Project name:
Integration and exploitation of networked Solar radiation Databases for environment monitoring
Project acronym:
SoDa
Contract number:
IST-1999-12245

Deliverable title:
Interpolation schemes - "Profile Method" (a process-based distance for interpolation schemes)
Deliverable number:
D5-1-1
Type of deliverable:
Report
Security:

Author:
ENSMP (Mireille Lefèvre, Lucien Wald)
Contributors:
Meteotest (Jan Remund)

Contractual date of delivery to the Commission:
July 2001
Actual date of delivery to the Commission:
July 2001

Abstract

This document reports on the activities held in order to select an interpolation method for the prototype of the SoDa service.

An important aspect in this service is the necessity of being able to perform a fast interpolation using an unknown number of measuring sites. Several resources necessitate estimates of meteorological fields at any point in the world. These estimates should be provided by a fast interpolation technique.

There is a wealth of publications dealing with the interpolation of meteorological fields. This work does not propose a new method but focuses on the definition of the distance between sites, a distance that is used in all methods. It was felt that the availability of geographical elements in digital form for the whole world, may lead to a better accuracy in interpolation through their integration in the interpolation technique. Orography and presence of water bodies are such geographical elements that influence the meteorological fields.

Several techniques are available to consider these elements. They lead to better results than the standard interpolation technique, the gravity method, at the expenses of more complexity, preparatory work and more computational time. This work explores the possibility of integrating the geographical elements as well as the latitudinal effects into an effective distance. This effective distance may then be used in the gravity method, a fast-running method, instead of the standard geodetic distance.

Approximately 700 sites are used in Europe to assess the benefits of this distance relative to the geodetic distance and another taking into account the difference in elevation. Each site is used in turn as the point where the meteorological parameter should be assessed. The estimates are compared to the actual values and biases and root mean square errors are computed. The meteorological parameters under concern are ten-year averages of monthly means of daily sum of horizontal global irradiation, daily sum of sunshine duration, daily extreme of air temperature, atmospheric pressure and water vapor pressure, and of monthly sums of precipitation.

This work demonstrates that taking into account the latitudinal effects in the distance brings an increase in the accuracy in interpolation. Such effects have been seldom mentioned in previous publications. The orographic effects may be partly corrected by adding the weighted difference in elevation to the geodetic distance. The following effective distance between the point $\boldsymbol{P}$ and each of the measuring sites $\boldsymbol{X_i}$ for all parameters, is found to give better results than the others:

$$d_{eff}^2 = f_{NS}^2\ (d_{geo}^2 + f_{oro}^2\ dh^2)$$

with $f_{NS} = [1 + 0.3\ ½\boldsymbol{F}_P - \boldsymbol{F}_X½\ [1 + (\sin\boldsymbol{F}_P + \sin\boldsymbol{F}_X) / 2]$, where d_{geo} is the geodetic distance in km, latitudes $\boldsymbol{F}_P$ and $\boldsymbol{F}_X$ are expressed in degrees, dh is the difference in elevation between $\boldsymbol{P}$ and $\boldsymbol{X_i}$ (expressed in km) and f_{oro} is set to 500.

The interpolation errors are in line with previous publications. Graphs are computed, which permit to provide to the customer of the service an assessment of the interpolation error, together with the interpolated value.

Introduction

In meteorology, as well as in other geophysical sciences, Earth surface processes are mostly known by the means of networks of ground-based instruments though increasing use is made of satellite observations. Each network is scarce and interpolation methods are necessary to assess the value $V(\mathbf{P})$ of the parameter under concern for any geographical point $\mathbf{P}$ located between the sites of measurements $\mathbf{X_i}$.

There is a wealth of publications dealing with the interpolation of meteorological fields. This work does not propose a new method but focuses on the definition of the distance between sites, a distance that is used in all methods. An important aspect of our work is the necessity of being able to perform a fast interpolation using an unknown number of measuring sites, especially in the framework of the Web service SoDa, which exploits a smart network of distributed resources to deliver information relating to the solar radiation (Rigollier *et al.* 2000). Several resources necessitate estimates of meteorological fields at any point in the world (SoDa 2001). These estimates should be provided by a fast interpolation technique.

The problem of interpolation may be seen as an adjustment mathematical problem, that is what is the hyper-surface that fits best the observed values $V(\mathbf{X_i})$? (Picinbono 1986). This surface is usually defined in an analytical form but piecewise polynomials, i.e. *B*-splines are often very appropriate (Hou, Andrews 1978). Once the parameters of the model known by adjustment over the observed values, it provides a value for any geographical point $\mathbf{P}$ for the area under concern. Least-square fitting of surfaces of polynomial type on a global or local scale, thin plate method or Hsieh-Clough-Tocher method, are examples of such methods (Hutchinson *et al.* 1984; Hulme *et al.* 1995). One of their advantages is the degree of continuity of the derivatives of the estimated field.

A more conventional approach consists in estimating the meteorological parameter at the single geographical point $\mathbf{P}$ using the nearest sites where measurements are made. The estimated value is a combination of the measurements $V(\mathbf{X_i})$ weighted by a function of the distance d_i between the point $\mathbf{P}$ and each of the measuring sites $\mathbf{X_i}$. The shorter the distance, the larger the influence of the site and the larger the weight of the site in the combination. The estimated value $V(\mathbf{P})$ is expressed as:

$$V(\mathbf{P}) = \sum_{i=1}^{N} w_i \, V(\mathbf{X_i}) \qquad (1)$$

where $V(\mathbf{X_i})$ is the value measured by the i^{th} site among the nearest N measuring sites and w_i is the weight for this site, with $\sum_{i=1}^{N} w_i = 1$.

In such a linear interpolation, there is no bias, that is that the estimated field V^* takes the measured values $V(\mathbf{X_i})$ at the geographical points $\mathbf{X_i}$ located over the measuring sites. This may not be the case for the fields resulting from a method based upon adjustment mechanisms. However, the first derivative of the field resulting from a linear interpolation is not continuous. Popular methods like kriging, objective analysis, and inverse squared distance are examples of linear interpolation methods (Journel, Huijbregts 1978; Thiébaux, Pedder 1987). In the inverse squared distance method, also called gravity method, the weights w_i are given by

$$w_i = \frac{1}{d_i^2} \Big/ \left(\sum_{j=1}^{N} \frac{1}{d_j^2} \right) \qquad (2)$$

where d_i, d_j are respectively the distances from respectively the site $\mathbf{X_i}$ and the sites $\mathbf{X_j}$ to the geographical point $\mathbf{P}$:

$$d_i = \left\| \overrightarrow{PX_i} \right\| \text{ and } d_i, d_j \text{ ¹ } 0 \qquad (3)$$

Different methods use different weights. Some methods make use of the estimated spatial structure of the field itself to compute the weights. This structure is expressed by e.g. the variogram or the correlation function.

Whatever the method, of adjustment type or linear type, a distance between the sites and the geographical point ***P*** is computed. It is usually the geodetic distance or the Euclidean distance if the data are located on a grid whose cells are regularly spaced. Given two sites ***P*** and ***X*** of geographical co-ordinates ($\boldsymbol{F}_P$, $\boldsymbol{l}_P$) and ($\boldsymbol{F}_X$, $\boldsymbol{l}_X$), where ***F*** and ***l*** stand for latitude and longitude, the geodetic distance d_{geo} is:

$$d_{geo} = R\ \boldsymbol{Q} \qquad (4)$$

where

$$\cos \boldsymbol{Q} = \sin \boldsymbol{F}_P \sin \boldsymbol{F}_X + \cos \boldsymbol{F}_P \cos \boldsymbol{F}_X \cos(\boldsymbol{l}_P - \boldsymbol{l}_X) \qquad (5)$$

where $\boldsymbol{Q}$ is expressed in radians and $(\boldsymbol{l}_P - \boldsymbol{l}_X)$ is the difference in longitude. The geodetic distance is the distance at the surface of the Earth if the Earth is considered as a sphere of radius *R* (R= 6371 km). It is also called the horizontal distance. Here, the latitude is counted positive from the equatorial plane northwards and negative southwards. The longitude is counted positive eastwards from the Greenwich meridian and negative westwards.

Taking into account the importance of natural processes

The spatial structure of the field of the meteorological parameter under concern is a function, usually complex, of several natural processes. Among others are the latitudinal effects and the orography. They should be taken into account for a better estimate of the field. Some methods, like the co-kriging, offer the possibility to take into account the cross-variation or the cross-structure of the parameter under concern and of another one. The latter should be positively correlated to the first one and sampled for a larger number of locations. The range of values of the second parameter taken by the measuring sites should be large, too. Finally, the set of values should represent the field under concern. This is not often the case, as shown by the example of France in the construction of the European Solar Radiation Atlas (ESRA, 2000). In this case, forty-four stations were used that measure global irradiation, sunshine duration, air temperature and pressure at ground level, precipitation and water vapor. Of these 44 sites, only 12 have elevations greater than 200 m (27 % of the total). Out of these 12, three have elevation greater than 400 m (7 %) and three have elevation between 300 and 400 m. This range of elevation is too narrow to establish an accurate description of the effect of the orography. In addition, approximately 40 % of the French territory exhibit elevation greater than 200 m, and more than 15 % has elevation larger than 500 m. It follows that the data set does not reproduce the distribution of the elevation values.

The use of detailed additional information may improve the results of the interpolation at the expenses of computational complexity. Co-kriging is often used in such cases, though other techniques prove as efficient (Beyer *et al.* 1997; Hudson, Wackernagel 1994; Lo 1989; Zelenka *et al.* 1992). In the construction of the European Solar Radiation Atlas (ESRA 2000), a segmentation of Europe was performed by the means of clustering analysis before proceeding to interpolation using a gravity method. In a similar way, Supit (1994) proposes to combine interpolated values of the Angström coefficients, representative sunshine duration for the area under concern and a regression model to estimate the solar irradiation. In their so-called climatologically aided interpolation method, Willmott and Robeson (1995) make a combined use of air temperature climatology and interpolated temperature deviations at measuring sites. Kunz, Remund (1995) used detailed models describing local relationships between orography and meteorological fields for Switzerland, a country with a very dense network of measuring stations. Some methods make use of multiple linear regression analysis of data from a set of nearby measuring sites to locally model changes in meteorological fields. Briefly said, the value at the point ***P*** in the neighborhood is then predicted from these sites using the regression curve, each site $\boldsymbol{X_i}$ being in addition inverse-weighted by its distance to ***P*** (Anonymous 1995; Jones, Thornton 1999; Nalder, Wein 1998; Price *et al.* 2000; Şen, Şahin 2001; Van der Voet *et al.* 1994).

Beside these efforts, other authors tried to integrate / model the natural processes into the definition of an effective distance. This process-based distance is then used in any interpolation method, instead of the geodetic distance. This is the approach selected for this work. Compared to the others, this technique permits to develop fast interpolation methods that can be launched within a Web service.

The effective distance

Beyer *et al.* (1997), MeteoNorm (2000) or Zelenka *et al.* (1992) used an effective distance:

$$d_{eff} = \sqrt{d^2_{geo} + (f\,\partial h)^2} \qquad (6)$$

where d_{geo} is the geodetic distance between the measuring site ***X*** and the current location ***P*** and ∂h is the difference in elevation of both locations. The geodetic distance and the elevation should be expressed in the same units, e.g. in km The parameter f controls the equivalence between horizontal and vertical distances. f is set to 100 by Zelenka *et al.* and to 300 by Beyer *et al.* The latter stress that the results are only weakly dependent upon f, for most of the elevation differences encountered in their work. This effective distance was then used in co-kriging or inverse squared distance interpolation.

Anonymous (1995) and Van der Voet *et al.* (1994) studied a effective distance taking into account the absolute difference in elevation of both locations, the possible presence of a climatic barrier between the site ***X*** and the point ***P*** empirically defined and the distance to the coastline.

The present work explores various ways to take into account the orography, the effect of the latitude and the presence of the large water bodies. This exploration of the possible modeling of the orographic effects and presence of the sea by an effective distance is possible because all necessary information is now available for the whole Earth in a digital form and with a sufficient accuracy. When launching the study, it was expected that a better representation of the orographic features and profiles would lead to a better accuracy in interpolation.

This work analyses the following effective distance d_{eff}. It is based on the geodetic distance, which is effectively increased to take into account some processes:

$$d_{eff}^2 = f_{lat}^2\, f_{NS}^2\, (d_{geo}^2 + f_{oro}^2\, d_{oro}^2 + f_{sea}^2\, d_{sea}^2) \qquad (7)$$

where $(f_{oro}\, d_{oro})$ is the additional distance taking into account the orography, f_{lat} and f_{NS} are taking into account the latitudinal effects and $(f_{sea}\, d_{sea})$ those induced by large water bodies.

Latitudinal effects

The meteorological fields have well-marked distributions along the latitudes. They are anisotropic: the fields tend to be homogeneous along the latitude with strong north-south gradients (see e.g., Anthes 1997; Atlas of hydrometeorological data 1991; ESRA 2000).

The meteorological phenomena are preferentially moving along the latitudes. Their size and lifetime are related to each other and the relationship is independent from the latitude. We suggest to normalize the distances by the perimeter of the latitude, that is to take into account that a phenomenon moving along the Equator or along a mid-latitude has the same lifetime. The poles are singular points. The perimeter of the latitude of the site ***P*** is $R\cos\Phi_P$. The geodetic distance is then corrected by the factor f_{lat}:

$$f_{lat} = 2 / (\cos F_P + \cos F_X) \quad (8)$$

with $(\cos F_P + \cos F_X)$ ¹ 0. From a practical point of view, f_{lat} is limited to $2\ 10^3$, *i.e.* $(\cos F_P + \cos F_X)$ should be more than 10^{-3}. Actually, we found out that this transformation has a negligible impact on the results. Price *et al.* (2000) reported a similar result. Accordingly, the factor f_{lat} is set to 1 in the following.

Another factor f_{NS} is introduced to model the anisotropy of the fields and to penalize the North-South distances:

$$f_{NS} = 1 + a_{NS}\ ½F_P - F_X½\ [1 + (\sin F_P + \sin F_X) / 2] \quad (9)$$

where a_{NS} is an arbitrary parameter.

Orography

In the Equation [6], the orographic effects are represented by a difference in elevation. This does not reproduce the case of an orographic barrier separating two sites ***P*** and ***X*** of equal elevation. It is proposed to use instead the profile of elevation.

Given an elevation profile extracted from a gridded digital elevation model (DEM), and approximating a grid cell by its center, the distance d_i at the ground surface between two cells i and $i+1$ of elevation h_i and h_{i+1} is:

$$d_i^2 = d_{geo}^2 + (h_{i+1} - h_i)^2 \quad (10)$$

We thus define the orographic distance as:

$$d_{oro}^2(\mathbf{P}, \mathbf{X_i}) = \sum_{i=1}^{N-1} (h_{i+1} - h_i)^2 \quad (11)$$

where N is the number of cells comprised between ***P*** and $\mathbf{X_i}$, along the shortest geodetic distance. For two profiles exhibiting the same overall difference in elevations, the distance d_{oro} penalizes that offering the largest local gradients.

The meteorological parameters strongly depend upon the elevation. We enforce the orographic distance by multiplying it by the parameter f_{oro}. It controls the equivalence between the horizontal and vertical distances. Setting f_{oro} to 100 means that a difference in elevation of 100 m is equivalent to a geodetic distance of 10 km.

The digital elevation model used is the model TerrainBase (TerrainBase 1995). The grid cell is 5' of arc angle (approximately 10 km at mid-latitude) and the elevation step is 1 m.

Other distances are computed for comparison with d_{oro}. We define:

the elevation difference distance: $f_{oro}\, dh$ (10)

where dh is the difference in elevation between ***P*** and $\mathbf{X_i}$, like in Equation 6

the elevation maximum difference distance: $f_{oro}\, DH$ (11)

where DH is the maximum difference in height that can be found in the orographic profile between ***P*** and $\mathbf{X_i}$: it is the absolute value of the difference between the highest point and the lowest one.

Water bodies effect

Large water bodies separating sites are assumed to create climatological barriers. The distance over such water bodies between ***P*** and $\mathbf{X_i}$ is called d_{sea}. Using TerrainBase in an appropriate manner, one may represent large water bodies and thus compute d_{sea} using the same cell size than for orography. Adding a quantity equal to $f_{sea}\, d_{sea}$, where f_{sea} is determined in an empirical way, will penalize the geodetic distance.

Database and tests

The various components of the effective distance were assessed for an area encompassing Europe, and ranging from 30° West to 70° East and from 25° to 75° North. Ten-year averages of monthly means of several parameters are available in the CD-ROM of the ESRA (2000) for a large number of sites:

- daily sums of horizontal global irradiation for 586 sites,
- daily sum of sunshine duration for 556 sites,
- daily extreme of air temperature for 435 sites,
- monthly sum of precipitation for 435 sites,
- atmospheric pressure for 266 sites,
- water vapor pressure for 274 sites.

The daily irradiations and the daily sums of sunshine duration were converted into respectively daily clearness indices *KT* and relative sunshine durations *S/S0*. Temperature, air pressure and water vapor pressure were not corrected for elevation. Standardizing measurements made at a given site to the elevation of the pixel containing this site read into the digital elevation model is possible if a model is available. Jones, Thornton (1999) used a lapse rate model for temperature.

For each parameter, each site is in turn assumed to be unknown and the value of the parameter is assessed by interpolation technique using the other sites for each month. The estimates are compared to the measured values and the discrepancies are computed. The set of discrepancies for all sites is then analyzed as a function of month, latitude, site, longitude and elevation. Two quantities are computed to synthesize the discrepancies: the bias and the root mean square.

Two interpolation techniques are used: the gravity method and the nearest-neighbor method. Having two methods permit to compare the results and the conclusions reached for each of them regarding the distances. Several distances are tested and compared. For each distance and if relevant, empirical parameters were adjusted in order to decrease the errors in interpolation:

- distance *D1* (geodetic): $D1 = d_{geo}$
- distance *D2* (equation [6]): $D2^2 = (d_{geo}^2 + f^2\, dh^2)$. According to MeteoNorm (2000), the quantity f is set to 300 to radiation parameters, 100 for temperature, 200 for precipitation and 100 for air pressure and water vapor pressure. Elevations are in km.
- distance *D3*: $D3 = f_{NS}\, d_{geo}$, with $f_{NS} = [1 + a_{NS} |\Phi_P - \Phi_X| [1 + (\sin\Phi_P + \sin\Phi_X) / 2]$. The value of a_{NS} leading to the smallest error in interpolated value was found to be 0.3. Latitudes are expressed in degrees. The range of values [0.2, 0.4] gives similar errors.
- distance *D4*: $D4^2 = f_{NS}^2\, (d_{geo}^2 + f_{oro}^2\, d_{oro}^2)$
- distance *D5*: $D5^2 = f_{NS}^2\, (d_{geo}^2 + f_{oro}^2\, dh^2)$
- distance *D6*: $D6^2 = f_{NS}^2\, (d_{geo}^2 + f_{oro}^2\, DH^2)$
- distance *D7*: $D7^2 = f_{NS}^2\, (d_{geo}^2 + f_{oro}^2\, dh^2 + f_{sea}^2\, d_{sea}^2)$

For distances *D4*, *D5* and *D6*, the quantity f_{oro} was set to 500, with elevation in km. The influence of the exact value of f_{oro} on the error is negligible, provided the quantity is approximately in the range [100, 1000]. This value is the same for all meteorological parameters.

The influence of the component $(f_{sea}^2\, d_{sea}^2)$ was assessed by using a selected number of sites of similar elevations, which are separated by large water bodies (e.g., sites in Azores Islands, Eire and Spain).

Results

The results show the large influence of the parameter f_{NS} on the error in interpolation. Compared to the distance *D1* (geodetic) and *D2*, the distance *D3* provides better results. Actually, taking into account the latitudinal effects is the major contributor to the increase in accuracy.

The influence of the water bodies component is negligible. Varying f_{sea} from 0 to 10 does not lead to a noticeable change in accuracy. Accordingly, this component is disregarded.

As for the orographic effects, we were disappointed by the results obtained by the distance *D4*. This distance takes into account the profile of elevation between the point $\boldsymbol{P}$ and the site $\boldsymbol{X_i}$. An improvement of accuracy was expected, which was not evidenced at all. Better results were attained by simpler formulae, such as the distances *D5* and *D6*. The reasons for that remain obscure to us. The results attained for the distances *D5* and *D6* are similar, with a slight advantage to *D5*. This distance has also a major advantage: it is faster to compute than *D6*. Finally, it should be noted that the influence of the orography correction is small compared to that of the latitudinal effects.

Table 1 gives the bias and rmse (root mean square error) for each meteorological parameter, for the two interpolation methods (nearest neighbor and gravity) and for the distances *D1*, *D2* and *D5*. It shows that in any case the smallest bias and rmse are observed for the distance *D5*. Radiation parameters are reproduced with an acceptable accuracy (approximately 10 %), which still remain low, given the fact that we are dealing with climatological means. Accuracy on rainfall is poor; it is well known that rainfall is a discrete field and that such interpolation methods are not appropriate to estimate precipitation. On the contrary, climatological means of air pressure are very continuous fields and these interpolation methods give very good results. As for the air temperature, bias is small and rmse amounts to approximately 2-3 °C. Such errors are similar to those found in previous publications for similar areas and similar types of data (Hulme *et al.* 1995). When compared to the results of Hulme *et al.* (1995), based on a fitting of approximately 800 stations using a thin-plate technique, covering the same area and dealing with similar types of data, our errors are similar for the precipitation and larger for the sunshine duration, air temperature and water vapor pressure. For irradiation, our results are similar to those reported in WMO (1981) for geodetic distances between sites $\boldsymbol{X_i}$ and point $\boldsymbol{P}$ less than 400 km, that is 8 to 10 %.

			Clearness index	*Relative sunshine duration*	*Rainfall (mm)*	*Air pressure (0.1 hPa)*	*Water vapor pressure (0.1 hPa)*	*Air temp. max. (°C)*	*Air temp. min. (°C)*
		Number of samples	6778	6521	4981	3166	3240	5220	5220
		Mean value	0.44	0.41	58	10154	99	14.4	6.5
Nearest-neighbor	*D1*	*Bias*	0.001	0.000	0.8	0.4	-1.2	-0.4	-0.2
		RMSE	0.046 (11 %)	0.058 (14 %)	27.9 (48 %)	17.8 (0.17 %)	19.4 (20 %)	4.0	2.8
	D2	*Bias*	-0.001	-0.002	-1.3	0.4	-0.6	0.0	0.1
		RMSE	0.040 (9 %)	0.054 (13 %)	24.6 (42 %)	17.8 (0.17 %)	18.6 (19 %)	3.2	2.1
	D5	*Bias*	-0.001	-0.001	-2.0	0.5	-1.0	0.0	0.0
		RMSE	0.040 (9 %)	0.052 (10 %)	23.9 (41 %)	17.1 (0.17 %)	15.4 (16 %)	2.7	2.0
Gravity	*D1*	*Bias*	-0.004	-0.008	3.3	1.4	0.2	-0.4	-0.2
		RMSE	0.041 (9 %)	0.063 (15 %)	23.2 (40 %)	20.4 (0.20 %)	19.6 (20 %)	3.6	2.8
	D2	*Bias*	-0.007	-0.009	1.9	1.4	0.4	-0.3	-0.1
		RMSE	0.041 (9 %)	0.062 (15 %)	23.2 (40 %)	20.4 (0.20 %)	19.5 (20 %)	3.5	1.9
	D5	*Bias*	-0.003	-0.004	0.3	-0.4	0.4	0.1	0.1
		RMSE	0.034 (8 %)	0.046 (11 %)	21.7 (37 %)	14.8 (0.15 %)	14.9 (15 %)	2.5	1.9

Table 1. Bias and rmse (root mean square error) for each meteorological parameter, for the two interpolation methods (nearest neighbor and gravity) and for the distances *D1*, *D2* and *D5*.

Figure 1 displays the bias and rmse observed as a function of the latitude for the clearness index, for the three distances *D1*, *D2* and *D5*. One may observe that the errors are constant with the latitude. The increase in bias and rmse with the lowest and highest latitudes is due to a border effect: there is no measuring sites southwards (respectively northwards) of the point ***P*** because we are working in a limited area and not on the whole Earth. This anisotropy in the distribution of sites used to interpolate leads to an increase in error.

Ideally, the bias should be equal to zero. The distance *D5* is that offering the bias that is the most constant with the latitude and the closest to zero. This distance is also that exhibiting the smallest rmse. In addition, this rmse is almost constant with latitude. The errors observed for the other distances are slightly larger than those observed with *D5*. The differences are not very large, as already reported in Table 1.

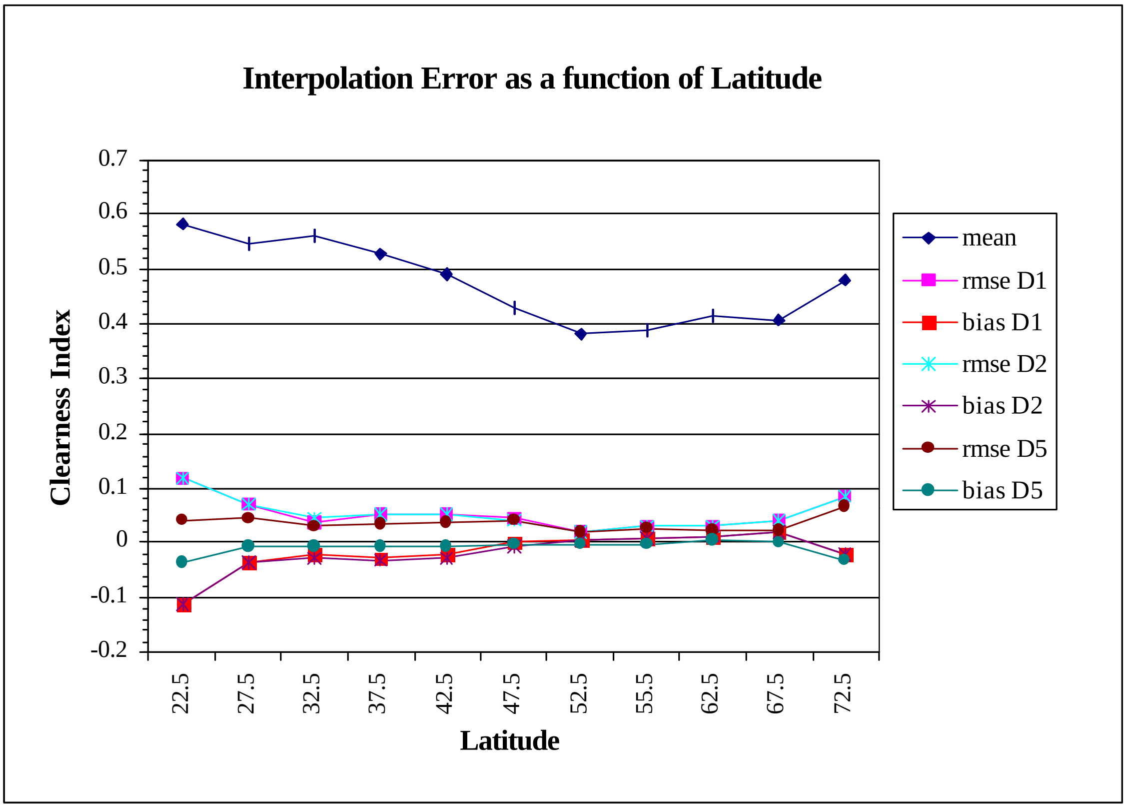


Figure 1. Interpolation errors (bias, rmse) as a function of the latitude (in degrees) for the clearness index, using three different distances *D1*, *D2* and *D5* and the gravity method. The mean value of the measured clearness index is also given.

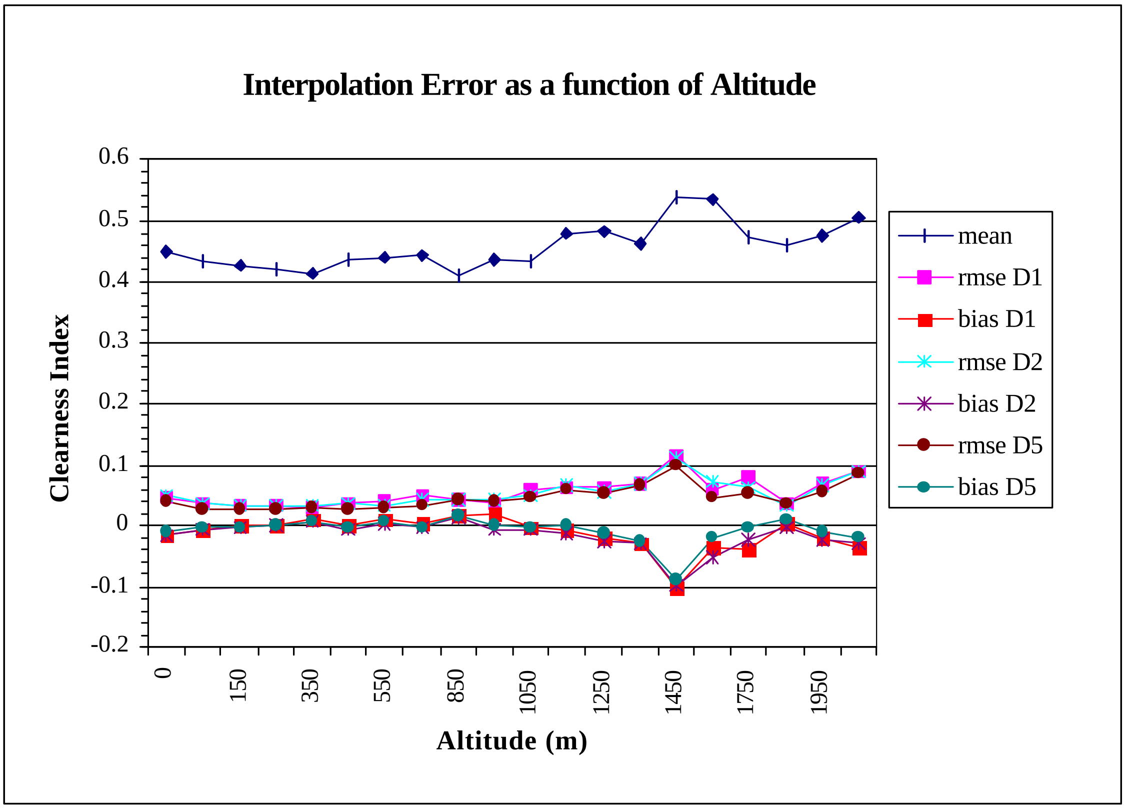
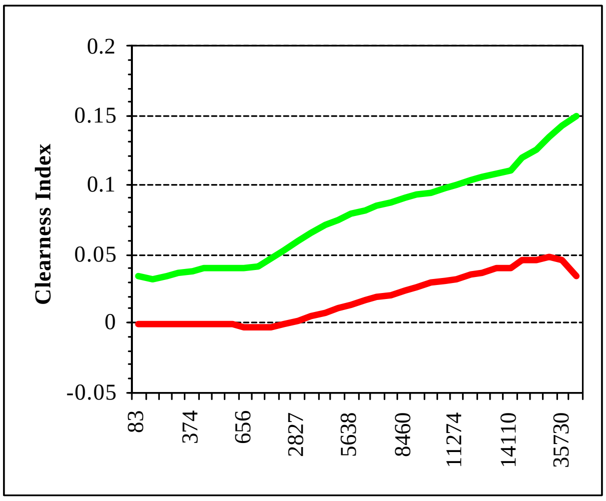


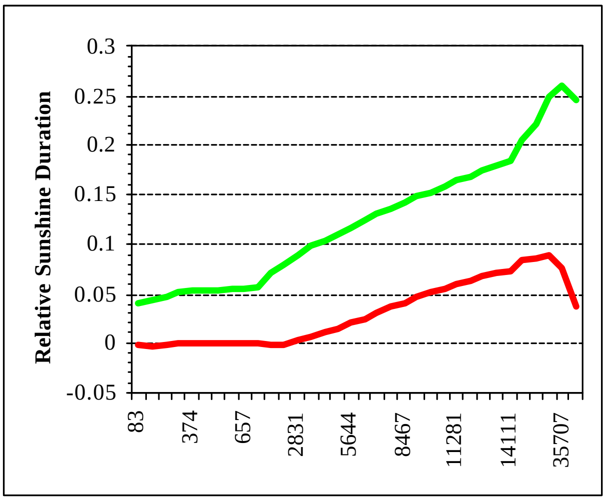
Figure 2. As for Figure 1, but as a function of the altitude (in m).

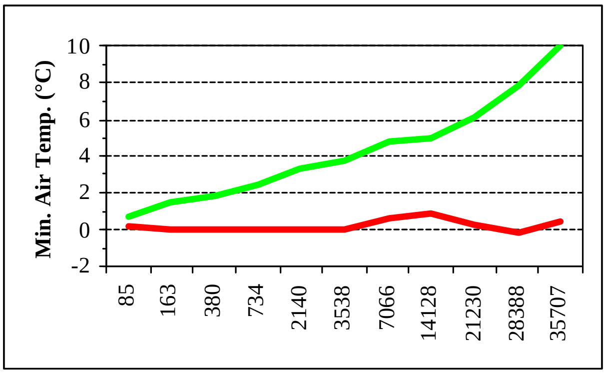
Figure 2 displays the bias and rmse observed as a function of the elevation of the point ***P*** for the clearness index, for the three distances *D1*, *D2* and *D5*. The errors are more or less constant with the altitude, except for sites with an elevation of approximately 1500 m. The DEM TerrainBase gives the elevation of the center of the grid cell. This elevation may differ from that of the measuring site contained within the cell. We observed that the larger the difference in these elevations, the larger the errors in interpolation. This mostly occurs for sites with elevation *circa* 1500 m and thus partly explains the increase of errors for this elevation.

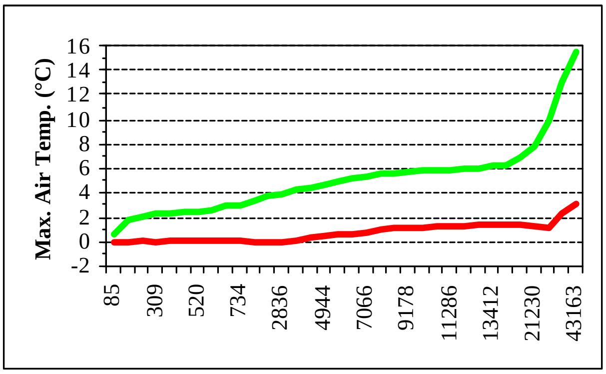
The bias is equal to zero for all distances, except around 1500 m. Indeed, the differences in errors between the distances are very small. Nevertheless, the distance *D5* attains the best results.

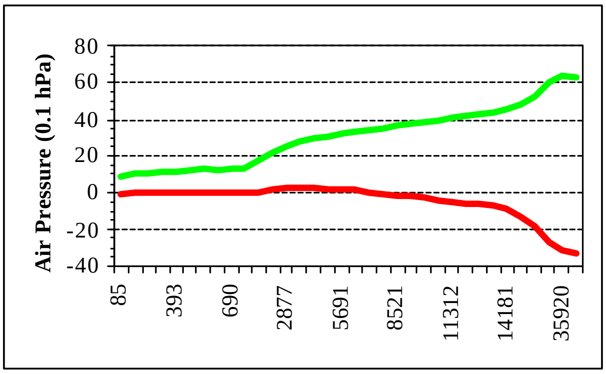
Figure 3 displays the bias and rmse that are expected as a function of the effective distance *D5* for each parameter. These curves may serve to provide an approximate rmse together with an assessment of the meteorological parameter at any geographical point. They have been obtained in the following way. For each point ***P*** and each meteorological parameter, the monthly values are estimated using measuring sites located within a ring at a given effective distance from ***P***, given a tolerance of ±30 % on the distance. This distance was set to various values in an iterative way. For each distance, the errors are computed and synthesized into bias and rmse. Finally, the curves of the bias and rmse as a function of the effective distance are displayed. The effective distance reported in the figure 3 is the mean effective distance of the sites located within the ring. The interpolation method is the gravity method.

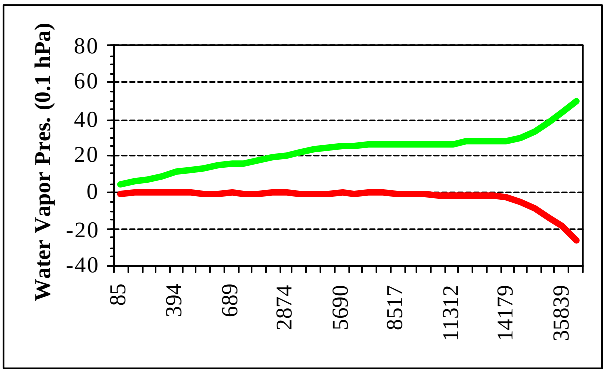

Clearness Index
0.2
0.15
0.1
0.05
0
-0.05
83
374
656
2827
5638
8460
11274
14110
35730


Relative Sunshine Duration
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
83
374
657
2831
5644
8467
11281
14111
35707


Min. Air Temp. (°C)
10
8
6
4
2
0
-2
85
163
380
734
2140
3538
7066
14128
21230
28388
35707


Max. Air Temp. (°C)
16
14
12
10
8
6
4
2
0
-2
85
309
520
734
2836
4944
7066
9178
11286
13412
21230
43163


Air Pressure (0.1 hPa)
80
60
40
20
0
-20
-40
85
393
690
2877
5691
8521
11312
14181
35920


Water Vapor Pres. (0.1 hPa)
80
60
40
20
0
-20
-40
85
394
689
2874
5690
8517
11312
14179
35839

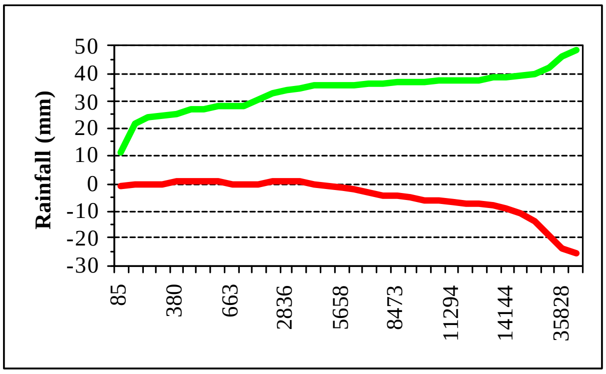


Figure 3. Interpolation errors as a function of the effective distance *D5* (in km) for several meteorological parameters. Bias is in red, rmse in green. The gravity method is used.

Conclusion

This work demonstrates that taking into account the latitudinal effects in the distance brings an increase in the accuracy in interpolation. Such effects have been seldom mentioned in previous publications. The orographic effects may be partly corrected by adding the weighted difference in elevation to the geodetic distance.

We recommend the use of the following effective distance between the point ***P*** and each of the measuring sites $\boldsymbol{X_i}$ for all parameters:

$$d_{eff}^2 = f_{NS}^2 \, (d_{geo}^2 + f_{oro}^2 \, dh^2) \qquad (11)$$

with f_{NS} = *[1 + 0.3 ½$\boldsymbol{F_P}$ -$\boldsymbol{F_X}$½ [1+ (sin$\boldsymbol{F_P}$ + sin$\boldsymbol{F_X}$) / 2]*, where d_{geo} is the geodetic distance in km, latitudes $\boldsymbol{F_P}$ and $\boldsymbol{F_X}$ are expressed in degrees, ***d**h* is the difference in elevation between ***P*** and $\boldsymbol{X_i}$ (expressed in km) and f_{oro} is set to 500.

This distance is identical to the geodetic distance if sites are on the same latitudes and have the same elevation. If the difference in latitude between the point ***P*** and a site $\boldsymbol{X_i}$ is 10° at 45°N, then the effective distance is equal to 6 times the geodetic distance. As for the difference in elevation, assuming no difference in latitude between the point ***P*** and a site $\boldsymbol{X_i}$ and assuming that the geodetic distance is much larger than (500 ***d**h*), the effective distance may be approximated by:

$$d_{eff} \text{ » } d_{geo} \, [1 + (f_{oro}^2 \, dh^2 / d_{geo}^2)/2] \qquad (12)$$

For a difference in elevation of 200 m and a geodetic distance of 200 km, then the effective distance is equal to 1.1 times the geodetic distance. The larger the geodetic distance, the smaller the influence of the difference in elevation.

The orography database TerrainBase, a resource of the prototype of the service SoDa, may provide the elevation for any geographical point. Nevertheless, it is recommended to use the known elevation of the measuring sites.

As for the operational implementation of an interpolation method running in real time, we suggest to narrow the number of stations to the nearest ones and then to use the gravity method.

Acknowledgements

The Commission of the European Communities under Contract Number IST-1999-12245 supported this work. Discussions with the colleagues participating in this project are gratefully acknowledged.

References

Anonymous, 1995. Spatial interpolation of daily meteorological data. Report to the Joint Research Centre, Institute for Remote Sensing, Commission of the European Communities.

Anthes R. A., 1997. *Meteorology*, 7th Edition. Prentice Hall, New Jersey, USA, 214 pp.

Atlas of hydrometeorological data – Europe, 1991. In Russian. Published by Army Publishing House, Moscow, Russia, 371p.

Beyer H.-G., Czeplak G., Terzenbach U., Wald L., 1997. Assessment of the method used to construct clearness index maps for the new european solar radiation atlas (ESRA). *Solar Energy*, 61, 6, 389-397.

ESRA, *European solar radiation atlas*, 2000. Fourth edition, includ. CD-ROM. Edited by J. Greif, K. Scharmer. Scientific advisors: R. Dogniaux, J. K. Page. Authors: L. Wald, M. Albuisson, G. Czeplak, B. Bourges, R. Aguiar, H. Lund, A. Joukoff, U. Terzenbach, H. G. Beyer, E. P. Borisenko. Published for the Commission of the European Communities by Presses de l'Ecole, Ecole des Mines de Paris, Paris, France.

Hou H. S., Andrews H. C., 1978. Cubic splines for image interpolation and digital filtering. *IEEE Transactions on Acoustic, Speech, and Signal Processing*, vol. ASSP-26, 508-517.

Hudson G., Wackernagel H., 1994. Mapping temperature using kriging with external drift: theory and an example from Scotland. *International Journal of Climatology*, 14, 77-91.

Hulme M., Conway D., Jones P. D., Jiang T., Barrow E.M., Turney C. , 1995. A 1961-1990 climatology for Europe for climate change modelling and impact applications, *International Journal of Climatology*, **15**, 1333-1364.

Hutchinson M.F., Booth T.H., McMahon L.P., Nin H. A., 1984. Estimating monthly mean values of daily total solar radiation for Australia, *Solar Energy* **32**, 277-290.

Jones P. G., Thornton P. K., 1999. Fitting a third-order Markov rainfall model to interpolated climate surfaces. *Agricultural and Forest Meteorology*, 97, 213-231.

Journel A., Huijbregts C., 1978. *Mining Geostatistics*. Academic Press, London.

Kunz S., Remund J., 1995. METEONORM - a comprehensive meteorological database and planing tool, *In . 13th European PV-Solar Energy Conference*, pp. 733-735, Nice, France.

Lo F. K., 1989. Estimation des précipitations à partir de l'analyse utilisant le relief pour l'hydrométéorologie (AURELHY). *Physio-Géo*, 19, 37-48.

MeteoNorm, version 4.0, 2000. Version 3.0, 1997. Global meteorological database for solar energy and applied climatology. Published by Meteotest, Bern, Switzerland.

Nalder I. A., Wein R. W., 1998. Spatial interpolation of climatic normals. *Agricultural and Forest Meteorology*, 92, 211-225.

Picinbono B., 1986. Remarques sur l'interpolation des signaux. *Traitement du Signal*, 3, 165-170.

Price D. T., McKenney D. W., Nalder I. A., Hutchinson M. F., Kesteven J. L., 2000. A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. *Agricultural and Forest Meteorology*, 101, 81-94.

Rigollier C., Albuisson M., Delamare C., Dumortier D., Fontoynont M., Gaboardi E., Gallino S., Heinemann D., Kleih M., Kunz S., Levermore G., Major G., Martinoli M., Page J., Ratto C., Reise C., Remund J., Rimoczi-Paal A., Wald L., Webb A., 2000. Exploitation of distributed solar radiation databases through a smart network: the project SoDa. EuroSun 2000, June 2000, Copenhagen, Denmark.

Şen Z., Şahin A. D., 2001. Spatial interpolation and estimation of solar irradiation by cumulative semivariograms. *Solar Energy*, 71, 11-21.

SoDa, 2001. SoDa – Integration and exploitation of networked Solar radiation Databases for environment. *http://soda.jrc.it.*

Supit I., 1994. Global radiation. Published by the European Commission, Office for Official Publications of the European Communities, Luxembourg, CL-NA-15745-EN-C, 194 p.

TerrainBase, 1995. TerrainBase: Worldwide Digital Terrain Data. Documentation Manual, CD-ROM Release 1.0, April 1995. NOAA, National Geophysical Data Center, Boulder, Colorado, USA.

Thiébaux H. J., Pedder M. A., 1987. *Spatial Objective Analysis: with applications in atmospheric science*. Academic Press, 295 p.

Van der Voet P., Van Diepen C.A., Oude Voshaar J., 1994. Spatial interpolation of daily meteorological data: a knowledge-based procedure for the region of the European Communities. Report 53.3, 118 pp, DLO Winand Staring Centre, Wageningen, The Netherlands.

Willmott C. J., Robeson S. M., 1995. Climatologically aided interpolation (CAI) of terrestrial air temperature. *International Journal of Climatology*, 15, 221-229.

WMO, 1981. Meteorological aspects of the utilization of solar radiation as an energy source. Technical Note 172, WMO- No 557, World Meteorological Organization, Geneva, Switzerland, 298 p.

Zelenka A., Czeplak G., d'Agostino V., Josefson W., Maxwell E., Perez R., 1992. Techniques for supplementing solar radiation network data, Technical Report (3 volumes), International Energy Agency, # IEA-SHCP-9D-1, Swiss Meteorological Institute, Krahbuhlstrasse, 58, CH-8044 Zurich, Switzerland.

Zelenka A., Lazic D., 1987. Supplementing network global irradiance data. *In* Advances in solar energy technology. *In* Proceedings of 1986 Biennial Congress ISES Hamburg, Pergamon Press, vol. 4, 3861-3865.