Meteorologisk institutt
met.no

Report 28/02

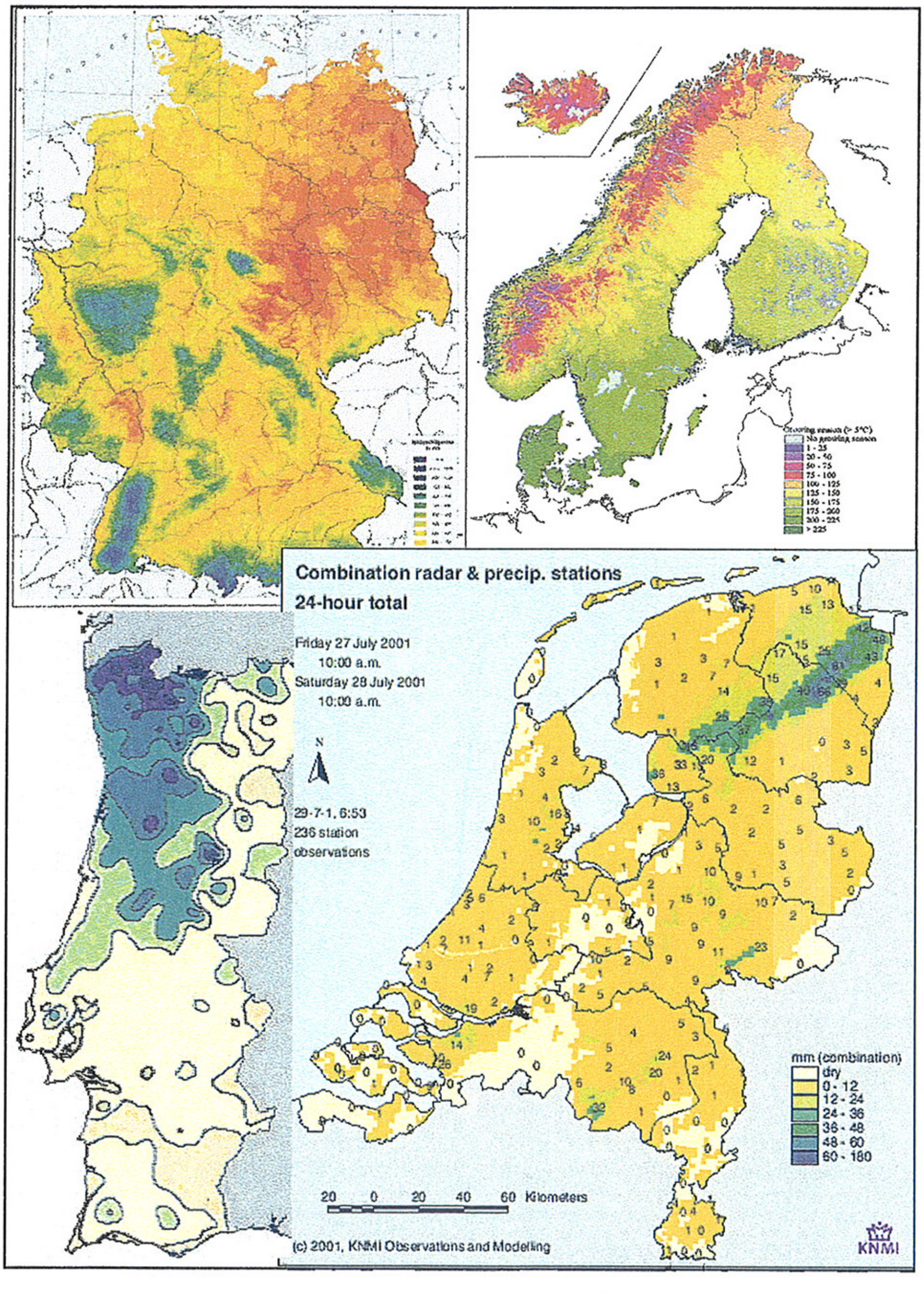
KLIMA

COST719 - The use of geographical information systems in climatology and meteorology

Report no. 1/WG2 Spatialisation/ COST-719

Applications of spatial interpolation of climatological and meteorological elements by the use of geographical information systems (GIS)

Ole Einar Tveito
Wolfgang Schöner
(editors)



ISSN 0805-9918

met.no REPORT

NORWEGIAN METEOROLOGICAL INSTITUTE

P.O. BOX 43 BLINDERN, N - 0313 OSLO

TELEPHONE: (+47) 22 96 30 00

REPORT NO.

28/02 KLIMA

DATE:

27.12.2002.

TITLE:

Applications of spatial interpolation of climatological and meteorological elements by the use of geographical information systems (GIS).

EDITORS:

Ole Einar Tveito (*met.no*) and Wolfgang Schöner (ZAMG)

PROJECT CONTRACTORS:

COST 719 "The use of geographical information systems in climatology and meteorology"

SUMMARY:

Within the frames of COST719, "The use of geographical information systems (GIS) in climatology and meteorology" an inventory on the use of GIS in spatial interpolation application is carried out. The inventory showed that GIS is used as a tool for a number of different spatialisation problems. This report presents different interpolation strategies, and which methods and geographical information these demands. GIS shows to be a good tool for administration of data and presentation of the final maps. However, for climatological and meteorological purposes are the possibilities incorporated in standard GIS software still not sufficient.

KEYWORDS:

GIS, Spatial interpolation

SIGNATURES:

Ole Einar Tveito	**Eirik J.Førland**
SENIOR SCIENTIST	HEAD (ACTING) OF
COST719/WG2 Chairman	*met.no* CLIMATOLOGY DIVISION

Contents

Acknowledgement.

This report is a result of an inventory among the participating countries in COST action 719. The editors would like to acknowledge all contributors to the inventory. Without their kind efforts this inventory could not been accomplished.

Country	Name	Institution
Austria	Wolfgang Schöner	ZAMG
	Reinhold Steinacker	University of Vienna
Belgium	Clemens Mensink	VITO (Flemish Institute of Applied Research)
	Liliane Janssen	VITO
Finland	Jaakko Helminen	FMI
France	Jean-Pierre Ceron	MeteoFrance
Germany	Rolf Ullrich	DWD
Greece	Athanasios Lukas	University of Thessaly
Hungary	Zita Bihari	HMS
Italy	Bernardo Gozzini	IATA – CNR
The Netherlands	Albert Klein-Tank	KNMI
	Jov Verkaik	KNMI
	Wim de Rooy	KNMI
Norway	Ole Einar Tveito	DNMI
Poland	Zbigniew Ustrnul	Institute of meteorology and water management
Portugal	Sofia Moita	Portuguese Meteorological Institute
	Antonio Perdigao	IHERA
Slovenia	Mojca Dolinar	Min. of Environment, Hydrometeorological Institute
	Damijana Kastelec	Agronomy Dep., Univ. of Ljubljana
	Tomaž Vrhovec	Chair of Meteorology, Univ of Ljubljana
Spain	Carlos Almarza Mata	Spanish Meteorological Institute
	Estrella Guttiérrez	Spanish Meteorological Institute
Sweden	Hasse Alexandersson	SMHI
Switzerland	Estelle Grüter	MeteoSwiss
	Christian Häberli	Univ. of Vienna
	Martin Hölzle	Univ. of Zürich
United Kingdom	John Thornes	Univ. of Birmingham.

1. Introduction

The general public and many authorities at various levels ask ever more questions or require more information and facts on various climatological, meteorological, hydrological and environmental issues. They generally require multifaceted, cross-discipline data or knowledge so that quantitative methodologies and tools are necessary to extract, analyse, shape and assess the required information.

Many of these questions require information about climatological or meteorological characteristics at locations where measurements are not available. It is therefore a challenge to climatologists and meteorologists to provide such information by interpolating the values obtained by measurements. This is however a very complicated task, since meteorological and climatological characteristics highly depends upon local conditions (physiography, hydrography, land-use, etc.). Also the representativity of the meteorological observation stations is a matter of concern, especially in regions with complex terrain. By the development of computers, the processing capability has increased considerably. This has again made it possible to develop new approaches within spatial interpolation, utilizing the different available information in a much better way than before.

One of the most powerful aids in this context can be Geographic Information Systems (GIS), which makes it possible to combine easily all necessary procedures for visualisation, integration, flexible combination and storage of information for different data sets. GIS is already widely accepted in the geo-sciences but their extra value for meteorological applications is still under-used. An exception can be made for climatological data that have been recognised as suitable for GIS processing and visualisation.

In COST719, one working group (WG2) consider the problem of spatial interpolation of climatological and meteorological variables. In this area the focus will be on examining interpolation functionality already present in GIS and statistical software and the recognition of gaps. In particular to:

- Study the potential and limitations of existing GIS (interpolation) functionality for spatialisation of meteorological and climate data;
- Compare with other spatialisation algorithms;
- Set up recommendations/specifications for future GIS tools for spatialisation, suitable for meteorological and climate applications, and take these up with industry.

The first topic for WG2 was to carry out an inventory in order to get an overview of existing approaches of spatialisation among the participating countries. The objective was two-fold:

- Which methods are applied, and with which degree of success?
- How these methods are are linked to GIS or to geographical datasets?

This report presents the findings of this inventory. In chapter two, an overview of the principles and different approaches for spatial interpolation are presented. Chapter three and four deal with the results of the inventory, which methods that are actually applied, and how they perform. In chapter five, the results of some of the applications are presented. In the conclusions, the findings are discussed, and recommendations for further development are given. This report is a basis for the further activities within WG2. In the next phase, an objective comparison between different methods and approaches will be carried out, and the recommendation given in chapter four is having this in mind.

2. Overview of spatialisation methods for climatological/ meteorological purposes

2.1. Spatialisation and methods of spatial analysis

Spatialisation is a term not strictly defined in the literature. It stands for a set of methods describing the dependency of neighbouring data of a data set in typically a Cartesian coordinates system. Major aims of spatialisation are data condensation and data visualisation. However, spatialisation can be also seen in the wider context of spatial analysis, which means a transformation to derive new or extended information from existing information. Several different methods of spatial analysis can be distinguished as shown in Table 2.1. In this classification system spatial analysis is seen as transformation between points, lines, areas, networks and surfaces.

Table 2.1 Types of spatial analysis seen in the context of transformation between points, lines, areas, networks and surfaces (translated from Strobl, 1998)

		TO				
		points	**lines**	**areas**	**networks**	**surface**
FROM	**points**	generalization	generate lines	e.gThiessen, allocation	define nodes	interpolation
	lines	resolve lines	generalization	generate areas	define breaks	interpolation
	areas	find centre point	derive perimeter	generalization	generalization	interpolation
	networks	identify nodes	identify breaks	allocation	generalisation	interpolation
	surfaces	point raster	Construct isopleths	define zones	extract network	generalization

In meteorology and climatology spatial information is based almost exclusively on measurements at certain points (e.g. climate stations). From there spatial interpolation that transforms points to surfaces plays a major role in meteorology and climatology. However, new measurement techniques like remote sensed data (radar and satellite data) offer the opportunity of quasi-continuous 2D or 3D data fields.

2.2 Criteria to distinguish spatialisation methods

For selection of an adequate spatialisation method it is very important to consider the methodical background and assumptions behind of each method from the toolbox of possible methods. The following chapter summarizes criteria to distinguish spatialisation methods and introduce some first criteria for selection of different methods.

2.2.1 Data elements type

a) Spatialisation scheme for scalar data (continuous data, discrete data)
b) Spatialisation scheme for vector data

The criterion of data elements type to classify different analysis methods is well known from applied statistics. Several methods of spatialisation are based on statistical concepts, too. From there the criterion of data elements type is also useful for classification of spatialisation methods. In Meteorology and climatology spatialisation of continuous scalar data is of major importance. For spatialisation of vector data (e.g. wind) separation into two scalar data is frequently used. Then these scalar data are interpolated independently by spatialisation schemes for scalar data.

2.2.2 Methodical background

a) Deterministic methods
b) Stochastic methods

Stochastic methods incorporate the concept of randomness that means the interpolated surface is conceptualised as one of many that might have been observed, all of which could have produced the known data points. They often allow the computation of statistical significance of the surface and uncertainty of the predicted values to be calculated. Deterministic methods do not use the probability theory. They offer the opportunity to explain spatial phenomena by physical reasons. In practice the user often prefers deterministic methods because they are conceptual and less abstract. This quality makes the deterministic approach special important.

2.2.3 Extension of region taken into account for interpolation

a) Global interpolation methods.
b) Local interpolation methods

Global interpolators determine a single function that is mapped across the whole region and therefore tends to produce smoother surfaces with less abrupt changes (a change in one input value affects the entire field). They are used when there is a hypothesis about the form of the surface e.g. a trend.

Local interpolators apply an algorithm repeatedly to a small portion of the total set of points (a change in an input value only affects the result within the window). Some local interpolators may be extended to include a large proportion of the data points in the set, thus making them in a sense global. The distinction between global and local interpolators is thus a continuum and not a dichotomy.

2.2.4 Uncertainty of data points

a) Approximate interpolators
b) Exact interpolators

Approximate interpolators are used when there is some uncertainty about the given values. This utilizes the belief that in many data sets there are global trends, which vary slowly, overlaid by local fluctuations, which vary rapidly and produces an error in the recorded values (it is a well-known fact in e.g surveying) . The effect of smoothing will therefore be to reduce the effects of error on the resulting surface. In the contrary, exact interpolators honour the data points upon which the interpolation is based.

2.3 Step by step decisions for selection of APPROPRIATE spatialisation method

STEP 1	Data type determination
STEP 2	To realise if data include barriers (this step is closely related to step 1, barriers can be semi-permeable e.g. weather fronts, or impermeable)
STEP 3	Uncertainty of data values (data quality, requirements for normative values) to choose either exact interpolation or approximation method
STEP 4	Decision of importance of local peculiarities to choose either a local or a global interpolation scheme
STEP 5	Good reasons to work either with a deterministic or a stochastic approach

2.4 Overview of methods

2.4.1 Deterministic methods

Deterministic interpolation techniques cover a wide range of mathematical methods varying from very simple algorithms to sophisticated procedures. Generally, these methods can be grouped in the following subgroups: methods based on use of the data from the one nearest station, methods based on the weighted linear combination of data from neighbour stations, approximate polynomial interpolators (Trend Surface Analysis), exact polynomial interpolators and radial basis functions.

Deterministic methods do not use the probability theory and they do not provide accuracy measures of the predicted (interpolated) values.

2.4.1.1 Methods based on use of the data from the one nearest station

This type of methods include widely used Thiessen polygons (Thiessen, 1911), Delauney triangulation technique etc. Using the Thiessen polygons the value of the closest observation is assigned to each point of the interpolation grid. The Delauney triangulation technique follows the well-known partition of Voronoï. Both techniques are computationally efficient, but each predicted value of the interpolation grid is based on only one measured location. The interpolation result depends strongly on the spatial distribution of the measurements and their spatial representativeness. Thiessen polygons technique was widely used by hydrologists for spatial interpolation of precipitation.

2.4.1.2 Methods based on the weighted linear combination of data from neighbouring stations

In this type of methods, the interpolated value is obtained from a linear combination of the values measured at neighbouring locations. The weights of the linear combination usually depend on the distances between the interpolated grid point and the measurement locations. The choise of the search neighborhood for interpolation plays an important role for this kind of methods. The Inverse Distance Weighting method (IDW) is one of the most popular example for this group of methods. Barnes (1973), Cressman (1959) and Shepard (1968) have made some modifications of this method for use in meteorology.

IDW (Inverse Distance Weighting)

Short description:
The weights of linear combination are proportional to the inverse of the distance between the interpolated and measured points, and are normalised making their sum over all stations within the search neighborhood equal to 1. IDW is an advanced nearest neighbours approach, allowing a number of neighbouring stations to be included in the estimation of interpolated value. The closer the stations the higher is the weight. The "cut-off" criteria may be either maximum distance from the interpolated point or maximum number of points within the search neighborhood. The latter is the most common. The weight of distance can be also raised to a weighting power p that has in general values ranging from 0 to 2.

Shepard (1968) extended the IDW method by using not only the distance but also the direction between neighbouring points.

Mathematical background:

$$\hat{z}(\mathbf{s}_0) = \frac{\sum_{i=1}^{n} d_{0,i}^{-p} z(\mathbf{s}_i) I(d_{0,i} \leq r)}{\sum_{i=1}^{n} d_{0,i}^{-p} I(d_{0,i} \leq r)}$$

where:

$\hat{z}(\mathbf{s}_0)$ interpolated value at location $\mathbf{s}_0$;

$z(\mathbf{s}_i)$ measured value at locations $\mathbf{s}_i$;

$d_{0,i}$ Euclidean distance between the i-th measured location and location $\mathbf{s}_0$;

$I(d_{0,i} \leq r)$ indicator function with value 1 if $d_{0,i} \leq r$ and with value 0 if $d_{0,i} > r$, where r denotes the radius of search neighborhood. The search neighborhood is the region around the predicted location, from which measured values are taken into account for spatial interpolation.

These predictions are simple, requiring no knowledge of spatial model parameters. The predictor is not resistant to outliers but it is an exact interpolator.

Fundamental assumptions:
IDW is based on the assumption that difference between the spatial data depends on the distance between two locations.

Measure of success of interpolation:
There is no such measure for IDW. Cross-validation can be used to find out the appropriate size of search neighborhoods (or the appropriate number of measured locations to be considered).

Implementation in GIS-software:
IDW is the method offered by all high level GIS software products. The extended method of Shepard is not implemented in major GIS software products (but in e.g. Golden Software Surfer).

2.4.1.3 Approximate polynomial functions (Trend Surface Analysis, TSA)

Polynomials are fitted to the sample points using e.g. the least squares adjustment (one polynomial is fitted to the entire surface). In general approximate polynomial functions are global interpolators. The trend surface (global trend) can be approximated by first to n-th order polynomials. By choosing the order of the polynomial the structure of surface can be affected (a polynomial of order n gives at most n-1 maxima and minima). In the simplest case trend surface analysis uses simple or multivariate linear regression to obtain a surface that gives the best fit to the measured data. This technique is usually applied in the first step of spatial data analysis as well as for de-trending purposes. Because of the physical background of meteorological and climatological variables, different explanatory variables in the multivariate regression model, for instance geographical coordinates, elevation, aspect, slope, distance to coast etc. can be used. Important to note here is that regression models, which have some physically based reasons and not just use the best fit between predicted and measured values are used.

Many climate elements show dependencies on elevation especially for the climate means, which could be explained either by physically reasons (e.g. air temperature, solar radiation) or as a result of frequency statistics (e.g. cloudiness, sunshine duration). Trend surface analysis offers a good possibility for spatialisation of monthly and annual values of various climate elements. However, in many cases this method fails for spatialisation of higher time scale of climate data (e.g. daily values).

Example of Multivariate Linear Regression method

Short description:
The purpose of linear regression is to express the relation between a predicted (interpolated) variable (predictand) and one or more explanatory variables (predictors). In its simplest form it is used to fit a straight line through points scattered in a plane. The linear regression models are often used in a global interpolation scheme.

By considering some statistical assumptions about the probability distribution of the interpolated variable ($Z(\mathbf{s})$), the standard error of regression model can be calculated and the inference about the regression parameter and about the predicted (interpolated) values can be done. In this case the regression model is not deterministic any more, but it is stochastic and we can calculate the prediction accuracy. In the case of meteorological or climatological variables stochastic residuals from the regression model are not independent between each other (serial correlation).

Mathematical background:

$$Z(\mathbf{s}) = \beta_0 + \beta_1 x_1(\mathbf{s}) + \beta_2 x_2(\mathbf{s}) + \ldots + \beta_p x_p(\mathbf{s}) + \delta$$

$Z(\mathbf{s})$ is a random variable (if the assumption of normal probability distribution of δ is assumed) at the location $\mathbf{s}$, $x_1(\mathbf{s})$, $x_2(\mathbf{s})$,..., $x_p(\mathbf{s})$ are explanatory variables, β_0, β_1, β_2,..., β_p are coefficients of linear combination, δ is a random error with normal distribution $N(0,1)$. In the contrast to the stochastic geostatistics methods (mentioned later in this text), the errors are assumed to be independent. Concerning deterministic model the δ and its probability distribution are not important.

Fundamental assumptions:
If we are using multivariate regression on a deterministic manner, then the only assumption should be that the regression model could be interpreted on the basis of physical reasons. In the case of a stochastic model we have to assume the identity normal distribution of δ and their spatial independence.

Using linear regression we must take care that only interpolations are allowed from the theoretical point of view, in practice it happened very quickly that we calculate extrapolations (predictions for the values of explanatory variables which are out of the observational (sample) range). On the basis of statistical theory we cannot say anything about accuracy of extrapolations.

Measure of success of interpolation:
In the case of stochastic regression model, success can be estimated by explained variance of predicted variable and by regression standard error. The regression standard error can be used for the calculation of predictions confidence intervals.

Implementation in GIS-software:
Linear Regression has to be elaborated with a standard statistical program and can be easily computed by map calculator functions.

2.4.1.4 Exact polynomial functions

Short description:
For exact polynomial interpolators many polynomials are fitted either to regions with overlapping neighbourhood (local interpolation) or to the entire region of investigation (global interpolation e.g. Splines). Polynomial functions offer a wide field of interpolation methods. Most popular is the group of spline functions due to its wide range of applications. 1-dimensional, 2-dimensional and 3-dimensional splines can be distinguished. Major differences between different spline functions are based on different smoothing algorithms to ensure smooth surfaces. If this algorithm is not used, the resulting surfaces will tend to have a strongly oscillating pattern in order to fulfil the criteria of exact interpolation. A very powerful polynomial interpolator is for instance the Akima interpolation (Akima, 1970), which belongs to the best interpolators of orography. It works with partial interpolating cubic polynomials with constant transitions of first order.

Mathematical background (example of a Spline function, ESRI 1997):

$$Z_j^* = T_j + \sum_{i=1}^{n} \lambda_i R(h_{ij})$$

with

$$T_j = a_1 + a_2 x_j + a_3 y_j$$

$$R(h_{ij}) = \frac{1}{2\pi}\left\{\frac{h_{ij}^2}{4}\left[\ln\left(\frac{h_{ij}}{2\tau}\right) + c - 1\right] + \tau^2\left[\kappa_0 \frac{h_{ij}}{\tau} + c + \ln\left(\frac{h_{ij}}{2\pi}\right)\right]\right\}$$

where: τ = weight parameter which produces smoother surfaces;

κ_0 = modified Bessel function;
c = a constant (0.577215);
$a_1,..a_3$ = coefficients calculated by solving a linear equation system.

Fundamental assumptions:
It is assumed that measurements are without error

Measure of success of interpolation:
Cross validation can be used as a measure of success.

Implementation in GIS-software:
Spline interpolation is offered by most high level GIS software products. Additional powerful polynomial methods are implemented in e.g ArcGis Geostatistical Analyst.

2.4.1.5 Radial basis functions RBF

Short description:
Radial basis functions are used for calculating smooth surfaces from large number of data points. This method is not adequate for sample points with large changes within small distances. RBF are

exact interpolators. Basic concept is to fit a surface through the data points while minimising the total curvature of the surface. The fit between the data points results from the selected basic function.

Mathematical background:
The radial Basis Function S(x) is given by:

$$S(x) = p(x) + \sum_{i=1}^{n} \lambda_i \phi(|x - xi|)$$

where

p(x)	is a polynomial of degree at most k
λ_i	are the weights
ϕ	is a basic function
xi	are distinct data points

There are several different choices for the basic function like:

Thin –plate spline:	$\phi(r) = r^2 \log(r)$
Gaussian:	$\phi(r) = \exp(-cr^2)$
Multiquadratic:	$\phi(r) = \sqrt{r^2 + c^2}$

Fundamental assumptions:
It is assumed that there are no large changes within small distances and that data points are without error.

Measure of success of interpolation:
There is no such measure for Spline interpolation.

Implementation in GIS-software:
Radial basic functions are implemented only in high-cost GIS software e.g ArcGis Geostatistical Analyst.

2.4.2 Stochastic methods

Stochastic methods incorporate the concept of randomness in spatial processes. The interpolated surface is conceptualised as one of many that might have been observed, all of which could have produced the known data points. Stochastic methods for spatial problems are often referred to as geostatistics. It is based on the idea that values measured at locations close together tend to be more alike than values measured at locations further apart. A variogram or covariance function provides a method for quantifying the spatial correlation. From there the spatial correlation is incorporated in spatial modelling. Detailed geostatistics theory can be found in Cressie (1991).

The mathematical development of kriging for use in geology was made in early 60-ies by Matheron (1963) in France. At the same time a similar method was independently developed in Soviet Union by Gandin (1963) for use in synoptic meteorology. He called this method optimum interpolation. It is a method widely used especially within the analysis of meterological fields. In contradiction to kriging, Gandin interpolation is based on the spatial correlation function and the assumption of 2.O stationarity. It also needs a first guess field as “background”. Optimum interpolation is therefore suited for analysis of fields on short time resolution, e.g. for meteorological applications where NWP-fields can be used as first guess fields.

In geostatistics it is assumed that the data $Z(\mathbf{s}_1),\ldots Z(\mathbf{s}_n)$ are partial realization of a random process $\{Z(\mathbf{s}):\mathbf{s}\in D\}$, where $\mathbf{s}$ denotes a location and D is a fixed domain in R^2 or R^3. To provide replication necessary for modelling and estimation, the intrinsic stationarity is assumed:

$$E[Z(\mathbf{s}+\mathbf{r})-Z(\mathbf{s})]=0 \text{ or } E[Z(\mathbf{s})]\equiv\mu$$
$$\mathrm{var}(Z(\mathbf{s}+\mathbf{r})-Z(\mathbf{s}))=2\gamma(\mathbf{r}), \quad \mathbf{s},\mathbf{s}+\mathbf{r}\in D.$$

The first equation implies that the mean of the process is a constant that does not depend on spatial location. The second equation defines a variogram $2\gamma(\mathbf{r})$ that depends only on the separation vector $\mathbf{r}$ ($\mathbf{r}=\mathbf{s}_i-\mathbf{s}_j$ is a distance vector between locations $\mathbf{s}_i$ and $\mathbf{s}_j$) and not on the location $\mathbf{s}$ itself. $\gamma(\mathbf{r})$ is often called a semivariogram. The variogram function satisfies certain properties related to continuity, symmetry, behaviour at infinity and nonnegative definiteness (Cressie 1991).

In the derivation of kriging equations, the semivariogram is assumed to be known. The semivariogram $\hat{\gamma}(\mathbf{r})$ is estimated from the data using one of many estimators such as the classical semivariogram estimator (Matheron 1963):

$$\hat{\gamma}(\mathbf{r})=\frac{1}{2|N(\mathbf{r})|}\sum_{N(\mathbf{r})}(Z(\mathbf{s}_i)-Z(\mathbf{s}_j))^2,$$

where $N(\mathbf{r})$ denotes the set of pairs of locations at distance $\mathbf{r}$ and $|N(\mathbf{r})|$ denotes the number of corresponding pairs of locations.

In practise, we fit a parametric semivariogram model to the values of the estimated semivariogram. The parametric semivariogram model must satisfy the necessary properties of the theoretical semivariogram mentioned above. A collection of valid models is given by Cressie (1991). A parametric semivariogram model is in general three parameter model: the parameter c_0 is the so-called "nugget effect", it represents the microscale variation, c_s is "partial sill", the variance of random process and a is "range", the longest distance with correlated values of the random process. The parameters of the semivariogram model may be estimated from the empirical semivariogram using several techniques such as weighted least squares or maximum likelihood.

2.4.2.1 Ordinary kriging

Short description:

The prediction obtained by ordinary kriging is a linear combination of measured values, with weights depending on the spatial correlation between the data.

Mathematical background:

The ordinary kriging model for spatial stochastic process $Z(\mathbf{s})$ is:

$$Z(\mathbf{s})=\mu+\delta(\mathbf{s})$$

μ is unknown expected value of random process, independent on location $\mathbf{s}$, $\delta(\mathbf{s})$ is a zero-mean intrinsically stationary random process with existing variogram $2\gamma(\mathbf{r})$. The predicted value $\hat{Z}(\mathbf{s}_0)$ can be expressed as: $\hat{Z}(\mathbf{s}_0)=\sum_{i=1}^{n}\lambda_i Z(\mathbf{s}_i)$. Linear coefficients λ_i, $i=1,\ldots,n$ are calculated under the condition for uniformly unbiased predictor:

$E(\hat{Z}(\mathbf{s}_0)) = E(Z(\mathbf{s}_0)) \quad \sum_{i=1}^{n} \lambda_i = 1$

and under the constraint of minimal prediction error variance (kriging variance) $\sigma^2(\mathbf{s}_0)$ at location $\mathbf{s}_0$:

$$\sigma^2(\mathbf{s}_0) = E(Z(\mathbf{s}_0) - \hat{Z}(\mathbf{s}_0))^2 .$$

The details of this theory can be found in Cressie (1991), Wackernagel (1995), Isaaks and Srivastava (1989) and also in other places.

Fundamental assumptions:
For the spatial process $Z(\mathbf{s})$ the intrinsic stationarity is assumed. The predictions are weighted linear combinations of the available data. Linear coefficients are calculated under the condition for uniformly unbiased predictor and under the constraint of minimal prediction error variance (kriging variance).

A disadvantage of this method using meteorological variables is that they can rarely be considered as intrinsic stationary random process. In some cases we can use different size and shape of search neighbourhood to eliminate this problem.

Measure of success of interpolation:
The ordinary kriging gives prediction errors, called kriging standard errors (square root of kriging variance).

Implementation in GIS-software:
Ordinary kriging is offered by all high level GIS software products (Arc GIS, Surfer, Idrisi, Gstat, GSLIB, Isatis, S-Plus Spatial module…).

2.4.2.2 Cokriging

Short description:
The cokriging procedure is an extension of kriging when multivariate variogram or covariance model and multivariate data are available. A variable of interest is cokriged at a specific location from data about itself and about auxiliary variables (covariables) in the neighbourhood. The data set may not cover all variables at all sample locations. If all variables have been measured at all sample locations and if the variables are intrinsically correlated, than cokriging is equivalent to kriging.

Mathematical background:

Suppose that the data are $k \times 1$ vectors (variables) measured on n locations. Multivariate process can be written with the $n \times k$ matrix:

$$\mathbf{Z} \equiv (\mathbf{Z}(\mathbf{s}_1),...,\mathbf{Z}(\mathbf{s}_n))^T$$

with i,j-th element $Z_j(\mathbf{s}_i)$, $i = 1,...,n$ in $j = 1,...,k$. We want to predict $Z_1(\mathbf{s}_0)$ based not only on $\mathbf{Z}_1 \equiv (Z_1(\mathbf{s}_1),...,Z_1(\mathbf{s}_n))^T$, but also based on the covariables $\mathbf{Z}_j \equiv (Z_j(\mathbf{s}_1),...,Z_j(\mathbf{s}_n))^T$, $j \neq 1$.

The same assumptions as by ordinary kriging should be expected for each of k variables. The kriging predictor of $Z_1(\mathbf{s}_0)$ is a linear combination of all the available data values of all k variables:

$$\hat{Z}_1(\mathbf{s}_0)=\sum_{i=1}^{n}\sum_{j=1}^{k}\lambda_{ij}Z_j(\mathbf{s}_i)$$

Assuming an uniformly unbiased predictor we get the conditions:

$\sum_{i=1}^{n}\lambda_{1i}=1$ in $\sum_{i=1}^{n}\lambda_{ji}=0$, $j=2,\ldots,k$.

Therefore, the best linear unbiased predictor is obtained by minimizing

$$E\left(Z_1(\mathbf{s}_0)-\sum_{i=1}^{n}\sum_{j=1}^{k}\lambda_{ij}Z_j(\mathbf{s}_i)\right)^2$$

By cokriging algebra becomes much more complicated than by kriging. For more details of theory see Cressie (1991).

In general, cross-variogram is defined as cross-products between differences of these variables for a pair of locations separated by a vector **h**:

$$2\gamma_{ij}(\mathbf{h})\equiv \mathrm{var}(Z_i(\mathbf{s}+\mathbf{h})-Z_j(\mathbf{s}))$$

In the case of cokriging we have to estimate the variogram model for the variable of interest, the variogram models for all covariables and also cross-variogram models for all pairs of primary variable and each covariable.

Cokriging becomes a highly complex method when more than one covariable is considered. From comparison studies of cokriging with other multivariate interpolation methods (mentioned later in this text) it was shown that cokriging gives better results only in the case when spatial correlation between covariables and variable of interest is high and when the covariables are oversampled with respect to the primary variable.

Fundamental assumptions:
Mathematically assumptions are the same as by ordinary kriging, additional assumptions exist with respect to cross-variogram model estimation (Cauchy-Schwarz inequality).

Measure of success of interpolation:
The cokriging gives prediction errors, called kriging standard errors (square root of kriging variance).

Implementation in GIS-software:
Cokriging is offered by all high level GIS software products (ArcGIS, Surfer, Idrisi, Gstat, GSLIB, Isatis, S-Plus Spatial module…).

2.4.3 Combined methods (deterministic and stochastic part)

2.4.3.1. Universal kriging

Short description:
Universal kriging is a spatial multiple regression implementing a model which splits the random function into a linear combination of deterministic functions (known at any point of the region) and

a random component, the residual random function. It turns out that in general this model needs additional specification because the variogram of the residual random function can only be inferred in exceptional situations (Wackernagel 1995).

Mathematical background:
In general, the mathematical model for a spatial random process $Z(\mathbf{s})$ can be written as:

$$Z(\mathbf{s}) = \mu(\mathbf{s}) + \delta(\mathbf{s}) = \sum_{j=1}^{p+1} f_{j-1}(\mathbf{s})\beta_{j-1} + \delta(\mathbf{s}), \quad \mathbf{s} \in D, \tag{1}$$

$Z(\mathbf{s})$ is a spatial random process value at the location $\mathbf{s}$, $\mu(\mathbf{s})$ is expected value of random process $E(Z(\mathbf{s}))$, for which it is assumed that is a linear combination of known functions $\{f_0(\mathbf{s}),\ldots,f_p(\mathbf{s})\}$, $\mathbf{s} \in D$, β_{j-1} are coefficients of the linear combination, $\delta(\mathbf{s})$ is a zero-mean intrinsically stationary random process with existing variogram $2\gamma(\mathbf{r})$,

The random process value at location $\mathbf{s}_0$, $Z(\mathbf{s}_0)$ is:

$$Z(\mathbf{s}_0) = \sum_{j=1}^{p+1} f_{j-1}(\mathbf{s}_0)\beta_{j-1} + \delta(\mathbf{s}_0). \tag{2}$$

According to the universal kriging theory, the predicted value $\hat{Z}(\mathbf{s}_0)$ is expressed as a linear combination of the measured values $Z(\mathbf{s}_i)$ on n locations:

$$\hat{Z}(\mathbf{s}_0) = \sum_{i=1}^{n} \lambda_i Z(\mathbf{s}_i). \tag{3}$$

Linear coefficients λ_i, $i = 1,\ldots,n$ are calculated under the condition for uniformly unbiased predictor:

$$E(\hat{Z}(\mathbf{s}_0)) = E(Z(\mathbf{s}_0)) \qquad \sum_{i=1}^{n} \lambda_i f_{j-1}(\mathbf{s}_i) = f_{j-1}(\mathbf{s}_0) \text{ for } j = 1,\ldots,p+1. \tag{4}$$

and under the constraint of minimal prediction error variance (kriging variance) $\sigma^2(\mathbf{s}_0)$ at location $\mathbf{s}_0$:

$$\sigma^2(\mathbf{s}_0) = E\left(Z(\mathbf{s}_0) - \hat{Z}(\mathbf{s}_0)\right)^2. \tag{5}$$

The constrained optimisation problem can be written equivalently as the unconstrained minimisation of:

$$E\left(Z(\mathbf{s}_0) - \sum_{i=1}^{n} \lambda_i Z(\mathbf{s}_i)\right)^2 - 2\sum_{j=1}^{p+1} m_{j-1}\left\{\sum_{i=1}^{n} \lambda_i f_{j-1}(\mathbf{s}_i) - f_{j-1}(\mathbf{s}_0)\right\} \tag{6}$$

with respect to $\lambda_1,\ldots,\lambda_n$, and $m_0,\ldots,m_p$ (Lagrange multipliers that ensure assumption (4)). Assuming

$$\mathrm{var}(\delta(\mathbf{s}+\mathbf{r}) - \delta(\mathbf{s})) = 2\gamma(\mathbf{r}), \tag{7}$$

Eq. (6) becomes

$$-\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma(\mathbf{s}_i-\mathbf{s}_j)+2\sum_{i=1}^{n}\lambda_i\gamma(\mathbf{s}_0-\mathbf{s}_i)-2\sum_{j=1}^{p+1}m_{j-1}\left\{\sum_{i=1}^{n}\lambda_i f_{j-1}(\mathbf{s}_i)-f_{j-1}(\mathbf{s}_0)\right\} \quad (8)$$

Upon differentiating (8) with respect to $\lambda_1,...,\lambda_n$, and $m_0,...,m_p$ and setting the result to zero, the optimal weights and Lagrange multipliers are obtained. On their basis the kriging prediction and minimal kriging variance at location $\mathbf{s}_0$ can be calculated. The details of this theory can be found in Cressie (1991) or Wackernagel (1995).

In general it is difficult and cumbersome to infer the underlying variogram (assuming it existence) because the variogram model is not independent on deterministic part of the model. Two steps lead to a more sound and workable model. First we must use translation-invariant and pair wise orthogonal functions for deterministic part of the model (Intrinsic random functions of order *k*). And second, we use generalized covariance function instead of variogram model, which is only one particular case of generalized covariance function.

Particular formulation of universal kriging is kriging with the external drift (Wackernagel, 1995).

2.4.3.2 Kriging with external drift

The kriging with external drift consists in integrating into the kriging system supplementary universality conditions about one or several external drift variables $g_i(\mathbf{s})$, $i = 1,\ldots, n$ measured exhaustively in the spatial domain. Actually the functions $g_i(\mathbf{s})$ need to be known at all locations of $Z(\mathbf{s})$ measurements as well as at the nodes of the estimation grid. In this case the deterministic functions stay "external" to the process of inferring the variogram.

Kriging with the external drift is very common method used in meteorology and climatology.

Fundamental assumptions:

For the random part $\delta(\mathbf{s})$ of the spatial process the intrinsic stationarity is assumed. External drift variables should be measured exhaustively in the spatial domain on all locations of $Z(\mathbf{s})$ measurements as well as at the nodes of the estimation grid.

Measure of success of interpolation:

Results of cross-validation, kriging variance. On the basis of statistical theory we can not say anything about accuracy of extrapolations in the deterministic part of the model.

Implementation in GIS-software:

Kriging with external drift (universal kriging) is offered by all high level GIS software products (Arc GIS, Surfer, Idrisi, Gstat, GSLIB, Isatis, S-Plus Spatial module…).

2.4.3.3 Residual kriging or detrended kriging

Short description:
In this case a regression (e.g. multivariate linear) is previously fitted between the spatial variable $Z(s)$ and the set of explanatory variables $x_i(s)$, the residuals of regression model $\delta(s)$ are calculated for each sample location and interpolated using ordinary kriging method.

Mathematical background:

$$Z(s) = \beta_0 + \beta_1 x_1(s) + \beta_2 x_2(s) + \ldots + \beta_p x_p(s) + \delta(s)$$

$$Z^*(s) = \beta_0 + \beta_1 x_1(s) + \beta_2 x_2(s) + \ldots + \beta_p x_p(s)$$

$$\sigma^*(s) = Z(s) - Z^*(s)$$

In the second step $\delta(s)$ is assumed as random variable that follows the intrinsic hypothesis and values of $\delta^*(s)$ is interpolated by a simple method like ordinary kriging. The final predicted values of $Z(s)$ are obtained by summation of $Z^*(s)$ and $\delta^*(s)$ over the kriging grid.

For the deterministic part various forms of regression models can be used (e.g. horizontal and vertical regionalised regression models, with filtering at borders). Predictors are e.g. elevation, inclination, aspect, distance to coast, latitude, etc.

Fundamental assumptions:
The same as for the kriging with external drift

Measure of success of interpolation:
Results of cross-validation, kriging variance, regression residuals analysis.

Implementation in GIS-software:
For the residual kriging calculation we can use different high level GIS software products (Arc GIS, Idrisi, Gstat, GSLIB, Isatis, S-Plus Spatial module...) or some standard statistical package for deterministic part and then GIS software for stochastic part kriging.

2.4.3.4 AURELHY (Analysis Using the RELief for Hydrometeorology)

A special example of residual (detrended) kriging is the AURELHY method introduced by Meteo France (Benichou, 1986). Meteorological fields are strongly influenced by topography, e.g rainfall. The AURELHY method considers the topography, originally described by the altitude at each grid point. In addition the elevation of neighbouring grid points (e.g 11x11 matrix) are condensed and reduced to the first principal components, which allow describing schematically how local topography is structured. These are calculated to find explicative variables, which allow deriving the field under consideration from topographical parameters using a linear multiple regression. The part of the field that is not explained by topography components is then interpolated using kriging techniques, and the result is finally obtained as the sum of the deterministic and stochastic parts. It is worth to mention that such methods based on statistical relationships DO NOT apply on an event basis.

2.4.4 Physically based methods

These methods are purely based on physical laws and/or empirical parameterisations. Some of these methods involve NWP model output, but others work in combination or solely with observations, like e.g. the energy balance method. Also reported are more statistical and mathematical based interpolation and analysis methods, which use observations, radar, or NWP model fields as background. All the former considerations, together with the wide range of variables and temporal and spatial scales, make it difficult to give a simple description of physical methods.

Some of the reported more physically based methods use NWP model output. Because NWP models usually produce large-scale output, the purpose of the so-called downscaling methods is to increase the spatial resolution of the NWP output by adding information of the topography (like e.g. roughness length or albedo) on a high resolution. Another solution would be to increase the resolution of the NWP model itself (down to a several kilometres horizontal resolution). However, this implies a relatively small calculation area and high computational costs. Besides, when e.g. site-specific wind speed in a heterogeneous terrain is needed the horizontal resolution has to be unrealistic high. On the other hand, disadvantage of the downscaling techniques is the vulnerability for the quality of the NWP model output. Moreover, some of the downscaling techniques can be inaccurate during very stable atmospheric conditions.

In principle, downscaling methods can also be applied on observations, but the inventory has not resulted in such contributions.

2.4.5 Methods of artificial neural networks

The methods of artificial neural networks (ANN) are relatively new as methods for spatial interpolation. They are learning systems both defining and adjusting necessary parameters without deterministic algorithms. To do this they have to be trained by examples of input data and output solutions. ANN can approximate unknown functions and probability distributions. From there they are a good alternative to the classical statistical modelling procedure. Major part of modelling procedure is construction of the network topology. Due to its complexity it is out of the scope of this paper to introduce the method of ANN in more detail. For detailed description of use of ANN in spatialisation see e.g. Sarközy F. (www.agt.bme.hu/public_e/funcint/funcint.html). The black box behaviour is an important feature of ANN which do not allow interpretation of relation between input variables and output data. An example of use of ANN for spatialisation of precipitation is given by Xingong (1997).

3. Inventory of spatialisation applications.

At the 2nd Management Committee meeting of COST719 in Vienna, 14-15.May 2001 it was decided (WG2M1/D1) to carry out an inventory about existing spatialisation applications within COST action 719. This inventory should focus on the methods, their strengths and limitations, and how they are related to GIS. A form (annex 1) was worked out, and distributed to all participants of the action. Cooperation towards the spatialisation group (WG1) of COST 718 has been established, and the inventory form was also distributed to them.

The contributions show a large variation of spatialisation applications, from advanced operational applications to loose plans of implementing such. Many of the applications are based upon the recent availability of spatialisation techniques within GIS software like the ESRI ArcView extension Spatial Analyst or IDRISI, or other software packages like SURFER or GSLIB.

This chapter presents the results of the inventory, focusing on the methods, predictor data/metadata and predicted variables. The advantages and shortcomings of the applications are emphasised, pointing out the direction where the needs for further development is necessary.

3.1 Methods used in spatialisation applications.

As presented in chapter two, there are a variety of methods available for spatial interpolation. Some of them have even several options, making it possible to include co-variables or trend expressions. In table 3.1, an overview of the use of the different methods is presented.

As it can be seen form this table, most of the methods available in the most applied software packages are used for spatialisation purposes.

3.1.1 Mathematical methods

These methods are as described in chapter two only based on coordinates of the input points (nearest neighbour approaches like Thiessen-polygons or more smoothed approaches like IDW-methods. These methods are used in about 12 applications.

Different inverse distance weighting approaches are reported in 6 applications. All report that the main advantages of the method are that it is quick and reliable. The main disadvantage is that the method not provides any direct measure of the uncertainty. If the variables to be interpolated not behave smooth in space, the method gives poor results.

Spline functions are curve-fitting techniques, fitting a curve or surface to the observation points. The application of such is reported also by six contributions. One reported advantage is as for IDW that this is a quick method. It is also reliable in the sense that it is an accurate interpolator. The resulting surfaces shows smooth patterns, which is reported as an advantage as this gives good graphical presentations. However the method may often give unrealistic patterns, and thereby poor, and even invalid estimates. When large variations over short distances occur in the input variables, unrealistic wave patterns have a tendency to occur.

3.1.2 Deterministic methods.

These methods are based upon deterministic relation between a set predictors and the predictand. These may be based on physical relations, or be established by the use of probability theory. The latter approach is the case for linear regression, which is one of the most used methods to set up such relations. In the inventory, a few cases are reported using only regression, but finding such relations is also widely used in combination with stochastic geostatistical methods, where the latter is used to interpolate local residuals (see next section). In some cases manual correction is used. The contributions on regression methods do not really discuss the advantages and disadvantages. One advantage frequently mentioned is that this is a quick method. The main disadvantage is that because it is a deterministic method, it is not valid for cases not following the same statistical distribution as the sample used to define the relation. It does not take anisotropy into account, and extremes will be smoothed (as for most interpolation methods).

3.1.3 Stochastic methods.

These methods utilize the spatial covariance structure of the process to be interpolated. In spatial interpolation geostatistics is another term for this approach. Kriging is the most known class of methods in this category, while its "cousin" optimum interpolation also must be regarded as a member. Different variants of kriging are applied among the contributors of this inventory.

In *ordinary kriging* the spatial covariance (expressed as a semivariogram assuming the intrinsic hypothesis) is considered. The direct use of kriging on the observed variables is not widely applied, as the intrinsic hypothesis is not likely to be valid. However a few examples exist. The most prominent advantage of kriging is the estimation of kriging variance, showing the uncertainty of the interpolation. The disadvantage of applying ordinary kriging is that the intrinsic hypothesis rarely is fulfilled.

Therefore other variants of kriging are developed to fulfil this important assumption. Two main approaches are used, *co-kriging* and *residual kriging*. While co-kriging include another co-variable in the establishment of the (cross-)semivariogram, residual kriging use a description of a trend, which is subtracted from the input variables before an ordinary kriging is used on the residuals. This approach is also named detrended kriging. In most cases the trend expression is so representative that the intrinsic hypothesis is closely fulfilled. The most used variables used to describe the trend expression are different topographical and physiographical characteristics. Usually linear regression is used to establish the trend. The reported advantages are that the uncertainty is known (at least for the stochastic component of the interpolation), and that the non-stationarity reduced. The approach is fast and reliable, and it objective and consistent. The results depend on the choice of the trend variables, and that they describe the desired climatology satisfactorily. E.g. are there problems describing minimum temperature reasonably well, since this element highly depends upon local terrain patterns.

A special version of detrended kriging is the Aurelhy method, developed at MeteoFrance. In this application, principal components of the terrain are used as trend variables. These PCs actually describe terrain shapes, which obviously should play a significant role on the local climatology. This approach is implemented in France and Hungary, and shows good results. The advantages are the same as for residual kriging, and the results are explicit functions of the orography. However sometimes wrong estimation of the PCA-loadings give some strange wave patterns.

Optimum interpolation is a method widely used especially within the analysis of meterological fields. It is also often referred to as Gandin interpolation after the Russian scientist Lev Gandin who published some major works on this issue in the 1960ies. In contradiction to kriging, Gandin interpolation is based on the spatial correlation function and the assumption of 2.O stationarity. It also needs a first guess field as "background". Optimum interpolation is therefore suited for analysis of fields on short time resolution, e.g. for meteorological applications where NWP-fields can be used as first guess fields.

3.1.4 Physical methods

Physical methods are based on the physical laws concerning energy exchange, motion and mass. Their input variables can be variables interpolated by the above-mentioned methods, or they are based on boundary conditions. NWP-models are the most advanced methods in this respect. The advantage of such methods is that they provide a detailed continuous description (both horizontally and vertically) for a number of variables. Their main disadvantage is the need of computer facilities, and sometime also their parameterisation of near surface conditions can be questioned. The most obvious disadvantage is however the spatial resolution, which still is too coarse to describe local patterns over a larger region. The topography is too smoothed to represent local terrain effects. Therefore downscaling or downtransformation techniques are developed to provide small-scale estimates, especially concerning wind speed. These methods are however vulnerable for the output of the dynamical models and their description of atmospheric stability. Some methods of statistical interpolation with the dynamically estimated field as background are reported.

Table 3.1 Application of different methods in spatialisation application within climatology and meteorology.

Method	Country	Elements	Spatial resolution	Temporal resolution
Mathematical				
IDWSA	Slovenia	Precipitation	1 x 1 km	Hourly Daily
IDW	Spain	?	1 x 1 km	
	Norway	Climate residuals	1 x 1 km	Monthly
	Belgium	Air pollution	Variable	10 min. – 6 hours
Weighted Average	Switzerland	Global horizontal insolation	2-4 x 2-4 km	?
Gaussian weighted ditstance	Sweden	Precipitation	5 x 5 km	Monthly
Thiessen	UK	Road weather Winter indices	20m 1km	Hourly-daily
Nearest neighbour (Thiessen)	Belgium	Air pollution	Variable	10 min. – 6 hours
Bilinear interpolation				
Spline (tension)	Portugal	Precipitation, Temperature, Wind, Pressure, Insolation	400 x 400 m 1 x 1 km	Daily Monthly (Seasonal, annual)
Spline	Netherlands	Precipitation Temperature Radiation Sunshine	10 x 10 km	Daily, ..., Annual
	Spain	?	1 x 1 km	
	Norway	Climate residuals	1 x 1 km	Monthly
	Switzerland	All continous	Point.	10 min – hourly
	Greece	Precipitation, Drought indices Evaporation Hail energy	~ 50 x 50 km	Daily Monthly
Deterministic				
Manual analysis	Sweden	Temperature Precipitation Snow depth	Variable	Monthly
Regression	Sweden	Precipitation	5 x 5 km	Monthly
	Italy	Precipitation Temperature	100-2000 x 100-2000 m	Daily, ..., Annual
(fuzzy)	Italy	Precipitation Temperature	100-2000 x 100-2000 m	Daily, ..., Annual
	Germany	Precipitation Temperature	1 x 1 km	> Month

Method	Country	Elements	Spatial resolution	Temporal resolution
		Sunshine Cloudiness….		
Stochastic (geostat)				
Kriging	Slovenia	Precipitation	1 x 1 km	> Hourly
	Italy	Precipitation Temperature	100-2000 x 100-2000 m	Daily, …., Annual
	Norway	Climate residuals Precipitation	1 x 1 km	Monthly
	Norway	Precipitation		
Universal kriging	Slovenia	Precipitation	1 x 1 km	> Hourly
Residual kriging (Detrended kriging)	Austria	Temperature, Precipitation, Sunshine, Snow characteristics.	25 – 250 m	> Monthly
	Finland	Temperature Precipitation	10 x 10 km	Monthly
	Poland	Temperature (days)	250 x 250 m	> Monthly
	Norway	Temperature	1 x 1 km	Monthly
	UK	Road icing	20 m	hourly
(Aurelhy)	France	Temperature Precipitation	1 x 1 km 5 x 5 km	10 days Monthly
	Hungary	Temperatures Precipitation Humidity	600 x 900 m	Monthly, Annual
Co-kriging	Switzerland	Insolation	2-4 km	?
	Poland	Temperature (days)	250 x 250 m	> Monthly
Physical				
Objective analysis	Sweden	All met. field data	22 km	3 hrs
(incl. Opt.interpolation)	France	Meteorological fields	100-1000 km^2	10 min – 6 hrs.
Downscaling	Netherlands	Wind Temperature Humidity	500 x 500 m	Diagnostic
	Switzerland/ Austria	Pressure Temperature Humidity Wind	Any	Any
	Switzerland	Energy balance	4 x 4 km 25 x 25 m	Daily
Radarbased scaling	Netherlands	Precipitation	500 x 500 m	Daily

3.2 Variables

The inventory reveal a wide spread also in the modelled variables. Almost all climatological and meteorological elements are represented. The answers reflect the community from where the inventory has been carried out, so there is a majority of descriptions on climatological variables compared to meteorological.

3.2.1 Temperature

Twenty applications are estimating the spatial distribution of temperature, including applications considering the whole energy balance. Among the applications for temperature, around half of them use geographical data sets like topography and terrain characteristics as predictors. Most of these applications are residual interpolation approaches where the climatological trend is removed. In some approaches output from NWP-models is used as first guess fields, and observations and other data are used to establish fine resolution descriptions. The other applications are using different methods not allowing additional predictors. These methods are spline function, IDW-methods or ordinary kriging directly on the observations. The contributions report a higher degree of satisfaction by the use of methods taking geographical information into account.

A few applications also estimate minimum and maximum temperatures. Other temperature related variables like number of degree-days, temperature seasons, potential evapotranspiration and drought indices are also reported.

3.2.2 Precipitation

Many applications for interpolating precipitation are also reported. Around half of the applications are running only based on the observation without considering extra information from geographical datasets or co-variables. However quite a number of contributions report the use of such information, mostly in detrended approaches. One interesting approach is to apply scaled fields based on the difference between observations and 24hrs accumulated precipitation estimates based on radar observations (Klein-Tank, 1999).

A few applications are used to map snow cover and snow depth.

3.2.3 Humidity

Humidity is a very complex element with respect to spatial variability. Only a few applications are reported, and they are mostly physical approaches.

3.2.4 Energy balance

Estimation of the energy balance is regarded by several contributions. Especially applications for estimation of radiation variables seems to be well suited to implement in a GIS-environment, utilizing the terrain analysis functionality in the GIS.

3.2.5 Wind

Wind is a difficult element to estimate locally, and the only reliable way to this is to apply a physical approach. Downtransformation or fine scale dynamical models are applied. Of the five reported applications, four are using such approaches.

3.2.6 Number of days.

One important climatological measure is number of days a certain element occurs above or below a certain threshold. This number of days is often conservative in the spatial dimension, and therefore well suited for GIS applications. To the inventory the following indices where reported:

- Number of days with precipitation above a given threshold
- Number of days with snow cover
- Number of days with temperature above/below a given threshold, including temperature degree days

3.3 Scale

3.3.1 Temporal scale

Most of the reported application have a temporal scale of one month or coarser. It reflects that most of the contributions are from the climatology sector. There is however a number of applications on daily scale, and a few with even finer resolution (hourly, 10 min.). The latter resolution is used in applications coupled with dynamic models.

3.3.2 Spatial scale

Most of the reported applications use a spatial resolution of about 1 x 1 km^2. But the applications show a large variability also here, from 10 m up to 50 km. The scale seems to a certain extent to be related to the purpose of the application.

Table 3.2 Number of application classified by elements and methods.

Element	*Total*	*Mathematical*	*Deterministic*	*Stochastic (raw)*	*Stochastic (incl covariables)*	*Physical*
Temperature	**25**	4	3	5	10	3
Precipitation	**20**	5	5	5	5	0
Radiation	**8**	3	1	2	2	0
Humidity	**4**	0	0	1	1	2
Energy balance	**1**	0	0	0	0	1
Number of days	**6**	1	1	2	2	0
Evaporation	**3**	1	0	2	0	0
Wind	**5**	1	0	0	0	4

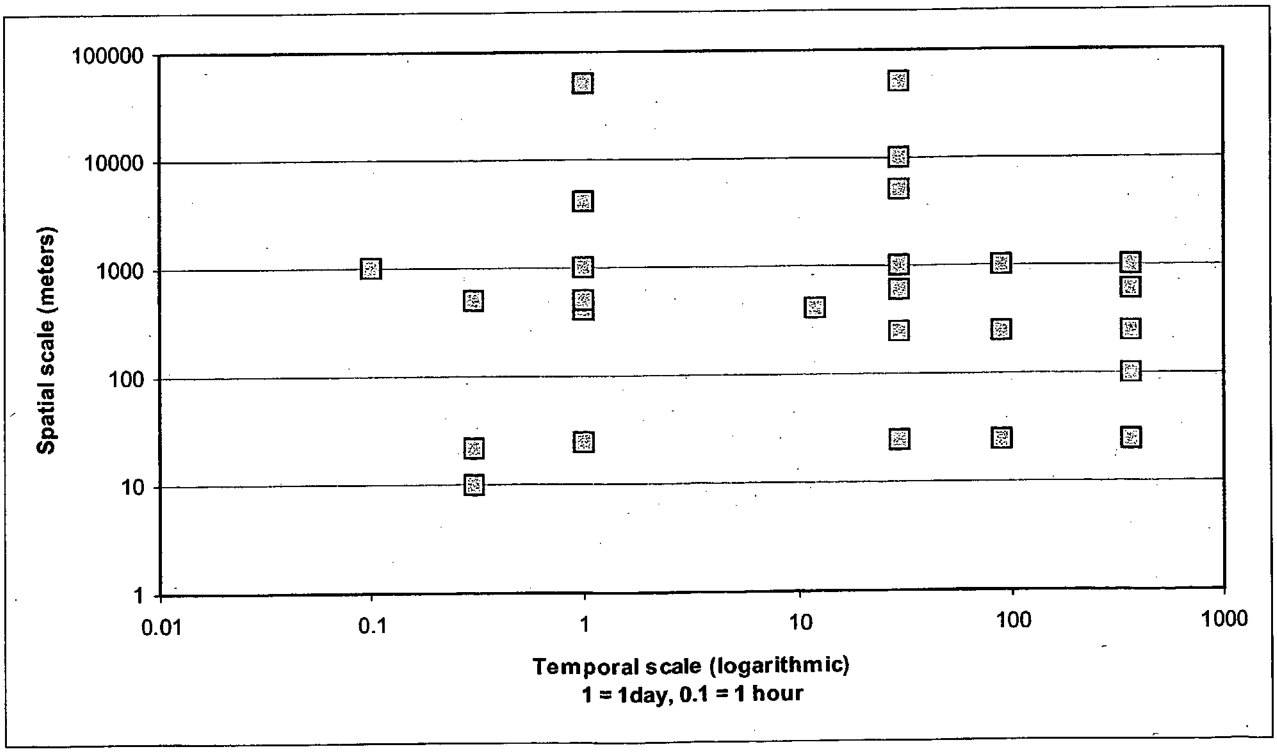


Fig 3.1 Spatial vs. temporal scale in the reported applications.

As can be seen from figure 3.1 there is no obvious link between the temporal and spatial scale. But it seems that the choice of scale is more related to the purpose of the application, resolution of available terrain models, and resolution of other models and not necessarily motivated by scientific reasons.

3.4 Predictors and metadata

The predictors used except meteorological and climatological information are mainly

- Digital terrainmodels (DTM)
- Derivations from DTMs
    - Slope
    - Aspect
    - Curvature
    - Principal components (Aurelhy)
    - Regional terrain statistics
- Coordinates (position)
- Land use
    - Roughness
    - Albedo
    - Hydrography
- Output from NWP-models.
- Satellite images
- Radar measurements

The most used predictor variable is (beside coordinates which is compulsory in spatial modelling) is elevation (digital terrain models). This is related to the purpose, and the strong connection between the variables to be interpolated and topography (temperature and precipitation).

4. Software

Most approaches are using software from ESRI[1]. ArcView and the Spatial Analyst extension is the software mostly used. A few institutions also have access to ArcINFO. These softwares are in the most recent versions (versions ≥ 8.1) merged into one source code. Other GIS-software applied is IDRISI and Surfer.

The limited interpolation capacities in commercial software have caused a wide use of external software libraries like IDL, IMSL, GSLIB, GSTAT, VarioWin etc., and many institutions have their own applications written in Fortran or C.

[1] ESRI: Environmental Systems Research Institute, http://www.esri.com

5. Examples

In this section a selection of contributed examples are presented. The presented examples reflect different methods, elements and scales.

5.1 Temperature

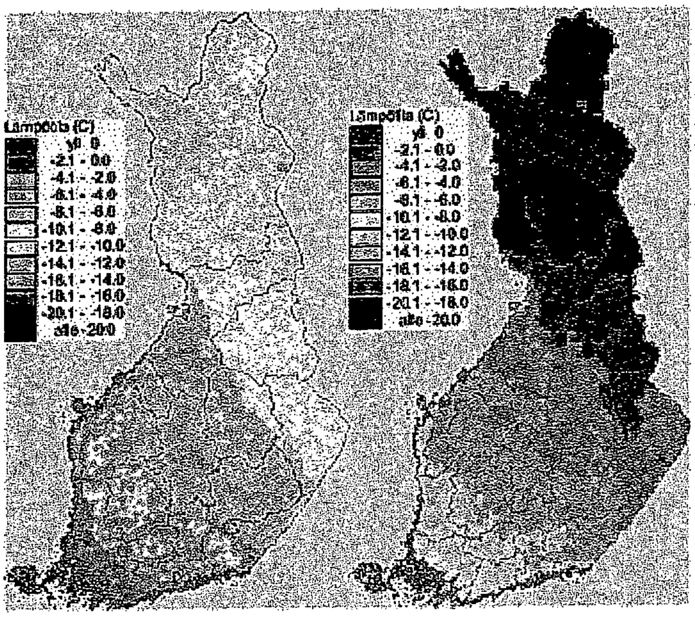
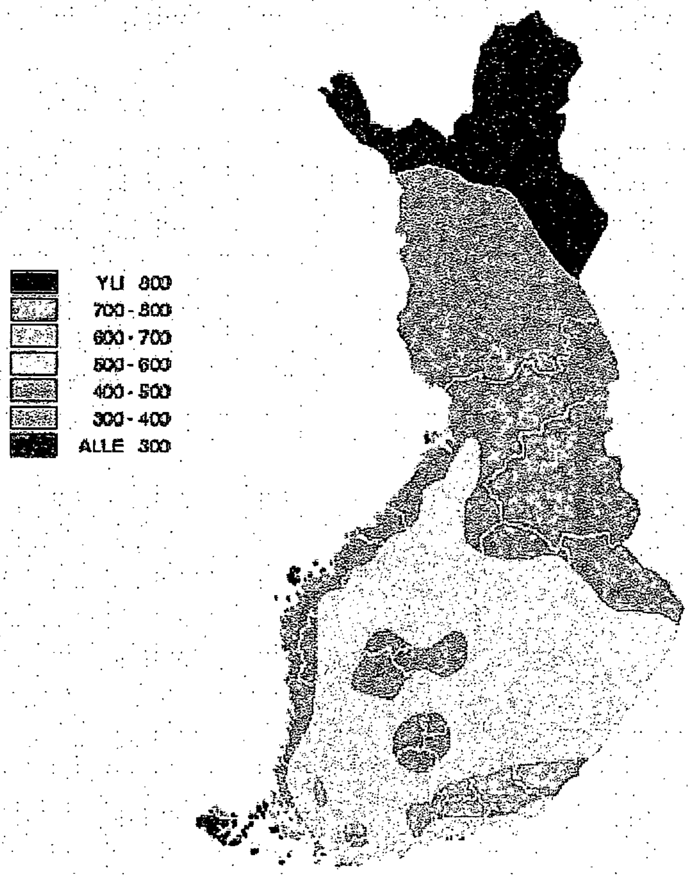


Fig. 5.1
Temperatures in Finland in January (average to the left, minimum to the right.)
Derived by residual kriging using terrain, distance to water bodies as independent variables. Resolution 10 x 10 km.

Kasvukauden tehoisa lamposumma 09.07.2001



copyright: Ilmatieteen laitos

Fig. 5.2 Effective temperature sum of the vegetation period 09.07.2001 (Translation: Yli = above, Alle = Below) (©Finnish Meteorological Institute)

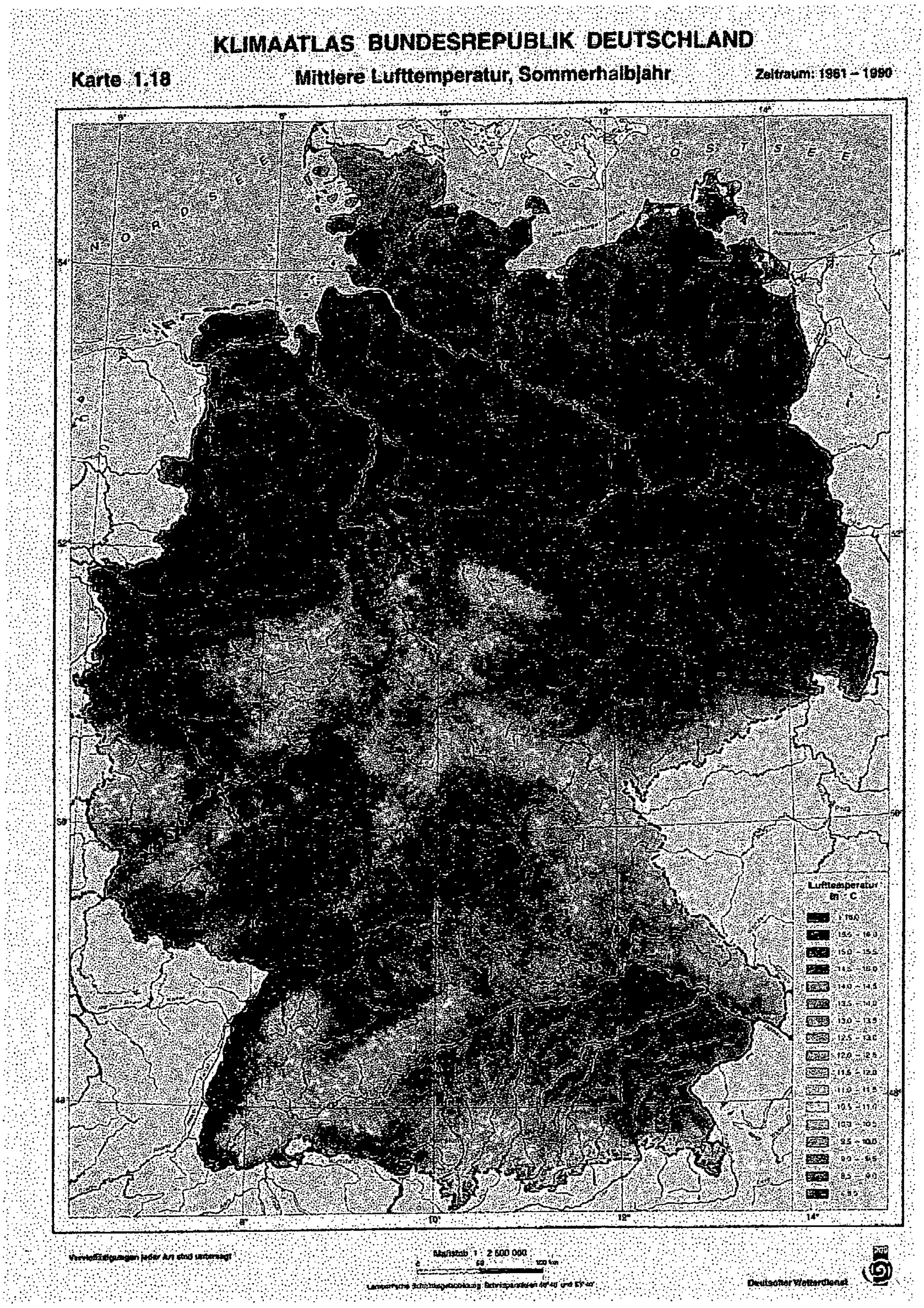


Fig.5.3 Mean May-October temperature in Germany, estimated by spatially interpolated regression expressions applying topography as predictor (DWD, 1999).

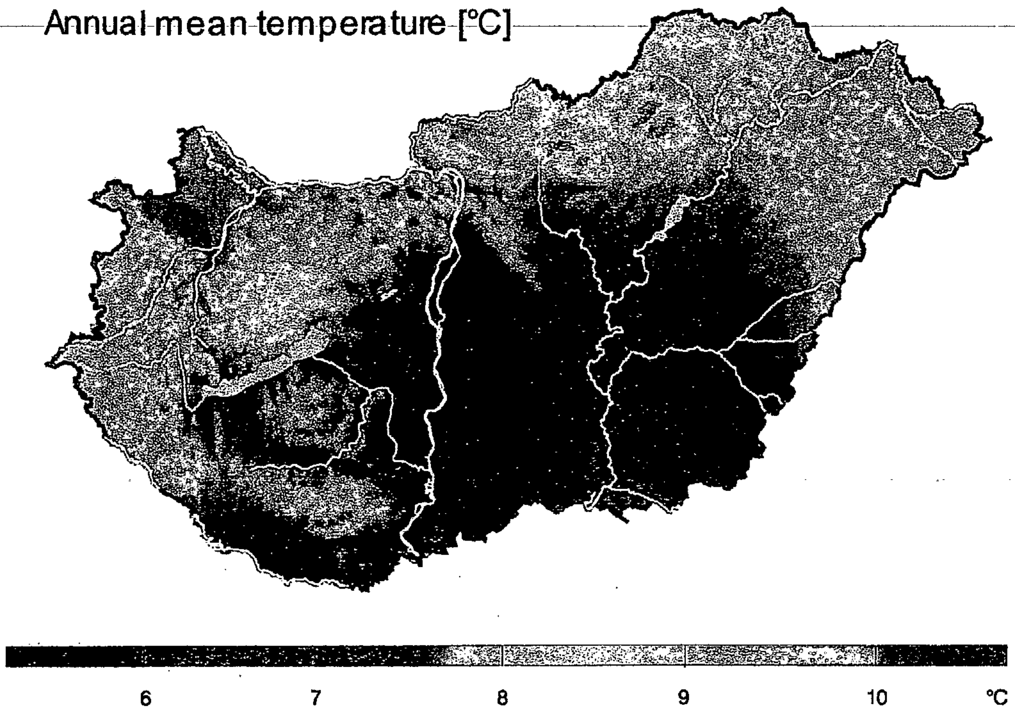


Fig. 5.4. Map of mean annual temperature in Hungary derived by the Aurelhy method.

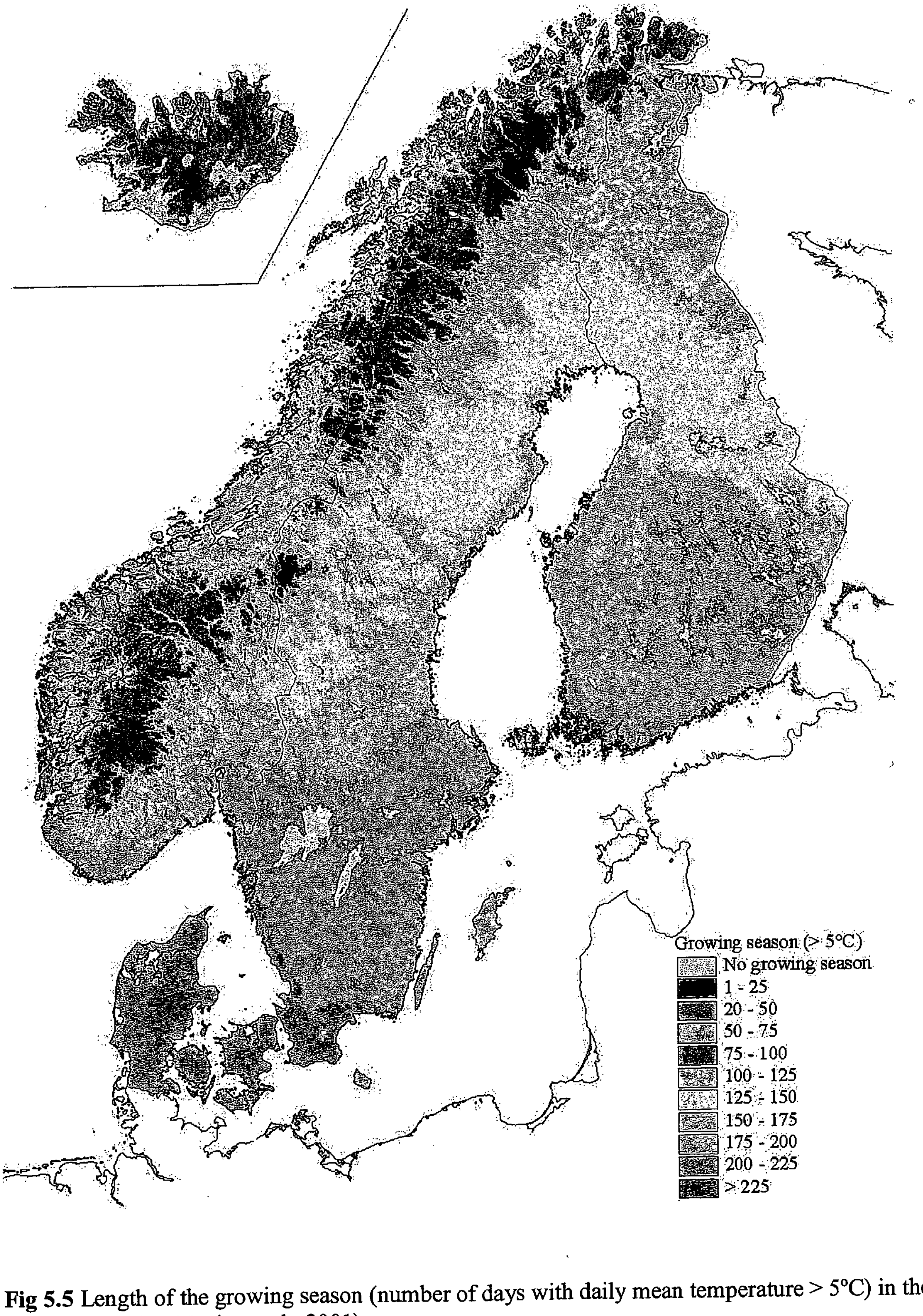


Fig 5.5 Length of the growing season (number of days with daily mean temperature > 5°C) in the Nordic countries (Tveito et al., 2001).

TOTAL AMOUNT OF PRECIPITATION. 1971 - 2000

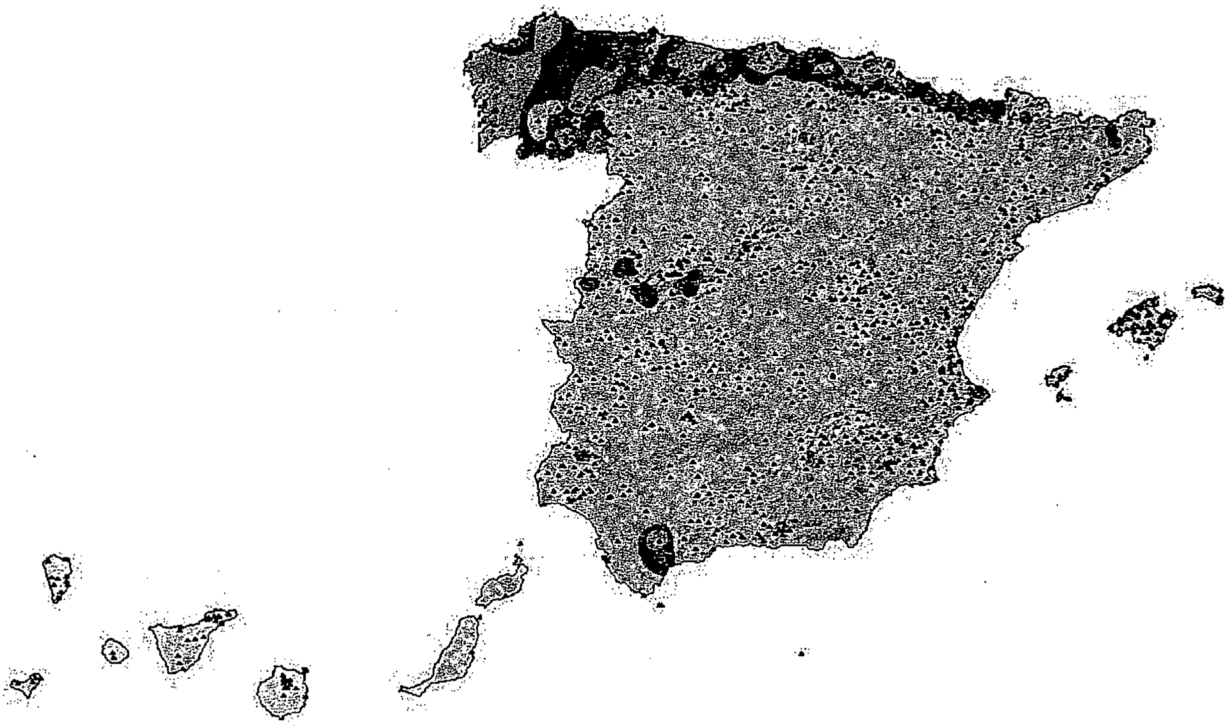


Fig. 5.6 Precipitation in Spain 1971-2000 (with shaded relief).

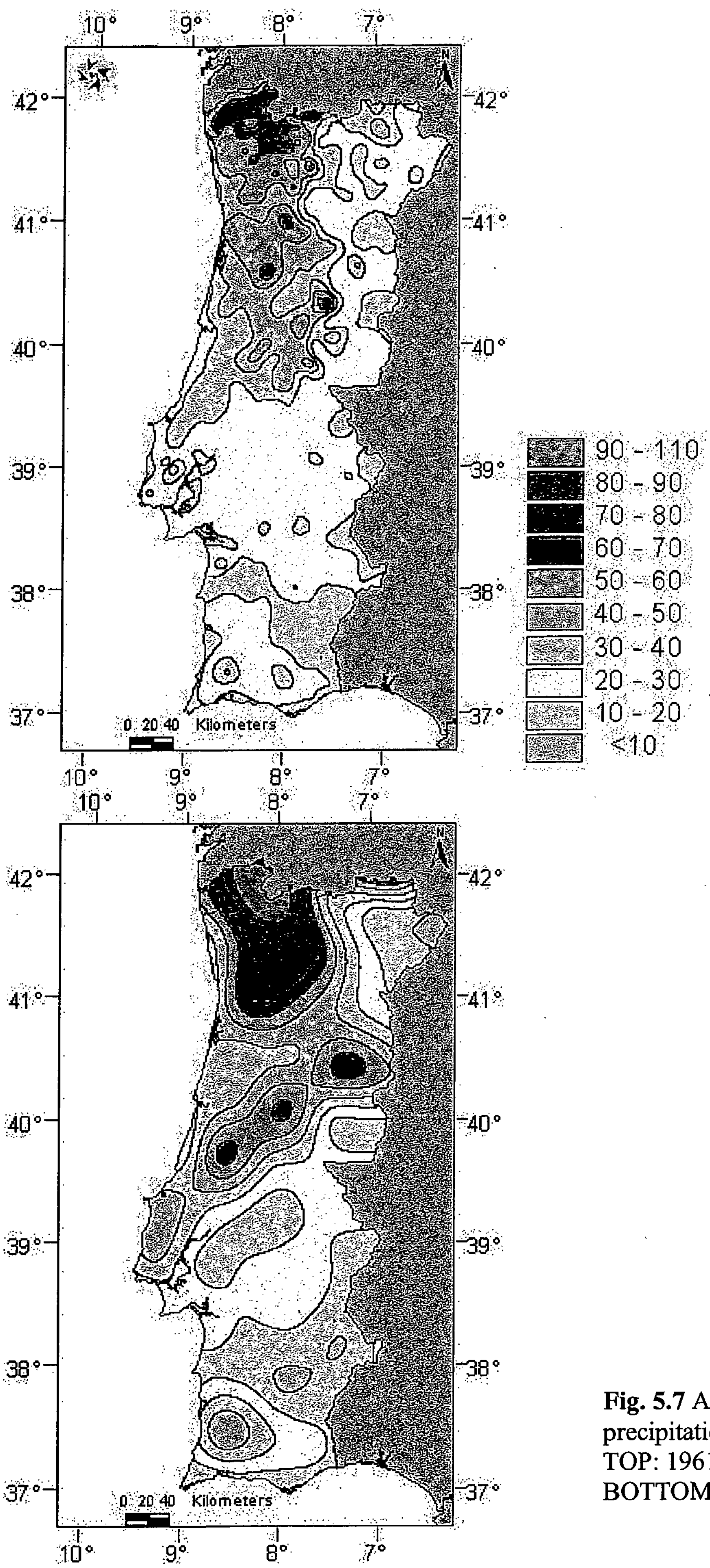


Fig. 5.7 Annual number of days with precipitation > 10 mm in Portugal
TOP: 1961-90
BOTTOM: Scenario 2080-99

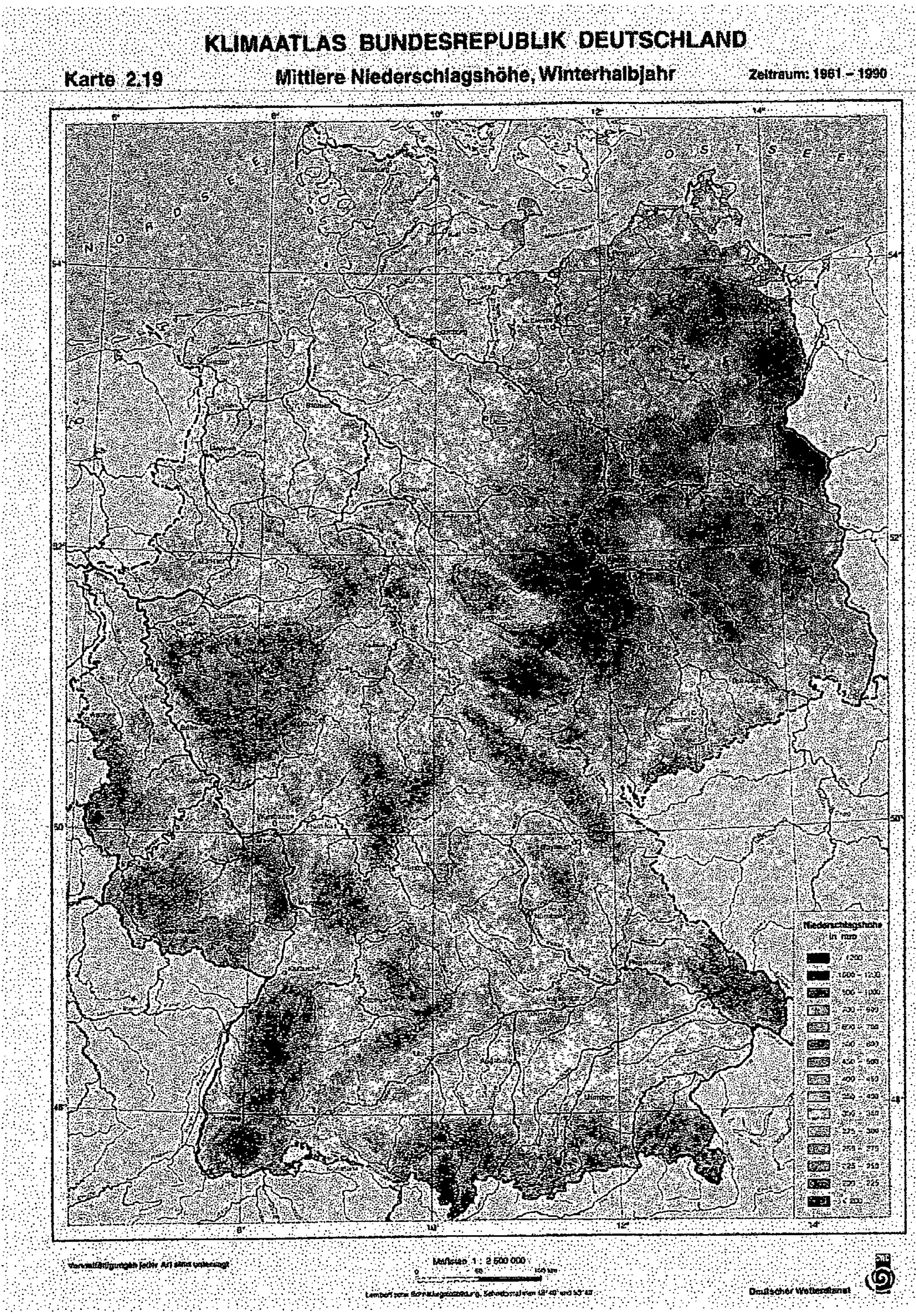


Fig 5.8 November-April precipitation in Germany, estimated by spatially interpolated regression (DWD, 1999)

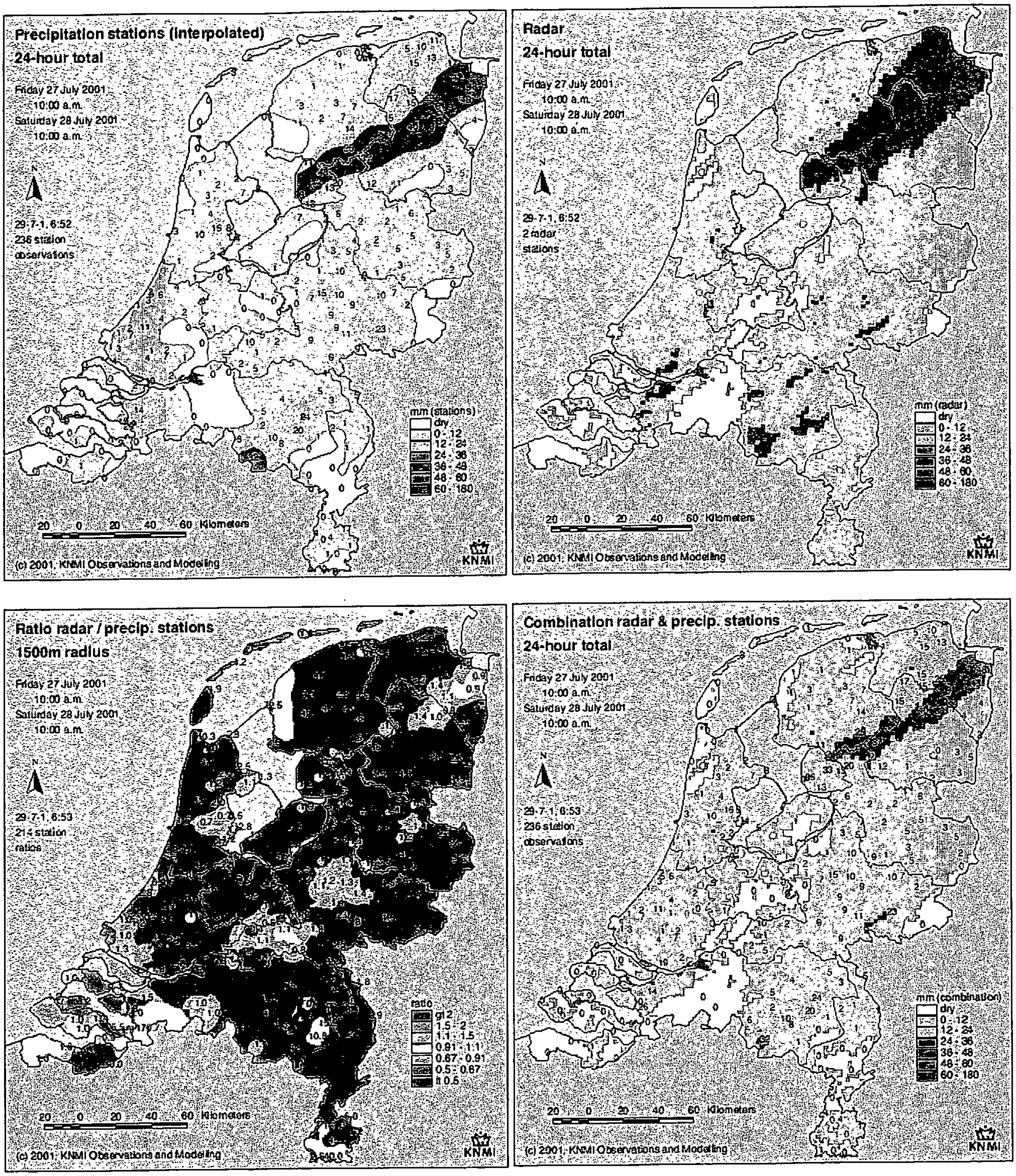


Fig 5.9 Radarbased scaling of daily precipitation, (Klein Tank, 1999)

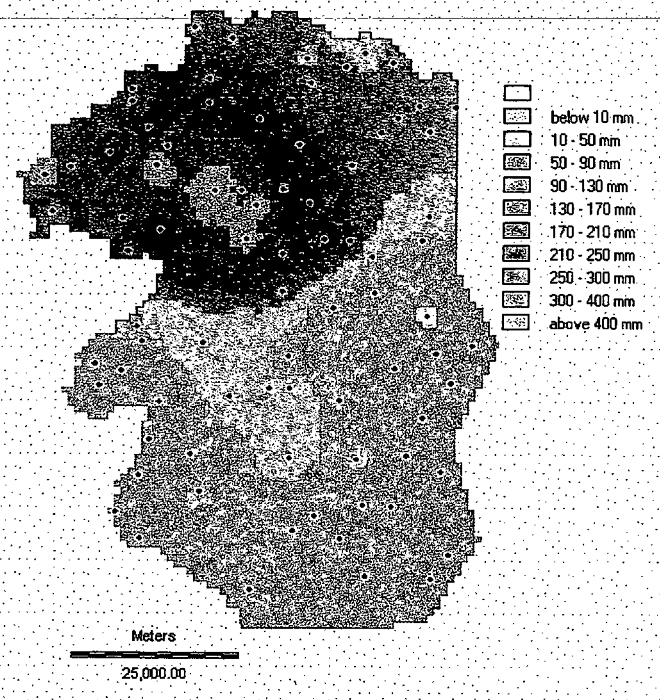


Fig. 5.10 Map of precipitation in Situation 2 using IDSWA method considering a circle with 15 km radius as influential surrounding. (Kastelec, 2000)

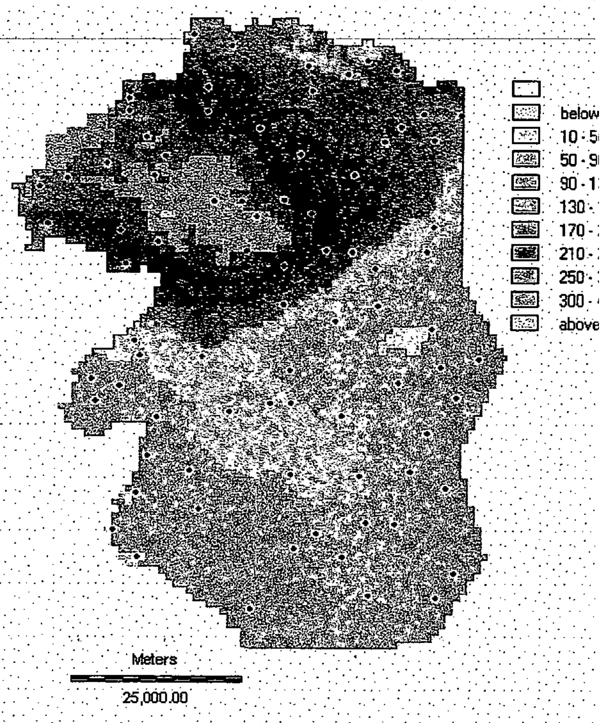


Fig 5.11 Map of precipitation in Situation 2 using universal kriging (y, z) considering a circle with 30 km radius as influential surrounding. (Kastelec, 2000)

5.3 Wind

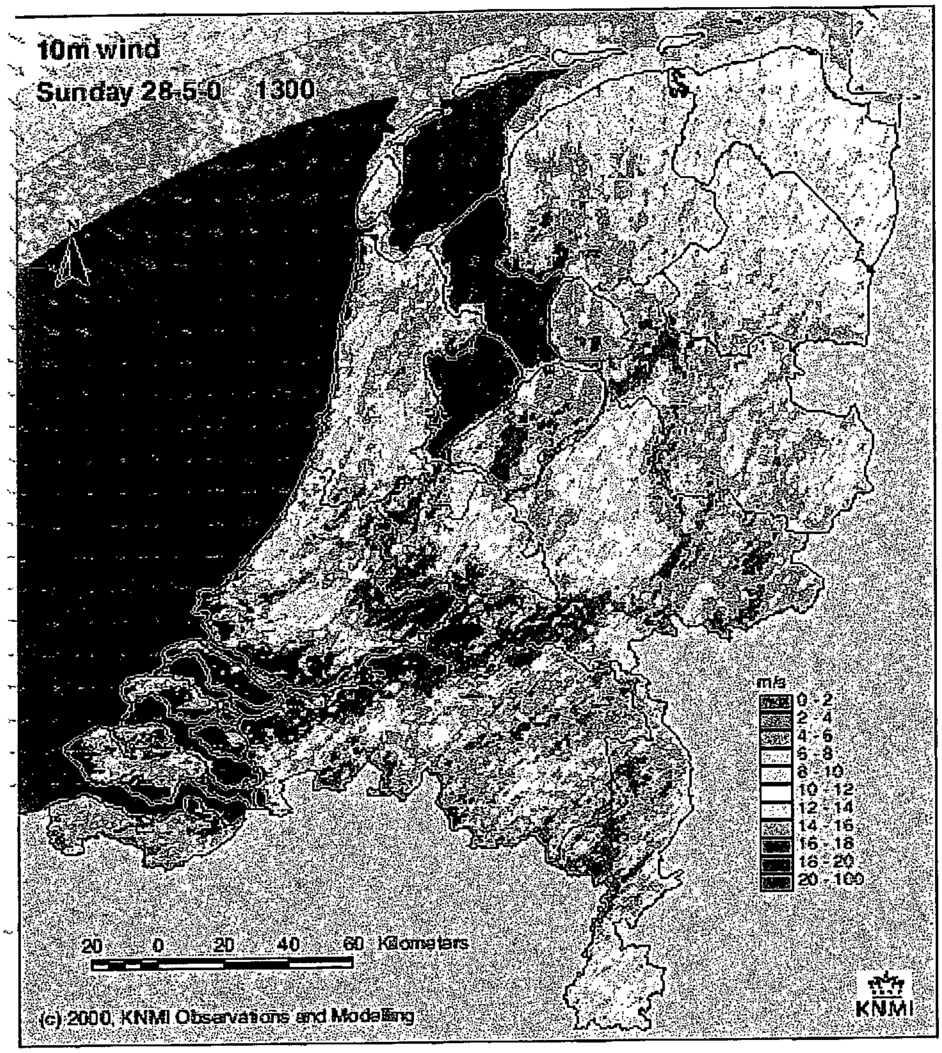


Fig.5.12 Wind in 10m-level in Netherlands obtained by a physical downtransformation from NWP-fields.

5.4 Other variables

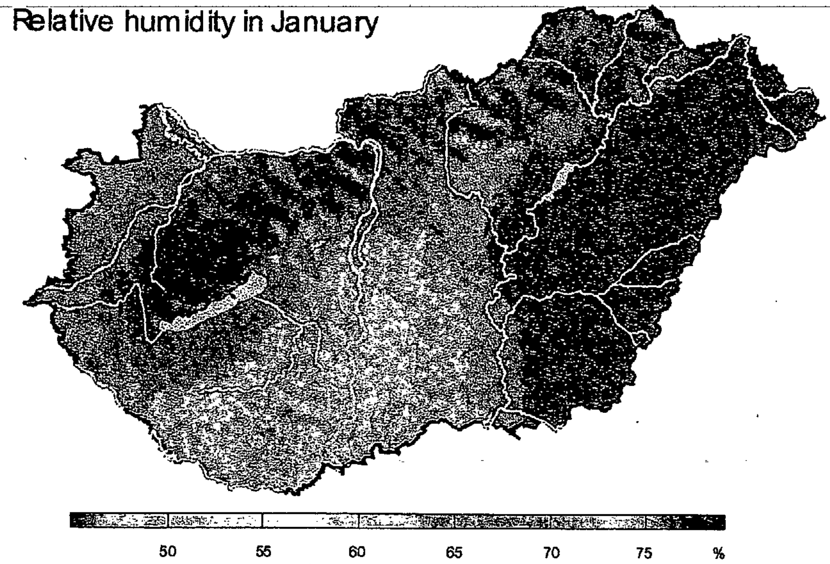


Fig.5.13 Mean relative humidity in Hungary.

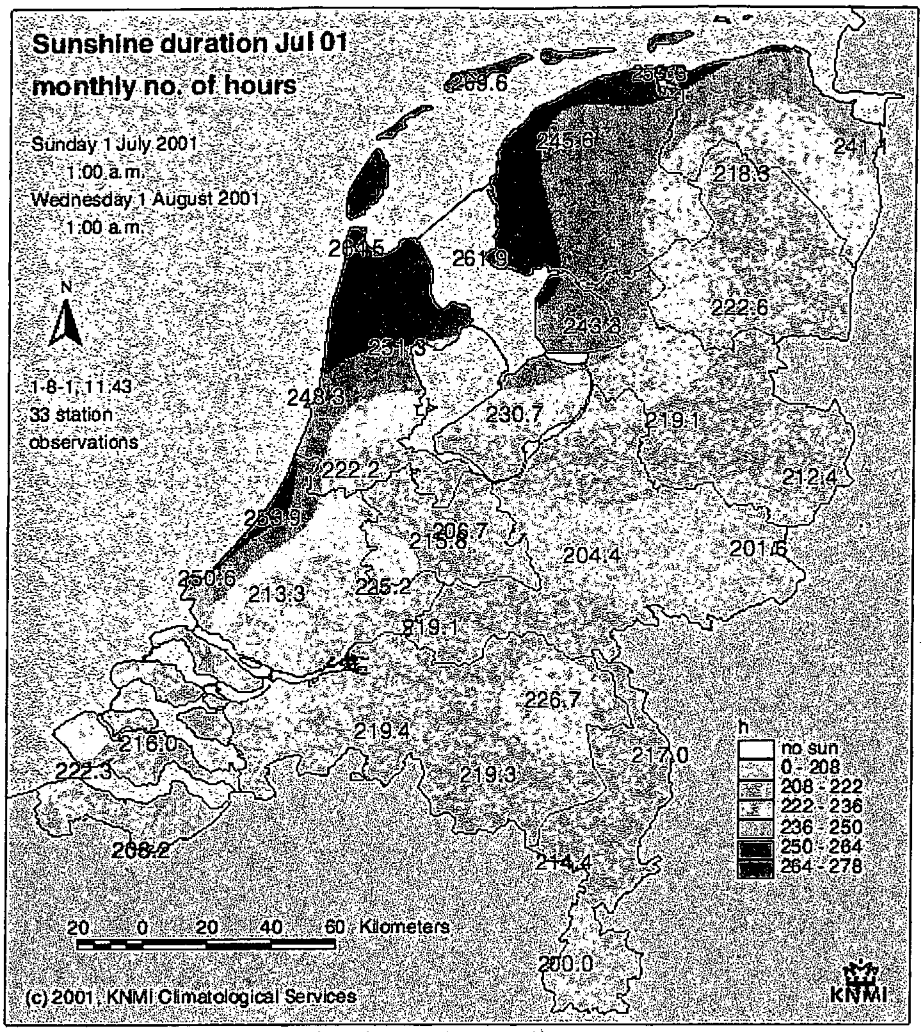


Fig.5.14 Sunshine duration in Netherlands.

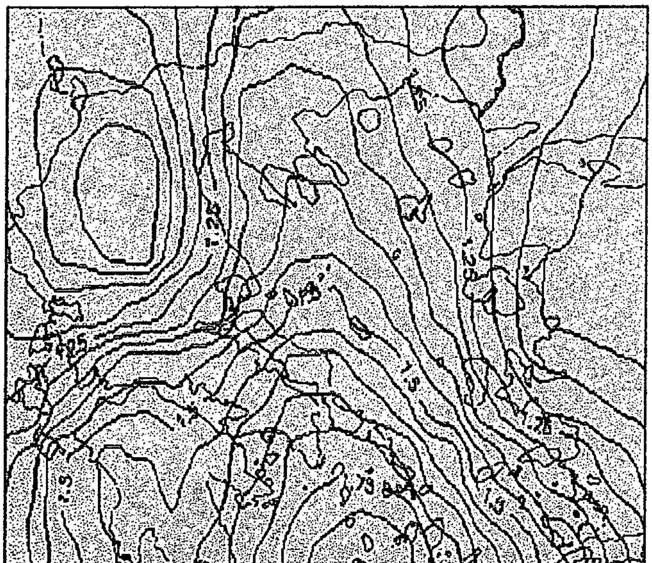


Fig 5.15 Evapotranspiration in Greece (Dalezios et al., 2001)

6. Discussion & Conclusions.

The use of spatialisation techniques in meteorology and climatology is wide spread. Using such methods in combination with GIS has made it possible to utilize geographical information in such applications much more effectively. It is no doubt that this combination is powerful, and GIS act as a bridge between meteorological data and geographical data.

There are however need for improvement and development before taking fully advantage of these possibilities:

1. The spatial interpolation schemes provided by most GIS-software are still limited, and need to be developed further.
2. There is a need for education and training of climatologists and meteorologist both to understand the possibilities, limitations and assumptions of different spatialisation methods. There is also a need for high skills in both GIS-techniques as well as spatialisation theory.
3. Data must be made available in usable formats, taking care of both the spatial and temporal dimensions of the data.
4. The implementation of spatialisation methods in climatological and meteorological applications must be done in close collaboration with the users, so that their needs can be met. This will be an everlasting iterative process.

Utilising GIS in spatial climatology and meteorology is still in an early phase. There are many challenges ahead. The potential using GIS and spatially distributed data is large, and there is a lot of research which is carried out and which may be implemented in a GIS-framework (ref.point 2). It is therefore of major importance to get spatialisation and GIS specialists together.

7 References

Akima H., 1970. A new method of interpolation and smooth curve fitting based on local procedures. Journal of Association for Computing Machinery, Vol. 17, No 4, pp. 589 – 602.

Barnes, S.L.1973. Mesoscale objective map analysis using weighting time series observations. NOAA. Tech. Memo. ERL. NSSL-62. U.S Department of Commerce, 60 pp.

Mestre Antonio. 2001. Report on rainfall spatialisation. Report COST 718 WG1 "Meteorological applications for agriculture", draft paper, 23 pp.

Cressie.N, 1991. Statistics for spatial data. Wiley, New York 900 pp.

Cressman G.P. 1959. An operational objective analysis system. Monthly Weather Review 87, 367-374.

Cornford D. 1998. An overview of interpolation. Seminar on Data Spatial Distribution in Meteorology and Climatology. EU Cost 79 publication. Edited by M.Bindi and B. Gozzini. Luxembourg, 1998. pp 45-56

Dalezios,N.R, Loukas,A., Bampzelis,D. 2001. Spatial variability of reference evapotranspiration in Greece, Submitted to Elsevier Journal Physics and Chemistry of the Earth-Part B.

Daly.C, Neilson R.P and Philips D.L, 1994. A statistical-topographical model for mapping climatological precipitation over mountainous terrain. Journal of Applied Meteorology 31, 661-676.

DWD, 1999. Klimaatlas der Bundesrepublik Deutschland, Teil 1, ISBN 3-88148-355-1, Offenbach am Main 1999

Frei and Schär C. 1998. A precipitation climatology of the Alps from high resolution rain gauge observations. Int. J. Climat. 18, 873-900.

Gandin, L. 1963 Objective analysis of meterological fields, /Gidrometeoizdatm, Leningrad), Translated into English by Israel program for scientific translation, 1965

Gandin et al., 1976. Statistical Structure of Meteorological Fields. Budapest 364 pp.

Kastelec, D. 2000. The spatial interpolation of the daily precipitation in the mountainous regions in Slovenia. *Geophys. res. abstr.*, 2000, vol. 2, CD-ROM, OA27

Klein-Tank,A.M.G 1999. Integration of precipitation data from various sources using a GIS. In Proceedings of the international conference on Quality, management and availability of data for hydrology and water resources management, 22-26 March 1999, Koblenz, Germany

Isaaks E.H and R.M. Srivastava ,1989. An introduction to applied geostatistics. Oxford University Press 561 pp.

Johnston K., Ver Hoef J. M., Krivoruchko K. and Lucas N. 2001. Using ArcGIS Geostatistical Analyst. ESRI. pp 300.

Matheron G. 1963. Principles of geostatistics. Economic Geology, 58: 1246-1266.

Matheron .G, 1981. Splines and Kriging: their formal equivalence. In Merriam D.F Eds. Down-to-Earth Statistics: Solutions Looking for Geological Problems. Pp 77-95. Syracuse University Geology Contribution N° 8 New York.

Schöner W. and H. Mohnl. 22000. A new map of accumulated new snow of Austria. 26th Int. Conf. o. Alpine Meteorology, Österreichische Beiträge zur Meteorologie und Gophysik 23: CD-ROM.

Sheppard.D., 1968. A two-dimensional interpolation function for irregularly spaced data. Proc. of the 1968 ACM National Conference. pp 517-524.

Strobl J., 1999. Universitätslehrgang "Geographische Informationssysteme", Modul 9, Räumliche Analysemethoden 1, Inst. f. Geography Univ. Salzburg, p 1-5.

Thiessen.A.H, 1911. Precipitation averages for large areas. Monthly Weather Review 39, pp 1082-1084.

Tveito, O.E., E.J.Førland, H.Alexandersson, A.Drebs, T.Jonsson, H.Tuomenvirta, E.Vaarby-Laursen (2001) Nordic climate maps, DNMI report 06/01 KLIMA

Wackernagel H. 1995. Multivariate Geostatistics. Springer-Verlag: 256 str.

Wackernagel H. 1998. Splines and kriging with drift. Seminar on Data Spatial Distribution in Meteorology and Climatology. EU Cost 79 publication. Edited by M.Bindi and B. Gozzini. Luxembourg, 1998. pp 57-64.

Xingong L., 1997. Development of a neural network spatial interpolator for precipitation estimation. GIS/LIS '97 Annual Conference, Oct. 28-30 1997, Cincinnatti, Ohio, Proceedings Cdrom, pp 667 – 676.