Solar irradiance forecast and assesment in the intertropical zone

Mouhamet Diallo

Thèse

En vue de l'obtention du grade de

Docteur de l'université de Guyane

Spécialité : *Sciences pour l'ingénieur, mention physique*

Présentée et soutenue publiquement le *18 Décembre 2018* par :

DIALLO MOUHAMET

Estimation et prédiction de l'ensoleillement en zone intertropicale

JURY

SOUBDHAN Ted	Professeur, Université des Antilles	Président du Jury
FOURNIER Richard	Professeur, Université Toulouse III	Examinateur
BLANC Phillipe	Directeur de recherche, Mines ParisTech	Rapporteur
DUDHIA Jimy	Directeur de recherche, National center for atmospheric research	Rapporteur
LINGUET Laurent	Professeur, Université de Guyane	Directeur de Thèse
SEYLER Frédérique	Directrice de recherche, Institut de recherche pour le développement	Co-Directrice de Thèse

Ecole Doctorale n°587 : *Diversités, Santé et développement en Amazonie*

Don’t inspire to make a living, inspire to make a difference. (D. Washington)

Genius is 1% inspiration and 99% perspiration (E. Thomas)

La vie se résume à une succession de choix: choisir ses amis pour qu’ils soient des amis qui vous poussent en direction de la réussite; choisir sa femme afin que dans les moment le s plus difficiles, qu’elles soit honnête envers nous; choisir comment investir son temps, ses efforts. Néanmoins le choix le plus important sur lequel on a aucune emprise, c’est le choix de ses parents. Merci papa, merci maman.

Remerciements

Je commencerai les remerciements par deux citations qui en résument le contenu. Celui qui ne sait pas d'où il vient ne peut savoir où il va ; tout seul on va plus vite ensemble on va plus loin…

À cet effet je tiens à remercier mes directeurs de thèse : Mr Linguet, Mme Seyler. Je les remercie tout d'abord de m'avoir choisi pour mener à bout ce projet de recherche. Je leur exprime ma reconnaissance et les remercie ensuite pour le concours qu'ils ont apportés pour que cette thèse arrive à terme.

Pour ne froisser personne pas faire de jaloux je tiens à préciser que je cite chaque individu par date de connaissance et en cas de doute par ordre alphabétique! Donc si vous vous retrouvez plus bas dans la liste, vous n'aviez qu'à me rencontrer plus tôt ☺

Je remercie ensuite mes collègues et amis doctorants de l'ird du cnrs : Marco, Tommy, Youven, Marjorie, Justine, Camilla, Sihem, Maha, Wilna et Sushi. Ils ont grandement contribué par leurs gaités et soutiens à traverser les moments difficiles de la thèse. J'espère que ces amitiés tissées dans les moments difficiles perdura encore longtemps. J'espère que vous vous reconnaitrez à travers ces mots-clés : « clic !, No stress, Mr steele, osef, espoir, ah les femmes !, inchallah !, all you need is love ☺

Même si j'ai passé une majeure de mon temps au bureau, j'ai aussi rencontré des gens à l'extérieure que je ne remercierai jamais assez des moments agréables que l'on a partagé. Même si pour certains nous avons perdu contact ou que nos relations se sont détériorées: rendons à César ce qui est à César :
Sevahnee, Mme Félix !, Manu, Juliette, Guilhem, Miléna, Clément, Pris, Mélanie, Thibault, Ronan, Djanella, Christophe, Selma, Iris, Abdoul lakhad, Bass, Nico, Sarah, Paulo, Mariano, Viviane et Roger, Katia, Stef, Leïla, Parrain !, Audrey, David, Babacar, Ahmadou, Moctar, Bill, Séverin, Mr Diaby, Charlotte. J'espère que vous vous reconnaitrez à travers ces mots-clés : J'attends toujours, tapisser, brega, ta copine elle est celib ?, non distribution, sms, dragué !, largué ☺!, chaleureux, simplicité !, l'espoir (je crois que c'est le restaurant ou on a mangé le jour de mon arrivée ?), joxia, regard des autres, planète, rorota, duffman, escroc (amicale), retard, frère , consiglieri, no-stress, travaillopathe.

Je remercie aussi le personnel administratif, technique, et de recherche aussi bien de l'ird du cnrs, l'UG et de Météo-France: Madame Duterville, Madame Robinel, Rosiane, Christophe, Pape, Michel, Annaig, Abdennebi, Max, Eric, Josi, Jeanine, Jean louis, Jean Claude, Rolland, Serge, Marie-claude, Ahmed, Phillipe, Mati, Allyx, Thomas, Olivier, Mme Allouache, Madame Aouizerate, Luc, Corinne, Mr Criton, Madame Ho-Coui-Youn. J'espère que vous vous reconnaitrez à travers ces mots-clés : bonté !, sincérité !, amitié, d'une très grande aide, coquin, Salsa, chaleureux, brute de décoffrage conciliant!

Let's talk about conference and or « training session » friends: Claire, Mireille, Dominique, Hadrien, Camille Craig, Rami ! Thank you for everything, I hope to see you and work with you sooner than later !

Merci pour tout …

Résumé de la thése

La Guyane est un territoire d'outre-mer, situé en zone intertropicale (ZIT). Cette zone est le lieu de phénomènes de convections intenses. De ce fait, l'énergie solaire incidente au sol est très variable ce qui constitue un frein à son exploitation à grande échelle.

La question de recherche étudiée dans ce manuscrit est: comment peut-on améliorer les estimations et prédictions de rayonnement au sol en ZIT de façon à augmenter le taux de pénétration dans le réseau électrique de cette énergie renouvelable intermittente? Afin de répondre à cette question, nous avons utilisé deux outils. Le code Héliosat-II (HII) et le modéle de prévisions météorologiques Weather and research forecast (WRF). Nous avons utilisé ces outils de manière à améliorer les estimations et prévisions de rayonnement global au sol (IGH) dans la ZIT.

La première partie de ce manuscrit présente le contexte de la thèse. La seconde présente une modification d'H-II permettant d'améliorer les estimations d'IGH par une modélisation explicite de l'absorption de nuages. Ces estimations améliorées donnent ainsi des outils décisionnels permettant de situer au mieux une centrale solaire en fonction du potentiel solaire du site et des systèmes services avoisinants. La seconde partie traite dans un premier précision des prévisions des modèles globaux IFS et GFS (i.e integrated forecast system, global forecast system GFS) en ZIT. Ces produits téléchargés sont validés par comparaison avec des mesures in situ de trois pays situés dans la ZIT et caractérisés par des climats tropicaux. Cette étude permet de combler un vide dans l'étude des prévisions d'IGH des modèles globaux en ZIT. Nous proposons ensuite une méthode générique permettant de calibrer le modèle WRF en ZIT. Cette méthodologie vise à limiter le nombre de simulations à effectuer en sélectionnant et en faisant varier uniquement les paramètres ayant le plus d'influence sur le rayonnement au sol en ZIT. Pour valider cette méthodologie nous avons comparé les prévisions d'IGH du modelé WRF calibré avec celle du modelé AROME ainsi qu'avec des mesures in situ en Guyane. La quatrième partie présente l'utilisation d'une méthode hybride ensembliste variationnelle d'assimilation de donnée permettant d'améliorer les prévisions de rayonnements en ZIT. Cette méthode initialement utilisée pour améliorer la description de phénomènes convectifs extrêmes tels que prévision de la trajectoire des cyclones est pour la première fois appliquée pour améliorer les prévisions d'IGH. Cette méthodologie appliquée à la ZIT fournie alors des prévisions améliorées d'IGH permettant ainsi une gestion améliorée de centrale solaire.

Les travaux effectués dans la seconde section ont donnés lieu à une publication dans solar energy journal. Les travaux effectués dans la troisième section ont été soumis, mais non publiés à la date de remise du manuscrit de thèse.

Nous présentons une conclusion en français de cette thése après celle rédigé en anglais.

Table of contents

CHAPTER I : Thesis framework and research issue 11
I. Introduction 12
I.1 Technical challenges of solar energy for electricity production 12
I.2 Thesis framework and goals 12
II. Manuscript outline 14
CHAPTER II : Improving ground irradiance assessment in the intertropical zone 16
I. Modeling and measuring the solar irradiance at the ground 17
I.1 Concepts and definitions 17
I.2 Modeling the extraterrestrial irradiance 23
I.3 Instruments for measuring the GHI and quality control procedures 23
I.4 Modeling the solar irradiance depletion due to the atmospheric constituent: the radiative transfer equation 27
I.5 Input datasets for radiative transfer models 30
I.6 Accurate Methods to solve the radiative transfer equation for shortwaves 31
I.7 Approximate methods to solve the radiative transfer equation for shortwaves 35
I.8 Validation metrics 42
II. Improving the GHI estimate from satellite method in the inter-tropical zone 45
II.1 Comparison between statistical, semi-physical and physical parameterizations of the radiative transfer equation 45
II.2 Improving the Heliosat-2 Method for Surface Solar Irradiation Estimation Under Cloudy Sky Areas 46
III. Summary and conclusion of Chapter II 65
CHAPTER III : Assessing the accuracy of numerical weather prediction model in the intertropical zone 66
I. Forecasting the solar irradiance at the ground 67
I.1 Recommendation for solar energy forecast method depending on the horizon 67
I.2 Introduction to Numerical weather prediction models 68
II. Studying the GHI forecast accuracy of global NWP models in the intertropical zone 73
III. Forecasting the GHI in the intertropical zone using mesoscale NWP models 95
IV. Summary and conclusion of chapter III 110
CHAPTER IV : Improving the irradiance forecast of numerical weather prediction model using data assimilation 111

I. Improving the solar irradiance forecast under tropical climate for long range forecast horizon 112
II. Data assimilation methods 113
II.1 Stochastic method and variational methods 114
II.2 Nudging methods 116
III. Comparison between DA algorithms 117
IV. Data assimilation for solar energy forecast purposes in the intertropical zone 120
IV.1 Description of the Hybrid 3D Ensemble variational 120
IV.2 Building the ensemble members 121
IV.3 Modeling of the background error covariances A and B 122
IV.4 Experiments 123
IV.5 Conclusions 129
CHAPTER V : Thesis summary and conclusions 130
Appendix A 134
I. Variational method 134
I.1 3D-Var 134
I.2 4D-Var 135
II. Stochastic method 136
II.1 Kalman filter 136
II.2 stochastic and deterministic ensemble Kalman filter 137
III. Back and forth nudging 139
Bibliography 140

List of figures

Fig. I.1 French Guiana electricity grid (Ministry of the environment, 2016a) 14
Fig. II.1 Earth radiation budget credit American meteorological society (AMS) 17
Fig. II.2 schematic representation of the solar zenith angle $\theta 0$, sun declination δ, the hour angle h (Liou, 2002) 18
Fig. II.3 schematic representation of the azimuth angle $\emptyset$ (Sidek et al., 2014) 18
Fig. II.4 schematic representation of angles involved in radiometric quantities (Liou, 2002) 19
Fig. II.5 schematic representation of scattering angle for an incoming ray (In) in the direction $\Omega(\mu, \phi)$that is deviated in the direction $\Omega'(\mu', \phi')$ (Liou, 2002) 22
Fig. II.6 Solar radiation spectrum credit geosciencebigpicture 23
Fig. II.7 thermopile (left) and photodiode (right) pyranometer. Source Wikipedia, Licor 24
Fig. II.8 Plane Parallel atmosphere 28
Fig. II.9 Transfer of solar radiation in plane-parallel layers, illustrating attenuation by extinction, a; multiple scattering, b; and single scattering of the unscattered solar flux, c. Source (Liou, 2002) 29
Fig. II.10 Configuration of the adding method for two layers of optical depth $\tau 1$ and $\tau 2$respectively. The upward transmission and reflection of the ith layer are $Ti*$and Ri , respectively whereas the downward transmission and reflection are Ti and $Ri*$. D and U are the the combined total transmission and reflection functions between layers 1 and 2. Source (Liou, 2002) 33
Fig. II.11 solar radiation flux at the atmosphere inspired after (Noia et al., 1993a) 37
Fig. II.12 atmospheric model of Gautier et al. (1980) adapted from (Noia et al., 1993b) 39
Fig. II.13 atmospheric model of Janjai, (2010) 40
Fig. III.1 Conceptual diagram of forecast skill as a function of forecast lead time for different forecast methods. Extracted from (Kleissl, 2013; Ruiz-Arias and Goenka, 2017) 67
Fig. III.2 Physical process accounted for in NWP models extracted from (Helmert, 2016) 69
Fig. III.3 Interaction between physical parameterization. Extracted from (Dudhia, 2014) 70
Fig. VII.1 Measurements instrument used in DA system credit WMO 113
Fig. VII.2 Observation network in ECMWF DA system 114
Fig. VII.3 Flowchart describing the hybrid ETKF–3DVAR DA system in WRF model; inspired after (Barker and Clayton, 2011; Kutty et al., 2018) 121
Fig. VII.4 Cdf of Kc values in March, May, September. The GHI forecasts from Rochambeau, Saint Georges, Maripasoula, Saint Laurent, Kourou and Ile Royale 124
Fig. VII.5 Cdf of V values in March, May, September. The GHI forecasts from Rochambeau, Saint Georges, Maripasoula, Saint Laurent, Kourou and Ile Royale 126
Fig. VII.6 WRF-M9, WRF-M9-3DEnVar, and IFS MAE as a function of K_c and V 127
Fig. VII.7 MAE (W/m²) and MBE (W/m²) of WRF-M9, WRF-M9-3DEnVar-c and IFS 128

List of tables

Tab. I.1Projection of population evolution and electricity needs. The electricity production is based on the BPEOD reference scenario. The population evolution is based on the INSEE projection as reported in the PPE (Ministry of the environment, 2016a) 12
Tab. II.1 WMO classification of pyranometers, extracted from (Sen, 2008)..... 25
Tab. II.2 literature review of limits within range QC procedure 26
Tab. II.3 input dataset for radiative transfer code Source (ECMWF, 2016a; Stensrud, 2009; Stephens, 1984)..... 30
Tab. II.4 spectral interval choose in RRTM, extracted from (ECMWF, 2016a)..... 35
Tab. II.5 Gaussian Points, Gaussian Weights and Legendre polynomials values for two-stream approximation. g is the asymmetry factor defined Eq. II.52. Inspired after (Liou, 2002)..... 41
Tab. III.1 spatial and temporal resolution of GHI forecast products from global model operated by forecast center for the first 48 hour forecasts. The temporal and spatial resolution of these products refers to those of datasets available for download. Source (BAM, 2018; CMC, 2016; DWD, 2018; ECMWF, 2016b; Figueroa et al., 2016; NOAA, 2016; Walters et al., 2017) 71
Tab. III.2 Comparison of commonly used mesoscale model Source: for numerics ((Burridge, 1975; Doms and Baldauf, 2015; Käellen, 1996; Majewski, 2009; NCAR, 2017; Termonia et al., 2018; Tudor et al., 2013; Undén and co-authors, 2002); for Radiation (Doms and Baldauf, 2015; Jimenez et al., 2016; Termonia et al., 2018; Walters et al., 2017; Zubler et al., 2011) 72
Tab. VII.1 comparison of advantage and disadvantage of existing DA algorithms used in NWP models. These DA algorithm are detailed Appendix A 118
Tab. VII.2 Comparison between WRF-M9 and WRF-M9-3DEnVar for all experiments. All six stations showed chapter III section III were merged..... 125
Tab. VII.3 Relative improvement of WRF-M9-3DEnVar-c over WRF-M9 and IFS 128
Tab. VII.4 Relative improvement of the MAE WRF with MOS over IFS. These relative improvement were computed using MAE values for a 48h forecast initialized at 00h UTC that were provided by Perez et al. (2013). Mediterranenan stations include Cordoba, Huelva,Granada, Carpentras; Desert rock is the Arid station considered; For Humid continental we considered Goodwin Creek and Penn state..... 128

Nomenclature

3D-Var: 3 dimension variational data assimilation
BEC: Background error covariance
cdf: cumulative distribution function
CVT: control variable transform
DA: Data assimilation
DENKF: deterministic ensemble Kalman filter
EAKF: ensemble adjust Kalman filter
EDF: électricité de France; French historical energy provider
ENKF: ensemble Kalman filter
ENTKF: ensemble transform Kalman filter
ESRF: ensemble square root filter
ETKF: ensemble transform Kalman filter
GFS: global forecast system
GHI: Global horizontal irradiance
GOES: Geostationary Operational Environmental Satellite
GTOA: top of atmosphere irradiance
IFS: integrated forecast system
ITCZ: inter tropical convergence zone
ITZ: intertropical zone
KF: Kalman filter
LCOE: The levelized cost of electricity assesses the energy production of a device weighted by its lifetime costs. It allows a comparing the cost of generating electricity from alternative energy technologies.
MME: Multi model ensemble
MOS: Model output statistic
NDA: Nudging Data Assimilation
NWP: numerical weather prediction model
OI: optimal interpolation
pdf: probability distribution function
PV: photovoltaic
QC: quality check
RE: renewable energies
RTE: radiative transfer equation
RTP: radiative transfer parameterization
SENKF: stochastic ensemble Kalman filter
TL: Linke trouble
WMO: World Meteorological Organization
WRF: Weather and research forecast model

α_w: water vapor absorption
α_{aer}: aerosol absorption
B background error covariance matrix
$\beta_{\mathrm{a},\lambda}$: absorption coefficient
$\beta_{\mathrm{e},\lambda}$: extinction coefficient
$\beta_{\mathrm{s},\lambda}$: scattering coefficient
C:cloud amount
d_0: innovation for the analysis at t_0
$\delta\vec{x}(0)$: analysis increment
ε : squared ratio between the mean earth-sun distance and the instaneous earth-sun
$\overrightarrow{\varepsilon_a}$: analysis error
$\overrightarrow{\varepsilon_b}$: background error
ε_e: the emissivity
$\overrightarrow{\varepsilon_m}$: model errors
ε_0: observation error
E_λ: irradiance
H: observation operator
$\widetilde{H}(t_0)$: the linearization of the observation observatory at the analysis time t_0
Io : Solar constant
K: Kalman gain
K_c: clear sky index
K_t: clearness index
L: radiance
$M(\vec{x})$: model operator
μ cosines of solar zenith angle
ω_λ: single scattering albedo
P: analysis error covariance matrix
Q: model error covariance matrix
R: observation error covariance matrix
ρ_c: cloud albedo
ρ_g: ground albedo
$\rho_{s,d}$: scattering coefficient for the upward diffuse radiation due to a Rayleigh atmosphere
$\rho_{s,b}$: scattering coefficient for the downward beam radiation due to a Rayleigh atmosphere
$\overline{\mathrm{T}}(\mu_0)$: broadband atmospheric transmission function
T: the transmissivity
T: the temperature
τ_λ :optical depth
$\vec{x}$: NWP model state vector
$\overrightarrow{x_b}$: background state
y_n : observation of the state variable over a given time interval n

CHAPTER I : Thesis framework and research issue

I. Introduction

I.1 Technical challenges of solar energy for electricity production

The sun is a star with a 6.96 10^5 km radius and a surface temperature of 5800K (Liou, 2002). According to Boltzmann law, it emits 3.94 10^{26} W.s toward space. Earth is located on average 1.5 10^8 km from the sun; having a 6378.137 km radius (Liou, 2002), it receives 1.8 10^{17} W.s. This energy is approximately 4000 times of the World's total primary energy consumption in 2016 (BP Statistical Review, 2017; Enerdata, 2017). First studies dealing with the conversion of solar energy into electricity date back to the mid-twenties (Albert Einstein, 1905; Bell labs, 1954; Fairley, 2008). However, at that time, the efficiency and cost of solar cells were not competitive with oil thermal energy. The interest in solar energy was renewed with the launch of the first satellites, as it was a reliable and renewable option for space applications (House et al., 1986). The combination of increasing energy needs (Birol, 2010), increasing concern for global warming (Pachauri and Meyer, 2015), fossil fuel depletion (Shafiee and Topal, 2009) and cost decrease (IRENA, 2016), has established solar energy as a viable solution to meet the needs. According to IEA (2017) the global average levelized cost of electricity (LCOE) from utility-scale solar photovoltaic (PV) projects declined by 70% from 2010 to 2016. Nevertheless, there are high discrepancies in the LCOE between regions. The LCOE in 2016 for solar PV (IRENA, 2018) was ~ 0.1 USD per kWh in Europe whereas in Africa it was ~0.17 USD per kWh. The LCOE in Africa for 2016 was higher than the fossil fuel cost range (i.e. [0.05:0.15] USD per kWh).

Regions with higher solar potential are located in the intertropical zone (ITZ) (Löf et al., 1966; Müller et al., 2015; Trieb et al., 2009). Nevertheless, the solar irradiance in the ITZ is highly variable due to the rapid cloud transformations that occur in this area (Aryaputera et al., 2015; Rossow et al., 2013; Wheeler and Kiladis, 1999). Consequently, the solar energy available at the ground in this area is highly variable. This variability causes economical and technical challenges to fully exploit this resource.

I.2 Thesis framework and goals

French Guiana is a French territory located in the ITZ. This territory has fragile supply and demand equilibrium because it has to produce its energy locally. From 2009 to 2014 French Guiana's population increased on average of 2.4% each year (INSEE: Rémi Charrier, 2017). The 2030 demography projection Tab. I.1 shows that this trend will be magnified. Based on these projections, EDF the historical energy provider in France assessed that there will be an increase of electricity consumption that need to be addressed (Ministry of the environment, 2016a).

Tab. I.1Projection of population evolution and electricity needs. The electricity production is based on the BPEOD reference scenario. The population evolution is based on the INSEE projection as reported in the PPE (Ministry of the environment, 2016a)

	2015	2020	2025	2030
Population number	~280 000	~330 000	~297 000	~380 000
Electricity needs GWh	879	1026	1158	1280

The multiannual energy plan-*programmation pluriannuelle de l'énergie* (PPE) (Ministry of the environment, 2016a) aims to meet the electricity needs using renewable energies (REs). Previous studies (Fillol et al., 2017) showed that French Guiana has a high solar potential. The global horizontal irradiance (GHI) ranges between 1650 and 2000 kWh/kWc/year; the direct normal irradiance (DNI) ranges between 1850 and 2250 kWh/kWc/year. Therefore, solar energy will be a driving force to meet the growing demand in electricity. In 2014, solar photovoltaic has a 6% (48 GWh) share in French Guiana electrical mix with an installed capacity of 34 MW (Ministry of the environment, 2016a). The goal set for French Guiana is to increase this installed capacity by 111.7% by 2023. As an intermittent RE source, the integration of high loads of solar photovoltaic output into the electricity grid introduces instabilities (Clean Energy Council, 2017; IER, 2013; Zoulias: CRES, 2016). This issue led the government to set in 2018 a 35% penetration threshold for intermittent RE sources. In 2012, a penetration rate of 22% was already reached in French Guiana (Dambreville, 2014). Increasing the intermittent REs installed capacity may, on the one hand, provide enough electricity but on the other may destabilize the grid. Various solutions were suggested to increase the penetration of intermittent REs and ensure, at the same time, grid safety. These solutions can be classified into five categories (Zoulias: CRES, 2016):

-use thermal fossil fuel power plant
-update the electricity grid
-use energy storage systems
-use geographical information systems
- forecast the incoming solar energy

The first solution deals with the use of fossil fuel power plants as backup production systems. As the electricity production from REs depend on the weather conditions, it cannot meet user's needs consistently. Fossil fuel power plants have low inertia, they can increase or decrease (i.e. ramp) the electricity production quickly to meet user's needs and ensure grid stability (Ministry of the environment, 2016a). In 2014, thermal fossil fuel represents in French Guiana an installed capacity of 127.4 MW and this capacity is to increase up to 160 MW by 2030 (Ministry of the environment, 2016a).

The second solution deals with extending the grid, improving transport lines and incorporating grid elements with "smart" functionalities in order to balance supply and demand. Extending the grid allows importing and exporting power with the neighboring countries. This extension helps in maintaining the equilibrium between supply and demand, which in turn helps in stabilizing the grid. Inverters are grid devices that feed solar power into the grid, converting the power from direct current to alternating current (Táczi and Szörényi, 2016). These devices cut off the solar plant output to the grid if they detect an abnormal state. As solar energy becomes more and more important if all grid inverter cut off all the solar plant output, it can result in system instabilities. Therefore, each inverter of the grid should vary on its own, adapting the cut-off frequency to grid state (IRENA, 2016). To best of the author knowledge this technology has not yet been implemented in French Guiana. Fig. I.1 shows the 2014 French Guiana's electricity grid; it is 414 km long and is not connected to neighboring countries. Grid lines extend from Saint Laurent to Cayenne. Feasibility studies of grid extension to neighboring countries are planned for 2018 (Ministry of the environment, 2016a).

The third solution is to increase the penetration of intermittent renewable energy into the electricity grid and ensure grid safety by using energy storage systems (IRENA, 2016). Energy storage systems allow storing electricity from variable renewable generation when the production exceeds the demand (IRENA, ETSAP, 2015). The energy previously stored can then be supplied upon demand,

when renewable energy production is not available or insufficient. Energy storage systems increase system flexibility and provide supply security. In 2014, the available energy storage systems in French Guiana are distributed amongst the 2 PV plant of 5 MW installed capacity. By 2023, out of the 53 MW intermittent REs installed capacity, 45 MW will be backed up by energy storage systems. However, energy storage systems require high investments. The cost of a PV facility with backup production systems is estimated to be 3100 k€/MW whereas the cost of a PV facility without backup is estimated to be 2900 k€/MW (Ministry of the environment, 2016a).

The fourth solution deal with the use of Geographical Information Systems (Zoulias: CRES, 2016). This tool is used to identify the most suitable locations to install REs power plants with respect to both energy potential and surrounding facilities that favor grid stability.

The fifth solution is to forecast the solar energy available to solar power plants in order to increase the penetration of solar energy into French Guiana's electricity grid.

In this thesis we consider the fourth and fifth solutions applied in the ITZ.

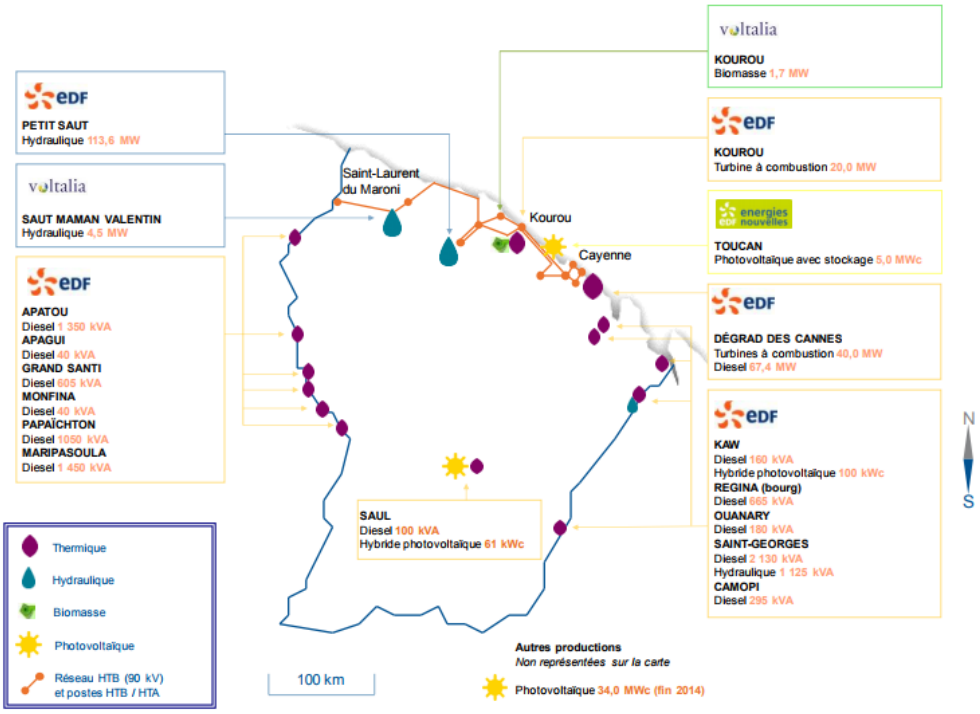

Fig. I.1 French Guiana electricity grid (Ministry of the environment, 2016a)

II. Manuscript outline

This thesis dissertation aims to answer the following scientific issue: how could the solar irradiance be assessed and forecasted in French Guiana to increase the penetration rate of this intermittent renewable energy into the electricity grid? Weather conditions in French Guiana are driven by the physical phenomena in the ITZ; consequently, methods applied in French Guiana can be extended to all the ITZ.

To answer this scientific issue, we use two tools: Heliosat-II method (H-II) and weather and research forecast (WRF) numerical weather prediction model (NWP). We used these tools in order to produce improved GHI estimates in French Guiana and the ITZ.

The first chapter introduces the thesis and the research issue. The second chapter presents a modification to H-II; with this modification H-II can account for cloud absorption. The GHI estimates from modified H-II provide tools for decision making in French Guiana. These tools allow identifying the most suitable locations to install solar facilities in French Guiana with respect to both solar potential and surrounding facilities favoring grid stability. In the third chapter we first study the accuracy of the GHI forecasts from integrated forecast system (IFS) and global forecast system (GFS) NWP models in French Guiana, La Reunion and Singapore. We validate the accuracy of these downloaded products by comparison with ground measurements from these three territories located in the ITZ that have tropical climate. This study aims to fill the gaps in the accuracy of global NWP model in the ITZ. Secondly, we propose a methodology to calibrate WRF to produce improved GHI forecasts in the ITZ. The goal is to restrain and select the minimum number of simulations to run, to obtain improved GHI forecasts compared to a non-calibrated NWP model. This methodology to calibrate WRF is validated in French Guiana by comparison with the GHI forecasts of AROME and ground measurements. The fourth chapter deals with the use of an hybrid 3D variational (3D-Var) ensemble transform Kalman filter (ENTKF) to further improve the GHI forecasts of WRF calibrated for French Guiana. This methodology originally used in the tracking of extreme convection events such as cyclones is applied for the first time for GHI forecasts. Its application in French Guiana allows us to obtain an improved GHI forecast which makes monitoring the electricity production from solar facilities easier. In the fifth chapter, we summarize and conclude thesis.

The research undergone in the second chapter is published in solar energy journal whereas the research on IFS, GFS, WRF and AROME from chapter three were submitted to journal for publication but were not published yet at the time of thesis dissertation submission.

Note: In this manuscript we aimed to remain consistent with notations and abbreviations. Consequently, notations and abbreviations are explained only when they are first introduced. For later mention to these notations and abbreviations, the reader is invited to refer to the nomenclature and definition sections.

CHAPTER II : Improving ground irradiance assessment in the intertropical zone

I. Modeling and measuring the solar irradiance at the ground

The solar irradiance at the ground that may be converted into electricity is a function of the irradiance at the outer edge of the atmosphere and the atmosphere constituents as seen in Fig. II.1. To describe this quantity we first define relevant radiometric and geometric variables. Second we describe how the outer edge of the atmosphere solar irradiance is computed. Third we discuss the measurement devices for the global component of the solar radiation and quality control procedures. Last we discuss the modeling of the solar radiation at the ground and the metrics used to validate the accuracy of a model.

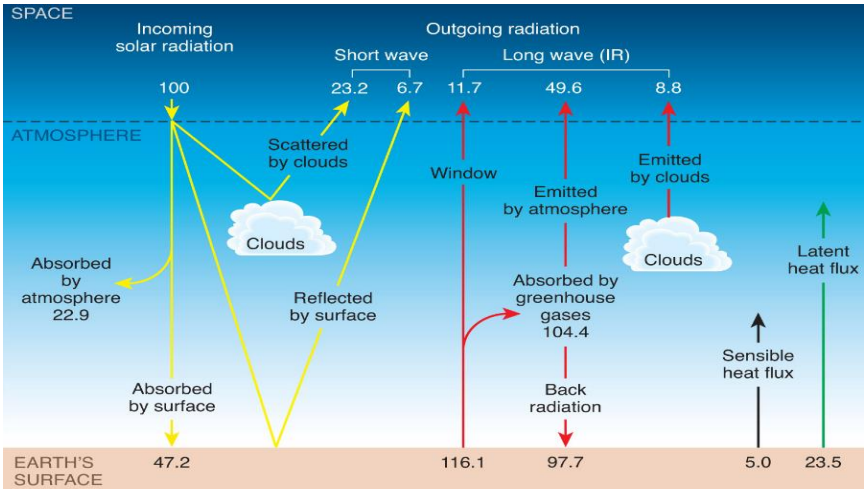

Fig. II.1 Earth radiation budget credit American meteorological society (AMS).

I.1 Concepts and definitions

I.1.a Solar constant I_o

The solar constant I_0 is the total amount of solar energy reaching the top of the atmosphere at the mean distance between the sun and the earth across a surface of unit area normal to the solar beam. The solar constant is computed from the energy F_0 emitted from the sun using an energy conservation principle (Liou, 2002):

$$\mathrm{I_0 = F_0 \left(\frac{r_s}{r_o}\right)^2}, \qquad \text{II.1}$$

where r_s and r_o are the sun radius and earth-sun mean distance respectively; these variables are expressed in m. F_0 is the energy emitted from the sun; its value is given by the Boltzmann law for black bodies. F_0 and I_0 are expressed in W/m². The ratio ε between the mean earth-sun distance and the instantaneous earth-sun distance is the sun correction factor, its approximation can be found in (Mather and Koch, 2011).

I.1.b Solar zenith angle θ_0, μ_0

The solar zenith angle is the angle between the local zenith and the center of the sun disc. It is a function of the hour angle h, the sun declination δ and the local latitude φ and writes:

$$\mu_0 = \cos\theta_0 = \sin\varphi \sin\delta + \cos\varphi \cos\delta \cosh \quad \text{II.2}$$

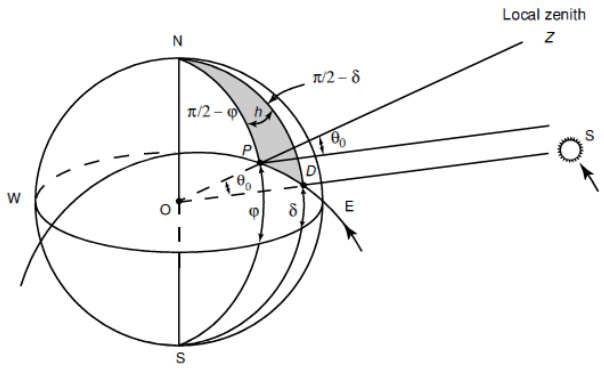

Fig. II.2 schematic representation of the solar zenith angle θ_0, sun declination δ, the hour angle h (Liou, 2002)

I.1.c Azimuth Ø

The azimuth Ø is the angle between a reference direction (here the north) and a line from the observer to a point of interest (here the sun) projected on the same plane as the reference direction orthogonal to the zenith.

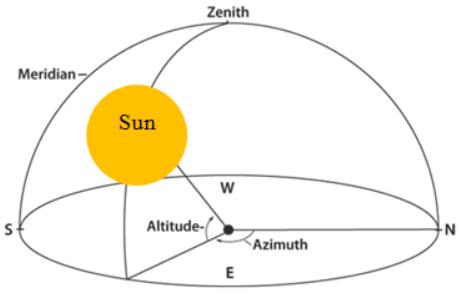

Fig. II.3 schematic representation of the azimuth angle Ø (Sidek et al., 2014)

I.1.d Radiance L, radiation

The monochromatic radiance L_λ or luminance is the amount of radiant energy dE_λemitted by an element of area dA in a specified wavelength interval $[\lambda, \lambda + d\lambda]$, in a direction $d\Omega$ and during a dt time span. The radiance L_λ is an instantaneous value, it is expressed in W/m^3.sr.

$$L_\lambda = \frac{dE_\lambda}{\cos\theta \, d\Omega \, d\lambda \, dt \, dA} \quad \text{II.3}$$

Radiation is the accumulation of radiance over a time span. If the radiance is monochromatic, it is expressed in J/m^3.sr. The total or polychromatic radiance (radiation) is the integration over all wavelengths of the monochromatic radiance (radiation); it is expressed in W/m^2.sr (J/m^2.sr).

I.1.e Irradiance E, Irradiation

The monochromatic irradiance or emittance is the radiant flux received from all directions by an element of area dA. The radiance E_λ is an instantaneous value; it is expressed in W/m^3.

$$\mathrm{E}_\lambda = \int_\Omega \mathrm{L}_\lambda \cos\theta \, \mathrm{d}\Omega \qquad \text{II.4}$$

Irradiation is the accumulation of irradiance over time. If the irradiance is monochromatic it is expressed in J/m^3. The total or polychromatic irradiance (irradiation) is the integration over all wavelengths of the monochromatic irradiance (irradiation); it is expressed in W/m^2 (J/m^2).

I.1.f Top of atmosphere irradiance, G_{TOA}

The top of atmosphere irradiance (G_{TOA}) is the solar flux density at the top of the atmosphere when the instantaneous and mean earth-sun are r_s and r_o respectively, and the sun is located by the solar zenith angle θ_0 in radians.

$$\mathrm{G}_{\mathrm{TOA}} = \mathrm{I}_0 \, \mu_0 \qquad \text{II.5}$$

$\mathrm{G}_{\mathrm{TOA}}$ and $\mathrm{I}_0$are expressed in W/m^2.

I.1.g Global horizontal irradiance, Global horizontal irradiation

The global horizontal irradiance (GHI) is the irradiance E_λ received on a horizontal surface located on the ground.

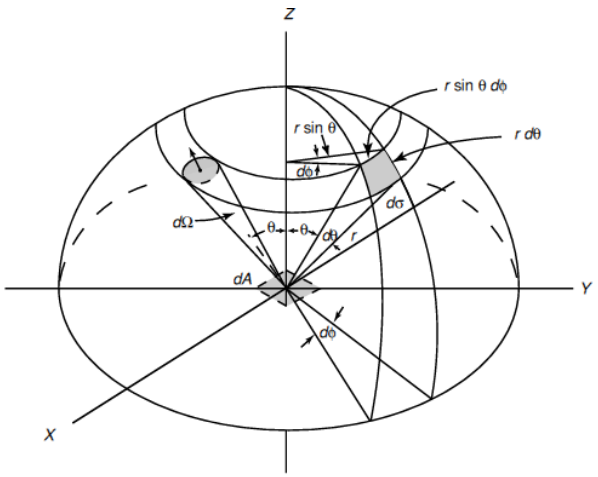

Fig. II.4 schematic representation of angles involved in radiometric quantities (Liou, 2002)

The global horizontal irradiance is composed by two components: the irradiance received directly in a straight line from the sun on a normal surface (DNI) and the irradiance received on the ground after multiple scattering in the atmosphere (DHI).

$$GHI = cos(\theta_0) \, . DNI + DHI \qquad \text{II.6}$$

The global horizontal irradiation is the accumulation of GHI over a time span. The GHI and global horizontal irradiation are expressed in W/m^2 and J/m^2 respectively.

I.1.h Clear sky model Gc

A clear sky model Gc approximates the maximal irradiance value available at the ground under cloudless sky conditions; it accounts for the depletion caused by aerosols and water vapor. Gc always depends on the solar zenith angle; it also requires a varying number of other inputs data that describe the local atmospheric conditions. According to Gueymard (2012a) these input parameters are typically less than or equal to eight.

I.1.i Clearness index K_t and Clear sky index K_c

The clearness index K_t is the ratio of the GHI to the irradiance outside the atmosphere. The clear sky index K_c instead divides the GHI by the clear sky atmosphere irradiance. Both K_c and K_t serve as proxies for the atmosphere transmission factor; these dimensionless numbers are expressed as follows:

$$\mathrm{K_t} = \frac{\mathrm{GHI}}{\mathrm{G_{TOA}}}\ and \qquad \text{II.7}$$

$$K_c = \frac{GHI}{Gc}. \qquad \text{II.8}$$

Clear sky conditions are defined for $\mathrm{K_c} > 0.65$ or $\mathrm{K_t} > 0.7$; cloudy sky conditions are defined for $0.4 < \mathrm{K_c} < 0.65$ or $0.2 < \mathrm{K_t} < 0.7$ and overcast conditions for $\mathrm{K_c} < 0.4$ or $\mathrm{K_t} < 0.2$ (Aryaputera et al., 2015; Yousif et al., 2013). There is no consensus whether to use clear sky index or clearness index (Smith et al., 2017). The clear sky index allows the removal of diurnal and seasonal signals from a given set of radiation data (Langella et al., 2016), besides it is less dependent to the air mass than the clearness index. Nevertheless, the clear-sky index subject to many errors due to the clear sky model input parameters.

I.1.j Albedo, reflectance and reflectivity

The monochromatic bidirectional reflectance is defined as the ratio of the radiant flux reflected by a surface to the incoming radiant flux (Schaepman-Strub et al., 2006). It is a dimensionless number and is expressed mathematically as follows:

$$\rho(\mathrm{T}, \lambda, \theta_\mathrm{i}, \phi_\mathrm{i}, \theta_\mathrm{r}, \phi_\mathrm{r}) = \frac{\mathrm{dL_r}(\mathrm{T}, \lambda, \theta_\mathrm{r}, \phi_\mathrm{r})}{\mathrm{dL_i}(\mathrm{T}, \lambda, \theta_\mathrm{i}, \phi_\mathrm{i})}. \qquad \text{II.9}$$

Where θ_i and ϕ_i give the direction of the incoming radiance dL_i; θ_r and ϕ_r the direction of the reflected radiance dL_r. T is the medium temperature. According to the International Commission on Illumination CIE (2011), reflectivity is distinguished from reflectance by the fact that reflectivity is a value that applies to thick reflecting objects.

The albedo is the ratio of the radiant flux reflected from a unit surface area into the whole hemisphere to the incident radiant flux of hemispherical angular extent. Consequently, in Eq.II.9, the albedo is defined when $\theta_r = \phi_r = 2\pi$ (Schaepman-Strub et al., 2006).

When the term surface reflectance is used for remote sensing it implies that narrow waveband or single wavelength is considered. Surface albedo, on the other hand, is used as a substitute for the integrated hemispherical albedo, of several wave-lengths (Duguay and LeDrew, 1992).

I.1.k Absorptance, absorption coefficient, absorptivity

The absorptance is defined as the ratio of the radiant flux absorbed by a surface to the incoming radiant flux. It is a dimensionless number and is expressed mathematically as follows:

$$\alpha(T, \lambda, \theta_i, \phi_i) = \frac{dL_a(T, \lambda, \theta_i, \phi_i)}{dL_i(T, \lambda, \theta_i, \phi_i)}. \quad \text{II.10}$$

According to Boulet et al., (2015) absorptance and absorption coefficients are inter changeable; yet absorptance refers to volumetric properties whereas absorptivity is used for surface properties. However, as showed by the literature review of Hu et al.(2002) there are still ambiguities in the usage of these optical terms.

I.1.l The emissivity

The emissivity is defined as the ratio of the ratio of the radiance of a specific object or surface to that of a standard black body. It is a dimensionless number and is expressed mathematically as follows:

$$\varepsilon_e(T, \lambda, \theta, \phi) = \frac{dL(T, \lambda, \theta, \phi)}{dL^0(T, \lambda)}. \quad \text{II.11}$$

I.1.m Scattering

Scattering (diffuse reflection) similarly to reflection redirects an incoming radiation stream. Contrary to the reflection, there is a total absorption and emission of the incoming particle or photon. Scattering phenomenon happens when the surface is rough relative to wavelengths whereas reflection happens when the surface is smooth. To characterize the angular distribution of the scattered particle or photon the phase function p is used. It describes photons with initial direction $\overrightarrow{\Omega'}(\mu', \phi')$ that follow the direction $\vec{\Omega}(\mu, \phi)$ after single or multiple scattering such that Eq. II.12 is verified.

$$\int_{4\pi} p\left(\lambda, \vec{\Omega}, \overrightarrow{\Omega'}, t\right) d\Omega' = 1. \quad \text{II.12}$$

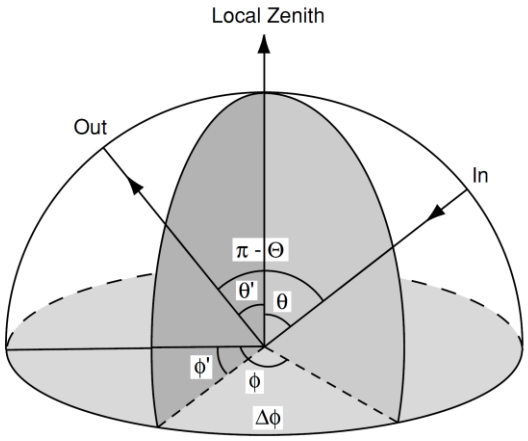

Fig. II.5 schematic representation of scattering angle for an incoming ray (In) in the direction $\vec{\Omega}(\mu, \phi)$that is deviated in the direction $\overrightarrow{\Omega'}(\mu', \phi')$ (Liou, 2002)

For spherical geometry, the scattering angle is related to the incoming and outgoing directions following Eq.II.13 as shown Fig. II.5

$$\cos\Theta = \mu\mu' + (1-\mu^2)^{\frac{1}{2}}\,(1-\mu'^2)^{\frac{1}{2}}\cos(\phi'-\phi). \qquad \text{II.13}$$

I.1.n The transmittance and transmissivity

The transmittance or transmissivity T is the ratio of the radiant flux directly transmitted after passing through a medium (atmosphere) to the amount of radiant flux that would have passed the same distance through a vacuum. It is the amount of light that remains after the absorption and scattering by the media. T is related to the absorptance and reflectance as follows:

$$T + \rho + \alpha = 1. \qquad \text{II.14}$$

I.2 Modeling the extraterrestrial irradiance

The top of atmosphere irradiance (G_{TOA}) is a function of astronomical geometric parameters (Rigollier et al., 2000), it is defined Eq.II.5. Fig. II.6 shows the spectral distribution of the top of atmosphere irradiance.

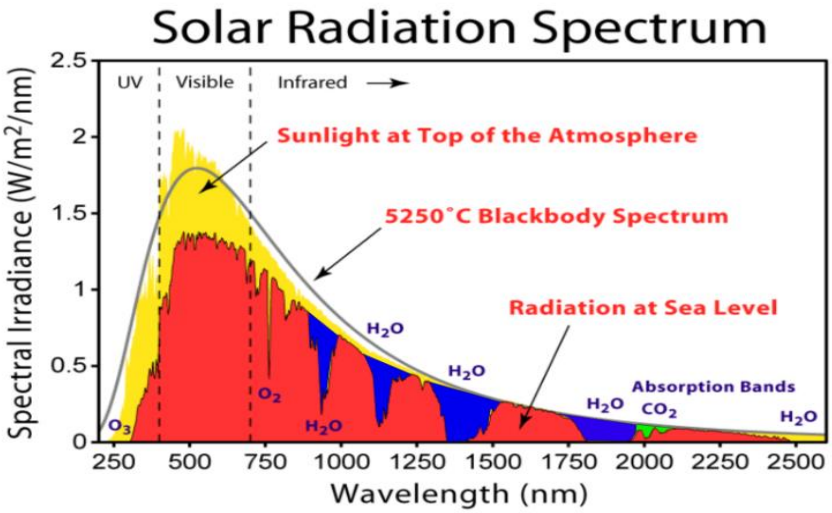

Fig. II.6 Solar radiation spectrum credit geosciencebigpicture

The accumulation over time of the top of atmosphere irradiance is the top of atmosphere irradiation. Its expression can be found below:

$$G_{TOA,t} = \int_{sunrise}^{sunset} G_{TOA}\, dt. \qquad \text{II.15}$$

Firstly, substituting the solar zenith angle expression Eq.II.2 and II.5 into Eq.II.15; secondly, expressing the time in function of the hour angle and earth angular velocity; lastly, assuming that the variation of the earth-sun distance in one day can be neglected yields (Liou, 2002):

$$G_{TOA,t} \cong \frac{Io\ \varepsilon}{\pi}\,(\sin\varphi \sin\delta\, H + \cos\varphi \cos\delta \sin H), \qquad \text{II.16}$$

where *H* represents a half-day expressed in radians. $G_{TOA,t}$ is expressed in J/m².

I.3 Instruments for measuring the GHI and quality control procedures

In manuscript only the global horizontal (GHI) component of the solar irradiance is studied. We do not consider the direct normal (DNI) and the diffuse horizontal parts (DHI) because the French weather services in French Guiana do not measure them.

I.3.a Instruments for measuring the GHI

The GHI is measured by two types of radiometers: thermopile and photodiode (e.g photoelectric or silicon) pyranometers (Driesse et al., 2016; Kleissl, 2013; Myers, 2013; Sengupta et al., 2015). They differ from each other by the type of sensor they use to infer the amount solar irradiance received.

Fig. II.7 thermopile (left) and photodiode (right) pyranometer. Source Wikipedia, Licor

Thermopile pyranometers measure indirectly the irradiance by quantifying the temperature difference between two surfaces: one absorbing the other non-absorbing. These surfaces are in contact with two metals with different heat capacities called thermocouple. When the absorbing surface of the thermopile is heated by the incident solar irradiance it creates a thermal flux upon the thermocouple junction. Therefore, this junction produces a voltage proportional to the difference in temperature (Badescu, 2008; Myers, 2013). A commonly used combination is the T type thermocouple, consisting of junctions of copper and constantan; it produces about 40 microvolts per degree centigrade.

Photodiode pyranometers measure indirectly the irradiance by quantifying the photocurrent generated in response to an incident irradiance flux. Photodiode pyranometers are less expensive and have a lower response time than thermopile pyranometers. The response time of thermopile pyranometers ranges between 1 and 240s (Sen, 2008) whereas the response time of photodiode pyranometers is approximately 10µs (Patil et al., 2013). Nevertheless, photodiode pyranometers are less accurate than most thermopile pyranometers (Badescu, 2008). A large part of photodiode pyranometers inaccuracies are related to their narrow spectral responses and or related to calibration issues (Kleissl, 2013). Thermopile pyranometers are sensitive to the whole shortwave spectrum whereas photodiode pyranometers have a narrow spectral sensitivity range: 350–1000 nm (Badescu, 2008). The most common solid-state detector used on solar photodiodes is crystalline silicon (Badescu, 2008; Myers, 2013).

According to Kleissl (2013), Younes et al. (2005), Myers (2013) and Sengupta et al. (2015) the measurement uncertainty of pyranometers are associated with: the calibration of the instrumentation, the data acquisition equipment, and the data processing step. The uncertainties related to the calibration process include the pyranometer spectral response, temperature response, non-linearity, aging and longwave radiation errors. The uncertainties related to the data acquisition include the error associated with the measurement of the sensor signals and environmental influences such as the cosine, azimuth and zenith response. The Uncertainties due to data processing results for instance from the average of sub-hourly irradiance measurements to hourly measurements.

Based on these criteria, the World Meteorological Organization (WMO) classifies pyranometers into three categories based on their accuracies: secondary standard, first class and second class. Tab. II.1 shows the characteristic of each category.

Tab. II.1 WMO classification of pyranometers, extracted from (Sen, 2008)

Characteristics	Secondary standard	First class	Second class
Resolution (smallest detectable change in W/m²)	±1	±5	±10
Stability (percentage of full scale, change per year)	±1	±2	±5
Cosine response (% deviation from ideal at 10 ° elevation on a clear day)	±3	±7	±15
Azimuth response (% deviation from ideal at 10 ° elevation on a clear day)	±3	±5	±10
Temperature response (% maximum error due to change of ambient temperature within the operating range)	±1	±2	±5
Non linearity (% of full scale)	±0.5	±2	±5
Spectral sensitivity (% deviation from mean absorbance 0.3 -3µm)	±2	±5	±10
Response time (99% response)	<25s	<60s	<240s

In French Guiana, French weather services use a First class pyranometers. They are thermopile of brand Kipp & zonen type CM6B.

I.3.b Quality control procedures

Before comparing global NWP GHI forecasts to ground measurements, one should quality control (QC, ore quality check) the ground data used as reference. QC procedures aim to detect faulty measurements related to calibration of the instrumentation, data acquisition equipment, and or data processing. QC procedures can be classified into four categories (Espinar et al., 2011):

-QC based on redundancy
-QC based on limits within range
-step check QC
-and consistency QC.

QC based on redundancy use two instruments for measuring the same meteorological variable. These measurements from two different instruments are then compared and should be equal within the coupled uncertainty of both instruments. However, this approach is expensive; measurement sites do not often have several instruments that measure simultaneously the GHI. QC based on limits within a range makes sure that GHI measurements comply to physical limits and statistical knowledge. Statistical knowledge allows accounting for rare observation cases (Espinar et al., 2011). Step checks QC aims to detect unrealistic jumps in values from two consecutive GHI measurements or stagnation within a given time interval. Consistency QC consists in measuring simultaneously at the same location the three components of the solar radiation, i.e. GHI, DNI, DHI (Eq.II.6) and verifying their inner consistencies. To achieve maximum benefit the WMO combining these QC procedure starting by the QC based on limits within a range then step and consistency check (WMO, 2007).

In French Guiana the GHI is measured by French weather services (Météo-France) every 15 min, then accumulated to hourly values. These hourly values are then stored and released to the users. As sub-hourly measurements are not available in 2016, this QC procedure is not applicable. Besides, only the GHI is measured and only one pyranometer is available in each measurement sites; therefore, the redundancy and consistency QC cannot be applied.

Tab. II.2 shows the different types of limits within range QC procedures found in the literature. Lower and higher limit values apply to daylight time: between sunrise and sunset. This table shows that Espinar et al. (2011) and Helioclim (Younes et al., 2005) are the most constrained QC. They share the same lowest GHI value; however, their highest observable values differ. The lowest value was set up according to the analysis of collected data and clearness index value from the European Solar Radiation Atlas for solar altitude greater than 2° (Geiger et al., 2002). Helioclim (Younes et al., 2005) maximum observable value involves the use of a clear sky model (Rigollier et al., 2000). Espinar et al. (2011) higher limit is defined as the minimum between Baseline Surface Radiation Network (BSRN) (Long and Dutton, 2002) and Muneer and Fairooz (2002) based limits. It has the advantage of accounting for rare observations using a constant.

Tab. II.2 literature review of limits within range QC procedure

QC	Minimum range	Maximum range
(Espinar et al., 2011)	$0.03\ GHI_{TOA}$	$\min(1.2\ I_0, 1.5\ I_0\ \cos(\theta_z)^{1.2} + 100)$
Helioclim (Younes et al., 2005)	$0.03\ GHI_{TOA}$	$\min(1.1.\ Gc, GHI_{TOA})$
CIE (Younes et al., 2005)	0	$1.2\ I_0$
Muneer (Muneer and Fairooz, 2002)	0	$\min(1.2\ I_0, GHI_{TOA})$
BSRN (Long and Dutton, 2002)	-4	$1.5\ I_0\ \cos(\theta_z)^{1.2} + 100$
WMO (WMO, 2007)	0	1600
Younes (Younes et al., 2005)	none	Gc

Most QC procedures shown Tab. II.2 have restrictions on the minimum solar elevation angle they should be used for. Espinar et al. (2011) and Helioclim (Younes et al., 2005) QC are valid for solar altitude greater than 2° (Geiger et al., 2002). CIE (Younes et al., 2005), Muneer and Fairooz (2002) QC are valid only for solar elevation greater than 4°. Younes (Younes et al., 2005) QC is the most restrictive, it assumes that the solar elevation is greater than 7°. Restrictions on the solar elevation aim to limit the cosine effect error at sunrise and sunset. The more acute the angle of the sun (at sunrise and sunset), the greater the error associated with the cosine effect is. Besides, at low sunrise and sunset angle the change of GHI value might be too small to be detectable change by pyranometers (Sen, 2008).

When disproportional number of observations are rejected by the QC it may either be an indication of strongly biased data that should not be considered as outliers or an indication that the QC is not adequate and should be further studied and possibly modified (Espinar et al., 2011).

I.4 Modeling the solar irradiance depletion due to the atmospheric constituent: the radiative transfer equation

The following equation is the radiative transfer equation (RTE); it describes the extinction and generation process of radiance through the atmosphere path (Liou, 2002; Qu, 2013).

$$\frac{1}{c}\frac{\partial L_\lambda(\vec{r},\vec{\Omega},t)}{\partial t}+(\vec{\Omega}.\vec{\nabla})L_\lambda(\vec{r},\vec{\Omega},t)=-\beta_{e,\lambda}\,L_\lambda(\vec{r},\vec{\Omega},t)+\beta_{e,\lambda}\,J_\lambda(\vec{r},\vec{\Omega},t), \qquad \text{II.17}$$

where $L_\lambda(\vec{r},\Omega,\mathrm{t})$ is the monochromatic radiance in the position $\vec{r}$ from direction $\vec{\Omega}$ at a time t.

The first term of the left-hand side (lhs) describes the radiance time evolution within a domain whereas the second term of the lhs describes the radiance exchange through the domain boundaries. The first term of the right hand side (rhs) describes the extinction causes by the absorption and scattering (Eq.II.18) whereas the second term of the rhs describes the generation of radiance by source terms (Eq.II.19). The extinction coefficient $\beta_{e,\lambda}$ and radiance source term J_λ are expressed as follows:

$$\beta_{e,\lambda}=(\beta_{s,\lambda}+\beta_{a,\lambda})\ and \qquad \text{II.18}$$

$$\mathrm{J}_\lambda(\vec{\mathrm{r}},\vec{\Omega},\mathrm{t})=\frac{\beta_{\mathrm{s},\lambda}}{4\pi\,\beta_{\mathrm{e},\lambda}}\int_{\overrightarrow{\Omega'}}\mathrm{L}_\lambda(\vec{\mathrm{r}},\Omega,\mathrm{t})\;\mathrm{p}(\lambda,\vec{\Omega},\overrightarrow{\Omega'})\mathrm{d}\,\Omega'+\frac{\beta_{\mathrm{a},\lambda}}{\beta_{\mathrm{e},\lambda}}\,\mathrm{B}_\lambda(\vec{\mathrm{r}},\vec{\Omega},\mathrm{t}). \qquad \text{II.19}$$

Where $\beta_{s,\lambda}$ and $\beta_{a,\lambda}$ are the scattering and absorption coefficients respectively. The source term J_λ includes photons scattered inward the domain and photons emitted by the medium. In Eq.II.19, $B_\lambda(\vec{r},\Omega,\mathrm{t})$ is the radiant flux emitted by the atmosphere, mostly in the thermal infrared (Liou, 2002); $p\left(\lambda,\vec{\Omega},,\overrightarrow{\Omega'}\right)$ is the phase function defined Eq.II.12.

The radiance stationarity approximation is often made to simplify the RTE (Liou, 2002; Qu, 2013). Applying these assumptions to Eq.II.17 yields:

$$\frac{(\vec{\Omega}.\vec{\nabla})\mathrm{L}_\lambda(\vec{\mathrm{r}},\vec{\Omega})}{\beta_{\mathrm{e},\lambda}}=-\mathrm{L}_\lambda(\vec{\mathrm{r}},\vec{\Omega})+\frac{\omega_\lambda}{4\pi}\int_{\overrightarrow{\Omega'}}\mathrm{L}_\lambda(\vec{\mathrm{r}},\vec{\Omega})\;\mathrm{p}(\lambda,\vec{\Omega},\overrightarrow{\Omega'})\mathrm{d}\,\Omega'+(1-\omega_\lambda)\mathrm{B}_\lambda(\vec{\mathrm{r}},\vec{\Omega}), \qquad \text{II.20}$$

where, $\omega_\lambda=\frac{\beta_{\mathrm{s},\lambda}}{\beta_{\mathrm{e},\lambda}}$ is the single scattering albedo. The scattering term may be written as a sum of single scattering and multiple scattering terms.

$$\frac{\omega_\lambda}{4\pi}\int_{\overrightarrow{\Omega'}}\mathrm{L}_\lambda(\vec{\mathrm{r}},\vec{\Omega})\;\mathrm{p}\left(\lambda,\vec{\Omega},\overrightarrow{\Omega'}\right)\mathrm{d}\,\Omega' \qquad \text{II.21}$$
$$=\frac{\omega_\lambda}{4\pi}\,\mathrm{Fo}\;\mathrm{p}\left(\lambda,\vec{\Omega}_0,\overrightarrow{\Omega'}\right)\mathrm{e}^{-\tau_\lambda}+\frac{\omega_\lambda}{4\pi}\int_{\overrightarrow{\Omega'}}\mathrm{L}_\lambda\left(\vec{\mathrm{r}},\overrightarrow{\Omega'}\right)\;\mathrm{p}\left(\lambda,\vec{\Omega},\overrightarrow{\Omega'}\right)\mathrm{d}\,\Omega',$$

where $\vec{\Omega}_0$ shown Fig. II.9 gives the direction of incoming TOA radiation Fo (Fig. II.9; Eq.II.1).For computations speed purposes the atmosphere is considered as plane-parallel in localized portions (Liou, 2002). The plane-parallel atmosphere hypothesis considers the atmosphere as a sum of vertical atmospheric layers. Each layer is characterized by homogeneous properties and bordered by the bottom and top infinite plates called boundaries.

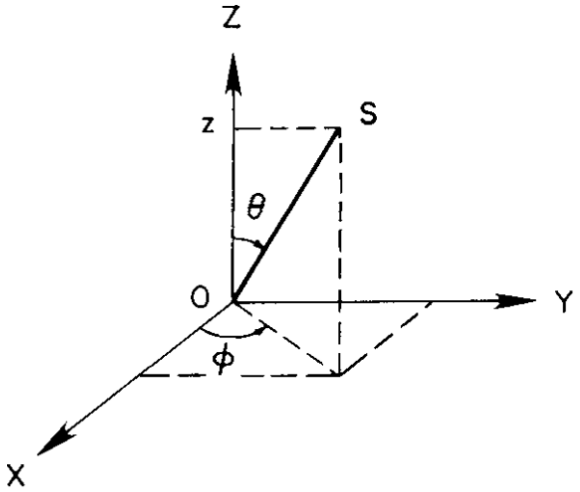

Fig. II.8 Plane Parallel atmosphere

Applying these assumptions to Eq.II.20 and assuming the angular symmetry θ of the radiance yields (Fouquart et al., 1991; Liou, 2002; Qu, 2013):

$$\mu\frac{dL_\lambda(\tau,\mu,\phi)}{d\tau} = L_\lambda(\tau,\mu,\phi) - \frac{\omega_{\lambda(\tau)}}{4\pi} F_0\, p(\tau,\mu_0,\phi_0,\mu',\phi')e^{\frac{-\tau_\lambda}{\mu}} - \frac{\omega_{\lambda(\tau)}}{4\pi}\int_0^{2\pi}\int_{-1}^{1} L_\lambda(\tau,\mu',\phi')\, p(\tau,\mu,\phi,\mu',\phi')\, d\mu' d\phi' + \big(1-\omega_\lambda(\tau)\big)B_\lambda(\tau,\mu), \quad \text{II.22}$$

where μ, $\mu'$and μ_0 angles are defined Fig. II.9. τ is the optical depth, it is expressed as follows:

$$\tau(z) = \int_{z_1}^{z_2} -\beta_{e,\lambda}\, dz. \quad \text{II.23}$$

Eq.II.22 describes the effect of the optical constituents of the atmosphere (cloud water droplets, cloud ice crystals, water vapor, ozone, aerosols, carbon dioxide and other minor trace gases) on the transfer of radiance for each atmospheric layer $d\tau$.

A common approximation to Eq.II.22 is to consider that the azimuthal dependence of the radiance can be neglected; then we can define the azimuthally average radiance and phase function as follows (Liou, 2002):

$$L_\lambda(\tau,\mu) = \frac{1}{2\pi}\int_0^{2\pi} L_\lambda(\tau,\mu,\phi)d\phi \;\; and \quad \text{II.24}$$

$$p(\tau,\mu,\mu') = \frac{1}{2\pi}\int_0^{2\pi} p(\tau,\mu,\phi,\mu',\phi')d\phi'. \quad \text{II.25}$$

For simplicity Eq. II.22 can be split in two uncoupled equations, one that describes the beam component of the radiation $L_{\lambda,b(\tau,\mu)}$ and the other the diffuse component $L_{\lambda,d(\tau,\mu)}$. They are written as follows (Liou, 2002):

$$\mu \frac{dL_{\lambda,b}(\tau,\mu)}{d\tau} = L_{\lambda,b}(\tau,\mu) \; and \qquad \text{II.26}$$

$$\mu \frac{dL_{\lambda,d}(\tau,\mu)}{d\tau} = L_{\lambda,d}(\tau,\mu) - \frac{\omega_{\lambda(\tau)}}{2} F_0 \, p(\tau,\mu_0,\mu) e^{\frac{-\tau_\lambda}{\mu}} - \frac{\omega_{\lambda(\tau)}}{2} \int_{-1}^{1} L_{\lambda,d}(\tau,\mu) p(\tau,\mu,\mu') \, d\mu' + \big(1-\omega_\lambda(\tau)\big) B_\lambda(\tau,\mu). \qquad \text{II.27}$$

The beam component equation can be solved analytically; its solution is written as follows:

$$L_{\lambda,b}(\tau) = \mu_0 I_0 \, e^{-\tau/\mu_0}, \qquad \text{II.28}$$

where L_0 is the TOA radiance and μ_0 is the azimuth angle of the incoming solar radiation.

A further simplification made to solve the diffuse component of RTE Eq.II.27 is to separate the contributions from shortwave and the contribution from longwave radiations. The overlap between these ranges is relatively small which allows treating the two types of radiative transfer and their source functions separately. Stephens (1984) indicates that Rayleigh scatter is dominant only for the shorter wavelengths, while liquid water absorption in clouds occurs only for the longer wavelengths. The transfer of shortwave radiation is not as complex as longwave radiation, since the problem of the simultaneous absorption and emission of radiation from layer to layer does not occur (Stensrud, 2009). These hypotheses simplify the complexity of the radiative transfer problem (Liou, 2002); for shortwave the thermal emission B_λ can be neglected since scattering process is dominant (Stephens, 1984). The following subsection introduces the different method to solve the RTE. Since the shortwave radiation is of interest for solar PV because it generates significant PV power (Kleissl, 2013), only the method dealing with downward shortwave radiation modeling is discussed in this manuscript.

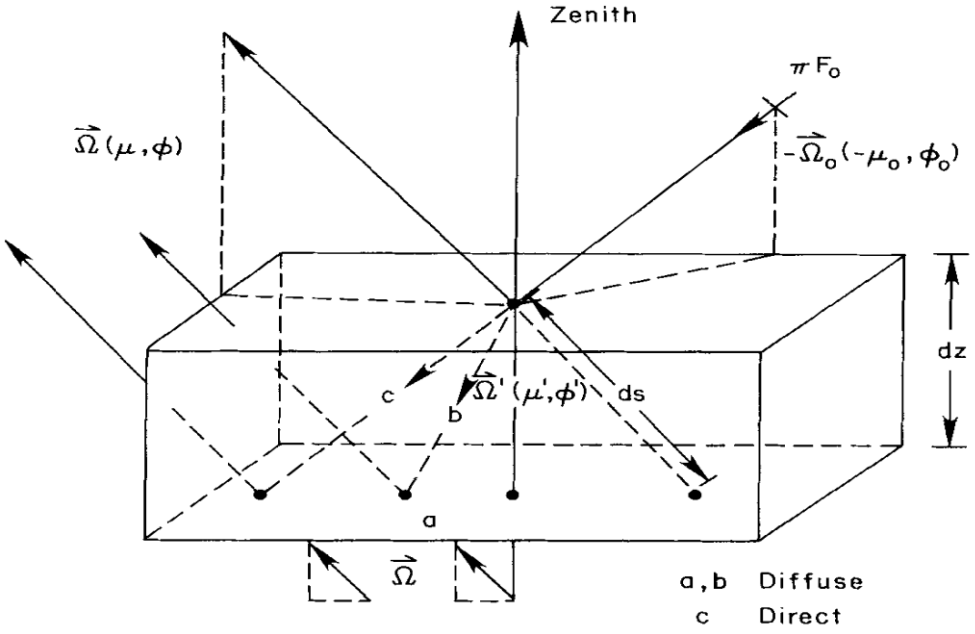

Fig. II.9 Transfer of solar radiation in plane-parallel layers, illustrating attenuation by extinction, a; multiple scattering, b; and single scattering of the unscattered solar flux, c. Source (Liou, 2002)

I.5 Input datasets for radiative transfer models

To solve the RTE Eq.II.26, II.27; a knowledge on the optical state of the atmosphere is required. Tab. II.3 shows the input datasets commonly required by radiative transfer codes (ECMWF, 2016a; Kneizys et al., 1996; Mayer and Kylling, 2005; Qu, 2013; Ricchiazzi et al., 1998; Stensrud, 2009; Vermote et al., 1997).

Tab. II.3 input dataset for radiative transfer code Source (ECMWF, 2016a; Stensrud, 2009; Stephens, 1984)

Variable	Source
Vertical profile of temperature	- NWP model forecasts - Or climatology from measurements campaigns (ANDERSON et al., 1986)
Vertical profile of pressure	- NWP model forecasts - Or climatology from measurements campaigns (ANDERSON et al., 1986)
Vertical profile of water vapor mixing ratio	- NWP model forecasts - Or climatology from measurements campaigns (ANDERSON et al., 1986)
Gas concentrations (Carbon dioxide, ozone and trace gases)	- Climatology from chemistry transport model and satellite observations (Inness et al., 2013)
Ground albedo	-Climatology from satellites measurements (Schaaf et al., 2002)
Ground emissivity	-Climatology value (ECMWF, 2016a)
Aerosol : Number concentration optical thickness	-Climatology from satellites measurements (Nabat et al., 2013; Zubler et al., 2011) - Climatology from chemistry transport models (Tegen et al., 1997)
Cloud properties: content (liquid, ice, snow, water), optical thickness (τ)	-Climatological value from satellites observations (Rossow and Schiffer, 1999) -or diagnosed from variable forecasted by NWP models (Bengtsson et al., 2017; Slingo, 1989; Stephens et al., 1990)
Effective radius of cloud particles	-Climatological value from satellites observations (Rossow and Schiffer, 1999) -or diagnosed from variable forecasted by NWP models (Bengtsson et al., 2017; ECMWF, 2016a; Troccoli and Morcrette, 2014)
Single scattering albedo (wo)	- approximated as a function the effective radius of particles sizes (Stephens, 1984)
Asymmetry parameter (g)	-approximated as a function the effective radius of particles sizes (Hansen and Travis, 1974)
Solar zenith angle	-Computed from trigonometric formulas (Liou, 2002)
Altitude	-From digital elevation model (USGS, 2018)

Climatological datasets shown Tab. II.3 are primarily inferred from satellite observations, atmospheric sounding and ground measurements that are combined using regressions (ANDERSON et al., 1986; Cionni et al., 2011). According to Cionni et al. (2011) this procedure introduces biases. For unobserved or sparsely observed variable multi-year simulations of global numerical weather prediction model may be used (Chin et al., 2000; Cionni et al., 2011).

These input datasets have different coverage, uneven and sometimes coarse resolutions. We found in the literature (ANDERSON et al., 1986) that vertical profiles of temperature, pressure and water vapor mixing ratio from climatology have a typical resolution of 1 km. NWP model have higher spatial and temporal resolution in; for Integrated Forecast System (IFS) the spatial resolution ranges between 20 and 6000 m (ECMWF, 2018). Nevertheless, the forecast evaluation of IFS (Richardson et al., 2013; Haiden et al., 2014, 2015, 2016) shows that its accuracy varies with the time of the year and the considered location. Besides the aerosol dataset currently used in a large number of operational center (ECMWF, 2016a; NOAA, 2016; Termonia et al., 2018) has a 5° by 4° spatial resolution and monthly temporal resolution. According to Nabat et al. (2013) the coarse spatial resolution of this dataset misrepresents the regional aerosol loads; this misrepresentation of aerosol loads lead to inaccurate irradiance estimates (Jimenez et al., 2015; Ruiz-Arias et al., 2013; Zhong et al., 2016). Consequently, to solve the RTE, lots of input datasets are required; these datasets despite being available worldwide have a coarse resolution and or rely on interpolation methods which misrepresent the regional variability and cause inaccurate assessment of the GHI.

I.6 Accurate Methods to solve the radiative transfer equation for shortwaves

Accurate methods to solve the RTE Eq.II.27 we found in the are the discrete ordinate (disort) and adding methods. However, these methods are not the preferred choice to assess the GHI operationally because they are computationally expensive and require a lot of input data. They are used in some NWP model to assess the GHI; however, the RTE is not solved for every model time step and or is solved in coarser grid than the NWP model resolution (ECMWF, 2016a; Ruiz-Arias et al., 2013).

Disort method uses Eq.II.27 as a starting point; the integral over all the directions μ is replaced by a summation over a finite number of quadrature points N (Liou, 2002). N is typically equal to 16 (Yang et al., 2016). Using Gaussian quadrature to approximate Eq. II.27 yields:

$$\mu_i \frac{dL_{\lambda,d}(\tau,\mu_i)}{d\tau} = L_{\lambda,d}(\tau,\mu_i) - \frac{\omega_{\lambda(\tau)}}{2} F_0\, p(\tau,\mu_0,\mu_i) e^{\frac{-\tau_\lambda}{\mu_0}} - \frac{\omega_{\lambda(\tau)}}{2} \sum_{j=-N}^{N} L_{\lambda,d}(\tau,\mu_i) a_j p(\tau,\mu_i,\mu_j), \quad \text{II.29}$$

where a_j are the Gaussian quadrature points.

The second level of approximation in the disort method consists in expanding the radiance and phase function with Fourier cosine series. In case both are independent of azimuth Eq.II.24-II.25, they are written as follows:

$$L_{\lambda,d}(\tau,\mu_i) = \sum_{m=0}^{N} L^m_{\lambda,d}(\tau,\mu_i)\ and \quad \text{II.30}$$

$$p(\tau,\mu_i,\mu_j) = \sum_{m=0}^{N} \sum_{l=m}^{N} \omega_l^m p_l^m(\tau,\mu_i) p_l^m(\tau,\mu_j). \quad \text{II.31}$$

Substituting Eq.II.30 and Eq.II.31 into Eq.II.29 yields a system of 2N coupled equations with non-constant coefficients. These equations describe the upwelling (Eq.II.32) and downwelling (Eq.II.33) radiance:

$$\mu_i \frac{dL^m_{\lambda,d}(\tau,\mu_i)}{d\tau} = L^m_{\lambda,d}(\tau,\mu_i) - \frac{\omega_{\lambda(\tau)}}{2} F_0 \sum_{m=0}^{N} \sum_{l=m}^{N} \omega_l^m p_l^m(\tau,\mu_i) p_l^m(\tau,\mu_0)\, e^{\frac{-\tau_\lambda}{\mu_0}} - \frac{\omega_{\lambda(\tau)}}{2} \sum_{j=-N}^{N} L^m_{\lambda,d}(\tau,\mu_i) a_j \sum_{m=0}^{N} \sum_{l=m}^{N} \omega_l^m p_l^m(\tau,\mu_i) p_l^m(\tau,\mu_j) \quad and \qquad \text{II.32}$$

$$-\mu_i \frac{dL^m_{\lambda,d}(\tau,-\mu_i)}{d\tau} = L^m_{\lambda,d}(\tau,-\mu_i) - \frac{\omega_{\lambda(\tau)}}{2} F_0 \sum_{m=0}^{N} \sum_{l=m}^{N} \omega_l^m p_l^m(\tau,-\mu_i) p_l^m(\tau,\mu_0)\, e^{\frac{-\tau_\lambda}{\mu_0}} - \frac{\omega_{\lambda(\tau)}}{2} \sum_{j=-N}^{N} L^m_{\lambda,d}(\tau,-\mu_i) a_j \sum_{m=0}^{N} \sum_{l=m}^{N} \omega_l^m p_l^m(\tau,-\mu_i) p_l^m(\tau,\mu_j). \qquad \text{II.33}$$

Eq.II.32 and II.33 assume that the medium consists of several adjacent homogeneous layers in which the single-scattering albedo and phase function are taken to be constant within each layer but allowed to vary from layer to layer. For large numbers of *N*, this system of equation is solved using numerical methods (Campbell, 1969; Stamnes et al., 1988; Stamnes and Conklin, 1984).

Adding is an accurate method that allows calculating the shortwave radiative flux across multiple vertical levels. The underlying idea of the adding method is that knowing the reflection and transmission of two individual layers, one can deduce the reflection and transmission of the combined layer by calculating the successive reflections and transmissions between these two layers (Stensrud, 2009). Accounting for multiple reflections of the light beam in the two layers, as shown in Fig. II.10, the combined reflection and transmission functions are given by:

$$\begin{cases} R_{12} = R_1 + \widetilde{T}_1{}^{*} R_1 \widetilde{T}_1 + \widetilde{T}_1{}^{*} R_2\, R_1{}^{*} R_2 \widetilde{T}_1 + \widetilde{T}_1{}^{*} R_2\, R_1{}^{*} R_2\, R_1{}^{*} R_2 \widetilde{T}_1 + \cdots \\ \widetilde{T}_{12} = \widetilde{T}_2 \widetilde{T}_1 + \widetilde{T}_2\, R_1{}^{*} R_2 \widetilde{T}_1 + \widetilde{T}_2\, R_1{}^{*} R_2\, R_1{}^{*} R_2 \widetilde{T}_1 + \cdots \\ U = R_2 \widetilde{T}_1 + R_2\, R_1{}^{*} R_2 \widetilde{T}_1 + R_2\, R_1{}^{*} R_2\, R_1{}^{*} R_2 \widetilde{T}_1 \\ \widetilde{D} = \widetilde{T}_1 + R_1{}^{*} R_2 \widetilde{T}_1 + R_1{}^{*} R_2\, R_1{}^{*} R_2 \widetilde{T}_1 + \cdots \end{cases} \qquad \text{II.34}$$

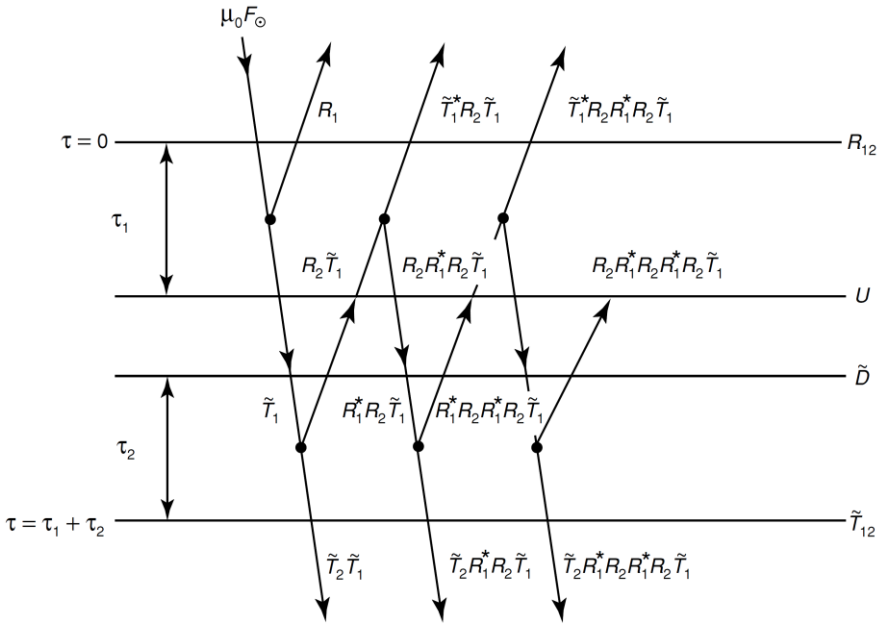

Fig. II.10 Configuration of the adding method for two layers of optical depth τ_1 and $\tau_2$respectively. The upward transmission and reflection of the i^{th} layer are $\tilde{T}_i^*$ and R_i, respectively whereas the downward transmission and reflection are $\tilde{T}_i$ and R_i^*. $\tilde{D}$ and U are the the combined total transmission and reflection functions between layers 1 and 2. Source (Liou, 2002)

In Eq.II.34 the infinite series can be replaced by a single inverse function such that:

$$\begin{cases} \sum_{n\geq 0} (R_1{}^* R_2)^n = \dfrac{1}{1 - R_1{}^* R_2} = 1 + S, with \\ S = \dfrac{R_2\, R_1{}^*}{(1 - R_2\, R_1{}^*)}. \end{cases} \qquad \text{II.35}$$

As seen in Eq.II.26-II.27 the diffuse and beam component of the radiance can be treated separately. We can similarly separate the diffuse and direct components of the total transmission as follows:

$$\tilde{\mathrm{T}}_\mathrm{i} = \mathrm{T}_\mathrm{i} + \mathrm{e}^{\frac{-\tau_\mathrm{i}}{\mu'}}, \qquad \text{II.36}$$

where $\mu' = \mu_0$ when transmission is associated with the incident solar beam and $\mu' = \mu$ when it is associated with the emergent light beam in the direction μ.

On the basis of Eq.II.34, II.35 and II.36, a set of iterative equations for the computation of diffuse transmission and reflection involving the two layers may be written as follows (Liou, 2002):

$$\begin{cases} \mathrm{Q} = \mathrm{R_1{}^* R_2} \\ \mathrm{S} = \mathrm{Q}(1 - \mathrm{Q})^{-1} \\ \mathrm{D} = \mathrm{T_1} + \mathrm{ST_1} + \mathrm{S}\mathrm{e}^{\frac{-\tau_1}{\mu_0}} \\ \mathrm{U} = \mathrm{R_2 D} + \mathrm{R_2}\mathrm{e}^{\frac{-\tau_1}{\mu_0}} \\ \mathrm{R_{12}} = \mathrm{R_1} + \mathrm{U}\mathrm{e}^{\frac{-\tau_1}{\mu_0}} + \mathrm{T_1^* U} \\ \mathrm{T_{12}} = \mathrm{D}\mathrm{e}^{\frac{-\tau_2}{\mu_0}} + \mathrm{T_2}\mathrm{e}^{\frac{-\tau_1}{\mu_0}} + \mathrm{T_2 D} \end{cases} \qquad \text{II.37}$$

In Eq.II.37 the product of any two parameters implies that integration over the solid angle (Fig. II.4) is to be performed so as to take into account all the possible multiple scattering contributions. Therefore, the product of *A* and *B* where they can be any of the parameters *R,*T, *U,* and *D* refers to the following equation:

$$A_1 B_2 = 2\int_0^1 A(\tau_1, \mu, \mu')B(\tau_2, \mu, \mu_0)\mu' d\mu'. \qquad \text{II.38}$$

In practice, one may begin with the computations for initial layers with small optical depths $\Delta\tau \approx 10^{-8}$ so that the single scattering approximation is sufficiently accurate. Using the single scattering approximation one can find an analytical solution to Eq.II.22; this solution is written as follows:

$$L_\lambda(\tau, \mu) = L(\tau_1, \mu)e^{-\frac{(\tau_1-\tau)}{\mu}} + \frac{\omega}{4}F_0 p(\mu, -\mu_0)\int_\tau^{\tau_1} e^{\left(-\frac{(\tau'-\tau)}{\mu}+\frac{\tau'}{\mu_0}\right)}\frac{d\tau'}{\mu} \text{ and} \qquad \text{II.39}$$

$$L_\lambda(\tau, -\mu) = L(0, -\mu)e^{-\frac{-\tau}{\mu}} + \frac{\omega}{4}F_0 p(\mu, -\mu_0)\int_\tau^{\tau_1} e^{\left(-\frac{(\tau'-\tau)}{\mu}+\frac{\tau'}{\mu_0}\right)}\frac{d\tau'}{\mu}. \qquad \text{II.40}$$

Where $L(\tau, \mu)$is the upward (reflected) and downward (transmitted) radiance for a finite atmosphere bounded on two sides at $\tau = 0$ and $\tau = \tau_1$. Eq.II.37 are then used to compute the reflection and transmission functions for an optical depth of $2\Delta\tau$. This procedure may be repeated until the desirable optical depth is reached (Liou, 2002).

Adding and disort methods that are described above assume a plane parallel homogeneous atmosphere; therefore, they cannot describe accurately the horizontal variations of optical properties induced by clouds (Barker et al., 2008; Oreopoulos et al., 2012; RÄISÄNEN et al., 2005). According to Oreopoulos et al. (2012) neglecting the horizontal variation in cloud optical depth leads to sizable bias on the estimated solar flux. To solve this issue, the fraction of cloud in the layer is modeled using method such as the maximum overlap, random overlap and Monte Carlo independent assumption column. They allow addressing columns whose layers are only partially filled by clouds which, in turn, overlapped vertically (Barker et al., 2008).

To compute the broadband radiation calculations the radiance is first written in a flux form as follow (Pincus and Stevens, 2013):

$$F_\lambda(\tau) = \int_\mu L_\lambda(\tau, \mu)\,\mu\, d\mu. \qquad \text{II.41}$$

Eq.II.26-II.27 are then discretized using a quadrature formula written as follows:

$$F(\tau) = \sum_g^G b_g\, F_{\lambda,g}(\tau), \qquad \text{II.42}$$

where $F_{\lambda,g}$ is the individual flux in the spectral interval defined by *g* and b_g the quadrature points for the corresponding spectral interval. Tab. II.3 shows the spectral interval used in Rapid radiative transfer model RRTM (ECMWF, 2016a).

Tab. II.4 spectral interval choose in RRTM, extracted from (ECMWF, 2016a)

Spectral intervals cm^{-1}	Number of g-points	Gases included	
		Troposphere	Stratosphere
800-2600	12	H_2O	CO_2
2600-3250	6	H_2O, CH_4	
3250-4000	12	H_2O, CO_2	H_2O, CO_2
4000-4650	8	H_2O, CH_4	CH_4
4650-5150	8	H_2O, CO_2	CO_2
5150-6150	10	H_2O, CH_4	H_2O, CH_4
6150-7700	10	H_2O, CO_2	H_2O, CO_2
7700-8050	2	H_2O, O_2	O_2
8050-12850	10	H_2O	
12850-16000	8	H_2O, O_2	O_2
16000-22650	6	H_2O	
22650-29000	6		
29000-38000	8	O_3	O_3
38000-50000	6	O_2, O_3	O_2, O_3

I.7 Approximate methods to solve the radiative transfer equation for shortwaves

Radiative transfer parametrizations (RTPs) are simplified version of the sophisticated, rigorous radiative transfer models Eq.II.22 (Gueymard, 2012a). RTP are physical, semi-physical or statistical based methods. These methods aim to find the GHI value with a limited accuracy loss and faster than accurate methods. RTP are favored for operational purposes.

I.7.a Statistical RTP

Statistical RTPs are non-spectral model. They either use regression between irradiance and meteorological parameters; regression between irradiance and the digital count of satellite images or use historical measurements of irradiance to assess ground irradiance (Badescu, 2008; Noia et al., 1993a, 1993b). Angstrom (1924) pioneered solar radiation statistical modeling; he used a linear relation to relate GHI and sunshine hours (Badescu, 2008; Sen, 2008):

$$\frac{\mathrm{GHI}}{\mathrm{G_{TOA}}} = \mathrm{a} + \mathrm{b}\,\frac{\mathrm{S}}{\mathrm{S_0}}. \qquad \text{II.43}$$

Where S is the monthly average of daily bright sunshine hours; which is the total hours when the sunlight is higher than a specified threshold value; a and b are Angstrom coefficients. The ratio $\frac{S}{S_0}$ is related to the cloud amount n as demonstrated by Davies et al. (1984).

$$\frac{\mathrm{S}}{\mathrm{S_0}} = 1 - \mathrm{n}. \qquad \text{II.44}$$

Angstrom coefficients were derived using yearly GHI and sunshine duration measurements in Stockholm. A major difficulty of Angstrom (1924) relationship arises from the unicity of its coefficients: these coefficients depend on physical and spatial properties of the atmosphere at the region of interest (Badescu, 2008). Akinoğlu (1991) tabulated Angstrom coefficient for 100 locations

and noticed a wide spread of values. Later, Ögelman et al. (1984) suggested the addition of a nonlinear term to Angstrom (1924) formulation. This model was validated in Turkey by Akinoğlu and Ecevit (1990); results found showed an improvement of the GHI estimates compared to Angstrom (1924) original model.

However, regression analyses are limited in their accuracies and the number of variables they can process. Studies showed that GHI can be correlated to several parameters (Elizondo et al., 1994). Therefore, artificial neural networks were used to correlate GHI, sunshine and or meteorological variable because of their ability to deal with linear and nonlinear correlations between several variables (Zarzalejo et al., 2005). One of the pioneer of this research area is (Elizondo et al., 1994), he correlated the GHI to: daily precipitation, maximum and minimum daily temperature, day length, daily total. Several other parameters were also suggested as predictor: the location coordinate (Al-Alawi and Al-Hinai, 1998; Reddy and Ranjan, 2003; Sözen et al., 2004), the pressure (Al-Alawi and Al-Hinai, 1998; López et al., 2005), the relative humidity (Al-Alawi and Al-Hinai, 1998; López et al., 2005; Reddy and Ranjan, 2003), the wind speed (Al-Alawi and Al-Hinai, 1998; Reddy and Ranjan, 2003), the clear sky radiation (Elizondo et al., 1994), solar geometry angle (López et al., 2005; Zarzalejo et al., 2005), the clearness index (López et al., 2005), the relative air mass (López et al., 2005), the dew point temperature (López et al., 2005) and the precipitable water (López et al., 2005). According to Mao et al. (1999), the predictor choice is related to the geographical location. A major difficulty arises from the choice of relevant predictors. Having redundant predictors causes over learning which prevent the model to be applied generally to other datasets than the training dataset (Elizondo et al., 1994). López et al. (2005) suggested that the selection process should be done using an automatic relevance determination method (ARD).

The launch of the first satellite promoted the use of statistical model using satellite images as input (House et al., 1986; Levanon, 1971). Statistical satellite methods establish a regression between the digital counts measured by the satellites radiometers and simultaneous ground irradiance values measured at the ground (Noia et al., 1993a). Tarpley (1979); Cano et al. (1986); Hay and Hanson (1978) pioneered the statistical modeling of solar irradiance using satellites images. Their methods were regularly enhanced, but their core principles remain unchanged. These methods may be generalized as follows:

$$\frac{\mathrm{GHI}}{\mathrm{G_c}} = \mathrm{a} + \mathrm{bn} + \mathrm{cn}^2. \quad \mathrm{a,b,c} \in \Re \qquad \text{II.45}$$

Where n is called the cloud index; it is a measure effective cloud cover obtained using satellite measurements. a, b and c are either fixed parameters or parameters varying with n; they are determined with sets of satellite images and ground-based radiation measurements covering the same period.

For example, Rigollier et al. (2004) improved the model of Cano (1985) by adding a calibration relationship between the atmosphere emerging radiance and the numerical count measured by satellite sensors. Rigollier et al. (2004) also improved the model of Cano (1985) by adding a physical parameterization to describe the cloud and ground albedo, and adding several parameters such as the Linke turbidity factor and air mass to account for atmospheric extinction. The updated calibration step for the satellite digital count allowed processing images taken by different sensors (Rigollier et al., 2002a). Besides, to improve GHI estimates the cloud and ground albedo formulations that were

previously defined empirically were updated so that their new formulations are expressed using physical laws.

I.7.b Semi-physical RTP

Semi-physical RTPs attempt to model the absorption and scattering physical processes occurring in the atmosphere using physical principle; however, with parameters that are defined empirically. Semi-physical RTPs are non-spectral models, they treat the atmosphere as plane parallel and account for scattering (Davies et al., 1984). They generally approximate the downward shortwave radiative transfers as follows (Davies et al., 1984; Stensrud, 2009; Noia et al., 1993b):

$$\mathrm{GHI} = \mathrm{G_{TOA}}\,\overline{\mathrm{T}}\,(\mu_0)\mathrm{f}(\rho_\mathrm{g}). \qquad \text{II.46}$$

$\bar{T}$ is the atmospheric transmission function averaged over the entire solar spectrum; it is defined as:

$$\bar{T}(\mu_0) = \frac{1}{G_{TOA}} \int_0^\infty G_{TOA,\lambda}\, T_\lambda(\mu_0) d\lambda, \qquad \text{II.47}$$

where $G_{TOA,\lambda}$and $T_\lambda$are the monochromatic top of atmosphere irradiance and transmission function, respectively. To compute the GHI, $\bar{T}$ is obtained using an energy conservation equation in an earth-atmosphere column. This energy conservation following Fig. II.11 writes (Hay, 1993):

$$\mathrm{IE}^{\uparrow} - \mathrm{IE}^{\downarrow} = \mathrm{E_a} + \mathrm{E_g^{\downarrow}}(1 - \rho_\mathrm{g}), \qquad \text{II.48}$$

where $IE^{\uparrow}$ and $IE^{\downarrow}$ are: the solar stream reflected to the space and the extraterrestrial solar irradiance (G_{TOA}) respectively. E_a is the solar stream absorbed by the atmosphere; $E_g^{\downarrow}$ is solar irradiance at the earth's surface (GHI). $f(\rho_g)$ describes the multiple reflections between the ground and the atmosphere; it is obtained using empirical relationships that correlate measurements made by the satellite in the satellite band to the broadband irradiance.

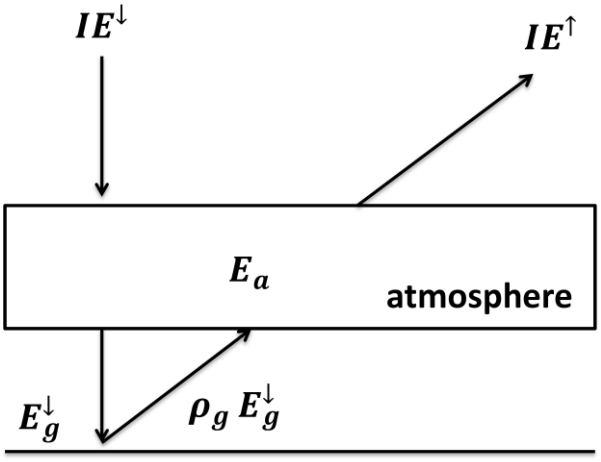

Fig. II.11 solar radiation flux at the atmosphere inspired after (Noia et al., 1993a)

The bibliographic review we carried out highlighted the following model: (Gautier et al., 1980; Janjai, 2010; Gautier and Landsfeld, 1997). These models use Eq.II.47 to define $\bar{T}$; they, however, differ in the atmospheric processes that are accounted for.

Gautier et al. (1980) model the atmosphere for the downward path of the solar radiation as a scattering layer on top of an absorbing layer whereas in the upward path these layers are inverted. These layers model the absorption due to the water vapor and the scattering from a Rayleigh atmosphere. A pure Rayleigh atmosphere contains only the permanent atmospheric gases that scatter radiation by Rayleigh scattering. It excludes the effects of water vapor, clouds, and aerosols (AMS, 2012). Clouds add an extra layer that both absorb and reflect the incoming radiation as shown Fig. II.12. Gautier et al. (1980) model does not account for the ozone absorption and multiple scattering. Besides, it assumes isotropic reflection, isotropic scattering and a single homogeneous layer of clouds (Diak and Gautier, 1983).

Under clear sky conditions the GHI is expressed as follows (Gautier et al., 1980; Noia et al., 1993b):

$$\begin{cases} \mathrm{GHI} = \mathrm{G_{TOA}}\,\overline{\mathrm{T}}(\mu_0)\left(1+\rho_{\mathrm{s,d}}(\theta)\rho_{\mathrm{g}}\right) \; and \\ \overline{\mathrm{T}}(\mu_0) = \left(1-\rho_{\mathrm{s,b}}\right)\left(1-\alpha_{\mathrm{w}}^{\downarrow}\right). \end{cases} \qquad \text{II.49}$$

Where $\rho_{s,d}$ and $\rho_{s,b}$ are the broadband scattering coefficients for the upward diffuse radiation and the broadband downward beam radiation coefficient due to a Rayleigh atmosphere. ρ_g is the broadband surface albedo; $\alpha_w^{\downarrow}$ is the broadband water vapor absorption coefficient for the downward path of the solar radiation. ρ_d,ρ_b and $\alpha_w^{\downarrow}$ values were extracted from Coulson (1959) study. ρ_g was computed using a conservation equation for the upwelling radiance. This equation related the ground albedo to the solar radiation measured by Geostationary Operational Environmental Satellite (GOES) in the visible band for cloud-free days. Gautier et al. (1980) assumed that there was no need for a conversion relation between satellite visible band measurements into a broadband equivalent (Diak and Gautier, 1983). The satellite images are also used to detect the occurrence of clouds by applying an empirical brightness threshold (Noia et al., 1993b).

Under cloudy sky conditions the atmospheric transmission function:

$$\begin{cases} \mathrm{GHI} = \mathrm{G_{TOA}}\,\overline{\mathrm{T}}(\mu_0) \; and \\ \overline{\mathrm{T}}(\mu_0) = \left(1-\rho_{\mathrm{s,b}}\right)\left(1-\alpha_{\mathrm{w,t}}^{\downarrow}\right)(1-\alpha_{\mathrm{c}})(1-\rho_{\mathrm{c}})\left(1-\alpha_{\mathrm{w,b}}^{\downarrow}\right). \end{cases} \qquad \text{II.50}$$

Where $\alpha_{w,t}^{\downarrow}$ and $\alpha_{w,b}^{\downarrow}$ are water vapor absorption coefficient above and below the cloud level respectively. Gautier et al., (1980) estimated empirically that for low and middle clouds there is an average of 30% of water vapor above the cloud level (i.e. $\alpha_{w,t}^{\downarrow} = 0.3\,\alpha^{\downarrow}$). Besides α_c was modeled empirically as linear function of ρ_c which was obtained through satellite measurements.

Janjai (2010) atmosphere model is more complex; it consists of two layers. The first layer where ozone absorption occurs is from the TOA to the top of the cloud layer. The second layer is from the top of the cloud layer to the earth's surface, containing air molecules, water vapor, aerosols and clouds. Contrary to the model of Gautier et al. (1980), Janjai (2010) includes the effect of ozone and aerosols and assumes that the optical state of the atmosphere on the downwelling path is identical to the upwelling path (Fig. II.13). Therefore, the GHI is expressed with the following equation:

$$\begin{cases} \mathrm{GHI} = \mathrm{G_{TOA}}\,\overline{\mathrm{T}}(\mu_0)\left[\left(1-\rho_\mathrm{g}\right)+\rho_\mathrm{g}(\alpha_\mathrm{w}+\alpha_\mathrm{aer})\right] and \\ \overline{\mathrm{T}}(\mu_0) = \tau_0(1-\rho_\mathrm{a}-\alpha_\mathrm{w}-\alpha_\mathrm{aer}). \end{cases} \quad \text{II.51}$$

Where ρ_g was determined empirically from several cloud free satellite images. α_w is the broadband water vapor absorption computed using Lacis and Hansen (1974) approximation for water vapor; α_{aer} the broadband aerosol absorption is computed using ground measurement of the visibility and aerosol single scattering albedo; τ_0 the broadband ozone absorption transmission coefficient is computed from Lacis and Hansen (1974) approximation for ozone assuming that the ozone layer is purely absorbing. ρ_a is the broadband earth-atmosphere reflectance obtained empirically using satellite band earth-atmosphere reflectance. The information about clouds are included in ρ_a.

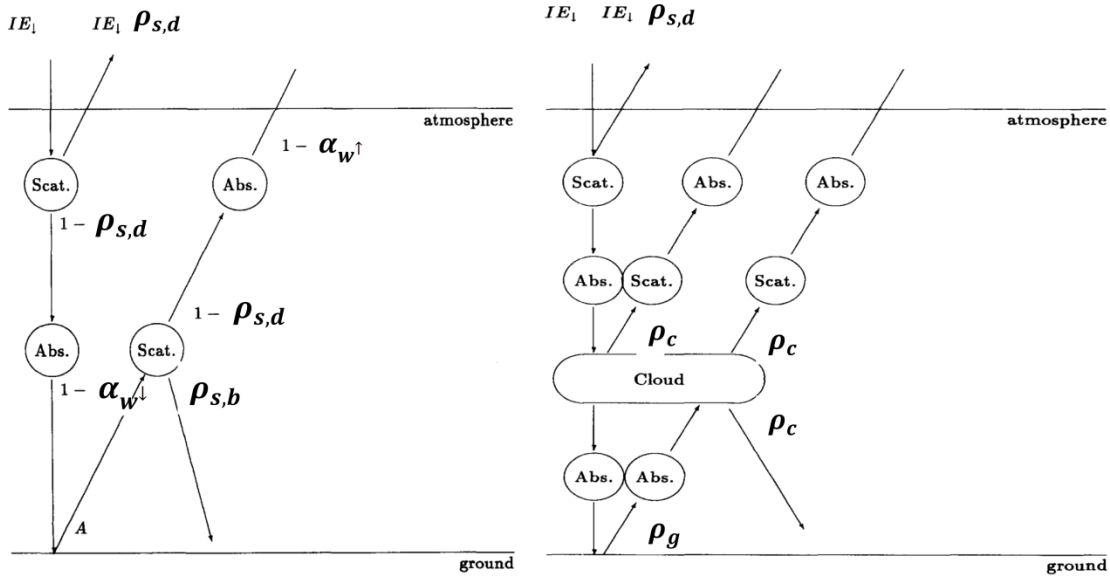

Fig. II.12 atmospheric model of Gautier et al. (1980) adapted from (Noia et al., 1993b)

The atmosphere model of Gautier and Landsfeld (1997) has a more detailed description of the atmospheric processes than the model of Janjai (2010). Similarly to the model of Janjai (2010) the water vapor absorption, ozone absorption, aerosol absorption and scattering are included. The atmosphere model of Gautier and Landsfeld (1997) also includes the absorption of carbon dioxide and oxygen; and the multiple scattering between the ground and an atmosphere that is clear or cloudy. Under clear sky the monochromatic irradiance $GHI_{c,\lambda}$ is expressed as follows (Frouin et al., 1989):

$$\begin{cases} \mathrm{dGHI_{c,\lambda}} = \mathrm{G_{TOA,\lambda}}\,\mathrm{T_\lambda}(\mu_0)\,\mathrm{e}^{-\left(1-\rho_{\mathrm{g},\lambda}(\theta)s_\lambda\right)}\mathrm{d}\lambda \text{ and} \\ \mathrm{T_\lambda}(\mu_0) = \mathrm{T_{g,\lambda}}(\mu_0)\mathrm{e}^{\left(\frac{-\tau}{\mu_0}+\mathrm{T_{d,\lambda}}(\mu_0)\right)}. \end{cases} \quad \text{II.52}$$

Where $T_{\mathrm{g},\lambda}$ and $T_{\mathrm{d},\lambda}$ are the transmittance due to absorbing gases and due to diffuse sky. $\rho_{g,\lambda}$ and s_λ are the ground reflectance and spherical albedo of the atmosphere. Under cloudy sky conditions the GHI writes:

$$\mathrm{GHI} = \frac{\mathrm{GHI_c}(1-\rho_\mathrm{c}-\alpha_\mathrm{c})}{1-\rho_\mathrm{c}\,\rho_\mathrm{g}}, \quad \text{II.53}$$

where GHI_c is the value of irradiance under clear sky for the shortwave frequency range. This model assumes that in the spectral range of interest, solar radiation is: scattered by air molecules and aerosols, absorbed by ozone, water vapor, carbon dioxide, oxygen, and aerosols.

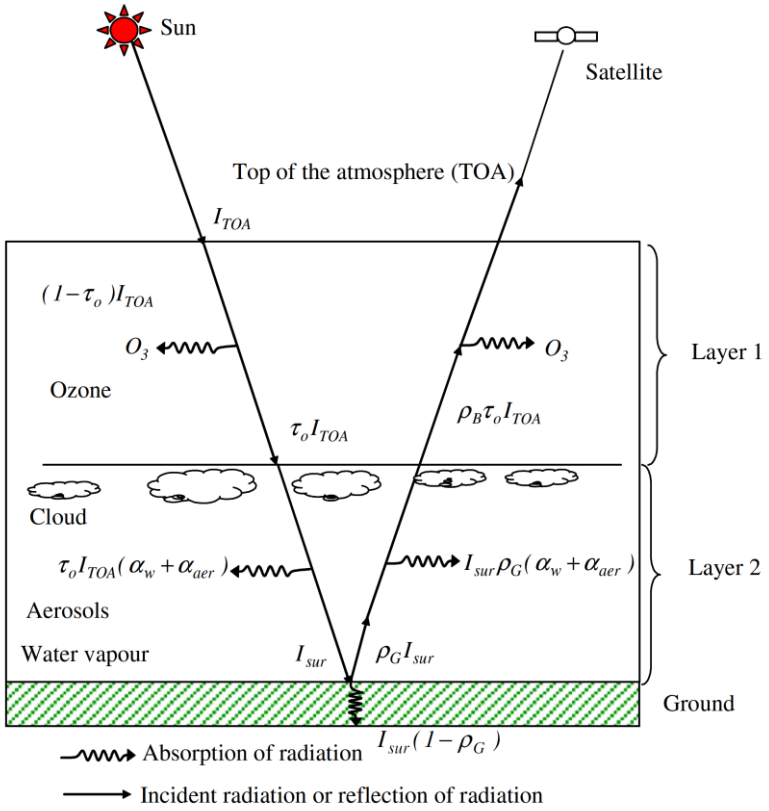

Fig. II.13 atmospheric model of Janjai, (2010)

Besides, reflection and absorption by clouds is supposed to occur in one layer. ρ_g is determined empirically from a time series of satellite images. The minimum brightness value for each satellite pixel during the daytime defines a threshold that is used to classify the satellite pixel as clear or cloudy. Under clear sky conditions, the satellite sensors measure the reflection from the ground. s_λ, $T_{d,\lambda}$ are approximated using the relationship suggested by Tanre et al. (1979). Similarly $T_{g,\lambda}$ is approximated using a relationship suggested in Gautier and Landsfeld (1997). α_c was modeled as linear function of ρ_c which is obtained through satellite measurements in the same way than Gautier et al. (1980).

I.7.c Physical RTP

Physical RTPs study the physical processes occurring in the atmosphere and influencing solar radiation; namely absorption and scattering. They make approximations to Eq.II.22 that simplify its resolution. The two stream and Eddington approximation are the most commonly used methods to assess the GHI (Anderson and co-authors, 2004; Brunner et al., 2015; Mathiesen and Kleissl, 2011; Morcrette, et al., 2007; Ritter and Geleyn, 1992; Slingo, 1989; Stuhlmann et al., 1990; Troccoli and Morcrette, 2014; Zhong et al., 2016).

Two-stream approximation is similar to the disort method; it assumes that there are only upward and downward radiance in the directions μ_1 and $-\mu_1$. Consequently, $N = 1$ in Eq.II.29, which

gives $j = -1$ for the downward stream and $j = 1$ for the upward stream. The upward and downward directions are given by the Gauss quadrature formula. The values of $\mu_{i,j}$, p_l and ω_l Eq.II.30 -II.31 for the two-stream approximation are shown Tab. II.5.

Tab. II.5 Gaussian Points, Gaussian Weights and Legendre polynomials values for two-stream approximation. g is the asymmetry factor defined Eq. II.52. Inspired after (Liou, 2002)

n	Radiation stream value $\pm\mu_n$	Integral Gaussian weights a_n	Phase function polynomial value p_n	Phase function polynomial weight ω_n
0			$P_0 = 1$	$\omega_0 = 1$
1	$\mu_1 = 0.5773503$	$a_1 = 1$	$P_1 = x$	$\omega_1 = 3g$

Consequently, we obtain two equations from Eq.II.29: one equations that describe the downward stream and the other describing the upward stream. After some mathematical manipulations Eq.II.29 yields (Liou, 2002):

$$\mu_1 \frac{dM(\tau,\mu_1)}{d\tau} = (1-\omega_\lambda(\tau)g)N(\tau,\mu_1) - (S^- - S^+)e^{\frac{-\tau_\lambda}{\mu_0}} \; and \qquad \text{II.54}$$

$$\mu_1 \frac{dN(\tau,\mu_1)}{d\tau} = \big(1-\omega_\lambda(\tau)\big)M(\tau,\mu_1) - (S^- + S^+)e^{\frac{-\tau_\lambda}{\mu_0}}. \qquad \text{II.55}$$

Where,

$$\begin{cases} M = L(\tau,\mu_1) + L(\tau,-\mu_1), \\ N = L(\tau,\mu_1) - L(\tau,-\mu_1), \\ g = \frac{\omega_1}{3} = \frac{1}{2}\int_{-1}^{1} p(\cos\Theta)\cos\Theta \, d\cos\Theta, \\ b = \frac{1-g}{2} \; and \\ S^{\pm} = \frac{F_0\omega_\lambda(\tau)}{4\pi} \; (1 \pm 3g\mu_1\mu_0). \end{cases} \qquad \text{II.56}$$

g is called the asymmetry factor, it is a function of ω_1, the 1^{th} moment of the phase function p. The two stream resolution of the RTE is used mostly in the radiative transfer code of NWP model: McRad (Mathiesen and Kleissl, 2011; Morcrette, et al., 2007; Troccoli and Morcrette, 2014), New Goddard (Zhong et al., 2016) and COSMO (Brunner et al., 2015; Ritter and Geleyn, 1992). It was also used in a satellite based method by Stuhlmann et al., (1990). Stuhlmann et al.(1990) used the Two stream method to relate the cloud index n Eq.II.45 measured in the satellite narrow band to a broadband cloud transmittance.

In Eddington's approximation the phase function is approximated similarly as the two-stream but the radiance is expanded using Legendre polynomial and Eq.II.27 is not discretized using Gaussian quadrature.

$$L_{\lambda,d}(\tau,\mu_i) = \sum_{l=0}^{N=1} L_{l\lambda,d}(\tau,)\, p_l(\mu_i) \qquad i = N, -N \text{ with } N \in \Re \qquad \text{II.57}$$

Substituting Eq.II.57 into in Eq.II.27 and rearranging it yields (Liou, 2002):

$$\frac{dL_{1\lambda,d}(\tau,\mu)}{d\tau} = 3\big(1-\omega_\lambda(\tau)\big)L_0(\tau,\mu) - \frac{3\omega}{4\pi}F_0 e^{e^{\frac{-\tau_\lambda}{\mu_0}}} \; and \qquad \text{II.58}$$

$$\frac{dL_{0\lambda,d}(\tau,\mu)}{d\tau} = (1-\omega_\lambda(\tau)g)L_1(\tau,\mu) + \frac{3\omega}{4\pi}F_0 e^{e^{\frac{-\tau_\lambda}{\mu_0}}}. \qquad \text{II.59}$$

The Eddington resolution of the RTE is used in GFDL-SW (Anderson and co-authors, 2004; Mathiesen and Kleissl, 2011) and (Slingo, 1989).

The two-stream and Eddington methods are good approximations for optically thick layers, but they produce inaccurate results for thin layers and when significant absorption is involved (Liou, 2002). These inaccuracies are related to simplicity of the phase function approximation Eq.II.31; according to Liou, (2002) a two term Legendre polynomial expansion is not adequate to describe the strong forward scattering of aerosols and cloud particles. To solve this issue Delta-Function adjustment was introduced; this procedure lead to the δ-Eddington and the δ-two-stream. Consequently, the optical depth, single-scattering albedo, and phase function were scaled such tat:

$$\begin{cases} f = \dfrac{\omega_2}{5}, \\ \tau' = (1-\omega f)\tau, \\ \omega' = \dfrac{(1-f)\omega}{1-\omega f} \; and \\ p'(\tau,\mu,\mu') = 1 + 3g'\mu\mu'. \end{cases} \qquad \text{II.60}$$

Where f is the fraction of the energy scattered in the forward direction; it is expressed a function of the phase function second moment. τ', ω',p' are the scaled optical depth, single scattering albedo and phase function respectively (Liou, 2002). Under clear sky both the δ-Eddington and δ-two-streams have relative errors of less than 1.5%. However, under cloudy skies relative errors might be as high as 10%, which indicates that higher order approximation scheme is necessary in order to obtain the accurate solar cloud absorption in weather and climate models (F. Zhang et al., 2013).

Method such as the Heliosat-4 (Qu, 2013; Qu et al., 2016) aim to find a tradeoff between the accuracy of method such as disort, adding and the computational efficiency of two-stream and Eddington. In the Heliosat-4 (Qu, 2013; Qu et al., 2016) the RTE was resolved for a selected number of atmosphere characterized by fixed optical properties. These computations define nodes that can be used to interpolate the GHI for other types of atmosphere characterized by other optical properties.

I.8 Validation metrics

Various metrics have been proposed and used to quantify the accuracy of solar irradiance forecasts. They can be broadly divided into five categories (J. Zhang et al., 2013): statistical metrics, variability estimation metrics, uncertainty quantification and propagation metrics, ramping characterization metrics and economic and reliability metrics. These metrics describe different characteristics of the solar irradiance obtained from a model. These metrics are insufficient on their own and must be combined to establish a complete, coherent comparison for benchmarking (Beyer et al., 2009). Determining which metrics are most appropriate depends on the user: system operators are interested in metrics that accurately reflect the costs of forecast errors whereas researchers are

interested in metrics that give relative performance of different forecast models under different conditions (Kleissl, 2013). In this section we describe conventional statistical metrics, temporal variability and ramping metric that are commonly used to assess models performances.

Statistical metrics commonly used (Jimenez et al., 2016; Perez et al., 2013, 2010; Rigollier et al., 2004; Zhong et al., 2016) are the correlation coefficient (R), the root mean square error (RMSE), the mean average error (MAE), the mean bias error (MBE) and the Kolmogorov Smirnov integral (KSI). The RMSE and MAE are indicators of the model spread (Kleissl, 2013), the RMSE allows to put more weight on large forecast errors whereas the MAE weighs linearly all deviations to the observations (Lorenz et al., 2009b). According to Perez et al. (2013) the RMSE is most important model validation metric in the context of renewable power forecasting where small errors are more tolerable and large forecast errors have a disproportionately high impact in the grid management. Nevertheless, according to Chai and Draxler (2014) there is no definite consensus on the most appropriate metric between MAE and RMSE to assess a model performance. The use of RMSE is recommended when the error distribution is expected to be Gaussian (Chai and Draxler, 2014). The MBE describes systematic deviations of the model GHI estimates (Perez et al., 2013); it indicates the model average bias or deviation. This metric contrary to the RMSE and MAE indicate if the model under or over-estimate the GHI. Contrary to the MAE and RMSE a deviation to the observation does not automatically increase the MBE value; low MBE may be caused by highly over-estimated values of GHI that are balanced by highly underestimated values of GHI or vice versa. KSI and R measure correlation between datasets; the KSI instead of comparing raw datasets, it compares their cumulative distribution function (CDF). The evaluation of CDF is helpful for applications where decisions are related to threshold values (Espinar et al., 2009; Perez et al., 2013). The KSI value is advantageous because it shows how dataset values are statistically distributed. It not only describes the distribution of the points around a unit line, similarly to MBE, but also allows the user to distinguish the behaviors of stations with similar RMSE values (Espinar et al., 2009). Statistical metrics discussed above are expressed as follows:

$$MAE = \frac{\sum_i |y_i - x_i|}{N \sum_i x_i}, \qquad \text{II.61}$$

$$\mathrm{MBE} = \frac{\sum_i y_i - x_i}{N \sum_i x_i}, \qquad \text{II.62}$$

$$\mathrm{RMSE} = \frac{\sqrt{\frac{1}{n}\sum_i (y_i - x_i)^2}}{N \sum_i x_i}, \qquad \text{II.63}$$

$$\mathrm{R} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})}\sqrt{\sum_i (y_i - \bar{y})}} = 1 - \frac{\sigma^2 (y_i - x_i)}{\sigma^2 (x_i)} \; and \qquad \text{II.64}$$

$$\mathrm{KSI} = \int_{xmin}^{xmax} |\mathrm{CDF}_{comp}(x_j) - \mathrm{CDF}_{ref}(x_j)| dx. \qquad \text{II.65}$$

Where $x_j \in [x_{min} + (n-1)p, x_{min} + np)$ with $p = {(x_{max} - x_{min})}/{m}$; n is the discretisation subdivision $n \in [1,.., m]$; m is the total number of discretisation subdivisions, taken as ≥ 100 for all integrations; and x_{max} and x_{min} are the extreme values of GHI; y_i is the model estimate and x_i is the ground measurement value. $\bar{x}$ and $\bar{y}$ are the model-estimated and ground-measured means, respectively; N is the total number of data points. RMSE and MBE are expressed in W/m². σ is the standard deviation error and σ^2 the variance.

These statistical metric are related to each other as follows (Lorenz et al., 2009a):

$$\mathrm{RMSE}^2 = \mathrm{MBE}^2 + \sigma^2. \qquad \text{II.66}$$

Temporal variability and ramping characterization metrics intend to address one of the biggest concerns associated with integrating a large amount of solar power into the grid (Bright et al., 2017). They give indication on the ability of a model to describe sudden and large variations of GHI values; therefore, sudden and large variations of electricity output caused by clouds s and extreme weather events. Variability metrics include the mean variability of the clearness index: M_{K_c} Eq.II.67, and the mean squared variability of the clearness index : V Eq.II.68 (Marquez and Coimbra, 2012). $\mathrm{M_{K_c}}$and V are expressed as follows:

$$V = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(K_c(k+\Delta t) - K_c(k))^2}. \qquad \text{II.68}$$

$$\mathrm{M_{K_c}} = \frac{1}{\mathrm{N}}\sum_{\mathrm{k=1}}^{\mathrm{N}}\mathrm{K_c(k)} \ and \qquad \text{II.67}$$

$$V = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(K_c(k+\Delta t) - K_c(k))^2}. \qquad \text{II.68}$$

Where N is the number of sample is the consider time interval. The GHI ramp rate is defined as the change in irradiance divided by the time over which it occurs; it is expressed as follows (Mathiesen et al., 2013):

$$\mathrm{RR(\Delta t)} = \frac{\mathrm{x_{i+\Delta t} - x_i}}{\mathrm{\Delta t}}. \qquad \text{II.69}$$

The ramp rate may also be used to describe the overall ability of a model to reproduce the variability of ground measurement datasets using the CDF as the show by Bright et al. (2017).

Consequently, we used the metric described here in this section to describe the ability of a model to reproduce specific characteristics of the GHI measured at the ground.

II. Improving the GHI estimate from satellite method in the inter-tropical zone

The first section presented the different methods used to solve the complex RTE. This section gives the advantage and drawback of statistical, semi-physical and physical parameterizations and how they can be combined to improve the GHI estimates in French Guiana and the ITZ using the Heliosat-2 method.

II.1 Comparison between statistical, semi-physical and physical parameterizations of the radiative transfer equation

The main advantage of statistical parameterization methods is their simplicity, they rely on linear or nonlinear relationship between the GHI and other measured variables (Noia et al., 1993a). Nevertheless, a major drawback of statistical parameterization methods arises from the unicity of the relationship coefficients. These coefficients depend on the physical and spatial properties of the atmosphere at the region of interest (Badescu, 2008; Raphael and Hay, 1984). Statistical methods such as the Heliosat-2 was validated in several type of climates ranging from oceanic, Mediterranean, desert to semi-arid climate regions (Dagestad, 2004; Lefèvre et al., 2007; Dürr and Zelenka, 2009; Moradi et al., 2009; Wahab et al., 2009; Jumaily et al., 2010; Blanc et al., 2011). These studies showed a varying level of accuracy and sizable biases under cloudy skies. Inaccuracies under cloudy sky conditions (Girodo et al., 2006; Suárez et al., 2012; Polo et al., 2014; Albarelo et al., 2015) were explained by the inability to model properly the optical properties of clouds such as cloud optical depth which influence cloud reflectance, absorptance and transmittance (Welch et al., 1980).

Semi-physical methods presented rely on physical principles. They are broadband and consider that for shortwave the most important radiative processes occurring within the atmosphere are: scattering and absorption by molecules, clouds, and aerosols. These methods use satellites images to provide information on the optical properties of the cloud and for some case other processes (Janjai, 2010). Nevertheless, semi-physical methods require an extensive knowledge of the optical properties of the atmosphere that might not be available. In that case they are obtained through empirical relationships (Hay, 1993; Lacis and Hansen, 1974; Noia et al., 1993b; Stephens, 1984). When these data are available they are climatological and do not vary with the atmospheric state (Gautier and Landsfeld, 1997; Stephens, 1984; Stephens et al., 1990). Nevertheless, satellites semi-physical method offer the advantage of: (1) being more computationally efficient, (2) having a wide geographical coverage, (3) having a high spatial and temporal resolution (Hay, 1993).

Physical parameterization methods similarly to semi-physical methods consider the most important radiative processes occurring within the atmosphere; besides, they include a spectral dependency. Physical methods have a more detailed description of the atmosphere than semi-physical methods; therefore, they require a lot more input datasets that might not be accurate or might be available sparsely. According to Stensrud (2009) inaccurate aerosol optical depths and ozone amounts inputs lead to high bias in GHI estimates. One advantage of physical parameterization methods is their potential to assess more accurately the GHI; however, with high computational cost (Pincus and Stevens, 2013). According to Fouquart et al., (1991) and Stensrud, (2009) there are two main causes of the uncertainty in the shortwave calculations from physical parameterization methods. The first is

related to the approximation made to compute vapor absorption and the lower spectral resolution used to reduce the computation time.

Recently, the Heliosat-2 statistical method was used to produce GHI estimates in French Guiana with GOES images (Albarelo et al., 2015). On cloudy sky days, the root mean square error (RMSE) of hourly GHI estimates attained values of up to 78%. These unsatisfactory results can be attributed to the significant cloud cover seen in intertropical zones, as clouds are known to be an important source of bias (Gautier and Landsfeld, 1997; Pereira et al., 2000; Janjai, 2010). To improve the GHI estimates from this statistical method we considered coupling it with semi-physical methods that account for cloud properties. This approach was chosen because: (1) satellite images can provide information on the optical properties of the cloud; according to Stuhlmann et al., (1990) the cloud fraction observed by satellite is used as a proxy of cloud transmittance. (2) Semi-physical methods are less computationally expensive than Physical parameterization methods and require less input parameters.

The article below describes how the statistical heliosat-2 method was coupled with a radiative transfer parameterization under cloudy sky conditions.

II.2 Improving the Heliosat-2 Method for Surface Solar Irradiation Estimation Under Cloudy Sky Areas

(Preprint version of: https://doi.org/10.1016/j.solener.2018.05.032)

AUTHORS:
Diallo Mouhamet[a], Albarelo Tommy[a], Antoine Primerose[a], Linguet Laurent[a]

[a]*Université de la Guyane – UMR Espace-Dev, BP 792, 97 337, Cayenne, France*

Corresponding author: laurent.linguet@univ-guyane.fr

Abstract

The purpose of this study was to improve global horizontal irradiation (GHI) estimates under cloudy sky conditions using an optimised version of the Heliosat-2 method calibrated with Geostationary Operational Environmental Satellite images. The optimised version was coupled with a radiative transfer parameterisation[1] (RTP) to better account for local cloud properties. The key element of this parameterisation is the cloud absorption coefficient, which was the only element to be computed. The obtained estimates were compared against GHI measurements from six meteorological stations in French Guiana over four years. We used root mean square error (RMSE), mean bias error (MBE), correlation coefficient, and the Kolmogorov Smirnov test integral to assess GHI estimate accuracy. The estimates were also compared with those obtained with the original optimised Heliosat-2 method. Our results show an improvement in the GHI estimates under cloudy, clear, and all sky conditions. Under cloudy skies, the RMSE and MBE of our method ranged from 43% to 63% and −2% to −22%, respectively, whereas those of the original optimised Heliosat-2 method ranged from 66%

[1]RTP: radiative transfer parameterisation

to 87% and −48% to −65%, respectively. Another effect of the new method was the improvement in clear sky GHI estimates: the RMSE and MBE ranged from 16% to 24% and −20% to −8%, respectively, while those of the optimised Heliosat-2 ranged from 19% to 28% and −23% to −10%, respectively. The improvement in GHI estimates under cloudy sky conditions led to better GHI estimates under all sky conditions: the RMSE and MBE ranged from 19% to 26% and −7% to −2%, respectively, while the corresponding values for the original optimised Heliosat-2 ranged from 22% to 32% and −15% to −8%, respectively. The suggested cloudy sky RTP offers the advantages of simple implementation and good computation speed. This method requires only one parameter; users need to configure a cloud absorption coefficient suitable for their local conditions. We conclude that coupling the optimised Heliosat-2 method with an RTP improves GHI estimates in French Guiana and reduces discrepancies between satellite-derived irradiation and ground measurements in areas with high cloudiness.

Keywords: Heliosat-2, GOES, cloudy sky, global horizontal irradiance, radiative transfer parameterization

1 Introduction

Accurate knowledge of global horizontal irradiation (GHI) is vital in several fields, including climatology, agriculture and energy. For instance, GHI is used in climatology to assess ocean heat flux balance and in agronomy to monitor the growth rate of crops. GHI mapping can be inferred through surface radiation network measurements using pyranometers (Perez et al., 1994). However, the high manufacturing cost, maintenance costs, and an uneven network density of pyranometers limit their use (Inman et al., 2013; Qu, 2013). Several studies (e.g., Perez et al., 1994; Zelenka et al., 1999) have shown that when distance between the studied site and the pyranometers is greater than 30 km, solar radiation obtained through interpolation methods is of lower quality than satellite-derived solar radiation. Satellite-derived GHI methods were developed with the first satellite launches (Levanon, 1971; House et al., 1986) and were either classified as statistical (Tarpley, 1979; Cano et al., 1985; Möser and Raschke, 1984; Noia et al., 1993a, 1993b) or physical (Ellis and Haar, 1976; Gautier et al., 1980).

Heliosat-2 is one of the most widely validated methods developed for estimating solar irradiation from satellite images. One of the assets this method is that it is an open source method and is modular enough to be employed worldwide. It has been validated in oceanic, Mediterranean, desert and semi-arid climate regions (Dagestad, 2004; Lefèvre et al., 2007; Dürr and Zelenka, 2009; Moradi et al., 2009; Wahab et al., 2009; Jumaily et al., 2010; Blanc et al., 2011). Furthermore, it has been used in tropical areas over which the Intertropical Convergence Zone (ITCZ) transits: Mozambique (Wald et al., 2011), Zimbabwe (Blanc et al., 2011) and French Guiana (Marie-Joseph et al., 2013). However, GHI estimates from the stations in these tropical areas show greater biases than those from stations in oceanic, Mediterranean and semi-arid climate regions (Blanc et al., 2011). Therefore, improvement of GHI estimates from Heliosat-2 in intertropical zones remains an open topic of discussion.

Recently, Heliosat-2 was used to produce GHI estimates in French Guiana (Albarelo et al., 2015) with Geostationary Operational Environmental Satellite (GOES) images. On cloudy sky days, the root mean square error (RMSE) of hourly GHI estimates attained values of up to 78%. These unsatisfactory results can be attributed to the significant cloud cover seen in intertropical zones, as clouds are known to be an important source of bias (Gautier and Landsfeld, 1997; Pereira et al., 2000; Janjai, 2010). Overall, satellite-derived GHI methods exhibit positive biases under cloudy sky conditions (Diak et al., 1982; Girodo et al., 2006; Suárez et al., 2012; Polo et al., 2014; Tarpley, 1979), reflecting the inability of the satellite sensor to observe the lowest layer of clouds in the visible spectrum (Heinle et al., 2010). In the Heliosat-2 method, cloud attenuation is modelled empirically using a polynomial function of a cloud index that tends to one when the sky is overcast and zero when the sky is clear. However, the cloud index does not model cloud properties such as cloud optical depth, which influences cloud reflectance, absorbance and transmittance (Welch et al., 1980). Therefore, Heliosat-2 GHI estimates under cloudy sky conditions need to be improved to provide accurate GHI estimates in tropical areas. Several corrections have been proposed to improve the Heliosat-2 method irrespective of whether the sky is clear or cloudy; however, to the best of our knowledge, no methods have been proposed to correct the GHI estimates obtained using the Heliosat-2 method in tropical zones with a high occurrence of cloudy skies. The objective of this study was to improve GHI estimates from the Heliosat-2 method and GOES images for areas with high cloudiness.

Two main approaches have been used to assess GHI under cloudy sky conditions. The first involves radiative transfer models (RTMs) that simultaneously compute absorbed, scattered and emitted electromagnetic radiation through the Earth's atmosphere. This approach provides the most accurate results; however, it is time consuming and requires many input parameters (Stephens et al., 2001). Oumbe et al. (2009) developed the Heliosat-4 method that relies on an approximation of the RTM equation using a product of clear sky irradiance and a parameter representing cloud extinction. This approximation allows faster GHI computation when abaci, previously computed by an RTM, are combined with interpolation functions. Under cloudy skies, the tested cases exhibited an overall overestimation of GHI; the greatest positive biases were found when cloud coverage was between 2% and 10%, where biases were sometimes greater than 60 W/m^2. High RMSE values were encountered for desert stations where cloud fractions were between 20% and 30% and RMSE was up to 220 W/m^2 .

The second approach involves radiative transfer parameterisations (RTPs), which are simpler spectral models based on parameterisations of transmittance and absorption functions for basic atmospheric constituents such as Rayleigh scattering and ozone, water vapour and aerosol transmittances (Khalil, 2008). These parameterisations provide faster results with accuracy similar to the more time-consuming RTMs (Fouquart et al., 1991).

Gautier et al. (1980) developed a method based on an RTP to estimate GHI. Overall, the parameterisation showed potential; however, the results under cloudy sky conditions were not optimal. This first RTP was updated by Gautier and Landsfeld (1997) to better account for cloud effects; they found that under all sky conditions, the first version's monthly RMSE was 9%, while in the improved version it decreased to 4%. Janjai et al. (2005; 2011; 2013) also developed a method based on an RTP to estimate GHI in a tropical climate region under all sky conditions. Contrary to Gautier et al. (1980), they did not develop a separate RTP for clear and cloudy sky conditions; the monthly RMSEs were 6.8% for Vietnam (Janjai et al., 2005), 6.3% for Cambodia (Janjai et al., 2011)

and 9.6% for Myanmar (Janjai et al., 2013). Geiger et al. (2008) parameterised GHI as top-of-atmosphere irradiance lowered by an atmospheric transmission factor. Under cloudy sky conditions, the atmospheric transmission factor was inferred from the Gautier et al. (1980) cloudy sky RTM. Geiger et al. (2008) noted the monthly means of surface solar estimated biases to be within ±5% in most cases with clear sky conditions and within ±15% under cloudy sky conditions. The standard deviation was noted to be between 13 and 111 W/m² under clear sky conditions and between 21 and 191 W/m² under cloudy sky conditions. Bisht and Bras (2010) also used an RTM under clear and cloudy sky conditions. To improve GHI estimates under cloudy sky conditions, they chose a cloudy sky RTM suggested by Slingo (1989); this parameterisation describes the effect of clouds based on their fraction and optical thickness, both provided by MODIS cloud data products. Under clear sky conditions, the daily averaged bias was about 17.82 W/m² and the daily averaged RMSE was about 42.05 W/m²; under cloudy sky conditions, the values were 25.64 W/m² and 66.52 W/m², respectively.

With increasing interest in the development of solar technologies in tropical areas, the merits of Heliosat-2 in these areas are questioned. The solar energy potential of tropical areas can be up to three times higher than that of extra-tropical countries (Fillol et al., 2017); however, variability in solar energy is also higher in these areas due to the dynamic cloud cover and the hot and humid weather (Laing and Evans, 2011; Galvin, 2015). Designing and sizing systems using solar energy input (such as solar water heaters, photovoltaic cells or solar thermal concentrators) require solar data estimates with high accuracy even under tropical zones with a high occurrence of cloudy skies.

We propose an improvement to the Heliosat-2 method so that it can be widely used to obtain GHI estimates of tropical climate regions with as high standards as those of other climate regions. To achieve this goal, we used an RTP model (Polo et al., 2016) and focused on adjusting the atmospheric input data and cloud absorption to better match the local cloud regime. Our new method (RTP_OPT_H2) achieves the following: (1) no significant increase in computation time and (2) no deterioration of clear sky GHI estimation accuracy.

This paper is structured as follows. Section 2 describes both the ground and satellite data used in the study. Section 3 introduces the Heliosat-1 and Heliosat-2 methods with GOES images and describes how RTP_OPT_H2 is implemented to improve GHI estimates under cloudy sky conditions. Section 4 compares the results obtained for cloudy, clear and all sky conditions using both the optimised Heliosat-2 method (OPT_H2) and RTP_OPT_H2. Finally, Section 5 summarises the paper.

2 Data

2.1 Ground measurements

We exploited hourly GHI data from six stations operated by the French national weather services (Météo-France) over four years (Table 1).

Table 1: Ground meteorological stations in French Guiana

Station	Latitude (°)	Longitude (°)	Height (m)	Period
Rochambeau	4.81	−52.37	4	2010–2013
Saint Georges	3.88	−51.80	6	2010–2013

Maripasoula	3.63	−54.03	104	2010–2013
Saint-Laurent	5,48	−53,90	4	2010–2013
Kourou	5.12	−52.44	12	2010–2013
Ile Royale	5.28	−52.58	48	2010–2013

Three stations are located on the Atlantic Coast, namely, Rochambeau, Ile Royale and Kourou. Rochambeau is located 13 km from the Atlantic Ocean and Ile Royale is on a 0.6-km² island, 7 km offshore from the coast. The other three are located inland; Saint-George, Saint-Laurent and Maripasoula are located between 30 and 230 km from the coast. These stations are equipped with Kipp and Zonen pyranometers of type CM6B and CMP11; both types are fitted with a ventilation fan. CM6B is a first-class pyranometer, and CMP11 fulfils the accuracy requirements of a secondary standard pyranometer defined in (WMO, 2007). Preventive maintenance is performed every two months, and the pyranometers are replaced every two years. Prior to final installation, each pyranometer was calibrated at the national radiometry centre in Météo-France; once installed, the coefficients of the new pyranometer were entered into the data acquisition system of the station (Albarelo et al., 2015).

Hourly GHI measurements were run through a quality control process based on extreme values, following the method of Geiger et al. (2002) and the Solar Radiation Data (SoDa[2]) website. For extreme values, hourly GHI (*Gh*, in Wh/m^2) was considered valid when it adhered to the following condition:

$$0.03\ Gh_{TOA} < Gh\ < min\ \Big((1.2\ I_0), \big((1.5\ I_0\ \cos(\theta_0)^{1.2}) + \ 100\big)\Big), \quad (1)$$

where Gh_{TOA} is the top-of-atmosphere hourly surface solar insolation, θ_0 is the sun zenith angle and I_0 is the solar constant (1367 W/m^2).

2.2 Satellite data

The satellite data used in this study were sourced from the Comprehensive Large-Array Stewardship System (CLASS), provided by the National Oceanic and Atmospheric Administration (NOAA). These data were collected by GOES 13, which was launched in May 2006 (Hillger and Schmit, 2009). GOES 13 has five band imagers; in this study, images from the visible band (0.55–0.75 μm) were considered. Images from GOES 13 were taken every 30 min with a 1 × 1 km spatial resolution. The downloaded dataset had 30-min time steps in a series of 16-bit images from January 2010 to December 2013.

2.3 Climate

The climate in French Guiana is regulated by the ITCZ, which is the meeting point of the northeast and southeast trade winds (Albarelo et al., 2015). The ITCZ is characterised by the rapid ascent of hot air associated with the development of large and homogeneous cumulonimbus clouds (Marie-Joseph

[2]Solar Radiation Data. http://www.soda-is.com/eng/index.html.

et al., 2013). It is mainly an oceanic phenomenon (Vasquez, 2009); however, its latitude changes over the years, resulting in high humidity clouds that bring continuous rain over continental areas.

The ITCZ moves over French Guiana twice a year, although variations in the timings of the start and end of the seasons show that there are annual variations in the ITCZ movement (Albarelo et al., 2015). In general, the ITCZ moves northward to ~7° N from May to July and southward to ~15° S from November to January (Albarelo et al., 2015). This shift creates a seasonal cycle with four periods (Bovolo et al., 2012): (1) July to November, when the ITCZ lies to the north, producing a dry season characterised by sky that is mostly clear, although weak precipitations may occur (Albarelo et al., 2015). During this period cirrus and cumulus fractus clouds are predominant; (2) November to January, a short rainy season when the ITCZ moves south of French Guiana (Albarelo et al., 2015); (3) February to March, a transition period before the ITCZ begins its northward motion, characterised by a short dry season (Albarelo et al., 2015); and (4) April to May, when the ITCZ moves northward and a rainy season occurs. During this time, cumulonimbus clouds are predominant.

3 Methods

3.1 The Heliosat method

For details of the principles governing the Heliosat method and the development of the optimised Heliosat-2 method (hereafter, OPT_H2) that exploits GOES images over French Guiana, readers are directed to Albarelo et al. (2015) and Fillol et al. (2017). In summary, the Heliosat method involves the construction of a 'cloud index', *n*, resulting from a comparison of what is observed by the sensor with what should be observed over that pixel if the sky were clear, which is related to the 'clearness' of the atmosphere (Rigollier et al., 2002b):

$$n^t(i,j) = \frac{\rho^t(i,j) - \rho_g^t(i,j)}{\rho_c^t(i,j) - \rho_g^t(i,j)}, \quad (2)$$

where $n\,(i,j)^t$ is the cloud index (unitless); $\rho(i,j)^t$ is the apparent albedo, observed by a space-borne sensor for time t (unitless); $\rho_g^t(i,j)$ is the apparent albedo of the ground under clear sky conditions (unitless); and $\rho_c^t(i,j)$ is the apparent albedo of the brightest clouds (unitless).

As per the Heliosat-1 method core hypothesis (Cano et al., 1985), GHI is a linear function of *n* and is expressed as follows:

$$GHI = n\,G_b + (1-n)G_c, \quad (3)$$

where *GHI* is the global horizontal irradiance (Wh/m^2), *n* is the cloud index (unitless), G_c is the clear sky RTP (Wh/m^2) and G_b is the cloudy sky RTP (Wh/m^2). *GHI* is also related to top-of-atmosphere irradiance as follows:

$$GHI = K_t\,G_{TOA}, \quad (4)$$

where K_t is the clearness index or atmospheric transmission factor and G_{TOA} is the top-of-atmosphere irradiance. Using equations (3) and (4), Cano et al., (1985) found that K_t is linearly dependent on the cloud index equation (5):

$$K_t = a\,n + b. \quad (5)$$

The coefficients *a* and *b* have to be determined empirically with sets of satellite images and ground-based radiation measurements covering the same period (Beyer et al., 1996).

Beyer et al. (1996) adapted the GHI formula to the following:

$$GHI = K_c\,G_c, \quad (6)$$

where K_c is the clear sky index or an atmospheric transmission factor and G_c is the clear sky RTP. Beyer et al. suggested the use of a clear sky RTP to better account for the irradiation dependence on the solar zenith angle and atmospheric aerosol and water vapour content.

Using the linear relationship (5) between the atmospheric transmission factor K_t and the cloud index used by Cano et al., (1985), Beyer et al. (1996) also suggested a linear relationship between the atmospheric transmission factor for clear sky, K_c, and the cloud index. Using sets of satellite images and ground-based radiation measurements covering the same period, Beyer et al. (1996) found that K_c is expressed as follows:

$$K_c = 1 - n, \quad (7)$$

where K_c tends to one for clear sky conditions and tends to zero for cloudy sky conditions. Later, (Rigollier et al., 2002c) showed that for overcast skies ($n > 0.8$), a linear relationship is inappropriate and underestimates K_c; therefore, they suggested a quadratic equation:

$$0.8 < n \leq 1.1 \; K_c = 2.0667 - 3.6667n + 1.6667n^2. \quad (8)$$

The Heliosat-2 method was developed by Rigollier et al. (2004), considering previous works (Cano et al., 1985; Beyer et al., 1996; Rigollier et al., 2002a; 2002b) and setting the maximal and minimal values of the clear sky index to 1.2 and 0.05 (Rigollier and Wald, 1998), respectively (6). The clear sky index K_C was defined as follows.

$$\begin{aligned} n \leq -0.2 \qquad & K_c = 1.2, \\ -0.2 < n \leq 0.8 \qquad & K_c = 1 - n, \\ 0.8 < n \leq 1.1 \qquad & K_c = 2.0667 - 3.6667n + 1.6667n^2, \\ 1.1 < n \qquad & K_c = 0.05. \end{aligned} \quad (9)$$

The European Solar Radiation Atlas clear sky model in the Heliosat-2 method (Rigollier et al., 2000) uses the Linke turbidity factor (TL) as an input. TL is a parameter that describes the attenuation of solar radiation by the atmosphere under clear skies (Remund et al., 2003). TL varies for each site and each month and is calculated using monthly values from the SoDa database. The methodology used to obtain this database of TL values is described in Remund et al. (2003).

The OPT_H2 was developed by Albarelo et al. (2015) to consider GOES images as input instead of

METEOSAT images. The main change involved the modification of the calibration step to process data from GOES. A sensitivity analysis was also conducted with different values of cloud albedo (maximum value, Rigollier value, quantile 95) and TL.
However, in this version, overcast sky conditions are associated with significant errors in GHI estimates. Since the studied area is subject to strong climatic instability linked to the regular presence of the ITCZ, the accuracy of GHI estimates when the sky is cloudy degrades quickly; thus, there is a need for a GHI calculation formula based only on a clear sky RTP. In the Heliosat-2 code (Albarelo et al., 2015; Rigollier et al., 2004), two types of solar irradiation attenuation are accounted for: attenuation by aerosols and attenuation by clouds.

TL models the attenuation caused by aerosol absorption and scattering, assesses the optical depth of aerosols and incorporates water vapour and NO_2 optical depths (Gueymard, 2012b). The value of this parameter increases as scattering increases (Rigollier et al., 2000).

Cloud attenuation is modelled empirically using a clear sky index (K_c) function of cloud index (n) (9). When the sky is overcast, n tends to one and K_c tends to zero, which means that clouds fully attenuate the incoming radiation. However, using parameters that account for both aerosol and cloud attenuation have drawbacks. One of which is that TL is subjected to daily variations (Gueymard, 1998) that are not accounted for in the monthly climatological values provided by the SoDa database used in the Heliosat-2 code. Another drawback is that K_c does not model cloud properties such as cloud optical depth, which influences cloud reflectance, absorbance and transmittance (Welch et al., 1980).

3.2 Optimised Heliosat-2 method coupled with cloudy sky RTP

The limited modelling of cloudy skies by the Heliosat-2 method generates inaccuracies in GHI estimation under cloudy sky conditions in tropical areas, where skies are regularly disturbed by cloud formations with significant variability throughout the year. Therefore, a GHI calculation formula that considers a cloudy sky parameterisation is required. We modified the Heliosat-2 method by using equation (3) instead of (6) to estimate the GHI values, where G_b is a cloudy sky model that represents the extinction of irradiance due to clouds. Equation (3) was used by Cano et al., (1985) in the form

$$\frac{GHI}{G_{TOA}} = n.\frac{G_b}{G_{TOA}} + (1-n)\frac{G_c}{G_{TOA}} = n\left(.\frac{G_b}{G_{TOA}} - \frac{G_c}{G_{TOA}}\right) + \frac{G_c}{G_{TOA}} \qquad (10)$$

to empirically determine the regression coefficients between *GHI* and *n* (4-5). However, Cano et al., (1985) did not explain or propose a cloudy sky model G_b.

To consider the significant presence of clouds in the atmosphere of the studied area, we integrated a spatiotemporally dependent cloudy sky RTP into the GHI estimate formula of Cano et al., (1985). For this, we used equation (3) proposed by Cano et al., (1985), in which G_b now represents a cloudy sky RTP and *1-n* is replaced by the clear sky coefficient K_c (9). The GHI equation transforms into the following:

$$GHI = n\,G_b + K_c\,G_c. \qquad (11)$$

A bibliographic review highlighted four cloudy sky RTPs used in Gautier et al. (1980), Gautier and Landsfeld (1997), Slingo (1989) and Bisht and Bras (2010). The following criteria were used when

choosing the appropriate model: (1) the RTP should not significantly increase the computation time, and (2) it should not worsen the accuracy of clear sky GHI estimates.

Considering these constraints, we kept the cloudy sky RTP proposed by Gautier and Landsfeld (1997). It is written as follows:

$$G_b = G_c \frac{(1-\rho_{eff}-\alpha_c)}{(1-\rho_{eff}\,\rho_g)}, \quad (12)$$

where G_b is the cloudy sky RTP (Wh/m^2), G_c is the clear sky RTP (Wh/m^2), ρ_{eff} is the effective cloud albedo (unitless), α_c is the cloud absorption (unitless) and ρ_g is the apparent ground albedo (unitless).

For the effective cloud albedo, we used ρ_{eff}, as defined by Lefèvre et al. (2007):

$$\rho_{eff}{}^{t}(i,j) = 0.85 - 0.13\left[1 - exp\left(-4\,cos(\theta_0)\right)^5\right], \quad (13)$$

where θ_0 is the solar zenith angle.

The apparent ground albedo ρ_g was used by Rigollier et al. (2004) in the Heliosat-2 model and is defined as the minimisation of the following equation, since it is assumed that the presence of clouds increases the albedo.

$$\rho_g{}^{t}(i,j) = \frac{min[\,\rho^t(i,j) - \rho^t_{atm}(\theta_s,\theta_v,\Psi)]}{T(\theta_s)T(\theta_v)}. \quad (14)$$

For further details, the reader is encouraged to refer to Rigollier et al. (2004).

The cloud absorption α_c of Gautier and Landsfeld (1997) is linearly dependent on the effective cloud albedo:

$$\alpha_c = a\,\rho_{eff}. \quad (15)$$

The proportionality coefficient a can be determined empirically, and Gautier and Landsfeld (1997) suggested a value of 0.07 for cirrus-type clouds. This RTP considers plane parallel clouds (Gautier and Landsfeld, 1997) as well as reflection and absorption by clouds, which are assumed to occur in one layer (Gautier et al., 1980).

To compute the cloudy sky RTP, a cloud absorption index that describes the local cloud regime was required. As suggested by Gautier and Landsfeld (1997), the tendency of the satellite-derived GHI to underestimate or overestimate the ground-measured surface solar irradiation may be linked to a misrepresentation of cloud absorption. According to Diak and Gautier (1983) and Welch et al. (1980), the proportionality coefficient a (15) between cloud absorption and cloud albedo ranges from 0.04 to 0.17.

To find a (15) the author minimised the bias between the parameterised cloud absorption (15) and a cloud absorption computed from a RTM (Liou 1976) for solar zenith angle between 0° and 75°. First, the cloud type that mainly represents the local cloud regime was chosen. Second, the selected cloud

type's absorption per solar zenith angle α_{ct} was extracted from Liou (1976) drawings, where the cloud absorption for five different types of clouds are available. Third, the effective cloud albedo ρ_{eff} was computed following (14). Finally, a was chosen to minimise the bias between the cloud absorption defined in equation (15) and the cloud absorption α_{ct} obtained previously in Liou (1976) for solar zenith angle θ_s between 0° and 75°. To fit these two cloud absorptions, a (15) was varied between 0.04 and 0.17, and the authors solved the following equation:

$$min_{j\in R}(\frac{1}{N}\sum_{i=1}^{N}\alpha_{ct}(\theta s_i) - a_j\,\rho_{eff}(\theta s_i)), \quad (16)$$

where a_j is the value that can be obtained from the proportionality coefficient a and discrete values of solar zenith angle θs_i between 0° and 75°, chosen according to Liou (1976) cloud type absorption per solar zenith curves. Equation (16) is minimised when $a_j = a$.

The optimised Heliosat-2 method used in this study (RTP_OPT_H2) was designed for 16-bit GOES images instead of the 8-bit GOES images used by Albarelo et al. (2015). Monthly TL values for each site were obtained from the SoDa online database. GHI estimates using the RTP_OPT_H2 method (10, 11) were computed as follows: the clear sky RTP G_c, effective cloud albedo ρ_{eff} and ground albedo ρ_g used to compute the cloudy sky RTP as well as G_b were considered as outputs of the optimised Heliosat-2 code. These parameters were then used to compute the cloud index and GHI estimates. Within the Heliosat-2 code, G_b, K_c and G_c were computed simultaneously and combined to produce the GHI estimates.

4 Results and discussion

The accuracies of the GHI estimates from RTP_OPT_H2 were benchmarked against those of GHI estimates from the optimised Heliosat-2 and of GHI from in situ measurements.

The GHI datasets were compared using four criteria: mean bias error (MBE), the root mean square (RMSE, the Pearson correlation coefficient (R), and the Kolmogorov Smirnov test integral (KSI). These accuracy criteria were defined as follows:

$$BIAS = \frac{\sum_i y_i - x_i}{N\ \sum_i x_i}, \quad (17)$$

$$RMSE = \frac{\sqrt{\frac{1}{n}\sum_i (y_i - x_i)^2}}{N \sum_i x_i}, \text{ and} \quad (18)$$

$$R = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})}\sqrt{\sum_i (y_i - \bar{y})}}. \quad (19)$$

where y_i is the model estimate and x_i is the ground measurement value; $\bar{x}$ and $\bar{y}$ are the model-estimated and ground-measured means, respectively; and N is the total number of data points. RMSE and MBE are expressed in Wh/m².

According to Espinar et al. (2009), RMSE, MBE and R describe different characteristics of the behaviour of a dataset but are often insufficient to establish a complete, coherent comparison for benchmarking. While RMSE describes how points are clustered around a regression line, MBE

describes the distribution of the points around a unit line (Espinar et al., 2009). Therefore, the KSI can be used in such cases as a measure of similarity between two datasets described by their cumulative distribution function (CDF) (Espinar et al., 2009); the lower the KSI value, the closer the CDFs of the two datasets. The KSI value is advantageous because it shows how dataset values are statistically distributed. It not only describes the distribution of the points around a unit line, similar to MBE, but also allows the user to distinguish the behaviours of stations with similar RMSE values (Espinar et al., 2009). KSI is defined as follows:

$$KSI = \int_{xmin}^{xmax} |CDF_{comp}(x_j) - CDF_{ref}(x_j)|dx, \quad (20)$$

where $x_j \in [x_{min} + (n-1)p, x_{min} + np)$ with $p = {(x_{max} - x_{min})}/{m}$; n is the discretisation subdivision $n \in [1,.., m]$; m is the total number of discretisation subdivisions, taken as ≥ 100 for all integrations; and x_{max} and x_{min} are the extreme values of GHI.

Clear and cloudy skies were split with respect to the clearness index K_t values:

$$K_t = \frac{GHI}{G_{TOA}}. \quad (21)$$

Although it would have been also possible to separate cloudy skies from clear skies using the clear sky index (see eq. 6), we used the clearness index (K_t) instead. This is because the calculation of the clear sky index implies the use of a clear sky model, and we did not want to add uncertainties related to this model to our calculation. While we are also aware that K_t is more related to the turbidity in the atmosphere, the calculation of this index only requires GHI and the top-of-atmosphere GHI (GHI_{TOA}), as opposed to that of the clear sky index. Furthermore, GHI used to calculate K_t is obtained from validated and quality-checked ground measurements. K_t has already been used in studies in tropical climate areas, with highly variable sky conditions during the day:
-Soubdhan et al. (2009) studied the daily distribution of K_t in Guadeloupe, French West Indies. They found 4 classes of days: clear sky days, intermittent clear sky days, cloudy sky days and intermittent cloudy sky days. The classes were separated according to the daily K_t value, a sunshine threshold, a cloudy level and a dynamic level.

-Marie-Joseph et al. (2013) studied the applicability of the Heliosat-2 method to French Guiana in the north-eastern part of South America. Cloudy skies were assumed when $K_t < 0.2$ and clear skies were assumed when $K_t > 0.7$. We calculated the daily K_t value from the daily averaged GHI measured at the ground stations, as well as the standard deviation of the hourly K_t for each day; only data for days with a standard deviation of less than 15% were retained. Because of this constraint, the K_t values ranged between 0.7 and 0.1; K_t for clear sky days was almost equal to 0.7, while that for cloudy sky days was less than 0.2 (Albarelo et al., 2015).

Over all the years and stations, the studied dataset includes 70480 hourly all sky values, 4052 hourly clear sky values and 3700 hourly cloudy sky values.

4.1 Quality of GHI estimates from the OPT_H2 method

To compare RTP_OPT_H2 GHI estimates against OPT_H2 GHI estimates, we first estimated the accuracy of the OPT_H2 GHI estimates using ground measurement data for clear, cloudy and all sky conditions (Table 2).

Table 2: Results of the comparison between hourly GHI estimates from OPT_H2 and measured GHI under cloudy, clear and all sky conditions from 2010 to 2013 at all stations. Relative RMSE and MBE are given in brackets.

Station	MBE Wh/m² (%)	RMSE Wh/m² (%)	R
cloudy sky conditions			
Saint Georges	−97.1 (−61)	133.6 (84)	0.62
Rochambeau	−113.5 (−65)	150.8 (86)	0.62
Kourou	−101.9 (−62)	147.0 (81)	0.66
Ile Royale	−109.0 (−62)	146.9 (84)	0.64
Saint Laurent	−130.1 (−64)	159.8 (82)	0.72
Maripasoula	−123.1 (−49)	162.3 (66)	0.78
clear sky conditions			
Saint Georges	−85.4 (−15)	123.3 (22)	0.94
Rochambeau	−103.5 (−19)	136.1 (24)	0.94
Kourou	−96.0 (−16)	118.6 (20)	0.96
Ile Royale	−109.4 (−18)	140.1 (22)	0.96
Saint Laurent	−133.6 (−23)	159.4 (28)	0.95
Maripasoula	−63.3 (−11)	113.3 (19)	0.94
all sky conditions			
Saint Georges	−44.5 (−11)	130.3 (31)	0.89
Rochambeau	−55.5 (−13)	133.1 (30)	0.90
Kourou	−48.2 (−10)	120.6 (26)	0.92
Ile Royale	−40.6 (−8)	112.0 (23)	0.94
Saint Laurent	−63.8 (−15)	134.3 (31)	0.90
Maripasoula	−37.3 (−8)	125.5 (27)	0.90

For clear, cloudy and all sky conditions, biases were negative, indicating that OPT_H2 underestimated the GHI. Under clear skies, MBE ranged from −23% to −10% for the study period, and RMSE ranged from 19% to 28%. High biases found in clear skies are explained by two phenomena: The first phenomenon is the occurrence of clouds on days that are on average clear. Fig 1 show time series of GHI for two days on each station. One can notice for instance that clouds occur over Saint Georges (Julian day 259) around 2 pm UTC because of the high discrepancies between the clear sky modelled GHI and the ground measured GHI. The bias between OPT_H2 and the ground measurement is approximately 50 Wh/m². The second phenomenon that explains high biases found in clear skies is TL. TL describes the optical thickness of the atmosphere due to both the absorption by the water vapour and the absorption and scattering by the aerosols. Tropical climate is characterised by hot and humid weather conditions; therefore, the water vapour modelling heavily influence the GHI (Janjai et al., 2005). However, available data are monthly climatological values. According to (Rigollier

et al., 2000) a change of 1 in *T*L leads to a relative change of approximately 10–15% in clear-sky GHI estimates.

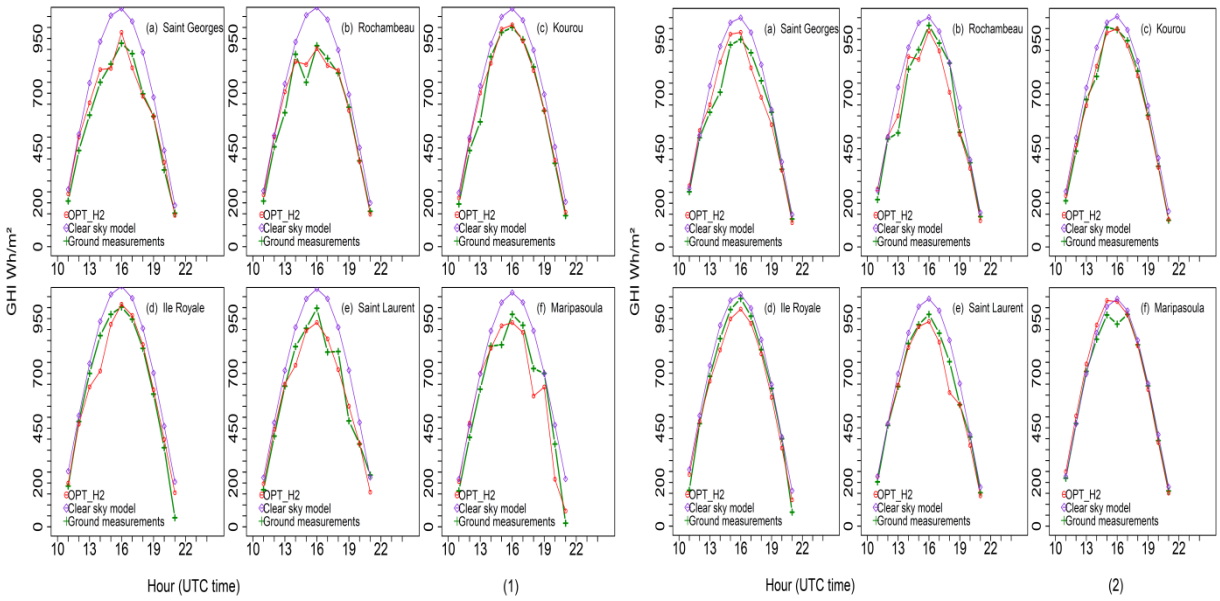

Fig. 1. GHI time series of two days where all stations experienced on average a clear sky day. (1) is for Julian day 259 (September 16th) on year 2010, (2) is for Julian day 277 (October 4th) on year 2013.

For all sky conditions, MBE and RMSE ranged from −15% to −8% and from 22% to 32%, respectively, while the correlation coefficient R ranged from 0.89 to 0.94. Under cloudy sky conditions, MBE increased by more than three times, with values ranging from −48% to −64%. RMSE also increased in this case, usually by almost four times, with values ranging from 66% to 84%. The correlation coefficient, a measure of the similarity between the irradiance estimated by the model and that measured by the pyranometers, was close to maximum under clear sky conditions. For each year and station, the correlation coefficient was better than 0.93 under clear sky conditions. Under cloudy sky conditions, there was a significant decrease in the correlation coefficient values: the highest value was 0.78 and the lowest value was 0.62.

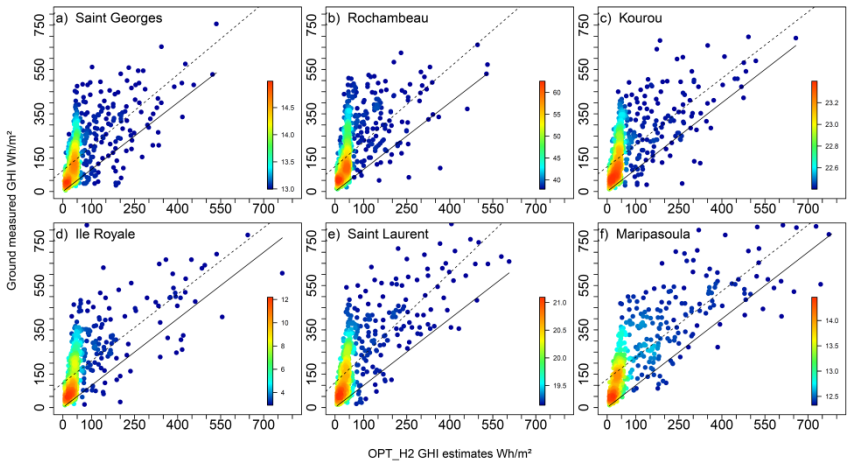

Fig. 2. Measured GHI versus OPT_H2 GHI estimates under cloudy conditions from 2010 to 2013 for all stations. The full line is the identity line and the dashed line is the linear regression between the

ground measurements and *OPT_H2*

As seen in Table 2 and Figure 2, under cloudy sky conditions, there was a large discrepancy between GHI estimates from the OPT_H2 model and the measured values. To address this discrepancy, we coupled the optimised model with an RTP, as described in Section 4.2.

4.2 Quality of GHI estimates from the OPT_H2 method coupled with an RTP

Using the methodology described in Section 3.2, the cloud absorption to effective cloud albedo proportionality coefficient, a, was computed to describe the effect of cumulonimbus clouds over French Guiana (Marie-Joseph et al., 2013). Equation (15) is minimised when $\alpha_j = a = 0.165$. This value was substituted in equation (14) to yield the following:

$$\alpha_c = 0.165\, \rho_{eff}. \quad (22)$$

This value of α_c was integrated in the cloudy sky RTP, and GHI estimates were generated with this parameterisation for the years 2010–2013 (Table 3).

Table 3: Results of the comparison between hourly GHI estimates from RTP_OPT_H2 and measured GHI under clear, cloudy and all sky conditions from 2010 to 2013. Relative RMSE and MBE are given in brackets.

Station	MBE Wh/m² (%)	RMSE Wh/m² (%)	R
cloudy sky conditions			
Saint Georges	−4.1 (−2)	99.1 (62)	0.62
Rochambeau	−19.4 (−11)	108.0 (62)	0.60
Kourou	−22.6 (−12)	103.2 (57)	0.66
Ile Royale	−20.1 (−11.14)	102.2 (58)	0.67
Saint Laurent	−42.9 (−21)	109.0 (54)	0.74
Maripasoula	−49.3 (−20)	108.1 (43)	0.84
clear sky conditions			
Saint Georges	−70.0 (−12)	106.4 (18)	0.94
Rochambeau	−86.8 (−15)	117.1 (20)	0.95
Kourou	−86.8 (−14)	105.7 (17)	0.97
Ile Royale	−100.0 (−16)	126.8 (20)	0.97
Saint Laurent	−116.2 (−20)	139.7 (24)	0.96
Maripasoula	−51.7 (−9)	101.3 (17)	0.94
all sky conditions			
Saint Georges	−9.4 (−2)	107.8 (26)	0.91
Rochambeau	−21.7 (−5)	108.1 (24)	0.92
Kourou	−20.7 (−4)	99.7 (21)	0.94
Ile Royale	−17.5 (−4)	94.2 (19)	0.95
Saint Laurent	−30.4 (−7)	109.0 (25)	0.92

Maripasoula	−11.7 (−3)	108.6 (24)	0.92

4.3 GHI estimates under cloudy sky conditions

Table 3 shows the impact of coupling a cloudy sky RTP with OPT_H2 under cloudy sky conditions.

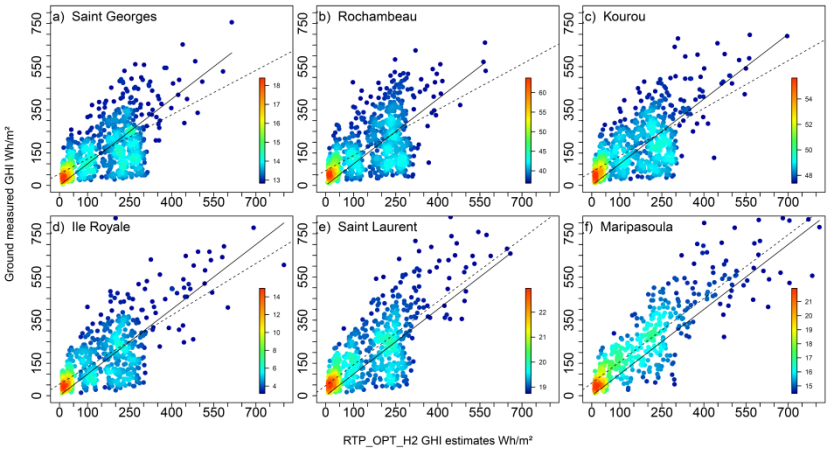

Fig. 3. All stations' GHI ground-measured data versus RTP_OPT_H2 estimates under cloudy sky conditions from 2010 to 2013. The full line is the identity line and the dashed line is the linear regression between the ground measurements and *RTP_OPT_H2*

Figure 3 shows the relationship between GHI from ground measurements and from OPT_H2. Comparing Figures 2 and 3, one can see that the gap between the *identity line and the* linear regression line between the ground measurements and OPT_H2 *has decreased, and in some cases, they intersect*.

There was an overall decrease in both MBE and RMSE, and the correlation coefficient improved for most stations. The RTP_OPT_H2 MBE ranged from −2% to −22% over the entire study period, while OPT_H2 MBE ranged from −48% to −65%. A noticeable decrease in the RMSE was also observed; OPT_H2 RMSE ranged from 66% to 87%, while RTP_OPT_H2 RMSE ranged from 43% to 63%.

Figure 4 shows a clear reduction in the gap between the ground-measured values and RTP_OPT_H2 values under cloudy sky conditions for all stations. The clearness index CDF (Fig. 4) show that for these days considered as cloudy on average, at least 50% of the hourly K_t values are lower than 0.2 for each station. K_t values higher than 0.4 represent less than 1%. The accuracy of the GHI estimates for this cloud absorption coefficient parameterisation (22) varies according to the location of the station, as seen in Table 3. One can classify stations into three groups with respect to the magnitude of the improvement: (1) Saint-Georges and Rochambeau, (2) Kourou and Ile Royale and (3) St-Laurent and Maripasoula. This classification is based on rainfall amount in the regions: the first group has a yearly mean rainfall between 3700 and 3500 mm, the second group has a yearly mean rainfall of approximately 3000 mm and the third group has a yearly mean rainfall between 2500 and 2650 mm according to Météo-France (Héritier, 2011). Hong et al. (2004) and Richards and Arkin (1980) noticed significant correlation between cloud groups and rainfall; therefore, the regions within each group may also have a similar cloud regime.

Figure 4 also shows that modelling cloud absorption with RTP_OPT_H2 improved the Heliosat-2 cloudy sky GHI CDF for irradiance values between 42 and 278 Wh/m². These values are obtained my multiplying the normalized value Fig. 4 by the maximum GHI value measured under cloudy sky condition. These values represent approximately the 85th percentile on average for all stations. The highest improvement was noted for irradiance of approximately 42 to 56 Wh/m² (Fig. 4), which, depending on the station, represents approximately 20% to 30% of the cloudy sky samples. However, the correction effect was less obvious for irradiance values of less than 42 Wh/m² or more than 333 Wh/m² (Fig. 4), except at Maripasoula, where noticeable improvements were found for values of up to 472 Wh/m².

The results for percentiles of values higher than 333 Wh/m² (Fig. 4) are explained by the occurrences of clear sky periods on days considered cloudy. The results for percentiles of values lower than 42 Wh/m² correspond to periods near sunrise or sunset (high θ_s), where Gautier and Landsfeld (1997) found discrepancies between their cloudy sky model (used in this study) and ground measurements. Rigollier et al. (2004) also found that Heliosat-2 performs poorly for high θ_s (values greater than or equal to 75°). Differences observed between stations in the intermediate percentiles can be explained by varying cloudiness conditions during the day. An analysis of satellite images containing the ground stations showed that cloud index values can vary rapidly for days considered cloudy; there were few days where the cloud index was constantly high.

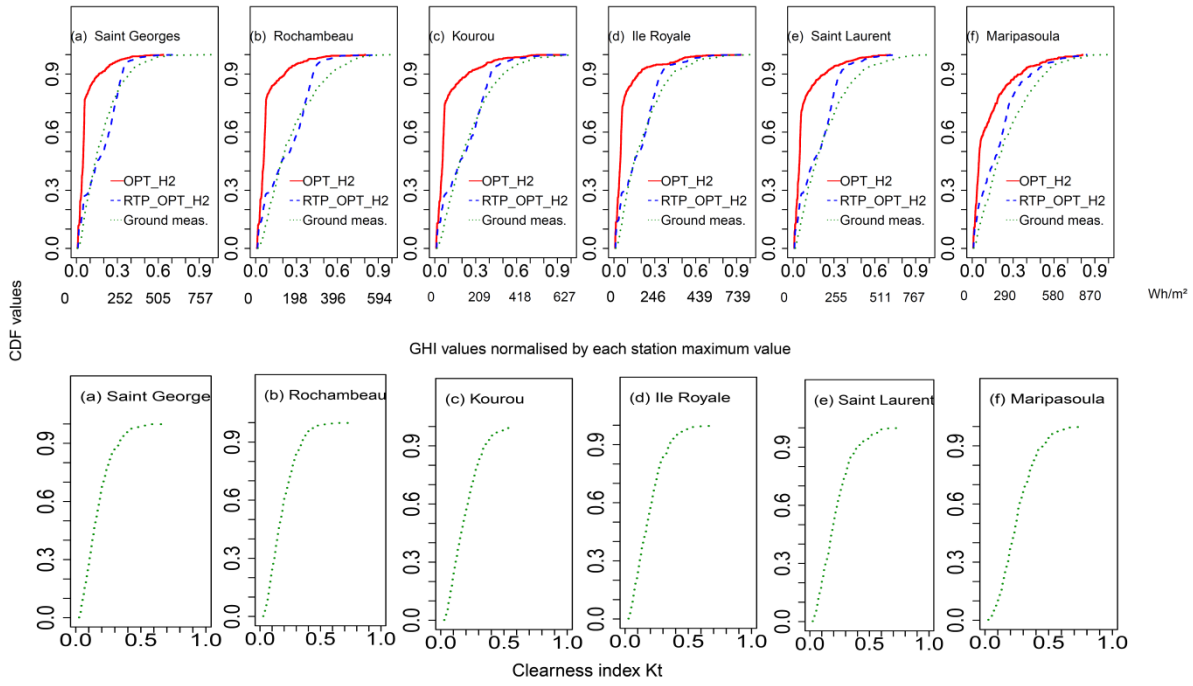

Fig. 4. CDF of GHI estimates from RTP_OPT_H2 and OPT_H2 compared with CDF of measured GHI from 2010 to 2013 under cloudy sky conditions. The clearness index CDF (computed using ground measurements) under cloudy sky conditions is shown alongside the GHI cdfs for each station. The normalized GHI CDFs are obtained by dividing the GHI values by the maximum ground GHI value measured under cloudy sky condition. These normalization values are: for Saint Georges 847 Wh/m², Rochambeau 661 Wh/m², Kourou 697 Wh/m², Iles Royale 822 Wh/m², Saint Laurent 852 Wh/m², Maripasoula 966 Wh/m²

4.4 Quality of GHI estimates under clear sky conditions

One of the objectives for RTP_OPT_H2 development was to not degrade the clear sky surface solar radiation estimates. Table 3 shows significant improvements in RMSE, MBE and correlation coefficients under clear sky conditions. MBE of the GHI estimates from OPT_H2 ranged from −23% to −10%, while that from RTP_OPT_H2 ranged from −8% to −20%. RMSE followed the same trend, ranging from 19% to 28% for OPT_H2 and 16% to 24% for RTP_OPT_H2. Overall, there was a minor increase (0.01) in the correlation coefficient. To further analyse the impact of RTP_OPT_H2 under clear sky conditions, we considered the CDF of the GHI estimated by RTP_OPT_H2, OPT_H2 and from ground measurements (Fig. 5).

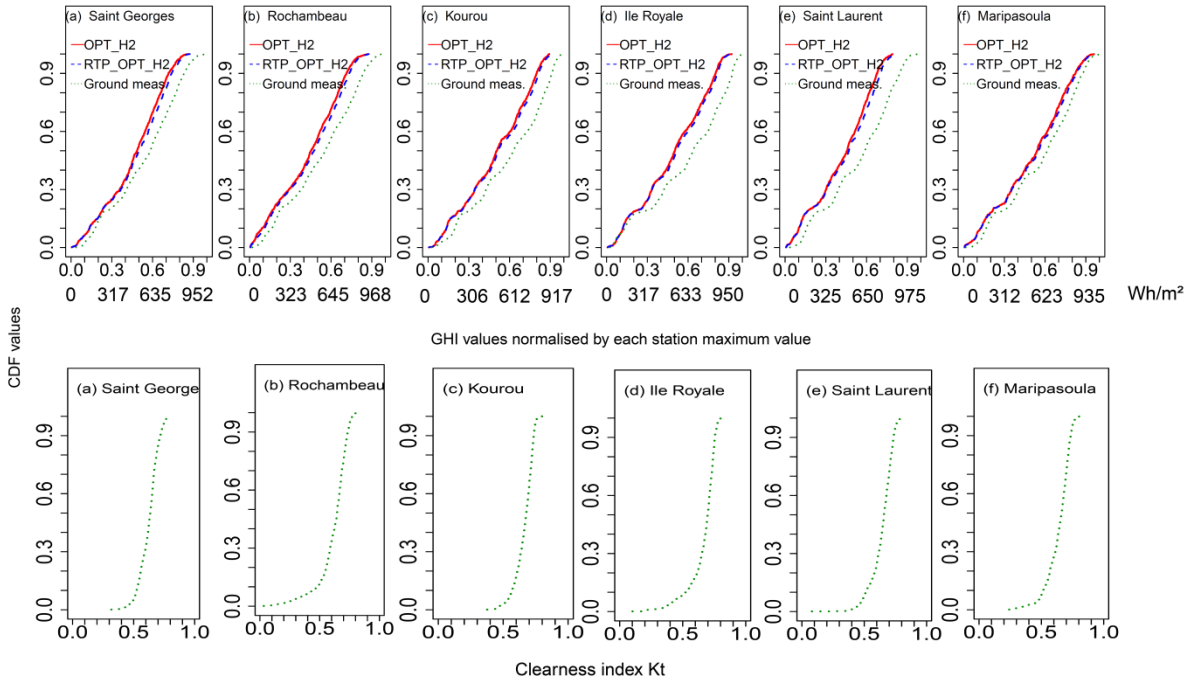

Fig. 5. CDF of GHI estimates from RTP_OPT_H2 and OPT_H2 compared with CDF of measured GHI from 2010 to 2013 under clear sky condition. The clearness index CDF (computed using ground measurements) under clear sky conditions is shown alongside the GHI cdfs for each station. The normalized GHI CDFs are obtained by dividing the GHI values by the maximum ground GHI value measured under clear sky condition. These normalization values are: for Saint Georges 1058 Wh/m², Rochambeau 1075 Wh/m², Kourou 1019 Wh/m², Iles Royale 1055 Wh/m², Saint Laurent 1083 Wh/m², Maripasoula 1038 Wh/m²

Figure 5 shows that, although RTP_OPT_H2 improved GHI estimation from each site similarly regardless of geographical location under clear sky conditions, this improvement was less noticeable than that observed under cloudy sky conditions (Fig. 4). The clearness index CDF (Fig. 5) show that for these days considered as clear on average, at least 50% of the hourly K_t values are higher than 0.65 for each station. K_t values lower than 0.5 represent less than 1%.

4.5 Quality of GHI estimates under all sky conditions

The computation results showed an overall improvement in GHI estimates under all sky conditions (Table 3). MBE of the GHI estimates from RTP_OPT_H2 ranged from −2% to −7% over the entire study period, while that from OPT_H2 ranged from −15% to −8%. RMSE also showed a decrease; RMSE of the GHI estimates from OPT_H2 ranged from 22% to 32%, while that from RTP_OPT_H2

ranged from 19% to 26%. Under all sky conditions, there was a noticeable improvement in the correlation coefficient; lower and higher values for RTP_OPT_H2 were 0.91 and 0.95, respectively, whereas those for OPT_H2 were 0.89 and 0.94, respectively. The clearness index CDF (Fig. 6) show that under all sky conditions there is a high occurrence of clouds, at least 50% of the hourly K_t values are lower than 0.5 for each station. The overall improvement under all sky conditions may be explained by improved cloud feature characterisation.

Previously, cloud optical depth was not modelled in OPT_H2 by K_c but is now modelled in RTP_OPT_H2 using cloud absorption α_c as a proxy for cloud optical depth (Gautier and Landsfeld, 1997). Welch et al. (1980) showed that when cloud optical depth decreases, cloud transmittance increases, and vice versa. Using the cloudy sky parameterisation allowed tweaking of the cloud optical depth to match the ITCZ regional cloud properties. To further analyse RTP_OPT_H2's effect under all sky conditions, the CDF of the GHI from RTP_OPT_H2, OPT_H2 and ground measurements was considered (Fig. 6).

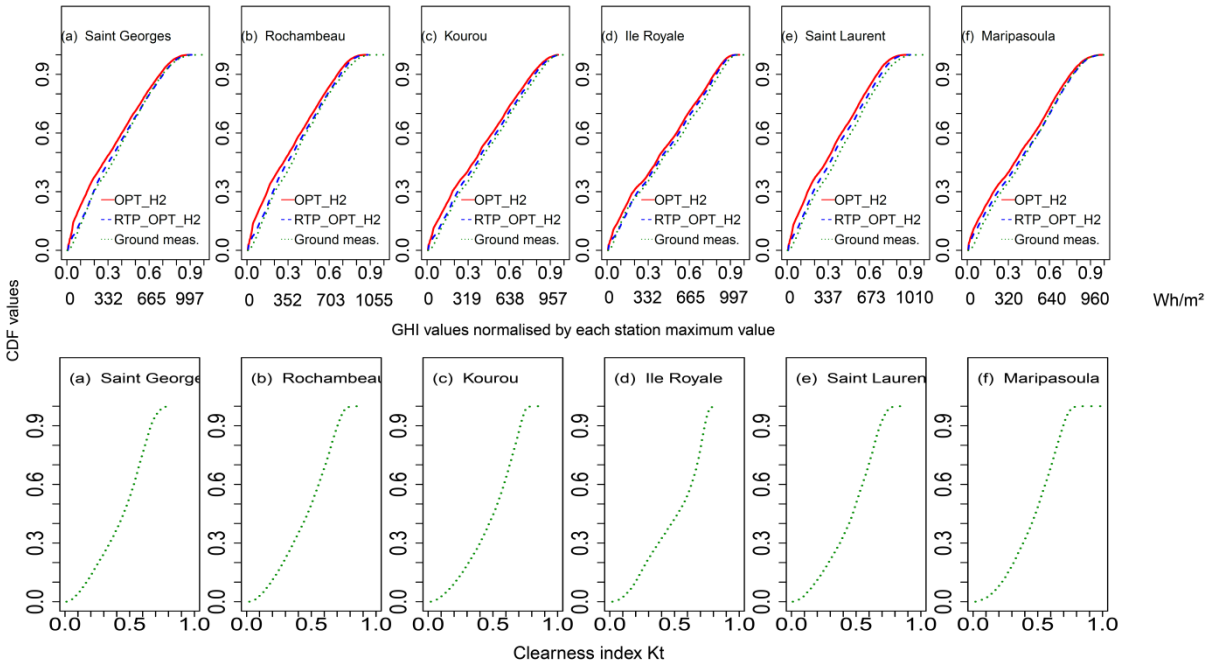

Fig. 6. CDF of GHI estimates from RTP_OPT_H2 and OPT_H2 compared with the CDF of measured GHI from 2010 to 2013 under all sky conditions. The clearness index CDF (computed using ground measurements) under all sky conditions is shown alongside the GHI cdfs for each station. The normalized GHI CDFs are obtained by dividing the GHI values by the maximum ground GHI value measured under all sky condition. These normalization values are: for Saint Georges 1108 Wh/m², Rochambeau 1172 Wh/m², Kourou 1063 Wh/m², Iles Royale 1108 Wh/m², Saint Laurent 1122 Wh/m², Maripasoula 1066 Wh/m²

Table 4: KSI between CDF of GHI from RTP_OPT_H2, OPT_H2 and ground measurements under clear, cloudy and all sky conditions from 2010 to 2013. These values are obtained my multiplying the normalized value Fig. 4-6 by the maximum GHI value measured in the station under the considered sky condition

Station	KSI between OPT_H2 GHI estimates and ground measurements 2010–2013 (Wh/m²)			KSI between RTP_OPT_H2 GHI estimates and ground measurements 2010–2013 (Wh/m²)		
	Cloudy sky	Clear sky	All sky	Cloudy sky	Clear sky	All sky
Saint Georges	63.7	36.1	18.1	15.9	29.1	4.7
Rochambeau	57.3	42.5	22.7	15.2	34.7	8.6
Kourou	58.8	40.9	18.1	11.6	35.8	7.6
Ile Royale	67.5	43.3	15.3	13.3	38.8	7.8
Saint Laurent	76.2	58.7	26.4	20.7	50.6	12.5
Maripasoula	72.7	25.8	13.6	28.8	21.3	5.4

Figure 6 shows a clear reduction in the discrepancy between ground-measured and RTP_OPT_H2 GHI values under all sky conditions for all stations, similar to the improvements noted for clear and cloudy sky conditions. Table 4 shows KSI for both OPT_H2 and RTP_OPT_H2 computed for the six stations under clear, cloudy and all sky conditions using CDF. One can see that the magnitude of the improvement clearly distinguishes the three groups described in Section 4.2.1. The first group (Saint Georges and Rochambeau) showed the greatest improvement, followed by the second and third groups. There was also an improvement under clear sky conditions, albeit lower than that under cloudy sky conditions. This is explained by the occurrence of clouds during days that are, on average, clear, as defined in equation (21).

5 Conclusions

In this article, we have attempted to improve GHI estimates under cloudy sky conditions encountered in tropical areas. We coupled a cloudy sky RTP that included cloud features in an optimised version of the Heliosat-2 method calibrated with GOES images. Using this method improved GHI estimation accuracy in the ITCZ, a zone characterised by high cloudiness.

The novelty of this modification to the Heliosat-2 method is its design, which (i) is based on a cloudy sky RTP with only one parameter—the cloud absorption and (ii) does not increase computation time. The bias improvement of GHI estimates under cloudy sky conditions lies between 25% and 39%. RTP_OPT_H2 had a greater effect on bias; for the entire dataset, under cloudy sky conditions, the bias changed from a mean of −60% for OPT_H2 to a mean of −16% for RTP_OPT_H2. This improvement under cloudy sky conditions has a direct effect on that under all sky conditions; the overall MBE and RMSE changed from mean values (based on all years and stations combined) of −11% and 28%, respectively, for OPT_H2 to −4% and 23%, respectively, for RTP_OPT_H2. This improvement is explained by the improved cloud feature characterisation resulting from the cloudy sky parameterisation integration in the GHI formula. KSI was also computed for all stations under clear, cloudy and all sky conditions.

RTP_OPT_H2 was found to be highly sensitive to cloud absorption parameterisation. We have shown that the magnitude of RTP_OPT_H2 bias improvement varied between three groups: (1) Saint-Georges and Rochambeau, (2) Kourou and Ile Royale and (3) St-Laurent and Maripasoula. This phenomenon can be explained by the similar cloud regimes amongst these groups. Experiments

showed that better results are obtained if cloud absorption is tweaked for each group of stations. Nevertheless, the cloud absorption value for cumulonimbus clouds was found to be good enough to describe the cloud regimes in French Guiana. Thus, we have shown that, despite having dynamic cloud cover, a single cloud albedo and cloud absorption proportionality coefficient can be used at different sites to provide good results.

The scope for future research involves creating a monthly map of cloud absorption, first, for French Guiana, and then, for other tropical areas using several satellite spectral channels and/or ground measurements to identify cloud types (Tapakis and Charalambides, 2013). Once the cloud type is known, the cloud absorption can be computed using Liou, (1976). To conclude, our Heliosat-2 method with RTP coupling can be applied universally and at various locations, provided the correct cloud absorption parameterisation, describing the cloud regime at the said location, is used.

Acknowledgments

The authors are grateful to Météo-France for supplying ground data and to NOAA for supplying satellite data via CLASS.

III. Summary and conclusion of Chapter II

In Chapter II we presented first the definition of radiometric quantities. Using these definitions we discussed methods used to model the GHI at the ground and how they can be validated against quality control ground measurements. These methods either solve accurately or approximate the RTE. Approximate methods answer a major concern that is to find a good compromise between the accuracy of the GHI estimates and the computation cost. Approximate method can either be classified as physical, semi-physical or statistical parameterizations. A key component for modelling the GHI in the ITZ is the ability of a model to describe cloud coverage and cloud content. Physical parameterizations account for cloud properties; however, they do not account for cloud coverage as statistical satellite methods. For this reason we presented in this chapter a new method that couple a statistical satellite methods (ie, Heliosat-II) and a physical parameterization for cloudy sky conditions. This methodology (RTP_OPT_H2) was validated against ground measurements and the native Heliosat-II (OPT_H2) calibrated with GOES images. We found that the bias improvement of GHI estimates under cloudy sky conditions lies between 25% and 39%. We also found that the improvement under cloudy sky conditions has a direct effect on the GHI estimates under all sky conditions The overall MBE and RMSE changed from mean values (based on all years and stations combined) of −11% and 28%, respectively, for OPT_H2 to −4% and 23%, respectively, for RTP_OPT_H2. Consequently, we have shown that our method improves the GHI estimates under cloudy and under all sky condition in French Guiana. The scope for future research includes extend this methodology to other location in the intertropical zone and provide accurate GHI estimate to improve solar facilities design and siting in in the ITZ.

CHAPTER III : Assessing the accuracy of numerical weather prediction model in the intertropical zone

I. Forecasting the solar irradiance at the ground

I.1 Recommendation for solar energy forecast method depending on the horizon

Knowledge of the global horizontal irradiance (GHI), its temporal and spatial distribution is crucial for many purposes, such as electricity production (Clack, 2017), crop growth monitoring (Campillo et al., 2012), and energy efficient building design (Oldewurtel et al., 2012). Several countries, including Australia (Clean Energy Council, 2017), Costa Rica (Nandwani, 2006), France (Ministry of the environment, 2016b), Germany (Lehr et al., 2006), the United States (O'Connor and Cleveland, 2015), and Sweden (Centeno López, 2016) have begun an energy transition. A primary goal of this energy transition is to increase the share of electricity produced from solar and or wind energy and decrease the share of electricity produced from fuel and or nuclear energy. However, the integration into the grid of electricity generated by intermittent renewable energies present economic and technical challenges (IER, 2013). To ensure grid stability and safety: accurate forecasts of solar irradiance and weather conditions up to 48 hours ahead are required (Dambreville et al., 2014). Many models have been developed to assess the GHI at the ground depending upon the forecast horizon (Diagne et al., 2013; Glassley et al., 2012; Inman et al., 2013; Kleissl, 2013; Pelland et al., 2013). These methods can be divided into three categories: ground based, satellite based and NWP based Fig. III.1.

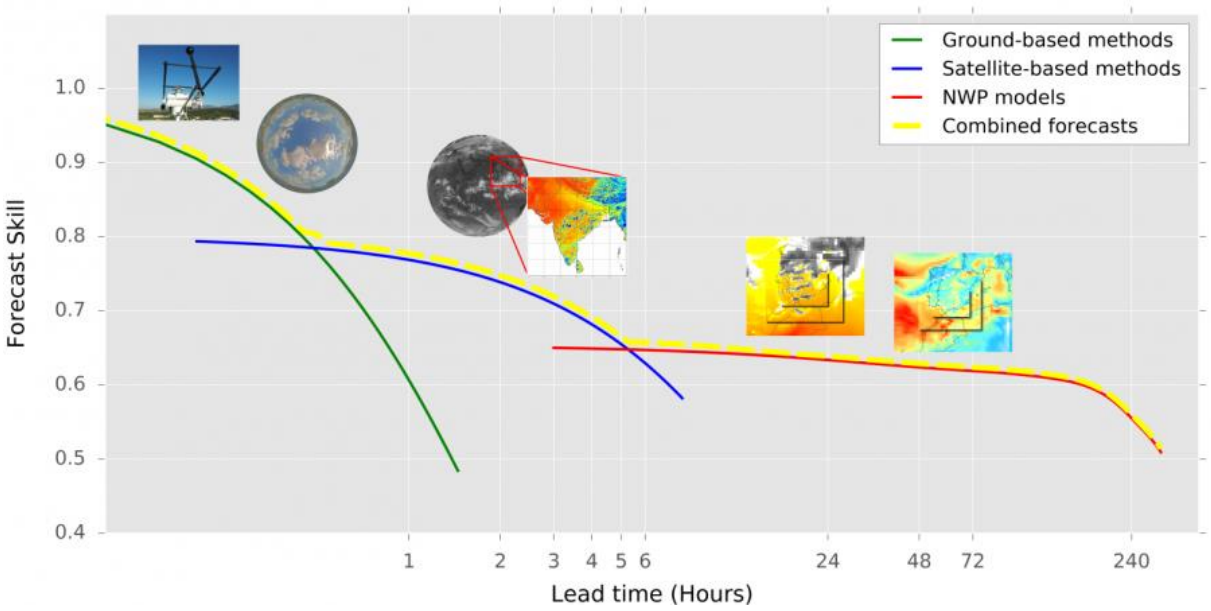

Fig. III.1 Conceptual diagram of forecast skill as a function of forecast lead time for different forecast methods. Extracted from (Kleissl, 2013; Ruiz-Arias and Goenka, 2017)

For intra hour GHI forecasts the use of sky imagers is suggested (Inman et al., 2013). Sky imagers forecast the GHI based on a cloud advection approach. First cloud motion vectors are generated by cross-correlating two consecutive sky images (Chow et al., 2011). Second the cloud images are then propagated forward in time using the motion vector previously defined. The second step results in a forecast of the cloud position (Quesada-Ruiz et al., 2014). Therefore, using a clear sky model and knowing the location of clouds, a GHI forecast can be obtained. The main challenge of these imagers forecasts are: the cloud detection method (Fu and Cheng, 2013; Yang et al., 2015) and the computation of the 3D-cloud map extrapolated from sky imagers images. The main advantage of

sky imagers forecasts is that they have higher spatial and temporal resolution than numerical weather prediction and satellite-based forecasts (Fu and Cheng, 2013). The temporal resolution is typically inferior or equal to 1 min whereas the spatial resolution is determined by the sky imager field of view (Chow et al., 2011; Marquez and Coimbra, 2013; Quesada-Ruiz et al., 2014; Schmidt et al., 2015). Using sky imagers to forecast the GHI Schmidt et al., (2015) found for a forecast horizon of 10 min a RMSE ranging between 70 and 250 W/m². for a 5 minute forecast horizon, Quesada-Ruiz et al.(2014) found that overall the RMSE was 105.26 W/m². For a 15-minute forecast horizon (Marquez and Coimbra, 2013) found a RMSE ranging between 299 and 401 W/m². Marquez et al. (2013) showed that intra hours forecasting using stochastic learning technic is possible, for a 30 min forecast horizon he found out a 1.78 W/m². The main disadvantage of this method is their reliance on training datasets (Mathiesen and Kleissl, 2011).

For intraday, between 1h and 6h forecast horizon, the use of satellite method is suggested (Diagne et al., 2013; Inman et al., 2013). Nevertheless, stochastic learning technics can be used forecast the GHI up to 2h ahead (Pedro and Coimbra, 2012). Satellite methods are also based on a cloud motion forecasts approach. They allow a greater forecast range because of their spatial resolution that is coarser than those of sky imagers (Diagne et al., 2014). Satellite advection forecast extends up to 6h at a resolution of 1 km² (C. W. Chow et al., 2011). The main challenges for satellite-based advection based forecasts are the cloud detection algorithm and cloud height determination (Qu, 2013). Perez et al. (2010) found that the RMSE of 6h GHI forecasts using a satellite method ranges between -31% and 38%.

Beyond 6h forecast horizon, cloud advection which is observed through satellite images become less important than cloud development and dissipation. NWP models describe the atmospheric state using conservation equations; therefore they account for dissipation phenomenon (Hamill and Nehrkorn, 1993). Consequently, for forecast horizon larger than 6h, the use of numerical weather prediction (NWP) model is suggested (Diagne et al., 2013; Inman et al., 2013). In this manuscript we aim to forecast the GHI up to 48 h ahead; therefore, we are interested in using NWP to assess the GHI. In the following section we present the different type of NWP models and how they describe the atmospheric state.

I.2 Introduction to Numerical weather prediction models

A numerical weather prediction model is a set of non-linear equations that are discretized and resolved using numerical algorithms (Coiffier, 2011). These equations model the physical process that occur in the atmosphere and its interaction with the ocean and land surface (Fig. III.2). These physical processes account for: the transfer of momentum and temperature by advection, convection, small-scale turbulent motions (diffusion) as well as the selective absorption and emission of radiation and the release of latent heat accompanying condensation (Schlesinger, 1988). According to Jacobson, (2005) Molecular diffusion occur over distances much smaller than 2 mm whereas advection motion such as thunderstorm occur over distances of 2–2000 km. This wide disparity of length scale introduces modeling and computation feasibility issues. To solve this issue NWP models were divided into two categories:

-global NWP models
-mesoscale models

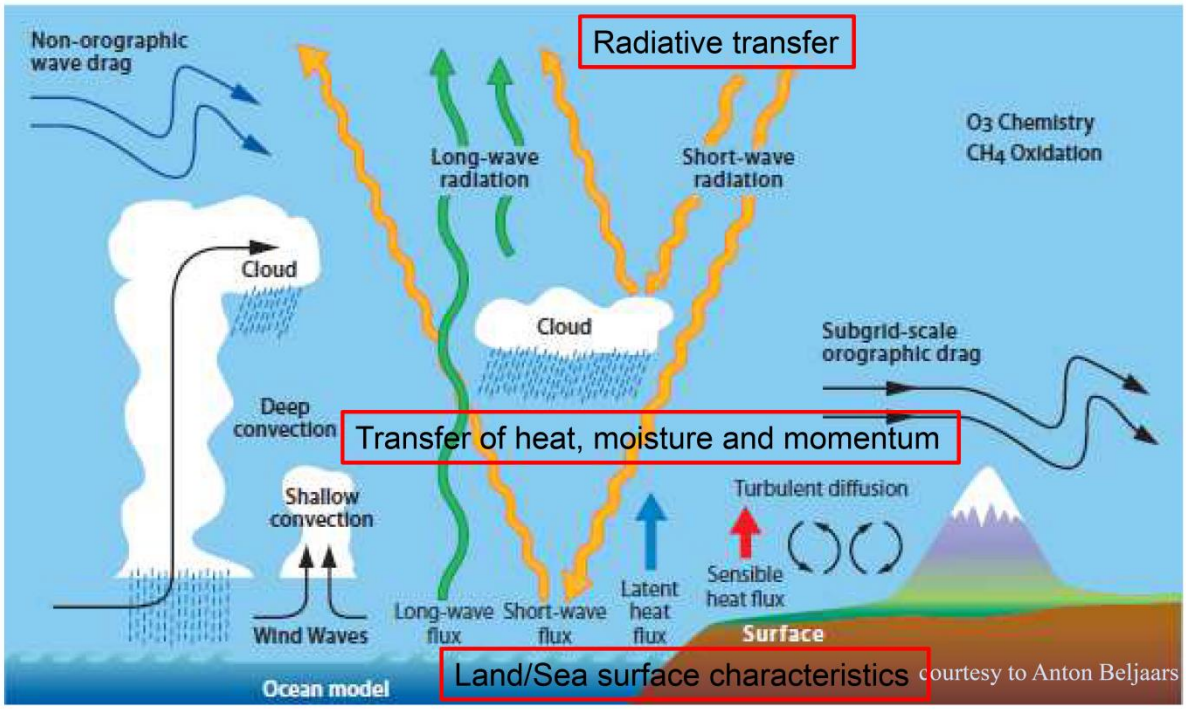

Fig. III.2 Physical process accounted for in NWP models extracted from (Helmert, 2016)

Global models, on one hand forecast the state of the entire earth atmosphere; they have a typical resolution of 16-50 km. Mesoscale models, on the other forecast the state of a portion of earth atmosphere. The spatial resolution of mesoscale models is defined by the user with respect to the computing power; it usually ranges between 5 and 20 km (Diagne et al., 2014). Mesoscale models take global model forecasts as initials and boundary conditions inputs so they can account for large scale phenomena; they do not include an ocean model. Yet the resolution of current NWP models is still too coarse to resolve the micro-scale physics associated with cloud formations (Inman et al., 2013). Because of NWP model coarse grid resolution, sub-grid scale (i.e unresolved) physical process must be approximated (hereafter, parameterized). Consequently, parameterization are used so that the physical effect of smaller scale processes are accounted for on larger scale (Kleissl, 2013). Physical parameterization scheme can be divided in five categories (Dudhia, 2014; Kleissl, 2013; Schlesinger, 1988; Stensrud, 2009):

-microphysics (or stratiform-microphysic)
-cumulus (or convective)
-radiation (or radiative transfer)
-surface (land surface-atmosphere)
-planetary boundary layer (or turbulent transport)

Microphysics scheme governs the formation of hydrometeors for stratiform clouds, and their growth to precipitation-size raindrops and snow particles. Cumulus scheme intend to represent vertical fluxes and latent heat due to unresolved upward or downward moving air current. Radiation scheme compute absorption by gases as well as scattering and absorption by particles. Surface scheme model the exchange of mass and energy between land surface and the atmosphere with respect to the land-surface properties. The planetary boundary layer scheme represents the small scale turbulence that transports heat and moisture in the vertical, thereby creating the conditions needed

for saturation and cloud formation. These parameterizations are not independent, they interact directly as shown Fig. III.3.

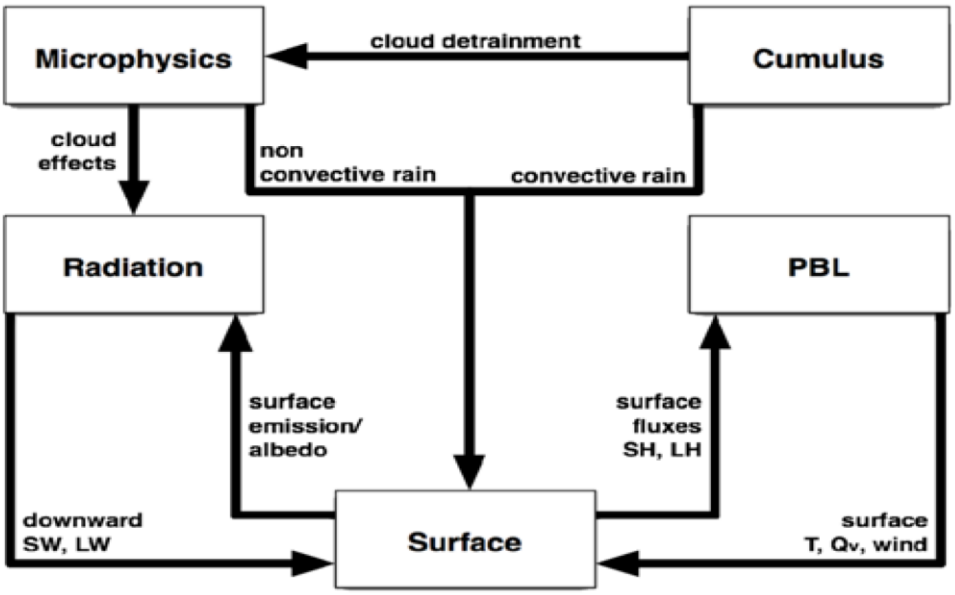

Fig. III.3 Interaction between physical parameterization. Extracted from (Dudhia, 2014)

The advantage of global models is that there GHI forecasts are available globally and freely for most cases; they do not require any computing facilities. There drawback is that these forecasts are available on grid resolutions greater than 10 km and with frequency greater or equal to 1-hour. Tab. III.1 shows the spatial and temporal resolution of GHI forecasts from global NWP for that are available for download. Among these models, the uncertainties of IFS and GFS GHI forecasts at the ground are the most well-documented (Mathiesen and Kleissl, 2011; Perez et al., 2013; Remund et al., 2008; Schroedter-Homscheidt et al., 2017; Troccoli and Morcrette, 2014). The aforementioned studies gathered data from a total of 43 stations. A large number of these validation stations represent Mediterranean, oceanic, continental, and arid climates, and are located in the United States, Canada. Nevertheless, these models have not been evaluated yet in the ITZ. Therefore, the accuracy the accuracy of IFS and GFS in French Guiana and in tropical climates is unknown.

Mesoscale model have finer resolution than global because they are ran for smaller scale domains. They have the potential to produce more realistic forecasts than lower resolution models by resolving smaller scale processes (Lorenz et al., 2016). Nevertheless, for a mesoscale model to be used it must be carefully calibrated, using physical parameterization that describes the region climate patterns (Ruiz-Arias et al., 2008). Most studies dealing with mesoscale model calibration concentrate on locations in the USA, Spain and Germany with continental climate(Lara-Fanego et al., 2012; Perez et al., 2013; Prabha and Hoogenboom, 2010; Ruiz-Arias et al., 2013; Zempila et al., 2015). These countries are amongst the world leaders in installed photovoltaic power capacity (Inman et al., 2013).

Tab. III.1 spatial and temporal resolution of GHI forecast products from global model operated by forecast center for the first 48 hour forecasts. The temporal and spatial resolution of these products refers to those of datasets available for download. Source (BAM, 2018; CMC, 2016; DWD, 2018; ECMWF, 2016b; Figueroa et al., 2016; NOAA, 2016; Walters et al., 2017)

NWP	Operational Center	Spatial resol. distrib. forecast	Temporal resol. distrib. forecast	Download link
GFS (GSM v13.0.2): global forecast system	National center for environmental prediction: NCEP	0.25° (~28 km), 46 levels	3 hours	https://rda.ucar.edu/datasets
IFS (cycle 41R2): Integrated forecast system	European center for medium range forecast: ECMWF	0.125° (~14km), 125 levels	1 hour	https://www.ecmwf.int/en/forecasts/datasets
UM : unified model v6.1	Meteorological office: Met Office	~17km, 70 level	1 hour	http://catalogue.ceda.ac.uk
GEM v4.0: Global environmental multi-scale	Canadian Meteorological center: CMC	0.240° (~25 km), 28 levels	1 hour or 3 hours	https://weather.gc.ca
ICON v1.2.3: ICOsahedral Non-hydrostatic	German weather Service: DWD	13 km, 90 levels	1 hour	https://opendata.dwd.de
ARPEGE cycle 43: Action de recherché petite echelle grande échelle	French Weather services: Météo-France	0.50° (~56 km),x levels	3 hours	https://donneespubliques.meteofrance.fr
BAM: Brazilian Global Atmospheric Model	Centro de previso de tempo estudos climaticos: CPTEC	0.18° (~20 km), 6 levels	6 hours	ftp://ftp.cptec.inpe.br/modelos/io/tempo/global

According to Warner, (2010) calibrating a numerical weather prediction model require the following step:

- determine the prevailing physical process in the geographic area of interest
- define the horizontal resolution so that physical process are resolved within the model grid
-define the vertical resolution so it describe accurately vertical structures such as the boundary layers gradients
-define the map projection
-validate the model accuracy for different season
- study the accuracy of the model with respect to the location of the computation domain and domain size.
- Perform tests to determine the sensitivity of the model accuracy to the vertical and horizontal grid increments

Tab. III.2 shows a comparison of commonly used mesoscale model (Dudhia, 2014; WMO, 2014). This comparison focuses on the numerical aspect and radiation parameterization options. Tab. III.2 shows that weather and research forecast model (WRF) (Skamarock et al., 2008) has the numerical methods with higher order of accuracy. According to Wang et al., (2013) high order numerical method achieve better accuracy than low order numerical method for the same grid resolution. WRF (Skamarock et

al., 2008) model is the most documented, besides it has physical parameterization that allows for feedbacks between aerosol, ozone, cloud and radiation which is not the case for all mesoscale NWP.

Tab. III.2 Comparison of commonly used mesoscale model Source: for numerics ((Burridge, 1975; Doms and Baldauf, 2015; Käellen, 1996; Majewski, 2009; NCAR, 2017; Termonia et al., 2018; Tudor et al., 2013; Undén and co-authors, 2002); for Radiation (Doms and Baldauf, 2015; Jimenez et al., 2016; Termonia et al., 2018; Walters et al., 2017; Zubler et al., 2011)

		WRF v3.8	ALADIN cycle 43	COSMO v5.1	HIRLAM V5	Unified model UM v6.0	HRM v2.5
Numerics	temporal integration order of accuracy	3rd order Runge Kutta	2nd order semi-implicit semi-Lagrangian	3rd order Runge Kutta	1st order Eulerian semi implicit	2nd order semi-implicit semi-Lagrangian	Explicit two step Lax-Wendroff 2nd order
	Horizontal advection order of accuracy	5th order default (6th order available)	Semi-Lagrangian. Order not mentionned	5th order default (6th order available)	Eulerian or semi-Lagrangian. Order not mentionned	Semi-Lagrangian. Order not found	Adjusted Lax-Wendroff. Order not mentioned
	Vertical advection order of accuracy	3rd order default (6th order available)	Semi-Lagrangian. Order not mentioned	2nd order Crank-Nicholson	Eulerian or semi-Lagrangian. Order not mentioned	Semi-Lagrangian. Order not mentioned	Adjusted Lax-Wendroff. Order not mentioned
	Horizontal Diffusion	6th order	4th order numerical diffusion	4th Order linear	6th order semi-Lagragian	2nd order	4th order linear diffusion
Radiation	Ozone dataset resolution available	CAM climatology: 2.82 ° latitudinal, monthly	UGAMP climatology: 2.5°*2.5° monthly	Not mentionned	Not mentionned	SPARC-II monthly variable longitude latitude	Not documented
	Climatologic Aerosol dataset available	-Tegen climatology: Monthly, 5°*4° -Thompson and Eidhammer climatology : Monthly, 0.5°* 1.25°	-Tegen climatology	-Tegen climatology -Tanré climatology: ~11°, constant in time	-Tegen climatology	CLASSIC aerosol climatology. resolution not found	Not documented
	Cloud radiation feedback	yes	yes	yes	Not mentioned	yes	Not documented
	Aerosol radiation feedback	yes	yes	yes	yes	yes	Not documented
	Ozone radiation feedback	yes	Yes	yes	yes	yes	Not documented
	cloud aerosol feedback	yes	Yes	Not mentioned	Not mentioned	yes	Not documented

Perez et al., (2013) studied the accuracy of global and mesoscale model for several location in Canada, Europe and the USA. They found that despite their higher resolution the accuracy of hourly GHI forecasts of mesoscale models are similar or worse than those of global models. According (Kotsopoulos et al., 2014; Morrison, 2010) describing more accurately physical phenomena may expose other NWP model deficiencies. For this reason we decided first study the accuracy of global NWP model in the ITZ and for tropical climates. Second, to propose a methodology to calibrate WRF in the ITZ so that it has improved GHI forecasted compared to non-calibrated mesoscale models. This methodology was validated in French Guiana using ground measurements and AROME (Seity et al., 2011) mesoscale model.

II. Studying the GHI forecast accuracy of global NWP models in the intertropical zone

Assessing GFS and IFS global weather prediction and numerical model forecast accuracy in the intertropical zone and for tropical climates[3]

Mouhamet Diallo[a,*], Hadrien Verbois[b, c, d], Fabrice Chane-Ming[e], Antoine Primerose[a], Wilfred Walsh[c], and Laurent Linguet[a]

[a]Université de la Guyane – UMR 228 Espace-Dev, BP 792, 0.275 km Route de Montabo, 97337 Cayenne, French Guiana, France
[b]Graduate School for Engineering and Integrative Science, National University of Singapore, Centre for Life Sciences (CeLS), #05-01, 28 Medical Drive, Singapore 117456
[c]Solar Energy Research Institute, National University of Singapore, PV Module Cluster, No. 1 CleanTech Loop, CleanTech One #06-01, Singapore 637141
[d]Department of Physics, National University of Singapore, Faculty of Science, 2 Science Drive 3, Blk S12, Level 2, Singapore 117551
[e] Université de la Réunion, Laboratoire de l'Atmosphère et des Cyclones, UMR8105, UMR CNRS-Météo France-Université, 15 avenue René Cassin - 97715 ST DENIS Messag Cedex 9, Ile de la Réunion, France

*Corresponding author: mouhametdiallo@gmail.com (M. Diallo)

Abstract

An understanding of available solar resources is vital for monitoring and predicting the yield of solar energy systems. The purpose of this study is to assess the global horizontal irradiance (GHI) forecast accuracy of the Global Forecast System (GFS) and the Integrated Forecast System (IFS) numerical weather prediction (NWP) models in the intertropical zone and for tropical climates. GFS (GSM v13.0.2) and IFS (cycle 41R2) forecast accuracy is validated against GHI measurements during 2016 from twelve meteorological stations located in French Guiana, Reunion Island, and Singapore. The mean average error (MAE), the mean bias error (MBE), the root mean squared error (RMSE), and

[3] Abbreviations: FS, Forecast Skill; GEM, Global Environmental Multiscale; GFS, Global Forecast System; GHI, Global Horizontal Irradiance; IFS, Integrated Forecast System; ITCZ, Intertropical-Convergence Zone; MAE, mean average error; NWP, numerical weather prediction; RMSE, root mean squared error; RRTMG, rapid radiative transfer model-G; UM, Unified Model.

the forecast skill (FS) were used to study: (1) the seasonal and yearly accuracy of IFS and GFS GHI forecasts (2) the bias distribution of both models as a function of the clear sky index and the clear sky index variability (3) the effect of temporally averaging IFS native hourly GHI outputs (IFS-1h) to 3-hour GHI outputs, and (4) the accuracy of both model forecasts relative to other climates. To compare GFS 3-hour forecast outputs to IFS-1h forecast outputs, we considered two approaches. (i) We averaged temporally IFS-1h forecasts following GFS cycle to obtain 3-hour forecasts; these 3-hour output forecasts were named IFS-3h-p. (ii)Starting from 00h UTC we kept every 3 hour forecasts of hourly IFS ; these 3-hour output forecasts were named IFS-3h. Results show that under all sky conditions the GHI forecasts of IFS-3h and IFS-3h-p outperform those of 3-hour GFS in all territories. We found that averaging temporally the GHI forecasts using GFS cycle increase the MAE value; under all sky conditions the relative improvement of the MAE of IFS-3h over the MAE of IFS-3h-p ranges between 21% and 46% depending on the territory. The magnitude of the improvement of the MAE of IFS-3h over the MAE of GFS-3h was found to be 48%, 57%, 37% in French Guiana, Reunion Island, and Singapore, respectively. GFS and IFS behave differently with respect to the sky conditions. IFS-3h and IFS-3h-p underestimate the GHI under clear sky conditions and overestimate the GHI under cloudy and overcast sky conditions. Under clear sky conditions GFS underestimates the GHI only in Reunion Island. Under cloudy and overcast sky conditions GFS overestimate the GHI except for Reunion Island where the GHI under cloudy sky is underestimated. We found that except for GFS in Reunion Island, the GHI forecasts of IFS-3h, IFS-3h-p and GFS are more accurate under clear sky conditions than under overcast and cloudy sky conditions; this result is similar to result found in previous study assessing IFS and GFS accuracies in extra tropical climate. However, we also find that the clear sky index variability has a key role on model accuracies which was not addressed previously. The FS computed for GHI 48 hours ahead in Mediterranean, oceanic, continental, and arid climates is superior, for both GFS and IFS, than in tropical regions. This study provides forecast accuracy metrics that aim to respond to the need for reliable solar power forecasts driven by increasing interest in the development of solar technologies in tropical areas.

Keywords: Global horizontal irradiance; Global Forecast System (GFS); Integrated Forecast System (IFS); French Guiana; La Réunion, Singapore

1 Introduction

The integration into the grid of electricity generated by intermittent renewable energies presents economic and technical challenges (IER, 2013). To deal with the uncontrollable nature of these resources, they are forecasted at different time scales so that the global production can be balanced to match users' consumption (Vallance et al., 2017).

Many models have been developed to assess solar radiation at the ground depending upon the forecast horizon (Diagne et al., 2013; Inman et al., 2013; Kleissl, 2013). For forecast horizons greater than six hours, it is widely accepted that numerical weather prediction (NWP) models are needed (Diagne et al., 2013; Inman et al., 2013). These models forecast either the state of the entirety of Earth's atmosphere, or some part of it. The following global NWP models were developed to forecast the state of the entire Earth atmosphere: the Global Forecast System (GFS) (Environmental Modeling Center, 2003), the Integrated Forecast System (IFS) (ECMWF, 2016b), the Unified Model (UM) (Staniforth et al., 2006), the Global Environmental Multiscale (GEM) (Côté et al., 1998), and the GME (Majewski et al., 2002). Among these models, IFS (Richardson et al., 2013; Haiden et al., 2014, 2015, 2016) and GFS forecast evaluations are the most well-documented (Fanglin, 2013, 2014, 2015).

Many studies have evaluated GFS and IFS solar irradiance forecast uncertainty at the ground by comparison with validation stations (Mathiesen and Kleissl, 2011; Perez et al., 2013; Remund et al., 2008; Schroedter-Homscheidt et al., 2017; Troccoli and Morcrette, 2014). The aforementioned studies gathered data from a total of 43 stations. A large number of these validation stations represent Mediterranean, oceanic, continental, and arid climates, and are located in the United States, Canada, and Europe as these regions are among the world leaders for installed photovoltaic power capacity (Inman et al., 2013).

Therefore, although many studies have evaluated GFS and IFS irradiance estimates at the ground, their performances under several specific climate zones still remain to be evaluated. The present work focuses on assessing the GHI forecast accuracy of GFS and IFS in the intertropical zone for tropical climates. This paper is structured as follows: section 2 describes both the ground and NWP forecast data used in the study; section 3 addresses GFS and IFS monthly and annual accuracy metrics; in section 4, we present our conclusions.

2 Datasets and Methods

In this section, we first describe the downloaded GHI data and the operative setup used in GFS and IFS models to produce these forecasts. Second, we describe the GHI measurements and the quality check procedure followed prior to using the data to validate IFS and GFS forecasts. Lastly, we describe the climates of French Guiana, Reunion Island, and Singapore as well as the microclimate at each measurement station.

2.1 Numerical Weather Prediction models

GFS (GSM v13.0.2) and IFS (cycle 41R2) forecasts used in this study were downloaded online respectively from the National Center for Atmospheric Research data archive ([dataset] NCEP-NWS-NOAA-USDC, 2015) and the European Center for Medium-Range Weather Forecasts (ECMWF) Meteorological Archival and Retrieval System (MARS) ([dataset] Maass, 2017). The downloaded datasets began from January 1st and ended December 31st of 2016. Both model GHI forecasts originated at 00:00 UTC, and were run for 48 hours. The primary IFS and GFS setups used to produce GHI datasets are presented below.

GFS was initialized using a hybrid 3D variational ensemble method (Buehner et al., 2013) with a 6-hour data assimilation window. GFS horizontal and vertical computational resolutions were T1534 (~13 km) and 64 levels. GHI forecasts were distributed with a lower resolution; the dataset exploited in this study was 0.25° (~28 km), 46 levels of spatial resolution, and a 3-hour temporal resolution. GFS GHI forecasts (W/m²) alternated between 3- and 6-hour averages (NOAA, 2016). GHI values (W/m²) at 03:00, 09:00, 15:00, and 18:00 UTC were the means of 3-hour irradiance values ending at forecast time. GHI values (W/m²) at 06:00, 12:00, 18:00, and 24:00 UTC were the means of 6-hour irradiance values ending at forecast time. IFS was initialized using a 4D variational method, with a 12-hour data assimilation window (ECMWF, 2016c). The horizontal and vertical computational resolution were TCo1279 (~9 km) and 137 levels, respectively. IFS GHI forecasts were also distributed with a lower resolution; the dataset exploited in this study was 0.125°, 125 levels of spatial resolution, and a 1-hour temporal resolution. GHI values were accumulated from the forecast initialization (Hogan, 2015). Both IFS and GFS used rapid radiative transfer model-G (RRTMG), (Mlawer et al., 1997) to parameterize radiative transfer for longwave and shortwave radiations; the scheme was called hourly by the dynamic core of each model.

Under cloudless skies, irradiance forecast accuracy is driven by aerosols (Jimenez et al., 2015). GFS models the effects of aerosols on the troposphere and stratosphere separately (DTC, 2016). In the troposphere, aerosol data are climatological values with a 5° horizontal resolution (Chin

et al., 2000; Hess et al., 1998). In the stratosphere, climatological aerosol values derived from Sato et al. (1993) were used, and they varied with latitude, and exhibited an uneven resolution greater than or equal to 30° depending upon the zonal band. GFS aerosols were divided into 11 types: insoluble, water soluble, soot, sea salt (2 types), minerals (3 types), transported minerals, sulfate, and stratospheric volcanic aerosols. IFS models aerosols using climatological values (Tegen et al., 1997); the stratosphere and troposphere are not treated separately as in GFS. The aerosol data used were of six types: organic carbon, soot, sulfate, sea salt, minerals, and stratospheric volcanic aerosols. Aerosol data had 5x4 longitudinal grids, latitudinal spatial resolution, height variation (59 levels), and monthly temporal variations.

Under cloudy skies, irradiance forecast accuracy is driven by cloud cover and content (Welch et al., 1980). The planetary boundary layer (i.e., turbulent transport) parameterization scheme and the cloud microphysics scheme have a strong influence on the distribution and physical properties of the simulated cloud fields (Cintineo et al., 2014; Otkin and Greenwald, 2008; Xie et al., 2012). The microphysics scheme governs the formation of hydrometeors that scatter and absorb radiation and describes their growth to precipitation-size raindrops and snow particles. In the NWP model, microphysics schemes are developed separately for stratiform and cumulus clouds (Kleissl, 2013). In this study, GFS modeled stratiform microphysics after the scheme of Zhao et al. (1997), which uses prognostic equations for cloud vapor, liquid water, and cloud ice. The precipitation types included rain, freezing rain, and snow. Liquid and ice phases do not coexist in this model. Deep cumulus clouds were parametrized after the bulk mass flux scheme of Arakawa and Schubert (1993). The shallow cumulus scheme used was modified from Arakawa and Schubert (1973) as described by Han and Pan (2011). Both deep and shallow cumulus schemes use prognostic equations to compute the cloud vapor and liquid water. IFS stratiform microphysics were evaluated after the methods of Forbes et al. (2011), and this scheme forecasts cloud liquid water, cloud ice, rain, snow and fraction. The precipitation types included rain, freezing rain, and snow. Deep and shallow cumulus microphysics were parametrized by a bulk mass flux scheme originally described by Tiedtke (1989). They both use a prognostic equation to compute the cloud vapor and liquid water.

The turbulent transport scheme used parameterizes the sub-grid-scale vertical transfer of heat, moisture, and momentum between the surface and atmosphere. Therefore, it interacts with cumulus microphysics to produce the cloud field (Xie et al., 2012). Vertical turbulent transport was treated differently in the surface layer and above. In the surface layer, the turbulence fluxes were based on the Monin Obukhov similarity theory both for IFS and GFS (DTC, 2016; ECMWF, 2016a). Above the surface layer, IFS used a weakly unstable boundary layer with a K-diffusion turbulence closure (ECMWF, 2016a), whereas GFS used an eddy diffusivity counter-gradient parameterization (Hong and Pan, 1996). For unstable boundary layers, both GFS and IFS used eddy diffusivity mass flux schemes, as described by Han et al. (2016) and Köhler et al. (2011), respectively.

Cloud radiation feedback was taken into account by GFS (GSM v13.0.2) and IFS (cycle 41R2) using the values for liquid, ice, and snow water contents from the microphysics scheme and a cloud fraction parameterization (DTC, 2016; ECMWF, 2016a). GFS and IFS assume that each grid box is either entirely filled with clouds or entirely clear for any given time step. A cloud fraction parameterization scheme assigns the fractional volume of a grid box that is occupied by clouds (Kleissl, 2013). IFS uses a prognostic equation to compute cloud fraction (Tiedtke, 1993), whereas GFS diagnoses cloud fraction from liquid-water content (Xu and Randall, 1996). Both GFS and IFS assume maximum cloud overlap.

2.2 Measurement stations

To assess GFS and IFS forecast accuracy for tropical climates, we selected GHI measurements from three tropical locations: (1) French Guiana, located in South America between the latitudes of 2° and 6° N, and longitudes of 51 ° and 55° W, (2) Reunion Island, located in the Southwest Indian Ocean (21° S, 55° E), and (3) Singapore, located in Southeast (1.37° N, 103.87° E). Our study exploits

hourly GHI data from twelve stations described in Table 1; these data extend from January 1st to December 31st of 2016. First the GHI value before sunrise and after sunset were removed; second the hourly GHI data between sunrise and sunset were quality checked after the methods of Espinar et al. (2011) shown in Equation (1):

$$0.03\ GHI_{TOA} < GHI < \min(1.2\ I_0, 1.5\ I_0\ \cos(\theta_z)^{1.2} + 100), \quad (1)$$

where GHI_{TOA} (W/m²) is the top of atmospheric irradiance, θ_z is the sun zenith angle (°), and I_0 is the solar constant (1367 W/m²). The top of atmospheric irradiance is defined as Equation (2):

$$GHI_{TOA} = \ I_0\ \varepsilon\ \cos(\theta_z), \quad (2)$$

where ε is the solar correction (Mather and Koch, 2011).

French Guiana and Reunion Island stations were equipped with type CM6B Kipp and Zonen pyranometers, whereas Singapore stations were equipped with Delta-T SPN1 pyranometers. French weather services in French Guiana and Reunion Island carry out preventive maintenance every two months, and pyranometers are replaced every two years. Singapore pyranometers are operated by the Solar Energy Research Institute of Singapore (SERIS) and are calibrated at the National Metrology Center of Singapore every two years. The GHI values resulting from the quality check procedure (Eq. 1) were then compared to the coincident IFS and GFS time stamps for a 48h forecast horizon. Figure 1 shows the monthly means GHI value for each station with hourly and 3 hourly frequency.

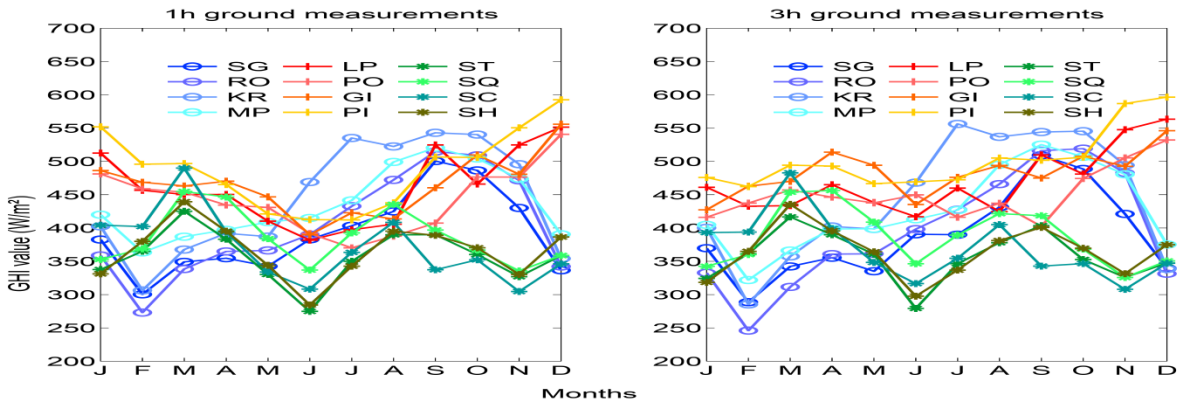

Figure 1: Monthly mean of global horizontal irradiance (GHI) values at 1 and 3-hour intervals. Values were computed using hourly quality checked GHI values. The abbreviations of the station name and location are shown Table 1.

Table 1: Validation station description.

Stations	Latitude (°)	Longitude (°)	Height (m)	Climate
French Guiana				
Saint Georges (SG)	3.890	- 51.804	6	Trop. monsoon
Rochambeau (RO)	4.822	-52.365	4	Trop. rainforest
Kourou (KR)	5.209	- 52.748	12	Trop. monsoon
Maripasoula (MP)	3.640	- 54.028	106	Trop. rainforest
Reunion Island				
Ligne-Paradis (LP)	-21.318	55.485	156	Trop. savanna
Le Port (PO)	-20.946	55.281	9	Trop. savanna
Gillot-Aéroport (GI)	-20.891	55.528	8	Trop. monsoon
Pierrefonds-Aéroport (PI)	-21.320	55.425	21	Trop. savanna

Singapore				
Seris 403 (ST)	1.443	103.784	45	Trop. rainforest
Seris 404 (SQ)	1.249	103.841	36	Trop. rainforest
Seris 405 (SC)	1.352	103.965	45	Trop. rainforest
Seris 408 (SH)	1.355	103.692	57	Trop. rainforest

2.3 Climate

In tropical regions, seasons are based on the precipitation amount (Hess and McKnight, 2013). Under the Köppen climate classification, a dry season month has an average monthly precipitation amount of less than 60 mm, whereas a rainy season month has precipitation amounts greater than 60 mm. Based on the Köppen climate classification system, tropical climates are divided into three categories: tropical rainforest, tropical monsoon, and tropical savanna (Hess and McKnight, 2013). Tropical rainforest climates have annual precipitation amounts between 150 and 250 cm, and precipitation amounts greater than 60 mm every month. Tropical monsoon and savanna climates have at least one month with precipitation amounts of less than 60 mm. Tropical monsoon climates have annual precipitation amounts between 250 and 500 cm, whereas the annual precipitation of tropical savanna climates is between 90 and 180 cm (Hess and McKnight, 2013). Table 1 shows the climates for stations of validation.

In the following subsections, we describe the climate of each territory, and the microclimates at each of the measurement sites.

2.3.1 French Guiana

The climate of French Guiana is regulated by the Intertropical-Convergence Zone (ITCZ), which is the meeting point of the northeast and southeast trade winds (Bovolo et al., 2012). The ITCZ moves over French Guiana twice a year, this shift from north to south creates a seasonal cycle with four periods (Philippe Héritier, 2011): (1) From July to November, the ITCZ lies north of French Guiana, and this period represents the dry season during which time the sky is mostly clear, although weak precipitation may occur (Albarelo et al., 2015). During this period cirrus and cumulus fractus clouds are predominant. (2) From November to January, there is a short rainy season when the ITCZ moves southward over French Guiana. (3) From February to March, there is a transition period between the southward and northward motions of the ITCZ, and depending on the year, this month behaves as a rainy or dry. (4) From April to May, the ITCZ moves northward, and a second rainy season begins. During this period, cumulonimbus clouds predominate. Rochambeau Station is located 13 km of the Atlantic Ocean. Maripasoula Station is located inland, ~230 km from the coast. Kourou Station is located, ~2 km from the Atlantic Coast. Saint Georges Station is located and inland, ~50 km from the Atlantic Ocean (Philippe Héritier, 2011).

2.3.2 Reunion Island

Reunion Island is a volcanic island with a complex orography driving several microclimates, but its overall climate is classified as tropical. There are two main seasons, the rainy season (January to March) and dry season (May to November). April and December are transitional months with rainy or dry periods. From November to April, clouds are observed at low and high altitude, below 3.5 km, and above 6.5 km, respectively. High altitude cirrus clouds represent ~13% of the cloud coverage (Bertrand Cadet et al., 2003). From May to October, there is a high occurrence of low altitude clouds. This cloud cover is three times greater than that observed from November to April, and clouds are denser, and observed at altitudes below 3.5 km (Badosa et al., 2013, 2015). Ligne-Paradis Station is located on the south of the island in the volcanic foothills, ~6 km from the Indian Ocean. Le Port Station is located in the northwest,

~100 m from the Indian Ocean. Gillot Aéroport is located in the north, ~300 m from the ocean. Pierrefonds-Aéroport is located in the southwest, ~500 m from the ocean, and has a tropical savanna climate.

2.3.3 Singapore

Singapore has a tropical rainfall climate characterized by high temperatures and humidity throughout the year. There is no distinct wet or dry season in Singapore; the main seasonal events are the Northeast (NE) and the Southwest (SW) monsoons, separated by two inter-monsoon periods. The NE monsoon lasts from December to early March and is characterized by prevailing north and northeastern winds and is divided into two phases. The wet phase lasts from December to January and sees rapid development of monsoon surges in the afternoon and early evening. The dry phase, from January to early March, is relatively drier. The SW monsoon extends from June to September with prevailing winds from the south and southwest. This period witnesses occasional Sumatra squall lines (SSL) that bring organized lines of thunderstorms to Singapore.

During the two inter-monsoon periods, there is little prevailing wind, and thunderstorms typically develop in the afternoon and the evening. The first inter-monsoon period (late March to May) is usually hotter and relatively drier than the second one (October to November). The four meteorological stations operated by SERIS are located in all cardinal directions of the island. Depending on their locations, each station is located less than 6 km from Singapore or the Johor Strait. Due to the small size of Singapore, the stations are relatively close and all experience the same tropical rainfall climate. However, there are small variations in the hourly GHI received by each station over the year, as illustrated in Figure 1.

3 Results and discussion

In this section we consider the GHI forecasts at the single grid points closest to each station; we do not average spatially the GHI forecasts of IFS and GFS. We present first the metrics used to assess the accuracy of both IFS and GFS; second we study their yearly accuracies and third their monthly accuracies. Lastly, we compare the accuracy of IFS and GFS in the different tropical climate to their respective accuracy in extra-tropical climates.

3.1 Error metrics

IFS and GFS GHI forecasts are compared with *in situ* measurements. Their performance is evaluated in terms of mean bias error (MBE), mean average error (MAE), root mean squared error (RMSE) and forecast skill (FS), defined as by Equations 3- 7, respectively:

$$MBE_{\%} = \frac{\sum_t GHI_{NWP}(t) - GHI_{ground}(t)}{N \sum_t GHI_{ground}(t)}; \quad (3)$$

$$MAE_{\%} = \frac{\sum_i |GHI_{NWP}(t) - GHI_{ground}(t)|}{N \sum_t GHI_{ground}(t)}; \quad (4)$$

$$RMSE_{\%} = \frac{\sqrt{\frac{1}{n}\sum_i (GHI_{NWP}(t) - GHI_{ground}(t))^2}}{N \sum_t GHI_{ground}}, \quad (5)$$

where GHI_{NWP} refers to the NWP model GHI estimate, GHI_{ground} to in-situ observed measurements and *N* is the total number of data points. RMSE is arguably the most important metric in renewable energy forecasting (Perez et al., 2013) because it puts more weight on large forecast errors that have a higher impact on electricity grid management. MAE provides a measurement of global error, and it is suitable for applications where forecast errors are proportional to the induced error cost (Perez et al., 2013). The MBE shows a model tendency to underestimate or to overestimate the GHI. $MBE_{\%}$, $MAE_{\%}$ and $RMSE_{\%}$ are unitless, they are expressed as percentages (%). Absolute MBE, MAE and

RMSE in W/m² are also used; they are obtained by multiplying their relative value by the mean GHI_{ground} value over the considered time span.

Persistence forecasts (Eq. 6) were implemented using the 24-hour previous irradiance values. Persistence forecasts serve as references to compute the forecast skills (FS) of the models (Perez et al., 2013):

$$GHI(t + 24h) = GHI(t) \tag{6}$$

$$FS = \frac{(RMSE_{persistence})^2 - (RMSE_{NWP})^2}{(RMSE_{persistence})^2}. \tag{7}$$

FS allows comparison between different forecast results from different locations, it is unitless. FS uses the absolute RMSE (in W/m²) computed from the GHI estimate of a NWP model. FS = 1 corresponds to a perfect model, while negative scores indicate performance worse than persistence (Perez et al., 2013).

The clear sky index K_c is used as a proxy for sky conditions, it is unitless. $K_c > 0.65$ indicates a clear sky, $0.4 < \mathrm{K_c} < 0.65$ a cloudy sky, and $K_c < 0.4$ overcast conditions (Aryaputera et al., 2015). K_c is defined as follows (Eq. 7):

$$K_c = \frac{GHI}{G_c}, \tag{8}$$

where G_c is the ESRA clear sky model as defined by Rigollier et al. (2000). Link turbidity, the clear sky attenuation factor input to G_c, is defined for each site and each month using monthly climatological values provided by the SoDa database ([dataset] Transvalor, 2014).

As rapid cloud transformations occur in tropical climates (Aryaputera et al., 2015), the mean value of $\mathrm{K_c}$, M_{K_c} (Eq. 9), and the mean absolute variability of $\mathrm{K_c}$, V (Eq. 10) (Marquez and Coimbra, 2012) are introduced to study the ability of IFS and GFS to reproduce the GHI variability in tropical climates; both $\mathrm{MK_c}$ and V are unitless:

$$M_{K_c} = \frac{1}{N} \sum_t^{t+N} K_c(t), \tag{9}$$

$$V = \sqrt{\frac{1}{N}\sum_t^{t+N}(K_c(t + dt) - K_c(t))^2}. \tag{10}$$

Where *dt* is the time step of ground measurements, i.e. hourly. For a model with hourly GHI outputs N=1 whereas for a model with a three hour GHI outputs N=2. To illustrate how $\mathrm{MK_c}$ and V were computed, we give the following example for a one hour and a three-hour output model. At 12 h UTC, and for a one hour output model, $\mathrm{MK_c}$ and V are computed using the $\mathrm{K_c}$ value at 11h UTC and 12 h UTC (N=1). At 12 h UTC, and for a three hour output model, $\mathrm{MK_c}$ and V are computed using the $\mathrm{K_c}$ value at 10h, 11h UTC and 12 h UTC (N=2).

3.2 Assessing the accuracy of hourly IFS forecasts

In this section we study the accuracy of hourly IFS forecasts; we name these forecasts IFS-1h. Table 2 gives the MAE and MBE for each territory under clear, cloudy, overcast and all sky conditions; figure 2 gives IFS-1h MAE distribution with respect to the mean clearness index (M_{K_c})and mean clearness index variability (*V)*.

Table 2 shows that the MAE of IFS-1h under clear sky condition is lower than the MAE of IFS-1h under cloudy and overcast sky conditions. Under clear sky conditions the MAE values range between

100 W/m² and 122 W/m² whereas under overcast and cloudy sky conditions the MAE values range between 127 W/m² and 225 W/m². Consequently IFS-1h forecasts are more accurate in clear sky conditions than in cloudy and overcast sky conditions. Nevertheless, Figure 2 shows that the accuracy of IFS forecasts does not decrease systematically as the sky conditions become clearer; IFS-1h is sensitive to the mean clear sky index and mean clear sky index variability. Table 2 also shows that under clear sky conditions, the MBE are negative for each territory. Oppositely, under cloudy and overcast sky conditions the MBE are positive. Consequently IFS-1h underestimates the GHI value under clear sky conditions whereas under cloudy and overcast sky conditions IFS-1h overestimate the GHI.

Table 2: IFS-1hour absolute MAE and MBE by territory under clear, cloudy, and overcast sky conditions. We computed M_{K_c} (Eq. 9) using N=1, and dt=1hour. Months and stations were merged by territory. The classification of sky conditions was based on ground measurements.

Territory	MAE of IFS 1-hour (W/m²)				MBE of IFS 1-hour (W/m²)			
	Sky conditions							
	Clear	Cloudy	Overcast	All sky	Clear	Cloudy	Overcast	All sky
French Guiana	104	127	173	119	-52	68	160	9
Reunion Island	100	137	216	117	-58	104	209	-4
Singapore	122	155	225	153	-60	107	215	41

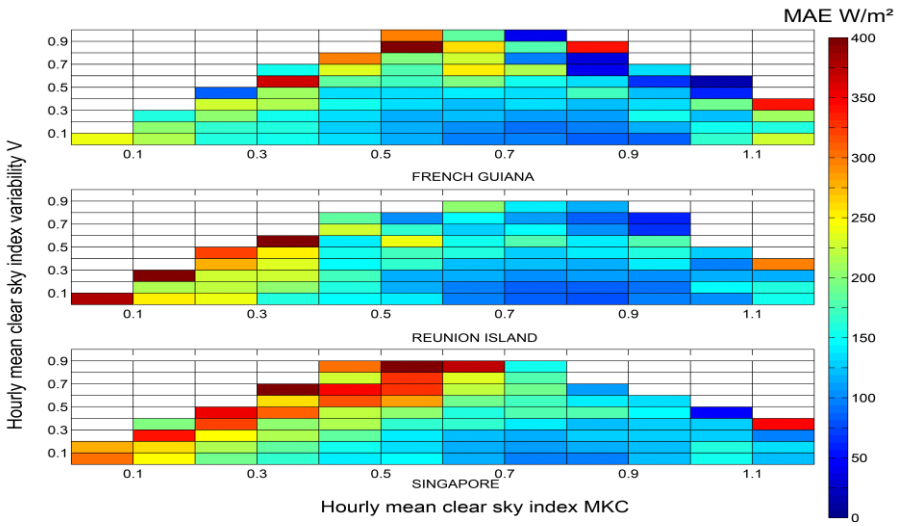

Figure 2: IFS-1h MAE distribution with respect to M_{K_c} (Eq. 9) and V (Eq. 10). We computed M_{K_c} (Eq. 9) and V (Eq. 10) using N=1, and dt=1hour. Months and stations were merged by territory. The classification of sky conditions was based on ground measurements

3.3 Comparing IFS and GFS yearly irradiance forecast accuracy in French Guiana, Reunion Island and Singapore

To compare IFS-1h outputs to GFS 3 hour outputs (GFS-3h), there are three alternatives. Firstly, the GHI forecasts of GFS can be interpolated to hourly outputs using K_c (Mathiesen and Kleissl, 2011). Verzijlbergh et al. (2015) found in his study by comparing hourly K_c values computed from ground measurement to hourly interpolated K_c values obtained using the perfect 3h average forecasts (also computed from ground measurements) that the RMSE was as high as 21%. Consequently, to compare IFS to GFS we did not consider this alternative. Secondly, IFS-1h GHI forecasts can be averaged temporally to obtain three hours forecasts outputs following GFS cycle; we named these forecasts IFS-3h-p. The GHI values of IFS-3h-p (W/m²) at 03:00, 09:00, 15:00, and 18:00 UTC are the means of 3-hour irradiance values ending at forecast time whereas the GHI values of IFS-3h-p (W/m²) at 06:00, 12:00, 18:00, and 24:00 UTC are the means of 6-hour irradiance values ending at forecast time. Lastly, from 00h UTC to 24h UTC we keep only every 3-hour forecasts of hourly IFS; we named these forecasts IFS-3h.

Figure 3 gives the MAE and MBE of IFS-3h-p, GFS-3h and IFS-3h for each territory under clear, cloudy, overcast and all sky conditions. Table 3 gives the Relative improvement of the MAE of IFS-3h over the MAE of IFS-3h-p and GFS-3h for each territory under clear, cloudy, overcast and all sky conditions. Figure 4 gives IFS-3h-p, GFS-3h and IFS-3h MAE distribution with respect to M_{K_c} and *V*.

Figure 3 shows that for all territories, the MBE of IFS-3h and IFS-3h-p are negative under clear sky conditions whereas under cloudy and overcast sky conditions the MBE of both are positive. Consequently, IFS-3h and IFS-3h-p under estimate the GHI under clear sky conditions whereas under cloudy and overcast sky conditions both overestimate consistently the GHI. Oppositely, the MBE of GFS-3h is consistently positive in Singapore and French Guiana. In Reunion Island, the MBE is positive under overcast sky conditions and negative under clear and cloudy sky conditions. Consequently, GFS-3h overestimates the GHI consistently in Singapore and French Guiana; in Reunion Island GFS-3h overestimates the GHI under overcast sky conditions whereas it underestimates the GHI under clear and cloudy sky conditions.

Figure 4 shows that the MAE of IFS-3h-p, GFS-3h and IFS-3h are highly correlated to M_{K_c} and *V*. This figure shows that except for GFS in Reunion Island, each model has a tendency to perform better under clear sky conditions. To conclude on the accuracy of each model with respect to the sky condition, we gather each GHI sample according to sky condition type that is reported by the ground measurements (Figure 3).

The MAE values (Fig. 3) show that for each territory IFS-3h and IFS-3h-p are more accurate in clear sky conditions and least accurate in cloudy and overcast sky conditions (Fig.3). Oppositely, we find that in Reunion Island, GFS-3h MAE of clear sky condition forecasts is higher than the MAE of cloudy and overcast sky conditions forecasts. Under clear sky conditions, the MAE of GFS-3h is 319 W/m² whereas under cloudy and overcast sky conditions it is 249 W/m² and 242 W/m², respectively. Under all sky conditions and for each territory, the MAE of IFS-3h-p is lower than the MAE of GFS-3h (Fig. 3). Figure 3 also shows that for each territory and under all sky conditions the MAE of IFS-3h is lower than the MAE of IFS-3h-p. Table 3 shows that under all sky conditions and for each territory, the relative improvement of the MAE of IFS-3h over IFS-3h-p ranges between 21% and 46%. This result contrasts with Lorenz et al., (2016) who found that averaging temporally improve the accuracy of NWP GHI outputs. Therefore, we conclude that (1) averaging temporally IFS-1h following GFS cycle to obtain IFS-3h-p worsen the MAE in tropical climate and (2) that both IFS-3h and IFS-3h-p are more accurate than GFS.

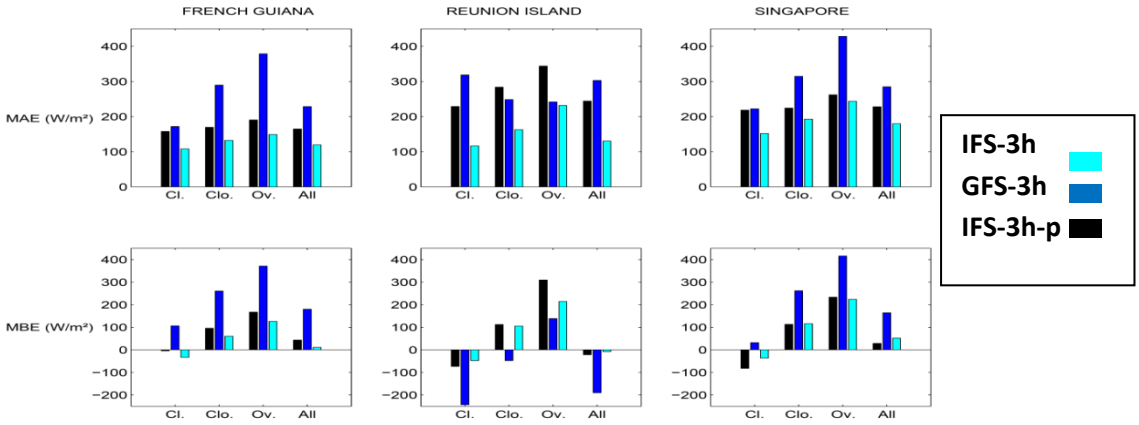

Figure 3: MAE and MBE of GHI forecast from IFS-3h, IFS-3hp and GFS-3h in French Guiana, Reunion Island and Singapore. Months and stations were merged by territory. The classification of sky conditions was based on ground measurements.

Table 3: Relative improvement of the MAE of IFS-3h over the MAE of IFS-3h-p and GFS-3h under clear, cloudy, and overcast sky conditions. The classification of sky the conditions for each 3-hour sample is based on M_{K_c} (Eq. 9) which is computed using ground measurements with N=2, and dt=1hour. Months and stations were merged by territory.

Territory	Relative improvement (%) of the MAE of IFS-3h (W/m²) over the MAE of IFS-3h-p (W/m²) and GFS-3h (W/m²)							
	Clear		Cloudy		Overcast		All	
	Vs IFS-3h-p	Vs GFS-3h	Vs IFS-3h-p	Vs GFS-3h	Vs IFS-3h-p	Vs GFS-3h	Vs IFS-3h-p	Vs GFS-3h
French Guiana	32%	37%	22%	54%	20%	61%	27%	48%
Reunion Island	49%	63%	43%	34%	33%	4%	46%	57%
Singapore	31%	32%	14%	39%	7%	43%	21%	37%

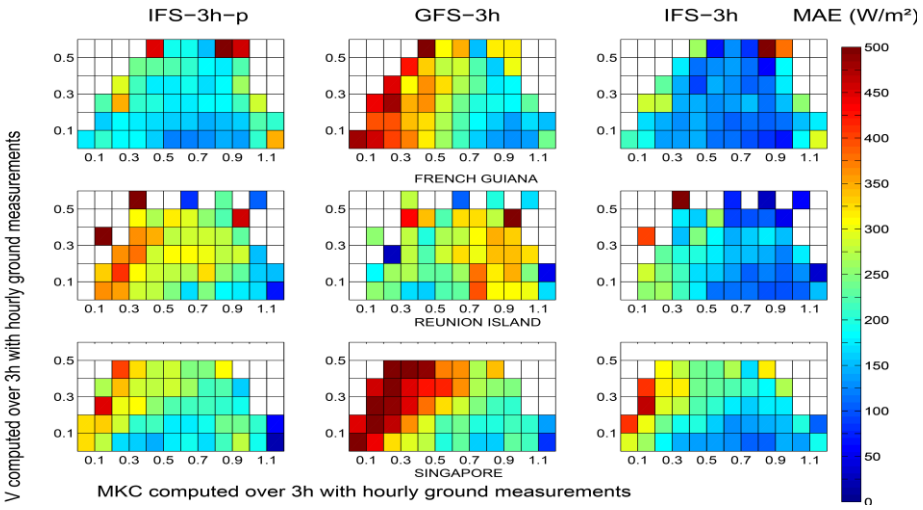

Figure 4: IFS and GFS bias distributions with respect to the mean K_c values (Mkc) of the previous three hours, and the mean of the K_c Variability (V). We computed M_{K_c} (Eq. 9) and V (Eq. 10) using N=2, and dt=1hour. Months and stations were merged by territory. The classification of sky conditions was based on ground measurements

3.4 Analyzing the sources of error of IFS and GFS GHI forecasts

We have shown in section 3.3 that the accuracy of IFS-3h-p, IFS-3h and GFS-3h forecasts was related to the $K_c$value. Consequently, to explain the sources of IFS and GFS GHI forecasts errors in tropical climates, we analyze and compare the $K_c$values computed using ground measurement and the $K_c$value computed using IFS-3h-p, IFS-3h and GFS-3h GHI forecasts. Figure 5 shows a comparison between the sky conditions measured at the ground i.e clear, cloudy or overcast to the sky conditions forecasted by the NWP model (i.e IFS-3h-p, IFS-3h or GFS-3h). This comparison uses the cdf of K_c values computed using ground measurements and the cdf K_c values forecasted by the NWP GHI forecasts.

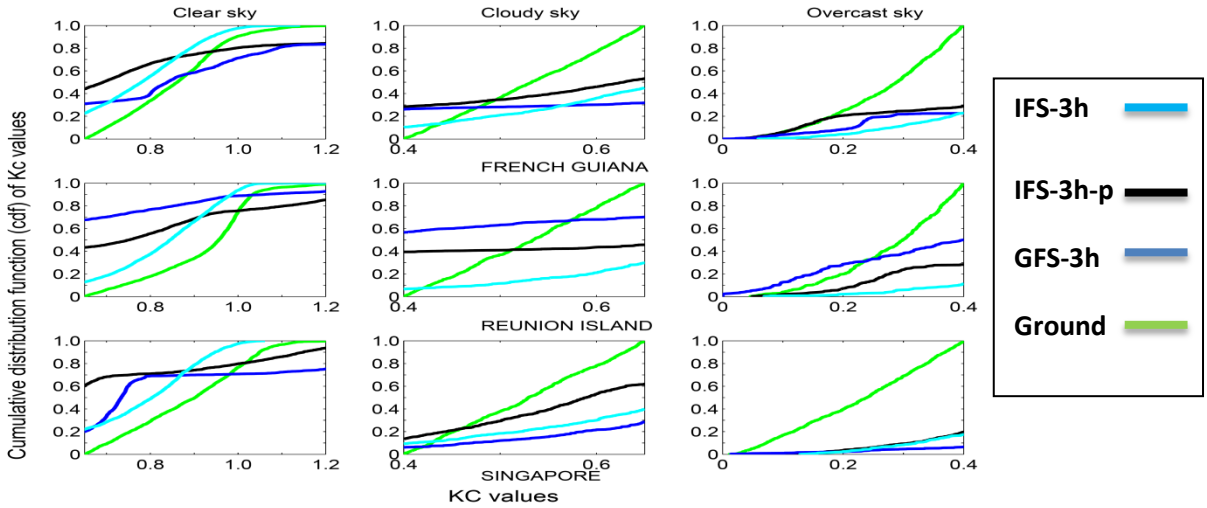

Figure 5: cdf of K_c computed using ground measurements and IFS-3h-p, GFS-3h and IFS-3h forecasts under clear, cloudy and overcast sky conditions. Months and stations were merged. The classification of sky conditions was based on ground measurements

3.4.1 Analyzing the sources of error under clear sky condition

Figure 5 shows that under measured clear sky conditions, for IFS-3h, there are 77%, 87% and 78% of the forecasted K_C values that are higher than 0.65 in French Guiana, Reunion Island and Singapore, respectively. Consequently, when the sky condition is reported as clear by the ground measurements, it is reported as clear by IFS-3h between 77% and 87% of the time.

For GFS-3h, under measured clear sky conditions, Figure 5 shows there are 69%, 32% and 80% of the forecasted K_C values that are higher than 0.65 in French Guiana, Reunion Island and Singapore, respectively. Consequently, when the sky condition is reported as clear by the ground measurements, it is reported by GFS-3h as clear between 32% and 80%.

For IFS-3h-p, under measured clear sky conditions, Figure 5 shows there are 56%, 57% and 40% of the forecasted K_C values that are higher than 0.65 in French Guiana, Reunion Island and Singapore, respectively. Consequently, when the sky condition is reported as clear by the ground measurements, it is reported by IFS-3h-p as clear between 40% and 57% of the time.

Therefore, IFS-3h-p, GFS-3h and IFS-3h are unable to forecast realistic cloud cover and or amount for sky conditions reported as clear by the ground measurements. Consequently, this inability contributes to the source of errors under clear sky. To quantify the source of errors due to sky conditions that are reported as clear by the ground measurements, but forecasted otherwise by IFS-3h-p, GFS-3h and IFS-3h; firstly, we computed first the MAE when IFS-3h-p forecast and the ground measurements assess simultaneously that the sky conditions are clear. Secondly, we computed the MAE when GFS-3h forecasts and the ground measurements assess simultaneously that the sky conditions are clear. Thirdly, we computed the MAE when IFS-3h forecast and the ground measurements assess simultaneously that the sky conditions are clear.

Firstly, we found that IFS-3h-p has a MAE of 123 W/m², 133 W/m² and 134 W/m² in French Guiana, Reunion Island and Singapore respectively. These values represent 78%, 58% and 61%, respectively of clear sky MAE values that were shown figure 3. Consequently, the source of errors due to sky conditions that are reported as clear by the ground measurements, but forecasted otherwise by IFS-3h-p represent between 22% and 39%.

Secondly, we found that GFS-3h has a MAE of 139 W/m², 131 W/m² and 172 W/m² in French Guiana, Reunion Island and Singapore respectively. These values represent 81%, 41% and 77%, respectively of clear sky MAE values that were shown figure 3. Consequently, the source of errors due to sky conditions that are reported as clear by the ground measurements, but forecasted otherwise by GFS-3h represent between 19% and 59%,

Thirdly, we found that IFS-3h has a MAE of 69 W/m², 80 W/m² and 96 W/m² in French Guiana, Reunion Island and Singapore respectively. These values represent 64%, 68% and 63%, respectively of clear sky MAE values that were shown figure 3. Consequently, the source of errors due to sky conditions that are reported as clear by the ground measurements, but forecasted otherwise by IFS-3h represent between 32% and 37%,

Therefore, the sky conditions that are reported as clear by the ground measurements, but forecasted otherwise by IFS-3h-p, GFS-3h and IFS-3h lead to sizable MAE values; for GFS-3h in Reunion Island they are as high as 59% of the overall MAE.

According to Zhong et al. (2016) : “often, but not always clear sky model biases are related to aerosol modeling issues”; for Jimenez et al. (2015) aerosols represent the largest source of uncertainty in the GHI forecasts under clear sky conditions. Consequently, we could explained the MAE when IFS-3h-p or GFS-3h or IFS-h forecasts and the ground measurements assess simultaneously that the sky conditions are clear by a misrepresentation of the aerosols.

To conclude IFS-3h-p, IFS-3h, and GFS-3h sources of error under clear sky conditions may include: (1) sky conditions reported as clear by the ground measurements that are inaccurately forecasted as cloudy or overcast and (2) an inaccurate representation of the aerosols.

3.4.2 Analyzing the sources of error under cloudy and overcast sky conditions

Figure 5 shows that for all territories considered, when the sky condition is reported as cloudy or overcast by the ground measurements, it is inaccurately forecasted by IFS-3h-p, GFS-3h and IFS-3h.

In French Guiana, under measured cloudy sky conditions and for IFS-3h-p, figure 5 shows that 26% of forecasted K_C values are lower than 0.4 whereas 32% of the forecasted K_C are higher than 0.65. Therefore, the match between sky conditions type that is measured at the ground and sky conditions type that is forecasted by IFS-3h-p is 6%. In French Guiana, under measured overcast sky conditions, figure 5 shows that 23% of forecasted K_C values are higher than 0.4. Therefore, the match between sky conditions type that is measured at the ground and sky conditions type that is forecasted by IFS-3h-p is 23%.

Using the same procedure described previously, for each territory and each NWP model: we find that under cloudy and overcast sky condition, the match between sky conditions measured at the ground and sky conditions forecasted by any NWP model (eg. IFS-3h-p, GFS-3h and IFS-3h) is lower than 50%. Consequently, these results may be explained by the inability of the NWP to either forecast accurately the cloud position and or extent. Similar conclusions were drawn by Mathiesen and al. (2013), Zempila et al. (2015) and Kleissl (2013); they explained high biases under cloudy and overcast skies were caused by the inability of NWP models to accurately predict cloud position and cloud extent. Lin et al. (2009) demonstrated that increasing the horizontal resolution of NWP models improves the accuracy of the simulated low cloud field; they concluded that clouds can be realistically simulated with horizontal resolutions higher than 4 km, but not with coarser resolutions. Consequently, we could explained the forecast accuracy of IFS-3h, IFS-3h-p and GFS-3h under cloudy and overcast sky conditions by their coarse resolution; this coarse resolution causes a misrepresentation of the sky conditions which lead to high MAE value.

3.5 IFS and GFS seasonal forecast accuracy assessment

In this section, seasonal forecast accuracies of both 3-hour IFS and GFS are assessed using the metrics described in section 3.1. To obtain the percent RMSE and MAE relative values, absolute value in W/m² were normalized using the monthly mean values shown in Figure 1, and following Equations 4 and 5.

3.5.1 French Guiana Sites

Figure 6 gives on the one hand IFS and GFS monthly RMSE and MAE, figure 7 gives on the other the monthly mean $K_c$value.

Figure 6 shows that both GFS and IFS GHI forecasts follow seasonal trends for French Guiana (i.e., there is a clear distinction between dry and rainy seasons). RMSE and MAE during the dry season

(i.e., July to November) are on average lower than those found in rainy season (i.e., from November to January and from April to May). The different forecast accuracies found may be explained by different clearness index value that varies according to the seasons (Fig. 7).

During the dry season, Kourou station exhibits the lowest average RMSE and MAE. Figure 7 shows that during this season, Kourou station exhibits the highest monthly mean K_c values. During the rainy seasons, Maripasoula exhibits lowest average RMSE and MAE. Figure 7 shows that during these seasons, Maripasoula have on average the highest monthly mean $K_c$values, respectively. These results suggest that there is a correlation between the monthly mean forecast accuracy and the monthly mean $K_c$value. Consequently IFS-3h and GFS-3h GHI forecasts are less accurate in cloudier sky conditions; the accuracy of IFS-3h and GFS-3h GHI forecasts are also influenced by the variability of the sky conditions.

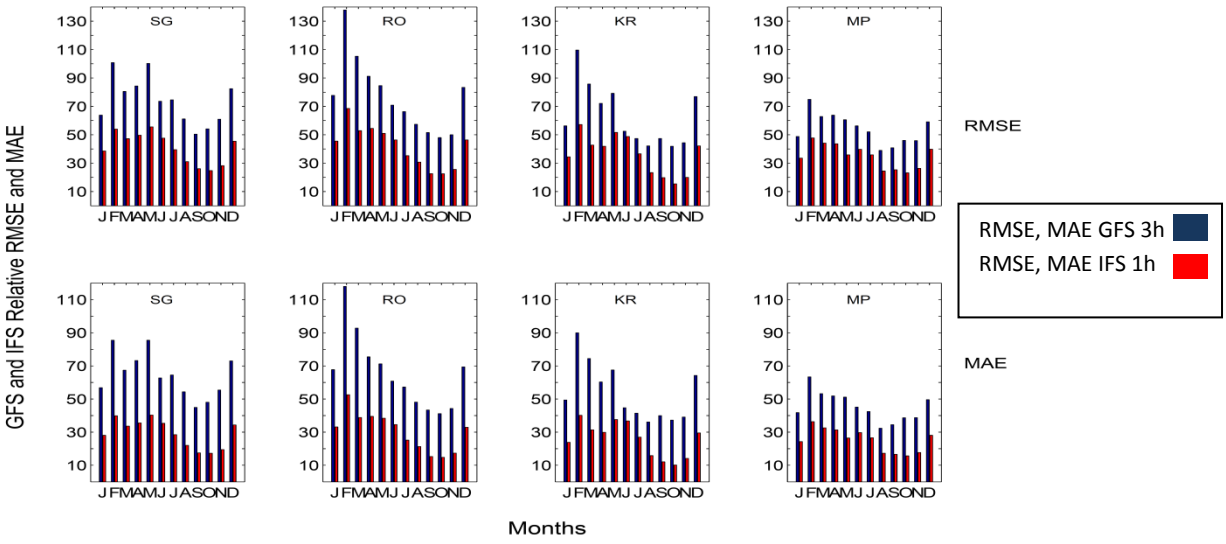

Figure 6: GFS-3h and IFS-1h monthly relative MAE and RMSE (%) for French Guiana stations. The monthly mean value used to normalized the MAE and RMSE are shown Figure 1

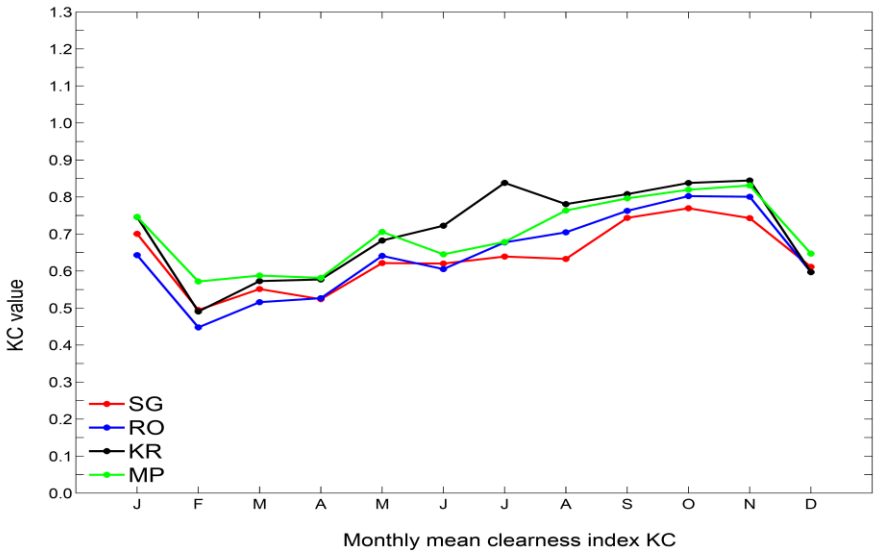

Figure 7: Monthly K_c value for all French Guiana stations. The Monthly K_c value were computed using hourly quality checked GHI measurement between sunrise and sunset

3.5.2 Reunion Island sites

Figure 8 gives on the one hand IFS and GFS monthly RMSE and MAE, figure 9 gives on the other the monthly mean $K_c$value.

When comparing figure 9 to Figure 7, it can be seen that, unlike in French Guiana, there is no distinct wet or dry season on Reunion Island. Figure 9 shows that the difference between monthly mean Kc values during the rainy and dry season is less visible. According to the French Weather Service, 2016 ranks 13th amongst the driest years since 1960. For Station PI, during the dry season, the lowest GFS RMSE and MAE (60% and 49%, respectively) are found in August; the lowest IFS RMSE and MAE (24% and 14%, respectively), they are found in June. For this station, the highest monthly mean K_c is found in June. Consequently, the monthly RMSE and MAE of the GHI forecasts of GFS are not as strongly correlated with monthly mean K_c as they are in French Guiana.

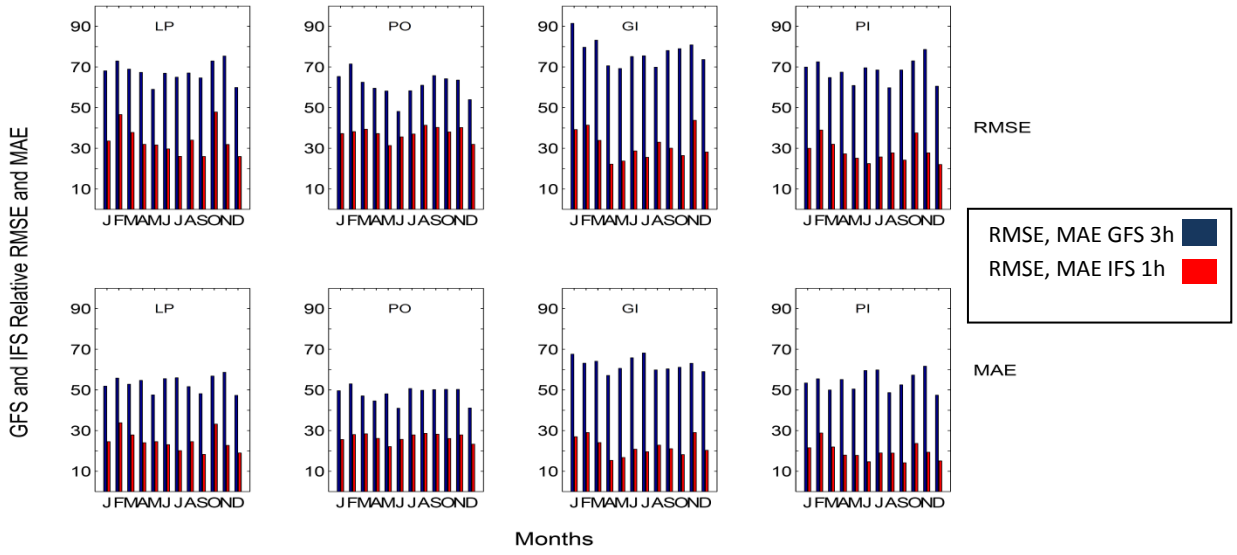

Figure 8: GFS-3h and IFS-1h monthly relative MAE and RMSE (%) for Reunion island stations. The monthly mean value used to normalized the MAE and RMSE are shown Figure 1

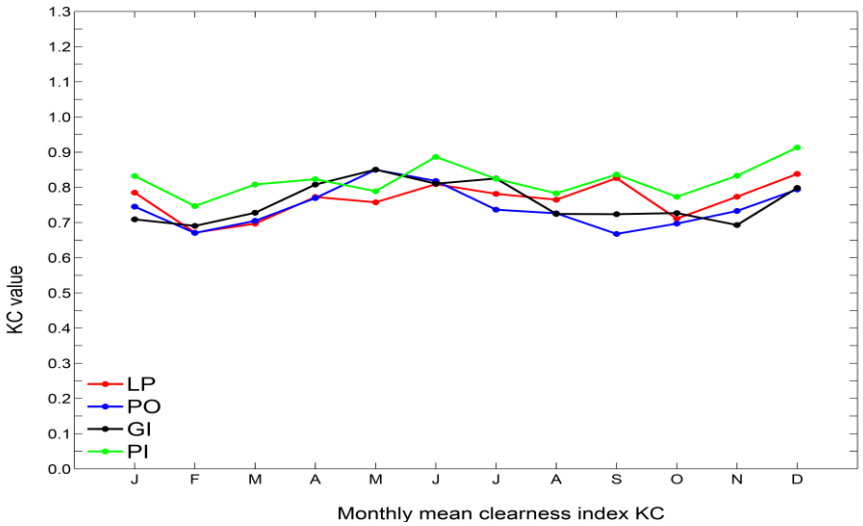

Figure 9: Monthly K_c value for all Reunion Island stations. The Monthly K_c value were computed using hourly quality checked GHI measurement between sunrise and sunset

3.5.3 Singapore sites

Figure 10 gives on the one hand IFS and GFS monthly RMSE and MAE, figure 11 gives on the other the monthly mean $K_c$value.

. From January to August, station SC has the highest monthly mean K_c value, as seen in figure 11. During this time span, this station exhibits overall the lowest RMSE and MAE. Consequently, IFS-3h and GFS-3h GHI forecasts are correlated to the monthly mean K_c value and they are less accurate in cloudier sky conditions.

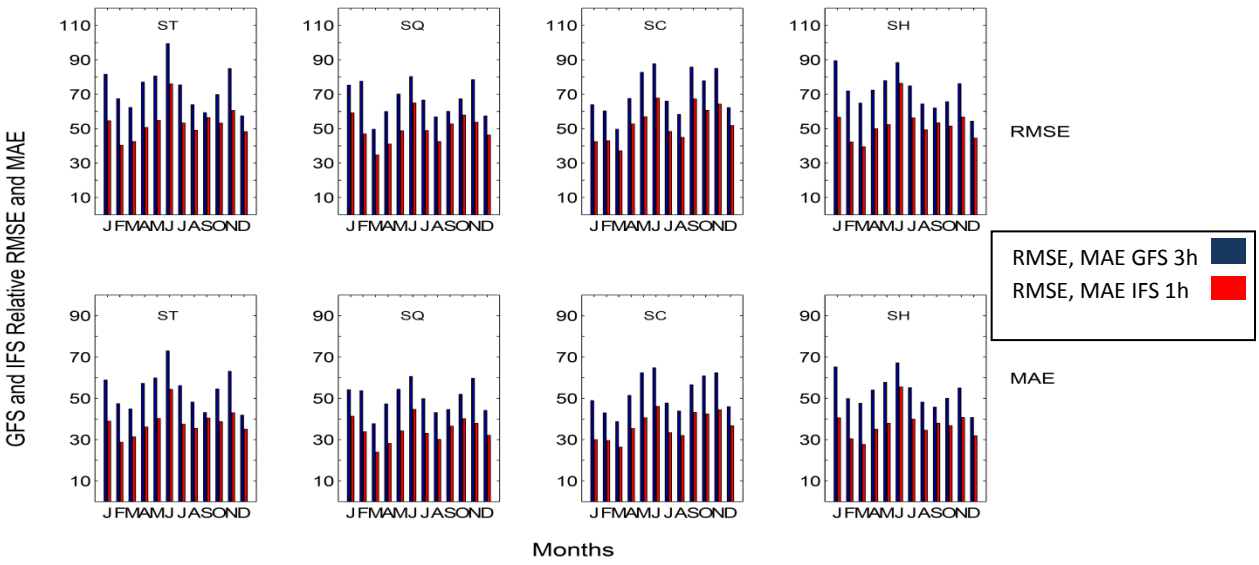

Figure 10: GFS-3h and IFS-1h monthly relative MAE and RMSE (%) for Singapore stations. The monthly mean value used to normalized the MAE and RMSE are shown Figure 1

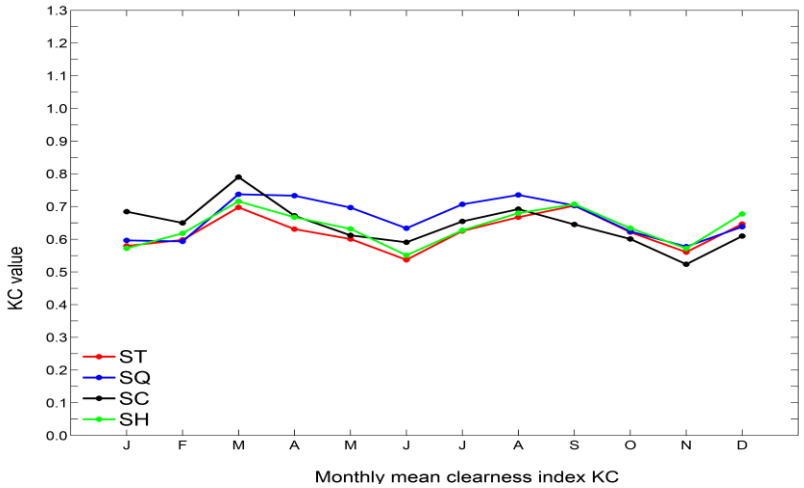

Figure 11: Monthly K_c value and associated standard deviations for Singapore stations. The Monthly K_c value were computed using hourly quality checked GHI measurement between sunrise and sunset

3.6 Comparing GFS and IFS forecast accuracy for tropical climates and comparing their accuracy in tropical climate to extra tropical climates

3.6.1 Comparing GFS and IFS forecast accuracy for the different tropical climates

*T*able 4 present the RMSE and FS for the different tropical climates. This table merged the yearly GHI values per tropical climate type; it allows comparing the forecast accuracy in tropical monsoon climate compared to tropical savanna and rainforest climates. Figure 12 presents the RMSE and FS per station. The yearly GHI values per stations were merged in order to analyze the variability amongst the same climate type. This approach allows us to obtain a range of RMSE value instead of a single value.

Table 4: RMSE of FS of IFS-1h GHI forecasts for tropical monsoon, savanna and rainforest climates. The yearly mean value used to normalized the RMSE are: 434 W/m² for stations with tropical monsoon climate; 468 W/m² for stations with tropical savanna climate, 392 W/m² for station with tropical rainforest climate

	RMSE	FS
Tropical monsoon	35%	0.39
Tropical savanna	33%	0.40
Tropical rainforest	46%	0.42

To compare the forecast accuracy for the different tropical climates we used IFS-1h forecasts because we showed that this model was able to reproduce the seasonal variation of GHI contrary to GFS-3h. Table 4 shows that overall the highest RMEs are found in tropical rainforest climate followed by tropical monsoon and tropical savanna climate. Nevertheless, there are discrepancies in the GHI forecast accuracies between stations sharing the same tropical climate type (Fig. 12). We have shown in section 3.3 that the forecast accuracy of IFS and GFS were highly correlated with the $\mathrm{MK_c}$ and V. Consequently, to explain the results shown table 4 we analyze the cdf of $\mathrm{MK_c}$ and V for each climate type (Fig. 13).

Figure 13 presents the cdf of $\mathrm{M_c}$ and V for stations with tropical monsoon, savanna and rainforest climate.We find, based on the analysis of the cdf $\mathrm{MK_c}$ shown figure 13, that stations with tropical rainforest stations have the highest occurrence of cloudy and overcast sky conditions sample followed by stations with tropical monsoon and savanna climate. For stations with tropical rainforest climate 50% of the sample has $\mathrm{MK_c}$ values lower than 0.70. For stations with tropical monsoon and tropical savanna climate 50% of the sample has $\mathrm{K_c}$ values lower than 0.75 and 0.85, respectively. Similarly the analysis of the cdf of V, shows that stations with tropical rainforest stations have the highest hourly mean variability followed by stations with tropical monsoon and savanna climate.

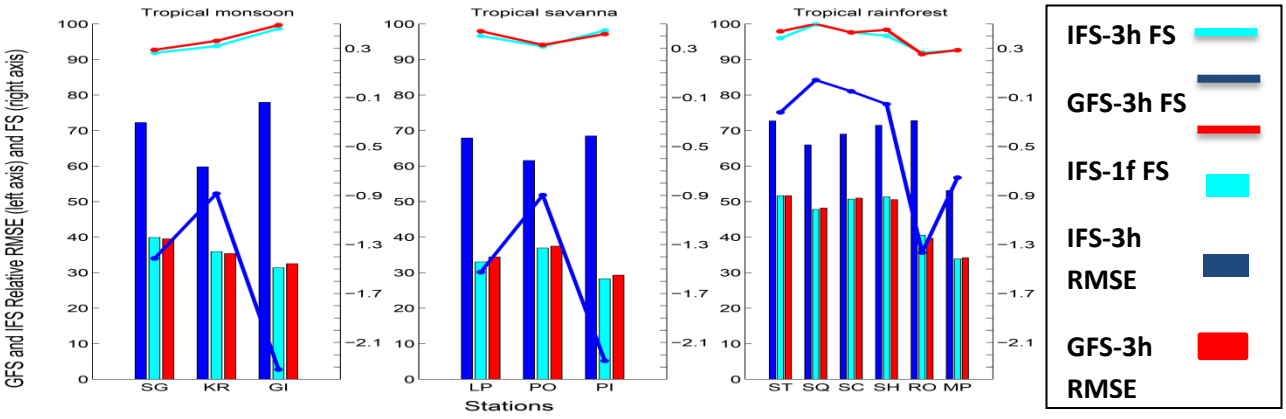

Figure 12: Yearly IFS and GFS FS and relative RMSE for all stations. The yearly GHI values used to normalize the RMSE were respectively: SG: 387 W/m², KR: 443W/m², GI: 482 W/m²; LP: 474 W/m², PO: 451 W/m², PI: 507 W/m²; ST: 359 W/m², SQ: 390 W/m², SC: 373 W/m², SH: 364 W/m², RO: 394 W/m², MP: 425 W/m²

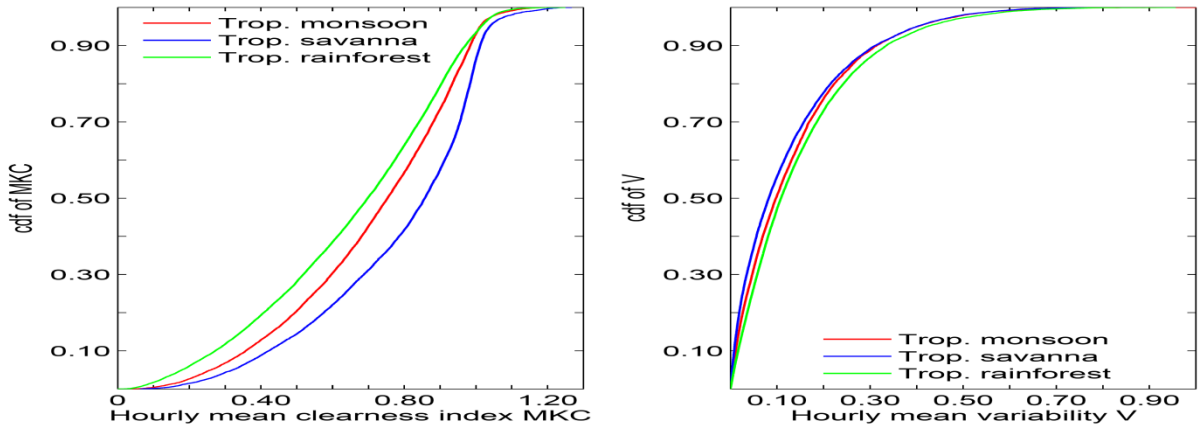

Figure 13: cdf of $\mathrm{MK_c}$ and V for stations with tropical monsoon, savanna and rainforest climate. We computed M_{K_c} (Eq. 9) and V (Eq. 10) using N=1, and dt=1hour

To compare the accuracy of the different tropical climates, the FS was also used. The FS allows the comparison of the forecast accuracy of a NWP model to that of a persistence model (Eq. 6) at different locations (Aryaputera et al., 2015). A negative FS value indicates that the NWP underperformed in comparison to the persistence model. A FS value close to 0 indicates that the NWP model forecast accuracy is similar to the minor persistence model forecast accuracy. A FS value close to 1 indicates that the NWP model outperforms the persistence model. Table 4 shows, by comparison with the persistence model, that it is more difficult for IFS-1h to forecast the GHI 48 hours in advance for tropical monsoon, savanna and rainforest in decreasing order.

3.6.2 *Comparing GFS and IFS forecast accuracy in tropical climates to their forecast accuracy in extra-tropical climates*

Table 5 gives the yearly FS and RMSE for different climate zones; we used this table to compare the GHI forecasts accuracy of IFS and GFS found in tropical climates to those found in extratropical climates.. The FS metric was not computed in all previous studies; when it was not available, it was computed if the absolute 24-hour persistence and 48-hour forecast RMSE were given by the authors (e.g., Perez et al., 2013; Schroedter-Homscheidt et al., 2017; Troccoli and Morcrette, 2014). FS and RMSE ranges were given for stations using the lowest and highest FS values, and the lowest and

highest RMSE values, respectively. GFS-based models NDFD and BLUE FORECAST are detailed in Perez et al. (2013). Every NWP was run for 48 hours, and was initialized at 00:00 UTC. The FS and RMSE values that were computed for tropical climates are shown Figure 12.

Table 5 shows that the 48h forecasts of GFS-3h have higher RMSE in tropical climates than in extra-tropical climate. The RMSEs of IFS-3h in tropical climate are in agreement with those found in extra-tropical climates. In Mediterranean, arid, semi-arid, continental, and humid continental stations, all GFS and IFS FS scores are positive. Among all tropical climate stations considered in Figure 12, GFS has a positive FS for only 1 station out of 12 (i.e., SQ, where GFS GHI forecast accuracy is similar to the accuracy of persistence models, and the FS score is 0.04). Conversely, IFS-3h FS is always positive; however, FS values found in tropical climates are lower than values found in other climates. Consequently, forecasting the GHI in tropical climate is more difficult for IFS and GFS. The lower GFS FS results might be explained by the coarser spatial resolution of GFS. This resolution may not be able to reproduce the higher variability of solar energy due to the dynamic cloud cover and hot and humid weather of tropical climates (Laing and Evans, 2011; Galvin, 2015).

Table 5: Yearly FS and RMSE for different climates. The native temporal resolution of IFS and GFS models used by Perez et al.(2013)are underlined.

Climate	Stations	GFS/ GFS based yearly RMSE, MAE, FS range	IFS yearly RMSE, MAE, FS range
Mediterranean	Huelva, Cordoba, Granada (Perez et al., 2013)	N.A RMSE: N.A FS: N.A	3-hour RMSE: [21:25]% FS: [0.45:0.63]
Arid	Desert Rock (Perez et al., 2013)	3-hour RMSE: ~27% FS: ~0.13	3-hour RMSE: ~22% FS:~0.42
Semi-arid	Boulder (Perez et al., 2013)	3-hour RMSE: ~45% FS:~0.22	3-hour RMSE: ~39% FS:0.41 ~
Continental	Bondville, Sioux Falls, Fort Peck, Furstenzell, Stuttgart, Wurzburg, Linz, and Wien (Perez et al., 2013)	3-hour RMSE: [39:49]% FS: [0.22:0.57]	3-hour RMSE: [34:52]% FS:[0.41:0.61]
Humid continental	Goodwin Creek, Penn State, Egbert, Bratt's, Varennes (Perez et al., 2013)	3-hour RMSE: [40:49]% FS: [0.32:0.43]	3-hour RMSE: [34:41]% FS: [0.54:0.60]
Tropical monsoon	Saint Georges (SG), Kourou (KR), Gillot-aéroport(GI)	3-hour RMSE: [59:78]% FS: [-2.32.:-1.41]	3-hour RMSE: [31:40]% FS: [0.26:0.46] hourly RMSE: [32:40]% FS: [0.28:0.49]
Tropical Savanna	Le port (PO), Pierrefonds aéroport(PI), Ligne-Paradis(LP)	3-hour RMSE: [61:69]% FS: [-2.24:-1.52]	3-hour RMSE: [28:37]% FS: [0.31:0.44] hourly RMSE: [29:38]% FS: [0.32:0.44]
Tropical rainforest	Rochambeau (RO), Maripasoula (MP), Station three (ST), Station four (SQ), Station Five (SC), Station eight (SH)	3-hour RMSE: [53:73]% FS: [-1.36:0.04]	3-hour RMSE: [33:52]% FS: [0.26:0.49] hourly RMSE: [34:52]%

Recently, Aryaputera et al. (2015) and Diagne et al. (2013) downscaled 0.5° GFS model forecasts using Weather and Research Forecast (Skamarock et al., 2005) limited-area NWP to forecast GHI on Reunion Island and in Singapore. Both studies forecasted the GHI hourly, on a 3 km resolution grid size. The yearly FS were computed from the RMSE values found by Aryaputera et al. (2015) and Diagne et al. (2013); these FS were -0.177, 0.07, and 0 for the tropical savanna station (Saint Pierre) and the tropical rainforest stations (station 305 and 500), respectively. These results are worse than those found in this study in tropical savannas (i.e., 0.32, 0.44), and tropical rainforest climates (i.e., 0.25, 0.50) using IFS-1h GHI forecasts, despite having a higher spatial resolution. Therefore, using IFS global NWP for GHI forecast is relevant in tropical climates. Using IFS and GFS global NWP forecasts, one might avoid calibrating a limited area NWP model, instead using physical parameterization that describes regional climate patterns. This process requires high computational resources, and does not guarantee better forecast results (Perez et al., 2013).

4 Conclusions

The goal of this study was to assess the accuracy of GFS and IFS global NWP model GHI forecasts for tropical climates. As the solar energy potential of tropical countries can be up to three times greater than that of extra-tropical countries, yet more variable due to the dynamic cloud cover (Fillol et al., 2017), predicting the yield of solar energy systems and their input into the electricity grid up to 48 hours in advance is crucial. Therefore, knowledge about the accuracy of the GHI forecasts from NWP models is needed.

To characterize GFS and IFS forecast accuracy for tropical climates, three countries were selected for study. Our results reveal that: (1) Higher RMSEs are found in tropical rainforest climate, followed by tropical monsoon, and tropical savanna climates. This result was explained by the higher occurrence of cloudy and overcast sky conditions sample found in stations with tropical rainforest climate followed by stations with tropical monsoon and savanna climates. (2) GFS and IFS behave differently with respect to the sky conditions. IFS underestimates the GHI under clear sky conditions and overestimate the GHI under cloudy and overcast sky conditions. Under clear sky conditions GFS underestimates the GHI only in Reunion Island. Under cloudy and overcast sky conditions GFS overestimate the GHI except for Reunion Island where the GHI under cloudy sky is underestimated. The inaccuracies under clear sky conditions were explained by the combination of two phenomena: the misrepresentation of aerosols and the inability of model forecast inaccurately clouds under measured clear sky. This inability to forecast realistic clouds cover also translates in cloudy and overcast sky conditions. (3) By comparing the MAE of IFS-3h and the MAE of IFS-3h-p we found that using the average cycle of GFS to obtain 3-hour worsen the model forecasts. (4) IFS and GFS model biases under clear sky conditions were lower than their biases under cloudy and overcast sky conditions. However, the biases of both models do not decrease systematically as the sky became clearer; the sky index variability also plays an important role in both model biases. (5) IFS-3h, IFS-3h-p outperforms the 3-hour GFS model in all countries. IFS-3h outperforms GFS-3h to a higher extent. Under all sky conditions, with all months and stations merged by territory, the magnitude of improvement for IFS-3h over GFS-3h were found to be 48%, 57%, 37% in French Guiana, Reunion Island, and Singapore, respectively.

This study fills the gap of global NWP models GHI forecast accuracy in tropical climates by studying two of the most widely used NWPs, IFS and GFS, for various types of tropical climates. Comparing the FS score for tropical climate stations to those of Mediterranean, arid, semi-arid, continental, and humid continental stations, we found that forecasting solar irradiance is more difficult in tropical climates. By comparing the FS found in this study with recent studies that calibrated limited-area NWP to forecast GHI on Reunion Island and in Singapore, we showed that using IFS global NWP for GHI forecast is relevant in tropical climates.

Nevertheless, GHI forecast accuracy in tropical climate regions is too poor to be exploited directly; they must be post-processed. Our future research will involve using post-processing techniques (Verbois et al., 2018) to further improve GHI forecast accuracy. The current study may be useful in the following ways: (1) as a benchmark tool to assess the forecast accuracy of a limited area NWP model GHI forecast, and (2) as a reference for post-processing or data assimilation methods to characterize the improvement obtained when forecasting GHI in tropical climates.

Acknowledgments

The authors are grateful to the National Environment Association (NEA), Solar Energy Research Institute of Singapore (SERIS), and Météo-France (French Guiana and Reunion Island) for supplying ground data. The authors also thank the European Center for Medium-Range Weather Forecasts (ECMWF) and the United States National Oceanic and Atmospheric Administration (NOAA) for providing IFS and GFS forecasts.

Funding

This research did receive grant from the following public agencies: “le Centre national d’étude spatiales (CNES)” and “la collectivité territorial de Guyane (CTG)”

Declaration of Interest

The authors declare that there are no competing interests.

III. Forecasting the GHI in the intertropical zone using mesoscale NWP models

Calibration of WRF irradiance in French Guiana and comparison with AROME forecasts

Mouhamet Diallo[a,*], Antoine Primerose[a], and Laurent Linguet[a]

[a]Université de la Guyane – UMR 228 Espace-Dev, BP 792, 0.275 km Route de Montabo, 97337 Cayenne, French Guiana, France

*Corresponding author: mouhametdiallo@gmail.com (M. Diallo)

Abstract

The purpose of this study is to propose a methodology to calibrate Weather and Research Forecast model (WRF) to produce improved GHI forecasts in the intertropical zone (ITZ). We designed a 34 ensemble members with high spatial and temporal resolution to find the set of parameterizations that better characterize sub-grid scale processes associated with clouds, aerosols and their interaction with radiation in the ITZ. This ensemble was constructed using five microphysics schemes (MP), four planetary boundary layer schemes (PBL), two land surface scheme (LSM), two radiation schemes (RAD) and two cumulus schemes (CU). This methodology is validated against GHI measurements during 2016 from six meteorological stations and AROME "outre-mer" (AROME-OM) GHI forecasts in French Guiana. To evaluate WRF and AROME-OM accuracy, the mean average error (MAE), the mean bias error (MBE) and the root mean square (RMSE), were used to study: (1) the seasonal and annual accuracy of WRF and AROME-OM GHI forecasts; and (2) the bias distribution of both models as a function of the clear sky index and the clear sky index variability (V). Results found show that the best WRF GHI forecasts in French Guiana are obtained using: Thomson aerosol aware MP, Mellor-Yamada Nakanishi and Niino PBL, Unified Noah Land LSM, Grell 3D CU, and Rapid Radiative Transfer Model-G RAD for both long and shortwave radiations. We also found that using set of physical parameterizations previously used in other tropical climate is not suitable, the RMSE discrepancy between our set of parameterization and those previously used are as high as 96 W/m² under all sky conditions. The WRF model we calibrated outperforms AROME-OM under clear cloudy and overcast sky conditions in French Guiana. All months and stations were merged. Under clear, cloudy, overcast and all sky conditions WRF improves AROME-OM by 44%, 26%, 8% respectively. This study provides a methodology to calibrate WRF GHI forecasts in the ITZ and cloudy sky areas; it aims to answer the need for reliable solar power forecast driven by the increasing interest in the development of solar technologies.

Keywords: Global horizontal irradiance, radiative transfer parameterization, WRF, tropical zone, AROME

1 Introduction

The world current energetic model is unsustainable; it relies primarily on fossil fuel powered energy (Shafiee and Topal, 2009). According to BP Statistical Review, (2017) oil make up the third of the 2016 primary energy consumption and coal share is approximately 28.1%. However, oil resources are expected to be depleted by 2040, the remaining fossil fuel resources by 2212 (Shafiee and Topal, 2009). The combination of increasing energy needs (Birol, 2010), increasing concern for global warming (Pachauri and Meyer, 2015) fossil fuel depletion (Shafiee and Topal, 2009) and cost decrease (IRENA, 2016) established the solar energy as a solution to meet the needs. Regions with higher solar potential are located in the intertropical zone (ITZ) (Löf et al., 1966; Müller et al., 2015; Trieb et al., 2009). Nevertheless, the solar irradiance in the ITZ is highly variable due to the rapid cloud transformations that occur in this area (Aryaputera et al., 2015; Rossow et al., 2013; Wheeler and Kiladis, 1999). To deal with the uncontrollable nature of the solar energy, the available resource must be forecasted at different time scales so that the global production can be balanced to match users' consumption (Vallance et al., 2017). It is commonly accepted that for forecast horizon larger than 6h, numerical weather prediction (NWP) models are needed (Diagne et al., 2013; Inman et al., 2013). They are two types of NWP model, global and mesoscale. Global model forecast the state of the entire earth atmosphere; they have a typical resolution of 16–50 km, with a forecast horizon up to 15 days ahead (Diagne et al., 2013). Global NWP models have a coarse resolution, to account for local effect mesoscale models were introduced. They have a typical resolution of 5–20 km (Diagne et al., 2013) and forecast the state of the atmosphere on a regional domain. They take global NWP model forecasts as initials and boundary conditions inputs. NWP models with high spatial resolution have the potential to produce more realistic cloud forecasts than lower resolution models (Lorenz et al., 2016) by resolving explicitly smaller scale processes. Yet the resolution of current NWP model is still too coarse to resolve the micro-scale physics associated with cloud formations (Inman et al., 2013; Mathiesen et al., 2013). Because of NWP model coarse grid resolution, sub-grid scale physical process must be approximated (hereafter, parameterized) so that effects of smaller scale processes are accounted for on larger scale (Kleissl, 2013). For a mesoscale model to be used it must be carefully calibrated, using physical parameterization that describes the region climate patterns (Ruiz-Arias et al., 2008). The parameterization schemes express the effect of sub grid, subscale processes on the variables resolved by the NWP spatial resolution (Kleissl, 2013). Most studies dealing with mesoscale models calibration concentrate on locations in the USA, Spain and Germany with continental climate(Lara-Fanego et al., 2012; Perez et al., 2013; Prabha and Hoogenboom, 2010; Ruiz-Arias et al., 2013; Zempila et al., 2015). These countries are amongst the world leaders in installed photovoltaic power capacity (Inman et al., 2013). Tropical regions have been less studied (Aryaputera et al., 2015; Diagne et al., 2014) and are more challenging due to the dynamic weather. Recently (Diallo et al., 2018) studied the accuracy of GFS and IFS global horizontal irradiance (GHI) forecast in the ITZ using measurement forecast from French Guiana, Reunion Island and Singapore. They showed that forecasting the GHI is more difficult in tropical climate than in extra tropical.

In this study is we propose a methodology to calibrate the Weather and Research Forecast (WRF) mesoscale model to produce improved GHI forecast in the intertropical zone. This methodology uses an ensemble of 34 members with high-resolution model simulations. It aims to restrain and select the minimum number of simulations to run, to obtain improved GHI forecasts in the ITZ compared to a non-calibrated model. We validated this methodology against GHI measurements during 2016 from six meteorological stations and AROME "outre-mer" (AROME-OM) GHI forecasts in French Guiana.

This paper is structured as follows: Section 2 describes the dataset used to validate WRF GHI forecast. Section 3 describes the methodology followed to build the ensemble members (EM). Section 4 presents the results of this study. Finally section 5 summarizes the paper and concludes.

2 Data

2.1 Ground Dataset

Our study exploits 2016 hourly data of GHI from six stations of the French national weather services. The characteristics of these stations are described table 1. Due computational capacity constraints associated with the computation of the 34 EM, three months representative of French Guiana seasons were selected. September represents the dry season, May the rainy season and March to describe the latency period between the ITCZ southward to northward motion over French Guiana. Before comparing the NWP GHI forecast to the ground data, ground measurements were quality checked following Espinar et al. (2011) range check formula as follows:

$$0.03GHI_{TOA} < GHI < min(1.2I_0, 1.5I_0 cos(\theta_z)^{1.2} + 100), \quad (1)$$

where GHI_{TOA} (W/m²) is the top of atmosphere irradiance, θ_z the sun zenith angle (°),I_0 the solar constant (1367 W/m²). The top of atmosphere irradiance is defined as follows:

$$GHI_{TOA} = I_0 \varepsilon cos(\theta_z), \quad (2)$$

where ε is the sun correction (Mather and Koch, 2011). Only valid hourly GHI measurements were used to assess the accuracy of WRF and AROME; when the hourly GHI measurements passed the quality check procedure it was then compared to the coincident NWP time stamp to compute the accuracy metrics.

Table 1: French Guiana GHI ground measurement stations

Stations	Latitude (°)	Longitude (°)	Height (m)	Instruments
Rochambeau	4.822	-52.365	4	Kipp & Zonen CM6B
Saint Georges	3.890	- 51.804	6	Kipp & Zonen CM6B
Maripasoula	3.640	- 54.028	106	Kipp & Zonen CM6B
Saint-Laurent	5.485	- 54.031	5	Kipp & Zonen CM6B
Kourou	5.209	- 52.748	12	Kipp & Zonen CM6B
Iles Royales	5.283	-52.583	48	Kipp & Zonen CM6B

2.2 AROME irradiance forecast

AROME France was developed by French weather services (Seity et al., 2011). It was developed to better characterize convection processes and used operationally since December 2008. In this study the authors used AROME “Outre mer” (AROME-OM). AROME-OM produces meteorological forecast

in French West Indies while AROME-France which produces meteorological forecasts in mainland France. Unlike AROME France it doesn't have its own data assimilation system, it is initialized using ARPEGE forecast as lateral boundary conditions and IFS forecast as initial conditions. AROME-OM GHI forecasts used in this study were downloaded online from French weather services catalogue servers (Meteo France, 2017). All forecasts originated at 00:00 UTC and were run for 36h. The main setups used to produce AROME-OM GHI forecasts are shown table 2.

Table 2: AROME-OM operational setup (Termonia et al., 2018; YESSAD, 2015)

Dynamical Core	
Initialization	Initial and boundary conditions provided by ARPEGE and IFS
Spatial resolution of Forecasts	-Horizontal: 2.5 km -Vertical: 90 levels
Temporal resolution of Forecasts	-Computation time step: 1 min -GHI output: hourly
Domain	-Single domain - Latitude: [1.05: 8.95] -Longitude: [-46.3:-56.75]
Physical Parameterization	
Radiation scheme (RAD)	-*RRTMG* (Iacono et al., 2008) *both for Longwave and Shortwave radiations* -*Call frequency: 15 min* -*Aerosol dataset:* (Tegen et al., 1997)
Microphysics scheme (MP)	-*ICE3 Single moment scheme, 5 types of hydrometeors* (PINTY et al., 1998) - *Call frequency: 1 min*
Planetary boundary layer scheme (PBL)	-Turbulent kinetic energy scheme (Cuxart et al., 2000), associated with a mass flux scheme for shallow convection (Bechtold et al., 2001) - *Call frequency: 1 min*
Land surface scheme (LSM) Cumulus scheme (Cu)	-*ISBA scheme* (Noilhan and Planton, 1989) -Eddy diffusivity max flux approach (Pergaud et al., 2009) - *Call frequency: 1 min*

3 Methodology

According to Warner, (2010) calibrating a NWP model involves the following steps: (1) determine the prevailing physical process in the geographic area of interest; (2) define the horizontal resolution so that physical processes are resolved within the model grid; (3) define the vertical resolution so it describes accurately vertical structures such as the boundary layers gradients; (4) validate the model accuracy for different seasons. Consequently, the following subsections explain the calibration process for WRF in the ITZ.

3.1 Physical parameterization

IFS and GFS NWP models GHI forecasts are inaccurate in the ITZ; their accuracies are highly sensitive to the cloudiness and its variability (Diallo et al., 2018). Results obtained under measured clear sky conditions were explained by a misrepresentation of aerosols and by clouds wrongly forecasted by the NWP. Under cloudy sky, they were explained by the inability of the NWP model to predict accurately cloud position, extent and content due to the NWP model coarse resolution. From this investigation on the origin of IFS and GFS inaccuracies, we concluded that to improve NWP GHI

forecast in the ITZ, sub-grid scale phenomena such as clouds, aerosols and their interaction with radiation should be better characterized.

Parameterization schemes allow modeling the effect of a sub-grid scale physical process; however, their accuracies highly depends on the weather regime and parameterization schemes interactions in the NWP model (Cossu and Hocke, 2014; Dudhia, 2014). Therefore, in order to find the set of parameterization that forecast the most accurately the GHI in the ITZ the authors built and computed an ensemble of several 48h GHI forecasts using WRF V3.8.1 (NCAR, 2017; Skamarock et al., 2008). The 34 EM used to calibrate WRF 48h GHI forecast, varied the MPs scheme, LSMs scheme, and PBLs schemes. The ITZ is an area with highly variable dynamic in which we encounter significant amounts of convective clouds (Galvin, 2015; Laing and Evans, 2011). The parameterization scheme we chose to vary have a strong influence on the distribution and physical properties of the simulated cloud field (Cintineo et al., 2014; Otkin and Greenwald, 2008; Xie et al., 2012). The EM also includes set of parameterizations previously used in Singapore and Reunion tropical climate by Aryaputera et al., (2015) and Diagne et al., (2014) respectively; they are shown table 3. The MPs scheme, LSMs scheme, and PBLs schemes varied in this study (members 1 to 32) were chosen with respect to the following constraint:

Chosen Microphysics schemes are either double or triple moment. For triple moment microphysics schemes, the mixing ratio of each hydrometeor, their number concentrations and reflectivities are independent whereas for double moment schemes only mixing ratio of each hydrometeor and number concentrations are independent. Triple and double moment schemes are more time consuming than single moment scheme but more computationally efficient than Bin scheme. They allow removing internal assumptions regarding reflectivity and number concentrations respectively which increases the flexibility of these schemes to adapt to the availability of cloud condensation or ice nuclei (Dudhia, 2014). According to (Chosson et al., 2014) using microphysics scheme with higher number of moments improve the representation of microphysical processes which may improve the radiative transfer computations. Yet as higher moment schemes are more realistic, they may expose other NWP model deficiencies and might not lead to better results (Morrison, 2010). Consequently, the following third and second moment microphysics schemes were selected: Thompson aerosol-aware (Thompson and Eidhammer, 2014), Milbrandt-Yau (Milbrandt and Yau, 2005), Morrison (Morrison et al., 2009), WDM6 (Lim and Hong, 2010).

Chosen PBL schemes equally represent local and non-local schemes. Local and non-local schemes differ in the determination of which model layers influence atmospheric condition at a given model level (Cohen et al., 2015). According to (Cohen et al., 2015) non-local schemes represent more accurately deep PBL circulations, though similar accuracy could be obtained using higher order of closure. Therefore, both local and non-local PBL with various orders of closure from 1 to 2 was accounted for. The major weakness of non-local scheme is their high sensitivity to diagnosed quantities that are difficult to define accurately in NWP models (Milovac et al., 2016). Consequently, the following local and non-local PBL schemes were selected: the non-local first order YSU (Hong et al., 2006), the local second order MYNN3 (Nakanishi and Niino, 2009) scheme, the local 1.5 order QNSE (Sukoriansky et al., 2005) scheme, and the mixed local, non-local first order Shin-Hong (Shin and Hong, 2015; Xu et al., 2018) scheme.

Chosen land surface scheme should at least represent vegetation, and soil moisture processes and computationally efficient. According to (Liang, 2012) LSMs that have more comprehensive treatments of physical processes of land reduce the uncertainties of regional climate model. However, Kotsopoulos et al. (2014) showed that more complex LSM may expose other NWP model deficiencies similarly to microphysics scheme (Morrison, 2010). Land surface schemes such as Noah MP (Niu et al., 2011), Pleim Xiu (Pleim and Xiu, 2003; Xiu and Pleim, 2001) and CLM4 (Dai et al., 2010) despite modeling more complex of land physical process were not used in this study because they were not computationally efficient; therefore, they could not be used for operational forecasting (Kotsopoulos et al., 2014; NCAR, 2017). Consequently, the following LSMs schemes were selected: Noah (Tewari et al., 2004), RUC (Benjamin et al., 2004). These LSMs communicate with the PBL schemes through the surface layer scheme. The surface layer scheme of each member were selected in accordance with (NCAR, 2017) recommendations, the one affiliated with the PBL was always preferred.

Due to computational constraints, the Cumulus scheme and radiation scheme may be fixed (members 1–32). RRTMG (Iacono et al., 2008) radiation schemes were chosen for both long wave and shortwave radiations. Previous work (Ruiz-Arias et al., 2013; Zempila et al., 2015) showed that this scheme outperform the other commonly used radiation scheme independently of the other parameterization. This scheme is one of the rare WRF radiative schemes to allow aerosols, ozone radiation feedback in WRF (Jimenez et al., 2016). According to (Jimenez et al., 2016; Ruiz-Arias et al., 2013) the inclusion of aerosols in the calculations shows that RRTMG outperform Dudhia's scheme (Dudhia, 1989) and is able to provide an excellent agreement with GHI observation. Grell 3D (Grell and Dévényi, 2002) cumulus scheme was chosen to model unresolved convective column physics. Previous work in the ITZ (Crétat et al., 2012; Raghavan et al., 2015) showed that this scheme has the ability to reproduce Inter Tropical Convergence Zone patterns. Besides Grell 3D (Grell and Dévényi, 2002) scheme accounts for the shallow convection which improves sub grid-scale clouds, shortwave irradiance feedback (NCAR, 2017).

3.2 Dynamical core setup

To initialize WRF initial and boundary conditions from a global NWP model (Côté et al., 1998; Déqué et al., 1994; ECMWF, 2016b; Environmental Modeling Center, 2003; Majewski et al., 2002). Among the existing global NWP model the usage of Global Forecast System (Environmental Modeling Center, 2003) is the most straightforward; there is no need for users pre-processing steps to ingest GFS forecasts into WRF. Besides GFS forecasts are available free of charge and archives dating back to 1997 are available (NCAR-RDA, 2017). Consequently, we chose in this study GFS forecast for initial and boundary conditions. To choose the initialization, the time needed for WRF to produce a balanced state from GFS coarser initial and boundary condition must be considered. According to (Aryaputera et al., 2015; Diagne et al., 2014) the spin up time in the ITZ ranges between 6 and 12 hours. The GFS model is initialized every 6 hour starting at 00h; consequently, we chose the initialization at 00 UTC to allow for a spin up of 12 hours before the sunrise in French Guiana[4].

To choose the spatial discretization of WRF the spatial resolution of the initial and boundary conditions (icbc) must be accounted for. GFS (GSM v13.0.2) forecasts used in this study ([dataset] NCEP-NWS-NOAA-USDC, 2015) have a spatial resolution of approximately 24 km. According to Lin et

[4] Reference point : WRF spin up

al. (2009) findings, clouds can be realistically simulated with horizontal resolutions higher than 4 km, but not with coarser resolutions. Therefore, there is a dimension mismatch between the resolution needed to simulate realistic cloud and the resolution of the icbcs. To solve this mismatch issue we used three ways nested domain; D01 with a 27 km horizontal resolution, D02 with a 9 km horizontal resolution and D03 with a 3 km horizontal resolution. To allow for each nest to produce accurate sink sand sources we have approximately $\frac{1}{3}$ of D01 surrounding each side of D02 and similarly $\frac{1}{3}$ of D02 surrounding each side of D01 (Gill, 2016). GHI forecasts resulting from the inner domain with a 3 km resolution centered on French Guiana are exploited (Fig. 1). We chose this resolution to find a balance between the computational power required and the ability of the model to produce realistic clouds. Previous studies in the ITZ used a vertical discretization ranging between 35 and 100 levels (Aryaputera et al., 2015; Diagne et al., 2013; Lima et al., 2016; Verbois et al., 2018) for solar energy forecasts. According to (Tselioudis, 2002) the cloud coverage and cloud properties are more accurately forecasted with increased vertical resolution. Consequently, in our study all domains have 100 levels vertical resolution.

Based on the horizontal and vertical resolution the time step of the WRF dynamic core was set to 50s to achieve convergence. All parameterization schemes are called every model time-step except for the radiation schemes called every 3 min.

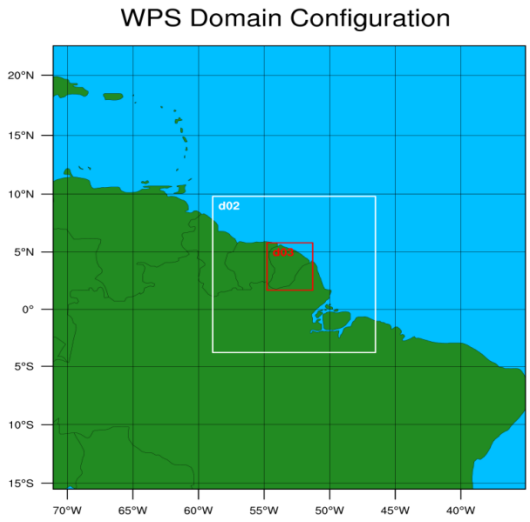

Figure 1: WRF computation domain[5]

4 Results founds

4.1 Error metrics

To study the forecast accuracy of the NWP models we used the root mean square error (RMSE), mean bias error (MAE), and mean bias error (MBE) respectively;

$$RMSE_{\%} = \frac{\sqrt{\frac{1}{n}\sum_i (GHI_{NWP}(t) - GHI_{ground}(t))^2}}{N \sum_t GHI_{ground}}, \quad (3)$$

[5] Reference point : *WRF computation domain*

$$MAE = \sum_t \frac{|GHI(t)_{forecasted} - GHI(t)_{measured}|}{N\ GHI(t)_{measured}} \text{ and} \quad (4)$$

$$MBE_{\%} = \frac{\sum_t GHI_{NWP}(t) - GHI_{ground}(t)}{N\ \sum_t GHI_{ground}\ (t)}. \quad (5)$$

In the metrics mention above $GHI(t)_{NWP}$ is the GHI forecasted by the NWP whereas $GHI(t)_{ground}$ is the ground measured GHI. They are both expressed in W/m².

To assess the accuracy of the NWP models with respect to the sky condition: the clearness index (K_c), clearness index mean value (M_{K_c}) and clearness index mean variability (V) over a time span was used. The sky condition were considered as clear when $K_c > 0.65$, cloudy when $0.4 < K_c < 0.65$ and overcast when $K_c < 0.4$ (Aryaputera et al., 2015).

$$K_c(t) = \frac{GHI\ ground(t)}{G_c(t)}, \quad (6)$$

$$M_{K_c} = \frac{1}{\Delta t}\ \sum_t^{t+\Delta t} K_c(t) \text{ and} \quad (7)$$

$$V = \sqrt{\frac{1}{\Delta t}\sum_{k=1}^{N}(K_c(t+\Delta t) - K_c(t))^2}. \quad (8)$$

where G_c is ESRA clear sky model as defined in (Rigollier et al., 2000). Link turbidity, the clear sky attenuation factor inputted to G_c is defined for each site and each month using monthly climatological values provided by the SoDa database. In equation 7-8, *dt* is the time step of ground measurements, i.e. hourly. For a model with hourly GHI outputs N=1. To illustrate how $\mathrm{MK_c}$ and V were computed, we give the following: at 12 h UTC, $\mathrm{MK_c}$ and V are computed using the $\mathrm{K_c}$ values at 11h UTC and 12 h UTC.

Using previously described metrics WRF and AROME GHI forecasts were validated against, quality checked in situ measurements.

4.2 Intercomparing the accuracies of the ensemble members

According to (Perez et al., 2013) the most important metric in renewable energy forecast is the RMSE because give more weight to the large forecasts errors that has higher impact on the electrical grid management. Therefore, to select the greatest ensemble member we used the RMSE as validation metric.

Figure 2 gives the RMSE of each member for all the month and stations merged under clear cloudy and overcast sky conditions. It shows that there are high discrepancies between the most and least accurate members. Under clear, cloudy, overcast and all sky conditions these discrepancies are 87 W/m², 157 W/m², 201W/m² and 96 W/m² respectively. They are observed between members 16 and 23; members 3 and 34; members 7 and 34; members 9 and 34 respectively. These results show that: (1) varying the PBL scheme, microphysics scheme and land surface model has a strong influence on the forecasted GHI. (2) Using parameterization previously defined in similar climate is not suitable. The member with the worst accuracy under all sky condition is 34, it was previously used by (Diagne et al., 2014) to forecast the GHI in Reunion Island tropical climate.

The best compromise between clear cloudy and overcast sky GHI forecast accuracy is obtained using member 9 (WRF-M9)[6] as show figure 2. This results may be explained by two phenomena**:** (1) Member 9 uses Thompson aerosol aware microphysics scheme (Thompson and Eidhammer, 2014); this scheme exploits an aerosol climatology derived from eight years (2001–2007) global model simulations instead of using Tegen (Tegen et al., 1997) aerosol datasets used in the other microphysics scheme. (Thompson and Eidhammer, 2014) aerosol dataset has a higher spatial resolution: 0.5-degree longitude by 1.25-degree latitude spacing whereas (Tegen et al., 1997) aerosol dataset has 5 by 4 longitudes, latitude spatial resolution. (2) This scheme is the only microphysics scheme up to WRF V3.8 that allows cloud radiation feedback; the effective cloud water, ice and snow radii from Thompson are fed into RRTMG (Jimenez et al., 2015). Consequently, WRF-M9 might improve aerosols and their interaction with radiation characterization which result in a better forecast under all sky conditions.

4.3 Comparing the GHI forecast of WRF Member against AROME French West Indies

In this section the GHI forecasts of WRF using Member 9 (WRF-M9) is compared against AROME French West Indies described section 2.2. We consider the GHI forecasts at the single grid points closest to each station and do not average spatially the GHI forecasts of IFS and GFS. First we study in section 4.3.1 the accuracy of WRF-M9 and AROME-OM when all stations and months are merged. Second we study in section 4.3.2 their accuracies for each station and each month.

4.3.1 Influence of the sky conditions

To analyze the influence of the sky conditions on the accuracy of WRF-M9 and AROME-OM: we give Figure 3, the MAE and MBE of each model under clear, cloudy and overcast sky conditions. Lastly, Figure 4 gives AROME-OM and WRF-M9 the MAE as a function of the clearness index mean value (M_{K_c}) and clearness index mean variability (V) over an hour time span.

Figure 3 shows that under all sky conditions WRF-M9 GHI forecast outperforms AROME-OM GHI forecast; the MAE of WRF-M9 is 139 W/m² whereas the MAE of AROME-OM is 206 W/m². To explain the behavior of WRF-M9 and AROME-OM under all sky, we study their individual performance under clear cloudy and overcast sky conditions.

Under measured clear sky conditions Figure 3 shows that the MAE and MBE of WRF-M9 and AROME-OM are the lowest. This result is in agreement with Diallo et al. (2018) findings in French Guiana. According to Jimenez et al. (2015) under clear skies, the forecast accuracy of the GHI is driven by aerosols. To analyze the effect of the different aerosol dataset used by WRF-M9 and AROME-OM on the GHI forecasts, we computed the MAE for samples that are reported as clear both by the NWP model and ground measurements. We found for all months and stations merged by territory that aerosols cause a MAE (MBE) of 80 W/m² (40 W/m²), for WRF-M9. For AROME-OM aerosols cause a MAE (MBE) of 167 W/m² (112 W/m²). From the positive MBE of WRF-M9 and AROME-OM we deduce that both WRF-M9 and AROME-OM overestimate of the aerosol optical depth (Jimenez et al., 2016); however, WRF-M9 to a lesser extent. WRF-M9 uses Thompson and Eidhammer, 2014 aerosol datasets which has a higher spatial than the one used in AROME-OM (Tegen et al., 1997) which is also available in WRF. Consequently, we may believe that WRF-M9

[6] Reference point : WRF-M9

aerosol dataset represent more accurately the aerosol load and distribution in French Guiana which in turn lead to better forecasts under clear sky condition.

Despite the ability of WRF-M9 and AROME-OM detect the occurrence of cloudy and overcast sky conditions being weak; they are in agreement with value reported in (Diallo et al., 2018). According to Lin et al. (2009) findings, clouds can be realistically simulated with horizontal resolutions higher than 4 km, but not with coarser resolutions. Consequently, both models have the potential to simulate realistic clouds; nevertheless, there are still high MAEs under cloudy and overcast sky conditions. The MAEs of WRF-M9 under cloudy and overcast sky conditions are 172 W/m² and 230 whereas the MAEs of AROME-OM are 185 W/m² and 231 W/m² respectively (Fig. 3). According to Mathiesen et al. (2013); Yang and Kleissl (2016) the model initialization is critical for NWP forecast accuracy; initial conditions derived from large-scale models will inherit the error of the parent model. We have showed in Diallo et al. (2018) that both IFS and GFS are inaccurate in cloudy and overcast sky conditions which may translate in AROME-OM and WRF-M9 forecasts respectively. As WRF-M9 and AROME-OM has quasi-similar horizontal (e.g 3 km vs 2.5 km) and vertical resolution (e.g 100 levels vs 90 levels); we may explained WRF-M9 improvements over AROME-OM (Tab.4) under cloudy and overcast sky conditions by its more frequent call of the radiation physics. WRF-M9 hourly GHI forecast is the result of 20 accumulated irradiance values divided by the accumulation time (eg, 1 hour) whereas AROME-OM hourly GHI forecast is the result of 4 accumulated irradiance values divided by the accumulation time (also 1 hour).

figure 4 shows that the MAE does not decrease systematically as the sky conditions become clearer. Figure 4 also show that the variability plays a key role on each model accuracy. For highly variable sky conditions $V \geq 0.4$, WRF-M9 outperforms AROME-OM (Fig. 4). We explain this result by the higher frequency of the radiation physics of WRF-M9.

Table 3: WRF 34 ensemble members' physical parameterization

	Member number																																
Physics	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
PBL																																		
YSU	■	■	■	■	■	■	■	■																									■	■
MYNN									■	■	■	■	■	■	■	■																		
QNSE																	■	■	■	■	■	■	■	■										
SHIN-HONG																									■	■	■	■	■	■	■	■		
MP																																		
Thompson	■				■				■				■				■				■				■				■					
Milbrandt		■				■				■				■				■				■				■				■				
Morisson			■				■				■				■				■				■				■				■			
WDM6				■				■				■				■				■				■				■				■		
WSM6																																	■	
WSM3																																		■
LSM																																		
Noah	■	■	■	■					■	■	■	■					■	■	■	■					■	■	■	■					■	■
RUC					■	■	■	■					■	■	■	■					■	■	■	■					■	■	■	■		
CU																																		
Grell-3D	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	
Grell-Devenyi																																	■	
Kain Frisch																																		■
RAD LW																																		
RRTMG	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
RAD SW																																		
RRTMG	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		
Dudhia																																	■	■

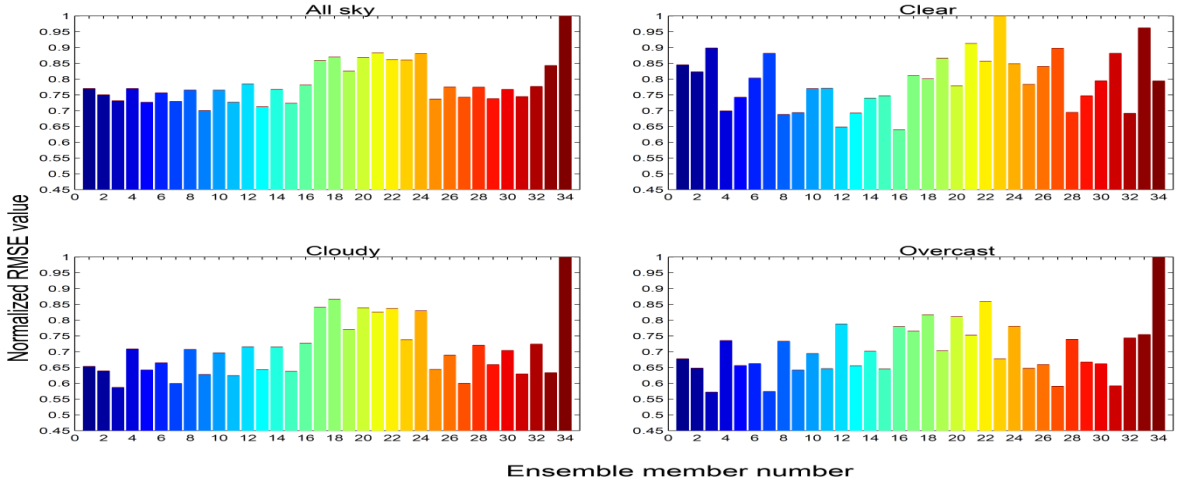

Figure 2: Normalized RMSE for each ensemble members all months and stations merged. The absolute RMSE (W/m²) of each member are normalized with the RMSE value of the less accurate member for the considered sky conditions. The normalization value under all, clear, cloudy and overcast sky conditions are 272, 208, 315, and 437 W/m² respectively. They are obtained for member 34, 23, 34, 34

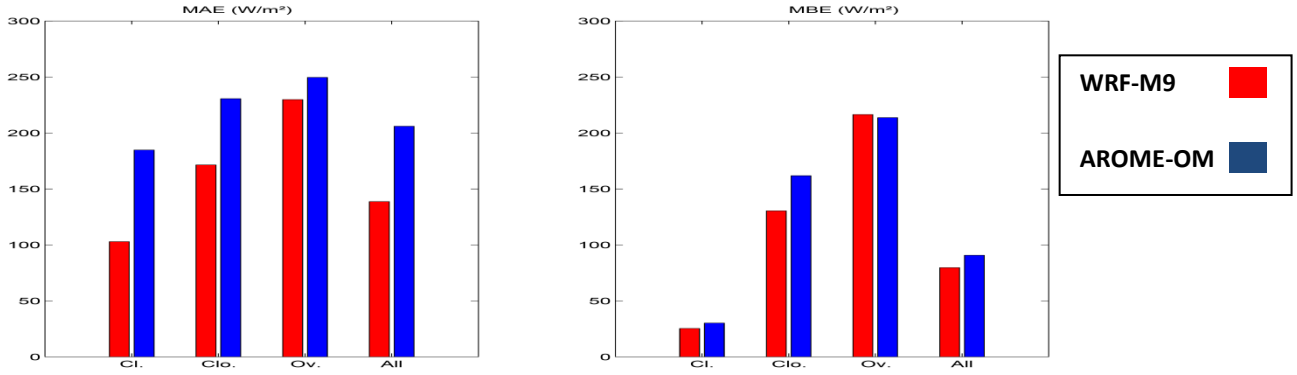

Figure 3: MAE and MBE of GHI forecast from AROME-OM and WRF-M9. Months and stations were merged. The classification of sky conditions is based on ground measurements.

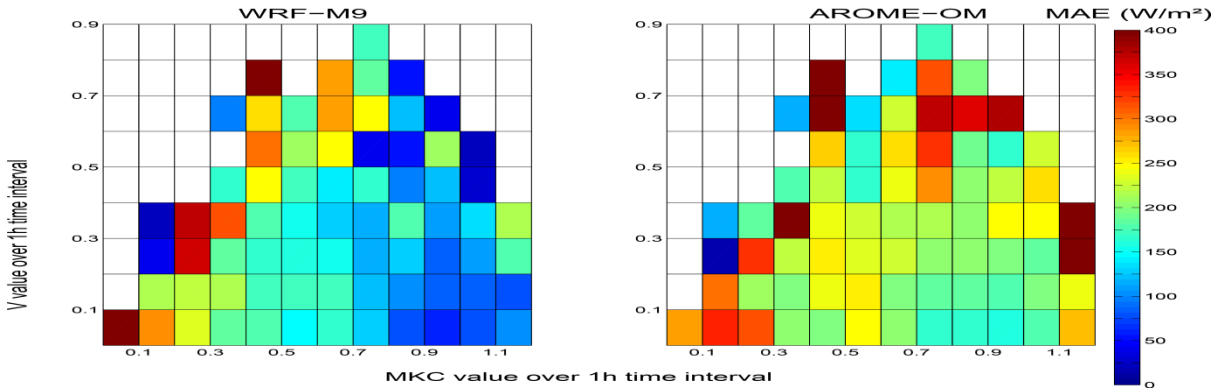

Figure 4: WRF and AROME-OM bias distributions with respect to the mean of the previous hours K_c (MK_c) values and the mean of the previous hour K_c variability (V). Months and stations were merged

Table 4: Relative improvement (%) of the MAE of WRF-M9 (W/m²) over the MAE of AROME-OM (W/m²). Months and stations were merged.

Territory	Relative improvement (%) of the MAE of WRF-M9 (W/m²) over the MAE of AROME-OM (W/m²)			
	Clear	Cloudy	Overcast	All
French Guiana	44%	26%	8%	33%

4.3.2 Seasonal forecast accuracy under all sky conditions

Figure 5 shows WRF-M9 and AROME-OM monthly MBE and MAE. Both AROME-OM and WRF-M9 GHI forecasts follow seasonal trends for French Guiana. There is a clear distinction between the MAE value dry season represented by September and the MAE value in the rainy seasons represented by May. March is a transition period between the southward and northward motions of the Inter Tropical Convergence Zone, and depending on the year, this month behaves as a rainy or dry. Consequently, except for Maripasoula the MAE in March is overall higher than the MAE value in the dry season but lower than the MAE values in the rainy season.

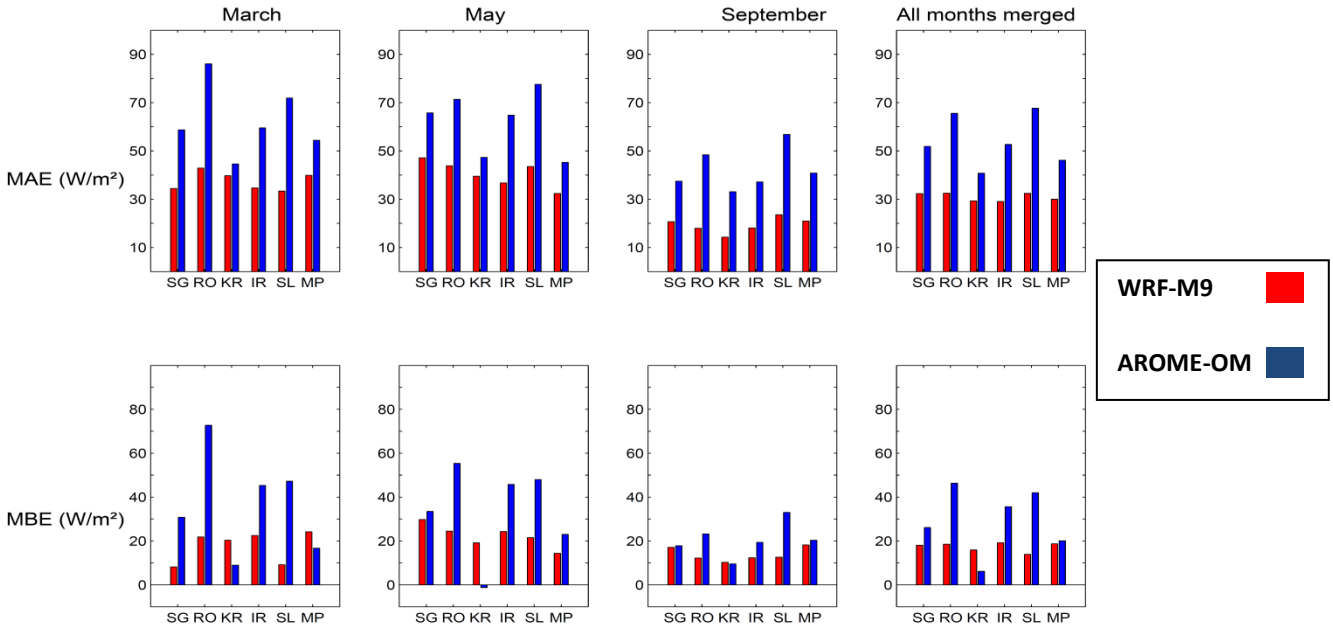

Figure 5: WRF and AROME-OM relative MBE and MAE for rainy season (May), dry season (September) and March transition month. The MAE and MBE were normalized using monthly mean values to obtain % values. Monthly GHI of SG in March, May, September and all months merged are 316, 312, 450 and 359 W/m² respectively; for RO they are 307, 342, 456 and 370 W/m² respectively; for KR they are 335, 362, 484 and 393 W/m² respectively; for IR they are 352, 358, 468 and 390 W/m² respectively. For SL they are 392, 324, 422 and 381 W/m² respectively; for MP they are 348, 386, 454 and 397 W/m² respectively

The different forecast accuracies found figure 5 may be explained by different cloud cover and cloud cover variability. The Lowest MAE are found during September for all stations (Fig. 5); Oppositely higher RMSE are found during May in all stations (Fig. 5) except for Rochambeau and Maripasoula. We explain results found in SG, KR, IR and SL by the higher variability of sky conditions found during May compare to March; figure 4 shows that both model underperform for highly variable sky conditions. For RO and MP May as higher occurrence of cloudy sky than March; we believe that it outweigh the high variable sky conditions and gives better RMSE.

Under all sky conditions and all months merged the MAE of WRF-M9 ranges between 29% and 33% whereas the MAE of AROME-OM ranges between 40% and 68% (Fig. 5). The improvement of WRF-M9 MAE over AROME-OM ranges between [35:52]% depending on the stations. WRF-M9 MAE values has similar standards as those of other climate regions despite forecast the GHI in tropical climate being more difficult (Diallo et al., 2018). Perez et al., (2013) showed that WRF 48 h GHI forecasts in the USA, Central Europe, Spain and Canada ranged between [17:39] %, [32:40] %, [16:14] %, [30:32] %.

5 Conclusions

The goal of this study was to propose a methodology to calibrate WRF so that it produced improved GHI forecasts in the ITZ compared to non-calibrated mesoscale model. This methodology was validated against ground measurements and AROME-OM GHI forecasts in French Guiana during 2016. French Guiana is a French territory located in the ITZ; its photovoltaic installed capacity is projected to increase by 112% by 2023. Because there are higher solar potential in the ITZ predicting the solar input to the solar facilities which will lead the electricity injected to the grid is vital. Therefore, knowledge of the forecast accuracy of NWP models is needed. We found in this study that using set of physical parameterization previously used in other similar climate is not suitable. It can worsen the all sky conditions RMSE by 96 W/m² compared to the calibrated model. The set of parameterization derived from our calibration (WRF-M9) outperform AROME-GHI forecast in clear cloudy and overcast sky conditions. Improvements under clear sky conditions were explained by the finer and more recent aerosol dataset used by WRF-M9 and the ability of WRF to reproduce more accurately the sky conditions. Improvements under cloudy and overcast sky conditions are explained by the higher call frequency of the radiation physics. The seasonal and accuracy study showed that MAE were lower in the dry season represented by September and higher in the rainy season represented by May. These different forecast accuracies was explained by different cloud cover and cloud cover variability found during these months. Under all sky conditions and month merged, the improvement of WRF-M9 MAE over AROME-OM ranges between [35:52]% depending on the stations.

To the best of the author's knowledge several studied showed that the MPs scheme, LSMs scheme, and PBLs schemes, as these schemes have a strong influence on the distribution and physical properties of the simulated cloud field (Cintineo et al., 2014; Otkin and Greenwald, 2008; Xie et al., 2012); however, no other studies analyzed the effect of varying these physics on the irradiance forecast in the ITZ. This study is the first dealing with the calibration of mesoscale models in French Guiana; results found that this calibration of WRF has similar standards as those of other climate regions.

Nevertheless, these GHI forecasts must be improved as PV managers have financial incentive to produce accurate forecasts of electricity production. Our future research will involve using data assimilation methods to improve the initialization of WRF forecasts (Mathiesen et al., 2013; Yang and Kleissl, 2016).

Acknowledgments

The authors are grateful to the Météo-France (French Guiana and Reunion Island) for supplying ground data and AROME forecasts. The authors also thank the European Center for Medium-Range Weather Forecasts (ECMWF) for providing access to their computing facilities to run WRF.

Funding

This research did receive grant from the following public agencies: “le Centre national d’étude spatiales (CNES)” and “la collectivité territorial de Guyane (CTG)”

Declaration of Interest

The authors declare that there are no competing interests.

IV. Summary and conclusion of chapter III

In Chapter III we presented first methods recommended for forecasting the GHI depending on the horizon of interest. Using this knowledge, we presented methods that forecast the GHI using NWP models. First we downloaded and studied the accuracy of IFS and GFS forecast products in the ITZ for tropical climate areas. That study filled the gap on the knowledge of global NWP models performance in the ITZ and for tropical climates. We studied two of the most widely used NWP models for various types of tropical climates located in French Guiana, Reunion Island and Singapore. Results found showed by comparison to the persistence model that forecasting the GHI in tropical climate is more difficult than forecasting the GHI in extra-tropical climates. Besides, under all sky conditions the GHI forecasts of IFS outperform those GFS in all territories; nevertheless, both models have high MAE values under clear, cloudy and overcast sky conditions. High MAE under measured clear sky conditions were explained by a misrepresentation of aerosols and clouds wrongly forecasted by the NWP whereas high MAE under cloudy and overcast sky conditions were explained by the coarse resolution of IFS and GFS that render them unable to produce realistic clouds. From this investigation on the origin of IFS and GFS inaccuracies, we concluded that to improve NWP GHI forecasts in the ITZ, sub-grid scale phenomena such as clouds, aerosols and their interaction with radiation should be better characterized. Consequently, the authors built and computed an ensemble of several 48h GHI forecasts using WRF. The aim of the second study was to suggest a generic method to calibrate mesoscale models to obtain improved GHI forecasts in the ITZ. The ensemble members used to calibrate WRF 48h GHI forecast varied the MPs scheme, LSMs scheme, and PBLs schemes. The ITZ is an area with highly variable dynamic in which we encounter significant amounts of convective clouds (Galvin, 2015; Laing and Evans, 2011). The parameterization scheme we chose to vary have a strong influence on the distribution and physical properties of the simulated cloud fields (Cintineo et al., 2014; Otkin and Greenwald, 2008; Xie et al., 2012). Results found showed that under all sky conditions, the improvement of WRF MAE over AROME-OM ranges between [35:52]% depending on the stations. We also found that MAE values have similar standards as those of extra-tropical regions despite forecast the GHI in tropical climate being more difficult.

The scope for future research includes extending this methodology to other locations in the inter-tropical zone and using data assimilation methods to improve the initialization of WRF forecasts (Mathiesen et al., 2013; Yang and Kleissl, 2016).

CHAPTER IV : Improving the irradiance forecast of numerical weather prediction model using data assimilation

I. Improving the solar irradiance forecast under tropical climate for long range forecast horizon

Solar radiation forecasts and especially GHI forecasts are needed for integrating into the grid and at a large scale the electricity production from PV plants. Increasing the installed capacity of PV power requires that solar irradiance forecasts be more and more accurate in terms of spatial and temporal resolutions (Thorey et al. 2015). However, NWP models are still highly biased in cloudy sky conditions; these conditions occur more frequently in the ITZ. According to Kleissl (2013), errors related to irradiance forecasts can have a wide variety of sources. These sources include: poor model initialization, excessively coarse vertical grid spacing, and inaccurate assumptions in the physical parameterizations. For this reason several methods were developed to increase the accuracy of NWP models irradiance forecasts. These methods could be classified into three categories:

-Model output statistics (MOS)
-Multi model ensemble (MME)
-Data assimilation (DA)

MOS is the most frequent mean to improve solar irradiance forecasts; it is a post-processing method, which consists in the determination of a relationship between the GHI and other ground measured variables in order to refine the output of NWP models (Glahn and Dale, 1972). This method was successfully used by Diagne et al. (2014), Lauret et al. (2014), Mathiesen et al. (2013), and Verbois et al. (2018). They found RMSE improvement of WRF GHI forecasts was high as 40%, 18%, 19%, and 33% respectively using a Kalman filter, a multivariate fourth-order regression and combination of a principal component analysis with a stepwise variable selection. However, these methods were exclusively based on statistical properties of the time series and do not use the physical properties of solar irradiance; therefore, they are site specific. Beside MOS need a training period require at least one year of archive. During the training and validation period the NWP setup and physics must not be changed (Kalnay, 2009).

MME combines several forecasts from one or several NWP models (e.g. an ensemble) in order to produce a single forecast hopefully more skillful than any individual model of the ensemble (Mallet et al., 2009). This methodology was used by Thorey et al. (2015) who combined linearly GHI forecasts from six NWP models. Thorey et al. (2015) found that for a 42-hour forecast the RMSE of the worst NWP is improved by as much as -24%. The advantage of multi-model ensemble methods is that they rely on several NWP; different models capture the same physical phenomenon differently which tend to reduce the uncertainty based on each individual model (Xue and Zhang, 2014). Nevertheless, multi-model ensemble has several limitations. The combination of the ensemble forecasts is determined to minimize its discrepancy with the observations; since the observations are not perfect, this approach is not entirely satisfactory. The second limitation is that the weights are computed only at the locations and for the variables that are observed. Computing weights for other locations and other do not guaranty similar accuracy (Mallet, 2010). Besides, similar to MOS methods, MME methods need an extensive training period and the NWP models must be frozen during the training and validation periods.

Data assimilation (DA) methods merge the physical information provided by numerical models and the information brought by the observations, in order to improve the forecast. These methods date back to the early seventies (Daley, 1999) and have been tested for several purposes such as improving storm forecasts (Fierro et al., 2014), improve wind forecasts (Portabella and Stoffelen,

2004), improve cloud forecasts (Benjamin et al., 2002), improving aerosol concentration forecasts. However, a literature study we carried out show that only a few publications using DA to improve solar irradiance forecasts were available (Mathiesen et al., 2013; Sahu et al., 2016; White et al., 2016; Yang and Kleissl, 2016). Nevertheless, these studies showed promising results. Clear advantages of DA over MME are that (1) DA methods take into account observational errors which make a difference when significant instrumental errors are involved, (2) DA methods are not site specific: DA improvements apply to the NWP forecast region (3) they are autonomous, they may be used with different NWP models and (4) they do not require a training period. Bourgin et al. (2014) investigated the interactions between data assimilation and post-processing in hydrological ensemble forecasting; he found that DA method outperform MOS especially for forecast horizon lower or equal to 12h. For these reasons we considered DA methods to improve solar irradiance forecasts in this thesis.

II. Data assimilation methods

DA is a technique in which an accurate image of the true state of the atmosphere at a given time (i.e analysis) is accumulated into a NWP model state to improve its forecasts (Bouttier and Courtier, 1999; Holm, 2008). This accurate image of the true state of the atmosphere is called the analysis. This analysis is constrained by the physical laws of the NWP model set of equations. Fig. IV.1 and IV.2 shows the type of observation used to build this accurate image, and their distribution around the globe.

DA methods can be divided into four categories (Bannister, 2008a; Daley, 1999; Kalnay, 2009; Lakshmivarahan and Lewis, 2013):

-interpolation method (e.g., function fitting)
-empirical methods (e.g. successive correction method)
-variational method (e.g., 4D-var, 3D-var)
-stochastic method (e.g., Kalman filter and ensemble Kalman filters)
-nudging methods

In this manuscript we focus only on real time assimilation algorithms systems because they only consider observations from the past (Bouttier and Courtier, 1999). Therefore, they are relevant in the operational management of solar energy systems. Consequently, interpolation and empirical methods are not discussed as they are outdated and no longer used in current NWP (ECMWF, 2016b; Environmental Modeling Center, 2003; Majewski et al., 2002; Staniforth et al., 2006). We present in the main body of the manuscript the general framework of stochastic, variational and nudging DA algorithms. Further details on the different DA methods, are given Appendix A.

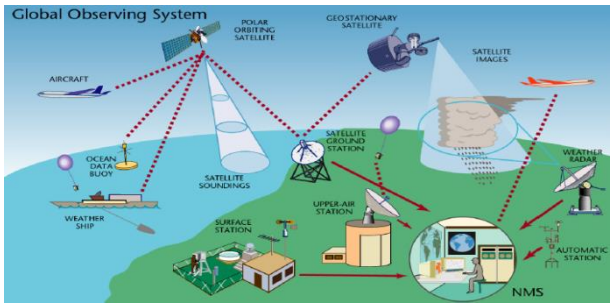

Fig. IV.1 Measurements instrument used in DA system credit WMO

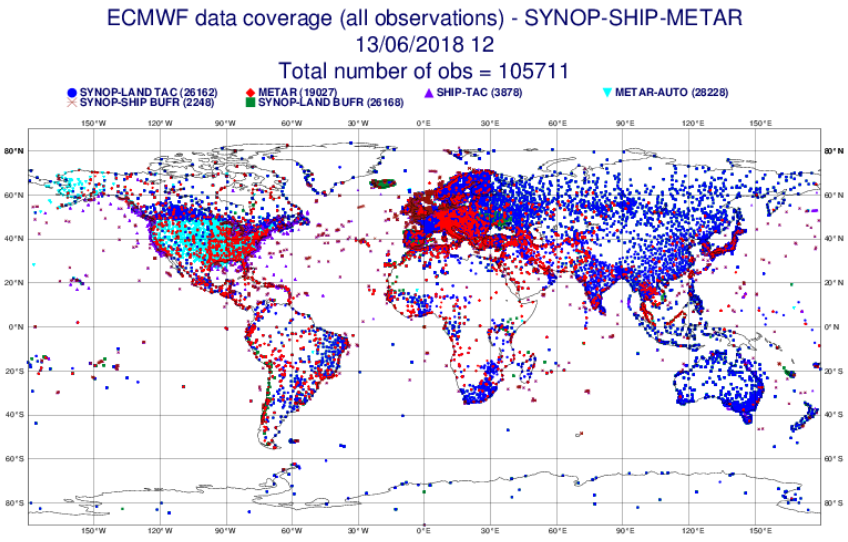

Fig. IV.2 Observation network in ECMWF DA system

II.1 Stochastic method and variational methods

Stochastic DA methods include commonly used in commonly used in atmospheric science (Kalnay, 2009) include ensemble square root filters (EnSRF) and ensemble Kalman filters (ENKF); while variational DA algorithms include 3D-Var and 4D-Var (Appendix A). Bennett (2004), Holm, (2008) and Kalnay (2009) showed that stochastic and variational DA algorithms can be derived from one common source. Consequently, The DA problem from the stochastic and variational framework is expressed as follows:

$$\vec{\mathrm{x}}(\mathrm{t}+1) = \vec{\mathrm{M}}\left(\vec{\mathrm{x}}(\mathrm{t})\right) + \overrightarrow{\varepsilon_{\mathrm{m}}}, \qquad \text{IV.1}$$

$$\vec{\mathrm{x}}(\mathrm{t}_0) = \overrightarrow{\mathrm{x}_{\mathrm{b}}}(\mathrm{t}_0) + \overrightarrow{\varepsilon_{\mathrm{b}}}\ and \qquad \text{IV.2}$$

$$\vec{\mathrm{y}}(\mathrm{t}_{\mathrm{n}}) = \vec{\mathcal{H}}\left(\vec{\mathrm{x}}(\mathrm{t}_{\mathrm{n}})\right) + \overrightarrow{\varepsilon_{0,\mathrm{n}}}. \qquad \text{IV.3}$$

Where $\vec{x} \in \Re^n$ is the NWP model state vector; $\overrightarrow{x_b} \in \Re^n$ the background state; $\vec{y}(t_n) \in \Re^p$ the observation of the state variable over a given time interval n; $\vec{M}(\vec{x}) \in \Re^n$ the model operator; $\vec{\mathcal{H}} \in \Re^{n \to p}$ the observation operator which computes the model equivalent of an observation $\vec{y}$ made a given time and location; $\overrightarrow{\varepsilon_m} \in \Re^n$, $\overrightarrow{\varepsilon_b} \in \Re^n$ and $\overrightarrow{\varepsilon_0} \in \Re^p$ approximate the model errors, the initial condition error (i.e., background error), and observation error respectively. $\overrightarrow{\varepsilon_{0,n}}$ is used a simplification for $\overrightarrow{\varepsilon_0}(t_n)$.

To model the uncertainty related to the model errors, background errors, and observation errors, probability distribution functions (pdfs) are used. Under the assumption that the background observation and model error pdfs are Gaussian, the probability to find the model state at x knowing the observation error is:

$$P(\vec{x}|\vec{y}) = \frac{1}{(2\pi)^{\frac{3}{2}}|B||R||Q|}\, e^{\frac{-1}{2}\left[\overrightarrow{\varepsilon_b}^{T} B^{-1} \overrightarrow{\varepsilon_b} + \overrightarrow{\varepsilon_m}^{T} Q^{-1} \overrightarrow{\varepsilon_m} + \sum_{n=0}^{N} \overrightarrow{\varepsilon_{0,n}}^{T} R^{-1} \overrightarrow{\varepsilon_{0,n}}\right]}, \quad \text{IV.4}$$

where $B \in \Re^{n.n}$,$Q \in \Re^{n.n}$ and $R \in \Re^{p.p}$ are the model error covariance matrix, background error covariance matrix and observation error covariance matrix respectively. They are defined as follows:

$$B = < (\overrightarrow{\varepsilon_b} - < \overrightarrow{\varepsilon_b} >) (\overrightarrow{\varepsilon_b} - < \overrightarrow{\varepsilon_b} >)^T >, \quad \text{IV.5}$$

$$R = < (\overrightarrow{\varepsilon_0} - < \overrightarrow{\varepsilon_0} >) (\overrightarrow{\varepsilon_0} - < \overrightarrow{\varepsilon_0} >)^T > and \quad \text{IV.6}$$

$$Q = < (\overrightarrow{\varepsilon_m} - < \overrightarrow{\varepsilon_m} >) (\overrightarrow{\varepsilon_m} - < \overrightarrow{\varepsilon_m} >)^T >. \quad \text{IV.7}$$

Similarly we define the analysis error $\overrightarrow{\varepsilon_a} \in \Re^n$ which measures the departure of the analysis to the true state and the analysis covariance matrix $P \in \Re^{n.n}$:

$$P = < (\overrightarrow{\varepsilon_a} - < \overrightarrow{\varepsilon_a} >) (\overrightarrow{\varepsilon_a} - < \overrightarrow{\varepsilon_a} >)^T >. \quad \text{IV.8}$$

In Eq.IV.5 to IV.8 < > and ^{T} represent the expectation and transposition operators, respectively.

The most important elements of DA methods are the covariance matrices *B*, *Q* and *R*. They determine to what extent the background fields will be corrected to match the observations (Bouttier and Courtier, 1999) and how the model error influence the analysis.

Q describes the correlations between errors in the model between several locations and for several time instances. Model errors can be caused by physical processes not described by the model equations, or by inaccurate physical parameterizations (Holm, 2008). In most assimilation algorithm the model is assumed to have no error; therefore, $Q = 0$.

B spreads the information vertically and horizontally with proper weights to observation y_n and the background $\overrightarrow{x_b}(t_0)$. Thus, if $|B|$ is very large compared to observation errors, the analysis is closer to observation, otherwise the analysis is closer to the background (Descombes et al., 2015; Liu, 2017; Rizvi, 2016). B can be either be approximated as static or dynamic. For the static case, B is approximated using long term simulations (i.e climatological); For the dynamic case B is approximated using a set of different numerical simulation (i.e an ensemble) that varies with the time. To approximate B several approaches may be used: the analysis of innovations method, the differences of varying length forecasts (NMC method), the forecast time lags method (CQ) and the ensemble method (Bannister, 2008a).

R includes the effects of measurement errors, errors in the design of the observation operators and representativeness errors (Bouttier and Courtier, 1999; Holm, 2008; Kalnay, 2009) . *R* is specified according to the knowledge of instrumental characteristics; in WRF (Skamarock et al.,

2008) for instance a standard table of observation error values provided by the US Air Force Weather Agency is used (Kavulich, 2017). This table contains error values broken down by observed variable, observation type, and pressure level.

The most likely state of the atmosphere (analysis) maximizes the joint probability Eq.IV.9. This maximum is attained when $\vec{x} = \overrightarrow{x_a}$, the analysis state vector. Consequently, the following cost function is minimized (Kalnay, 2009):

$$J(\vec{x}(t_0)) = \frac{1}{2}\left[\overrightarrow{\varepsilon_b}^T B^{-1}\overrightarrow{\varepsilon_b} + \overrightarrow{\varepsilon_m}^T Q^{-1}\overrightarrow{\varepsilon_m} + \sum_{n=1}^{N} \overrightarrow{\varepsilon_{0,n}}^T R^{-1}\overrightarrow{\varepsilon_{0,n}}\right] = J_b + J_m + J_0. \quad \text{IV.9}$$

Where J_b, J_m and J_0 are the background, observation and model cost functions respectively. The cost function J aims to find an optimal estimate for the initial state of the system $\vec{x}(t_0)$ given a prior estimate $\overrightarrow{x_b}(t_0)$ of the initial state and observations at several points t_n; $n = 0, \dots, N$ distributed over a time window$[t_0 : t_N]$.

Variational DA algorithms minimize Eq.IV.9 to obtain an analysis whereas stochastic DA solve directly the analysis equation (Eq.IV.9) directly by inversion (Appendix A). In operational NWP centers, to reduce computational costs, a sequence of linear approximations to the nonlinear minimization problem Eq.IV.9 using successive small increments $\delta\vec{x}$, given an estimated analysis $\vec{x}(t_0)$.

$$\delta\vec{x}(t_0) = \vec{x}(t_0) - \vec{x}_b(t_0) \quad \text{IV.10}$$

In practice the assimilation cycle has two step, the update and prediction step. If observations $y(t_n)$ are available, they are used in the update step to compute the analysis state vector $\overrightarrow{x_a}$ and its error covariance matrix P is then propagated into time during the prediction step.

II.2 Nudging methods

The nudging DA (NDA) algorithm is originally an empirical method that consists in adding to the prognostic equations a term that relaxes (nudges) the solution towards the observations. Current nudging DA assimilation method are no longer empirical, they derive the nudging coefficient from variational or ensemble methods. They are two types of nudging: the observation nudging and the analysis nudging (Auroux and Blum, 2008; Lakshmivarahan and Lewis, 2013; Lei, 2011; Lei and Hacker, 2015). In the case of analysis nudging, the model state is nudged toward gridded analysis based on the analysis error whereas in observation nudging, the model predictions are nudged to match better with observations at individual locations both on the surface and above (Xiangshang Li et al., 2016) based on the model error. The DA problem from the observation nudging framework is expressed as follows (Lakshmivarahan and Lewis, 2013):

$$\begin{cases} \vec{x}(t+1) = \overrightarrow{M}(\vec{x}(t)) + G_0(t)\overrightarrow{\varepsilon_0}(t) \; and \\ \vec{x}(t_0) = \vec{x}_0 \, . \end{cases} \quad \text{IV.11}$$

Whereas from analysis nudging, it is express as:

$$\begin{cases} \vec{x}(t+1) = \overrightarrow{M}\left(\vec{x}(t)\right) + G_a(t)\,\overrightarrow{\varepsilon_a}(t) \; and \\ \vec{x}(t_0) = \vec{x}_0 \,. \end{cases} \quad \text{IV.12}$$

Where $\overrightarrow{\varepsilon_0}, G_0, \vec{\varepsilon}_a, G_a$ are: the observation error, observation nudging coefficient matrix, analysis error and analysis nudging coefficient matrix respectively.

A variational variant of nudging method named back and forth nudging method was developed recently by Auroux and Blum (2008), (2005). The back and forth nudging algorithm consists in first solving the forward nudging equation (Eq.IV.11) and then the backward nudging equation. The initial condition of the backward integration is the final state obtained after integration of the forward nudging equation. At the end of this process, one obtains an estimate of the initial state of the system (Appendix A). The main difficulty in the NDA scheme resides in the estimation of the nudging coefficient G. They could be either computed empirically, statistically or using a variational approach (Kalnay 2009; Zou et al. 1992). The latter approach gives the optimal nudging method (Lakshmivarahan and Lewis, 2013).

III. Comparison between DA algorithms

A literature study we carried out showed that among DA methods commonly use in atmospheric sciences, only 3D-Var and nudging were used to improve solar irradiance forecasts. Sahu et al. (2016) recently used 3D-var data assimilation to improve irradiance forecast over southern California. Sahu et al. (2016) ran WRF for two consecutive days and noticed that 3-DVAR improved the MAE by 65%. Most available studies used nudging type method (Mathiesen et al., 2013; White et al., 2016; Yang and Kleissl, 2016). Mathiesen et al. (2013) and Yang and Kleissl (2016) used conjointly WRF (Skamarock et al., 2008) NWP model with GOES cloud cover images to increase or decrease artificially the water vapor mixing ratio to produce or dissipate clouds simulated in the NWP models so that they match GOES imagery. Yang and Kleissl (2016) found that their nudging method improved by 47% the solar irradiance forecast of WRF; yet under clear sky conditions, the model was negatively biased (Mathiesen et al., 2013); indicating that for some forecasts clouds were incorrectly dissipated. White et al., 2016 also used conjointly WRF NWP model with GOES cloud cover images but nudged the horizontal wind fields positively to produce cloud and negatively to dissipate clouds. This methodology was validated against two stations and for two months; it was found that the correlation coefficient was found to be greater for the cloud assimilation simulation, 0.71 compared to 0.65 and 0.67 compared to 0.47 at both locations. The main disadvantage of these methods is: these methods introduce fake sink/source term which creates stability issues (Mathiesen et al., 2013). Beside on their current stage of development, they rely on site specific parameters. Mathiesen et al. (2013) used a mixing ratio optically similar to the marine stratocumulus clouds observed in coastal California.

For these reasons a bibliographic review highlighting the advantage and drawback of each DA method commonly used in atmospheric science is presented in the next section. This bibliographic review was used to select the DA method to improve the solar irradiance forecast in the ITZ.

Tab. IV.1 comparison of advantage and disadvantage of existing DA algorithms used in NWP models. These DA algorithm are detailed Appendix A

DA methods	Nudging	3D-Var	4D-Var	Ensemble filters (ENKF,ENSRF)
Nudging		-Nudging does not need to linearize the model (Asch et al., 2016) - Nudging does not require a to compute B (Asch et al., 2016) ➔ less computational burden -Observations are treated at the correct time (Lorenc and Rawlins, 2005)➔ improve the forecast -can model implicitly the model error (Lakshmivarahan and Lewis, 2013) ➔improve the forecast (Howes et al., 2017)	-Nudging does not need to linearize the model (Asch et al., 2016) - Nudging does not need the development of an adjoint model (Asch et al., 2016) ➔ easier to code - Nudging does not require a covariance matrices B (Asch et al., 2016) ➔ less computational burden -can model implicitly the model error (Lakshmivarahan and Lewis, 2013) ➔improve the forecast (Howes et al., 2017)	- Nudging does not require a to compute B (Asch et al., 2016) ➔ less computational burden - Nudging is more adapted to observing networks with heterogeneous reporting times (Lei and Hacker, 2015)
3D-Var	-3DVar can process a wider range of observations ➔ In nudging only variables in the prognostic equations can be nudged (Asch et al., 2016) -Perfect model assumption ➔wrong analysis when model error is high relative to observation and background errors (Holm, 2008)		-3DVAR does not requires the development of linear and adjoint models(i.e TLM) (Bannister, 2018) ➔computationally cheaper but worst forecast (Lorenc and Rawlins, 2005)	-Perfect model assumption ➔wrong analysis when model error is high relative to observation and background errors (Holm, 2008)

4D-Var	-4DVar requires the development of linear and adjoint models (Bannister, 2018) ➔ more difficult to code and maintain -4DVar can process a wider range of observations (Skamarock et al., 2008)	- 4Dvar treats observations at the correct time (Bannister, 2018) ➔improve analysis -4Dvar evolves B is implicitly (Bannister, 2018; Schwartz et al., 2014) - 4D-Var computationally expensive compared ➔ depending on the computational facilities cannot be used operationally (Lorenc and Rawlins, 2005)		- 4DVAR analysis is consistent with the model (Fowler, 2016; Lorenc, 2003) ➔appropriate B - 4DVAR relies on the validity of: TL and perfect model (Fowler, 2016; Lorenc, 2003) ➔restrict assimilation window -B not evolved explicitly (Fowler, 2016; Schwartz et al., 2014) -4DVAR assumes no model error ➔ give the same credence to older observations at the beginning of the interval as to newer observations at the end of the interval (Kalnay, 2009)
Ensemble filters	-ENKF is more computationally demanding (Kalnay, 2009) - ENKF treat explicitly the model error ➔ improve the forecast (Mitchell et al., 2002) -ENKF better spread information in space and across variables (Fritzner et al., 2018)	-Ensembles are propagated with full nonlinear numerical (Janjic, n.d.) Model➔ no need to linearize -ENKF is more computationally demanding (Kalnay, 2009)	-B is flow dependent (Fowler, 2016; Lorenc, 2003) -No need to linearize the model (Fowler, 2016; Lorenc, 2003). However, due to computational cost we can only afford small ensemble sizes thus offering minimal if any benefit for strongly non-linear problems (Vetra-Carvalho, 2018) -Sensitive to ensemble size (Fowler, 2016) ➔filter divergence -ENKF is more computationally demanding (Kalnay, 2009)	

As shown Tab. IV.1 each DA method as its own set of advantages and drawbacks. Main drawbacks of variational method are that they do not natively take into account the model error and that they do not evolve explicitly B within the assimilation window. ENKF main drawback deal with the tradeoff between ensemble size, computational affordability and accuracy. For these reasons hybrid method were developed such as ETKF-3DVar (Wang et al., 2008), ENKF-3DVAR (Gao et al., 2013), ENSRF-3DVAR (Schwartz et al., 2014), ETKF-4DVar (Clayton et al., 2013), ENKF with nudging (Lei and Hacker, 2015), to benefit from the strengths of each individual DA methods. Most methods that couple several DA methods involve variational and ENKF methods; they are called EnVar (Bannister, 2017). Bannister (2017) classified in his study EnVar DA methods in three categories:

- pure EnVar system that uses a tangent linear model (TLM) similarly to variational methods (e.g En4dvar,En3dvar)
- pure EnVar system that does not use a TLM similarly to ENKF (4DEnvar, 3DEnvar)
-hybrid EnVar system whose B is computed with the contribution of a static part similarly to variational methods and a contribution from a flow dependent part similar to ENKF methods

According to Bannister (2017) hybrid methods tend to be better and more robust for smaller ensembles size than pure variational method, ensemble, or EnVar methods. Bannister (2017) also found that pure EnVar system that uses a TLM outperforms pure EnVar system that does not use a TLM. However, an EnVar system that uses a TLM requires more significant computational resources.

IV. Data assimilation for solar energy forecast purposes in the intertropical zone

The ITZ is an area with highly variable dynamic in which we encounter significant amounts of convective clouds. Consequently, the solar energy available at the ground is highly variable. Recently Schwartz (2016); and Schwartz and Liu (2014) showed that using continuously cycling an hybrid 3DEnVar DA systems improve the characterization of convective phenomena. As clouds convective phenomena have a sizeable effect on accuracy of NWP model GHI forecasts; therefore, hybrid 3DEnVar DA method was selected to improve WRF irradiance forecasts in the ITZ.

IV.1 Description of the Hybrid 3D Ensemble variational

The hybrid 3DEnVar DA algorithm is available in WRF-DA; it aims to minimize a cost function similar to that of Eq.IV.9; However, in the hybrid 3DEnVar the analysis increment ($\delta\vec{x}$) is partitioned into a weighted linear combination of the ensemble and static contributions; it is expressed as follows (Li et al., 2015; Schwartz et al., 2014):

$$\delta\vec{x}(t_0) = \delta\vec{x}_s(t_0) + \frac{1}{\sqrt{N-1}}\sum_{i=1}^{N} a_i\left(\overrightarrow{\delta x^{(i)}}(t_0) - < \overrightarrow{\delta x^{(i)}}(t_0) >\right) \quad \text{IV.13}$$

Where $\delta\vec{x}_s$ is the increment associated with the static BECs, and a_i the ith member's extended control variable (Lorenc, 2003) that give proper weight to the i^{th} member's departure from the ensemble mean. Therefore, the Hybrid 3DEnVar minimization problem is written as follows (Li et al., 2015; Schwartz et al., 2014):

$$J(\delta\vec{x}(t_0)) = \frac{\beta_1}{2}\delta\vec{x}_s(t_0)\,B^{-1}\delta\vec{x}_s(t_0) + \frac{\beta_2}{2}\vec{a}^T A^{-1}\vec{a} + \frac{1}{2}(d_0 - H\delta\vec{x}(t_0))^T R^{-1}(d_0 - H\delta\vec{x}(t_0)), \quad \text{IV.14}$$

where $\vec{a}$ is an array containing the extended control variable for each ensemble member. A is the ensemble covariance localization matrix; it controls the spatial correlation of $\vec{a}$; $\beta_1$and β_2 are user defined parameters that determine how much weight is given to the ensemble and static BECs. $\beta_1$and β_2 are constrained such that:

$$1 = \frac{1}{\beta_1} + \frac{1}{\beta_2}. \quad \text{IV.15}$$

Fig. IV.3 describes the hybrid 3DEnVar algorithm. In this figure we can see that the coupling between the ETKF and the 3DVAR algorithm. First the ensemble perturbation supplies an estimate of the flow dependent forecast error *A* to the 3DVAR analysis. Second the ensemble member analysis is re-centered about the hybrid analysis. The high-resolution hybrid analysis is then used as initial and boundary conditions for a WRF forecast instead of the GFS coarse resolution analysis. In our Studies we performed the hybrid analysis for the three domains of WRF showed previously chapter III section III.

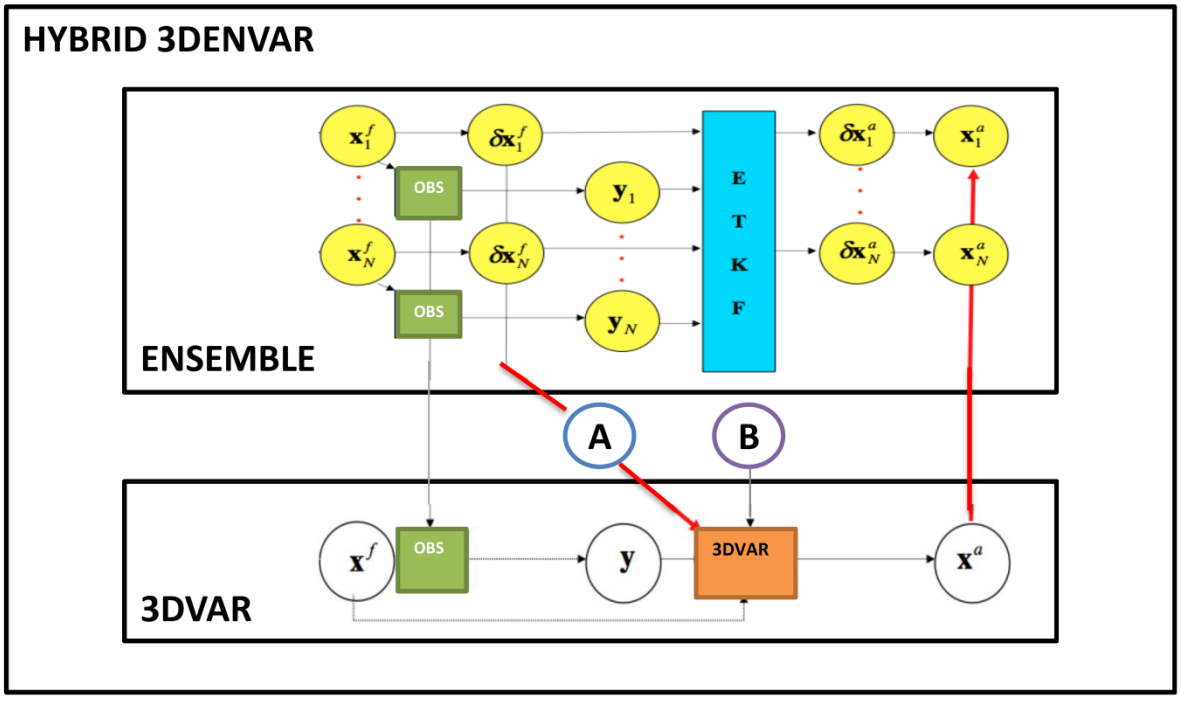

Fig. IV.3 Flowchart describing the hybrid ETKF–3DVAR DA system in WRF model; inspired after (Barker and Clayton, 2011; Kutty et al., 2018)

IV.2 Building the ensemble members

As stated in earlier chapter III section III, we chose the initialization at 00 UTC to allow for a spin up of 12 hours before the sunrise in French Guiana. Consequently, each member should be valid at 00h

UTC and allows for 12 hours spin-up at least. To build the ensemble valid for the day d at 00h UTC; first we perturbed GFS initial and boundary condition valid for the day d-1 at 00h UTC using random-cv utility provided in WRFDA (NCAR, 2017). Second we use this perturbed set of initial and boundary conditions to run WRF for a 24-hour forecast, which is valid for day d at 00h UTC. These forecasts were computed using WRF-M9 setup for the same three way nested domain shown chapter III section III. We repeated this cycle 40 times to obtain a 40 ensemble member. This ensemble size was based on Vetra-Carvalho (2013) finding that ETKF is able predicts the initial cloud growth for ensemble containing at least 30 members.

IV.3 Modeling of the background error covariances A and B

B plays a key role in Var, ENKF, pure and hybrid EnVar DA algorithms. It weights the importance of the background state by spreading information from observation points and imposing balance between analysis variables $\overrightarrow{x_a}(t)$ (Wang et al., 2014). Current NWP models have a typical dimension of 10^8 which gives B matrix with 10^{16} elements (Descombes et al., 2015). This number of elements is too important to be calculated explicitly and stored. Therefore, B needs to be preconditioned such that it becomes numerically manageable (Wang et al., 2014). To precondition B, first a control variable transform (CVT) is applied (Descombes et al., 2015; Wang et al., 2014). Instead of using the state variable $\vec{x}$ of the NWP model, other variables (control variables) that are likely to be mutually uncorrelated are selected. Second B is decomposed in a series of sub-matrices expressed in the control variables space. Each matrix corresponds to an elemental transform that can be individually modeled. These two steps aim to renders B diagonal, they are expressed mathematically as follows (Descombes et al., 2015; Liu, 2017; Wang et al., 2014):

$$\begin{cases} \delta\vec{\mathrm{x}} = \mathrm{U}\vec{\mathrm{u}}, \\ \mathrm{B} = \mathrm{U}\mathrm{U}^{\mathrm{T}} \; and \\ \mathrm{U} = \mathrm{U_p}\mathrm{U_v}\mathrm{U_h}\,. \end{cases} \qquad \text{IV.16}$$

where $\vec{u}$ is the control variable array; U the transformation matrix; U_p, U_v and $U_h$are the physical transform operator, vertical transform matrix, and the horizontal transform matrix respectively. They aim to remove the remaining correlation (e.g, balanced part) between the control variable using linear regressions, define the vertical correlation between control variables using recursive filters and horizontal correlations via empirical orthogonal decomposition of vertical covariance. S is a diagonal matrix composed of the standard deviations of the background error.

According to Dhanya and Chandrasekar (2016); Xin Li et al. (2016); Wang et al. (2014) the formulation and choice of control variables have an impact on the analysis as well as model forecasts. Bannister (2008b) and Gustafsson et al.(2018) listed in their study the control variable used in NWP forecast centers. In WRF V3.8 (NCAR, 2017) four choices of control variable are available (cv3, cv5, cv6, cv7). In cv3 the control variables are in physical space; i.e. the state vector variable are used. It is assumed that there is no spatial and multivariate correlation (Liu, 2017). This option uses stream function, unbalanced velocity potential, unbalanced temperature, mixing ratio and unbalanced logarithm surface pressure. Cv3 provides a generic B that has a global coverage, however, it needs to be tuned to get a an improved analysis (Guo et al., 2008; Radi et al., n.d.). In cv5, cv6, cv7 the control variable are instead in eigen-vector space (NCAR, 2017). cv5 and cv6 use quasi-similar control variables; cv5

uses stream function, unbalanced velocity potential, unbalanced temperature, pseudo-relative humidity and unbalanced surface pressure. Cv6 uses an unbalanced pseudo relative humidity instead of using the pseudo-relative humidity. A major difference between cv5 and cv6 is that cv5 correlation used between the velocity potential and temperature or the velocity potential and surface pressure (Chen et al., 2013). Therefore, neither temperature nor surface pressure observations can directly influence the divergent part of the wind. Similarly the moisture observations will not have any impact on other variables like wind, temperature, and surface pressure. To overcome these limitations, cv6 includes linear relationship between temperature and surface pressure with unbalanced velocity potential as well as the inclusion of moisture correlations with all other analysis variables (Chen et al., 2013). Chen et al. (2013) showed that in tropical regions the inclusion of these extra correlations improved the forecast accuracy. According to (Dhanya and Chandrasekar, 2016) Moisture is a highly variable component in the atmosphere especially in tropical regions where convection is a major driving factor for weather systems, better formulation of moisture fields in the analysis system can improve the analysis and consequently the model predictions. Cv7 uses a different set of control variables than cv5 and cv6; it uses horizontal wind components as momentum control variable temperature, pseudo-relative humidity and surface pressure. For this cv7 the transformation matrix U_p Eq.IV.16 becomes an identity matrix (Sun et al., 2016). Recent studies (Xin Li et al., 2016; Sun et al., 2016; Xie and MacDonald, 2012) have shown that using stream function and velocity potential instead of horizontal wind components tends to produce non-physical wind increments with opposite direction to the observed wind in the neighborhood around the observation point. White et al., 2016 showed in this study that nudging horizontal velocity has the ability to improve GHI forecasts. Therefore, a CVT that describes accurately wind increments is considered because in our study there are a significant number of wind velocity observations.

Consequently, for the static background B, the NMC method (Parrish and Derber, 1992) Eq.IV.17 is performed using the horizontal wind component as momentum control variable temperature, pseudo-relative humidity and surface pressure control variables. For the flow dependent background we use the ensemble method Eq.IV.18.

$$\mathrm{B_{NMC}} = \vec{\mathrm{x}}_{24} - \vec{\mathrm{x}}_{12} \; and \tag{IV.17}$$

$$\mathrm{B_{ENS}} = \frac{1}{\mathrm{N}-1}\sum_{\mathrm{i}=1}^{\mathrm{N}}(\vec{\mathrm{x}}_{\mathrm{i}} - \langle\vec{\mathrm{x}}_{\mathrm{i}}\rangle)(\vec{\mathrm{x}}_{\mathrm{i}} - \langle\vec{\mathrm{x}}_{\mathrm{i}}\rangle)^{\mathrm{T}}. \tag{IV.18}$$

Where $\vec{x}_{24}$ and $\vec{x}_{12}$ are forecasts which are valid at the same time for different initialization times; N and $\vec{x}_i$ are the ensemble size and an ensemble member respectively.

IV.4 Experiments

Due to computational constraints we selected a limited number of experiments for a limited time span to validate 3DEnVar GHI forecasts. In CHAPTER IIIIII we have calibrated and validated WRF GHI forecasts in French for March, May and September. To choose which month to use we set the following constraint. The GHI forecasts obtained using 3DEnVar should not deteriorate significantly the GHI forecast under clear sky conditions and should improve the GHI forecasts under cloudy and overcast sky conditions. To meet this constraint we validated the GHI forecasts 3DEnVar over

September because this month has the highest number of clear sky conditions samples. Fig. IV.4 shows the cdf of K_c values when the GHI measurements of all stations are merged from September 1st to September 29th. In September 5% of K_c value are lower than 0.4; approximately 12% range between 0.4 and 0.65 and 83% of the sample size have K_c values higher than 0.65.

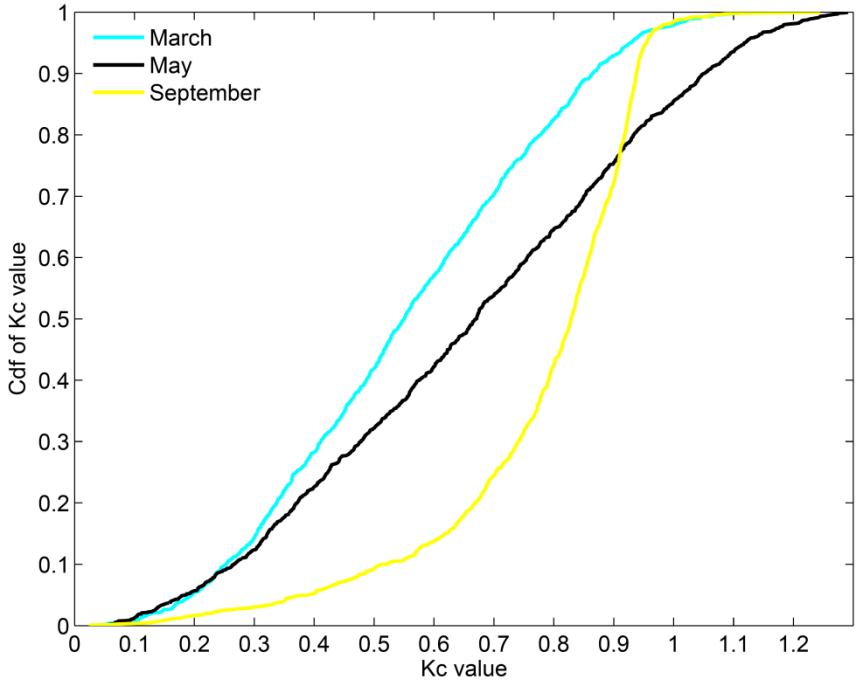

Fig. IV.4 Cdf of Kc values in March, May, September. The GHI forecasts from Rochambeau, Saint Georges, Maripasoula, Saint Laurent, Kourou and Ile Royale

In the 3DEnVar DA algorithm two parameters can be adjusted: the ensemble weight β_2 Eq.IV.14 and the correlation length scale. The correlation length scale was defined to 500 km, the default value. The ensemble weight was defined so that there is first a 50% and second a 75% contribution from the flow dependent background i.e. $\beta_2 = 2$ and $\beta_2 = \frac{4}{3}$ respectively. The experiments performed are described in this study are described as follows:

(i) The control experiment; WRF-M9 described chapter III. In this experiment there is no data assimilation; WRF is initialized using GFS forecasts at 00h UTC and ran for 48h. WRF forecasts are forced every 3h by GFS. Once a 48h forecast is obtained, a new one is launched so that the last 48h is valid for September 29th at 00h UTC.

(ii) 3DEnVar DA with no cycling (WRF-M9-3DEnVar). In this experiment $\beta_2 = 2$ and WRF is initialized at 00h UTC using the analysis produce by 3DEnVar and ran for 48h. From +3h to +45h WRF forecasts are forced using GFS. Once a 48h forecast is obtained, a new one is launched so that the last 48h is valid for September 29th at 00h UTC.

(iii) 3DEnVar DA with a 6h cycling (WRF-M9-3DEnVar-c). In this experiment $\beta_2 = 2$; the hybrid data assimilation is performed once at 00h UTC and WRF forecasts are cycled every 6 hours. WRF is run for 6h and the forecast obtained at t+6h is used a background for another 6h forecast. This forecast cycle is repeated so at the end we obtain a 48h

forecasts. Once a 48h forecast is obtained, a new one is launched so that the last 48h is valid for September 29th at 00h UTC.

(iv) 3DEnVar DA with a 6h cycling (WRF-M9-3DEnVar-c2). In this experiment $\beta_2 = \frac{4}{3}$.

The observations we assimilated were conventional observation downloaded online from NCEP servers (NCEP, 2018), they are Global Upper Air and Surface Weather Observations in PREPBUFR format.

IV.4.a Studying the impact of the hybrid DA on the GHI forecasts

Tab. IV.2 Comparison between WRF-M9 and WRF-M9-3DEnVar for all experiments. All six stations showed chapter III section III were merged

Ensemble background contribution	MAE (W/m²)				MBE (W/m²)			
	Clear	Cloudy	Overcast	All	Clear	Cloudy	Overcast	All
WRF-M9 (i)	79	134	236	102	17	128	233	55
WRF-M9-3DEnVar (ii)	81	132	227	103	6	121	224	45
WRF-M9-3DEnVar-c (iii)	79	130	221	100	4	121	217	43
WRF-M9-3DEnVar-c2 (iv)	83	137	231	106	2	122	231	43

Tab. IV.2 shows that under overcast sky conditions each experiment (ii, iii, iv) involving the use of the hybrid DA improves the GHI forecasts. The greatest improvement of the GHI forecasts under cloudy and overcast sky conditions is obtained using WRF-M9-3DEnVar-c; the relative improvement of the MAE and MBE of WRF-M9-3DEnVar-c compared to WRF-M9 are 6% and 7% respectively. We can also see Tab. IV.2 that cycling WRF forecasts as initial and boundary conditions improve the accuracy. Nevertheless, in September, increasing β_2 outweigh this gain; under cloudy sky conditions, the MAE of WRF-M9-3DEnVar-c2 GHI forecasts are worse than those of WRF-M9.

Under all sky conditions the lowest MAE is obtained using WRF-M9-3DEnVar-c (Tab. IV.2). The improvement of WRF-M9-3DEnVar-c MAE and MBE over the MAE and MBE of WRF-M9 are 2% and 22% respectively. Under clear sky conditions WRF-M9-3DEnVar-c and WRF-M9 have similar MAE values (Tab. IV.2); consequently, we explain this improvement by the better GHI forecasts under cloudy and overcast sky conditions. The magnitude of the improvement under all sky conditions (e.g. 2%) is explained by the fewer occurrence of cloudy and overcast conditions in September; 5% and

12% respectively. Nevertheless, the greater results of WRF-M9-3DEnVar-c compared to WRF-M9 meet the constraint initially set: not deteriorate significantly the GHI forecast under clear sky conditions and should improve the GHI forecasts under cloudy and overcast sky conditions. Consequently these results show that WRF-M9-3DEnVar-c is able to improve the forecast of cloudy and overcast sky conditions during September a rainy season month.

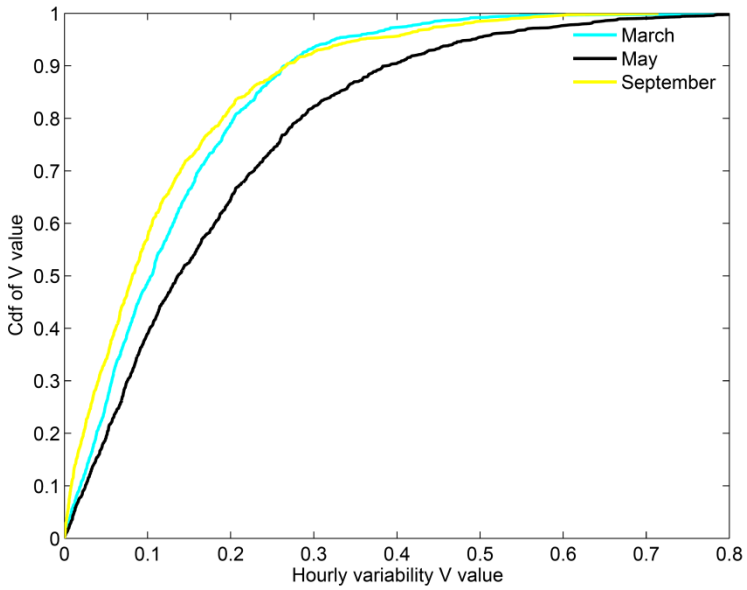

Fig. IV.5 Cdf of V values in March, May, September. The GHI forecasts from Rochambeau, Saint Georges, Maripasoula, Saint Laurent, Kourou and Ile Royale

Tab. IV.2 also shows that the hybrid DA impact the GHI forecasts under clear sky conditions. The GHI forecasts increase by as much as 4 W/m² (+5%) whereas the MBE decreases as much as 15 W/m² (-88%).

IV.4.b Comparison of WRF-M9-3DEnVar-c against WRF-M9 and IFS forecast in September

Perez et al. (2013) showed in this study that IFS is the most accurate global model in extra-tropical climate. This model even outperformed GHI forecasts from optimized mesoscale model including WRF. We have shown in Chapter III section that similarly to extra-tropical climate IFS is more accurate than GFS in tropical climate. Consequently, in this section we compare our calibrated WRF model with data assimilation (e.g., WRF-M9-3DEnVar-c) to IFS forecasts in French Guiana. According to Lorenz et al. (2016) spatial average improve the forecast metrics. Consequently, we averaged WRF GHI forecasts over a 12² km area and compared it against IFS hourly forecast available over a 14² km grid.

We have shown in chapter III sections II and III that IFS and WRF accuracy varied with K_c and V values. Consequently, we represented Fig. IV.6 WRF-M9, WRF-M9-3DEnVar-c and IFS MAE as a function of K_c and V. We can see Fig. IV.6 and Fig. IV.7 that averaging spatially the GHI forecasts of WRF-M9 and WRF-M9-3DEnVar-c increase the MAE improvement of WRF-M9-3DEnVar-c over WRF-M9 obtained previously for a single grid cell. Under all sky conditions and for a single grid cell the MAE improvement of WRF-M9-3DEnVar-c over WRF-M9 was 2%. For an average over a 12² km area the

MAE improvement of WRF-M9-3DEnVar-c over WRF-M9 is 6%. Under cloudy and overcast sky conditions the improvement cause by the spatial average is 9% and 7% respectively.

When comparing the GHI forecasts of WRF-M9-3DEnVar-c to those of IFS we find that under clear sky conditions WRF-M9-3DEnVar-c outperforms the GHI forecasts of IFS-1h. The relative improvement of WRF-M9-3DEnVar-c MAE over IFS MAE is 10% (Fig. IV.7). However, under cloudy and overcast sky conditions it is the opposite; the relative worsening of WRF-M9-3DEnVar-c MAE over IFS MAE is -37% and -45% respectively. Nevertheless, due to the higher occurrence of clear sky conditions during September the accuracies of WRF-M9-3DEnVar-c and IFS are close. Under all the relative worsening of WRF-M9-3DEnVar-c MAE over the MAE of IFS is -5%. Tab. IV.4 shows a comparison of IFS and post-processed WRF for various types of extra-tropical climate. For arid and humid continental climates the GHI forecasts of WRF are worse than the GHI forecasts of IFS; the relative worsening of WRF over IFS is -25% and -30 %respectively. In Mediterranean WRF with MOS has a similar accuracy than IFS. These results showed that the hybrid data we used performed successfully comparatively to MOS methods.

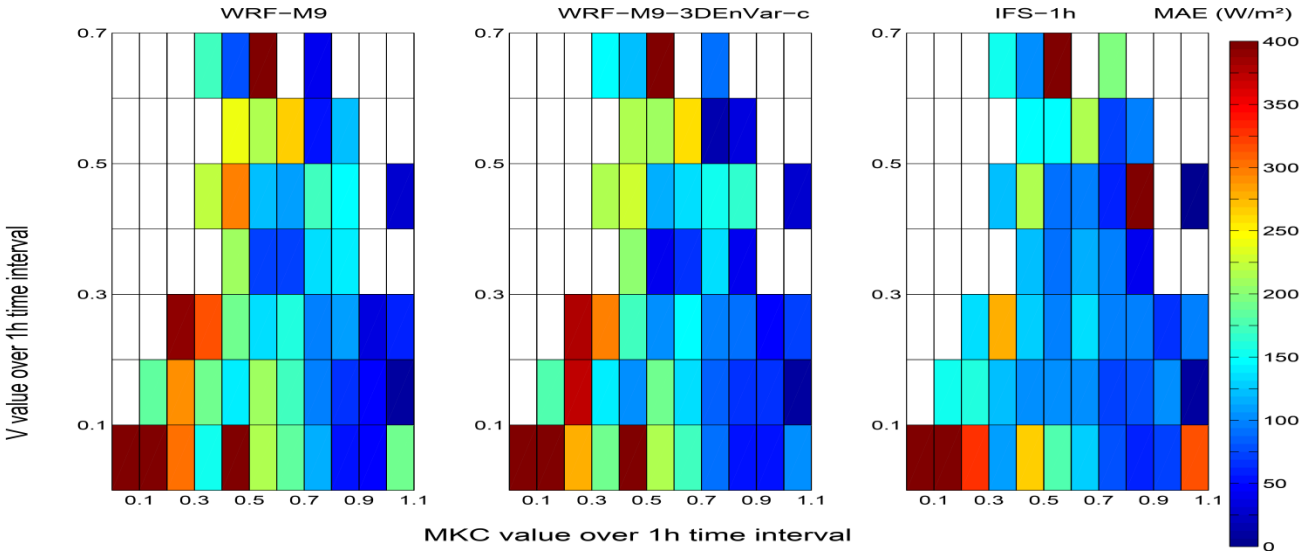

Fig. IV.6 WRF-M9, WRF-M9-3DEnVar, and IFS MAE as a function of K_c and V

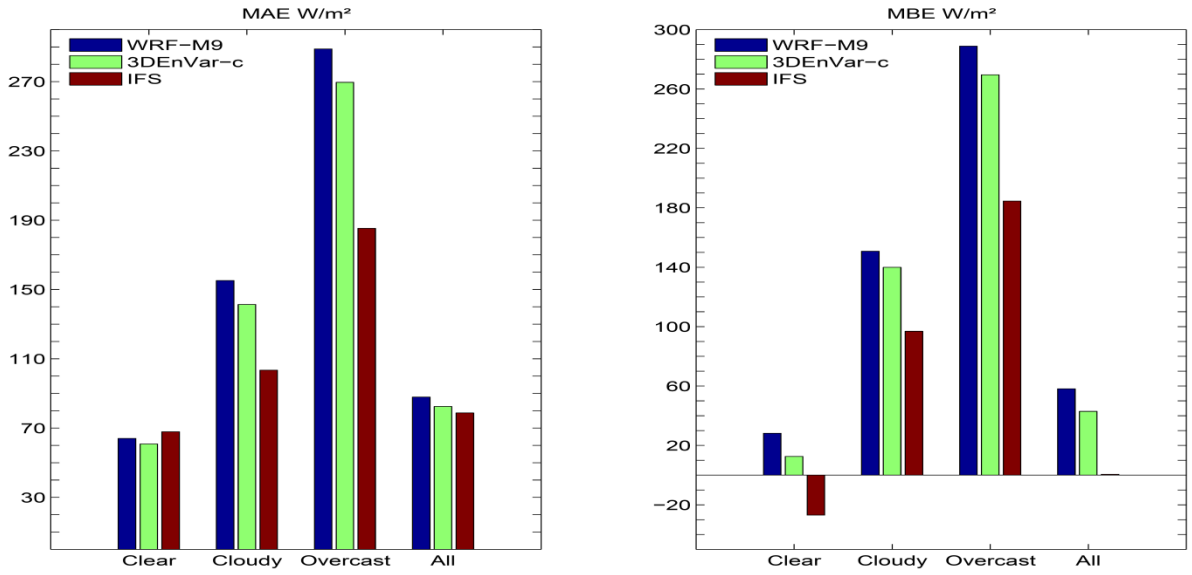

Fig. IV.7 MAE (W/m²) and MBE (W/m²) of WRF-M9, WRF-M9-3DEnVar-c and IFS

Tab. IV.3 Relative improvement of WRF-M9-3DEnVar-c over WRF-M9 and IFS

	Relative improvement (%) of the MAE of WRF-M9-3DEnVar-c (W/m²) over the MAE of WRF-M9 (W/m²)				Relative improvement (%) of the MAE of WRF-M9-3DEnVar-c (W/m²) over the MAE of IFS (W/m²)			
	Clear	Cloudy	Overcast	All	Clear	Cloudy	Overcast	All
September	5%	9%	7%	6%	10%	-37%	-45%	-5%

Tab. IV.4 Relative improvement of the MAE WRF with MOS over IFS. These relative improvement were computed using MAE values for a 48h forecast initialized at 00h UTC that were provided by Perez et al. (2013). Mediterranenan stations include Cordoba, Huelva,Granada, Carpentras; Desert rock is the Arid station considered; For Humid continental we considered Goodwin Creek and Penn state

	Mediterranean	Arid	Humid-Continental
Mean IFS MAE (W/m²)	60	56	78
Mean WRF (W/m²)	60	70	103
Relative improvement (%) of the MAE of WRF-MOS over the MAE of IFS (W/m²)	0%	-25%	-32%

IV.5 Conclusions

In this section we have suggested a method to improve the GHI forecasts of mesoscale models in French Guiana and in the ITZ. This method is based on an hybrid ensemble variational ensemble data assimilation: 3DEnVar; it was validated in French Guiana. Due to computational constraints we selected a limited number of experiments and over a limited time span to validate 3DEnVar GHI forecasts. The validation period was chosen so that it guarantees the ability of the method to not deteriorate significantly the GHI forecast under clear sky conditions and to improve the GHI forecasts under cloudy and overcast sky conditions. We found that : (i) our method meet the constraint initially set: In September during the dry season and under clear sky conditions WRF-M9 and WRF-M9-3DEnVar have similar MAE values; however, under cloudy sky and overcast conditions the MAE of WRF-M9 is improved by 6% and 7% respectively. (ii) the balance between the flow dependent background and static background highly influence the forecast accuracy. Due to lack of computational resources we did not optimized this balance. Nevertheless, the fact that the forecasts under cloudy and overcasts sky conditions are improved whereas the GHI forecasts under clear sky are not deteriorated shows that this method is promising. (iii) Cycling WRF GHI forecasts improve the GHI forecasts. The GHI forecasts of WRF-M9-3DEnVar were also compared to the GHI of IFS. We found that under clear sky conditions WRF-M9-3DEnVar-c outperforms the GHI forecasts of IFS-1h. The relative improvement of WRF-M9-3DEnVar-c MAE over IFS MAE is 10%. However, under cloudy and overcast sky conditions it is the opposite; the relative worsening of WRF-M9-3DEnVar-c MAE over IFS MAE is -37% and -45% respectively. Nevertheless, due to the higher occurrence of clear sky conditions during September the accuracies of WRF-M9-3DEnVar-c and IFS are close; the relative improvement of WRF-M9-3DEnVar-c MAE over the MAE of WRF-M9 is -5%. Our future research will involve optimizing the balance between the flow dependent background and static background and validating the method in the rainy season.

CHAPTER V : Thesis summary and conclusions

The contribution of this thesis to the scientific community was: to use existing and well-known methods (i.e Heliosat-II (H-II) and weather and research forecast (WRF)) and adapt them so that they produce improved Global Horizontal Irradiance (GHI) estimates in French Guiana. Weather conditions in French Guiana are driven by the physical phenomena in the Intertropical zone (ITZ); consequently, methods developed in this manuscript for French Guiana aimed to be extended to all the ITZ.

Our first contribution to the scientific community was to modify the H-II method so it can account for cloud absorption. The rationale behind this choice came from a literature review of the physical properties of the ITZ. The ITZ is an area with highly variable cloud cover dynamic. In this area we encounter significant amounts of convective clouds. Consequently, to improve the GHI forecasts in the ITZ the effect of clouds must be modeled accurately. Native H-II method accounts for cloud attenuation empirically by using a clear sky index (K_c) that is a function of a cloud index derived from satellite images. A drawback of using K_c is that it does not model cloud properties such as cloud optical depth, which influences cloud reflectance, absorbance and transmittance. Therefore, to consider the significant presence of clouds we integrated a spatiotemporally dependent cloudy sky RTP into the GHI estimate formula of native H-II. This cloudy sky model integrates the cloud albedo and the cloud absorption which is a proxy for the cloud optical depth. This study showed that the GHI forecasts of the H-II method can be improved by accounting for the cloud absorption. Our method does not increase computation time and only require one parameter that must be computed beforehand using cloud absorption abacus available in the literature. The scope for future research involves validating this methodology to other locations; for this purpose a monthly map of cloud absorption must be computed, first, for French Guiana, and then, for other tropical areas.

Our second contribution to the scientific community was to fill the gap on the knowledge about the forecasts accuracy of GHI products from global NWP models in the ITZ. We compared IFS and GFS against GHI measurements from three territories located in the ITZ and with tropical climates. That study provided benchmark tools to assess the forecast accuracy of a limited area NWP model GHI forecast and for assessing the quality of post-processing methods. It also identified the source of inaccuracies of the GHI forecasts from NWP in the ITZ. This study showed the inability of model forecast realistic clouds cover; this inability impact cloudy, overcast and clear sky conditions. Under sky conditions reported as clear by the ground measurements, IFS and GFS forecasts cloudy and overcast sky conditions a significant amount of time. We compared results in tropical climates to results in extra tropical climates (Mediterranean, oceanic, continental, and arid climates) and found by comparison with the persistence that is more difficult to forecasts the GHI in tropical climates. The scope for future research involves studying for other locations and other climate in the ITZ. The scope for future research involves validating this methodology to other locations in the ITZ.

Our third contribution to the scientific community was to suggest based on the analysis of IFS and GFS a generic methodology to WRF mesoscale model in the ITZ. The aim of this methodology was to restrain and select the minimum number of simulations to run, to obtain improved GHI forecasts in the ITZ compared to a non-calibrated model. Consequently, we built 34 Ensemble Members (EM) by

varying the microphysics, land surface model, and planetary boundary layer parameterization schemes. These parameterizations were varied because they have a strong influence on the distribution and physical properties of the simulated cloud field. This methodology was validated against GHI forecasts from AROME mesoscale model and ground measurements in French Guiana. This study showed that the calibration process of a mesoscale model must be carefully considered and that using physical parameterization previously selected in similar climate is not suitable.

Our last contribution to the scientific community was to use an hybrid ensemble variational data assimilation method to improve the GHI forecasts of WRF in the ITZ. The rationale behind this choice was to find a method that does not require an extensive training dataset; that is applicable for the whole computation domain and that does not require that the setup of the NWP model to be frozen. Ensemble variational data assimilation methods have proved their abilities to improve the characterization of severe convection events but have not been used yet for solar energy purposes. This method used the greatest ensemble member previously computed on the calibration process of WRF first to compute the static background and to build the initial ensemble needed for the assimilation method. Results found showed that this method improves the GHI forecast under cloudy and overcast sky conditions without lowering significantly the accuracy under clear sky conditions. The accuracy of the GHI forecasts depend highly on the balance between the flow dependent and static background which needs to be optimized. Nevertheless, this method is computationally demanding: it requires computing the background and the initial ensemble. The scope for future research is to find the adequate balance between the static and the flow dependent background covariance and to modify our method so it can be used operationally. In order the computational time the following solution may be considered: (i) instead of computing the initial ensemble, we may use ensemble forecasts available from forecasts center at coarser resolution and or (ii) use and optimize background covariance also provided globally for coarser resolutions. Using these two suggestions may one hand reduce the computing time but on the other reduce the accuracy of the GHI forecasts.

Conclusions et apport de la thése

La contribution de ce travail de thèse à la communauté scientifique a été : d'utiliser des outils connus et réputés et de les adapter pour qu'ils fournissent des estimations (avec H-II) et des prévisions (avec WRF) améliorées en Guyane. Les conditions météorologiques en Guyane sont influencées par les phénomènes physiques se déroulant dans la zone intertropicale; en conséquence, les méthodes appliquées en Guyane française ont pour vocation à être étendues à toute la zone intertropicale.

Notre première contribution à la communauté scientifique a été de modifier H-II afin qu'il puisse tenir compte de l'absorption des nuages. Ce choix a été motivé par une étude des propriétés physiques de l'ITZ. Cette zone est caractérisée par une dynamique nuageuse importante; en conséquence les nuages et leurs propriétés physiques doivent être modélisés pour obtenir de meilleurs estimés de GHI. La méthode H-II originelle modèle l'atténuation des nuages en utilisant une relation empirique fonction d'un index nuageux déduit à partir d'images satellites. Un inconvénient de cette méthode est que les propriétés physiques des nuages tels leurs épaisseurs optiques ne sont pas modélisées. L'épaisseur optique influence à la fois l'absorption, la réflexion et la transmission des nuages. En conséquence, afin de tenir compte de la présence significative de nuage, nous intégrons un modèle de transfert radiatif en ciel couvert dans la formulation de GHI de la méthode H-II originelle. Ce modèle de ciel couvert intègre l'albédo des nuages qui est un proxy de l'épaisseur optique. Cette étude a montré que la prise en compte de l'absorption des nuages permet une amélioration des estimés de GHI. Notre méthode n'augmente pas de manière significative le temps de calcul et ne nécessite qu'un seul paramètre supplémentaire ; celui-ci est calculé à priori à partir d'abaque d'absorption des nuages disponible dans la littérature. Les recherches futures consisteraient à valider cette méthode dans d'autre pays de la zone intertropicale.

Notre deuxième contribution à la communauté scientifique d'étudier les performances des modèles IFS et GFS en zone intertropicale. Nous avons comparé les prévisions d'IFS et de GFS avec des mesures au sol provenant de trois territoires dans l'ITZ de climat tropical. Cette étude a permis des outils comparatifs permettant de valider la calibration du modèle WRF et la validation de méthode de post-processing. De plus ils ont permis d'identifier les sources d'erreurs des prévisions de GHI en zone intertropicale. Cette étude a montré que les modèles IFS et GFS étaient incapables de simuler une couverture nuageuse réaliste ; cette inhabilité se répercute aussi bien en ciel clair que couvert. Nous avons comparé les résultats obtenus en climat tropical aux résultats obtenus en climat extratropical (Méditerranéen, océanique, continental, et aride); ceci nous a permis de montrer par comparaison avec le modèle de persistance que la prévision est climat extratropical est plus aisé. Les recherches futures consisteraient à valider cette méthode dans d'autre pays de la zone intertropicale.

Notre troisième contribution à la communauté scientifique a été de suggéré une méthode générique permettant de calibrer le modèle mésoéchelle WRF dans l'ITZ. La méthodologie développée pour calibrer WRF se base sur l'identification des sources d'erreurs des prévisions de GHI en zone intertropicale. L'objectif de cette étude a été de choisir restreindre et choisir le nombre minimum de simulations à effectuer afin que le modèle WRF produise des prévisions de GHI de meilleures qualités qu'un modèle numérique non calibré. En conséquence nous avons construit un ensemble de simulation comprenant 34 membres. Cet ensemble fut construit en faisant varier le schéma de microphysique, le schéma de surface et le schéma de couche limite. Ces éléments ont été variés car une étude de la littérature a montré qu'ils avaient le plus d'influence sur les propriétés physiques de la couche nuageuse simulée. Cette méthode a été validée avec des mesures au sol ainsi que des prévisions de rayonnement du modèle AROME. Cette méthode a montré l'intérêt du processus de calibration avant l'utilisation d'un modèle de prévisions météorologiques mésoéchelle.

Notre dernière contribution à la communauté scientifique a porté sur l'utilisation d'une méthode hybride ensembliste variationnelle afin d'améliorer les prévisions de GHI de WRF dans l'ITZ. Le choix de cette méthode a été basé sur la recherche d'une méthode : (1) ne nécessitant pas de base de données (2) applicable sur tout le domaine de simulation et (3) ne nécessitant pas que les réglages physiques du modèle soient inchangés. Notre étude de la littérature a montré que l'habilité des méthodes ensemblistes d'améliorer la caractérisation des phénomènes convectifs ; néanmoins, à notre connaissance ces méthodes n'ont pas été utilisées pour améliorer les prévisions de GHI. L'utilisation d'une méthode hybride (ETKF-3DVAR) a montré une amélioration des estimés de GHI en ciel couvert sans pour autant dégrader de manière significative les estimés de GHI en ciel clair. Nous avons montré que les estimés de GHI sont sensibles à l'équilibre entre la partie statique (3DVAR) et la partie dynamique (ETKF). Néanmoins cette méthode nécessite une puissance de calcul significative pour définir les matrices de « Background » ainsi que les ensembles. Les recherches futures consisteraient à (i) optimiser l'équilibre entre la partie statique et la partie dynamique et (ii) rendre opérationnelle cette méthode en réduisant le temps de calcul. Pour cela il peut être envisagé : (1) d'utiliser des ensembles fournis pas des centres de prévisions météorologiques et de quantifier l'utilisation de ceux-ci sur la prévision des estimés de GHI. (2) Utiliser les matrices de Background à l'échelle planétaire et les optimiser pour notre échelle régionale.

Appendix A

We present in this section how the stochastic and variational DA algorithms that are commonly used in atmospheric science can be derived from one common source Eq.IV.9. Each DA methods has two step; when observations $y(t_n)$ are available, they are used in the update step to compute the analysis state vector $\overrightarrow{x_a}$ and its error covariance matrix P then propagated into time during the prediction step.

I. Variational method

I.1 3D-Var

In practice, the 3D-Var method attempts to find an approximate solution of the minimum of Eq.IV.9 under the following hypothesis(Holm, 2008):

-no model error $Q = 0$
-all observations in the time window $[t_0 - h: t_0 + h]$ are treated as if they were at t_0, the analysis time which is close to their average time.
-linearity of $\overrightarrow{\mathrm{M}}$ and $\overrightarrow{\mathcal{H}}$
-unbiased errors $< \overrightarrow{\varepsilon_b} >=< \overrightarrow{\varepsilon_0} >=< \overrightarrow{\varepsilon_m} >= 0$
-errors are uncorrelated $< \overrightarrow{\varepsilon_b}\,\overrightarrow{\varepsilon_0} >=< \overrightarrow{\varepsilon_0}\,\overrightarrow{\varepsilon_m} >=< \overrightarrow{\varepsilon_m}\,\overrightarrow{\varepsilon_0} >= 0$

Under the above assumptions Eq.IV.9 yields (Rihan et al., 2005):

$$\mathrm{J}\left(\vec{\mathrm{x}}(0)\right) = \frac{1}{2}\left[\overrightarrow{\varepsilon_{\mathrm{b}}}^{\mathrm{T}}\,\mathrm{B}^{-1}\overrightarrow{\varepsilon_{\mathrm{b}}} + \overrightarrow{\varepsilon_{0,\mathrm{n}}}^{\mathrm{T}}\mathrm{R}^{-1}\overrightarrow{\varepsilon_{0,\mathrm{n}}}\right] = \mathrm{J_b} + \mathrm{J}_0. \qquad 0.1$$

B is stationary, it is computed beforehand using the one of the algorithm described Eq.IV.17 (Bouttier and Courtier, 1999). As one hypothesis of the 3D-VAR is linearity, the forward model $\vec{\mathrm{M}}$ should be linearized to evolve the uncertainty in the state from one observation time to the next.

In operational NWP centers, to reduce computational costs, a sequence of linear approximations to the nonlinear minimization problem Eq.0.1 using successive small increments $\delta\vec{\mathrm{x}}$. Given an estimated analysis $\vec{x}(t_0)$; the nonlinear cost function J is then linearized about the corresponding model trajectory $\vec{x}(t_i), i = 1..n$ satisfying the nonlinear forecast model $\overrightarrow{\mathrm{M}}$. This incremental variational3D-Var DA scheme is written as follows (Haben et al., 2009; Lawless et al., 2005):

$$\begin{cases} \tilde{\mathrm{J}}\left(\delta\vec{\mathrm{x}}(\mathrm{t}_0)\right) = \frac{1}{2}\left[\delta\vec{\mathrm{x}}(\mathrm{t}_0) - \left(\vec{\mathrm{x}}_{\mathrm{b}}(\mathrm{t}_0) - \vec{\mathrm{x}}(\mathrm{t}_0)\right)\right]^{\mathrm{T}}\mathrm{B}^{-1}\left[\delta\vec{\mathrm{x}}(\mathrm{t}_0) - \left(\vec{\mathrm{x}}_{\mathrm{b}}(\mathrm{t}_0) - \vec{\mathrm{x}}(\mathrm{t}_0)\right)\right] + \\ \qquad + \frac{1}{2}\left[\tilde{\mathrm{H}}(\mathrm{t}_0)\delta\vec{\mathrm{x}}(\mathrm{t}_0) - {\mathrm{d}_0}^{\mathrm{T}}\right]^{\mathrm{T}}\mathrm{R}^{-1}\left[\tilde{\mathrm{H}}(\mathrm{t}_0)\delta\vec{\mathrm{x}}(\mathrm{t}_0) - {\mathrm{d}_0}^{\mathrm{T}}\right] \textit{ and} \\ \qquad\qquad \text{with } \mathrm{d}_0 = \vec{\mathrm{y}}(\mathrm{t}_0) - \vec{\mathrm{H}}\left(\vec{\mathrm{x}}(\mathrm{t}_0)\right). \end{cases} \qquad 0.2$$

Where $\delta\vec{x}(0) \in \Re^n$ is the analysis increment; $\widetilde{H}(t_0) \in \Re^{n\to p}$ the linearization of the observation operator at the analysis time t_0 and d_0 the innovation for the analysis att_0. Therefore the Linearize model $\widetilde{M}$ prediction step follows Eq.0.3:

$$\overrightarrow{\delta \mathrm{x}}(\mathrm{t}+1) = \widetilde{\mathrm{M}}(\mathrm{t}+1,\mathrm{t})\overrightarrow{\delta \mathrm{x}_\mathrm{a}}(\mathrm{t}), \qquad 0.3$$

Where $\overrightarrow{\delta x} \in \Re^n$ is the forecast model increment and $\widetilde{M} \in \Re^{n.n}$ the linearized model between t+1 and t.

I.2 4D-Var

4D-Var is an extension of the 3D-Var which allows for observations distributed within a time interval $[t_1 : t_N]$. It attempts to find an approximate solution of the minimum of Eq.IV.9, under the following approximations (Holm, 2008):

-no model error
-linearity of $\overrightarrow{\mathrm{M}}$ and $\overrightarrow{\mathcal{H}}$
-unbiased errors
-errors are uncorrelated

Under the above assumptions Eq.IV.9 becomes:

$$\mathrm{J} = \frac{1}{2}\left[\overrightarrow{\varepsilon_\mathrm{b}}^{\mathrm{T}}\,\mathrm{B}^{-1}\overrightarrow{\varepsilon_\mathrm{b}} + \sum_{\mathrm{n}=1}^{\mathrm{N}} \overrightarrow{\varepsilon_{0,\mathrm{n}}}^{\mathrm{T}}\,\mathrm{R}^{-1}\,\overrightarrow{\varepsilon_{0,\mathrm{n}}}\right] = \mathrm{J_b} + \mathrm{J_0}. \qquad 0.4$$

4D-Var contrary to 3D-Var does not assimilate observation at a particular point within the assimilation window (Rihan et al., 2005), instead the assimilation window is divided in several sub interval (Laroche et al., 2007) to assimilate observation at their correct time. Therefore, 4D-Var has the ability to implicitly evolve the initial B over the length of the assimilation window with the tangent linear dynamics:

$$\mathrm{B(t)} \approx \mathrm{MB(t=t_0)M^T}, \qquad 0.5$$

where $B(t = t_0)$ is the background at the beginning of the assimilation windows and M the linearized forecast model value between the beginning and the end of the assimilation window. However, B does not propagate error information from one assimilation cycle to the next; it revert to the stationary value at the beginning of each assimilation window (Miyoshi and Kadowaki, 2008; Raynaud et al., 2011; Schwartz et al., 2014; THEPAUT et al., 1996). Only the information about the state $\vec{x}$ is propagated from one cycle to the next following. Similarly to 3D-Var is exploited in operational NWP centers in the incremental form; it is expressed as follows (Haben et al., 2009; Lawless et al., 2005):

$$\begin{cases} \tilde{J}\left(\delta\vec{x}(t_0)\right) = \frac{1}{2}\left[\delta\vec{x}(t_0) - \left(\vec{x}_b(t_0) - \vec{x}(t_0)\right)\right]^T B^{-1}\left[\delta\vec{x}(t_0) - \left(\vec{x}_b(t_0) - \vec{x}(t_0)\right)\right] + \\ + \frac{1}{2}\left[\tilde{H}\delta\vec{x}(t_0) - \vec{d}\right]^T R^{-1}\left[\tilde{H}\delta\vec{x}(t_0) - \vec{d}\right], \\ \text{with } \tilde{H} = [\tilde{H}(t_0), \tilde{H}(t_1)\tilde{M}(t_1, t_0)^T, \dots, \tilde{H}(t_n)\tilde{M}(t_n, t_0)^T]^T \textit{ and} \\ \vec{d} = \left[{d_0}^T, {d_1}^T, \dots, {d_N}^T\right]. \end{cases} \quad 0.6$$

The information about the state $\vec{x}$ is propagated from one assimilation cycle similar to 3-DVar Eq. 0.3.

II. Stochastic method

II.1 Kalman filter

In the Kalman filter (KF) DA, the analysis equation is solved directly by inversion.

$$\nabla J(\vec{x} \to \overrightarrow{x_a}) = 0, \quad 0.7$$

The KF DA assumes the following:

-the analysis is performed at each time step of the model, it uses only the observations available during that time step.
-linearity of $\overrightarrow{M}$ and $\overrightarrow{\mathcal{H}}$
-unbiased errors
-errors are uncorrelated

Under the assumptions listed above (Holm, 2008), Eq.0.7 yields:

$$\overrightarrow{x_a}(t) = \overrightarrow{x_b}(t) + K(t)\left(y(t) - H\left(\vec{x}_b(t)\right)\right) \text{ with} \quad 0.8$$

$$K(t) = BH^T\left(HBH^T + R\right)^{-1} \text{ and} \quad 0.9$$

$$P(t) = (I - KH)B. \quad 0.10$$

Where *K* is the Kalman gain. Contrary to 3DVar and 4DVar *B* is propagated from one cycle to the next alongside the information about the state $\vec{x}$. Therefore, the prediction step follows the following equations:

$$\vec{x}(t+1) = M\overrightarrow{x_a}(t) \textit{ and} \quad 0.11$$

$$B(t+1) = MP(t)M^T + Q. \quad 0.12$$

However, KF is very expensive to evaluate directly for atmospheric science due to size of *K* e.g $O(10^7)$ (Ghil and Malanotte-Rizzoli, 1991). Therefore, several approximations were made to achieve a computationally efficient DA algorithm. One simplification of KF is the ensemble Kalman filter (ENKF) is presented in the following section.

II.2 stochastic and deterministic ensemble Kalman filter

ENKF is an approximate version of the KF where the distribution of possible states is represented by an ensemble. This ensemble is propagated forward through time and updated when new data become available (Katzfuss et al., 2016). The ensemble representation is a form of dimension reduction, instead of specifying *B*, the EnKF uses an ensemble of possible forecasts that contains valuable flow-dependent information about the background-error statistics (Bannister, 2017). This dimension reduction leads to computational feasibility even for very high-dimensional systems. ENKF DA algorithms are either classified as stochastics (SENKF) or deterministic (DENKF). They differ on the methodology used to get the updated ensemble; they share the same prediction step but have different update state. SENKF directly use the Kalman gain together with random perturbations whereas DENKF uses a non-random transformation on the forecast ensemble (Lei et al., 2010).

Stochastic ENKF methods were first introduced by (Evensen, 1994), since then many variations on the original algorithm have been developed (Evensen, 2003; Houtekamer and Zhang, 2016; Lei et al., 2010). Nevertheless, the update step from the SENKF framework can be expressed generally as follows:

$$\overrightarrow{\mathrm{x_a^{(i)}}} = \overrightarrow{\mathrm{x_b^{(i)}}} + \mathrm{K}\left(\tilde{\mathrm{y}} + \vec{\mathrm{H}}\left(\overrightarrow{\mathrm{x_b^{(i)}}}\right)\right) \; and \tag{0.13}$$

$$\mathrm{K} = \mathrm{BH^T}\left(\mathrm{HBH^T} + \mathrm{R_e}\right)^{-1}. \tag{0.14}$$

Where $\overrightarrow{x_a^{(i)}}$ and $\overrightarrow{x_b^{(i)}}$ are the analysis and state vector of the ensemble member *i*. $\tilde{y}$ is the perturbed observations and R_e the observation error covariance matrix consistent with$\tilde{y}$. Eq.0.13-0.14 differ from Eq. 0.8-0.9 because of the addition of a perturbation to the observation. Burgers (1998) showed that the observations must be treated as random variables at the analysis steps in order for *P*, the variance of the ensemble to be consistent with Eq.0.10 after the update step. Burgers (1998) found that for the SENKF that P is expressed as follows:

$$\mathrm{P(t)} = (\mathrm{I} - \mathrm{KH})\mathrm{B} + \mathcal{O}\left(\mathrm{N}^{-\frac{1}{2}}\right), \tag{0.15}$$

where N is the ensemble size and $\mathcal{O}(N^{-1/2})$ a negligible term that decreases as the ensemble size increase. Therefore, $\tilde{y}$ and R_e were defined as follows:

$$\begin{cases}\tilde{\mathrm{y}} = \vec{\mathrm{y}} + \vec{\varepsilon}\ \ and \\ \mathrm{R_e} = < \vec{\varepsilon}\vec{\varepsilon}^{\mathrm{T}} >. \end{cases} \quad 0.16$$

Where $\vec{y}$ and $\vec{\varepsilon}$ are the observation as defined in the KF and ε the observation perturbation. R_e assumes that errors are unbiased. In SENKF *K* is not computed directly as in KF but using ensemble-based approximations (Houtekamer and Zhang, 2016):

$$\begin{cases} \mathrm{BH}^{\mathrm{T}} = \dfrac{1}{\mathrm{N}-1}\sum_{\mathrm{i}=1}^{\mathrm{N}}\left(\overrightarrow{\mathrm{x_b^{(i)}}} - < \overrightarrow{\mathrm{x_b^{(i)}}} >\right)\left(\mathrm{H}\overrightarrow{\mathrm{x_b^{(i)}}} - < H\overrightarrow{\mathrm{x_b^{(i)}}} >\right)^{\mathrm{T}}\ and \\ \mathrm{HBH}^{\mathrm{T}} = \dfrac{1}{\mathrm{N}-1}\sum_{\mathrm{i}=1}^{\mathrm{N}}\left(\mathrm{H}\overrightarrow{\mathrm{x_b^{(i)}}} - < H\overrightarrow{\mathrm{x_b^{(i)}}} >\right)\left(\mathrm{H}\overrightarrow{\mathrm{x_b^{(i)}}} - < H\overrightarrow{\mathrm{x_b^{(i)}}} >\right)^{\mathrm{T}}. \\ \text{with } < \overrightarrow{\mathrm{x_b^{(i)}}} >= \dfrac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}\overrightarrow{\mathrm{x_b^{(i)}}}. \end{cases} \quad 0.17$$

Similarly to KF, *B* is evolved explicitly propagated from one cycle to the next alongside the information about the state $\vec{x}$. First the state of the member number is evolved:

$$\overrightarrow{\mathrm{x^{(i)}}}(\mathrm{t}+1) = \mathrm{M}\left(\overrightarrow{\mathrm{x_a{}^{(i)}}}(\mathrm{t})\right) + \overrightarrow{\varepsilon_{\mathrm{m}}}. \quad 0.18$$

Then when all members are evolved, the ensemble mean and its covariance are reconstructed as follows:

$$\begin{cases} < \overrightarrow{\mathrm{x^{(i)}}} >= \dfrac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}\overrightarrow{\mathrm{x^{(i)}}}\ \ and \\ \mathrm{B} = \dfrac{1}{\mathrm{N}-1}\sum_{\mathrm{i}=1}^{\mathrm{N}}(\overrightarrow{\mathrm{x^{(i)}}} - < \overrightarrow{\mathrm{x^{(i)}}} >)(\overrightarrow{\mathrm{x^{(i)}}} - < \overrightarrow{\mathrm{x^{(i)}}} >)^{\mathrm{T}} = \mathrm{XX}^{\mathrm{T}}. \end{cases} \quad 0.19$$

Where X is an array that contains the ensemble perturbations. In the SENKF if the ensemble number is not sufficient due to the computational resources, perturbing the observation with random perturbations will cause sampling errors, which makes the filter suboptimal (Sakov and Oke, 2008). Statistics will not be necessarily representative of the state with estimated variances that are systematically too small (Bannister, 2017). For this reason DENKF methods such as the ensemble square root filter (EnSRF), the ensemble adjustment Kalman filter (EAKF), and the ensemble transform Kalman filter (ETKF) were proposed (Houtekamer and Zhang, 2016). The idea of DENKF is to create an updated ensemble with covariance consistent with Eq.0.10 using a transformation matrixT_M. Instead of updating each ensemble member separately as in Eq.0.13, the ESRF generates the new ensemble simultaneously as follows (Sakov and Oke, 2008; Zhang et al., 2009):

$$< \overrightarrow{x_a^{(i)}} >=< \overrightarrow{x_b^{(i)}} > +K\left(y+< \vec{H}\left(\overrightarrow{x_b^{(i)}}\right) >\right), \quad 0.20$$

$$K = BY_b^T\left(Y_b Y_b^T + R\right)^{-1} \; and \quad 0.21$$

$$P = BT_M. \quad 0.22$$

Where Y_b and T are the perturbation and transformation matrices. Y_b is expressed as follows:

$$Y_b = \frac{1}{\sqrt{N-1}}\left(\vec{H}\left(\overrightarrow{x_b^{(1)}}\right) - < \vec{H}\left(\overrightarrow{x_b^{(i)}}\right) >, \ldots, \vec{H}\left(\overrightarrow{x_b^{(N)}}\right) - < \vec{H}\left(\overrightarrow{x_b^{(i)}}\right) >\right). \quad 0.23$$

T_M is not unique, its different values lead to the EnSRF, EAKF and ETKF. To the best of the author knowledge no previous studies using ENKF to improve the solar irradiance forecast was found in the literature similarly to 4D-var.

III. Back and forth nudging

A variational variant of nudging method named back and forth nudging method was developed recently by Auroux and Blum (2008), (2005). The back and forth nudging algorithm consists in first solving the forward nudging equation (Eq.IV.11) and then the backward nudging equation. The initial condition of the backward integration is the final state obtained after integration of the forward nudging equation. At the end of this process, one obtains an estimate of the initial state of the system. These forward and backward integrations are then cycled until convergence of the algorithm. Therefore, the DA problem from the back and forth nudging (BFN) framework is expressed as follows:

$$\begin{cases} \vec{x}_k(t+1) = \vec{M}\left(\vec{x}_k(t)\right) + G\,\overrightarrow{\varepsilon_{0,k}}(t), \\ \vec{x}_k(t_0) = \vec{\tilde{x}}_{k-1}(t_0), \end{cases} \text{ and } \begin{cases} \vec{\tilde{x}}_k(t+1) = \vec{M}\left(\vec{\tilde{x}}_k(t)\right) + \tilde{G}\,\overrightarrow{\widetilde{\varepsilon_{0,k}}}(t), \\ \vec{\tilde{x}}_k(t) = \vec{x}_k(t). \end{cases} \quad 0.24$$

where k is the iteration step; $\vec{\tilde{x}}_k$ and $\vec{x}_k$ are the state vector in the backward and forward NDA respectively; $\overrightarrow{\widetilde{\varepsilon_{0,k}}}$ and $\overrightarrow{\varepsilon_{0,k}}$ and are he observation error in the backward and forward NDA respectively. Auroux and Blum (2008) tested his algorithm by comparison with the 4D-var using the Lorenz, Burgers and quasi-geostrophic model non-linear systems. Result found showed that BFN algorithm is better than the variational method for the same number of iterations (and hence for the same computing time). It converges in a small number of iterations. However, the BFN was not tested using more complex models such as primitive equations with various types of observations

Bibliography

Akinoğlu, B.G., 1991. A review of sunshine-based models used to estimate monthly average global solar radiation. Renew. Energy 1, 479–497.

Akinoğlu, B.G., Ecevit, A., 1990. Construction of a quadratic model using modified \AAngstrom coefficients to estimate global solar radiation. Sol. Energy 45, 85–92.

Al-Alawi, S.M., Al-Hinai, H.A., 1998. An ANN-based approach for predicting global radiation in locations with no direct measurement instrumentation. Renew. Energy 14, 199–204.

Albarelo, T., Marie-Joseph, I., Primerose, A., Seyler, F., Wald, L., Linguet, L., 2015. Optimizing the Heliosat-II Method for Surface Solar Irradiation Estimation with GOES Images. Can. J. Remote Sens. 41, 86–100. https://doi.org/10.1080/07038992.2015.1040876

Albert Einstein, 1905. The photoelectric effect.

AMS, 2012. glossary meteorology [WWW Document]. glossary.ametsoc.org.

ANDERSON, G.P., CHETWYND, J.H., SHE, E.P., 1986. AFGL Atmospheric Constituent Profiles (0.120km).

Anderson, J.L., co-authors, 2004. The New GFDL Global Atmosphere and Land Model AM2–LM2: Evaluation with Prescribed SST Simulations. J. Clim. 17, 4641–4673. https://doi.org/10.1175/JCLI-3223.1

Angstrom, A., 1924. Solar and terrestrial radiation. Report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation. Q. J. R. Meteorol. Soc. 50, 121–126.

Aryaputera, A.W., Yang, D., Walsh, W.M., 2015. Day-Ahead Solar Irradiance Forecasting in a Tropical Environment. J. Sol. Energy Eng. 137, 051009.

Asch, M., Bocquet, M., Nodet, M., 2016. Data Assimilation: Methods, Algorithms, and Applications. Society for Industrial and Applied Mathematics, Philadelphia, PA. https://doi.org/10.1137/1.9781611974546

Auroux, D., Blum, J., 2008. A nudging-based data assimilation method: the Back and Forth Nudging (BFN) algorithm. Nonlinear Process. Geophys. 15, 305–319. https://doi.org/10.5194/npg-15-305-2008

Auroux, D., Blum, J., 2005. Back and forth nudging algorithm for data assimilation problems. Comptes Rendus Math. 340, 873–878. https://doi.org/10.1016/j.crma.2005.05.006

Badescu, V. (Ed.), 2008. Modeling solar radiation at the earth's surface: recent advances ; [with CD-ROM]. Springer, Berlin Heidelberg.

Badosa, J., Haeffelin, M., Chepfer, H., 2013. Scales of spatial and temporal variation of solar irradiance on Reunion tropical island. Sol. Energy 88, 42–56. https://doi.org/10.1016/j.solener.2012.11.007

Badosa, J., Haeffelin, M., Kalecinski, N., Bonnardot, F., Jumaux, G., 2015. Reliability of day-ahead solar irradiance forecasts on Reunion Island depending on synoptic wind and humidity conditions. Sol. Energy 115, 306–321. https://doi.org/10.1016/j.solener.2015.02.039

Bannister, R., 2018. Variational data assimilation Background and method.

Bannister, R.N., 2017. A review of operational methods of variational and ensemble-variational data assimilation: Ensemble-variational Data Assimilation. Q. J. R. Meteorol. Soc. 143, 607–633. https://doi.org/10.1002/qj.2982

Bannister, R.N., 2008a. A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances. Q. J. R. Meteorol. Soc. 134, 1951–1970. https://doi.org/10.1002/qj.339

Bannister, R.N., 2008b. A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics. Q. J. R. Meteorol. Soc. 134, 1971–1996. https://doi.org/10.1002/qj.340

Barker, D.M., Clayton, A.M., 2011. Hybrid Variational-Ensemble Data Assimilation.

Barker, H.W., Cole, J.N.S., Morcrette, J.-J., Pincus, R., Räisänen, P., von Salzen, K., Vaillancourt, P.A., 2008. The Monte Carlo Independent Column Approximation: an assessment using several global atmospheric models. Q. J. R. Meteorol. Soc. 134, 1463–1478. https://doi.org/10.1002/qj.303

Bechtold, P., Bazile, E., Guichard, F., Mascart, P., Richard, E., 2001. A mass-flux convection scheme for regional and global models. Q. J. R. Meteorol. Soc. 127, 869–886. https://doi.org/10.1002/qj.49712757309

Bell labs, 1954. Bell Labs Demonstrates the First Practical Silicon Solar Cell.

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., Gleeson, E., Hansen-Sass, B., Homleid, M., Hortal, M., Ivarsson, K.-I., Lenderink, G., Niemelä, S., Nielsen, K.P., Onvlee, J., Rontu, L., Samuelsson, P., Muñoz, D.S., Subias, A., Tijm, S., Toll, V., Yang, X., Køltzow, M.Ø., 2017. The HARMONIE–AROME Model Configuration in the ALADIN–HIRLAM NWP System. Mon. Weather Rev. 145, 1919–1935. https://doi.org/10.1175/MWR-D-16-0417.1

Benjamin, S., Kim, D., Brown, J., 2002. Cloud/hydrometeor initialization in the 20-km RUC using GOES and radar data, in: Proceedings of 10th Conf. on Aviation, Range, and Aerospace Meteorology.

Benjamin, S.G., Grell, G.A., Brown, J.M., Smirnova, T.G., Bleck, R., 2004. Mesoscale Weather Prediction with the RUC Hybrid Isentropic–Terrain-Following Coordinate Model. Mon. Weather Rev. 132, 473–494. https://doi.org/10.1175/1520-0493(2004)132<0473:MWPWTR>2.0.CO;2

Bennett, A.F., 2004. Inverse Modeling of the Ocean and Atmosphere, Cambridge university press. ed.

Bertrand Cadet, Leah Goldfarb, Anne Réchou, 2003. A sub-tropical cirrus clouds climatology from Reunion Island (21°S, 55°E) lidar data set.

Beyer, H.G., Costanzo, C., Heinemann, D., 1996. Modifications of the Heliosat procedure for irradiance estimates from satellite images. Sol. Energy 56, 207–212. https://doi.org/10.1016/0038-092X(95)00092-6

Beyer, H.G., Martinez, J.P., Suri, M., Torres, J.L., Lorenz, E., Müller, S.C., Hoyer-Klick, C., Ineichen, P., 2009. D 1.1.3 Report on Benchmarking of Radiation Products 160.

Birol, F., 2010. World energy outlook 2010. Int. Energy Agency 1.

Bisht, G., Bras, R.L., 2010. Estimation of net radiation from the MODIS data under all sky conditions: Southern Great Plains case study. Remote Sens. Environ. 114, 1522–1534. https://doi.org/10.1016/j.rse.2010.02.007

Blanc, P., Gschwind, B., Lefèvre, M., Wald, L., 2011. The HelioClim project: Surface solar irradiance data for climate applications. Remote Sens. 3, 343–361.

Boulet, P., Brissinger, D., Collin, A., Acem, Z., Parent, G., 2015. On the Influence of the Sample Absorptivity when Studying the Thermal Degradation of Materials. Materials 8, 5398–5413. https://doi.org/10.3390/ma8085251

Bourgin, F., Ramos, M.H., Thirel, G., Andréassian, V., 2014. Investigating the interactions between data assimilation and post-processing in hydrological ensemble forecasting. J. Hydrol. 519, 2775–2784. https://doi.org/10.1016/j.jhydrol.2014.07.054

Bouttier, F., Courtier, P., 1999. Data Assimilation Concepts and Methods.

Bovolo, C.I., Pereira, R., Parkin, G., Kilsby, C., Wagner, T., 2012. Fine-scale regional climate patterns in the Guianas, tropical South America, based on observations and reanalysis data. Int. J. Climatol. 32, 1665–1689. https://doi.org/10.1002/joc.2387

BP Statistical Review, 2017. bp-statistical-review-of-world-energy-2017-full-report.pdf.

Bright, J.M., Babacan, O., Kleissl, J., Taylor, P.G., Crook, R., 2017. A synthetic, spatially decorrelating solar irradiance generator and application to a LV grid model with high PV penetration. Sol. Energy 147, 83–98. https://doi.org/10.1016/j.solener.2017.03.018

Brunner, D., Savage, N., Jorba, O., Eder, B., Giordano, L., Badia, A., Balzarini, A., Baró, R., Bianconi, R., Chemel, C., Curci, G., Forkel, R., Jiménez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L., Im, U., Knote, C., Makar, P., Manders-Groot, A., van Meijgaard, E., Neal, L., Pérez, J.L., Pirovano, G., San Jose, R., Schröder, W., Sokhi, R.S., Syrakov, D., Torian, A., Tuccella, P., Werhahn, J., Wolke, R., Yahya, K., Zabkar, R., Zhang, Y., Hogrefe, C., Galmarini, S., 2015. Comparative analysis of meteorological performance of coupled chemistry-meteorology models in the context of AQMEII phase 2. Atmos. Environ. 115, 470–498. https://doi.org/10.1016/j.atmosenv.2014.12.032

Buehner, M., Morneau, J., Charette, C., 2013. Four-dimensional ensemble-variational data assimilation for global deterministic weather prediction. Nonlinear Process. Geophys. 20, 669–682. https://doi.org/10.5194/npg-20-669-2013

Burgers, G., 1998. Analysis Scheme in the Ensemble Kalman Filter. Mon. WEATHER Rev. 126, 6.

Burridge, D.M., 1975. A split semi-implict reformulation of the Bushby-Timpson 10-level model. Q. J. R. Meteorol. Soc. 101, 777–792. https://doi.org/10.1002/qj.49710143006

Campbell, P.M., 1969. A numerical method for discrete ordinate and moment equations in radiative transfer. Int. J. Heat Mass Transf. 12, 497–507. https://doi.org/10.1016/0017-9310(69)90145-8

Campillo, C., Fortes, R., del Henar Prieto, M., 2012. Solar radiation effect on crop production, in: Solar Radiation. InTech.

Cano, D., Monget, J., Albuisson, M., Guillard, H., Regas, N., Wald, L., 1986. A method for the determination of the global solar radiation from meteorological satellites data. Sol Energy 37 31–39.

Centeno López, E., 2016. energy2016_EvaCentenoLopez.pdf.

Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. Geosci. Model Dev. 7, 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Chen, Y., Rizvi, S.R.H., Huang, X.-Y., Min, J., Zhang, X., 2013. Balance characteristics of multivariate background error covariances and their impact on analyses and forecasts in tropical and Arctic regions. Meteorol. Atmospheric Phys. 121, 79–98. https://doi.org/10.1007/s00703-013-0251-y

Chin, M., Rood, B., Thompson, A., 2000. Atmospheric sulfur cycle simulated in the global model GOCART: Model description and global properties.

Chosson, F., Vaillancourt, P.A., Milbrandt, J.A., Yau, M.K., Zadra, A., 2014. Adapting Two-Moment Microphysics Schemes across Model Resolutions: Subgrid Cloud and Precipitation Fraction and Microphysical Sub–Time Step. J. Atmospheric Sci. 71, 2635–2653. https://doi.org/10.1175/JAS-D-13-0367.1

Chow, C.W., Urquhart, B., Lave, M., Dominguez, A., Kleissl, J., Shields, J., Washom, B., 2011. Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. Sol. Energy 85, 2881–2893. https://doi.org/10.1016/j.solener.2011.08.025

CIE, 2011. ILV: International lighting vocabulary, new. Commission Internationale de L'Eclairage.

Cintineo, R., Otkin, J.A., Xue, M., Kong, F., 2014. Evaluating the performance of planetary boundary layer and cloud microphysical parameterization schemes in convection-permitting ensemble forecasts using synthetic GOES-13 satellite observations. Mon. Weather Rev. 142, 163–182.

Cionni, I., Eyring, V., Lamarque, J.F., Randel, W.J., Stevenson, D.S., Wu, F., Bodeker, G.E., Shepherd, T.G., Shindell, D.T., Waugh, D.W., 2011. Ozone database in support of CMIP5 simulations: results and corresponding radiative forcing. Atmospheric Chem. Phys. 11, 11267–11292. https://doi.org/10.5194/acp-11-11267-2011

Clack, C.T.M., 2017. Modeling Solar Irradiance and Solar PV Power Output to Create a Resource Assessment Using Linear Multiple Multivariate Regression. J. Appl. Meteorol. Climatol. 56, 109–125. https://doi.org/10.1175/JAMC-D-16-0175.1

Clayton, A.M., Lorenc, A.C., Barker, D.M., 2013. Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. Q. J. R. Meteorol. Soc. 139, 1445–1461. https://doi.org/10.1002/qj.2054

Clean Energy Council, 2017. CLEAN ENERGY AUSTRALIA.

Cohen, A.E., Cavallo, S.M., Coniglio, M.C., Brooks, H.E., 2015. A Review of Planetary Boundary Layer Parameterization Schemes and Their Sensitivity in Simulating Southeastern U.S. Cold Season Severe Weather Environments. Weather Forecast. 30, 591–612. https://doi.org/10.1175/WAF-D-14-00105.1

Coiffier, J., 2011. Fundamentals of Numerical Weather Prediction. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511734458

Cossu, F., Hocke, K., 2014. Influence of microphysical schemes on atmospheric water in the Weather Research and Forecasting model. Geosci Model Dev 147–160. https://doi.org/10.5194/gmd-7-147-2014

Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M., Staniforth, A., 1998. The Operational CMC–MRB Global Environmental Multiscale (GEM) Model. Part I: Design Considerations and Formulation. Mon. Weather Rev. 126, 1373–1395. https://doi.org/10.1175/1520-0493(1998)126<1373:TOCMGE>2.0.CO;2

Coulson, K.L., 1959. Characteristics of the radiation emerging from the top of a Rayleigh atmosphere—II. Planet. Space Sci. 1, 277–284. https://doi.org/10.1016/0032-0633(59)90032-7

Crétat, J., Pohl, B., Richard, Y., Drobinski, P., 2012. Uncertainties in simulating regional climate of Southern Africa: sensitivity to physical parameterizations using WRF. Clim. Dyn. 38, 613–634.

Cuxart, J., Bougeault, P., Redelsperger, J.-L., 2000. A turbulence scheme allowing for mesoscale and large-eddy simulations. Q. J. R. Meteorol. Soc. 126, 1–30. https://doi.org/10.1002/qj.49712656202

Dagestad, K.-F., 2004. Mean bias deviation of the Heliosat algorithm for varying cloud properties and sun-ground-satellite geometry. Theor. Appl. Climatol. 79, 215–224. https://doi.org/10.1007/s00704-004-0072-5

Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C.L., 2010. Technical Description of version 4.0 of the Community Land Model (CLM) 266.

Daley, R., 1999. Atmospheric data analysis, 1. paperback ed., reprinted, transferred to digital print. ed, Cambridge atmospheric and space science series. Cambridge Univ. Press, Cambridge.

Dambreville, R., 2014. Prévision du rayonnement solaire global par télédétection pour la gestion de la production d'énergie photovoltaïque. Grenoble.

Dambreville, R., Blanc, P., Chanussot, J., Boldo, D., 2014. Very short term forecasting of the Global Horizontal Irradiance using a spatio-temporal autoregressive model. Renew. Energy 72, 291–300. https://doi.org/10.1016/j.renene.2014.07.012

[dataset] Maass, C., 2017. MARS content [WWW Document]. software.ecmwf.int. URL https://software.ecmwf.int/wiki/display/UDOC/MARS+content (accessed 5.14.18).

[dataset] NCEP-NWS-NOAA-USDC, 2015. NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive.

[dataset] Transvalor, M.P., 2014. LINKE TURBIDITY [WWW Document]. Soda-Procom. URL http://www.soda-pro.com/web-services/atmosphere/linke-turbidity-factor-ozone-water-vapor-and-angstroembeta

Davies, J.A., Abdel-Wahab, M., McKAY, D.C., 1984. Estimating Solar Irradiation on Horizontal Surfaces. Int. J. Sol. Energy 2, 405–424. https://doi.org/10.1080/01425918408909940

Déqué, M., Dreveton, C., Braun, A., Cariolle, D., 1994. The ARPEGE/IFS atmosphere model: a contribution to the French community climate modelling. Clim. Dyn. 10, 249–266.

Descombes, G., Auligné, T., Vandenberghe, F., Barker, D.M., Barré, J., 2015. Generalized background error covariance matrix model (GEN_BE v2.0). Geosci. Model Dev. 8, 669–696. https://doi.org/10.5194/gmd-8-669-2015

Dhanya, M., Chandrasekar, A., 2016. Impact of variational assimilation using multivariate background error covariances on the simulation of monsoon depressions over India. Ann. Geophys. 34, 187–201. https://doi.org/10.5194/angeo-34-187-2016
Diagne, M., David, M., Boland, J., Schmutz, N., Lauret, P., 2014. Post-processing of solar irradiance forecasts from WRF model at Reunion Island. Sol. Energy 105, 99–108. https://doi.org/10.1016/j.solener.2014.03.016
Diagne, M., David, M., Lauret, P., Boland, J., Schmutz, N., 2013. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renew. Sustain. Energy Rev. 27, 65–76. https://doi.org/10.1016/j.rser.2013.06.042
Diak, G.R., Gautier, C., 1983. Improvements to a Simple Physical Model for Estimating Insolation from GOES Data. J. Clim. Appl. Meteorol. 22, 505–508. https://doi.org/10.1175/1520-0450(1983)022<0505:ITASPM>2.0.CO;2
Diak, G.R., Gautier, C., Masse, S., 1982. An operational system for mapping insolation from goes satellite data. Sol. Energy 28, 371–376. https://doi.org/10.1016/0038-092X(82)90254-7
Diallo, M., Verbois, H., Chane-Ming, F., Walsh, W., Linguet, L., 2018. Assessing GFS and IFS global numerical weather prediction model forecast accuracy under tropical climate. Submitt. Sol. Energy.
Doms, G., Baldauf, M., 2015. A Description of the Nonhydrostatic Regional COSMO-Model 164.
Driesse, A., Zaaiman, W., Riley, D., Taylor, N., Stein, J.S., 2016. Investigation of pyranometer and photodiode calibrations under different conditions. IEEE, pp. 0127–0132. https://doi.org/10.1109/PVSC.2016.7749562
DTC, 2016. GFS Operational Physics Documentation [WWW Document]. URL https://dtcenter.org/GMTB/gfs_phys_doc_dev/index.html
Dudhia, J., 2014. A history of mesoscale model development. Asia-Pac. J. Atmospheric Sci. 50, 121–131. https://doi.org/10.1007/s13143-014-0031-8
Dudhia, J., 1989. Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. J. Atmospheric Sci. 46, 3077–3107. https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2
Duguay, C.R., LeDrew, E.F., 1992. Estimating Surface Reflectance and Albedo from Landsat-5 Thematic Mapper over Rugged Terrain. Photogramm. Eng. 9.
Dürr, B., Zelenka, A., 2009. Deriving surface global irradiance over the Alpine region from METEOSAT Second Generation data by supplementing the HELIOSAT method. Int. J. Remote Sens. 30, 5821–5841. https://doi.org/10.1080/01431160902744829
ECMWF, 2018. L137 model level definitions [WWW Document]. ecmwf.int. URL https://www.ecmwf.int/en/forecasts/documentation-and-support/137-model-levels
ECMWF, 2016a. IFS DOCUMENTATION – Cy41r2 Operational implementation 8 March 2016 PART IV: PHYSICAL PROCESSES.
ECMWF, 2016b. Cycle 41r2 summary of changes [WWW Document]. URL https://www.ecmwf.int/en/forecasts/documentation-and-support/evolution-ifs/cycles/cycle-43r1-summary-changes
ECMWF, 2016c. IFS DOCUMENTATION – Cy41r2 Operational implementation 8 March 2016 PART II: DATA ASSIMILATION.
Elizondo, D., Hoogenboom, G., McClendon, R.W., 1994. Development of a neural network model to predict daily solar radiation. Agric. For. Meteorol. 71, 115–132.
Ellis, J.S., Haar, T.V., 1976. Zonal average earth radiation budget measurements from satellites for climate studies. NASA-CR 149319, Atmospheric Science Paper 240, 1–63.
Enerdata, 2017. Enerdata_Energy_Statistical_Yearbook_2017.xlsx.
Environmental Modeling Center, 2003. NCEP Office Note 442 [WWW Document]. URL http://www.lib.ncep.noaa.gov/ncepofficenotes/files/on442.pdf
Espinar, B., Ramírez, L., Drews, A., Beyer, H.G., Zarzalejo, L.F., Polo, J., Martín, L., 2009. Analysis of different comparison parameters applied to solar radiation data from satellite and German radiometric stations. Sol. Energy 83, 118–125. https://doi.org/10.1016/j.solener.2008.07.009

Espinar, B., Wald, L., Blanc, P., Hoyer-Klick, C., Homscheidt, M.S., Wanderer, T., 2011. Project ENDORSE-Excerpt of the report on the harmonization and qualification of meteorological data: Procedures for quality check of meteorological data. Mines ParisTech.
Evensen, G., 2003. The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean Dyn. 53, 343–367. https://doi.org/10.1007/s10236-003-0036-9
Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. 99, 10143. https://doi.org/10.1029/94JC00572
Fairley, P., 2008. Solar-Cell Squabble. IEEE Spectr. 45, 36–40. https://doi.org/10.1109/MSPEC.2008.4476444
Fanglin, Y., 2015. Review of GFS Forecast Skills in 2015.
Fanglin, Y., 2014. Review of GFS Forecast Skills in 2014 [WWW Document]. URL https://vlab.ncep.noaa.gov/documents/95651/849649/GFS.performance.review.2014.pdf/e0713131-f5ef-4512-b517-61561338542d?version=1.0&download=true
Fanglin, Y., 2013. Review of GFS Forecast Skills in 2013 [WWW Document]. URL https://vlab.ncep.noaa.gov/documents/95651/849649/GFS.performance.review.2013.pdf/07d5b70c-0dc6-4e6e-a6fb-22b276a50fed?version=1.0&download=true
Fierro, A.O., Gao, J., Ziegler, C.L., Mansell, E.R., MacGorman, D.R., Dembek, S.R., 2014. Evaluation of a Cloud-Scale Lightning Data Assimilation Technique and a 3DVAR Method for the Analysis and Short-Term Forecast of the 29 June 2012 Derecho Event. Mon. Weather Rev. 142, 183–202. https://doi.org/10.1175/MWR-D-13-00142.1
Fillol, E., Albarelo, T., Primerose, A., Wald, L., Linguet, L., 2017. Spatiotemporal indicators of solar energy potential in the Guiana Shield using GOES images. Renew. Energy 111, 11–25. https://doi.org/10.1016/j.renene.2017.03.081
Forbes, R.M., Tompkins, A.M., Untch, A., 2011. A New Prognastic Bulk Microphysics Scheme for the IFS. European Centre for Medium-Range Weather Forecasts.
Fouquart, Y., Bonnel, B., Ramaswamy, V., 1991. Intercomparing shortwave radiation codes for climate studies. J. Geophys. Res. Atmospheres 96, 8955–8968.
Fowler, A., 2016. The Ensemble Kalman filter.
Fritzner, S.M., Graversen, R.G., Wang, K., Christensen, K.H., 2018. Comparison between a multi-variate nudging method and the ensemble Kalman filter for sea-ice data assimilation. J. Glaciol. 64, 387–396. https://doi.org/10.1017/jog.2018.33
Frouin, R., Lingner, D.W., Gautier, C., Baker, K.S., Smith, R.C., 1989. A simple analytical formula to compute clear sky total and photosynthetically available solar irradiance at the ocean surface. J. Geophys. Res. Oceans 94, 9731–9742.
Fu, C.-L., Cheng, H.-Y., 2013. Predicting solar irradiance with all-sky image features via regression. Sol. Energy 97, 537–550. https://doi.org/10.1016/j.solener.2013.09.016
Galvin, J.F.., 2015. An Introduction to the Meteorology and Climate of the Tropics.
Gao, J., Xue, M., Stensrud, D.J., 2013. The Development of a Hybrid EnKF-3DVAR Algorithm for Storm-Scale Data Assimilation. Adv. Meteorol. 2013, 1–12. https://doi.org/10.1155/2013/512656
Gautier, C., Diak, G., Masse, S., 1980. A simple physical model to estimate incident solar radiation at the surface from GOES satellite Data. J. Appl. Meteorol. 19, 1005–1012. https://doi.org/10.1175/1520-0450(1980)019<1005:ASPMTE>2.0.CO;2
Gautier, C., Landsfeld, M., 1997. Surface solar radiation flux and cloud radiative forcing for the Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP): A satellite, surface observations, and radiative transfer model study. J. Atmospheric Sci. 54, 1289–1307.
Geiger, B., Meurey, C., Lajas, D., Franchistéguy, L., Carrer, D., Roujean, J.-L., 2008. Near real-time provision of downwelling shortwave radiation estimates derived from satellite observations. Meteorol. Appl. 15, 411–420. https://doi.org/10.1002/met.84
Geiger, M., Diabaté, L., Ménard, L., Wald, L., 2002. A web service for controlling the quality of measurements of global solar irradiation. Sol. Energy 73, 475–480. https://doi.org/10.1016/S0038-092X(02)00121-4

Ghil, M., Malanotte-Rizzoli, P., 1991. Data Assimilation in Meteorology and Oceanography, in: Advances in Geophysics. Elsevier, pp. 141–266. https://doi.org/10.1016/S0065-2687(08)60442-2

Gill, D., 2016. WRF tutorial: nesting in WRF [WWW Document]. URL http://www2.mmm.ucar.edu/wrf/users/wrfda/tutorial.html

Girodo, M., Mueller, R.W., Heinemann, D., 2006. Influence of three-dimensional cloud effects on satellite derived solar irradiance estimation—First approaches to improve the Heliosat method. Sol. Energy 80, 1145–1159. https://doi.org/10.1016/j.solener.2005.09.005

Glahn, H., Dale, L., 1972. (MOS) The use of model output statistic in objective weather forecasting.

Glassley, W., Kleissl, J., Van Dam, C.., Shiu, H., Huang, J., Braun, G., Holland, R., 2012. CALIFORNIA RENEWABLE ENERGY FORECASTING, RESOURCE DATA, AND MAPPING: CURRENT STATE OF THE ART IN SOLAR FORECASTING.

Grell, G.A., Dévényi, D., 2002. A generalized approach to parameterizing convection combining ensemble and data assimilation techniques: PARAMETERIZING CONVECTION COMBINING ENSEMBLE AND DATA ASSIMILATION TECHNIQUES. Geophys. Res. Lett. 29, 38-1-38–4. https://doi.org/10.1029/2002GL015311

Gueymard, C.A., 2012a. Clear-sky irradiance predictions for solar resource mapping and large-scale applications: Improved validation methodology and detailed performance analysis of 18 broadband radiative models. Sol. Energy 86, 2145–2169. https://doi.org/10.1016/j.solener.2011.11.011

Gueymard, C.A., 2012b. Clear-sky irradiance predictions for solar resource mapping and large-scale applications: Improved validation methodology and detailed performance analysis of 18 broadband radiative models. Sol. Energy 86, 2145–2169. https://doi.org/10.1016/j.solener.2011.11.011

Gueymard, C.A., 1998. Turbidity determination from broadband irradiance measurements: a detailed multicoefficient approach. J. Appl. Meteorol. 37, 414–435.

Guo, Y., H, L., X, M., Y, H., T, T., H, K., 2008. Impact of WRF-Var (3DVar) Background Error Statistics on Typhoon analysis and Forecast. 7th WRF Users' Workshop.

Gustafsson, N., Janjić, T., Schraff, C., Leuenberger, D., Weissmann, M., Reich, H., Brousseau, P., Montmerle, T., Wattrelot, E., Bučánek, A., Mile, M., Hamdi, R., Lindskog, M., Barkmeijer, J., Dahlbom, M., Macpherson, B., Ballard, S., Inverarity, G., Carley, J., Alexander, C., Dowell, D., Liu, S., Ikuta, Y., Fujita, T., 2018. Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres: GUSTAFSSON *et al.* . Q. J. R. Meteorol. Soc. https://doi.org/10.1002/qj.3179

Haben, S.A., Lawless, A.S., Nichols, N.K., 2009. Conditioning of the 3DVAR Data Assimilation Problem.

Haiden, T., Janousek, M., F, 2015. Evaluation of ECMWF forecasts, including 2014-2015 upgrades [WWW Document]. URL https://www.ecmwf.int/sites/default/files/elibrary/2015/15275-evaluation-ecmwf-forecasts-including-2014-2015-upgrades.pdf

Haiden, T., Janousek, M., Richardson, D., 2016. Evaluation of ECMWF forecasts, including the 2016 resolution [WWW Document]. URL https://www.ecmwf.int/en/elibrary/16924-evaluation-ecmwf-forecasts-including-2016-resolution-upgrade

Haiden, T., Janousek, M., Vitart, F., 2014. Evaluation of ECMWF forecasts, including 2013-2014 upgrades [WWW Document]. URL https://www.ecmwf.int/sites/default/files/elibrary/2014/12525-evaluation-ecmwf-forecasts-including-2013-2014-upgrades.pdf

Hamill, T.M., Nehrkorn, T., 1993. A Short-Term Cloud Forecast Scheme Using Cross Correlations. Weather Forecast. 8, 401–411. https://doi.org/10.1175/1520-0434(1993)008<0401:ASTCFS>2.0.CO;2

Han, J., Pan, H.-L., 2011. Revision of Convection and Vertical Diffusion Schemes in the NCEP Global Forecast System. Weather Forecast. 26, 520–533. https://doi.org/10.1175/WAF-D-10-05038.1

Han, J., Witek, M.L., Teixeira, J., Sun, R., Pan, H.-L., Fletcher, J.K., Bretherton, C.S., 2016. Implementation in the NCEP GFS of a Hybrid Eddy-Diffusivity Mass-Flux (EDMF) Boundary Layer Parameterization with Dissipative Heating and Modified Stable Boundary Layer Mixing. Weather Forecast. 31, 341–352. https://doi.org/10.1175/WAF-D-15-0053.1

Hansen, J.E., Travis, L.., 1974. Light scattering in planetary atmospheres, Kluwer Academic Publishers. ed, Space Sci Rev (1974).

Hay, J.E., 1993. Satellite based estimates of solar irradiance at the earth's surface—I. Modelling approaches. Renew. Energy 3, 381–393. https://doi.org/10.1016/0960-1481(93)90105-P

Hay, J.E., Hanson, K., 1978. A satellite-based methodology for determining solar irradiance at the ocean surface during GATE. Bull Am Meteorol Soc 59, 1549.

Heinle, A., Macke, A., Srivastav, A., 2010. Automatic cloud classification of whole sky images. Atmospheric Meas. Tech. 3, 557–567. https://doi.org/10.5194/amt-3-557-2010

Hess, D., McKnight, T.L., 2013. McKnight's physical geography: a landscape appreciation, Eleventh ed. ed. Pearson Prentice Hall, Upper Saddle River, NJ.

Hess, M., Koepke, P., Schult, I., 1998. Optical properties of aerosols and clouds: The software package OPAC. Bull. Am. Meteorol. Soc. 79, 831–844.

Hillger, D.W., Schmit, T.J., 2009. Observing Systems: The GOES-13 Science Test: A Synopsis. Bull. Am. Meteorol. Soc. 90, 592–597. https://doi.org/10.1175/2008BAMS2564.1

Hogan, R., 2015. Radiation Quantities in the ECMWF model and MARS.

Holm, E.V., 2008. Lecture notes on assimilation algorithms.

Hong, S.-Y., Noh, Y., Dudhia, J., 2006. A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. Mon. Weather Rev. 134, 2318–2341. https://doi.org/10.1175/MWR3199.1

Hong, S.-Y., Pan, H.-L., 1996. Nonlocal Boundary Layer Vertical Diffusion in a Medium-Range Forecast Model.

Hong, Y., Hsu, K.-L., Sorooshian, S., Gao, X., 2004. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. J. Appl. Meteorol. 43, 1834–1853.

House, F.B., Gruber, A., Hunt, G.E., Mecherikunnel, A.T., 1986. History of satellite missions and measurements of the Earth Radiation Budget (1957–1984). Rev. Geophys. 24, 357. https://doi.org/10.1029/RG024i002p00357

Houtekamer, P.L., Zhang, F., 2016. Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation. Mon. Weather Rev. 144, 4489–4532. https://doi.org/10.1175/MWR-D-15-0440.1

Howes, K.E., Fowler, A.M., Lawless, A.S., 2017. Accounting for model error in strong-constraint 4D-Var data assimilation: Accounting for Model Error in 4D-Var. Q. J. R. Meteorol. Soc. 143, 1227–1240. https://doi.org/10.1002/qj.2996

Hu, C., Muller-Karger, F.E., Zepp, R.G., 2002. Absorbance, absorption coefficient, and apparent quantum yield: A comment on common ambiguity in the use of these optical concepts. Limnol. Oceanogr. 47, 1261–1267. https://doi.org/10.4319/lo.2002.47.4.1261

Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shephard, M.W., Clough, S.A., Collins, W.D., 2008. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. J. Geophys. Res. 113. https://doi.org/10.1029/2008JD009944

IEA, 2017. World Energy Outlook-2017.

IER, 2013. Germany's Green Energy Destabilizing Electric Grids [WWW Document]. URL https://www.instituteforenergyresearch.org/analysis/germanys-green-energy-destabilizing-electric-grids/

Inman, R.H., Pedro, H.T.C., Coimbra, C.F.M., 2013. Solar forecasting methods for renewable energy integration. Prog. Energy Combust. Sci. 39, 535–576. https://doi.org/10.1016/j.pecs.2013.06.002

Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R.J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V.,

Hurtmans, D., Jones, L., Kaiser, J.W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M.G., Simmons, A.J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., the MACC team, 2013. The MACC reanalysis: an 8 yr data set of atmospheric composition. Atmospheric Chem. Phys. 13, 4073–4109. https://doi.org/10.5194/acp-13-4073-2013

INSEE: Rémi Charrier, 2017. Synthèse démographique de la Guyane Une démographie toujours dynamique.

IRENA, 2018. Renewable power generation costs in 2017.

IRENA, 2016. renewable cost IRENA_Power_to_Change_2016.pdf.

IRENA, ETSAP, 2015. Renewable Energy Integration in Power Grids.

Jacobson, M.Z., 2005. Fundamentals of atmospheric modeling. Cambridge university press.

Janjai, S., 2010. A method for estimating direct normal solar irradiation from satellite data for a tropical environment. Sol. Energy 84, 1685–1695. https://doi.org/10.1016/j.solener.2010.05.017

Janjai, S., Laksanaboonsong, J., Nunez, M., Thongsathitya, A., 2005. Development of a method for generating operational solar radiation maps from satellite data for a tropical environment. Sol. Energy 78, 739–751.

Janjai, S., Masiri, I., Laksanaboonsong, J., 2013. Satellite-derived solar resource maps for Myanmar. Renew. Energy 53, 132–140. https://doi.org/10.1016/j.renene.2012.11.014

Janjai, S., Pankaew, P., Laksanaboonsong, J., Kitichantaropas, P., 2011. Estimation of solar radiation over Cambodia from long-term satellite data. Renew. Energy 36, 1214–1220. https://doi.org/10.1016/j.renene.2010.09.023

Janjic, T., n.d. Ensemble Kalman filter methods.

Jimenez, P.A., Hacker, J.P., Dudhia, J., Ellen Haupt, S., Ruiz-Arias, J.A., Gueymard, C.A., Thompson, G., Eidhammer, T., Deng, A., 2015. WRF-Solar: An augmented NWP model for solar power prediction. Model description and clear sky assessment. Bull. Am. Meteorol. Soc. 151029072131009. https://doi.org/10.1175/BAMS-D-14-00279.1

Jimenez, P.A., Hacker, J.P., Dudhia, J., Haupt, S.E., Ruiz-Arias, J.A., Gueymard, C.A., Thompson, G., Eidhammer, T., Deng, A., 2016. WRF-Solar: Description and Clear-Sky Assessment of an Augmented NWP Model for Solar Power Prediction. Bull. Am. Meteorol. Soc. 97, 1249–1264. https://doi.org/10.1175/BAMS-D-14-00279.1

Jumaily, K., AL-Salihi, A.M., Al-Tai, O.T., 2010. Evaluation of Meteosat-8 measurements using daily global solar radiation for two stations in Iraq. Int J Energy Env. 1 635–642.

Käellen, E., 1996. HIRLAM DOCUMENTATION [WWW Document]. URL http://hirlam.org/index.php/component/docman/cat_view/114-model-and-system-documentation/131-hirlam-documentation?Itemid=70

Kalnay, E., 2009. Atmospheric modeling, data assimilation and predictability, 5. print. ed. Cambridge Univ. Press, Cambridge.

Katzfuss, M., Stroud, J.R., Wikle, C.K., 2016. Understanding the Ensemble Kalman Filter. Am. Stat. 70, 350–357. https://doi.org/10.1080/00031305.2016.1141709

Kavulich, M.J., 2017. Observations Lecture 1: Conventional observations and Observation Pre-processing for WRFDA (OBSPROC).

Khalil, S.A., 2008. Parameterization models for solar radiation and solar technology applications. Energy Convers. Manag. 49, 2384–2391. https://doi.org/10.1016/j.enconman.2008.01.023

Kleissl, J., 2013. Solar Energy Forecasting and Resource Assessment, 1. ed. ed. Elsevier Science, Burlington.

Kneizys, F.., Abreu, L.., Anderson, G.., Chetwynd, J.., Shettle, E.., Berk, A., Bernstein, L.., Robertson, D.., Acharya, P., Rothman, L.., Selby, J.., Gallery, W.., Clough, S.., 1996. The MODTRAN 2/3 Report and LOWTRAN 7 MODEL.

Köhler, M., Ahlgrimm, M., Beljaars, A., 2011. Unified treatment of dry convective and stratocumulus-topped boundary layers in the ECMWF model. Q. J. R. Meteorol. Soc. 137, 43–57. https://doi.org/10.1002/qj.713

Kotsopoulos, S., Pytharoulis, I., Tegoulias, I., Karacostas, T., 2014. A Comparative Study of Two Land Surface Schemes in WRF-ARW Model. Presented at the COMECAP 2014 e-book of proceedings, p. 33.

Kutty, G., Muraleedharan, R., Kesarkar, A.P., 2018. Impact of Representing Model Error in a Hybrid Ensemble-Variational Data Assimilation System for Track Forecast of Tropical Cyclones over the Bay of Bengal. Pure Appl. Geophys. 175, 1155–1167. https://doi.org/10.1007/s00024-017-1747-z

Lacis, A.A., Hansen, J., 1974. A Parameterization for the Absorption of Solar Radiation in the Earth's Atmosphere. J. Atmospheric Sci. 31, 118–133. https://doi.org/10.1175/1520-0469(1974)031<0118:APFTAO>2.0.CO;2

Laing, Evans, 2011. Introduction to tropical meteorology, Educational material from The COMET Program.

Lakshmivarahan, S., Lewis, J.M., 2013. Nudging Methods: A Critical Overview, in: Park, S.K., Xu, L. (Eds.), Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II). Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 27–57. https://doi.org/10.1007/978-3-642-35088-7_2

Langella, R., Proto, D., Testa, A., 2016. Solar Radiation Forecasting, Accounting for Daily Variability. Energies 9, 200. https://doi.org/10.3390/en9030200

Lara-Fanego, V., Ruiz-Arias, J.A., Pozo-Vázquez, D., Santos-Alamillos, F.J., Tovar-Pescador, J., 2012. Evaluation of the WRF model solar irradiance forecasts in Andalusia (southern Spain). Sol. Energy 86, 2200–2217. https://doi.org/10.1016/j.solener.2011.02.014

Laroche, S., Gauthier, P., Tanguay, M., Pellerin, S., Morneau, J., 2007. Impact of the Different Components of 4DVAR on the Global Forecast System of the Meteorological Service of Canada. Mon. Weather Rev. 135, 2355–2364. https://doi.org/10.1175/MWR3408.1

Lauret, P., Diagne, M., David, M., 2014. A Neural Network Post-processing Approach to Improving NWP Solar Radiation Forecasts. Energy Procedia 57, 1044–1052. https://doi.org/10.1016/j.egypro.2014.10.089

Lawless, A.S., Gratton, S., Nichols, N.K., 2005. An investigation of incremental 4D-Var using non-tangent linear models. Q. J. R. Meteorol. Soc. 131, 459–476. https://doi.org/10.1256/qj.04.20

Lefèvre, M., Wald, L., Diabaté, L., 2007. Using reduced data sets ISCCP-B2 from the Meteosat satellites to assess surface solar irradiance. Sol. Energy 81, 240–253. https://doi.org/10.1016/j.solener.2006.03.008

Lehr, U., Nitsch, J., Kratzat, M., Lutz, C., Edler, D., 2006. Renewable energy and employment in Germany. Energy Policy 36, 108–117.

Lei, J., Bickel, P., Snyder, C., 2010. Comparison of Ensemble Kalman Filters under Non-Gaussianity. Mon. Weather Rev. 138, 1293–1306. https://doi.org/10.1175/2009MWR3133.1

Lei, L., 2011. A hybrid nudging-ensemble Kalman filter approach to data assimilation. The Pennsylvania State University.

Lei, L., Hacker, J.P., 2015. Nudging, Ensemble, and Nudging Ensembles for Data Assimilation in the Presence of Model Error. Mon. Weather Rev. 143, 2600–2610. https://doi.org/10.1175/MWR-D-14-00295.1

Levanon, N., 1971. Determination of the sea surface slope distribution and wind velocity using sun glitter viewed from a synchronous satellite. J. Phys. Oceanogr. 1, 214–220.

Li, Xiangshang, Choi, Y., Czader, B., Roy, A., Kim, H., Lefer, B., Pan, S., 2016. The impact of observation nudging on simulated meteorology and ozone concentrations during DISCOVER-AQ 2013 Texas campaign. Atmospheric Chem. Phys. 16, 3127–3144. https://doi.org/10.5194/acp-16-3127-2016

Li, X., Ming, J., Xue, M., Wang, Y., Zhao, K., 2015. Implementation of a dynamic equation constraint based on the steady state momentum equations within the WRF hybrid ensemble-3DVar data assimilation system and test with radar T-TREC wind assimilation for tropical Cyclone Chanthu (2010): HYBRID DATA ASSIMILATION WITH CONSTRAINT. J. Geophys. Res. Atmospheres 120, 4017–4039. https://doi.org/10.1002/2014JD022706

Li, Xin, Zeng, M., Wang, Y., Wang, W., Wu, H., Mei, H., 2016. Evaluation of two momentum control variable schemes and their impact on the variational assimilation of radarwind data: Case study of a squall line. Adv. Atmospheric Sci. 33, 1143–1157. https://doi.org/10.1007/s00376-016-5255-3
Liang, C., 2012. A COMPARATIVE STUDY OF TWO LAND SURFACE SCHEMES IN WRF MODEL OVER EASTERN CHINA. J. Trop. Meteorol. 12.
Lim, K.-S.S., Hong, S.-Y., 2010. Development of an Effective Double-Moment Cloud Microphysics Scheme with Prognostic Cloud Condensation Nuclei (CCN) for Weather and Climate Models. Mon. Weather Rev. 138, 1587–1612. https://doi.org/10.1175/2009MWR2968.1
Lima, F.J.L., Martins, F.R., Pereira, E.B., Lorenz, E., Heinemann, D., 2016. Forecast for surface solar irradiance at the Brazilian Northeastern region using NWP model and artificial neural networks. Renew. Energy 87, 807–818. https://doi.org/10.1016/j.renene.2015.11.005
Lin, W., Zhang, M., Wu, J., 2009. Simulation of low clouds from the CAM and the regional WRF with multiple nested resolutions. Geophys. Res. Lett. 36. https://doi.org/10.1029/2008GL037088
Liou, K.-N., 2002. An introduction to atmospheric radiation. Academic press.
Liou, K.-N., 1976. On the absorption, reflection and transmission of solar radiation in Cloudy atmospheres. J. Atmospheric Sci. 33, 798–805. https://doi.org/10.1175/1520-0469(1976)033<0798:OTARAT>2.0.CO;2
Liu, Z., 2017. Algorithm (2): Background Error Modeling and Estimation.
Löf, G.O.G., Duffie, J.A., Smith, C.O., 1966. World distribution of solar radiation. Sol. Energy 10, 27–37. https://doi.org/10.1016/0038-092X(66)90069-7
Long, C.., Dutton, E.., 2002. BSRN recommended QC tests V2.
López, G., Batlles, F.J., Tovar-Pescador, J., 2005. Selection of input parameters to model direct solar irradiance by using artificial neural networks. Energy 30, 1675–1684. https://doi.org/10.1016/j.energy.2004.04.035
Lorenc, A.C., 2003. The potential of the ensemble Kalman filter for NWP—a comparison with 4D-Var. Q. J. R. Meteorol. Soc. 129, 3183–3203. https://doi.org/10.1256/qj.02.132
Lorenc, A.C., Rawlins, F., 2005. Why does 4D-Var beat 3D-Var? Q. J. R. Meteorol. Soc. 131, 3247–3257. https://doi.org/10.1256/qj.05.85
Lorenz, E., Hurka, J., Heinemann, D., Beyer, H.G., 2009a. Irradiance Forecasting for the Power Prediction of Grid-Connected Photovoltaic Systems. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2, 2–10. https://doi.org/10.1109/JSTARS.2009.2020300
Lorenz, E., Kühnert, J., Heinemann, D., Nielsen, K.P., Remund, J., Müller, S.C., 2016. Comparison of global horizontal irradiance forecasts based on numerical weather prediction models with different spatio-temporal resolutions: Comparsion of irradiance forecasts based on NWP models. Prog. Photovolt. Res. Appl. 24, 1626–1640. https://doi.org/10.1002/pip.2799
Lorenz, E., Remund, J., Müller, S.C., Traunmüller, W., Steinmaurer, G., Pozo, D., Ruiz-Arias, J.A., Fanego, V.L., Ramirez, L., Romeo, M.G., others, 2009b. Benchmarking of different approaches to forecast solar irradiance, in: 24th European Photovoltaic Solar Energy Conference, Hamburg, Germany. p. 25.
Majewski, D., 2009. HRM_users_guide v2.5.
Majewski, D., Liermann, D., Prohl, P., Ritter, B., Buchhold, M., Hanisch, T., Paul, G., Wergen, W., Baumgardner, J., 2002. The operational global icosahedral-hexagonal gridpoint model GME: Description and high-resolution tests. Mon. Weather Rev. 130, 319–338.
Mallet, V., 2010. Ensemble forecast of analyses: Coupling data assimilation and sequential aggregation: ENSEMBLE FORECAST OF ANALYSES. J. Geophys. Res. Atmospheres 115. https://doi.org/10.1029/2010JD014259
Mallet, V., Stoltz, G., Mauricette, B., 2009. Ozone ensemble forecast with machine learning algorithms. J. Geophys. Res. 114. https://doi.org/10.1029/2008JD009978
Mao, Q., McNider, R.T., Mueller, S.F., Juang, H.-M.H., 1999. An optimal model output calibration algorithm suitable for objective temperature forecasting. Weather Forecast. 14, 190–202.

Marie-Joseph, I., Linguet, L., Gobinddass, M.-L., Wald, L., 2013. On the applicability of the Heliosat-2 method to assess surface solar irradiance in the Intertropical Convergence Zone, French Guiana. Int. J. Remote Sens. 34, 3012–3027. https://doi.org/10.1080/01431161.2012.756598
Marquez, R., Coimbra, C.F.M., 2013. Intra-hour DNI forecasting based on cloud tracking image analysis. Sol. Energy 91, 327–336. https://doi.org/10.1016/j.solener.2012.09.018
Marquez, R., Coimbra, C.F.M., 2012. Proposed Metric for Evaluation of Solar Forecasting Models. J. Sol. Energy Eng. 135, 011016. https://doi.org/10.1115/1.4007496
Marquez, R., Pedro, H.T.C., Coimbra, C.F.M., 2013. Hybrid solar forecasting method uses satellite imaging and ground telemetry as inputs to ANNs. Sol. Energy 92, 176–188. https://doi.org/10.1016/j.solener.2013.02.023
Mather, P.M., Koch, M., 2011. Computer processing of remotely-sensed images: an introduction, 4th ed. ed. Wiley-Blackwell, Chichester, West Sussex, UK ; Hoboken, NJ.
Mathiesen, P., Collier, C., Kleissl, J., 2013. A high-resolution, cloud-assimilating numerical weather prediction model for solar irradiance forecasting. Sol. Energy 92, 47–61. https://doi.org/10.1016/j.solener.2013.02.018
Mathiesen, P., Kleissl, J., 2011. Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States. Sol. Energy 85, 967–977. https://doi.org/10.1016/j.solener.2011.02.013
Mayer, B., Kylling, A., 2005. Technical note: The libRadtran software package for radiative transfer calculations - description and examples of use. Atmos Chem Phys 24.
Meteo France, 2017. DONNÉES DE MODÈLE ATMOSPHÉRIQUE À AIRE LIMITÉE À HAUTE RÉSOLUTION [WWW Document]. URL https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=131&id_rubrique=51
Milbrandt, J.A., Yau, M.K., 2005. A Multimoment Bulk Microphysics Parameterization. Part II: A Proposed Three-Moment Closure and Scheme Description. J. Atmospheric Sci. 62, 3065–3081. https://doi.org/10.1175/JAS3535.1
Milovac, J., Warrach-Sagi, K., Behrendt, A., Späth, F., Ingwersen, J., Wulfmeyer, V., 2016. Investigation of PBL schemes combining the WRF model simulations with scanning water vapor differential absorption lidar measurements: WRF Sensitivity to PBL Schemes and LSMs. J. Geophys. Res. Atmospheres 121, 624–649. https://doi.org/10.1002/2015JD023927
Ministry of the environment, 2016a. Projet_PPE_Guyane.
Ministry of the environment, 2016b. PPE_France.
Mitchell, H.L., HOUTEKAMER, P.L., PELLERIN, G.R., 2002. Ensemble Size, Balance, and Model-Error Representation in an Ensemble Kalman Filter. Mon. WEATHER Rev. 130, 18.
Miyoshi, T., Kadowaki, T., 2008. Accounting for Flow-dependence in the Background Error Variance within the JMA Global Four-dimensional Variational Data Assimilation System 4, 4.
Mlawer, E.J., Taubman, S.J., Brown, P.D., Iacono, M.J., Clough, S.A., 1997. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J. Geophys. Res. Atmospheres 102, 16663–16682. https://doi.org/10.1029/97JD00237
Moradi, I., Mueller, R., Alijani, B., Kamali, G.A., 2009. Evaluation of the Heliosat-II method using daily irradiation data for four stations in Iran. Sol. Energy 83, 150–156. https://doi.org/10.1016/j.solener.2008.07.010
Morcrette, J., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Doblas-Reyes, F., Hague, J., Weisheimer, A., 2007. Recent advances in radiation transfer parameterizations.
Morrison, H., 2010. An overview of cloud and precipitation microphysics and its parameterization in models. WRF Workshop 46.
Morrison, H., Thompson, G., Tatarskii, V., 2009. Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One- and Two-Moment Schemes. Mon. Weather Rev. 137, 991–1007. https://doi.org/10.1175/2008MWR2556.1

Möser, W., Raschke, E., 1984. Incident Solar Radiation over Europe Estimated from METEOSAT Data. J. Clim. Appl. Meteorol. 23, 166–170. https://doi.org/10.1175/1520-0450(1984)023<0166:ISROEE>2.0.CO;2
Müller, R., Pfeifroth, U., Träger-Chatterjee, C., Cremer, R., Trentmann, J., Hollmann, R., 2015. Surface Solar Radiation Data Set - Heliosat (SARAH) - Edition 1. https://doi.org/10.5676/EUM_SAF_CM/SARAH/V001
Muneer, T., Fairooz, F., 2002. Quality control of solar radiation and sunshine measurements – lessons learnt from processing worldwide databases. Build. Serv. Eng. Res. Technol. 23, 151–166. https://doi.org/10.1191/0143624402bt038oa
Myers, D.R., 2013. Solar radiation: practical modeling for renewable energy applications. CRC Press.
Nabat, P., Somot, S., Mallet, M., Chiapello, I., Morcrette, J.J., Solmon, F., Szopa, S., Dulac, F., Collins, W., Ghan, S., Horowitz, L.W., Lamarque, J.F., Lee, Y.H., Naik, V., Nagashima, T., Shindell, D., Skeie, R., 2013. A 4-D climatology (1979–2009) of the monthly tropospheric aerosol optical depth distribution over the Mediterranean region from a comparative evaluation and blending of remote sensing and model products. Atmospheric Meas. Tech. 6, 1287–1314. https://doi.org/10.5194/amt-6-1287-2013
Nakanishi, M., Niino, H., 2009. Development of an Improved Turbulence Closure Model for the Atmospheric Boundary Layer. J. Meteorol. Soc. Jpn. 87, 895–912. https://doi.org/10.2151/jmsj.87.895
Nandwani, S.S., 2006. Uses of solar energy in Costa Rica. Renew. Energy 31, 689–701. https://doi.org/10.1016/j.renene.2005.08.008
NCAR, N.C. for A.R., 2017. ARW Version 3.8 Modeling System User's Guide January 2017.
NCAR-RDA, 2017. Available GRIB Datasets from NCAR.
NCEP, 2018. NCEP ADP Global Upper Air and Surface Weather Observations (PREPBUFR format), May 1997 - Continuing.
Niu, G.-Y., Yang, Z.-L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., Xia, Y., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. J. Geophys. Res. 116. https://doi.org/10.1029/2010JD015139
NOAA, E.M. center, 2016. The Global Forecast System (GFS) - Global Spectral Model (GSM) [WWW Document]. URL http://www.emc.ncep.noaa.gov/GFS/doc.php
Noia, M., Ratto, C.F., Festa, R., 1993a. Solar irradiance estimation from geostationary satellite data: I. Statistical models. Sol. Energy 51, 449–456. https://doi.org/10.1016/0038-092X(93)90130-G
Noia, M., Ratto, C.F., Festa, R., 1993b. Solar irradiance estimation from geostationary satellite data: II. Physical models. Sol. Energy 51, 457–465. https://doi.org/10.1016/0038-092X(93)90131-7
Noilhan, J., Planton, S., 1989. A Simple Parameterization of Land Surface Processes for Meteorological Models. Mon. Weather Rev. 117, 536–549. https://doi.org/10.1175/1520-0493(1989)117<0536:ASPOLS>2.0.CO;2
O'Connor, P.A., Cleveland, C.J., 2015. U.S. Energy Transitions 1780–2010.
Ögelman, H., Ecevit, A., Tasdemiroğlu, E., 1984. A new method for estimating solar radiation from bright sunshine data. Sol. Energy 33, 619–625.
Oldewurtel, F., Parisio, A., Jones, C.N., Gyalistras, D., Gwerder, M., Stauch, V., Lehmann, B., Morari, M., 2012. Use of model predictive control and weather forecasts for energy efficient building climate control. Energy Build. 45, 15–27. https://doi.org/10.1016/j.enbuild.2011.09.022
Oreopoulos, L., Lee, D., Sud, Y.C., Suarez, M.J., 2012. Radiative impacts of cloud heterogeneity and overlap in an atmospheric General Circulation Model. Atmospheric Chem. Phys. 12, 9097–9111. https://doi.org/10.5194/acp-12-9097-2012
Otkin, J.A., Greenwald, T.J., 2008. Comparison of WRF Model-Simulated and MODIS-Derived Cloud Data. Mon. Weather Rev. 136, 1957–1970. https://doi.org/10.1175/2007MWR2293.1
Oumbe, A., Blanc, P., Wald, L., 2009. new method for estimating solar energy resource.

Pachauri, R.K., Meyer, L.A., 2015. Changements climatiques 2014: rapport de synthèse : contribution des Groupes de travail I, II et III au cinquième Rapport d'évaluation du Groupe d'experts intergouvernemental sur l'évolution du climat. GIEC, Genève (Suisse).

Parrish, D.F., Derber, J.C., 1992. The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. Mon. Weather Rev. 120, 1747–1763. https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2

Patil, A., HARIA, K., PASHTE, P., 2013. PHOTODIODE BASED PYRANOMETER. Int. J. Adv. Sci. Eng. Technol. 1, 5.

Pedro, H.T.C., Coimbra, C.F.M., 2012. Assessment of forecasting techniques for solar power production with no exogenous inputs. Sol. Energy 86, 2017–2028. https://doi.org/10.1016/j.solener.2012.04.004

Pelland, S., Remund, J., Kleissl, J., Oozeki, T., DeBrabandere, K., 2013. Photovoltaic and Solar Forecasting.

Pereira, E.B., Martins, F.R., Abreu, S.L., Couto, P., Stuhlmann, R., Colle, S., 2000. Effects of burning of biomass on satellite estimations of solar irradiation in Brazil. Sol. Energy 68, 91–107.

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renné, D., Hoff, T.E., 2010. Validation of short and medium term operational solar radiation forecasts in the US. Sol. Energy 84, 2161–2172. https://doi.org/10.1016/j.solener.2010.08.014

Perez, R., Lorenz, E., Pelland, S., Beauharnois, M., Van Knowe, G., Hemker, K., Heinemann, D., Remund, J., Müller, S.C., Traunmüller, W., Steinmauer, G., Pozo, D., Ruiz-Arias, J.A., Lara-Fanego, V., Ramirez-Santigosa, L., Gaston-Romero, M., Pomares, L.M., 2013. Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. Sol. Energy 94, 305–326. https://doi.org/10.1016/j.solener.2013.05.005

Perez, R., Seals, R., Stewart, R., Zelenka, A., Estrada-Cajigal, V., 1994. Using satellite-derived insolation data for the site/time specific simulation of solar energy systems. Sol. Energy 53, 491–495. https://doi.org/10.1016/0038-092X(94)90128-O

Pergaud, J., Masson, V., Malardel, S., Couvreux, F., 2009. A Parameterization of Dry Thermals and Shallow Cumuli for Mesoscale Numerical Weather Prediction. Bound.-Layer Meteorol. 132, 83–106. https://doi.org/10.1007/s10546-009-9388-0

Philippe Héritier, 2011. Le climat guyanais ; petit atlas climatique de la Guyane française, Météo-France. ed.

Pincus, R., Stevens, B., 2013. Paths to accuracy for radiation parameterizations in atmospheric models: PATHS TO PARAMETERIZATION ACCURACY. J. Adv. Model. Earth Syst. 5, 225–233. https://doi.org/10.1002/jame.20027

PINTY, J.-P., JABOUILLE, P., d'Aerologie, L., 1998. A MIXED-PHASE CLOUD PARAMETERIZATION FOR USE IN A MESOSCALE NON-HYDROSTATIC MODEL: SIMULATIONS OF A SQUALL LINE AND OF OROGRAPHIC PRECIPITATION 4.

Pleim, J.E., Xiu, A., 2003. Development of a Land Surface Model. Part II: Data Assimilation. J. Appl. Meteorol. 42, 1811–1822. https://doi.org/10.1175/1520-0450(2003)042<1811:DOALSM>2.0.CO;2

Polo, J., Antonanzas-Torres, F., Vindel, J.M., Ramirez, L., 2014. Sensitivity of satellite-based methods for deriving solar radiation to different choice of aerosol input and models. Renew. Energy 68, 785–792. https://doi.org/10.1016/j.renene.2014.03.022

Polo, J., Wilbert, S., Ruiz-Arias, J.A., Meyer, R., Gueymard, C., Súri, M., Martín, L., Mieslinger, T., Blanc, P., Grant, I., Boland, J., Ineichen, P., Remund, J., Escobar, R., Troccoli, A., Sengupta, M., Nielsen, K.P., Renne, D., Geuder, N., Cebecauer, T., 2016. Preliminary survey on site-adaptation techniques for satellite-derived and reanalysis solar radiation datasets. Sol. Energy 132, 25–37. https://doi.org/10.1016/j.solener.2016.03.001

Portabella, M., Stoffelen, A., 2004. A probabilistic approach for SeaWinds data assimilation: SEAWINDS DATA ASSIMILATION. Q. J. R. Meteorol. Soc. 130, 127–152. https://doi.org/10.1256/qj.02.205

Prabha, T., Hoogenboom, G., 2010. Evaluation of solar irradiance at the surface—inferences from in situ and satellite observations and a mesoscale model. Theor. Appl. Climatol. 102, 455–469. https://doi.org/10.1007/s00704-010-0329-0

Qu, Z., 2013. La nouvelle méthode Heliosat-4 pour l'évaluation du rayonnement solaire au sol. Ecole Nationale Supérieure des Mines de Paris.

Qu, Z., Oumbe, A., Wald, L., 2016. Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat‐4 method. Meteorol. Z. https://doi.org/10.1127/metz/2016/0781

Quesada-Ruiz, S., Chu, Y., Tovar-Pescador, J., Pedro, H.T.C., Coimbra, C.F.M., 2014. Cloud-tracking methodology for intra-hour DNI forecasting. Sol. Energy 102, 267–275. https://doi.org/10.1016/j.solener.2014.01.030

Radi, A., AL-KATHERI, A.A., DHANHANI, A., n.d. Tuning of WRF 3D-Var data assimilation system over Middle-East and Arabian Peninsula 10.

Raghavan, S.V., Vu, M.T., Liong, S.Y., 2015. Regional climate simulations over Vietnam using the WRF model. Theor. Appl. Climatol. https://doi.org/10.1007/s00704-015-1557-0

RÄISÄNEN, P., BARKER, H.W., COLE, J.N.S., 2005. The Monte Carlo Independent Column Approximation's Conditional Random Noise: Impact on Simulated Climate. J. Clim. 18, 16.

Raphael, C., Hay, J.E., 1984. An Assessment of Models which use Satellite Data to Estimate Solar Irradiance at the Earth's Surface. J. Clim. Appl. Meteorol. 23, 832–844. https://doi.org/10.1175/1520-0450(1984)023<0832:AAOMWU>2.0.CO;2

Raynaud, L., Berre, L., Desroziers, G., 2011. An extended specification of flow-dependent background error variances in the Météo-France global 4D-Var system. Q. J. R. Meteorol. Soc. 137, 607–619. https://doi.org/10.1002/qj.795

Reddy, K.., Ranjan, M., 2003. Solar resource estimation using artificial neural networks and comparison with other correlation models. Energy Convers. Manag. 44, 2519–2530. https://doi.org/10.1016/S0196-8904(03)00009-8

Remund, J., Perez, R., Lorenz, E., 2008. Comparison of solar radiation forecasts for the USA, in: Proc. of the 23rd European PV Conference. pp. 1–9.

Remund, J., Wald, L., Lefèvre, M., Ranchin, T., Page, J., 2003. Worldwide Linke turbidity information, in: ISES Solar World Congress 2003. International Solar Energy Society (ISES), p. p.13.

Ricchiazzi, P., Yang, S., Gautier, C., Sowle, D., 1998. SBDART: A Research and Teaching Software Tool for Plane-Parallel Radiative Transfer in the Earth's Atmosphere. Bull. Am. Meteorol. Soc. 79, 2101–2114. https://doi.org/10.1175/1520-0477(1998)079<2101:SARATS>2.0.CO;2

Richards, Arkin, 1980. on the relationship between satellite observed cloud cover and precipitation. Mon Weather Rev 109, 1081–1093.

Richardson, D., Bidlot, J., Vitart, F., 2013. Evaluation of ECMWF forecasts, including 2012–2013 upgrades [WWW Document]. URL https://www.ecmwf.int/en/elibrary/11921-evaluation-ecmwf-forecasts-including-2012-2013-upgrades

Rigollier, C., Bauer, O., Wald, L., 2000. On the clear sky model of the ESRA — European Solar Radiation Atlas — with respect to the heliosat method. Sol. Energy 68, 33–48. https://doi.org/10.1016/S0038-092X(99)00055-9

Rigollier, C., Lefèvre, M., Blanc, P., Wald, L., 2002a. The Operational Calibration of Images Taken in the Visible Channel of the Meteosat Series of Satellites. J. Atmospheric Ocean. Technol. 19, 1285–1293. https://doi.org/10.1175/1520-0426(2002)019<1285:TOCOIT>2.0.CO;2

Rigollier, C., Lefevre, M., Blanc, P., Wald, L., 2002b. The operational calibration of images taken in the visible channel of the meteosat series of satellites. J. Atmospheric Ocean. Technol. 19, 1285–1293.

Rigollier, C., Lefevre, M., Cros, S., Wald, L., 2002c. Heliosat 2: an improved method for the mapping of the solar radiation from Meteosat imagery, in: 2002 EUMETSAT Meteorological Satellite Conference. EUMETSAT, Darmstadt, Germany, pp. 585–592.

Rigollier, C., Lefèvre, M., Wald, L., 2004. The method Heliosat-2 for deriving shortwave solar radiation from satellite images. Sol. Energy 77, 159–169. https://doi.org/10.1016/j.solener.2004.04.017
Rigollier, C., Wald, L., 1998. Towards operational mapping of solar radiation from Meteosat images, HAL.
Rihan, F.A., Collier, C.G., Roulstone, I., 2005. Four-dimensional variational data assimilation for Doppler radar wind data. J. Comput. Appl. Math. 176, 15–34. https://doi.org/10.1016/j.cam.2004.07.003
Ritter, B., Geleyn, J.-F., 1992. A Comprehensive Radiation Scheme for Numerical Weather Prediction Models with Potential Applications in Climate Simulations. Mon. Weather Rev. 120, 303–325. https://doi.org/10.1175/1520-0493(1992)120<0303:ACRSFN>2.0.CO;2
Rizvi, S.R., 2016. WRFDA Background Error (Modeling and Estimation).
Rossow, W.B., Mekonnen, A., Pearl, C., Goncalves, W., 2013. Tropical Precipitation Extremes. J. Clim. 26, 1457–1466. https://doi.org/10.1175/JCLI-D-11-00725.1
Rossow, W.B., Schiffer, R.A., 1999. Advances in Understanding Clouds from ISCCP. Bull. Am. Meteorol. Soc. 80, 2261–2287. https://doi.org/10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2
Ruiz-Arias, J.A., Dudhia, J., Santos-Alamillos, F.J., Pozo-Vázquez, D., 2013. Surface clear-sky shortwave radiative closure intercomparisons in the Weather Research and Forecasting model: SW CLOSURE INTERCOMPARISONS IN WRF. J. Geophys. Res. Atmospheres 118, 9901–9913. https://doi.org/10.1002/jgrd.50778
Ruiz-Arias, J.A., Pozo-Vázquez, D., Sánchez-Sánchez, N., Montávez, J.P., Hayas-Barrú, A., Tovar-Pescador, J., 2008. Evaluation of two MM5-PBL parameterizations for solar radiation and temperature estimation in the South-Eastern area of the Iberian Peninsula. Il Nuovo Cimento C 31, 825–842.
Sahu, D.K., Kim, C.K., Zhong, X., Kleissl, J., 2016. Assimilating in-situ observations over Southern California for improved solar forecasting. IEEE, pp. 1007–1012. https://doi.org/10.1109/PVSC.2016.7749762
Sakov, P., Oke, P.R., 2008. A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. Tellus Dyn. Meteorol. Oceanogr. 60, 361–371. https://doi.org/10.1111/j.1600-0870.2007.00299.x
Sato, M., Hansen, J.E., McCormick, M.P., Pollack, J.B., 1993. Stratospheric aerosol optical depths, 1850–1990. J. Geophys. Res. Atmospheres 98, 22987–22994.
Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., Strugnell, N.C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R.P., Hu, B., Liang, S., Privette, J.L., Roy, D., 2002. First operational BRDF, albedo nadir reflectance products from MODIS. Remote Sens. Environ. 83, 135–148. https://doi.org/10.1016/S0034-4257(02)00091-3
Schaepman-Strub, G., Schaepman, M.E., Painter, T.H., Dangel, S., Martonchik, J.V., 2006. Reflectance quantities in optical remote sensing—definitions and case studies. Remote Sens. Environ. 103, 27–42. https://doi.org/10.1016/j.rse.2006.03.002
Schlesinger, M.E. (Ed.), 1988. Physically-Based Modelling and Simulation of Climate and Climatic Change. Springer Netherlands, Dordrecht.
Schmidt, T., Kalisch, J., Lorenz, E., Heinemann, D., 2015. Evaluating the spatio-temporal performance of sky imager based solar irradiance analysis and forecasts. Atmospheric Chem. Phys. Discuss. 15, 26997–27039. https://doi.org/10.5194/acpd-15-26997-2015
Schroedter-Homscheidt, M., Benedetti, A., Killius, N., 2017. Verification of ECMWF and ECMWF/MACC's global and direct irradiance forecasts with respect to solar electricity production forecasts. Meteorol. Z. 26, 1–19. https://doi.org/10.1127/metz/2016/0676
Schwartz, C.S., 2016. Improving Large-Domain Convection-Allowing Forecasts with High-Resolution Analyses and Ensemble Data Assimilation. Mon. Weather Rev. 144, 1777–1803. https://doi.org/10.1175/MWR-D-15-0286.1

Schwartz, C.S., Liu, Z., 2014. Convection-Permitting Forecasts Initialized with Continuously Cycling Limited-Area 3DVAR, Ensemble Kalman Filter, and “Hybrid” Variational–Ensemble Data Assimilation Systems. Mon. Weather Rev. 142, 716–738. https://doi.org/10.1175/MWR-D-13-00100.1

Schwartz, C.S., Liu, Z., Lin, H.-C., Cetola, J.D., 2014. Assimilating aerosol observations with a “hybrid” variational-ensemble data assimilation system: Hybrid aerosol data assimilation. J. Geophys. Res. Atmospheres 119, 4043–4069. https://doi.org/10.1002/2013JD020937

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., Masson, V., 2011. The AROME-France Convective-Scale Operational Model. Mon. Weather Rev. 139, 976–991. https://doi.org/10.1175/2010MWR3425.1

Sen, Z., 2008. Solar Energy Fundamentals and Modeling Techniques. Springer London, London.

Sengupta, M., Kurtz, S., Dobos, A., Wilbert, S., Lorenz, E., Renné, D., Myers, D., Wilcox, S., Blanc, P., Perez, R., 2015. Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications. IEA Solar Heating and Cooling Programme. https://doi.org/10.18777/ieashc-task46-2015-0001

Shafiee, S., Topal, E., 2009. When will fossil fuel reserves be diminished? Energy Policy 37, 181–189. https://doi.org/10.1016/j.enpol.2008.08.016

Shin, H.H., Hong, S.-Y., 2015. Representation of the Subgrid-Scale Turbulent Transport in Convective Boundary Layers at Gray-Zone Resolutions. Mon. Weather Rev. 143, 250–271. https://doi.org/10.1175/MWR-D-14-00116.1

Sidek, M.H.., Hasan, W.Z.., Kadir, M.Z.A.A., Shafie, S., Radzi, M.A.M., Ahmad, S.A., Marhaban, M.H., 2014. GPS based portable dual-axis solar tracking system using astronomical equation. IEEE, pp. 245–249. https://doi.org/10.1109/PECON.2014.7062450

Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., Huang, X., Duda, M., 2008. A Description of the Advanced Research WRF Version 3. https://doi.org/10.5065/D68S4MVH

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Wang, W., Powers, J.G., 2005. A description of the advanced research WRF version 2. DTIC Document.

Slingo, A., 1989. A GCM parameterization for the shortwave radiative properties of water clouds. J. Atmospheric Sci. 46, 1419–1427. https://doi.org/10.1175/1520-0469(1989)046<1419:AGPFTS>2.0.CO;2

Smith, C.J., Bright, J.M., Crook, R., 2017. Cloud cover effect of clear-sky index distributions and differences between human and automatic cloud observations. Sol. Energy 144, 10–21. https://doi.org/10.1016/j.solener.2016.12.055

Soubdhan, T., Emilion, R., Calif, R., 2009. Classification of daily solar radiation distributions using a mixture of Dirichlet distributions. Sol. Energy 83, 1056–1063. https://doi.org/10.1016/j.solener.2009.01.010

Sözen, A., Arcaklioğlu, E., Özalp, M., 2004. Estimation of solar potential in Turkey by artificial neural networks using meteorological and geographical data. Energy Convers. Manag. 45, 3033–3052. https://doi.org/10.1016/j.enconman.2003.12.020

Stamnes, K., Conklin, P., 1984. A new multi-layer discrete ordinate approach to radiative transfer in vertically inhomogeneous atmospheres. J. Quant. Spectrosc. Radiat. Transf. 31, 273–282. https://doi.org/10.1016/0022-4073(84)90031-1

Stamnes, K., Tsay, S.-C., Wiscombe, W., Jayaweera, K., 1988. Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. Appl. Opt. 27, 2502. https://doi.org/10.1364/AO.27.002502

Staniforth, A., White, A., Diamantakis, M., co-authors, 2006. Unified model documentation paper N°15 [WWW Document]. URL http://research.metoffice.gov.uk/research/nwp/publications/papers/unified_model/umdp15_v6.3.pdf

Stensrud, D., 2009. Parameterization schemes: keys to understanding numerical weather prediction models. Cambridge University Press, New York, NY.

Stephens, G., 1984. The parameterization of radiation for numerical weather prediction and climate models. Mon. Weather Rev. 112, 826–860.
Stephens, G.L., Gabriel, P.M., Partain, P.T., 2001. Parameterization of atmospheric radiative transfer. Part I: Validity of simple models. J. Atmospheric Sci. 58, 3391–3409. https://doi.org/10.1175/1520-0469(2001)058<3391:POARTP>2.0.CO;2
Stephens, G.L., Tsay, S.., Stackhouse, P.., Flatau, P., 1990. The Relevance of the Microphysical and Radiative Properties of Cirrus Clouds to Climate and Climatic Feedback. Am. Meteorol. Soc.
Stuhlmann, R., Rieland, M., Paschke, E., 1990. An Improvement of the IGMK Model to Derive Total and Diffuse Solar Radiation at the Surface from Satellite Data. J. Appl. Meteorol. 29, 586–603. https://doi.org/10.1175/1520-0450(1990)029<0586:AIOTIM>2.0.CO;2
Suárez, R.A., Abal, G., Siri, R., Musé, P., 2012. Brightness-dependent Tarpley model for global solar radiation estimation using GOES satellite images: Application to Uruguay. Sol. Energy 86, 3205–3215. https://doi.org/10.1016/j.solener.2012.08.012
Sukoriansky, S., Galperin, B., Perov, V., 2005. Application of a New Spectral Theory of Stably Stratified Turbulence to the Atmospheric Boundary Layer over Sea Ice'. Bound.-Layer Meteorol. 117, 231–257. https://doi.org/10.1007/s10546-004-6848-4
Sun, J., Wang, H., Tong, W., Zhang, Y., Lin, C.-Y., Xu, D., 2016. Comparison of the Impacts of Momentum Control Variables on High-Resolution Variational Data Assimilation and Precipitation Forecasting. Mon. Weather Rev. 144, 149–169. https://doi.org/10.1175/MWR-D-14-00205.1
Táczi, I., Szörényi, G., 2016. System Effects of Intermittent Renewable Generators (Wind, Solar)–Balancing.
Tanre, D., Herman, M., Deschamps, P.Y., de Leffe, A., 1979. Atmospheric modeling for space measurements of ground reflectances, including bidirectional properties. Appl. Opt. 18, 3587. https://doi.org/10.1364/AO.18.003587
Tapakis, R., Charalambides, A.G., 2013. Equipment and methodologies for cloud detection and classification: A review. Sol. Energy 95, 392–430.
Tarpley, J.D., 1979. Estimating incident solar radiation at the surface from geostationary satellite Data. J. Appl. Meteorol. 18, 1172–1181. https://doi.org/10.1175/1520-0450(1979)018<1172:EISRAT>2.0.CO;2
Tegen, I., Hollrig, P., Joyce, P., 1997. Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results.
Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., Bénard, P., Bochenek, B., Degrauwe, D., Derková, M., El Khatib, R., Hamdi, R., Mašek, J., Pottier, P., Pristov, N., Seity, Y., Smolíková, P., Španiel, O., Tudor, M., Wang, Y., Wittmann, C., Joly, A., 2018. The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1. Geosci. Model Dev. 11, 257–281. https://doi.org/10.5194/gmd-11-257-2018
Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M.A., 2004. Implementation and verification of the unified NOAH land surface model in the WRF model. Presented at the 20th conference on weather analysis and forecasting, pp. 11–15.
THEPAUT, J.-N., COURTIER, Phil., BELAUD, G., LEMAiTRE, G., 1996. Dynamical structure functions in a four-dimensional variational assimilation: A case study 27.
Thompson, G., Eidhammer, T., 2014. A Study of Aerosol Impacts on Clouds and Precipitation Development in a Large Winter Cyclone. J. Atmospheric Sci. 71, 3636–3658. https://doi.org/10.1175/JAS-D-13-0305.1
Thorey, J., Mallet, V., Chaussin, C., Descamps, L., Blanc, P., 2015. Ensemble forecast of solar radiation using TIGGE weather forecasts and HelioClim database. Sol. Energy 120, 232–243. https://doi.org/10.1016/j.solener.2015.06.049
Tiedtke, M., 1993. Representation of clouds in large scale models.
Tiedtke, M., 1989. A comprehensive mass flux scheme for cumulus parameterization in large-scale models.

Trieb, F., Schillings, C., O'Sullivan, M., Pregger, T., Hoyer-Klick, C., 2009. Global Potential of Concentrating Solar Power. SolarPaces Conf. Berl. Sept. 2009 11.

Troccoli, A., Morcrette, J.-J., 2014. Skill of Direct Solar Radiation Predicted by the ECMWF Global Atmospheric Model over Australia. J. Appl. Meteorol. Climatol. 53, 2571–2588. https://doi.org/10.1175/JAMC-D-14-0074.1

Tselioudis, G., 2002. Evaluation of midlatitude cloud properties in a weather and a climate model: Dependence on dynamic regime and spatial resolution. J. Geophys. Res. 107. https://doi.org/10.1029/2002JD002259

Tudor, M., Ivatek-ahdan, S., Stanei, A., Horvath, K., Baji, A., 2013. Forecasting Weather in Croatia Using ALADIN Numerical Weather Prediction Model, in: Ray, P. (Ed.), Climate Change and Regional/Local Responses. InTech. https://doi.org/10.5772/55698

Undén, P., co-authors, 2002. HIRLAM-5 Scientific Documentation.

USGS, 2018. Shuttle Radar Topography Mission (SRTM) [WWW Document]. URL https://lta.cr.usgs.gov/SRTM

Vallance, L., Charbonnier, B., Paul, N., Dubost, S., Blanc, P., 2017. Towards a standardized procedure to assess solar forecast accuracy: A new ramp and time alignment metric. Sol. Energy 150, 408–422. https://doi.org/10.1016/j.solener.2017.04.064

Vasquez, T., 2009. The Intertropical Convergence Zone [WWW Document]. Weather. Mag. URL http://www.weatherwise.org/Archives/Back%20Issues/2009/Nov-Dec%202009/full-Intertropical-Converge.html (accessed 11.21.16).

Verbois, H., Huva, R., Rusydi, A., Walsh, W., 2018. Solar irradiance forecasting in the tropics using numerical weather prediction and statistical learning. Sol. Energy 162, 265–277. https://doi.org/10.1016/j.solener.2018.01.007

Vermote, E.F., Tanré, D., Deuzé, J.L., Herman, M., Morcette, J.-J., 1997. Second simulation of the satellite signal in the solar spectrum, 6S: An overview. Geosci. Remote Sens. IEEE Trans. On 35, 675–686.

Verzijlbergh, R.A., Heijnen, P.W., de Roode, S.R., Los, A., Jonker, H.J.J., 2015. Improved model output statistics of numerical weather prediction based irradiance forecasts for solar power applications. Sol. Energy 118, 634–645. https://doi.org/10.1016/j.solener.2015.06.005

Vetra-Carvalho, D.S., 2018. The Ensemble Kalman filter.

Vetra-Carvalho, S., 2013. Properties of the Ensemble Kalman Filter for Convective-Scale Numerical Weather Prediction.

Wahab, M.A., El-Metwally, M., Hassan, R., Lefevre, M., Oumbe, A., Wald, L., 2009. Assessing surface solar irradiance in Northern Africa desert climate and its long-term variations from Meteosat images. Int. J. Remote Sens. 31, 261–280.

Wald, L., Blanc, P., Lefevre, M., Gschwind, B., 2011. The performances of the HelioClim databases in Mozambique, in: ISES Solar World Congress 2011. pp. 268–275.

Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T., Bushell, A., Copsey, D., Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Martin, G., Milton, S., Mittermaier, M., Morcrette, C., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P.L., Wilkinson, J., Willett, M., Woolnough, S., Xavier, P., 2017. The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations. Geosci Model Dev 34.

Wang, H., Huang, X.-Y., Sun, J., Xu, D., Zhang, M., Fan, S., Zhong, J., 2014. Inhomogeneous Background Error Modeling for WRF-Var Using the NMC Method. J. Appl. Meteorol. Climatol. 53, 2287–2309. https://doi.org/10.1175/JAMC-D-13-0281.1

Wang, X., Barker, D.M., Snyder, C., Hamill, T.M., 2008. A Hybrid ETKF–3DVAR Data Assimilation Scheme for the WRF Model. Part I: Observing System Simulation Experiment. Mon. Weather Rev. 136, 5116–5131. https://doi.org/10.1175/2008MWR2444.1

Wang, Z.J., Fidkowski, K., Abgrall, R., Bassi, F., Caraeni, D., Cary, A., Deconinck, H., Hartmann, R., Hillewaert, K., Huynh, H.T., Kroll, N., May, G., Persson, P.-O., van Leer, B., Visbal, M., 2013.

High-order CFD methods: current status and perspective: HIGH-ORDER CFD METHODS. Int. J. Numer. Methods Fluids 72, 811–845. https://doi.org/10.1002/fld.3767
Warner, T.T., 2010. Numerical Weather and Climate Prediction. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511763243
Welch, R.M., Cox, S.K., Davis, J.M., 1980. Solar radiation and clouds, Meteorological Monographs. Amer. Meteorol. Soc, Boston, MA.
Wheeler, M., Kiladis, G.N., 1999. Convectively Coupled Equatorial Waves: Analysis of Clouds and Temperature in the Wavenumber–Frequency Domain. J. Atmospheric Sci. 56, 374–399. https://doi.org/10.1175/1520-0469(1999)056<0374:CCEWAO>2.0.CO;2
White, A.T., Biazar, A.P., McNider, R.T., Doty, K., Khan, M., Dornblaser, B., 2016. Improving cloud prediction in WRF through the use of GOES satellite assimilation.
WMO, 2014. Status of WMO Forecasting Centres relative to Numerical Models for 2014.
WMO, 2007. WMO Guide to the Global Observing System.
Xie, B., Fung, J.C.H., Chan, A., Lau, A., 2012. Evaluation of nonlocal and local planetary boundary layer schemes in the WRF model: EVALUATION OF PBL SCHEMES IN WRF. J. Geophys. Res. Atmospheres 117, n/a-n/a. https://doi.org/10.1029/2011JD017080
Xie, Y., MacDonald, A.E., 2012. Selection of Momentum Variables for a Three-Dimensional Variational Analysis. Pure Appl. Geophys. 169, 335–351. https://doi.org/10.1007/s00024-011-0374-3
Xiu, A., Pleim, J.E., 2001. Development of a Land Surface Model. Part I: Application in a Mesoscale Meteorological Model. J. Appl. Meteorol. 40, 192–209. https://doi.org/10.1175/1520-0450(2001)040<0192:DOALSM>2.0.CO;2
Xu, H., Wang, Y., Wang, M., 2018. The Performance of a Scale-Aware Nonlocal PBL Scheme for the Subkilometer Simulation of a Deep CBL over the Taklimakan Desert. Adv. Meteorol. 2018, 1–12. https://doi.org/10.1155/2018/8759594
Xu, K.., Randall, D.., 1996. A semiempirical cloudiness parameterization for use in climate models.
Xue, L., Zhang, D., 2014. A multimodel data assimilation framework via the ensemble Kalman filter. Water Resour. Res. 50, 4197–4219. https://doi.org/10.1002/2013WR014525
Yang, H., Kleissl, J., 2016. Preprocessing WRF initial conditions for coastal stratocumulus forecasting. Sol. Energy 133, 180–193. https://doi.org/10.1016/j.solener.2016.04.003
Yang, J., Min, Q., Lu, W., Yao, W., Ma, Y., Du, J., Lu, T., Liu, G., 2015. An automated cloud detection method based on the green channel of total-sky visible images. Atmospheric Meas. Tech. 8, 4671–4679. https://doi.org/10.5194/amt-8-4671-2015
Yang, Q., Liu, X., Wu, W., Kizer, S., Baize, R.R., 2016. Fast and accurate hybrid stream PCRTM-SOLAR radiative transfer model for reflected solar spectrum simulation in the cloudy atmosphere. Opt. Express 24, A1514. https://doi.org/10.1364/OE.24.0A1514
YESSAD, K., 2015. BASICS ABOUT ARPEGE/IFS, ALADIN AND AROME IN THE CYCLE 42 OF ARPEGE/IFS.
Younes, S., Claywell, R., Muneer, T., 2005. Quality control of solar radiation data: Present status and proposed new approaches. Energy 30, 1533–1549. https://doi.org/10.1016/j.energy.2004.04.031
Yousif, C., Quecedo, G.O., Santos, J.B., 2013. Comparison of solar radiation in Marsaxlokk, Malta and Valladolid, Spain. Renew. Energy 49, 203–206. https://doi.org/10.1016/j.renene.2012.01.031
Zarzalejo, L.F., Ramirez, L., Polo, J., 2005. Artificial intelligence techniques applied to hourly global irradiance estimation from satellite-derived cloud index. Energy 30, 1685–1697. https://doi.org/10.1016/j.energy.2004.04.047
Zelenka, A., Perez, R., Seals, R., Renné, D., 1999. Effective accuracy of satellite-derived hourly irradiances. Theor. Appl. Climatol. 62, 199–207.
Zempila, Melas, Kazantzidis, Giannaros, 2015. Evaluation of Wrf shortwave radiation parametrization in predicting global horizontal irradiance in Greece.
Zhang, F., Shen, Z., Li, J., Zhou, X., Ma, L., 2013. Analytical Delta-Four-Stream Doubling–Adding Method for Radiative Transfer Parameterizations. J. Atmospheric Sci. 70, 794–808. https://doi.org/10.1175/JAS-D-12-0122.1

Zhang, J., Hodge, B.-M., Florita, A., Lu, S., Hamann, H.F., Banunarayanan, V., 2013. Metrics for evaluating the accuracy of solar power forecasting, in: 3rd International Workshop on Integration of Solar Power into Power Systems, London, England.

Zhang, Y., Liu, N., Oliver, D.S., 2009. Ensemble filter methods with perturbed observations applied to nonlinear problems. Comput. Geosci. 14, 249–261. https://doi.org/10.1007/s10596-009-9149-7

Zhong, X., Ruiz-Arias, J.A., Kleissl, J., 2016. Dissecting surface clear sky irradiance bias in numerical weather prediction: Application and corrections to the New Goddard Shortwave Scheme. Sol. Energy 132, 103–113. https://doi.org/10.1016/j.solener.2016.03.009

Zou, X., Navon, I.M., Ledimet, F.X., 1992. An Optimal Nudging Data Assimilation Scheme Using Parameter Estimation. Q. J. R. Meteorol. Soc. 118, 1163–1186. https://doi.org/10.1002/qj.49711850808

Zoulias: CRES, 2016. MAXIMIZATION OF THE PENETRATION OF RES IN ISLANDS.

Zubler, E.M., Lohmann, U., Lüthi, D., Schär, C., 2011. Intercomparison of aerosol climatologies for use in a regional climate model over Europe: INTERCOMPARISON OF AEROSOL CLIMATOLOGIES. Geophys. Res. Lett. 38, n/a-n/a. https://doi.org/10.1029/2011GL048081