

# GIS-based construction of baseline climatologies for the Mediterranean using terrain variables

Maureen D. Agnew\*, Jean P. Palutikof

Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, United Kingdom

**ABSTRACT:** A GIS-based method for constructing high-resolution (in space) maps of mean seasonal temperature and precipitation is developed for the Mediterranean Basin. Terrain variables and geographical location are used as predictors of the climate variables at all points on a grid with a 1 km resolution, using a regression-based approach. Variables used for model development include: longitude, latitude, elevation, distance from the nearest coast, direction to the nearest coast, slope, aspect, and the ratio of land to sea within given radii. Seasonal mean temperature and precipitation data, for the observation period 1952 to 1989, were assembled from 248 temperature sites and 285 precipitation sites in order to initialise the regression model. Temperature data from 36 stations and precipitation data from 35 stations were retained for model validation. Climate surfaces were constructed using the regression equations, and refined by kriging the residuals from the regression model and subtracting the result from the predicted 'observation' surface. Latitude, elevation and distance from the sea are found to be the most effective predictors of local seasonal climate. Validation determined that regression plus kriging predicts mean seasonal temperatures with a coefficient of determination ( $R^2$ ), between the expected and observed values, of 0.87 (summer) and 0.97 (winter), and mean seasonal precipitation with an  $R^2$  of 0.46 (autumn) and 0.94 (summer). A simple regression model without kriging yields less accurate results in all seasons except for the temperature data in spring.

**KEY WORDS:** Temperature · Precipitation · Regression · Terrain · Kriging

## 1. INTRODUCTION

The Mediterranean Basin is vulnerable to climate change particularly through the effects on the water balance and the implications for agriculture and domestic and industrial water supply (Brandt & Thornes 1996, Jeftic et al. 1996). However, when scenarios are constructed directly from GCM (general circulation model) output, they are unsuitable for use in impact studies because the spatial resolution is too coarse (von Storch et al. 1993, Palutikof & Wigley 1996). Downscaling techniques are therefore required to generate scenarios with a finer spatial resolution. The first step is to estimate the baseline climatology at a high spatial resolution. Here, as a basis for downscaling, the possibilities of using a GIS (Geographic Information System) to spatially interpolate climate data

from point sources (either station observations or GCM grid-points) in the Mediterranean Basin are assessed, using information such as height above sea level, distance to the sea, and latitude/longitude as predictors.

A GIS is an effective tool for data integration and spatial analysis (Rhind 1991) and has been used to model climate conditions in Scotland (Matthews et al. 1993, Aspinall & Matthews 1994, Cornford 1997), in Great Britain (Lennon & Turner 1995), and in Ireland (Goodale et al. 1998) with some success. The current work uses a grid-based analysis in which the attributes (represented by temperature and precipitation) are a function of the location, expressed as an absolute (latitude, longitude, elevation, etc.) and as a relative (distance to sea, slope, etc.). The underlying hypothesis is that climate at any location is influenced by the environmental attributes of the surroundings. The Arc/Info GIS is used to map (1) observed seasonal means of temperature and (2) precipitation, at an approximate resolution of 1 km.

\*E-mail: m.agnew@uea.ac.uk

High spatial resolution climatologies are required for hydrological and geomorphological applications. The Mediterranean is an area subject to soil erosion and land degradation, problems which may become more severe if global warming trends continue (for example, Poesen & Hooke 1997, Kirkby et al. 1998, Mulligan 1998). Hydrological and geomorphological impact models, at the level of a catchment or slope, require climate variables at a spatial resolution of 1 km or less as input. We therefore use a 1 km resolution DEM to construct a 1 km resolution baseline climatology for temperature and precipitation.

## 2. DATA ACQUISITION

The climate data (seasonal mean temperature and precipitation) were assembled from the CRU (Climatic Research Unit, UK) archives of climate observations, resulting in a network of 248 temperature sites and 285 precipitation sites. Station data were averaged over the period 1952 to 1989 (being the period with the most dense network). Standard seasons were used: December, January, and February for winter, March, April, and May for spring, etc. A point coverage of the climate stations used for parameterising the model was created in Arc/Info (see Fig. 1). The study area is of rectangular dimension (south-west corner 35° N, 10° W; north-east corner 45° N, 30° E). A 1:1 million-scale vector coastline of this area and the administrative boundaries was obtained from the ESRI (1992) Digital Chart of the World (DCW). DCW data are available in 10° × 10° tiles, which were joined with the internal boundaries dissolved using Arc/Info. The vector coastline was then converted to raster (gridded) data at the same resolution as the elevation data, 0.5 dm (decimal minutes). Land cells were assigned a value of 1, and sea cells a value of 0. The digital elevation model (DEM) 30ASDEM forms the basis for spatial interpolation between point sources, and is the preliminary release of

30 arc second (approximately 1 km) digital elevation data from the US Geological Survey (available from USGS 1997).

The terrain and location variables listed in Table 1 were offered as candidate variables in the regression models and were derived from the 1 km resolution DEM. Elevation (ELEV) is included because it is known to be a strong determinant of climate. Vertical lapse rates average 6°C km<sup>-1</sup> in the free troposphere, although there are seasonal and geographical variations (Meteorological Office 1991). Precipitation is in general positively correlated with altitude, although the rate of increase is modified by local factors such as the proximity to large water bodies or rainshadows (Dingman 1981, Daly et al. 1994). The mean elevation within chosen radii (1, 2, 3, 4, 5, 10 km) was also derived using the FOCALMEAN function in Arc/Info to measure the wider influence of elevation at any one location. The maximum elevation (using FOCALMAX) within a wedge of given orientation and radius (ZX<sub>x</sub>) was included to measure the influence of orographic forcing on precipitation, since cells located on the leeward side of mountains should be drier than those on the windward side. This variable will be most appropriate where the topography is fairly regular and the prevailing wind is from a clearly defined direction.

Distance to the nearest coastline (DISTANCE) is included to account for maritime/landmass influences and is generated using a proximity function in Arc/Info. The logarithm of distance was also obtained since Cornford (1997) suggests that maritime influences may not penetrate far inland. Further measures of coastality, i.e. the mean number of land cells (LAN<sub>x</sub>) within chosen radii ( $x = 5, 10, 20, 50, 100, 150$  km), were computed using FOCALMEAN. When wind is from a clearly defined prevailing direction, only windward coasts should be affected by the maritime influence (Cornford 1999). To incorporate this effect, the direction of the nearest coast was computed for 8 compass points.

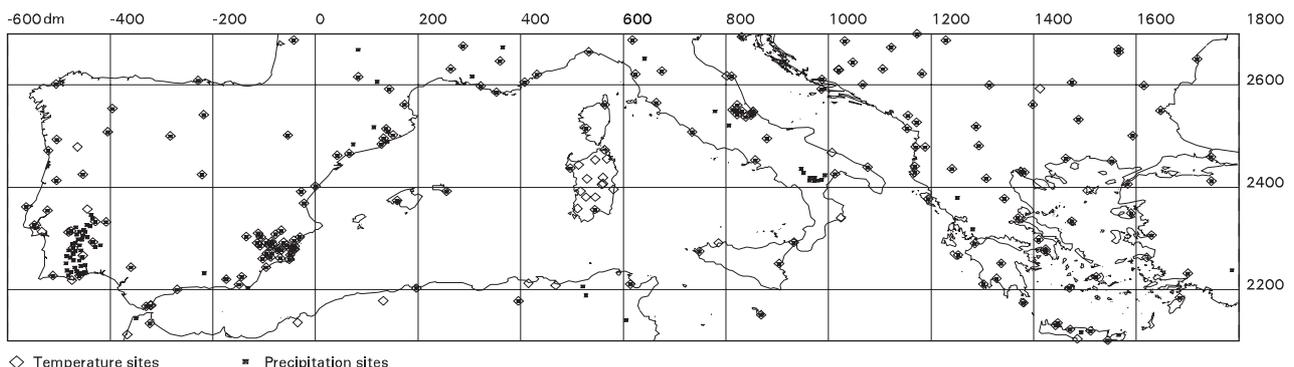


Fig. 1. Geographical location of temperature and precipitation stations used in the study

The longitude and latitude variables (X and Y) are included to parameterise large-scale gradients of temperature and precipitation within the Mediterranean. The direction in which each cell is facing (aspect, ASP<sub>x</sub>) may be important because of the effect on the microclimatology of the wind, and its SLOPE will determine the amount of solar radiation received and hence affect temperature. Differential heating of mountain slopes can trigger convective instability with spatially distinct rainfall patterns. Slope and aspect were derived using functions of the same name in Arc/Info (see Table 1 for a description of these variables).

### 3. METHODOLOGY FOR CONSTRUCTING HIGH SPATIAL RESOLUTION BASELINE CLIMATOLOGIES

A robust algorithm was used to map seasonal temperature and precipitation, based on regression modelling followed by kriging of residuals. This 2-stage methodology involves both deterministic and stochastic components. In the first stage, the observed temperature and rainfall data are interpolated from the station locations to a resolution of 0.5 dm using the terrain and location predictors (Fig. 2). These predictor variables, in the form of digital maps, are used to parameterise the regression model and represent the deterministic component of the model. In the second stage, the residuals from the regression model, treated as random spatial variables, represent the stochastic components and are interpolated using the geostatistical model known as kriging.

Table 1. Location and terrain variables derived from the 1 km resolution digital elevation model (DEM) and used in the regression models

Variable	Description
X	Longitude
Y	Latitude
DISTANCE	Distance to the nearest coast
DIR <sub>x</sub>	Direction of the nearest coast in 8 compass points, where <i>x</i> is N, NE, SE, SW, W, NW. For example, DIR <sub>N</sub> is given a value of 1 if the direction of the nearest coast is to the N, and 0 if the nearest coast is in any other direction. This approach is used to avoid discontinuity between 359° and 0°
LAN <sub>x</sub>	A land-to-sea ratio within a radius <i>x</i> , where <i>x</i> is 5, 10, 20, 50, 100, and 150 km
ELEV	Elevation in metres
ELEV <sub>x</sub>	Mean elevation within <i>x</i> , where <i>x</i> is a radius of 1, 2, 3, 4, 5, and 10 km
ZX <sub>x</sub>	Max elevation within <i>x</i> , where <i>x</i> is a wedge with a radius of 1, 2, 3, 4, 5, 10, 20, 50, and 100 km, with direction N, NE, E, SE, S, SW, W, and NW. For example, ZX <sub>100SE</sub> is the maximum elevation in a SE direction within a 45° wedge with a radius of 100 km
SLOPE	Slope in degrees, i.e. the maximum rate of change in elevation from each cell to its neighbour (where the neighbour is selected to maximise the slope)
ASP <sub>x</sub>	Aspect in degrees, i.e. the down-slope direction of the maximum rate of change of elevation between each cell and its neighbours, where <i>x</i> is N, NE, E, SE, S, SW, W, and NW. For example, ASP <sub>N</sub> is given a value of 1 if the cell slopes to the north, and 0 if the slope of the cell is in any other direction

Stage I: Developing the regression model

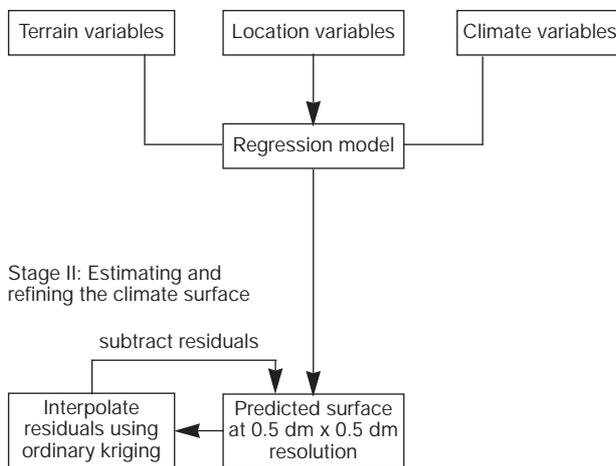


Fig. 2. Two-stage methodology for constructing high-resolution baseline climatologies

#### Stage I: Terrain/climate model parameterisation

(1) The climate and terrain data were converted to raster format within the GIS. Each variable was tested for normality and a functional transformation was applied on a seasonal basis where necessary. For example, the variable representing the ratio of land-to-sea cells in a 150 km radius (LAN150) was exponentially transformed to conform more closely to a normal distribution.

(2) Stepwise multiple linear regression was performed with climate as the dependent variable and terrain/location variables as the predictors. Regression has been used as a satisfactory method for identifying variables used to model climate surfaces in Britain (Matthews et al. 1994, Lennon & Turner 1995, Cornford 1997) and Ireland (Goodale et al. 1998). In total 8 equations were developed for each season and for each climate variable.

## Stage II: Estimating and refining the climate surfaces using geostatistics

(1) Climate fields for each season were estimated using the developed regression models. The residuals from these models were rasterised and subsequently analysed for spatial autocorrelation and trend using a variogram (see Isaaks & Srivastava 1989 for an overview of variograms). The experimental variogram shows semi-variance ( $\gamma$ ) against the separation distance ( $h$ ) between point sources, and can be estimated by:

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^n \{ [z(x_i) - z(x_i + h)]^2 \} \quad (1)$$

where  $n$  is the number of pairs of sample points ( $x$ ), and  $z$  represents the residuals of temperature and rainfall. The variogram was fitted to the semi-variance points using the Levenberg-Marquardt method (Press et al. 1988) of non-linear least squares approximation. The residuals, treated as a spatially dependent variable (Burrough 1986), can be used in a geostatistical sense to partially account for any spatial correlation in the residual temperature and precipitation fields. The residuals were interpolated using local ordinary kriging (see Oliver et al. 1989, Aspinall & Matthews 1994, Cornford 1997) and subtracted from the predicted 'observation' surfaces. Kriging was performed on the data using several different methods (spherical, circular, exponential, Gaussian, and linear). Each method was then tested by constructing the variogram. Models that provided optimal fit to the semi-variance points were selected. The final selection was made on the basis of the RMSE (root mean square error) calculated for the validation samples. In fact, the choice of interpolation algorithm is considered to be less important when a full set of terrain data is available (Cornford 1999), i.e. the choice of mapping algorithm has less effect on the accuracy of the interpolated surface than in the case where little or no additional information is available.

(2) Validation sites were selected to cover a wide range of Mediterranean environments. The models were validated by computing the RMSE and correlation coefficients between the predicted and observed climate for the sample (36 temperature validation stations and 35 precipitation validation stations).

## 4. TERRAIN/CLIMATE RELATIONSHIPS

Deterministic relationships between the observed climate data and the location and terrain variables have been identified on a seasonal basis and are summarised in Tables 2 & 3. For both temperature (Table 2)

Table 2. Temperature correlations with important environment and location variables. Correlation significance: \*0.05 level (2-tailed); \*\*0.01 level (2-tailed). LN: natural log transformation; EX: exponential transformation

	DJF	MAM	JJA	SON
X	-0.35**	-0.15*	0.19**	-0.16*
Y	-0.74**	-0.68**	-0.59**	-0.74**
ELEV	-0.51**	-0.65**	-0.49**	-0.59**
DISTANCE	-0.65**	-0.48**	-0.36**	-0.61**
LAN5	-0.36**	-0.29**	-0.13	-0.34**
LAN20	-0.51**	-0.42**	-0.24**	-0.49**
LAN50	-0.56**	-0.47**	-0.31**	-0.55**
LAN100	-0.63**	-0.51**	-0.37**	-0.62**
LAN150EX	-0.68**	-0.54**	-0.41**	-0.66**
SLOPELN	-0.12	-0.14*	-0.05	-0.14*
DIR_S	-0.02	0.02	0.02	-0.00
DIR_SE	-0.03	0.03	0.10	0.02
DIR_SW	-0.29**	-0.21**	-0.19**	-0.27**
DIR_W	0.06	0.02	-0.01	0.05
DIR_N	0.12	0.06	0.04	0.09
ASP_S	0.05	0.02	0.00	0.02
ASP_SW	0.01	0.02	0.00	0.01

and precipitation (Table 3), the highest correlation coefficients are with latitude (Y). With higher latitudes, temperature decreases and precipitation increases, although the strength of the relationship varies from season to season. For temperature, the relationship is

Table 3. Precipitation correlations with important environment and location variables. Correlation significance: \*0.05 level (2-tailed); \*\*0.01 level (2-tailed). LN: natural log transformation; EX: exponential transformation

	DJF	MAM	JJA	SON
X	0.17**	0.12	0.34**	0.21**
Y	-0.01	0.50**	0.87**	0.46**
ELEV	0.02	0.21**	0.28**	0.08
ZX100W	-0.19**	-0.01	0.22**	-0.01
ZX100SE	0.25**	0.44**	0.41**	0.31**
ZX100S	0.05	0.28**	0.42**	0.12
ZX100SW	-0.12*	0.10	0.35**	0.01
ZX100E	0.34**	0.40**	0.25**	0.41**
ZX50S	0.08	0.27**	0.37**	0.17**
ZX20W	0.02	0.20**	0.27**	0.11
DISTANCE	-0.17**	0.12	0.37**	-0.16**
LAN5	-0.09	0.09	0.11	-0.11
LAN20	-0.11	0.18**	0.22**	-0.10
LAN50	-0.13*	0.20**	0.27**	-0.10
LAN100	-0.15*	0.19**	0.31**	-0.11
LAN150EX	-0.14*	0.20**	0.35**	-0.11
SLOPELN	0.10	0.15*	0.11	0.08
DIR_S	-0.06	-0.05	-0.07	-0.07
DIR_SE	-0.26**	-0.18**	0.01	-0.20**
DIR_SW	0.11	0.21**	0.30**	0.21**
DIR_W	0.18**	0.15*	-0.07	0.15*
DIR_N	0.14*	-0.03	-0.07	0.02
ASP_S	-0.14*	-0.11	-0.06	-0.09
ASP_SW	0.09	0.08	0.04	0.19**

strongest in autumn and winter ( $r = -0.74$ ) and weakest in summer, while for precipitation the relationship is strongest in summer ( $r = 0.87$ ) and weakest (close to zero) in winter.

Elevation is also a strong determinant of climate. Correlations between temperature and elevation are negative as expected ( $r = -0.49$  to  $-0.65$ ), assuming a typical atmospheric lapse rate (Table 2), while correlations between precipitation and elevation are generally positive, although very weak in autumn and winter, and attributable to orographic forcing (Table 3). It is interesting to note that maximum elevation within a wider area (generally radius of 100 km, ZX100x) explains more of the variance than site elevation (ELEV) for precipitation. In spring, for example, the correlation coefficient is 0.21 with ELEV and 0.44 with ZX100SE. Generally there is a positive correlation between rainfall and the maximum elevation in a wedge of given radius and direction. However, if the direction of the wedge is to the west, the relationship is negative in winter. The selected direction of maximum elevation, and the sign of the association, should be related to the direction of the prevailing winds. The strong differences between the circulation patterns of, on the one hand, the western and eastern parts of the Mediterranean Basin and, on the other, winter and summer (Meteorological Office 1962, 1964, Goosens 1985, Bartzokas 1989) help to account for the complexity of the relationship between rainfall and measures of elevation.

Distance from the sea and the land/sea ratio are negatively associated with temperature throughout the year (Table 2). In winter, spring and autumn this may be attributed to the warming influence of the sea. However, in summer we would expect the association

to be positive. In fact, the correlations are lower in summer but still negative. It is possible that the relationship between latitude and distance to the sea has some influence in this case. Precipitation decreases with distance inland (DISTANCE) in winter and autumn, but increases in summer when convective forcing has a greater rôle (Table 3). The correlation coefficients are generally small, the largest being the summer value of 0.37.

Longitude is a significant but less important predictor of seasonal climate on a local scale. In winter, high pressure and therefore cold continental air dominates over central Europe, and hence the eastern basin of the Mediterranean. Simultaneously, temperatures over the western basin of the Mediterranean are moderated by mild maritime airmasses (Wallén 1970). These zonal differences in circulation patterns, of a warmer airmass in the western Mediterranean and a colder airmass in the east, most likely explain the negative relationship with longitude (X), which is strongest in winter but also present in spring and autumn. In summer, when the relationship is positive, the Atlantic Ocean is relatively cold, with the cooling influence of the Atlantic decreasing eastwards in favour of warm continental air. The positive relationship between rainfall and longitude in all seasons is unexpected (Table 3).

For the regression models (Tables 4 & 5), outliers with a strong influence (tested using Cook's Distance) were removed and the models refined. The amount of variance accounted for by the terrain and location variables is higher for temperature than for precipitation, ranging from 65% in JJA (June, July, August) to 92% in DJF (December, January, February). The models do not perform as well for the precipitation data, with the exception of JJA, which has an adjusted coefficient of

Table 4. Summary multiple regression results for temperature. +: positive association, -: negative association. Variable abbreviations are described in Table 1. LN: natural log transformation, EX: exponential transformation. Predictors are given in order of importance

Season	Adjusted R <sup>2</sup>	Variables
DJF	0.92	-Y/-DISTANCE/-X/-LAN150/-LAN5/+SLOPELN/-ELEV/+DIR_W
MAM	0.85	-Y/-ELEV/+SLOPELN/-X/-ASP_NW/-LAN150EX
JJA	0.65	-Y/-ELEV/+X/+LAN20EX/+SLOPELN/-DIR_SW
SON	0.89	-Y/-ELEV/-DISTANCE/-X/-LAN150EX/+SLOPELN

Table 5. Summary multiple regression results for precipitation. +: positive association, -: negative association. Variable abbreviations are described in Table 1. LN: natural log transformation, EX: exponential transformation. Predictors are given in order of importance

Season	Adjusted R <sup>2</sup>	Variables
DJF	0.36	-ZX100W/+ZX100SE/-DISTANCE/-ASP_N/-DIR_SE/+SLOPELN/-DIR_S
MAM	0.55	+Y/+ZX100SE/-ZX100W/+ZX20W/-DISTANCE/+DIR_W/-X/+SLOPELN/-ASP_S
JJA	0.85	+Y/+ZX20W/+ZX100S/-DIR_N/+X/+DISTLN
SON	0.53	+Y/-DISTANCE/+ZX100E/-ZX100W/+ZX50S/+DIR_W/-ZX100SW/+SLOPELN

determination ( $R^2$ ) of 0.85. The poorer results for precipitation are not surprising given the extreme spatial variability inherent in precipitation data and a bias in the distribution of precipitation sites towards lowland locations. The powerful relationship in summer is principally due to the influence of latitude (correlation between rainfall and latitude in this season 0.87) and largely reflects the fact that sites in the south of the study area will have close to zero summer rainfall.

Latitude (Y) occurs consistently in the regression models, and with the exception of the winter (DJF) precipitation model is the most important predictor of climate. Elevation is found to be the second (first in the case of DJF precipitation) most important predictor variable in 6 of the 8 models, either the elevation of the grid cell (ELEV) for temperature or the maximum elevation within a wedge of some radius and direction (ZXx) for precipitation. The most useful radius appears to be 100 km in either a southerly, south-easterly, or easterly direction (ZX100S, ZX100SE, ZX100E).

A measure of maritime influence greatly improves the fit of the models, in agreement with the results of Lennon & Turner (1995), and illustrates the considerable influence of the sea on the Mediterranean climate. Distance from the sea (DISTANCE) appears in the regression equations for autumn and winter temperatures. In both cases, greater distances are associated with lower temperatures, as expected in these seasons. Similarly, an increase in the land/sea ratio is associated with a decrease in temperature in all seasons except summer. Distance to the nearest coastline is included in all precipitation models, but is a more powerful indicator of autumn and winter precipitation (respectively the second and third variable entered).

Longitude, direction to the nearest coast, aspect, and slope are secondary predictors of climate, but can be important in particular seasons. Longitude (X) is included as a predictor of temperature for every season (entered either as the third or fourth variable), but for precipitation is only included in spring and summer as a minor predictor.

The natural logarithm of slope is included in 7 of the models. Aspect is a more significant predictor in the winter and spring precipitation models than in the temperature models. Land sloping to the north or south is drier in winter and spring. The direction to the nearest coast has more significance in the precipitation models. In essence, rainfall is higher if the nearest coast lies between  $247^\circ$  and  $292^\circ$ , the direction of the prevailing wind for western parts of the Mediterranean Basin between mid-October and mid-May (Wallén 1970).

The precipitation regression models are generally constructed using a higher number of explanatory variables than the temperature regression models.

This complexity may be in part a consequence of the smaller-scale systems governing spatial patterns of precipitation (Corte-Real et al. 1995).

It might be expected that a model derived using observational data from a more geographically restricted area, e.g. Iberia, would generate more accurate results than a more general model for the Mediterranean Basin. For example, within a single season the direction of the prevailing winds varies over the present domain, and therefore the impact of variables such as direction to the nearest coast, and direction of areas of high elevation, will also vary across the domain. This being the case, non-linearities in the relationships between the geographical predictor variables and the climate predictands must exist in this geographically complex region, which cannot be properly modelled by linear regression equations. It is perhaps not surprising that the variables selected by the stepwise procedures tend to be those which vary in the 2 spatial dimensions of the domain, either smoothly (e.g. latitude and longitude) or not at all (e.g. elevation).

## 5. REFINING THE REGRESSION MODEL USING GEOSTATISTICS

The KRIGING command in ARC/INFO was used to generate variograms and thus kriged surfaces of the temperature and precipitation residuals. The sample variogram was used to select the 'best' variogram model. Where the choice of method was not immediately apparent, the residuals were kriged using several apparently reasonable kriging models, and the most appropriate method was then selected on the basis of the RMSE calculated for the validation sample.

The variograms computed from the regression model residuals, for each season and climate variable, revealed smaller values of semi-variance for shorter distances between pairs of sample input points (the lags). This suggests that the climate residuals are more similar for stations located closer together than for those stations further apart. Fig. 3 shows the sample variogram generated for summer temperatures. There is evidence of a positive correlation between semi-variance and separating distances up to a range of approximately 80 to 100 dm (160 to 200 km). The range defines the limit of spatial dependence, and is indicated by the point at which the variogram rises linearly to its upper boundary (sill). At distances greater than the range there is no discernible spatial correlation in the residuals. Both Gaussian and spherical methods of estimated semi-variance seem to provide a reasonable fit to the sample data. The experimental variograms did not reveal a gently parabolic concave-upward shape near the origin, which would have indicated the

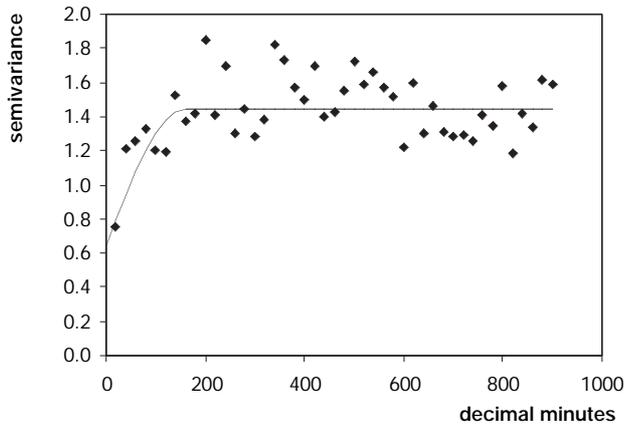


Fig. 3. Sample variogram for spring temperature ( $^{\circ}\text{C}^2$ ). The fit uses a Gaussian model

presence of locally changing linear drift (Journal & Rossi 1989, ESRI 1997). Drift is a systematic change in  $z$  value, here the residuals, in a particular direction. Ordinary kriging was therefore used (as opposed to universal kriging) since the variogram appeared to be free of any structural component (i.e. drift or local trends), and it can be assumed that the variance depends only on the distance between points.

The range found to be most appropriate for fitting the temperature residuals was greater than that for the precipitation residuals by an average distance of 190 km. This implies that spatial dependence exists over a wider area for temperature than for precipitation. This is in agreement with the results of Corte-Real et al. (1995), who found that precipitation in the Mediterranean, being governed by smaller-scale systems, is less coherent in space than temperature.

The kriging variance surfaces produced can be used to assign a degree of certainty to the kriged climate surfaces. Lower values indicate a higher degree of confidence in the interpolated residual field, although this also depends on the accuracy of the variogram model. Table 6 suggests that more certainty can be attached to the interpolated autumn temperature surface (mean kriging variance of  $0.1^{\circ}\text{C}^2$ ) than the summer temperature surface (mean kriging variance of  $1.8^{\circ}\text{C}^2$ ). In contrast, a higher degree of confidence can be attributed

to the interpolated summer precipitation surface (mean kriging variance of  $11\text{ mm}^2$ ) than the winter precipitation surface (mean kriging variance of  $804\text{ mm}^2$ ). The spatial pattern in the variance for spring (MAM) precipitation is shown in Fig. 4, and is typical of that for all climate surfaces generated. Kriging variance is higher (i.e. lower certainty) in central Spain, where there are few climate stations, and around the edges of the domain beyond which there are no stations from which to interpolate.

## 6. VALIDATION OF THE ‘OBSERVATION’ SURFACES

The ‘observation’ surfaces for each climate variable were generated using 2 methods: predicting the climate at every point in the surface using (1) regression alone, and (2) regression followed by kriging of residuals. Then, for the locations of the validation set, the predicted values of temperature and precipitation were extracted for comparison with observations. The results of the validation are shown for the ‘best’ interpolation model from each method in Tables 7 & 8. Overall, validation reveals the temperature surfaces to be more accurate than the precipitation surfaces, although accuracy varies by season and by method of generation. After kriging of residuals, the RMSE for temperature ranges from approximately  $0.8$  (DJF) to  $1.4^{\circ}\text{C}$  (JJA). Values of  $R^2$  show a seasonal pattern similar to the RMSE, varying from  $0.87$  for summer to  $0.97$  for winter. Kriging the residuals improves the RMSE in all seasons except spring (MAM). The greatest improvement is apparent in summer (JJA), when kriging reduces the RMSE by  $0.37^{\circ}\text{C}$ . As measured by  $R^2$ , the improvement varies from an increase of  $0.02$  in spring to an increase of  $0.11$  in summer.

There seems to be greater seasonal variability in the accuracy of the precipitation surfaces (Table 8). After kriging, the RMSE varies from approximately  $4$  (JJA) to  $27\text{ mm}$  (DJF). Values of  $R^2$  display a seasonal variation from  $0.46$  for autumn to  $0.94$  for summer. In agreement with the temperature validation, kriging the residuals improves the RMSE and  $R^2$  in all seasons.

Table 6. Descriptive statistics for the kriging variance surfaces

	Temperature ( $^{\circ}\text{C}^2$ )				Precipitation ( $\text{mm}^2$ )			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
Minimum	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Maximum	1.0	0.8	2.6	0.2	2064.8	435.2	23.3	841.1
Mean	0.4	0.4	1.8	0.1	804.0	304.2	11.3	166.5
Standard deviation	0.11	0.05	0.16	0.02	343.84	68.88	5.23	109.78

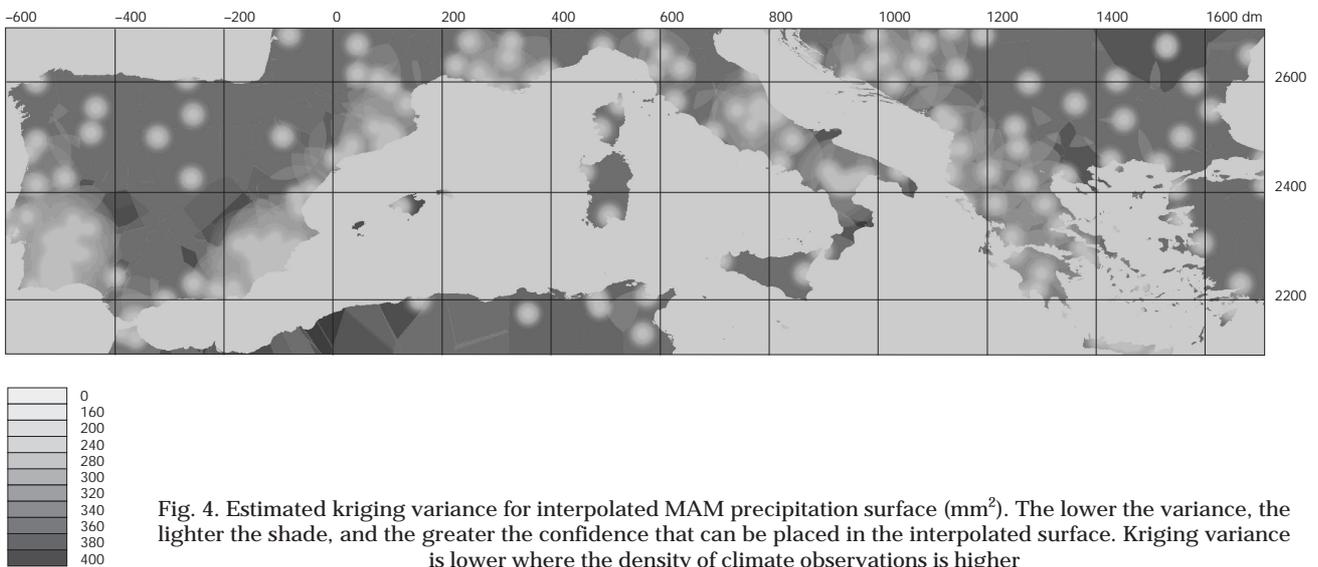


Fig. 4. Estimated kriging variance for interpolated MAM precipitation surface ( $\text{mm}^2$ ). The lower the variance, the lighter the shade, and the greater the confidence that can be placed in the interpolated surface. Kriging variance is lower where the density of climate observations is higher

The greatest improvement is apparent in winter, when the RMSE decreases from 35.47 to 27.10 mm (a reduction of 23.6%), and  $R^2$  increases by 0.35.

Expected values (as predicted by the terrain-climate model) were plotted against observed values for the validation sample. The examples for the winter season are shown in Fig. 5. For temperature, the sample points cluster closely along the diagonal. However, the seasonal precipitation validations were less successful, with greater dispersion from the diagonal. The only exception is for summer precipitation (not shown).

Table 7. Validation of the observed temperature values (1952 to 1989). RMSE: root mean square error

Season	Method 1: Regression		Method 2: Regression + kriging	
	RMSE ( $^{\circ}\text{C}$ )	$R^2$	RMSE ( $^{\circ}\text{C}$ )	$R^2$
DJF	1.16	0.94	0.80	0.97
MAM	1.19	0.90	1.21	0.92
JJA	1.76	0.76	1.39	0.87
SON	1.26	0.91	0.99	0.95

Table 8. Validation of the observed precipitation values (1952 to 1989)

Season	Method 1: Regression		Method 2: Regression + kriging	
	RMSE (mm)	$R^2$	RMSE (mm)	$R^2$
DJF	35.47	0.16	27.10	0.51
MAM	16.79	0.24	13.61	0.50
JJA	7.52	0.82	4.39	0.94
SON	21.05	0.29	19.64	0.46

## 7. TERRAIN-INTERPOLATED CLIMATE SURFACES

The spatial descriptive statistics (minimum, mean, maximum and standard deviation) shown in Table 9 are computed for each climate surface on a cell-by-cell basis. Mean monthly temperatures range from  $5.3^{\circ}\text{C}$  in winter (DJF) to  $21.3^{\circ}\text{C}$  in summer (JJA), with a spatial variance which is greatest in DJF.

Fig. 6 shows the high-resolution temperature surfaces for each season generated by estimating the temperature for each cell (using the terrain regression model) and subtracting the kriged residuals from the regression model. Mountainous areas are clearly depicted due to the strong negative relationship between elevation and temperature. The Pyrenees, for example, are shown to be considerably colder than the surrounding lower elevation cells. This altitudinal temperature gradient is superimposed on a large-scale latitudinal gradient with temperatures decreasing northwards.

Mean monthly precipitation for the surfaces ranges from 34.7 mm in JJA to 74.5 mm in DJF (Table 9). Similar to the temperature surfaces, the spatial variance is greatest in DJF (standard deviation of 39.1 mm). The most prominent spatial patterns of precipitation in the Mediterranean Basin appear to be similar for the autumn, winter and spring surfaces (Fig. 7). Precipitation is greatest along western coasts and is lowest in eastern Spain, northern Africa (approximately 50 km inland), and the interior of Turkey and Romania (i.e. inland from the influence of moisture-laden westerlies) (Fig. 7). This west-east gradient (within each land-mass) is most apparent in autumn and winter. A positive relationship with elevation is also clear with, for example, higher rainfall in the Pyrenees, Cordillera

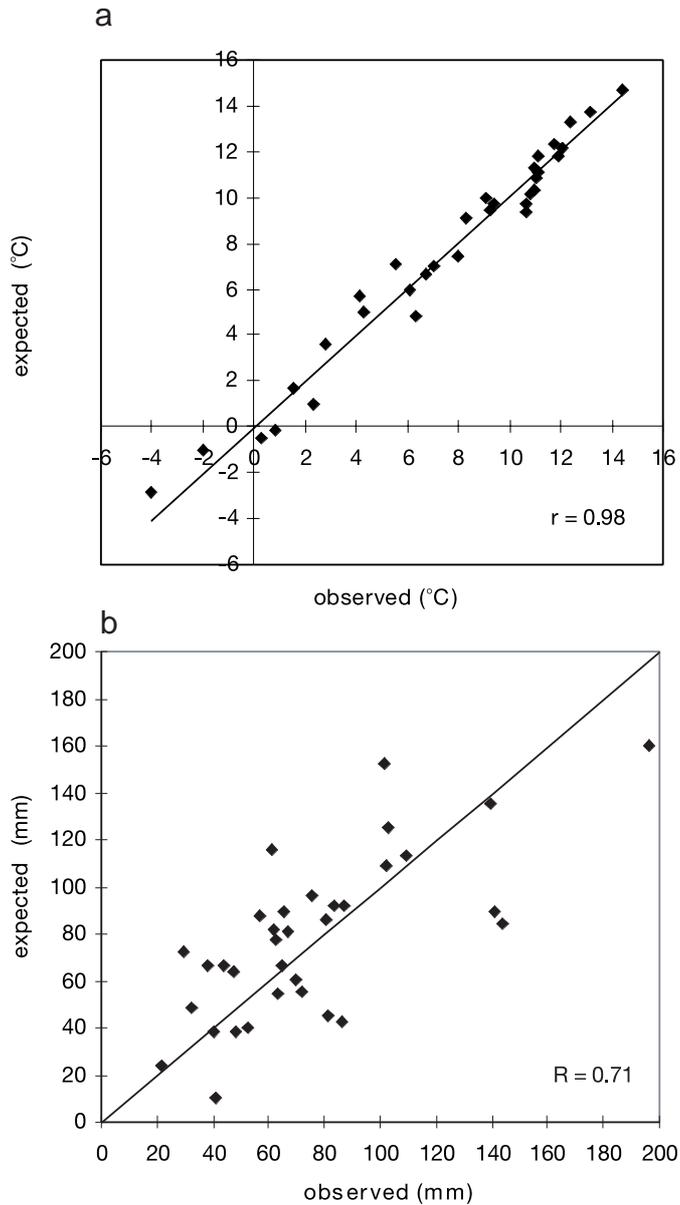


Fig. 5. Scatterplots of observed and expected values for the validation sample. (a) DJF temperature model; (b) DJF precipitation model

Table 9. Basin-wide descriptive statistics for the high-resolution climate fields

	Mean monthly temperature (°C)				Mean monthly precipitation (mm)			
	Min.	Max.	Mean	SD	Min.	Max.	Mean	SD
DJF	-9.9	15.2	5.3	3.6	0.0	272.7	74.5	39.1
MAM	-4.9	18.3	11.6	2.4	1.9	166.0	57.3	19.8
JJA	5.6	27.7	21.3	2.4	0.0	107.2	34.7	21.8
SON	-1.4	21.6	14.0	2.8	0.0	232.5	66.5	30.2

Cantabrica and Cordillera Central in Spain, the southern Italian Alps and the Apennines. Summer rainfall displays a strong latitudinal gradient, increasing northwards, and superimposed on this pattern is a relationship with elevation which is more moderate than in the other 3 seasons but which can be seen, for example, in an increase in precipitation with altitude in the Pyrenees. Relative to the summer surface, the large-scale latitudinal gradient is weaker in spring and autumn and non-existent in winter, when broad-scale elevation patterns are more important in determining precipitation.

**8. MODEL LIMITATIONS AND FUTURE RESEARCH**

A GIS-based technique has been outlined for generating seasonal temperature and precipitation baseline climatologies at a high spatial resolution. This approach involves regression using terrain variables as predictors followed by kriging of the residuals. There are certain limitations to the technique. (1) The terrain variables are all derived from a DEM which is known to contain errors. The estimated absolute vertical accuracy of the DEM (at 90% confidence) is 30 m linear error, or a RMSE of 18 m (assuming a Gaussian distribution with a mean of zero) (DMA 1986). (2) The generation of terrain variables from the DEM is another source of error which varies geographically. Florinsky (1998) found that errors for DEM-derived local topographic variables (gradient, aspect, horizontal and vertical landsurface curvatures) are typically highest for flat areas.

Two main limitations are noted in the selection of candidate variables for regression: (1) The relationships between climate and terrain variables may vary under different synoptic situations (see Laughlin & Kalma 1987). For example, lapse rates vary with cloud cover and humidity, which are synoptically driven. (2) Several terrain variables have not been included in the analysis which are expected to have an influence on the spatial distribution of temperature and precipitation, e.g. landcover and proximity to large bodies of freshwater.

There are also potential areas of weakness in the kriging component of the algorithm. If the assumptions required to krig a surface are fully met, then kriging by definition will be the best linear unbiased predictor (Weber & Englund 1994, Cornford 1999). However, the assumption of stationarity (i.e. that the same pattern of variation can be observed at all locations on the surface) is rarely met in reality (Fedorov 1989). To test this assumption, the mean and variance of the residuals in blocks of approximately 600 km (roughly equivalent to the range of the variogram plots) were considered. These were found to be broadly similar although there

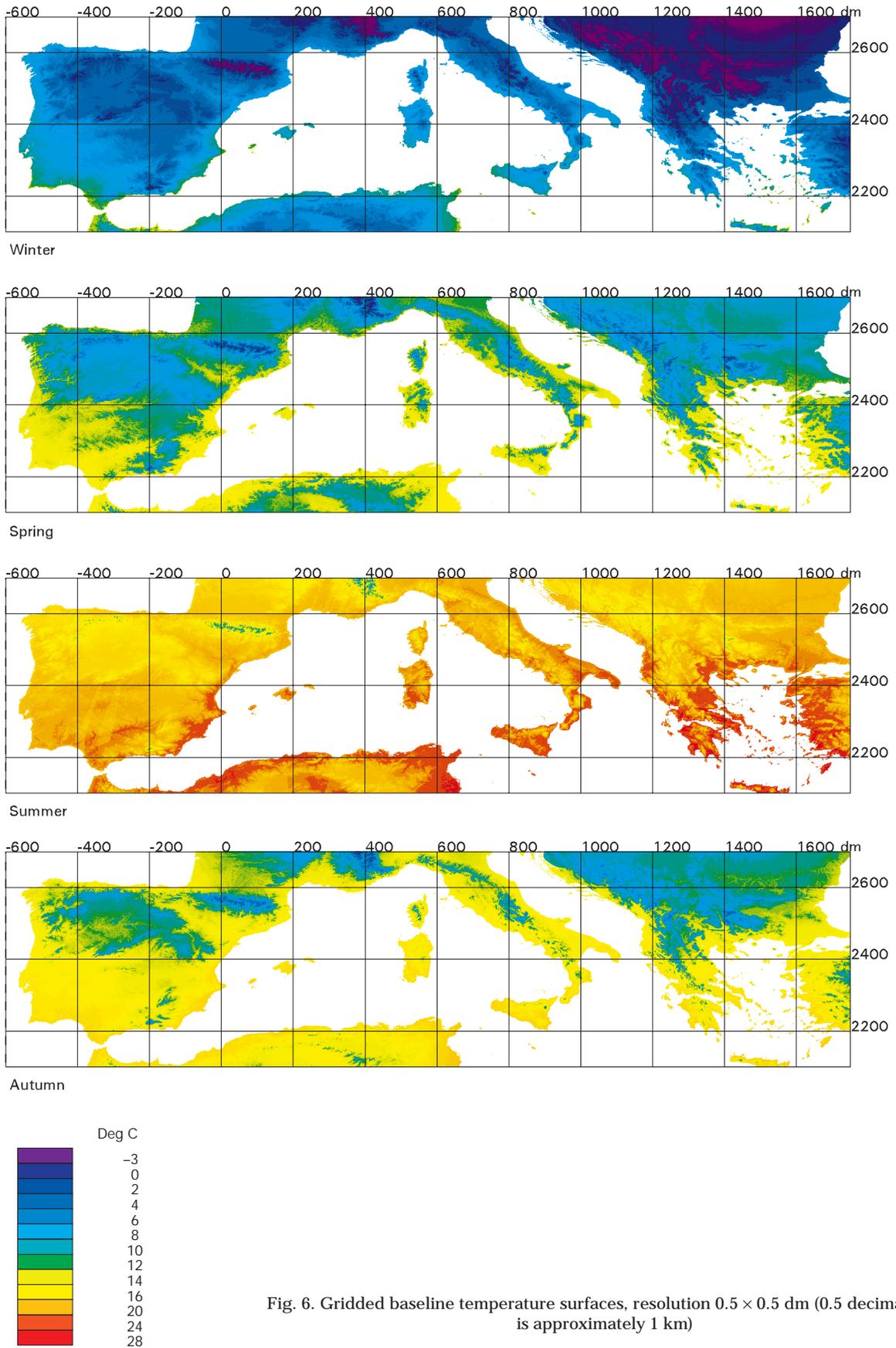
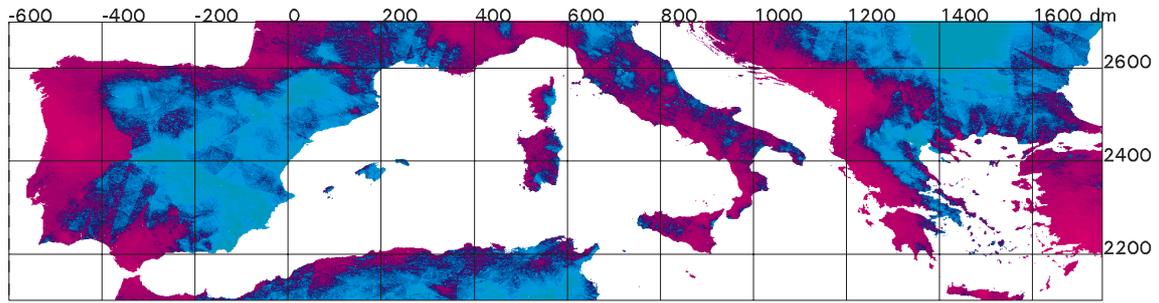
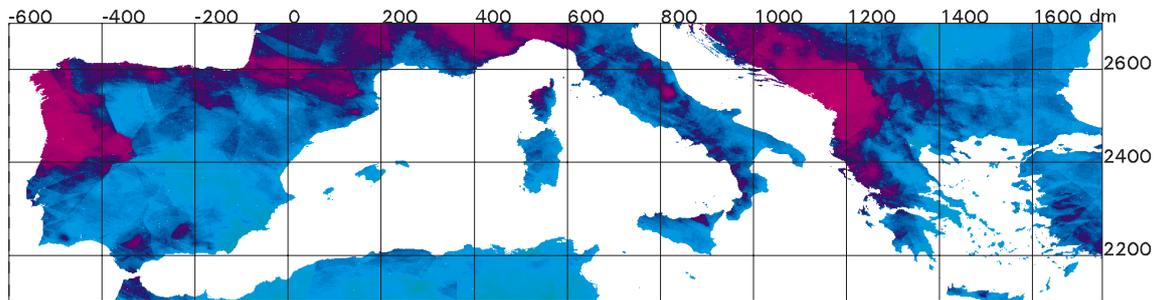


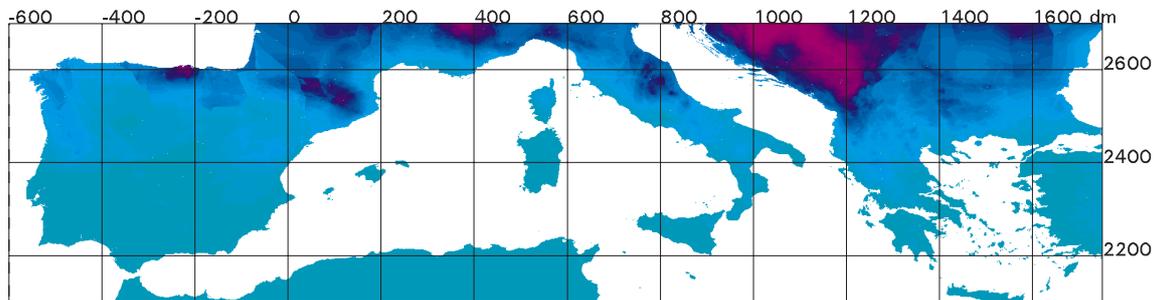
Fig. 6. Gridded baseline temperature surfaces, resolution  $0.5 \times 0.5$  dm (0.5 decimal min is approximately 1 km)



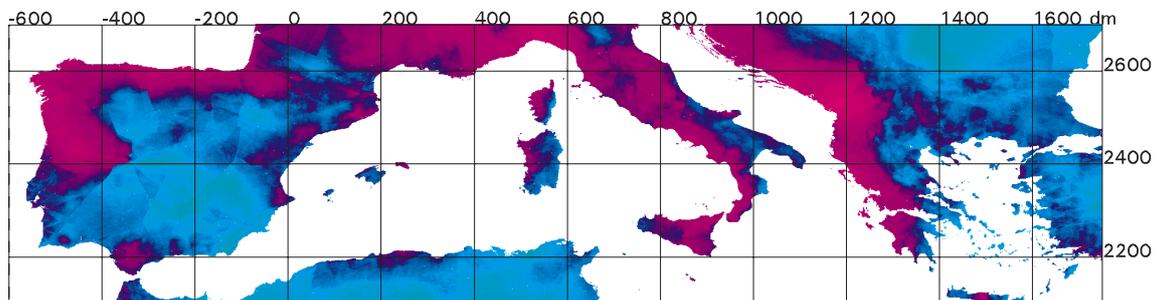
Winter



Spring



Summer



Autumn

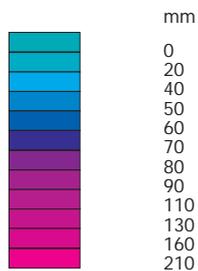


Fig. 7. Gridded baseline precipitation surfaces, resolution  $0.5 \times 0.5$  dm (0.5 decimal min is approximately 1 km)

were notable increases in the means of the residuals in blocks with a high proportion of sea cells as opposed to land cells. In addition, the accuracy of the kriged surface depends to some extent on the density of the network of climate stations. Plots of the semi-variance revealed that there were several areas (particularly central Spain) where more climate sites would improve the accuracy of the interpolation.

Despite the limitations outlined above, a GIS-based approach has considerable potential, especially given the complex land-sea configuration, and the range of topography, characteristic of the Mediterranean. The use of terrain variables to predict climate point data provides a means of constructing a climate surface with a physical basis. Moreover, the results of the validation provide clear evidence of the usefulness of kriging in the spatial interpolation of climate variables.

With respect to future developments of the method, construction of regional regression models may prove useful, since there are well-recognised differences in the circulation patterns for western and eastern parts of the Mediterranean Basin (Wallén 1970, Flocas 1984, Reddaway & Bigg 1996). Related to this, the interpolation of the regression residuals may be improved if the semi-variogram is estimated using a moving-window approach. A further possibility for future research is to use average seasonal synoptic circulation patterns as predictors in the regression model, in addition to the terrain-based explanatory variables. Further research is required to assess the transferability of the GIS-based interpolation method between regions.

## 9. CONCLUSIONS

This paper investigates the application of a GIS to construct DEM-derived terrain variables which are subsequently used to interpolate climate data for every 0.5 by 0.5 km grid cell within the Mediterranean Basin. The results presented suggest that regression model building using terrain variables, followed by kriging of the regression residuals, is a useful tool for interpolating seasonal temperature and precipitation data. The use of geographical and topographical predictors provides an estimation of the local spatial structure of climate which is physically plausible. Interpolation results can be transferred to any grid point in the domain for which the characteristics of the predictors variables are known. Latitude and elevation are found to be the most powerful predictors of local climate. A measure of maritime influence also improves the model fit, a reflection of the importance of oceanic effects in the Mediterranean region. Validation demonstrates that kriging the residuals improves the RMSE for 7 out of the 8 models developed.

The methodology for generating high-resolution climate surfaces has potential application to statistical downscaling of general circulation model (GCM) output. Methods used for downscaling are strongly influenced by the systematic errors of GCMs. The use of a technique which involves both deterministic (terrain and location parameterisation) and stochastic (interpolation of spatially random variables) components may contribute to the reduction of this systematic error.

*Acknowledgements.* This research was funded by the Commission of the European Union under contracts ENV4-CT95-0121 (MEDALUS III) and ENV4-CT97-0448 (WISE). We would like to thank Dan Cornford and anonymous reviewers for their helpful and constructive comments and advice.

## LITERATURE CITED

- Aspinall RJ, Matthews KB (1994) Climate-change impact on distribution and abundance of wildlife species—an analytical approach using GIS. *Environ Pollut* 86:217–223
- Bartzokas A (1989) Annual variation of pressure over the Mediterranean Sea. *Theor Appl Climatol* 40:135–146
- Brandt CJ, Thornes JB (eds) (1996) *Mediterranean desertification and land use*. Wiley, London
- Burrough PA (1986) *Principles of Geographic Information Systems for land resource assessments*. Clarendon Press, Oxford
- Cornford D (1997) The development and application of techniques for mapping daily minimum air temperatures. PhD thesis, The University of Birmingham
- Cornford D (1999) Terrain information and interpolation. In: Conference proceedings EUR 18472 EN: Cost Actions 77-79-711, 28 September–3 October 1997, Volterra
- Corte-Real J, Zhang X, Wang X (1995) Large-scale circulation regimes and surface climatic anomalies over the Mediterranean. *Int J Climatol* 15:1135–1150
- Daly C, Neilson RP, Phillips DL (1994) A statistical-topographical model for mapping climatological precipitation over mountainous terrain. *J Appl Meteorol* 33:140–158
- Dingman SL (1981) Elevation: a major influence on the hydrology of New Hampshire and Vermont, USA. *Hydrol Sci Bull* 24:329–339
- DMA (1986) Defense Mapping Agency product specifications for Digital Terrain Elevation Data (DTED), 2nd edn. Defense Mapping Agency Aerospace Center, St Louis, MO
- ESRI (1992) *Digital chart of the World*. Environmental Systems Research Institute Inc, New York
- ESRI (1997) *GRID command references*. Environmental Systems Research Institute Inc, Redlands, CA
- Fedorov VV (1989) Kriging and other estimators of spatial field characteristics (with special reference to environmental studies). *Atmo Environ* 23:175–184
- Flocas AA (1984) The annual and seasonal distribution of fronts over central-southern Europe and the Mediterranean. *J Climatol* 4:255–267
- Florinsky IV (1998) Accuracy of local topographic variables derived from digital elevation models. *Int J Geogra Inf Sci* 12(1):47–61
- Goodale CL, Aber JD, Ollinger SVV (1998) Mapping monthly precipitation temperature and solar radiation for Ireland with polynomial regression and a digital elevation model. *Clim Res* 10:35–49

- Goosens C (1985) Principal component analysis of Mediterranean rainfall. *J Climatol* 5:379–388
- Isaaks EH, Srivastava RM (1989) An introduction to applied geostatistics. Oxford University Press, Oxford
- Jeftić L, Keckes S, Pernetta JC (eds) (1996) Climate change and the Mediterranean. Arnold, London
- Kirkby MJ, Abrahart R, McMahon MD, Shao J, Thornes JB (1998) MEDALUS soil erosion models for global change. *Geomorphology* 24:35–49
- Journal AG, Rossi ME (1989) When do we need a trend model in kriging? *Math Geol* 21:715–739
- Laughlin GP, Kalma JD (1987) Frost hazard assessment from local weather and terrain data. *Agric For Meteorol* 40: 1–16
- Lennon JL, Turner JRG (1995) Predicting the spatial distribution of climate: temperature in Great Britain. *J Anim Ecol* 64:370–392
- Matthews KB, Allison S, MacDonald AM (1993) Climate change data directory. MLURI Occasional Publication, Macaulay Land Use Research Institute, Aberdeen
- Matthews KB, Macdonald A, Aspinall RJ, Hudson G, Law ANR, Paterson E (1994) Climate soil moisture deficit—Climate and soil integration in a GIS. *Clim Change* 28: 273–287
- Meteorological Office (1962) Weather in the Mediterranean. I. General meteorology, 2nd edn. MO391, HMSO, London
- Meteorological Office (1964) Weather in the Mediterranean. II. General meteorology 2nd edn. MO391b, HMSO, London
- Meteorological Office (1991) Meteorological glossary. HMSO, London
- Mulligan M (1998) Modelling the geomorphological impact of climatic variability and extreme events in a semi-arid environment. *Geomorphology* 24:59–78
- Oliver M, Webster R, Gerrard J (1989) Geostatistics in physical geography. II. Applications. *Trans Inst Brit Geog* 14: 269–286
- Palutikof JP, Wigley TML (1996) Developing climate change scenarios for the Mediterranean region. In: Jeftić L, Keckes S, Pernetta JC (eds) Climate change and the Mediterranean, Vol 2. Arnold, London, p 27–56
- Poesen JWA, Hooke JM (1997) Erosion, flooding and channel management in Mediterranean environments of southern Europe. *Prog Phys Geogr* 21:157–199
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1988) Numerical recipes in C. The art of scientific computing. Cambridge University Press, New York
- Reddaway JM, Bigg GR (1996) Climate change over the Mediterranean and links to the more general circulation. *Int J Climatol* 16:651–661
- Rhind D (1991) Geographic information systems and environmental problems. *Int Social Sci J* 43:649–668
- USGS (1997) 30 arc-sec digital elevation model data. US Geological Survey, Virginia (accessed 3 December 1997); available at <http://edcwww.cr.usgs.gov/glis/hyper/guide/30asdem#30asdem>
- von Storch H, Zorita E, Cubash U (1993) Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime. *J Clim* 6: 1161–71
- Wallén CC (ed) (1970) World survey of climatology, Vol V. Climate of Northern and Western Europe. Elsevier, London, p 1–21
- Weber DD, Englund EJ (1994) Evaluation and comparison of spatial interpolation II. *Math Geol* 26:589–603

*Editorial responsibility: Hans von Storch, Geesthacht, Germany*

*Submitted: January 8, 1999; Accepted: September 19, 1999  
Proofs received from author(s): January 20, 2000*