Meteorol. Appl. 7, 75–82 (2000)

A statistical method to downscale temperature forecasts. A case study in Catalonia

Carme Hervada-Sala, *Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya, Colom 1, 08222 Terrassa, Spain*

Vera Pawlowsky-Glahn, Eusebi Jarauta-Bragulat, *Departament de Matemàtica Aplicada III, Universitat Politècnica de Catalunya, Campus Nord, 08034 Barcelona, Spain*

The aim of this work is to present a methodology to downscale weather forecasts (i.e. to give regional-scale forecasts based on synoptic-scale forecasts from a numerical model). This methodology consists of the regionalisation of a large zone into thermally homogeneous meteorological regions followed by a study of how different types of weather affect them. Once the meteorological behaviour of each region is determined, cokriging is used to predict its temperature knowing the current temperature and type of weather expected (i.e. based on forecasts from a numerical model). Thus, from a methodological point of view, it is shown that geostatistical prediction techniques can be used in the meteorological sciences by combining classical multivariate statistics and space-time prediction techniques. To illustrate this methodology, the maximum and minimum temperatures at 54 observatories in Catalonia in winter and synoptic data from Barcelona's airport have been used. The method proposed in this paper, which represents an improvement in temperature prediction compared to that given by classical statistical methods, allows the forecasting of temperature. This case study shows that forecasts can be easily downscaled and that, to produce daily forecasts, just a pocket calculator is needed. The results are presented for a one-day forecast, but this technique can be applied to forecasts over a four-day period.

1. Introduction

Many different physical phenomena affect the vertical and horizontal temperature distributions, and they interact in a non-linear way. Thus, complex atmospheric equations have to be solved numerically to forecast the weather. Models for different scales of motion have to take into account different physical phenomena. The best-tested and most reliable models are the synoptic ones, which take into account only large-scale motion and so neglect mesoscale factors, which are of great importance at a regional scale, as in the Mediterranean zone. Therefore, there is wide interest in downscaling weather forecasts. All the existing techniques, including mesoscale numerical models, need powerful computers or have operationally limited time constraints. Hayden & Purser (1995) developed a recursive filter objective analysis that allows the downscaling of variable fields; it is not difficult to apply, but it needs long calculations. In this paper a new method is presented for downscaling forecasts to give a value for each region (or zone). It is easy to apply and, once preliminary multivariate calculations are made, forecasts can be produced with a pocket calculator.

2. Methodology

The proposed methodology consists in combining multivariate statistical clustering techniques with geostatistical techniques. Cluster analysis is a standard method in multivariate statistics (Anderberg, 1973), while geostatistics is an approach that has been seldom used in fields other than mining or the geological sciences. Therefore, a brief resume of geostatistics has been included in the Appendix. Readers not acquainted with the basic hypothesis and purpose, as well as with the standard terminology of this field, should refer to it whenever necessary.

Clustering techniques are used to divide a large area into homogeneous meteorological regions, as well as to classify meteorological situations into weather types (Steps 1 and 2; see later). Those techniques need powerful computers and involve a large number of calculations. However, meteorological regions and weather types do not change frequently, so clustering could be applied once and then the results updated whenever it seems necessary.

Geostatistical techniques are used to determine each region's meteorological behaviour and to forecast the daily temperature in each region (Steps 3, 4 and 5; see later). Geostatistical tools used to characterise each region's meteorological behaviour also need a large amount of data and a good graphics device, but they can be produced once and then updated if necessary. Daily forecasts of weather type are based on the output of a numerical model. Finally the daily temperature

forecast is obtained as the solution to an equation depending on temperature and weather type.

There are five steps that have to be followed.

- *Step 1*. Divide a large area into homogeneous meteorological regions and find their representative stations.
- *Step 2*. Group meteorological situations into weather types.
- *Step 3*. For each representative station and weather type, find the average temperature (called the expected temperature), and construct the real series (i.e. a time series of daily temperature) and the expected series (i.e. a time series of expected temperature).
- *Step 4*. For each representative station, describe the behaviour of the two time series through the semivariograms.
- *Step 5*. Predict the next day's temperature at each representative station by solving an equation involving the real temperature and expected temperature (the next day's expected temperatures are inferred from forecasts given by numerical models).

Steps 1 to 4 are seldom used so their results need only be updated whenever it seems necessary.

3. Cluster analysis

3.1. Dividing a large area (Step 1)

To divide the whole region into thermally homogeneous zones, a classical clustering method has been applied to histograms of maximum (T_{max}), minimum (T_{min}), average $\{(T_{max} + T_{min})/2\}$ and oscillation ($T_{max} - T_{min}$) temperatures. No geographical restriction (such as height above sea level or other orographic features) has been imposed. This allows for the presence of geographically dispersed points inside a group.

Regionalisation, or clustering, of different stations on the basis of their temperatures allows the division of a synoptic zone into smaller regions – at the mesoscale – that are thermally homogeneous. To do this, five-day temperature histograms are built for each station. The temperature data used are daily maximum, minimum, average and oscillation values. Then all observatories are grouped to obtain smaller regions with a homogeneous meteorological behaviour. This method has been used in France by SCAB (1987) and in Catalonia by Hervada *et al.* (1990).

Then, representative stations for each temperature region have to be identified.

3.2. Identifying weather types (Step 2)

As there are many synoptic variables (i.e. temperature, wet bulb temperature, wind components, pressure and sky cover) and they are not independent of each other, a principal component analysis was performed to give a smaller number of independent factors prior to classifying synoptic situations into weather types. The basic idea is that days with similar factor scores will be meteorologically alike. Then, Euclidean distance and average linkage are used to cluster them into weather types. Using this analysis it has been possible to determine 13 different types of winter weather in Catalonia (see Hervada & Fernandez-Mills, 1989).

4. Geostatistical techniques

4.1. Building two time series (Step 3)

All days have been grouped according to weather types, and then the average temperature for each weather type and station has been calculated. Then, two time series of average temperatures have to be constructed for each station:

- The first series, referred to as the real one, assigns to every day its average temperature.
- The second series, referred to as the expected one, assigns to every day the average temperature for the station based on its type of weather.

The real and expected series show the behaviour of temperature in every zone by means of their semivariogram and cross-semivariogram models.

4.2. Modelling the behaviour (Step 4)

Three semivariograms have to be built for each region for each year: two belong to the real and expected series, and the last is the cross-semivariogram between the two series. The first two semivariograms account for the auto-correlation inside each series and the last one accounts for the cross-correlation between them. The semivariogram function must be modelled from experimental data, as described in the Appendix.

To prove that semivariogram models are the correct ones, cross-validation is used. Cross-validation takes away a known value and estimates it with neighbouring values and the chosen semivariogram model.

Deutsch & Journel (1992) give a program to construct and validate semivariograms; the models have to be adjusted by experience and by trial and error (Clark, 1982; Olea, 1995). If a model for all semivariograms is known, then it is possible to estimate the temperature value from values at neighbouring points.

4.3. Forecasting (Step 5)

To forecast the temperature for each of the next four days in every zone, ordinary cokriging has been used. Cokriging (Deutsch & Journel, 1992) is a geostatistical

tool consisting of a set of generalised linear regression techniques for minimising an estimation variance defined from a prior model for semivariograms using more than one variable. In our case, temperature will be estimated from past temperature values ('real data') and past and future weather types ('expected data').

To solve ordinary cokriging equations, real data for the last five days, and expected data for the last three and next four days are required. The main reasons for using cokriging are that, assuming stationarity of the increments, it produces a linear, unbiased estimator with minimum error variance which, performed on a regular grid, keeps the same coefficients for each prediction. In addition, cokriging optimises the use of available information because it uses not only semivariograms but also cross-semivariograms to account for the auto-correlation as well as for the cross-correlation in time of the two variables.

Then, to forecast temperature T, the following first-order equation has to be solved:

$$T = \sum_{i=-4}^{0} R_i \lambda_i + \sum_{j=-2}^{4} E_j \mu_j \qquad (1)$$

where R_i and E_j are the real and expected series and λ_i and μ_j are weights.

Cokriging consists in finding the weights by using the information contained in the semivariograms. Weights used in equation (1) must satisfy two conditions. They must be such that:

- The estimate given by equation (1) is unbiased.
- The estimation error variances are the smallest possible.

Therefore, this is a classical minimisation problem subject to two constraints. The Lagrange multiplier method may be used to minimise a function with two constraints. For a detailed explanation see Isaaks & Srivastava (1989). Some subroutines to solve the equations are given by Deutsch & Journel (1992).

5. Case study

5.1. Database

Naturally, the accuracy of this method increases with the number of data, so to begin with, as many measuring points as possible are needed. For each station, the maximum and minimum temperatures and other meteorological variables are registered daily. Once they have been grouped, just representative points (i.e. one for each cluster) are used. Each representative point is the station whose thermal histogram is closest to the average histogram for the group. A synoptic station, where temperature, humidity, pressure, sky cover and wind are registered four times per day, is also needed to characterise the meteorological situations. The data from every measuring point has to cover a period of 30–50 years. The period has to be the same for all stations and taken at the same time and height (Conrad & Pollak, 1950; Baker, 1975; Edwards, 1982). Every season is analysed separately to avoid seasonal cycles. To group the stations and the meteorological situations five days' average values have been used because they allow the simplification of the data records and do not add much error – the average cyclone life is more than five days.

The available database consists of data from 54 observatories spread over Catalonia (the north-eastern part of the Iberian peninsula) plus information from the synoptic station at Barcelona's airport (see Figure 1). These observatories belong to the INM, the Spanish Institute of Meteorology. Daily maximum and minimum temperatures have been recorded from 1957 to 1986. Missing data have not been considered because they were less than 5% of the total. If days without data had been significant, Alexanderson's method would have been used (see Hanssen-Bauer & Forland, 1994). To avoid periodicities, only winter temperatures have been considered in the present study.

5.2. Results

The weights used in equation (1) have been found for each of the nine regions by solving the cokriging equations. They are shown in Table 1. For example, the equation for the cool coastal region takes the form:

$$\begin{aligned} T_1 = {} & 0.119R_{-4} + 0.004R_{-3} + 0.138R_{-2} - 0.031R_{-1} + \\ & 0.771R_0 - 0.175E_{-2} - 0.005E_{-1} - 0.466E_0 + \\ & 0.589E_1 + 0.021E_2 + 0.031E_3 - 0.003E_4 \end{aligned}$$

where T is the forecast temperature, R and E are the real and expected series, and the numbers are the weights. The subscripts indicate chronological position: 1 stands for the day for which temperature is going to be estimated and 2, 3, and 4 refer to the following days; 0, –1, –2, –3, and –4 refer to the previous days.

These weights have been used to reconstruct all 29 studied winters and to compare the results with real temperatures. Forecasting errors, as shown in Table 2, provide evidence that cokriging is a better estimator for daily temperature than the classic statistical technique based on the prediction of weather types (i.e. estimating daily temperature as the average belonging to the type of weather that is forecast). In fact, cokriging accounts for the type of weather expected as well as for the atmospheric temperature from some previous days. Cokriging errors – the differences between real and forecast temperatures – depend mainly on the observing site. On the other hand, errors using the classic statistical method can be large, and they depend on both the weather type and the observing site.

Figure 1. *Location of the 55 observing stations in Catalonia.*

Table 1. *Weights for the real series (λ_i) and for the expected series (μ_j) for each region to be used in equation (1).*

Region	Weights											
	λ_{-4}	λ_{-3}	λ_{-2}	λ_{-1}	λ_0	μ_{-2}	μ_{-1}	μ_0	μ_1	μ_2	μ_3	μ_4
High mountain	0.127	–0.035	0.019	0.021	0.868	–0.146	–0.069	–0.716	0.839	0.051	0.050	–0.008
Mean mountain	0.145	0.027	0.150	–0.044	0.722	–0.145	–0.047	–0.288	0.456	0.049	–0.010	–0.014
Plateau	0.128	0.035	0.153	–0.004	0.688	–0.145	–0.047	–0.288	0.456	0.049	–0.010	–0.014
Central basin	0.098	0.012	0.156	–0.033	0.766	–0.200	0.018	–0.630	0.833	0.003	–0.009	–0.014
Hollow	0.121	0.017	0.156	–0.017	0.724	–0.190	–0.020	–0.232	0.352	0.044	0.053	–0.007
Prelittoral basin	0.097	0.010	0.180	–0.003	0.716	–0.270	–0.016	–0.385	0.575	0.040	0.070	–0.014
Cool coast	0.119	0.004	0.138	–0.031	0.771	–0.175	–0.005	–0.466	0.598	0.021	0.031	–0.003
Mean coast	0.111	0.010	0.164	–0.017	0.732	–0.153	–0.049	–0.370	0.518	0.053	0.004	–0.004
Warm coast	0.114	0.032	0.185	–0.002	0.672	–0.193	–0.047	–0.283	0.449	0.055	0.024	–0.006

Table 2. *Forecasting errors for the cokriging technique using the NWP forecast of weather type and the classic statistical method which assigns average temperature for each kind of types of weather for every region.*

Region	Cokriging error (°C)	Types of weather error (°C)
High mountain	2.6	4.9 to 3.2
Mean mountain	2.2	3.6 to 2.7
Plateau	2.1	4.0 to 2.4
Central basin	1.6	10.9 to 2.2
Hollow	1.9	4.0 to 1.7
Prelittoral basin	1.8	3.2 to 1.7
Cool coast	1.9	3.1 to 2.0
Mean coast	1.7	3.1 to 1.5
Warm coast	1.6	3.1 to 1.8

Sometimes it is not possible to receive the forecasts from an NWP model. Therefore the cokriging technique has been tested using the previous day's weather in place of the NWP forecast to assess whether this change increases the error of the prediction. These errors have also been compared with those found using the classic statistical method based on types of weather; the results are shown in Figure 2. There is evidence that:

- The classic statistical method always gives a greater error than use of the cokriging technique with NWP forecasts.
- Cokriging using the previous day's synoptic situation gives a larger error, but not significantly so.

These conclusions are supported by the errors at a cumulative frequency of 50% and 90% for the three methods given in Table 3.

Table 3. *Error at the 50% and 90% levels of cumulative frequency for the three techniques tested.*

Technique	Error at various levels of cumulative frequency (°C)	
	50%	90%
Cokriging using the NWP forecast of weather type	1.4	3.8
Cokriging using the previous day's weather type	1.5	4.0
Classic statistical method	2.5	5.0

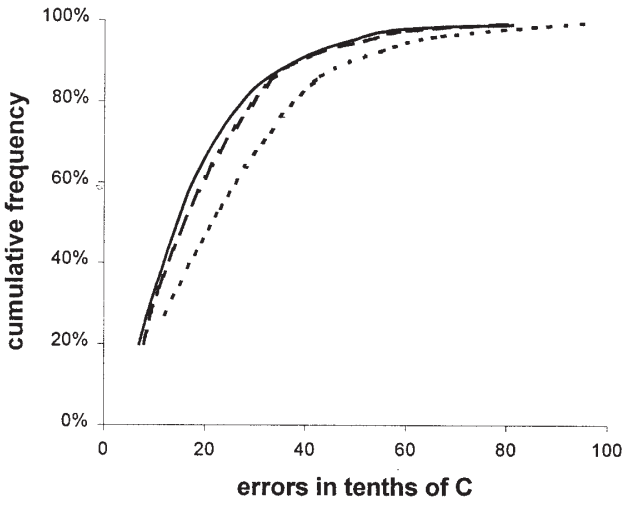

Figure 2. *Cumulative frequency of the forecasting errors for the cokriging technique using the NWP forecast of weather type (full line), the cokriging technique using the previous day's weather type (dashed line) and the classic statistical method (dotted line).*

6. Conclusions

The main conclusions from this study are:

- The methodology proposed in this paper allows the downscaling of forecasts.
- The cokriging technique is a good method for forecasting temperatures because it takes into account past temperatures as well as past and expected types of weather.
- If the weather type of the previous day is used to forecast, errors do not affect significantly the prediction values.
- Multivariate calculations and meteorological behaviour, which determine the weights in the cokriging equations, have to be calculated just once and then updated if necessary.
- Once weights for cokriging have been determined at each region, the forecast of temperatures at any other time can easily be computed.

However, there is a restriction: as semivariograms show a short-term range (maximum range is five days), only short-term forecasts will be improved. It is not advisable to use geostatistical tools to make predictions for more than a semivariogram's range. So this methodology can be applied to temperature forecasts over a four-day period.

Acknowledgements

The authors are deeply indebted to Dr Riddaway and the referees for their advice and critical review of the manuscript.

Appendix. Basic concepts in geostatistics

A1. Definitions

(a) Regionalised variables and random functions

Geostatistics (or surface pattern analysis) are used to study continuously varying phenomena (i.e. phenomena that take a value at each point of space or time). These phenomena are termed regionalised variables in geostatistics.

A regionalised variable can be defined as a function taking a value $z(x)$ at each point of the space R^n, where $z(x)$ depends only on x (the geographical-time location). In this paper $n = 1$; that is we consider only one dimension of time per station.

The set of realisations of $z(x)$ is called a random function $\{Z(x), x \in A\}$ (Journel & Huijgbregts, 1978), which is defined by its distribution function:

$$F_Z(x) = \mathrm{P}[Z(x_1) \leq z_1, \ldots, Z(x_n) \leq z_n]$$

It is not possible to make a statistical inference with a single realisation of the random function because its distribution function cannot, in general, be known. Thus some assumptions about the statistical behaviour of the random function have to be made in order to proceed with the statistical analysis. Geostatistical models can become increasingly elaborate by increasing the number of assumptions, but it is better to keep to the minimum set of assumptions in order to describe in a realistic way the analysed phenomenon.

Most assumptions suppose that the regionalised variable repeats itself in space-time and that the sampling is representative of the regionalised variable; i.e. if we were able to repeat the sampling over the same study area A we would obtain the same results (statistically). In linear geostatistics the assumptions made are stationarity and isotropy.

(b) Semivariogram

The semivariogram $\gamma(h)$ is an alternative to the auto-covariance function and is calculated as half the variance of the difference between temperature at two days, separated by a lag h:

$$2\gamma(h) = \mathrm{Var}[T(u+h) - T(u)]$$

The alternative to the cross-covariance function between two different random functions is known as the cross-semivariogram and measures the cross-variability of the two functions.

Traditionally, the semivariogram has been used for modelling spatial variability rather than temporal variability and is the key to any geostatistical study. In essence, the semivariogram replaces the Euclidean distance h by a structural distance $2\gamma(h)$ that is specific to the attribute and the field under study. The semivariogram distance measures the average degree of dissimilarity between an unsampled value $T(u)$ and a nearby data value. The more 'dissimilar' sample value should receive a lesser weight in the estimation of $T(u)$.

(c) Stationarity

The stationarity assumptions are taken on the first- and second-order moments of the distribution function (Journel & Huijgbregts, 1978).

- The first-order moment does not depend on location:

  $$\mathrm{E}[Z(x)] = m \qquad \forall x \in A$$

- For the second-order moments, the variance, covariogram function and semivariogram function are given by:

  $$Var[Z(x)] = \mathrm{E}[(Z(x) - m(x))^2]$$

  $$C(h) = \mathrm{E}[\,Z(x)\,Z(x+h)] - m^2$$

  $$\gamma(h) = \tfrac{1}{2}\mathrm{E}[(Z(x) - Z(x+h))^2]$$

- The following identities are useful:

  $$Var[Z(x)] = C(0)$$

  $$\gamma(h) = C(0) - C(h).$$

Both $\gamma(h)$ and $C(h)$ can be used to characterise the structure of spatial variability of $Z(x)$ by means of their estimators (see below). Variations of these functions are also employed in the geostatistical literature, such as correlogram, relative variogram and madogram (Cressie, 1991; Maynou *et al.*, 1996).

The function $\gamma(h)$ is more general than $C(h)$ because it only assumes that the variance of the increments is finite. The concept of finite variance of the increments in the semivariogram is called the intrinsic hypothesis (Matheron, 1971). This hypothesis is weaker than the assumption of second-order stationarity and includes it.

A2. Structure functions

(a) Experimental semivariogram

The underlying autocorrelation function of the random function $Z(x)$ is often unknown, but under the stationarity conditions (see above) it can be estimated from the observations $Z(x_i)$. Experimental semivariograms $\overline{\gamma}(h)$ are computed and later fitted to a theoretical semivariogram model. The spatial auto-correlation descriptor is computed as follows:

$$\overline{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (x_i - x_j)^2$$

where $N(h) = \mathrm{card}\{(x_i, x_j)\colon x_i - x_j = h;\ i, j = 1, \ldots, n\}$ is the number of pairs used to compute the experimental semivariogram for each distance class.

(b) Experimental cross-semivariogram

This measure of cross-variability is defined as half of the average product of h increments relative to two different functions:

$$\overline{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (z_i - z_j)(y_i - y_j)$$

where $N(h)$ is the number of pairs used to compute the experimental semivariogram, z_i is the value of function Z at the tail of the pair, z_j is the corresponding head value, and $(y_i - y_j)$ is the corresponding increment of the other function Y.

(c) Semivariogram models

To proceed with the geostatistical modelling the experimental semivariogram – or cross-semivariogram – needs to be fitted to a theoretical semivariogram model. The latter has to comply with some mathematical conditions (Matheron, 1971). Some models grow continuously up to a certain range a where they stabilise around the sill $C(0)$. The range a is the distance beyond which there is no spatial correlation between samples.

The function $\gamma(h)$ reaches the sill asymptotically only in the exponential model and oscillates around the sill in the wave (or hole-effect) model.

There exist a number of methods for fitting the experimental semivariogram to a theoretical model. The most common consists in graphically fitting the two or three parameters that define the theoretical model. This simple method is very useful for well-defined experi-

mental semivariograms that easily suggest the semivariogram function to use.

(c) Nugget effect

Although strictly $\gamma(0) = 0$, the function $\gamma(h)$ often shows a discontinuity at the origin owing to small-scale phenomena: this is the so-called nugget effect. It is due to the fact that, at a scale smaller than the minimum distance between samples, $\gamma(h)$ cannot be studied from the experimental data set. It can be interpreted in a number of ways, depending on the phenomenon under study, as white noise, measurement error, or the existence of spatial structure at distances smaller than minimum value of $|h|$ (Cressie, 1991). When the nugget effect is observed over the entire range of h (flat semivariogram), the absence of spatial structure at the scale of study can be assumed.

A3. Spatial prediction

One of the most important aspects of spatial modelling concerns the possibility of making a statistical inference of a spatial autocorrelated variable. The process of estimating the value of $Z(x)$ at unknown locations is called kriging. The predictor (estimator) of $Z(x)$ at x_0 is taken to be $Z^*(x_0)$.

If $Z(x)$ is a stationary random function, then the difference $R(x) = Z(x_0) - Z^*(x_0)$ is also a random function. As $Z(x_0)$ is unknown, $R(x)$ is also unknown, but under the stationary hypothesis (see above) the computation of its two first moments becomes possible: $m_R = \mathrm{E}[R(x)] = 0$ and the estimation variance $Var[R(x)]$ is small.

A linear unbiased estimator $Z^*(x_0)$ of $Z(x_0)$ from the observed values $Z(x_i)$ is proposed. This estimator can be computed analytically with:

$$Z^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i)$$

where λ_i are the weights attributed to each $Z(x_i)$ from $i = 1, 2, ..., n$, subject to $\Sigma_i \lambda_i = 1$, in order to guarantee the weights are unbiased. The vector of weights to be given to each observed value is obtained by solving a linear kriging system of equations. Kriging is then an optimised weighting of the samples using its autocorrelation structure (Matheron, 1971). This system is:

$$-\sum_{j=1}^{n} \lambda_j \gamma(x_i, x_j) - \mu + \gamma(x_0, x_i) = 0, \;\; i = 1 \ldots n$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

where μ is a Lagrange parameter and $\gamma(a, b)$ is the value of the fitted semivariogram model from point a to point b. The kriging (or estimation) variance is:

$$\sigma^2(x_0) = \sum_{i=1}^{n} \lambda_{ii} \gamma(x_0, x_i) + \mu$$

and can be used to give a precision index to the point estimates or to construct confidence intervals for the estimate, under the assumption of normality of the data.

The term cokriging is used for linear regression that also uses data defined on different attributes. In the case of a single secondary variable y, the ordinary cokriging estimator of $Z(u)$ is written as follows:

$$Z^*(x_0) = \sum_{i=1}^{N_1} \lambda_i \; Z(x_i) + \sum_{j=1}^{N_2} \lambda'_j \; Y(x_j)$$

where the λ_i are the weights applied to the N_1 z-samples and the λ'_j are the weights applied to the N_2 y-samples. The sum of the weights applied to the primary variable is set to 1, and the sum of the weights applied to any other variable is set to 0. In the case of two variables, these two conditions are:

$$\sum_i \lambda_i = 1 \quad \sum_j \lambda'_j = 0$$

While kriging requires a model for the Z covariance, cokriging requires a joint model for the matrix of covariance. That is, it needs both Z and Y semivariograms and the cross-semivariogram.

References

Anderberg, M. R. (1973). *Cluster Analysis for Applications*. Academic Press, 359 pp.

Baker, D. G. (1975). The effect of observation time on mean temperature estimation. *J. Appl. Meteorol.*, **14:** 471–476.

Clark, I. (1982). *Practical Geostatistics*. Applied Science Publishers (1st edition 1979), 129 pp.

Conrad, V. & Pollak, L. W. (1950). *Methods in Climatology*. Harvard University Press, 459 pp.

Cressie, N. (1991). *Statistics for Spatial Data*. Wiley-Interscience, 900 pp.

Deutsch, C. V. & Journel, A. G (1992). *Geostatistical Software Library and User's Guide – GSLIB*. Oxford University Press, 2 diskettes + 340 pp.

Edwards, H. B. (1982). Measurements of true mean temperature for determination of climatic trends. *J. Appl. Meteorol.* **21:** 1763–1766.

Hanssen-Bauer, I. & Forland, E. J. (1994). Homogenising long Norwegian precipitation series. *J. Climate*, **7:** 1001–1013.

Hayden, C. M. & Purser, R. J. (1995). Recursive filter objective analysis of meteorological fields : applications to NESDIS operational processing. *J. Appl. Meteorol.* **34:** 3–15.

Hervada, C. & Fernandez Mills, G. (1989). An air mass classification method for Barcelona. *Annales Geophysicae*, Special Issue, 177–177.

Hervada, C., Serra, C., Periago, M. C. & Fernández Mills, G. (1990). La régionalisation thermique à partir de deux méthodes de classement. *La Météorologie*, **35:** 7–16.

Isaaks, E. H. & Srivastava, R. M. (1989). *Applied Geostatistics*. Oxford University Press, 561 pp.

Journel, A. G. & Huijgbregts, C. J. (1978). *Mining Geostatistics*. Academic Press, 600 pp.

Matheron, G. (1971). *The Theory of Regionalised Variables and its Applications*. Cahiers du Centre de Morphologie Mathématique, Fontainebleau, France, 211 pp.

Maynou, F., Conan, G. Y., Cartes, J. E., Company, J. B. & Sardà, F. (1996). Spatial structure and seasonality of decapod crustacean populations on the north-western Mediterranean slope. *Limnol. Oceanogr*. **41:** 113–125.

Olea, R. A. (1995). Fundamentals of semivariogram estimation, modelling and usage. In *Stochastic Modelling and Geostatistics* (edited by J. M. Yarus & R. L. Chalmers), AAPG Computer Applications in Geology, No. 3, 380 pp.

SCAB (Statistiques, Climatologie, Agrométéorologie, Biométéorologie de l'École Nationale de la Météorologie) (1987). Les régions termiques en France. *La Météorologie*, **18:** 9–6.