

## MODÉLISATION DE LA DISTRIBUTION SPATIALE DES PRÉCIPITATIONS EN SICILE

Yves CORNET, Catherine FITSCHY, Oscar GOBBATO et Michel ERPICUM

### Résumé

Nous présentons une technique de cartographie de la distribution spatiale des précipitations en Sicile. Il s'agit d'une modélisation statistique permettant d'estimer ces précipitations en fonction de variables topographiques. La composante stochastique résultant de l'ajustement par moindres carrés a été soumise à une interpolation par krigeage. Les résultats de cette modélisation et son exactitude sont discutés en regard des résultats d'autres recherches du même type, d'une part, et par rapport à une interpolation portant sur les données brutes, d'autre part. Si l'on compare les résultats de notre modèle aux résultats de cette interpolation, le principal intérêt de notre approche est plus l'identification des facteurs agissant de façon déterminante sur la pluviométrie selon la saison que l'exactitude à l'échelle du territoire étudié. Cette identification est particulièrement intéressante pour mieux circonscrire les zones qui présentent une sensibilité particulière à la désertification sur base du seul critère considéré, la pluviométrie. On peut néanmoins envisager de développer des modèles semblables pour d'autres éléments du climat. Il est évident que le processus de désertification est influencé par d'autres facteurs que les seules conditions climatiques, mais ils n'ont pas été envisagés dans cette recherche.

### Mots-clés

Modèle statistique, cartographie, précipitations, Sicile

### Abstract

*We present a method for the mapping of the spatial distribution of precipitations in Sicily. This is a statistical modeling of the precipitations depending on topographical variables. The stochastic component resulting from the least squares adjustment has been submitted to a kriging interpolation process. The results of this modeling and its accuracy are firstly compared to the results obtained in other researches using the same approach and secondly to the results obtained from interpolation applied to the raw data. Regarding the last comparison, the main interest of our modeling is the identification of the factors acting in a determining way on pluviometry according to the season rather than a better accuracy at the spatial scale of the analyzed territory. This identification is particularly interesting for better circumscribing the zones which have a particular sensitivity to the desertification on the basis of the only considered criterion, pluviometry. One can nevertheless develop similar models for other climatic elements. Of course the desertification process is not only influenced by climatic conditions. Other factors must be taken into account to explain this process, but they were not considered in this research.*

### Keywords

*Statistic model, cartography, precipitations, Sicily*

## I. INTRODUCTION

Les méthodes exposées dans cette publication ont été développées dans le but d'affiner la cartographie des précipitations en Sicile. Les résultats obtenus sont présentés et discutés d'un point de vue climatologique en relation avec la problématique de la désertification (Gobbato, 2000). Dans le contexte de la Sicile, située en marge des zones affectées par la désertification, l'amélioration de l'exactitude de la cartographie des précipitations est nécessaire à l'interprétation évolutive de ce processus.

Nous avons volontairement concentré notre recherche sur un seul des éléments du climat, la pluviométrie. L'approche méthodologique envisagée dans cette étude reposant sur les principes de la modélisation statistique, celle-ci pourra être transférable à l'étude de la température.

Nous avons exploité les variables physiographiques connues de façon continue sur le territoire et intervenant dans la détermination des lames d'eau mensuelles. Ces variables sont les variables explicatives qui ont été traitées en mode image à une résolution de 250 m. Les la-

mes d'eau n'étant connues qu'en 176 points de l'espace étudié, elles ont été exploitées pour ajuster un modèle statistique exprimant la relation entre variables physiographiques et lames d'eau précipitées. Vingt de ces stations ont cependant été extraites de l'échantillon avant ajustement pour permettre la validation de la modélisation. Le modèle statistique obtenu a ensuite été exploité pour spatialiser les précipitations en exploitant d'abord sa composante déterministe et ensuite sa composante stochastique. La composante déterministe traduit les relations existant entre les précipitations et les variables physiographiques retenues parmi celles envisagées par le processus statistique de sélection utilisé. La composante stochastique est connue sous la forme de résidus aux stations qui peuvent être expliqués par des facteurs ignorés dans le modèle. Ces résidus ont fait l'objet d'une analyse de continuité spatiale et d'une interpolation par krigeage de façon à pouvoir corriger les surfaces de précipitations attendues par l'exploitation de la seule composante déterministe.

Cette approche permet de prendre en compte l'hétérogénéité de la variabilité spatiale du climat à l'échelle topoclimatique, échelle d'analyse qui se justifie pleinement vis-à-vis de la complexité de la désertification observée dans cette région.

La continuité spatiale des résidus du modèle statistique est donc une condition nécessaire à l'application de la méthode. Elle ne peut être vérifiée que dans certaines situations d'espacement des données et de conditions de variabilité spatiale du phénomène considéré par rapport à cet espacement. Ainsi, lorsque les stations sont trop distantes les unes des autres par rapport à l'échelle spatiale de variabilité, le krigeage est difficilement applicable. Dans ce cas, la modélisation de la relation entre semi-variance et décalage spatial (*lag*) ne permet pas de rendre compte de la façon dont varie cette semi-variance dans le voisinage proche d'un point quelconque de l'espace considéré. Par ailleurs, lorsque d'éventuels facteurs, négligés dans le modèle et pouvant expliquer une part des résidus, ne présentent plus aucune corrélation spatiale à l'échelle de l'espacement des stations pour les premiers ordres de voisinage, l'interpolation sur les résidus ne se justifie plus.

En 1986, Benichou & Lebreton ont appliqué une méthode similaire de cartographie de champs pluviométriques. Lennon & Turner (1995) et Goodale *et al.* (1998) ont également mené, au Royaume-Uni, des recherches portant sur ce type de modélisation climatique. En Belgique, de tels types de modèles ont été développés pour estimer les précipitations et les températures à une résolution hectométrique à kilométrique en se basant sur un réseau très dense de stations (Alexandre *et al.*, 1999). Dans cette dernière étude, la seule variable physiographique exploitée était l'altitude obtenue grâce au Modèle Numérique de Terrain (MNT) de niveau 1 produit par l'IGN belge.

Sa résolution est de 6" de longitude et 3" de latitude, au nord de 50° Nord, et 3" de longitude et 3" de latitude, au sud de 50° Nord. Ce MNT fut rééchantillonné dans le référentiel cartographique Lambert belge à la résolution de travail. L'échelle temporelle considérée dans ce cas était mensuelle (normales climatiques établies sur 30 années d'observation).

Dans le bassin méditerranéen, Carrega & Garcia (2000) et Assaba & Laborde (2000) ont réalisé des recherches du même type, respectivement en Corse et en Toscane. Sur la totalité du bassin méditerranéen, Agnew & Palutikof (2000) ont, dans le cadre du projet MEDALUS, exploité les données de 248 stations thermométriques et 285 stations pluviométriques pour modéliser la distribution spatiale des températures et précipitations. Cette modélisation fut validée en exploitant 36 autres stations pour les températures et 35 pour les précipitations. Les variables explicatives introduites dans le modèle par ces chercheurs étaient les suivantes : la longitude ; la latitude ; la plus courte distance à la côte ; la direction associée à cette distance était estimée par secteur de 45° (8 variables binaires) ; le rapport entre surface de terre émergée et surface de mer dans des zones circulaires de 5, 10, 20, 50, 100 et 150 km ; l'altitude ; l'altitude maximum dans un rayon de 1, 2, 3, 4, 5, 10, 20, 50 et 100 km en fonction de la direction par secteur de 45° ; la pente et la direction de ligne de plus grande pente par secteur de 45° également (8 variables binaires).

La zone d'analyse étudiée par Agnew & Palutikof (2000) était, quant à elle, beaucoup plus vaste puisque leur recherche couvrait tout le bassin méditerranéen. À cause de cette grande étendue, ces chercheurs ont exploité un système de référence spatiale de type Plate-Carrée. Ce référentiel n'est toutefois pas particulièrement adapté à l'interpolation sur une aussi vaste zone. L'interpolation repose, en effet, sur des mesures de distances qui, dans ce référentiel, ne sont pas comparables sur toute la surface considérée et dépendent de leur direction de mesure. Ainsi, pour illustrer les situations les plus défavorables, une distance angulaire de 80 minutes décimales, par exemple, selon l'arc du parallèle de 45° N correspond, sur une sphère terrestre de 6 375 km de rayon, à une longueur de 105 km. Sur le parallèle de 35° N, cette même distance angulaire correspond, sur une même sphère, à une distance de 121 km. Des distances semblables dans le référentiel choisi présentent en fait une différence maximale de l'ordre de 16 km, soit environ 15 %, dans la réalité. Par ailleurs, étant donné la configuration du bassin méditerranéen, les stations sont généralement concentrées sur sa périphérie. Leur distribution spatiale et leur nombre relativement petit par rapport à la grandeur de cet espace sont défavorables à l'exploitation de la corrélation spatiale des résidus pour leur interpolation par krigeage, méthode utilisée. L'exemple de semi-variogramme donné par ces chercheurs atteste d'ailleurs de cette faiblesse. Les semi-variances faibles

correspondant aux petites distances sont trop peu fiables et trop peu nombreuses pour rendre compte de la forme du semi-variogramme modélisé. De plus, l'hétérogénéité spatiale de la densité de stations implique que la partie du semi-variogramme modélisé correspondant au voisinage proche ne traduit que la continuité spatiale des zones à forte concentration de stations.

Notre modélisation de la distribution spatiale des précipitations en Sicile s'inspire de notre expérience belge (Alexandre *et al.*, 1998 et 1999) et de celle de Agnew & Palutikof (2000) en Méditerranée. L'île sicilienne constitue un cas idéal d'un point de vue méthodologique et très intéressant d'un point de vue climatologique.

En effet, du point de vue méthodologique, dans ce petit espace du bassin méditerranéen, nous pouvons nous affranchir de l'inconvénient lié au référentiel spatial et à la densité spatiale des stations pluviométriques (figure 1) largement plus élevée que celle exploitée par Agnew & Palutikof (2000). Elle reste cependant plus faible que celle rencontrée en Belgique (Alexandre *et al.*, 1999), mais doit quand même permettre de modéliser précisément le semi-variogramme. Il est ainsi possible de contrôler si le type de modèle mis au point par Agnew & Palutikof (2000) à l'échelle du bassin méditerranéen est transposable à une échelle spatiale plus compatible avec une analyse de la désertification.

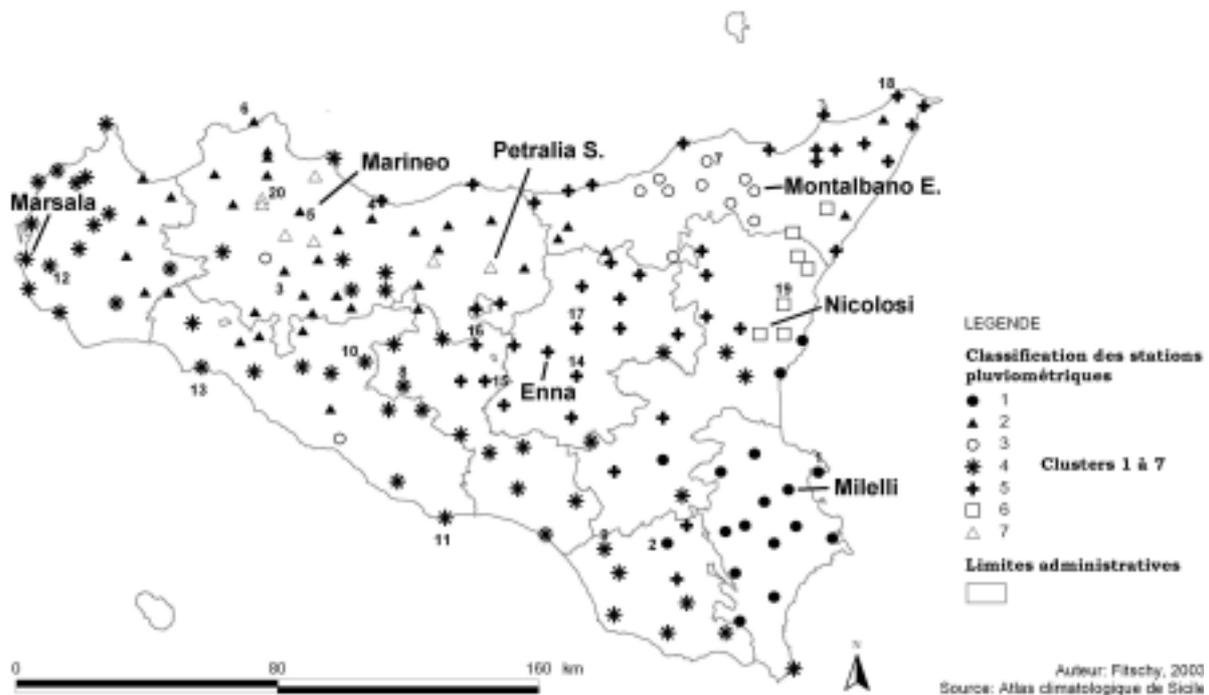
Du point de vue climatologique, la position de la Sicile,

située entre les deux grands bassins de Méditerranée occidentale et orientale en font une zone d'étude particulièrement intéressante, car les caractéristiques thermiques et hydrodynamiques des deux bassins sont très contrastées. De plus, la disposition des principales chaînes de montagnes, les Madonie, Nebrodi et Peloritani, avec leur impact sur l'influence relative de ces deux bassins marins sur la continentalité constitue un autre argument favorable au choix de cette île. Par ailleurs, une partie de la Sicile (province d'Agrigente) a déjà fait l'objet d'une étude détaillée de l'un d'entre nous (Gobbato, 2001), en collaboration avec l'ENEA, dans laquelle les critères établis par Kosmas *et al.* (1999) ont été appliqués pour analyser les effets du climat sur la désertification.

Après avoir expliqué le développement du modèle statistique, nous validerons ses résultats sur base d'un échantillon stratifié de stations de contrôle retirées *a priori* de l'échantillon utilisé pour l'ajustement par moindres carrés. L'exactitude de la seule composante déterministe sera comparée à celle du modèle après ajout de la composante stochastique et à celle du résultat d'une interpolation portant sur les données brutes de précipitations.

## II. LES DONNÉES

Les données exploitées sont celles de l'Atlas climatologique de Sicile (2001) et un MNT à 250 m de réso-



**Figure 1.** Localisation et typologie des stations selon leur régime pluviométrique (cf. § IIB)

Les 7 stations nommées sont les plus représentatives de chaque *cluster* ; leurs régimes pluviométriques sont illustrés par la figure 3. Les stations numérotées de 1 à 20 sont les stations de validation.

lution gracieusement mis à notre disposition par l'ENEA (*Ente per le Nuove tecnologie, l'Energia e l'Ambiente*).

Le MNT a été produit à une résolution de 250 m dans le référentiel cartographique UTM. L'Atlas climatologique de Sicile a été réalisé sous la forme de CDROM par l'INTEA S.r.l. de Palerme sur base des méthodes de traitement des données et de leur représentation cartographique élaborées par l'*Unità Operativa di Agrometeorologia* (Drago *et al.*, 1998) de l'*Assessorato Agricoltura e Foreste, Servizi allo Sviluppo* de la région autonome de Sicile. Le CDROM est constitué d'une base de données collectées sur une période d'observation de 30 ans (1965-1994) aux 55 stations thermo-pluviométriques et aux 124 stations pluviométriques du service hydrographique régional à partir des données publiées dans les annales hydrologiques.

## A. Données pluviométriques

### 1. Localisation des stations

Seule la pluviométrie a été étudiée, car le nombre de stations pluviométriques autorise une approche à l'échelle désirée et une modélisation plus complète (nombre et variabilité des facteurs explicatifs), alors que les stations thermométriques sont trop peu nombreuses.

La projection cartographique des cartes italiennes est la projection de Gauss-Boaga et le système géodésique national italien est défini par rapport à l'ellipsoïde international orienté à M. Mario (Rome) en 1940. En surimpression, les cartes italiennes disposent d'un canevas kilométrique UTM du système géodésique européen (ED 1950). Nous avons projeté l'ensemble des données (*shapefiles* exploités par l'Atlas climatologique et MNT) dans ce pseudo-référentiel cartographique. La position des stations est donc repérée dans ce référentiel en unités métriques.

Parmi les 179 stations disponibles, nous en avons d'emblée éliminé 3 qui sont situées sur les îles éoliennes et l'île du seuil de Tunisie, soit Salina, Lipari et Pantelleria. En effet, les relations entre précipitations et variables explicatives risquent d'y être très différentes de celles qui existent sur l'île principale.

### 2. Précipitations moyennes ou précipitations médianes

Comme le rappellent Drago *et al.* (1998), l'analyse climatologique des précipitations à partir de moyennes annuelles calculées sur une période de 30 ans n'est pas recommandée dans les régions méditerranéennes. En effet, les moyennes peuvent être influencées par des événements climatiques exceptionnels et marginaux qui,

par ailleurs, n'ont pas toujours affecté l'ensemble des stations de la zone d'étude. Ainsi, contrairement à Agnew & Palutikof (2000) et à la plupart des autres auteurs déjà cités, nous avons donc exploité les médianes qui sont moins sensibles à ces événements.

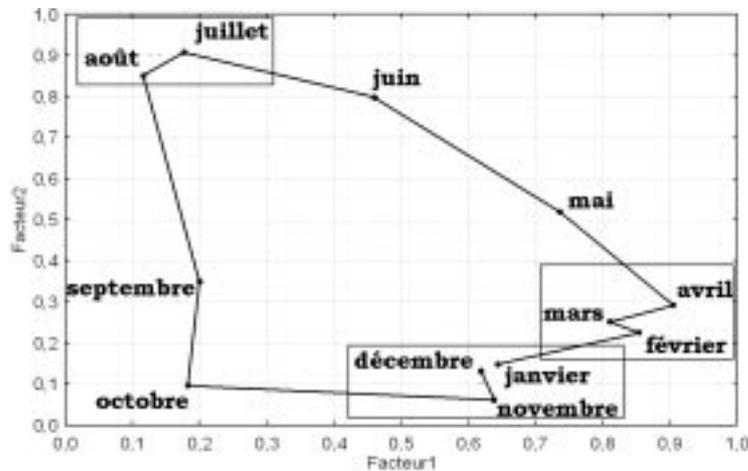
### 3. Qualité des données pluviométriques

En ce qui concerne la qualité des données exploitées pour calculer ces médianes, Drago *et al.* (1998) précisent la procédure exploitée pour identifier d'éventuelles données erronées. Cette méthode est basée sur une première comparaison, pour une même station, de chaque valeur à la distribution des valeurs de la période correspondante. Dans le cas où une valeur anormale est ainsi repérée à un moment donné de la série temporelle, une seconde comparaison aux valeurs des stations voisines est réalisée à ce moment. Les erreurs confirmées par cette seconde comparaison ont été corrigées et les valeurs manquantes ont été estimées en se fondant sur les valeurs observées aux stations voisines au moment considéré. Cette méthode standard de contrôle de qualité devrait assurer la validité des médianes extraites de la base de données en exploitant les outils du CDROM réalisé par l'INTEA S.r.l.

### 4. Résolution temporelle

Agnew & Palutikof (2000) ont créé un modèle pour chaque trimestre de l'année. Cette façon de procéder permet de réduire l'influence éventuelle d'événements exceptionnels sur les moyennes exploitées par ces chercheurs. Nous avons préféré l'approche d'Alexandre *et al.* (1998) qui ont développé un modèle pour chaque mois de l'année. Il semble d'ailleurs que, dans notre cas, le découpage de l'année en 4 trimestres ne soit pas justifié d'un point climatologique.

Nous avons classé les stations (*cf.* § II.B.) pour stratifier l'échantillon destiné à la validation. Leur regroupement a été effectué par classification hiérarchique appliquée au résultat d'une ACP (Analyse en Composantes Principales) des 176 valeurs de précipitations médianes pour chacun des 12 mois de l'année. Les poids factoriels de chacune des 12 variables correspondant aux précipitations mensuelles médianes font clairement ressortir un regroupement de certains mois dans l'espace des facteurs caractérisant les différents régimes pluviométriques (figure 2). Ce regroupement ne correspond pas à un découpage de l'année en 4 trimestres. Par ailleurs, les mois de novembre, décembre, janvier, février, mars et avril sont assez semblables en constituant toutefois deux sous-groupes proches. Les autres mois ne montrent qu'assez peu de similitudes entre eux, sauf les deux mois typiquement estivaux : juillet et août. Une classification hiérarchique des précipitations mensuelles dans l'espace des stations a, par ailleurs, confirmé cette constatation.



**Figure 2.** Extraction des composantes principales. Poids factoriels des précipitations médianes des 12 mois de l'année pour les 2 premiers facteurs après rotation de type « varimax normalisé » sur 3 facteurs

Ces arguments ont confirmé la nécessité de calculer un modèle de distribution spatiale des précipitations spécifique pour chaque mois de l'année.

### B. Classification des stations sur base du régime pluviométrique annuel

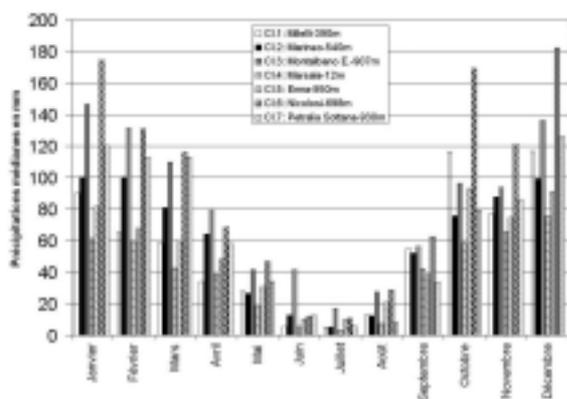
Une analyse préliminaire des stations a été menée afin de déterminer les similitudes entre les régimes pluviométriques annuels des stations météorologiques considérées. Pour cela, nous avons exploité les résultats de l'ACP portant sur les 12 variables que constituent les précipitations médianes de chaque mois sur les 176 cas (les stations). La matrice d'inertie exploitée est celle des coefficients de corrélation. Trois facteurs ont été extraits. Les deux premiers expliquent respectivement 64,5 et 14,8 % de la variance totale, soit déjà près de 80 % de celle-ci, alors que le troisième en explique encore 8,9 %. L'analyse des communautés nous renseignant sur la représentativité des précipitations médianes par les facteurs calculés montre l'importance de ce troisième facteur afin que tous les mois soient représentés dans l'espace factoriel. Une rotation a enfin été appliquée au nouveau système de référence obtenu. L'analyse des poids factoriels exprimant les corrélations entre les variables de départ et les facteurs calculés (figure 2) permet d'interpréter la signification de ces facteurs et surtout de reconnaître les similitudes entre les précipitations médianes mensuelles des différents mois.

Le facteur 1 traduit essentiellement l'importance des quantités de précipitations printanières. Le facteur 2 traduit l'importance des précipitations estivales par rapport aux précipitations hivernales. Le facteur 3, non reporté sur le graphique de la figure 2, traduit quant à lui l'importance des précipitations automnales.

Les stations caractérisées par leur score factoriel obtenu selon ces 3 facteurs ont ensuite été classées. Les distances euclidiennes de regroupement entre stations et groupes de stations dans l'espace factoriel ont été analysées selon les principes de la classification hiérarchique de Ward. Cette approche a ainsi permis de mettre en évidence l'existence de 7 classes de stations à régime pluviométrique semblable. Une classification par nuée dynamique de type *ISOData (Iterative Self Organizing Data)* en 7 classes a ensuite été appliquée aux données. Son résultat est représenté sur la carte de la figure 1.

On constate que la plupart des stations se regroupent dans l'espace géographique selon le *cluster* auquel elles appartiennent. Sur cette figure 1, nous avons également reporté une étiquette numérique allant de 1 à 20 à côté de certains symboles représentant les stations. Il s'agit des 20 stations qui ont été extraites du set de données pour valider les modèles de distribution spatiale de précipitations. Ces 20 stations ont été extraites du set de données par échantillonnage aléatoire stratifié au sein de chaque *cluster*.

Nous avons reporté sur la figure 3 l'histogramme des précipitations mensuelles médianes pour chacune des stations caractéristiques de chaque *cluster*. Par rapport au territoire sicilien, les stations du *cluster* 1, principalement localisées dans la province de Syracuse à l'est des Mts Iblei, enregistrent des précipitations relativement faibles pendant les mois hivernaux à partir de janvier. Cette faiblesse des précipitations est aussi observée au printemps et en été. Par contre, l'automne et, dans une moindre mesure, le début de l'hiver y sont plus pluvieux que dans la majorité des autres stations. Par ailleurs, la sécheresse de l'été apparaît très intense et très tôt à partir de juin jusqu'à août.



**Figure 3.** Régimes pluviométriques typiques des 7 stations caractéristiques de chaque *cluster*

Par rapport aux autres types de stations, celles du *cluster* 2, réparties au nord et au nord-ouest de l'île, se caractérisent par des précipitations encore importantes au printemps. Le minimum de juillet est bien marqué et l'automne se caractérise par un accroissement important des précipitations dès septembre. Ensuite, par rapport au *cluster* 1, le volume de précipitations croît plus régulièrement et le maximum relatif d'octobre n'existe pas.

Les stations du *cluster* 3 sont, à l'exception de deux cas, situées au nord-est des Mts Nebrodi. Il s'agit des stations parmi les plus pluvieuses après celles du *cluster* 6. La période pluvieuse se prolonge jusqu'en juin alors que les stations des autres *clusters*, même celles du 6, enregistrent, pendant ce mois, des précipitations beaucoup plus faibles. Des précipitations estivales moins faibles que dans les autres types de stations sont observées en juillet et août. Comme dans le cas du *cluster* 1, l'automne est pluvieux bien que novembre enregistre une diminution des précipitations par rapport à octobre.

Les stations du *cluster* 4 sont assez dispersées sur l'entièreté du territoire du nord-ouest au sud-est, mais elles sont rarement localisées dans la moitié nord du territoire. Elles enregistrent un volume de précipitations qui est souvent le plus faible pendant la grande majorité des mois de l'année. L'évolution des précipitations d'un mois à l'autre est plus lente et régulière que dans les autres stations. Ce sont les stations qui connaissent la période estivale la moins arrosée, la plus longue et la plus marquée.

Les stations du *cluster* 5 sont presque toutes situées à l'extrémité nord-est des Mts Peloritani et à l'ouest de l'Etna au sud des Mts Nebrodi et Madonie ou le long de la côte septentrionale. Elles connaissent, de décembre à juin-juillet, une évolution des précipitations assez comparable à celle des stations du *cluster* 4, mais leur volume y est un peu plus important. L'augmentation des

précipitations en début d'automne est relativement lente, mais un pic est enregistré en octobre avant une légère diminution en novembre. Par cette distribution temporelle des précipitations automnales, ce *cluster* est semblable aux *clusters* 1 et 3. L'amplitude du cycle annuel est relativement faible en raison de pluies hivernales plutôt faibles et d'un été un peu moins sec que dans la majorité des autres stations.

Les stations du *cluster* 6, situées sur les versants orientaux de l'Etna et des Mts Peloritani, ont dès septembre et jusqu'en mai une saison très pluvieuse. En novembre, ces stations enregistrent également une diminution des précipitations. Il s'agit des stations les plus pluvieuses où l'amplitude du cycle annuel est la plus importante.

Enfin, les stations du *cluster* 7 sont localisées au nord et au nord-est des Mts Sicani. Elles ont un hiver et un printemps pluvieux qui se prolonge jusqu'en avril. Juin, juillet et août sont secs. L'augmentation de précipitations en automne est relativement faible et le mois de novembre se marque par un arrêt de leur augmentation vers l'hiver. Le maximum d'octobre n'y est pas constaté.

### C. Les variables explicatives

À l'exception des coordonnées cartographiques des stations, les variables explicatives retenues ont été extraites du MNT. Elles sont décrites et discutées ci-dessous.

#### 1. X\_COORD et Y\_COORD

Ces deux premières variables sont les coordonnées des stations exprimées en mètres, dans le référentiel UTM. Elles sont introduites pour modéliser la tendance spatiale. Étant donné que le modèle adopté est du premier degré, cette tendance répondra à l'équation d'un plan ou d'une droite selon que les deux ou une seule des deux coordonnées seront retenues par la procédure de choix des variables explicatives.

Ces coordonnées proviennent de la base de données géographiques de l'Atlas climatologique de Sicile. Les champs X\_COORD et Y\_COORD ont été ajoutés sous ArcView (*SHAPE.ADDXYCOORD*) dans la table d'attributs de la couverture (*shapefile*) de points correspondant aux stations météorologiques.

#### 2. ALT

Il s'agit de l'altitude des stations exprimée en mètres. Bien que présente dans la base de données de l'Atlas climatologique de Sicile, cette variable a été recalculée sous ArcInfo (*LATTICESPOT*) par interpolation bilinéaire dans le MNT à l'endroit de chaque station pluviomé-

trique. Les altitudes utilisées lors de l'ajustement du modèle statistique sont celles de la base de données de l'Atlas.

### 3. DZMA<sub>ij</sub>

Ces variables au nombre de 56 ont été introduites dans le modèle pour prendre en considération, dans un voisinage donné, l'influence des chaînes de montagnes sur la progression et la modification des propriétés des masses d'air responsables des précipitations. Il s'agit de la différence d'altitude entre chaque station et le point culminant dans chacun des 8 secteurs circulaires de 45° et de rayon donné. L'indice *i* représente le rayon en kilomètres qui varie de la façon suivante : 1, 2, 3, 4, 5, 10 et 20. L'indice *j* représente le secteur. Les secteurs considérés sont les suivants : N, NE, E, SE, S, SW, W et NW.

### 4. ELEVi

Il s'agit de 5 variables calculées sous ArcInfo (FOCALMEAN). Elles permettent de prendre en considération l'influence de l'altitude moyenne autour de la station dans un rayon de *i* km. Les rayons considérés sont 1, 2, 3, 4, 5 et 10 km.

Si toutes les autres influences restent constantes, il est évident que l'accroissement de l'altitude moyenne se traduit généralement par un accroissement des précipitations lié au phénomène de détente (pseudo-)adiabatique des masses d'air. Le refroidissement qui en découle est à l'origine d'un rapprochement du point de saturation. Les précipitations sont donc plus abondantes puisque cet état est plus fréquemment atteint. Cependant, l'influence de l'altitude sur les précipitations ne sera pas nécessairement la même pour une station située dans une région d'altitude moyenne supérieure à la station (site concave, vallée, cuvette) que dans une région d'altitude moyenne plus faible (site convexe, crête, sommet) ou dans une région d'altitude moyenne comparable (versant régulier à l'échelle du voisinage considéré, plaine ou plateau). La prise en compte simultanée de l'altitude (ALT) et de l'élévation moyenne dans le voisinage (ELEVi) permet d'évaluer cette influence sur les précipitations.

### 5. X et Y

X et Y ( $e_x$  et  $e_y$  dans les formules ci-dessous) sont les coordonnées orthogonales de l'extrémité du vecteur axial (hémisphère inférieur,  $e_z \leq 0$ ) du plan tangent à la surface topographique. Ces variables permettent d'introduire dans le modèle l'influence de l'orientation et de la pente du site à l'échelle d'un voisinage de 750 m sur 750 m (3x3 pixels du MNT).

Ces deux composantes ont été calculées sous ArcInfo à

l'aide des fonctions focales et des opérateurs arithmétiques du module GRID.

La formulation mathématique de ces composantes dans un voisinage 3x3 est la suivante.

Soit l'équation d'un paraboloïde passant par les 9 points d'un kernel 3x3 (système de référence planimétrique local ayant pour origine le centre de ce kernel) :

$$Z = aX^2Y^2 + bX^2Y + cXY^2 + dX^2 + eY^2 + fXY + gX + hY + i \quad (1)$$

Soit les dérivées partielles de cette fonction par rapport à X et Y en (0,0,i) :

$$\left(\frac{\partial Z}{\partial X}\right)_{en(0,0)} = g \quad \text{et} \quad \left(\frac{\partial Z}{\partial Y}\right)_{en(0,0)} = h \quad (2)$$

Soit l'équation d'un plan parallèle au plan tangent au paraboloïde au centre du kernel et passant par l'origine (0,0,0) :

$$\pi \equiv e_x X + e_y Y + e_z Z = 0 \quad (3)$$

avec  $e_x^2 + e_y^2 + e_z^2 = 1$

On en déduit que

$$Z = -\frac{e_x}{e_z} X - \frac{e_y}{e_z} Y \quad (4)$$

$$\text{avec} \quad -\frac{e_x}{e_z} = g \quad \text{et} \quad -\frac{e_y}{e_z} = h$$

On peut ainsi déterminer les 3 composantes du vecteur axial du plan *p* à condition de connaître *g* et *h* :

$$e_z = \pm \sqrt{\frac{1}{g^2 + h^2 + 1}} \quad (5)$$

$$e_x = -g e_z$$

$$e_y = -h e_z$$

$e_z$  est > ou < 0 selon que l'on considère l'hémisphère supérieur ou inférieur.

calcul de g	calcul de h						
<table border="1" style="margin: auto;"> <tr><td>-1</td><td>0</td><td>1</td></tr> </table> $/ 2*r$	-1	0	1	<table border="1" style="margin: auto;"> <tr><td>1</td></tr> <tr><td>0</td></tr> <tr><td>-1</td></tr> </table> $/ 2*r$	1	0	-1
-1	0	1					
1							
0							
-1							

*g* et *h* sont calculés par convolution de chacun des deux kernels suivants, puis en divisant le résultat par le double de la résolution *r* du MNT :

### 6. X9 et Y9

X9 et Y9 sont les composantes du vecteur unitaire donnant l'orientation de la résultante de l'addition des 81 vecteurs axiaux des plans tangents à la surface topogra-

phique calculés dans un voisinage de 2 250 m sur 2 250 m (9x9 pixels) selon la procédure expliquée ci-dessus. Ces deux variables représentent donc l'orientation moyenne de la zone dans laquelle se trouve la station à l'échelle définie par le voisinage considéré. Leur calcul a été effectué en exploitant la fonction FOCALSUM du module GRID de ArcInfo puis en divisant chaque composante par la norme de cette résultante.

### 7. X19 et Y19

Ces deux variables peuvent être définies comme X9 et Y9. La seule différence est que le voisinage considéré est de 4 750 m sur 4 750 m (19x19 pixels).

### 8. DISTCOST

*DISTCOST* est la distance euclidienne au point le plus proche à la côte en m. Il s'agit d'une mesure de la continentalité. Elle est le résultat de l'application de la fonction EUCDISTANCE du module GRID de ArcInfo au *grid* résultat de la rasterisation du *shapefile* de la ligne de côte à la résolution de 250 m (*source\_grid*).

### 9. XDC1000 et YDC1000

Il s'agit des deux composantes du vecteur unitaire donnant la direction du point le plus proche de la côte. Elles sont exprimées en millièmes et permettent de pondérer l'influence de la continentalité en fonction de la direction. Elles ont été calculées dans le module GRID ArcInfo de la façon suivante :

$$XDC1000 = 1000 * \sin((EUCDIRECTION(source\_grid)) \text{ DIV } \text{DEG}) \quad (6)$$

$$YDC1000 = 1000 * \cos((EUCDIRECTION(source\_grid)) \text{ DIV } \text{DEG}) \quad (7)$$

### 10. LSRi (Land/Sea surface Ratio)

Ces variables expriment l'influence de la proportion entre les surfaces des terres émergées et de mer dans un voisinage circulaire de rayon de longueur  $i$ ,  $i$  prenant les valeurs de 5, 10 et 20 km. Ces variables permettent de pondérer, dans le modèle, l'effet de continentalité en fonction de l'importance de la surface des terres émergées par rapport à la surface de mer présente dans le voisinage de chaque station.

La procédure de calcul mise en œuvre sous ArcInfo est la suivante. Un premier masque binaire (*grid\_1*) a été calculé à l'aide de la fonction conditionnelle CON du module GRID et du MNT. Il s'agit d'un masque identifiant les zones émergées (altitude  $\geq 0$ ). Le négatif de ce masque a été calculé à l'aide de la même fonction pour identifier les zones de mer (*grid\_2*). Ensuite, la fonction FOCALSUM a été appliquée à chacun de ces masques pour obtenir dans le rayon désiré les surfaces (en pixels) cou-

vertes par les terres émergées et la mer. L'opérateur arithmétique « / » a ensuite permis d'obtenir le rapport désiré par division de la première image obtenue par la seconde.

## III. MODÈLE STATISTIQUE DE PRÉDICTION DES PRÉCIPITATIONS

Le modèle statistique de prédiction des précipitations en fonction des variables explicatives est un modèle linéaire du type suivant :

$$P\_i\_50P = \sum_{j=1}^r (\alpha_j * Vp_j) + C + e(x, y) \quad (8)$$

$P\_i\_50P$  50° centile ou médiane du mois  $i$  avec  $i$  variant de 01 à 12, janvier à décembre ;

$$\sum_{j=1}^r (\alpha_j * Vp_j) + C \quad \text{composante déterministe,}$$

où  $j$  indice des variables prédictives significatives variant de 1 à  $r$ ,  $r$  étant le nombre de variables introduites,  $\alpha_j$  coefficient de régression de la variable prédictive retenue en ordre  $j$ ,  $Vp_j$  variable prédictive retenue en ordre  $j$ ,  $C$  constante ;  $e(x, y)$  erreur d'ajustement (résidu) – il s'agit en fait de la composante stochastique qui est fonction de  $X$  et  $Y$  -  $c$ 'est sur ce résidu que portera l'interpolation.

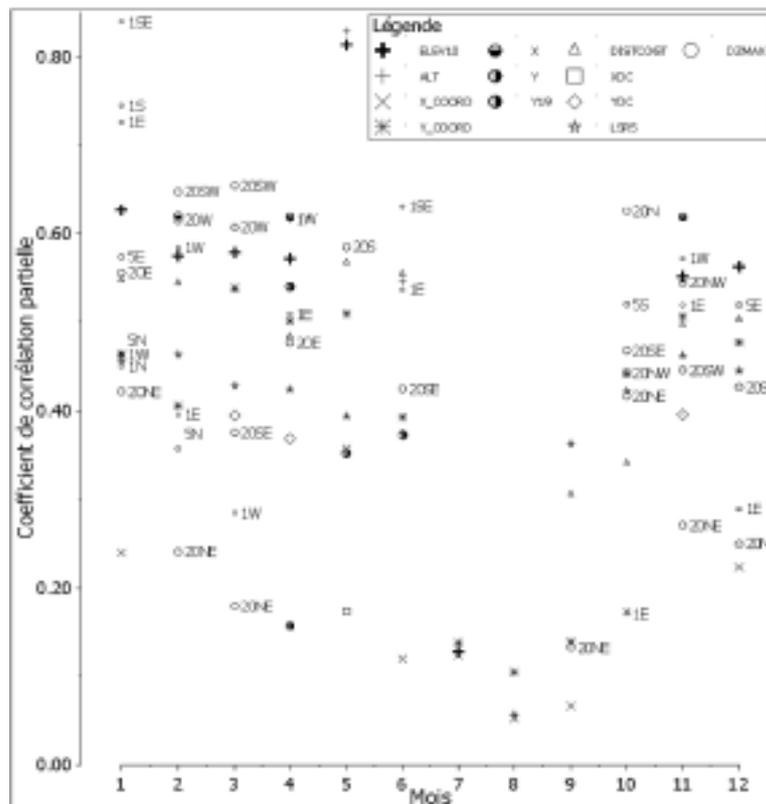
Ce modèle a été déterminé par régression incrémentielle ascendante.

La procédure adoptée est la même pour chaque mois étudié individuellement. La règle de décision pour l'introduction des variables est basée sur la part décroissante de la variance expliquée par chaque nouvelle variable. Les variables pour lesquelles le rapport  $F$  de Fischer passe sous le seuil de tolérance d'introduction de 0.1, ne sont pas introduites. Sur base du test  $t$  de Student, seules les variables significatives au niveau 5 % ont été retenues.

Cette procédure a été exécutée de façon itérative sur les seules variables significatives retenues à l'étape précédente, jusqu'à ce que toutes les variables retenues soient significatives. Le modèle obtenu pour chaque mois, l'ordre d'introduction des variables, la part de la variance expliquée par le modèle, le test de signification du coefficient de corrélation multiple et l'erreur standard d'estimation sont reportés dans le tableau 1. La figure 4 montre les valeurs des coefficients de corrélation partielle de chacune des variables introduites pour chaque mois.

**Tableau 1.** Paramètres et significativité statistique des différents modèles d'estimation des précipitations mensuelles

MOIS	Janvier	Février	Mars	Avril	Mai	Juin	Juillet	Août	Septembre	Octobre	Novembre	Décembre
INTERCEPT	-688.18	-533.84	-873.37	-952.30	-307.65	-212.24	-190.13	-312.04	-381.22	-542.71	-642.87	-783.92
ALT					0.0082(7)	0.0107(3)						
X_COORD	0.000086(6)				0.000015(4)	0.000023(2)	0.000031(1)	0.000058(1)	0.000052(1)	0.000128(1)		0.000115(3)
Y_COORD	0.000166(7)	0.000134(5)	0.000215(5)	0.000233(1)	0.000076(2)	0.000050(1)	0.000044(2)	0.000071(2)	0.000095(2)	0.000132(6)	0.000164(6)	0.000189(7)
X		0.038469(10)		0.030604(10)							0.037491(11)	
Y				0.017478(9)								
Y19				-0.090824(5)	-0.023105(9)	-0.031722(6)						
DISTCOST	-0.000993(2)	-0.000964(2)	-0.000680(3)	-0.000370(3)	-0.000140(6)	-0.000148(5)			-0.000198(4)	-0.000649(3)	-0.000698(4)	-0.001309(2)
XDC1000					-0.001423(8)							
YDC1000			-0.004169(9)	-0.002449(7)								-0.005893(8)
LSR5	52.57(3)	44.90(4)	35.04(4)	25.34(4)	5.65(5)			6.74(3)	13.48(5)	39.84(5)	39.97(5)	58.26(4)
DZMAX1N	-0.08897(9)											
DZMAX1W	0.10926(5)	0.09402(6)	0.05200(7)	0.06088(8)							0.05667(3)	
DZMAX1S	-0.13975(11)											
DZMAX1SE	0.16938(12)					-0.02658(7)						
DZMAX1E	-0.13745(13)	-0.06972(9)		-0.03001(11)		0.02478(8)				-0.05584(9)	-0.05233(10)	-0.06470(8)
DZMAX5N	0.03773(8)	0.01351(11)										
DZMAX5S										0.01814(10)		
DZMAX5E	0.02772(14)											0.02198(9)
DZMA20N										0.01311(8)		
DZMA20NW										0.00926(2)	0.00823(2)	
DZMA20W	-0.01246(8)	-0.00870(8)										
DZMA20SW	0.01451(3)	0.01705(1)									0.01067(1)	
DZMA20S					0.00347(3)							
DZMA20SE			-0.01011(6)			0.00328(4)				-0.01116(7)		-0.01980(6)
DZMA20E	-0.01791(4)			-0.004430(6)								
DZMA20NE	-0.01204(10)	-0.00984(7)	-0.00591(10)						-0.00763(3)	-0.02041(4)	-0.01250(7)	-0.01755(5)
ELEV10	0.05678(1)	0.06405(1)	0.05601(2)	0.03372(2)	0.01405(1)		0.00462(3)				0.02346(9)	0.06400(1)
Nombre var.	14+1	11+1	10+1	11+1	9+1	8+1	3+1	3+1	5+1	10+1	11+1	9+1
R <sup>2</sup>	0.658	0.773	0.794	0.794	0.748	0.557	0.418	0.493	0.403	0.586	0.623	0.644
F	19.4	44.7	56.0	50.4	48.1	23.1	36.4	49.2	20.2	20.5	21.6	29.4
p	<0.0000	<0.0000	<0.0000	<0.0000	<0.0000	<0.0000	<0.0000	<0.0000	<0.0000	<0.0000	<0.0000	<0.0000
ESE abs.	18.7	12.7	10.6	7.9	4.3	4.4	3.7	4.8	7.7	15.6	13.0	17.4
ESE rel.	0.21	0.16	0.16	0.16	0.15	0.41	0.51	0.32	0.18	0.18	0.17	0.18
N	156	156	156	156	156	156	156	156	156	156	156	156



**Figure 4.** Coefficients de corrélation partielle entre les variables explicatives du modèle calculé et le volume de précipitations médian pour chaque mois de l'année  
 Variables retenues par régression incrémentielle ascendante ; nombre de variables explicatives ; coefficients de régression et ordre d'introduction des variables (entre parenthèses) ; coefficient de détermination multiple et test de signification de ce coefficient ; erreur standard d'estimation exprimée de façons absolue et relative (ESE abs. et ESE rel.) ; nombre de stations prises en compte lors de l'ajustement.

#### IV. DISCUSSION DES MODÈLES CALCULÉS

La part de la variance expliquée par le modèle de chaque mois ( $R^2$ ) montre que les estimations sont nettement meilleures pour les mois de février à mai. La composante stochastique est donc moins importante pendant cette période. Par contre, les variables explicatives retenues expliquent généralement une part un peu plus réduite de la variance pour le début de l'été (juin), la fin de l'automne et l'hiver (novembre à janvier). Enfin, pour les mois estivaux, la composante stochastique est beaucoup plus importante.

Ces constatations se traduisent sur les erreurs standard d'estimation absolue et relative des précipitations par la seule prise en compte de la composante déterministe de cette modélisation. En effet, l'erreur relative d'estimation est assez faible en automne, en hiver et au printemps (15 et 21 %). Par contre, elle est beaucoup plus importante en été (32 à 51 %).

Par ailleurs, les modèles obtenus mettent en évidence l'existence d'une grande variabilité du nombre et des types de facteurs expliquant la distribution spatiale des précipitations en fonction du mois.

Ainsi, on constate que la tendance spatiale existe pour tous les mois de l'année. Elle est caractérisée par une augmentation des précipitations du sud vers le nord ou du sud-ouest vers le nord-est selon le mois. L'analyse des coefficients de corrélation partielle (figure 4) permet par ailleurs d'affirmer qu'après élimination de l'effet des autres facteurs pour chaque modèle mensuel, la composante méridienne de cette tendance est toujours plus importante que l'autre. Cette différence d'importance est cependant très faible pendant les 3 mois estivaux (juillet, août et septembre).

L'effet de l'altitude locale n'est déterminant que pour estimer les précipitations des mois de mai et juin. Pendant ces mois, on constate une augmentation des précipitations de l'ordre de 0,8 à 1 mm/hm. On peut même affirmer qu'indépendamment des autres facteurs, cet effet est déterminant pour ces deux mois (figure 4). Quant à l'effet de l'altitude moyenne de la région, il n'est important que si on considère l'altitude moyenne dans un rayon assez grand (10 km). Cet effet est pris en compte pour tous les mois, sauf juin, août, septembre et octobre. Il est toujours plus marqué en hiver et, en général, la corrélation partielle de cette variable avec le volume médian de précipitations est toujours très importante par rapport aux autres variables explicatives retenues.

L'orientation locale influence les précipitations pendant les mois hivernaux, sauf en décembre, et le mois d'avril. On constate que les précipitations sont plus importantes si la station est exposée à l'ouest et au SSW. L'orienta-

tion générale de la région dans laquelle se trouve la station joue un certain rôle entre avril et juin et les précipitations sont plus faibles, si le versant est exposé au S. On constate par ailleurs que c'est la composante est-ouest de l'orientation locale qui a l'effet le plus marqué lorsqu'elle est prise en compte (figure 4). Après élimination de l'influence des autres facteurs, cet effet est particulièrement important par rapport à l'ensemble des autres variables retenues pour les 3 mois (février, avril et novembre) pour lesquels cette variable est prise en compte. En ce qui concerne la composante méridienne de l'orientation, elle joue plus un rôle à une échelle spatiale plus vaste qu'à l'échelle locale.

Les variables de modélisation de la continentalité jouent un rôle sur les précipitations pendant toute l'année, avec, cependant, une nette réduction d'influence en été (de juin à août). Ainsi, l'augmentation de la distance à la côte se traduit par une diminution des précipitations toute l'année, sauf en juillet et août. Cette diminution est de l'ordre de 0,20 mm/km en septembre. Elle croît fortement en octobre, puis plus faiblement et régulièrement jusque décembre, quand elle atteint un maximum de 1,31 mm/km. Ensuite, elle décroît régulièrement jusque mai-juin pour atteindre sa valeur minimale de 0,14 mm/km. Incontestablement, cette distance à la côte est la variable la plus importante des 4 variables permettant de modéliser l'effet de continentalité (figure 4). La direction dans laquelle se trouve cette côte est la variable la moins importante des 4. Ainsi, la direction dans laquelle la côte la plus proche se trouve ne joue donc qu'un rôle mineur de mars à mai et en novembre. Quand la côte la plus proche est au nord, on constate d'ailleurs, en mars, avril et novembre, une décroissance des précipitations. En mai, c'est plutôt le fait de se trouver plus près d'une côte située à l'est qui entraînerait une diminution des précipitations. Enfin, la dernière variable permettant de modéliser la continentalité est le rapport terre/mer dans un rayon de 5 km (*LSR5*). Si l'on considère sa relation avec la variable à expliquer après élimination des autres facteurs (figure 4), la variable *LSR5* a un effet un peu moins important que la distance à la côte la plus proche. Elle influence les précipitations tous les mois de l'année sauf en juin et juillet. On constate que l'accroissement de ce rapport *LSR5* se traduit par des précipitations plus importantes. Cette variable est absente du modèle en juin et juillet et son effet est très faible en août.

En ce qui concerne l'effet « barrière » produit par le relief, on constate que ce sont plutôt les précipitations des mois hivernaux et, dans une moindre mesure, printaniers et automnaux qui subissent cette influence. Généralement, ce sont des effets qui se manifestent à deux échelles spatiales à l'exception des mois de janvier, février, octobre et décembre. Il s'agit d'un effet à petite distance (1 km) et d'un effet à très grande distance (20 km). Pour

les 4 mois cités, une échelle intermédiaire (5 km) joue aussi un certain rôle. En ce qui concerne la direction dans laquelle se trouve la barrière, on constate qu'à petite distance et, pour le mois de janvier, une barrière se trouvant au N, au S ou à l'E, entraîne une diminution des précipitations. Par contre, quand la barrière se trouve à l'W ou au SE, c'est l'inverse. L'effet des barrières situées à l'W et à l'E se remarque aussi pour les autres mois hivernaux, printaniers et automnaux. En ce qui concerne cet effet à plus grande distance, on constate que ce sont aussi les mois hivernaux, de fin d'automne et de début de printemps qui se distinguent par la diversité des influences selon la direction dans laquelle se trouvent les barrières. Les secteurs W et SW sont pris en compte dans les modèles de février, mars et novembre. Les secteurs SE à NE le sont dans les modèles de janvier à avril et de septembre à décembre. En octobre et novembre, les secteurs N et NW interviennent aussi. Les mois de mai et juin sont des mois de transition pour lesquels l'effet « barrière » joue un rôle qui reste plus important que les autres facteurs (figure 4), mais la directionnalité de cet effet est plus nettement marquée puisque seuls les secteurs E, SE et S sont retenus. Si l'on considère les coefficients de corrélation partielle entre ces variables  $DZMAX_i$  et la variable à expliquer (figure 4), on peut constater que l'effet des barrières situées à l'E, au SE et au S est également très marqué en janvier. En février, mars et avril, ce sont plutôt les barrières situées à l'ouest et au sud-ouest qui ont une influence plus prononcée. En octobre, novembre et décembre, l'effet « barrière » est plus variable selon la direction et, à l'exception du mois d'octobre, il est moins prédominant par rapport à d'autres facteurs.

## V. COMPOSANTE STOCHASTIQUE

Après ajustement du modèle statistique, la composante stochastique est connue sous forme de résidu en chaque station. Les résidus ont été soumis à une procédure d'interpolation de type krigeage par point qui, contrairement au krigeage par bloc, ne permet pas d'estimer la variabilité locale (Isaaks et Srivastava, 1989). Aucune dérive spatiale (krigeage ordinaire) n'a été envisagée, puisque celle-ci a été prise en compte par les variables de localisation testées par la procédure de régression incrémentielle ascendante. La première étape de cette procédure est le calcul d'un semi-variogramme expérimental et l'ajustement d'un semi-variogramme modélisant la relation entre semi-variance et décalage spatial (*lag*). Ce semi-variogramme permet de pondérer l'influence d'une donnée dans l'interpolation en un point sur base de la distance entre cette donnée et ce point, d'une part, et de la continuité spatiale de la variable étudiée, d'autre part. Deux conditions doivent être vérifiées pour effectuer l'interpolation : la représentativité spatiale des données, testée à l'aide du paramètre

statistique R de Clark et Evans (Collet, 1992, Reich & Davis, 2000), et la continuité spatiale.

En ce qui concerne la représentativité spatiale, le paramètre R, compare, pour des ordres de voisinage  $k$  allant de 1 à 20, la distance moyenne entre les points et leur voisin d'ordre  $k$  à une distance théorique qui serait observée si la distribution était aléatoire. Nos tests ont montré que les valeurs du rapport R, sont toujours significativement supérieures à 1. La distribution de nos stations doit donc être considérée comme dispersée. Ce type de distribution n'autorise l'interpolation que si la continuité spatiale de la variable soumise à cette interpolation est vérifiée. La construction des semi-variogrammes a mis cette continuité en évidence.

Les modèles de semi-variogrammes adoptés présentent les caractéristiques détaillées au tableau 2. Dans le voisinage proche, on constate que les semi-variogrammes modélisés sont toujours quasiment linéaires. La plus nette différence entre les mois de l'année, est la dimension du voisinage dans lequel cette linéarité est observée. Une seconde spécificité observée pendant le mois d'avril est la forme ondulatoire du semi-variogramme (figure 5) qui traduit probablement l'influence d'un facteur plus global sur l'explication de la variation spatiale des résidus peut-être partiellement lié à la circulation atmosphérique de cette période de l'année. Ce type de structure est également perceptible, mais de façon moins nette, sur le semi-variogramme expérimental du mois de mars. Ces mois de mars et avril sont aussi parmi les mois pour lesquels la corrélation spatiale dans le voisinage très proche est la plus faible et varie le plus lentement dans l'espace jusqu'à une distance de l'ordre de 40 km. Cette variation est beaucoup plus rapide pour les mois de mai à septembre en étant très forte dans le voisinage très proche et en s'annulant à environ 20 km de distance. Pour les mois suivants jusqu'à décembre, la variation spatiale de cette corrélation est plus lente que pour les 4 mois précédents, mais cette corrélation y est beaucoup plus importante dans le voisinage très proche que dans les cas des mois de mars et avril et elle se prolonge jusqu'à une distance comparable à celle de ces deux mois, soit 40 km. Sa variation spatiale y est donc généralement plus rapide en restant linéaire à proximité, mais en s'incurvant souvent à partir d'une vingtaine de km.

Les cartes de pluviométrie obtenues pour les mois représentatifs de chaque saison (figure 6) montrent, d'une part, les précipitations attendues après correction de la composante déterministe par la composante stochastique (fond de carte coloré) et, d'autre part, la valeur de correction correspondant à la seule composante stochastique (données ponctuelles et isolignes). On observe sur ces cartes que la composante stochastique prend une plus grande importance pour les mois typiquement estivaux et, par contre, qu'elle est très faible pour les mois typi-

Tableau 2. Caractéristiques structurales des semi-variogrammes modélisés

	Type de fonction	Angle d'anisotropie	Longueur du grand axe de l'ellipse d'anisotropie	Rapport entre grand et petit axe de l'ellipse d'anisotropie	Échelle de semi-variogramme (mm <sup>2</sup> )	Puissance du semi-variogramme	Effet pépité	Erreur (mm <sup>2</sup> )
Janvier*	Linéaire	0.00		1.00			Effet pépité	312,80
Février*	Linéaire	0.00		1.00			Effet pépité	149,00
Mars	Circulaire	90.00	51600	2.00	35,61		Effet pépité	71,51
Avril	Ondulatoire	153,90	11530	1,44	20,28		Effet pépité	37,91
Mai	Exponentiel	169	7385	1,60	17,86			
Juin	Circulaire	69,29	37400	2.00	19.00			
Juillet	Linéaire	165.90					Effet pépité	1.00
Aout	Exponentiel	159,70	46800	2.00	26.00		Effet pépité	3,54
Septembre	Circulaire	8,44	25230	1,57	58,02			
Octobre	Exponentiel	174,20	11650	1,56	224,2			
Novembre	Puissance	150,10	40000	2.00	138.00	0,129		
Décembre	Exponentiel	90.00	14570	1,01	223,90		Effet pépité	68,11

\* Pour les mois de janvier et février, la semi-variance est constante en fonction du décalage spatial (*lag*), l'interpolation sur base de tous les points du set de données fournit donc une composante stochastique constante pour tous les points de l'espace considéré.

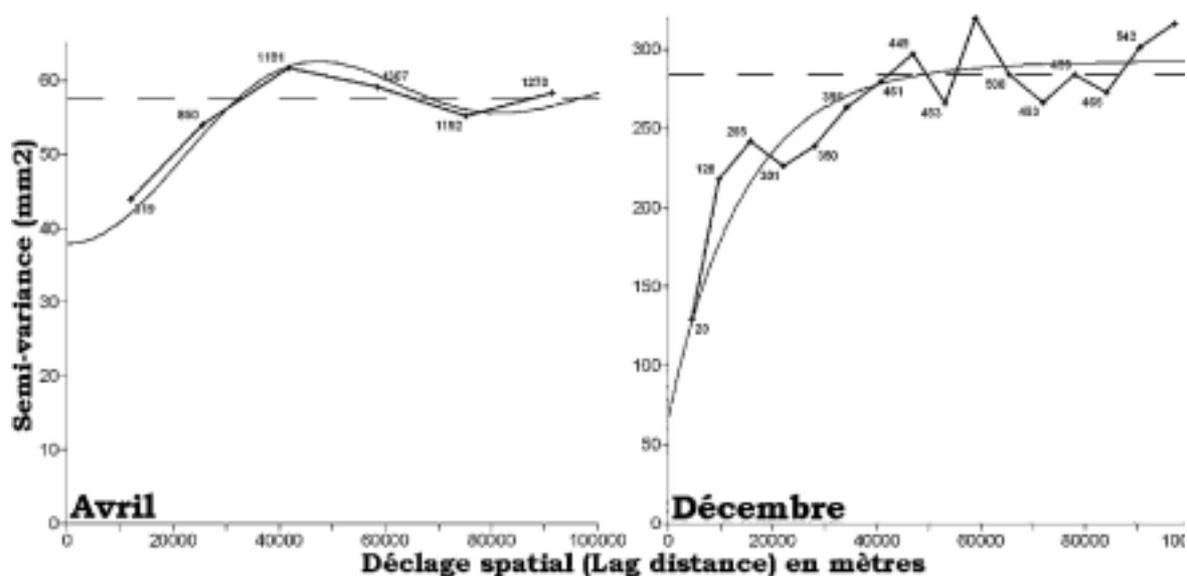


Figure 5. Deux exemples de semi-variogrammes expérimentaux et modélisés

**Tableau 3.** Validation de la modélisation RMSe (mm) : erreur quadratique moyenne exprimée en valeur absolue en mm de précipitations ; RMSe rel. (%) : erreur quadratique moyenne exprimée en valeur relative (%) par rapport à la valeur de la moyenne des précipitations médianes du set de données de validation

MOIS	RMSe (mm)				RMSe rel. (%)			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
1	15,7	15,8	21,7	35,5	17,2	17,3	23,8	47,6
2	12,1	12,1	14,6		15,9	15,9	19,2	
3	12,9	11,4	14,5	16,8	19,6	17,3	22,1	29,3
4	9,4	8,7	8,4		18,4	16,9	16,5	
5	4,9	4,6	4,5		17,6	16,7	16,1	
6	4,7	4,5	3,7	7,5	47,1	44,8	37,0	21,7
7	2,9	2,4	2,5		42,4	35,5	37,6	
8	4,3	3,8	3,4		28,9	25,5	23,0	
9	9,7	8,3	8,8	21,1	22,6	19,2	20,3	31,7
10	23,2	21,6	20,6		26,6	24,8	23,7	
11	16,8	15,8	15,2		20,8	19,5	18,8	
12	19,9	16,6	18,4	35,5	20,6	17,2	19,0	47,6

- (1) Notre modèle avant introduction de la composante stochastique.  
(2) *Idem* mais après introduction de la composante stochastique.  
(3) Interpolation par krigeage sur les données brutes.  
(4) Modèle développé par Agnew & Palutikof (2000).

quement hivernaux. Les autres mois présentent des caractéristiques intermédiaires.

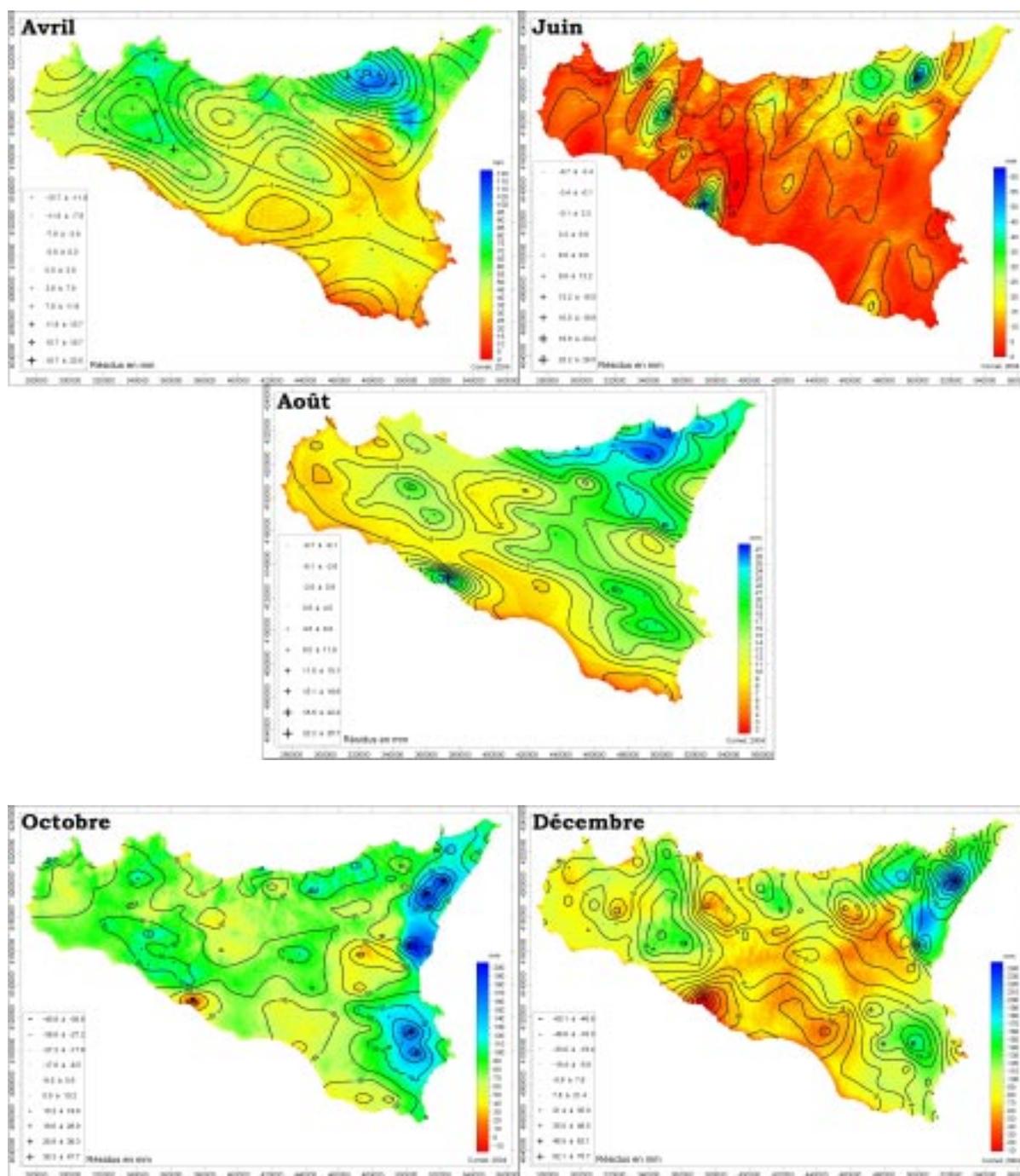
## VI. VALIDATION DU MODÈLE STATISTIQUE

Une validation du modèle obtenu pour chaque mois a été entreprise indépendamment de l'ajustement. Parmi les 176 stations pluviométriques, 20 (figure 1) ont été sélectionnées par échantillonnage aléatoire stratifié pour effectuer cette validation. La stratification a été réalisée en fonction du résultat de la classification discutée au § II.B. En pratiquant de la sorte, nous n'avons pas adapté notre validation à la variabilité des modèles obtenus pour chaque mois et à la variabilité des environnements rencontrés sur base des variables explicatives, car seules la variable dépendante et sa variabilité temporelle au cours d'une année moyenne ont été prises en compte dans la stratification.

La comparaison des valeurs de précipitations observées et calculées pour ces 20 stations en exploitant uniquement la composante déterministe du modèle statistique, d'une part, et cette composante déterministe corrigée de la composante stochastique, d'autre part, a permis d'estimer une erreur quadratique moyenne (RMSe) exprimée en grandeurs absolue et relative. Le tableau 3 montre les résultats de cette validation. La valeur de la RMSe relative a été estimée par rapport à la valeur moyenne des précipitations médianes du mois considéré sur l'ensemble des stations de validation.

Dans ce tableau, nous avons également reporté la RMSe correspondant à une interpolation par krigeage des valeurs de précipitations médianes brutes de Sicile, ainsi que celle du modèle de Agnew et Palutikov (2000) qui ont analysé les précipitations moyennes à d'autres échelles spatiale et temporelle. La technique adoptée pour l'interpolation sur les données brutes est le krigeage ordinaire par point en utilisant des semi-variogrammes modélisés linéaires avec un effet « pépite ». Cet effet produit une interpolation non-exacte et un lissage de la surface obtenue et peut être expliqué par le fait que l'on exploite des normales climatiques mensuelles qui sont affectées d'une variance traduisant une variabilité au cours du mois de chaque année et une variabilité inter-annuelle. Il est néanmoins surprenant de constater que cet effet est moins présent lorsque l'on analyse la continuité spatiale de la composante stochastique du modèle statistique développé.

Si, dans une première étape d'analyse, on compare la composante déterministe de ce modèle à sa composante stochastique, on constate que l'amélioration des résultats après introduction de la composante stochastique est très faible. Dans le meilleur des cas, elle atteint 3,4 % et, exceptionnellement 6,9 %, en juillet par exemple. Pour janvier et février, cette amélioration est nulle, on peut même dire que la prise en compte de la composante stochastique produit des résultats moins exacts. Ceci est probablement lié à la constance de la semi-variance des résidus en fonction du décalage spatial.



**Figure 6.** Résultat de la modélisation de la distribution spatiale des précipitations en Sicile pour les mois les plus représentatifs de chaque saison : avril, juin, août, octobre et décembre

L'image colorée représente la composante déterministe corrigée des résidus interpolés (composante stochastique). Une même palette de couleur a été appliquée à toutes les situations. Les précipitations y sont exprimées en mm.

Les symboles affectés aux points correspondant à chaque station représentent les résidus.

Les isolignes représentent les courbes d'égal résidu (composante stochastique). Ces courbes ont été tracées à partir de l'image obtenue par interpolation sur les résidus calculés en chaque station et exprimés en mm.

Si, dans une seconde étape d'analyse, on compare les résultats de notre modèle statistique à l'interpolation sur les données brutes, on constate qu'ils sont légèrement à nettement plus exacts que la spatialisation par interpolation sur les données brutes pour les mois de décembre à mars, ainsi que pour juillet et septembre. Par contre, les précipitations estimées par notre modèle statistique sont moins exactes que celles estimées par la seconde approche pour les autres mois et plus particulièrement pour juin et octobre.

Enfin, si nous comparons les valeurs absolues, exprimées en mm, des RMSe obtenues en Sicile et des RMSe de la modélisation de Agnew & Palutikof (2000), on constate qu'elles sont plus faibles, sauf pour le mois d'octobre. En valeurs relatives, les RMSe de nos modèles sont beaucoup plus petites que celles calculées par ces auteurs pour les mois automnaux, hivernaux et printaniers, mais pas pour les mois d'été. Cependant, il faut signaler que, dans la plupart des cas, la variance des données utilisées par Agnew & Palutikof est supérieure à celle de nos données.

## CONCLUSION

Dans notre recherche, nous avons tenté d'améliorer les techniques de spatialisation des précipitations en Sicile par la construction d'un modèle statistique permettant d'estimer les précipitations médianes mois par mois en fonction de facteurs physiographiques. Une attention particulière a été portée à la résolution temporelle et à la stratification de la validation de nos résultats. Nous avons exploité des normales climatiques établies sur plusieurs années d'observations. Ainsi, la variabilité interannuelle, qui constitue une des composantes déterminantes de l'évolution spatiale de certains processus partiellement liés au climat comme la désertification, par exemple, a donc été négligée. Par ailleurs, nous n'avons considéré qu'un élément du climat, la pluviométrie. Notre approche méthodologique pourrait néanmoins s'appliquer à d'autres éléments, comme la température, par exemple. Une telle approche permettrait de mieux comprendre les processus affectant certaines régions de l'île en intégrant les résultats de cette recherche à l'analyse des autres facteurs liés au sol et au sous-sol voire aux actions anthropiques.

En ce qui concerne plus spécifiquement la validation de nos résultats, le type de modélisation statistique mis en œuvre ne produit pas nécessairement une plus grande exactitude dans l'estimation du champs continu de la quantité de précipitations par rapport aux interpolations portant sur les données brutes, par exemple. Néanmoins, ce type d'approche permet de mieux identifier les facteurs déterminant les précipitations en fonction du mois considéré. Il s'agit principalement des facteurs qui interagissent intensément avec la circulation atmosphé-

rique générale et qui provoquent ces précipitations.

Parmi ceux-ci, certains traduisent des situations de plus forte sensibilité au processus de désertification. La spatialité de ces facteurs est maîtrisée à une très bonne résolution par rapport aux données climatologiques connues de façon discrète. Cette maîtrise doit être exploitée dans l'analyse de ce type de processus. À ce sujet, il est aussi intéressant d'envisager une analyse de la série temporelle d'images AVHRR de NOAA et une comparaison des indicateurs de l'état sanitaire de la végétation calculés à partir de ces images avec le modèle climatique que nous avons développé.

## REMERCIEMENTS

Nous remercions Mademoiselle Francesca Giordano de l'ENEA (*Ente per le Nuove tecnologie, l'Energia e l'Ambiente*) qui a gracieusement mis le MNT de Sicile à notre disposition. Nous tenons aussi à exprimer toute notre reconnaissance aux *reviewers* de notre article pour ses nombreux commentaires critiques et suggestions.

## BIBLIOGRAPHIE

- AGNEW M. & PALUTIKOF J., 2000. GIS-based construction of baseline climatologies for the Mediterranean using terrain variables, *Climate Research*, 14, pp. 115-127.
- ALEXANDRE J., ERPICUM M., MABILLE G. & CORNET Y., 1998. Précipitations atmosphériques et altitude. Prélude à une cartographie des montants annuels et mensuels en Belgique. *Publications de l'Association Internationale de Climatologie*, 11, pp. 219-226.
- ALEXANDRE J., ERPICUM M. & VANDIEPENBEECK M., 1999. Atlas de Belgique. Planche III.1. Climat, Commentaires au dos. Ed. *Commission de l'Atlas National de Belgique*.
- ASSABA M. & LABORDE J.-P., 2000. La prise en compte du relief dans l'estimation des pluies mensuelles : le cas de la Corse. *Publications de l'Association Internationale de Climatologie*, 13, pp. 149-157.
- BENICHOU P. & LEBRETON O., 1986. Prise en compte de la topographie pour la cartographie de champs pluviométriques statistiques : la méthode Aurelhy ». *Les colloques de l'INRA*, 39, Ed. INRA, Paris, pp. 51-69.
- CARREGAP. & GARCIA E., 2000. Les fortes précipitations quotidiennes en Toscane et leur cartographie. *Publications de l'Association Internationale de Climatologie*, 13, pp. 169-180.
- COLLET C., 1992. *Systèmes d'information géographique en mode image*. Coll. « Gérer l'environnement », Presses polytechniques et universitaires romandes, 186 p.

- DRAGO A., Cartabellotta D., Lo Bianco B. & Lombardo M., 1998. Guida dell'atlante climatologico della Sicilia. CDROM réalisé par l'INTEA S.r.l. pour l'Assessorato Agricoltura e Foreste (Regione Siciliana) – Servizi allo Sviluppo – Unità Operativa di Agrometeorologia.
- FITSCHY C., 2003. *Modélisation de la distribution spatiale des précipitations en Sicile*. TFE, inédit Université de Liège, 95 p.
- GOBBATO O., 2000. L'uso dei sistemi informatici territoriali per lo studio delle aree sensibili alla desertificazione nella provincia di Agrigento. *Tesi di Laurea*, Università degli studi di Roma « La Sapienza », 104 p.
- GOODALE C.L., ABER, J.D. & OLLINGER, S.V.V., 1998. Mapping monthly precipitation temperature and solar radiation for Ireland with polynomial regression and digital elevation model. *Clim. Res.*, 10, pp. 35-49.
- ISSAKS E.H. & SRIVASTAVA R.M., 1989. *An introduction to applied geostatistics*. Oxford University Press, New York, Oxford, 561 p.
- KOSMAS C., KIRKBY M. & GEESON N., 1999. *The Medalus project – Mediterranean desertification and land use*. Office for Official Publications of the European Communities, Luxembourg, 87 p.
- LENNON J.J. & TURNER J.R.G., 1995. Predicting the spatial distribution of climate – temperature in Great-Britain. *J. Anim. Ecol.*, 64, pp. 370-392.
- REICH R.M. & DAVIS R., 2000. *Quantitative spatial analysis (Course notes for NR/ST 523)*. Colorado State University, Fort Collins, Colorado 80523, 540 p.

*Adresses des auteurs :*

Yves CORNET  
Département de Géographie  
Université de Liège, 17 allée du 6-Août (B5)  
B-4000 Liège

Tél : +32 4 3665371

Fax : +32 4 3665693

ycornet@ulg.ac.be

Michel ERPICUM  
Département de Géographie  
Université de Liège, 2 allée du 6-Août (B11)  
B-4000 Liège

michel.erpicum@ulg.ac.be

Catherine FITSCHY  
Licenciée en Sciences géographiques  
Rue E. Grandprez, 46  
B-4020 Liège

Oscar GOBBATO  
Licencié en Géographie  
Université « La Sapienza » de Rome  
Via Monte Cavo 32  
00040 Rocca di Papa (Rm)