatial sources of information.

As a summary, the procedure to evaluate the uncertainty in temporal validation metrics due to the SUL should follow the following steps:

1. Find the models that match both the temporal dynamics of measured SM series and the estimated spatial SM distribution:
   1.1. Find an empirical temporal model $f_i^{temp}(t)$
   1.2. Find a spatial model $g_A^{spat}(t)$. The spatial model varies in time because the spatial distribution of SM also varies in time
2. Simulate point time series $x_i(t)$ according to both the temporal and the spatial models
3. Compare each of the simulated series to the large-support series (satellite) $y(t)$ and derive confidence intervals for the statistics $\Delta R, \Delta B, \Delta STDD$

This is illustrated in Figure 5.2 and described in detail in the following sections. It is important to highlight that the temporal model should reproduce quite accurately the dynamics of the *in situ* measurements so that the simulated series can replace *in situ* measurements in the validation of satellite products.

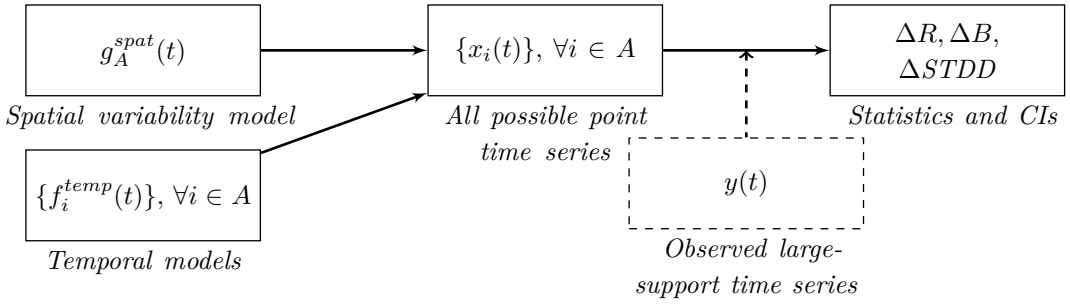


Figure 5.2 – Procedure to estimate the effect of the sampling uncertainty in validation statistics.

5.1.1.2 Datasets

The feasibility of the whole procedure is tested for the Little Washita validation region. Examples with the SM datasets acquired for this region will be included at each of the steps of the procedure.

The *in situ* dataset consists in the SM time series measured at **5 Little Washita stations (#4, #6, #9, #10 and #13)** for the period 2013/01 - 2013/06 (6 months). Only 5 out of 20 were selected because the aim is to simply show the feasibility of the procedure presented in this chapter. These 5 specific stations were selected because they did not present any gaps for the

period of study and temporal semi-empirical models with a reasonably good fit were easily found for them. I dit not try to find a model for the other stations of the network, so they could be used for further validation in future studies. As satellite datasets, both SMOS L2 and AMSR2 LPRM soil moisture time series for the same time period and region are used. The series are filtered with the same quality thresholds as in previous chapters (SMOS: RFI $\leq$ 10 % and DQX $\leq$ 20 %; AMSR2: quality mask $<$ 68). A detailed description of the SMOS and AMSR2 soil moisture products was included in sections 3.3.3 and 4.4.1.1, respectively.

5.1.2 Temporal model

The scope of this section is to find an empirical or semi-empirical temporal model to generate *all the possible* point time series $x_i(t)$ for the Little Washita region. SM time series have two characteristic temporal patterns: sudden peaks due to precipitation and exponential-like dry-down decays due to various factors (e.g. infiltration, soil type, vegetation). These two features are well represented in the temporal semi-physical model of Pan et al. (2003), already used in chapter 3. Unfortunately, although Pan's model was able to represent well the characteristic SM dynamics, the fit between the output modelled SM and the measured SM was not satisfactory enough for the current application (they differed in decay patterns and the minimum and maximum values, see Figure 3.6). The scope is now to find a model able to exactly reproduce the measured SM series. In other words, over-fitting is desired in this particular context so that the modelled series $x_i(t)$ can be considered as true *in situ* series.

I propose to develop an empirical polynomial model with a similar structure to that of Pan's: at time $t+1$, the increase of SM depends on the previous precipitation $prec_t$ and decays follow an exponential function whose exponent depends on the soil depth Z, the sampling period Δt and the previous vegetation leaf area index LAI_t. The Pan's model depended on additional physical parameters like drainage, SM loss and vegetation interception. The model developed here is considered semi-empirical because it relies in a reduced number of physical parameters and variables and their functional relationships are not physically-based.

The general structure of the **temporal model** $f(\cdot)$ is given by the following two equations:

$$f(SM_t) = SM_{t+1} = \begin{cases} SM_t + f_p(prec_t, SM_t), & \text{if } prec_t > 0 \\ SM_t \cdot e^{-f_d(LAI_t, SM_t)\cdot \frac{\Delta t}{Z}}, & \text{otherwise} \end{cases} \tag{5.1}$$

where $f_p(\cdot)$ and $f_d(\cdot)$ are polynomial functions, also referred as the *peak model* and the *decay model*, respectively. The $f_p(\cdot)$ and $f_d(\cdot)$ models are obtained by least-squares fitting (Levenberg-Marquardt method) of different polynomial functions to the observed SM, precipitation and LAI data. The equations obtained are:

$$f_p(prec_t, SM_t) = a_{00} + a_{10}SM_t + a_{01}prec_t + a_{20}SM_t^2 + a_{02}prec_t^2 + a_{11}prec_tSM_t \tag{5.2a}$$

$$f_d(LAI_t, SM_t) = b_{00} + b_{10}SM_t + b_{01}LAI_t + b_{20}SM_t^2 + b_{11}LAI_tSM_t \tag{5.2b}$$

The same model structure (but with different a and b coefficients values) was found for the 5 Little Washita stations The construction of the models and the obtention of the a and b coefficients for the peak and decay models is described in detail in Appendix E. As an example, the modelled series for station #9 is shown in Figure 5.3. The fit is not perfect, but represents with sufficient accuracy the measured SM series.

Finally, it is important to highlight that the temporal models here are designed for the finest temporal sampling available, Δt=30 min in the case of Little Washita series.

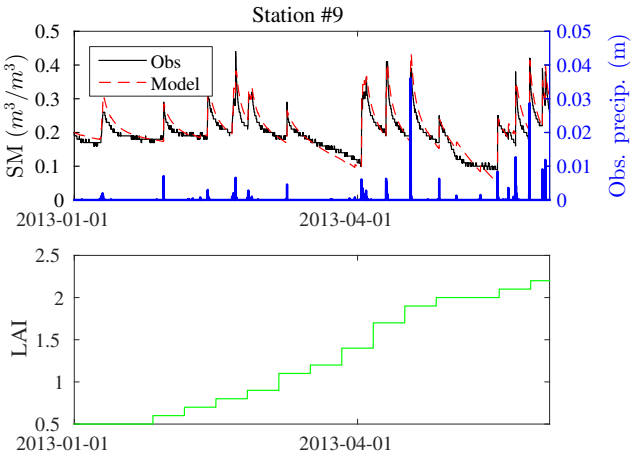


Figure 5.3 – Final model fit for station #9

5.1.3 Spatial model

The only physical variables used by the temporal model are precipitation, LAI and SM. A way to spatialize the temporal model would consist in providing high-resolution maps of the 3 variables, SM included. Obviously, if we knew the true SM field for each time step, there would not exist any sampling uncertainty problem to solve. Instead, we can estimate the statistical spatial distribution of SM. On the one hand, the spatial PDF of SM can be estimated from intensive ground sampling, which is usually conducted when a new *in situ* network is deployed or during specific temporary validation campaigns. On the other hand, the spatial PDF of SM has been studied several times in literature. This literature was reviewed in Chapter 3. It was concluded that **the spatial distribution of SM can be inferred from the spatial mean SM level**: the distribution function is similar to a beta, right- and left-skewed to the limits of the SM range (∼0-0.5 m^3/m^3) and symmetric bell-shaped in the middle SM range. This was true at least for areas similar to the Little Washita region, with no abrupt topography and low homogeneous vegetation.

I propose to use this knowledge to build the spatial model. In the case of the Little Washita region, I developed a simplistic model based on existing literature since I did not have access to intensive ground sampling data. Three important **assumptions** are made:

(i) The spatial PDF at each time step is modelled as a **beta function**
(ii) The mean of the available *in situ* samples at each time step (the network average, **NAvg**) is a good representation the most probable value of the population (mode)
(iii) The standard deviation of the available *in situ* samples at each time step ($\hat{\boldsymbol{\sigma}}_{\boldsymbol{t,insitu}}$) is a good representation of the standard deviation of the population

Assumption (i) is valid except if fractional precipitation occurs within the footprint, in which case the distribution is bi- or multi-modal (Ryu and Famiglietti, 2005), depending on the contrast between the areas affected and not affected by precipitation. Regarding assumption (iii), the variance calculated with 5 stations is highly uncertain (the sample variance could be far from the population variance), more uncertain than the respective sample mean. However, the level of uncertainty is unknown. To my knowledge, there is not any study relating the sample size and the spatial SM variance within satellite footprints. The only few related publications are dedicated to the spatial mean or the coefficient of variation (CV) at best, and their results are very different: the necessary spatial sample size ranges from 50 locations in 160x160 fields (Wang et al., 2008) to only 6 locations in 10^2 km^2 (Brocca et al., 2012).

Finally, the spatial distribution of SM at time t can be expressed as:

$$X_t \sim B_{\alpha,\beta}(mode = NAvg, \sigma = \hat{\sigma}_{t,insitu}) \tag{5.3}$$

The α and β parameters of the beta distribution $B_{\alpha,\beta}$ are derived from the values of the mode and the standard deviation (the calculation is not shown here).

Figure 5.4 shows the spatial PDFs produced with a standard deviation of 0.05 m^3/m^3 and modes 0.1, 0.25 and 0.4, from left to right. The minimum and maximum possible values of the beta

distributions have been rescaled to 0 and 0.5, respectively. These are considered as the minimum and maximum possible SM values for the region of study.

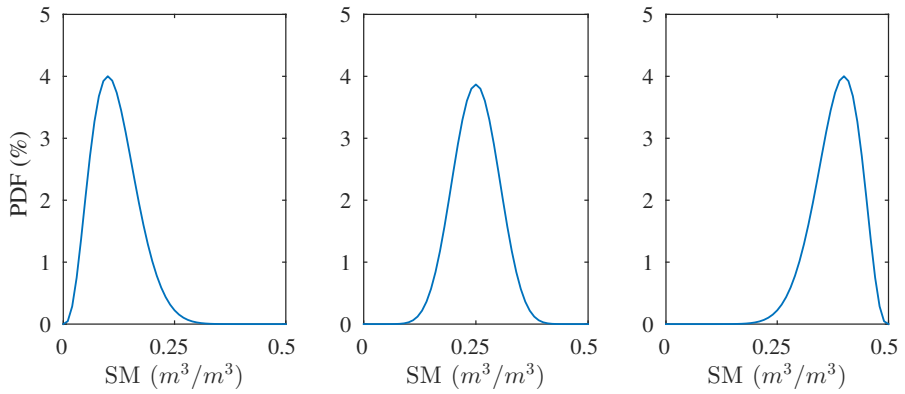


Figure 5.4 – Spatial model: spatial PDFs produced with a standard deviation of 0.05 m^3/m^3 and modes 0.1, 0.25 and 0.4, from left to right

5.1.4 Spatio-temporal simulations

Once the temporal and spatial models are obtained for the region of study, the simulation of $x_i(t)$ time series can be undertaken. The scope is to produce a Monte Carlo simulation of M point soil moisture time series that respect both the statistical spatial law and semi-empirical temporal law:

$$\text{M simulated point time series:} \quad \{x_1(t), x_2(t), \ldots, x_M(t)\} \qquad t = 0, 1, \ldots, N-1$$

A possible solution is to randomly sample from the modelled spatial distributions at each time step. Unfortunately, nothing ensures that the dynamics of the temporal model $f(\cdot)$ are respected: the output time series will have a very high level of white noise and they will not exhibit typical SM dynamics. A second alternative would consist in obtaining M samples from the spatial distribution at t=0, $\vec{x}(t=0) = \{x_1, x_2, \ldots, x_M\}_{t=0}$, then propagate them with the temporal model. This ensures that the spatial distribution at t=0 and the temporal laws are respected, but it does not ensure that the spatial distributions at t≠0 are respected.

$$\text{M simulated samples at time } t: \quad \vec{x}(t) = \{x_1, x_2, \ldots, x_M\}_{t=0}$$

The second alternative could work if a way to harmonize the propagated samples with the spatial distribution at t≠0 were found. This is actually possible with statistical mapping approaches. I propose to match the distribution of the propagated samples, $f(\vec{x}(t))$ or $\vec{x}_{temporal}$, with the distribution of the spatial model at the same time step, $\vec{x}(t+1)$ or $\vec{x}_{spatial}$. This mapping approach can be performed through **CDF matching** or *quantile mapping*. The CDF of a random variable X is defined as the probability of X to have a value smaller or equal to x:

$$CDF(X) = Pr[X \leq x] \tag{5.4}$$

The CDF matching is a non-linear transformation that can be applied to any distribution type. The transformation is suitable when the distributions have similar shapes. In the case of $\vec{x}_{temporal}$ and $\vec{x}_{spatial}$, the CDF matching is expressed as

$$\vec{x}_{temporal,NEW} = CDF^{-1}_{spatial}(eCDF_{temporal}(\vec{x}_{temporal})) \tag{5.5}$$

where $eCDF_{temporal}(\cdot)$ denotes the empirical CDF function of the distribution of propagated samples and $CDF^{-1}(\cdot)$ is the inverse of the theoretical CDF of the spatial distribution.

Here, the eCDFs are approximated by dividing the range of $\vec{x}_{temporal}$ values in as many steps (quantiles) as there are values. Then, values in between the quantiles are approximated to the closer lower quantile. The inverse CDFs (spatial model) are obtained analytically for a beta distribution function whose parameters are derived from *in situ* measurements (Equation 5.3).

The spatio-temporal simulation is done sequentially as follows.

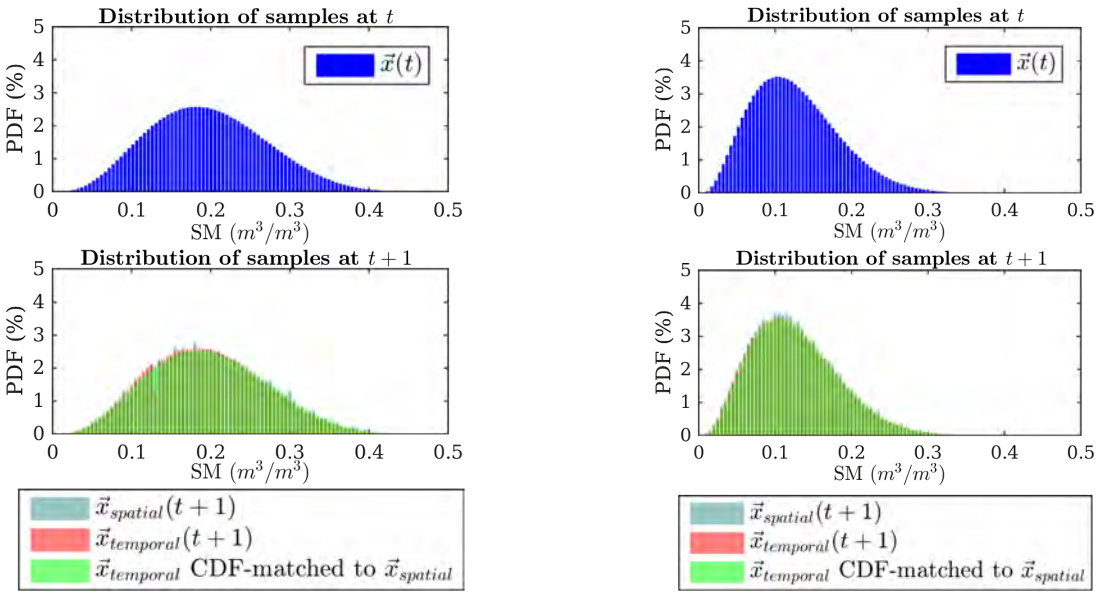
While $t < N$:

1. At t=0, randomly sample M times from the spatial PDF $X_{t=0}$ (Equation 5.3):
$$\vec{x}(t=0) = \{x_1, x_2, \ldots, x_M\}_{t=0}$$
2. Propagate $\vec{x}(t)$ to t+1 with the temporal model $f(\cdot)$ (Equation 5.1):
$$\vec{x}_{temporal} = f(\vec{x}(t))$$
3. CDF matching: transform the propagated samples to match the theoretical spatial distribution at t+1 (Equation 5.5):
$$\vec{x}_{temporal} \rightarrow \vec{x}_{temporal,NEW}$$
4. Update the time series:
$$\vec{x}_{t+1} = \vec{x}_{temporal,NEW}$$
5. Update the time step and go back to step 2
$$t = t + 1$$

There is still a remaining issue. If the network is represented by K stations, then K temporal models are available, with their respective coefficients and the input variables for the concerned locations. In the Little Washita sample case, K=5 stations were considered (#4, #6, #9, #10 and #13). **Which of the 5 temporal models should be used** to propagate the samples **in step 2**? In this experiment, the models are assigned randomly to the samples generated at t=0 following an uniform distribution. The assignation is done only at t=0: at the following time steps, the samples that have been propagated with a specific model, will be again propagated with the same model. This implicitly assumes that each of the stations represents 1/K of the surface. Note that if the relative representativeness weights of the stations were known, then the random selection could be done based on that distribution. For example, if station #4 represented 30 % of the area, then 30 % of the spatial samples could be randomly selected and propagated with the temporal model of station #4.

The full procedure is tested with the Little Washita example for M=10,000 simulations. Figure 5.5 displays the distributions obtained at t=3000 and t=6000 (Figures 5.5a and 5.5b, respectively). The distribution of SM samples at time t is in the top row, while the CDF-matching operation is illustrated in the bottom row. The differences between $\vec{x}_{spatial}$ (red) and $\vec{x}_{temporal}$ (grey-blue) are small. This is expected in the absence of precipitation since the difference in SM is very small between 2 consecutive time steps (Δt is 30 min in the Little Washita series). The similarity between $\vec{x}_{spatial}$ and $\vec{x}_{temporal}$ is also beneficial for the CDF-matching approach, which succeeds to match the $\vec{x}_{temporal}$ samples to the theoretical spatial CDF. For example, in Figure 5.5a, the frequency of the x=0.14 m^3/m^3 value is corrected from 1.25 % to 2 %.

The M=10,000 simulated series for the Little Washita region are shown in the 2 top rows of Figure 5.6: in the top row, each series is represented by a different coloured line and in the 2nd row, the density of series is represented in a degraded grey colour scheme. The simulated series seem to successfully follow the typical SM dynamics in the region. The average simulated series and the average measured series differ a little as expected, since the average of the 5 measured series was used as the mode of the spatial distributions. The spatial standard deviation (3rd row)



(a) At time step t=3000

(b) At time step t=6000

Figure 5.5 – Distribution of samples before and after CDF-matching

is mostly respected by the simulation. The largest differences between the standard deviation of the simulated and that of the measured series appear when the fit between the simulated and the real dynamics is the worst. This is evident when Figure 5.6 is compared against Figure E.4 in Appendix E. When comparing the simulated and the measured series, the periods from the 10th to the 31st of March and from the 25th of April to the 15th of May are the ones that show the largest differences in SM (Figure E.4) and thus, in standard deviation (Figure 5.6).

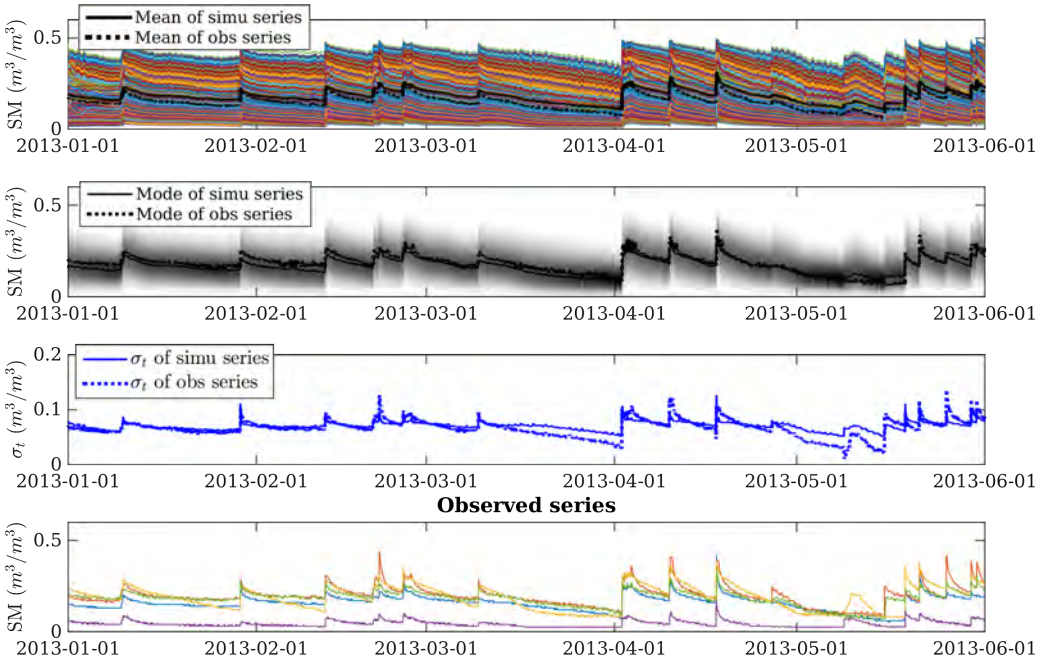


Figure 5.6 – Simulation of M=10,000 point series for the Little Washita region. In the top row each coloured line is one different simulated series and in the 2nd-top row, the density of series is represented in gray. The 3rd row shows the spatial standard deviations of the 5 measured series and of the 10,000 simulated series. The bottom row shows the 5 measured *in situ* series.

5.1.5 Calculation of statistics and final results

Once the simulation of the point series $x_i(t)$ has finished, they can be compared to the large-support (satellite) time series $y(t)$ and the respective temporal statistics can be computed. For example, in the case of correlation (R), M possible scores will be produced as follows:

$$R_i = R\{x_i(t), y(t)\}, \quad i = 1, 2, \ldots, M \tag{5.6}$$

Considering the Little Washita experiment, the correlation (R), bias (B) and standard deviation of the differences (STDD) are computed for 3 cases, depending on the large-support dataset used: a) the spatial average of *in situ* series (network average, NAvg), b) the SMOS time series and c) the AMSR2 time series. The STDD (also called *unbiased* RMSE), is chosen instead of the root mean square error (RMSE) because it is not affected by the bias:

$$STDD = \sqrt{E\left\{\left[(y(t) - \hat{\mu}_y) - (x_i(t) - \hat{\mu}_{x_i})\right]^2\right\}} = \sqrt{RMSE^2 - B^2} \tag{5.7}$$

where $y(t)$ is the large-support SM series, $x_i(t)$ one of the simulated point SM series and $\hat{\mu}$ is the sample temporal mean.

In the example, only the samples with simultaneous availability of *in situ*, SMOS and AMSR2 observations are kept, so that the statistics obtained for each pair are comparable. The final total number of temporal samples is 52. The selected large-support time series are shown in Figure 5.7, together with the simulated series (top row) and the 5 *in situ* series considered (bottom row). The temporal dynamics of the NAvg, SMOS and AMSR2 seem in agreement witht those of the measured and simulated point series.

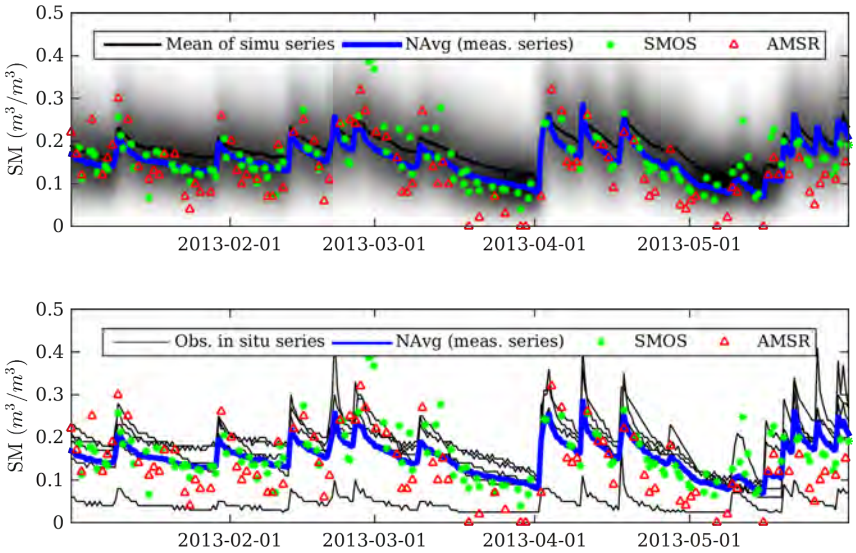
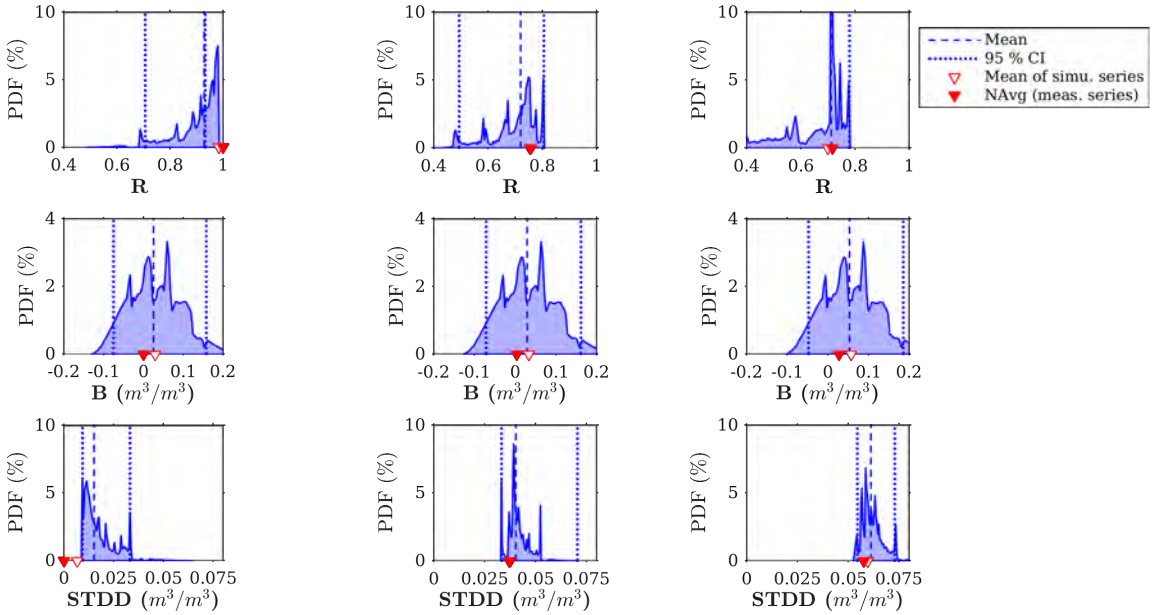


Figure 5.7 – The network average (NAvg), SMOS and AMSR2 series on top of the simulated point series for the Little Washita region

Figure 5.8 shows the distributions of statistics obtained from the comparison of the simulated time series and the 3 large-support datasets (the NAvg (a), SMOS (b) or AMSR2 (c)). In addition, each large-support dataset is compared to the average of the simulated series (vertical red line) and to the NAvg (vertical red bold line). The distance between these 2 lines is small in all cases. This supports the use of **the simulated series as a reasonable proxy of the overall temporal dynamics of the area**. The other relevant aspect is that the distributions here are far from normal, which justifies the use of non-parametric methods like Monte Carlo for assessing the uncertainty in temporal metrics. This is not surprising since the statistical formulas of the correlation and of the STDD are not linear, so even if the compared datasets

followed a normal distribution, the outputs would not necessarily do the same. Finally, it is also striking that most of the distributions in Figure 5.8 present **peaks**. They are associated with the specific temporal dynamics of the 5 *in situ* series used as modelling references. Most probably, increasing the number of reference series would help produce smoother distributions.



(a) Comparison with the NAvg (b) Comparison with SMOS (c) Comparison with AMSR2

Figure 5.8 – Comparison of the network *in situ* average (NAvg), SMOS and AMSR2 time series with the simulated point series

The spread of the distributions is considerable in all the plots of Figure 5.8. Regarding **correlation**, the simulated series show correlation values with the NAvg that range between 0.7 and 1, and with SMOS between 0.45 and 0.8. In the AMSR2 case, most of the correlation samples are concentrated around a peak between 0.7 and 0.8. This peak is hidden in the picture by the vertical lines and accounts for around 20 % of the samples. It is worth reminding that the distributions presented are not exact representations of the spatial variability of the region, but are good **representations of the range of statistical possibilities** (assuming beta distributions). This implies that, in the case of SMOS and AMSR2, we could rarely expect correlations above 0.8, supposing that the *in situ* series used as reference were sufficiently representative of the temporal dynamics of the region.

Regarding the **bias** (Figure 5.8, middle row), the shape of the distributions is the same for the 3 large-support datasets, which is expected since the distribution of means of the simulated series is the same for the 3 cases. The bias uncertainty is large, much more than the STDD uncertainty, of approximately 0.2 m^3/m^3 between the 5-95 % CI. This may be an artefact of the models and the simulation. Whether the same distribution of biases could be found in reality needs further investigations and dedicated *in situ* campaigns.

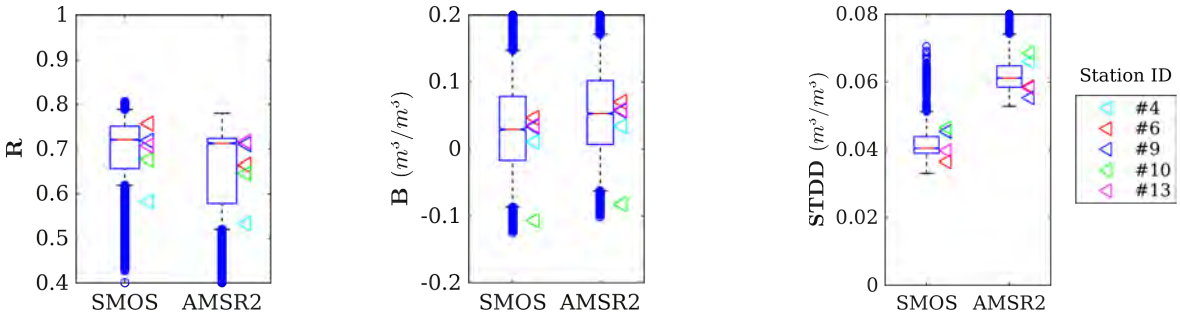
The dispersion of **STDD** is small although non-negligible: around 0.025 m^3/m^3 for the 3 cases (Figure 5.8, bottom row). It seems that the models used produce variations mainly in correlation and bias. As suggested before, further field-work studies are needed to verify these results. The NAvg is the large-support dataset that whose error structure is the most similar to the *in situ* dataset. For this reason, the comparison between the simulated series and the NAvg serve as a **minimum bound for the SUL**. The lower limit of the STDD distribution for the NAvg is 0.01 m^3/m^3, so it can be stated that the minimum error we could expect due to sampling uncertainty is 0.01 m^3/m^3.

5.1.6 Inter-comparison of validation statistics affected by the SUL

The final objective is to state whether the difference between the validation statistics of two different SM products are significant or not given the SUL. The two SM products under validation can concern different satellites, models or dataset versions. Although the terms *significance* and *confidence interval* (CI) are not equivalent, CIs provide a good first approximation. More about significance and CIs can be found in the shaded box below.

A way to ease the interpretation is to present the information in Figure 5.8 as box plots. The statistics for the SMOS and AMSR2 cases are presented under the form of box plots in Figure 5.9, where the median and the 5 %, 25 %, 75 % and 95 % percentiles are depicted. Although trivial, this representation is important since it compares directly the validation of two satellites taking into account the sampling uncertainty. For further information, the statistics obtained for the 5 reference *in situ* series when compared to the satellite datasets are depicted with triangle markers. The markers are within the limits of the distributions, which confirms that the simulations were reliable enough. Moreover, a part of the simulated series present better statistics than those of the measured *in situ* series, confirming that the simulation did not simply consist in injecting uncorrelated noise.

Figure 5.9 indicates that the differences in correlation and bias between SMOS and AMSR2 analyses are not significant. For example, supposing that only station #6 was available, its correlation with SMOS is higher (0.75) than with AMSR2 (0.67) but given the spread of the distributions in Figure 5.9a, it cannot be assumed that this difference in correlation is significant: if we had sampled at another location, SMOS would not have been necessarily better than AMSR2. Similarly, no meaningful differences are found for the bias. On the contrary, the distributions of STDD are well separated and so the almost 0.02 m^3/m^3 of separation between the STDD of station #6 and SMOS and that of station #6 and AMSR2 is significant, even after taking into account the sampling uncertainty in this region.



(a) Correlation (b) Bias (c) STDD

Figure 5.9 – Validation of SMOS and AMSR2 datasets with 5 *in situ* series with simoultaneous assessment of the statistical sampling uncertainty (box plots). In the box plots, the middle red line is the median, the limits of the box are the 25 % and the 75 % quartiles and the whiskers are the 5 % and the 95 % percentiles of the uncertainty distribution. Statistics falling outside the 90 % percentile are represented as blue circles.

A note about confidence intervals (CIs) and significance. CIs and statistical significance are intrinsically related, although they are not the same. A 90 % CI, for example, is the range that covers the mean value 90 % of the time, not the probability of the mean to fall inside that range. If the CIs of two statistics do not overlap, then their mean values are significantly different. The opposite is not necessarily true (Schenker and Gentleman, 2001): two statistics can have overlapping CIs, yet be significantly different. The significance of the difference in the statistics is evaluated by obtaining the probability of finding a null difference by chance (hypothesis testing). This is not straightforward when the statistics are obtained for serially correlated data (e.g. Ebisuzaki, 1997; Holm, 2015) and should deserve a full chapter. For this reason, herein significance is evaluated with CIs instead of with hypothesis testing. The aim of this chapter is to simply provide general guidelines on how to assess the impact of the sampling uncertainty in the statistics used in satellite validation campaigns.

5.1.7 Discussion

The procedure presented in this section related the sampling uncertainty of the locations (SUL) to the uncertainty in temporal validation metrics. It has for main advantage to be a **standard procedure** applicable to any region. Empirical temporal models and statistical spatial models can be found for any area under study. The approach only requires some *in situ* SM and precipitation time series measured in the region of study as prerequisites.

However, the procedure presented has some limitations. First, it requires that the ensemble of measured *in situ* SM and precipitation series represent well the temporal dynamics of the region. If they are not sufficient to represent the dynamics, then the confidence intervals obtained for the validation statistics are underestimated. This is particularly important in the case of sparse networks, where only one station per satellite footprint is available to validate satellite observations. Although underestimated, the confidence intervals could still be interpreted as the minimum bound of the uncertainty. For this reason, I recommend to use the procedure also for sparse networks.

Unlike typical model-regression approaches, here **over-fitting** is a desired characteristic. The temporal models should provide outputs that fit the *in situ* dynamics as much as possible. Moreover, the same period is used for constructing the model and for simulation. This seeks to have a distribution of point-satellite statistics that are a good representation of the reality. It is not the first time that satellite SM products are validated with models. However, in such cases one particular remark is always addressed: it is not known whether the model is a good representation of the true SM and measurements are considered better options if they are available. The over-fitting here seeks to circumvent this weak point in this particular context.

Even with over-fitting, the models are still sources of uncertainty, since they are representations of the reality and not the reality itself. The temporal model error can be assessed by direct comparison to the reference *in situ* series used to build it. However, it is not possible to assess the **error of the statistical spatial model** since the region is too large to be densely sampled. Here, I have adopted a simple model that relies on literature review, where spatial distributions are more or less skewed depending on the mean SM level. However, in many occasions, SM spatial distributions exhibit more complex shapes such as bi-modal shapes.

When more than one station is available within the satellite footprint, more than one temporal model is possible. The question about how they should be assigned to each of the simulated samples at t=0 is non-trivial. By default, I have used an uniform random scheme where each model is applied to 1/K of the simulated series, being K the number of *in situ* stations. Other sampling schemes could be used if the representativeness of each station was known.

Another limitation of the procedure is related to the significance of the *in situ* spatial mean

(NAvg) and standard deviation, which are used to build the spatial model. They are calculated based on a small sample size (≤ 20), 5 in the Little Washita sample case, so their uncertainty is high. Their uncertainties can be statistically estimated with non-parametric approaches like bootstrap. The following section covers this aspect.

Finally, it is important to remind that the approach presented here is a **geo-statistical simulation** that generates SM values according to a prescribed statistical spatial law and the typical temporal dynamics of the region. The uncertainty distributions should not be interpreted as the *actual* statistics but as the *possible* statistics for the region of study from a statistical perspective.

5.2 Assessment of the sampling uncertainty of the mean (SUM)

When satellite SM time series are validated with the aid of *in situ* dense networks, they are compared to the network average (NAvg), which is the spatial average of the SM sampled at a set of defined locations. The NAvg is an estimation of the mean SM of the area. The NAvg suffers from the same issues as the SUL and the estimated mean is different depending on the number of samples and their locations. This is usually referred in the statistical literature as the *uncertainty of the sample mean*. Herein, the term **sampling uncertainty of the mean (SUM)** is adopted to stress that the uncertainty is due to sampling. Intuitively, for a given number of sampled locations, the more important the spatial variability is, the higher the SUM should be. Similarly, for a given spatial variability level, the lower is the number of sampled locations, the higher the SUM should be. This section seeks to find a standard procedure for assessing the SUM and its influence on temporal validation statistics.

5.2.1 How to assess the uncertainty of the mean?

The evaluation of the uncertainty in temporal validation statistics due to the SUM concerns two aspects:

1. The assessment of the SUM
2. The propagation of the SUM to the temporal validation statistics

It is necessary to decide first the methodology for assessing the SUM (1^{st} step). The implementation of the 2^{nd} step, the propagation of the SUM to the temporal validation statistics, will be different depending on that 1^{st} step. This section describes the approaches available for assessing the SUM.

There are two main approaches to assess the SUM: geostatistical and purely statistical. **Geostatistical approaches** consist in building a model of the surface based on statistical relationships between different geophysical variables, the output variable (soil moisture) included. The most common example of geostatistical approaches is *kriging*, which was first introduced in the 1950s by Daniel Krige. In short, kriging provides an estimate of the variable under analysis at a specific location based on a linear interpolation of some available measurements. Each measurement location contributes differently to the interpolation depending on the separation distance and the degree of spatial correlation. The two parameters are summarised in a *variogram* that represents the mean variance between two points as a function of their separation distance. **Block-kriging** is a type of kriging that, instead of estimating the value at a location, estimates the average value of a whole area (Cressie, 1993, pp. 284-286; Webster and Oliver, 1992, pp. 150-154). It has been used, for example, for upscaling SM point measurements to the field scale (Wang et al., 2015b). The main advantage of block-kriging is that it provides simultaneously an estimate of the spatial average $\hat{\mu}$, the spatial variance $\hat{\sigma}^2$ and the interpolation error or SUM. To do that, the experimental variogram for the variable of interest (soil moisture) is first computed based on measurements. Then, a model is found for the variogram and the block-kriging equations can be

applied. The estimation of the variogram is based on the underlying assumptions of second-order stationarity (Webster and Oliver, 1992, p. 72) and on a good sample size and spacing (Webster and Oliver, 1992, p. 75-93). Second-order stationarity in space means that the spatial mean and variance of the variable of interest are constant over the area.

Block-kriging could be used to estimate the spatial mean, variability and SUM at the satellite footprint scale from *in situ* measurements. However, SM is not second-order stationary at such extents and the available number of samples is too low (the sampling size in permanent dense *in situ* networks is between 10 and 20 locations). Second-order stationarity could be circumvented with detrending approaches but the sampling size problem is unavoidable. I conducted some tests to evaluate the necessary sampling size to build reliable SM variograms in the Little Washita region. I obtained the empirical variograms of both measured and modelled (DISPATCH) SM data for different sample sizes and I found that a minimum sampling size of around 100 was required for moderately humid conditions. Smaller number of samples produced empirical variograms that were very "noisy" so fitting a variogram model was difficult. The 100 sampling size is consistent with former studies (Webster and Oliver, 1992). Other kriging approaches like block co-kriging could provide smoother variograms, because they estimate the unknown SM information from other related geophysical variables (e.g. temperature, vegetation) available with higher density. However, the related variables used and the detrending are region-dependent. The main premise for the approaches presented in this chapter is that they should be as generic as possible and easily applicable to any validation site. For this reason, block co-kriging is not considered here.

Statistical methods can also provide estimates of the sampling uncertainty of the mean. If the spatial distribution of SM is known, then analytical approaches can be used. When the distribution is normal, the calculation of the confidence interval (CI) for a desired coverage level $1-\alpha$ is straightforward. The $[100 \cdot \alpha/2, 100 \cdot (1-\alpha/2)]$ % CI defines the range that is likely to contain the mean $100 \cdot (1-\alpha)$ % of the times. For example, if $\alpha = 0.10$, then the $[5, 95]$ % CI interval contains the population mean μ with a 90 % probability. The lower and upper values of the CI are calculated as a function of the sample mean $\hat{\mu}$, the value of the *student-t* distribution at $\alpha/2$ confidence level with K-1 degrees of freedom, the sample spatial variance $\hat{\sigma}^2$ and the number of spatial samples K (Helsel and Hirsch, 2002, p. 75):

$$[CI_l, \quad CI_u] = \left[\hat{\mu} - t_{(\alpha/2,K-1)}\sqrt{\hat{\sigma}^2/K}, \quad \hat{\mu} + t_{(\alpha/2,K-1)}\sqrt{\hat{\sigma}^2/K}\right] \tag{5.8}$$

If the population distribution is not normal, the samples have to be transformed to match a normal distribution. This is possible in the case of skewed distributions thanks to log-normal transformations (Helsel and Hirsch, 2002, p. 76) or Box-Cox transformations (Wang, 2001).

When applied to SM samples, the analytical method presents two major drawbacks: First, the spatial distribution of SM is not necessarily known and it is not rare that it exhibits shapes that cannot be transformed to normal (e.g. bimodal distributions). Secondly, in analytical methods, the centre point of the CI is the sample mean, implying that there is no bias between the sample and the population mean, which is not necessarily true. As a consequence, analytical methods will not be considered here.

Non-parametric bootstrap approaches allow to approximate the population distribution and to obtain the respective CI and bias without requiring preliminary knowledge on the population distribution. The population distribution is reconstructed by resampling the available sample. Since no preliminary assumptions are needed, bootstrap is a generic approach that can be used in any validation area to obtain the uncertainty of the spatial SM mean. Since bootstrap generates an ensemble of possible spatial means, these can be directly used for computing the ensemble of possible satellite validation statistics (see next section). For these reasons, bootstrap will be used for assessing the SUM. Bootstrap has also been used in the SM literature to assess the spatial mean and the SUM (Ryu and Famiglietti, 2005; Wang et al., 2008; Cosh et al., 2006)

and to assess the statistical uncertainty of the error variance in validation studies (Yilmaz and Crow, 2014; McColl et al., 2014; Draper et al., 2013).

5.2.2 Bootstrap approaches

Efron (1979) introduced the bootstrap approach for assessing the statistical uncertainty of a summary statistic. The bootstrap is an empirical method, applicable to any population distribution and that does not require preliminary knowledge of the underlying distribution of the statistic. The sample itself is used to estimate the population, which is approximated by resampling the available sample. For this reason, the accuracy of the error estimate is highly dependent on the sample size. The underlying hypothesis of this method is that the sample is representative of the population.

Let us suppose we want to estimate the statistical uncertainty of a statistic $\hat{q}$ that has been computed on a sample $(x_1, x_2, \ldots, x_K)$ derived from the (unknown) population X. A first bootstrap sample is built by randomly resampling it with replacement: $X_b^* = (x_1^*, x_2^*, \ldots, x_K^*)$. Then, the summary statistic q_b^* for that bootstrap sample can be computed: $q_b^* = q(X_b^*)$. The procedure is repeated N_B times so a total of N_B bootstrapped samples and thus, N_B bootstrapped statistics, are obtained. For example, if K=5 and N_B=100, we can have a first bootstrap sample that is $X_1^* = (x_2, x_2, x_5, x_3, x_3)$, a second bootstrap sample that is $X_2^* = (x_1, x_3, x_4, x_5, x_5)$, and so on. This provides an ensemble of bootstrap statistics $(q_1^*, q_2^*, \ldots q_{100}^*)$ that describes a distribution.

The bootstrapped statistics form a distribution whose mean and standard deviation can also be estimated:

$$\hat{\mu}^* = \frac{\sum_{b=1}^{N_B} q_b^*}{N_B} \tag{5.9}$$

$$\hat{\sigma}^* = \sqrt{\frac{\sum_{b=1}^{N_B} (q_b^* - \hat{\mu}^*)}{N_B - 1}} \tag{5.10}$$

The difference $\hat{\mu}^* - \hat{q}$ is the bias of the estimate and $\hat{\sigma}^*$ represents the random error of the estimate.

Confidence intervals (CIs) can also be obtained from the bootstrapped distributions. The method selected here for computing CIs is the *percentile* method (Carpenter and Bithell, 2000), which consists in a simple calculation of the percentiles of the observed distribution of the bootstrapped statistic. The advantages of the percentile method is that it is easy to implement and is transformation invariant. "Transformation invariant" implies that if the CI obtained for X is $[x_{CI_l}, x_{CI_u}]$ and if $g(\cdot)$ is a monotonic transformation, the CI of $g(X)$ equals $[-g(x_{CI_l}), g(x_{CI_u})]$. A good review of the available CI methods can be found in (DiCiccio and Efron, 1996; Carpenter and Bithell, 2000).

In this section, bootstrap is used for assessing the statistical uncertainty of the NAvg series ($\hat{q}$ in Equations 5.9 and 5.10) and how this uncertainty affects the statistics obtained between the NAvg and the satellite time series. This is done as follows:

1. Randomly select the station time series of the *in situ* network under study. This conforms the initial sample $(x_1, x_2, \ldots, x_K)$. Note that each x_i is a time series measured at a different station i and that K is the total number of available stations.
2. Apply bootstrap to obtain N_B *in situ* NAvg series: $\text{NAvg}^* = (\text{NAvg}_1^*(t), \ldots, \text{NAvg}_{N_B}^*)$. Note that the bootstrap is spatial, not temporal: each full series x_i is an input sample for the bootstrap.
3. Compare the bootstrapped *in situ* average series, NAvg*, with the satellite time series $y(t)$ and obtain N_B values of correlation (R*), bias (B*) and standard deviation of the differences (STDD*). These are obtained as follows (example for the correlation only):

$$R^* = (R_1^*, R_2^*, \ldots, R_{N_B}^*) \quad / \quad R_b^* = R\{NAvg_b^*(t), y(t)\}, \quad b = 1, 2, \ldots, N_B \tag{5.11}$$

5.2.3 Dependency between the SUM and the sample size

One of the critical aspects of bootstrap is the number of samples available. The lower the number of samples is, the more probable the bootstrap estimate is wrong. For example, the distribution of bootstrapped means calculated with only 2 samples is probably further from the true mean than the distribution of the bootstrapped means calculated with 100 samples. Is the number of stations in dense networks good enough to produce accurate bootstrap statistics? The aim of this section is to answer this question by evaluating the width of the confidence intervals of the validation statistics (R, B, STDD) as a function of the number of stations.

5.2.3.1 Datasets

The datasets are similar to those in the SUL procedure (section 5.1.1.2). The 20 stations of the Little Washita region are considered here and the period 2013/01 - 2013/06 is again selected. As satellite datasets only the SMOS L2 SM product is considered. SMOS series are filtered with the same quality thresholds as in section 5.1.1.2.

Initially, the experiment was conceived so that the bootstrapped series were also compared with AMSR2 time series. For this reason, only the samples with simultaneous availability of *in situ*, SMOS and AMSR2 observations are kept (52 in total). AMSR2-*in situ* comparisons are not shown here because the SMOS-*in situ* comparisons are sufficient for illustrating the effect of the number of stations in the uncertainty of the validation metrics.

5.2.3.2 Methodology

In overall, the simulation takes place in 3 stages:

For K'=2 to K:

1. Select K' time series out of the K available *in situ* series and obtain the respective bootstrapped NAvg*. K equals 20 in the Little Washita network.
2. Compare with the satellite time series and obtain the bootstrapped R*, B* and STDD*.
3. Increment K' by one and go back to step 1.

The number of selected samples K' starts at 2 because K'=1 simply represents the individual stations of the network and not any spatial average series.

The **first step** of the simulation is itself a small selection algorithm where multiple different subsets of size K' are used. Given K available stations, there are multiple combinations of stations that can produce subsets of size K' < K. Ideally, we should take all the possible subsets. However, this is computationally very expensive when K = 20 like in the Little Washita case. The number of combinations without replacement (subsets) is given by the expression $\binom{K}{K'}$. The maximum number of subsets is reached at K' = 10 and is $\binom{20}{10} = 184,756$. Supposing that the number of bootstrap NAvg* time series were N_B = 10,000 and that each sample had 50 values, the final number of values would be $\sim 92.378 \cdot 10^9$, which makes 369.512 giga-bytes. This would require a large hard-disk space and a extremely long execution time on a personal computer. Since the aim of the experiment here is simply to serve for illustrative purposes, **a reduced number of subsets $N_{subsets}$ is selected**: if the number of possible subsets is smaller than 1000, then $N_{subsets}$ is set to this number; otherwise, $N_{subsets}$ is set to 1000. **The algorithm for step 1 is integrated in the simulation and the complete approach is as follows**:

Complete method

For K'=2 to K:

1. Obtain the bootstrap NAvg* for $N_{subsets}$ of size K' as follows.

   For n_s=1 to $N_{subsets}$:

   1.1. Randomly select one subset of K' time series (different from any previous one)

   1.2. Obtain the N_B bootstrapped NAvg* samples for that subset

   1.3. Increment n_s by one and go back to step 1.1.

   A total of $N_{subsets}$ x N_B bootstrapped NAvg* samples is available at the end of this step.

2. Compare with the satellite time series and obtain the bootstrapped R*, B* and STDD*.
3. Increment K' by one and go back to step 1.

In step 1.2, the number of bootstrapped samples N_B has been set so it changes depending on K'. This is because the number of possible combinations *with* replacement increases with K'. Small subset sizes exhibit low number of possible combinations, so it is not efficient to use too large numbers of bootstrap simulations. For example, for K' = 2, only 3 combinations are possible so it is not efficient to use 10,000 bootstrap samples. The number of bootstrap samples is arbitrarily decided as follows: If K' is smaller than 4, then N_B is set to 100. If K' is between 4 and 6 samples, N_B is set to 1,000. For larger number of samples, N_B is set to 10,000.

5.2.3.3 Experiment and results

The Little Washita *in situ* time series are bootstrapped and compared to the respective SMOS series following the procedure explained before. The SMOS time series were described in section 5.2.3.1. Figure 5.10 shows the distributions obtained for the temporal correlation metric. The distributions for the temporal bias and the temporal STDD are included in Appendix F.

As expected, Figure 5.10 reveals that the distributions of the statistics become narrower with increasing sample sizes, i.e. the statistical uncertainty of the NAvg, **the SUM, decreases with increasing number of *in situ* stations**. When only 4 or less series are available, the shape of the distributions is noisy and present several sudden peaks. This is because the number of possible combinations with replacement is low (3, 10 and 35 for K' = 2, 3 and 4, respectively).

In Figure 5.10, the increase of the SUM becomes less evident for sample sizes larger than 14. In order to ease the analysis, Figure 5.11 directly presents the width between the vertical dashed lines in Figure 5.10 (the 90 % CI) as a function of the sample size K', for each of the validation metrics. The width of the CI decreases first very quickly with K', then slowly for larger values of K' (K' $\geq$ 14). In this region and period, the original sample size of K = 20 seems a good

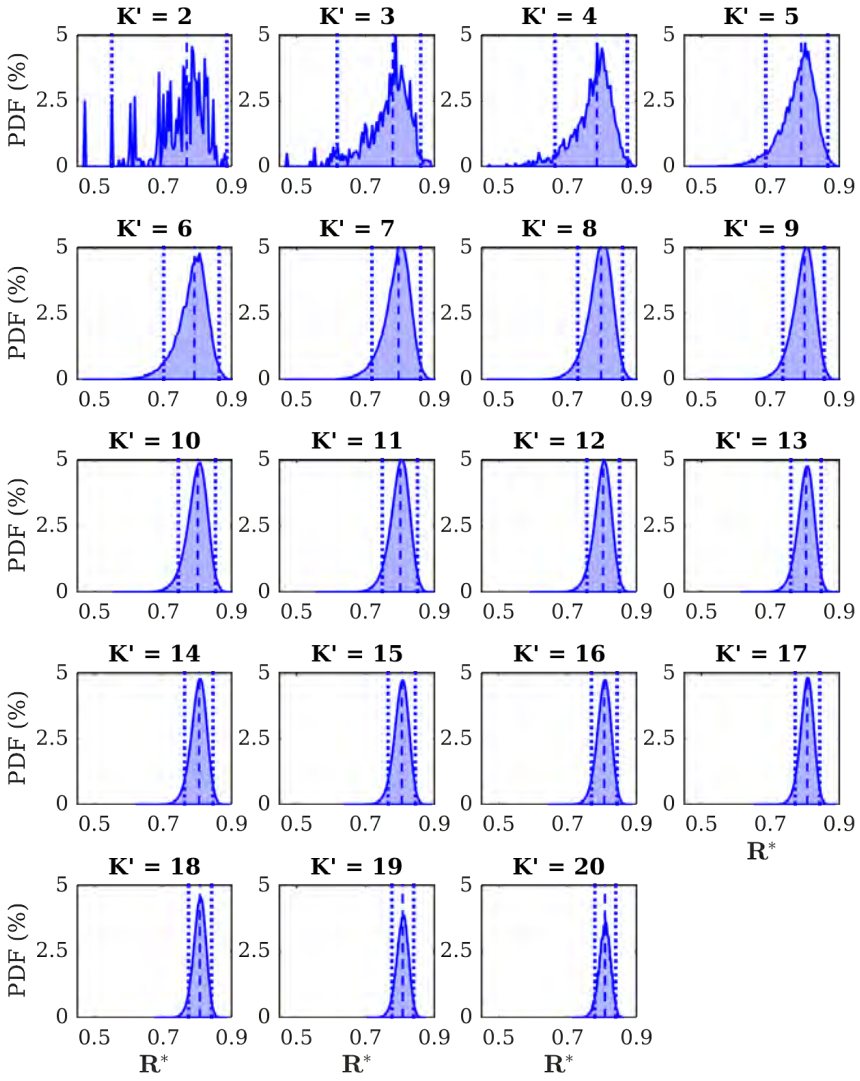
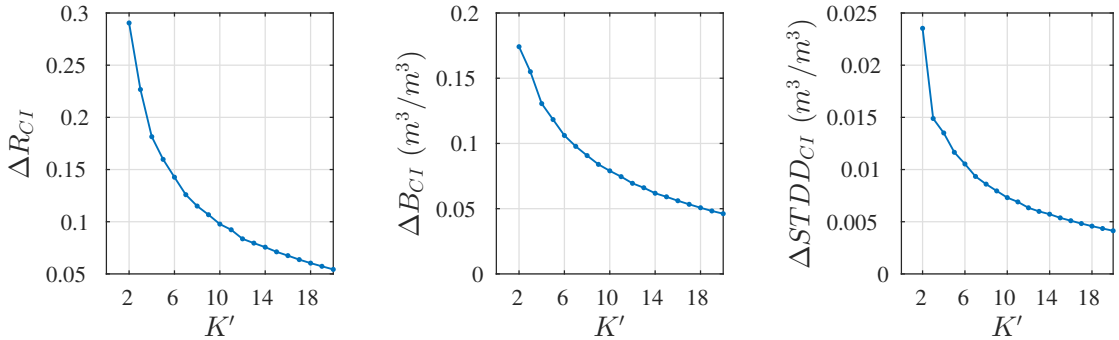


Figure 5.10 – Bootstrap distribution of the correlation between SMOS and the *in situ* NAvg, as a function of the number of *in situ* time series taken to compute the NAvg. The dashed line correspond to the median and the dotted lines to the 5 % and 95 % confidence limits

compromise between the statistical uncertainty of the NAvg and the number of stations that need to be maintained. When compared to SMOS, the uncertainty in correlation is 0.05, in bias is 0.05 m^3/m^3 and in STDD is 0.004 m^3/m^3.



(a) Correlation (b) Bias (c) STDD

Figure 5.11 – Width of the 90 % CI of the bootstrap distributions of the validation statistics, as a function of the number of *in situ* time series taken to compute the NAvg in Little Washita

It is important to highlight that this experiment implicitly assesses the SUL and the SUM at the same time. When bootstrapping is performed within an only selected subset of K' stations, then only the SUM is evaluated for that particular subset. When multiple different subsets of K' < K stations are tested we are also implicitly answering the question "what would have happened if the surface had been sampled at other different locations?" In other words, we are evaluating the contribution of the SUL to the SUM. It should be noted however, that this procedure underestimates the SUL because K=20 is far from being a dense sampling of the surface and because it analyses the influence of the SUL in the SUM, not the SUL itself. For a complete assessment of the SUL, the approach presented in section 5.1 is recommended.

5.2.4 Comparison of different SUM estimation approaches

The previous section has evaluated the uncertainty in the validation statistics that is due to the SUM, as a function of the number of *in situ* stations used to estimate the NAvg. I suggested that, when the number of selected stations was lower than the number of available stations (K' < K), the evaluations of the SUM were implicitly affected by the SUL. However, the SUL was underestimated because K=20 seemed an insufficient number of stations to represent it. A possible solution is to use the point time series simulated in section 5.1 to assess the SUL, since they are supposed to be a full sampling of the (statistical) SM field.

It is expected that the magnitude of the estimated SUM changes with both the number of selected point series K' and the number of available point series K. The previous section showed the dependence with K'. This section will address the dependence with the number of available point series K.

5.2.4.1 Datasets and methods

Two datasets are used: the 10,000 point SM series simulated in section 5.1 and the 20 *in situ* SM series of the Little Washita network. The period analysed is the same as in previous sections, 2013/01 - 2013/06 (6 months).

Only the dependence with the number of available point series K will be evaluated, so the number of selected locations K' is fixed. K' is set to 5 to match the number of reference stations used in the modelling of section 5.1. Regarding the number of available series K, three cases are considered:

- **Case A**: K' = 5, K = 20
- **Case B**: K' = 5, K = 10,000
- **Case C**: K' = K = 5

The approaches to evaluate the SUM in each case are different:

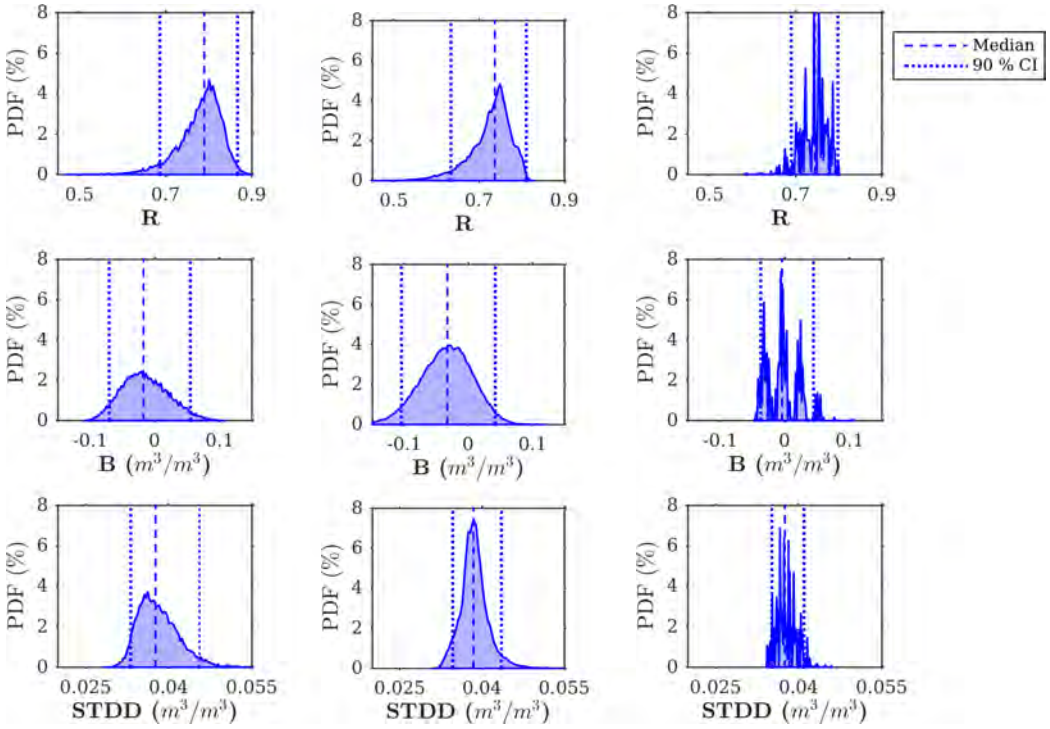
- **Case A**: 1,000 different subsets of 5 series out of 20 *in situ* series are bootstrapped
- **Case B**: 1,000 different subsets of 5 series out of 10,000 *simulated* series are bootstrapped
- **Case C**: 1 only subset of 5 *in situ* series is bootstrapped (stations #4, #6, #9, #10, #13)

In case A, the SM field is under-sampled (20 locations), while B represents the case where the field is fully sampled. In both cases, the influence of the SUL is implicitly assessed through the random selection of different subsets of stations. Case C cannot account for the SUL since the full set of available stations is used.

A total of N_B=1000 bootstrap average series (NAvg*) are produced for each subset. Therefore, cases A and B provide 1e9 NAvg* series and case C, 1000 NAvg* series. As in the previous sections, these bootstrap average series are compared to the SMOS soil moisture series for the same region and the respective temporal statistics are obtained (R^*, B^* and $STDD^*$).

5.2.4.2 Results

Figure 5.12 shows the temporal statistics obtained for each of the cases described in section 5.2.4.1. Each column corresponds to a different case and cases are ordered from left to right. The distributions of case C contrast with those of the other cases: they are noisy and irregular. This is probably due to the low number of stations that makes the NAvg very unstable and sensitive to each of the individual series. One of the *in situ* series is very different from the others (#4), specially in terms of bias (this can be seen in the lowest plot of Figure 5.6). It is not surprising then that the most important spikes appear in the bias distribution (2nd row of Figure 5.12C). The distributions of the A and B cases are similar except for the STDD. It seems that, by chance, the subsets obtained from the 10,000 simulated series are more similar in terms of STDD than those obtained from the 20 *in situ* series.



(A) K'=5 from K=20 *in situ* series (B) K'=5 from K=10,000 simulated series (C) K' = K = 5 *in situ* series

Figure 5.12 – Temporal statistics and their respective uncertainty distributions obtained when validating SMOS soil moisture series with only 5 ground series (K'=5). The uncertainty distributions are obtained with 3 different approaches: A) Bootstrap 1,000 subsets of K'=5 series out of 20 *in situ* series; ; B) Bootstrap 1,000 subsets of K'=5 series out of 10,000 simulated series; C) Bootstrap one group of K'=K=5 *in situ* series.

It seems that the CIs of the distributions of case C are slightly narrower than those of the other cases. This supports the argument that case C only accounts for the SUM and not the SUL, while cases A and B are also affected by the SUL. The exact CI widths of the distributions are included in Table 5.1, together with their medians and a short description of the different cases. Table 5.1 confirms that the C experiment slightly underestimates the width of the CIs with respect to the other two approaches, but only in the case of the correlation and the bias. No important biases are found between the median of the C statistics and those of the A and B cases. Therefore, the subset of stations of case C (#4, #6, #9, #10 and #13) is sufficient to provide a non-biased estimation of the validation statistics.

Table 5.1 – Characteristics of the distributions of the bootstrapped statistics obtained with different approaches. The approaches (A, B, C) assess the sampling uncertainty in groups of 5 stations.

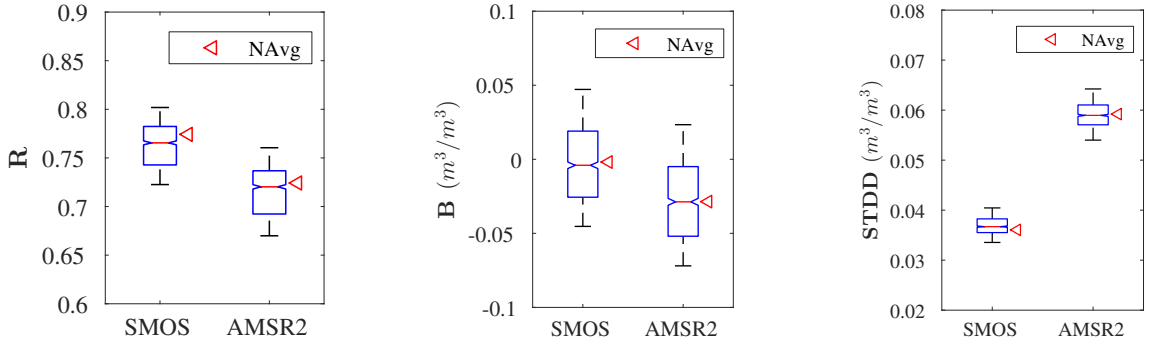
Case					**Correlation**		**Bias** (m^3/m^3)		**STDD** (m^3/m^3)	
Code	SU	Series	K'	K	ΔCI	Median	ΔCI	Median	ΔCI	Median
A	SUL + SUM	*in situ*	5	20	0.16	0.79	0.12	-0.02	0.012	0.038
B	SUL + SUM	simulated	5	10,000	0.17	0.74	0.14	-0.03	0.009	0.038
C	SUM	*in situ*	5	5	0.1	0.75	0.1	0	0.007	0.037

The medians and CI widths of cases A and B are mostly in agreement, except for the STDD. This suggests that the simulation approach of section 5.1 provides an ensemble of series that is consistent with the mean SM process represented by the 20 *in situ* stations. Given the similarity of the statistics of cases A and B, there is no gain on simulating the 10,000 point series to estimate the SUM. The SUM can be effectively estimated with the 20 available stations in the Little Washita network. Finally, the approach of section 5.1 can be reserved to analyse the SUL when no NAvg is used.

5.2.5 Inter-comparison of validation statistics affected by the SUM

Following the same rationale of section 5.1.6 box plots can ease the interpretation of the uncertainty information when comparing the validation of two different satellite products. In this section, I show how this can be done when validating SMOS and AMSR2 soil moisture products with the Little Washita NAvg. Only the SUM is assessed here and two cases are tested: when only the stations #4, #6, #9, #10 and #13 are available, and when the 20 Little Washita stations are available.

Figure 5.13 shows the intercomparison of SMOS and AMSR2 validations in the presence of SUM when only the 5 designated stations are available. In the box plots, the horizontal red line is the median, the box limits the 25 % and 75 % quantiles and the whiskers the 5 and 95 % percentiles. The triangle marker denotes the statistic obtained in the comparison of the measured NAvg (each station has a weight of 1/5 in the NAvg) and the satellite series. The difference in STDD between the 2 satellites is significant, however it is clearly not the case of the bias: the boxes almost totally overlap. SMOS seems better correlated to the *in situ* average than AMSR2, although it is hard to evaluate whether the difference is significant: the lower confidence limit of the SMOS correlation coincides with the median of the AMSR2 correlation.

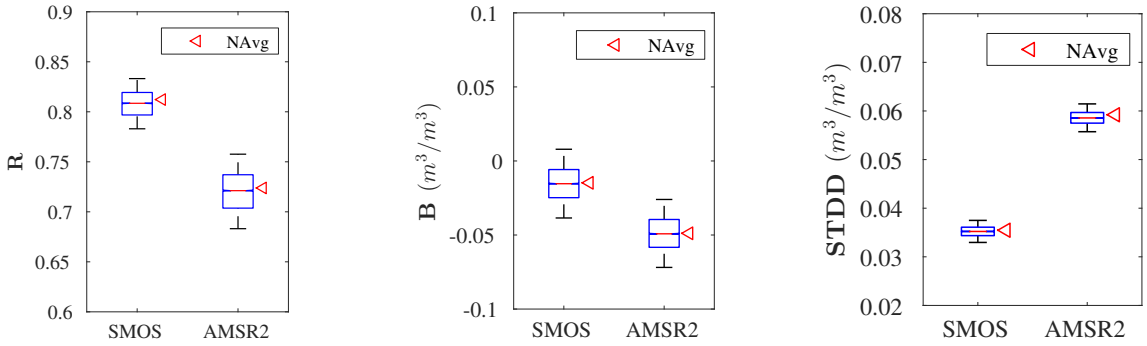


(a) Correlation (b) Bias (c) STDD

Figure 5.13 – Validation of SMOS and AMSR2 datasets with the NAvg of 5 *in situ* series with simultaneous assessment of the sampling uncertainty of the NAvg (box plots). The triangle represents the comparison with the measured NAvg

It is expected that, with more available stations, the SUM will be smaller and consequently the comparison between the performances of the two satellites might be affected. Figure 5.14 shows the intercomparison of SMOS and AMSR2 validations when the 20 Little Washita stations

are considered. Now, the correlation of SMOS is clearly better than that of AMSR2 even considering the SUM. The same is applicable for the STDD. A part of the bias distributions of the two satellites overlap but the difference in bias seems significant: the confusion surface only concerns the lower and the upper 25 % quartile of the SMOS and AMSR2 distributions, respectively.



(a) Correlation (b) Bias (c) STDD

Figure 5.14 – Validation of SMOS and AMSR2 datasets with the NAvg of 20 *in situ* series with simultaneous assessment of the statistical sampling uncertainty (box plots). The triangle represents the comparison with the measured NAvg

5.3 Conclusions

This chapter has addressed the sampling uncertainty of *in situ* measurements and its influence in satellite temporal validation statistics. The sampling uncertainty has been divided in two components, the sampling uncertainty of the locations (SUL) and the sampling uncertainty of the mean (SUM). Specific procedures have been developed to assess each of them separately.

The premise of this chapter was that the approaches proposed should be easily applicable to any validation area. In other words, they should not be region-specific and as generic as possible so they could be operationally used in satellite validation campaigns. The approaches presented fulfil this requirement. The approach presented for the SUL is based on empirical and statistical models that can be easily built for any validation area. The simulations are based on **Monte Carlo** samplings, which are well suited to any distribution type. Finally, the confidence intervals of the validation statistics are obtained empirically from the obtained distribution of samples. The approach presented for the SUM is also generic. The uncertainty of the mean was evaluated through **bootstrap** resampling and no assumptions are made about the underlying distribution of SM.

However the approaches here present some drawbacks. The simulations produced to assess the **SUL** are constrained by different factors: (i) temporal models need to perfectly mimic the *in situ* series, (ii) the assignation of the temporal models to the simulated samples is arbitrary and (iii) the accuracy of the spatial model is difficult to assess. The modelling and simulation procedures were tested with 5 *in situ* Little Washita series spanning 6 months. The results were satisfactory and the 10,000 point simulated series allowed to build confidence intervals in the statistics obtained when validating SMOS and AMSR2 soil moisture over the Little Washita region. Future research should verify that the procedure is applicable to longer time series and other validations regions. Long series increase the probability of finding long gaps as well as the difficulty of model fitting. Finally, it should be noted that, when applied to sparse validation networks (1 station per satellite footprint), the procedure most probably underestimates the sampling uncertainty. However, it is still useful as it could serve to indicate the minimum geostatistical uncertainty for the region under study.

Regarding the assessment of the **SUM**, the procedure proposed is easily applicable to any validation site since it simply consists in bootstrapping the available *in situ* series. However, it is worth highlighting that this is a purely statistical procedure: its reliability is based on the representativeness of the available sample series. For example, if the region of study corresponds to a satellite footprint ($\sim 40^2$ km^2) and we have 20 sample locations, but they all are located together in a small area of 100 x 100 m^2, they will probably be very similar so the SUM will be evaluated artificially small. The sample locations should be located carefully so that the statistical uncertainty can be interpreted as representative of the sampling uncertainty of the region of study.

A useful aspect of this chapter is that the contribution of the SUL to the NAvg can be assessed if the SUM bootstrap approach is applied to the simulated series of the SUL procedure. This was tested with the Little Washita case and the results were satisfactory. Further studies are needed to analyse whether this is applicable to other validation networks.

Finally, the assessment of the sampling uncertainty covers one of the components of the spatial scale mismatch. In previous chapters, the spatial scale mismatch was addressed as a whole. It remains to know how the sampling uncertainty evolves along time scales so that the specific contributions of the other two components of the spatial scale mismatch (the scaling uncertainty and the geophysical mismatch) can be addressed. This is totally feasible and is kept as one of the future lines of investigation.

Chapter 6

Conclusions and perspectives

Contents

6.1 Summary . . . 132
6.1.1 Improving the knowledge on the spatial scale mismatch . . . 132
6.1.2 Quantifying the spatial scale mismatch . . . 134
6.1.3 Methods . . . 134
6.1.4 Towards an integrated assessment of the spatial scale mismatch in satellite validation . . . 136
6.2 Perspectives . . . 137
6.2.1 Perspectives for the proposed approaches . . . 137
6.2.2 Application to other geophysical variables . . . 138
6.3 Open questions . . . 138
6.4 Final note . . . 139

Soil moisture (SM) is an essential climate variable whose continuous and global monitoring is necessary for several applications, like meteorological forecasting, hydrological modelling, agricultural assessment or prevention of natural hazards. Both continuous and global monitoring are achieved through instruments on board satellites. In particular, the retrieval of SM from satellite microwave radiometers at low frequency bands has been highly advanced and surface SM products can be obtained from various satellites equipped with such sensors, such as AMSR-E/2, SMOS and SMAP.

Satellite SM products can contain errors from various sources, the instrument, the models and/or the auxiliary parameters. For this reason, their accuracy is evaluated through comparison with ground SM measurements, a process called validation. One of the difficulties is that the size of the surfaces integrated by each of the measurement systems, or **supports**, is different: the satellite support (footprint) ranges between 27^2 and 55^2 km^2, depending on the satellite, and the *in situ* support concerns only some few centimetres. This **spatial scale mismatch** between satellite (large-support) and *in situ* (point-support) SM estimates hampers the validation process and the derived statistics. Different methods have been proposed to overcome the spatial scale mismatch, but they are usually limited by various physical and statistical conditions. Moreover, the magnitude of the contribution of the spatial scale mismatch to the validation statistics remains unknown.

What is a good definition for the SM at the satellite resolution? What is exactly behind the spatial scale mismatch? Is the spatial scale mismatch measurable? Is it possible to quantify its impact in the satellite validation metrics? The answers to these questions have been investigated through the following objectives, which were already stated at the end of the introduction (chapter 1) and that will articulate this chapter of conclusions:

- Improve the current knowledge on the spatial scale mismatch.
- Quantify the spatial scale mismatch if possible.
- Provide novel or improved methods for its analysis. The methods should be as generic as possible and easily applicable to all validation regions.
- Prepare the path for an integrated assessment of the spatial scale mismatch in satellite validation campaigns.

6.1 Summary

6.1.1 Improving the knowledge on the spatial scale mismatch

One of the intrinsic uncertainties of the spatial scale mismatch has a epistemic nature: we do not know what is the SM at the large-support scale. Two definitions are possible:

- **Definition I**: The large-support SM is the average of all the infinitesimal point SM values contained within the support
- **Definition II**: The large-support SM is the SM level associated with the integral of the energy emitted within the support

In situ measurements provide an estimate of the first definition through the network average (NAvg), which is the spatial mean of the SM measured by different stations within the satellite footprint. The satellite provides an estimate of the second definition. The true (unknown) SM values of definition I and of definition II are fundamentally different due to a number of factors (models, non-linearities, etc.).

In chapter 1, I defined the spatial scale mismatch as made of 3 components: the scaling uncertainty, the geophysical mismatch and the sampling uncertainty. The *scaling uncertainty* refers to the difference in the aforementioned definitions. A *geophysical mismatch* between *in situ* and satellite measurements is present because *in situ* stations record geophysical dynamics at a particular point while satellite sensors provide an integrated view of the geophysical dynamics. The *sampling uncertainty* refers to the unknown capacity of the *in situ* stations to represent the footprint SM of definition I: this unknown capacity depends on the location and the number of spatial samples. These definitions are important because **this is the first time that the definition of the spatial scale mismatch is stated with precision** in the case of SM. Moreover, they provide the basis of an heuristic framework for the assessment of the spatial scale mismatch.

Chapters 2 and 3 helped to understand the geophysical elements behind the spatial scale mismatch. In chapter 2, the analysis of previous literature indicated that differences between the point and the footprint SM exist because of within-footprint variability of geophysical factors such as precipitation, soil characteristics, vegetation and topography. These are drivers of the SM spatial variability and their importance varies with the region landscape and climatological conditions. Data about these drivers can help reduce the support of satellite products (a process called *disaggregation*) and thus, reducing the spatial scale mismatch. I evaluated the DISPATCH disaggregation approach in 3 regions with contrasting climate conditions. The experiment confirmed that **the principal drivers of SM spatial variability are different depending on the region** landscape and climatological conditions. The disaggregation supposed an improvement with respect to the original satellite SM products (SMOS) in semi-arid and arid regions, like in the Yanco area (Australia) and the Walnut Gulch catchment (U.S.), but not in humid regions (Little Washita catchment, U.S.). This was expected because DISPATCH is based on the assumption of a sub-footprint SM distribution controlled by the soil temperature and the vegetation status (NDVI) through evapotranspiration. The experiment also confirmed

that **disaggregation approaches can be used to overcome the spatial scale mismatch, but they are not applicable globally**.

Chapter 3 approached the geophysical elements behind the spatial scale mismatch from a temporal perspective. It was shown that, when two different SM series are compared (two series measured and different locations or provided by sensors with different supports), they exhibit differences at specific time scales depending on the geophysical element causing the difference. In other works, **SM time scales and SM-related geophysical variables are connected**:

- precipitation is usually a small time scale factor (sub-weekly scales)
- surface characteristics like texture are mostly associated to middle scales (sub-weekly to sub-seasonal scales)
- seasonal factors like vegetation are restricted to the SM seasonal scales

Additionally, the influence of precipitation and surface characteristics was occasionally propagated to seasonal scales.

Finally, the study of the spatial scale mismatch was completed in chapters 4 and 5 through the analysis of point and large-support SM series from a spatio-temporal perspective. In chapter 4, I linked SM time scales with the **spatial representativeness** of point series. The *spatial representativeness* of a location that works as an *effective support* and is defined as the surrounding extension where the observed SM is "sufficiently similar" to the SM at that location. The degree of similarity was expressed here with correlation-based metrics that compared the time series measured at the study locations with other SM datasets. From the analysis of the spatial representativeness at a range of time scales, from 0.5 to 128 days, I concluded that:

- The spatial scale mismatch tends to decrease with the time scale.
- In contrast, the dispersion of mismatch scores increases with the time scale: at seasonal scales some locations are highly representative while others are extremely poorly representative of the footprint SM.
- A sub-weekly scales, the spatial scale mismatch is very important and uniform for all locations.

The first of the conclusions of the list was derived from the fact that the spatial representativeness scores tend to increase with the time scale. From this it can be inferred that **SM spatial and temporal scales are connected**: small spatial scales are connected with small time scales and large spatial scales with large time scales. This is a relevant outcome of this thesis, since it is the first time that such a study has concerned footprint extents.

While chapter 4 was dedicated to the spatial scale mismatch as a whole, chapter 5 was specifically dedicated to the study of the **sampling uncertainty**. This was analysed in two cases: a) when the satellite is compared (validated) with only one *in situ* stations, and b) when the satellite is compared with the average of a set of stations (the network average, NAvg). The sampling uncertainty of the first case was referred as the sampling uncertainty of the locations (SUL), and that of the second ad the sampling uncertainty of the mean (SUM). The SMOS and the AMSR2 soil moisture products were validated in the Little Washita region and the influence of both the SUL and the SUM was evaluated. The experiment showed that **the uncertainty in validation statistics due to the SUL is largely more important than the uncertainty due to the SUM**. This is true when the NAvg is calculated using the typical number of stations of dense *in situ* networks (10-20). Another relevant conclusions is that **the contributions of the SUL and the SUM were non-negligible**. When the two different satellite SM products

were compared, their differences in performance were more or less significant depending on the degree of sampling uncertainty. For this reason, the SUL and the SUM should be systematically estimated when comparing the validation statistics of two different satellite SM products.

6.1.2 Quantifying the spatial scale mismatch

If the first question of this thesis was "what is the spatial scale mismatch?", the second question stems naturally from it: to what extent the spatial scale mismatch affects the satellite validation statistics? The quantification of the spatial scale mismatch was demonstrated feasible when only the **sampling uncertainty** was considered. In chapter 5, the uncertainty of the correlation, the bias and the standard deviation of the differences (STDD) between satellite and *in situ* measurements was quantitatively assessed. For example, the uncertainty in correlation due to the SUL when SMOS was validated in the Little Washita region for a particular period was estimated of the order of 0.15. Most importantly, confidence intervals were derived for all metrics, which allowed to evaluate the significance of the differences observed between SMOS and AMSR2 products for the same region.

It should be noted that the estimation of the SUL was based on the statistical spatial distribution of SM in the study area, which is not always known. This is usually the case of sparse *in situ* networks, where only one station is available per satellite footprint and where additional temporary sampling campaigns are rare. According to the literature, the statistical spatial distribution of SM is similar for many regions: bell-shaped when the mean SM is in its middle range and skewed towards its minimum and maximum values. This information can be used to build a minimal statistical model and to provide a **minimum threshold for the SUL in sparse networks**.

Regarding the whole spatial scale mismatch, its quantitative assessment was revealed more difficult. In chapter 4 I hypothesized that since small spatial scales seemed connected to small time scales, then sub-footprint heterogeneity should be connected to small time scales. Consequently, the differences in temporal dynamics between *in situ* and satellite SM series should only be present in the smallest time scales. The analyses indicated that this was only partly true: it is true that sub-weekly time scales were greatly and equally affected by the spatial scale mismatch but other time scales are not independent from it. Indeed, **differences in typical small time scale factors, such as precipitation and soil texture, seem to be propagated up to seasonal scales**, specially if the seasonal component of the SM signal is weak. Moreover, the spatial scale mismatch is more or less important depending on the *in situ* location. Nevertheless, two useful ideas can be derived from chapter 4:

1. Small time scales could serve to build a **lower bound for amount of spatial scale mismatch**
2. The dependence between spatial scale mismatch and time scales is strong, so the spatial scale mismatch could never be totally solved without taking time scales into account

6.1.3 Methods

The main premise for the methods developed during this thesis was that they should be as generic as possible so they could be easily applied to all validation regions. Until now the approaches used to study or evaluate the spatial scale mismatch were constrained by geophysical conditions (e.g. disaggregation) or by statistical prerequisites (e.g. triple collocation) that were difficult to fulfil everywhere. As a consequence, the approaches proposed should be somehow novel or different from the existing ones.

Wavelet transforms were one of the core tools used in this thesis. They have allowed to study SM time series on a time scale basis. *In situ* and model time series were analysed in scales ranging from 3 h to 128 days and satellite time series from 12 h to 128 days. It was shown that wavelet transforms are well suited to analyse SM signals since they succeed to reveal their characteristic temporal patterns (peaks, decays, seasonality, trends). For example, they allowed to link SM time scales with SM-related factors (precipitation, soil and vegetation) in chapter 3.

One of the main constrains of wavelet transforms is that they have to be applied to regularly sampled series. In practice, soil moisture time series contain gaps and they have been filled by using an approach based on the discrete cosinus transform (DCT). This approach has been revealed appropriate for the case of satellite soil moisture series where gaps are not localised in long periods of unavailability but evenly distributed along the series ; these represent between 50 % and a 70 % of the total number of samples in the series, depending on the satellite and the experiment. The most significant conclusions concerning the application of the DCT to soil moisture series are: 1) DCT gapfilling is not the root cause of the higher variance observed in satellite datasets with respect to *in situ* datasets at short time scales, 2) the contribution of the DCT gapfilling to the decorrelation between satellite and *in situ* series is small, and 3) DCT gapfilling can foster correlation between two time series where gaps are arranged in a similar way.

The study of the spatial scales required using methods to "measure" them. A way to measure the spatial scales is to evaluate the spatial representativeness of different locations within the footprint. Moreover, as mentioned before, I was interested in linking SM spatial and temporal scales, so the assessment of the spatial representativeness required methods adapted to operate on a per-time scale basis. I proposed the **correlated area (CArea)** method, an adaptation of the inverse footprint method of Orlowsky and Seneviratne (2014), specially suited for SM gridded data. In addition, wavelet transforms led me to the development of the **wavelet-based correlation (WCor)** approach, which consists in comparing the wavelet decompositions (detail series) of the *in situ* SM series against the NAvg. In the case of sparse networks, the same can be done using another large-support dataset different from the NAvg, such as a satellite. In this case, the choice of the large-support dataset has to be carefully addressed. By using modelled gridded SM data, I showed that both methods, CArea and WCor, performed equally well and provided consistent representativeness scores.

In the case of actual *in situ* SM data, the CArea method cannot be applied so WCor was compared against other approaches that were typically used in SM studies to assess the spatial representativeness: temporal stability analysis (TStab) and triple collocation (TC). TStab could not be applied to decomposed series because of their multiple zero-crossings. Regarding TC, it is suitable for week and month scales where it provides consistent results with WCor. However, at sub-weeekly scales where datasets cross-correlations were found low, and at seasonal scales, where the number of independent samples was also low, TC scores were erratic. As a conclusion, WCor was more robust and accurate than the other approaches. I recommend WCor when analysing real measured data (dense and sparse *in situ* networks) and when time scale decompositions are desired.

The evaluation of the sampling uncertainty in chapter 5 was possible thanks to a number of generic approaches: **statistical modelling**, **empirical modelling**, **Monte Carlo** simulations, **CDF matching** and **bootstrap**. In order to estimate the influence of the SUL in temporal validation statistics, a complete procedure based on these tools was designed. In overall, for a given study region, the **SUL procedure** consists in 1) building an empirical temporal model for the point SM series, 2) building a statistical spatial model for the area, 3) simulating point SM series constrained by both models through Monte Carlo simulations and CDF matching, 4) comparing the simulated series against the satellite SM series to obtain the uncertainty distribution of the statistics.

The SUL procedure gave reasonable results for the validation of SMOS over the Little Washita

region: the Little Washita *in situ* series fall within the distribution of simulated series, similarly to the SMOS-*in situ* statistics, which fall also within the uncertainty distributions obtained. These should not be interpreted as the actual statistics but as the possible statistics for the region of study from a geo-statistical perspective. Finally, the SUL procedure can be applied to other validation regions different from the one tested here. **This is the first procedure dedicated to the evaluation of the SM sampling uncertainty.**

Bootstrap was used to assess the uncertainty in validation statistics due to the sampling uncertainty of the NAvg: first, the *in situ* station series were combined 1,000 times with replacement to generate 1,000 possible NAvg series. Then, these series were compared to the satellite SM series under validation. From this, an empirical distribution of statistics was obtained. The method was easy to implement and did not require any prerequisites, so I recommend its systematic use when validating satellite products with dense *in situ* networks.

Bootstrap can be combined to the SUL procedure to estimate not only the SUM but also the influence of the SUL in the SUM. This was described in section 5.2.4 and I concluded that, in the case of "well-sampled" networks like Little Washita (20 stations), there is no gain of combining bootstrap and the SUL procedure with respect to the simple bootstrap.

6.1.4 Towards an integrated assessment of the spatial scale mismatch in satellite validation

By answering the initial questions (what is SM at the footprint resolution, what is the spatial scale mismatch and how to measure it), this thesis has provided some useful elements for the improvement of validation studies.

The most direct consequence is that **the approaches for assessing the SUL and the SUM** (chapter 5) **are readily applicable** to validation studies. Naturally, further research can bring some improvements, but there is no objection to apply them as they are right now.

Regarding the investigations about SM spatial and temporal scales, as previously mentioned, the extraction of the spatial scale mismatch from specific time scales has not been possible. However, the investigation have provided relevant aspects of the SM signal that can be taken into consideration when validating SM products. First, **sub-weekly time scales** should probably be omitted when validating satellite SM products with *in situ* measurements. These scales are fundamentally different between the two observation systems due to the spatial scale mismatch, so it does not make sense to evaluate the performance of the satellite product taking them into account.

An additional lesson from the investigations on spatial and temporal scales is that we should be careful with seasonal scales. Studies using **triple collocation (TC)** to validate satellite estimates or to assess the representativeness of *in situ* stations, systematically remove the seasonal SM component before conducting the analyses. As showed in chapters 3 and 4, the seasonal scale can be affected or not by the spatial scale mismatch and its influence depends on the location studied. By removing seasonal components not affected by the spatial scale mismatch, TC-based validations would omit a part of the SM signal that indeed, fosters similarity between the satellite and the *in situ* measurements. If affected by the spatial scale mismatch, TC-based representativeness assessments would **artificially benefit** those stations not spatially representative at seasonal scales.

Finally, wavelet decompositions could be used to improve **detrending** in validation studies. In the SM community, detrending is systematically accomplished with moving average filters of 30-day length. This approach presents multiple drawbacks. For example, the window length is arbitrarily fixed without any preliminary investigation. Continuous wavelet transforms could be used to choose the best window length. Additionally, the frequency response of moving average filters is much wider than most of the wavelet filter types, so the output series are still

considerably affected by slower variations of the signal (>30 days). Wavelet decompositions could help select the best window length and providing cleaner detrending.

6.2 Perspectives

6.2.1 Perspectives for the proposed approaches

In this thesis, the analysis of the SM spatial and temporal scales could not result in the development of a method to evaluate the full spatial scale mismatch. However, it showed that the spatial scale mismatch has an indissoluble link with the SM time scales and that the first cannot be understood without the second. In this context, there are some lines of research concerning SM time scales that deserve to be explored.

In sparse networks, only one *in situ* location is compared to the satellite SM. In this case, the spatial scale mismatch is made of only two components: the sampling uncertainty and the geophysical mismatch. The first can be estimated with the approaches in chapter 5. The second cannot be fully estimated, as mentioned in previous sections, but a minimum bound can be extracted from the sub-weekly scales. If this minimum mismatch could be estimated in terms of variance, then it could be incorporated to the spatial model in the SUL procedure of chapter 5. This would provide **the minimum uncertainty in validation statistics due to the spatial scale mismatch**.

When multiple *in situ* stations are available (dense networks), it is the network average (NAvg) that is directly compared to the satellite SM. Two uncertainty components are added to those of the sparse network case: the scaling uncertainty and the sampling uncertainty of the mean (SUM), which is part of the sampling uncertainty. **The evaluation of the scaling uncertainty is the most difficult problem**. Possible hints for its resolution could be found in sensibility studies and studies specifically addressing the gap between the linear average of the observations and the linear average of the retrievals (e.g. Crow et al., 2001; Pellarin et al., 2003; Crow et al., 2005; Zhan et al., 2008; Crosson et al., 2010). How to translate the information obtained in these studies to the operational assessment of the scaling uncertainty needs dedicated research.

Regarding the assessment of the spatial representativeness of particular locations, new methods can still be proposed. A recent communication of Rodríguez-Fernández et al. (2017b) suggested that **neural networks could be used to analyse the representativeness** of *in situ* locations. Neural networks are efficient non-linear regression tools and they have been already used with success to retrieve SM from observations (brightness temperatures). In Rodríguez-Fernández et al. (2017b), *in situ*-like SM time series were produced by a neural network using SMOS observations (brightness temperatures) and MODIS NDVI values as input vectors. They suggested that strong differences between the neural network SM series and the actual *in situ* series would indicate poor spatial representativeness of this one. This is a promising approach since it can be also used for all validation sites. However, how to use it to evaluate the uncertainty in the validation metrics needs to be investigated. Finally, it would be of great value to analyse the wavelet decomposition of the neural network SM series: this can provide more elements to the understanding of the connection between SM spatial and time scales.

Some of the approaches presented in this thesis deserve further research. Those presented in chapter 5 for assessing the sampling uncertainty are ready to be applied to other validation networks. This is desirable and necessary to validate them. One of the aspects observed is that the temporal modelling of the SUL procedure seemed to inflate the bias of output statistics. The comparison to additional *in situ* measurements could help to determine if this bias is consistent with the reality. In addition, a promising line of work is to use neural networks to model the SM dynamics in the SUL procedure.

The conclusions reached on SM spatial and time scales are highly valuable for modelling and upscaling/downscaling methods. Models designed on a per time scale basis might show improved skills. For example, at sub-weekly time scales, disaggregation could consider that the redistribution of SM only needs to be based on small spatial scale factors, such as rain or texture. At seasonal scales, disaggregation might only require seasonal varying signals, such as vegetation. This kind of schemes has already been used for disaggregating other geophysical variables like land surface temperature.

6.2.2 Application to other geophysical variables

This thesis has addressed the spatial scale mismatch in the case of the SM variable. However, the spatial scale mismatch is a common problem in satellite remote sensing that affects many other geophysical variables. The approaches developed here are generic enough to be applied to other variables. Naturally, the approaches should be adapted and interpreted according to the characteristics of the variable under study.

Snow-related variables, like snow cover or snow water equivalent (SWE), are good candidates for the approaches of this thesis because they present several similarities with the SM variable. First, snow can be highly variable within the satellite footprint. MODIS, which is usually preferred for its good temporal sampling (1 day), provides snow products at 250 m resolution and snow can vary considerably within that range. The spatial redistribution of snow cover varies in time: while in some periods of the year the snow cover is thin and mostly homogeneous, in other periods, the snow is arranged in patches with more important heights. Most snow variables are bounded by its minimum zero value so they also exhibit skewed distributions like SM when approaching zero (Grünewald et al., 2010). Finally, wavelet decomposition and simple simulation approaches can be applied because there also exist permanent *in situ* stations measuring snow variables continuously in time and snow is a consequence of precipitation events.

Rainfall is another geophysical variable suffering from similar observational scaling problems. For example, rainfall products from TRMM satellite are provided at a 0.25° resolution (∼25 km) and rain gauges measure at a particular point. Moreover, similar questions to those rose for the footprint SM are present in the case of the rainfall variable. What is the rainfall rate at 25 km resolution? The average of the rainfall in the area? Wavelet decompositions could be applied to rainfall since there exist both long satellite and *in situ* series.

Another variable that could be studied with the approaches developed in this thesis is the **land surface temperature (LST)**. There exist permanent *in situ* stations that measure the LST at resolutions up to 10 m. Although there are satellite missions providing LST products at similar resolutions (e.g. Landsat at 30 m), coarser resolution sensors like MODIS (1 km) are still used and need to be validated. The LST is fundamentally heterogeneous within the 1 km footprint, so its validation also suffers from the spatial scale mismatch problem.

6.3 Open questions

Research starts but never ends and this thesis is not an exception. This thesis leaves some open questions, either because answering them was not the objective of this thesis, either because the answers were more complex than expected.

The main practical question of this thesis was whether the spatial scale mismatch between satellite and *in situ* SM measurements could be assessed. I showed that this can be done, at least partially. The sampling uncertainty can be evaluated, as demonstrated in chapter 5. The evaluation of the whole spatial scale mismatch needs however further research. In section 6.2.1,

I have suggested some lines that can be explored. As a conclusion, the obtention of **a minimum bound for the contribution of the spatial scale mismatch in satellite validation statistics seems feasible**.

More general and almost philosophical questions do not have yet a clear explanation. For example, the notion of **spatial representativeness** remains open. All the methods used for assessing the spatial representativeness were based on correlation metrics (chapter 4). However, the representativeness could be expressed with other metrics like the variance. Can we affirm that a station time series is representative of the footprint time series because they correlate well although they are very different in variance? Usually, multiple metrics cannot be maximised at the same time, so some discussion in the SM community is needed about what are the metrics that should be prioritized.

In chapter 1, I presented the two possible **definitions of footprint-support SM**. Are we eager to accept that the footprint SM is a translation of the integrated energy and that it should not be obligatorily equal to the SM average? Or do we intend the satellite SM to perfectly match the (unknown) SM average of the region of study? In the first case, we should adapt our large spatial scale models to what the satellites see. Again, this is a important issue that should be discussed by the SM community.

Finally, apart from the spatial scale mismatch, there are **other sources of mismatch** between the satellite and the *in situ* measurements, the temporal and the vertical mismatch. The temporal mismatch is produced because unlike the satellite, the *in situ* measurements may not be taken exactly at the satellite overpass. The vertical mismatch appears when the sensing depths of the satellite and the *in situ* sensors do not coincide. It is known that the sensing depth of the satellite sensor varies in time due to changes in surface conditions (SM level, texture, vegetation). In contrast, the *in situ* station is installed at a constant depth, unless changes in the soil surface take place, which are not so uncommon. The vertical mismatch is mainly translated into decorrelation. These sources of mismatch were not within the objectives of this thesis, but it is important to consider them when validating satellite SM products since they also contribute to the uncertainty in the statistics.

6.4 Final note

This thesis has contributed to a better understanding of the spatial scale mismatch between satellite and *in situ* SM measurements. It has shown that SM spatial and temporal scales are inter-connected and that the spatial scale mismatch can be partially estimated. This has been possible with the adaptation and development of approaches based on generic signal processing and statistical tools. Such approaches can be used globally and adapted to other geophysical variables.

The uncertainty in satellite validation statistics due to the spatial scale mismatch has been usually overlooked, most probably, because it is difficult to estimate. This thesis has helped to raise the awareness of the magnitude of the spatial scale mismatch in satellite validation. Its influence can be dramatic at sub-weekly time scales and one of its components, the sampling uncertainty, can completely change the comparison of the relative performances of two different satellite SM products. It is essential that, when comparing satellite SM products, everyone ask him/herself if the difference is significant or can be simply obtained by chance due to the spatial scale mismatch.

Finally, there is a common general belief that improving the resolution of satellite SM products will make negligible the spatial scale mismatch. Different solutions have been proposed (SMAP) or are under development (SMOS-Next, SMOS-HR) to provide global SM products at higher resolutions. The target resolutions are limited to kilometric scales (1 to 10 km) because the solutions are based on L-band sensors, which are the ones that show a better sensibility to SM

under a wide range of vegetation conditions. However, reaching kilometric resolutions will not remove the spatial scale mismatch: SM can vary rapidly in just some few meters. The spatial scale mismatch between satellite and *in situ* SM measurements has still a long life to run.

Chapitre 6

Conclusions et perspectives (français)

Contents

6.1 Résumé 142
6.1.1 Améliorer les connaissances sur la différence d'échelle spatiale 142
6.1.2 Quantification de la différence d'échelle spatiale 144
6.1.3 Méthodes 145
6.1.4 Vers une évaluation intégrée de la différence d'échelle spatiale dans la validation satellitaire 147
6.2 Perspectives 148
6.2.1 Perspectives pour les approches proposées 148
6.2.2 Application à d'autres variables géophysiques 149
6.3 Questions ouvertes 150
6.4 Note finale 150

L'humidité du sol est une variable climatique essentielle dont la surveillance continue et globale est nécessaire pour plusieurs applications, comme la prévision météorologique, la modélisation hydrologique, la gestion de ressources en agriculture ou la prévention des risques naturels. La surveillance continue et globale de l'humidité du sol est possible grâce à des instruments embarqués à bord des satellites. En particulier, les radiomètres micro-ondes à basse freéquence fournissent des estimations de bonne qualité et des produits d'humidité du sol de surface sont obtenus grâce à divers satellites équipés de tels capteurs, comme AMSR-E / 2, SMOS et SMAP.

Les produits satellitaires d'humidité du sol sont entachés d'erreurs provenant de diverses sources (instrument, modèles et / ou paramètres auxiliaires). Pour cette raison, leur précision est évaluée au moyen de mesures terrain, une opération appelée validation. L'une des difficultés de la validation est que les dimensions des surfaces intégrées par chacun des systèmes de mesure, ou **supports**, sont différentes : le support satellitaire (empreinte) varie entre 27^2 et 55^2 km^2, selon le satellite, et le support de mesures *in situ* concerne seulement quelques centimètres. Cette **différence d'échelle spatiale** entre l'estimation du satellite (grand support) et la mesure *in situ* (support ponctuel) rend très délicat le processus de validation et les statistiques qui en découlent. Différentes méthodes ont été proposées pour surmonter la différence d'échelle spatiale, mais elles sont généralement limitées par diverses conditions physiques et statistiques. De plus, la contribution de la différence d'échelle spatiale aux statistiques de validation reste inconnue.

Quelle est la bonne définition pour l'humidité du sol à la résolution satellitaire ? Quelle est la vraie nature de la différence d'échelle spatiale ? La différence d'échelle spatiale est mesurable ? Est-il

possible de quantifier son impact sur les statistiques de validation du satellite ? Les réponses à ces questions ont été étudiées par le biais des objectifs qui suivent, énoncés dans l'introduction (chapitre 1) et qui articuleront également ce chapitre des conclusions :

- Améliorer les connaissances actuelles sur la différence d'échelle spatiale.
- Quantifier la différence d'échelle spatiale si possible.
- Fournir des méthodes novatrices ou améliorées pour son analyse. Les méthodes devront être aussi génériques que possible et facilement applicables à toutes les régions de validation.
- Mettre en place les éléments pour aller vers une évaluation intégrée de la différence d'échelle spatiale dans la validation de produits satellitaires d'humidité du sol.

6.1 Résumé

6.1.1 Améliorer les connaissances sur la différence d'échelle spatiale

L'une des incertitudes intrinsèques de la différence d'échelle spatiale possède un caractère épistémique : on n'a pas une définition univoque pour l'humidité du sol à grand support. Deux définitions sont possibles :

– **Définition I** : L'humidité du sol à grand support est la moyenne de toutes les valeurs d'humidité du sol ponctuelles contenues dans le support

– **Définition II** : L'humidité du sol à grand support est l'humidité associée à l'intégrale de l'énergie émise dans le support.

Les mesures *in situ* fournissent une estimation de l'humidité du sol selon la première définition comme le network average (NAvg) qui est la moyenne spatiale de l'humidité du sol mesurée par différentes stations au sein de l'empreinte satellitaire. Le satellite fournit une estimation de l'humidité du sol selon la deuxième définition. Les vraies valeurs d'humidité du sol (inconnues en fait) selon les définitions I et II sont fondamentalement différentes en raison d'un certain nombre de facteurs (modèles, non-linéarités, échantillonnage, etc.).

Au chapitre 1, j'ai décomposé la différence d'échelle spatiale en trois composantes : l'incertitude d'échelle, l'écart géophysique et l'incertitude d'échantillonnage. L'*incertitude d'échelle* se réfère à la différence entre les définitions mentionnées dans le paragraphe précédent. L'*écart géophysique* entre les mesures *in situ* et les estimations par satellite existe parce que les stations *in situ* enregistrent la dynamique géophysique à des points particuliers tandis que les capteurs satellitaires offrent une vue intégrée de la dynamique géophysique. Enfin, l'*incertitude d'échantillonnage* fait référence au manque de connaissance sur la capacité des stations *in situ* à représenter l'humidité du sol de la définition I ; ce degré de connaissance dépend de l'emplacement et du nombre de stations. Ces définitions sont importantes car **c'est la première fois que la définition de la différence d'échelle spatiale est décrite avec précision** dans le cas de l'humidité du sol. En outre, elles fournissent la base d'un cadre heuristique pour l'évaluation de la différence d'échelle spatiale.

Les chapitres 2 et 3 ont aidé à comprendre les éléments géophysiques conditionnant la différence d'échelle spatiale. Au chapitre 2, l'analyse de la littérature a indiqué que les différences entre l'humidité du sol ponctuelle et l'humidité du sol à grand support existent du fait de la variabilité intra-empreinte des facteurs géophysiques tels que la précipitation, les caractéristiques du sol, la végétation et la topographie. Ce sont les moteurs de la variabilité spatiale de l'humidité du sol et leur importance varie selon le paysage et la climatologie. Disposer d'informations sur ces "moteurs géophysiques" aide à réduire le support des produits satellites (dans unprocessus

appelé *désagrégation* ou *downscaling*) et ainsi réduire la différence d'échelle spatiale. J'ai évalué l'approche de désagrégation DISPATCH sur 3 régions possédant des conditions climatiques contrastées. L'expérience a confirmé que **les principaux facteurs de variabilité spatiale de l'humidité du sol sont différents selon l'environnement et les conditions climatologiques**. La désagrégation a montré une amélioration par rapport aux produits satellitaires d'humidité du sol originels (SMOS) dans des régions semi-arides et régions arides, comme Yanco (Australie) et le bassin versant de Walnut Gulch (États-Unis), mais pas dans les régions humides (bassin versant de Little Washita, US). Cela était attendu car DISPATCH est basé sur l'hypothèse d'une distribution spatiale de l'humidité du sol contrôlée par la température du sol et le statut végétal (NDVI) par évapotranspiration. L'expérience a également confirmé que **les approches de désagrégation peuvent être utilisées pour réduire la différence d'échelle spatiale, mais elles ne sont pas applicables à l'échelle mondiale**.

Le chapitre 3 a abordé les éléments géophysiques conditionnant la différence d'échelle spatiale à partir d'une perspective temporelle. On a montré que lorsque deux séries différentes d'humidité du sol sont comparées (deux séries correspondant à des emplacements différents ou correspondant à des capteurs ayant des supports différents), elles présentent des différences à des échelles de temps spécifiques liées aux caractéristiques des éléments géophysiques. **Les échelles de temps de l'humidité du sol et les variables géophysiques qui lui sont liées sont connectées** :

- la précipitation est habituellement un facteur à petite échelle de temps (échelles sous-hebdomadaires)
- les caractéristiques de surface comme la texture sont principalement associées aux échelles intermédiaires (échelles sous-hebdomadaires et sous-saisonnières)
- les facteurs saisonniers comme la végétation sont limités aux échelles saisonnières de l'humidité du sol

En outre, l'influence de la précipitation et les caractéristiques de surface sont occasionnellement propagées vers les échelles saisonnières. Enfin, l'étude de la différence d'échelle spatiale a été conduite dans les chapitres 4 et 5 à travers l'analyse des séries d'humidité du sol à petit et à grand support depuis une perspective spatio-temporelle. Au chapitre 4, j'ai connecté les échelles de temps de l'humidité du sol avec la **représentativité spatiale** des séries ponctuelles. La représentativité spatiale ou *support effectif* d'un point dans l'espace est défini comme étant l'étendue où l'humidité du sol observée est "suffisamment similaire" à l'humidité du sol à cet endroit. Le dégrée de similitude a été exprimée ici avec des indicateurs basés sur la corrélation qui ont comparé des séries temporelles à petit support avec d'autres séries temporelles à petit et grand support. À partir de l'analyse de la représentativité spatiale à des différentes échelles de temps, de 0,5 à 128 jours, j'ai conclu que :

- La différence d'échelle spatiale tend à diminuer avec l'échelle de temps.
- En revanche, la dispersion des scores obtenus augmente avec l'échelle de temps : à l'échelle saisonnière certains points sont très représentatifs du grand support tandis que d'autres sont extrêmement peu représentatifs.
- À l'échelle sous-hebdomadaire, la différence d'échelle spatiale est très importante et uniforme pour tous les emplacements.

La première des conclusions de la liste découle des scores obtenus de représentativité spatiale qui ont tendance à augmenter avec l'échelle de temps. À partir de là, on peut déduire que **les échelles spatiales et temporelles de l'humidité du sol sont reliées** : les petites échelles spatiales sont reliées aux petites échelles de temps et les grandes échelles spatiales aux grandes

échelles de temps. Ceci est un résultat pertinent de cette thèse, car c'est la première fois qu'une telle étude porte sur l'étendue de l'empreinte.

Alors que le chapitre 4 était consacré à la différence d'échelle spatiale dans son ensemble, le chapitre 5 a été spécifiquement dédié à l'étude de l'**incertitude d'échantillonnage**. Celle-ci a été analysée dans deux cas : a) lorsque le satellite est comparé (validé) avec une seule station de mesure *in situ*, et b) lorsque le satellite est comparé à la moyenne d'un ensemble de stations de mesures (network average, NAvg). L'incertitude d'échantillonnage du premier cas a été identifiée comme l'incertitude d'échantillonnage des positions (sampling uncertainty of locations, SUL), et celle du deuxième cas comme l'incertitude d'échantillonnage de la moyenne (SUM). Des produits d'humidité du sol du satellite SMOS et AMSR2 ont été validés dans la région de Little Washita et l'influence de la SUL et de la SUM a été évaluée. L'expérience a montré que **l'incertitude dans les statistiques de validation due à la SUL est plus importante que l'incertitude due à la SUM**. Ceci est vrai lorsque le NAvg est calculé en utilisant le nombre typique de stations dans les réseaux de mesures *in situ* de type dense (10-20 stations). Une autre conclusion majeure est que **les contributions de la SUL et de la SUM n'étaient pas négligeables**. Lorsque les deux produits d'humidité du sol satellitaires différents ont été comparés, leurs différences de performance étaient plus ou moins importantes en fonction du degré d'incertitude de l'échantillonnage. Pour cette raison, la SUL et la SUM devraient être systématiquement estimées quand on compare les statistiques temporelles de validation de deux produits d'humidité du sol issus de satellites différents.

6.1.2 Quantification de la différence d'échelle spatiale

Si la première question de cette thèse était "qu'est-ce que la différence d'échelle spatiale ?", la deuxième question en découle naturellement : dans quelle mesure la différence d'échelle spatiale affecte la validation des statistiques de validation des satellites ? La quantification de la différence d'échelle spatiale a été démontrée faisable uniquement lorsque l'**incertitude d'échantillonnage** a été prise en compte. Au chapitre 5, l'incertitude de la corrélation, du biais et de l'écart type des différences (STDD) entre les mesures satellitaires et *in situ* d'humidité du sol ont été évaluées quantitativement. Par exemple, l'incertitude dans la corrélation due à la SUL lorsque SMOS a été validé dans la région de Little Washita pendant une période donnée a été estimée de l'ordre de 0,15. Plus important encore, des intervalles de confiance ont été obtenus pour toutes les métriques, ce qui a permis d'évaluer la significativité statistique des différences observées entre SMOS et les produits AMSR2 pour la même région.

Il convient de noter que l'estimation de la SUL était basée sur la distribution spatiale statistique de l'humidité du sol dans la zone d'étude, ce qui n'est pas toujours connu. Ceci est généralement le cas des réseaux de mesure *in situ* de type épars, où une seule station est disponible par empreinte satellitaire et où les campagnes d'échantillonnage plus dense sont rares. Selon la littérature, la distribution spatiale statistique de l'humidité du sol est similaire pour de nombreuses régions : en forme de cloche lorsque l'humidité du sol moyenne est en milieu de gamme et asymétrique vers ses valeurs minimales et maximales. Cette information peut être utilisée pour construire un modèle statistique minimal et fournir un **seuil minimum pour la SUL des réseaux de type épars**.

En ce qui concerne la différence d'échelle spatiale totale, son évaluation quantitative s'est révélée plus difficile. Au chapitre 4, j'ai émis l'hypothèse que, si les petites échelles spatiales étaient liées à des petites échelles de temps, l'hétérogénéité à l'intérieur de l'empreinte satellitaire devrait être reliée à de petites échelles de temps. Par conséquent, les différences de la dynamique temporelle entre les séries d'humidité du sol *in situ* et satellite ne devraient être présentes que pour les plus petites échelles de temps. Les analyses indiquent que cela n'est que partiellement vrai : s'il est vrai que les échelles de temps sous-hebdomadaires apparaissent largement et uniformément affectées par la différence d'échelle spatiale, d'autres échelles de temps ne sont pas épargnées.

En effet, **des différences dans des éléments géophysiques à petite échelle temporelle, tels que les précipitations et la texture du sol, semblent se propager jusqu'aux échelles saisonnières**, surtout si la composante saisonnière du signal humidité est de faible amplitude. De plus, la différence d'échelle spatiale est plus ou moins importante en fonction de l'emplacement de la station de mesure *in situ*. Néanmoins, deux idées utiles ressortent du chapitre 4 :

1. Les petites échelles de temps pourraient servir à construire une limite inférieure pour la différence d'échelle spatiale
2. La différence d'échelle spatiale dépend fortement des échelles de temps ; vouloir la réduire ou la résoudre totalement ne pourrait pas se faire sans tenir compte du temps.

6.1.3 Méthodes

Le prérequis principal pour les méthodes développées au cours de cette thèse était qu'elles devraient être aussi génériques que possible afin d'être facilement applicable à toute région de validation. Jusqu'à présent, les approches utilisées pour étudier ou évaluer la différence d'échelle spatiale étaient contraintes par des conditions géophysiques (comme par exemple dans le cas de la désagrégation) ou par des prérequis statistiques (par exemple, la triple collocation) qui étaient difficiles à respecter de manière générale voire à vérifier. Par conséquent, pour pallier ces contraintes, les approches proposées devaient être nouvelles ou différentes.

Les **transformées en ondelettes** ont été l'un des outils de base utilisés dans cette thèse. Elles ont permis d'étudier les séries temporelles d'humidité du sol sur une base d'échelles de temps. Des séries temporelles de mesures *in situ* et modélisées ont été analysées en échelles de temps allant de 3 h à 128 jours et des séries temporelles satellitaires de 12 h à 128 jours. Il a été montré que les transformées en ondelettes sont bien adaptées pour analyser les signaux humidité car elles réussissent à révéler leur dynamiques temporelles caractéristiques (pics, décroissances par assèchement, saisonnalité). Ainsi, elles ont permis de relier les échelles de temps de l'humidité du sol avec les facteurs géophysiques liés à l'humidité du sol (précipitations, sol et végétation) au chapitre 3.

Les transformées en ondelettes ont pour contrainte principale de nécessiter un échantillonnage temporel régulier pour être applicables. Dans la pratique, des trous existent et ont été comblés avec une approche en transformée en cosinus discrète (TCD). Cette méthode s'est révélé adéquate pour le cas des séries satellitaires d'humidité du sol où les trous sont de courtes durées et sont répartis uniformément dans la série ; ils représentent entre 50 % et 70 % du nombre total d'échantillons, suivant les satellites et les conditions expérimentales. Les conclusions les plus significatives sur l'impact de la TCD, malgré un taux de remplissage élevé, sont : 1) elle n'est pas à l'origine de la plus grande variance des séries satellitaires dans les petites échelles de temps comparée aux séries de mesure *in situ*, 2) sa contribution à la décorrélation entre les séries satellitaires et *in situ* reste faible, 3) elle peut cependant induire de la corrélation entre deux séries dont les trous sont placés de manière similaire.

L'étude des échelles spatiales a nécessité l'utilisation et le développement de méthodes pour leur "mesurer". Un moyen de mesurer l'échelle spatiale consiste à évaluer la représentativité spatiale de différents sites au sein de l'empreinte satellitaire. Cependant, mon principal intérêt était de relier les échelles spatiales et temporelles de l'humidité du sol, et donc l'évaluation de la représentativité spatiale a exigé des méthodes applicables sur chaque échelle de temps. J'ai proposé la **méthode de l'aire corrélée (CArea)**, une adaptation de la méthode de l'empreinte inverse (inverse footprint) d'Orlowsky et Seneviratne (2014), spécialement adaptée à des données d'humidité du sol densément échantillonnées dans l'espace. En outre, les transformées en ondelettes m'ont amené au développement de la **corrélation basé sur des ondelettes** ou **wavelet-based**

correlation (WCor), qui consiste à comparer les décompositions en ondelettes (séries de détail) de la série d'humidité *in situ* avec la série NAvg. Dans le cas de réseaux épars, la même méthode peut être appliquée mais en utilisant une autre donnée à grand support se substituant à NAvg, tel qu'une humidité satellitaire. Dans ce cas, le choix de la donnée à grand support doit être soigneusement considéré. En utilisant des données maillées modélisées d'humidité du sol, j'ai montré que les performances des deux méthodes, CArea et WCor, étaient bonnes et similaires.

Dans le cas des données d'humidité du sol *in situ* réelles, la méthode CArea ne peut pas être appliquée, donc WCor a été comparée à d'autres approches habituellement utilisées dans les études d'humidité du sol pour évaluer la représentativité spatiale : l'analyse de la stabilité temporelle (TStab) et la triple collocation (TC). Lorsque l'on s'intéresse aux séries décomposées par ondelettes TStab devient inapplicable, les séries de détails étant par construction de moyenne nulle. En ce qui concerne TC, la méthode est appropriée pour les échelles hebdomadaires et mensuelles où elle fourni des résultats cohérents avec WCor. Cependant, à des échelles sous-hebdomadaires où les corrélations croisées des jeux de données ont été trouvées faibles, et aux échelles saisonnières où le nombre d'échantillons indépendants était petit, les scores de TC se sont avérés erratiques. En conclusion, WCor s'est montré plus robuste et plus précise que les autres approches. Je la recommande donc lorsqu'on analyse des données réelles mesurées (réseaux *in situ*de type dense et épars) et lorsque la décomposition en échelles de temps est souhaitée.

L'évaluation de l'incertitude d'échantillonnage au chapitre 5 a été possible grâce à un certain nombre d'approches génériques : **modélisation statistique**, **modélisation empirique**, **simulations de Monte Carlo**, **CDF *matching*** et **bootstrap**. Afin d'estimer l'influence de la SUL dans les statistiques temporelles de validation des produits satellitaires, une procédure complète basée sur ces outils a été conçue. D'une manière générale, pour une région d'étude donnée, la **procédure évaluant l'impact de la SUL** consiste en 1) la construction d'un modèle temporel empiriquepour les séries d'humidité du sol ponctuelles, 2) la construction d'un modèle spatial statistique pour la zone, 3) la simulation des séries d'humidité du sol contraintes par les deux modèles à l'aide de simulations Monte Carlo et de CDF matching, et 4) la comparaison des séries simulées avec la série satellite d'humidité du sol pour obtenir les distributions d'incertitude des statistiques.

La procédure d'estimation de la SUL et de son impact a donné des résultats raisonnables dans la validation des produits d'humidité du sol de SMOS sur la région de Little Washita : les séries *in situ* de Little Washita se trouvent à l'intérieur de la distribution de séries simulées, de manière similaire aux statistiques SMOS-*in situ*, qui se trouvent aussi à l'intérieur des distributionsd'incertitude obtenues. Les distributions ne doivent pas être interprétées comme des statistiques réelles mais comme les statistiques possibles pour la région d'étude d'un point de vue géostatistique. Enfin, procédure d'évaluation de l'impact de la SUL peut être appliquée à d'autres régions de validation sans difficulté majeure. **C'est la première procédure dédiée à l'évaluation de l'incertitude d'échantillonnage de l'humidité du sol**.

L'approche par bootstrap a été utilisée pour évaluer l'incertitude dans les statistiques de validation induite par la SUM : Dans un premier temps, les séries de stations de mesures *in situ* ont été combinées 1000 fois avec remplacement pour générer 1000 séries possibles de NAvg* (séries bootstrap). Ensuite, ces séries ont été comparées à la série satellitaire en cours de validation. Finalement, une distribution empirique des statistiques a été obtenue. La méthode est facile à mettre en œuvre et ne nécessite aucun prérequis, donc je recommande son utilisation systématique lors de la validation de produits satellites avec des réseaux de mesure *in situ* de type dense.

L'approche par **bootstrap** peut être combinée à la procédure d'estimation de la SUL afin d'estimer non seulement la SUM mais aussi l'influence de la SUL dans la SUM. Ceci a été décrit dans la section 5.2.4 et j'ai conclu que, dans le cas de réseaux «bien échantillonnés» comme Little Washita (20 stations), la combinaison bootstrap-SUL n'apporte pas de gain par rapport

au bootstrap simple.

6.1.4 Vers une évaluation intégrée de la différence d'échelle spatiale dans la validation satellitaire

En répondant aux questions initiales (qu'est-ce que l'humidité du sol à la résolution satellitaire, quel est la différence d'échelle spatiale et comment la mesurer), cette thèse a fourni des éléments utiles pour l'amélioration des études de validation.

La conséquence la plus directe est que **les approches pour évaluer la SUL et la SUM** (chapitre 5) **sont facilement applicables dans les études de validation**. Naturellement, d'autres recherches peuvent apporter des améliorations, mais il n'y a aucune objection à les appliquer comme elles le sont actuellement.

En ce qui concerne ma recherche sur les échelles spatiales et temporelles d'humidité du sol, il n'a pas été possible d'extraire la différence d'échelle spatiale à partir d'échelles de temps spécifiques. Cependant, des caractéristiques pertinentes du signal humidité ont été identifiées et peuvent être prises en compte lors de la validation des produits d'humidité du sol. Premièrement, **les échelles de temps sous-hebdomadaires** devraient probablement être omises lors de la validation de produits satellitaires avec des mesures *in situ*. Ces échelles sont fondamentalement différentes entre les deux systèmes d'observation en raison d'une forte différence d'échelle spatiale. Il n'est donc pas pertinent d'évaluer la performance d'un produit satellitaire à ces échelles qui reflèteront davantage les écarts de géophysiques et d'échantillonnage.

Une leçon supplémentaire de ma recherche sur les échelles spatiales et temporelles est que nous devrions faire attention aux échelles saisonnières. Les études utilisant la **triple collocation (TC)**, soit pour valider les estimations satellitaires ou pour évaluer la représentativité des stations de mesures *in situ*, éliminent systématiquement l'échelle saisonnière de l'humidité du sol avant de procéder aux analyses. Comme le montrent les chapitres 3 et 4, les échelles saisonnières peuvent être affectées ou non par la différence d'échelle spatiale et son influence dépend de l'emplacement étudié. En supprimant les composants saisonniers qui ne sont pas affectés par la différence d'échelle spatiale, les validations basées sur la TC omettraient une partie du signal humidité du sol qui, en effet, favorise la similitude entre le satellite et les mesures *in situ*. Si les composants saisonniers sont en revanche affectés par la différence d'échelle spatiale, les évaluations de représentativité basées sur la TC **bénéficieraient artificiellement** aux stations avec une mauvaise représentativité spatiale à l'échelle saisonnière.

Enfin, les décompositions en ondelettes pourraient être utilisées pour **éliminer la tendance** dans les séries temporelles utilisées dans les études de validation lorsque c'est adapté. Dans la communauté scientifique s'intéressant à la thématique de l'humidité du sol, l'élimination de la tendance est systématiquement réalisée avec des filtres de moyenne mobile de longueur voisine de 30 jours. Cette approche présente de multiples inconvénients. Par exemple, la longueur de la fenêtre est arbitrairement fixée sans aucune enquête préliminaire. Les transformations continues d'ondelettes pourraient être utilisées pour choisir la meilleure longueur de fenêtre. En outre, la réponse fréquentielle des filtres de moyenne mobile sont beaucoup plus larges que la plupart des types de filtres d'ondelettes, de sorte que les séries filtrées sont encore considérablement affectées par des variations plus lentes du signal (>30 jours). Les décompositions en ondelettes pourraient aider à sélectionner la meilleure longueur de fenêtre et utiliser la série de lissage de l'échelle sélectionnée pour une élimination plus propre de la tendance.

6.2 Perspectives

6.2.1 Perspectives pour les approches proposées

Dans cette thèse, l'analyse des échelles spatiales et temporelles de l'humidité du sol n'a pas pu se concrétiser dans le développement d'une méthode pour évaluer la différence d'échelle spatiale de manière globale. Cependant, elle a montré que la différence d'échelle spatiale a un lien indissoluble avec les échelles de temps de l'humidité du sol et que la première ne peut pas être comprise sans le deuxième. Dans ce contexte, il existerait des lignes de recherche concernant les échelles de l'humidité du sol qui méritent d'être explorées.

Dans les réseaux épars, un seul emplacement *in situ* est comparé à l'humidité du sol du satellite. Dans ce cas, la différence d'échelle spatiale n'est constituée que de deux composantes : l'incertitude d'échantillonnage et l'écart géophysique. La première peut être estimée avec les approches du chapitre 5. La deuxième ne peut pas être entièrement estimée, comme mentionné dans les sections précédentes, mais une limite inférieure peut être extraite des échelles sous-hebdomadaires. Si cette différence minimale était estimée en termes de variance, elle pourrait être incorporée au modèle spatial dans la procédure basée sur la SUL du chapitre 5. Cela fournirait **l'incertitude minimale dans les statistiques de validation induite par la différence d'échelle spatiale**.

Lorsque plusieurs stations *in situ* sont disponibles (réseaux denses), c'est la moyenne du réseau (NAvg) qui est directement comparée à l'humidité du sol du satellite. Deux composants d'incertitude sont ajoutés au cas des réseaux épars : l'incertitude d'échelle et l'incertitude d'échantillonnage de la moyenne(SUM), qui font partie de l'incertitude d'échantillonnage. **L'évaluation de l'incertitude d'échelle est le problème le plus difficile**. Des idées pour sa résolution peuvent être trouvées dans les études de sensibilité et les études traitant spécifiquement l'écart entre la moyenne des observationset la moyenne des inversions d'humidité du sol (e.g. Crow et al., 2001 ; Pellarin et al., 2003 ; Crow et al., 2005 ; Zhan et al., 2008 ; Crosson et al., 2010). Comment traduire les informations obtenues dans ces études à l'évaluation opérationnelle de l'incertitude d'échelle nécessite d'une recherche dédiée.

En ce qui concerne l'évaluation de la représentativité spatiale de points particuliers dans l'espace, de nouvelles méthodes peuvent encore être proposées. Une communication récente de Rodríguez-Fernández et al. (2017b) propose d'utiliser **les réseaux de neurones pour analyser la représentativité** des séries de mesures *in situ*. Les réseaux de neurones sont des outils de régression non linéaire efficaces et ils ont déjà été utilisés avec succès pour inverser l'humidité du sol à partir des observations de températures de brillance. Rodríguez-Fernández et al. (2017b) a produit des séries temporelles d'humidité du sol avec un réseau de neurones en utilisant la température de brillance d'humidité du sol de SMOS et le NDVI du satellite MODIS comme vecteurs d'entrée entraînés sur des mesures *in situ*. Ils ont suggéré que les séries *in situ* réelles auraient une représentativité spatiale médiocre quand elles présenteraient de fortes différences avec les séries générées par le réseau de neurones. C'est une approche prometteuse car elle peut également être utilisée pour tous les sites de validation. Cependant, leur utilisation pour évaluer l'incertitude dans les statistiques temporelles de validation doit être étudiée. Enfin, il serait très utile d'analyser la décomposition en ondelettes des séries d'humidité du sol issues des inversions par réseau de neurones : cela peut fournir plus d'éléments à la compréhension de la connexion entre les échelles spatiales et temporelles de l'humidité du sol.

Certaines des approches présentées dans cette thèse méritent d'être approfondies. Celles présentés dans le chapitre 5 pour évaluer l'incertitude d'échantillonnage sont prêtes à être appliquées à d'autres réseaux de validation. Ceci est souhaitable et nécessaire pour les valider. L'un des aspects observés est que la modélisation temporelle de la procédure SUL semblait gonfler le biais des statistiques de sortie. La comparaison avec des mesures *in situ* supplémentaires pourrait aider à déterminer si ce biais est cohérent avec la réalité. En outre, une ligne de travail

prometteuse consisterait à utiliser des réseaux de neurones pour générer des séries d'humidité du sol dans la procédure SUL.

Les conclusions sur les échelles spatiales et temporelles de l'humidité du sol sont très utiles pour la modélisation pour les méthodes d'upscaling et de downscaling. Des modèles conçus en stratifiant la modélisation par échelles de temps pourraient permettre d'obtenir de meilleures performances. Par exemple, à des échelles de temps sous-hebdomadaires, la désagrégation pourrait considérer que la redistribution d'humidité du sol doit uniquement être basée sur de facteurs à petite échelle spatiale, tels que la pluie ou la texture. Aux échelles saisonnières, la désagrégation ne nécessiterait que des signaux saisonniers, comme la végétation. Ce type de schémas a déjà été utilisé pour désagréger d'autres variables géophysiques comme la température de surface de la terre.

6.2.2 Application à d'autres variables géophysiques

Cette thèse a abordé la différence d'échelle spatiale dans le cas de la variable humidité du sol. Cependant, la différence d'échelle spatiale est un problème commun dans la télédétection par satellite qui affecte aussi d'autres variables géophysiques. Les approches développées ici sont suffisamment génériques pour être appliquées à d'autres variables. Naturellement, les approches devraient être adaptées et interprétées selon les caractéristiques de la variable étudiée et les propriétés des systèmes d'observation associés.

Les variables liées à la **neige**, comme la couverture de neige ou le contenu en eau équivalente de la neige (SWE), sont de bonnes candidates pour les approches présentées dans cette thèse parce qu'elles présentent plusieurs similitudes avec la variable humidité du sol. Tout d'abord, la neige peut être très variable dans l'empreinte satellitaire. Le satellite MODIS, qui est généralement préféré pour son bon échantillonnage temporel (1 jour), fournit des produits de neige à une résolution de 250 m et la neige peut varier considérablement dans cette empreinte. La redistribution spatiale de la couverture neigeuse varie aussi dans le temps : à certaines périodes de l'année, la couverture de neige est mince et surtout homogène, et à d'autres périodes, la neige est disposée en tache avec des hauteurs plus importantes. La plupart des variables associées à la neige sont bornées par la disparition totale de la neige, elles présentent des distributions asymétriques comme l'humidité du sol lorsqu'elles s'approchent de zéro (Grünewald et al., 2010). Enfin, la décomposition en ondelettes et les approches de simulation simples peuvent être appliquées car il existe également des stations de mesures *in situ* permanentes mesurant les variables de neige continuellement dans le temps et la neige est une conséquence des événements de précipitation.

La **précipitation** est une autre variable géophysique qui souffre de problèmes similaires d'échelle d'observation. Par exemple, les produits pluviométriques du satellite TRMM sont fournis à une résolution de 0,25°(∼25 km) et les pluviomètres mesurent à un point particulier dans l'espace. En outre, des questions similaires à celles posées sur l'humidité du sol à grand support sont présentes dans le cas de la variable pluie. Qu'est-ce que le taux de précipitations à la résolution de 25 km ? La moyenne des précipitations dans la région ? Des décompositions en ondelettes pourraient être appliquées aux précipitations car il existe à la fois des longues séries de mesures *in situ* et satellitaires de pluie.

Une autre variable qui pourrait être étudiée avec les approches développées dans cette thèse est la **température de surface de la terre** (**LST**, pour son acronyme en anglais). Il existe des stations de mesures *in situ* permanentes qui mesurent la LST à des résolutions allant jusqu'à 10 m. Bien qu'il existe des missions satellites offrant des produits LST à des résolutions similaires (e.g. Landsat à 30 m), des capteurs de résolution plus grossiers comme MODIS (1 km) sont toujours utiliséset doivent être validés. La LST est fondamentalement hétérogène dans une empreinte de 1 km ; sa validation souffre également de la différence d'échelle spatiale.

6.3 Questions ouvertes

La recherche commence mais ne se termine jamais et cette thèse n'est pas une exception. Cette thèse laisse certaines questions ouvertes, soit parce que leurs réponses n'étaient pas incluses dans ses objectifs, soit parce que les réponses étaient plus complexes que prévu, soit encore parce qu'elles sont apparues à la lueur de ces travaux.

La principale question pratique de ce travail de thèse était de savoir si la différence d'échelle spatiale entre les deux mesures d'humidité du sol, satellitaire et *in situ*, pouvait être évaluée. J'ai montré que c'est faisable, au moins partiellement. L'incertitude d'échantillonnage peut être évaluée comme le montre le chapitre 5. En revanche, l'évaluation de la différence d'échelle spatiale totale nécessite des recherches supplémentaires. Dans la section 6.2.1, j'ai suggéré quelques voies qui peuvent être explorées. En conclusion, **l'obtention de la différence d'échelle spatiale minimum présente dans les statistiques temporelles de validation satellitaires semble possible**.

Des questions plus générales et presque philosophiques n'ont pas encore de réponses claires. Par exemple, la notion de **représentativité spatiale** reste ouverte. Toutes les méthodes utilisées pour l'évaluation de la représentativité spatiale sont basées sur des mesures de corrélation (chapitre 4). Cependant, la représentativité pourrait s'exprimer avec d'autres paramètres comme la variance. Peut-on affirmer qu'une série temporelle de mesure *in situ* est représentative des séries temporelles à grand support lorsque leurs dynamiques respectives sont bien corrélées alors qu'elles sont très différentes en variance ? En général, variance et corrélation ne peuvent pas être maximisées en même temps. Par conséquent, la communauté scientifique s'intéressant à la thématique de l'humidité des sols doit décider quelles sont les métriques les plus pertinentes à utiliser pour l'évaluation de la représentativité spatiale.

Au chapitre 1, j'ai présenté les **deux définitions possibles pour l'humidité du sol à grand support**. Serait-on prêt à accepter que l'humidité du sol à grand support soit une traduction de l'énergie intégrée et qu'elle ne devrait pas être obligatoirement égale à l'humidité du sol moyenne ? Ou considérerait-on que l'humidité du sol issue d'un satellite corresponde parfaitement à l'humidité du sol moyenne (inconnue) de la région d'étude ? Dans le premier cas, les modèles opérant sur des grandes échelles spatiales devraient s'adapter à ce que les satellites voient. C'est une question importante qui doit être discuté par la communauté scientifique.

Enfin, en dehors de la différence d'échelle spatiale, il existe **d'autres sources d'écart** entre le satellite et les mesures *in situ* : l'écart temporel et le d'écart vertical. L'écart temporel apparaît quand les mesures *in situ* ne sont pas prises exactement au moment du survol du satellite. L'écart vertical apparaît lorsque la profondeur de détection du satellite et celle des capteurs *in situ* ne coïncident pas. On sait que la profondeur de détection du capteur satellitaire varie dans le temps en raison des changements dans les conditions de surface (niveau d'humidité du sol, texture,végétation). En revanche, la station *in situ* est installée à une profondeur constante, à moins que des changements dans le sol aient lieu, ce qui n'est pas si rare. Cette discordance verticale se traduit principalement en décorrélation et, dans une moindre mesure sans doute, en variabilité. La prise en compte de cet effet n'entrait pas dans les objectifs de cette thèse, mais il serait important de les considérer lors de la validation des produits d'humidité du sol satellitaires ou du moins d'évaluer leur impact car ils contribuent nécessairement à l'incertitude dans les statistiques.

6.4 Note finale

Cette thèse avait pour objectif de contribuer à une meilleure compréhension de la différence d'échelle spatiale entre les mesures d'humidité du sol par satellite et *in situ*. Elle a montré que les échelles spatiales et temporelles de l'humidité du sol sont interconnectées et que la différence

d'échelle spatiale peut être partiellement estimée. Cela a été possible grâce à l'adaptation et aux développements des approches génériques basées sur des outils de traitement du signal et des outils statistiques. Ces approches peuvent être utilisées globalement et adaptées à d'autres variables géophysiques.

L'incertitude dans les statistiques temporelles de validation de produits satellitaires induite par la différence d'échelle spatiale a été généralement négligée, très probablement parce qu'elle est difficile à estimer. J'espère que cette thèse aura aussi sensibilisé sinon convaincu le lecteur de l'influence majeure de la différence d'échelle spatiale dans la validation des produits satellitaires. Cette différence d'échelle peut être très significative dans les dynamiques sous-hebdomadaires. L'une de ses composantes, l'incertitude d'échantillonnage, peut complètement changer la comparaison des performances relatives de deux produits satellitaires. Il est donc essentiel lorsque l'on compare des produits satellitaires d'humidité du sol de réfléchir à la signification statistique des différences : rendent-elles compte des performances réelles ou bien sont-elles obtenues par hasard en raison de la différence d'échelle spatiale ?

Enfin, il existe une croyance générale selon laquelle l'amélioration de la résolution des produits satellitaires d'humidité du sol rendra négligeable la différence d'échelle spatiale. Différentes solutions ont été proposées (SMAP) ou sont en cours de développement (SMOS-Next, SMOS-HR) pour fournir des produits d'humidité du sol globaux à des résolutions spatiales améliorées. Ces résolutions cibles sont limitées aux échelles kilométriques (1 à 10 km) parce que les solutions sont basées sur des capteurs en bande L caractérisés par une grande sensibilité à l'humidité du sol dans une large gamme de conditions de végétation. Cependant, les hautes résolutions kilométriques ne supprimeront pas les différences spatiales : l'humidité du sol peut varier rapidement dans quelques mètres seulement. La problématique de la différence d'échelle spatiale entre les mesures d'humidité du sol par satellite et les mesures *in situ* restera un sujet d'actualité pour encore de nombreuses années.

Appendix A

The DISPATCH processor

The descriptions herein concern he operational version of DISPATCH integrated in the Level-4 (L4) of the French ground segment of the SMOS mission. The DISPATCH processor, called CATDS Level-4 Disaggregation processor (C4DIS), disaggregates SMOS Level-3 soil moisture products to a final resolution of 0.01°.

A.1 DISPATCH algorithm

DISPATCH distributes the SM within the low-resolution pixel according to the spatial distribution of the soil evaporative efficiency (SEE). The SEE is estimated at the target higher resolution (HR) following the model:

$$SEE_{HR} = \frac{T_{s,max} - T_s}{T_{s,max} - T_{s,min}} \tag{A.1}$$

with T_s being the soil temperature and $T_{s,max}$ and $T_{s,min}$ being the soil temperatures at minimum and maximum soil moisture, respectively. They are derived from the optical observations (MODIS). Using the triangle approach (Carlson et al., 1994), T_s can be expressed as:

$$T_s = \frac{T_{surf} - f_v T_v}{1 - f_v} \tag{A.2}$$

with T_{surf} being the LST from the higher-resolution sensor (MODIS), f_v the fractional vegetation cover and T_v the vegetation temperature, both to be estimated. The triangle method is illustrated in Figure A.1

T_{surf} has been previously corrected for elevation effects (Merlin et al., 2013):

$$T_{surf} = T_{surf,original} + \gamma \cdot (H_{HR} - H_{LR}) \tag{A.3}$$

with γ ($°C/m$) the mean lapse rate (set to 0.006 $°C/m$ and valid locally), H_{HR} the altitude of the HR pixel and H_{LR} the mean altitude of within the low-resolution pixel.

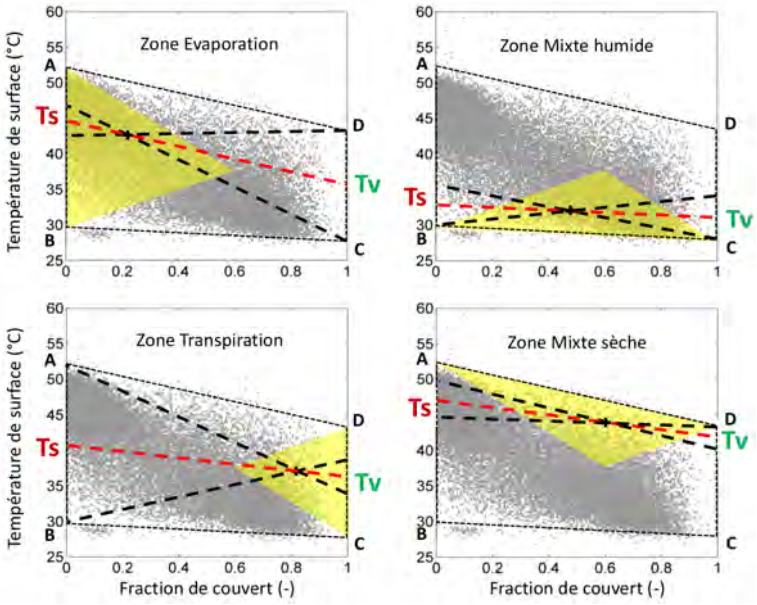


Figure A.1 – Example of the triangle method of Moran et al. (1994), which is based on the position of the pixel in the LST-f_v space, applied to the ASTER data acquired on the 17/08/2011 in Urgell, Spain. Source: Merlin (2016)

The fractional vegetation cover is calculated as:

$$f_v = \frac{NDVI - NDVI_s}{NDVI_v - NDVI_s} \tag{A.4}$$

where $NDVI$ is obtained from the higher-resolution sensor (MODIS), and $NDVI_s$ and $NDVI_v$ are set to arbitrary NDVI values for bare soil (0.15) and full-vegetated surface (0.90), respectively.

Vegetation temperature T_v in Equation A.2 is estimated with the "hourglass" approach in Moran et al. (1994). Temperature end-members in Equation A.1 ($T_{s,min}$, $T_{s,max}$) are obtained from the minimum and the maximum surface temperatures in the scene ($T_{surf,min}$, $T_{surf,max}$), respectively. This is done through approximations in the LST-f_v space (A.1 and Merlin et al. (2013)):

1. At the pixel with minimum temperature $T_{surf,min}$
   (a) if the vegetated portion is low ($f_v < 0.5$), we assume $f_v \approx 0$ to derive the minimum soil temperature, $T_{s,min} = T_{surf,min}$. $T_{v,min}$ is set to $T_{surf,min}$ as well
   (b) if the vegetated portion is large ($f_v >= 0.5$), we assume $f_v \approx 1$ to derive the minimum vegetation temperature, $T_{v,min} = T_{surf,min}$. Then, $T_{s,min}$ is derived as the minimum T_s obtained from Equation A.2 with $T_v = T_{v,min}$ in poorly vegetated pixels ($f_v < 0.5$)
2. At the pixel with maximum temperature $T_{surf,max}$
   (a) if $f_v < 0.5$, we assume $f_v \approx 0$ and $T_{s,max} = T_{surf,max}$. $T_{v,max}$ is derived as the maximum T_v obtained from Equation A.2 with $T_s = T_{s,max}$ in highly vegetated pixels ($f_v >= 0.5$)
   (b) if $f_v >= 0.5$, we assume $f_v \approx 1$ and $T_{v,max} = T_{surf,max}$. Then, $T_{s,max}$ is derived as the maximum T_s obtained from Equation A.2 with $T_v = T_{v,max}$ under low vegetated conditions ($f_v < 0.5$)

Once SEE_{HR} has been obtained for all the HR pixels within the large-support pixel, the downscaling relationship can be applied to obtain soil moisture at the target resolution (SM_{HR}):

$$SM_{HR} = SM_{LR} + \frac{\partial SM}{\partial SEE_{LR}} \cdot (SEE_{HR} - SEE_{LR}) \tag{A.5}$$

with SM_{LR} being the soil moisture observed by the low resolution sensor (SMOS) and SEE_{LR} the spatial average of the SEE_{HR} values within the LR pixel. The partial derivative in Equation A.5 equals SM_p, a paramater that is calculated at LR following the linear SEE-SM model in Budyko (1961):

$$SM_p = SM_{LR}/SEE_{LR} \tag{A.6}$$

A.2 The C4DIS processor

Herein, the basic features of the C4DIS processor are described. A more detailed description of the processor can be found in Molero et al. (2016).

The pre-processor prepares the data for the application of DISPATCH. Up to 24 possible combinations of LR and HR data are generated as follows: First, the pre-processor resamples the SMOS L3 dataset from the original 25-km grid to four 50-km grids. This produces four LR observations per input SMOS dataset. The procedure is further described in Merlin et al. (2012) and Malbéteau et al. (2016). Then, the pre-processor selects the MODIS LST datasets available within a **3-day window** centred on the SMOS date, both Aqua and Terra observations included. In consequence, each of the 4 resampled SMOS observations can be disaggregated using the surface temperature of 3x2 MODIS products, which makes 24 possible combinations.

DISPATCH is applied to each of the (up to) 24 sets of observations. Less than 24 sets are possible if some of the products are not available or if some of the LST/NDVI products exceed the cloud-free and sea-free thresholds (67 % and 90 %, respectively). Moreover, the operational version filters those HR pixels that do not have maximum quality (quality flags 0 and 17, Wan (2006) and Solano et al. (2010)).

The (up to) 24 disaggregated SM datasets are averaged together in one final dataset called **SM_HR**. This helps to reduce independent random errors. However, the 3-day window is based on the assumption that SM fields are stable for periods of at least 1 day around the SMOS overpass time. This assumption is not satisfied if a precipitation event occurs in between, in which case additional noise will be introduced.

The final L4 disaggregated product contains 3 datasets: the SM_HR dataset and two other that could serve as indicators of the method uncertainty. These are the **STD** dataset, which contains the standard deviation of the 24-disaggregated output ensemble, and the **COUNT** dataset, which is the size of the ensemble. SM_HR values are masked if COUNT values are lower than 3.

Appendix B

Confidence intervals (CI) for the wavelet variance

The CIs of the biased estimator of the wavelet variance $\hat{\sigma}_j^2$ can be calculated following two methods according to Percival and Walden (2000, pp. 311-315). The first assumes that $\{W_j(t)\}$ is a Gaussian process, so for a large number of samples N, the sample variance $\hat{\sigma}_j^2$ is approximately Gaussian distributed with mean σ_j^2 and variance $2A_j/N$, A_j a constant. The CIs obtained with this method are often inconvenient because they can have a negative lower limit. The other method does not suffer from this drawback by assuming that $df \cdot \hat{\sigma}_j^2/\sigma_j^2$ is a chi-square random variable with df equivalent degrees of freedom (χ_{df}^2). The resultant CI is:

$$[CI_l, CI_u] = \left[\frac{df \cdot \hat{\sigma}_j^2}{Q_{df}(1-p)}, \quad \frac{df \cdot \hat{\sigma}_j^2}{Q_{df}(p)} \right] \tag{B.1}$$

where $Q_{df}(p)$ is the p probability point of the χ_{df}^2 distribution. The above equation requires a value for df that can be calculated following different strategies. If the sample size is large enough ($N = 128$ is already considered sufficient) the following expression for df gives CIs that are "reasonable accurate" (Percival and Walden, 2000, p. 315):

$$\hat{df} = \frac{N\hat{\sigma}_j^4}{\hat{A}_j} \tag{B.2}$$

where $\hat{A}_j$ is a constant derived from the spectral density function $\hat{S}_j(f)$ of the wavelet coefficients $\{W_j(t)\}$:

$$\hat{A}_j = \frac{1}{2}\int_{-1/2}^{1/2} (\hat{S}_j(f))^2 df \tag{B.3}$$

Given that the length of the time series analysed in this thesis is largely bigger than 128, the Equations B.1, B.2 and B.3 are used to derive the CIs for the biased wavelet variance. The following figure shows the wavelet variance of the Little Washita *in situ* series described in section 3.3 with their respective CIs calculated with the selected method. The Little Washita series had a length of 3 years and were sampled at 3 h interval (8760 samples).

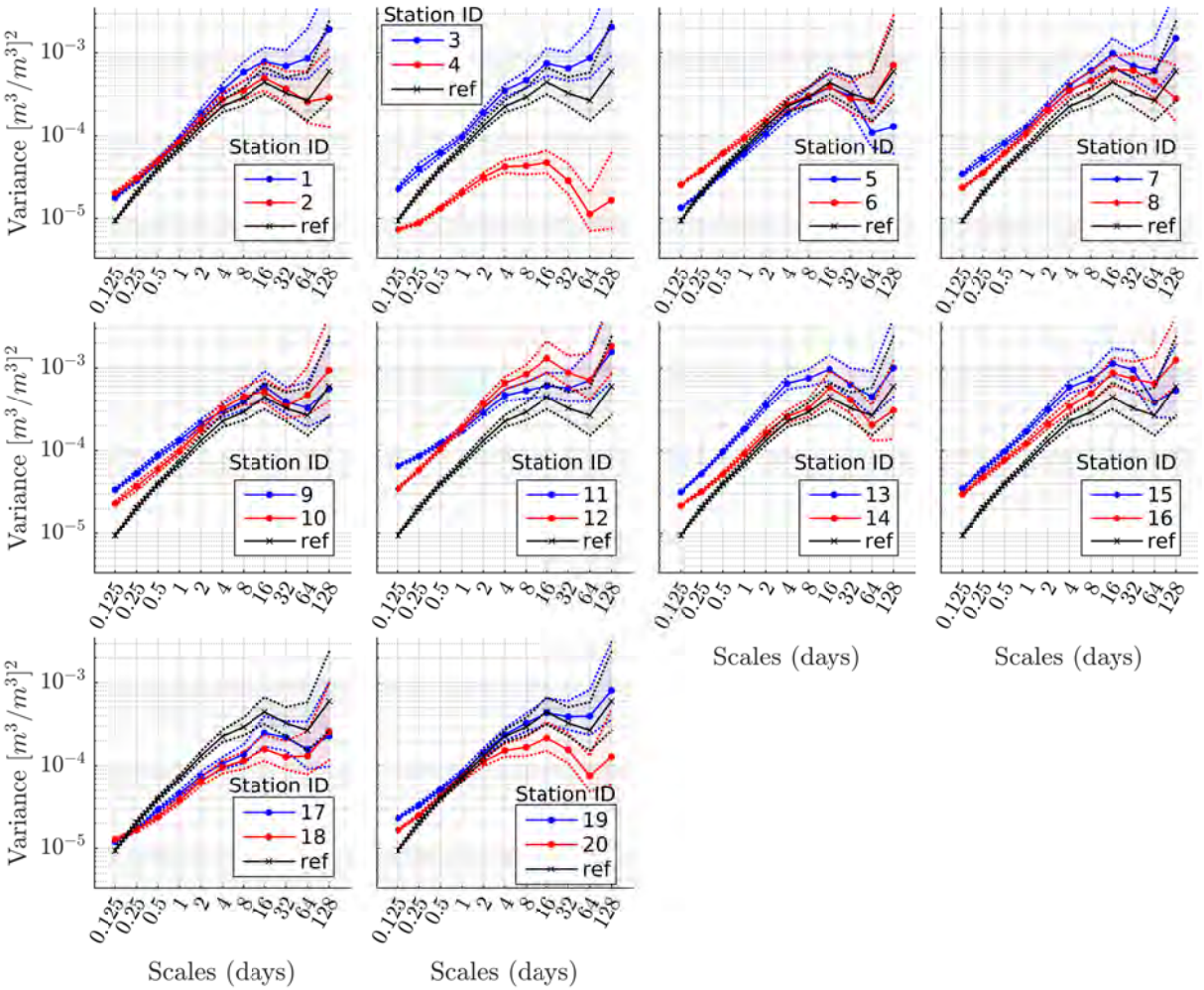


Figure B.1 – Variance decomposition of the *in situ* time series (color lines) and the average *in situ* series (black line) of the Little Washita network, with the respective confidence intervals

Appendix C

Confidence intervals (CI) for the wavelet correlation

There exists one available method to compute the confidence intervals (CIs) of the sample wavelet correlation. However, it produces CIs that are excessively large at large time scales. Herein, I present the theory on CIs for the correlation coefficient and I propose a new approach to produce CIs at large time scales that are more tractable.

The sample correlation $R_{X,Y}$ measures the strength of the linear relationship between a set of N bivariate observations, $(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)$:

$$R = \frac{1}{N} \sum_{n=1}^{N} \frac{x_n - \hat{\mu}_X}{\hat{\sigma}_X} \cdot \frac{y_n - \hat{\mu}_Y}{\hat{\sigma}_Y} \tag{C.1}$$

If the observations are independent and follow the same bivariate distribution, R is a consistent and approximately unbiased estimator of the population correlation $\rho_{X,Y}$ (Rodriguez, 2006). The Fisher's z transformation (Equation C.2) helps for the construction of confidence intervals for R, since it approaches normality much faster than R. Rodriguez (2006) showed that the standard deviation of z is approximately $(N-3)^{-1/2}$ and that the variance of R is inversely dependent on the sample size N (Equation C.3).

$$z = tanh^{-1} R = \frac{1}{2} log\left(\frac{1+R}{1-R}\right) \tag{C.2}$$

$$Var\{R_{x,y}\} = \frac{(1-\rho_{xy}^2)^2}{N_{eff}-1} \cdot \left(1 + \frac{11\rho_{xy}^2}{2N_{eff}}\right) + \mathcal{O}(N_{eff}^{-3}) \tag{C.3}$$

This estimation of the statistical significance of R is adequate when the data have a bivariate normal distribution and do not exhibit serial auto-correlation. However, most typical geophysical series, like soil moisture, have serial dependence (Ebisuzaki, 1997). Serial dependence reduces the effective sample size from N to N_{eff} independent samples (Mudelsee, 2003). As can be inferred from Equations C.2 and C.3, the reduction in sample size widens the uncertainty limits for R.

For time series with some degree of auto-correlation, Percival et al. (2000) identified two possible ways to compute the uncertainty of the wavelet autocorrelation defined in Equation C.4. The first option is to estimate N_{eff} and replace N by N_{eff} in Equation C.3. The second option consists in bootstrap approaches.

$$R_{X,Y,j} = \frac{Cov\{W_{X,j}, W_{Y,j}\}}{\sqrt{\hat{\sigma}^2_{X,j} \cdot \hat{\sigma}^2_{Y,j}}} \tag{C.4}$$

CIs with the analytical method (wavelet coefficients) Whitcher et al. (2000) proposed an analytical solution to approximate the $100(1-2p)\%$ confidence intervals (CIs) for the MODWT wavelet correlation at scale j:

$$[CI_l, CI_u]_j = \left[tanh\left\{ z_j - \frac{\Phi^{-1}(1-p)}{\sqrt{N_{\mathit{eff},j} - 3}} \right\}, \quad tanh\left\{ z_j + \frac{\Phi^{-1}(1-p)}{\sqrt{N_{\mathit{eff},j} - 3}} \right\} \right] \tag{C.5}$$

Where z_j is the z-transform of $R_{X,Y,j}$, $\Phi^{-1}(1-p)$ is the p probability point for the standard normal distribution and $N_{\mathit{eff},j}$ is the effective sample size at scale j. They propose to approximate $N_{eff,j}$ as the number of DWT coefficients at scale j (Equation C.6). The rationale is that DWT coefficients can be regarded as uncorrelated if the original series is equal or similar to a long-memory process (Percival et al., 2000) (MODWT are always serially correlated).

$$N_{\mathit{eff},j} = \left\lfloor \frac{N}{w_{a,j}} \right\rfloor \tag{C.6}$$

where $w_{a,j} = 2^j$ is the autocorrelation length for the equivalent wavelet filter, i.e. the number of samples that contribute the most in the construction of one wavelet coefficient (Cornish et al., 2006).

The drawback of this approach is that former research showed that the estimation of N_{eff} is not trivial and the results are not always consistent (Ebisuzaki, 1997). Maybe because of this, the resultant CIs of the analytical method seem too pessimistic for large scales. To illustrate this, I applied this method to the wavelet correlation between the point *in situ* series and the average *in situ* series of the Little Washita dataset described in section 3.3. The series expanded 3 years and were sampled at 3 h interval. Eleven levels of decomposition were produced ($J0 = 11$, τ_j=128 days, 8760 samples). I found that the comparison of the wavelet correlations obtained was intractable from scale 7 (8 days) onwards due to too wide CIs (not shown here). This was explained by the very rapid decrease in the effective number of samples with scale. For example, at scale 11, the autocorrelation width was $w_{a,11} = 2^{11} = 2048$ samples so the effective sample size was reduced to $N_{eff,9} = \lfloor 8760/2048 \rfloor = 4$.

CIs with bootstrap approaches (detail series) I propose a new method to compute the CIs for the wavelet-based correlation metric defined in section 3.2.4. The method is based on bootstrap approaches. In contrast to the analytical method, bootstrap is non-parametric, i.e. it does not require any assumption on the distribution of $R_{X,Y,j}$. However, standard bootstrap does not take into account dependence between the samples. A simple way to account for the serial dependence is not to break blocks of samples that are strongly auto-correlated. In other words, the bootstrapped samples are not samples but blocks of samples. This method is called *block bootstrap* (Davison and Hinkley, 1997). The block length must be chosen not smaller than the temporal auto-correlation length so that the output bootstrapped series is not "whitened". As explained in the description of the analytical method, the estimation of auto-correlation length is not trivial so the value $block\ length = w_{a,j} = 2^j$ seems a good preliminary approximation.

An important drawback when applying block bootstrap is that the number of blocks is reduced significantly in the last scales. If we consider again the Little Washita 3-years time series (2190 samples) as example, at scale 9 the number of blocks would be $\lfloor 2190/512 \rfloor = 4$. A way to increase the sample space is to allow the blocks to overlap (Davison and Hinkley, 1997).

The CIs for Figure C.1 shows the correlation and the 95 % CIs calculated with the *overlapping block bootstrap* on the Little Washita series. The CIs here exhibit a strong correspondence between correlation levels and uncertainty: the lower the correlation, the higher the uncertainty, and *vice versa*. The effect in the last scale is striking: the series exhibiting correlation close to 1 have extremely narrow CIs. However, the dependence with $N_{eff,j}$ is always present. For moderate values of correlation (0.5-0.9), same correlation values at different time scales do not have the same CIs: the larger scale has wider CIs (e.g. series #12, the CIs of the 64-day scale are wider that those of the 0.5-day scale). For a more detailed view of the bootstrap process, Figures C.2 and C.3 show the bootstrapped distributions at each time scale for series #1 and #3, respectively.

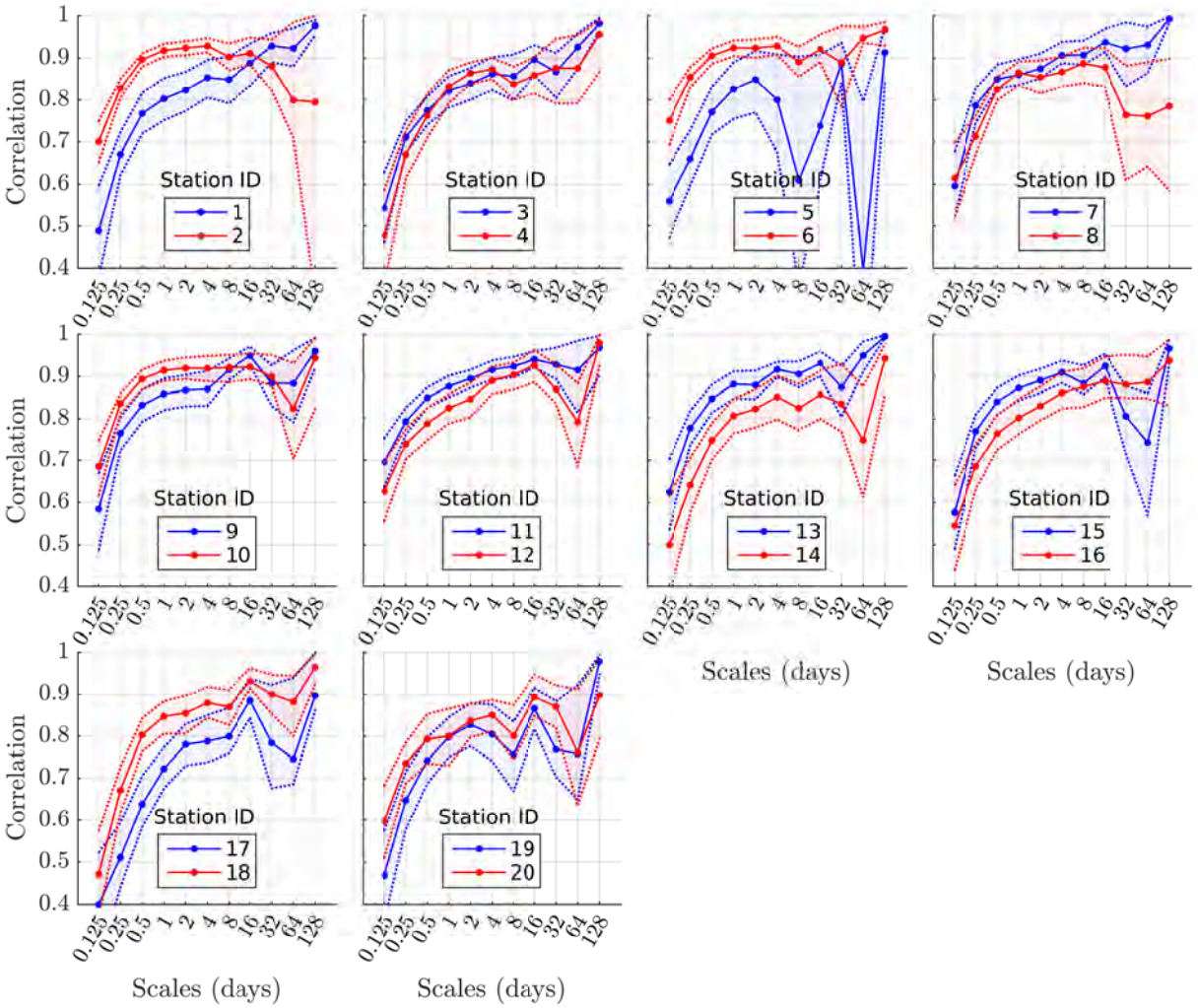


Figure C.1 – Correlation between the detail series of the point *in situ* series and the average *in situ* series of Little Washita, with the respective 95 % CIs calculated with the overlapping block bootstrap method

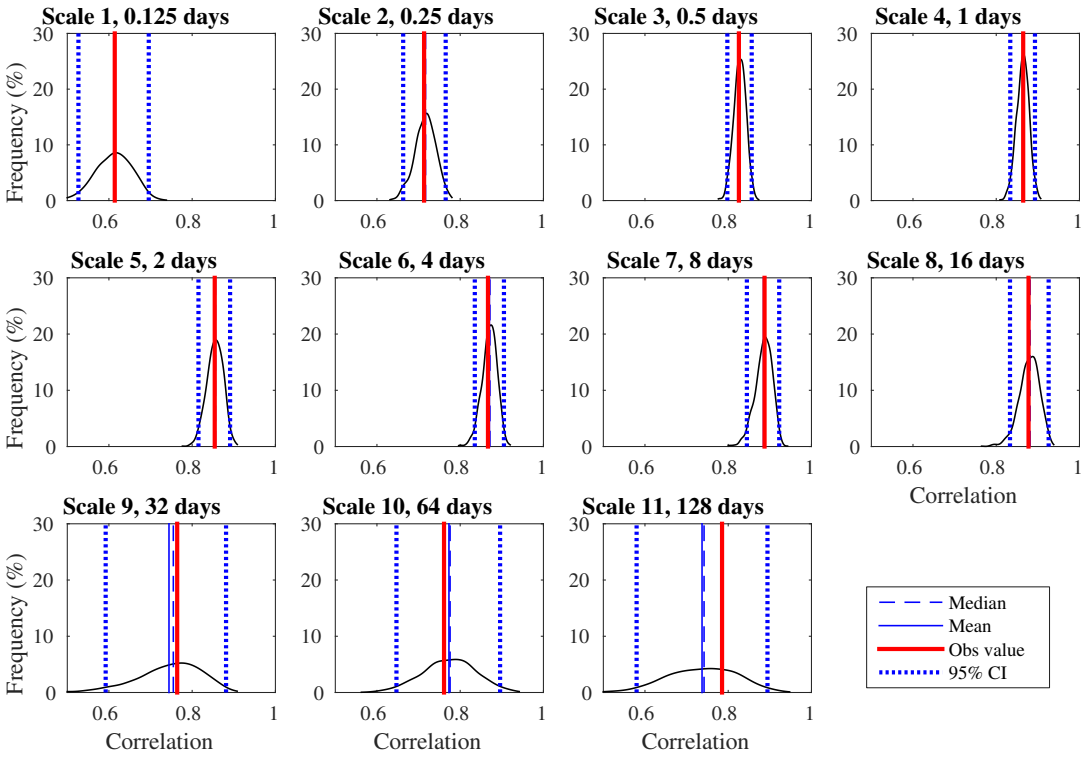


Figure C.2 – Bootstrap distributions of the correlation between **station #8** series and the average *in situ* series of the Little Washita network

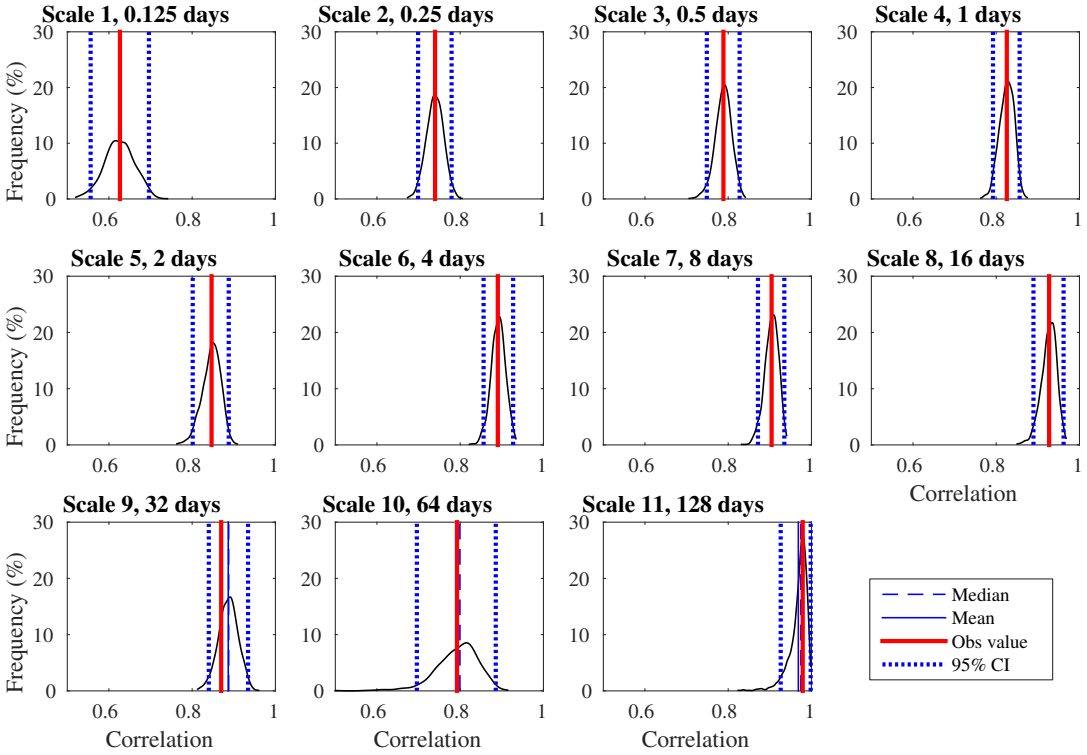


Figure C.3 – Bootstrap distributions of the correlation between **station #12** series and the average *in situ* series of the Little Washita network

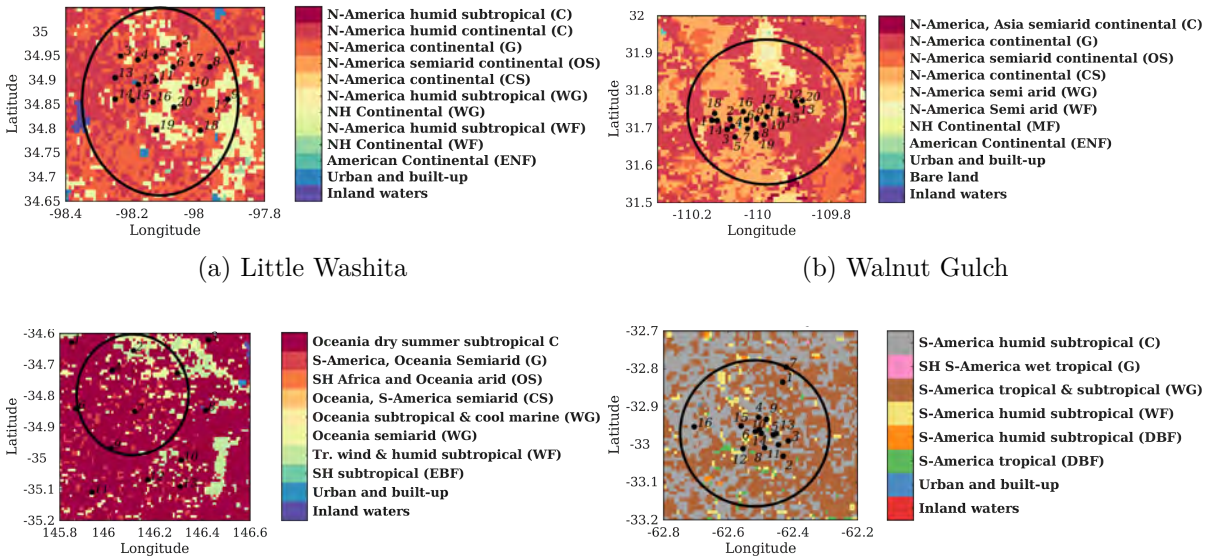
Appendix D

In situ networks: land cover maps

Figure D.1 shows the Ecoclimap land cover map (Masson et al., 2003) for the four dense *in situ* networks used in this thesis. It is presented in the Discrete Flexible Fine Grid (DFFG), 4 by 4 km grid used in the SMOS Level-2 processor for weighting land cover classes (Richaume et al., 2006; Kerr et al., 2014).

The Ecoclimap database accounts for a total of 218 ecosystems. They can be summarised in a smaller number of classes for modeling purposes, as they are in the SMOS Level-2 processor. Figure D.1 shows the maps of the aggregated land cover classes for the same four dense *in situ* networks as before.

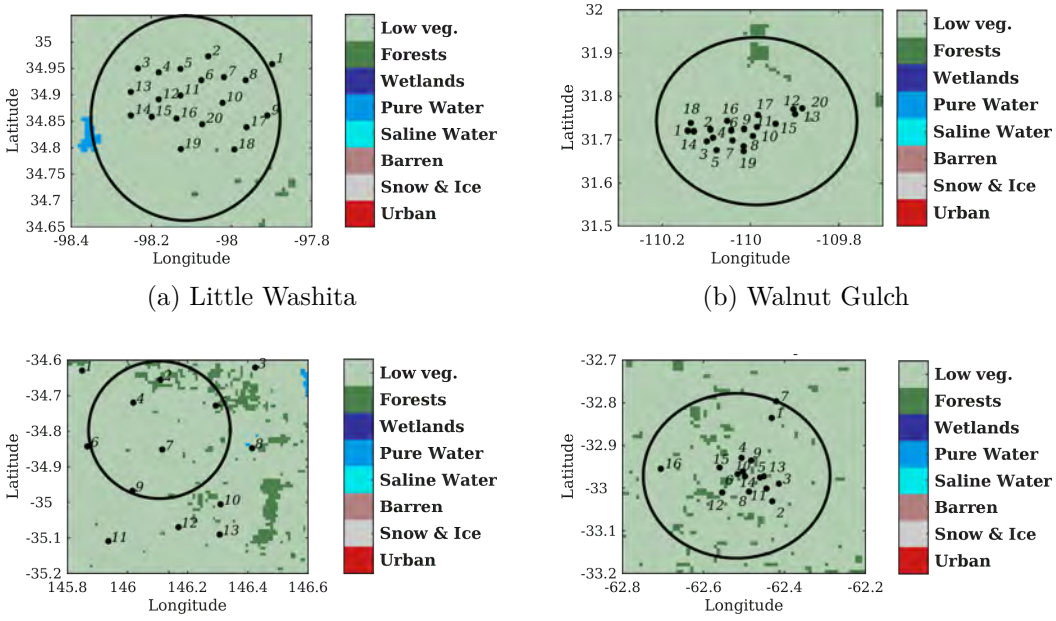
By looking at both Figures D.1 and Figures D.2, we can observe that the networks are homogeneous in terms of big land cover classes and low vegetation is the main class. However, we may find different low vegetation types within each network. In the absence of finer classification data, Figure D.3 illustrates well this situation in Yanco. In this figure, it can be seen that the NDVI time series differ between stations in the Yanco network, implying different agricultural practices and/or crops.



(a) Little Washita (b) Walnut Gulch

(c) Yanco (d) Monte Buey

Figure D.1 – Ecoclimap land cover maps for the four dense *in situ* networks used in this thesis. The circle represents a mean SMOS footprint size of 40 km.



(a) Little Washita (b) Walnut Gulch

(c) Yanco (d) Monte Buey

Figure D.2 – Aggregated Ecoclimap land cover maps for the four dense *in situ* networks used in this thesis. The aggregated classes are those considered in the SMOS retrieval algorithms. The circle represents a mean SMOS footprint size of 40 km.

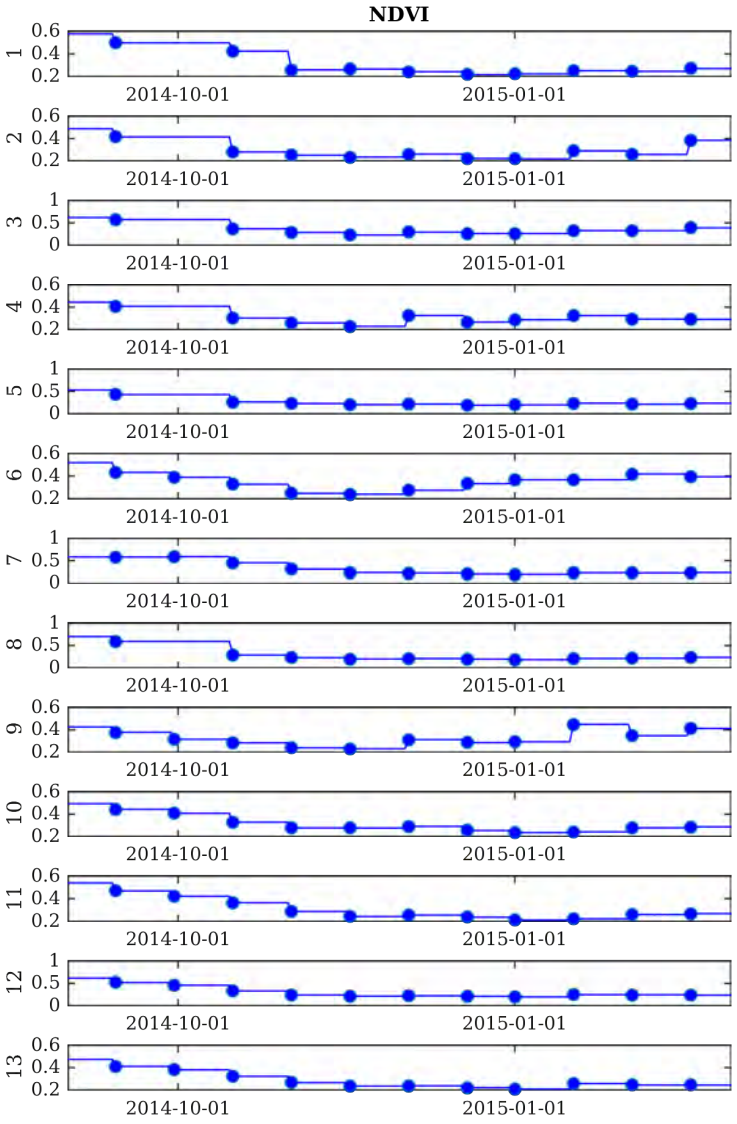


Figure D.3 – NDVI time series observed at each of the station locations (the station ID is indicated in the axis labels at the left). The values are extracted from the NDVI raster, input to the disaggregation algorithm in the C4DIS processor (Molero et al., 2016).

Appendix E

Temporal empirical models of 5 *in situ* series

The following model is obtained for the temporal dynamics of the SM measured at stations #4, #6, #9, #10, and #13 of the Little Washita network during the 2013/01 - 2013/06 period:

$$f(SM_t) = SM_{t+1} = \begin{cases} SM_t + f_p(prec_t, SM_t), & \text{if } prec_t > 0 \\ SM_t \cdot e^{-f_d(LAI_t, SM_t)\cdot \frac{\Delta t}{Z}}, & \text{otherwise} \end{cases} \tag{E.1}$$

where Δt is the sampling period, Z is the soil depth (5 cm), and $f_p(\cdot)$ and $f_d(\cdot)$ are the peak and decay models that are expressed as:

$$f_p(prec_t, SM_t) = a_{00} + a_{10}SM_t + a_{01}prec_t + a_{20}SM_t^2 + a_{02}prec_t^2 + a_{11}prec_tSM_t \tag{E.2a}$$

$$f_d(LAI_t, SM_t) = b_{00} + b_{10}SM_t + b_{01}LAI_t + b_{20}SM_t^2 + b_{11}LAI_tSM_t \tag{E.2b}$$

The a and b coefficients are obtained by least-squares fit (Levenberg-Marquardt method). The least-squares fit minimises the sum of the squares of the deviations between the observation and the model. In the case of the peak and decay models, the least-squares initial equations are:

$$\hat{a}_{00,\ldots,11} = \arg\min_a \sum_{i=1}^{m} [\Delta SM_i - f_p(prec_i, SM_i, a_{00,\ldots,11})] \tag{E.3a}$$

$$\hat{b}_{00,\ldots,11} = \arg\min_b \sum_{i=1}^{m} [E_i - f_d(LAI_i, SM_i, b_{00,\ldots,11})] \tag{E.3b}$$

where m is the number of samples used for fitting and $E = -ln(\Delta SM)\frac{Z}{\Delta t}$. Normally, m should have been equal the total number of time intervals $N-1$ and Δt should have denoted the sampling period (30 min) and ΔSM, the difference between two consecutive SM values, $SM_{t+1} - SM_t$. However, the models found did not lead to simulated SM series reasonably close to the measured SM series. This can be observed in Figure 3.6, where the modelled and the measured series differ in the decay patterns (slightly) and in the minimum and maximum values reached (moderately).

Better models were finally found by selecting only some of the samples of the series for the fit. For fitting the peak model, only the samples at the beginning and at the end of a peak interval were considered. A *peak interval* is defined as an interval of increasing SM leading to a peak. For fitting the decay model, only the samples at the beginning and at the end of a dry-down

decay were considered. This implies that prior to the application of Equation E.3 the peak and decay intervals had to be identified. This was not straightforward since the precipitation did not always preceded by one time step the increase in SM. The algorithm for the detection of the peak intervals consisted in i) selecting the local maxima (peaks) preceding a decay, ii) localizing the minimum SM value preceding each peak by 1 time step or more (and after any other local maximum). The selection of the peak intervals is illustrated in Figure E.1 for station #9. The algorithm for the detection of the decay intervals repeated i) and changed the 2$^{\text{nd}}$ step by: ii') localizing the minimum SM value *after* each peak by 1 time step or more (and before any other local maximum).

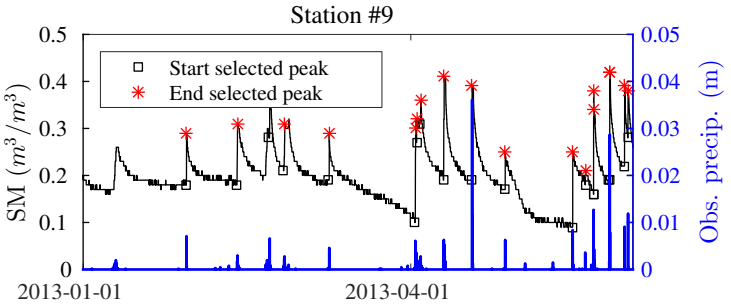
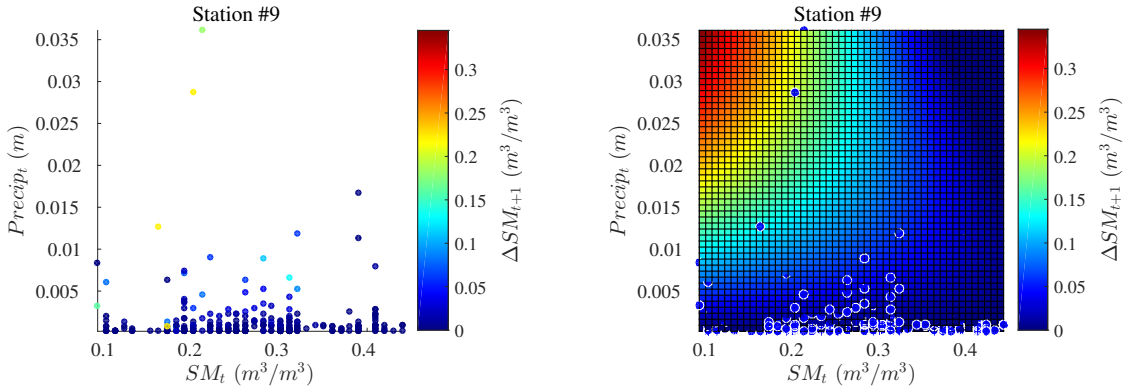


Figure E.1 – Pre-selection of peaks and decays for the fit of the peak model for station #9

After the selection of m intervals, Equation E.3 could be solved by replacing Δt by the interval length (in time units) and ΔSM by the difference in SM between the end and the start of the interval. As an example, Figures E.2 and E.3 show the fit obtained for the peak and decay models for station #9.

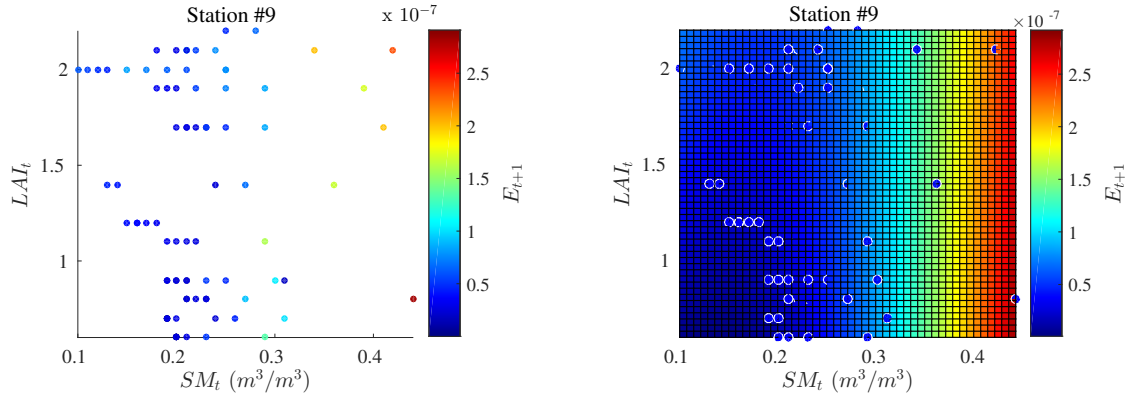


(a) Observed data

(b) Model fit

Figure E.2 – Fit of the peaks model to the observed data at station #9

The fitted a and b coefficients were found different depending on the *in situ* series used as a reference. The final modelled series for each of the stations are displayed in the following figures, together with their reference *in situ* SM series and the explanatory variables used, LAI and precipitation.



(a) Observed data (b) Model fit

Figure E.3 – Fit of the decay model to the observed data at station #9

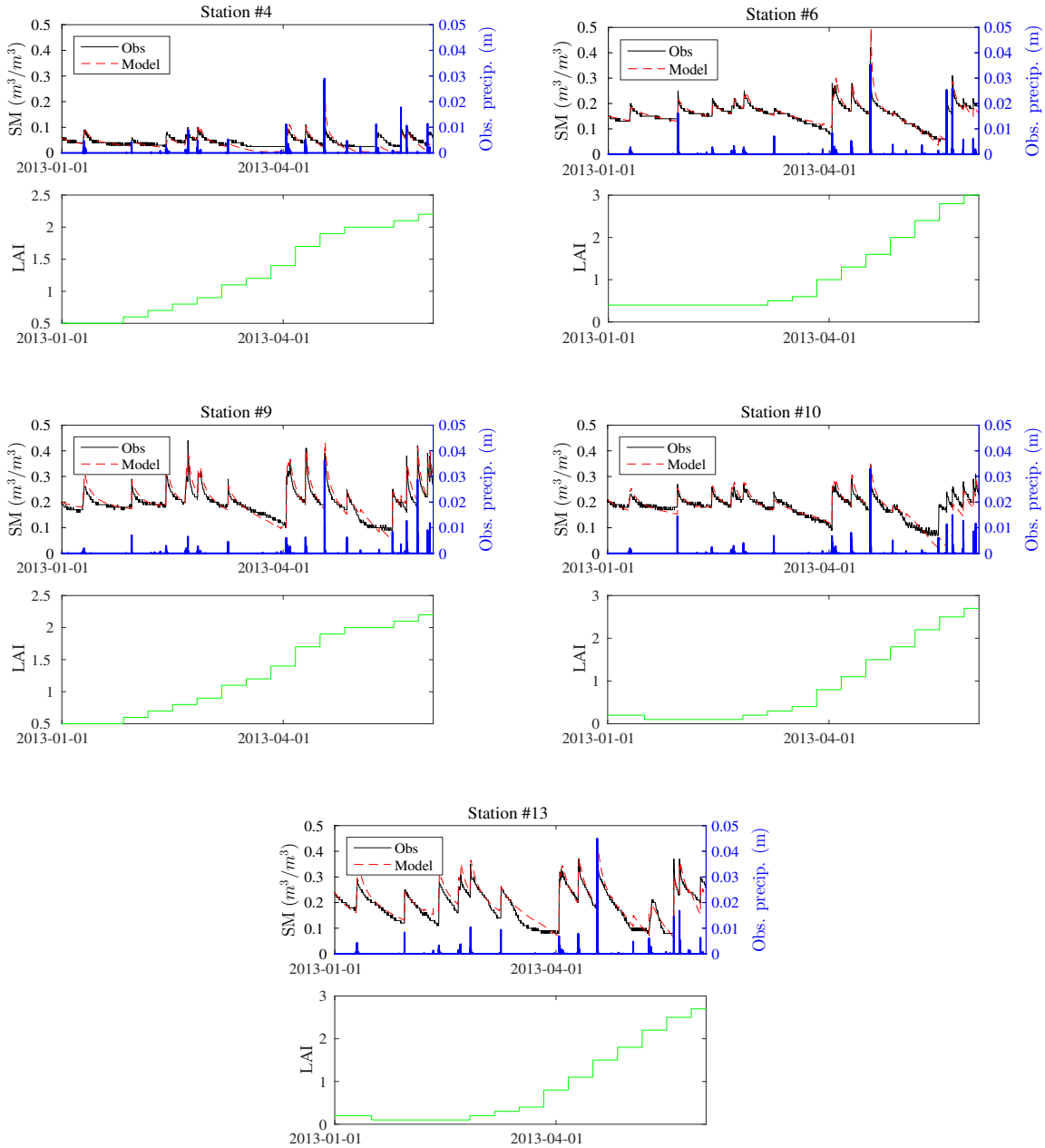


Figure E.4 – Observed and modelled time series at each of the 5 selected stations locations: #4, #6, #9, #10 and #13

Appendix F

Bootstrap distributions and sample size

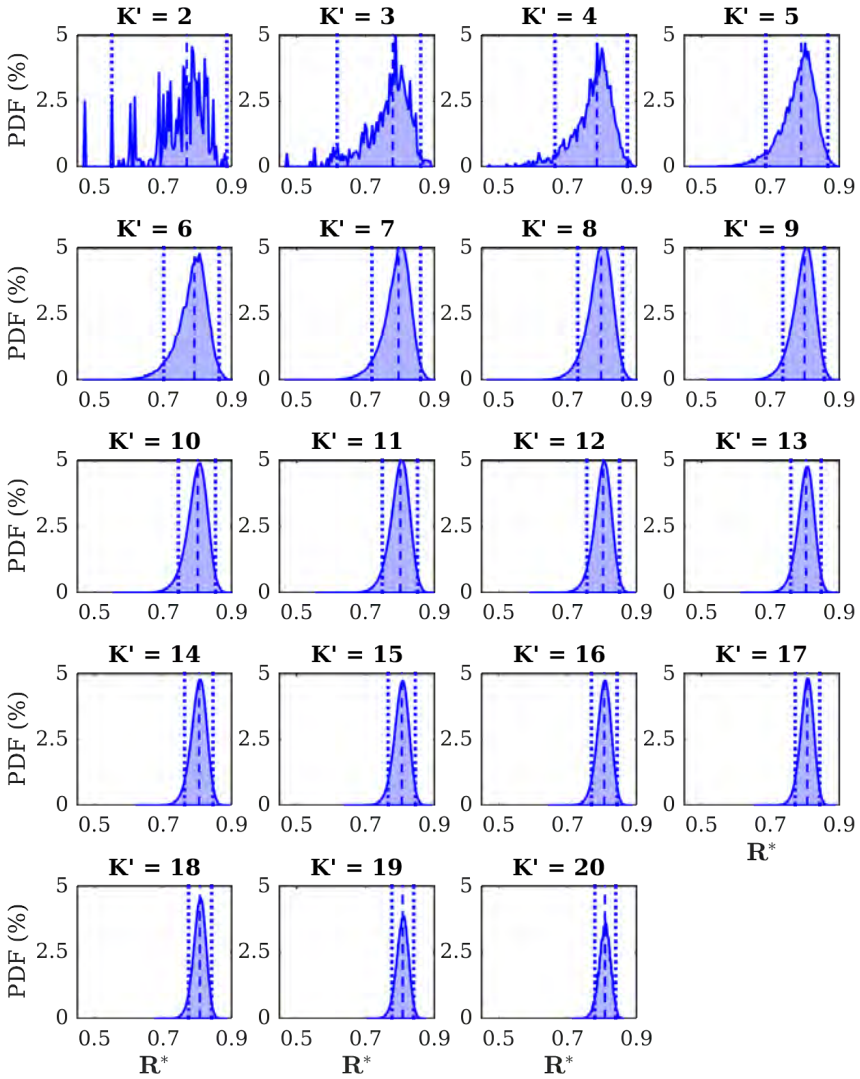


Figure F.1 – Bootstrap distribution of the correlation between SMOS and the *in situ* NAvg, as a function of the number of *in situ* time series taken to compute the NAvg. The dashed line correspond to the median and the dotted lines to the 5 % and 95 % confidence limits

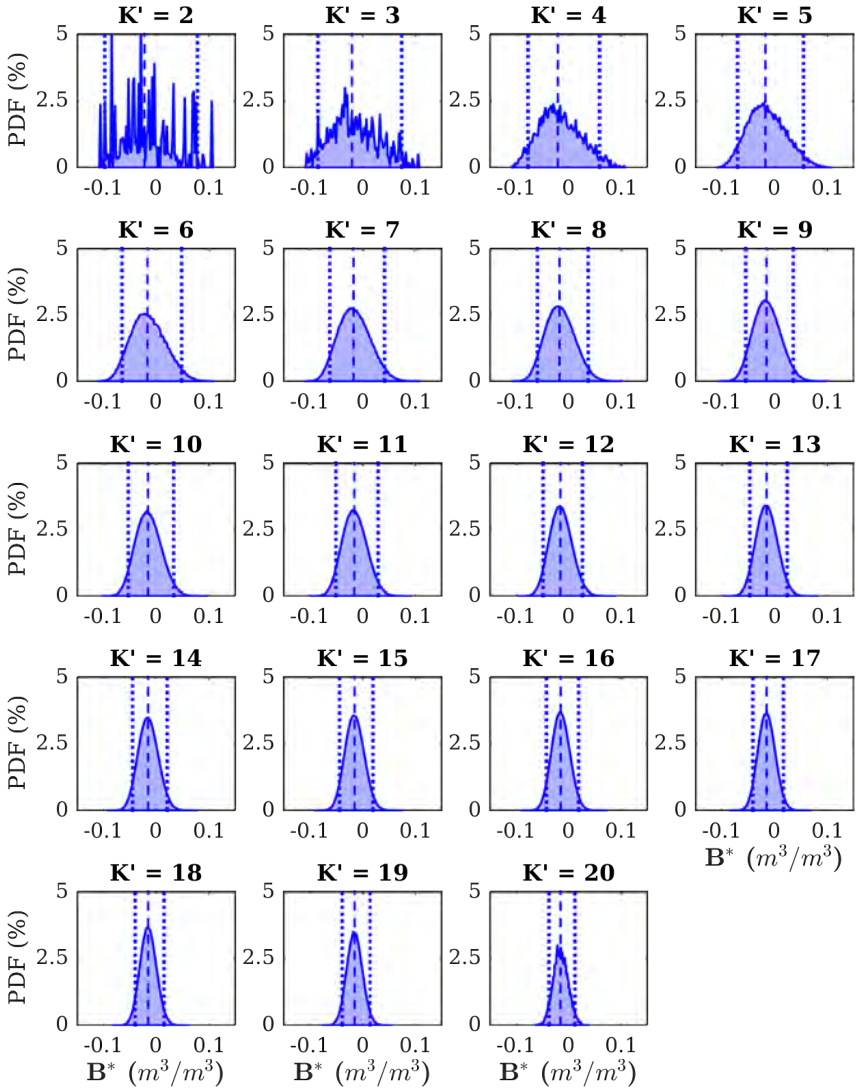


Figure F.2 – Bootstrap distribution of the bias between SMOS and the *in situ* NAvg, as a function of the number of *in situ* time series taken to compute the NAvg. The dashed line correspond to the median and the dotted lines to the 5 % and 95 % confidence limits

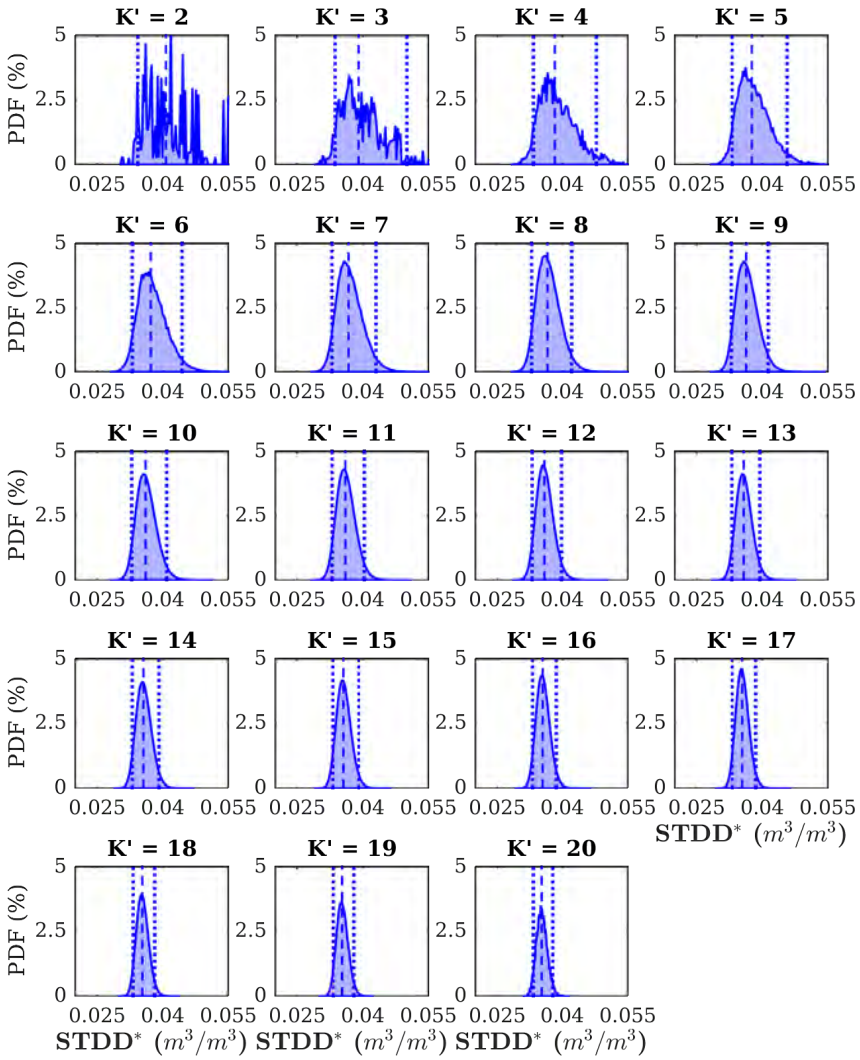


Figure F.3 – Bootstrap distribution of the STDD between SMOS and the *in situ* NAvg, as a function of the number of *in situ* time series taken to compute the NAvg. The dashed line correspond to the median and the dotted lines to the 5 % and 95 % confidence limits

Bibliography

Al Bitar, Ahmad, Delphine J. Leroux, Yann H. Kerr, Olivier Merlin, Philippe Richaume, Alok K. Sahoo, et al. (2012). "Evaluation of SMOS Soil Moisture Products Over Continental U.S. Using the SCAN / SNOTEL Network". In: *IEEE Transactions on Geoscience and Remote Sensing* 50.5, pp. 1572–1586.

Al Bitar, Ahmad, Yann H. Kerr, Olivier Merlin, François Cabot, and Jean-Pierre Wigneron (2013). "Root zone soil moisture and drought index from SMOS". In: *Satellite Soil Moisture Validation and Application Workshop*. Frascati, Italy.

Albergel, Clément, Patricia de Rosnay, Claire Gruhier, Joaquin Muñoz-Sabater, Stefan Hasenauer, Lars Isaksen, et al. (2012). "Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations". In: *Remote Sensing of Environment* 118, pp. 215–226. ISSN: 00344257. DOI: 10.1016/j.rse.2011.11.017.

Allen, Paul B. and James W. Naney (1991). *Hydrology of the Little Washita River Watershed, Oklahoma: Data and Analyses*. United States Department of Agriculture, Agricultural Research Service, ARS-90.

Barford, Lee A., R. Shane Fazzio, and David R. Smith (1992). *An introduction to wavelets*. HPL-92-124. Hewlett-Packard Company. ISBN: 0121745848. DOI: 10.1109/MMW.2005.1491268.

Berger, Richard A and Arthur B Sweney (1965). "Variance and Correlation Coefficients". In: *Research Quarterly. American Association for Health, Physical Education and Recreation* 36.3, pp. 368–370. ISSN: 10671188. DOI: 10.1080/10671188.1965.10614705.

Bindlish, Rajat, Thomas J. Jackson, Michael H. Cosh, Tianjie Zhao, and Peggy E. O'Neill (2015). "Global Soil Moisture From the Aquarius/SAC-D Satellite: Description and Initial Assessment". In: *IEEE Geoscience and Remote Sensing Letters* 12.5, pp. 923–927. DOI: 10.1109/LGRS.2014.2364151.

Bircher, Simone, N Skou, K H Jensen, Jeffrey P. Walker, and L Rasmussen (2012). "A soil moisture and temperature network for SMOS validation in Western Denmark". In: *Hydrology and Earth System Sciences* 16.5, pp. 1445–1463. ISSN: 10275606. DOI: 10.5194/hess-16-1445-2012.

Bircher, Simone, Mie Andreasen, Johanna Vuollet, Juho Vehviläinen, Kimmo Rautiainen, François Jonard, et al. (2016). "Soil moisture sensor calibration for organic soil surface layers". In: *Geosci. Instrum. Method. Data Syst* 5, pp. 109–125. DOI: 10.5194/gi-5-109-2016.

Blinowska, K J and P J Durka (1997). "An introduction to wavelet analysis with applications to vegetation time series". In: *British journal of audiology* 31.6, pp. 449–459. ISSN: 0300-5364. DOI: 10.3109/03005364000000039.

Blöschl, Günter and Murugesu Sivapalan (1995). "Scale issues in hydrological modelling: A review". In: *Hydrological Processes* 9.September 1994, pp. 251–290. ISSN: 08856087. DOI: 10.1002/hyp.3360090305. arXiv: 95/030251-40 [0885-6087].

Bolten, John D., Iliana E. Mladenova, Wade T. Crow, and Curt Reynolds (2016). "Enhancing the USDA global crop assessment decision support system using SMAP L3 Soil Moisture data". In: *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, pp. 5241–5243. ISBN: 978-1-5090-3332-4. DOI: 10.1109/IGARSS.2016.7730365.

Brocca, Luca, R. Morbidelli, F. Melone, and Tommaso Moramarco (2007). "Soil moisture spatial variability in experimental areas of central Italy". In: *Journal of Hydrology* 333.2-4, pp. 356–373. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2006.09.004.

Brocca, Luca, T. Tullo, F. Melone, Tommaso Moramarco, and R. Morbidelli (2012). "Catchment scale soil moisture spatial–temporal variability". In: *Journal of Hydrology* 422-423, pp. 63–75. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2011.12.039.

Brocca, Luca, Luca Ciabatta, Christian Massari, Tommaso Moramarco, Sebastian Hahn, Stefan Hasenauer, et al. (2014). "Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data". In: *Journal of Geophysical Research: Atmospheres* 119.9, pp. 5128–5141. ISSN: 2169897X. DOI: 10.1002/2014JD021489.

Brocca, Luca, Wade T. Crow, Luca Ciabatta, Christian Massari, Patricia De Rosnay, Markus Enenkel, et al. (2017). "A Review of the Applications of ASCAT Soil Moisture Products". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10.5, pp. 2285–2306. ISSN: 21511535. DOI: 10.1109/JSTARS.2017.2651140.

Brouwer, C., A. Goffeau, and M. Heibloem (1985). *Irrigation Water Management: Training Manual No. 1 - Introduction to Irrigation.* Tech. rep. 1. Food and Agriculture Organization (FAO) of the United Nations, p. 152.

Budyko, M. I. (1961). "The heat balance of the Earth's surface". In: *Soviet Geography* 2.4, pp. 3–13.

Calvet, Jean-Christophe, Noureddine Fritz, Françis Froissard, David Suquia, Alain Petitpa, and Bruno Piguet (2007). "In situ soil moisture observations for the CAL/VAL of SMOS: The SMOSMANIA network". In: *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1196–1199. ISBN: 1424412129. DOI: 10.1109/IGARSS.2007.4423019.

Carlson, Toby N. (2007). "An Overview of the "Triangle Method" for Estimating Surface Evapotranspiration and Soil Moisture from Satellite Imagery". In: *Sensors* 7.8, pp. 1612–1629. ISSN: 1424-8220. DOI: 10.3390/s7081612.

Carlson, Toby N., Robert R. Gillies, and Eileen M. Perry (1994). *A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover.* DOI: 10.1080/02757259409532220.

Carpenter, J and J Bithell (2000). "Bootstrap intervals: when, which, what? A practical guide for medical statisticians". In: *Statistics in Medicine* 19.August 1999, pp. 1141–1164. ISSN: 02776715. DOI: 10.1002/(SICI)1097-0258(20000515)19:9<1141::AID-SIM479>3.0.CO;2-F.

Casagrande, Erik, Brigitte Mueller, Diego G. Miralles, Dara Entekhabi, and Annalisa Molini (2015). "Wavelet correlations to reveal multiscale coupling in geophysical systems". In: *Journal of Geophysical Research Atmospheres* 120.15, pp. 7555–7572. ISSN: 21698996. DOI: 10.1002/2015JD023265. arXiv: arXiv:1502.05255v1.

Cayan, Daniel R. and Konstantine P. Georgakakos (1995). "Hydroclimatology of Continental Watersheds: 2. Spatial Analyses". In: *Water Resources Research* 31.3, pp. 677–697. ISSN: 19447973. DOI: 10.1029/94WR02376.

Chaney, Nathaniel W., Joshua K. Roundy, Julio E. Herrera-Estrada, and Eric F. Wood (2014). "High-resolution modeling of the spatial heterogeneity of soil moisture: Applications in network design". In: *Water Resources Research*, pp. 1–20. DOI: 10.1002/2013WR014964.

Chauhan, N. S., S. Miller, and P. Ardanuy (2003). "Spaceborne soil moisture estimation at high resolution: a microwave-optical/IR synergistic approach". In: *International Journal of Remote Sensing* 24.22, pp. 4599–4622. ISSN: 0143-1161. DOI: 10.1080/0143116031000156837.

Chen, Fan, Wade T. Crow, A. Colliander, Michael H. Cosh, Thomas J. Jackson, Rajat Bindlish, et al. (2016). "Application of Triple Collocation in Ground-Based Validation of Soil Moisture Active/Passive (SMAP) Level 2 Data Products". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 1–14. ISSN: 21511535 19391404. DOI: 10.1109/JSTARS.2016.2569998.

Cho, Eunsang, Chun-Hsu Su, Dongryeol Ryu, Hyunglok Kim, and Minha Choi (2017). "Does AMSR2 produce better soil moisture retrievals than AMSR-E over Australia?" In: *Remote Sensing of Environment* 188, pp. 95–105. ISSN: 00344257. DOI: 10.1016/j.rse.2016.10.050.

Choi, Minha, Jennifer M. Jacobs, and Michael H. Cosh (2007). "Scaled spatial variability of soil moisture fields". In: *Geophysical Research Letters* 34.1, pp. 1–6. ISSN: 00948276. DOI: 10.1029/2006GL028247.

Colliander, A., Thomas J. Jackson, Rajat Bindlish, Steven K. Chan, Narendra Narayan Das, S.B. Kim, et al. (2017). "Validation of SMAP surface soil moisture products with core validation sites". In: *Remote Sensing of Environment* 191, pp. 215–231. ISSN: 00344257. DOI: 10.1016/j.rse.2017.01.021.

Coopersmith, E J, B S Minsker, and M Sivapalan (2014). "Using similarity of soil texture and hydroclimate to enhance soil moisture estimation". In: *Hydrology and Earth System Sciences* 18, pp. 3095–3107. DOI: 10.5194/hess-18-3095-2014.

Coopersmith, E. J., J. E. Bell, K. Benedict, J. Shriber, O. McCotter, and Michael H. Cosh (2017). "Relating coccidioidomycosis (valley fever) incidence to soil moisture conditions". In: *GeoHealth* 1.1, pp. 51–63. ISSN: 2471-1403. DOI: 10.1002/2016GH000033.

Coopersmith, Evan J., Michael H. Cosh, Jesse E. Bell, and Ryan Boyles (2016). "Using machine learning to produce near surface soil moisture estimates from deeper in situ records at U.S. Climate Reference Network (USCRN) locations: Analysis and applications to AMSR-E satellite validation". In: *Advances in Water Resources* 98, pp. 122–131. ISSN: 0309-1708. DOI: 10.1016/J.ADVWATRES.2016.10.007.

Cornish, Charles R., Christopher S. Bretherton, and Donald B. Percival (2006). "Maximal overlap wavelet statistical analysis with application to atmospheric turbulence". In: *Boundary-Layer Meteorology* 119.2, pp. 339–374. ISSN: 00068314. DOI: 10.1007/s10546-005-9011-y.

Cosh, Michael H., Thomas J. Jackson, Patrick J. Starks, and Gary Heathman (2006). "Temporal stability of surface soil moisture in the Little Washita River watershed and its applications in satellite soil moisture product validation". In: *Journal of Hydrology* 323.1-4, pp. 168–177. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2005.08.020.

Cosh, Michael H., Thomas J. Jackson, M.S. Moran, and Rajat Bindlish (2008). "Temporal persistence and stability of surface soil moisture in a semi-arid watershed". In: *Remote Sensing of Environment* 112.2, pp. 304–313. ISSN: 00344257. DOI: 10.1016/j.rse.2007.07.001.

Craigmile, Peter F. and Peter Guttorp (2011). "Space-time modelling of trends in temperature series". In: *Journal of Time Series Analysis* 32, pp. 378–395. ISSN: 01439782. DOI: 10.1111/j.1467-9892.2011.00733.x.

Cressie, Noel A. C. (1993). *Statistics for Spatial Data (revised edition)*, p. 928. ISBN: 978-0-471-00255-0. DOI: 10.2307/2533238.

Crosson, William L., Ashutosh S. Limaye, and C. A. Laymon (2010). "Impacts of spatial scaling errors on soil moisture retrieval accuracy at L-band". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 3.1, pp. 67–80. ISSN: 1939-1404. DOI: 10.1109/JSTARS.2010.2041636.

Crow, Wade T., Matthias Drusch, and Eric F. Wood (2001). "An observation system simulation experiment for the impact of land surface heterogeneity on AMSR-E soil moisture retrieval". In: *IEEE Transactions on Geoscience and Remote Sensing* 39.8, pp. 1622–1631. ISSN: 01962892. DOI: 10.1109/36.942540.

Crow, Wade T., Steven Tsz K. Chan, Dara Entekhabi, Paul R. Houser, Ann Y. Hsu, Thomas J. Jackson, et al. (2005). "An observing system simulation experiment for hydros radiometer-only soil moisture and freeze-thaw products". In: *IEEE Transactions on Geoscience and Remote Sensing* 43.6, pp. 1289–1303. DOI: 10.1109/TGRS.2005.845645.

Crow, Wade T., Diego G. Miralles, and Michael H. Cosh (2010). "A quasi-global evaluation system for satellite-based surface soil moisture retrievals". In: *IEEE Transactions on Geoscience and Remote Sensing* 48.6, pp. 2516–2527. ISSN: 01962892. DOI: 10.1109/TGRS.2010.2040481.

Crow, Wade T., Aaron A. Berg, Michael H. Cosh, Alexander Loew, Binayak P. Mohanty, Rocco Panciera, et al. (2012). "Upscaling sparse ground-based soil moisture observations for

the validation of coarse-resolution satellite soil moisture products". In: *Reviews of Geophysics* 50.2, pp. 1–20. DOI: 10.1029/2011RG000372.

Crow, Wade T., Fan Chen, Rolf H. Reichle, and Q. Liu (2017). "L band microwave remote sensing and land data assimilation improve the representation of prestorm soil moisture conditions for hydrologic forecasting". In: *Geophysical Research Letters* 44.11, pp. 5495–5503. ISSN: 00948276. DOI: 10.1002/2017GL073642.

Daly, Edoardo and Amilcare Porporato (2005). "A Review of Soil Moisture Dynamics: From Rainfall Infiltration to Ecosystem Response". In: *Environmental Engineering Science* 22.1, pp. 9–24. ISSN: 1092-8758. DOI: 10.1089/ees.2005.22.9.

Das, Narendra Narayan, Dara Entekhabi, and Eni G. Njoku (2011). "An algorithm for merging SMAP radiometer and radar data for high-resolution soil-moisture retrieval". In: *IEEE Transactions on Geoscience and Remote Sensing* 49.5, pp. 1504–1512. ISSN: 01962892. DOI: 10.1109/TGRS.2010.2089526.

Daubechies, Ingrid (1992). "6. Orthonormal Bases of Compactly Supported Wavelets". In: *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, pp. 167–213. DOI: 10.1137/1.9781611970104.ch6.

Davison, A.C. and D.V. Hinkley (1997). "Complex dependence". In: *Bootstrap methods and their application*. Cambridge: Cambridge University Press. Chap. 8. ISBN: 978-0521574716.

De Lannoy, Gabriëlle J M, Niko E. C. Verhoest, Paul R. Houser, Timothy J. Gish, and Marc Van Meirvenne (2006). "Spatial and temporal characteristics of soil moisture in an intensively monitored agricultural field (OPE3)". In: *Journal of Hydrology* 331.3-4, pp. 719–730. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2006.06.016.

De Rosnay, Patricia, Matthias Drusch, Drasko Vasiljevic, Gianpaolo Balsamo, Clément Albergel, and Lars Isaksen (2013). "A simplified extended kalman filter for the global operational soil moisture analysis at ECMWF". In: *Quarterly Journal of the Royal Meteorological Society* 139.674, pp. 1199–1213. ISSN: 00359009. DOI: 10.1002/qj.2023.

Delworth, Thomas L. and Syukuro Manabe (1988). *The Influence of Potential Evaporation on the Variabilities of Simulated Soil Wetness and Climate*. DOI: 10.1175/1520-0442(1988)001<0523:TIOPEO>2.0.CO;2.

Dharssi, I., K. J. Bovis, B. Macpherson, and C. P. Jones (2011). "Operational assimilation of ASCAT surface soil wetness at the Met Office". In: *Hydrology and Earth System Sciences* 15.8, pp. 2729–2746. ISSN: 10275606. DOI: 10.5194/hess-15-2729-2011.

DiCiccio, Thomas J. and Bradley Efron (1996). "Bootstrap confidence intervals". In: *Statistical Science* 11.3, pp. 189–228. ISSN: 01677152. DOI: 10.1214/ss/1032280214.

Diem, Jeremy E. (2013). "Comments on "Changes to the North Atlantic Subtropical High and Its Role in the Intensification of Summer Rainfall Variability in the Southeastern United States"". In: *Journal of Climate* 26.2, pp. 679–682. ISSN: 0894-8755. DOI: 10.1175/JCLI-D-11-00390.1.

Dorigo, Wouter A., Klaus Scipal, Robert M. Parinussa, Yi Y Liu, W. Wagner, Richard A. M. de Jeu, et al. (2010). "Error characterisation of global active and passive microwave soil moisture datasets". In: *Hydrology and Earth System Sciences*, pp. 2605–2616. DOI: 10.5194/hess-14-2605-2010.

Douville, H. (2004). "Relevance of soil moisture for seasonal atmospheric predictions: Is it an initial value problem?" In: *Climate Dynamics* 22.4, pp. 429–446. ISSN: 09307575. DOI: 10.1007/s00382-003-0386-5.

Draper, Clara S., Jeffrey P. Walker, Peter J. Steinle, Richard A. M. de Jeu, and Thomas R. H. Holmes (2009). "An evaluation of AMSR-E derived soil moisture over Australia". In: *Remote Sensing of Environment* 113.4, pp. 703–710. ISSN: 00344257. DOI: 10.1016/j.rse.2008.11.011.

Draper, Clara S., Rolf H. Reichle, Richard A. M. de Jeu, Vahid Naeimi, Robert M. Parinussa, and Wolfgang Wagner (2013). "Estimating root mean square errors in remotely sensed soil moisture over continental scale domains". In: *Remote Sensing of Environment* 137, pp. 288–298. ISSN: 0034-4257. DOI: 10.1016/j.rse.2013.06.013.

Ebisuzaki, Wesley (1997). "A method to estimate the statistical significance of a correlation when the data are serially correlated". In: *Journal of Climate* 10.9, pp. 2147–2153. ISSN: 08948755. DOI: 10.1175/1520-0442(1997)010<2147:AMTETS>2.0.CO;2.

Efron, B. (1979). "Bootstrap Methods: Another Look at the Jackknife". In: *The Annals of Statistics* 7.1, pp. 1–26. ISSN: 0090-5364. DOI: 10.1214/aos/1176344552. arXiv: arXiv:1306.3979v1.

Entin, Jared K., Alan Robock, Konstantin Y. Vinnikov, Steven E. Hollinger, Suxia Liu, and A. Namkhai (2000). "Temporal and spatial scales of observed soil moisture variations in the extratropics". In: *Journal of Geophysical Research* 105.D9, pp. 865–877. ISSN: 0148-0227. DOI: 10.1029/2000JD900051.

Famiglietti, James S., J. W. Rudnicki, and Matthew Rodell (1998). "Variability in surface moisture content along a hillslope transect: Rattlesnake Hill, Texas". In: *Journal of Hydrology* 210.1-4, pp. 259–281. ISSN: 00221694. DOI: 10.1016/S0022-1694(98)00187-5.

Famiglietti, James S., J. A. Devereaux, C. A. Laymon, T. Tsecgaye, Paul R. Houser, Thomas J. Jackson, et al. (1999). "Ground-based investigation of soil moisture variability within remote sensing footprints during the Southern Great Plains 1997 (SGP97) Hydrology Experiment". In: *Water Resources Research* 35.6, pp. 1839–1851.

Famiglietti, James S., Dongryeol Ryu, Aaron a. Berg, Matthew Rodell, and Thomas J. Jackson (2008). "Field observations of soil moisture variability across scales". In: *Water Resources Research* 44.1, pp. 1–16. ISSN: 00431397. DOI: 10.1029/2006WR005804.

Fang, Bin, Venkat Lakshmi, Rajat Bindlish, Thomas J. Jackson, Michael H. Cosh, and Jeffrey Basara (2013). "Passive Microwave Soil Moisture Downscaling Using Vegetation Index and Skin Surface Temperature". In: *Vadose Zone Journal* 12.4. ISSN: 1539-1663. DOI: 10.2136/vzj2013.05.0089.

Fortak, Heinz (1982). *Meteorologie*. Berlin: Dietrich Reimer.

Foufoula-Georgiou, Efi and Praveen Kumar (1994). *Wavelets in Geophysics, Volume 4*. San Diego, California, USA: Academic Press, p. 373. ISBN: 0122628500.

Gibon, François, Thierry Pellarin, A Alhassane, S Traoré, and C Baron (2017). "Crop yield monitoring in the Sahel using root zone soil moisture anomalies derived from SMOS soil moisture data assimilation". In: *EGU General Assembly Conference*. Vol. 19. EGU General Assembly Conference Abstracts. Vienna, Austria.

Goodrich, David C., Jean Marc Faurès, David A. Woolhiser, Leonard J. Lane, and Soroosh Sorooshian (1995). "Measurement and analysis of small-scale convective storm rainfall variability". In: *Journal of Hydrology* 173.1-4, pp. 283–308. ISSN: 00221694. DOI: 10.1016/0022-1694(95)02703-R.

Goodwin, Laura D. and Nancy L. Leech (2006). "Understanding Correlation: Factors That Affect the Size of r". In: *The Journal of Experimental Education* 74.3, pp. 249–266. ISSN: 0022-0973. DOI: 10.3200/JEXE.74.3.249-266.

Grinsted, a., J. C. Moore, and S. Jevrejeva (2004). "Application of the cross wavelet transform and wavelet coherence to geophysical time series". In: *Nonlinear Processes in Geophysics* 11.5/6, pp. 561–566. ISSN: 1607-7946. DOI: 10.5194/npg-11-561-2004. arXiv: hal-00302394,version1.

Grossmann, A. and J. Morlet (1984). "Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape". In: *SIAM Journal on Mathematical Analysis* 15.4, pp. 723–736. ISSN: 0036-1410. DOI: 10.1137/0515056.

Gruber, Alexander, Wouter A. Dorigo, Simon Zwieback, A. Xaver, and W. Wagner (2013a). "Characterizing Coarse-Scale Representativeness of in situ Soil Moisture Measurements from the International Soil Moisture Network". In: *Vadose Zone Journal* 12.2, pp. 1–16. DOI: 10.2136/vzj2012.0170.

Gruber, Alexander, Wolfgang Wagner, Alena Hegyiova, Felix Greifeneder, and Stefan Schlaffer (2013b). "Potential of Sentinel-1 for high-resolution soil moisture monitoring". In: *IEEE International Symposium on Geoscience and Remote Sensing (IGARSS)*, pp. 4030–4033. DOI: 10.1109/IGARSS.2013.6723717. arXiv: 13 [978-1-4799-1114-1].

Gruber, Alexander, Chun-Hsu Su, Simon Zwieback, Wade T. Crow, Wouter A. Dorigo, and W. Wagner (2016). "Recent advances in (soil moisture) triple collocation analysis". In: *International Journal of Applied Earth Observation and Geoinformation* 45, pp. 200–211.

Grünewald, T., M. Schirmer, R. Mott, and M. Lehning (2010). "Spatial and temporal variability of snow depth and SWE in a small mountain catchment". In: *The Cryosphere Discussions* 4.1, pp. 1–30. ISSN: 1994-0440. DOI: 10.5194/tcd-4-1-2010.

Haar, Alfred (1910). "Zur Theorie der orthogonalen Funktionensysteme". In: *Mathematische Annalen* 69.3, pp. 331–371. ISSN: 0025-5831. DOI: 10.1007/BF01456326.

Hasselmann, K. (1976). "Stochastic climate models: Part I. Theory". In: *Tellus A* 28.6, pp. 473–485. ISSN: 1600-0870. DOI: 10.3402/tellusa.v28i6.11316.

Hawley, Mark E., Thomas J. Jackson, and Richard H. McCuen (1983). "Surface soil moisture variation on small agricultural watersheds". In: *Journal of Hydrology* 62, pp. 179–200.

Helsel, D. R. and R. M. Hirsch (2002). *Statistical Methods in Water Resources Techniques of Water Resources Investigations, Book 4, Chapter A3*. U.S. Geological Survey (USGS), p. 522.

Holm, Sverre (2015). "Prudence in estimating coherence between planetary, solar and climate oscillations". In: *Astrophysics and Space Science* 357.2. ISSN: 1572946X. DOI: 10.1007/s10509-015-2334-3. arXiv: arXiv:1504.01611v1.

Ines, Amor V. M., Narendra N. Das, James W. Hansen, and Eni G. Njoku (2013). "Assimilation of remotely sensed soil moisture and vegetation with a crop simulation model for maize yield prediction". In: *Remote Sensing of Environment* 138, pp. 149–164. ISSN: 00344257. DOI: 10.1016/j.rse.2013.07.018.

Isaaks, E H and R M Srivastava (1989). *Applied Geostatistics*. Vol. 21. New York: Oxford University Press, p. 592. ISBN: 0195050134. DOI: 10.1016/0040-1951(74)90006-7. arXiv: arXiv:1011.1669v3.

Jackson, Thomas J., D.M. Le Vine, Ann Y. Hsu, Anna Oldak, Patrick J. Starks, C.T. Swift, et al. (1999). "Soil moisture mapping at regional scales using microwave radiometry: the Southern Great Plains Hydrology Experiment". In: *IEEE Transactions on Geoscience and Remote Sensing* 37.5, pp. 2136–2151. ISSN: 01962892. DOI: 10.1109/36.789610.

Jackson, Thomas J., Michael H. Cosh, Rajat Bindlish, Patrick J. Starks, David D. Bosch, Mark S. Seyfried, et al. (2010). "Validation of advanced microwave scanning radiometer soil moisture products". In: *IEEE Transactions on Geoscience and Remote Sensing* 48.12, pp. 4256–4272. ISSN: 01962892. DOI: 10.1109/TGRS.2010.2051035.

Jackson, Thomas J., Rajat Bindlish, Michael H. Cosh, Tianjie Zhao, Patrick J. Starks, David D. Bosch, et al. (2012). "Validation of soil moisture and Ocean Salinity (SMOS) soil moisture over watershed networks in the U.S." In: *IEEE Transactions on Geoscience and Remote Sensing* 50.5, pp. 1530–1543. ISSN: 01962892. DOI: 10.1109/TGRS.2011.2168533.

Jacobs, Jennifer M., Binayak P. Mohanty, En-Ching Hsu, and Douglas Miller (2004). "SMEX02: Field scale variability, time stability and similarity of soil moisture". In: *Remote Sensing of Environment* 92.4, pp. 436–446. DOI: 10.1016/j.rse.2004.02.017.

JAXA (2013). *GCOM-W1 "SHIZUKU" Data Users Handbook*. Tech. rep. January.

Kerr, Yann H., Philippe Waldteufel, Jean-Pierre Wigneron, Jean Michel Martinuzzi, Jordi Font, and Michael Berger (2001). "Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission". In: *IEEE Transactions on Geoscience and Remote Sensing* 39.8, pp. 1729–1735. ISSN: 01962892. DOI: 10.1109/36.942551.

Kerr, Yann H., Philippe Waldteufel, Jean-Pierre Wigneron, Steven Delwart, François Cabot, Jacqueline Boutin, et al. (2010). "The SMOS Mission: New Tool for Monitoring Key Elements of the Global Water Cycle". In: *Proceedings of the IEEE* 98.5, pp. 666–687. ISSN: 0018-9219. DOI: 10.1109/JPROC.2010.2043032.

Kerr, Yann H., Philippe Waldteufel, Philippe Richaume, Jean-Pierre Wigneron, Paolo Ferrazzoli, Ali Mahmoodi, et al. (2012). "The SMOS Soil Moisture Retrieval Algorithm". In: *Geoscience and Remote Sensing* 50.5, pp. 1384–1403.

Kerr, Yann H., Elsa Jacquette, Ahmad Al Bitar, François Cabot, Arnaud Mialon, and Philippe Richaume (2013a). *CATDS L3 Soil Moisture Retrieval Processor, Algorithm*

Theoretical Baseline Document (ATBD). Tech. rep. 2.0. ACRI-ST, CESBIO, CNES, LATT and SA.

Kerr, Yann H., Lucie Berthon, Arnaud Mialon, François Cabot, Ahmad Al Bitar, Philippe Richaume, et al. (2013b). *CATDS Level 3 - Data product description - Soil Moisture and Brightness Temperature*. Tech. rep. CESBIO.

Kerr, Yann H., Philippe Waldteufel, Philippe Richaume, Paolo Ferrazzoli, and Jean-Pierre Wigneron (2014). *SMOS Level-2 Processor Soil Moisture Algorithm Theoretical Basis Document (ATBD)*. V4.a. Toulouse: SM-ESL (CBSA).

Kerr, Yann H., Amen Al-Yaari, Nemesio J. Rodríguez-Fernández, Marie Parrens, Beatriz Molero, Delphine J. Leroux, et al. (2016). "Overview of SMOS performance in terms of global soil moisture monitoring after six years in operation". In: *Remote Sensing of Environment* 180, pp. 40–63. ISSN: 00344257. DOI: 10.1016/j.rse.2016.02.042.

Kolassa, Jana, R.H. Reichle, and Clara S. Draper (2017). "Merging active and passive microwave observations in soil moisture data assimilation". In: *Remote Sensing of Environment* 191, pp. 117–130. ISSN: 00344257. DOI: 10.1016/j.rse.2017.01.015.

Kornelsen, Kurt C. and Paulin Coulibaly (2013). "McMaster Mesonet soil moisture dataset: Description and spatio-temporal variability analysis". In: *Hydrology and Earth System Sciences* 17.4, pp. 1589–1606. ISSN: 10275606. DOI: 10.5194/hess-17-1589-2013.

Korres, W., T. G. Reichenau, P. Fiener, C. N. Koyama, Heye R. Bogena, T. Cornelissen, et al. (2015). "Spatio-temporal soil moisture patterns - A meta-analysis using plot to catchment scale data". In: *Journal of Hydrology* 520, pp. 326–341. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2014.11.042.

Koster, Randal D, P. A. Dirmeyer, Zhichang Guo, Gordon Bonan, Edmond Chan, Peter Cox, et al. (2004). "Regions of Strong Coupling Between Soil Moisture and Precipitation". In: *Science* 305.2004, pp. 1138–1140. DOI: 10.1126/science.1100217.

Koster, Randal D., Sarith P. P. Mahanama, Ben Livneh, Dennis P. Lettenmaier, and Rolf H. Reichle (2010). "Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow". In: *Nature Geoscience* 3.9, pp. 613–616. ISSN: 1752-0894. DOI: 10.1038/ngeo944.

Laiolo, P., S. Gabellani, L. Campo, F. Silvestro, F. Delogu, R. Rudari, et al. (2016). "Impact of different satellite soil moisture products on the predictions of a continuous distributed hydrological model". In: *International Journal of Applied Earth Observation and Geoinformation* 48, pp. 131–145. ISSN: 03032434. DOI: 10.1016/j.jag.2015.06.002.

Legates, D. R., R. Mahmood, D. F. Levia, T. L. DeLiberty, S. M. Quiring, C. Houser, et al. (2011). "Soil moisture: A central and unifying theme in physical geography". In: *Progress in Physical Geography* 35.1, pp. 65–86. ISSN: 0309-1333. DOI: 10.1177/0309133310386514.

Lei, Fangni, Wade T. Crow, Huanfeng Shen, Robert M. Parinussa, and Thomas R. H. Holmes (2015). "The impact of local acquisition time on the accuracy of microwave surface soil moisture retrievals over the contiguous United States". In: *Remote Sensing* 7.10, pp. 13448–13465. ISSN: 20724292. DOI: 10.3390/rs71013448.

Leroux, Delphine J., Yann H. Kerr, Ahmad Al Bitar, Rajat Bindlish, Thomas J. Jackson, Beatrice Berthelot, et al. (2013a). "Comparison Between SMOS, VUA, ASCAT, and ECMWF Soil Moisture Products Over Four Watersheds in U.S." In: *IEEE Transactions on Geoscience and Remote Sensing* 52.3, pp. 1–10.

Leroux, Delphine J., Yann H. Kerr, Philippe Richaume, and Remy Fieuzal (2013b). "Spatial distribution and possible sources of SMOS errors at the global scale". In: *Remote Sensing of Environment* 133, pp. 240–250. ISSN: 00344257. DOI: 10.1016/j.rse.2013.02.017.

Leroux, Delphine J., Thierry Pellarin, Théo Vischel, Jean Martial Cohard, Tania Gascon, François Gibon, et al. (2016). "Assimilation of SMOS soil moisture into a distributed hydrological model and impacts on the water cycle variables over the Ouémé catchment in Benin". In: *Hydrology and Earth System Sciences* 20.7, pp. 2827–2840. ISSN: 16077938. DOI: 10.5194/hess-20-2827-2016.

Li, Wenhong, Laifang Li, Rong Fu, Yi Deng, and Hui Wang (2011). "Changes to the North Atlantic Subtropical High and Its Role in the Intensification of Summer Rainfall Variability in the Southeastern United States". In: *Journal of Climate* 24.5, pp. 1499–1506. ISSN: 0894-8755. DOI: 10.1175/2010JCLI3829.1.

Lievens, H., Sat Kumar Tomer, Ahmad Al Bitar, Gabriëlle J M De Lannoy, Matthias Drusch, G. Dumedah, et al. (2015). "SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia". In: *Remote Sensing of Environment* 168, pp. 146–162. ISSN: 00344257. DOI: 10.1016/j.rse.2015.06.025.

Lorenz, Ruth, Eric B. Jaeger, and Sonia I. Seneviratne (2010). "Persistence of heat waves and its link to soil moisture memory". In: *Geophysical Research Letters* 37.9. ISSN: 00948276. DOI: 10.1029/2010GL042764.

Malbéteau, Yoann, Olivier Merlin, Beatriz Molero, Christoph Rüdiger, and Stephan Bacon (2016). "DisPATCh as a tool to evaluate coarse-scale remotely sensed soil moisture using localized in situ measurements: Application to SMOS and AMSR-E data in Southeastern Australia". In: *International Journal of Applied Earth Observation and Geoinformation* 45, Part B, pp. 221–234. ISSN: 0303-2434. DOI: 10.1016/j.jag.2015.10.002. arXiv: arXiv:1011.1669v3.

Malbéteau, Yoann, Olivier Merlin, Simon Gascoin, Jean Philippe Gastellu, L. Olivera, Cristian Mattar, et al. (2017). "Correcting land surface temperature data for elevation and illumination effects in mountainous areas: a case study using ASTER data over the Imlil valley, Morocco". In: *Remote Sensing of Environment* 189, pp. 25–39. DOI: 10.1016/j.rse.2016.11.010.

Manolakis, Dimitris G., Vinay K. Ingle, and Stephen M. Kogon (2005). *Statistical and Adaptive Signal Processing*. Norwood: Artech House. ISBN: 9780874216561. DOI: 10.1007/s13398-014-0173-7.2. arXiv: arXiv:1011.1669v3.

Martínez-Fernández, José and A. Ceballos (2005). "Mean soil moisture estimation using temporal stability analysis". In: *Journal of Hydrology* 312.1-4, pp. 28–38. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2005.02.007.

Martínez-Fernández, José, A. González-Zamora, Nilda Sánchez, A. Gumuzzio, and C.M. Herrero-Jiménez (2016). "Satellite soil moisture for agricultural drought monitoring: Assessment of the SMOS derived Soil Water Deficit Index". In: *Remote Sensing of Environment* 177, pp. 277–286. ISSN: 00344257. DOI: 10.1016/j.rse.2016.02.064.

Massari, Christian, Luca Brocca, Tommaso Moramarco, Yves Tramblay, and Jean Francois Didon Lescot (2014). "Potential of soil moisture observations in flood modelling: Estimating initial conditions and correcting rainfall". In: *Advances in Water Resources* 74, pp. 44–53. ISSN: 03091708. DOI: 10.1016/j.advwatres.2014.08.004.

Masson, Valéry, Jean Louis Champeaux, Fabrice Chauvin, Christelle Meriguet, and Roselyne Lacaze (2003). "A global database of land surface parameters at 1-km resolution in meteorological and climate models". In: *Journal of Climate* 16.9, pp. 1261–1282. ISSN: 08948755. DOI: 10.1175/1520-0442(2003)16<1261:AGDOLS>2.0.CO;2.

McColl, Kaighin a., Jur Vogelzang, Alexandra G. Konings, Dara Entekhabi, María Piles, and Ad Stoffelen (2014). "Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target". In: *Geophysical Research Letters* 41.17, pp. 6229–6236. ISSN: 19448007. DOI: 10.1002/2014GL061322.

Merlin, Olivier, Jeffrey P. Walker, Jetse D. Kalma, Edward J. Kim, Jorg Hacker, Rocco Panciera, et al. (2008a). "The NAFE'06 data set: Towards soil moisture retrieval at intermediate resolution". In: *Advances in Water Resources* 31.11, pp. 1444–1455. ISSN: 03091708. DOI: 10.1016/j.advwatres.2008.01.018.

Merlin, Olivier, Jeffrey P. Walker, Abdelghani Chehbouni, and Yann H. Kerr (2008b). "Towards deterministic downscaling of SMOS soil moisture using MODIS derived soil evaporative efficiency". In: *Remote Sensing of Environment* 112, pp. 3935–3946. DOI: 10.1016/j.rse.2008.06.012.

Merlin, Olivier, Christoph Rüdiger, Philippe Richaume, Ahmad Al Bitar, Arnaud Mialon, Jeffrey P. Walker, et al. (2010). "Disaggregation as a top-down approach for evaluating 40 km resolution SMOS data using point-scale measurements: application to AACES-1". In: *Proceedings of SPIE*. Vol. 7824. DOI: 10.1117/12.865751.

Merlin, Olivier, Christoph Rüdiger, Ahmad Al Bitar, Philippe Richaume, Jeffrey P. Walker, and Yann H. Kerr (2012). "Disaggregation of SMOS Soil Moisture in Southeastern Australia". In: *IEEE Transactions on Geoscience and Remote Sensing* 50.5, pp. 1556–1571. ISSN: 0196-2892. DOI: 10.1109/TGRS.2011.2175000.

Merlin, Olivier, Maria José Escorihuela, Miquel Aran Mayoral, Olivier Hagolle, Ahmad Al Bitar, and Yann H. Kerr (2013). "Self-calibrated evaporation-based disaggregation of SMOS soil moisture: An evaluation study at 3 km and 100 m resolution in Catalunya, Spain". In: *Remote Sensing of Environment* 130.2013, pp. 25–38.

Merlin, Olivier, Yoann Malbéteau, Y. Notfi, Stephan Bacon, Salah Er-Raki, Said Khabba, et al. (2015). "Performance metrics for soil moisture downscaling methods: Application to DISPATCH data in central Morocco". In: *Remote Sensing* 7.4, pp. 3783–3807. DOI: 10.3390/rs70403783.

Merlin, Olivier (2016). "Suivi des ressources en eau par télédétection multi-capteur: désagrégation de données spatiales et modélisation descendante des processus hydrologiques (HDR)". HDR. Toulouse: Université Paul Sabatier, Toulouse III.

Merlin, Olivier, Vivien Georgiana Stefan, A Amazirh, A Chanzy, E Ceschia, Salah Er-Raki, et al. (2016). "Modeling soil evaporation efficiency in a range of soil and atmospheric conditions using a meta-analysis approach". In: *Water Resources Research* 52.5, pp. 3663–3684. DOI: 10.1002/2015WR018233.

Miralles, Diego G., Wade T. Crow, and Michael H. Cosh (2010). "Estimating Spatial Sampling Errors in Coarse-Scale Soil Moisture Estimates Derived from Point-Scale Observations". In: *Journal of Hydrometeorology* 11, pp. 1423–1429. ISSN: 1525-755X. DOI: 10.1175/2010JHM1285.1.

Mironov, Valery L. and Sergey V. Fomin (2009). "Temperature and Mineralogy Dependable Model for Microwave Dielectric Spectra of Moist Soils". In: *Progress In Electromagnetics Research Symposium*. Beijing. DOI: 10.2529/PIERS090220054025.

Mironov, Valery L., Lyudmila G. Kosolapova, and Sergej V. Fomin (2009). "Physically and mineralogically based spectroscopic dielectric model for moist soils". In: *IEEE Transactions on Geoscience and Remote Sensing* 47.7, pp. 2059–2070. ISSN: 01962892. DOI: 10.1109/TGRS.2008.2011631.

Mishra, Ashok K. and Vijay P. Singh (2010). "A review of drought concepts". In: *Journal of Hydrology* 391.1-2, pp. 202–216. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2010.07.012.

Mittelbach, Heidi and Sonia I. Seneviratne (2012). "A new perspective on the spatio-temporal variability of soil moisture: Temporal dynamics versus time-invariant contributions". In: *Hydrology and Earth System Sciences* 16.7, pp. 2169–2179. ISSN: 10275606. DOI: 10.5194/hess-16-2169-2012.

Mladenova, Iliana E., Venkat Lakshmi, Thomas J. Jackson, Jeffrey P. Walker, Olivier Merlin, and Richard A. M. de Jeu (2011). "Validation of AMSR-E soil moisture using L-band airborne radiometer data from National Airborne Field Experiment 2006". In: *Remote Sensing of Environment* 115.8, pp. 2096–2103. ISSN: 00344257. DOI: 10.1016/j.rse.2011.04.011.

Mohanty, Binayak P. and T.H Skaggs (2001). "Spatio-temporal evolution and time-stable characteristics of soil moisture within remote sensing footprints with varying soil, slope, and vegetation". In: *Advances in Water Resources* 24.9-10, pp. 1051–1067. ISSN: 03091708. DOI: 10.1016/S0309-1708(01)00034-3.

Molero, Beatriz, Olivier Merlin, Yoann Malbéteau, Ahmad Al Bitar, François Cabot, Vivien Georgiana Stefan, et al. (2016). "SMOS disaggregated soil moisture product at 1 km resolution: Processor overview and first validation results". In: *Remote Sensing of Environment* 180. ISSN: 00344257. DOI: 10.1016/j.rse.2016.02.045.

Molero, Beatriz, D. J. Leroux, P. Richaume, Y. H. Kerr, O. Merlin, M. H. Cosh, et al. (2018). "Multi-Timescale Analysis of the Spatial Representativeness of In Situ Soil Moisture Data within Satellite Footprints". In: *Journal of Geophysical Research: Atmospheres*, pp. 1–19. ISSN: 2169897X. DOI: 10.1002/2017JD027478.

Molini, Annalisa, Gabriel G Katul, and Amilcare Porporato (2010). "Causality across rainfall time scales revealed by continuous wavelet transforms". In: *Journal of Geophysical Research* 115, pp. 1–16. DOI: 10.1029/2009JD013016.

Mood, A M, F A Graybill, and D C Boes (1974). *Introduction to the Theory of Statistics*. Vol. 3, p. 480. ISBN: 0070854653.

Moors, J. J. A. and J. Muilwijk (1971). "An Inequality for the Variance of a Discrete Random Variable". In: *The Indian Journal of Statistics, Series B (1960-2002)* 33.3/4, pp. 385–388. ISSN: 0028-0836. DOI: 10.1038/157869b0.

Moran, M.S., T.R. Clarke, Y. Inoue, and A. Vidal (1994). *Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index*. DOI: 10.1016/0034-4257(94)90020-5.

Mudelsee, Manfred (2003). "Estimating Pearson's Correlation Coefficient With Bootstrap Confidence Interval From Serially Dependent Time Series". In: *Mathematical Geology* 35.6, pp. 651–665.

Muilwijk, J. (1966). "Note on a Theorem of M. N. Murthy and V. K. Sethi". In: *The Indian Journal of Statistics, Series B (1960-2002)* 28.1/2, p. 183. ISSN: 0028-0836. DOI: 10.1038/157869b0.

Narayan, Ujjwal, Venkataraman Lakshmi, and Thomas J. Jackson (2006). "High-resolution change estimation of soil moisture using L-band radiometer and radar observations made during the SMEX02 experiments". In: *IEEE Transactions on Geoscience and Remote Sensing* 44.6, pp. 1545–1554. ISSN: 01962892. DOI: 10.1109/TGRS.2006.871199.

Nicolai-Shaw, Nadine, Martin Hirschi, Heidi Mittelbach, and Sonia I. Seneviratne (2015). "Spatial representativeness of soil moisture using in situ, remote sensing, and land reanalysis data". In: *Journal Geophysical Research: Atmospheres* 120.19, pp. 9955 –9964. DOI: 10.1002/2015JD023305.

Nicolai-Shaw, Nadine, Lukas Gudmundsson, Martin Hirschi, and Sonia I. Seneviratne (2016). "Long-term predictability of soil moisture dynamics at the global scale: Persistence versus large-scale drivers". In: *Geophysical Research Letters*. ISSN: 19448007. DOI: 10.1002/2016GL069847.

Nishida, Kenlo, Ramakrishna R. Nemani, Joseph M. Glassy, and Steven W. Running (2003). "Development of an evapotranspiration index from Aqua/MODIS for monitoring surface moisture status". In: *IEEE Transactions on Geoscience and Remote Sensing* 41.2, pp. 493–500. ISSN: 01962892. DOI: 10.1109/TGRS.2003.811744.

Nykanen, Deborah K. and Efi Foufoula-Georgiou (2001). "Soil moisture variability and scale-dependency of nonlinear parameterizations in coupled land-atmosphere models". In: *Advances in Water Resources* 24.9-10, pp. 1143–1157. ISSN: 03091708. DOI: 10.1016/S0309-1708(01)00046-X.

Oki, Taikan and S. Kanae (2006). "Global Hydrological Cycles and World Water Resources". In: *Science* 313.5790, pp. 1068–1072. ISSN: 0036-8075. DOI: 10.1126/science.1128845.

Oliva, Roger, Elena Daganzo-Eusebio, Yann H. Kerr, Susanne Mecklenburg, Sara Nieto, Philippe Richaume, et al. (2012). "SMOS radio frequency interference scenario: Status and actions taken to improve the RFI environment in the 1400-1427-MHZ passive band". In: *IEEE Transactions on Geoscience and Remote Sensing* 50.5, pp. 1427–1439. ISSN: 01962892. DOI: 10.1109/TGRS.2012.2182775.

Orlowsky, Boris and Sonia I. Seneviratne (2014). "On the spatial representativeness of temporal dynamics at European weather stations". In: *International Journal of Climatology* 34.10, pp. 3154–3160. ISSN: 08998418. DOI: 10.1002/joc.3903.

Owe, Manfred, Richard A. M. de Jeu, and Thomas R. H. Holmes (2008). "Multisensor historical climatology of satellite-derived global land surface moisture". In: *Journal of*

Geophysical Research: Earth Surface 113.1, pp. 1–17. ISSN: 21699011. DOI: 10.1029/2007JF000769.

Pan, Feifei (2012). "Estimating Daily Surface Soil Moisture Using a Daily Diagnostic Soil Moisture Equation". In: *Journal of Irrigation and Drainage Engineering* 138.7, pp. 625–631. ISSN: 0733-9437. DOI: 10.1061/(ASCE)IR.1943-4774.0000450.

Pan, Feifei, Christa D. Peters-Lidard, and Michael J. Sale (2003). "An analytical method for predicting surface soil moisture from rainfall observations". In: *Water Resources Research* 39.11. ISSN: 00431397. DOI: 10.1029/2003WR002142.

Panciera, Rocco, Jeffrey P. Walker, Thomas J. Jackson, Douglas a. Gray, Mihai a. Tanase, Dongryeol Ryu, et al. (2014). "The soil moisture active passive experiments (SMAPEx): Toward soil moisture retrieval from the SMAP mission". In: *IEEE Transactions on Geoscience and Remote Sensing* 52.1, pp. 490–507. ISSN: 01962892. DOI: 10.1109/TGRS.2013.2241774.

Parinussa, Robert M., Thomas R. H. Holmes, and Richard A. M. de Jeu (2012). "Soil moisture retrievals from the WindSat spaceborne polarimetric microwave radiometer". In: *IEEE Transactions on Geoscience and Remote Sensing* 50.7, pp. 2683–2694. ISSN: 01962892. DOI: 10.1109/TGRS.2011.2174643.

Peischl, S, Jeffrey P. Walker, Christoph Rüdiger, N Ye, Yann H. Kerr, Edward J. Kim, et al. (2012). "The AACES field experiments: SMOS calibration and validation across the Murrumbidgee River catchment". In: *Hydrology and Earth System Sciences* 16.6, pp. 1697–1708. ISSN: 1607-7938. DOI: 10.5194/hess-16-1697-2012.

Pellarin, Thierry, Jean-Pierre Wigneron, Jean-Christophe Calvet, and Philippe Waldteufel (2003). "Global soil moisture retrieval from a synthetic L-band brightness temperature data set". In: *Journal of Geophysical Research* 108.D12. ISSN: 0148-0227. DOI: 10.1029/2002JD003086.

Pellarin, Thierry, Luca Brocca, Wade T. Crow, Yann H. Kerr, C Massari, C Román-Cascón, et al. (2017). "SMOS+RAINFALL: Evaluating the ability of different methodologies to improve rainfall estimations using soil moisture data from SMOS". In: *EGU General Assembly Conference Abstracts*. Vol. 19. EGU General Assembly Conference Abstracts, p. 18932.

Percival, Donald B. and Andrew T Walden (2000). *Wavelet methods for time series analysis*. Cambridge University Press. ISBN: 0-521-64068-7.

Percival, Donald B., S Sardy, and A.C. Davison (2000). "Wavestrapping time series: Adaptive wavelet-based bootstrapping". In: *Nonlinear and Nonstationary Signal Processing*. Cambridge University Press, pp. 442–471.

Petropoulos, George P., Gareth Ireland, and Brian Barrett (2015). "Surface soil moisture retrievals from remote sensing: Current status, products & future trends". In: *Physics and Chemistry of the Earth* 83-84, pp. 36–56. ISSN: 14747065. DOI: 10.1016/j.pce.2015.02.009.

Piles, María, Adriano Camps, Mercè Vall-Llossera, Ignasi Corbella, Rocco Panciera, Christoph Rüdiger, et al. (2011). "Downscaling SMOS-derived soil moisture using MODIS visible/infrared data". In: *IEEE Transactions on Geoscience and Remote Sensing* 49.9, pp. 3156–3166. ISSN: 01962892. DOI: 10.1109/TGRS.2011.2120615.

Qiu, Yang, Bojie Fu, Jun Wang, and Liding Chen (2001). "Soil moisture variation in relation to topography and land use in a hillslope catchment of the Loess Plateau, China". In: *Journal of Hydrology* 240.3-4, pp. 243–263. ISSN: 00221694. DOI: 10.1016/S0022-1694(00)00362-0.

Richaume, Philippe, Yann H. Kerr, Philippe Waldteufel, S. Dai, Ali Mahmoodi, and A. Hahne (2006). *SMOS L2 Processor Discrete Flexible Fine Grid Definition (SO-TN-CBSA-GS-0011)*. Tech. rep. 1.c, p. 42.

Robinson, D. A., C. S. Campbell, J. W. Hopmans, Brian K. Hornbuckle, S. B. Jones, R. Knight, et al. (2008). "Soil Moisture Measurement for Ecological and Hydrological Watershed-Scale Observatories: A Review". In: *Vadose Zone Journal* 7.1, p. 358. ISSN: 1539-1663. DOI: 10.2136/vzj2007.0143.

Rodriguez, R. N. (2006). "Correlation". In: *Encyclopedia of statistical sciences, Vol. 2*. Ed. by S. Kotz and N. L. Johnson. 2nd Ed. New York: John Wiley & Sons, Inc., pp. 193–204. ISBN: 0-471-15044-4. DOI: 10.1002/0471667196.ess0423.pub2.

Rodríguez-Fernández, Nemesio J., P De Rosnay, Clément Albergel, Filipe Aires, Catherine Prigent, Philippe Richaume, et al. (2017a). *SMOS Neural Network Soil Moisture Data Assimilation, ESA contract report*. Tech. rep. European Centre for Medium-Range Weather Forecasts (ECMWF), p. 35.

Rodríguez-Fernández, Nemesio J., Veronica de Souza, Yann H. Kerr, Philippe Richaume, and Ahmad Al Bitar (2017b). "Soil moisture retrieval using SMOS brightness temperatures and a neural network trained on in situ measurements". In: *IGARSS*. Fort Worth, Texas. ISBN: 9781509049516.

Rodriguez-Iturbe, Ignacio (2000). "Ecohydrology: A hydrologic perspective of climate-soil-vegetation dynamics". In: *Water Resources Research* 36.1, pp. 3–9. ISSN: 00431397. DOI: 10.1029/1999WR900210.

Rosenbaum, U., Heye R. Bogena, M. Herbst, J. a. Huisman, T. J. Peterson, a. Weuthen, et al. (2012). "Seasonal and event dynamics of spatial soil moisture patterns at the small catchment scale". In: *Water Resources Research* 48.10, pp. 1–22. ISSN: 00431397. DOI: 10.1029/2011WR011518.

Rüdiger, Christoph, Jean-Christophe Calvet, Claire Gruhier, Thomas R. H. Holmes, Richard A. M. de Jeu, and Wolfgang Wagner (2009). "An Intercomparison of ERS-Scat and AMSR-E Soil Moisture Observations with Model Simulations over France". In: *Journal of Hydrometeorology* 10.2, pp. 431–447. ISSN: 1525-755X. DOI: 10.1175/2008JHM997.1.

Ryu, Dongryeol and James S. Famiglietti (2005). "Characterization of footprint-scale surface soil moisture variability using Gaussian and beta distribution functions during the Southern Great Plains 1997 (SGP97) hydrology experiment". In: *Water Resources Research* 41, pp. 1–13. ISSN: 00431397. DOI: 10.1029/2004WR003835.

Salkind, Neil J (2010). "Standard error of the estimate". In: *Encyclopedia of research design. Volume 3*. Ed. by SAGE. London, pp. 1426–1430.

Schenker, Nathaniel and Jane F. Gentleman (2001). "On Judging the Significance of Differences by Examining the Overlap Between Confidence Intervals". In: *The American Statistician* 55.3, pp. 182–186.

Schoonover, Jon E. and Jackie F. Crim (2015). "An Introduction to Soil Concepts and the Role of Soils in Watershed Management". In: *Journal of Contemporary Water Research & Education* 154.1, pp. 21–47. ISSN: 19367031. DOI: 10.1111/j.1936-704X.2015.03186.x.

Seneviratne, Sonia I., Thierry Corti, Edouard L. Davin, Martin Hirschi, Eric B. Jaeger, Irene Lehner, et al. (2010). "Investigating soil moisture-climate interactions in a changing climate: A review". In: *Earth-Science Reviews* 99.3-4, pp. 125–161. ISSN: 00128252. DOI: 10.1016/j.earscirev.2010.02.004. arXiv: WebofScience.

Seyfried, Mark S. and B. P. Wilcox (1995). "Scale and the Nature of Spatial Variability: Field Examples Having Implications for Hydrologic Modeling". In: *Water Resources Research* 31.1, pp. 173–184. ISSN: 19447973. DOI: 10.1029/94WR02025.

Sharma, R., R. G. Shandil, and G. Kapoor (2011). "A Note on Karl Pearson's Coefficient of Dispersion". In: *Himachal Pradesh University Journal* 2011.July, pp. 1–3.

Singh, Dharmendra, P. K. Gupta, R. Pradhan, A. K. Dubey, and R. P. Singh (2016). "Discerning shifting irrigation practices from passive microwave radiometry over Punjab and Haryana". In: *Journal of Water and Climate Change* 8.2.

Skøien, J. O., Günter Blöschl, and Andrew W. Western (2003). "Characteristic space scales and timescales in hydrology". In: *Water Resources Research* 39.10. ISSN: 00431397. DOI: 10.1029/2002WR001736.

Smith, Adam, Jeffrey P. Walker, Andrew W. Western, R. I. Young, Kevin Ellett, R. C. Pipunic, et al. (2012). "The Murrumbidgee soil moisture monitoring network data set". In: *Water Resources Research* 48.7, W07701. ISSN: 0043-1397. DOI: 10.1029/2012WR011976.

Smith, Steven W. (1997). "Moving Average Filters". In: *The Scientist and Engineer's Guide to Digital Signal Processing*. San Diego, California, USA: California Technical Publishing. Chap. 15, pp. 277–284. ISBN: 0966017633. DOI: 10.1016/0165-1765(91)90077-X.

Solano, Ramon, Kamel Didan, Andree Jacobson, and Alfredo Huete (2010). *MODIS Vegetation Index User's Guide (MOD13 Series)*. Tech. rep. May. University of Arizona, p. 38.

Stoffelen, Ad (1998). "Toward the true near-surface wind speed: Error modeling and calibration using triple collocation". In: *Journal of Geophysical Research* 103.C4, p. 7755. ISSN: 0148-0227. DOI: 10.1029/97JC03180.

Stommel, Henry (1963). "Varieties of oceanographic experience". In: *Science* 139.3555, pp. 572–576. ISSN: 00368075. DOI: 10.1126/science.139.3555.572.

Su, Chun-Hsu and Dongryeol Ryu (2015). "Multi-scale analysis of bias correction of soil moisture". In: *Hydrology and Earth System Sciences* 19.1, pp. 17–31. ISSN: 1607-7938. DOI: 10.5194/hess-19-17-2015.

Svoboda, Mark, Doug LeComte, Mike Hayes, Richard R. Heim, Karin L. Gleason, Jim Angel, et al. (2002). "The drought monitor". In: *Bulletin of the American Meteorological Society* 83.8, pp. 1181–1190. ISSN: 00030007. DOI: 10.1175/1520-0477(2002)083<1181:TDM>2.3.CO;2.

Taylor, Karl E. (2001). "Summarizing multiple aspects of model performance in a single diagram". In: *Journal of Geophysical Research: Atmospheres* 106.D7, pp. 7183–7192. DOI: 10.1029/2000JD900719.

Thibeault, M., J. M. Caceres, D. Dadamia, A. G. Soldano, M. Uriburu Quirno, J. M. Guerrieri, et al. (2015). "Spatial and temporal analysis of the Monte Buey SAOCOM and SMAP core site". In: *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, pp. 969–971. ISBN: 978-1-4799-7929-5. DOI: 10.1109/IGARSS.2015.7325929.

Tomer, Sat Kumar, Ahmad Al Bitar, Muddu Sekhar, Mehrez Zribi, Soumya Bandyopadhyay, and Yann H. Kerr (2016). "MAPSM: A Spatio-Temporal Algorithm for Merging Soil Moisture from Active and Passive Microwave Remote Sensing". In: *Remote Sensing* 8.12, p. 990. ISSN: 2072-4292. DOI: 10.3390/rs8120990.

Topp, G. Clarke (2003). "State of the art of measuring soil water content". In: *Hydrological Processes* 17.14, pp. 2993–2996. ISSN: 0885-6087. DOI: 10.1002/hyp.5148.

Torres, Guilherme M., Romulo P. Lollato, and Tyson E. Ochsner (2013). "Comparison of drought probability assessments based on atmospheric water deficit and soil water deficit". In: *Agronomy Journal* 105.2, pp. 428–436. ISSN: 00021962. DOI: 10.2134/agronj2012.0295.

Trenberth, K E (2009). "Precipitation in a changing climate - more floods and droughts in the future". In: *Water in a changing climate, GEWEX News*. Vol. 19. 2. Melbourne, pp. 8–10.

Trenberth, K E, L Smith, T T Qian, Aiguo Dai, and J Fasullo (2007). "Estimates of the global water budget and its annual cycle using observational and model data". In: *Journal of Hydrometeorology* 8.4, pp. 758–769. ISSN: 1525-755X. DOI: 10.1175/jhm600.1.

Vachaud, G., A. Passerat De Silans, P. Balabanis, and M. Vauclin (1985). "Temporal Stability of Spatially Measured Soil Water Probability Density Function". In: *Soil Science Society of America Journal* 49.4, p. 822. ISSN: 03615995. DOI: 10.2136/sssaj1985.03615995004900040006x.

Velpuri, Naga Manohar, Gabriel B. Senay, and Jeffrey T. Morisette (2016). "Evaluating New SMAP Soil Moisture for Drought Monitoring in the Rangelands of the US High Plains". In: *Rangelands* 38.4, pp. 183–190. ISSN: 01900528. DOI: 10.1016/j.rala.2016.06.002.

Vinnikov, Konstantin Y., Alan Robock, Nina A. Speranskaya, and C. Adam Schlosser (1996). "Scales of temporal and spatial variability of midlatitude soil moisture". In: *Journal of Geophysical Research: Atmospheres* 101.D3, pp. 7163–7174. ISSN: 01480227. DOI: 10.1029/95JD02753.

Vogelzang, Jur and Ad Stoffelen (2012). *Triple collocation, EUTMETSAT report*. Tech. rep.

Wagner, Wolfgang, Vahid Naeimi, Klaus Scipal, Richard A. M. de Jeu, and José Martínez-Fernández (2007). "Soil moisture from operational meteorological satellites".

In: *Hydrogeology Journal* 15.1, pp. 121–131. ISSN: 14312174. DOI: 10.1007/s10040-006-0104-6.

Walker, Jeffrey P. and Paul R. Houser (2004). "Requirements of a global near-surface soil moisture satellite mission: Accuracy, repeat time, and spatial resolution". In: *Advances in Water Resources* 27.8, pp. 785–801. ISSN: 03091708. DOI: 10.1016/j.advwatres.2004.05.006.

Wan, Zhengming (1999). *MODIS Land-Surface Temperature Algorithm Theoretical Basis Document (LST ATBD)*. Tech. rep. v. 3.3. Institute for Computational Earth System Science, University of California, Santa Barbara.

– (2006). *MODIS Land Surface Temperature Products Users' Guide - Collection 5*. Tech. rep. September. Sioux Falls, South Dakota: USGS/Earth Resources Observation and Science (EROS) Center.

Wanders, Niko, Marc Bierkens, Steven de Jong, A de Roo, and Derek Karssenberg (2014a). "The benefits of using remotely sensed soil moisture in parameter identification of large-scale hydrological models". In: *Water Resources Research* 50.8, pp. 6874–6891. ISSN: 1944-7973. DOI: 10.1002/2013WR014639.

Wanders, Niko, Derek Karssenberg, A De Roo, Steven de Jong, and M F P Bierkens (2014b). "The suitability of remotely sensed soil moisture for improving operational flood forecasting". In: *Hydrol. Earth Syst. Sci* 18, pp. 2343–2357. DOI: 10.5194/hess-18-2343-2014.

Wang, Chunmei, Qiang Zuo, and Renduo Zhang (2008). "Estimating the necessary sampling size of surface soil moisture at different scales using a random combination method". In: *Journal of Hydrology* 352.3-4, pp. 309–321. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2008.01.011.

Wang, Chunmei, Qingyan Meng, Yulin Zhan, Jing Peng, Xiangqin Wei, Jian Yang, et al. (2015a). "Ground sampling methods for surface soil moisture in heterogeneous pixels". In: *Environmental Earth Sciences* 73.10, pp. 6427–6436. ISSN: 1866-6280. DOI: 10.1007/s12665-014-3866-5.

Wang, F. K. (2001). "Confidence interval for the mean of non-normal data". In: *Quality and Reliability Engineering International* 17, pp. 257–267. DOI: 10.1002/qre400.

Wang, Guojie, Damien Garcia, Yi Y Liu, Richard A. M. de Jeu, and a. Johannes Dolman (2012). "A three-dimensional gap filling method for large geophysical datasets: Application to global satellite soil moisture observations". In: *Environmental Modelling & Software* 30, pp. 139–142. ISSN: 13648152. DOI: 10.1016/j.envsoft.2011.10.015.

Wang, Jianghao, Yong Ge, Gerard B M Heuvelink, and Chenghu Zhou (2015b). "Upscaling In Situ Soil Moisture Observations To Pixel Averages With Spatio-Temporal Geostatistics". In: *Remote Sensing* 7.9, pp. 11372–11388. ISSN: 20724292. DOI: 10.3390/rs70911372.

Webster, R. and M. A. Oliver (1992). "Sample adequately to estimate variograms of soil properties". In: *Journal of Soil Science* 43.1, pp. 177–192. ISSN: 00224588. DOI: 10.1111/j.1365-2389.1992.tb00128.x.

Western, Andrew W. and Günter Blöschl (1999). "On the spatial scaling of soil moisture". In: *Journal of Hydrology* 217.3-4, pp. 203–224. ISSN: 00221694. DOI: 10.1016/S0022-1694(98)00232-7.

Western, Andrew W., Günter Blöschl, and Rodger B. Grayson (1998). "Geostatistical characterisation of soil moisture patterns in the Tarrawarra catchment". In: *Journal of Hydrology* 205.1-2, pp. 20–37. ISSN: 00221694. DOI: 10.1016/S0022-1694(97)00142-X.

Western, Andrew W., Rodger B. Grayson, and Günter Blöschl (2002). "Scaling of Soil Moisture: A Hydrologic Perspective". In: *Annual Review of Earth and Planetary Sciences* 30.1, pp. 149–80. ISSN: 0084-6597. DOI: 10.1146/annurev.earth.30.091201.140434.

Western, Andrew W., Sen Lin Zhou, Rodger B. Grayson, Thomas A. McMahon, Günter Blöschl, and David J. Wilson (2004). "Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes". In: *Journal of Hydrology* 286.1-4, pp. 113–134. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2003.09.014.

Whitcher, Brandon, Peter Guttorp, and Donald B. Percival (2000). "Wavelet analysis of covariance with application to atmospheric time series". In: *Journal of Geophysical Research* 105.D11, p. 14941. ISSN: 0148-0227. DOI: 10.1029/2000JD900110.

Wigneron, Jean-Pierre, Yann H. Kerr, Philippe Waldteufel, Kauzar Saleh, Maria José Escorihuela, Philippe Richaume, et al. (2007). "L-band Microwave Emission of the Biosphere (L-MEB) Model: Description and calibration against experimental data sets over crop fields". In: *Remote Sensing of Environment* 107, pp. 639–655. ISSN: 00344257. DOI: 10.1016/j.rse.2006.10.014.

Yee, Mei Sun, Jeffrey P. Walker, Alessandra Monerris, Christoph Rüdiger, and Thomas J. Jackson (2016). "On the identification of representative in situ soil moisture monitoring stations for the validation of SMAP soil moisture products in Australia". In: *Journal of Hydrology* 537, pp. 367–381. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2016.03.060.

Yilmaz, M. Tugrul and Wade T. Crow (2014). "Evaluation of Assumptions in Soil Moisture Triple Collocation Analysis". In: *American Meteorological Society*, pp. 1293–1302. ISSN: 15257541. DOI: 10.1175/JHM-D-13-0158.1.

Zhan, Xiwu, Paul R. Houser, Jeffrey P. Walker, and Wade T. Crow (2006). "A method for retrieving high-resolution surface soil moisture from hydros L-band radiometer and radar observations". In: *IEEE Transactions on Geoscience and Remote Sensing* 44.6, pp. 1534–1544. ISSN: 01962892. DOI: 10.1109/TGRS.2005.863319.

Zhan, Xiwu, Wade T. Crow, Thomas J. Jackson, and Peggy E. O'Neill (2008). "Improving spaceborne radiometer soil moisture retrievals with alternative aggregation rules for ancillary parameters in highly heterogeneous vegetated areas". In: *IEEE Geoscience and Remote Sensing Letters* 5.2, pp. 261–265. ISSN: 1545598X. DOI: 10.1109/LGRS.2008.915931.

Zreda, M., W. J. Shuttleworth, X. Zeng, C. Zweck, D. Desilets, T. Franz, et al. (2012). "COSMOS: the COsmic-ray Soil Moisture Observing System". In: *Hydrology and Earth System Sciences* 16.11, pp. 4079–4099. ISSN: 1607-7938. DOI: 10.5194/hess-16-4079-2012.

Zwieback, Simon, Klaus Scipal, Wouter A. Dorigo, and Wolfgang Wagner (2012). "Structural and statistical properties of the collocation technique for error characterization". In: *Nonlinear Processes in Geophysics* 19.1, pp. 69–80. ISSN: 1607-7946. DOI: 10.5194/npg-19-69-2012.

AUTEUR : Beatriz MOLERO RODENAS

TITRE : Différence d'échelle spatiale entre les mesures satellitaires et *in situ* d'humidité du sol : analyse par des approches spatio-temporelles

RÉSUMÉ : L'humidité du sol est une variable climatique essentielle dont le suivi à l'échelle globale est possible grâce à des instruments micro-ondes à bord des satellites. La précision de ces estimations est validée par comparaison directe aux mesures au sol. Tandis que les estimations satellitaires ont des résolutions allant de 30 à 100 km, les capteurs *in situ* sont généralement représentatifs d'une zone de quelques centimètres (résolution ponctuelle). Cette différence entre l'échelle spatiale des estimations satellitaires et *in situ* impacte le processus de validation et les statistiques obtenues à un niveau qui n'est pas connu actuellement.

Cette thèse vise à améliorer la connaissance de l'impact du changement d'échelle spatiale, ainsi qu'à fournir des méthodes d'évaluation de celle-ci applicables à toute zone de validation. Pour ce faire, la relation entre les échelles spatiales et temporelles a été étudiée. Des séries modélisées et mesurées sur des régions différentes du globe ont été décomposées en échelles de temps allant de 0,5 et 128 jours, en utilisant des transformées en ondelettes. La représentativité spatiale des mesures à résolution ponctuelle a ensuite été évaluée, par échelle de temps, avec 4 approches différentes : l'analyse de la stabilité temporelle, la triple colocation, le pourcentage de zones corrélées (CArea) et une nouvelle approche utilisant des corrélations basées sur des ondelettes (WCor). De plus, l'incertitude d'échantillonnage a été évaluée séparément avec des approches bootstrap et des simulations de Monte Carlo de séries à résolution ponctuelle.

À l'issue de ces expériences, il y a été constaté que la moyenne des valeurs de représentativité spatiale obtenues tend à augmenter avec l'échelle de temps, mais aussi leur dispersion. Cela implique que certaines stations ont de vastes zones de représentativité à des échelles saisonnières, tandis que d'autres ne l'ont pas. Aux échelles sous-hebdomadaires, toutes les stations présentaient de très petites zones de représentativité. Enfin, l'impact de l'incertitude d'échantillonnage s'est avéré assez important dans les métriques de validation satellitaire.

MOTS-CLÉS : Humidité du sol, différence d'échelle spatiale, résolution, ondelettes, approches spatio-temporelles

---

TITLE: Analysis of the spatial scale mismatch between satellite and ground measurements of soil moisture using spatio-temporal approaches

ABSTRACT: Soil moisture is an essential climate variable that is globally monitored with the help of satellite borne microwave instruments. The accuracy of satellite soil moisture estimations is assessed by direct comparison to *in situ* measurements. While satellite estimates have a resolution ranging between 30 and 100 km, *in situ* sensors typically measure over an area of a few centimetres (point resolution). This spatial scale mismatch between satellite and *in situ* soil moisture estimates impairs the validation process and the respective summary statistics to an extent that is not currently known.

This thesis aims at improving the knowledge of the spatial scale mismatch, as well as providing methods for its assessment applicable to any validation area. To this end, the connection between the SM spatial and time scales was investigated. Modelled and measured soil moisture series at different regions of the globe were decomposed into time scales ranging from 0.5 to 128 days, using wavelet transforms. The spatial representativeness of the point measurements was then assessed, on a per time scale basis, with 4 different approaches: temporal stability analysis, triple collocation, the percentage of correlated areas (CArea) and a new approach that uses wavelet-based correlations (WCor). Moreover, one of the components of the mismatch, the sampling uncertainty, has been assessed separately with bootstrap and Monte Carlo simulations of point-support series.

It was found that the average of the spatial representativeness values tends to increase with the time scales but so does their dispersion. This implies that some stations had large representativeness areas at seasonal scales, while others do not. At sub-weekly scales, all stations exhibited very small representativeness areas. Finally, the sampling uncertainty has been observed to have a considerable impact on satellite validation statistics.

KEY-WORDS: Soil moisture, spatial scale mismatch, resolution, wavelets, spatio-temporal approaches

---

DIRECTEUR DE THESE : Yann KERR

LIEU ET DATE DE SOUTENANCE : CESBIO, 18 de novembre de 2018

DISCIPLINE ADMINISTRATIVE : Surfaces et interfaces continentales, Hydrologie

INTITULÉ ET ADRESSE DE L'U.F.R. OU DU LABORATOIRE : Centre d'Études Spatiales de la BIOsphère - CESBIO (UMR 5126), 18 avenue. Édouard Belin, bpi 2801, 31401 Toulouse cedex 9