



**HAL**  
open science

# Détection de zones brûlées après un feu de forêt à partir d'une seule image satellitaire SPOT 5 par techniques SVM

Olivier Zammit

► **To cite this version:**

Olivier Zammit. Détection de zones brûlées après un feu de forêt à partir d'une seule image satellitaire SPOT 5 par techniques SVM. Mathématiques [math]. Université Nice Sophia Antipolis, 2008. Français. NNT: . tel-00345683

**HAL Id: tel-00345683**

**<https://theses.hal.science/tel-00345683>**

Submitted on 9 Dec 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE NICE - SOPHIA ANTIPOLIS

## ÉCOLE DOCTORALE STIC

SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

# THÈSE

pour obtenir le titre de

## Docteur en Sciences

de l'Université de Nice-Sophia Antipolis

Mention : Automatique, Traitement du Signal et des Images

présentée et soutenue par

Olivier ZAMMIT

## DÉTECTION DE ZONES BRÛLÉES APRÈS UN FEU DE FORÊT À PARTIR D'UNE SEULE IMAGE SATELLITAIRE SPOT 5 PAR TECHNIQUES SVM

Thèse dirigée par Josiane ZERUBIA et Xavier DESCOMBES  
et préparée à l'INRIA Sophia Antipolis, dans le projet ARIANA

soutenue le 26 Septembre 2008

### Jury :

Joël LE ROUX	Professeur, UNSA	Président
Jocelyn CHANUSSOT	Professeur, INPG	Rapporteur
Lorenzo BRUZZONE	Professeur, Università degli Studi di Trento, Italie	Rapporteur
Michel DESHAYES	Enseignant-Chercheur, AgroParisTech	Examinateur
Jean-Guy BOUREAU	Ingénieur de Recherche, IFN	Examinateur
Hervé YÉSOU	Ingénieur de Recherche, SERTIT	Examinateur
Josiane ZERUBIA	Directrice de Recherche, INRIA	Directrice
Xavier DESCOMBES	Directeur de Recherche, INRIA	Co-encadrant

### Membres Invités :

Yves DESMAZIÈRES	Ingénieur, Infoterra France	Invité
Luc BARTHÉLÉMY	Ingénieur, AKKA Technologies	Invité



# Remerciements

La page des remerciements, celle que j'ai écrite en dernier mais qui figure en premier dans le manuscrit, celle qui permet de se remémorer ces trois années passées, celle qui rassemble les personnes qui m'ont aidé, soutenu et fait avancer.

Tout d'abord, je souhaite remercier mes directeurs de thèse Josiane Zerubia et Xavier Descombes qui m'ont accueilli dans leur projet. Merci à Josiane pour ses encouragements, son soutien, sa patience et pour m'avoir permis de faire de nombreuses rencontres scientifiques lors des démonstrations, séminaires et conférences. Merci à Xavier pour ses conseils et idées, et pour m'avoir guidé tout au long de ma thèse. J'espère avoir le plaisir de rencontrer à nouveau sur les terrains de rugby et lors des troisièmes mi-temps.

Je tiens aussi à remercier Joël Le Roux d'avoir accepté de présider mon jury de thèse, Jocelyn Chanussot et Lorenzo Bruzzone d'avoir accepté le travail consciencieux de rapporteur. Je remercie également Michel Deshayes, Jean-Guy Boureau et Luc Barthélémy pour leurs remarques pertinentes. Un grand merci à Hervé Yésou et Yves Desmazières pour avoir suivi régulièrement l'avancement de cette thèse, pour leurs nombreux conseils et commentaires ; et au Commandant Jean-Claude Poppi pour les discussions et échanges très instructifs lors de nos différentes rencontres.

Je remercie également AKKA Technologies (ex-SILOGIC) pour le co-financement de ma thèse, le CNES pour la fourniture d'images satellitaires via le programme ISIS et de cartographies ; Infoterra, le SERTIT, l'ONF, les Services Départementaux d'Incendie et de Secours du Var, des Alpes-Maritimes et de Haute-Corse pour la fourniture de cartographies de zones brûlées.

Je remercie chaleureusement les nombreux stagiaires, thésards, post-doctorants ou ingénieurs que j'ai rencontrés pendant ces trois années au sein du projet Ariana : Alexandre, Alexis, Avik, Aymen, Ayoub, Carlo, Caroline, Dan, Daniele, Éric, Fatih, Florent C, Florent L, Gabriel, Giovanni, Giuseppe, Grégory, Guillaume, Maria, Mats, Meena, Mikael, Nabil, Nadia, Neismon, Peter, Pierre, Praveen, Raffaele, Saloua, Sotiris, Stig, Ting, Vladimir... J'ajouterais des mentions spéciales à ceux qui m'ont soutenu tout au long de ma thèse : Pierre et Alexis, pour leurs conseils scientifiques et informatiques, leur enthousiasme et les thés ; et Alexandre, avec qui j'ai partagé bien plus qu'un bureau pendant trois ans. Je remercie aussi Ian et Laure, chercheurs permanents du projet qui m'ont accompagné tout

au long de ma thèse et Corinne qui m'a énormément facilité les démarches administratives.

Je remercie également mes autres collègues et amis de l'INRIA et de l'IS3 avec qui j'ai partagé des très agréables moments lors des parties de football, basket-ball ou autres activités sportives ou culturelles.

Et merci aussi à toute l'équipe du Riviera Rugby Club.

Merci à ma famille pour m'avoir permis d'arriver jusque là.

Merci à Céline.

# Table des matières

<b>Table des matières</b>	<b>iii</b>
<b>Table des figures</b>	<b>vii</b>
<b>Liste des tableaux</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Feux de Forêt . . . . .	1
1.1.1 Définition . . . . .	1
1.1.2 Généralités et Statistiques . . . . .	1
1.1.3 Conséquences des feux de forêt . . . . .	3
1.1.4 Pourquoi évaluer les dégâts après un feu de forêt ? . . . . .	3
1.2 Comment détecter les zones brûlées . . . . .	4
1.2.1 Méthodes manuelles . . . . .	4
1.2.2 Intérêt de la télédétection . . . . .	5
1.2.3 Méthodes automatiques utilisant la télédétection . . . . .	9
1.3 Méthode proposée et son intérêt . . . . .	14
<b>2 Classification</b>	<b>17</b>
2.1 Techniques de classification classiques . . . . .	17
2.1.1 K-moyennes . . . . .	17
2.1.2 K-Plus Proches Voisins . . . . .	19
2.1.3 Maximum de Vraisemblance . . . . .	19
2.2 SVM . . . . .	20
2.2.1 Cas linéairement séparable . . . . .	21
2.2.2 Cas non-linéairement séparable . . . . .	23
2.2.3 Quelques noyaux . . . . .	26
2.2.4 One-Class SVM . . . . .	27
2.2.5 Implantation . . . . .	29
2.3 Conclusion sur les techniques SVM . . . . .	31
<b>3 Expérimentation et comparaison</b>	<b>33</b>
3.1 Test sur des "exemples jouets" . . . . .	33
3.2 Test sur des images SPOT 5 . . . . .	36
3.2.1 Réglage des SVM . . . . .	36

3.2.2	Réglage des OC-SVM . . . . .	41
3.2.3	Résultats des différentes techniques de classification . . . . .	43
3.3	Analyse des erreurs . . . . .	46
3.4	Conclusion . . . . .	52
<b>4</b>	<b>Post-traitement</b>	<b>55</b>
4.1	Introduction . . . . .	56
4.2	Régularisation par morphologie mathématique . . . . .	56
4.2.1	Présentation de la méthode . . . . .	56
4.2.2	Expérimentation et comparaison . . . . .	57
4.3	Régularisation par Champs de Markov . . . . .	62
4.3.1	Présentation de la méthode . . . . .	62
4.3.2	Expérimentation et comparaison . . . . .	64
4.4	Technique de croissance de régions . . . . .	66
4.4.1	Présentation de la méthode . . . . .	66
4.4.2	Expérimentation et comparaison . . . . .	69
4.5	Conclusion . . . . .	77
<b>5</b>	<b>Sélection de la base d'apprentissage</b>	<b>79</b>
5.1	Peut-on garder la même base d'apprentissage avec une autre image? . . . . .	79
5.2	Sélection par K-moyennes . . . . .	80
5.2.1	Présentation de la méthode . . . . .	80
5.2.2	Expérimentation et comparaison . . . . .	80
5.3	Sélection à partir de l'histogramme par ligne de partage des eaux . . . . .	85
5.3.1	Présentation de la méthode . . . . .	86
5.3.2	Expérimentation et comparaison . . . . .	86
5.4	Conclusion . . . . .	89
	<b>Conclusion et Perspectives</b>	<b>97</b>
	<b>Annexes :</b>	<b>100</b>
<b>A</b>	<b>Séparabilité spectrale des zones brûlées</b>	<b>101</b>
A.1	Utilisation d'une seule bande spectrale . . . . .	101
A.2	Utilisation de deux bandes spectrales . . . . .	102
A.3	Utilisation d'au moins trois bandes spectrales . . . . .	102
A.4	Conclusion . . . . .	104
<b>B</b>	<b>Algorithme SMO</b>	<b>107</b>
B.1	Optimisation analytique . . . . .	107
B.2	Choix des multiplicateurs à optimiser . . . . .	109
B.3	Mise à jour des paramètres . . . . .	110
<b>C</b>	<b>Champs de Markov</b>	<b>111</b>
C.1	Définitions . . . . .	111
C.2	Propriétés d'un champ de Markov . . . . .	111

<b>D Morphologie mathématique</b>	<b>113</b>
D.1 Opérations de base de la morphologie mathématique . . . . .	113
D.2 Ligne de partage des eaux . . . . .	116
<b>E Algorithme ACP</b>	<b>119</b>
E.1 Description de la méthode . . . . .	119
E.2 Application aux images satellitaires . . . . .	121
E.3 Conclusion . . . . .	122
<b>F Calcul de la précision</b>	<b>127</b>
F.1 Partition de l'image . . . . .	127
F.2 Taux de classification . . . . .	128
F.3 Conclusion . . . . .	129
<b>G Cas particulier des OC-SVM</b>	<b>131</b>
G.1 Calcul de l'hypersphère englobant la plupart des données et de taille minimale	131
G.2 Comparaison avec les OC-SVM . . . . .	132
<b>Glossaire</b>	<b>135</b>
<b>Bibliographie</b>	<b>137</b>
<b>Publications et activités de recherche</b>	<b>143</b>
<b>Résumé</b>	<b>145</b>
<b>Abstract</b>	<b>147</b>





# Table des figures

1.1	Réflectance . . . . .	7
1.2	Est de la Corse : Image SPOT 5, bandes spectrales et estimation satellitaire . . . . .	8
1.3	Est de la Corse : Indices de végétation et seuillages . . . . .	10
2.1	SVM : cas linéaire . . . . .	24
2.2	SVM : cas des marges relaxées . . . . .	24
2.3	Rôle de la fonction $\Phi$ . . . . .	25
2.4	SVM avec un noyau Gaussien . . . . .	27
2.5	One-Class SVM . . . . .	28
3.1	Base d'apprentissage et base de test pour un mélange de gaussiennes . . . . .	34
3.2	Résultat de la classification pour les quatre algorithmes . . . . .	34
3.3	Choix des hyperparamètres des SVM . . . . .	39
3.4	Validation croisée. . . . .	41
3.5	Choix des hyperparamètres des OC-SVM. . . . .	42
3.6	Fonctions de décisions des OC-SVM en fonction du couple de paramètres . . . . .	44
3.7	Massif des Maures - 2003 ( 2500 hectares) . . . . .	47
3.8	Ouest de la Corse - 2005 (environ 2800 hectares) . . . . .	48
3.9	Est de la Corse - 2005 (environ 1000 hectares) . . . . .	49
3.10	Alpes Maritimes - 2007 (environ 450 hectares) . . . . .	50
4.1	Régularisation par morphologie mathématique . . . . .	57
4.2	Précision Totale avant et après régularisation par morphologie mathématique . . . . .	58
4.3	Massif des Maures - 2003 (environ 2500 hectares) . . . . .	59
4.4	Est de la Corse - 2005 (environ 1000 hectares) . . . . .	60
4.5	Régularisation par champs de Markov . . . . .	64
4.6	Ouest de la Corse - 2005 ( 2800 ha) . . . . .	65
4.7	Est de la Corse - 2005 ( 1000 ha) . . . . .	65
4.8	Alpes-Maritimes - 2007 (environ 450 ha) . . . . .	65
4.9	Lignes de niveau pour les OC-SVM . . . . .	69
4.10	Cannes-Mandelieu - 2007 (environ 450 ha) . . . . .	71
4.11	Ouest de la Corse - 2005 (environ 2800 ha) . . . . .	72
4.12	Est de la Corse - 2005 (environ 1000 ha) . . . . .	73
4.13	Lucéram - 2003 (50 hectares) . . . . .	74
4.14	Est de la Corse - courbes ROC . . . . .	75

4.15	Alpes-Maritimes - courbes ROC . . . . .	75
5.1	Nord de la Corse (1000 ha) . . . . .	81
5.2	Est de la Corse (1000 ha) . . . . .	82
5.3	Massif des Maures - 2003 (2500 hectares) . . . . .	83
5.4	Nord de la Corse (1000 hectares) . . . . .	84
5.5	Sections d'histogrammes lors de différentes étapes . . . . .	87
5.6	Ouest de la Corse, 2005 (environ 2800 ha) . . . . .	88
5.7	Sections d'histogrammes lors de différentes étapes . . . . .	90
5.8	Alpes Maritimes, 2007 (environ 450 ha) . . . . .	91
5.9	Est de la Corse, 2005 (environ 1000 ha) . . . . .	92
5.10	Alpes Maritimes, 2007 (environ 7 ha) . . . . .	93
A.1	Histogrammes des zones brûlées et non-brûlées pour les 4 canaux de SPOT 5 pour 4 images différentes . . . . .	103
A.2	Scattergrammes des zones brûlées et non-brûlées en fonction de différentes bandes . . . . .	104
C.1	Systèmes de Voisinages et cliques associées en dimension 2 . . . . .	112
D.1	Opérations de base de morphologie mathématique. . . . .	115
D.2	Ligne de Partage des Eaux . . . . .	117
D.3	Algorithme de Ligne de Partage des Eaux . . . . .	117
E.1	Analyse en Composantes Principales d'un nuage de points de dimension 2 . . . . .	119
E.2	Bandes de "Nord de la Corse" © CNES 2003, Distribution SPOT Image . . . . .	123
E.3	Composantes Principales de "Nord de la Corse" . . . . .	123
E.4	Histogrammes des 4 bandes de l'image "Nord de la Corse" . . . . .	124
E.5	Histogrammes des 4 Composantes Principales de "Nord de la Corse" . . . . .	124
E.6	Bandes de l'image "Ouest de la Corse" © CNES 2005, Distribution SPOT Image . . . . .	125
E.7	Composantes Principales de l'image "Ouest de la Corse" . . . . .	125
E.8	Histogrammes des 4 bandes de l'image "Ouest de la Corse" . . . . .	126
E.9	Histogrammes des 4 Composantes Principales de l'image "Ouest de la Corse" . . . . .	126
G.1	Cas particulier des "One-Class SVM" . . . . .	133

# Liste des tableaux

1.1	Statistiques sur les feux de forêt . . . . .	2
3.1	Moyenne, variance de la précision et temps de calcul des différentes techniques pour 100 expérimentations . . . . .	35
3.2	Données satellitaires et cartes de référence . . . . .	37
3.3	Types de végétation . . . . .	37
3.4	Informations sur les images et bases d'apprentissage (BA) . . . . .	37
3.5	Performances des SVM pour différents noyaux . . . . .	38
3.6	Choix des hyperparamètres des OC-SVM . . . . .	42
3.7	Massif des Maures - 2003 : Comparaison avec la carte de référence . . . . .	51
3.8	Ouest de la Corse - 2005 : Comparaison avec la carte de référence . . . . .	51
3.9	Est de la Corse - 2005 : Comparaison avec la carte de référence . . . . .	51
3.10	Alpes Maritimes - 2007 : Comparaison avec la carte de référence . . . . .	51
4.1	Massif des Maures - 2003 : Comparaison avec la carte de référence . . . . .	61
4.2	Ouest de la Corse - 2005 : Comparaison avec la carte de référence . . . . .	61
4.3	Est de la Corse - 2005 : Comparaison avec la carte de référence . . . . .	61
4.4	Alpes Maritimes - 2007 : Comparaison avec la carte de référence . . . . .	61
4.5	Régularisation par champs de Markov des classifications par OC-SVM . . . . .	66
4.6	Croissance de régions des OC-SVM . . . . .	76
5.1	Massif des Maures - 2003 : Comparaison avec la carte de référence . . . . .	83
5.2	Ouest de la Corse - 2005 : Comparaison avec la carte de référence . . . . .	95
5.3	Est de la Corse - 2005 : Comparaison avec la carte de référence . . . . .	95
5.4	Alpes Maritimes - 2007 : Comparaison avec la carte de référence . . . . .	95
A.1	Valeurs de l'indice de séparabilité $M$ . . . . .	103
E.1	ACP de l'image "Nord de la Corse" . . . . .	122
E.2	ACP de l'image "Corse Ouest" . . . . .	122



# Chapitre 1

## Introduction

### 1.1 Feux de Forêt

#### 1.1.1 Définition

La dénomination exacte de "feux de forêt" concerne normalement les incendies qui se propagent sur une surface minimale d'un hectare, d'un seul tenant. Néanmoins, nous ne ferons pas de restriction sur la taille des zones brûlées.

La notion de "feux de forêt" ne se rapporte pas seulement aux forêts au sens strict mais aussi aux formations subforestières de petite taille : le maquis (formation végétale dense, plus basse qu'une forêt, sur sol siliceux), la garrigue (formation végétale plutôt ouverte sur sol calcaire, composée principalement d'arbrisseaux) et les landes (formations végétales sur sols acides, composées de genêts et de petits arbustes).

#### 1.1.2 Généralités et Statistiques

Les feux de forêt font partie des risques majeurs dans les écosystèmes forestiers. Chaque année, de nombreux hectares partent en fumée. Dans les pays méditerranéens, la période de l'année la plus propice aux feux de forêt est l'été à cause de la sécheresse et d'une faible teneur en eau des sols. Le Tableau 1.1 fournit quelques statistiques : superficie brûlée totale en hectares et nombre de feux concernant les 5 pays européens les plus touchés [Schmuck et al., 2007].

Ces statistiques montrent des fluctuations différentes entre les pays. Les conditions climatiques exceptionnelles de l'été 2003 ont largement contribué au bilan dramatique cette année là.

En France, les feux de petite taille (moins de 5 ha) représentent 90 % du nombre d'éclosions pour 8 % de la surface incendiée. En contrepartie, les feux de grande taille (plus de 50 ha) représentent 2 % du nombre d'éclosions pour 77 % de la surface incendiée, d'après la base de données Prométhée <sup>1</sup>.

---

<sup>1</sup>La base de données Prométhée (consultable sur <http://www.promethee.com>) recense les statistiques sur les feux (lieu et date du sinistre, surface brûlée, types de végétation...) identifiés par les sapeurs-pompiers depuis 1973 pour 15 départements du sud-est de la France (zones les plus vulnérables en France).

Année	Portugal	Espagne	France	Italie	Grèce
Surface totale des zones brûlées (en hectares)					
Moyenne (1980-1989)	73 484	244 788	39 157	147 150	52 417
Moyenne (1990-1999)	102 203	161 319	22 735	118 573	44 108
2000	159 605	188 586	24 078	114 678	145 033
2001	111 850	93 297	20 642	76 427	18 221
2002	124 411	107 464	30 160	40 791	6 013
2003	425 726	148 172	73 278	91 805	3 517
2004	129 539	134 193	13 711	60 176	10 267
2005	338 262	188 697	22 135	47 575	6 437
Nombre de feux					
Moyenne (1980-1989)	7 381	9 515	4 910	11 575	1 264
Moyenne (1990-1999)	22 250	18 152	5 538	11 164	1 748
2000	34 109	24 118	4 603	8 595	2 581
2001	26 533	19 547	4 309	7 134	2 535
2002	26 488	19 929	4 097	4 601	1 141
2003	26 195	18 616	7 023	9 697	1 452
2004	21 870	21 394	3 775	6 428	1 748
2005	35 697	25 492	4 698	7 951	1 544

TAB. 1.1: Statistiques sur les feux de forêt (source : Commission Européenne - <http://forest.jrc.it>)

### 1.1.3 Conséquences des feux de forêt

Les feux de forêt peuvent s'avérer bénéfiques à la santé de nombreux écosystèmes forestiers. Le brûlage dirigé et l'écobuage (débroussaillage par le feu) constituent deux opérations d'aménagement et d'entretien utilisées dans le monde entier.

En effet, des insectes, des champignons et des bactéries en tous genres peuvent proliférer dans les amas de feuillage et de branches mortes qui recouvrent le sol. L'incendie de forêt (dans des conditions modérées) est un moyen naturel pour stériliser le sol, libérer de l'espace pour la croissance des nouveaux plants et stimuler la croissance des plantes herbacées.

De plus, sans feux de forêt, certaines espèces d'arbres et certains écosystèmes de la forêt boréale seraient condamnés (les pommes de certains pins ne s'ouvrent pour libérer leur graines qu'à des températures supérieures à 100°C).

Un incendie n'est destructif qu'en période de sécheresse lors des saisons sèches. Dans ce cas, il est très coûteux, tant au niveau des moyens matériels et humains mis en oeuvre, que des conséquences environnementales et économiques qui en découlent [Colin et al., 2001].

Les feux de forêt ont des conséquences immédiates, mais également à plus long terme. Les premiers effets découlent directement du brûlage du combustible ou de la chaleur. Il s'agit de la destruction, partielle ou totale, de la végétation et de la disparition de la faune. Le sol est lui aussi affecté par une perte d'éléments minéraux (fertilisants, azote) et la diminution de ses réserves d'eaux.

Les effets secondaires résultants des incendies sont aussi destructifs que les effets primaires. La fumée et les cendres qu'ils produisent peuvent causer des problèmes de santé. De plus, les régions méditerranéennes sont constituées par des terrains fortement pentus et sont touchées par de fortes pluies en automne. Après un incendie, il y a un risque important d'érosion ou de coulée de boue, consécutif à l'augmentation du ruissellement sur un sol dénudé de sa couverture végétale.

Tous les arbres touchés dans la forêt ne sont pas morts. Depuis quelques années, le bois brûlé est ramassé pour les raisons suivantes :

- il peut être facilement emporté par le vent et représente un danger,
- la coupe de branches brûlées mais pas mortes améliore le rétablissement de la forêt,
- même si elles ont perdu en valeur, ces branches peuvent être utilisées comme combustible,
- dans le cas de terrain fortement pentu, la disposition du bois horizontalement permet de diminuer le risque d'érosion.

### 1.1.4 Pourquoi évaluer les dégâts après un feu de forêt ?

La cartographie des zones brûlées est un maillon essentiel dans la gestion de la situation post-crise. Elle fournit des renseignements précieux pour tous les acteurs responsables de la gestion de ces territoires (Ministère de l'Agriculture, Préfectures, Services Départementaux d'Incendie et de Secours, Collectivités Locales, Inventaire Forestier National, Office National des Forêts, etc.).



Son importance réside, d'abord, dans l'inventaire des dégâts afin d'évaluer les pertes écologiques et économiques. L'inventaire permet l'estimation et l'allocation des ressources (matérielles, humaines et financières) nécessaires à la mise en place d'un programme de restauration et de réhabilitation des zones touchées par les feux [Ezzine et al., 2005].

Des opérations de reboisement peuvent être envisagées afin de reconstituer la forêt. Elles ne sont pas systématiques, mais permettent de restabiliser le sol et ainsi d'éviter une possible érosion. Pour le reboisement, des peuplements moins combustibles, par leur structure et leur composition, sont privilégiés.

D'autre part, la cartographie des zones brûlées permet la localisation spatiale de la région sinistrée. Elle facilite l'analyse des différents facteurs impliqués dans le processus de feu, contribuant ainsi à la compréhension du comportement de l'incendie (qui dépend du combustible, des conditions météorologiques et de la topographie de la région) et aide à l'identification de la cause de l'incendie. Sur un plus long terme, elle permet de suivre l'évolution des zones brûlées, qui présentent un risque d'inflammabilité important après repousse et contribue à l'élaboration et à la mise à jour des plans de prévention des risques. Elle constitue donc un outil important d'aide à la gestion des futurs incendies dans une région donnée.

De plus, les propriétaires des habitations, des terrains endommagés ou détruits par un incendie peuvent réclamer des indemnités à leurs compagnies d'assurance. Les compagnies d'assurance doivent avoir recours aux bases de données concernant l'étendue de l'incendie pour identifier rapidement qui a été touché et ce qui peut donner lieu à des compensations financières.

## 1.2 Comment détecter les zones brûlées

### 1.2.1 Méthodes manuelles

Les méthodes traditionnelles de cartographie des zones brûlées sont généralement obtenues par relevé du contour sur le terrain. Les contours des zones brûlées sont généralement établis à partir d'un certain nombre de points référencés avec des systèmes GPS (Global Positioning System) [Henry, 2008]. Ces méthodes sont coûteuses et pas forcément très précises. En effet, les zones forestières vastes sont souvent inaccessibles par la route et/ou localisées sur des montagnes de fortes pentes, il faut donc avoir recours à un hélicoptère. De plus, les relevés de contours ne prennent généralement pas en compte les petites zones de végétation épargnées par le feu qui se situent à l'intérieur de la zone brûlée [Garcia and Chuvieco, 2004].

Dans le cas de grand feux, les contours peuvent être obtenus par interprétation visuelle d'images aériennes [Deshayes, 2004].

Plus récemment, de nouvelles expérimentations sont testées par les acteurs opérationnels de la lutte contre les incendies : une caméra couplée à un système GPS et associée à un Système d'Information Géographique (SIG) à bord d'un avion Cessna ou d'un drone permettent de cartographier le sinistre.

### 1.2.2 Intérêt de la télédétection

La télédétection désigne l'ensemble des techniques qui permettent d'obtenir des informations sur des objets ou des phénomènes, par l'intermédiaire d'instruments de mesure (un satellite par exemple), sans contact direct avec les cibles étudiées, en utilisant les propriétés des ondes qu'ils émettent ou réfléchissent.

De nombreux domaines utilisent la télédétection : la cartographie, l'océanographie, l'agriculture, la sylviculture, l'environnement, l'écologie, la météorologie...

Elle peut contribuer à la gestion des incendies de forêts au cours des trois phases : avant (prévention, gestion du risque) [Chuvienco, 2003], pendant (gestion de la crise) [Flannigan and Vonder Haar, 1986] et après l'incendie (estimation des dégâts et réhabilitation) [Gonzalez-Alonzo et al., 2004].

Concernant la troisième phase, l'utilisation de la télédétection par satellite permet de réduire le coût et le temps nécessaire à l'évaluation des dégâts après un feu de forêt. De plus, elle fournit périodiquement et automatiquement des informations, sur de très grandes surfaces et sur plusieurs bandes spectrales.

Quand les rayons du soleil frappent un objet, certaines longueurs d'onde du spectre sont réfléchies, d'autres sont absorbées. Les différents objets n'absorbent pas tous la même partie du rayonnement spectral. En conséquence, le spectre du rayonnement émis diffère selon l'objet. La télédétection mesure le rayonnement réfléchi par les objets et permet donc de déterminer certaines propriétés des objets. Elle utilise le même principe que la vision humaine : exploiter la couleur et parfois la texture, pour identifier les objets, mais son champ d'analyse est élargi à des parties du spectre électromagnétique allant au-delà du domaine de la lumière visible.

Par exemple, les capteurs utilisés par le satellite SPOT 5 peuvent mesurer le rayonnement réfléchi dans le domaine visible (de 0.50 à 0.68  $\mu m$ ), du proche et moyen infra-rouge (0.78 à 0.89  $\mu m$  et de 1.58 à 1.75  $\mu m$ ).

La Figure 1.1 donne la réflectance (proportion de rayonnement réfléchi) en fonction de la longueur d'onde pour 3 types d'objet : la végétation (alfafla et érable), le sol (sec et humide) et l'eau (turbide et claire).

Dans les bandes Bleu et Rouge, le rayonnement est en grande partie absorbé par les pigments de feuilles vertes (la chlorophylle) pour la photosynthèse (utilisation de l'énergie solaire pour la synthèse de matière organique) et il est plus réfléchi dans la bande Vert, ce qui explique pourquoi les feuilles des arbres nous apparaissent en vert.

Le satellite SPOT 5 capte le rayonnement dans les bandes Vert, Rouge, Proche Infra-Rouge (PIR) et Moyen Infra-Rouge (MIR), mais pas dans la bande Bleu. Dans une image satellitaire fausse couleur, le canal Vert est habituellement représenté en nuance de bleu, le Rouge en nuance de vert et l'Infra-Rouge en rouge. Ainsi, sur l'image satellitaire, la végétation apparaît en rouge (forte réflectance dans l'Infra-Rouge), la mer en bleu (plus grande réflectance dans le Vert), les nuages en blanc (grande réflectance dans les 3 bandes). Le sol, les routes et les zones urbaines apparaissent plutôt en gris, plus ou moins clair selon leur composition.

Par exemple, la Figure 1.2 donne la composition fausse couleur d'une image multi-spectrale SPOT 5 (PIR, Rouge et Vert) de résolution 10 mètres, les 4 bandes séparément et la carte de référence correspondante (cartographie des zones brûlées réalisée par Infoterra - ESA) pour la partie Est de la Corse.

Dans la bande Rouge, la végétation apparaît bien en foncé (un peu moins dans la bande Vert). Dans la bande Proche Infra-Rouge, la végétation apparaît plus claire (haute réflectance dans le PIR, faible dans le Rouge). En comparant les réflectances dans le Rouge et l'Infra-Rouge, il est alors possible de mesurer la quantité relative de végétation. Généralement, si la réflectance est grande dans l'IR, la végétation est plus dense et est composée de forêt. Par contre, si elle est faible, la végétation est alors plus probablement composée de végétation moins dense, comme du maquis.

Les zones brûlées apparaissent en gris foncé. En effet, la végétation saine, en pleine activité chlorophyllienne réfléchit une très grande partie du spectre infra-rouge. Une végétation brûlée présente des réflectances faibles dans le domaine du visible et de l'infra-rouge, en raison de la présence de carbone et de cendres.

### Indices de végétation

Les *indices de végétation*, très largement utilisés en télédétection, donnent une estimation de variables géophysiques : couverture de la végétation (densité, santé), teneur en eau des sols, etc. Ils sont obtenus par un calcul faisant généralement intervenir deux bandes spectrales. Ils permettent ainsi d'estimer rapidement et simplement des variables environnementales.

Le NDVI ("Normalized Difference Vegetation Index") [Rouse et al., 1973] est l'indice le plus connu en télédétection. Il est calculé à partir des bandes Rouge (R) et Proche Infra-Rouge (PIR). La végétation saine absorbe une grande quantité de lumière Rouge et réfléchit une grande quantité de lumière IR. Au contraire, une végétation moins saine réfléchira plus de Rouge et absorbera plus d'IR. Cet indice est fondé sur la différence de réflectances entre ces deux bandes et est normalisé :

$$NDVI = \frac{NIR - R}{NIR + R}$$

La plupart des satellites permettent de calculer cet indice afin de quantifier la densité et la vigueur de la végétation sur une zone donnée. Le NDVI fournit une valeur entre -1 et +1. Une valeur de NDVI nulle indique une faible densité de végétation saine, une valeur proche de 1 indique une grande densité de végétation saine. Les nuages, l'eau, la neige, le sol et les roches ont généralement des valeurs de NDVI proches de 0 (même réflectance dans le Rouge et l'IR).

Le NDVI est très utile pour la surveillance de la végétation à grande échelle puisqu'il est normalisé. Ainsi, il subit moins l'influence des changements de conditions d'illumination, de pentes de terrain et d'angle d'incidence de la visée satellitaire. Il est ainsi largement utilisé pour détecter les variations de biomasse et les zones brûlées.

Il existe d'autres indices de végétation :

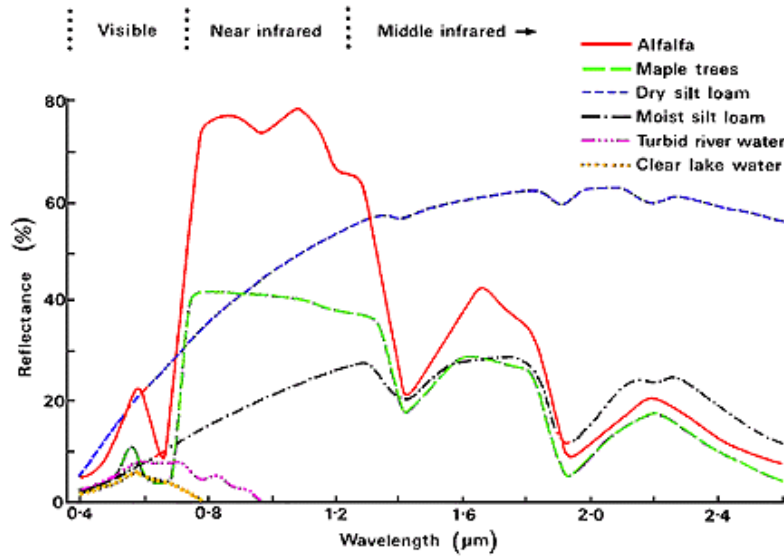


FIG. 1.1: Réflectance - ref. : <http://ceos.cnes.fr>

- le SAVI ("Soil Adjusted Vegetation Index") [Huete, 1988], qui a démontré une plus grande sensibilité dans le cas de végétation éparse. Sa formule ressemble à celle du NDVI mais comprend un facteur qui permet d'ajuster la quantité de réflectance du sol.

$$SAVI = (1 + L) * \frac{NIR - R}{NIR + R + L}$$

où L est un paramètre qui varie entre 0 et 1 selon la quantité de sol visible. L=1 est généralement choisi si cette quantité est inconnue, L=0 lorsque la végétation est dense.

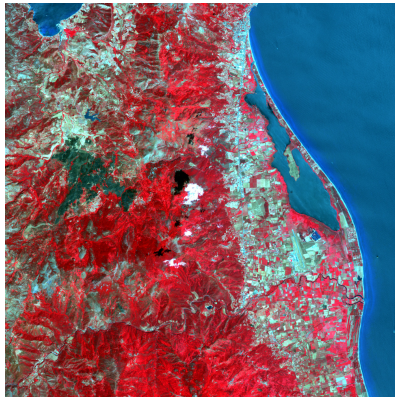
- le GEMI ("Global Environmental Monitoring Index"), qui est moins affecté par les variations atmosphériques [Pinty and Verstraete, 1992] et plus pertinent pour la détection de zones brûlées que le NDVI [Pereira, 1999]. Il distingue mieux les faibles valeurs de réflectance.

$$GEMI = \eta(1 - 0.25\eta) - \frac{(R - 0.125)}{1 - R}$$

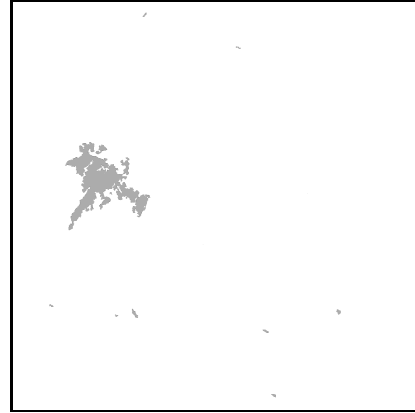
$$\text{où } \eta = \frac{2(NIR^2 - R^2) + 1.5NIR + 0.5R}{NIR + R + 0.5}$$

- le BAI ("Burnt Area Index") [Chuvieco et al., 2002] défini spécifiquement pour la détection de zones brûlées et qui met en valeur la quantité de carbone et cendres des zones brûlées. Cet index est donné par l'inverse de la distance euclidienne entre chaque pixel et un point de référence :

$$BAI = \frac{1}{(NIR - \rho_{NIR})^2 + (R - \rho_R)^2}$$



(a) Image SPOT 5 fausse couleur



(b) Estimation satellitaire



(c) Bande Vert



(d) Bande Rouge



(e) Bande Proche Infra-rouge



(f) Bande Moyen Infra-Rouge

FIG. 1.2: Est de la Corse (Image SPOT 5 © CNES 2005 Distribution SPOT Image - Estimation satellitaire projet RISK-EOS © Infoterra - ESA).

où  $\rho_{PIR}$  et  $\rho_R$  sont, respectivement, les réflectances de référence dans le PIR et le Rouge. Ces paramètres valent 0.06 et 0.1, valeurs fondées sur l'analyse empirique d'images satellitaires AVHRR.

- le NDII ("Normalized Difference Infrared Index") ou NBR ("Normalized Burn Ratio") [Key and Benson, 1999], défini par analogie avec le NDVI en remplaçant la bande Rouge par la bande MIR :

$$NDII = \frac{NIR - MIR}{NIR + MIR}$$

La Figure 1.3 montre la bande spectrale PIR (la plus pertinente pour la végétation) et différents indices de végétation (NDVI et NDII) pour l'image SPOT 5 précédente ainsi que leur seuillage. Les seuils ont été déterminés manuellement afin de mettre en évidence les zones brûlées ainsi que les régions ayant des indices de végétation équivalents.

Ces indices de végétation ne distinguent pas à proprement parler les zones brûlées/non-brûlées mais les zones de végétation/non-végétation.

### 1.2.3 Méthodes automatiques utilisant la télédétection

Les méthodes et les techniques utilisées pour la cartographie des zones brûlées par télédétection sont très diversifiées. Généralement, elles diffèrent en fonction de deux facteurs : le type et le nombre d'images utilisées. Le choix du type et du nombre d'images (bandes spectrales, résolution, couverture spatiale, temps de revisite du satellite, ...) doit prendre en compte les objectifs, les ressources financières, les ressources humaines ainsi que les délais.

Selon le choix du nombre d'images, les méthodes de cartographie de zones brûlées peuvent être divisées en deux catégories : les méthodes multi-temporelles et celles uni-temporelles.

#### Les approches multi-temporelles

Les méthodes multi-temporelles sont fondées sur des techniques de détection de changements, i.e. sur les différences des caractéristiques spectrales d'un même objet situé dans des images acquises avant et après le feu.

Elles sont généralement fondées sur l'analyse d'images satellitaires basse résolution (de l'ordre du kilomètre - par exemple, les satellites NOAA AVHRR, SPOT Végétation et MODIS pour le domaine de l'infra-rouge thermique) ou moyenne résolution (de l'ordre de la centaine de mètres - MODIS pour le domaine du visible et proche infra-rouge, IRS WiFS, SAC-C/MMRS). En effet, dans ce cas, le satellite nécessite un temps de revisite très court ainsi qu'une large couverture spatiale ce qui implique une résolution grossière.

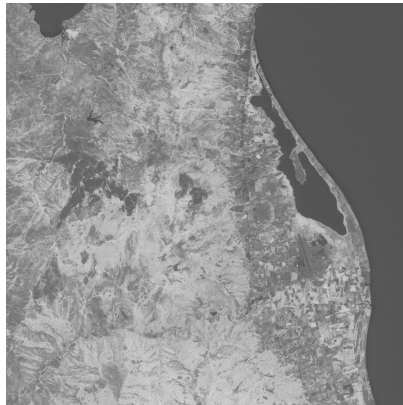
Une première approche très simple consiste à comparer directement certaines bandes spectrales des images acquises avant et après le feu. En effet, la végétation brûlée réfléchit beaucoup moins dans le domaine de l'infra-rouge. Cette approche très simpliste ne donne



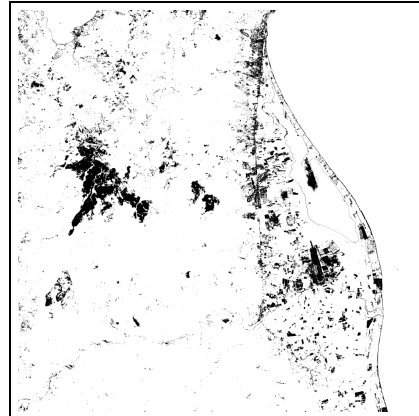
(a) Bande PIR



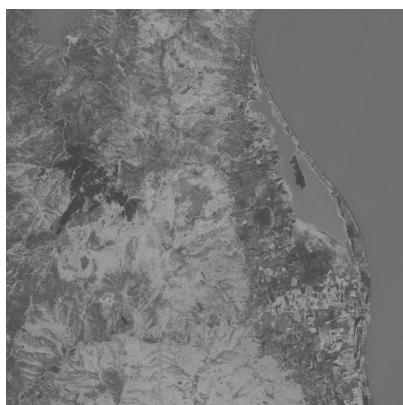
(b) Seuillage de la Bande PIR



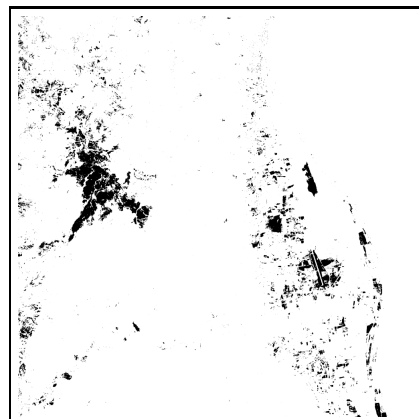
(c) NDVI



(d) Seuillage du NDVI



(e) NDII



(f) Seuillage du NDII

FIG. 1.3: Est de la Corse (Image SPOT 5 © CNES 2005, Distribution Spot Image).

pas de bons résultats puisque les deux images peuvent présenter des différences d'illumination et/ou d'angles d'incidence, qui peuvent engendrer des erreurs lors du calcul des variations des valeurs spectrales.

De nombreuses méthodes sont fondées sur l'analyse de la variation d'indices de végétation entre les images avant-feu et après-feu, acquises par le même satellite afin de tester l'efficacité des indices en fonction des bandes spectrales des différents satellites (IRS WiFS [Barbosa et al., 2002], Landsat TM [Conese and Checcacci, 2006]).

Le NDVI, dont les intérêts ont été énoncés précédemment, est l'indice le plus utilisé pour la cartographie des zones brûlées [Kasischke and French, 1995]. Cependant, de nombreux auteurs [Pereira, 1999; Chuvieco et al., 2002] ont montré ses limites en ce qui concerne la cartographie de zones brûlées, notamment pour la végétation éparsée des régions méditerranéennes (le NDVI est affecté par la réflectance du sol).

Le BAI et le NDII présentent une meilleure sensibilité pour discriminer les zones brûlées que le NDVI ou le GEMI et évitent la confusion avec l'eau et les ombres [Garcia and Chuvieco, 2004].

Le NDII donne de très bons résultats pour la détection de zones brûlées avec des images Landsat. En effet, la réflectance diminue dans le PIR (0.75-0.9  $\mu\text{m}$ ) et augmente dans le MIR (bande 7 : 2.08-2.35  $\mu\text{m}$ ). Ces variations sont un peu moins sensibles pour l'instrument Végétation des satellites SPOT 4 et SPOT 5 (PIR : 0,78 - 0,89  $\mu$  et MIR : 1,58 - 1,75  $\mu\text{m}$ ). Néanmoins, [Lasaponara and Lanorte, 2004] montrent aussi une meilleure sensibilité du NDII pour la capteur Végétation.

Cependant, la comparaison d'une seule image avant feu avec une seule image après feu peut induire des erreurs si l'une des deux images comporte du "bruit" (nuages ou autres perturbations atmosphériques, ombre des nuages, ...) [Chuvieco et al., 2008]. Par exemple, une zone brûlée qui avait été survolée par un nuage dans l'image avant-feu peut être considérée comme une zone de végétation saine dans l'image après-feu puisque son NDVI n'a pas diminué. De même, une zone de végétation saine survolée par un nuage dans l'image après-feu peut être considérée comme une zone de végétation brûlée puisque son NDVI est proche de zéro.

Une alternative à ce problème est d'utiliser plusieurs images avant-feu et plusieurs images après-feu.

Dans le cas de l'utilisation de séries temporelles d'images, la méthode la plus répandue est la méthode dite "*Maximum Value Composite*" (MVC). Pour chaque pixel, on calcule le maximum de NDVI sur une période de 10 jours ou d'un mois. La région dont la valeur de NDVI est maximale sur un tel intervalle de temps ne devrait pas avoir de nuage. La période à choisir dépend du compromis suivant : une longue période implique une diminution des interférences des nuages mais implique aussi une non-détection des variations à court terme de la zone considérée (ce qui n'est pas très gênant dans notre problème puisque la végétation ne change pas radicalement et naturellement en quelques jours).

Cette technique ne permet malheureusement pas d'éliminer tous les nuages et ombres parasites, qui possèdent de faibles valeurs de NDVI comme la végétation brûlée. De plus, elle n'est pas efficace dans le cas de régions qui présentent une végétation peu dense [Chuvieco et al., 2005a].



Fraser [Fraser et al., 2002] propose d'utiliser un algorithme de régression logistique fondé sur un calcul de probabilité en fonction des variations d'indices NDVI et NDII entre les images avant-feu et après-feu acquises par l'instrument Végétation du satellite SPOT.

Dans [Chuvieco et al., 2005b], Chuvieco teste plusieurs techniques composites fondées sur la sélection de maxima ou minima de certaines bandes spectrales et montre l'efficacité de l'utilisation de la température de surface des séries multi-temporelles d'images NOAA AVHRR (bande 4 : 10.30 - 11.30  $\mu m$ ) ou MODIS (10.78 - 11.28  $\mu m$ ) pour la cartographie de zones brûlées. Cette bande spectrale fournit une meilleure distinction de la présence de carbone et de cendres.

Chuvieco [Chuvieco et al., 2008] développe un algorithme de détection fondé sur la comparaison conjointe de deux indices (GEMI et BAI) de deux séries temporelles d'images AVHRR.

Cependant, ces techniques nécessitent l'acquisition de deux séries temporelles d'images (l'une avant-feu et l'autre après-feu), ce qui implique le choix d'un satellite ayant un temps de revisite très court et donc de très faible résolution (1 km). Elles permettent donc de détecter uniquement et de façon grossière de très larges zones brûlées.

Pour les méthodes fondées sur les indices de végétation, les seuils sont généralement réglés manuellement. De plus, l'efficacité des différents indices de végétation dépend des capteurs du satellite choisi et peut être sensible aux conditions d'illumination (cas des indices non normalisés).

#### **Méthode utilisée par Infoterra France - ESA :**

Infoterra France, filiale de EADS Astrium, est spécialisée dans la fourniture de produits et services de géo-information développés à partir de données d'observation de la Terre. Dans le domaine de la gestion des risques et des situations de crises, Infoterra France contribue à la mise en place du programme GMES ("Global Monitoring for Environment and Security") pour le compte de la Commission Européenne et de l'Agence Spatiale Européenne.

Les cartographies de zones brûlées fournies par Infoterra France sont obtenues quasi-automatiquement (réglage des seuils et reprise manuelle) à partir de la comparaison de paramètres biophysiques (liés à la végétation, à l'eau, aux nuages, etc.) calculés à partir d'images avant et après feu. Cette méthode utilise maintenant des images SPOT 2 et 4 (images multispectrale de 20m de résolution et panchromatique de 10m de résolution) et avait été testée avec succès sur des images Landsat (de résolution 30m).

Les cartographies (à partir d'images Landsat) dont nous disposons ne sont cependant pas exhaustives : problème de présence de nuages dans les images, détection de zones brûlées uniquement supérieures à 10 hectares (1 hectare dans le cas d'images SPOT).

#### **Méthode utilisée par le SERTIT :**

Le SERTIT (SErvice Régional de Traitement d'Image et de Télédétection) est spécialisé dans l'analyse de données de télédétection de crise. Il est soutenu par l'Agence Spatiale Européenne (ESA), et a été retenu par le Centre National d'Etudes Spatiales (CNES) pour produire la valeur ajoutée aux images satellitaires acquises lors de catastrophes majeures

(inondations, tempêtes, séismes, incendies, etc.) dans le cadre de la charte internationale "Espace et Catastrophes Majeures".

Dans le cas de grands feux, la sécurité civile peut activer la charte internationale qui est un accord international entre les différentes agences spatiales pour mettre les documents à disposition, gratuitement, en cas de catastrophe d'origine naturelle ou humaine.

Les procédures de traitement de cartographie rapide des dégâts sont fondées sur l'exploitation d'indices de changement tenant compte principalement des variations de luminosité.

### Les approches uni-temporelles

Dans le cas de l'utilisation de plusieurs images, les indices de végétation sont très utiles puisqu'ils permettent de mettre en valeur les changements de la végétation et sont moins sensibles aux différences d'illumination des images. Cependant, dans le cas de l'utilisation d'une seule image, le seuillage simple de bandes spectrales ou d'indices de végétation ne suffit pas si l'on désire détecter les zones brûlées (cf. Fig. 1.3). En effet, ces indices fournissent une information sur la densité de la végétation ; une faible densité ne signifie pas forcément que la zone est brûlée.

Beaucoup moins utilisées, les méthodes uni-temporelles sont fondées sur les différences des signatures spectrales des différents objets présents dans l'image.

L'annexe A propose d'étudier les caractéristiques spectrales des zones brûlées et non brûlées, pour chaque bande (Vert, Rouge, PIR et MIR), puis dans le cas de l'utilisation conjointe de deux ou plusieurs bandes.

La végétation brûlée est caractérisée par une réflectance un peu plus élevée dans le domaine visible et le MIR, et beaucoup plus faible dans le domaine PIR. Ce qui signifie que la courbe de réflectance pour la végétation brûlée est beaucoup plus horizontale que celle de la végétation saine, mais plus proche de celle des zones de non-végétation (comme les zones urbaines, les routes ou l'eau) [Quintano et al., 2006].

Ces approches sont généralement fondées sur des algorithmes standard de classification supervisée (qui consiste à identifier tous les pixels brûlés de l'image à partir d'échantillons de pixels brûlés et non-brûlés).

[Hudak and Brockett, 2004] utilisent l'algorithme du parallélépipède après avoir réalisé une Analyse en Composantes Principales (ACP) sur une image Landsat (l'annexe E présente l'algorithme de l'ACP). Cependant, cette technique de classification du parallélépipède ne permet pas de classer tous les pixels de l'image (les parallélépipèdes définis par les échantillons de pixels brûlés et non-brûlés peuvent se chevaucher).

L'algorithme ACP consiste à rechercher les directions de l'espace qui représentent statistiquement au mieux les données et représente les données dans le nouvel espace défini par ces directions. [Rogan and Franklin, 2001] et [Quintano et al., 2006] proposent d'utiliser le "Spectral Mixture Analysis" (SMA) qui représente les données comme une combinaison linéaire des réflectances de tous les éléments présents dans le pixel. Cette méthode estime le pourcentage de végétation, sol ou ombre contenu dans les caractéristiques spectrales des pixels. Puis, un arbre de décision [Rogan and Franklin, 2001] ou un algorithme de croissance de régions [Quintano et al., 2006] sont utilisés afin de détecter les zones brûlées.

Cependant, le SMA, qui est simplement une transformation linéaire rapide, donne de moins bons résultats que les réseaux de neurones [Quintano et al., 2006].

Aussi, [Sunar and Ozkan, 2001] et [Henry, 2008] proposent d'appliquer des algorithmes de classification plus traditionnels comme le maximum de vraisemblance ou les réseaux de neurones directement sur les composantes spectrales des pixels.

Bien évidemment, les approches multi-temporelles sont au moins aussi précises que les approches uni-temporelles puisqu'elles réduisent la confusion entre les zones brûlées et d'autres zones qui présentent des caractéristiques spectrales assez similaires. Elles nécessitent cependant l'acquisition d'une autre image (ou plusieurs), ce qui implique l'obligation de disposer d'une base de données récente, un coût supérieur (achat de l'image avant feu), des traitements plus compliqués (recalage entre les images avant et après feu, corrections géométrique, radiométrique et atmosphérique des images) [Henry, 2008].

En dépit des avantages apportés par la télédétection pour la cartographie des zones brûlées, certaines contraintes doivent être considérées comme la résolution temporelle du capteur (délai entre le feu et l'acquisition de(s) l'image(s)), la résolution spatiale (qui définit la taille minimale des zones pouvant être détectées) et la résolution spectrale (bandes spectrales des satellites qui évitent plus ou moins la confusion entre les zones brûlées et les zones de réponse spectrale légèrement différente).

### 1.3 Méthode proposée et son intérêt

Le grand intérêt de notre approche est d'utiliser une seule image haute résolution SPOT 5 afin de détecter les zones brûlées. Le but étant de réaliser une cartographie très exhaustive des zones brûlées. Le satellite SPOT 5 fournit des images multispectrales (Vert, Rouge, PIR et MIR) de résolution 10m ou des images Supermode (Vert, Rouge et PIR) de résolution 2.5m, et possède un temps de revisite de 2 ou 3 jours.

Le traitement d'une seule image après-feu évite l'acquisition d'une deuxième image SPOT 5, les problèmes de traitement liés à la comparaison et permet aussi de choisir un satellite haute résolution avec un temps de revisite supérieur à une journée.

La technique de classification est fondée sur les Séparateurs à Vaste Marge ou "Support Vector Machines" (SVM) [Burges, 1998]. Cette technique de classification supervisée assez récente a été développée par [Vapnik, 1998] et est utilisée dans de nombreux domaines tels que la reconnaissance de chiffres écrits à la main [Schölkopf et al., 1999], l'imagerie biomédicale [Schölkopf et al., 2004] et plus récemment la classification d'images en télédétection [Roli and Fumera, 2001] [Melgani and Bruzzone, 2004]. De plus, elle a démontré de meilleures précision et capacité de généralisation que les autres algorithmes de classification plus traditionnels comme la méthode du maximum de vraisemblance ou les réseaux de neurones [Pal and Mather, 2005; Fauvel, 2007] par exemple.

Les SVM ont pour but de séparer l'ensemble des données en deux classes en trouvant la fonction séparatrice optimale à partir d'observations, un sous-ensemble de données dont on connaît les classes, appelé *base d'apprentissage*. Ils considèrent l'apprentissage des observations et le processus de classification en traitant de la même façon les observations des

deux classes. Concernant notre problème de classification de zones brûlées, les pixels brûlés possèdent des réponses spectrales assez similaires (faible réflectance dans les 4 bandes) alors que celles des pixels non brûlés sont bien plus différentes puisqu'elles caractérisent des objets appartenant à des sous-classes variées comme les feuillus, les conifères, les champs, les zones urbaines, l'eau, les routes, etc.

Nous proposons ainsi d'utiliser les *One-Class SVM* (OC-SVM) [Schölkopf et al., 2001], cas particulier des SVM qui permet l'apprentissage et la classification à l'aide d'observations d'une seule classe, l'idée étant de définir le support de la distribution des observations. Les OC-SVM permettent donc de réduire les coûts de mémoire et de calcul des tâches d'apprentissage et de classification.

Dans le **chapitre 2**, nous introduisons différentes méthodes de classification dont les SVM et OC-SVM.

Le **chapitre 3** compare ces différentes méthodes avec les différentes cartes de référence fournies par le CNES, Infoterra France, le SERTIT, les différents Services Départementaux d'Incendie et de Secours (SDIS) et l'Office National des Forêts. Ces méthodes de classification utilisent uniquement l'information couleur du pixel et ne prennent pas en compte la cohérence spatiale des feux.

Le **chapitre 4** présente alors plusieurs approches de prise en compte de l'information spatiale des données telles que la morphologie mathématique [Soille, 1999] ou les champs de Markov [Winkler, 1995; Li, 1995]. Cependant, ces approches de régularisation ne donnent pas de bons résultats sur les contours des zones brûlées : elles lissent les bords des régions. Une autre méthode fondée sur le seuillage par hystérésis [Canny, 1986] permet d'éviter ce problème en considérant les SVM et OC-SVM comme des techniques de croissance de régions.

Un inconvénient des méthodes supervisées réside dans la sélection de la base d'apprentissage. En effet, il faut fournir des pixels déjà classés afin que ces méthodes apprennent à classer de nouveaux pixels. Les éléments de la base d'apprentissage ne peuvent être choisis aléatoirement et leur choix est crucial puisqu'ils définissent complètement la classification. Habituellement, cette phase de sélection est faite manuellement et nécessite du temps afin de choisir les pixels les plus discriminants de chaque classe. Dans le **chapitre 5**, nous proposons des processus de sélection automatique de la base d'apprentissage.

Enfin, le **chapitre 6** présente les conclusions ainsi que des améliorations possibles des modèles proposés.



## Chapitre 2

# Classification - Prise en compte de l'information spectrale

Une technique de classification est une méthode d'analyse de données qui a pour but de séparer l'image en plusieurs classes d'intérêt, i.e de regrouper les données en sous-ensembles homogènes, qui présentent des caractéristiques communes. Elle peut être supervisée ou automatique (i.e. non supervisée). Dans le cas supervisé, elle nécessite un sous-ensemble de données dont on connaît les classes, appelé *base d'apprentissage*, afin de pouvoir classifier de nouvelles données.

Concernant notre problème de cartographie de zones brûlées, nous devons déterminer deux classes, la classe des pixels brûlés et celle des pixels non brûlés.

Ce chapitre présente quelques techniques de classification en traitement d'image. Nous présentons tout d'abord des méthodes très classiques et largement utilisées en télédétection comme les algorithmes des K-Moyennes, des K-Plus Proches Voisins et du Maximum de Vraisemblance [Mather, 2004; Richards and Jia, 1999]. Puis, nous présentons les Séparateurs à Vaste Marge (SVM) et le cas particulier des "One-Class SVM".

## 2.1 Techniques de classification classiques

### 2.1.1 K-moyennes

L'algorithme des K-moyennes [Schowengerdt, 1997], aussi appelé "nuées dynamiques", est une méthode de classification automatique qui sépare les données en  $K$  classes. Sa grande popularité provient de sa grande simplicité.

On définit (aléatoirement si aucune information n'est donnée)  $K$  centres de gravité appelés centroïdes. Puis, on assigne chaque élément à la classe dont le centroïde est le plus proche (par rapport à une distance, euclidienne par exemple). Après chaque affectation, on calcule le centre de gravité des éléments de chaque classe. Les deux dernières étapes sont répétées de façon itérative jusqu'à la convergence des centroïdes.

Cet algorithme présente néanmoins quelques inconvénients : il faut fixer au préalable le nombre de classes  $K$ , il converge vers un minimum local et n'est pas adapté au cas de classes de structures non convexes et de tailles différentes.

[Palubinskas et al., 1998] modifient l'algorithme des K-moyennes en lui ajoutant un terme entropique afin d'estimer automatiquement le nombre de classes.

L'algorithme classique des K-moyennes propose de classifier les données en K ensembles en minimisant l'énergie suivante :

$$U = \sum_{i=1}^K \sum_{\vec{x}_j \in C_i} \|\vec{x}_j - \vec{\mu}_i\|^2$$

où  $\vec{\mu}_i$  désigne la moyenne des éléments de la classe  $C_i$ . Dans ce cas, le nombre de classes est fixé et représente le seul paramètre de l'algorithme, les moyennes  $\vec{\mu}_i$  étant estimées récursivement.

On définit l'entropie par  $E = - \sum_{i=1}^K p_i \cdot \log(p_i)$ , où  $p_i = \frac{\text{Card}(C_i)}{N}$  représente la probabilité qu'un pixel de l'image de taille N appartienne à la classe  $C_i$ .

En ajoutant ce terme d'entropie à l'énergie précédente, on obtient :

$$U = \sum_{i=1}^K \sum_{\vec{x}_j \in C_i} \|\vec{x}_j - \vec{\mu}_i\|^2 - \alpha_E \cdot \sum_{i=1}^K p_i \cdot \log(p_i)$$

On peut remarquer que le premier terme énergétique est minimal lorsque K est égal au nombre de niveaux de gris, mais que le second terme est minimal lorsqu'il n'y a qu'une seule classe. Le paramètre  $\alpha_E$  joue le rôle de compromis entre les deux termes. Après reformulation, on obtient :

$$U = \sum_{i=1}^K \sum_{\vec{x}_j \in C_i} (\|\vec{x}_j - \vec{\mu}_i\|^2 - \alpha \cdot \log(p_i)) \text{ avec } \alpha_E = \alpha \cdot N$$

L'algorithme est initialisé avec un nombre élevé de classes. A chaque itération, on assigne chaque pixel  $x$  à la classe  $i$  la plus proche au sens de la distance définie par l'équation précédente :  $d(\vec{x}, \vec{\mu}_i) = \|\vec{x} - \vec{\mu}_i\|^2 - \alpha \cdot \log(p_i)$ .

Il est à noter que le paramètre  $\alpha$  décroît de manière exponentielle à chaque itération. Ainsi le terme entropique détermine dès les premières itérations le nombre optimal de classes, mais n'intervient presque plus lorsque l'algorithme a convergé [Lorette, 1998]. L'algorithme des K-moyennes entropique se comporte alors comme un simple algorithme des K-moyennes.

L'algorithme des K-moyennes est donc le suivant :

---

**Algorithme des K-moyennes entropique :**

---

- Initialisation : choix d'un grand nombre de classe ( $K = 50$  par exemple) et positionnement aléatoire des  $K$  centroïdes sur des points de l'image,
  - Itérations : tant que les centroïdes changent de position :
    - on assigne chaque pixel à la classe du centroïde le plus proche au sens de la distance définie par :  $d(\vec{x}, \vec{\mu}_i) = \|\vec{x} - \vec{\mu}_i\|^2 - \alpha \cdot \log(p_i)$ ,
    - si une classe ne comporte plus d'élément, elle est éliminée,
    - on recalcule le centre de gravité  $\vec{\mu}_i$  des éléments de chaque classe
-

### 2.1.2 K-Plus Proches Voisins

L'algorithme des K Plus Proches Voisins (K-PPV) [Schowengerdt, 1997] est une méthode de classification supervisée. Il affecte un élément à la classe la plus représentée parmi les K plus proches éléments de la base d'apprentissage.

Le paramètre  $K$  permet de réduire les effets du bruit sur la classification. Cependant, si  $K$  est trop grand, l'algorithme aura tendance à assigner un nouvel objet à la classe de la base d'apprentissage la plus représentée. La vitesse de cet algorithme est proportionnelle au nombre d'éléments de la base d'apprentissage.

L'algorithme des K-Plus Proches Voisins est donc le suivant :

---

Algorithme des K-Plus Proches Voisins :

---

Pour chaque pixel  $\vec{x}$  de l'image :

- On calcule les distances entre le pixel et tous les vecteurs  $\vec{x}_i$  de la base d'apprentissage :  $d(\vec{x}, \vec{x}_i)$
  - On cherche les K vecteurs de la base d'apprentissage les plus proches du pixel, i.e, ceux qui ont les K plus petites distances
  - On assigne le pixel à la classe la plus représentée parmi ces K vecteurs.
- 

### 2.1.3 Maximum de Vraisemblance

L'estimation du Maximum de Vraisemblance (MV) est une méthode statistique supervisée [Mather, 2004]. Au lieu d'affecter un objet à la classe dont le centre de gravité est le plus proche (K-moyennes) ou dont les éléments sont les plus proches (K-PPV), il se base sur l'analyse statistique de la distribution des éléments de la base d'apprentissage pour définir des probabilités d'appartenance à chaque classe. Le nouvel objet est assigné à la classe pour laquelle la probabilité d'appartenance est la plus élevée.

Un avantage de cette méthode est qu'elle fournit, en plus de la classe, un degré de confiance lié à ce choix. Il faut cependant faire une hypothèse sur le type de distribution des éléments de la base d'apprentissage. Dans le cas d'une distribution gaussienne, on cherche à maximiser pour chaque nouvel objet  $\vec{x} \in R^m$  la probabilité d'appartenance à la classe  $y_i$  :

$$\arg \max_{y_i} P(\vec{x}/y_i) = \arg \max_{y_i} \frac{1}{\sqrt{2\pi^m |Q_i|}} \exp \left( -\frac{1}{2} (\vec{x} - \vec{\mu}_i)^T \cdot Q_i^{-1} \cdot (\vec{x} - \vec{\mu}_i) \right)$$

où  $\vec{\mu}_i$  et  $Q_i$  désigne respectivement la moyenne et la matrice de covariance associées à la classe  $y_i$ .

L'algorithme du Maximum de Vraisemblance est donc le suivant :

---

Algorithme du Maximum de Vraisemblance :

---

- On calcule des statistiques pour chaque classe de la base d'apprentissage : moyennes  $\vec{\mu}_i$  et matrices de co-variance  $Q_i$ ,
  - Pour chaque pixel  $\vec{x}$  :
    - on calcule les probabilités d'appartenance à chaque classe :  $P(\vec{x}/y_i)$
    - on assigne le pixel à la classe ayant la plus grande probabilité.
-



## 2.2 SVM

Le but d'un algorithme supervisé est de prédire le label d'éléments de  $R^m$  à partir d'un ensemble d'observations appelé *base d'apprentissage*  $\{(\vec{x}_i, y_i)\}_{i \in \langle 1, N \rangle}$ , composé par des objets  $\vec{x}_i \in R^m$  et les labels correspondant à la classe d'appartenance ( $y_i = \pm 1$ ) des objets. Le problème est donc de trouver une fonction  $f$  qui assigne le label +1 (respectivement -1) aux éléments  $\vec{x}$  tels que  $f(\vec{x}) \geq 0$  (respectivement  $f(\vec{x}) < 0$ ). La surface de séparation est donnée par l'équation  $f(\vec{x}) = 0$  qui divise l'espace en deux parties, l'une positive, l'autre négative.

Une solution évidente est de minimiser l'erreur d'apprentissage ou *risque empirique*, i.e. le taux de mauvaise classification sur la base d'apprentissage :

$$R_{emp}(f) = \frac{1}{2N} \sum_{i=1}^N |y_i - f(\vec{x}_i)|$$

Ce principe, appelé Minimisation du Risque Empirique ("Empirical Risk Minimization" - ERM) a néanmoins quelques inconvénients [Vert, 2001]. Premièrement, cela ne permet pas de définir une unique solution. Deuxièmement, le taux d'erreur sur tout l'espace peut être beaucoup plus grand que le taux d'erreur d'apprentissage (minimiser l'erreur sur un sous-ensemble d'éléments n'est pas équivalent à minimiser l'erreur sur tous les éléments). Dans ce cas, la fonction n'a pas une bonne capacité de généralisation : on parle de sur-apprentissage. Ce phénomène de sur-apprentissage est obtenu lorsque la surface de séparation a une forme très complexe et trop liée à la base d'apprentissage.

Le risque total, i.e. le taux de mauvaise classification sur  $R^m$  vaut :

$$R(f) = \int \frac{1}{2} |y - f(\vec{x})| dP(x, y)$$

La théorie de l'apprentissage statistique [Vapnik, 1998] fournit une relation entre le risque empirique et le risque total, i.e. le taux de mauvaise classification sur  $R^m$  : soit  $\eta \in [0, 1]$ , avec une probabilité de  $1 - \eta$ , on a :

$$R(f) \leq R_{emp}(f) + \sqrt{\frac{h[\log(2N/h) + 1] - \log(\eta/4)}{l}} \quad (2.1)$$

où  $h$  est un entier positif appelé "*VC-dimension*" (dimension de Vapnik-Chervonenkis). Pour un ensemble de fonctions  $\{f\}$ , il est défini comme le plus grand nombre de points qui peuvent être séparés par l'ensemble  $\{f\}$ , quelles que soient les étiquettes des points. On peut donc noter que les deux risques dépendent de la fonction de décision choisie  $f$ , alors que le deuxième terme de l'inégalité précédente est monotone croissant par rapport à  $h$  et dépend donc de l'ensemble de fonctions choisi.

A la différence des autres algorithmes supervisés, les SVM sont fondés sur le principe de la Minimisation du Risque Structurel ("Structural Risk Minimization" - SRM) [Vapnik, 1998] dont l'idée fondamentale est de minimiser la borne supérieure du risque total (cf Eq. 2.1). Ainsi, l'ensemble de fonction  $\{f\}$  est limité, ce qui implique une meilleure capacité de généralisation des SVM. Ce principe privilégie donc la capacité de généralisation par rapport à la classification sur la base d'apprentissage, ce qui permet d'éviter un sur-apprentissage.

### 2.2.1 Cas linéairement séparable

Une base d'apprentissage est dite linéairement séparable s'il existe au moins une fonction linéaire  $f$  qui classe correctement tous les objets de la base d'apprentissage. Evidemment, ce cas est très restrictif mais il permet de présenter le principe fondateur des SVM. Le cas non linéairement séparable sera étudié dans la prochaine partie.

Dans le cas linéairement séparable, l'hyperplan séparateur (voir Figure 2.1) est défini par l'équation  $f(\vec{x}) = \vec{w} \cdot \vec{x} + b = 0$  avec  $\vec{w}$  un vecteur normal à l'hyperplan et  $b$  le biais. Afin de classifier correctement la base d'apprentissage, nous devons avoir :

$$\begin{cases} \vec{w} \cdot \vec{x}_i + b \geq \epsilon > 0 & \text{si } y_i = +1 \\ \vec{w} \cdot \vec{x}_i + b \leq -\epsilon < 0 & \text{si } y_i = -1 \end{cases} \text{ ce qui est équivalent à } \begin{cases} \vec{w} \cdot \vec{x}_i + b \geq 1 & \text{si } y_i = +1 \\ \vec{w} \cdot \vec{x}_i + b \leq -1 & \text{si } y_i = -1 \end{cases}$$

Ces contraintes impliquent que la *marge*, i.e. deux fois la plus petite distance entre un point de la base d'apprentissage et l'hyperplan séparateur, définie par la distance entre  $\vec{w} \cdot \vec{x} + b = -1$  et  $\vec{w} \cdot \vec{x} + b = +1$ , est égale à  $\frac{2}{\|\vec{w}\|}$ . Les éléments de la base d'apprentissage doivent être du bon côté du tube défini par les équations  $\vec{w} \cdot \vec{x} + b = -1$  et  $\vec{w} \cdot \vec{x} + b = +1$ .

Selon la théorie de l'apprentissage statistique, dans le cas de classifieurs linéaires de risque empirique nul, la "VC-dimension" est directement liée à la marge. En effet, Vapnik montre que, pour une famille de classifieurs linéaires  $\{f\}$  telle que  $\|\vec{w}\| \leq A$  :

$$h \leq \min([R^2 A^2], m) + 1 \quad (2.2)$$

où  $R$  désigne le rayon de la plus petite sphère contenant les vecteurs  $\vec{x}_i$  et  $[\cdot]$  désigne la partie entière [Schölkopf et al., 1999].

L'hyperplan choisi doit donc faire partie de la famille d'hyperplans à marge maximale afin de minimiser la "VC-dimension". Cet hyperplan est appelé *Hyperplan Séparateur Optimal* (HSO).

Comme maximiser la marge est équivalent à minimiser l'inverse de la marge, les paramètres optimaux  $\vec{w}^*$  and  $b^*$  sont obtenus après résolution du Problème d'Optimisation Quadratique (POQ) convexe sous contraintes linéaires :

$$\min_{(w,b)} \frac{\|\vec{w}\|^2}{2} \quad \text{tels que : } y_i(\vec{w} \cdot \vec{x}_i + b) \geq 1, \forall i \in \langle 1, N \rangle \quad (2.3)$$

En introduisant les multiplicateurs de Lagrange  $\{\lambda_i\}_{i \in \langle 1, N \rangle}$  correspondant aux contraintes, nous obtenons le problème dual :

$$L(\vec{w}, b, \vec{\lambda}) = \frac{\|\vec{w}\|^2}{2} - \sum_{i=1}^N \lambda_i \cdot [y_i(\vec{w} \cdot \vec{x}_i + b) - 1] \quad (2.4)$$

L'unique point d'inflexion du Lagrangien  $L$ , qui est un minimum par rapport à  $(\vec{w}, b)$  et un maximum par rapport à  $\lambda$ , détermine la solution du POQ. Tout d'abord, minimisons

le Lagrangien par rapport à  $(\vec{w}, b)$  :

$$\frac{\partial L}{\partial \vec{w}} = \vec{w} - \sum_{i=1}^N \lambda_i y_i \vec{x}_i = 0 \quad (2.5)$$

$$\frac{\partial L}{\partial b} = -\sum_{i=1}^N \lambda_i y_i = 0 \quad (2.6)$$

En remplaçant ces équations dans Eq. (2.4), nous obtenons :

$$\begin{aligned} L(\vec{w}, b, \vec{\lambda}) &= \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j \vec{x}_i \cdot \vec{x}_j - \sum_{i=1}^N \lambda_i \cdot [y_i (\sum_{j=1}^N \lambda_j y_j \vec{x}_j \cdot \vec{x}_i + b) - 1] \\ &= \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j \vec{x}_i \cdot \vec{x}_j - \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j \vec{x}_j \cdot \vec{x}_i + \sum_{i=1}^N \lambda_i \\ &= -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j \vec{x}_i \cdot \vec{x}_j + \sum_{i=1}^N \lambda_i \end{aligned}$$

Le problème dual est donc réduit au POQ convexe avec des contraintes linéaires :

$$\begin{aligned} \max_{\lambda} W(\lambda) &= -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j \vec{x}_i \cdot \vec{x}_j + \sum_{i=1}^N \lambda_i \\ \text{tels que : } &\sum_{i=1}^N \lambda_i y_i = 0, \quad 0 \leq \lambda_i, \quad \forall i \in \langle 1, N \rangle \end{aligned}$$

Si nous appelons  $\lambda^*$  la solution du problème, le paramètre optimal  $\vec{w}^*$  est obtenu à partir de l'équation (2.5) :

$$\vec{w}^* = \sum_{i=1}^N \lambda_i^* y_i \vec{x}_i \quad (2.7)$$

et  $b^*$  est obtenu à partir des conditions de Karush-Kuhn-Tucker (KKT) [Bishop, 2006] qui dérivent des contraintes de (2.3) :

$$\lambda_i^* \cdot [y_i (\vec{w}^* \cdot \vec{x}_i + b^*) - 1] = 0$$

Si  $\lambda_i^*$  est non nul, le vecteur correspondant  $\vec{x}_i$  est positionné sur le bord du tube (à une distance égale à la moitié de la marge de l'HSO :  $y_i (\vec{w}^* \cdot \vec{x}_i + b^*) = 1$ ). Sinon, le vecteur correspondant  $\vec{x}_i$  est positionné du bon côté du tube et sa distance par rapport à l'HSO est supérieure à la moitié de la marge.

Remarquons que  $\vec{w}^*$  est une combinaison linéaire des vecteurs  $\vec{x}_i$  tels que les multiplicateurs de Lagrange  $\lambda_i^*$  soient non nuls. Ces vecteurs sont appelés *Vecteurs Supports* (VS) et sont les éléments de la base d'apprentissage les plus proches de l'HSO. De plus, les paramètres optimaux  $(\vec{w}^*, b^*, \vec{\lambda}^*)$  ne changent pas si l'on enlève des vecteurs non-supports de la base d'apprentissage.

La fonction de décision  $f$  est donnée par :

$$f(\vec{x}) = \vec{w}^* \cdot \vec{x} + b^* = \sum_{i=1}^N \lambda_i^* y_i \vec{x}_i \cdot \vec{x} + b^*$$

La classification de nouveaux objets  $\vec{x}$  nécessite le calcul du produit scalaire entre  $\vec{x}$  et les VS. De plus, la fonction de décision  $f$  est définie complètement par les VS et ne dépend pas des vecteurs non-supports.

### 2.2.2 Cas non-linéairement séparable

Dans la partie précédente, nous avons supposé la base d'apprentissage linéairement séparable et nous avons montré comment obtenir le classifieur à marge maximale. Dans cette partie, nous allons généraliser au cas de bases d'apprentissage non-linéairement séparables. En effet, l'hypothèse de cas linéairement séparable est trop restrictive pour de nombreuses applications réelles, surtout lorsque les données sont bruitées.

**Marge relaxée :**

L'analyse précédente est étendue afin de permettre des erreurs sur la base d'apprentissage en introduisant des variables dites *souples*  $\{\xi_i\}_{i \in \langle 1, N \rangle}$  dans le modèle précédent (voir Figure 2.2).

Ces variables quantifient les erreurs réalisées sur les éléments de la base d'apprentissage. En effet, la variable souple  $\xi_i$  correspondant au vecteur  $\vec{x}_i$  est nulle si le vecteur est du bon côté du tube défini par  $\{\vec{w}, b\}$ , sinon il est égal à la distance entre la frontière du tube  $\vec{w} \cdot \vec{x} + b = \pm 1$  et le vecteur  $\vec{x}_i$ .

$$\xi_i(\vec{w}, b) = \begin{cases} 0 & \text{si } y_i \cdot (\vec{w} \cdot \vec{x}_i + b) \geq 1 \\ 1 - y_i \cdot (\vec{w} \cdot \vec{x}_i + b) & \text{si } y_i \cdot (\vec{w} \cdot \vec{x}_i + b) \leq 1 \end{cases}$$

Les SVM à marge relaxée ont pour but de maximiser la marge et minimiser l'erreur d'apprentissage, ce qui implique un nouveau POQ :

$$\min_{(\vec{w}, b, \xi)} \frac{\|\vec{w}\|^2}{2} + C \sum_{i=1}^N \xi_i \quad (2.8)$$

$$\text{tel que : } y_i(\vec{w} \cdot \vec{x}_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad \forall i \in \langle 1, N \rangle \quad (2.9)$$

où  $C$  est appelé *paramètre de régularisation* et est positif. Lorsque  $C$  est grand, le problème pénalise plus l'erreur d'apprentissage alors qu'un paramètre  $C$  plus petit autorise une marge et une erreur d'apprentissage plus grande. Ce paramètre permet donc de définir un compromis entre une grande marge et un faible nombre d'erreurs sur la base d'apprentissage.

Comme précédemment (voir Eq. (2.3)), introduisons les multiplicateurs de Lagrange  $\{\lambda_i\}_{i \in \langle 1, N \rangle}$  et  $\{\mu_i\}_{i \in \langle 1, N \rangle}$  associés respectivement aux contraintes  $y_i(\vec{w} \cdot \vec{x}_i + b) \geq 1 - \xi_i$  et  $\xi_i \geq 0$  :

$$L(\vec{w}, b, \vec{\xi}, \vec{\lambda}, \vec{\mu}) = \frac{\|\vec{w}\|^2}{2} + C \sum_{i=1}^N \xi_i - \sum_{i=1}^N \lambda_i \cdot [y_i(\vec{w} \cdot \vec{x}_i + b) - 1 + \xi_i] - \sum_{i=1}^N \mu_i \cdot \xi_i \quad (2.10)$$

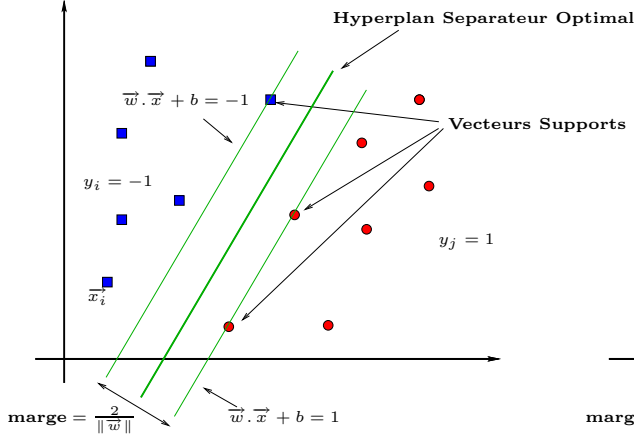


FIG. 2.1: SVM : cas linéaire

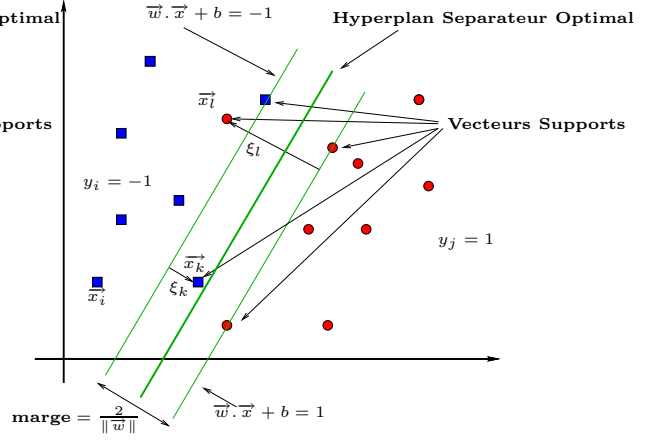


FIG. 2.2: SVM : cas des marges relaxées

Afin de minimiser le Lagrangien par rapport à  $\{\vec{w}, b, \xi\}$ , calculons les dérivées partielles :

$$\frac{\partial L}{\partial \vec{w}} = \vec{w} - \sum_{i=1}^N \lambda_i y_i \vec{x}_i = 0 \quad (2.11)$$

$$\frac{\partial L}{\partial b} = - \sum_{i=1}^N \lambda_i y_i = 0 \quad (2.12)$$

$$\frac{\partial L}{\partial \xi_i} = C - \lambda_i - \mu_i = 0 \quad (2.13)$$

En remplaçant ces équations dans Eq. (2.10), nous obtenons :

$$\max_{\lambda} W(\lambda) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j \vec{x}_i \cdot \vec{x}_j + \sum_{i=1}^N \lambda_i$$

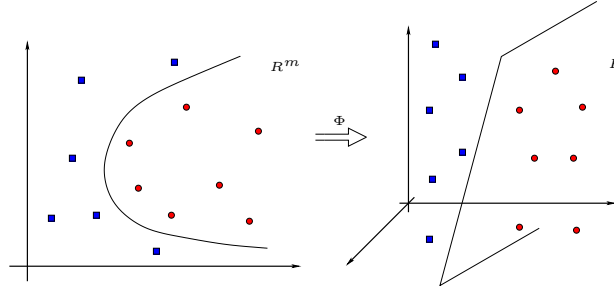
tel que :  $\sum_{i=1}^N \lambda_i y_i = 0$ ,  $0 \leq \lambda_i \leq C$ ,  $\forall i \in \langle 1, N \rangle$

La fonction à minimiser est identique à celle du problème dans le cas linéairement séparable, nous avons néanmoins une contrainte supplémentaire :  $\lambda$  doit être borné supérieurement par  $C$ . Après avoir résolu ce problème, nous retrouvons la solution  $\vec{w}^*$  à partir de la même équation (2.7) que précédemment.

Remarquons que les conditions KKT permettent de calculer  $b$  :

$$\begin{aligned} \lambda_i^* \cdot [y_i (\vec{w}^* \cdot \vec{x}_i + b^*) - 1 + \xi_i^*] &= 0 \\ C - \lambda_i^* - \mu_i^* &= 0 \\ \mu_i^* \cdot \xi_i^* &= 0 \end{aligned}$$

ce qui signifie que :

FIG. 2.3: Rôle de la fonction  $\Phi$ 

- si  $\lambda_i^* = 0$ , alors  $\mu_i^* = C$ ,  $\xi_i^* = 0$  et  $[y_i(\vec{w}^* \cdot \vec{x}_i + b^*) - 1] > 0$ ; le vecteur  $\vec{x}_i$  est du bon côté du tube,
- si  $0 < \lambda_i^* < C$ , alors  $0 < \mu_i^* < C$ ,  $\xi_i^* = 0$  et  $[y_i(\vec{w}^* \cdot \vec{x}_i + b^*) - 1] = 0$ ; le vecteur  $\vec{x}_i$  est sur la frontière du tube et est un VS,
- si  $\lambda_i^* = C$ , alors  $\mu_i^* = 0$ ,  $\xi_i^* > 0$  et  $[y_i(\vec{w}^* \cdot \vec{x}_i + b) - 1 + \xi_i^*] = 0$ ;  $\vec{x}_i$  n'est pas du bon côté du tube (il est soit bien classifié du bon côté de l'HSO mais à l'intérieur du tube, soit du mauvais côté de l'HSO et donc mal classifié). Dans ce cas, il est toujours un Vecteur Support et intervient dans la définition de  $\vec{w}$  et donc dans la position de l'HSO.

### SVM non linéaires :

Dans la plupart des cas, la base d'apprentissage nécessite une surface de décision plus compliquée qu'un simple hyperplan linéaire. Pour pouvoir prendre en compte un séparateur non linéaire, les SVM linéaires peuvent être généralisés par l'introduction d'une fonction  $\Phi$  qui transfère les données de  $R^m$  vers un ensemble de dimension supérieure  $F$  dans lequel elles deviennent linéairement séparables (voir Figure 2.3) :

$$\begin{aligned} \Phi : R^m &\longrightarrow F \\ \vec{x} &\longmapsto \Phi(\vec{x}) \end{aligned}$$

L'idée est donc de séparer linéairement les données transférées dans l'espace  $F$  de dimension supérieure grâce à la fonction  $\Phi : \{(\Phi(\vec{x}_i), y_i)\}_{i \in \langle 1, N \rangle}$ . Par analogie avec l'analyse précédente, les paramètres de l'HSO sont donnés par la résolution du problème d'optimisation suivant :

$$\max_{\lambda} W(\lambda) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j \Phi(\vec{x}_i) \cdot \Phi(\vec{x}_j) + \sum_{i=1}^N \lambda_i \quad (2.14)$$

$$\text{tels que : } \sum_{i=1}^N \lambda_i y_i = 0, \quad 0 \leq \lambda_i \leq C, \quad \forall i \in \langle 1, N \rangle \quad (2.15)$$

ce qui donne :

$$\vec{w}^* = \sum_{i=1}^N \lambda_i^* y_i \Phi(\vec{x}_i) \quad (2.16)$$

$$f(\vec{x}) = \sum_{i=1}^N \lambda_i^* y_i \Phi(\vec{x}_i) \cdot \Phi(\vec{x}) + b^* \quad (2.17)$$

Ces formules sont obtenues en remplaçant les vecteurs  $\vec{x}$  par leur valeur dans l'ensemble des images  $\Phi(\vec{x})$ .

Remarquons que, dans l'équation (2.14) et (2.17),  $\Phi$  n'intervient que dans le produit scalaire de deux éléments de l'espace des images. Il suffit donc de connaître la valeur du produit scalaire pour n'importe quel couple de points, qui est appelé *noyau* :

$$K(\vec{x}, \vec{x}') = \Phi(\vec{x}) \cdot \Phi(\vec{x}')$$

Par conséquent, si l'on trouve une fonction  $K$  qui s'écrit comme un produit scalaire, la connaissance explicite de  $\Phi$  n'est pas nécessaire : il s'agit du "*truc du noyau*".

Le théorème de Mercer fournit un critère très efficace afin de savoir si une fonction peut être considérée comme un noyau : une fonction symétrique  $K$  est un noyau si et seulement si, quelle que soit la famille de vecteurs  $\{\vec{x}_i\}_{i \in \langle 1, \dots \rangle}$ ,  $K(\vec{x}_i, \vec{x}_j)$  est une matrice définie positive.

Ce théorème permet de calculer directement le produit scalaire  $K(\vec{x}_1, \vec{x}_2)$  sans avoir à calculer  $\Phi(\vec{x}_1)$  et  $\Phi(\vec{x}_2)$ . Ainsi, le calcul et la définition de  $\Phi$  sont donc évités.

Un noyau correspond en fait à un produit scalaire (fonction symétrique définie positive) dans l'espace des images. Le problème d'optimisation reste donc quadratique convexe avec des contraintes d'égalité linéaires puisque le noyau est défini positif.

### 2.2.3 Quelques noyaux

Un noyau caractérise le degré de similarité ou la corrélation entre deux éléments. Il a une valeur élevée pour deux vecteurs qui se ressemblent et faible pour deux vecteurs plus différents.

Les noyaux les plus classiques (qui sont aussi largement utilisés en télédétection) sont :

- le noyau linéaire :  $K(\vec{x}, \vec{x}') = \vec{x} \cdot \vec{x}'$ ,
- le noyau polynomial :  $K(\vec{x}, \vec{x}') = (\vec{x} \cdot \vec{x}' + 1)^q$ . Plus le degré  $q$  est élevé, plus la forme de l'HSO est complexe,
- le noyau RBF (Radial Basis Function) :  $K(\vec{x}, \vec{x}') = \exp(-\frac{\|\vec{x} - \vec{x}'\|}{\sigma})$ ,
- le noyau Gaussien :  $K(\vec{x}, \vec{x}') = \exp(-\frac{\|\vec{x} - \vec{x}'\|^2}{2\sigma^2})$ .

Presque toutes les formes peuvent être obtenues à partir de ces deux derniers noyaux. Plus  $\sigma$  est proche de zéro, plus les gaussiennes centrées sur les VS seront pointues et plus complexe sera l'OSH. Ces noyaux correspondent à un transfert dans des espaces de "VC-dimensions" infinies.

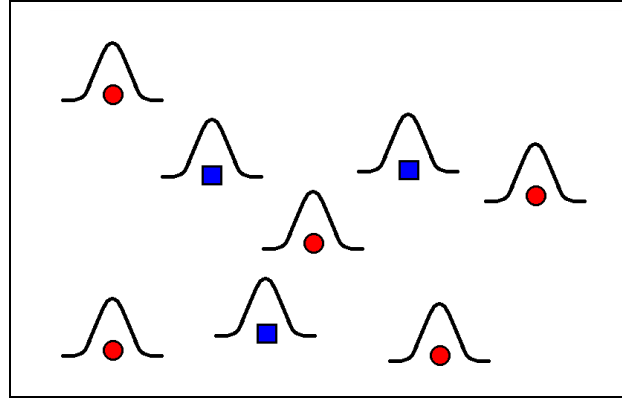


FIG. 2.4: SVM avec un noyau Gaussien

En effet, la fonction de décision est égale à une somme de gaussiennes centrées sur les Vecteurs Supports  $\vec{x}_i$  pondérées par les coefficients  $y_i \lambda_i$ . Comme le montre la figure 2.4, quel que soit le nombre d'éléments de la base d'apprentissage, il est possible de trouver une petite valeur de  $\sigma$  telle que les gaussiennes soient fines et que tous les éléments soient correctement classifiés.

D'autres noyaux ont démontré de bons résultats de classification :

- le noyau KMOD ("Kernel with MODerate decrease") [Ayat et al., 2002] :

$K(\vec{x}, \vec{x}') = a[\exp(\frac{\gamma}{\|\vec{x} - \vec{x}'\|^2 + \alpha^2}) - 1]$  où  $a = \frac{1}{\exp(\frac{\gamma}{\alpha^2}) - 1}$ . Ce noyau montre une décroissance plus rapide proche de zéro et une décroissance plus modérée vers l'infini que le noyau gaussien. Il a donc tendance à différencier plus facilement des vecteurs proches et conserve une information pour des vecteurs plus éloignés,

- le noyau SAM ("Spectral Angle Mapper") [Mercier and Lennon, 2003] :

$K(\vec{x}, \vec{x}') = \exp(-\gamma \arccos(\frac{\vec{x} \cdot \vec{x}'}{\|\vec{x}\| \|\vec{x}'\|}))$ . Ce noyau est invariant par changement d'échelle et n'est pas fonction de la distance quadratique. Il semble être pertinent dans le cas de classification d'image puisqu'il est plus robuste aux différences d'illumination ou d'ombres entre deux vecteurs.

Remarquons enfin que de nouveaux noyaux peuvent être facilement obtenus via des opérations simples sur des noyaux. Par exemple, une combinaison linéaire de noyaux est un noyau (il satisfait la condition de Mercer). Ainsi, il peut être efficace de combiner deux noyaux afin de tirer parti de leurs avantages respectifs :

$$K(\vec{x}, \vec{x}') = \mu.K_1(\vec{x}, \vec{x}') + (1 - \mu).K_2(\vec{x}, \vec{x}')$$

#### 2.2.4 One-Class SVM

[Schölkopf et al., 2001] ont récemment étendu la théorie des SVM afin de réaliser l'apprentissage et la classification en utilisant uniquement la base d'apprentissage positive. Les "One-Class SVM" (OC-SVM) ont pour but d'estimer le support qui inclut la plupart des éléments positifs de la base d'apprentissage, i.e. trouver une fonction positive dans une région qui englobe la plupart des données et négative ailleurs.



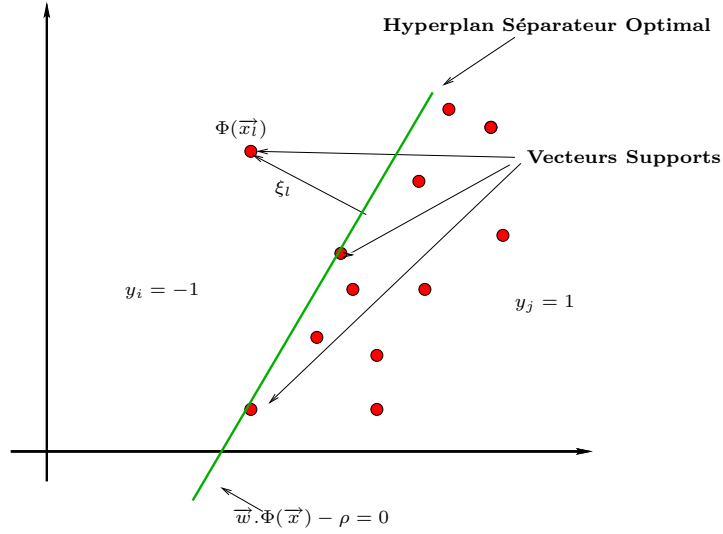


FIG. 2.5: One-Class SVM

Cette approche est équivalente à trouver la surface qui sépare les données positives et l'origine pour un certain seuil  $\rho$  (cf. figure 2.5) :

$$f(\vec{x}) = \text{sign}(\vec{w} \cdot \Phi(\vec{x}) - \rho)$$

Soit  $l$  le nombre d'éléments (appartenant tous à la classe positive) de la base d'apprentissage  $\{\vec{x}_i\}_{i \in \langle 1, l \rangle}$ , le problème à maximiser est :

$$\min_{(w, \rho, \xi)} \left[ \frac{\|\vec{w}\|^2}{2} - \rho + \frac{1}{\nu \cdot l} \sum_{i=1}^l \xi_i \right]$$

$$\text{tels que : } \vec{w} \cdot \Phi(\vec{x}_i) \geq \rho - \xi_i, \quad \xi_i \geq 0, \quad \forall i \in \langle 1, l \rangle$$

où  $\nu \in [0, 1]$  définit un compromis entre la marge (distance entre l'hyperplan et l'origine défini par  $\frac{\rho}{\|\vec{w}\|}$ ) et le nombre d'erreurs d'apprentissage. Lorsque  $\nu$  est petit, le problème interdit de grosses erreurs d'apprentissage ; lorsque  $\nu$  est plus grand, le problème permet une plus grande marge.

Remarquons que le paramètre de régularisation  $\frac{1}{\nu \cdot l}$  (appelé  $C$  dans le cas des SVM classiques) est défini en fonction de la taille de la base d'apprentissage et de  $\nu$ , dont nous verrons l'intérêt par la suite.

En introduisant les multiplicateurs de Lagrange, le problème quadratique précédent est équivalent à :

$$\max_{\lambda} W(\lambda) = -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \lambda_i \lambda_j K(\vec{x}_i, \vec{x}_j) \quad (2.18)$$

$$\text{tels que : } \sum_{i=1}^l \lambda_i = 1, \quad 0 \leq \lambda_i \leq \frac{1}{\nu \cdot l}, \quad \forall i \in \langle 1, l \rangle \quad (2.19)$$

Ce problème ressemble énormément à celui des SVM classiques (cf. Eq. 2.14). La solution  $\vec{w}^*$  est donnée par :  $\vec{w}^* = \sum_{i=1}^l \lambda_i^* \Phi(\vec{x}_i)$  et  $\rho^*$  est obtenu à partir des contraintes :  $\lambda_i^* \cdot [\vec{w}^* \cdot \Phi(\vec{x}_i) - \rho^*] = 0$

Finalement, nous obtenons :

$$f(\vec{x}) = \sum_{i=1}^l \lambda_i^* K(\vec{x}_i, \vec{x}) - \rho^* \quad (2.20)$$

La classification de données non classées  $\vec{x}$  est donc déduite à partir des éléments de la base d'apprentissage tels que  $\lambda_i$  est différent de 0. Ces vecteurs sont aussi appelés Vecteurs Supports (VS). De plus, si  $0 < \lambda_i < \frac{1}{\nu l}$ , alors le VS correspondant se situe sur la surface séparatrice :  $f(\vec{x}_i) = 0$ . Les vecteurs de la base d'apprentissage tels que  $\lambda_i = 0$  sont classifiés du bon côté de la surface (à l'intérieur de la surface). Ceux tels que  $\lambda_i = \frac{1}{\nu l}$  sont mal classifiés, à l'extérieur de la surface :  $f(\vec{x}_i) < 0$ .

**Remarque :** Dans l'annexe G, nous proposons de résoudre une variante du problème des OC-SVM : trouver le cercle de taille minimal qui englobe la plupart des éléments positifs de la base d'apprentissage. Dans cette annexe, on démontre aussi que ces deux problèmes sont équivalents dans le cas de certains noyaux.

## 2.2.5 Implantation

### Cas des SVM

Il existe de nombreuses méthodes pour résoudre le problème d'optimisation quadratique sous contraintes linéaires des SVM (voir Eq. (2.14) et (2.15)) telles que la méthode du gradient conjugué, ou celle des points intérieurs. Cependant, ces méthodes ne sont pas très efficaces lorsque la taille de la base d'apprentissage est élevée. Des algorithmes plus spécialisés ont été développés pour résoudre plus rapidement le POQ.

Ces algorithmes exploitent les avantages du problème d'optimisation (unicité de la solution et représentation parcimonieuse de la solution qui n'utilise que les Vecteurs Supports) afin de s'affranchir de quelques inconvénients (stockage en mémoire d'une matrice de taille  $N^2$ ).

L'algorithme de "chunking" proposé par [Vapnik, 1998] utilise le fait que la solution du POQ reste la même si l'on enlève les vecteurs non supports. Cet algorithme casse le complexe POQ en une suite de POQs plus petits et a pour but d'identifier les multiplicateurs de Lagrange non nuls et rejeter les multiplicateurs nuls.

Une méthode simple et fréquemment utilisée pour résoudre le POQ est l'algorithme "Sequential Minimal Optimization" (SMO) de Platt [Schölkopf et al., 1999] qui réduit le POQ à un sous-problème de dimension deux qui peut être résolu analytiquement et de façon très rapide. La présentation de l'algorithme est faite dans l'annexe B. [Keerthi et al., 2001] proposent une modification du SMO afin d'améliorer son temps de convergence.

Il n'y a cependant pas de critère concernant le choix des paramètres impliqués dans le POQ (paramètre  $C$  pour les marges relaxées et paramètre du noyau  $\sigma$  pour un noyau Gaussien par exemple). Un moyen de choisir ces paramètres est de définir une grille de valeurs, de tester leur efficacité et de prendre le plus pertinent. Une autre méthode habituelle est la technique de *validation croisée* [Schölkopf et al., 2004] : la base d'apprentissage est coupée en plusieurs sous-parties et on teste la classification obtenue sur chaque sous-partie après avoir réalisé l'apprentissage des SVM sur l'ensemble des autres sous-parties, pour chaque couple de paramètres. Le couple de paramètres choisi est celui qui fournit la plus faible erreur sur l'ensemble des tests.

La technique de validation croisée fournit une estimation de l'erreur totale d'une classification comme étant l'erreur de cette classification sur un sous-ensemble de données qui n'a pas été utilisé lors de la phase d'apprentissage. Une autre méthode consiste à utiliser la "VC-dimension" (Eq. 2.2) qui définit une borne supérieure de l'erreur totale (Eq. 2.1). L'idée est donc de choisir le couple d'hyperparamètres qui minimise la borne supérieure de l'erreur totale en utilisant une descente de gradient par rapport aux hyperparamètres [Chapelle et al., 2002; Keerthi, 2002].

### Cas des OC-SVM

Le POQ des OC-SVM est très semblable à celui des OC-SVM. [Schölkopf et al., 2001] ont donc légèrement modifié l'algorithme SMO afin de résoudre ce problème d'optimisation.

Concernant le choix des paramètres des OC-SVM ( $\sigma$  et  $\nu$ ), la technique de validation croisée ne peut être appliquée puisque la base d'apprentissage ne contient que des éléments de la classe positive. En effet, un couple de paramètres qui génère un classifieur positif pour tout point produit une erreur nulle sur des sous-ensembles de la base d'apprentissage, mais n'est évidemment pas pertinent.

Notons que, d'après les contraintes 2.19 :

$$1 = \sum_{i=1}^l \lambda_i \leq \frac{1}{\nu} \text{ et } 0 \leq \lambda_i \leq \frac{1}{\nu.l}$$

ce qui implique la contrainte :

$$0 \leq \nu \leq 1$$

De plus,  $\sum_{i=1}^l \lambda_i = 1$ , on a alors  $0 \leq \lambda_i \leq 1$ . Il ne sert donc à rien de choisir  $\nu$  tel que  $\lambda_i \leq 1 \leq \frac{1}{\nu.l}$ . Il convient donc de choisir :

$$\frac{1}{l} \leq \nu \leq 1$$

Pour les Vecteurs Supports,  $\lambda_i^{VS} \neq 0$ . De plus, pour les "outliers" (Vecteurs Supports de la base d'apprentissage mal classifiés),  $\lambda_i^{mc} = \frac{1}{\nu.l}$ . Il vient donc :

$$1 = \sum_{i=1}^{l_{VS}} \lambda_i^{VS} \leq \frac{l_{VS}}{\nu.l} \quad \text{et} \quad \sum_{i=1}^{l_{mc}} \lambda_i^{mc} = \frac{l_{mc}}{\nu.l} \leq \sum_{i=1}^l \lambda_i = 1$$

en notant respectivement  $l_{VS}$  et  $l_{mc}$  les nombres de Vecteurs Supports et d'outliers.

Finalement, on obtient :

$$\nu \leq \frac{l_{VS}}{l} \quad \text{et} \quad \frac{l_{mc}}{l} \leq \nu \quad (2.21)$$

ce qui signifie que  $\nu$  est une limite supérieure du taux d'outliers et une limite inférieure du taux de VS contenu dans la base d'apprentissage. Ce paramètre peut donc être choisi en fonction de l'estimation du nombre d'outliers de la base d'apprentissage.

Remarquons que, si  $\nu = 1$ , il vient nécessairement  $\lambda_i = \frac{1}{l} = \frac{1}{\nu l}$ ,  $\forall i \in \{1, \dots, l\}$ . Ce qui implique que tous les vecteurs de la base d'apprentissage sont des outliers et sont donc tous mal classifiés !

[Rätsch et al., 2000] proposent une méthode de sélection des paramètres des OC-SVM : il suffit d'ajouter quelques éléments de la classe négative dans la base d'apprentissage et de choisir le couple de paramètres qui n'englobe que des éléments positifs et sépare le mieux les deux classes.

## 2.3 Conclusion sur les techniques SVM

Les techniques SVM et OC-SVM ont de nombreux avantages par rapport aux algorithmes de classification usuels :

- la phase d'apprentissage consiste en la résolution d'un problème d'optimisation quadratique convexe (cf. Eq. 2.19) sous contraintes linéaires, qui a donc un unique maximum.
- la surface de décision est définie entièrement par la position des Vecteurs Supports, dont le nombre peut être beaucoup plus petit que la taille de la base d'apprentissage,
- la complexité du problème d'optimisation (Eq. 2.19) ne dépend pas de la dimension de l'espace des données, mais de la taille de la base d'apprentissage. Ce qui permet d'utiliser avec succès les SVM pour les problèmes de grandes dimensions, notamment dans les problèmes de classification d'images hyperspectrales [Melgani and Bruzzone, 2004; Fauvel, 2007].



## Chapitre 3

# Expérimentation et comparaison des techniques de classification

Les différentes méthodes de classification que nous venons d'exposer sont testées dans ce chapitre. Dans un premier temps, elles sont testées sur des données simulées afin de comprendre leur comportement. Puis, elles sont utilisées avec des données réelles.

### 3.1 Test sur des "exemples jouets"

En pratique, les formes géométriques des nuages de points appartenant à une même classe peuvent souvent être décrites par des ellipsoïdes. L'hypothèse de distribution gaussienne constitue donc une bonne approximation de données réelles.

Les quatre techniques de classification présentées précédemment sont testées en dimension 2 sur des mélanges de gaussiennes de moyennes et variances différentes. La figure 3.1 représente la base d'apprentissage (à gauche) et la base de test (à droite), générées toutes les deux par le mélange de gaussiennes. La première classe est composée d'une seule gaussienne (en bleu, de centre  $\{0,0\}$  et de variance  $\{1,1\}$ ) alors que la seconde est composée de deux gaussiennes (en vert, de centres  $\{2,2\}$ ,  $\{4,-1\}$  et de variances  $\{0.25,0.75\}$ ,  $\{2, 0.5\}$ ). Dans un premier cas, 200 points sont tirés aléatoirement pour chaque classe afin de former la base d'apprentissage pour les algorithmes supervisés. La base de test est composée de 2000 points par classe.

La figure 3.2 montre les résultats de classification pour chaque technique : K-moyennes (en haut à gauche), K-PPV (en haut à droite), MV (en bas à gauche), SVM (en bas à droite). La classification par OC-SVM n'est pas présentée ici car cette technique se comporte de la même façon que les SVM. De plus, le tableau 3.1 donne les moyennes et variances des précisions de chaque classification sur la base de test pour 100 expérimentations (tirages aléatoires de la base d'apprentissage).

Concernant l'algorithme de K-moyennes, le paramètre  $K$  a été fixé à 3 afin de déterminer les trois gaussiennes. Les trois gaussiennes sont grossièrement détectées. Néanmoins, les frontières de décision de l'algorithme, déterminées automatiquement, sont mal évaluées

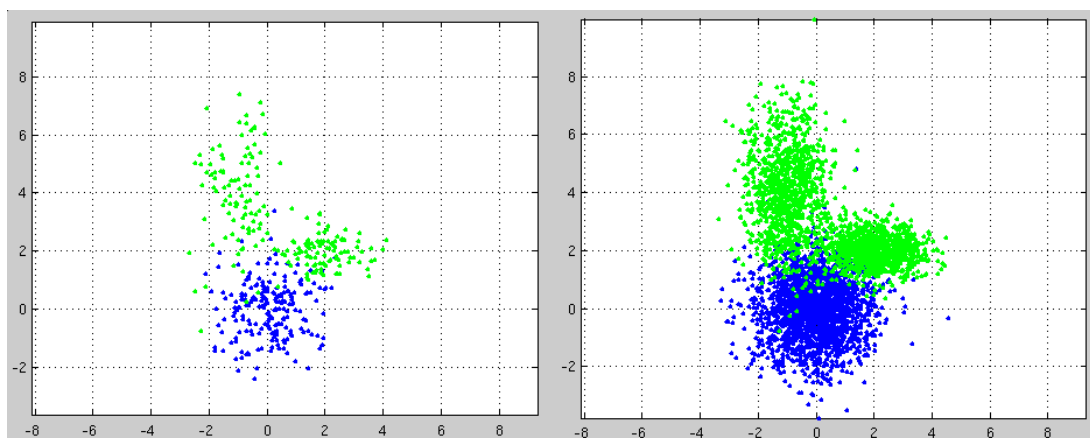


FIG. 3.1: Base d'apprentissage et base de test pour un mélange de gaussiennes

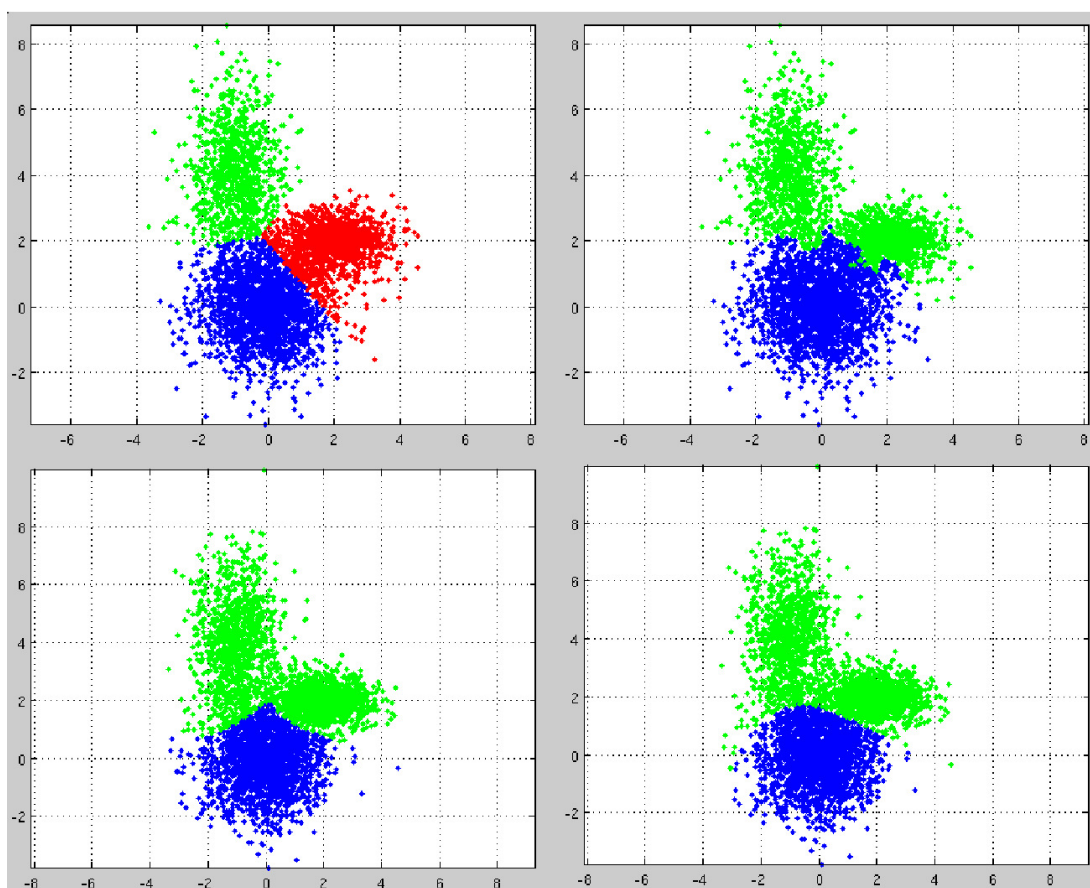


FIG. 3.2: Résultat de la classification pour les algorithmes K-moyennes (en haut à gauche), K-PPV (en haut à droite), MV (en bas à gauche), SVM (en bas à droite)

Taille de la base d'apprentissage :		10	20	50	100	200
K-moyennes	moyenne (%)	91.49	91.54	91.52	91.58	91.43
	variance (%)	0.54	0.48	0.56	0.55	0.52
	temps (s)	3	3	3	3	3
K-PPV	moyenne (%)	90.49	92.13	92.60	93.11	93.15
	variance (%)	3.30	1.84	1.24	0.94	0.80
	temps (s)	45	47	52	60	75
MV	moyenne (%)	88.76	93.02	94.06	94.28	94.25
	variance (%)	4.47	1.38	0.60	0.46	0.43
	temps (s)	36	36	36	36	36
SVM	moyenne (%)	<b>92.95</b>	<b>93.73</b>	<b>94.11</b>	<b>94.36</b>	<b>94.43</b>
	variance (%)	1.29	0.73	0.44	0.30	0.38
	temps (s)	11	14	22	44	108

TAB. 3.1: Moyenne, variance de la précision et temps de calcul des différentes techniques pour 100 expérimentations

et sont des droites. Comme cet algorithme est automatique et ne nécessite pas de base d'apprentissage, il est normal que la moyenne de sa précision et son temps de calcul soient indépendants de la taille de la base d'apprentissage. Ils varient cependant au cours des 100 itérations à cause de l'initialisation aléatoire des centroïdes.

Concernant l'algorithme de K-PPV, la classification n'est pas trop affectée par le chevauchement des gaussiennes au centre de la figure car un paramètre  $K$  élevé considère les éléments isolés de la base d'apprentissage comme du bruit (par exemple, l'élément de la classe bleue de coordonnées  $[0.25, 3.35]$ ). Il a été fixé à 5 lors de ces expérimentations. Cependant, les frontières dépendent énormément de la position des points de la base d'apprentissage et ne sont pas du tout régulières. Le temps de calcul est proportionnel à la taille de la base d'apprentissage, ce qui est logique puisque l'algorithme calcule la distance entre chaque nouvel objet et chaque élément de la base d'apprentissage.

Concernant l'algorithme du Maximum de Vraisemblance, on cherche à déterminer les trois gaussiennes à partir des moyennes et matrices de covariance estimées grâce à la base d'apprentissage. Cette technique de classification donne de bons résultats, à peu près équivalents à ceux des SVM dans le cas d'une base d'apprentissage de grande taille. Cependant, elle nécessite une hypothèse sur le type de distribution des données. Dans le cas d'une base d'apprentissage de petite taille, les estimations des paramètres de la distribution supposée ne sont pas assez précises, ce qui rend les résultats plus instables et moins précis. Le temps de calcul de cette méthode ne dépend presque pas de la taille de la base d'apprentissage puisqu'elle ne sert qu'à calculer la moyenne et la matrice de covariance pour chaque classe.

Concernant les SVM, le noyau gaussien a été utilisé dans les expérimentations et les paramètres des SVM ont été choisis de façon à permettre des erreurs de classification de la



base d'apprentissage ( $\sigma$  grand et  $C$  petit). Cette technique fournit les meilleurs résultats quelle que soit la taille de la base d'apprentissage. Son temps de calcul varie cependant exponentiellement avec la taille de la base d'apprentissage.

## 3.2 Test sur des images SPOT 5

Les images utilisées dans les expérimentations sont des images SPOT 5 représentant des régions du Sud de la France. Elles ont, soit une résolution de 10 mètres et sont alors composées des bandes Vert, Rouge, Proche Infra-Rouge (PIR) et Moyen Infra-Rouge (MIR), soit une résolution de 2,5 mètres et sont composés des bandes Vert, Rouge et PIR. Dans ce cas, elles sont réalisées par le CNES en fusionnant une image panchromatique Supermode de résolution 2,5 mètres avec l'image 4 bandes correspondante de résolution 10 mètres.

Le tableau 3.2 donne des informations sur les données : le lieu, la superficie brûlée en hectares, l'année, l'ordre du délai entre l'extinction du feu et l'acquisition de l'image par le satellite, la résolution de l'image ainsi que le type de carte de référence correspondant à l'image, i.e. si elle a été obtenue par GPS ou par Photo-Interprétation (PI, dans ce cas, la résolution de l'image dont dérive la carte est donnée) ou si l'on dispose des deux. L'image n°1 dont nous disposons et qui nous a été fournie par le CNES a cependant subi une rehaussement de contraste.

La dénomination "carte de référence" désigne ici une cartographie des zones brûlées réalisée par des experts, soit directement sur le terrain, soit par interprétation visuelle et/ou automatique d'une image de télédétection.

La carte de référence de l'image n°1 nous a été fournie par le CNES et dérive de l'image SPOT 5 correspondante. Celle de l'image n°2 nous a été fournie par l'ONF et a été obtenue par GPS. Pour les images de la Corse, le SDIS2B nous a fourni les contours GPS des plus grandes zones brûlées et Infoterra France nous a fourni leurs estimations des zones brûlées de l'année 2005 obtenues à partir d'images satellitaires Landsat (30 mètres de résolution). Pour la dernière image, le SDIS06 nous a fourni les contours GPS de certaines zones brûlées; la cartographie réalisée par le SERTIT <sup>1</sup> dérive de l'image SPOT 5 correspondante.

Le tableau 3.3 donne les différents types de végétation qui composaient les régions brûlées pour chaque image, selon la base de données Prométhée. Concernant les feux du Massif des Maures et de Corse de l'année 2003, les cartographies des types de peuplement des zones incendiées sont données dans la revue l'If [Inventaire Forestier National, 2003].

### 3.2.1 Réglage des SVM

Cette partie traite du réglage des SVM (choix de la base d'apprentissage, du noyau, des paramètres).

---

<sup>1</sup>disponible à l'adresse internet :

[http://sertit.u-strasbg.fr/SITE\\_Charte/2007/3\\_france\\_feux\\_juillet\\_2007/sud\\_france\\_feux\\_2007.html](http://sertit.u-strasbg.fr/SITE_Charte/2007/3_france_feux_juillet_2007/sud_france_feux_2007.html)

$n^{\circ}$	Zone	ha	Année	Délai	Résolution	Carte de référence
1	Massif des Maures	19.000	2003	1 jour	2.5 m	PI - 2.5m
2	Lucéram	2.000	2003	2 mois	10 m	GPS
3	Nord de la Corse	20.000	2003	1 jour	10 m	GPS
4	Ouest de la Corse	2.000	2005	1 mois	10 m	GPS et PI - 30m
5	Est de la Corse	2.000	2005	1 mois	10 m	GPS et PI - 30m
6	Alpes-Maritimes	500	2007	1 jour	10 m	GPS et PI - 10m

TAB. 3.2: Données satellitaires et cartes de référence

$n^{\circ}$	Nom de l'image	Type de végétation brûlée
1	Massif des Maures	landes, garrigues, maquis, futaie (feuillue, résineuse, mélangée)
2	Lucéram	futaie résineuse
3	Nord de la Corse	landes, garrigues, maquis
4	Ouest de la Corse	landes, garrigues, maquis, futaie feuillue
5	Est de la Corse	futaie résineuse
6	Alpes-Maritimes	landes, garrigues, maquis, futaie résineuse

TAB. 3.3: Types de végétation

$n^{\circ}$	Nom de l'image	Taille de l'image	Nombre bandes	Résolution	Taille BA pos.	Taille BA neg.
1	Massif des Maures	10.600*4.500	3	2.5 m	220	330
2	Lucéram	256*256	4	10 m	50	50
3	Nord de la Corse	1.400*1.000	4	10 m	110	150
4	Ouest de la Corse	2.200*2.000	4	10 m	90	110
5	Est de la Corse	2.500*2.500	4	10 m	110	150
6	Alpes-Maritimes	2.200*1.200	4	10 m	120	130

TAB. 3.4: Informations sur les images et bases d'apprentissage (BA)

### Base d'apprentissage

La base d'apprentissage doit décrire de façon représentative et complète les deux classes. Plusieurs itérations peuvent être nécessaires afin de la sélectionner manuellement [Mitra et al., 2004]. A chaque itération, les SVM sont appliqués sur une petite zone de l'image et les pixels les plus caractéristiques et mal classés sont ajoutés à la base d'apprentissage. Comme les résultats dépendent des paramètres des SVM, les pixels qui viennent d'être ajoutés peuvent être mal classés lors de l'itération suivante et ne doivent pas être ajoutés à la base d'apprentissage une nouvelle fois : ils peuvent être bien classés avec un autre couple de paramètres.

La base d'apprentissage est formée d'ensembles connexes de 10 pixels. Pour chaque image, seules des zones (brûlées et non-brûlées) visiblement distinctes ont été sélectionnées afin de former la base d'apprentissage.

Le tableau 3.4 donne des informations concernant les images (nom, nombre de pixels, nombre de bandes spectrales, résolution) et bases d'apprentissage utilisées dans cette thèse (taille des bases d'apprentissage positives et négatives).

### Choix du noyau

Un deuxième point-clé des SVM réside dans le choix du noyau et de ses paramètres. Il n'y a malheureusement pas un noyau optimal unique quels que soient les problèmes. Plusieurs noyaux ont été testés. Pour chaque noyau, leurs paramètres optimaux ont été déterminés empiriquement. Le tableau (3.5) fournit la Précision Totale (PT), les taux de Faux Positifs (FP) et Faux Négatifs (FN) entre la classification obtenue par SVM et la carte de référence pour plusieurs noyaux : polynomial (pol), Gaussien (Gaussien), combinaison linéaire de Gaussien et "Spectral Angle Mapper" (SAM+G) et KMOD. Le tableau donne aussi les statistiques obtenues à partir d'un SVM linéaire.

La définition des différentes statistiques du tableau sont donnés dans l'annexe F. Les temps de calcul sont donnés à titre indicatif. L'ordinateur utilisé possède un processeur de 2 Ghz avec 2 Go de RAM.

Carte de référence (2550 ha)	Surface brûlée (ha)	PT (%)	FP (%)	FN (%)	Temps (min)
SVM linéaire	8867	75.08	25.19	22.03	35
SVM pol.	2859	97.65	1.85	7.71	40
SVM Gaussien	2744	97.69	1.62	9.74	60
SVM SAM+G.	2745	97.50	1.72	10.80	90
SVM KMOD	2782	97.68	1.69	8.99	60

TAB. 3.5: Performances des SVM pour différents noyaux

Les différents noyaux donnent approximativement les mêmes résultats. Dans la suite, nous avons choisi d'utiliser le noyau Gaussien. Ce noyau, très populaire car très simple,

présente l'avantage de ne dépendre que d'un seul paramètre et ne nécessite pas un temps de calcul très long. De plus, il fait passer les données dans un espace de dimension infinie, ce qui permet de séparer linéairement n'importe quelle base d'apprentissage, contrairement au noyau polynomial.

### Sélection des hyperparamètres

Les paramètres du noyau et de régularisation sont appelés *hyperparamètres* des SVM. Dans cette partie, les hyperparamètres des SVM avec un noyau gaussien ( $C$  pour les marges relaxés et  $\sigma$  pour le noyau Gaussien) sont testés.

**Résultats expérimentaux :** La figure 3.3 donne le taux de bonne classification des éléments de la base d'apprentissage (précision d'apprentissage - figure de gauche) et la précision totale entre la classification par SVM et la carte de référence (figure de droite) en fonction des hyperparamètres. D'après cette figure, il existe une plage de couples d'hyperparamètres qui fournissent une bonne concordance entre la classification par SVM et la carte de référence.

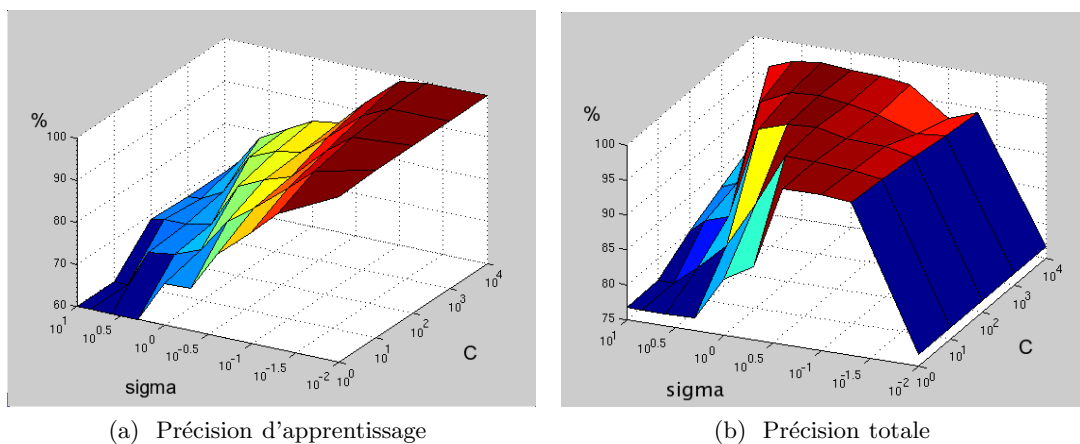


FIG. 3.3: Choix des hyperparamètres des SVM

L'hyperparamètre  $\sigma$  correspond à la variance du noyau Gaussien. Etudions plus précisément l'influence de chaque paramètre lors des phases d'apprentissage et de classification. Le paramètre  $C$  est fixé à 100 et nous faisons varier  $\sigma$  :

- si  $\sigma$  est trop petit, le nombre de Vecteurs Supports (VS) est grand et la forme du classifieur est très complexe puisqu'elle est excessivement corrélée avec les VS. Il s'agit du phénomène de *sur-apprentissage* : coller trop correctement aux données de la base d'apprentissage force le modèle à ne caractériser correctement que la base d'apprentissage et ne sera pas généralisable à l'espace tout entier des données. Le taux d'erreur d'apprentissage est presque nul, mais la précision totale est faible.
- lorsque  $\sigma$  augmente, la forme du séparateur devient moins complexe, ce qui permet de moins coller aux données de la base d'apprentissage et donc d'obtenir une

meilleure généralisation lorsque l'on désire classer de nouveaux objets. Le taux d'erreur d'apprentissage n'est plus nul, mais la précision totale est meilleure.

- si  $\sigma$  devient trop élevé, la forme du séparateur devient trop simple et ne prend pas assez en compte les informations apportées par la base d'apprentissage. Le taux d'erreur d'apprentissage n'est pas nul et la précision totale est faible.

Le paramètre de régularisation  $C$  définit le compromis entre la minimisation de la marge et l'erreur sur la base d'apprentissage. Il contrôle aussi l'équilibre entre la précision sur la base d'apprentissage et la complexité du classifieur. Maintenant, le paramètre  $\sigma$  est fixé à 1 et nous faisons varier  $C$  :

- si  $C$  est trop petit, le modèle permet une mauvaise classification des éléments de la base d'apprentissage. Les taux d'erreur d'apprentissage et de classification sont faibles puisque le modèle ne prend pas en compte les données.
- lorsque  $C$  augmente, le modèle pénalise plus les erreurs sur la base d'apprentissage. La précision totale augmente.
- si  $C$  devient très élevé, le modèle interdit, de plus en plus, les erreurs sur la base d'apprentissage, ce qui peut impliquer le phénomène de sur-apprentissage. La précision totale peut donc diminuer. Cependant, lorsque  $C$  croît vers l'infini, la précision totale devient constante puisque le maximum du problème de programmation quadratique a été atteint (cf Equ. (2.14)).

**Validation croisée :** Une manière courante pour choisir les paramètres est la technique de *validation croisée* ("K-fold cross validation") [Schölkopf et al., 2004]. Elle consiste à diviser la base d'apprentissage en  $K$  sous-parties, de faire l'apprentissage des SVM sur  $K-1$  sous-parties et de tester le classifieur obtenu sur la  $K^{me}$  sous-partie. Ce processus est répété  $K$  fois, avec chacune des  $K$  sous-parties. Les  $K$  tests donnent le taux d'erreur de validation croisée pour un couple de paramètres. Le couple de paramètres choisi est celui qui fournit la meilleure précision de validation croisée.

Les valeurs des hyperparamètres à tester doivent couvrir un très large domaine. [Hsu et al., 2003] proposent de choisir des valeurs croissantes exponentiellement pour  $\sigma$  et  $C$  (par exemple  $\sigma = 2^{-5}, 2^{-4}, \dots, 2^5$ ), les couples testés correspondant aux intersections de la grille définie par ces paramètres.

Si le noyau choisi utilise plusieurs paramètres (ce qui n'est pas le cas du noyau Gaussien que nous avons choisi), la grille de recherche devient trop complexe puisque le nombre de points d'intersection de la grille croît exponentiellement avec le nombre de paramètres. Dans ce cas, les méthodes fondées sur la minimisation de la "VC-dimension" [Chapelle et al., 2002] permettent de fournir de bons résultats.

La figure 3.4 donne la précision de validation croisée et la précision totale en fonction des hyperparamètres. La technique de validation croisée ne fournit pas le meilleur couple de paramètres. Minimiser l'erreur sur une sous-partie de la base d'apprentissage n'est pas équivalent à minimiser l'erreur totale.

Si le nombre de parties  $K$  est trop élevé : la base de test sera trop petite, le biais du test sera faible, mais sa variance sera grande et la validation croisée nécessitera un grand temps de calcul.

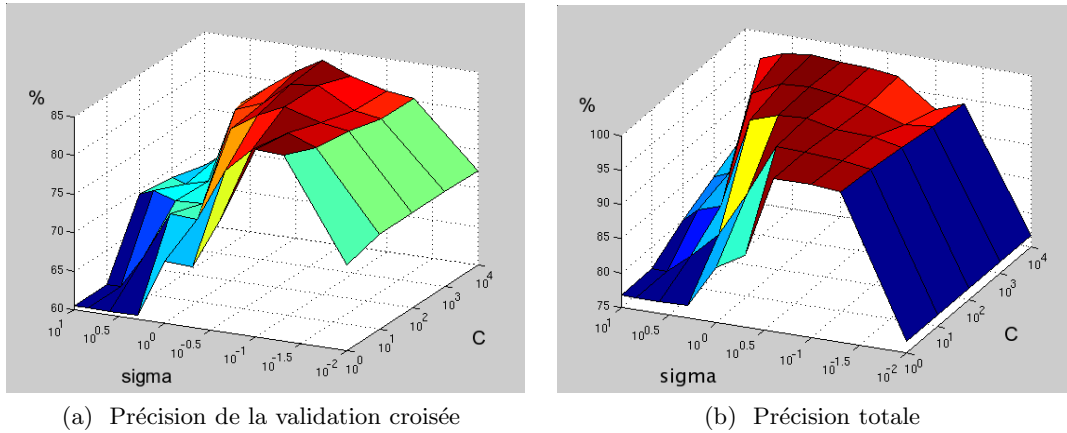


FIG. 3.4: Validation croisée.

Si le nombre de parties  $K$  est faible : le temps de calcul sera réduit, de même que la variance du test mais le biais sera grand (le nombre d'éléments de la base d'apprentissage n'est pas suffisant comparé aux nombres d'éléments de la base de test).

En pratique, le nombre de parties  $K$  dépend de la taille de la base d'apprentissage. Pour de grandes bases, la "3-fold cross validation" est précise. Pour des bases de petite taille, il faut couper la base d'apprentissage en de plus petites parties afin de réaliser l'apprentissage sur beaucoup d'éléments.

Les statistiques et figures présentées précédemment ont été obtenues avec l'image satellitaire "Massif des Maures". Les autres images satellitaires fournissent les mêmes conclusions concernant le choix du noyau et des hyperparamètres.

### 3.2.2 Réglage des OC-SVM

Concernant la taille de la base d'apprentissage ou le choix du noyau, les OC-SVM se comportent de la même façon que les SVM, ce qui est logique puisque l'algorithme des OC-SVM dérive de celui des SVM.

Cependant, les OC-SVM se comportent un peu différemment en fonction du couple d'hyperparamètres utilisé.

La figure 3.5 fournit le taux de bonne classification des éléments de la base d'apprentissage (figure de gauche) et la précision totale entre la classification par OC-SVM et la carte de référence (figure de droite) en fonction des hyperparamètres ( $\nu$  pour les marges relaxées et  $\sigma$  pour le noyau Gaussien).

D'après l'équation 2.21,  $\nu$  est une limite supérieure du taux d'outliers et une limite inférieure du taux de Vecteurs Supports (VS) contenu dans la base d'apprentissage [Schölkopf et al., 2001] :

$$\nu \leq \frac{l_{VS}}{l} \quad \text{et} \quad \frac{l_{mc}}{l} \leq \nu \quad (3.1)$$

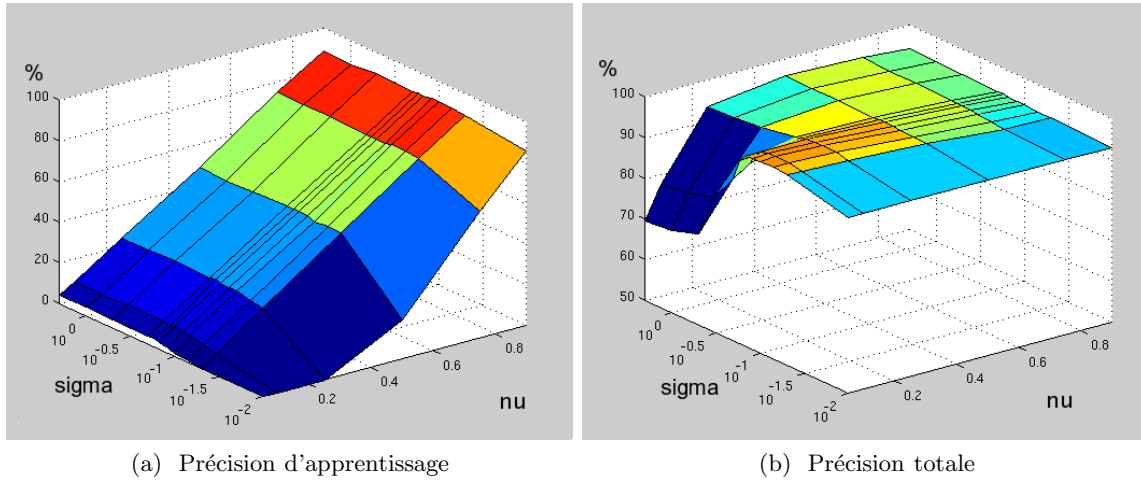


FIG. 3.5: Choix des hyperparamètres des OC-SVM.

La figure 3.5a et le tableau 3.6, qui fournit les taux d'outliers et de Vecteurs Supports (i.e. le pourcentage d'éléments mal classifiés et de VS dans la base d'apprentissage) en fonction de  $\nu$ , montrent que  $\nu$  permet bien de contrôler le taux d'outliers. D'après la figure 3.5a, le taux d'outliers est quasiment constant lorsque  $\sigma$  varie à  $\nu$  fixé et il varie proportionnellement à  $\nu$  lorsque  $\sigma$  est fixé.

Taux d'outliers (%)	1.67	8.33	24.17	47.50	75.00	90.00	100.00
$100 * \nu$	5	10	25	50	75	90	100
Taux de VS (%)	9.16	14.16	28.33	51.66	75.83	91.66	100.00

TAB. 3.6: Choix des hyperparamètres des OC-SVM

D'après la figure 3.5b, la plage de couples d'hyperparamètres qui fournissent une bonne concordance entre la classification par OC-SVM et la carte de référence est beaucoup plus restreinte que dans le cas de la classification par SVM (figure 3.3).

On peut remarquer que cette plage correspond :

- à des faibles valeurs de  $\nu$ , qui est une borne supérieure du taux d'outliers contenu dans la base d'apprentissage,
- à des valeurs de  $\sigma$  proches de la variance des vecteurs de la base d'apprentissage.

La fonction de décision des OC-SVM est une somme de gaussiennes (centrées sur  $\vec{x}_i$ , de variance  $\sigma$  et pondérées par des coefficients positifs ou nuls  $\lambda_i$ ) et d'un biais  $\rho$  tel que la fonction soit nulle aux points  $\vec{x}_i$  pour  $0 < \lambda_i < \frac{1}{l}$  :

$$f(\vec{x}) = \sum_{i=1}^l \lambda_i \exp\left(-\frac{\|\vec{x} - \vec{x}_i\|^2}{2\sigma^2}\right) - \rho$$

Lorsque  $\sigma$  est petit, la fonction de décision est complexe car les gaussiennes sont fines. Lorsque  $\sigma$  augmente, les gaussiennes grossissent et la forme de la fonction de décision

se rapproche de celle d'une sphère qui contient au moins  $(1 - \nu) * l$  éléments de la base d'apprentissage. Lorsque la valeur de  $\sigma$  est éloignée de celle de la variance des vecteurs de la base d'apprentissage, les gaussiennes sont soit trop fines, soit trop larges.

La figure 3.6 permet de comprendre la forme de la fonction de décision des OC-SVM en fonction du couple de paramètres utilisés :  $\sigma = \{0.02, 0.1, 0.2\}$  et  $\nu = \{0.01, 0.2, 0.5\}$ . Les points et petits cercles bleus correspondent respectivement aux éléments de la base d'apprentissage classés positivement et négativement par le contour en rouge. La valeur  $\sigma = 0.1$  (figures de la deuxième ligne) correspond à la variance des éléments de la base d'apprentissage.

Afin de sélectionner le couple d'hyperparamètres, une autre méthode proposée par [Rätsch et al., 2000] consiste à ajouter quelques éléments de la classe négative dans la base d'apprentissage et de choisir le couple de paramètres qui n'englobe que des éléments positifs et sépare le mieux les deux classes.

Les statistiques et figures présentées dans cette partie ont été obtenues avec l'image satellitaire "Alpes-Maritimes". Les autres images satellitaires fournissent les mêmes conclusions concernant le choix des hyperparamètres.

### 3.2.3 Résultats des différentes techniques de classification

Concernant l'algorithme des K-moyennes, le choix de  $K = 2$  ne donne pas de bons résultats. En effet, les lacs, rivières et routes présentent des caractéristiques spectrales très proches de celles des zones brûlées, un algorithme K-moyennes à 2 classes regroupe ces objets avec les zones brûlées. Cependant, si  $K$  est très élevé, l'algorithme est très long à converger. De plus, comme plusieurs classes correspondent aux zones brûlées, il faudra ensuite les détecter visuellement puis les regrouper afin d'obtenir la classification binaire en "brûlé" et "non-brûlé".

Le nombre de classes  $K$  est déterminé automatiquement selon le modèle proposé par [Palubinskas et al., 1998]. Il varie généralement entre 10 et 15 selon l'image utilisée. Ce nombre de classes semble raisonnable en comparaison avec le nombre de catégories d'objets dans l'image (zones brûlées, végétation saine, végétation moins saine, champs, zones urbaines, route, eau turbide, eau claire, nuage, ombres). Si l'on souhaite utiliser l'algorithme des K-moyennes classique, l'initialisation aléatoire de l'algorithme ou un choix de valeurs de  $K$  entre 10 et 15 ne changent pas considérablement la classification obtenue.

Concernant l'algorithme des K-Plus Proches Voisins (K-PPV), le nombre de voisins  $K$  est choisi empiriquement afin d'obtenir les meilleurs résultats possibles avec la base d'apprentissage utilisée par les SVM.

La base d'apprentissage utilisée est la même pour les 4 méthodes de classification supervisée (mais seuls les éléments positifs sont utilisés pour les OC-SVM).

#### - Massif des Maures :

La figure 3.7 montre les résultats obtenus à partir des 5 techniques de classification présentées dans le chapitre précédent. Elles classifient un pixel de l'image uniquement à



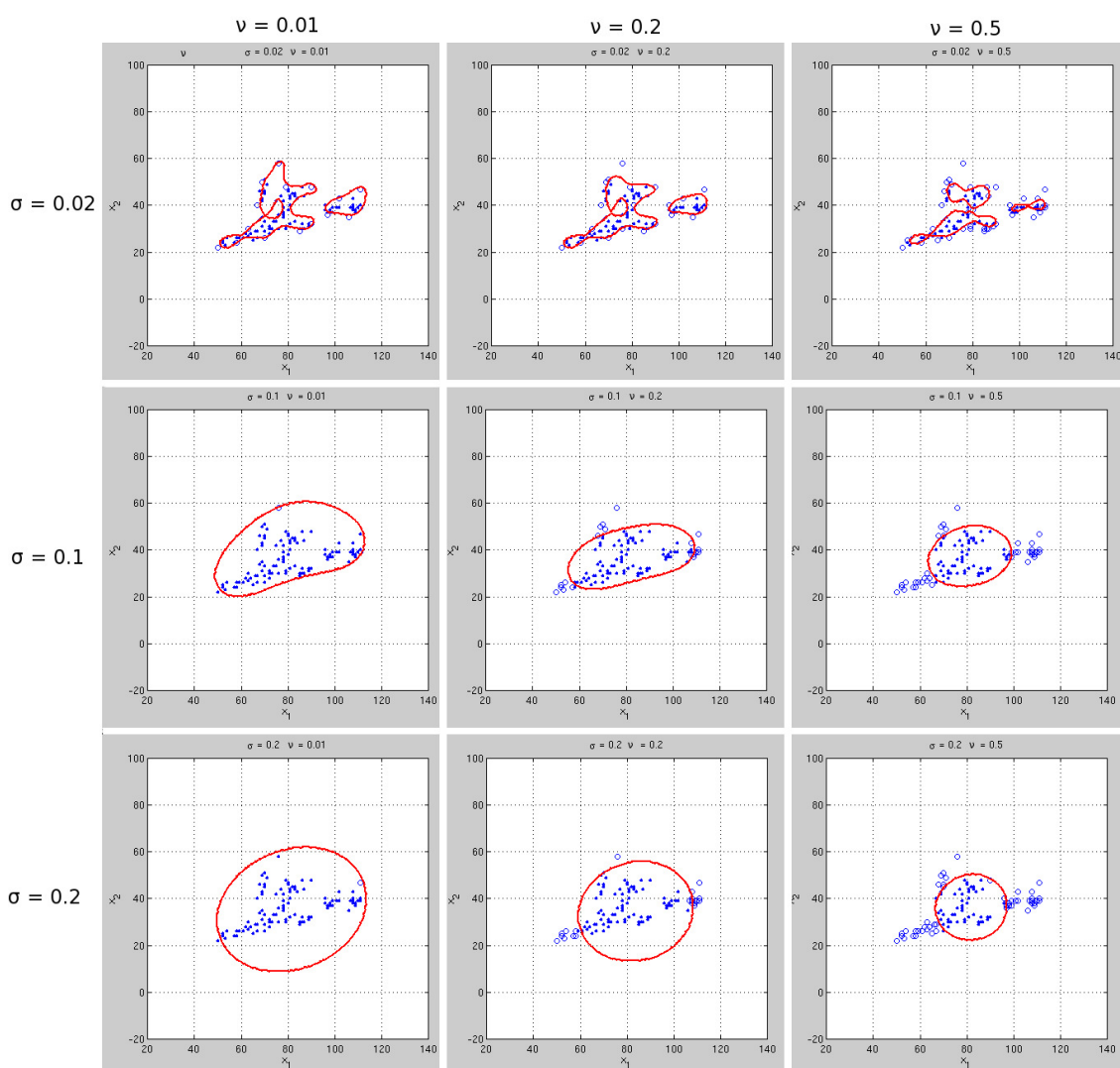


FIG. 3.6: Fonctions de décisions des OC-SVM en fonction du couple de paramètres

partir des trois réflectances spectrales du pixel.

Le tableau 3.7 donne quelques statistiques sur les résultats obtenus à partir des cinq techniques de classification. Il s'agit de la surface brûlée extraite, de la Précision Totale (PT : taux de pixels classés en accord avec la carte de référence dans toute l'image), de la Précision Moyenne (PM : moyenne des taux de précision sur chaque classe), des taux de Faux Positifs (FP : taux de pixels mal classés positifs parmi les pixels négatifs de la carte de référence) et de Faux Négatifs (FN : taux de pixels mal classés négatifs parmi les pixels positifs de la carte de référence), et du temps de calcul (les formules permettant de calculer ces statistiques sont données dans l'annexe F).

Concernant les quatre algorithmes de classification supervisée, les SVM fournissent les meilleurs taux de précisions totale et moyenne. Les OC-SVM ont une meilleure précision totale que les algorithmes du Maximum de Vraisemblance et des K Plus Proches Voisins. Pour cette image, l'algorithme des K-moyennes fournit de très bons résultats car la taille de la zone brûlée est très grande comparée à celle de l'image satellitaire, et ses caractéristiques spectrales sont assez éloignées de celles des autres objets présents dans l'image.

#### - Ouest de la Corse :

La figure 3.8 montre les résultats obtenus à partir des cinq techniques pour l'image "Ouest de la Corse".

Le tableau 3.8 donne quelques statistiques sur les résultats obtenus à partir des cinq techniques de classification.

Concernant les quatre algorithmes de classification supervisée, les SVM fournissent les meilleurs taux de précisions totale et moyenne. Les OC-SVM montrent une précision totale moins bonne que celles des algorithmes du Maximum de Vraisemblance et des K Plus Proches Voisins. Pour cette image, l'algorithme des K-moyennes fournit aussi beaucoup de faux positifs (la taille de la zone brûlée est assez petite comparée à celle de l'image et le recouvrement spectral des zones brûlées et non brûlées est assez important).

On peut remarquer que les taux de Faux Positifs sont très faibles contrairement aux taux de Faux Négatifs alors que les différentes cartographies semblent montrer le contraire. En effet, le nombre de pixels faux positifs est de 44.500, celui de faux négatifs est de 76.700 par exemple pour les SVM. Les différences de taux s'expliquent par le fait que le nombre de pixels négatifs (non-brûlés dans la carte de référence) est très élevé (environ 4.160.000 pixels), alors que le nombre de pixels positifs (brûlés dans la carte de référence) est 14 fois plus faible (environ 240.000 pixels) pour cette image.

#### - Est de la Corse :

La figure 3.9 montre les résultats obtenus à partir des cinq techniques pour l'image "Est de la Corse".

Le tableau 3.9 donne quelques statistiques sur les résultats obtenus à partir des cinq techniques de classification.

Concernant les quatre algorithmes de classification supervisée, les SVM fournissent les meilleurs taux de précisions totale et moyenne. Les OC-SVM ont encore une meilleure précision totale que les algorithmes du Maximum de Vraisemblance et des K Plus Proches Voisins. Pour cette image, l'algorithme des K-moyennes fournit beaucoup de faux positifs.

On peut noter que l'algorithme du MV classe l'ombre du nuage comme une zone brûlée. L'algorithme du MV estime l'appartenance de l'ombre du nuage avec les deux gaussiennes définies par les bases d'apprentissage positive et négative. En considérant les caractéristiques spectrales de l'objet, sa probabilité d'appartenance à la gaussienne "brûlée" est supérieure à sa probabilité d'appartenance à la gaussienne "non brûlée".

#### - Alpes-Maritimes :

La figure 3.10 montre les résultats obtenus à partir des cinq techniques pour l'image "Alpes-Maritimes"

Le tableau 3.10 donne quelques statistiques sur les résultats obtenus à partir des cinq techniques de classification.

Concernant les quatre algorithmes de classification supervisée, les SVM fournissent les meilleurs taux de précisions totale et moyenne. Les OC-SVM montre une précision totale moins bonne que celle des algorithmes du Maximum de Vraisemblance, mais supérieure à celle des K Plus Proches Voisins. On peut noter que l'algorithme du MV classe l'aéroport comme une zone brûlée (l'aéroport est spectralement plus proche du nuage de points, supposé gaussien, de la zone brûlée que de celui de la zone non brûlée).

### 3.3 Analyse des erreurs

Les zones brûlées extraites montrent une bonne concordance avec les cartes de référence. Cependant, il existe des différences qui peuvent être expliquées de plusieurs manières :

- erreurs sur les zones brûlées extraites :
  - une mauvaise classification d'objets qui ont les mêmes caractéristiques spectrales que les zones brûlées (exemple : bords de mer ou de routes cf. fig 3.7),
  - certaines zones de végétation sont faiblement touchées par les feux et ressemblent à de la végétation non saine ou sèche,
  - des erreurs sur les frontières des zones brûlées : un pixel d'une image est classifiée comme brûlé si les houppiers (ensemble des parties aériennes d'un arbre : branches, rameaux et feuillage d'un arbre) sont brûlés, un arbre est classifié comme brûlé dans le relevé de terrain si son tronc a été touché par les flammes.
- erreurs sur les cartes de référence :
  - le contour est obtenu à partir d'un certain nombre de points relevés par GPS ; opération délicate si la zone brûlée est dans une région montagneuse,
  - certaines zones non brûlées à l'intérieur de la zone brûlée sont comptées comme brûlées. Ces zones peuvent être de la végétation non brûlée (en fonction des conditions météo, de la topographie du terrain, du type de végétation, de l'activité des pompiers), mais aussi des routes, des bâtiments, etc.

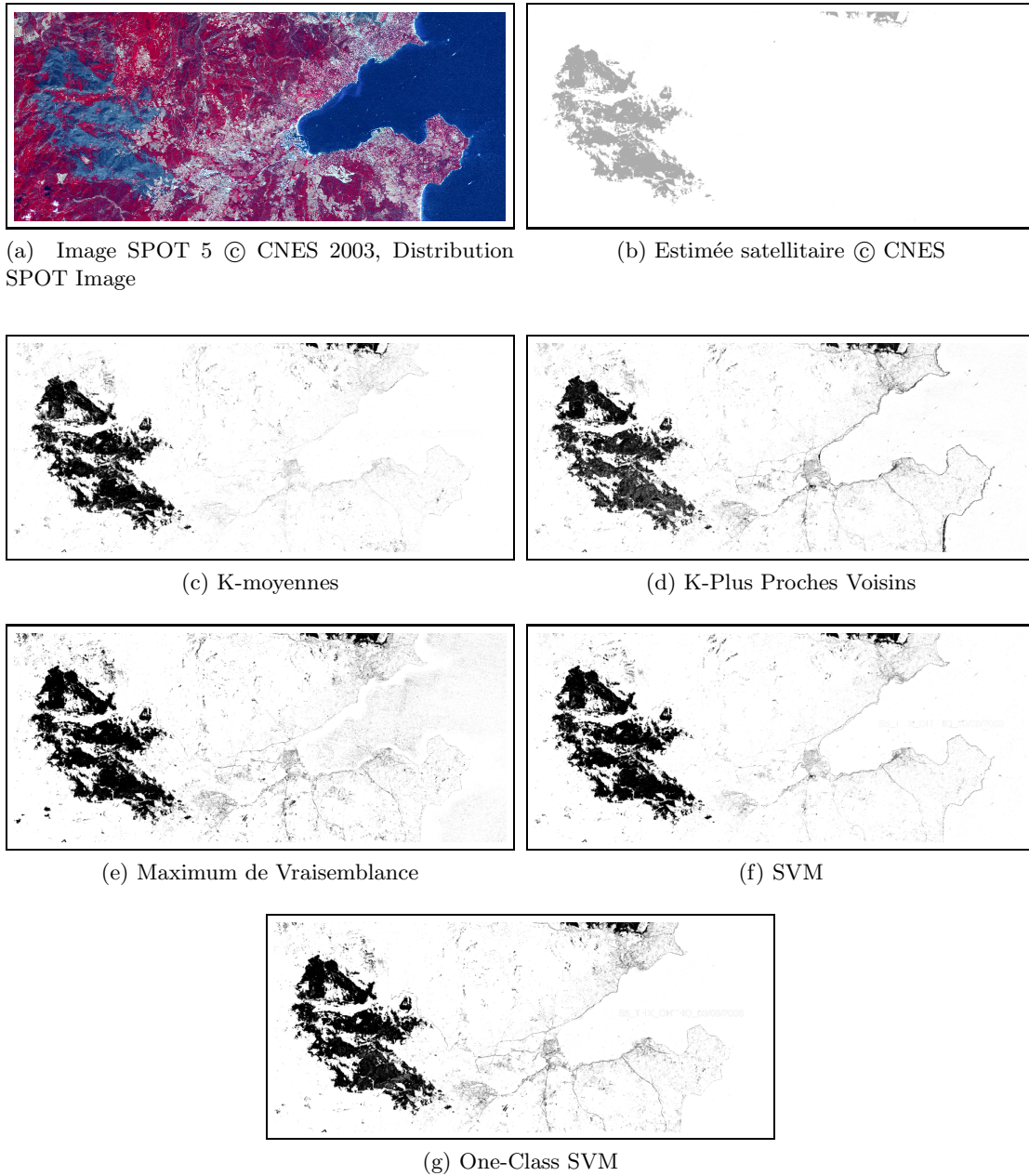
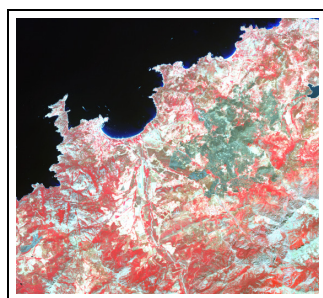


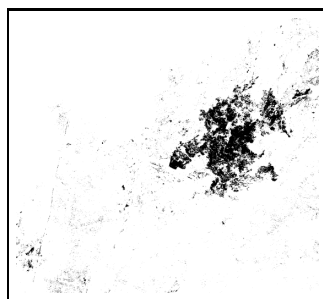
FIG. 3.7: Massif des Maures - 2003 ( 2500 hectares)



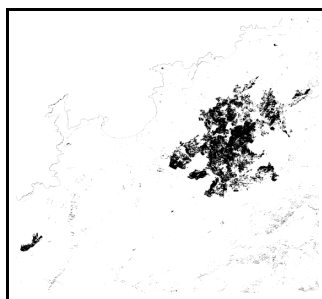
(a) Image SPOT 5 © CNES 2005, Distribution SPOT Image



(b) Estimée satellitaire - projet RISK-EOS, © Infoterra - ESA



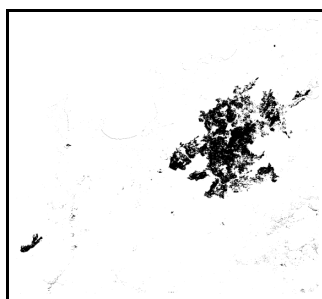
(c) K-moyennes



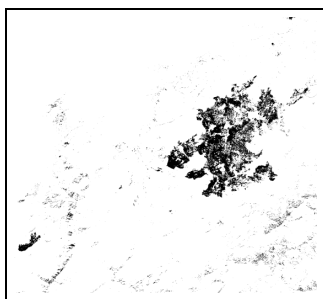
(d) K-Plus Proches Voisins



(e) Maximum de Vraisemblance

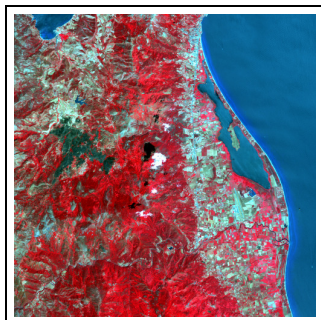


(f) SVM

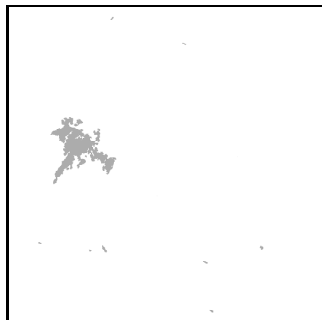


(g) One-Class SVM

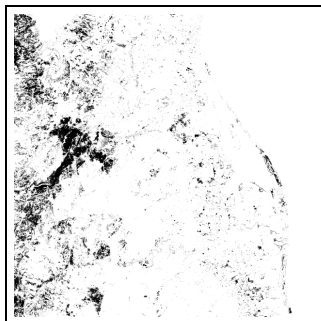
FIG. 3.8: Ouest de la Corse - 2005 (environ 2800 hectares)



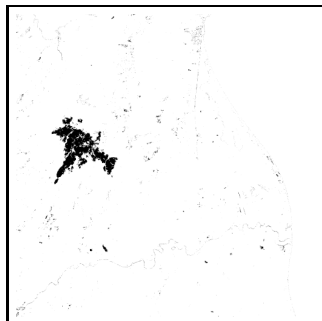
(a) Image SPOT 5 © CNES 2005, Distribution SPOT Image



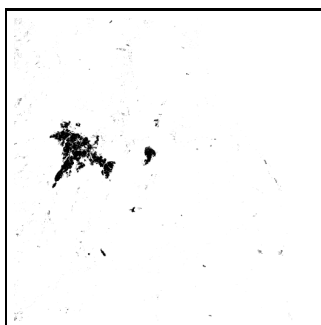
(b) Estimée satellitaire - projet RISK-EOS, © Infoterra - ESA



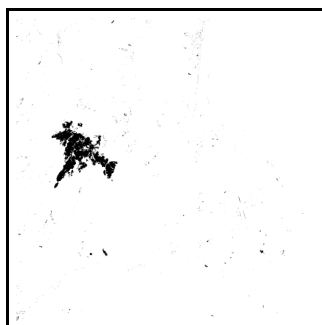
(c) K-moyennes



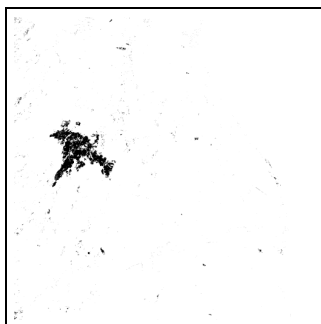
(d) K-Plus Proches Voisins



(e) Maximum de Vraisemblance

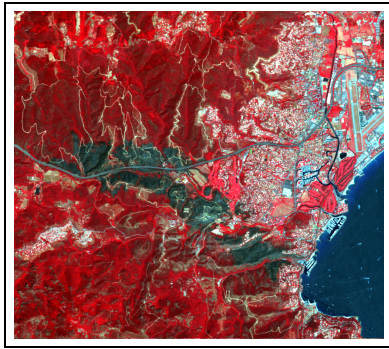


(f) SVM



(g) One-Class SVM

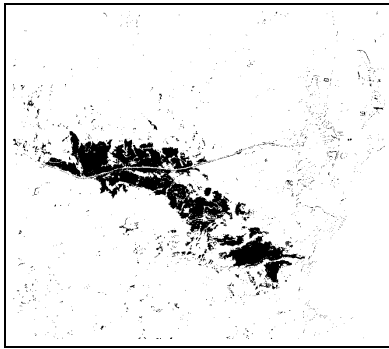
FIG. 3.9: Est de la Corse - 2005 (environ 1000 hectares)



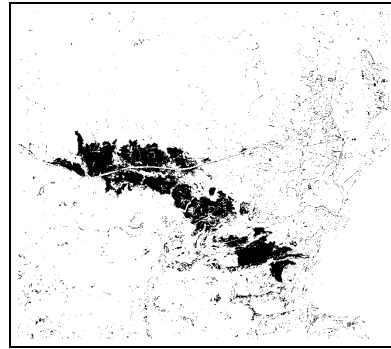
(a) Image SPOT 5 © CNES 2007,  
Distribution SPOT Image



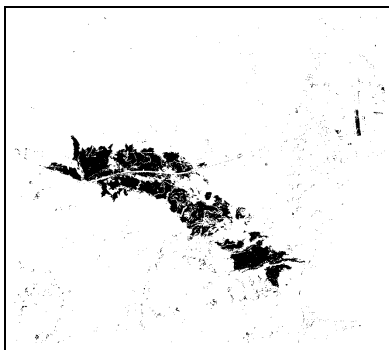
(b) Estimée satellitaire © SERTIT



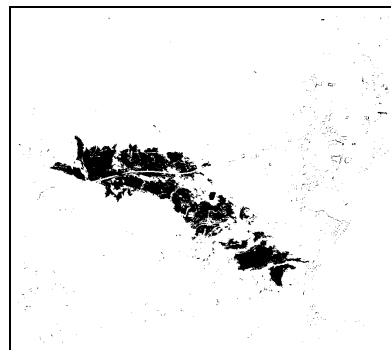
(c) K-moyennes



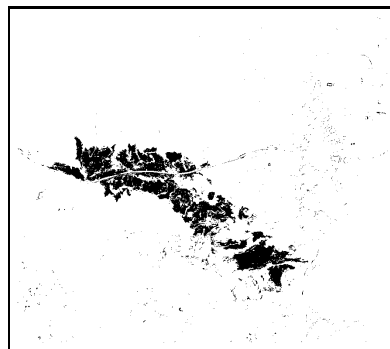
(d) K-Plus Proches Voisins



(e) Maximum de Vraisemblance



(f) SVM



(g) One-Class SVM

FIG. 3.10: Alpes Maritimes - 2007 (environ 450 hectares)

Carte de référence (2550 ha)	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
K-moyennes	2489	97.49	91.42	1.26	15.89	45
K-PPV	2666	95.98	88.19	2.41	21.21	30
MV	3201	96.32	94.00	3.21	8.78	5
SVM	2744	97.69	94.32	1.62	9.74	60
OC-SVM	2960	96.96	93.93	2.41	9.74	15

TAB. 3.7: Massif des Maures - 2003 : Comparaison avec la carte de référence

Carte de référence (2847 ha)	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
K-moyennes	2406	96.43	81.64	1.37	35.34	20
K-PPV	2243	97.00	82.68	0.87	33.77	3
MV	1824	97.04	79.43	0.34	40.81	1
SVM	2384	97.56	86.15	0.74	26.96	5
OC-SVM	2847	96.39	80.74	1.28	37.25	2

TAB. 3.8: Ouest de la Corse - 2005 : Comparaison avec la carte de référence

Carte de référence (992 ha)	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
K-moyennes	4638	93.42	85.14	6.31	23.41	30
K-PPV	1061	99.31	90.78	0.40	18.03	3
MV	1041	99.28	89.71	0.40	20.17	1
SVM	994	99.48	91.76	0.26	16.22	5
OC-SVM	924	99.33	87.51	0.27	24.70	2

TAB. 3.9: Est de la Corse - 2005 : Comparaison avec la carte de référence

Carte de référence (457 ha)	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
K-moyennes	461	97.34	90.19	1.46	18.16	20
K-PPV	483	96.29	87.37	2.21	23.04	3
MV	375	97.63	86.99	0.58	25.43	1
SVM	382	98.08	88.96	0.39	21.69	5
OC-SVM	362	97.54	85.96	0.52	27.54	2

TAB. 3.10: Alpes Maritimes - 2007 : Comparaison avec la carte de référence



- autres erreurs possibles :
  - le décalage temporel entre l’extinction du feu et l’acquisition de l’image satellitaire SPOT 5. Les caractéristiques spectrales des zones brûlées peuvent avoir changé pendant cette période, surtout s’il y a eu de grosses intempéries. La fréquence de revisite du satellite SPOT 5 est d’environ 3 jours mais certaines zones ne sont pas régulièrement couvertes (le satellite SPOT 5 possède une visée OBLIQUE et peut être programmé afin de couvrir une autre zone et/ou l’on doit choisir une image acquise plus tard à cause de la présence de nuages),
  - la différence de résolution entre la carte de référence et l’image satellitaire,
  - le recalage entre la carte de référence et l’image satellitaire.

Ces erreurs sont regroupées en deux catégories :

- l’*omission* : zone détectée comme non brûlée dans l’image satellitaire, mais identifiée comme brûlée dans la carte de référence (correspond aux faux négatifs).
- la *commission* : zone détectée comme brûlée dans l’image satellitaire, mais identifiée comme non brûlée dans la carte de référence (correspond aux faux positifs).

Lors de la réalisation de cartographies par imagerie satellitaire, l’erreur de commission est généralement plus critique que l’erreur d’omission [Chuviéco et al., 2002]. Comme les contours sont utilisés à des fins d’arrêtés d’interdiction de chasse et de pâturage, une surcartographie des zones brûlées est plus gênante. Il en est de même pour les compagnies d’assurances.

Néanmoins, pour les acteurs de la sécurité civile, qui connaissent la localisation des zones brûlées, il vaut mieux commettre des erreurs de cartographies que d’en oublier.

### 3.4 Conclusion

Dans ce chapitre, nous avons comparé les différentes techniques de classification qui ont été présentées dans le chapitre précédent. Les différentes techniques ont été testées d’abord sur un cas simple de mélanges de gaussiennes afin de comprendre leur comportement (en fonction de la base d’apprentissage, de la précision de la classification, du temps de calcul), puis sur des données réelles.

Les différentes méthodes classifient chaque pixel de l’image à partir de la seule information apportée par les valeurs spectrales du pixel considéré.

Dans les différentes expérimentations, à partir de la même base d’apprentissage, les SVM surpassent toutes les autres techniques de classification en terme de précision mais sont cependant plus lents en terme de temps de calcul. Les OC-SVM et MV paraissent assez équivalents en terme de précision globale et légèrement supérieurs aux K-PPV. Cependant, l’algorithme du MV semble effectuer de mauvaises classifications sur des objets non-brûlés comme les ombres ou l’aéroport (commission plus importante). Excepté pour la première image satellitaire, l’algorithme automatique des K-moyennes ne fournit pas de

bons résultats.

Les erreurs d'omission sont généralement localisées sur certains bords des zones brûlées qui présentent des caractéristiques spectrales assez proches de la végétation plus saine, mais atteinte par la sécheresse.

Les erreurs de commission sont dues à la similarité des caractéristiques spectrales des zones brûlées avec de la végétation non saine, certaines routes, des bords de mer...

Les statistiques des classifications ont été obtenues en les comparant avec des estimations satellitaires des zones brûlées dont les contours ne sont pas très précis : les cartographies d'Infoterra présentées sont des données vecteur obtenues généralement à partir d'images satellitaires de 30m de résolution, celle du SERTIT est aussi au format vecteur, celle du CNES a été obtenue de manière automatique à partir de l'image satellitaire SPOT 5 de résolution 2.5m.

Notre attention ne doit cependant pas se porter uniquement sur les zones brûlées de grande taille (supérieure à 50 hectares) même si leur cartographie exige énormément de temps, est plus contraignante et plus prioritaire pour les acteurs de la sécurité civile.

En effet, les zones brûlées de grande taille représentent un très faible pourcentage du nombre de zones brûlées (cf. chapitre 1). Dans les différentes cartographies, toutes les zones brûlées de taille supérieure à 10 hectares (seuil minimal des cartographies d'Infoterra) sont détectées.

Concernant les zones brûlées de taille inférieure, nous pouvons comparer les zones extraites avec celles de la base de donnée Prométhée, qui fournit la position et la taille de toutes les zones brûlées du Sud-Est de la France (certaines zones brûlées de taille inférieure au seuil de 10 hectares sont aussi présentes dans les cartographies d'Infoterra).

Les zones brûlées de petite taille (inférieure à 10 hectares) visuellement distinctes dans les images satellitaires sont détectées par les différents algorithmes. Cependant, certaines sont grossièrement détectées et leurs formes ne correspondent pas visuellement à la zone brûlée des images satellitaires.

Les classifications obtenues par les différentes méthodes à partir des composantes spectrales du pixel considéré, présentent toutes du "bruit", i.e. des pixels isolés mal classifiés comme brûlés ou non-brûlés. Classifier un pixel uniquement à partir de sa "couleur" ne fournit donc pas une très bonne cartographie. Dans le chapitre suivant, nous proposons d'améliorer les classifications obtenues en prenant en compte l'information spatiale apportée par les pixels voisins dans l'image.



## Chapitre 4

# Post-traitement - Prise en compte de l'information spatiale

Il existe de nombreuses techniques de classification qui permettent de prendre en compte à la fois l'information spectrale apportée par les différents canaux de l'image mais aussi l'information spatiale, i.e l'information apportée par les pixels voisins de chaque pixel.

Ces méthodes analysent généralement les différentes textures de l'image, qui sont caractérisées par une certaine homogénéité de groupements connexes de pixels dans l'image. La texture décrit la répétition spatiale de motifs selon une règle précise (écailles de reptile, mur de brique) ou de façon aléatoire (herbe, sable). On distingue donc deux classes de textures : les macrotextures et le microtextures [Cocquerez and Philipp, 1995].

La technique d'analyse de texture la plus populaire consiste à extraire les paramètres caractéristiques de la texture observée à partir des matrices de cooccurrence [Haralick, 1979]. D'autres approches proposent d'utiliser des filtres de Gabor [Jain and Farrokhnia, 1991] ou des ondelettes [Unser, 1995] afin de segmenter l'image. Les champs de Markov constituent aussi un moyen élégant afin de prendre en compte les interactions spatiales entre les pixels [Cocquerez and Philipp, 1995; Li, 1995; Winkler, 1995].

Plus récemment, [Kim et al., 2002] ou [Camps-Valls et al., 2006] proposent de concaténer l'information spectrale (valeurs spectrales du pixel) et une information spatiale (valeurs spectrales des voisins ou moyenne et variance sur le voisinage du pixel) pour former les données qui vont être utilisées par les SVM.

Cependant, ces méthodes fournissent une bonne classification à l'intérieur des structures de texture homogène mais ne donnent pas de bons résultats sur les frontières des différentes structures puisque l'information spatiale, définie par un voisinage fixe (qui ne dépend pas du pixel de l'image), est erronée sur les frontières. [Fauvel et al., 2007] proposent alors d'utiliser des filtres morphologiques [Soille, 1999] afin de définir des voisinages qui s'adaptent au mieux aux contours des structures.

Dans le chapitre précédent, l'utilisation seule de l'information spectrale permettait d'obtenir une bonne précision sur l'intérieur des zones brûlées, mais était insuffisante pour déterminer correctement certains contours des différentes zones. De plus, les pixels correspondant aux zones brûlées ont des caractéristiques spectrales assez homogènes et ne

nécessitent pas réellement une modélisation de la texture. [Smith et al., 2002] montrent que l'analyse de la texture de zones brûlées, pour une image infra-rouge de résolution 1km, ne fournit pas de meilleurs résultats qu'une simple technique de classification automatique utilisant uniquement l'intensité des pixels.

Dans ce chapitre, nous nous affranchissons donc de l'analyse de la texture. Nous nous attachons cependant à améliorer la connexité des zones détectées à partir des techniques de classification présentées dans le chapitre précédent, en effectuant une régularisation spatiale de la carte segmentée.

## 4.1 Introduction

Après une classification, il subsiste généralement des pixels isolés, mal classés à cause du bruit présent dans l'image satellitaire : les zones contiennent de petites discontinuités. Afin d'améliorer la cartographie, il faut la régulariser en réaffectant ces pixels isolés.

Dans notre problème de détection, les zones brûlées sont généralement caractérisées par des ensembles de pixels connexes et homogènes. Ainsi, des pixels isolés et classés comme brûlés ne peuvent être considérés comme réellement brûlés. De même, des pixels non brûlés isolés et à l'intérieur d'une zone brûlée ne peuvent être considérés comme non endommagés.

Dans ce chapitre, nous présentons trois techniques de post-classification afin de prendre en compte l'information spatiale apportée par les bandes spectrales de SPOT 5. Les deux premières sont des techniques très utilisées pour la régularisation (ou lissage) de classification. Elles sont fondées sur la morphologie mathématique [Serra, 1983; Soille, 1999] et la théorie des champs de Markov [Li, 1995; Winkler, 1995]. La dernière méthode décrite régularise aussi la classification, mais permet également d'améliorer les contours des différentes régions.

## 4.2 Régularisation par morphologie mathématique

Une technique de régularisation très simple consiste à utiliser des opérations de Morphologie Mathématique (MM).

### 4.2.1 Présentation de la méthode

Pour régulariser les classifications, deux opérations classiques en morphologie mathématique sont appliquées sur les classifications. Dans un premier temps, une fermeture est réalisée sur la classification afin d'éliminer les clusters brûlés de petites tailles (ce qui retire principalement des faux positifs). Puis, une ouverture est appliquée à la nouvelle classification afin de remplir les trous à l'intérieur des zones brûlées (ce qui retire principalement des faux négatifs). Les détails de ces deux opérations sont donnés dans l'annexe D.

Le disque est choisi comme élément structurant car il permet d'obtenir des contours plus réguliers que les autres éléments structurants. La taille de l'élément structurant dépend de la résolution de l'image satellitaire utilisée. Pour une image SPOT 5 (multispectral) de résolution 10 m, un cercle de rayon d'un pixel, ce qui correspond à une croix en 4-connexité (équivalent à 0.05 hectare) est utilisé. Pour une image SPOT 5 (supermode)

de résolution 2.5 m, un cercle de rayon égal à 4 pixels (équivalent à 0.03 hectare) est utilisé.

La figure 4.1 montre l'effet de la régularisation par morphologie mathématique pour la classification par SVM d'une partie de l'image SPOT 5 "Alpes-Maritimes", de résolution 10m.

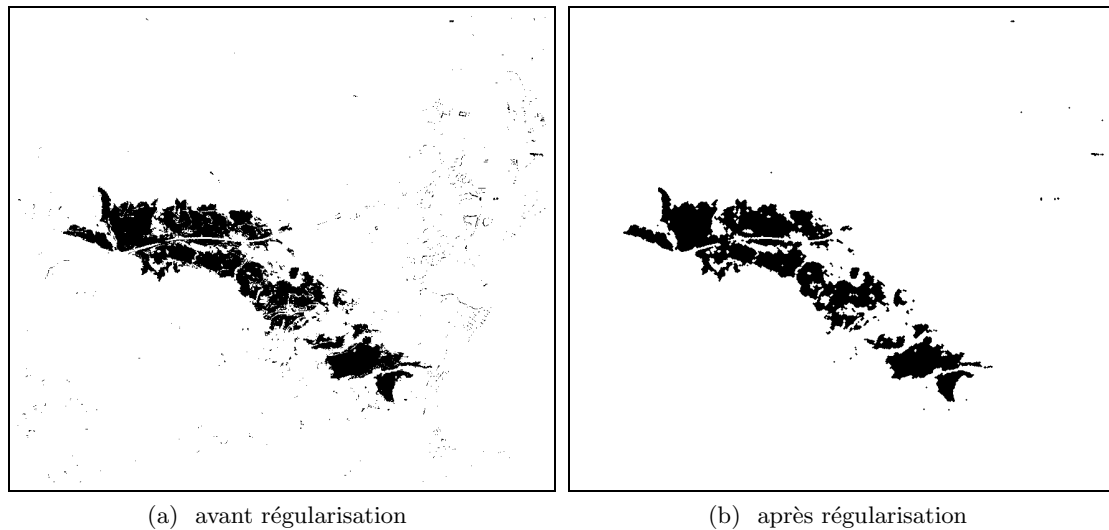


FIG. 4.1: Régularisation par morphologie mathématique

Concernant la cartographie de zones brûlées, l'erreur de commission (zones non brûlées détectées comme brûlées) est généralement plus critique que l'erreur d'omission (zones brûlées détectées comme non brûlées) [Chuvienco et al., 2002]. C'est pourquoi l'opération de fermeture est d'abord réalisée afin de diminuer l'erreur de commission (élimination des petits clusters de pixels brûlés) avant l'opération d'ouverture. Réaliser les deux opérations dans le sens inverse peut d'abord faire fusionner des petits clusters isolés et classés comme brûlés qui ne disparaîtront plus après la fermeture.

### 4.2.2 Expérimentation et comparaison

#### Sélection des hyperparamètres

La figure 4.2 donne la précision totale de la classification par SVM avant (à gauche) et après régularisation par morphologie mathématique (à droite) en fonction des hyperparamètres ( $C$  pour les marges relaxées et  $\sigma$  pour le noyau Gaussien).

D'après cette figure, la régularisation de la classification permet d'améliorer la classification par SVM (plus la couleur est proche du rouge, meilleure est la classification pour ce couple de paramètres). De plus, elle augmente la plage de couples d'hyperparamètres qui fournissent une bonne concordance entre la classification obtenue et la carte de référence, ce qui rend l'approche plus robuste.

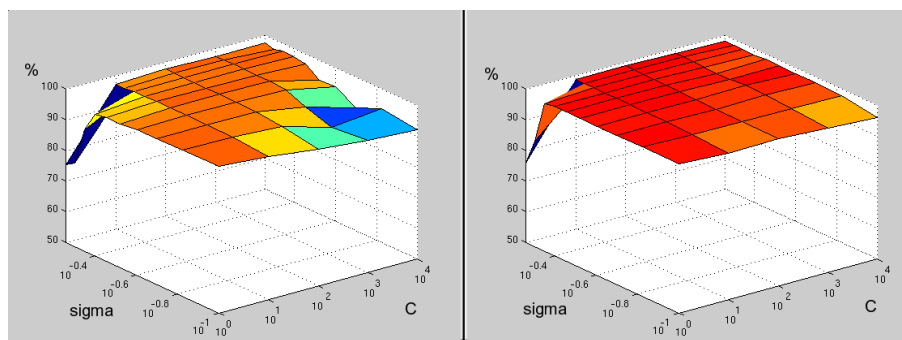


FIG. 4.2: Précision Totale avant et après régularisation par morphologie mathématique

### Régularisation des différentes techniques de classification

#### - Massif des Maures :

La figure 4.3 montre les résultats obtenus après régularisation par morphologie mathématique des cinq techniques de classification présentées dans le chapitre précédent. La résolution de l'image étant de 2.5m, l'élément structurant considéré est un cercle de rayon 4 pixels.

Le tableau 4.1 ci-après donne quelques statistiques sur les résultats obtenus après régularisation par morphologie mathématique des cinq techniques de classification. Il s'agit de la surface brûlée extraite, de la Précision Totale (PT), de la Précision Moyenne (PM), des taux de faux positifs (FP) et de faux négatifs (FN), et du temps de calcul.

#### - Est de la Corse :

La figure 4.4 montre les résultats obtenus après régularisation par morphologie mathématique des cinq techniques de classification présentées dans le chapitre précédent. La résolution de l'image étant de 10m, l'élément structurant considéré est un cercle de rayon 1 pixel. Un élément structurant de taille supérieure permettrait d'obtenir des images encore plus régularisées, mais effacerait des petites régions réellement brûlées et modifierait trop le contour des grandes régions.

Le tableau 4.3 donne quelques statistiques sur les résultats obtenus après régularisation par morphologie mathématique des cinq techniques de classification.

#### - Autres images :

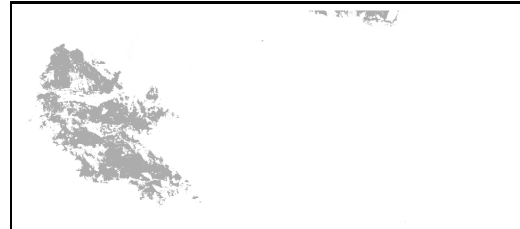
Les tableaux 4.2 et 4.4 donnent quelques statistiques sur les résultats obtenus après régularisation par Morphologie Mathématique (MM) des cinq techniques de classification pour les images "Ouest de la Corse" et "Alpes Maritimes".

### Conclusion

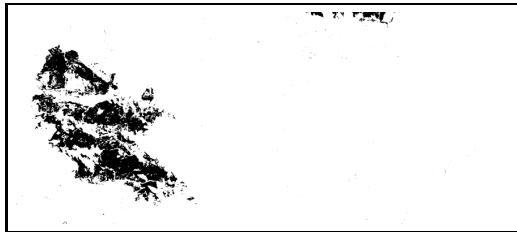
La régularisation par morphologie mathématique permet donc d'améliorer les cartographies obtenues par les différentes techniques de classification. Cependant, elle modifie



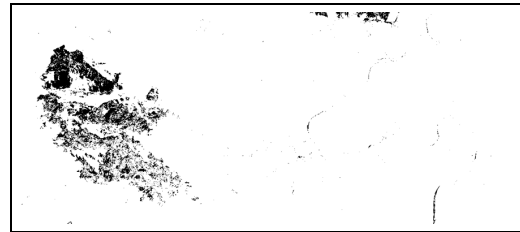
(a) Image SPOT 5 © CNES 2003, Distribution SPOT Image



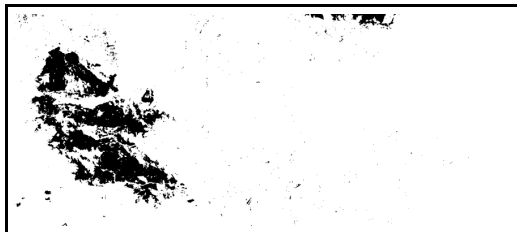
(b) Estimée Satellitaire © CNES



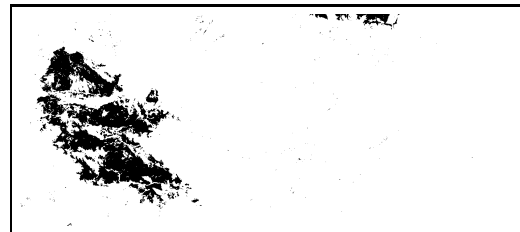
(c) K-moyennes + MM



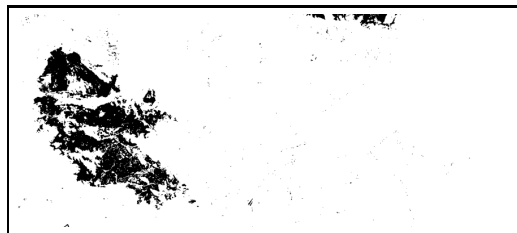
(d) K-Plus Proches Voisins + MM



(e) Maximum de Vraisemblance + MM



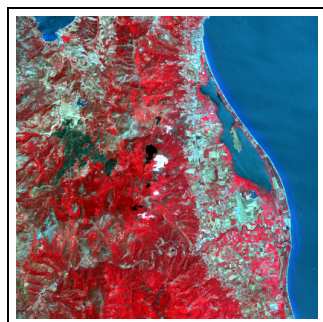
(f) SVM + MM



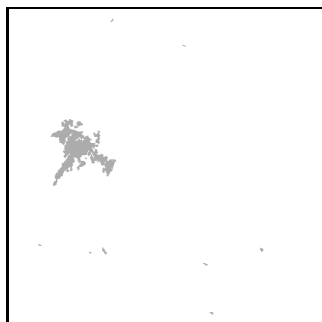
(g) One-Class SVM + MM

FIG. 4.3: Massif des Maures - 2003 (environ 2500 hectares)

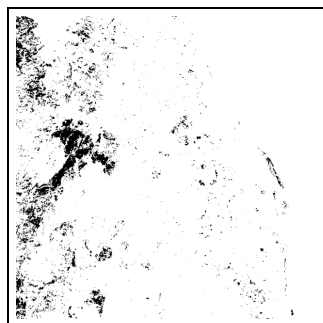




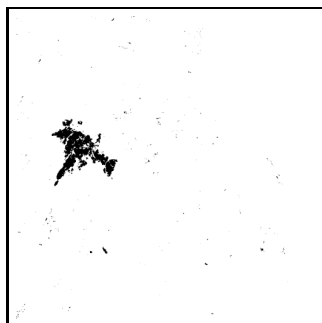
(a) Image SPOT 5 © CNES 2005, Distribution SPOT Image



(b) Estimée satellitaire - projet RISK-EOS, © Infoterra - ESA



(c) K-moyennes + MM



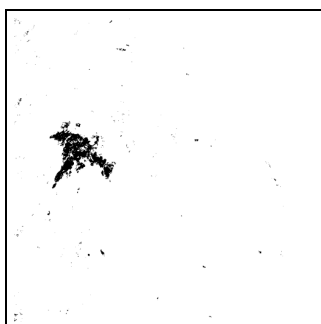
(d) K-PPV + MM



(e) MV + MM



(f) SVM + MM



(g) One-Class SVM + MM

FIG. 4.4: Est de la Corse - 2005 (environ 1000 hectares)

Carte de référence (2550 ha)	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
K-Moyennes + MM	1967	97.90	88.09	0.08	23.79	55
K-PPV + MM	1366	95.75	75.87	0.15	48.10	40
MV + MM	2417	98.62	94.42	0.51	10.64	15
SVM + MM	2298	98.71	93.65	0.24	12.45	65
OC-SVM + MM	2347	98.59	93.70	0.40	12.20	25

TAB. 4.1: Massif des Maures - 2003 : Comparaison avec la carte de référence

Carte de référence (2847 ha)	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
K-Moyennes + MM	2073	96.99	81.25	0.67	36.83	21
K-PPV + MM	2008	97.38	82.31	0.38	34.99	4
MV + MM	1721	97.03	78.55	0.21	42.67	2
SVM + MM	2281	97.77	86.15	0.51	27.18	6
OC-SVM + MM	2372	97.02	83.79	1.02	31.39	3

TAB. 4.2: Ouest de la Corse - 2005 : Comparaison avec la carte de référence

Carte de référence (992 ha)	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
K-Moyennes + MM	3290	95.55	85.83	4.13	24.22	31
K-PPV + MM	927	99.54	91.03	0.18	17.76	4
MV + MM	976	99.40	89.99	0.29	19.73	2
SVM + MM	929	99.60	91.97	0.15	15.91	6
OC-SVM + MM	906	99.45	89.05	0.21	21.69	3

TAB. 4.3: Est de la Corse - 2005 : Comparaison avec la carte de référence

Carte de référence (457 ha)	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
K-Moyennes + MM	408	98.25	90.95	0.53	17.57	21
K-PPV + MM	378	97.81	87.78	0.51	23.93	4
MV + MM	346	97.95	86.69	0.15	26.47	2
SVM + MM	361	98.36	89.00	0.07	21.91	6
OC-SVM + MM	367	98.18	88.63	0.22	22.52	3

TAB. 4.4: Alpes Maritimes - 2007 : Comparaison avec la carte de référence

le contour des zones détectées. De plus, il reste encore des fausses détections et des sur-détections.

Des opérateurs par reconstruction géodésique ou des filtres d'aires pourraient résoudre cet inconvénient. Néanmoins, ils ne rectifient pas la mauvaise classification des chemins endommagés (classés comme non brûlés) ou les bords de routes (classés comme brûlés) qui traversent la zone brûlée.

Les résultats obtenus par SVM fournissent encore les meilleurs résultats après régularisation. L'algorithme des OC-SVM donne des résultats équivalents à ceux obtenus par Maximum de Vraisemblance en termes de précision totale. Comme les SVM, les OC-SVM distinguent plus facilement les objets de caractéristiques assez similaires à celles des zones brûlées (ombres des nuages, aéroports, etc.). En revanche, ils n'utilisent que la base d'apprentissage positive.

C'est pourquoi, dans la suite de cette thèse, nous n'utiliserons désormais que les deux techniques SVM.

### 4.3 Régularisation par Champs de Markov

Les Champs de Markov (CM) permettent de tenir compte des interactions spatiales et sont largement utilisés en traitement d'image dans les domaines de la segmentation et de la restauration d'images [Geman and Geman, 1984].

Afin de ne pas surcharger cette partie, les définitions et propriétés des champs de Markov sont données dans l'annexe C.

#### 4.3.1 Présentation de la méthode

L'image segmentée obtenue après classification est considérée comme une réalisation  $y$  d'un champ  $Y$ , version bruitée de la réalisation  $x$  d'un champ markovien  $X$ . L'idée est de restaurer l'image  $y$ , i.e. d'estimer l'image  $x$  qui correspond le plus probablement à l'image  $y$  observée :

$$\arg \max_x P(X = x/Y = y)$$

D'après le théorème de Bayes, il vient :

$$P(X = x/Y = y) = \frac{P(Y = y, X = x)}{P(Y = y)} = \frac{P(Y = y/X = x)P(X = x)}{P(Y = y)} \quad (4.1)$$

Pour simplifier les notations, nous noterons dans la suite :  $P(y)$  pour  $P(Y = y)$  et  $P(x)$  pour  $P(X = x)$ .

Dans l'équation 4.1,  $P(y)$  ne dépend pas de  $x$  et ne joue donc pas de rôle dans la recherche du maximum a posteriori. Supposons que  $X$  soit un champ de Markov. D'après le théorème de Hammersley-Clifford, il existe une fonction appelée énergie  $U_r$  telle que :

$$P(x) = \frac{1}{Z} \exp(-U_r(x)) \text{ avec } U_r(x) = \sum_{c \in C} U_c(x)$$

où  $U_c$  est appelée fonction potentiel et ne dépend que des pixels  $x$  intervenant dans la clique  $c$ .  $C$  est l'ensemble des cliques. Supposons également que le bruit sur un pixel soit indépendant du bruit sur un autre pixel :

$$P(y/x) = \prod_{s \in S} P(y_s/x_s) = \sum_{s \in S} \exp(-U_f(y/x))$$

On obtient alors :

$$P(X = x/Y = y) = \frac{1}{Z.P(y)} \exp \left( - \sum_{s \in S} U_f(y/x) - \sum_{c \in C} U_c(x) \right)$$

Le premier terme de l'exponentielle représente l'attache aux données (fidélité à l'image segmentée), le second est un terme de régularisation qui prend en compte le voisinage de chaque pixel.

Pour le terme de régularisation, nous choisissons le modèle d'Ising qui favorise les configurations qui contiennent des voisins avec la même classe. Le voisinage est constitué par les 4 ou 8 plus proches voisins dans un espace bidimensionnel. Les potentiels sont définis uniquement pour les cliques d'ordre 2 :

$$U_c(x) = \beta \sum_{\{s,t\} \in C_2} (1 - 2\delta_{x_s=x_t})$$

Pour le terme d'attache aux données, il convient de choisir :

$$U_f(y/x) = g(x_s - y_s)$$

où  $g$  est une fonction paire, positive et nulle en 0.

Un choix très répandu consiste à supposer que la quantité  $x_s - y_s$ , différence entre la classe après régularisation et celle obtenue par la segmentation bruitée, suit une loi gaussienne centrée. Cela revient donc à choisir :  $U_f(y/x) = \frac{(x_s - y_s)^2}{\sigma^2}$

Afin de rechercher le maximum, nous utilisons l'algorithme *Iterative Conditional Mode* (ICM) [Besag, 1986]. Cet algorithme modifie, à chaque itération, les valeurs de tous les sites de l'image. Il construit, à partir d'une configuration initiale  $X^0$ , une suite d'image  $X^k$  convergeant vers une approximation locale du Maximum a Posteriori :

---

Algorithme ICM :

- Choix d'une image initiale  $X^0$
- Construction de  $X^{k+1}$  à partir de  $X^k$  :
  - pour tout site  $s$  de l'image  $X^k$ , on calcule :

$$P(X_s = z/Y_{t \in V_s} = X_{t \in V_s}^k)$$

- pour tout  $z = \{0, 1\}$ , on choisit :

$$X_s^{k+1} = \arg \max_z P(X_s = z/Y_{t \in V_s} = X_{t \in V_s}^k)$$


---

L'algorithme s'arrête lorsque le nombre de changements d'une itération à l'autre est nul ou suffisamment faible [Cocquerez and Philipp, 1995]. Cet algorithme est très rapide, une trentaine d'itérations suffit généralement. Cependant, il converge vers un maximum local et ses performances dépendent donc de son initialisation. En choisissant l'image segmentée comme configuration initiale, l'algorithme converge vers une version régularisée de cette image.

La figure 4.5 montre l'effet de la régularisation par champs de Markov pour la classification par SVM d'une partie de l'image "Alpes-Maritimes", de résolution 10m, avec un voisinage 8-connextité.

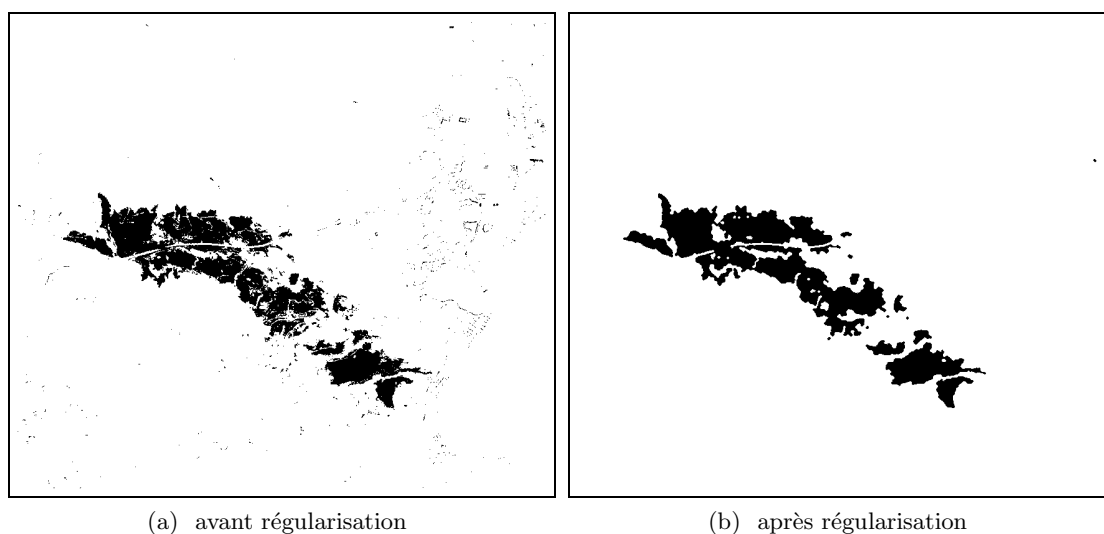


FIG. 4.5: Régularisation par champs de Markov

### 4.3.2 Expérimentation et comparaison

La régularisation par champs de Markov est une méthode largement utilisée afin d'améliorer la cohérence spatiale après segmentation d'une image.

Les figures 4.6, 4.7 et 4.8 donnent les régularisations des classifications par OC-SVM pour trois images satellitaires, par Morphologie Mathématique (MM - à gauche) et par Champs de Markov (CM - à droite).

Le tableau 4.5 compare les régularisations des classifications par OC-SVM pour les trois images avec les estimations satellitaires fournies par Infoterra France ou le SERTIT. Les statistiques concernant les régularisations par morphologie mathématique pour ces 3 images ont été précédemment donnés (cf. tableaux 4.2 4.3 4.4).

La régularisation par champs de Markov permet d'obtenir des cartographies avec moins de fausses alarmes que celles par morphologie mathématique. En effet, pour les champs de Markov, contrairement à la morphologie mathématique, la 8-connextité est utilisée car elle permet d'obtenir une classification plus régulière, au prix d'une complexité calculatoire

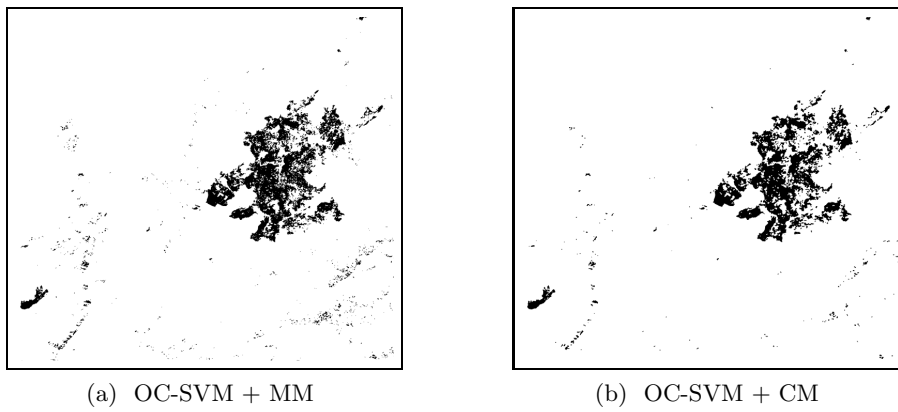


FIG. 4.6: Ouest de la Corse - 2005 ( 2800 ha)

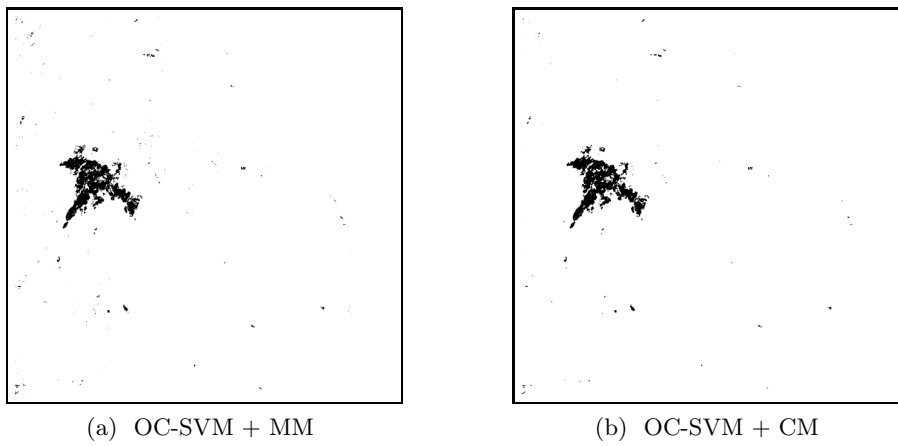


FIG. 4.7: Est de la Corse - 2005 ( 1000 ha)

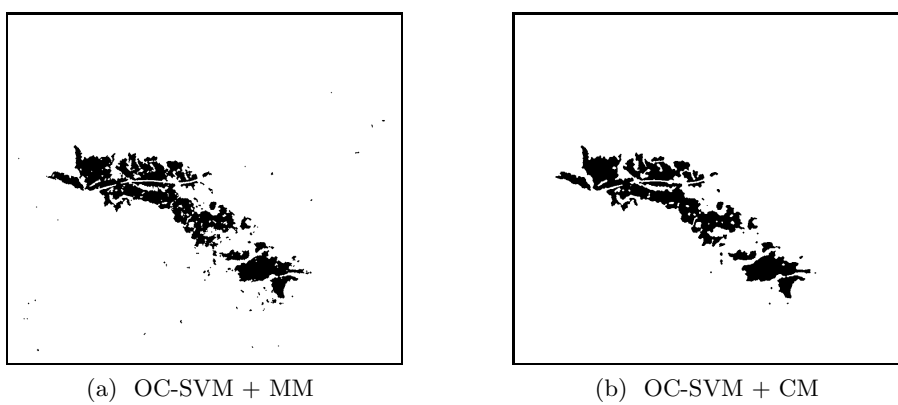


FIG. 4.8: Alpes-Maritimes - 2007 (environ 450 ha)

Image	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
Ouest de la Corse	2124	97.19	82.49	0.62	34.40	4
Est de la Corse	868	99.51	89.04	0.15	21.77	4
Alpes-Maritimes	349	98.18	87.68	0.07	24.57	4

TAB. 4.5: Régularisation par champs de Markov des classifications par OC-SVM

légèrement plus élevée que la 4-connexité.

Il reste néanmoins toujours des fausses alarmes. De plus, les contours des zones brûlées ne sont pas très précis.

## 4.4 Technique de croissance de régions

Les techniques de détection de zones brûlées sont généralement fondées sur les techniques de détection de changements entre images avant-feu et après-feu, qui font appel à des seuillages stricts [Kasischke and French, 1995; Pereira, 1999; Chuvieco et al., 2002]. Le choix du seuil est critique puisqu'il définit complètement la loi de décision et est déterminé empiriquement.

Les caractéristiques spectrales des zones brûlées peuvent cependant être très similaires à celles de zones non brûlées (elles varient selon le type de végétation, la sévérité du feu et le décalage temporel entre l'extinction du feu et l'acquisition de l'image) [Pereira, 1999]. Un seuillage strict impliquera donc des erreurs d'omission et de commission.

Plus récemment, certains auteurs [Stroppiana et al., 2000; Garcia and Chuvieco, 2004; Quintano et al., 2006; Chuvieco et al., 2008] ont démontré les avantages des techniques de croissance de régions afin de diminuer les erreurs de classification de zones brûlées et non-brûlées. En effet, les classes brûlées et non brûlées ne sont pas spectralement disjointes : il y a un recouvrement spectral entre les deux classes que nous cherchons à séparer (cf annexe A). Sur le contour d'une zone brûlée, la végétation peut être moins gravement touchée qu'à l'intérieur de la zone et peut présenter ainsi des caractéristiques spectrales très proches de celle de régions épargnées par le feu.

L'idée des techniques de croissance de régions est :

- dans un premier temps, de déterminer les zones brûlées ayant un haut taux de confiance. Ces zones sont celles qui sont les plus gravement touchées (appelées "coeurs" des zones brûlées),
- puis de grossir ces zones, considérées comme des graines, afin d'affiner les bordures des zones brûlées en ajoutant les pixels voisins s'ils présentent un critère de similarité suffisant.

### 4.4.1 Présentation de la méthode

Les méthodes de classification SVM et OC-SVM utilisées fournissent des lois de décisions dures. Les pixels sont soit classés positifs, soit classés négatifs, en fonction de la position du pixel par rapport à l'Hyperplan Séparateur Optimal (HSO). Pour chaque pixel,

la classification est, en fait, déterminée par le signe de  $f(\vec{x}) = \vec{w} \cdot \phi(\vec{x}) + b$  pour les SVM, celui de  $f(\vec{x}) = \vec{w} \cdot \phi(\vec{x}) - \rho$  pour les OC-SVM, ce qui est équivalent à utiliser un unique seuil égal à zéro.

Nous proposons ici d'utiliser l'information apportée par la distance  $f(\vec{x})$  entre le pixel et l'HSO en la considérant comme entrée d'un seuillage par hystérésis [Canny, 1986]. Le seuillage par hystérésis fait appel à deux seuils,  $S_{haut}$  et  $S_{bas}$ . L'image donnant la distance des pixels à l'HSO est d'abord segmentée par le seuil haut ( $S_{haut}$ ) afin d'obtenir des pixels positifs avec un haut taux de confiance (mais aussi quelques faux positifs). Pour notre application, ces pixels sont immédiatement classés comme brûlés. Un deuxième seuillage, cette fois avec le seuil bas ( $S_{bas}$ ), augmente la taille des régions obtenues précédemment par le seuil haut, mais fait aussi apparaître de nouveaux clusters. La segmentation finale est obtenue en sélectionnant, parmi tous les clusters, ceux qui contiennent des régions de confiance élevée.

Le seuillage par hystérésis est un algorithme itératif :

- La première étape consiste à détecter les graines, i.e, les pixels  $\vec{x}_s$  de l'image tels que  $f(\vec{x}_s) \geq S_{haut}$ . Cette phase permet d'identifier les pixels les plus sévèrement touchés par les feux afin d'éviter une grande erreur de commission (pixels sains classifiés comme brûlés), mais donne une grande erreur d'omission.
- Puis, à chaque itération, les 4 pixels voisins  $\{\vec{x}_t\}_{t \in V_s}$  des graines sont testés : si  $f(\vec{x}_t) \geq S_{bas}$ , alors ce pixel est ajouté à la graine. Cette deuxième phase permet d'affiner les bords des zones brûlées en combinant l'information spatiale et spectrale des pixels. Ainsi l'erreur d'omission diminue.

Le processus s'arrête si les graines n'ont pas grossi entre deux itérations.

Des pixels isolés, d'amplitude supérieure à  $S_{haut}$  peuvent être des faux positifs et engendrent donc des clusters faux positifs après le deuxième seuillage. Aussi, une opération d'érosion (cf annexe D) est appliquée après le premier seuil afin d'éliminer ces pixels isolés, considérés comme du bruit. Cette étape est logique car la taille des zones brûlées est généralement supérieure à un ou deux pixels pour des images satellitaires haute résolution. [Garcia and Chuvieco, 2004] et [Chuvieco et al., 2008] proposent d'éliminer les clusters dont la taille est respectivement inférieure à 3 pixels (pour des images satellitaires SAC-C/MMRS de résolution 175m), ou 5 pixels (pour des images satellitaires AVHRR de résolution 1000m), ce qui correspond à notre érosion avec un élément structurant circulaire de rayon 1 pixel.

De plus, il reste des pixels isolés d'amplitude inférieure à  $S_{bas}$  à l'intérieur des zones brûlées après le second seuillage. Une opération de fermeture (cf annexe D) est donc appliquée afin d'éviter des petits "trous" à l'intérieur des zones brûlées. Cette étape est aussi logique puisqu'un ou deux pixels à l'intérieur d'une zone brûlée ne peuvent pas être complètement sains pour des images satellitaires haute résolution.

L'algorithme utilisé est donc le suivant :



---

Algorithme Seuillage par hystérésis :

---

- Segmentation de l'image des distances avec le seuil haut  $S_{haut}$  :  
 $X_s^0 = 1$  si  $f(\vec{x}_s) \geq S_{haut}$
  - Erosion de  $X^0$  afin d'éliminer les pixels isolés.
  - Construction de  $X^{k+1}$  à partir de  $X^k$  : pour tout site  $X_s^k$  de l'image  $X^k$ ,
    - si  $X_s^k = 1$ , alors  $X_s^{k+1} = 1$
    - si  $S_{bas} \leq f(\vec{x}_s)$  et s'il existe  $X_t^k \in V_s$  tel que  $X_t^k = 1$ , alors  $X_s^{k+1} = 1$
  - S'il n'y a pas de changements entre  $X^{k+1}$  et  $X^k$ , on réalise une fermeture sur  $X^{k+1}$  et l'algorithme se termine.
- 

Comme le premier seuil définit le nombre de zones brûlées qui seront détectées et le deuxième définit la taille des zones brûlées, le choix de ces deux seuils est critique. Il n'y a malheureusement pas de règles concernant ce choix. Les différents seuils sont généralement obtenus expérimentalement à partir d'analyse de séries temporelles d'images satellitaires sur plusieurs années [Chuvieco et al., 2008].

Néanmoins, les seuils doivent respecter plusieurs critères exposés ci-après, ce qui restreint finalement le nombre de choix possibles.

#### Cas des SVM :

Pour les SVM, il est évident que  $S_{haut}$  doit être positif,  $S_{bas}$  doit être négatif. De plus, pour chaque vecteur  $\vec{x}$  de la base d'apprentissage, nous savons que :

- si  $y_i \cdot f(\vec{x}_i) \geq 1$ , alors le vecteur est bien classé :  $sign(f(\vec{x}_i)) = y_i$
- si  $y_i \cdot f(\vec{x}_i) = 1$ , alors le vecteur est bien classé et il se situe à une bonne distance de l'HSO
- si  $y_i \cdot f(\vec{x}_i) \leq 1$ , alors, soit le vecteur est bien classé, mais est proche de l'HSO ; soit il est mal classé et se situe du mauvais côté de l'HSO.

La plage de valeurs correctes pour chaque seuil se situe donc entre 0 et  $\pm 1$ .

De plus, si la base d'apprentissage ne comporte pas d'erreur, on peut choisir :

$$S_{haut} = \min_i \{ y_i \cdot f(\vec{x}_i) \text{ tel que } y_i \cdot f(\vec{x}_i) > 0 \}$$

$$S_{bas} = \max_i \{ y_i \cdot f(\vec{x}_i) \text{ tel que } y_i \cdot f(\vec{x}_i) < 0 \}$$

#### Cas des OC-SVM :

Pour les OC-SVM, pour chaque vecteur  $\vec{x}$  de la base d'apprentissage, nous savons seulement que :

- si  $f(\vec{x}_i) \geq 0$ , alors le vecteur est bien classé,
- si  $f(\vec{x}_i) < 0$ , alors il est mal classé, i.e. du mauvais côté de l'HSO.

Les bornes inférieure  $-1$  et supérieure  $+1$  ne sont plus valables pour les seuils. Cependant, les seuils doivent être compris dans une certaine plage de valeurs définie par les éléments de la base d'apprentissage.

La figure 4.9 montre la position de la base d'apprentissage et des lignes de niveau du classifieur OC-SVM, i.e les lignes de niveau de  $\vec{w} \cdot \phi(\vec{x}) - \rho$ , dans le cas 2-dimensions, pour

un couple d'hyperparamètres donné  $(\sigma, \nu)$ . La ligne de niveau en rouge représente l'HSO (ensemble de points tels que  $\vec{w} \cdot \phi(\vec{x}) - \rho = 0$ ), les croix et les ronds en bleu correspondent respectivement aux Vecteurs non Supports (à l'intérieur de la ligne de niveau rouge) et Vecteurs Supports (sur ou à l'extérieur de la ligne de niveau rouge).

Le seuillage par hystérésis est appliqué au classifieur OC-SVM. Des valeurs acceptables des seuils  $S_{haut}$  et  $S_{bas}$  correspondent respectivement aux lignes de niveau qui incluent un peu moins et un peu plus d'éléments de la base d'apprentissage que la ligne de niveau zéro (les premières courbes à l'intérieur et à l'extérieur de la ligne rouge de la figure 4.9). Le seuil défini par la deuxième ligne de niveau à l'intérieur de la ligne de niveau zéro est trop élevé car il n'inclut pas assez d'éléments de la base d'apprentissage et est donc trop restrictif. De même, le seuil défini par la deuxième ligne de niveau à l'extérieur de la ligne de niveau zéro est trop bas car la ligne de niveau inclut tous les éléments de la base d'apprentissage et est trop éloignée des Vecteurs Supports.

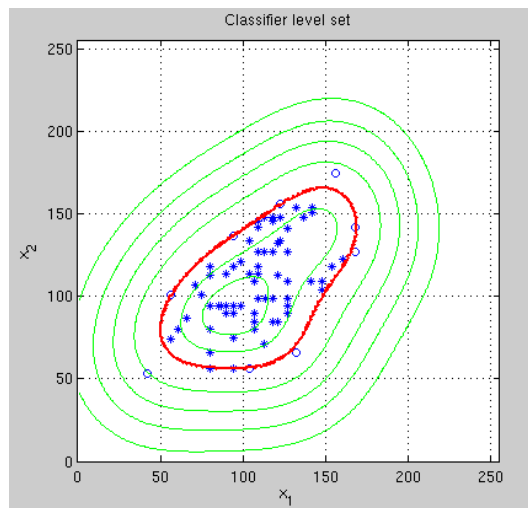


FIG. 4.9: Lignes de niveau pour les OC-SVM

L'élément structurant utilisé lors des opérations d'érosion et de fermeture doit être de petite taille afin de ne pas effacer des vrais positifs ni d'ajouter de faux négatifs. En effet, à cause des conditions météorologiques, du relief, du type de végétation ou de l'activité des pompiers, le feu peut laisser des zones de végétation non brûlée à l'intérieur de la région brûlée. Dans notre étude, nous choisissons un voisinage 4 ou 8-connexité pour une image satellitaire de résolution 10m. Pour une telle image, une érosion avec un élément structurant 8-connexe élimine les clusters dont la taille est inférieure à 9 pixels, i.e. 0.1 hectare environ.

#### 4.4.2 Expérimentation et comparaison

On peut remarquer que le choix des seuils dépend de la classification obtenue par techniques SVM et donc du choix des hyperparamètres.

Le choix des hyperparamètres est important puisqu'ils déterminent directement la position de l'HSO des SVM ou OC-SVM. La sélection automatique des paramètres par la

technique de validation croisée donne de bons résultats mais n'est pas optimale.

Le seuillage par hystérésis n'utilise pas la position des pixels par rapport à l'HSO mais la distance entre les pixels et l'HSO, ce qui permet d'être un peu plus libre sur la sélection des hyperparamètres (il en était de même pour les classifications régularisées par morphologie mathématique ou par champ de Markov - cf. Figure 4.2). Un couple d'hyperparamètres non optimal peut aussi donner de très bons résultats avec un bon couple de seuils.

La figure 4.10 montre la zone de Cannes-Mandelieu (Alpes-Maritimes), qui a brûlé en 2007. Elle donne l'image satellitaire SPOT 5 (a), la carte de référence (c), la détection de zone brûlée par classification OC-SVM classique, puis régularisée par Champs de Markov (e) et quelques étapes de l'approche par hystérésis (le classifieur OC-SVM, la première segmentation avec  $S_{haut}$  et la classification finale). L'estimation des zones brûlées a été réalisée par le SERTIT <sup>1</sup> à partir d'images satellitaires SPOT 5.

La figure 4.10b donne les valeurs du classifieur OC-SVM dont les données en entrée sont définies par les 4 bandes spectrales des pixels. La figure 4.10d représente la segmentation du classifieur SVM après seuillage par  $S_{haut}$ . Puis, les pixels isolés sont éliminés (car considérés comme du bruit) par le processus d'érosion et les pixels restants forment les graines qui vont engendrer la classification par hystérésis après le seuillage par  $S_{bas}$ . La figure 4.10f donne la classification finale obtenue après fermeture de la classification par hystérésis.

Les figures 4.11 et 4.12 donnent les images satellitaires "Ouest" et "Est de la Corse", les estimations satellitaires fournies par Infoterra France, les classifications par OC-SVM régularisées par Champs de Markov (CM) et les cartographies réalisées par l'approche de croissance de régions proposée et fondée sur les OC-SVM. Elles donnent aussi des zooms sur la superposition des contours obtenus par ces deux méthodes avec l'image satellitaire correspondante.

La figure 4.13 donne le relevé de terrain fourni par l'ONF des Alpes-Maritimes et les contours des zones brûlées (en vert) obtenus par OC-SVM régularisé par Champs de Markov (à gauche), par technique de croissance de régions (à droite) superposés avec l'image satellitaire correspondante.

Si l'on compare visuellement les classifications obtenues avec les cartes de référence, le seuillage par hystérésis du classifieur OC-SVM surpasse le simple OC-SVM régularisé. En effet, en combinant les OC-SVM avec le seuillage par hystérésis, le nombre de faux positifs est plus faible (cf. Fig. 4.11 et 4.13) et le contour des zones brûlées est plus précis (cf. Fig. 4.11, 4.12 et 4.13). De plus, le seuillage par hystérésis est très simple et rapide (son temps de calcul est de l'ordre de quelques secondes pour une image de 2000\*2000 pixels).

Le tableau 4.6 compare les classifications obtenues par la technique de croissance de régions pour les trois images satellitaires avec les estimations satellitaires fournies par

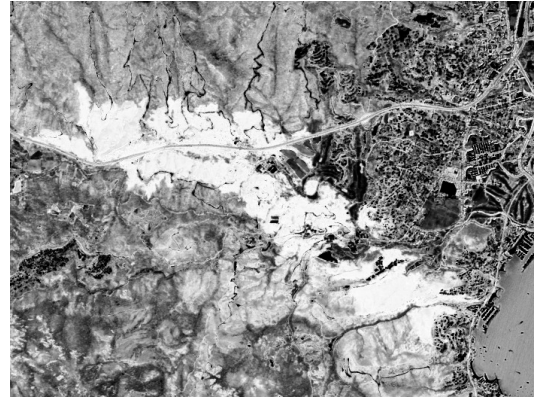
---

<sup>1</sup>disponible à l'adresse internet :

[http://sertit.u-strasbg.fr/SITE\\_Charte/2007/3\\_france\\_feux\\_juillet\\_2007/sud\\_france\\_feux\\_2007.html](http://sertit.u-strasbg.fr/SITE_Charte/2007/3_france_feux_juillet_2007/sud_france_feux_2007.html)



(a) Image SPOT 5 © CNES 2007, Distribution SPOT Image



(b) Classifieur OC-SVM



(c) Estimée Satellitaire ©SERTIT



(d) seuillage par  $S_{haut}$  du classifieur OC-SVM

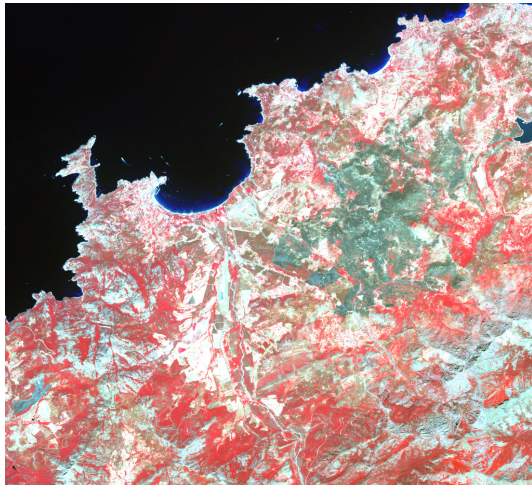


(e) OC-SVM + reg par Champs de Markov



(f) Zones brûlées extraites

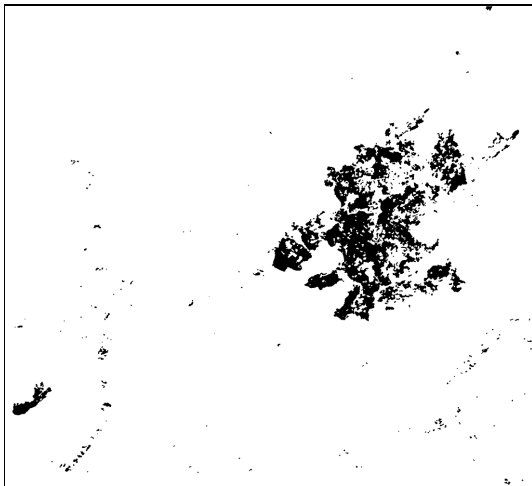
FIG. 4.10: Cannes-Mandelieu - 2007 (environ 450 ha)



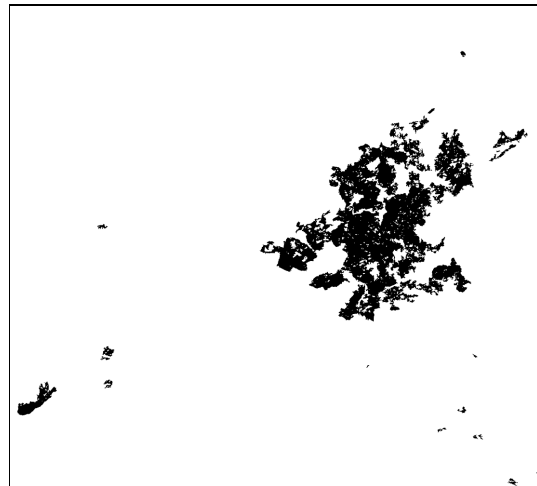
(a) Image SPOT 5 © CNES 2005, Distribution SPOT Image



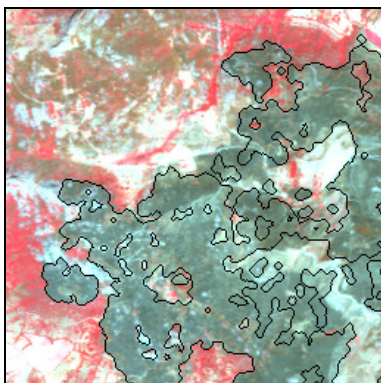
(b) Estimée satellite - projet RISK-EOS © Infoterra-ESA



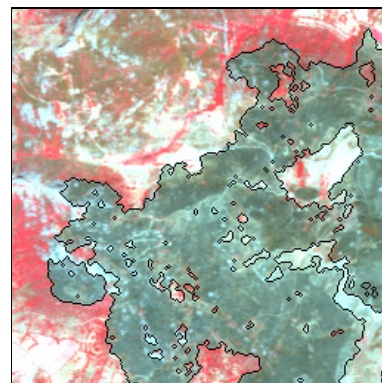
(c) OC-SVM + CM



(d) OC-SVM + croissance de régions



(e) zoom : Superposition de l'image satellitaire et du contour obtenu par OC-SVM + CM (en noir)

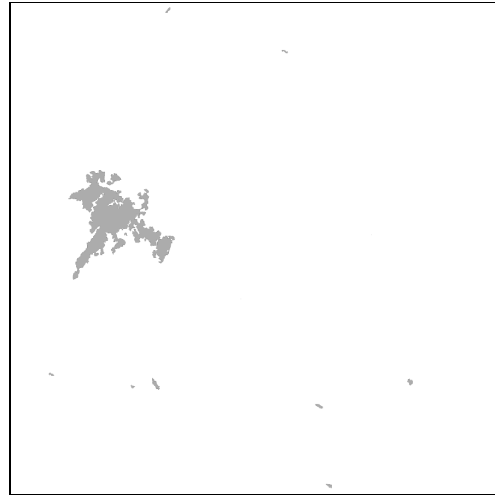


(f) zoom : Superposition de l'image satellitaire et du contour obtenu par OC-SVM + croissance de régions (en noir)

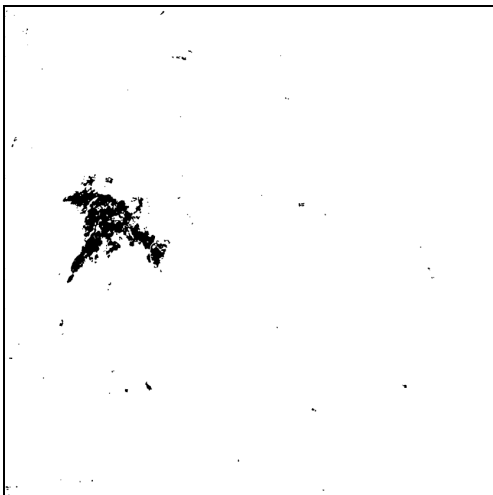
FIG. 4.11: Ouest de la Corse - 2005 (environ 2800 ha)



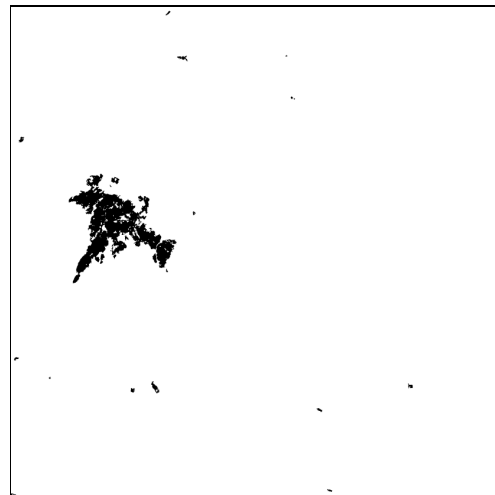
(a) Image SPOT 5 © CNES 2005, Distribution SPOT Image



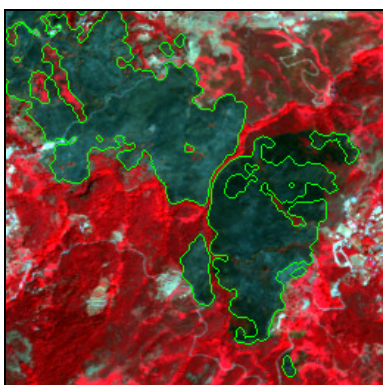
(b) Estimée satellite - projet RISK-EOS © Infoterra-ESA



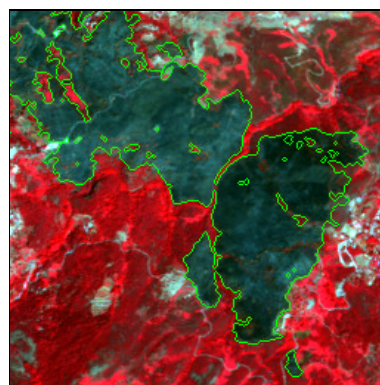
(c) OC-SVM + CM



(d) OC-SVM + croissance de régions



(e) zoom : Superposition de l'image satellitaire et du contour obtenu par OC-SVM + CM (en vert)

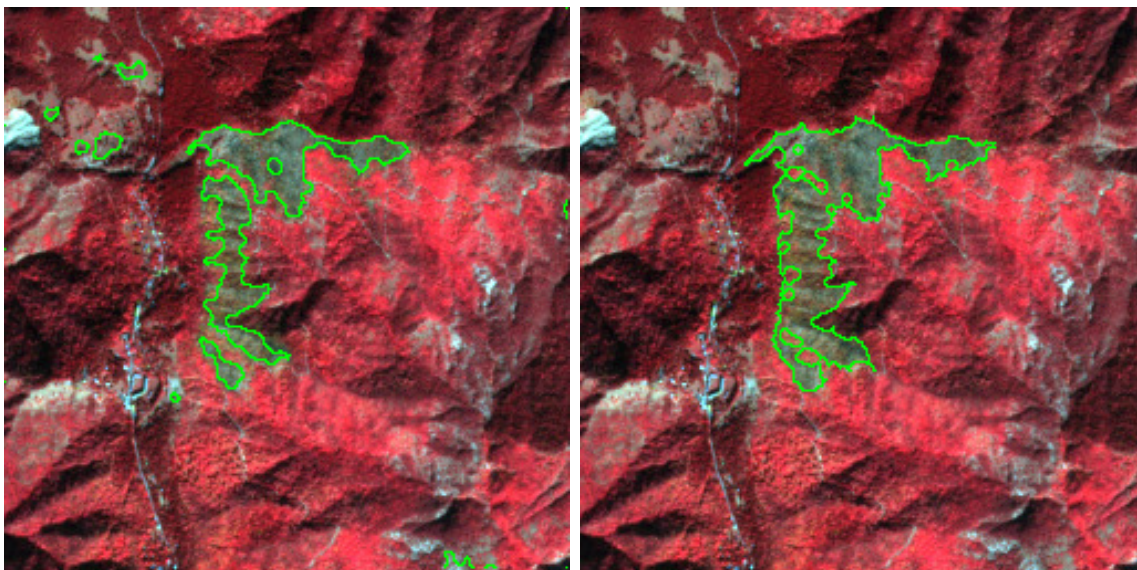


(f) zoom : Superposition de l'image satellitaire et du contour obtenu par OC-SVM + croissance de régions (en vert)

FIG. 4.12: Est de la Corse - 2005 (environ 1000 ha)



(a) Relevé de terrain © ONF des Alpes Maritimes



(b) Superposition de l'Image SPOT 5 © CNES 2003, Distribution SPOT Image et du contour obtenu par OC-SVM + CM  
 (c) Superposition de l'Image SPOT 5 © CNES 2003, Distribution SPOT Image, et du contour obtenu par OCSVM + croissance de régions

FIG. 4.13: Lucéram - 2003 (50 hectares)

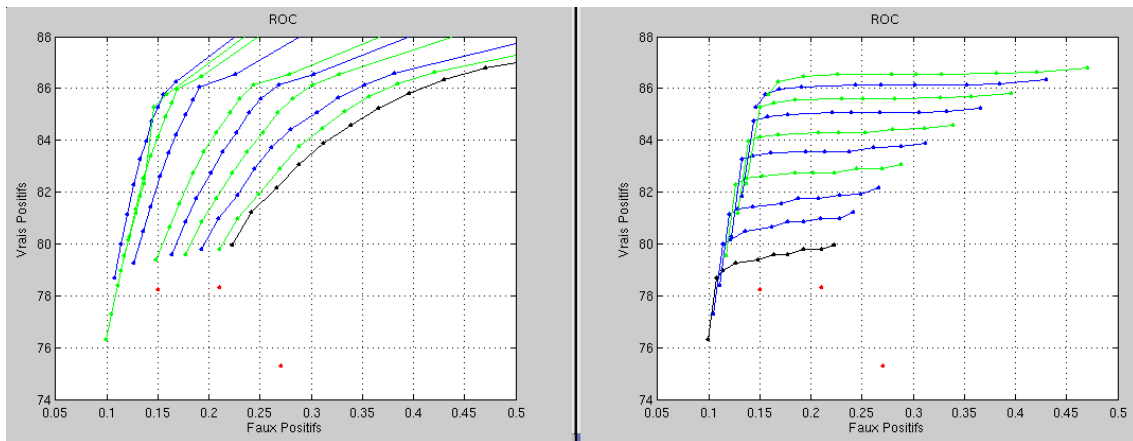


FIG. 4.14: Est de la Corse - courbes ROC

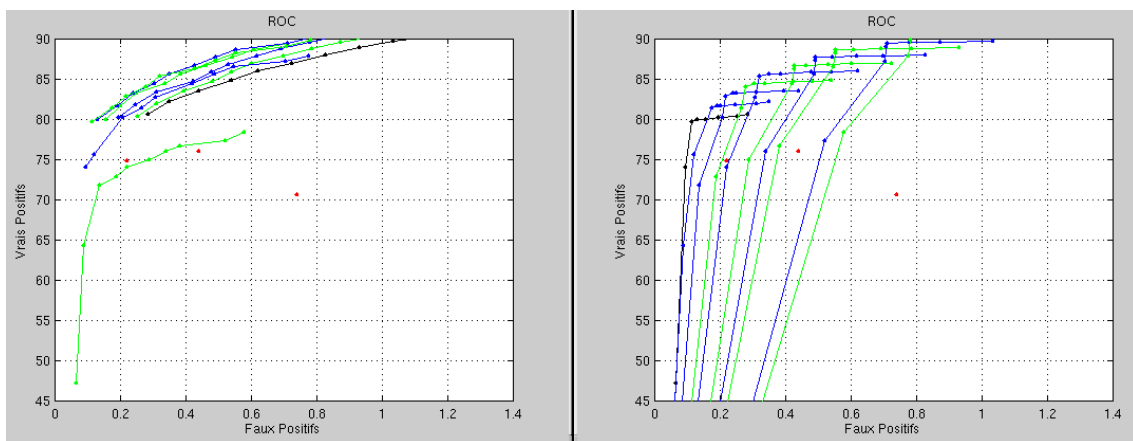


FIG. 4.15: Alpes-Maritimes - courbes ROC



Infoterra France ou par le SERTIT.

Image	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
Ouest de la Corse	2620	97.91	89.49	0.84	20.17	4
Est de la Corse	971	99.59	92.93	0.19	13.95	4
Alpes-Maritimes	378	98.49	90.32	0.14	19.21	4

TAB. 4.6: Croissance de régions des OC-SVM

Les figures 4.14 et 4.15 représentent les courbes "Receiver Operating Characteristic" (ROC), i.e. la sensibilité de la méthode par seuillage par hystérésis en fonction des deux seuils. Chaque point donne le taux de vrais positifs en pourcentage (égal à 100 moins le taux de faux négatifs) en fonction de celui des faux positifs pour un couple de seuils. Les figures représentent respectivement les courbes ROC pour différentes valeurs de  $S_{haut}$  (figure de gauche),  $S_{bas}$  (figure de droite).

#### - Est de la Corse :

La fonction de décision de OC-SVM varie entre  $-0.484$  et  $0.061$ .

Concernant la figure de gauche 4.14, les courbes correspondent aux seuils hauts suivants :  $\{0.005, 0.010, 0.015, 0.020, 0.025, 0.030, 0.035, 0.040, 0.45, 0.50\}$ . Elles sont obtenues en faisant varier le seuil bas qui prend les valeurs suivantes :  $\{-0.005, -0.010, -0.015, -0.020, -0.025, -0.030, -0.035, -0.040, -0.045, -0.050\}$ . Lorsque le seuil bas diminue (augmente en valeur absolue), à seuil haut constant, la taille des zones détectées augmente, les nombre de vrais et faux positifs augmentent, les courbes sont donc croissantes avec le seuil bas. Pour indication, la courbe noire correspond au premier seuil haut  $0.005$ .

Concernant la figure de droite 4.14, les courbes correspondent aux seuils bas et sont obtenues en faisant varier le seuil haut pour les valeurs énoncées ci dessus. Pour indication, la courbe noire correspond au premier seuil bas  $-0.005$ . Lorsque le seuil haut augmente, à seuil bas constant, le nombre de zones détectées diminue, les nombre de vrais et faux positifs diminuent, les courbes sont donc décroissantes avec le seuil haut.

Les trois points rouges correspondent respectivement aux classifications par OC-SVM sans régularisation (droite), avec régularisation par morphologie mathématique (milieu) et par champs de Markov (gauche).

Si l'on considère ces trois points rouges, à taux de vrais positifs ou faux négatifs constants, il est possible de trouver de nombreux couples de seuils qui donnent respectivement un meilleur taux de faux négatifs ou de vrais positifs.

#### - Alpes Maritimes :

La figure 4.15 fournit les courbes ROC, qui permettent de comparer, en fonction des seuils utilisés, les classifications obtenues avec la cartographie (données vecteur) réalisée à partir de l'image SPOT 5 par le SERTIT.

La fonction de décision de OC-SVM varie entre  $-0.501$  et  $0.046$ . Les figures représentent les mêmes informations que précédemment, les mêmes seuils sont utilisés.

Les courbes passent toutes par l'origine : un seuil haut supérieur à la valeur maximale

de la fonction de décision des SVM n'engendre pas de graines, il n'y a donc aucune détection de pixels positifs et donc des taux de vrais et faux positifs nuls.

Le point rouge de gauche correspond à la régularisation des OC-SVM par champs de Markov. Pour les taux de vrais positifs ou faux négatifs définis par ce point, il est possible de trouver un couple de seuils qui donne respectivement un meilleur taux de faux négatifs ou de vrais positifs.

Les statistiques fournies par la comparaison des différentes méthodes avec les cartes de référence sont légèrement faussées :

- Les estimations des zones brûlées réalisées par Infoterra France - ESA, considérées comme cartes de référence, dérivent de l'interprétation d'images satellitaires Landsat dont la résolution (30m) est supérieure à celle des images que nous utilisons. L'estimation fournie par le SERTIT, dérive de l'image SPOT 5 que nous utilisons. Ces données sont des données vecteur, ce qui implique une incertitude sur les contours des zones brûlées,
- De plus, la cartographie du SERTIT donne uniquement le contour des zones brûlées. Les routes et les parcelles de végétation qui sont à l'intérieur des zones brûlées, mais qui semblent visuellement non brûlées, sont comptabilisées comme brûlées.

Les zones brûlées des images "Ouest de la Corse" et "Alpes-Maritimes" comportent de nombreux chemins et zones de végétation qui semblent non touchées, ce qui explique pourquoi les taux de faux négatifs sont moins bons que pour les autres images.

Si l'on compare les taux de faux positifs et faux négatifs obtenus par la technique de croissance de régions (tableau 4.6) avec ceux obtenus par régularisation par champs de Markov (tableau 4.5), on peut remarquer que la méthode proposée donne plus de faux positifs, mais moins de faux négatifs. La technique par croissance de régions augmente la superficie des zones brûlées, il est donc normal de détecter plus de pixels brûlés et donc moins de pixels non brûlés. Cependant, elle fournit de meilleurs résultats visuellement.

Lorsque l'on utilise l'algorithme classique de SVM, on peut s'attendre à de meilleurs résultats qu'avec les OC-SVM ; ce qui est évident puisque cet algorithme utilise une base d'apprentissage négative en plus de la base d'apprentissage positive. Si l'on combine l'algorithme des SVM avec le seuillage par hystérésis, le résultat obtenu est visuellement à peine supérieur à celui obtenu par combinaison des OC-SVM et du seuillage par hystérésis, mais avec un temps de calcul largement supérieur (entre 2 à 4 fois plus long) afin de réaliser les phases d'apprentissage et de classification, sans compter le temps nécessaire à la sélection de la base d'apprentissage négative.

## 4.5 Conclusion

Trois méthodes ont été testées afin d'améliorer les classifications utilisant uniquement l'information spectrale apportée par l'image satellitaire. Ces méthodes permettent de combiner une information spatiale avec les classifications obtenues afin d'améliorer la forme des zones détectées et de diminuer le nombre de fausses alarmes.

La première méthode, la plus simple, consiste à utiliser des opérations de morphologie mathématique. La deuxième méthode, plus classique, est fondée sur la théorie des champs de Markov et fournit une classification plus régulière que la première. Ces deux méthodes permettent d'améliorer la cohérence des zones détectées et utilisent la cartographie binaire obtenue par classification par techniques SVM. Cependant, ces techniques de régularisation modifient les bordures des différentes zones brûlées et laissent quelques fausses alarmes.

La méthode que nous proposons considère la distance entre chaque pixel de l'image et la surface de séparation obtenue par les techniques SVM, qui sépare l'espace spectral en deux parties. Les pixels sont considérés comme brûlés en fonction de leur distance par rapport à l'hyperplan (utilisation de l'information spectrale) et en fonction des pixels voisins s'ils sont voisins de l'hyperplan (utilisation de l'information spatiale). Dans ce cas, le voisinage considéré dépend de la distance à la surface de séparation. Une première étape localise les ensembles de pixels ayant une grande probabilité d'être brûlés. Une deuxième étape permet d'augmenter la taille et d'améliorer la forme des clusters détectés dans la première phase (croissance de régions).

La comparaison des classifications obtenues à partir d'images SPOT 5 et des cartographies réalisées par Infoterra France à partir d'images satellitaires Landsat, ainsi que les statistiques qui en découlent, ne sont pas toujours pertinents à cause de la différence de résolution entre les deux images satellitaires. Il en est de même avec ceux obtenus à partir de la cartographie réalisée par le SERTIT, qui ne donne que le contour extérieur des zones brûlées.

Cependant, l'analyse visuelle des cartes de classification permet de constater que la méthode proposée fournit moins de fausses alarmes (erreur de commission) et des contours plus précis que les deux premières méthodes, quelle que soit la taille des zones brûlées.

Le point clé de la technique de croissance de régions utilisée concerne le choix des deux seuils. Il n'existe pas de méthode automatique permettant de sélectionner les seuils optimaux. De nombreux couples de seuils fournissent de bons résultats. Comme l'algorithme est très rapide, les seuils optimaux peuvent être obtenus empiriquement.

Le couple de seuils peut aussi être choisi en fonction des spécifications exigées par l'utilisateur : un seuil haut élevé (ou faible) implique une faible commission (ou omission) par exemple.

Récemment, certains auteurs imposent des contraintes supplémentaires afin de modérer l'effet des seuils. [Pilar Martin et al., 2006] considèrent aussi comme graines des pixels de la deuxième phase et distants de moins de 1000 mètres (2 pixels pour une image MODIS) de pixels de la première phase (les graines). [Chuvieco et al., 2008] fixent une distance maximale de 20 km (20 pixels pour une image satellitaire AVHRR) entre les pixels de la deuxième phase et ceux de la première phase.

La méthode proposée reste cependant supervisée, et nécessite donc des exemples afin de pouvoir classifier tous les pixels de l'image. Dans le chapitre suivant, nous nous intéressons au problème de la sélection automatique de la base d'apprentissage.

## Chapitre 5

# Sélection de la base d'apprentissage

Les SVM et OC-SVM, comme tous les algorithmes de classification supervisée, nécessitent un ensemble d'observations (appelé base d'apprentissage) afin de pouvoir prédire la classe de nouveaux objets. La sélection manuelle de la base est coûteuse et cruciale puisque des éléments de chaque classe brûlée ou non-brûlée (forêts, champs, mer, lacs, zones urbaines, routes) doivent être spécifiés afin d'avoir de bons résultats [Richards and Jia, 1999; Mather, 2004]. De plus, les caractéristiques spectrales dépendent des capteurs du satellite, de l'illumination, de l'orientation du terrain et diffèrent malheureusement d'une image à l'autre.

### 5.1 Peut-on garder la même base d'apprentissage avec une autre image ?

Dans cette partie, la base d'apprentissage et les hyperparamètres optimaux définis pour une image satellitaire sont testés sur une autre image afin de voir s'il est possible de garder la même fonction de décision lorsque l'on change d'image.

Les figures 5.1 et 5.2 montrent deux images satellitaires ("Nord de la Corse" et "Est de la Corse"), représentant la même région, acquises approximativement à la même heure mais à des jours différents, et les cartes de référence correspondantes. Elles fournissent aussi les classifications obtenues par SVM avec les bases d'apprentissage des deux images.

En raison des différences d'illumination entre 2 images satellitaire acquises à des dates différentes, il n'est pas possible de garder exactement la même fonction de décision lorsque l'on change d'image.

La théorie de l'apprentissage semi-supervisé propose d'adapter une fonction de décision en utilisant des éléments dont on ne connaît pas la classe (base de test) [Zhu, 2005]. L'idée est de tirer profit des éléments non classés, qui sont plus abondants que les éléments classés (base d'apprentissage) et sont disponibles. Par exemple, les "Transductive SVM" [Vapnik, 1998] ont pour but de maximiser la fonction objectif des SVM (cf. Eq. 2.14) par

rapport aux étiquettes ("labels") des éléments de la base de test. Concernant le choix de ces éléments, [Bruzzone et al., 2007] proposent, par exemple, d'utiliser les voisins des pixels de la base d'apprentissage et d'ajouter une contrainte d'homogénéité des labels entre voisins.

L'apprentissage semi-supervisé fournit des résultats très intéressants. Cependant, il nécessite néanmoins une base d'apprentissage propre à l'image considérée. En effet, les éléments de la base d'apprentissage servent à définir un premier classifieur, qui est ensuite amélioré grâce aux éléments de la base de test.

L'objectif de ce chapitre est de définir entièrement et automatiquement la base d'apprentissage à partir d'une image donnée.

Dans la suite, les méthodes proposées combinent un algorithme non-supervisé avec un algorithme supervisé afin d'utiliser leurs avantages respectifs. La combinaison, bien que plus lente, fournit bien évidemment des résultats moins précis que ceux des algorithmes supervisés puisque ces derniers bénéficient de l'information apportée par une base d'apprentissage choisie manuellement. Cependant, elle fournit de meilleurs résultats que ceux des algorithmes automatiques et sont utiles dans le cas d'absence de base d'apprentissage.

## 5.2 Sélection par K-moyennes

### 5.2.1 Présentation de la méthode

En premier lieu, l'algorithme des K-moyennes est utilisé pour fournir une première estimation des classes de chaque pixel de l'image.

Cet algorithme divise donc l'image en K classes, quelques-unes correspondant majoritairement aux zones brûlées, toutes les autres correspondant majoritairement aux zones non brûlées. L'idée est d'appliquer une érosion [Soille, 1999] pour chacune des K classes afin d'éliminer les petites régions connexes qui peuvent être mal classifiées, ce qui élimine aussi les contours des plus grosses régions connexes.

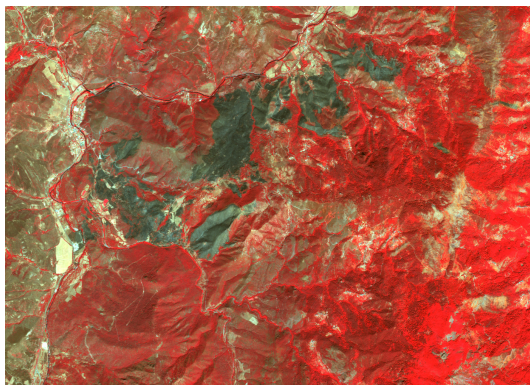
Après l'érosion, les pixels restants sont regroupés en deux catégories : zones brûlées et non brûlées. Ces pixels constituent alors la base d'apprentissage des SVM. Cette opération peut être effectuée manuellement. Un SVM multi-classes peut aussi être utilisé à partir des K classes érodées afin d'éviter le regroupement manuel des classes à cet instant et de le repousser en fin de classification. Cela permet aussi d'avoir une classification finale multi-classes.

### 5.2.2 Expérimentation et comparaison

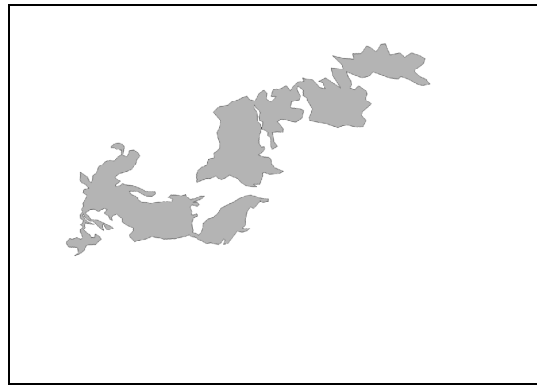
#### - Massif des Maures :

La figure 5.3 donne l'image SPOT 5, la carte de référence associée, les classifications obtenues par K-moyennes, par SVM et par la méthode précédemment exposée, régularisée par morphologie mathématique. Le tableau 5.1 fournit les statistiques correspondants aux trois cartographies.

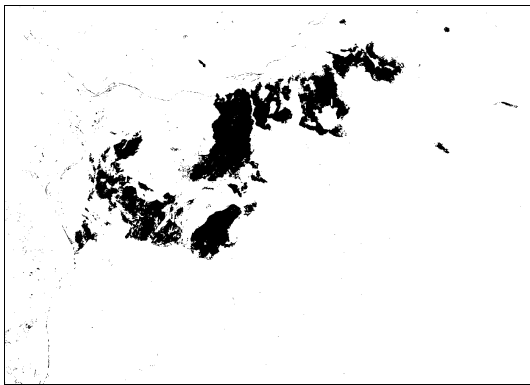
La combinaison des K-moyennes et des SVM fournit de meilleurs résultats que ceux obtenus par simple K-moyennes, mais moins bons que ceux obtenus par SVM classique avec une base d'apprentissage sélectionnée manuellement, ce qui est logique.



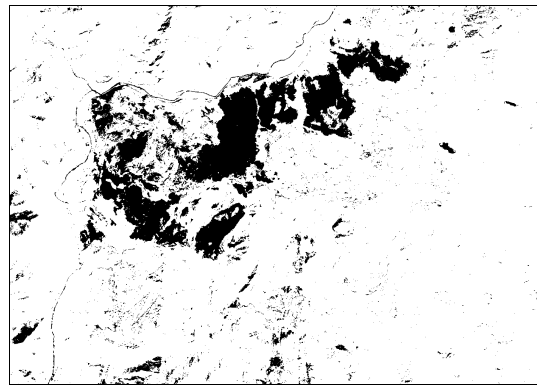
(a) Image SPOT 5 © CNES 2003, Distribution SPOT Image



(b) Relevé de terrain © SDIS2B



(c) SVM - base d'apprentissage "Nord de la Corse"



(d) SVM - base d'apprentissage "Est de la Corse"

FIG. 5.1: Nord de la Corse (1000 ha)

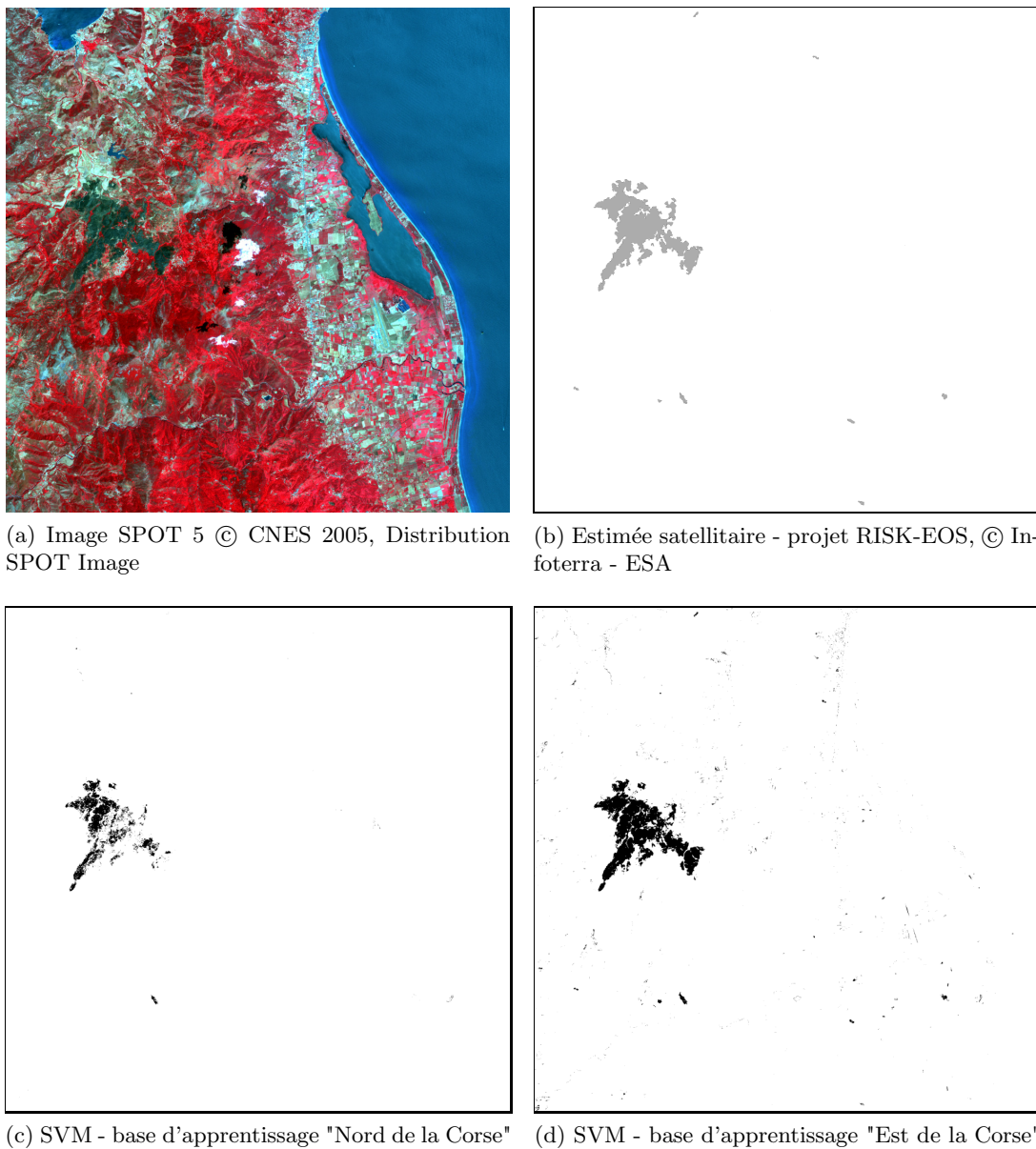


FIG. 5.2: Est de la Corse (1000 ha)

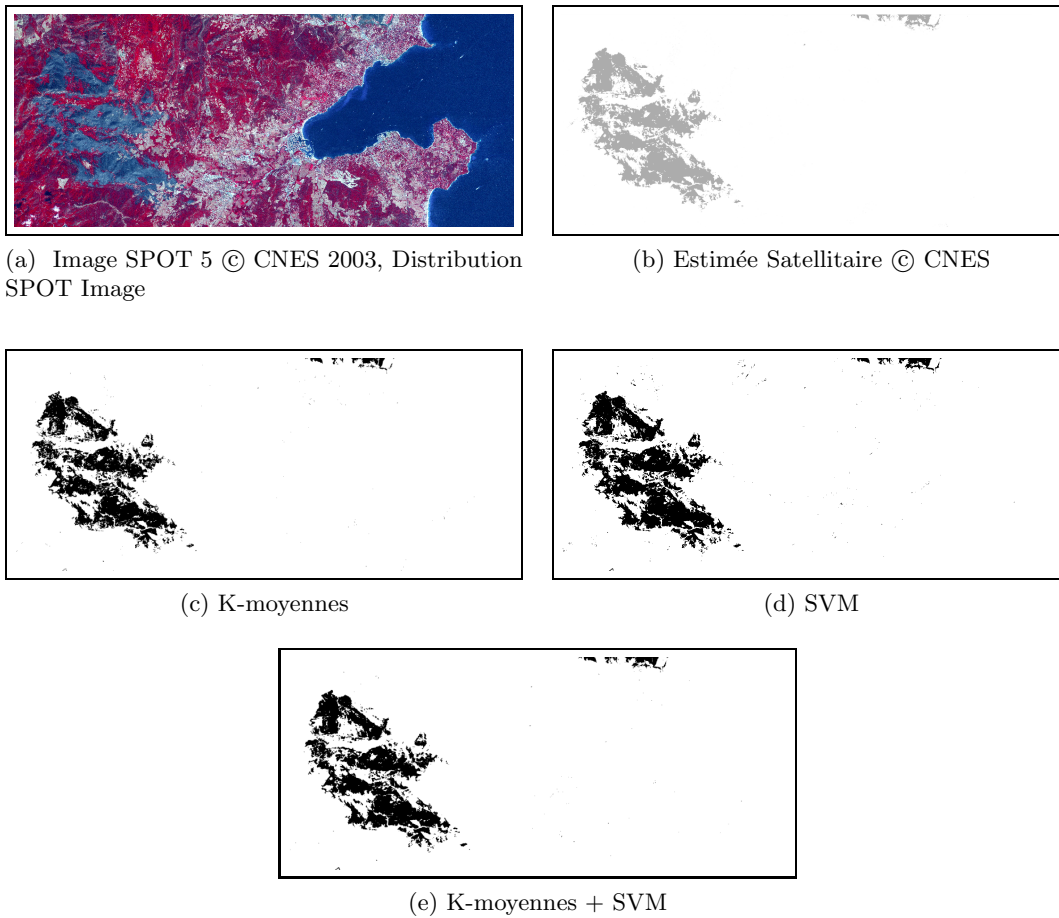
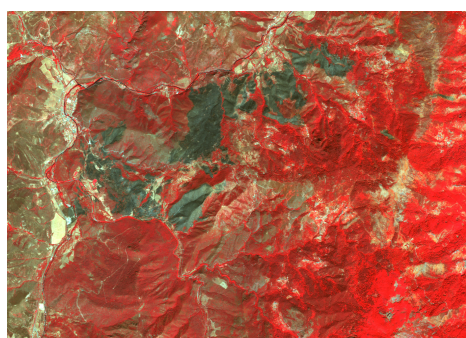


FIG. 5.3: Massif des Maures - 2003 (2500 hectares)

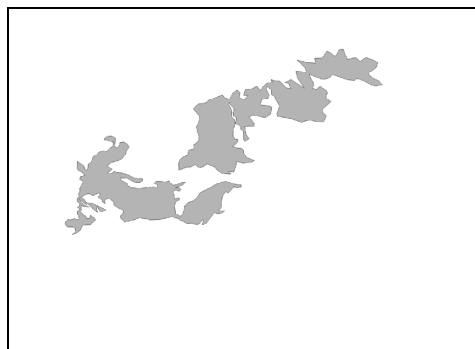
Carte de référence (2550 ha)	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
K-moyennes	1967	97.90	88.09	0.08	23.79	45
SVM	2298	98.71	93.65	0.24	12.45	60
K-moyennes + SVM	2101	98.37	90.78	0.07	18.36	90

TAB. 5.1: Massif des Maures - 2003 : Comparaison avec la carte de référence

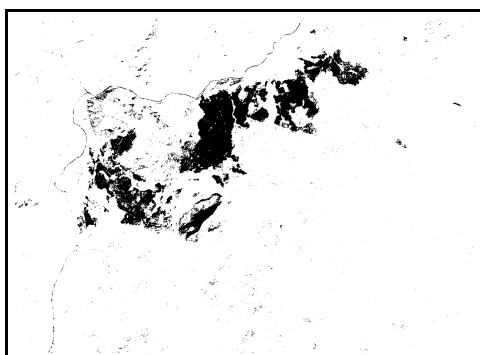




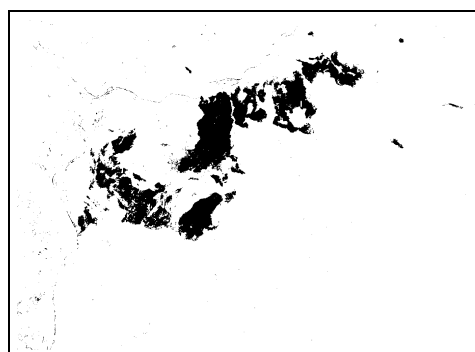
(a) Image SPOT 5 © CNES 2003, Distribution SPOT Image



(b) Relevé de terrain © SDIS2B



(c) K-moyennes



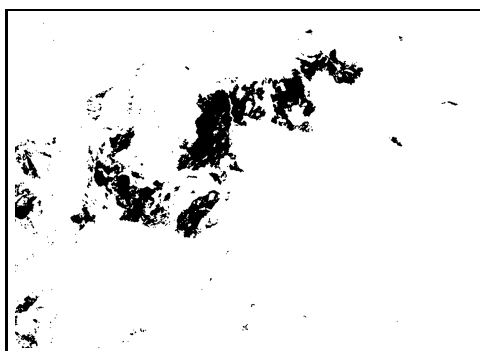
(d) simple SVM



(e) K-moyennes + SVM



(f) K-moyennes + SVM + croissance de régions



(g) K-moyennes + OCSVM



(h) K-moyennes + OCSVM + croissance de régions

FIG. 5.4: Nord de la Corse (1000 hectares)

#### - Nord de la Corse :

La figure 5.4 fournit l'image SPOT 5, le relevé de terrain associé, les classifications obtenues par K-moyennes, par SVM, par la combinaison des K-moyennes et SVM seule, puis avec le seuillage par hystérésis afin de prendre en compte l'information spatiale. Le relevé de terrain ne fournit cependant pas la position et la forme des cinq petites zones brûlées (quelques hectares) présentes dans l'image satellitaire.

La classification par K-moyennes+SVM (e) fournit de meilleurs résultats que par simple K-moyennes (c), mais moins bons que par simple SVM (d). On peut aussi remarquer qu'elle ne réussit pas à détecter toutes les zones brûlées de petites tailles détectées par la classification par SVM (d). Le seuillage par hystérésis est appliqué à la classification par K-moyennes+SVM et permet d'obtenir une classification comportant moins de fausses alarmes et qui respecte mieux les bords des zones détectées.

La base d'apprentissage positive sélectionnée par cette technique peut être utilisée par les OC-SVM. Les résultats sont donnés par les figures (g) et (h).

La méthode combinant K-moyennes et SVM ne fournit malheureusement de bons résultats que dans le cas de très grandes zones brûlées. En effet, l'algorithme des K-moyennes ne fournit une bonne estimation des zones brûlées que si elles ne sont pas minoritaires dans l'image.

Choisir un nombre de classes très élevé peut permettre de résoudre ce problème, mais cela augmente alors énormément la complexité de la méthode et exige une intervention manuelle non négligeable afin d'identifier les zones brûlées et non-brûlées.

### 5.3 Sélection à partir de l'histogramme par ligne de partage des eaux

L'annexe A montre qu'en se basant sur les histogrammes 3-Dimensions ou 4-Dimensions, il est possible de segmenter en grande partie les zones brûlées. L'idée est de sélectionner le petit pic de l'histogramme correspondant à la zone brûlée afin de définir la base d'apprentissage.

La segmentation de l'histogramme est effectuée par Ligne de Partage des Eaux (LPE) [Vincent and Soille, 1991].

Ce processus de segmentation d'histogramme est assez similaire à ceux de [Géraud et al., 2001] et [Fang et al., 2005]. [Géraud et al., 2001] segmentent l'histogramme d'une image couleur par LPE, puis régularisent la segmentation obtenue par champs de Markov (modèle de Potts). [Fang et al., 2005] proposent de segmenter un histogramme en 2 Dimensions (en fait, le projeté de l'histogramme couleur), puis d'utiliser la segmentation obtenue pour former la base d'apprentissage des SVM.

Comme le montrent les résultats de l'annexe E, les bandes Vert et Rouge d'une image satellitaire se ressemblent énormément et sont très corrélées. Nous n'utilisons donc que trois des quatre bandes spectrales (Rouge, PIR et MIR) afin de simplifier le modèle et diminuer le temps de calcul sans perte d'information. Afin de garder les vraies valeurs discrètes des données, nous n'utilisons pas les composantes principales fournies par l'ACP

De plus, n'utiliser que trois des quatre bandes de SPOT 5 fournit un histogramme plus compact, avec moins de valeurs isolées.

### 5.3.1 Présentation de la méthode

Dans un premier temps, un filtre Gaussien [Cocquerez and Philipp, 1995], avec une variance élevée, est appliqué à l'histogramme 3D afin de le lisser et d'éliminer les extrema locaux non significatifs.

Puis, un algorithme de LPE [Vincent and Soille, 1991] est appliqué sur l'opposé de l'histogramme afin de segmenter ses différents pics (une présentation de l'algorithme de LPE est faite dans l'annexe D). En utilisant la LPE sur l'opposé de l'histogramme, le pic recherché correspond à un minimum local, la frontière du "cluster" segmenté correspond à la base du pic.

En donnant les valeurs des 3 bandes spectrales d'un pixel brûlé de l'image, le "cluster" 3D correspondant (pic de l'histogramme) constitue la base d'apprentissage. Bien que le choix du pixel de départ soit crucial pour obtenir le bon pic de l'histogramme, il est très facile de sélectionner visuellement un pixel brûlé.

### 5.3.2 Expérimentation et comparaison

#### - Ouest de la Corse :

La figure 5.5 donne une section des histogrammes à une valeur de MIR donnée pour les étapes suivantes : l'histogramme original (en haut à gauche), celui du "cluster" sélectionné après LPE et filtrage Gaussien (en haut à droite), celui de l'OC-SVM (en bas à gauche), et celui après le seuillage par hystérésis (en bas à droite).

L'algorithme de LPE entraîne une sur-segmentation de l'espace couleur à cause de minima locaux non éliminés par le filtrage Gaussien. L'algorithme d'OC-SVM permet de fusionner quelques bassins et fournit une prédiction quantifiable de la positivité des pixels. Le processus d'hystérésis utilise cette prédiction mesurable afin de changer le contour du "cluster". On peut remarquer que le "cluster" sélectionné possède de faibles valeurs de PIR et de Rouge.

La figure 5.6 représente l'image "Ouest de la Corse", zone qui a brûlé en 2005. Elle donne l'image satellitaire (les bandes spectrales Rouge, PIR et MIR sont respectivement représentées en nuance de bleu, vert et rouge), la carte de référence correspondante (réalisée par Infoterra France - ESA à partir d'images Landsat de résolution 30m) et les classifications obtenues à partir des mêmes étapes que dans la figure 5.5.

Le tableau 5.2 compare les cartographies obtenues à partir de la segmentation de l'histogramme, des OC-SVM utilisant la base d'apprentissage définie par la segmentation de l'histogramme, de la Croissance de Régions (CR) fondée sur les OC-SVM utilisant la base d'apprentissage définie par la segmentation de l'histogramme. Les deux premières cartographies sont régularisées par champs de Markov afin d'améliorer leur cohérence spatiale et pouvoir être comparées avec la classification finale.

L'algorithme de OC-SVM avec une base d'apprentissage sélectionnée à partir de l'histogramme fournit une meilleure classification que par simple LPE sur l'histogramme. De

5.3. SÉLECTION À PARTIR DE L'HISTOGRAMME PAR LIGNE DE PARTAGE DES EAUX<sup>87</sup>

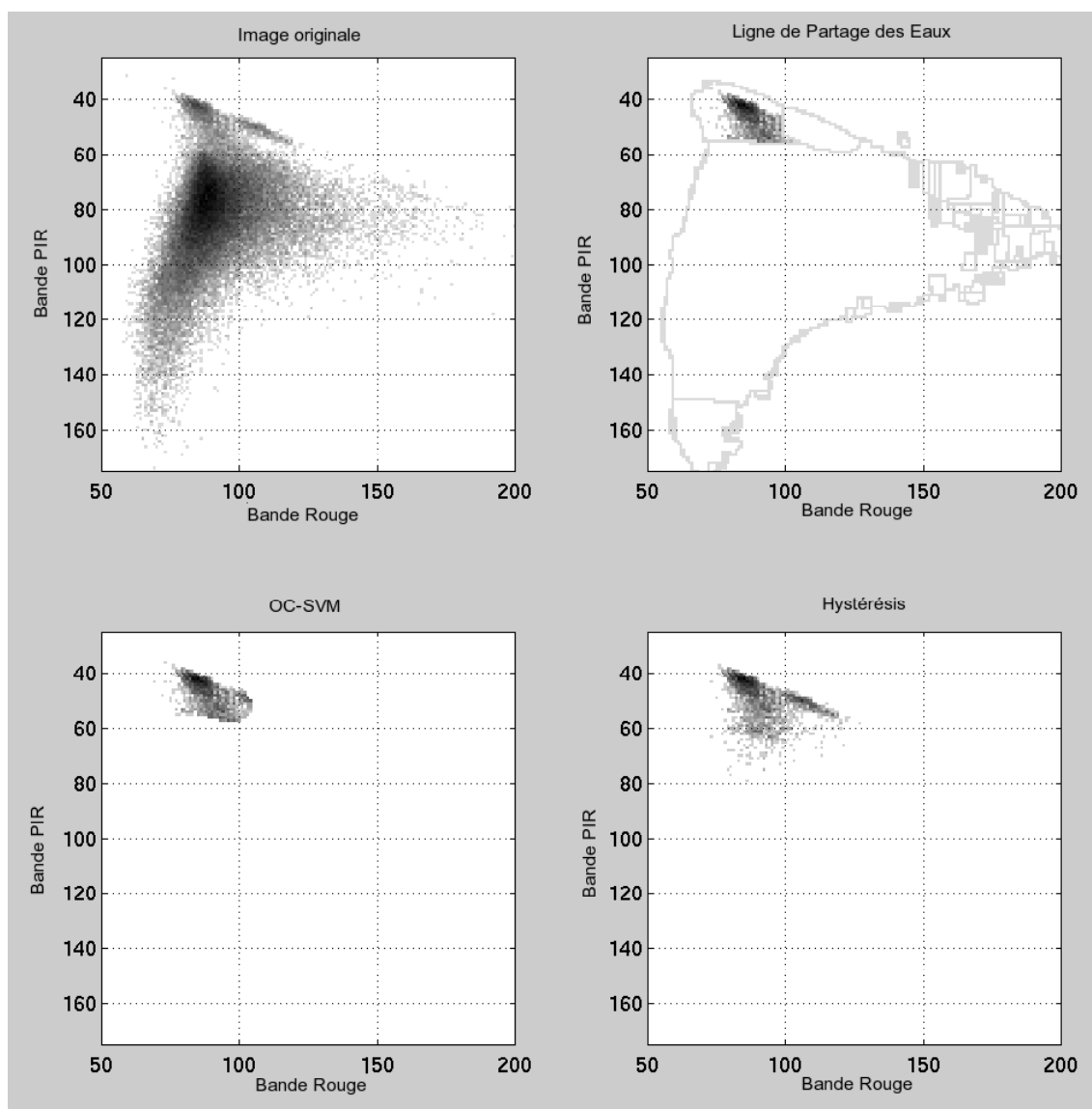


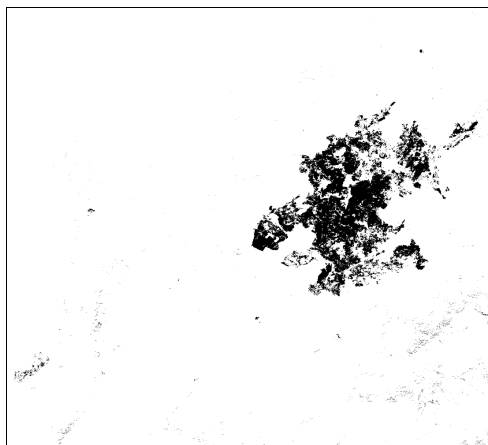
FIG. 5.5: Sections d'histogrammes lors de différentes étapes



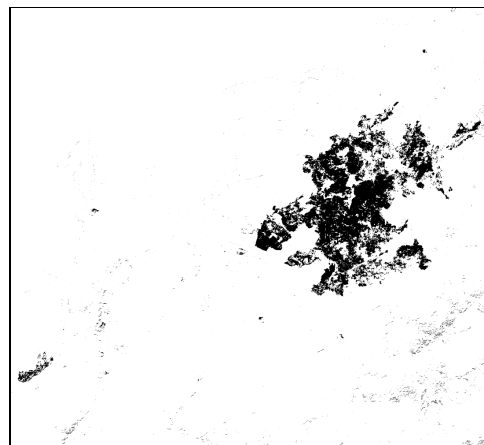
(a) Image SPOT 5 © CNES 2005, Distribution SPOT Image



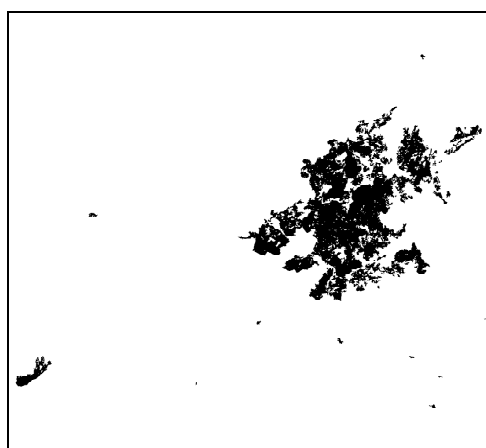
(b) Estimée Satellites - projet RISK-EOS ©Infoterra-ESA



(c) Histogramme



(d) Histogramme + OC-SVM



(e) Histogramme + OC-SVM + croissance de régions

FIG. 5.6: Ouest de la Corse, 2005 (environ 2800 ha)

plus, le seuillage par hystérésis permet de diminuer le nombre de faux positifs et d'améliorer les contours des zones brûlées. Les différences entre la cartographie finale et l'estimation satellitaire sont principalement composées de pixels situés sur les frontières des zones brûlées (du fait du recalage entre les deux images).

#### - Alpes Maritimes :

La figure 5.7 donne une section des histogrammes à une valeur de MIR donnée pour les étapes suivantes : histogramme obtenu après filtrage Gaussien (en haut à gauche), celui du "cluster" sélectionné après LPE (en haut à droite), celui de l'OC-SVM (en bas à gauche), et celui après seuillage par hystérésis (en bas à droite).

La figure 5.8 donne l'image satellitaire (les bandes spectrales Rouge, PIR et MIR sont respectivement représentées en nuance de bleu, vert et rouge), la carte de référence correspondante (réalisée par le SERTIT à partir d'images SPOT 5 de résolution 10m) et les classifications obtenues à partir des mêmes étapes que dans la figure 5.5.

Le tableau 5.4 compare les cartographies obtenues à partir de la segmentation de l'histogramme, des OC-SVM utilisant la base d'apprentissage définie par la segmentation de l'histogramme, de la Croissance de Régions (CR) fondée sur les OC-SVM utilisant la base d'apprentissage définie par la segmentation de l'histogramme. Les deux premières cartographies sont régularisées par champs de Markov afin d'améliorer leur cohérence spatiale et pouvoir être comparées avec la classification finale.

#### - Est de la Corse :

La figure 5.9 donne les mêmes images que précédemment.

Le tableau 5.3 compare les mêmes cartographies que précédemment. Les deux premières cartographies sont également régularisées par champs de Markov pour les mêmes raisons.

#### - Cas de petites zones brûlées :

La méthode proposée est testée dans le cas où l'image ne contient que des petites zones brûlées (de surface totale inférieure à 10 ha). La figure 5.10 fournit une partie de l'image Alpes-Maritimes, les cartographies des différentes étapes de la méthode proposée, ainsi que la superposition du contour et de l'image satellitaire.

La méthode proposée fournit de très bons résultats et ne dépend pas énormément de la taille de la zone brûlée, contrairement à celle fondée sur l'apprentissage par K-moyennes.

## 5.4 Conclusion

Les techniques SVM nécessitent une base d'apprentissage afin de pouvoir réaliser les phases d'apprentissage et de classification. Le choix de cette base est crucial car elle définit la fonction de décision finale. Elle est généralement déterminée visuellement sur le terrain ou par interprétation d'images [Mather, 2004] par des experts. Sa sélection est donc difficile et coûteuse.

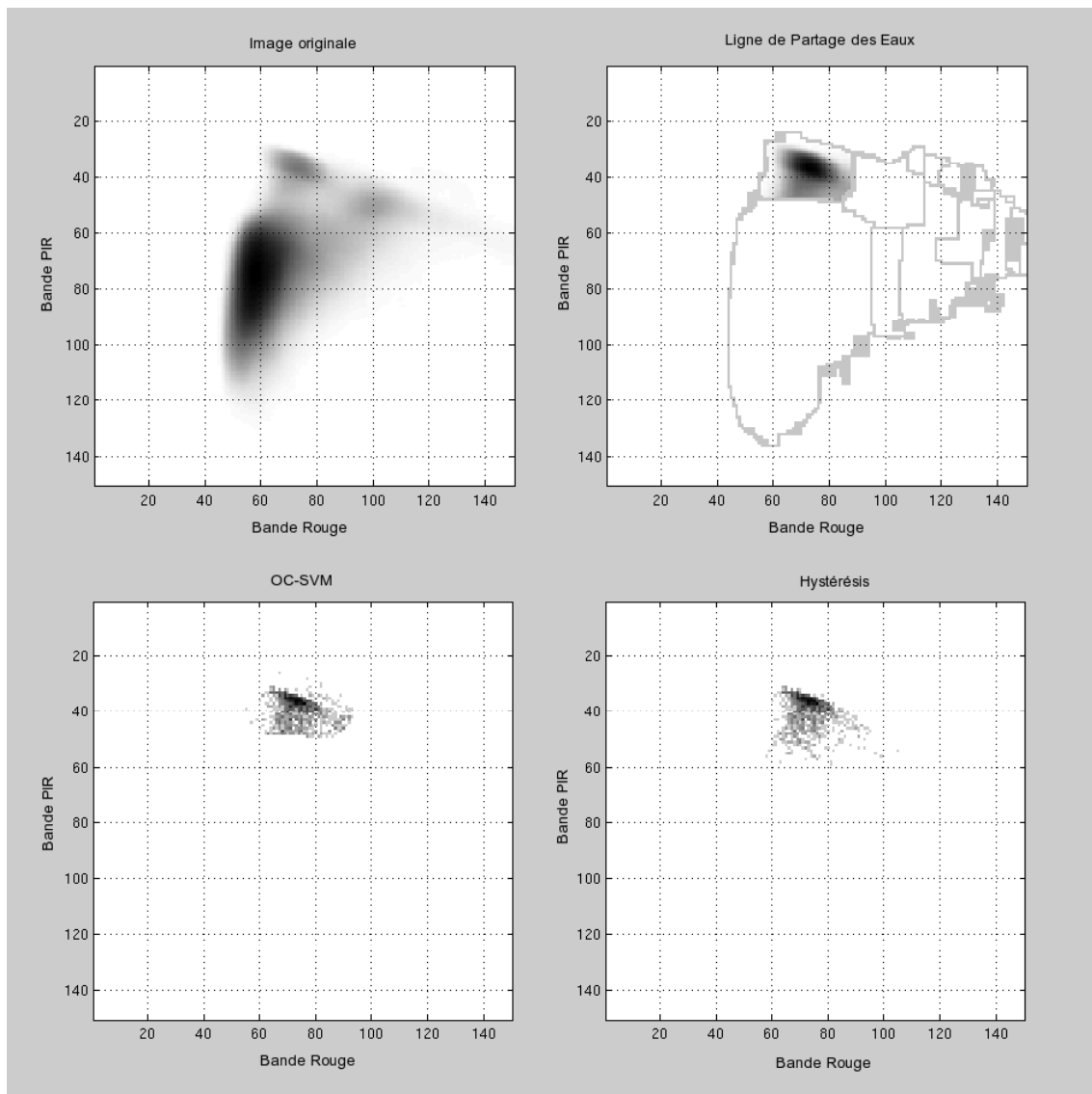
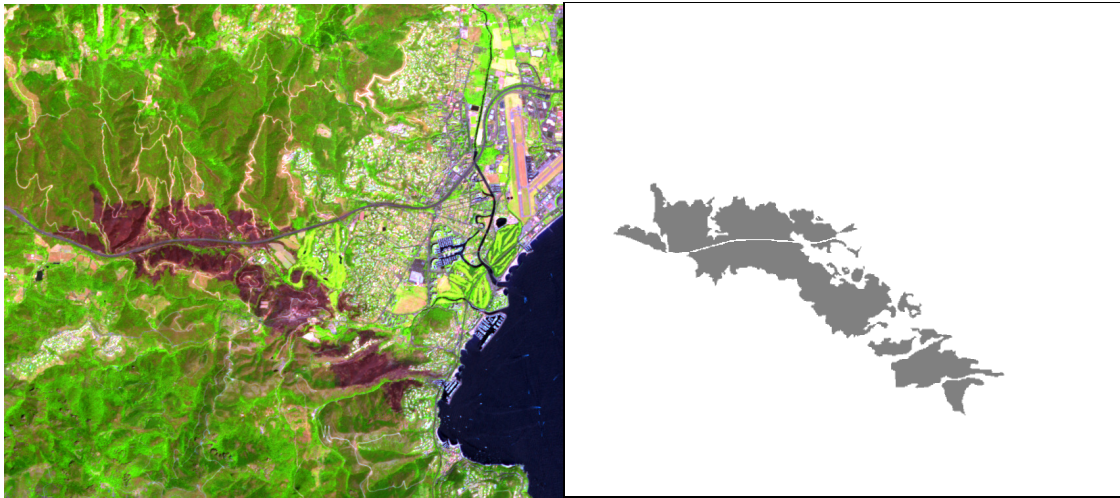
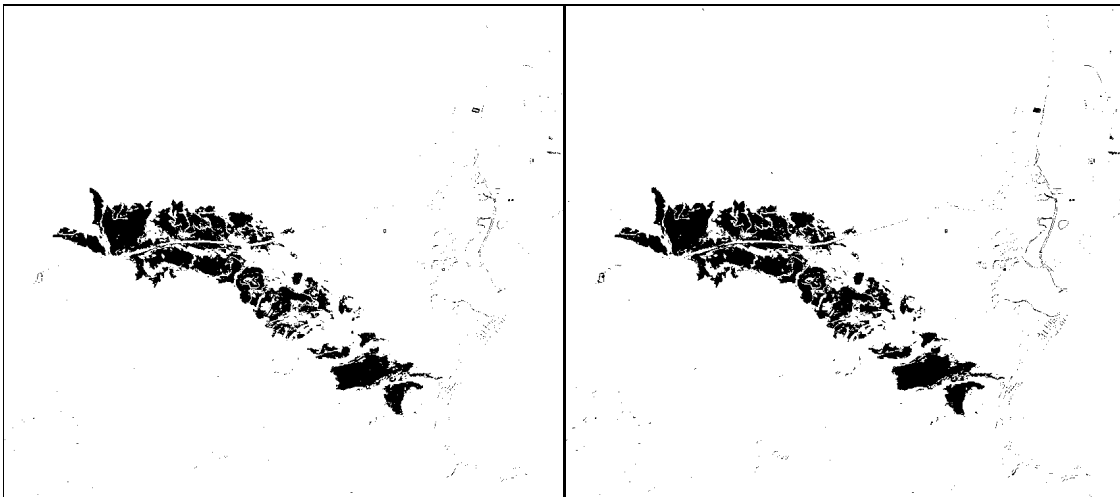


FIG. 5.7: Sections d'histogrammes lors de différentes étapes



(a) Image SPOT 5 © CNES 2005, Distribution SPOT Image

(b) Estimée Satellitaire ©SERTIT



(c) Histogramme

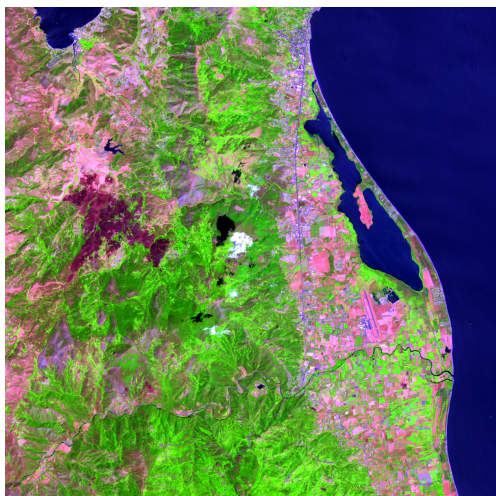
(d) Histogramme + OC-SVM



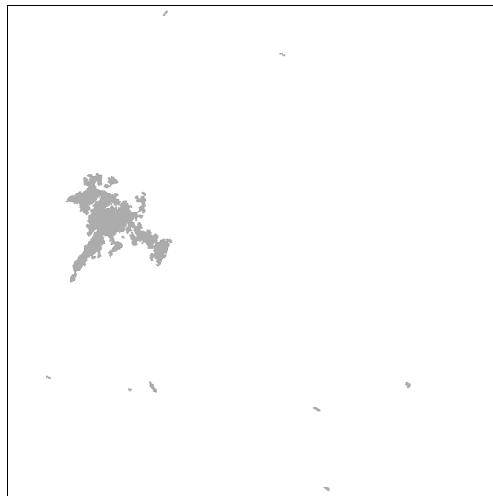
(e) Histogramme + OC-SVM + croissance de régions

FIG. 5.8: Alpes Maritimes, 2007 (environ 450 ha)

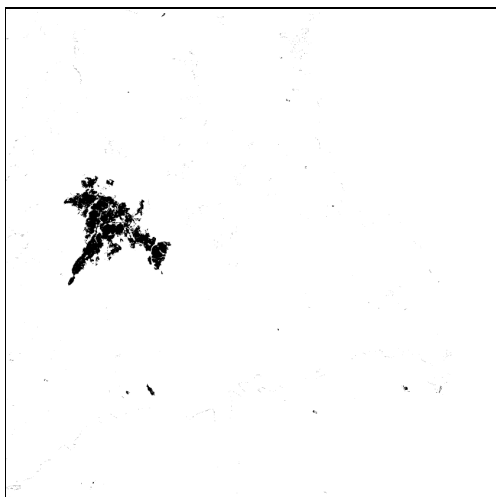




(a) Image SPOT 5 © CNES 2005, Distribution SPOT Image



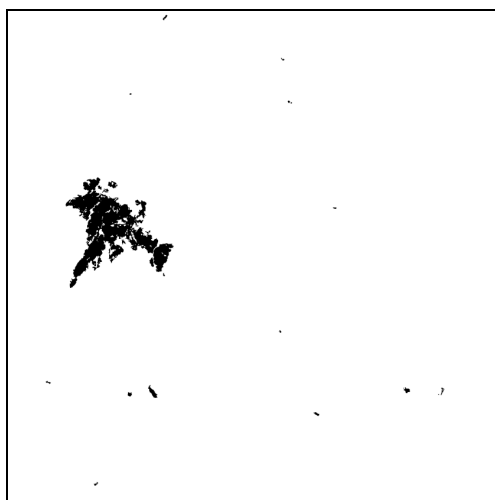
(b) Estimée Satellitaire - projet RISK-EOS ©Infoterra-ESA



(c) Histogramme



(d) Histogramme + OC-SVM

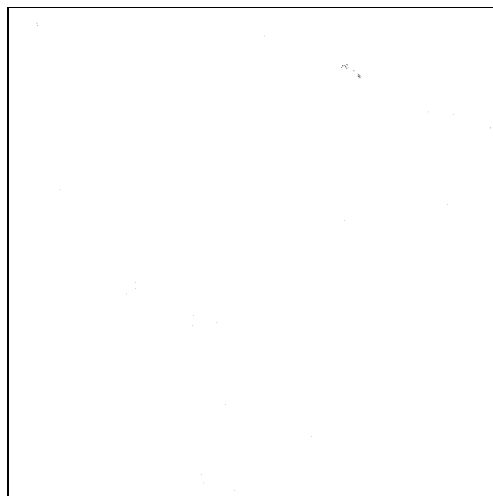


(e) Histogramme + OC-SVM + croissance de régions

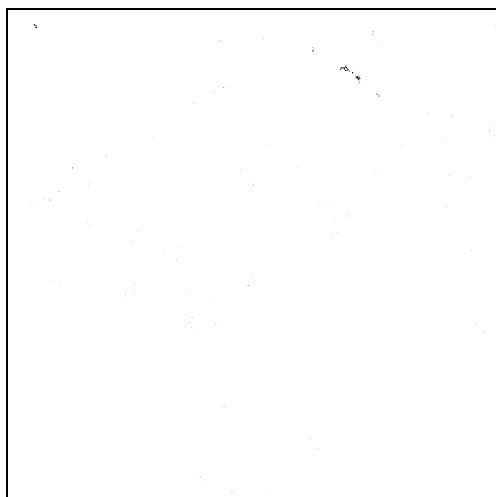
FIG. 5.9: Est de la Corse, 2005 (environ 1000 ha)



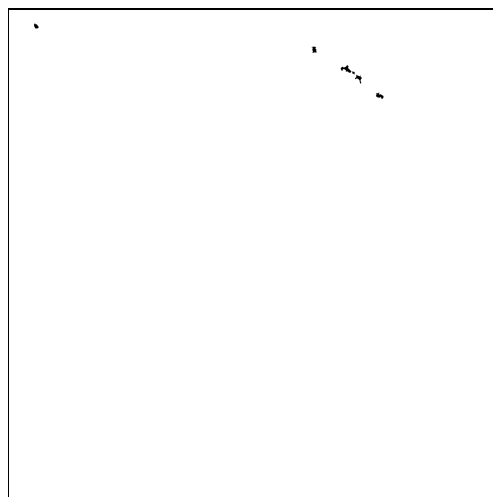
(a) Image SPOT 5 © CNES 2007, Distribution SPOT Image



(b) Histogramme



(c) Histogramme + OC-SVM



(d) Histogramme + OC-SVM + croissance de régions



(e) Superposition d'une partie de l'image SPOT 5 © CNES 2007, Distribution SPOT Image, et du contour obtenu par Histogramme + OC-SVM + croissance de régions

FIG. 5.10: Alpes Maritimes, 2007 (environ 7 ha)

Dans ce chapitre, nous avons proposé deux méthodes de sélection quasi-automatique de la base d'apprentissage.

La première utilise l'algorithme des K-moyennes afin de définir une première estimation de la classification qui va servir à déterminer une base d'apprentissage. L'utilisation, par les techniques SVM, de la base d'apprentissage obtenue permet d'améliorer la classification initiale par K-moyennes et fournit une bonne cartographie finale dans le cas où cette classification initiale comporte peu d'erreurs. Cependant, si la zone brûlée est de petite taille par rapport à la taille de l'image, la classification par K-moyennes ne donne pas de bons résultats.

La deuxième méthode est fondée directement sur les valeurs spectrales des pixels brûlés de l'image, ce qui permet une première estimation de la zone brûlée plus précise que par la méthode précédente et indépendante de la taille des zones brûlées. L'histogramme de l'image est segmenté, ce qui délimite les différents pics de l'histogramme. Le pic correspondant à la zone brûlée constitue la base d'apprentissage. La classification fondée sur les OC-SVM à partir de cette base d'apprentissage, puis utilisée comme technique de croissance de régions afin d'ajouter une information spatiale, donne une très bonne concordance entre la cartographie extraite et l'image satellitaire. En effet, visuellement, les contours des zones brûlées sont très précis. Les taux de faux positifs plus élevés sont dus aux problèmes liés au recalage des cartographies avec les estimations obtenues à partir d'un satellite de résolution supérieure.

La base d'apprentissage définie automatiquement peut être utilisée directement par les techniques SVM et fournit ainsi de très bons résultats. L'utilisateur peut aussi ajouter à cette base des pixels pertinents afin d'améliorer la cartographie finale. Cette base d'apprentissage peut également servir comme point de départ pour les algorithmes de classification semi-supervisée.

Carte de référence (2847 ha)	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
Histogram + MM	2068	97.55	83.50	0.36	32.63	5
Histogram + OC SVM + MM	2256	97.80	86.10	0.45	27.34	15
Histogram + OC SVM + CR	2740	98.14	91.45	0.86	16.23	16

TAB. 5.2: Ouest de la Corse - 2005 : Comparaison avec la carte de référence

Carte de référence (992 ha)	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
Histogram + MM	842	99.61	90.02	0.08	19.88	5
Histogram + OC SVM + MM	886	99.63	91.49	0.10	16.93	15
Histogram + OC SVM + CR	1004	99.65	94.67	0.19	10.47	16

TAB. 5.3: Est de la Corse - 2005 : Comparaison avec la carte de référence

Carte de référence (449 ha)	Surface brûlée (ha)	PT (%)	PM (%)	FP (%)	FN (%)	temps (min)
Histogram + MM	303	97.53	82.95	0.03	34.07	5
Histogram + OC SVM + MM	331	97.92	85.82	0.05	28.30	15
Histogram + OC SVM + CR	381	98.54	90.65	0.14	18.57	16

TAB. 5.4: Alpes Maritimes - 2007 : Comparaison avec la carte de référence



# Conclusion et Perspectives

La cartographie des zones brûlées après un feu de forêt constitue une étape importante pour la gestion de la situation post-crise. La plupart des méthodes d'évaluation des dégâts sont fondées sur l'analyse des changements d'indices de végétation entre images satellitaires basse ou moyenne résolution.

Dans ce manuscrit de thèse, nous abordons le problème de détection des zones brûlées à partir d'une seule image satellitaire haute résolution.

Tout d'abord, la classification en zones brûlées et non brûlées a été réalisée uniquement à partir de l'information spectrale apportée par les différents canaux du satellite SPOT 5. En effet, les zones brûlées sont caractérisées par une réflectance faible dans les bandes Vert et Rouge, très faible dans le Proche Infra-Rouge et moyenne dans le Moyen-Infra-Rouge, ce qui permet de les différencier de la plupart des objets contenus dans l'image satellitaire. Différentes techniques de classification supervisée ont été testées comme l'algorithme des K-Plus Proches Voisins, celui du Maximum de Vraisemblance ou les "Support Vector Machines" (SVM).

Les SVM, qui ont démontré d'excellentes capacités de classification et de généralisation dans de nombreux domaines, fournissent aussi de meilleurs résultats pour notre application.

Les zones brûlées sont caractérisées par des valeurs spectrales assez homogènes, contrairement aux zones non brûlées dont les caractéristiques spectrales varient énormément selon leur catégorie (végétation, routes, eau, zones urbaines, nuages, ombre...). Nous avons donc étudié les "One-Class SVM", une variante de la méthode SVM qui n'utilise que des exemples positifs (i.e. brûlés) afin de définir la classification. Evidemment, cette technique fournit des résultats un peu moins bons que ceux obtenus par SVM classique, mais est beaucoup plus rapide et permet de s'affranchir de la sélection d'éléments négatifs (non-brûlés) nécessaire dans les phases d'apprentissage et de classification.

L'analyse simple de la "couleur" donne une très bonne cartographie des dégâts pour les zones brûlées de très grandes tailles, mais est moins précise sur les contours des zones brûlées de petites tailles, qui sont généralement moins gravement touchées. En effet, les caractéristiques spectrales de certains bords des zones brûlées sont similaires à celles de zones de végétation peu saine. De plus, en raison du bruit présent dans l'image satellitaire, il existe de nombreux petits groupes de pixels isolés.

Une régularisation classique, par morphologie mathématique ou par champs de Markov, permet de supprimer ce bruit des cartographies mais ne permet pas d'améliorer les contours

des zones brûlés.

Le seuillage unique des techniques SVM fournit une loi de décision trop stricte. Nous proposons donc d'utiliser un seuillage par hystérésis :

- l'information apportée par les techniques SVM quantifie l'incertitude liée à l'affectation du pixel de l'image à chaque classe donnée,
- pour les pixels ayant un faible taux de confiance (i.e. dont la valeur se situe entre deux seuils), leur affectation dépend de celles des pixels voisins.

Ainsi, la classification est fondée à la fois sur l'information spectrale et sur l'information spatiale. Les techniques SVM sont utilisés comme des techniques de croissance de régions, ce qui améliore les contours des zones détectées.

Les deux seuils définissent un compromis entre les erreurs de commission et d'omission (i.e. entre sur-cartographier ou sous-cartographier les zones brûlées).

Les techniques SVM nécessitent une base d'exemples afin de pouvoir classifier l'image entière. Une base d'exemple optimale pour une image ne fournit généralement pas de bons résultats pour une autre image, en raison de la différence d'illumination entre les deux images par exemple. Deux techniques de sélection automatique de la base d'apprentissage ont été proposées.

La première est fondée sur l'algorithme des K-moyennes afin de définir grossièrement une première classification des données. Cette méthode fournit de bons résultats dans le cas où la zone brûlée est de grande taille, en comparaison avec celle de l'image utilisée.

La deuxième méthode est fondée sur la segmentation de l'histogramme 3-Dimensions de l'image par ligne de partage des eaux, ce qui permet de fournir une meilleure première classification des données que par K-moyennes dans le cas où la totalité des surfaces brûlées est de petite taille, comparée à celle de l'image utilisée. En effet, la ligne de partage des eaux permet de segmenter tous les pics de l'histogramme, contrairement à l'algorithme des K-Moyennes. Cette première classification n'est pas assez précise, mais permet de fournir une base d'apprentissage aux techniques SVM. Combinée avec le seuillage par hystérésis, cette technique fournit une cartographie avec une très bonne précision, même pour des surfaces brûlées de quelques hectares.

La méthode de classification finale n'est pas complètement automatique : les deux seuils optimaux et le pic de l'histogramme correspondant à la zone brûlée sont obtenus manuellement. Cependant, le seuillage par hystérésis est très rapide et de nombreux couples de seuils fournissent de bons résultats ; donner un seul point brûlé dans l'image permet de déterminer le bon pic de l'histogramme. A des fins opérationnelles, les cartographies réalisées doivent être au format vecteur et géoréférencées afin d'être exportables vers un Système d'Information Géographique (SIG) utilisateur.

De nombreuses images ont été testées afin de valider l'algorithme final par rapport au type de végétation brûlée (feuillus, résineux, landes, garrigues, maquis, ...), et à la taille de la zone brûlée.

Néanmoins, il faudrait aussi tester l'algorithme sur des feux dans des régions non méditerranéennes (feux de brousse en Afrique, par exemple), ce que nous n'avons pu faire

par manque de données satellitaires et cartes de référence associées.

Il serait intéressant de tester l'algorithme sur des images provenant de satellites futurs, de résolution spatiale et spectrale différentes (Pléiades, Sentinelle). L'utilisation d'images Radar à Synthèse d'Ouverture (RSO), ce qui permettrait d'acquérir des images quelle que soit l'heure même en présence de nuages, pourrait être investiguée.

Dans l'optique de l'évaluation d'une éventuelle reprise de végétation, [Brewer et al., 2005; Lentile et al., 2006] proposent d'utiliser l'indice de végétation NBR afin d'évaluer la sévérité du feu et donc d'estimer les zones plus ou moins endommagées par le feu.

A partir des cartographies, l'utilisation d'images satellitaires acquises après l'extinction de l'incendie permettrait de réaliser le suivi de la régénération de la végétation.

Dans ce manuscrit de thèse, le but recherché était de cartographier les zones brûlées après un feu de forêt. L'algorithme non supervisé de classification binaire, fondé sur l'utilisation des informations spectrale et spatiale, qui a été développé, a été testé avec succès sur une seule application. Il serait intéressant d'essayer l'algorithme sur d'autres applications de classification, dans le cas binaire ou multi-classes.

Nous pouvons également entrevoir des perspectives méthodologiques :

- L'algorithme supervisé "Relevant Vector Machines" (RVM) [Tipping, 2000; Tipping, 2001] semble être une alternative intéressante aux SVM : il ne dépend que d'un seul paramètre, fournit une loi de décision qui utilise moins de "Vecteurs Supports" (il est donc très rapide de classifier de nouveaux éléments) pour une même précision. Son seul inconvénient réside néanmoins dans la complexité de la phase d'apprentissage.
- Il existe parfois des petites zones brûlées de très petites tailles (quelques pixels) à proximité de plus grandes zones brûlées. Ces petites zones sont effacées par la première étape ou par le processus d'érosion lors du seuillage par hystérésis. Ajouter une liberté sur la distance entre ces petites zones et la zone brûlée principale [Pilar Martin et al., 2006] peut permettre de conserver ces zones et ainsi, d'améliorer la cartographie.
- La base d'apprentissage sélectionnée automatiquement n'est pas forcément optimale. Un apprentissage semi-supervisé pourrait permettre de construire un meilleur classifieur en utilisant des pixels non classés de l'image satellitaire.





## Annexe A

# Séparabilité spectrale des zones brûlées

Cette annexe présente les caractéristiques spectrales des zones brûlées et non brûlées et permet d'évaluer les performances des différentes bandes spectrales (utilisation d'une seule ou plusieurs bandes) pour la détection des zones brûlées.

### A.1 Utilisation d'une seule bande spectrale

Un indice largement utilisé et qui permet de mesurer la séparabilité des zones brûlées des zones non brûlées pour chaque bande est donné par [Pereira, 1999] :

$$M = \frac{|\mu_b - \mu_{nb}|}{\sigma_b + \sigma_{nb}}$$

avec, pour la bande spectrale considérée :

$\mu_b$  et  $\mu_{nb}$  : valeurs moyennes de la classe brûlée et non-brûlée,

$\sigma_b$  et  $\sigma_{nb}$  : écart-types de la classe brûlée et non-brûlée.

Des valeurs de  $M$  supérieures à 1 indiquent une bonne séparabilité pour la bande spectrale considérée. Au contraire, des valeurs proches de 0 indiquent une grande superposition des deux classes.

La figure A.1 montre les histogrammes des zones brûlées (en bleu) et non-brûlées (en vert) pour les 4 canaux de SPOT 5 (Vert, Rouge, PIR et MIR) pour 4 images satellitaires SPOT 5 (Centre, Ouest, Est de la Corse et Alpes-Maritimes). Le tableau A.1 donne l'indice de séparabilité pour chaque bande de 4 zones brûlées différentes (des 4 images SPOT 5).

Les pics composés par de nombreux pixels correspondant à la mer ont été retirés de l'image avant le calcul des histogrammes et des indices de séparabilité. La mer a une très faible réflectance dans les 4 bandes, ce qui aurait faussé le calcul des moyennes et écarts-types dans les indices de séparabilité [Garcia and Chuvieco, 2004].

Les histogrammes et les indices de séparabilité correspondants montrent que la bande spectrale PIR est la plus pertinente pour la détection de zones brûlées. Pour les autres bandes, le degré de recouvrement est très important.

Les histogrammes montrent aussi que les zones brûlées réfléchissent moins de Proche Infra-Rouge, mais plus de Moyen Infra-Rouge que les zones non brûlées. Les réflectances des deux autres bandes semblent identiques.

## A.2 Utilisation de deux bandes spectrales

La figure A.2 donnent les scattergrammes (graphes qui montrent la répartition des points selon 2-dimensions) des bandes Rouge (à gauche) et MIR (à droite) en fonction de la bande PIR pour deux images satellitaires SPOT 5 (Est et Ouest de la Corse). Contrairement aux histogrammes et indices de séparabilité précédemment étudiés, les scattergrammes présentés ici prennent en compte les pixels correspondant à la mer. Le scattergramme de la bande Vert en fonction de la bande PIR n'est pas donné, car elle ressemble énormément à celle de la bande Rouge.

Dans les scattergrammes, les points correspondants uniquement aux zones brûlées sont affichés en bleu, ceux qui correspondent uniquement aux zones non brûlées sont affichés en vert et ceux qui appartiennent aux deux zones sont affichés en rouge.

### Image Est de la Corse :

Dans le scattergramme Rouge/PIR, tous les pixels correspondant aux zones brûlées correspondent aussi aux zones non-brûlées, ils ont les mêmes caractéristiques spectrales dans le PIR et le Rouge. Dans le scattergramme MIR/PIR, on peut distinguer de nombreux points correspondant uniquement aux zones brûlées. Cependant, les nuages de points brûlés et non brûlés ne sont pas complètement distincts.

### Image Ouest de la Corse :

Dans le scattergramme Rouge/PIR, presque tous les pixels correspondant aux zones brûlées correspondent aussi aux zones non-brûlées. Dans le scattergramme MIR/PIR, on peut tirer les mêmes conclusions que pour l'image précédente.

En ne considérant que l'image après feux, l'information spectrale apportée par l'association des bandes MIR et PIR peut donc permettre de détecter grossièrement les zones brûlées au contraire de celles apportées par l'association des bandes Rouge (ou Vert) et PIR.

## A.3 Utilisation d'au moins trois bandes spectrales

Si l'on couple maintenant les bande Vert, Rouge et PIR, le nuage de points brûlés apparaît. Il est alors possible de distinguer des zones brûlées et non-brûlées. Cependant, il existe toujours une petite zone de recouvrement des deux zones (brûlée/non brûlée). En effet, les caractéristiques spectrales des zones brûlées varient selon du type de la végétation,

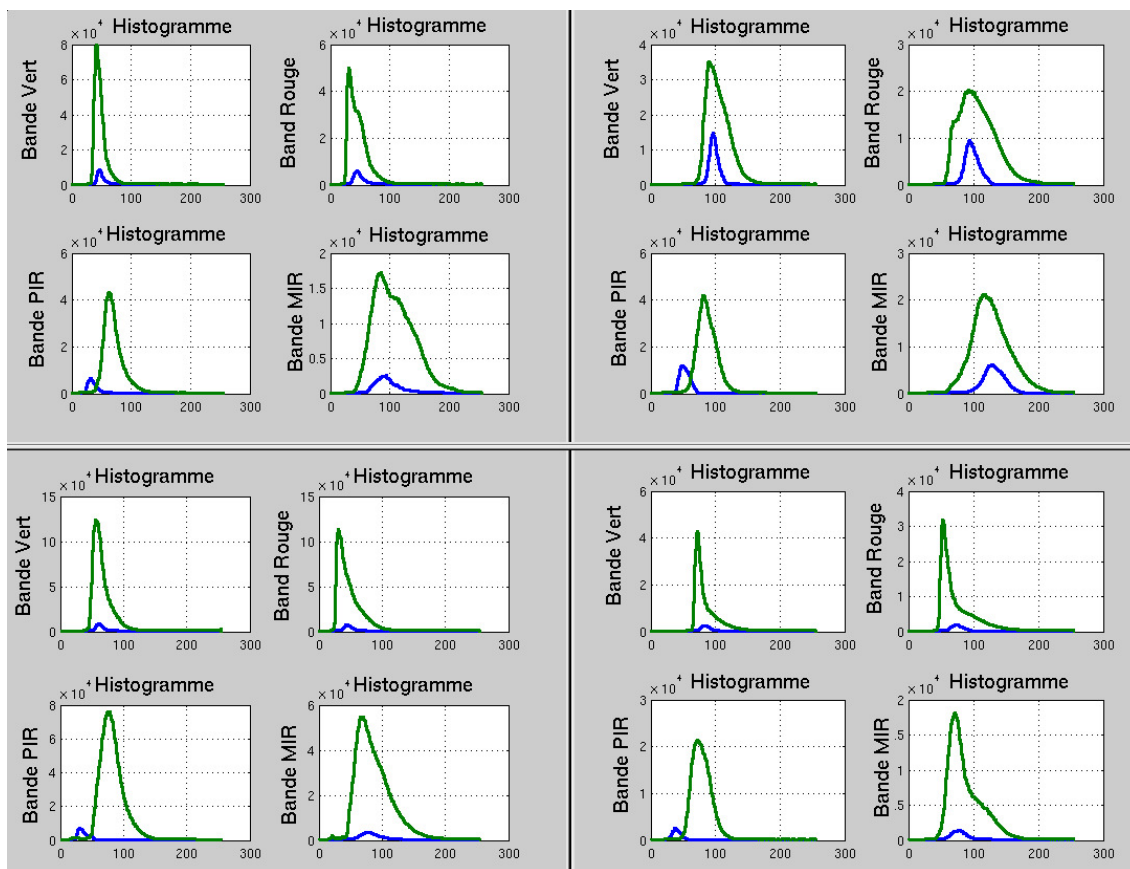
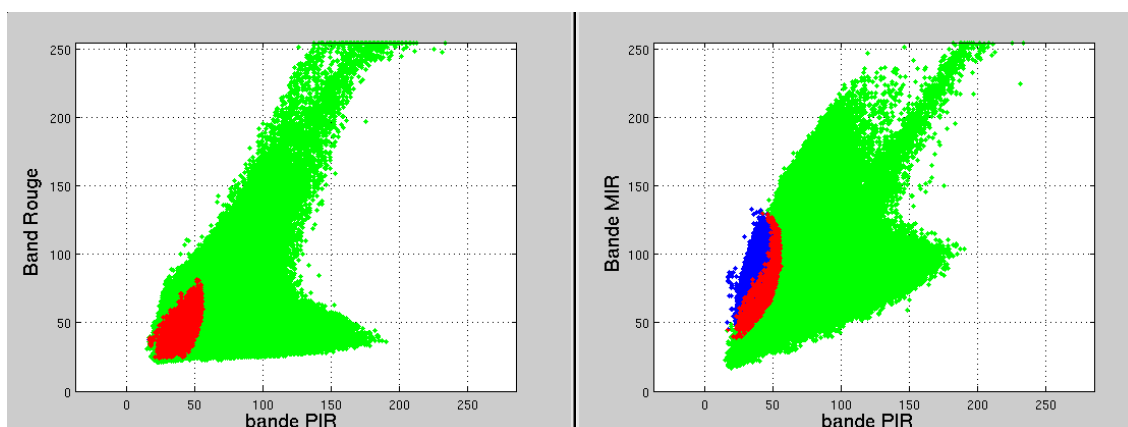


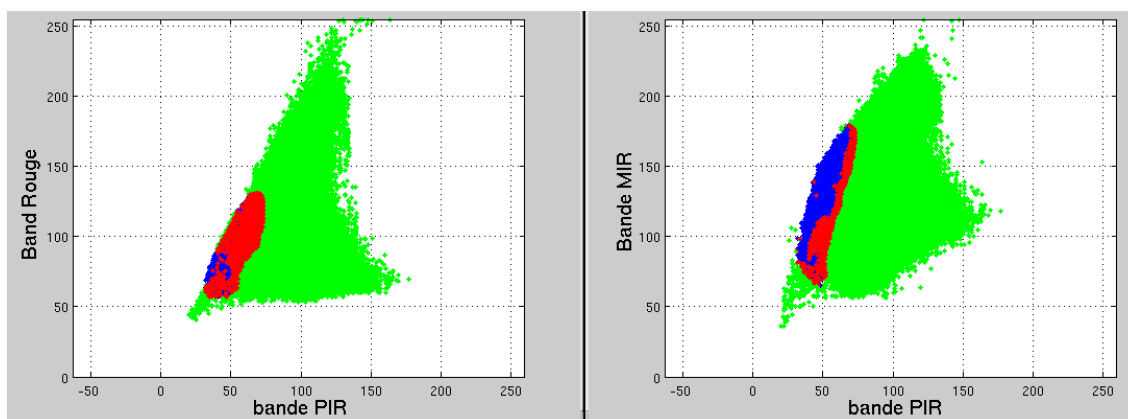
FIG. A.1: Histogrammes des zones brûlées (en bleu) et non-brûlées (en vert) pour les 4 canaux de SPOT 5 pour 4 images différentes

	Bande Vert	Bande Rouge	Bande PIR	Bande MIR
Nord de la Corse	0.1099	0.1070	1.5425	0.2137
Ouest de la Corse	0.0117	0.0488	1.9660	0.246
Est de la Corse	0.1668	0.0184	1.8572	0.1219
Alpes-Maritimes	0.2775	0.2129	1.5911	0.0956
Moyenne	0.1415	0.0968	<b>1.7392</b>	0.1693

TAB. A.1: Valeurs de l'indice de séparabilité  $M$



(a) Image Est de la Corse



(b) Image Ouest de la Corse

FIG. A.2: Scattergrammes des zones brûlées (en bleu), non-brûlées (en vert) et recouvrement des deux zones (en rouge) en fonction de différentes bandes

de la sévérité du feu et le décalage temporel entre l'extinction du feu et l'acquisition de l'image. Elles ne sont donc pas toutes entièrement identifiables par leurs caractéristiques spectrales.

## A.4 Conclusion

La bande spectrale la plus pertinente concernant la détection des zones brûlées est la bande Proche Infra-Rouge. Elle ne permet pas, cependant, de différencier directement les zones brûlées et non brûlées.

L'utilisation de deux bandes spectrales ne permet pas non plus de déterminer complètement les zones brûlées. On peut néanmoins remarquer que l'utilisation des bandes PIR et MIR permet de séparer spectralement une partie des zones brûlées et non brûlées. C'est pourquoi, pour les méthodes de classification fondées sur la détection de changements

entre images avant et après feux, les indices de végétation qui utilisent ces deux bandes donnent les meilleurs résultats.

Au moins trois bandes doivent donc être utilisées pour notre problème de cartographie de zones brûlées. Dans les images fausses couleurs présentées dans ce manuscrit de thèse et qui représente les bandes Vert, Rouge et PIR, les zones brûlées sont identifiables visuellement, il n'est donc pas nécessaire de disposer obligatoirement de la bande MIR.

Quel que soit le nombre de bandes et la combinaison utilisés, les caractéristiques spectrales des zones brûlées et non brûlées se chevauchent, il convient donc d'utiliser à la fois les informations spectrale et spatiale afin de détecter les dégâts après un feu de forêt.



## Annexe B

# Algorithme "Sequential Minimal Optimization"

L'algorithme *Sequential Minimal Optimization* (SMO) permet de résoudre rapidement et simplement le Problème de Programmation Quadratique (PPQ) des SVM :

$$\max_{\lambda} W(\lambda) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j K(\vec{x}_i, \vec{x}_j) + \sum_{i=1}^N \lambda_i \quad (\text{B.1})$$

$$\text{tels que : } \sum_{i=1}^N \lambda_i y_i = 0, \quad 0 \leq \lambda_i \leq C, \quad \forall i \in \langle 1, N \rangle \quad (\text{B.2})$$

Il décompose le PPQ de dimension  $N$  en une suite de sous-problèmes de dimension plus petite, qui peuvent être résolus analytiquement et ne nécessitent plus d'optimisation quadratique.

Dans le cas des SVM, le plus petit problème d'optimisation implique deux multiplicateurs de Lagrange à cause de la contrainte d'égalité linéaire. A chaque itération, l'algorithme choisit deux multiplicateurs à optimiser conjointement. Même si le nombre d'itérations est très important, l'optimisation est beaucoup plus rapide. De plus, il ne nécessite pas la manipulation d'une matrice de grande taille, très gourmande en mémoire.

Les trois étapes de l'algorithme sont : la méthode analytique pour résoudre un problème de dimension 2, la mise à jour du paramètre  $b$  tel que  $f(\vec{x}) = \sum_{i=1}^N \lambda_i y_i K(\vec{x}, \vec{x}_i) - b$  et le choix des multiplicateurs à optimiser.

### B.1 Optimisation analytique

#### Recherche du maximum :

Chaque étape du SMO optimise deux multiplicateurs que nous appellerons  $\lambda_1$  et  $\lambda_2$



dans la suite. La fonction à maximiser est :

$$W(\lambda_1, \lambda_2) = -\frac{1}{2}K_{11}\lambda_1^2 - \frac{1}{2}K_{22}\lambda_2^2 - sK_{12}\lambda_1\lambda_2 - y_1\lambda_1v_1 - y_2\lambda_2v_2 + \lambda_1 + \lambda_2 + W_{constant}$$

tels que :  $\lambda_1 + s\lambda_2 = \gamma$

où  $K_{ij} = K(\vec{x}_i, \vec{x}_j)$ ,  $s = y_1y_2$ ,  $v_i = \sum_{j=3}^N \lambda_j y_j K_{ij} = f(\vec{x}_i) + b - \lambda_1 y_1 K_{1i} - \lambda_2 y_2 K_{2i}$

En remplaçant  $\lambda_1$  par  $\gamma - s\lambda_2$ , il vient :

$$W = -\frac{1}{2}K_{11}(\gamma - s\lambda_2)^2 - \frac{1}{2}K_{22}\lambda_2^2 - sK_{12}(\gamma - s\lambda_2)\lambda_2 - y_1(\gamma - s\lambda_2)v_1 - y_2\lambda_2v_2 + \gamma - s\lambda_2 + \lambda_2 + W_{constant}$$

Le point stationnaire est obtenu en annulant la dérivée :

$$\begin{aligned} \frac{dW}{d\lambda_2} &= sK_{11}(\gamma - s\lambda_2) - K_{22}\lambda_2 + K_{12}\lambda_2 - sK_{12}(\gamma - s\lambda_2) + y_2v_1 - y_2v_2 - s + 1 \\ &= 0 \end{aligned}$$

ce qui donne :

$$(K_{11} + K_{22} - 2K_{12})\lambda_2^{new} = s(K_{11} - K_{12})\gamma + y_2(v_2 - v_1) + 1 - s$$

En remplaçant  $\gamma$  et  $v$  par leurs valeurs, on obtient :

$$(K_{11} + K_{22} - 2K_{12})\lambda_2^{new} = (K_{11} - 2K_{12} + K_{22})\lambda_2^{old} + y_2(f(\vec{x}_1) - f(\vec{x}_2)) + y_2 - y_1$$

Finalement,

$$\lambda_2^{new} = \lambda_2^{old} - \frac{y_2(E_1 - E_2)}{\eta} \quad (\text{B.3})$$

où  $0 > \eta = 2K_{12} - K_{11} - K_{22}$  correspond à l'opposée de la dérivée seconde de la fonction  $W$  définie positive et  $E_i = f^{old}(x_i) - y_i$  désigne l'erreur sur le  $i$ ème élément de la base d'apprentissage.

### Recherche de la solution qui satisfait les contraintes :

Si  $y_1 = y_2$ , alors  $s = 1$  et  $\lambda_1^{new} + \lambda_2^{new} = \lambda_1^{old} + \lambda_2^{old}$ . On obtient donc :

$$\begin{aligned} 0 &\leq \lambda_2^{new} \leq C \\ \text{et } \lambda_1^{old} + \lambda_2^{old} - C &\leq \lambda_2^{new} \leq \lambda_1^{old} + \lambda_2^{old} \end{aligned}$$

Si  $y_1 \neq y_2$ , alors  $s = -1$  et  $\lambda_2^{new} - \lambda_1^{new} = \lambda_2^{old} - \lambda_1^{old}$ . On obtient donc :

$$\begin{aligned} 0 &\leq \lambda_2^{new} \leq C \\ \text{et } \lambda_2^{old} - \lambda_1^{old} &\leq \lambda_2^{new} \leq \lambda_2^{old} - \lambda_1^{old} + C \end{aligned}$$

Le maximum de la fonction sous contraintes est obtenu en bornant la valeur de  $\lambda_2^{new}$  :

$$\lambda_2^{new,clipped} = \begin{cases} H & \text{si } H \leq \lambda_2^{new} \\ \lambda_2^{new} & \text{si } L < \lambda_2^{new} < H \\ L & \text{si } \lambda_2^{new} \leq L \end{cases}$$

en notant

$$\begin{aligned} L &= \max(0, \lambda_1^{old} + \lambda_2^{old} - C), & M &= \max(C, \lambda_1^{old} + \lambda_2^{old}) & \text{si } y_1 = y_2 \\ L &= \max(0, \lambda_2^{old} - \lambda_1^{old}), & M &= \max(C, C + \lambda_2^{old} - \lambda_1^{old}) & \text{sinon} \end{aligned}$$

La valeur de  $\lambda_1^{new}$  est calculée directement à partir de celle de  $\lambda_2^{new}$  :

$$\lambda_1^{new} = \lambda_1^{old} + s(\lambda_2^{old} - \lambda_2^{new,clipped})$$

## B.2 Choix des multiplicateurs à optimiser

À chaque itération, l'algorithme SMO optimise deux multiplicateurs de Lagrange. Si au moins l'un des deux multiplicateurs ne satisfait pas les conditions de Karush-Kuhn-Tucker (KKT) avant l'itération, la fonction objectif augmente et l'algorithme converge asymptotiquement. De façon à accélérer la convergence, le SMO utilise des heuristiques afin d'optimiser le choix du couple de multiplicateurs à sélectionner.

Les conditions de KKT sont :

$$\begin{aligned} \lambda_i = 0 &\Leftrightarrow y_i \cdot f(\vec{x}_i) > 1 \\ 0 < \lambda_i < C &\Leftrightarrow y_i \cdot f(\vec{x}_i) = 1 \\ \lambda_i = C &\Leftrightarrow y_i \cdot f(\vec{x}_i) < 1 \end{aligned}$$

Une première boucle itère sur tous les vecteurs de la base d'apprentissage et s'arrête dès qu'un vecteur ne satisfait pas les conditions de KKT.

Dans un second temps, une deuxième boucle itère seulement sur les Vecteurs Supports (VS) de la base d'apprentissage (dont les multiplicateurs sont strictement compris entre 0 et C) et s'arrête aussi dès qu'un VS ne satisfait pas les conditions de KKT.

La deuxième boucle réitère plusieurs fois sur les VS tant que la base d'apprentissage se satisfait pas les conditions de KKT.

Cette heuristique se concentre surtout sur les éléments les plus susceptibles de violer les conditions de KKT, i.e. les VS. En effet, lors de l'optimisation de deux multiplicateurs, les multiplicateurs qui valent 0 ou C sont plus susceptibles de ne pas évoluer, au contraire des multiplicateurs correspondant aux VS. C'est pourquoi, l'algorithme SMO itère de nombreuses fois sur les SV jusqu'à ce qu'ils satisfassent tous les conditions de KKT, puis il itère sur les éléments qui ne satisfont pas les conditions de KKT.

Les conditions de KKT doivent être satisfaites avec une erreur  $\epsilon$ . Généralement,  $\epsilon$  vaut  $10^{-3}$  : il est convenable que les valeurs de  $y_i \cdot f(\vec{x}_i)$  soient comprises entre 0.999 et 1.001. Si l'on désire une précision plus fine, l'algorithme convergera bien évidemment plus lentement.

Une fois le premier multiplicateur de Lagrange choisi, l'algorithme SMO sélectionne le deuxième afin de maximiser leurs variations lors de l'optimisation conjointe. Comme le calcul du noyau  $K$  demande beaucoup de temps, le SMO évalue juste le numérateur de l'équation B.3 :  $|E_1 - E_2|$ .

Lorsque tous les éléments sont en accord avec les conditions de KKT, l'algorithme est terminé.

### B.3 Mise à jour des paramètres

La valeur du paramètre  $b$  est mise à jour après chaque itération afin que le couple de multiplicateurs optimisé vérifie les conditions de KKT :  $b$  est calculé à partir des VS de la base d'apprentissage.  $b_1$  est valide si  $\lambda_1^{new}$  est strictement compris entre 0 et C car il oblige la sortie  $f(\vec{x}_1)$  à valoir  $y_1$  :

$$b_1 = E_1 + y_1(\lambda_1^{new} - \lambda_1^{old})K_{11} + y_2(\lambda_2^{new,clipped} - \lambda_2^{old})K_{12} + b^{old}$$

De même,  $b_2$  est valide si  $\lambda_2^{new}$  est strictement compris entre 0 et C car il oblige la sortie  $f(\vec{x}_2)$  à valoir  $y_2$  :

$$b_2 = E_2 + y_1(\lambda_1^{new} - \lambda_1^{old})K_{12} + y_2(\lambda_2^{new,clipped} - \lambda_2^{old})K_{22} + b^{old}$$

Lorsque  $b_1$  et  $b_2$  sont valides, ils sont égaux, sinon  $b^{new} = \frac{b_1 + b_2}{2}$ .

L'erreur  $E_i$  est mise à zéro si le multiplicateur correspondant  $\lambda_i$  est optimisé. Sinon elle est mise à jour pour tous les VS :

$$E_k^{new} = E_k^{old} + y_1(\lambda_1^{new} - \lambda_1^{old})K_{1k} + y_2(\lambda_2^{new,clipped} - \lambda_2^{old})K_{2k} + b^{old} - b^{new}$$

Si besoin, pour les autres éléments de la base d'apprentissage, l'erreur est calculée directement à partir du  $\vec{\lambda}$  en cours.

## Annexe C

# Champs de Markov

Cette annexe introduit brièvement les Champs de Markov. Pour avoir une présentation complète sur les champs de Markov, le lecteur intéressé peut se reporter à [Li, 1995; Winkler, 1995; Cocquerez and Philipp, 1995].

### C.1 Définitions

L'image  $x \in \Omega$  est composée par un ensemble fini  $S$  de pixels appelés *sites*  $s$  appartenant à  $Z^m$ .

**Voisinage :**

On définit un système de *voisinage*  $V_s$  :

$$V_s = \{t \in S\} \text{ tels que } \begin{cases} s \notin V_s \\ t \in V_s \Rightarrow s \in V_t \end{cases}$$

**Clique :**

On appelle *clique*, un sous-ensemble  $c$  de  $S$  dont les sites sont voisins deux à deux. Selon le voisinage utilisé, l'ensemble des cliques sera différent et fera intervenir plus ou moins de sites (cf. figure C.1). On note  $C$  l'ensemble des cliques relatives au voisinage  $V$ . L'ordre d'une clique correspond au nombre de pixels intervenant dans la clique.

### C.2 Propriétés d'un champ de Markov

**Champ de Markov :**

$X$  est un *champ de Markov* si et seulement si la valeur d'un pixel  $s$  ne dépend que de ses voisins  $V_s$ , ce qui s'exprime par :

$$P(X_s = x_s / X_r = x_r, r \in S - s) = P(X_s = x_s / X_r = x_r, r \in V_s)$$

**Champ de Gibbs :**

Le théorème de Hammersley-Clifford permet de caractériser une telle distribution :  $X$

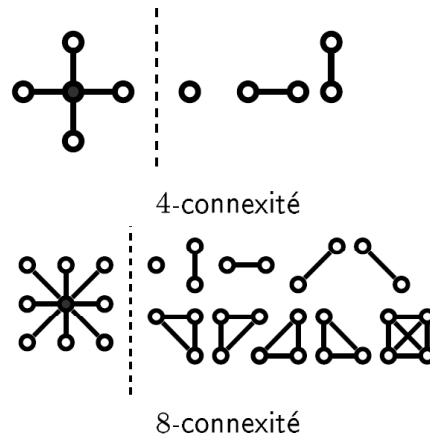


FIG. C.1: Systèmes de Voisinages et cliques associées en dimension 2

est un champ de Markov relativement à  $V$  et  $P(X = x) > 0$  si et seulement si  $X$  est un champ de Gibbs :

$$P(X = x) = \frac{1}{Z} \exp(-U(x))$$

avec  $U(x) = \sum_{c \in C} V(x)$  appelée énergie et  $Z = \sum_x \exp(-U(x))$  appelée fonction de partition.

## Annexe D

# Morphologie mathématique

Elle est appelée ainsi car elle analyse la forme des objets et s'appuie sur la théorie des ensembles topologiques. Le développement de la morphologie mathématique, dans les années 60, a été inspiré par les problèmes de traitement d'image, domaine qui constitue son principal champ d'application. Elle fournit en particulier des outils de filtrage, de segmentation et de quantification [Serra, 1983; Soille, 1999].

### D.1 Opérations de base de la morphologie mathématique

Dans cette partie, on considère des éléments de  $E = Z^2$ .

La morphologie mathématique repose sur la comparaison de l'image avec un élément structurant. Un *élément structurant* est un sous-ensemble de  $E$  dont on se donne la forme (disque, carré, losange, segment...) et la taille, et dont on particularise un point que l'on appelle le centre ou l'origine.

Soit  $B$  un sous-ensemble de  $E$  appelé élément structurant et  $x$  un élément de  $E$ . On note  $B_x$  l'ensemble  $B$  centré sur  $x$  :

$$B_x = \{b + x \text{ tel que } b \in B\}$$

#### L'érosion :

L'érosion de l'ensemble  $X \subset E$  par l'élément structurant  $B$  est définie par l'ensemble :

$$E_B(X) = \{x \text{ tel que } B_x \subset X\}$$

L'élément  $B$  est déplacé pour occuper successivement toute l'image. Pour chaque pixel, on se pose la question :  $B_x$  est-il complètement inclus dans  $X$ ? Si oui, le pixel fait partie de l'ensemble érodé.

C'est une transformation anti-extensive :  $E_B(X) \subset X$ .

- Les objets de taille inférieure à l'élément structurant disparaissent,
- Les contours des autres sont "rognés" d'une partie dont la taille correspond à la moitié de celle de l'élément structurant,
- les trous dans les objets grossissent,

- les objets fins sont coupés en plusieurs parties.

Le résultat obtenu après une érosion avec un élément structurant de taille  $n$  est équivalent à celui obtenu après  $n$  érosions de taille 1.

### La dilatation :

C'est l'opération duale de l'érosion :

$$D_B(X) = \{x \text{ tel que } B_x \cap X \neq \emptyset\} = \cup_{x \in X} B_x$$

L'élément  $B$  est déplacé pour occuper successivement toute l'image. Pour chaque pixel, on se pose la question :  $B_x$  intersecte-t-il  $X$  ? Si oui, le pixel fait partie de l'ensemble dilaté.

C'est donc une transformation extensive :  $X \subset D_B(X)$ .

- Les objets grossissent d'une partie dont la taille correspond à la moitié de celle de l'élément structurant,
- les trous dans les objets deviennent plus petits ou disparaissent,
- les objets situés à une distance inférieure à la taille de l'élément structurant fusionnent.

### L'ouverture et la fermeture :

La dilatation ne correspond pas à l'opération inverse de l'érosion. De même, l'érosion ne correspond pas à l'opération inverse de la dilatation.

La composition d'une érosion suivie d'une dilatation avec le symétrique de l'élément structurant utilisé lors de l'érosion s'appelle une *ouverture* :

$$O_B(X) = D_B(E_B(X))$$

La composition d'une dilatation suivie d'une érosion avec le symétrique de l'élément structurant utilisé lors de la dilatation s'appelle une *fermeture* :

$$F_B(X) = E_B(D_B(X))$$

L'ouverture élimine toutes les parties des objets qui ne peuvent pas contenir l'élément structurant. Elle est anti-extensive. La fermeture, opération duale de l'ouverture, comble les zones de taille inférieure à celle de l'élément structurant. Elle est extensive.

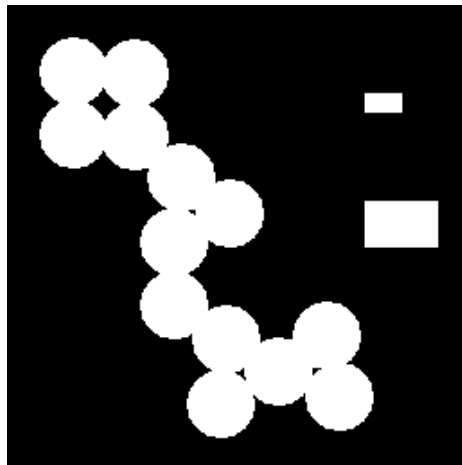
Ces deux opérations sont idempotentes :

$$O_B(O_B(X)) = O_B(X) \text{ et } F_B(F_B(X)) = F_B(X).$$

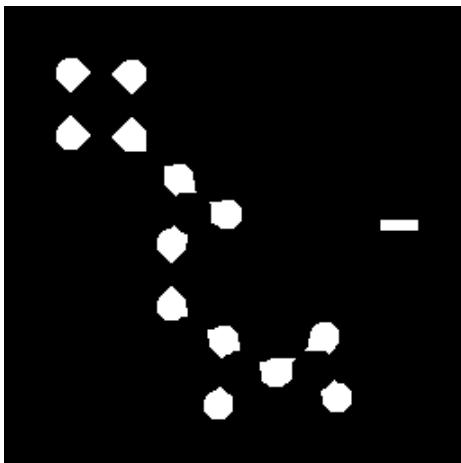
La figure D.1 représente une image binaire de taille 256\*256 pixels et les 4 opérations de morphologie mathématiques suivantes : érosion, ouverture, dilatation et fermeture avec le même élément structurant (disque de rayon 10 pixels).

L'ouverture et la fermeture morphologiques constituent des outils très intéressants de filtrage en traitement d'image. Cependant, comme le montre la figure D.1, ils peuvent modifier le contour des objets.

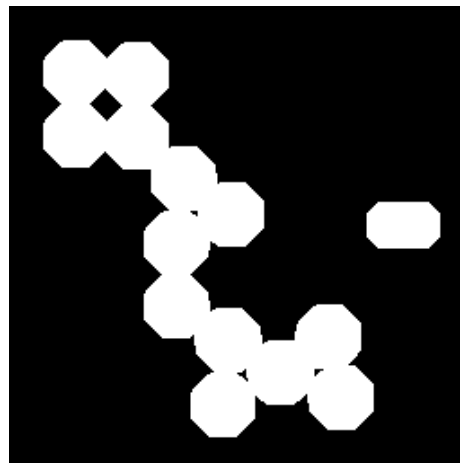
Des opérateurs par reconstruction géodésique ou des filtres d'aires permettent de résoudre cet inconvénient.



(a) Image originale



(b) Erosion



(c) Ouverture



(d) Dilatation



(e) Fermeture

FIG. D.1: Opérations de base de morphologie mathématique.



## D.2 Ligne de partage des eaux

L'algorithme de *Ligne de Partage des Eaux* (LPE) [Vincent and Soille, 1991] est l'algorithme morphologique le plus connu en segmentation. Il considère l'amplitude des données comme un relief topographique et fournit une partition en régions (appelées bassins versants) définies par les minima locaux et dont les frontières (appelées lignes de partage des eaux) correspondent aux crêtes de la surface (cf figure D.2).

Afin d'obtenir les frontières des régions, il faut imaginer l'immersion du relief dans l'eau qui ne peut pénétrer que par les minima des vallées. La ligne de partage des eaux est représentée par l'ensemble des points où deux lacs disjoints se rejoignent au cours de l'immersion (cf. figure D.3).

L'algorithme de calcul de la LPE proposé par [Vincent and Soille, 1991] a une complexité linéaire en fonction du nombre de pixels de l'image, ce qui le place parmi les méthodes de segmentation les plus rapides.

L'efficacité de cette méthode en tant qu'outil de segmentation dépend essentiellement du nombre de minima. Sans pré-traitement, les données seront le plus souvent sur-segmentées.

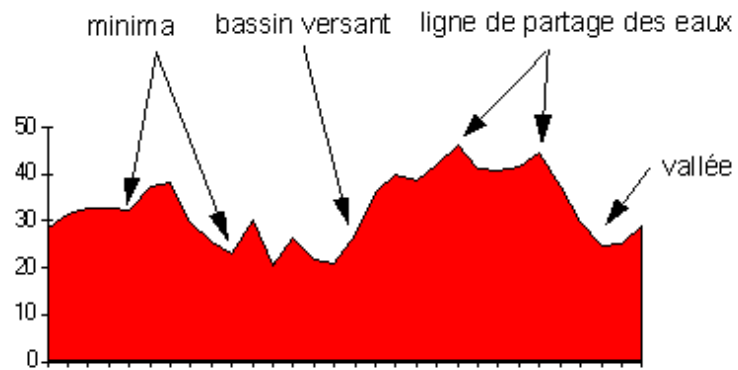


FIG. D.2: Ligne de Partage des Eaux

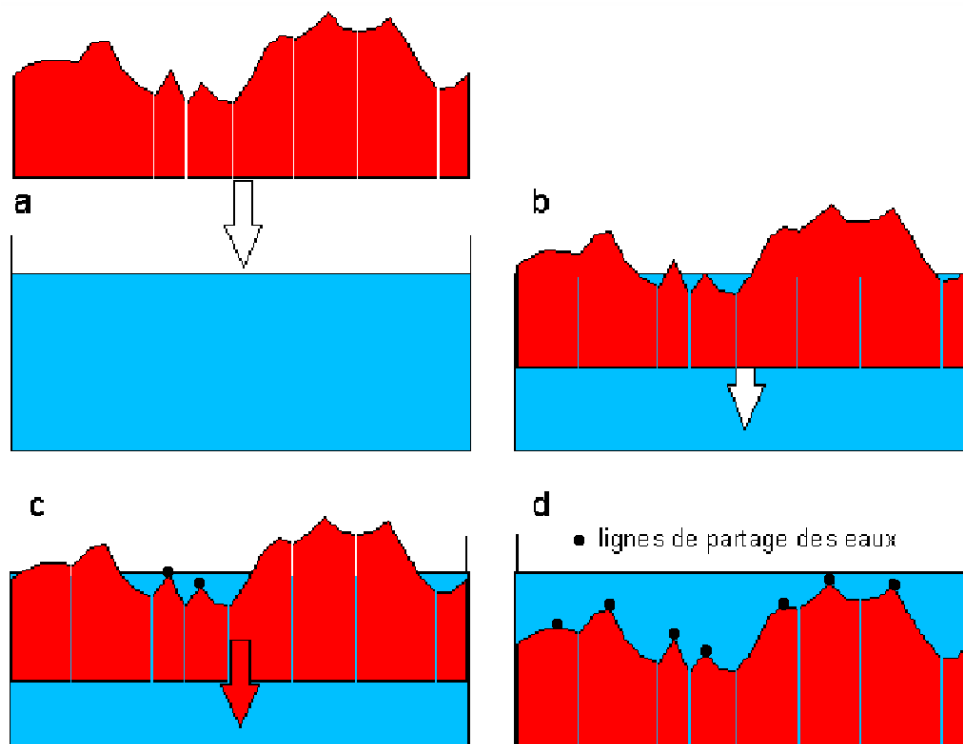


FIG. D.3: Algorithme de Ligne de Partage des Eaux



## Annexe E

# Algorithme "Analyse en Composantes Principales"

L'algorithme d'*Analyse en Composantes Principales* (ACP) est une méthode classique utilisée afin d'étudier les directions qui représentent le mieux les données appartenant à  $R^m$ . C'est une transformation linéaire orthogonale qui transporte les données vers un nouvel espace, de même dimension, défini par les directions des maxima de la variance.

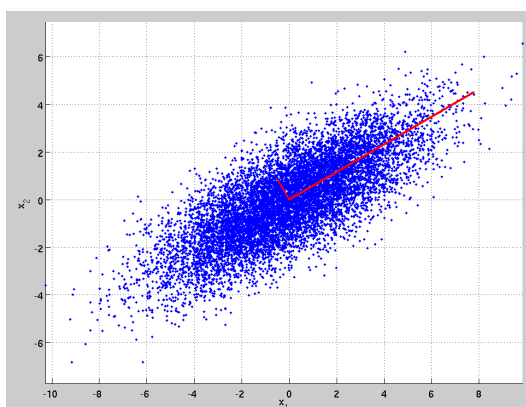


FIG. E.1: Analyse en Composantes Principales d'un nuage de points de dimension 2. Les vecteurs en rouge donnent les directions principales du nuage.

### E.1 Description de la méthode

On recherche donc les Composantes Principales  $C_1, C_2, \dots, C_m$  telles que les  $C_i$  soient des combinaisons linéaires des vecteurs de la base canonique de  $R^m$ , deux à deux non corrélées, de variance maximale et d'importance décroissante.

La variance est donnée par :

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - g)^T (x_i - g)$$

L'opérateur de projection orthogonale  $\pi$ , sur la droite de vecteur directeur  $v$  de norme unitaire s'écrit :

$$\pi = vv^T$$

La variance des données projetées vaut :

$$\begin{aligned} \sigma_v^2 &= \frac{1}{n} \sum_{i=1}^n [\pi(x_i - g)]^T [\pi(x_i - g)] \\ &= \frac{1}{n} \sum_{i=1}^n [vv^T(x_i - g)]^T [vv^T(x_i - g)] \\ &= \frac{1}{n} \sum_{i=1}^n (x_i - g)^T (vv^T vv^T) (x_i - g) \\ &= \frac{1}{n} \sum_{i=1}^n (x_i - g)^T vv^T (x_i - g) \\ &= \frac{1}{n} \sum_{i=1}^n v^T (x_i - g) (x_i - g)^T v \\ &= \frac{1}{n} v^T \left[ \sum_{i=1}^n (x_i - g) (x_i - g)^T \right] v \\ &= v^T \Sigma v \end{aligned}$$

avec  $\Sigma$  matrice de co-variance, définie positive.

Le but est de maximiser cette variance :

$$C_1 = \arg \max_v \sigma_v^2 = \arg \max_v (v^T \Sigma v) \text{ tel que } v^T v = 1$$

Il s'agit d'un problème d'optimisation sous contraintes. Le problème dual s'écrit :

$$L(v, \lambda) = v^T \Sigma v + \lambda(1 - v^T v)$$

En maximisant le lagrangien  $L$  selon  $v$ , il vient :

$$\frac{dL}{dv} = \Sigma v - \lambda v = 0$$

Le problème est donc équivalent à rechercher les valeurs propres de  $\Sigma$ , qui sont réelles positives puisque la matrice  $\sigma$  est définie positive. Les vecteurs propres sont orthonormés. La variance vaut alors :

$$\sigma_v^2 = v^T \Sigma v = v^T \lambda v = \lambda$$

Le vecteur  $C_1$  est donc obtenu en choisissant le vecteur propre de  $\Sigma$  ayant la valeur propre  $\lambda$  la plus élevée.

La seconde direction de variance maximale est donnée par :

$$C_2 = \arg \max_v (v^T \Sigma v) \text{ tel que } v^T v = 1 \text{ et } v^T C_1 = 0$$

Comme les vecteurs propres sont orthonormés, il suffit de choisir le deuxième vecteur propre de  $\Sigma$  de valeur propre maximale. Par itération, nous obtenons tous les vecteurs  $C_i$  qui forment la nouvelle base de  $R^m$ .

La somme des valeurs propres est égale à la variance totale :

$$tr(\Sigma) = \sigma^2 = \sum_{i=1}^n \lambda_i$$

Chaque valeur propre de  $\Sigma$  donne la proportion de la variance selon la direction engendrée par le vecteur propre correspondant.

L'algorithme ACP est donc le suivant :

---

Algorithme ACP :

- Calcul du centre de gravité :  $g = \sum_{i=1}^N x_i$
  - Calcul de la matrice de covariance :  $\Sigma = \frac{1}{N} \sum_{i=1}^N (x_i - g) \cdot (x_i - g)^T$
  - Recherche des valeurs propres et vecteurs propres de  $\Sigma$  :  $\lambda_i$  et  $v_i$
  - Projection des données sur les nouveaux axes :  $x_{ACP} = [v_1 v_2 \dots v_m]^T \cdot x$
- 

## E.2 Application aux images satellitaires

Dans cette partie, nous appliquons l'algorithme de l'ACP sur des images satellitaires SPOT 5 de résolution 10m, composées de 4 bandes spectrales (Vert, Rouge, PIR et MIR).

### Image "Nord de la Corse" :

Les figures E.2 et E.3 donnent respectivement les 4 bandes spectrales de l'image satellitaire et les 4 composantes principales de l'ACP pour l'image "Nord de la Corse". Les images correspondant aux bandes Vert et Rouge sont très ressemblantes. Après ACP, la bande 4 comporte visuellement très peu d'information.

Les figures E.4 et E.5 donnent respectivement les histogrammes des 4 bandes de l'image satellitaire et ceux des 4 composantes principales de l'ACP pour l'image "Nord de la Corse". Les bandes Vert et Rouge comportent moins d'information que les deux autres bandes. Après ACP, les bandes sont bien rangées selon la quantité d'information qu'elles possèdent.

Le tableau E.1 donne la proportion de la variance pour chaque bande de l'image satellitaire originale, puis celle de l'image transformée après ACP (égale aux valeurs propres de  $\Sigma$ ), et la décomposition de chaque composante principale selon les bandes Vert, Rouge, PIR et MIR de l'image satellitaire (vecteurs propres de la matrice de co-variance  $\Sigma$ ).

La 4ème composante principale est principalement une combinaison linéaire des bandes Vert et Rouge et sa variance est très proche de zéro. Les bandes Vert et Rouge de l'image satellitaire se ressemblent donc très fortement.

	Bande Vert	Bande Rouge	Bande PIR	Bande MIR
% Variance	0.1878	0.1410	0.0476	0.6236
	Bande ACP 1	Bande ACP 2	Bande ACP 3	Bande ACP 4
% Variance	0.7768	0.1880	0.0339	0.0013
Vecteur propre	0.3874	0.1966	0.7027	0.5635
	0.2166	0.0803	0.5184	-0.8234
	0.0957	-0.9772	0.1841	0.0457
	0.8910	-0.0001	-0.4513	-0.0498

TAB. E.1: ACP de l'image "Nord de la Corse"

**Image "Ouest de la Corse" :**

De même, l'ACP est appliquée sur l'image "Ouest de la Corse". Les figures E.8, E.9 et tableau E.2 donnent les mêmes informations que précédemment.

Sur les figures E.8 et E.9, les histogrammes des bandes PIR, MIR de l'image satellitaire, ainsi que celui de la première composante principale possèdent chacun un pic près de zéro. Ces pics correspondent au grand nombre de pixels composés par la mer dans les bandes spectrales correspondantes des figures E.6 et E.7.

	Bande Vert	Bande Rouge	Bande PIR	Bande MIR
% Variance	0.1826	0.2381	0.0886	0.4908
	Bande ACP 1	Bande ACP 2	Bande ACP 3	Bande ACP 4
% Variance	0.9216	0.0562	0.0209	0.0014
Vecteur propre	0.4874	0.5425	-0.3149	-0.6074
	0.2876	0.3760	-0.4104	0.7794
	0.3975	-0.7436	-0.5332	-0.0688
	0.7223	-0.1066	0.6694	0.1373

TAB. E.2: ACP de l'image "Corse Ouest"

**E.3 Conclusion**

L'ACP permet de trouver les directions principales des données. La quatrième bande principale, qui est en grande partie une combinaison linéaire des bandes Vert et Rouge, n'apportent quasiment pas d'information. Si l'on dispose des quatre bandes spectrales (Vert, Rouge, PIR et MIR), il n'est donc pas très utile de l'utiliser.

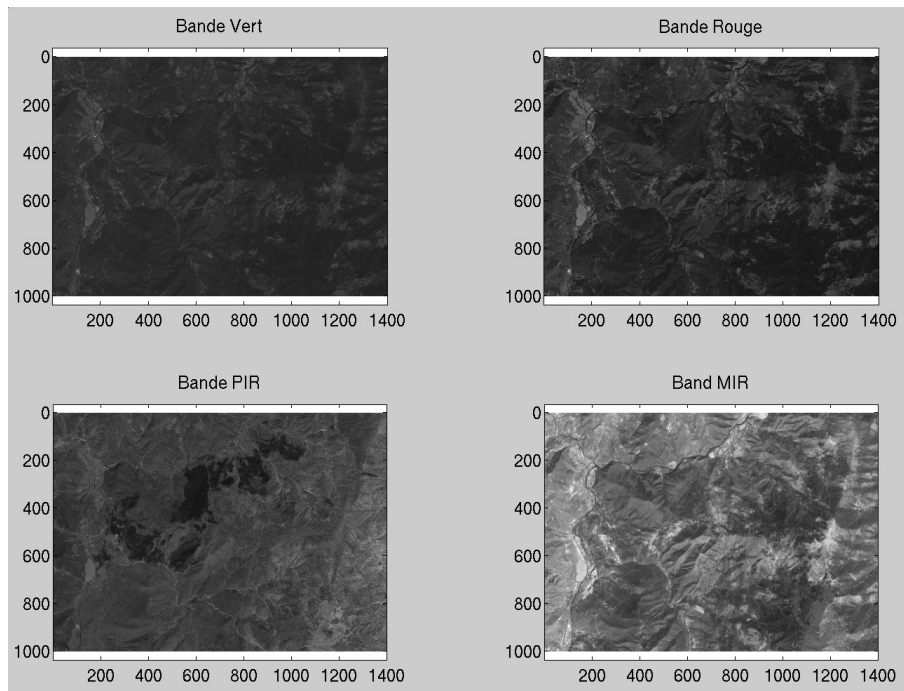


FIG. E.2: Bandes de "Nord de la Corse" © CNES 2003, Distribution SPOT Image

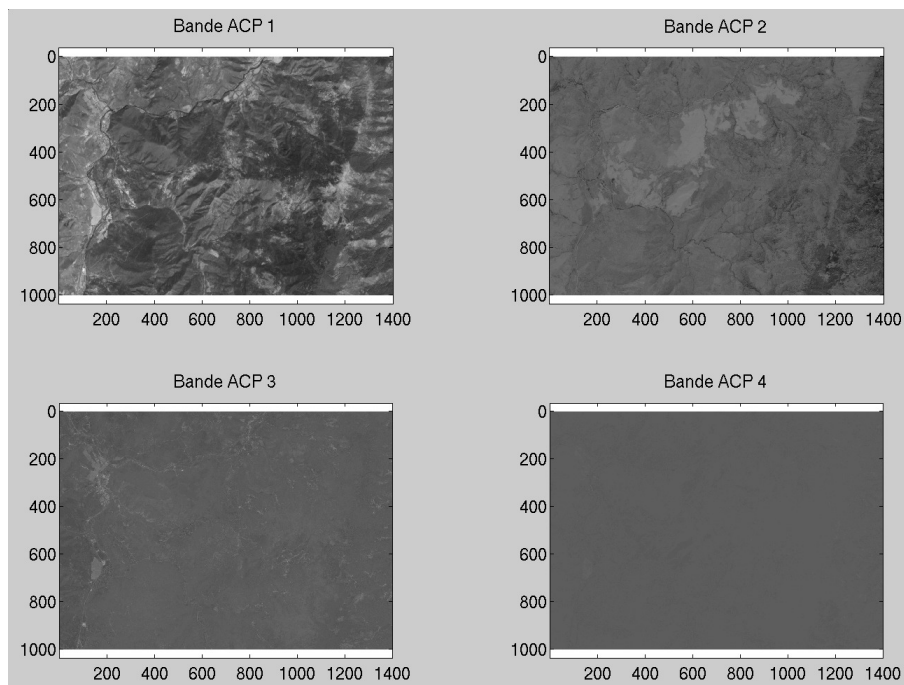


FIG. E.3: Composantes Principales de "Nord de la Corse"



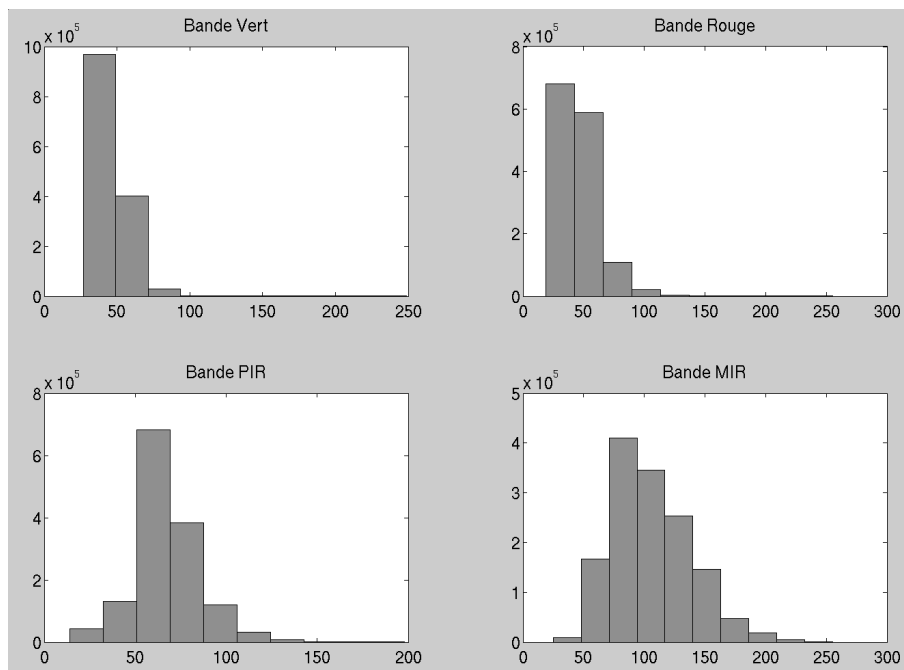


FIG. E.4: Histogrammes des 4 bandes de l'image "Nord de la Corse"

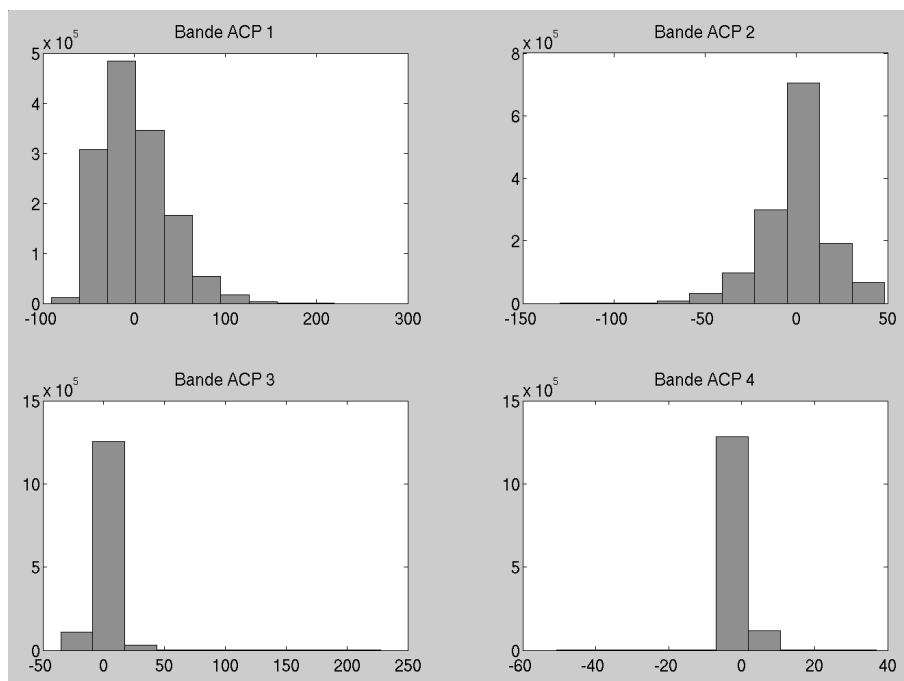


FIG. E.5: Histogrammes des 4 Composantes Principales de "Nord de la Corse"

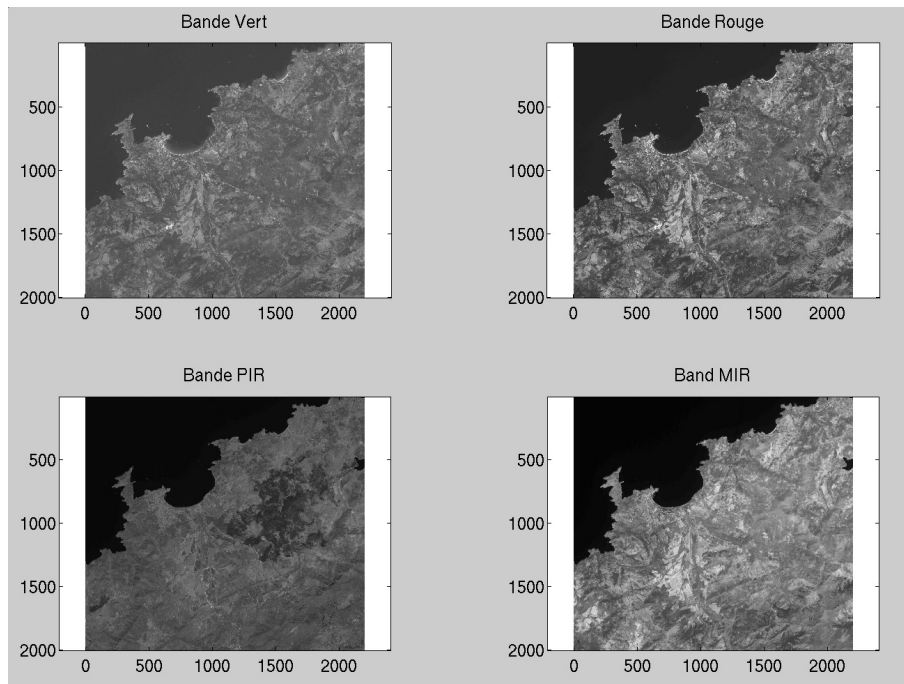


FIG. E.6: Bandes de l'image "Ouest de la Corse" © CNES 2005, Distribution SPOT Image

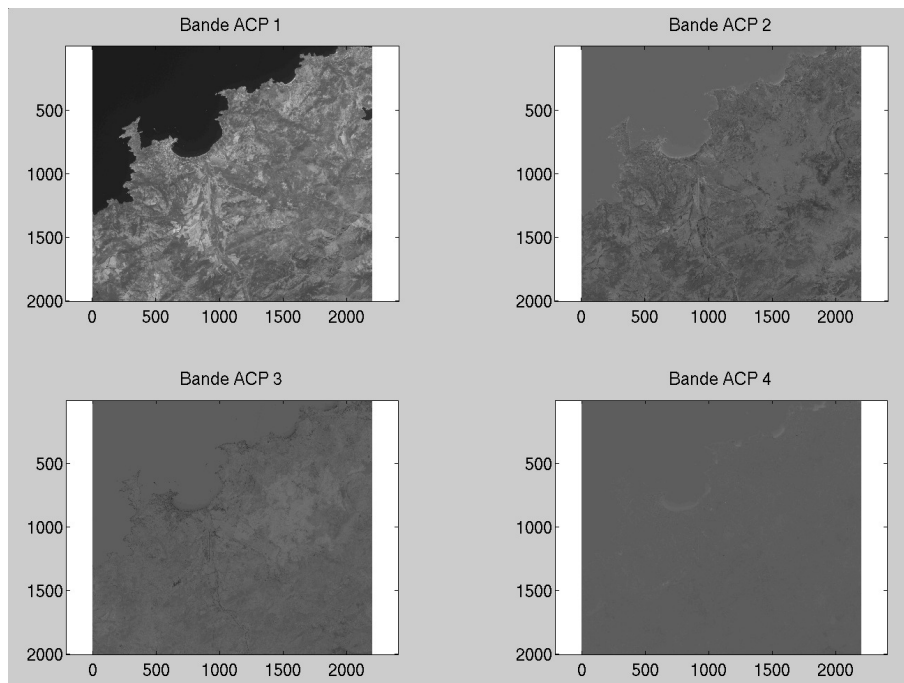


FIG. E.7: Composantes Principales de l'image "Ouest de la Corse"

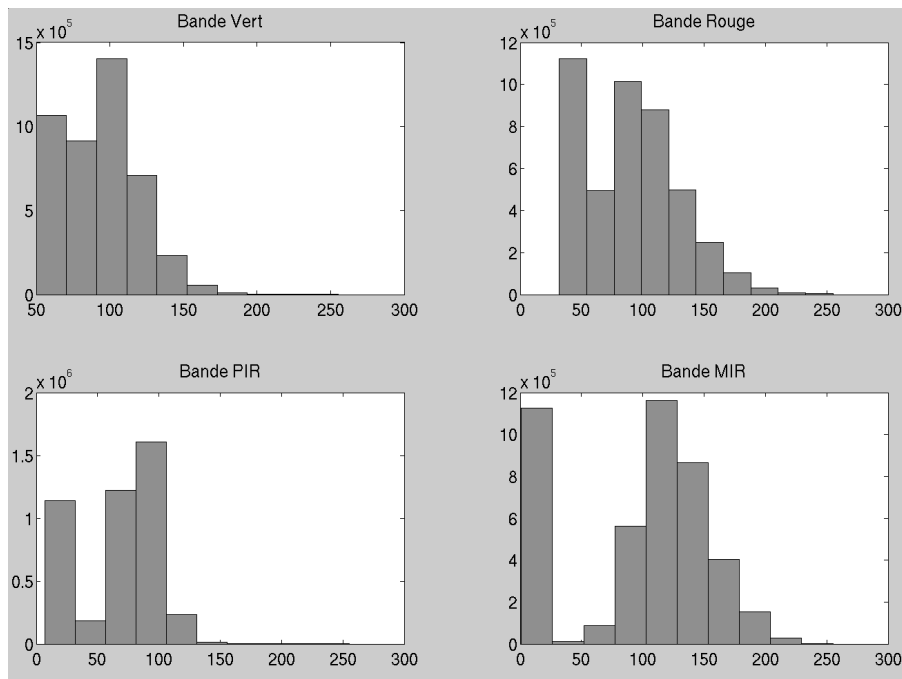


FIG. E.8: Histogrammes des 4 bandes de l'image "Ouest de la Corse"

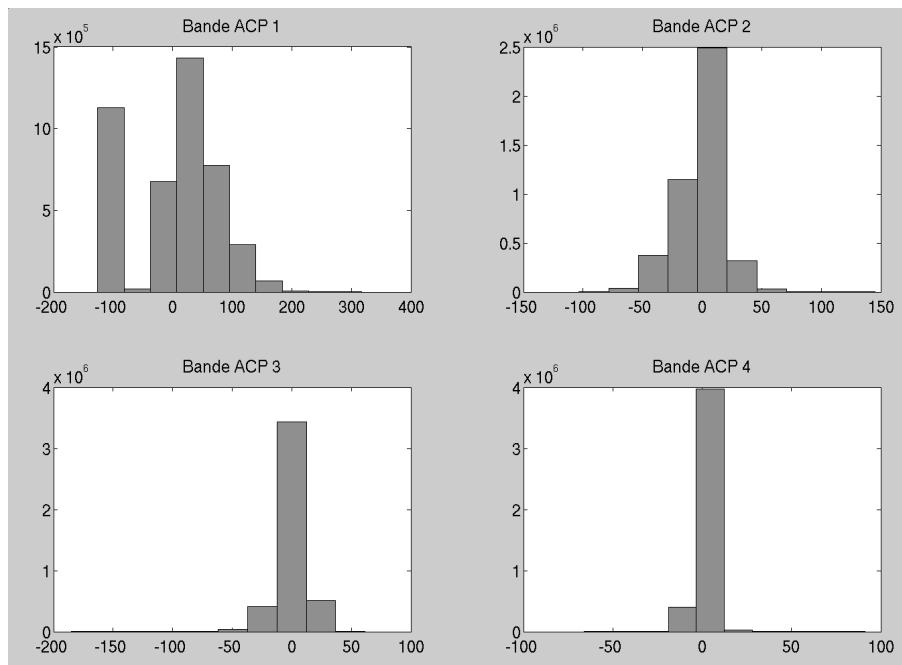


FIG. E.9: Histogrammes des 4 Composantes Principales de l'image "Ouest de la Corse"

## Annexe F

# Comment évaluer la précision d'un algorithme de classification binaire ?

Afin de quantifier les performances d'un algorithme de classification binaire (2 classes), les résultats de l'algorithme sont comparés à des cartes de référence.

Dans ce manuscrit de thèse, la dénomination "carte de référence" désigne une cartographie réalisée par des experts. Cette cartographie peut être obtenue, soit sur le terrain, soit par interprétation visuelle et/ou automatique d'une image de télédétection.

### F.1 Partition de l'image

Les pixels de l'image classifiée sont rangés en 4 catégories :

- **Vrais Positifs** ( $VP$ ) : pixels correctement classés comme positifs, i.e. pixels classés comme brûlés par l'algorithme et classés comme brûlés dans la carte de référence.
- **Vrais Négatifs** ( $VN$ ) : pixels correctement classés comme négatifs, i.e. pixels classés comme non brûlés par l'algorithme et classés comme non brûlés dans la carte de référence.
- **Faux Positifs** ( $FP$ ) : pixels non correctement classés comme positifs, i.e. pixels classés comme brûlés par l'algorithme et classés comme non brûlés dans la carte de référence.
- **Faux Négatifs** ( $FN$ ) : pixels non correctement classés comme négatifs, i.e. pixels classés comme non brûlés par l'algorithme et classés comme brûlés dans la carte de référence.

Pour simplifier les notations, dans la suite, nous noterons  $VP$ ,  $VN$ ,  $FP$  et  $FN$  le cardinal (nombre d'éléments) de chaque catégorie.

Les 4 catégories forment une partition de l'image :  $VP + VN + FP + FN$  représente le nombre total de pixels dans l'image.

## F.2 Taux de classification

**Précision Totale** ( $PT$ ) : taux de pixels bien classés dans l'image

$$PT = 100 * \frac{VP + VN}{VP + VN + FP + FN}$$

Une Précision Totale élevée signifie une bonne précision de l'algorithme.

**Taux de Faux Positifs** ( $TFP$ ) : taux de pixels mal classés positifs parmi les pixels négatifs de la carte de référence

$$TFP = 100 * \frac{FP}{TN + FP}$$

Un Taux de Faux Positifs faible signifie une bonne classification des pixels négatifs de la carte de référence.

La **spécificité** ( $TVN$ ) représente la proportion de pixels classés négatifs parmi les pixels négatifs de la carte de référence (précision sur la classe négative). Elle peut être vue comme la probabilité qu'un pixel soit classé négatif sachant que le pixel est négatif. Une spécificité élevée signifie une faible fausse détection de négatifs.  $TVN = 100 * \frac{TN}{TN+FP} = 100 - TFP$

**Taux de Faux Négatifs** ( $TFN$ ) : taux de pixels mal classés négatifs parmi les pixels positifs de la carte de référence

$$TFN = 100 * \frac{FN}{TP + FN}$$

Un Taux de Faux Négatifs faible indique une bonne classification des pixels positifs de la carte de référence.

La **sensibilité** ( $TVP$ ) représente la proportion de pixels classés positifs parmi les pixels positifs de la carte de référence (précision sur la classe positive). Elle peut être vue comme la probabilité qu'un pixel soit classé positif sachant que le pixel est positif. Une sensibilité élevée signifie une faible fausse détection de positifs.  $TVP = 100 * \frac{TP}{TP+FN} = 100 - TFN$

La mesure de la sensibilité (ou du taux de Faux Positifs) s'accompagne toujours de celle de la spécificité (ou du taux de Faux Négatifs). En effet, pris séparément, les deux taux ne quantifient pas entièrement l'algorithme de classification. Par exemple, un algorithme de classification avec une spécificité de 95% a peu de valeur si sa sensibilité est de 5% : un pixel est classé négatif pour 95% des pixels réellement négatifs, mais un pixel est classé positif pour 5% des pixels réellement positifs.

**Précision Moyenne** ( $PM$ ) : moyenne des taux de précision sur chaque classe

$$\begin{aligned} PM &= 100 * \frac{1}{2} * (TVP + TVN) \\ &= 100 - 100 * \frac{1}{2} * (TFP + TFN) \end{aligned}$$

La Précision Moyenne est égale à la moyenne de la sensibilité et de la spécificité. Elle peut être obtenue à partir des Taux de Faux Positifs et Taux de Faux Négatifs.

### F.3 Conclusion

La Précision Totale fournit la meilleure estimation d'un algorithme de classification. Cependant, si une des deux classes est minoritaire (classe brûlée généralement minoritaire dans les images satellitaires), la Précision Totale ne permet pas de mesurer la bonne ou mauvaise détection de cette classe. En effet, cette classe est peu prise en compte dans le calcul de la Précision Totale, mais elle a plus d'influence dans le calcul de la Précision Moyenne. Une grande différence entre ces deux précisions signifie qu'une des deux classes a été mal classée.

La comparaison d'algorithmes de classification mesure les taux de concordance entre les résultats de chaque algorithme et une carte de référence. Si la carte de référence présente des erreurs (ce qui est malheureusement très souvent le cas, en particulier dans cette thèse), la comparaison est biaisée.



## Annexe G

# Cas particulier des OC-SVM

Dans un premier temps, nous cherchons à résoudre le problème suivant : étant donné une fonction  $\Phi : R^m \rightarrow F$ , des éléments  $\{\Phi(\vec{x}_i)\}_{i \in \langle 1, l \rangle}$  de  $F$ , quelle est la plus petite hypersphère qui englobe la plupart de ces éléments ?

Dans un second temps, nous verrons pourquoi ce problème est un cas particulier des OC-SVM [Schölkopf et al., 2001], dont le but est de séparer les éléments de l'origine à l'aide d'un hyperplan.

### G.1 Calcul de l'hypersphère englobant la plupart des données et de taille minimale

Le problème est donc le suivant :

$$\min_{(R, \vec{c}, \xi)} \left[ R^2 + \frac{1}{\nu \cdot l} \sum_{i=1}^l \xi_i \right]$$

$$\text{tels que : } \|\Phi(\vec{x}_i) - \vec{c}\| \leq R^2 + \xi_i \quad , \quad \xi_i \geq 0, \forall i \in \langle 1, l \rangle$$

où :

- $\vec{c}$  et  $R$  représentent respectivement le centre et le rayon de l'hypersphère,
- $\xi_i$  sont les variables qui quantifient l'erreur donnée par la distance entre le vecteur  $\vec{x}_i$  et l'hypersphère (cf. partie 2.2.4),
- $\frac{1}{\nu \cdot l}$  le paramètre de régularisation.

Ce problème d'optimisation peut être résolu en utilisant les multiplicateurs de Lagrange :

$$L(R, \vec{c}, \xi, \lambda, \mu) = R^2 + \frac{1}{\nu \cdot l} \sum_{i=1}^l \xi_i + \sum_{i=1}^l \lambda_i \cdot [\|\Phi(\vec{x}_i) - \vec{c}\| - R^2 - \xi_i] - \sum_{i=1}^l \mu_i \cdot \xi_i \quad (\text{G.1})$$



ce qui donne :

$$\frac{\partial L}{\partial R} = 2R \left(1 - \sum_{i=1}^l \lambda_i\right) = 0 \quad \Rightarrow \quad \sum_{i=1}^l \lambda_i = 1 \quad (\text{G.2})$$

$$\frac{\partial L}{\partial \vec{c}} = -\sum_{i=1}^l 2\lambda_i \cdot (\Phi(\vec{x}_i) - \vec{c}) = 0 \quad \Rightarrow \quad \vec{c} = \sum_{i=1}^l 2\lambda_i \cdot \Phi(\vec{x}_i) \quad (\text{G.3})$$

$$\frac{\partial L}{\partial \xi_i} = \frac{1}{\nu \cdot l} - \lambda_i - \mu_i = 0 \quad \Rightarrow \quad 0 \leq \lambda_i \leq \frac{1}{\nu \cdot l} \quad (\text{G.4})$$

Les première et deuxième équations représentent les contraintes alors que la troisième implique que le centre  $\vec{c}$  du cercle est une combinaison linéaire des  $\Phi(\vec{x}_i)$ . En remplaçant ces équations dans le Lagrangien G.1, on obtient :

$$\max_{\lambda} W(\lambda) = -\sum_{i=1}^l \sum_{j=1}^l \lambda_i \lambda_j K(\vec{x}_i, \vec{x}_j) + \sum_{i=1}^l \lambda_i K(\vec{x}_i, \vec{x}_i) \quad (\text{G.5})$$

$$\text{tels que : } \sum_{i=1}^l \lambda_i = 1, \quad 0 \leq \lambda_i \leq \frac{1}{\nu \cdot l}, \quad \forall i \in \langle 1, l \rangle \quad (\text{G.6})$$

avec  $K(\vec{x}_i, \vec{x}_j) = \Phi(\vec{x}_i) \cdot \Phi(\vec{x}_j)$ .

La fonction de décision est donc la suivante :

$$f(\vec{x}) = \text{sign} \left( R^2 - \sum_{i=1}^l \sum_{j=1}^l \lambda_i \lambda_j K(\vec{x}_i, \vec{x}_j) + 2 \sum_{i=1}^l \lambda_i K(\vec{x}_i, \vec{x}) - K(\vec{x}, \vec{x}) \right)$$

## G.2 Comparaison avec les OC-SVM

On rappelle le problème d'optimisation des OC-SVM :

$$\max_{\lambda} W(\lambda) = -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \lambda_i \lambda_j K(\vec{x}_i, \vec{x}_j)$$

$$\text{tels que : } \sum_{i=1}^l \lambda_i = 1, \quad 0 \leq \lambda_i \leq \frac{1}{\nu \cdot l}, \quad \forall i \in \langle 1, l \rangle$$

Pour des noyaux  $K(\vec{x}, \vec{y})$  qui ne dépendent que de  $\vec{x} - \vec{y}$  (comme le noyau gaussien),  $K(\vec{x}, \vec{x})$  est constant. La contrainte égalité G.6 implique que le deuxième terme du problème dual G.5 ne dépend plus de  $\lambda$ . Le problème d'optimisation G.5 sous contraintes G.6 est donc équivalent au problème des OC-SVM.

De plus, les deux fonctions de décision sont aussi équivalentes. La figure G.1 permet de le comprendre géométriquement en dimension 2.

Les points rouges représentent les données de la base d'apprentissage. Comme  $K(\vec{x}, \vec{x}) = \Phi(\vec{x}) \cdot \Phi(\vec{x}) = \|\Phi(\vec{x})\|^2$  est constant, les éléments  $\Phi(\vec{x})$  sont tous situés sur la même

sphère (en pointillé noir) centrée sur l'origine, dans l'espace des images  $F$ . L'hyperplan séparateur optimal des OC-SVM, de distance maximale par rapport à l'origine, est donné par la droite verte. L'hypersphère, de taille minimale qui englobe tous les points de la base d'apprentissage, est donnée par le cercle bleu. Les points jaunes représentent de nouvelles données à classer. L'hypersphère et l'hyperplan classifient ces nouvelles données de la même façon : les points à l'intérieur de l'hypersphère sont tous du même côté de l'hyperplan, les points à l'extérieur de l'hypersphère sont tous de l'autre côté de l'hyperplan.

Chercher l'hypersphère de taille minimale qui englobe tous les points de la base d'apprentissage est donc équivalent à chercher l'hyperplan séparateur optimal des OC-SVM.

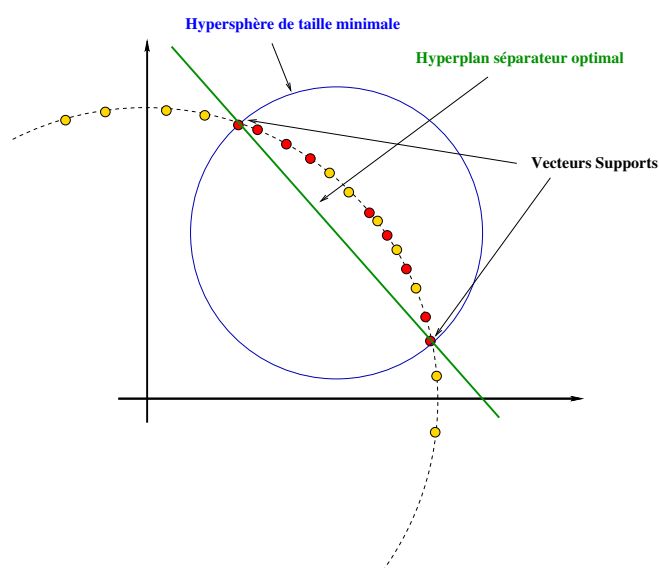


FIG. G.1: Cas particulier des "One-Class SVM"



# Glossaire

ACP : Analyse en Composantes Principales  
AVHRR : Advanced Very High Resolution Radiometer

BAI : Burnt Area Index

CM : Champs de Markov  
CNES : Centre National d'Etudes Spatiales

EADS : European Aeronautic Defence and Space company  
ERM : Empirical Risk Minimization  
ESA : European Space Agency

FP : Faux Positifs  
FN : Faux Négatifs

GEMI : Global Environmental Monitoring Index  
GMES : Global Monitoring for Environment and Security  
GPS : Global Positioning System

HSO : Hyperplan Séparateur Optimal

ICM : Iterative Conditional Mode  
IR : Infra-Rouge  
IRS - WiFS : Indian Remote Sensing - Wide Field Sensor

KMOD : Kernel with Moderate Decrease  
KKT : Karush-Kuhn-Tucker  
K-PPV : K Plus Proches Voisins

Landsat TM : Landsat Thematic Mapper  
LPE : Ligne de Partage des Eaux

MIR : Moyen Infra-Rouge  
MODIS : Moderate-resolution Imaging Spectroradiometer  
MCV : Maximum Composite Value

MM : Morphologie Mathématique  
MV : Maximum de Vraisemblance

NBR : Normalized Burn Ratio  
NDII : Normalized Difference Infrared Index  
NDVI : Normalized Difference Vegetation Index  
NOAA : National Oceanic and Atmospheric Administration

OC-SVM : One-Class SVM  
ONF : Office National des Forêts

POQ : Problème d'Optimisation Quadratique  
PIR : Proche Infra-Rouge  
PM : Précision Moyenne  
PT : Précision Totale

RBF : Radial Basis Function  
ROC : Receiver Operating Characteristic

SAC-C/MMRS : Argentinian Satellite for Scientific Application - Multispectral Medium Resolution Scanner  
SAVI : Soil Adjusted Vegetation Index  
SAM : Spectral Angle Mapper  
SDIS2B : Service Départemental d'Incendie et de Secours de Haute Corse  
SDIS06 : Service Départemental d'Incendie et de Secours des Alpes Maritimes  
SDIS83 : Service Départemental d'Incendie et de Secours du Var  
SERTIT : Service Régional de Traitement d'Image et de Télédétection  
SIG : Système d'Information Géographique  
SMA : Spectral Mixture Analysis  
SMO : Sequential Minimal Optimization  
SPOT : Satellite Pour l'Observation de la Terre  
SRM : Structural Risk Minimization  
SVM : Séparateurs à Vaste Marge

VC-dimension : dimension de Vapnik-Chervonenkis  
VS : Vecteur Support

# Bibliographie

- [Ayat et al., 2002]Ayat, N., Cheriet, M., and Suen, C. (2002). KMOD - A two parameter SVM kernel for pattern recognition. In *International Conference on Pattern Recognition (ICPR)*, Quebec City, Canada.
- [Barbosa et al., 2002]Barbosa, P. M., San-Miguel Ayanz, J., Martinez, B., and Schmuck, G. (2002). Burnt area mapping in Southern Europe using IRS-WiFS. In Viegas, editor, *Forest Fire Research & Wildland Fire Safety*, Rotterdam. Millpress.
- [Besag, 1986]Besag, J. (1986). On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, 48(3) :259–302.
- [Bishop, 2006]Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer-Verlag New York, Inc.
- [Brewer et al., 2005]Brewer, C. K., Winne, J. C., Redmond, R. L., Opitz, D. W., and Mangrich, M. V. (2005). Classifying and mapping wildfire severity : A comparison of methods. *Photogrammetric Engineering and Remote Sensing*, 71(11) :1311–1320.
- [Bruzzone et al., 2007]Bruzzone, L., Marconcini, M., and Persello, C. (2007). Fusion of spectral and spatial information by a novel SVM classification technique. In *IEEE International Geoscience And Remote Sensing Symposium (IGARSS)*, Barcelona, Spain.
- [Burges, 1998]Burges, C. J. C. (1998). A tutorial on Support Vector Machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2) :121–167.
- [Camps-Valls et al., 2006]Camps-Valls, G., Gomez-Chova, L., Munoz-Mari, J., Vila-Frances, J., and Calpe-Maravilla, J. (2006). Composite kernels for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 3(1) :93–97.
- [Canny, 1986]Canny, J. (1986). A computational approach to edge detection. *Pattern Analysis and Machine Intelligence*, 8(6) :679–698.
- [Chapelle et al., 2002]Chapelle, O., Vapnik, V. N., Bousquet, O., and Mukherjee, S. (2002). Choosing multiple parameters for Support Vector Machines. *Machine Learning*, 46(1–3) :131–159.
- [Chuvieco, 2003]Chuvieco, E. (2003). *Wildland Fire Danger Estimation and Mapping : The role of Remote Sensing data*, volume 4 of *Remote Sensing*. World Scientific.
- [Chuvieco et al., 2008]Chuvieco, E., Englefield, P., Trishchenko, A. P., and Luo, Y. (2008). Generation of long time series of burn area maps of the boreal forest from NOAA-AVHRR composite data. *Remote Sensing of Environment*, 112 :2381–2396.
- [Chuvieco et al., 2002]Chuvieco, E., Martin, M. P., and Palacios, A. (2002). Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. *International Journal of Remote Sensing*, 23(23) :5103–5110.

- [Chuvieco et al., 2005a]Chuvieco, E., Ventura, G., and Martin, M. P. (2005a). AVHRR multitemporal compositing techniques for burned land mapping. *International Journal of Remote Sensing*, 26(5) :1013–1018.
- [Chuvieco et al., 2005b]Chuvieco, E., Ventura, G., Martin, M. P., and Gomez, I. (2005b). Assessment of multitemporal compositing techniques of MODIS and AVHRR images for burned land mapping. *Remote Sensing of Environment*, 94 :450–462.
- [Cocquerez and Philipp, 1995]Cocquerez, J. P. and Philipp, S. (Editions 1995). *Analyse d'image : filtrage et segmentation*. Masson.
- [Colin et al., 2001]Colin, P.-Y., Jappiot, M., and Mariel, A. (2001). *Protection des forêts contre l'incendie. Fiches techniques pour les pays du bassin méditerranéen*. 36. Cahiers FAO Conservation.
- [Conese and Checcacci, 2006]Conese, C. and Checcacci, E. (2006). Burned forest mapping from satellite images. In *International Symposium on Environment Identities and Mediterranean Area (ISEIM)*, Corte-Ajaccio, France.
- [Deshayes, 2004]Deshayes, M. (2004). Feux de forêt : de la donnée spatiale à la protection civile. In *Colloque Espace et Education*, Paris, France.
- [Ezzine et al., 2005]Ezzine, H., Saidi, M., Casanova, J. L., Montes, A., and Ghribe, A. (2005). Conception et mise en place d'un système à base de données NOAA-AVHRR, pour la prévention et le suivi des incendies de forêts au Maroc. *Geo observateur*, (14).
- [Fang et al., 2005]Fang, Y., Pan, C., Liu, L., and Fang, L. (2005). Fast training of SVM via morphological clustering for color image segmentation. In *International Conference on Intelligent Computing*, pages 263–271.
- [Fauvel, 2007]Fauvel, M. (2007). Spectral and spatial methods for the classification of urban remote sensing data. Thèse de doctorat, Institut National Polytechnique de Grenoble.
- [Fauvel et al., 2007]Fauvel, M., Chanussot, J., and Benediktsson, J. A. (2007). Classification en imagerie hyperspectrale haute résolution : utilisation conjointe de l'information spatiale et spectrale. In *Colloque sur le Traitement du Signal et des Images - GRETSI*, Troyes, France.
- [Flannigan and Vonder Haar, 1986]Flannigan, M. D. and Vonder Haar, T. H. (1986). Forest fire monitoring using NOAA satellite AVHRR. *Canadian Journal of Forest Research*, 16 :975–982.
- [Fraser et al., 2002]Fraser, R. H., Fernandes, R., and Latifovic, R. (2002). Multi-temporal burned area mapping using logistic regression analysis and change metrics. In *IEEE International Geoscience And Remote Sensing Symposium (IGARSS)*, Toronto, Canada.
- [Garcia and Chuvieco, 2004]Garcia, M. and Chuvieco, E. (2004). Assessment of the potential of SAC-C/MMRS imagery for mapping burned areas in Spain. *Remote Sensing of Environment*, 92 :414–423.
- [Geman and Geman, 1984]Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6 :721–741.

- [Gonzalez-Alonzo et al., 2004]Gonzalez-Alonzo, F., Merino-de-Miguel, S., Garcia-Gigorro, S., Roldan-Zamarron, A., Cueva, J. M., and Arino, O. (2004). Mapping forest fire damage with Envisat. *ESA bulletin*, (120) :23–26.
- [Géraud et al., 2001]Géraud, T., Strub, P.-Y., and Darbon, J. (2001). Color image segmentation based on automatic morphological clustering. In *International Conference on Image Processing (ICIP)*, pages 70–73.
- [Haralick, 1979]Haralick, R. M. (1979). Statistical and structural approaches to texture. *IEEE*, 67 :786–804.
- [Henry, 2008]Henry, M. C. (2008). Comparison of single- and multi-date landsat data for mapping wildfire scars in Ocala national forest, Florida. *Photogrammetric Engineering and Remote Sensing*. In press.
- [Hsu et al., 2003]Hsu, C. W., Chang, C. C., and Lin, C. J. (2003). A practical guide to support vector classification. Technical report, Department of Computer Science and Information Engineering, National Taiwan University, Taipei.
- [Hudak and Brockett, 2004]Hudak, A. T. and Brockett, B. H. (2004). Mapping fire scars in a Southern African savannah using landsat imagery. *International Journal of Remote Sensing*, 25 :3231–3243.
- [Huete, 1988]Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3) :295–309.
- [Inventaire Forestier National, 2003]Inventaire Forestier National (2003). 2003, année marquée par les incendies. IF n°1.
- [Jain and Farrokhnia, 1991]Jain, A. K. and Farrokhnia, F. (1991). Unsupervised texture segmentation using Gabor filters. *Pattern Recognition*, 24(12) :1167–1186.
- [Kasischke and French, 1995]Kasischke, E. S. and French, N. H. F. (1995). Locating and estimating the areal extent of wildfires in Alaskan boreal forests using multiple-season AVHRR NDVI composite data. *Remote Sensing of Environment*, 51 :263–275.
- [Keerthi, 2002]Keerthi, S. S. (2002). Efficient tuning of SVM hyperparameters using radius/margin bound and iterative algorithms. *IEEE Trans. on Neural Networks*, 13 :1225–1229.
- [Keerthi et al., 2001]Keerthi, S. S., Shevade, S. K., Bhattacharyya, C., and Murthy, K. R. K. (2001). Improvements to Platt’s SMO algorithm for SVM classifier design. *Neural Computation*, 13(3) :637–649.
- [Key and Benson, 1999]Key, C. H. and Benson, N. C. (1999). Measuring and remote sensing of burn severity. In *Joint Fire Science Conference and Workshop*, Moscow, Russia.
- [Kim et al., 2002]Kim, K. I., Jung, K., Park, S. H., and Kim, H. J. (2002). Support Vector Machines for texture classification. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(11) :1542–1550.
- [Lasaponara and Lanorte, 2004]Lasaponara, R. and Lanorte, A. (2004). Spectral analysis of burned areas observed in the Italian Peninsula by using SPOT-Vegetation data. In *Proc. of the 2nd International Vegetation User Conference*, Antwerp, Belgium.



- [Lentile et al., 2006]Lentile, L. B., Holden, Z. A., Smith, A. M. S., Falkowski, M. J., Hudak, A. T., Morgan, P., Lewis, S. A., Gessler, P. E., and Benson, N. C. (2006). Remote sensing techniques to assess active fire characteristics and post-fire effects. *International Journal of Wildland Fire*, 15(3) :319–345.
- [Li, 1995]Li, S. Z. (1995). *Markov Random Field Modeling in Computer Vision*. Springer-Verlag.
- [Lorette, 1998]Lorette, A. (1998). Analyse de textures par méthodes markoviennes et par morphologie mathématiques : application à l’analyse des zones urbaines sur des images satellitales. Thèse de doctorat, Université de Sophia-Antipolis.
- [Mather, 2004]Mather, P. (2004). *Computer processing of remotely-sensed images. An introduction*. John Wiley & Sons, Ltd, third edition edition.
- [Melgani and Bruzzone, 2004]Melgani, F. and Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with Support Vector Machines. *IEEE Trans. Geoscience and Remote Sensing*, 42(8) :1778–1790.
- [Mercier and Lennon, 2003]Mercier, G. and Lennon, M. (2003). Support Vector Machines for hyperspectral image classification with spectral-based kernels. *Geoscience and Remote Sensing Symposium*, 1 :288–290.
- [Mitra et al., 2004]Mitra, P., Shankar, B. U., and Pal, S. K. (2004). Segmentation of multispectral remote sensing images using active Support Vector Machines. *Pattern Recognition Letters*, 25(9) :1067–1074.
- [Pal and Mather, 2005]Pal, M. and Mather, P. M. (March 2005). Support Vector Machines for classification in remote sensing. *International Journal of Remote Sensing*, 26 :1007–1011.
- [Palubinskas et al., 1998]Palubinskas, G., Descombes, X., and Kruggel, F. (1998). An unsupervised clustering method using the entropy minimization. In *International Conference on Pattern Recognition (ICPR)*, Brisbane, Australie.
- [Pereira, 1999]Pereira, J. M. C. (1999). A comparative evaluation of NOAA/AVHRR vegetation indices for burned surface detection and mapping. *IEEE Trans. on Geoscience and Remote Sensing*, 37 :217–226.
- [Pilar Martin et al., 2006]Pilar Martin, M., Gomez, I., and Chuvieco, E. (2006). Burnt area index (BAIM) for burned area discrimination at regional scale using MODIS data. In *International Conference on Forest Fire Research*, Figueira da Foz, Portugal.
- [Pinty and Verstraete, 1992]Pinty, B. and Verstraete, M. M. (1992). GEMI : A nonlinear index to monitor global vegetation from satellites. *Vegetation*, 101 :15–20.
- [Quintano et al., 2006]Quintano, C., Fernandez-Manso, A., Fernandez-Manso, O., and Shimabukuro, Y. E. (2006). Mapping burned areas in mediterranean countries using spectral mixture analysis from a uni-temporal perspective. *International Journal of Remote Sensing*, 27(4) :645–662.
- [Rätsch et al., 2000]Rätsch, G., Schölkopf, B., Mika, S., and Müller, K.-R. (2000). SVM and boosting : One class. Technical report 119, GMD FIRST.
- [Richards and Jia, 1999]Richards, J. A. and Jia, X. (1999). *Remote Sensing Digital Image Analysis : An Introduction*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

- [Rogan and Franklin, 2001]Rogan, J. and Franklin, J. (2001). Mapping burn severity in Southern California using spectral mixture analysis. In *IEEE International Geoscience And Remote Sensing Symposium (IGARSS)*, Sydney, Australie.
- [Roli and Fumera, 2001]Roli, F. and Fumera, G. (2001). Support Vector Machines for remote-sensing image classification. In Serpico, S. B., editor, *Image and Signal Processing for Remote Sensing VI*, volume 4170, pages 160–166, Bellingham. SPIE Proceedings.
- [Rouse et al., 1973]Rouse, J. W., Haas, R. H., Schell, J. A., and Deering, D. W. (1973). Monitoring vegetation systems in the great plains with ERTS. In *Third ERTS Symposium*.
- [Schmuck et al., 2007]Schmuck, G., San-Miguel Ayanz, J., Barbosa, P., Camia, A., Kucera, J., Amatulli, G., Boca, R., Schulte, E., and Dierks, H.-H. (2007). Forest fires in europe 2006. - EUR 22931 EN, JRC37598,.
- [Schowengerdt, 1997]Schowengerdt, R. A. (1997). *Remote sensing : models and methods for image processing. Second edition*. Academic Press.
- [Schölkopf et al., 1999]Schölkopf, B., Burges, C. J. C., and Smola, A. J. (1999). *Advances in Kernel Methods : Support Vector Learning*. MIT Press.
- [Schölkopf et al., 2001]Schölkopf, B., Platt, J. C., Shawe-Taylor, J. C., Smola, A. J., and Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. *Neural Computation*, 13(7) :1443–1471.
- [Schölkopf et al., 2004]Schölkopf, B., Tsuda, K., and Vert, J.-P. (2004). *Kernel Methods in computational biology*. MIT Press.
- [Serra, 1983]Serra, J. (1983). *Image Analysis and Mathematical Morphology*. Academic Press, Inc., Orlando, FL, USA.
- [Smith et al., 2002]Smith, A. M. S., Wooster, M. J., Powell, A. K., and Usher, D. (2002). Texture based feature extraction : application to burn scar detection in Earth observation satellite sensor imagery. *International Journal of Remote Sensing*, 23 :1733–1739.
- [Soille, 1999]Soille, P. (1999). *Morphological Image Analysis : Principles and Applications*. Springer-Verlag.
- [Stroppiana et al., 2000]Stroppiana, D., Maggi, M., Pereira, J. M., Graetz, D., Grégoire, J. M., Silva, J., Sà, A., Henry, P., Gond, V., and Bartholomé, E. (2000). Detection and mapping of burnt areas and active fires in tropical woodland ecosystems with the VEGETATION sensor : the SMOKO-FRACTAL case study over northern australia. In *Vegetation 2000 Conference*, Lake Maggiore, Italy.
- [Sunar and Ozkan, 2001]Sunar, F. and Ozkan, C. (2001). Forest fire analysis with remote sensing data. *International Journal of Remote Sensing*, 22(12) :2265–2277.
- [Tipping, 2000]Tipping, M. E. (2000). The relevance vector machine. In *Advances in Neural Information Processing Systems, San Mateo, CA*.
- [Tipping, 2001]Tipping, M. E. (2001). Sparse Bayesian Learning and the Relevance Vector Machine. *Journal of Machine Learning Research*, 1 :211–244.
- [Unser, 1995]Unser, M. (1995). Texture classification and segmentation using wavelet frames. *IEEE Trans. on Image Processing*, 4(11) :1549–1560.
- [Vapnik, 1998]Vapnik, V. N. (1998). *Statistical Learning Theory*. John Wiley and sons, inc.

- [**Vert, 2001**]Vert, J.-P. (2001). Introduction to Support Vector Machines and applications to computational biology. course notes, Ecole Nationale Supérieure des Mines de Paris, France.
- [**Vincent and Soille, 1991**]Vincent, L. and Soille, P. (1991). Watersheds in digital spaces : an efficient algorithm based on immersion simulations. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(6) :583–598.
- [**Winkler, 1995**]Winkler, G. (1995). *Image analysis, random fields and dynamic Monte Carlo methods : a mathematical introduction*. Springer.
- [**Zhu, 2005**]Zhu, X. (2005). Semi-supervised learning literature survey. Technical Report 1530, Computer Sciences, University of Wisconsin-Madison, USA.

# Publications et activités de recherche

## Journaux internationaux

1. O. Zammit, X. Descombes et J. Zerubia. Mapping burnt areas using Support Vector Machines. soumis à International Journal of Remote Sensing.

## Conférences internationales avec actes

1. O. Zammit, X. Descombes et J. Zerubia. Unsupervised One-Class SVM Using a Watershed Algorithm and Hysteresis Thresholding to Detect Burnt Areas. Pattern Recognition and Image Analysis 2008 (PRIA), Nijni Novgorod, Russie, Septembre 2008.
2. O. Zammit, X. Descombes et J. Zerubia. Combining One-Class Support Vector Machines and hysteresis thresholding : application to burnt area mapping. European Signal Processing Conference 2008 (EUSIPCO), Lausanne, Suisse, Août 2008.
3. O. Zammit, X. Descombes et J. Zerubia. Assessment of different classification algorithms for burnt land discrimination. IEEE International Geoscience and Remote Sensing Symposium 2007 (IGARSS), Barcelone, Espagne, Juillet 2007.
4. O. Zammit, X. Descombes et J. Zerubia. Burnt area mapping using Support Vector Machines. International Conference on Forest Fire Research 2006 (ICFFR), Figueira da Foz, Portugal, Novembre 2006.

## Conférences francophones avec actes

1. O. Zammit, X. Descombes et J. Zerubia. Apprentissage non supervisé des SVM par un algorithme des K-moyennes entropique pour la détection de zones brûlées. Colloques sur le Traitement du Signal et des Images - GRETSI, Troyes, France, Septembre 2007.

## Rapports de recherche

1. O. Zammit, X. Descombes et J. Zerubia. Support Vector Machines for burnt area discrimination. Rapport de Recherche INRIA n° 6343, Novembre 2007.

## **Séminaires**

1. Séminaire InTech'Sophia à Sophia Antipolis, France, le 21 juin 2007.
2. Séminaires à l'Université de Szeged, Hongrie, le 26 septembre 2006, au MTA SZTAKI (Académie des Sciences), Budapest, Hongrie, le 28 septembre 2006 et à l'Université Catholique Peter Pazmany, Budapest, Hongrie, le 29 septembre 2006.

## **Autres activités scientifiques**

1. Présentation des travaux du projet Ariana en environnement (détection des dégâts des feux, comptage des houppiers dans des plantations) le 12 octobre 2006 à Sophia Antipolis dans le cadre de la Fête de la Science.

# Résumé

Cette thèse aborde le problème de cartographie de zones brûlées à partir d'images satellitaires haute résolution. Nos modèles reposent sur le traitement d'une seule image SPOT 5, acquise après le feu afin de détecter automatiquement les zones brûlées.

Le modèle est fondé sur les Séparateurs à Vaste Marge (SVM), une technique de classification supervisée qui a démontré une meilleure précision et une meilleure capacité de généralisation que les algorithmes de classification plus traditionnels. Concernant notre problème de détection, les différentes zones brûlées possèdent des caractéristiques spectrales assez similaires, au contraire des zones non brûlées (végétation, routes, eau, zones urbaines, nuage, ombre...) dont les caractéristiques spectrales varient énormément. Nous proposons donc d'utiliser les One-Class SVM (OC-SVM), une technique qui dérive des SVM mais qui n'utilise que des exemples de pixels brûlés pour les phases d'apprentissage et de classification.

Afin de prendre en compte l'information spatiale de l'image, l'algorithme OC-SVM est utilisé comme une technique de croissance de régions, ce qui permet de diminuer les fausses alarmes et d'améliorer les contours des zones brûlées.

De plus, la base d'exemple de pixels brûlés nécessaire à l'apprentissage des techniques SVM est déterminée automatiquement à partir de l'histogramme de l'image.

Finalement, la méthode de classification proposée est testée sur plusieurs images satellitaires afin de valider son efficacité selon le type de végétation et la surface des zones brûlées. Les zones brûlées obtenues sont comparées à des cartes de référence fournies par le CNES, Infoterra France, le SERTIT, les Services Départementaux d'Incendies et de Secours ou l'Office National des Forêts.

## **Mots clés :**

Classification, image satellitaire, zones brûlées, feux de forêt, Séparateurs à Vaste Marge, croissance de régions, apprentissage.



# Abstract

This PhD investigates the problem of burnt area mapping from high-resolution satellite images. Our approach is based on the use of a single SPOT 5 image, acquired after the fire, to detect automatically the burnt areas.

The method is based on Support Vector Machines (SVM), a supervised classification technique that has greater accuracy and better generalization ability than traditional classifiers. Because all burnt pixels have similar spectral characteristics, while unburnt pixels vary greatly since they belong to different classes (forest, water, urban areas, roads, fields,...), we propose to use the One-Class SVM (OC-SVM), an extension of the original two-class SVM that uses only positive examples for the training and classification steps.

In order to take into account the spatial information provided by the image, the OC-SVM algorithm is used as a region-growing technique, thereby decreasing false positives and improving the boundaries of burnt areas.

Moreover, the samples of burnt areas required for the training step of the SVM are automatically selected from the image histogram.

Finally, the proposed classification approach is tested on several satellite images to validate its effectiveness with respect to vegetation type and burnt surface area. The burnt areas extracted are compared to reference maps provided by the French Space Agency, Infoterra France, SERTIT, local fire brigades, and the French Forestry Office.

**Keywords :**

Classification, satellite image, burnt areas, forest fire, Support Vector Machines, region growing, learning.