HYDROLOGICAL PROCESSES
Hydrol. Process. **20**, 2217–2232 (2006)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hyp.6199

Interpolating local snow depth data: an evaluation of methods

J. I. López-Moreno[1]* and D. Nogués-Bravo[2]

[1] *Instituto Pirenaico de Ecología, CSIC, Campus de Aula Dei, Apartado 202, 50080 Zaragoza, Spain*
[2] *Macroecology and Conservation Unit, University of Évora, Estrada dos Leoes, 7000-730 Évora, Portugal*

Abstract:

Snow depth measurements have been taken since 1986 at 106 snow poles distributed in the Spanish Pyrenees. Here, we compared the capacity of several local, geostatistical and global interpolator methods for mapping the spatial distribution of averaged snowpack (1986–2000) and the snowpack distribution in two single years with different climatic conditions. The error estimators indicate that the terrain complexity of the area makes it difficult to apply local and geostatistical methods satisfactorily. Regression-tree models provide an accurate description of the data set used (the calibration phase), but they show a relatively low predictive capability for the study case (the validation phase). Using linear regression and generalized additive models (GAMs), we achieved more robust estimations than by means of a regression-tree model. The GAMs give the most accurate prediction because they consider the non-linear relationships between snowpack and the external characteristics (physical features) of the sampling points. Copyright © 2006 John Wiley & Sons, Ltd.

KEY WORDS snowpack; spatial interpolation; error estimators; central Spanish Pyrenees

INTRODUCTION

Snow accumulation in the Spanish Pyrenees determines fluvial regime and explains the interannual variability of spring discharge (López-Moreno and García-Ruiz, 2004). Given that melting flow contributes to filling reservoirs, thereby ensuring water supply to irrigated areas in lowlands (López-Moreno, 2005; López-Moreno *et al.*, 2004), accurate estimates of the snow accumulated in the basins improve water resource management. Furthermore, snow depth maps are useful tools for other fields, such as avalanche- and flood-risk estimation, the planning of tourist activities, climate variability assessment, etc. (Carroll and Cressie, 1996; Haefner *et al.*, 1997).

One of the most reliable methods for mapping the spatial distribution of snow is the collection of data on local snow depth and density and subsequent interpolation. Therefore, in the last years several studies have focused on testing methods to map snowpack and snow water equivalent (Elder *et al.*, 1998; Yang and Woo, 1999; Balk and Elder, 2000; Chang and Li, 2000; Erxleben *et al.*, 2002; Anderton *et al.*, 2004). Recently, López-Moreno and Nogués-Bravo (2005) applied a generalized additive model (GAM) (Hastie and Tibshirani, 1987) to map snow depth in the Pyrenees. They proposed that GAMs (used mainly in other fields, such as ecology and biogeography) could be a promising tool to interpolate snow data and other climatic or environmental variables.

Here, we compare the most commonly used interpolation methods in snow studies in order to assess their capacity to predict the snow distribution at the end of April for the average of the period 1986–2000 in the central Spanish Pyrenees. In addition to the analysis of a multiyear average data set, single-year data sets were used for interpolating. Thus, a better understanding of the predictive capability of the models is obtained, and

* Correspondence to: J. I. López-Moreno, Instituto Pirenaico de Ecología, Campus de Aula Dei, Apartado 202, E-50080 Zaragoza, Spain. E-mail: nlopez@ipe.csic.es

Received 10 May 2005
Accepted 13 September 2005

it allows for a more balanced comparison with prior studies that usually consider single-year data sets. Two years, the 1994 and 1995 data sets, were chosen since they represent the most contrasting winters relating to the amount of snow accumulated (López-Moreno, 2005). Our approach is featured by a limited number of observations to model snowpack at regional scales. The scarcity of available observations at the large scale in snow studies is a frequent drawback owing to the difficulty in carrying out data collection surveys during wintertime.

THE STUDY AREA

The Pyrenees is an alpine range located in the northeast of the Iberian Peninsula (Figure 1). Altitude increases eastward, exceeding 3000 m a.s.l. in the headwaters of the Gállego, Cinca, Ésera and Noguera Ribagorzana rivers. The relief is organized in parallel bands following a west–east axis, which causes a strong barrier effect against Atlantic fronts, thereby enhancing the transition from northwest (Atlantic) to southeast (Mediterranean) influences. Precipitation ranges between 1000 and 2500 mm year^{-1} (García-Ruiz *et al.*, 2001). Temperature is mainly governed by the altitudinal gradient (Del Barrio *et al.*, 1990). During winter, the 0 °C isotherm is located around 1600 m a.s.l. (García Ruiz *et al.*, 1986). This isotherm allows the persistence of an extensive snowpack until May and isolated snow patches until the end of summer. However, the snowpack shows noticeable interannual variability, which constitutes a source of uncertainty when forecasting the availability of spring water resources (López-Moreno, 2005).

DATA AND METHODS

Data

In 1985, interest in assessing snow accumulation in the Pyrenees led the Office of Hydraulic Works (Spanish Ministry of Public Works, Transports and Environment) to fund the ERHIN programme (Estudio de los

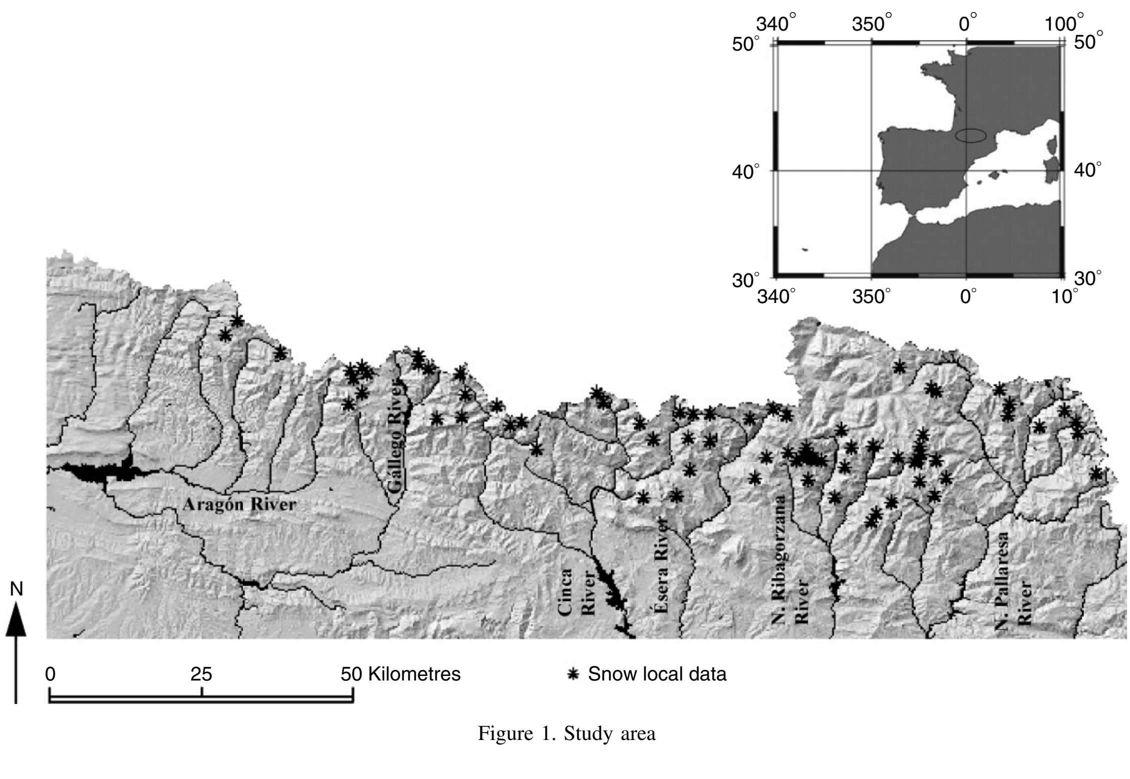

Figure 1. Study area

Recursos Hídricos producidos por la Innivación [Study of Water Resources from Snow Accumulation]). This programme involved the installation of 106 snow poles in the Pyrenees to collect three measurements per year on snow depth and, occasionally, on density. The programme was later extended to other mountain ranges in Spain (Cantabrian range and Sierra Nevada). The snow poles are located in flat areas, without shrub or tree cover, and they are relatively sheltered from wind-drift processes. Of all the sampling points, those with a full data record were selected (74) for this study. Samplings were carried out in January, at the end of March and at the end of April. This study used data from April because of their hydrological implications during the melting period (López-Moreno and García-Ruiz, 2004).

Interpolation methods

The interpolation methods compared in this study can be classified as detailed below (Burrough and McDonnell, 1998).

Local methods. These use only the information from the nearest sampling points to predict the unsampled areas (Burrough and McDonnell, 1998; Erxleben *et al*., 2002; Vicente-Serrano *et al*., 2003). These methods are based on the premise that the values between two points change progressively with distance. Two local methods were applied, namely inverse distance weighting (IDW) and splines.

With IDW, the weight of each point is inversely proportional to the distance between the samples. An exponent of the distance r can be applied to make the weight inversely proportional to any power of distance.

The splines method is based on a family of continuous, regular and derived functions adapted to the local variations of the interpolated data (Burrough and McDonnell, 1998). Data are obtained from the sampling points located in a radius around the unsampled point. The functions are adapted to the sampling points and there is no loss of continuity properties since each function has an important role in a given region and a null weight outside (Serrano-Vicente *et al*., 2003).

Geostatistical methods. These consider that the spatial variation of a variable can be described by a stochastic surface. Predictions are made from the distance-based relation between pairs of observations quantified by means of semi-variances between points. Two geostatistical methods were applied, namely ordinary kriging and cokriging.

With ordinary kriging, a variogram that summarizes the spatial structure of semi-variances is used to calculate the weighting factors that correspond to each point in order to provide estimates of the variable at the unsampled locations (Carroll and Cressie, 1996; Burrough and McDonnell, 1998; Vicente-Serrano *et al*., 2003).

The cokriging method allows the inclusion of secondary variables in the relationship between semi-variances with the distance between the pairs of observations. To calculate the weighting factors for each point, cross-variograms are used (Burrough and McDonnell, 1998; Vicente-Serrano *et al*., 2003; Diodato and Ceccarelli, 2005).

In order to optimize these approaches, sensitivity analyses were done considering different models, parameters and number of neighbours used in the calculations.

Global methods. The variability of the parameter to be interpolated is explained by means of the external characteristics of the sampling points (Burrough and McDonnell, 1998). The following global methods were applied: regression-tree models, linear models, and GAMs.

Regression-tree models are non-parametric methods based on the recursive splitting of the information from the predictor variables in order to minimize the sum of the squared residuals obtained in each group (Breiman *et al*., 1984). Generally, the tree size is selected according to a threshold in the change of the unexplained variance when a new group is obtained. Tree models are one of the methods most used for snow modelling and they provide an alternative to the assumption of linearity in relationships between the snowpack and the physical characteristics of the terrain (Erxleben *et al.*, 2002; Anderton *et al.*, 2004; Molotch *et al.*, 2005). Usually, non-linear relationships hinder the capacity to predict snow distribution (Anderton, 2000) and other environmental variables (Nogués-Bravo, 2003).

Linear models give predictions on the basis of the linear relationships between the response and the predictor variables using the following transference function:

$$z(x) = b_0 + b_1P_1 + b_2P_2 + \ldots + b_nP_n \tag{1}$$

where z is the predicted value at point x, $b_0, \ldots, b_n$ are the regression coefficients, and $P_1, \ldots, P_n$ are the values of the predictor variables at point x. The level of significance selected was $p < 0{\cdot}05$.

GAMs have been called data-driven approaches (Guisan *et al.* 2002), since modellers do not assume a special type of relationship (linear, quadratic, power, logarithmic, etc.) before model development. In GAMs (Guisan and Zimmermann 2000), the vector of parametric regression coefficients b is changed by a vector of non-parametric smoothers or functions. In other words, each regression coefficient b_p of a linear model is changed by a non-parametric smoother s_p. A GAM can be stated as

$$g(E(Y)) = PL = \alpha + sX + \varepsilon \tag{2}$$

or put another way

$$g(E(Y)) = PL = \alpha + f_1(X_1) + f_2(X_2) + \ldots + \varepsilon \tag{3}$$

where each predictor variable X_n is fitted by means of a function $f_n(\,)$. So, a GAM is the addition of different functions fitted to the independent variables in order to predict Y values. Data are fitted with respect to the partial residuals, i.e. the residuals after removing the effect of all predictor variables (Figure 1). Hastie and Tibshirani (1987) discuss various general scatter-plot smoothers that can be applied to the X variable values, with the target criterion to maximize the prediction quality of the (transformed) Y variable values. One such scatter-plot smoother are the cubic smoothing splines (Wood and Augustin 2002), which generally produce a smooth generalization of the relationship between the two variables in the scatter plot. A detailed description of how GAMs are fit to the data in relation to the algorithms used (outer and inner loop) can be founded in Hastie and Tibshirani (1987). In terms of the degrees of freedom, in a parametric regression one degree of freedom is lost when a single coefficient is estimated. Similarly, the more complex the spline, the greater the number of degrees of freedom that are lost. Degrees of freedom can be forced by the modeller to reduce the complexity of the adjusted spline, avoiding overfitting and obtaining response curves with an easier interpretation. There are computationally effective ways to choose the amount of smoothing, such as the general cross-validation (GCV) procedure that penalizes the complexity of the model (see Wood and Augustin (2002) for a technical exposition). The GCV score is used to find the model with the highest accounted deviance using the simplest splines (e.g. the GCV procedure tries to maximize the trade-off between model fit and the overall smoothness: when splines are forced to a maximum of four degrees of freedom the maximum smooth considered for each variable is four or less). The application of a GAM to snow accumulation data in the Pyrenees is explained in detail by López-Moreno and Nogués-Bravo (2005).

Variables

The topographic and locational-climate variables considered as potential predictors of snow depth were obtained from a digital elevation model (DEM) with a resolution of 100 m cell size.

- Altitude. This determines the type of precipitation (solid or liquid) and the evolution of melting in a given area (Caine 1975; Balk and Elder, 2000). Furthermore, precipitation in the Pyrenees follows a positive gradient, leading to an increase of snow accumulation in the highest sectors.
- Altitudinal range (RANG) observed between a cell and its neighbours or in radius of 500, 1000 or 2000 m (RANG500, RANG1000, RANG2000). RANG informs about the energy of the relief. The three RANGs

used here with lower spatial resolution (RANG500, RANG1000, RANG2000) are indicators of the degree of massivity of relief (Lorente and Beguería, 2002). Massive mountain sectors constitute a major obstacle for perturbations and, in general, receive more snowfall than isolated relief areas.

- Slope of a cell (SLOPE) or the mean slope in a radius of 500, 1000 or 2000 m (SLOPE500, SLOPE1000, SLOPE2000). The slope of a cell may affect snow redistribution (Mittaz *et al.*, 2002). When this variable is measured in a large radius, it provides data on the energy of the relief (Lorente and Beguería, 2002).
- Solar radiation (RAD) received by a cell. This is obtained by means of a model that uses the DEM to assess terrain complexity (aspect, slope and topographical shadows) and daily sun position. RAD is implemented in the MIRAMON geographical information system program (Pons, 1996, 1998). In addition, it was calculated for two radii: 500 and 1000 m around each pixel (RAD500, RAD1000).
- The distances to the Mediterranean Sea (DISMED) and to the Atlantic Ocean (DISATLAN) inform about the influence of Mediterranean and Atlantic air masses throughout the study area.
- The distance to the main divide (DMD), which generally coincides with the French border, allows the inclusion of the barrier effect of the distinct relief alignments, which follow a west–east axis.

Accuracy estimators

To assess the accuracy of the models, cross-validation was used to compare the estimated and observed values (Guisan and Zimmerman, 2000). This technique works by omitting one of the cases, fitting the model to the remainder and then applying the equation obtained to the omitted case in order to calculate its predicted value. This procedure was repeated for all cases in the data set. A range of error estimators were obtained from the relation between predicted and observed values (1986–2000 data set), namely mean bias error (MBE), mean absolute error (MAE), root-mean-square error (RMSE) and Willmott's D (Willmott, 1982):

$$\mathrm{MBE} = N^{-1}\sum_{i=1}^{N}(P_i - O_i) \tag{4}$$

$$\mathrm{MAE} = N^{-1}\sum_{i=1}^{N}|P_i - O_i| \tag{5}$$

$$\mathrm{RMSE} = \left[N^{-1}\sum_{i=1}^{N}(P_i - O_i)^2\right]^{0\cdot 5} \tag{6}$$

$$D = 1 - \frac{\sum_{i=1}^{N}(P_i - O_i)^2}{\sum_{i=1}^{N}(|P_i'| + |O_i'|)^2} \tag{7}$$

where N is the number of observations, O is the observed value, P is the predicted value, i is the counter for the individual observed and predicted values, $P_i' = P_i - \overline{O}$ and $O_i' = O_i - \overline{O}$, where $\overline{O}$ is the mean of observed values.

Willmott's D was the only error estimator selected to compare the accuracy of the predictions between single and multiyear average data sets, because the other error estimators fall short in providing information about the relative magnitude of the average difference (e.g. Willmott's D allows us a comparison of models with different ranges of values for snow accumulation).

RESULTS

Local and geostatistical methods

Figure 2 shows the averaged snow depth (1986–2000) predicted using local and geostatistical methods versus observations. The predictions considered here were those with highest accuracy shown in a sensitivity analysis. Thus, we considered for each method different models, parameters and number of neighbours used in the calculations. Figure 2a corresponds to the predictions obtained by means of a spline with tension considering 30 neighbours (parameter: 1165·6); Figure 2b results from applying an IDW (power: 2; number of neighbours: 30); Figure 2c shows the predictions obtained by means of spherical kriging (lag size: 20 000; number of lags: 15; number of neighbours: 30); and cokriging (Figure 2d) was calculated using altitude as covariable (lag size: 10 000; number of lags: 10; number of neighbours: 30). The coefficient of determination and the distinct error estimators indicate a very low capacity of these methods to predict the multi-average snow accumulation in the study area (Figure 2). Cokriging and IDW provided the highest correlation coefficients. However, they did not explain more than 15% of the variability of the snowpack. Furthermore, the distinct error estimators did not indicate that these methods were better than kriging and splines. Finally, single-year analyses provided worse results than the multiyear average interpolations. Willmott's *D* values never reached 0·4 when observed and predicted values were compared.

Regression-tree model

The first step to applying a regression-tree model is to select the size of the tree, since fitted trees may be more complex than is actually warranted by the data available (Sherrod, 2003; Anderton *et al.*, 2004). If two

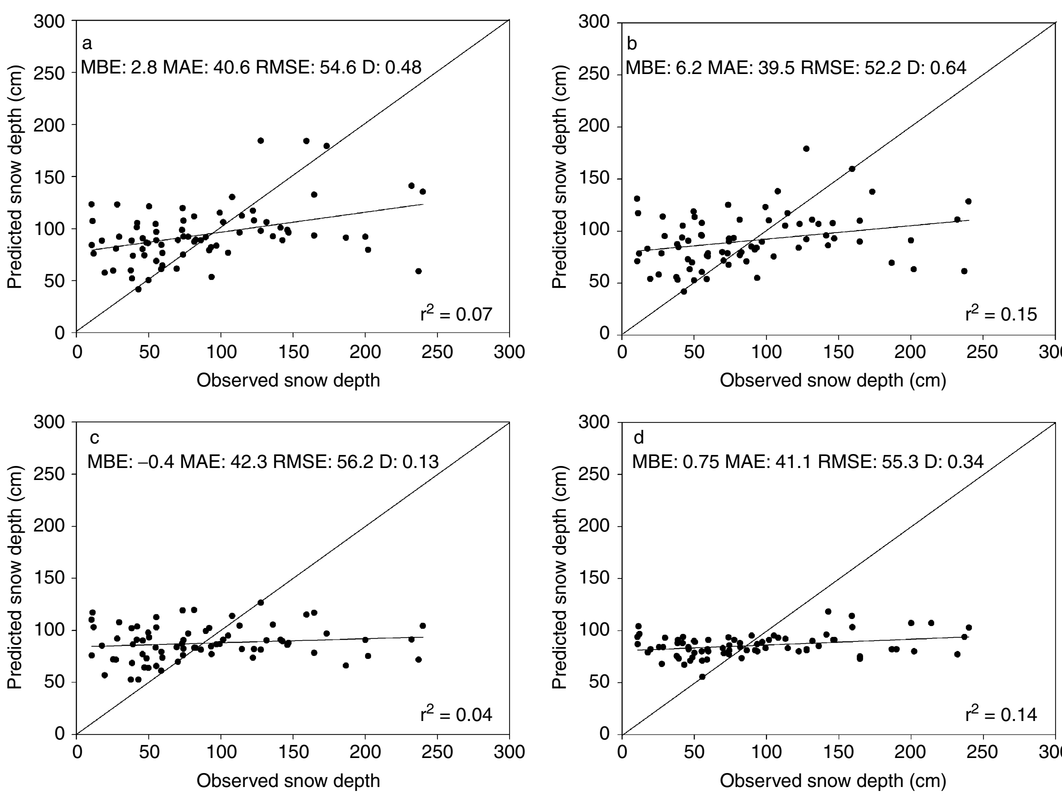

Figure 2. Correlation between measured and predicted snow depth, and accuracy estimators obtained using local and geostatistical methods: (a) splines; (b) IDW; (c) kriging; (d) cokriging

trees provide equivalent predictive accuracy, then the simple tree is preferred because it is easier to understand and faster to use for making predictions. Furthermore, smaller trees may provide greater predictive accuracy for unseen data than larger trees (Sherrod, 2003). The tree model was pruned to minimal cross-validated error. Figure 3 shows the change of error estimators (RMSE and Willmott's D) relating to tree size. It allowed us to improve the parsimony of the model accounting for the simplest tree with the highest predictive ability. Thus, the accuracy of the model increased until a six-node tree size, and then remained constant after this threshold. Figure 4 shows the six-node regression tree obtained. According to the splits, the spatial variance of snow depth (1986–2000) was explained by ALT, DISNORT, DISATLAN and RAD1000. The greatest snow depth (207 cm) was predicted in sectors above 2245 m a.s.l. and highly exposed to Atlantic air masses (DISATLAN $<$188 161 m). Snow accumulation is lower in eastern sectors. In this part of the Pyrenees the predicted snow depth reaches 126·9 cm in areas that receive low solar radiation (RAD1000 $<$1885 MJ day^{-1} m^{-2}) and close to the main divide (DMD $<$6200 m).

Figure 5 shows the relation between observed and predicted (1986–2000) snow depth values. When the prediction was obtained using all the observations, the predictive capacity was high ($r^2 = 0{\cdot}72$). However, the coefficient of determination decreased to 0·38 and the error estimators showed a noticeable lower accuracy when the predictions were obtained by means of cross-validation. When snow depth was interpolated for the single-year data set, Willmott's D showed a lower accuracy for 1994 ($D = 0{\cdot}59$) and 1995 ($D = 0{\cdot}61$) than for the averaged period 1986–2000 ($D = 0{\cdot}77$).

Linear regression

Several topographic and climate-locational variables were significantly related to snowpack distribution, 1986–2000 (Table I). Altitude was the variable with the highest correlation coefficient because of its direct control over the temperature pattern. Radiation showed a significant negative correlation with snow depth because of its positive relationship with melting (Gustafsson *et al.*, 2001). Slope and altitudinal range variables were significantly correlated with snowpack when considered at larger radius. Neighbour analysis at larger radius indicates the energy of relief or massivity, whereas at the smallest radius it shows micro-topographical variations. A stepwise regression model selected ALT, DISMED, DMD, RANG2000 and RAD1000 as significant predictor variables of snow depth ranged by their weight in the model, as the standardized

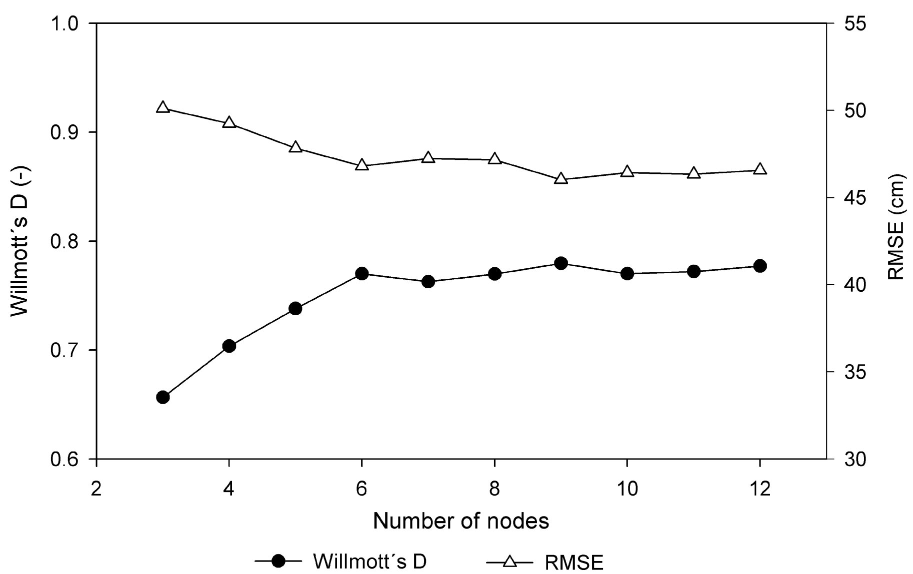

Figure 3. Relationship between number of nodes and regression-tree model accuracy

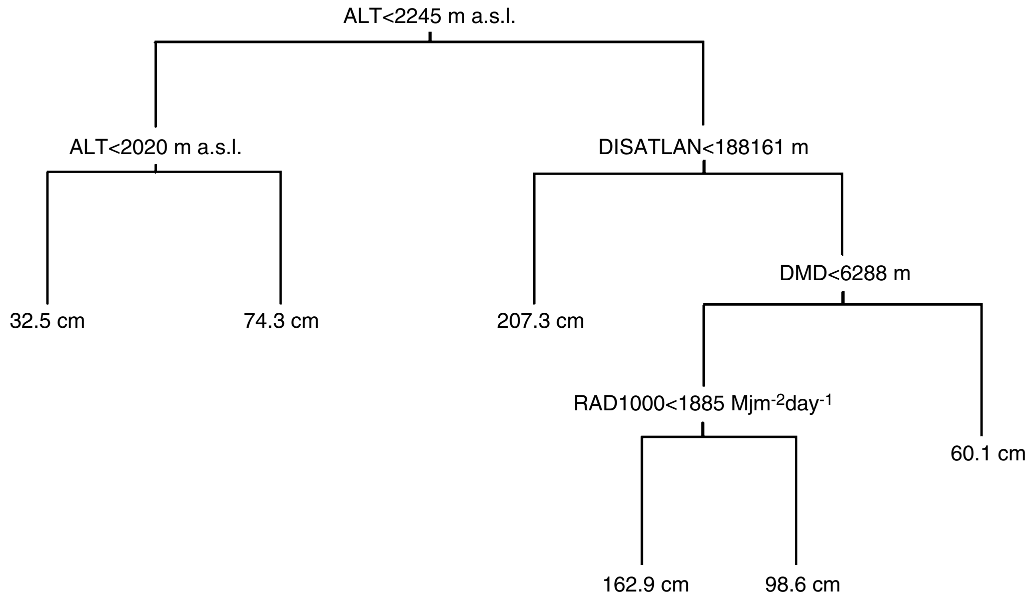

Figure 4. Regression-tree model

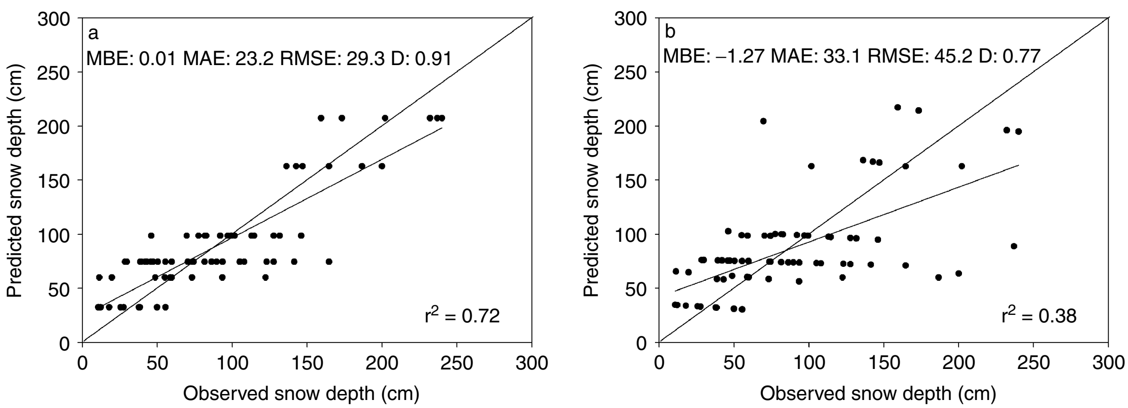

Figure 5. Correlation between measured and predicted snow depth, and accuracy estimators obtained using a regression-tree model: (a) predicted snow depth using all observations; (b) predicted snow depth obtained by cross-validation

coefficients (beta coefficients) indicate. These variables were very similar to those used to split the information in the regression-tree models, except that they included DISMED instead of DISATLAN.

According to the unstandardized coefficients, the following equation was used to map the snow depth distribution (1986–2000):

$$\text{Snow depth} = -318{\cdot}8 + 0{\cdot}19\text{ALT (m a.s.l.)} - 5 \times 10^{-4}\text{DISMED (m)} - 3 \times 10^{-3}\text{DMD (m)} + 7 \times 10^{-3}\text{RANG2000(m)} - 6 \times 10^{-3}\text{RAD1000 (MJ m}^{-2}\text{ day}^{-1}) \quad (8)$$

The sign of the unstandardized coefficients indicates the same direction of the response of snowpack to the independent variables as in the regression-tree model. An increase in depth was observed with height, as well as in shadow slopes near to the main divide, far from the Mediterranean Sea.

Table I. Correlation coefficients between predictor variables and snow depth and coefficients obtained in the stepwise linear regression model

	Correlation coefficient	Coefficients of the stepwise regression model		
		Unstandardized coefficients	Standardized (beta) coefficients	Significance
DISMD	−0·33[a]	−0·003	−0·290	<0·01
DISATLAN	0·21	—	—	—
DISMED	−0·27	−0·0005	0·374	<0·01
SLOPE	0·08	—	—	—
SLOPE500	0·13	—	—	—
SLOPE1000	0·21	—	—	—
SLOPE2000	0·25[a]	—	—	—
RANG	0·07	—	—	—
RANG500	0·17	—	—	—
RANG1000	0·20	—	—	—
RANG2000	0·28[a]	0·007	0·261	0·01
RAD	−0·24	—	—	—
RAD500	−0·26[a]	—	—	—
RAD1000	−0·27[a]	−0·006	−0·196	0·05
ALT	0·58[a]	0·189	0·754	<0·01

[a] Correlation is significant ($\alpha < 0{\cdot}05$)

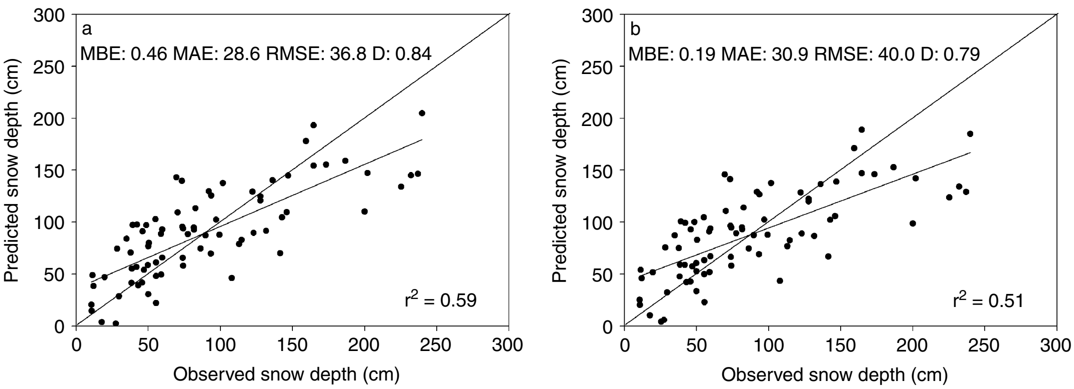

Figure 6. Correlation between measured and predicted snow depth, and accuracy estimators obtained using a stepwise linear regression model: (a) predicted snow depth using all observations; (b) predicted snow depth obtained by cross-validation

Figure 6 shows the relationships between observed and predicted values (1986–2000) estimated using all the observations, and by means of cross-validation. The coefficient of determination were $r^2 = 0{\cdot}58$ and $r^2 = 0{\cdot}51$ respectively. The error estimators also showed a clear improvement in accuracy compared with the local and geostatistical methods. The slight decrease in accuracy when the prediction was done by cross-validation indicates that linear regression is a robust method with regard to overfitting observations. When snow depth was interpolated for the single-year data set, Willmott's D showed a decrease on accuracy for 1994 ($D = 0{\cdot}61$) and a slight increase for 1995 ($D = 0{\cdot}83$) relating to the 1986–2000 period ($D = 0{\cdot}79$).

The GAM

Figure 7 shows the response curves of ALT, RAD, DA and DMD, all considered significant predictor variables by the GAM, to model snow depth (1986–2000). These graphs relate the magnitude of the response

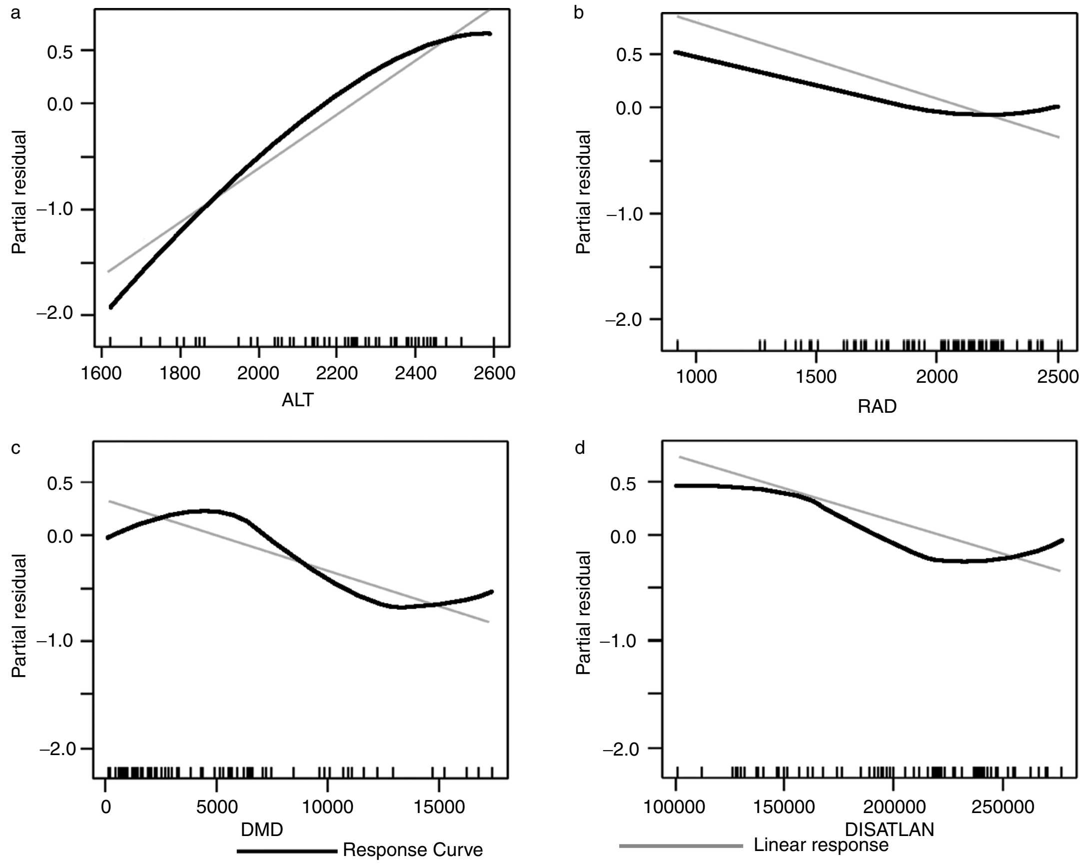

Figure 7. Curve and linear response of the predictor variables: (a) altitude; (b) solar radiation; (c) distance to the main divide; (d) distance to the Atlantic Ocean

variables against the partial residuals. Partial residuals were obtained after removing the effect of all other predictor variables. The linear adjustment (a standard linear regression) of the response variables to the partial residuals is also shown. The assumption of linearity implies a simplification of the observed response of the snowpack to the topographical and locational variables, since the GAM detects non-linear relationships (López-Moreno and Nogués-Bravo, 2005). Thus, noticeable differences in the response of a snowpack can be detected when non-linear relationships are considered, which could have great implications in the accuracy of the model obtained. In contrast to the regression tree and linear regression model, the GAM does not consider the altitudinal range at low spatial resolution as a significant variable for explaining snow distribution.

Coefficients of determination (0·73) and error estimators obtained after cross-validation indicate that the GAM provides the most accurate and robust prediction of the 1986–2000 snow depth distribution of the methods studied (Figure 8). Also, GAMs provide the best estimations when single-year data sets of snow depth are interpolated. Thus, results for 1994 ($D = 0{\cdot}83$) and 1995 ($D = 0{\cdot}87$), slightly lower than for the period 1986–2000 ($D = 0{\cdot}89$), improved the accuracy of the predictions obtained with regression tree and linear regression models.

Comparison of snow depth maps provided by global methods

Figure 9 shows the snow distribution maps (multiyear average of the period 1986–2000) obtained by means of the three global methods. The regression-tree model provides a map in discrete units, since all the

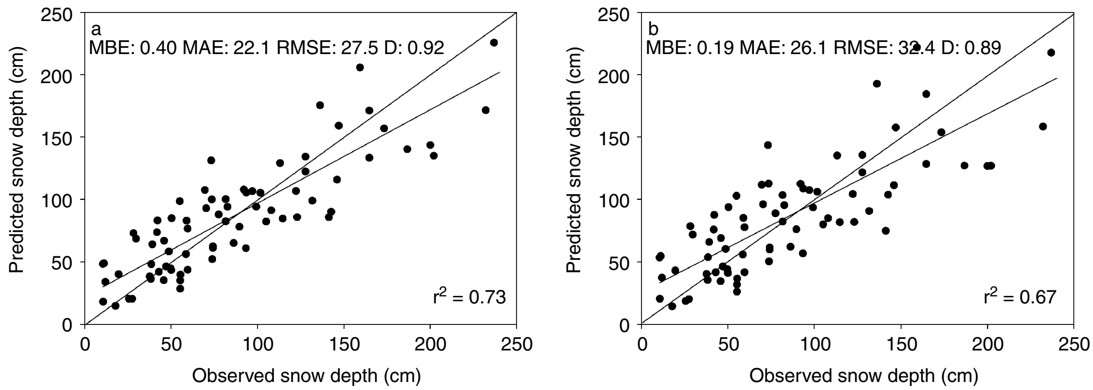

Figure 8. Correlation between measured and predicted snow depth, and accuracy estimators obtained using a GAM. (a) predicted snow depth using all observations; (b) predicted snow depth obtained by cross-validation

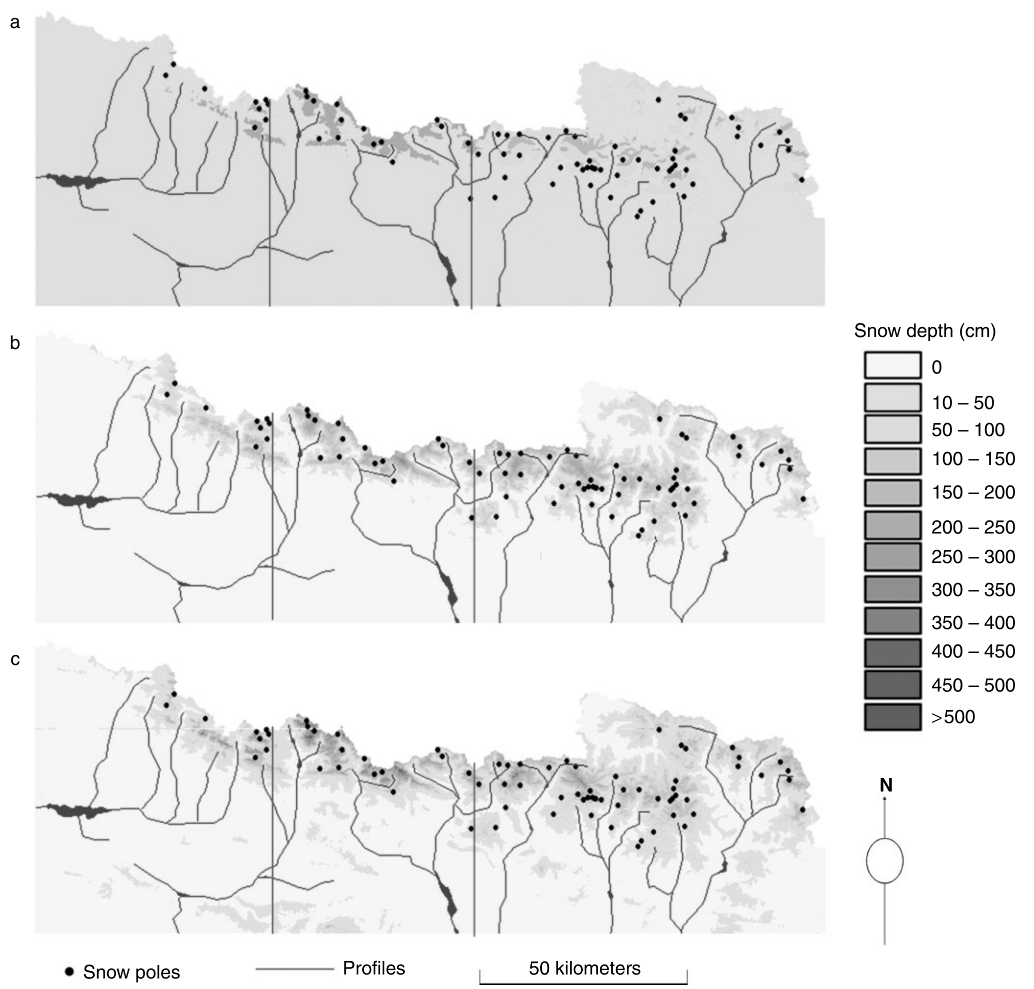

Figure 9. Snow depth maps obtained by means of global methods: (a) regression-tree model; (b) linear regression model; (c) GAM

cells with characteristics that agree with one of the six subgroups receive the value of the corresponding terminal node. It lacks a realistic representation of snow distribution; the minimum value of snow depth for the whole area is the lower terminal node (32·5 cm), thereby hindering the identification of the snow-covered area. The maximum snow depth predicted corresponds to the value of the node most favourable for snow accumulation (207·3 cm). Thus, the snow depth is underestimated in areas where relief is more favourable for the accumulation of snow than those of the sampled points. Linear regression and the GAM provide maps in continuous units since the predicted values result from the combination of the topographic and locational characteristics of each cell. Also, in the areas most prone to snow accumulation, these methods have the capacity to extrapolate greater depths than those observed. Linear regression models may provide negative depth values in areas less prone to snow accumulation. These values are converted to zero in order to obtain a realistic representation. Both methods predict snow cover from around 1600 m a.s.l., which coincides with the altitude of the winter 0 °C isotherm (García-Ruiz *et al*., 1986).

Two main differences were observed between the maps obtained by means of the linear regression and the GAM: (i) the GAM predicted snow cover in relatively low mountainous areas in the southernmost part of the study area (summits barely reach 2000 m a.s.l.), whereas the linear model maintained these areas free of snow; (ii) the greatest depths predicted by the GAM exceeded 500 cm, whereas the linear regression predicted that no cell in the study area would accumulate more than 350 cm of snow.

The differences in the maps provided by global methods can be analysed in more detail in Figure 10. This shows the predicted snow depth by the three global methods in the two profiles drawn in Figure 9, and their relation with altitude and solar radiation. The snow depth profiles provided by the regression-tree model are a broad representation of the linear regression and GAM predictions. This tree model showed an overestimation of the areas free of snow and an underestimation in the highest and most favourable areas for snow accumulation. The linear regression model predicted snow cover within the nearest 15 km to the main divide, whereas GAM predicted snow cover in southern, lower altitude massifs. Although snow depth follows a similar pattern to altitude, the remaining variables introduce clear modifications in the predictions, especially in the GAM. Thus, in GAM predictions of snow depth profiles, the greatest depth did not coincide with the highest altitude because of the effect of distance to the main divide and incident solar radiation. The prediction of the regression linear model was more directly influenced by altitude. In the GAM prediction, the influence of Atlantic or Mediterranean air masses was clearly visible when comparing the two profiles. Thus, the westernmost profile, which did not exceed 3000 m a.s.l., showed greater snow depth than the easternmost profile, though in the latter the altitude exceeded 3200 m and solar radiation was lower. The accumulation predicted by the linear regression model was more affected by altitude, without clear differences between the eastern and western profiles.

DISCUSSION AND CONCLUSIONS

Here, we assessed the capacity of several local, geostatistical and global interpolator methods to predict and map snow depth in unsampled areas of a large mountainous sector of the central Spanish Pyrenees.

Our results indicate a low predictive capacity of the local and geostatistical methods. Several studies report a low percentage of explained variance of snow distribution using these methods in mountainous areas (Erxleben *et al*., 2002). This low ability of local methods to predict snowpack values in our study is mainly related to the limited number of snowpack measurements in the central Pyrenees and their topographical complexity. Local interpolators based on the continuous change of response variable (temperature, rainfall, etc.) across space do not fit adequately in mountain areas with important changes of controlling variables such as altitude, aspect or slope in a reduced zone. However, other climatic variables, such as temperature and precipitation (both related to snow accumulation and melting), have been interpolated successfully using the closest observations or probability functions in areas where a dense network of observations is available and

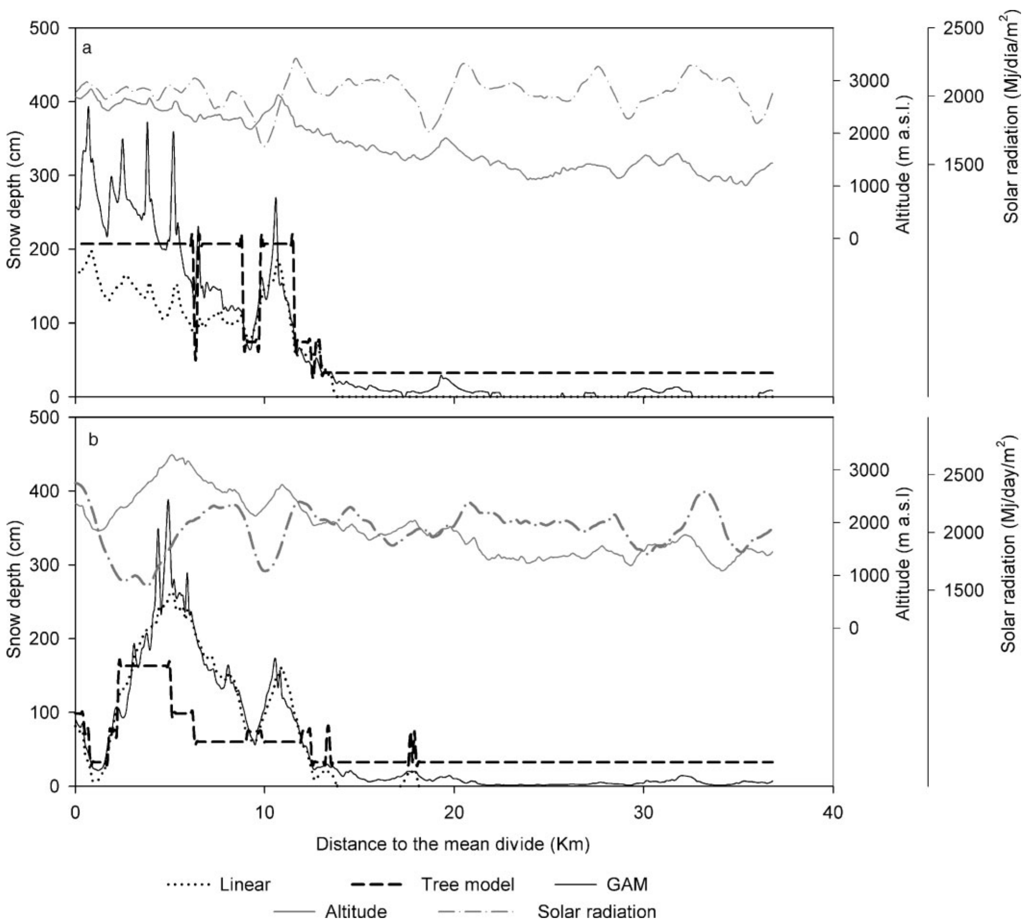

Figure 10. Predicted snow depth by the three global methods in two north–south profiles, and their relation with altitude and solar radiation: (a) western profile; (a) eastern profile

where the variable to be interpolated changes according to a marked spatial pattern (Burrough and McDonnell, 1998; Vicente-Serrano *et al*., 2003; Diodato and Ceccarelli, 2005).

In this study, global methods provided better results than local and geostatistical methods, since the predictions are based on the response of the snowpack to variables that summarize the terrain complexity and the climate conditions of the study area (Balk and Elder, 2000; Anderton *et al*., 2004; López-Moreno and Nogués-Bravo, 2005). The linear regression model considers the massivity of a sector (in contrast to isolated reliefs) to explain snow distribution. However, the regression tree and GAM do not include this variable as significant. The three global methods consider altitude, distance to the main divide, exposure to Atlantic or Mediterranean climatic influences and incident solar radiation as predictor variables of snowpack.

In recent years, regression-tree models have been commonly used for interpolating snow depth data. Here, we have shown that this method is useful for identifying the variables that explain snow accumulation and their non-linear responses with the dependent variable. Also, the graphical representations of trees facilitate the understanding of the relationships between controlling factors and response variable. However, we have identified several drawbacks of this method in unsampled areas:

1. Snow cover predictions were highly overfitted to the observations even though the complexity of the tree model was reduced to six nodes. Consequently, we observed a noticeable decrease in the explained variance

when the predictions were obtained by means of cross-validation. The tendency of the tree models to overfit their predictions has been reported elsewhere (Chambers and Hastie, 1993; Guisan and Zimmerman, 2000; Willfried *et al.*, 2003). In the upper Marble Fork basin, California (Leydecker and Sickman, 1999), the variance of snow distribution explained decreased from 60% to only 15% when non-selected points were modelled.
2. The snowpack map provided was poor, since it had only six discrete classes, which correspond to the terminal nodes of the tree regression. A greater number of nodes to improve cartography would lead to an increase in overfitting to the observations and meaningless relationships between predictors and response variable.
3. It cannot extrapolate predicted values above or below the observed values. This regression makes it difficult to distinguish between the snow-covered and non-snow-covered areas. Furthermore, the snow depth of the sectors with the greatest tendency to accumulate snow was underestimated. This is especially relevant, since the highest observation is below 2800 m, and the method used should also have the capacity to extrapolate a deeper snowpack in the highest and more favourable sectors.

The drawbacks commented here affect strongly our results since the specific conditions of this study: regional scale with a low number of observations. These constraints could be minimized with a sampling strategy that covers a wider range of locations and topographic characteristics. Nevertheless, measurement surveys are determined, in many cases, by the inaccessibility of some mountain areas.

In spite of the assumption of a linear response of the independent variables to snow depth, the stepwise linear regression model showed a high capacity to predict snow distribution and a low level of overfit to the observations. The use of linear regression for interpolating climatic data (Ninyerola *et al.*, 2000; Vicente-Serrano *et al.*, 2003) or snow depth data (Anderton, 2000; Chang and Li, 2000) has also provided good predictive and cartographic results.

The GAM explained 73% of the variance of snow depth (1986–2000). The increase in the accuracy of the prediction with regard to the linear model is due to the capacity of the GAM to include the non-linear relationships between the predictive variables and the snow distribution in the study area. Furthermore, the slight decrease in explained variance confirms the robustness of the model produced when the prediction is obtained by means of cross-validation. Also, GAMs provided the most accurate predictions of snow depth distribution for climatically different single years (1994 and 1995).

According to the field observations and available meteorological data in our study area, the amount of snow predicted by the GAM in the areas most prone to snow accumulation is more realistic than that predicted by the linear model, which clearly underestimates the results. The latter does not provide estimations greater than 350 cm, whereas field records confirm the occurrence of deeper accumulations. Thus, in Balneario de Panticosa, which is located at 1600 m a.s.l. and influenced by an oceanic climate, snowfall in the summit areas usually begins in October and precipitation from December to April exceeds 700 mm. Rijckborst (1967) reported 400 cm of snow accumulated around 3000 m a.s.l. in the headwater of the Esera River during the three months with highest accumulation of snow (i.e. December, February and April). Davy (1978) estimated that the average (1955–1965) snowfall over the Aneto Glacier during winter and early spring was 802 cm.

The opposite occurs with the snow depth predicted by the GAM in the southernmost sectors of the study area, where it introduces a remarkable overestimation. This overestimation may be an artefact, because the slight increase in snow depths observed furthest from the main divide leads to the GAM overestimating depths in some mountainous sectors in the south of the study area, where no observations are available. Therefore, this type of study requires a set of measurements that are well distributed and that extensively cover the topographic and locational situations of the study area.

On the basis of our results, we conclude that:

1. Given the density of available snow depth observations and the topographic complexity of the study area, local and geostatistical methods do not provide proper predictions of snowpack in unsampled areas.

2. Global methods show a noticeable increase in the capacity to predict snow distribution and allow the study of the effect of predictor variables on snow accumulation.
3. Regression trees tend to overfit their predictions to observations. On the contrary, linear regression and the GAM provide more robust predictions.
4. Tree-regression models cannot extrapolate their predictions in areas where the independent variables are out of the range of the observations, unlike linear regressions and the GAM. However, important degrees of uncertainty must be considered in extrapolated values out of the observed range of snow depth measurements. It is highly recommended to have a data set of observations covering the extent of the area of interest and the wider range of topographic conditions.
5. GAMs are useful to predict and to study the spatial distribution of snow and other climate variables. GAMs explain variance better and provide the most robust predictions for the spatial scale considered and the data set used. Finally, these models permit a better understanding of the non-linearity of the relationships between snowpack and predictor variables

ACKNOWLEDGEMENTS

This study was supported by the following research projects: ‘Hydrologic and erosive processes in Pyrenean catchments, related to land use changes and climate variability’ (PIRIHEROS, REN 2003-08678/HID) and ‘Characterisation and modelling of hydrological processes in gauged basins for the prediction of ungauged basins’ (CANOA, CGL 2004-04919-c02-01), both funded by the CICYT, Spanish Ministry of Science and Technology. David Nogués-Bravo is a post-doc researcher of the ALARM project (GEOCE-CT-2003-506675). We gratefully acknowledge the valuable comments of Dr Jose María García-Ruiz, Dr Sergio Vicente-Serrano and two anonymous referees.

REFERENCES

Anderton SP. 2000. *An analysis of spatial variability in snow processes in a high mountain catchment*. PhD thesis, School of Engineering, University of Durham.

Anderton SP, White SM, Alvera B. 2004. Evaluation of spatial variability in snow water equivalent for a high mountain catchment. *Hydrological Processes* **18**: 435–453.

Balk B, Elder K. 2000. Combining binary decision and geostatistical methods to estimate snow distribution in a mountain watershed. *Water Resources Research* **36**(1): 13–26.

Breiman L, Friedman JH, Olshen RA, Stone CJ. 1984. *Classification and Regression Trees.* Chapman and Hall: New York.

Burrough PA, McDonnell RA. 1998. *Principles of Geographical Information Systems.* Oxford University Press: Oxford.

Caine N. 1975. An elevational control of peak snowpack variability. *Water Resources Bulletin* **11**(3): 613–621.

Carroll SS, Cressie N. 1996. A comparison of geostatistical methodologies used to estimate snow water equivalent. *Water Resources Bulletin* **32**(2): 267–278.

Chambers JM, Hastie TJ. 1993. *Statistical Models*. Chapman and Hall: London.

Chang KT, Li Z. 2000. Modelling snow accumulation with a geographic information system. *International Journal of Geographical Information Science* **14**(7): 693–707.

Davy L. 1978. *L’Ebre Etude Hydrologique.* Universidad de Montpellier: Montpellier.

Del Barrio G, Creus J, Puigdefábregas J. 1990. Thermal seasonality of the high mountain belts of the Pyrenees. *Mountain Research and Development* **10**: 227–233.

Diodato N, Ceccarelli M. 2005. Interpolation processes using multivariate geostatistics for mapping of climatological precipitation mean in the Sannio Mountains (southern Italy). *Earth Surface Processes and Landforms* **30**: 259–268.

Elder K, Rosenthal W, Davis R. 1998. Estimating the spatial distribution of snow water equivalence in a montane watershed. *Hydrological Processes* **12**: 1793–1808.

Erxleben J, Elder K, Davis R. 2002. Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky Mountains. *Hydrological Processes* **16**: 3627–3649.

García-Ruiz JM, Puigdefábregas J, Creus J. 1986. La acumulación de la nieve en el Pirineo Central y su influencia. *Pirineos* **17**: 27–72.

García-Ruiz JM, Beguería S, López-Moreno JI, Lorente A, Seeger M. 2001. *Los Recursos Hídricos Superficiales del Pirineo Aragonés y su Evolución Reciente*. Geoforma: Logroño.

Guisan A, Zimmermann NE. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* **135**: 147–186.

Guisan A, Edwards JTC, Hastie T. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling* **157**(2–3): 89–100.

Gustafsson D, Stähli M, Jansson PE. 2001. The surface energy balance of a snow cover: comparing measurements to two different simulation models. *Theoretical and Applied Climatology* **70**: 81–96.
Haefner H, Seidel K, Ehrler. 1997. Applications of snow cover mapping in high mountain regions. *Physics and Chemistry of the Earth* **22**(3–4): 275–278.
Hastie T, Tibshirani R. 1987. Generalised additive model: some applications. *Journal of the American Statistical Association* **82**: 371–386.
Leydecker A, Sickman J. 1999. Distribution of snow in the upper Marble Fork basin, California. In *1999 AGU Fall Meeting*, San Francisco, CA (abstract presented as a poster). www.icess.ucsb.edu/swg/projects/Snow-Distribution.pdf [July 2005].
López-Moreno JI. 2005. Recent variations of snowpack depth in the central Spanish Pyrenees. *Arctic, Antarctic, and Alpine Research* **37**(2): 253–260.
López-Moreno JI, García-Ruiz JM. 2004. Influence of snow accumulation and snowmelt processes on the distribution of streamflow in the central Spanish Pyrenees. *Journal of Hydrological Sciences* **49**: 787–802.
López-Moreno JI, Nogués-Bravo D. 2005. A generalized additive model for the spatial distribution of snowpack in the Spanish Pyrenees. *Hydrological Processes* **19**: 3167–3176.
López-Moreno JI, Beguería S, García-Ruiz JM. 2004. The management of a large Mediterranean reservoir: storage regimes of the Yesa reservoir, upper Aragón River basin, central Spanish Pyrenees. *Environmental Management* **34**(4): 508–515.
Lorente A, Beguería S. 2002. Variation saisonnière de l'intensité des precipitations maximales dans les Pyrenées Centrales. Analyse spatiale et cartographique. *Publications de l'Association International de Climatologie* **14**: 327–335.
Mittaz C, Imhof M, Hoelze M, Haeberli W. 2002. Snowmelt evolution mapping using an energy balance over an alpine terrain. *Arctic, Antarctic and Alpine Research* **34**(3): 274–281.
Molotch NP, Colee MT, Bales RC, Dozier J. 2005. Estimating the spatial distribution of snow water equivalent in an alpine basin using binary regression tree models: the impact of digital elevation data and independent variable selection. *Hydrological Processes* **19**: 1459–1479.
Ninyerola M, Pons X, Roure JM. 2000. A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. *International Journal of Climatology* **20**: 1823–1841.
Nogués-Bravo D. 2003. El estudio de la biodiversidad: conceptos y métodos. *Cuadernos de Investigación Geográfica* **29**: 67–82.
Pons J. 1996. Estimación de la radiación solar a partir de modelos digitales de elevaciones. Propuesta metodológica. In *VII Coloquio de Geografía Cuantitativa, Sistemas de Informaciòn Geográfica y Teledetección*, Juaristi J, Moro I (eds), Vitoria, 87–97.
Pons X. 1998. *Manual of Miramon. Geographic information system and remote sensing software*. Centre de Recerca Ecològica I Aplicacions Forestals (CREAF), Bellaterra. http://www.creaf.uab.es/miramon [accessed 7 April 2006].
Rijckborst H. 1967. Hydrology of the upper Garone basin (Valle de Arán, Spain). *Leidse Geologische Mededelingen* **40**: 1–74.
Sherrod PH. 2003. DTREG. *Classification and regression trees for data mining and modelling.* www.dtreg.com/DTREG.pdf [July 2005].
Vicente SM, Saz MA, Cuadrat JM. 2003. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. *Climate Research* **24**: 161–180.
Willfried T, Araújo MB, Lavorel S. 2003. Generalized models vs. classification tree analysis: predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science* **14**: 669–680.
Willmott CT. 1982. Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society* **63**(11): 1309–1313.
Wood SN, Augustin NH. 2002. GAMs with integrated model selection using penalized regression splines applications to environmental modelling. *Ecological Modelling* **157**: 157–177.
Yang D, Woo MK. 1999. Representativeness of local snow data for large scale hydrologic investigations. *Hydrological Processes* **13**: 1977–1988.

