Assessing Snow Accumulation from Local Data in Alpine Areas: The Response of Regression-Based Snow Interpolation Methods to Sample and Grid Size

J. IGNACIO LÓPEZ MORENO,[1] J. LATRON,[2] AND A. LEHMANN[3]

ABSTRACT

This work analyses the response of four regression-based interpolation methods to changes in the number of cases and in the resolution of the digital elevation model (DEM). For this purpose, data obtained from an intensive random snow depth sampling (991 measurements) in a small catchment (6 km^2) in the Pyrenees, Spain, were used. Linear regression, classification trees, generalized additive models (GAMs), and a new method based on a correction by applying tree classification to residuals of GAMs, were used to calculate snow depth distribution from terrain characteristics under different combinations of sample size (100, 200…, 991 cases) and DEM spatial resolution, (from 5x5m to 95x95m every 10m of grid size).

Application of a tree classification to residuals obtained from GAMs yields the best accuracy scores. The other tested methods yield rather similar accuracy scores but different levels of robustness when a cross-validation procedure is applied. Accuracy of the model predictions declines as resolution of DEMs and sample size decreases. However, the sensitivity of the models to the number of cases used shows different thresholds, which has relevant implications to optimise the relation between the effort involved and the quality of the results, when fieldwork is planned.

Keywords: regression-based methods; spatial interpolation; sample size; DEM resolution; snow

INTRODUCTION

One of the most reliable procedures to assess snow accumulation in a given area is based on the transfer of data from point measurements to neighboring unsampled areas. Accordingly, in recent decades many research efforts have sought to develop, test, and compare different techniques of interpolating local snow data (Tyler et al. 2005 and the references therein).

Most such methods belong to the family of regression-based models, which empirically relate the amount of snow measured at the sampled points to their terrain characteristics. In this approach, different topographical and geographical features, generally derived from a digital elevation model (DEM), are used as predictor variables on the basis that they are closely related to the accumulation, redistribution, and ablation of the snow cover (Tyler et al. 2005; Molotch et al., 2005; Jost et al., 2007; López-Moreno and Stahli, 2008). Once the relationships between snow

[1] Instituto Pirenaico de Ecología, CSC, Campus de Aula Dei, P.O. Box 202, Zaragoza, 50.080, Spain.
[2] Soil Science Unit, University of Girona, Campus de Montilivi 17071, Girona, Spain.
[3] Climatic Change and Climate Impacts Group. University of Geneve. Battelle-D.7, Chemin de Drize. CH1227, Carouge (geneve), Switzerland.

accumulation and predictor variables have been established, the spatial distribution of snowpack can be determined for areas with known terrain characteristics.

Accuracy in the results can potentially be related to many different factors, including the choice of regression technique, characteristics of the terrain and climatic conditions of the study area, the employed spatial scale, quality of topographic data, sampling strategy, and selection of predictor variables. Among these factors, the resolution of the DEM (grid size) and sample size have been widely recognized to strongly influence the reliability and stability of predictions made using regression models for other disciplines (Tang et al., 2001; David et al., 2002; Cohen et al., 2003; Kienzle, 2004; Wechsler, 2006); however, the effects on model performance of the number of observations or DEM resolution have been mentioned only occasionally in snow studies, and have yet to be studied in detail.

This lack of analysis is surprising given the relevance of these parameters in snow research. The measuring of snow depth or snow water equivalent is not a trivial task: snow sampling requires a noticeable investment of human resources. Consequently, a consistent set of criteria should be established to aid in determining the required number of measurements. Moreover, it is well known that terrain attributes derived from a DEM change with the resolution of the underlying grid-cell size, affecting subsequent modeling of surface processes (Zhang and Montgomery, 1994; Kienzle, 2004).

To investigate further the effects of the number of observations and DEM resolution on modeled snowpack distribution, this study assesses the effect of sample and grid size on the accuracy and robustness of four regression-based methods used to interpolate punctual snow-depth data obtained during an intensive snow-sampling survey.

STUDY AREA

The study area is located on the southern side of the Pyrenees, close to the main divide (Spanish–French border) in the headwaters of the Gallego River, Spain (Figure 1).

The catchment is close to 6 km^2 in area, ranging in altitude between 1700 and 2400 m a.s.l.; all snow measurements were taken below 2300 m. The vegetation cover consists of high mountain meadows and rocky outcrops in steeper areas. Except for the existence of cliffs beneath some ridges, the landscape tends to be relatively gentle, with important variability in terrain curvature and slopes of less than 45%. The mean slope of the sampled points is 14%.

The study area, located in a transition zone between Atlantic and Mediterranean conditions, has mixed climatic influences and is exposed to winds from all directions. The mean annual temperature is 3°C, with 130 days per year with a mean daily temperature below 0°C. Mean annual precipitation is around 2000 mm, with more than half falling as snow (Anderton et al., 2004). Although the mean winter temperature is below 0°C, the area is subject to intense wintertime warm periods that trigger melting events and major metamorphosis of the snowpack; these warm spells may occur at any time throughout the snow season.

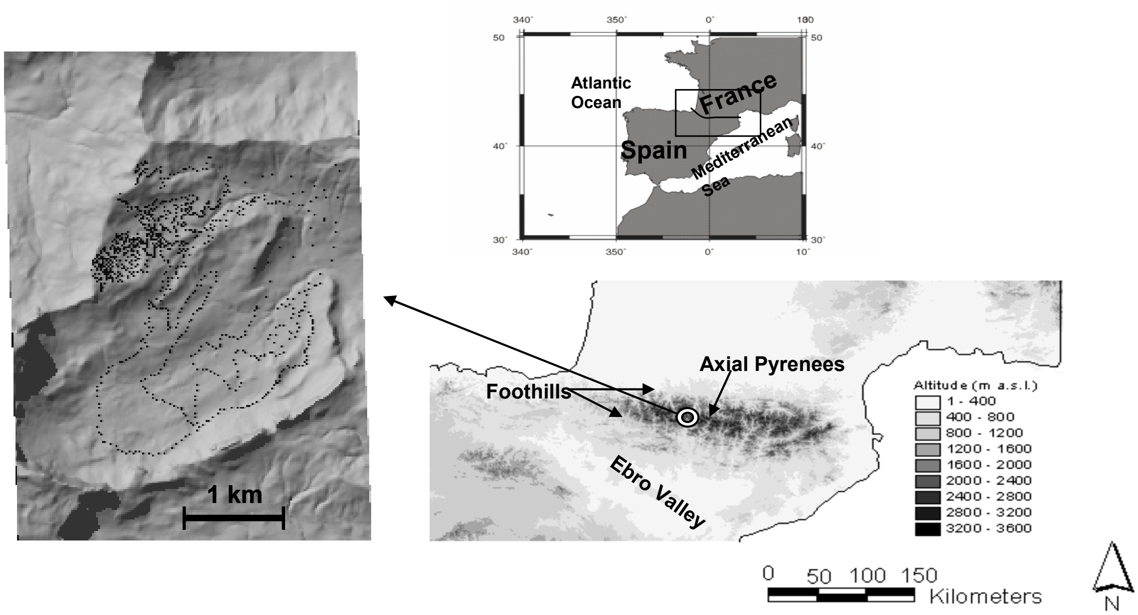

Figure 1. Study area. Dots indicate the locations of snow-depth measurements.

DATA AND METHODS

Snow sampling survey

A snow survey was carried out in the spring of 2006 (April 18–21). Fieldwork was planned for this time because snowpack in the catchment usually exhibits large spatial variability in spring, and terrain characteristics exert a strong control on its distribution. A total of 991 snow-depth measurements were collected manually using a steel probe. Each depth measurement involved four replicates within 50 cm of the first measurement. A random sampling strategy was adopted to obtain a large number of measurements (avoiding sectors with difficult access due to topography) and provide greater flexibility in handling the extreme heterogeneity of the snowpack. The locations of snow-depth measurements were accurately recorded using a submetric GPS (Geoexplorer XTTM handheld with HurricaneTM antenna), later to be translated into the DEM.

Digital elevation model and terrain characteristics

The DEM was compiled from a high-quality digitized topographic cartography at a scale of 1:5000, provided by the Aragonaise Government, Spain (http://sitar.aragon.es/). Isolines accounted for a vertical resolution of 5 m. The application of IsoMDE implemented in the GIS software MIRAMON (http://www.creaf.uab.es/MiraMon/) enabled the generation of DEMs from isolines and additional information such as watercourses, depressions, ridge lines, and spot heights. Constrained by the scale of the topography and the distance between isolines, the highest spatial resolution considered in our analysis was a grid-cell size of 5 × 5 m.

The original 5 × 5 m DEM was subsequently degraded to grid-cell resolutions of 15, 25, 35, 45, 55, 65, 75, 85, and 95 m. Terrain parameters to be used as predictor variables of snowpack were subsequently derived from each DEM at different grid sizes.

The selection of potential predictors was based on their ability to affect the rain/snow limit, motion of fresh snow (i.e., wind drift or avalanches), and snow ablation. The selected predictors were as follows:

-Altitude, which determines the type of precipitation (solid or liquid) and the evolution of melting in a given area.

-Slope of the cell, which may affect snow redistribution.

-Average solar radiation (RAD) received by each cell of the DEM from December to April under clear-sky conditions. This parameter was obtained from a physically based computational model (implemented in the MIRAMON GIS software, Pons and Ninyerola 2008).

-Easting and Northing, which informs on the east–west and north–south orientations of the slopes respectively. They were quantified via the sines and cosines of the aspect, respectively. Both variables, which logically have colinearity with potential incoming solar radiation, were introduced as predictors because they potentially reflect the effects of snow drift or deposition by wind.

-Mean curvature, which identifies concave and convex areas of the catchment.

-Topographic Position Index (TPI) at 100 m resolution, which informs on the cell position in relation to surrounding relief (Jennes, 2006).

-Compound topographic index (CTI), which is a function of both the slope and the upstream contributing area per unit width orthogonal to the flow direction (Gessler et al., 1995). This variable is commonly used as a wetness index, and it can inform on the position of a cell within a slope.

Regression-based methods and estimation of model accuracy

Regression-based methods rest on the creation of dependence models between snow data and other independent variables (terrain characteristics) for predicting the values of snow depth in unsampled regions. The four methods compared in this study are outlined below :

- 1: Linear Models give predictions based on the linear relationships between the response and predictor variables according to the following transference function:

$$z(x)= b_o + b_1P_1 + b_2P_2 + ... + b_nP_n \qquad (1)$$

where z is the predicted value at point x, $bo...bn$ are the regression coefficients, and $P_1...P_n$ are the values of the predictor variables at point x. The level of significance selected in this study was $p < 0.05$.

- 2: Classification tree models are non-parametric methods based on recursive splitting of the information from the predictor variables to minimize the sum of the squared residuals obtained in each group (Breiman *et al.*, 1984). The tree size is generally 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 commonly used for snow modeling, providing an alternative to the assumption of linearity in the relationships between snowpack and the physical characteristics of the terrain (i.e. Anderton *et al.,* 2004; Molotch *et al.,* 2005).

- 3: Generalized additive models (GAMs) are non-parametric extensions of generalized linear models (GLMs) that estimate response curves with a non-parametric smoothing function rather than with parametric terms (Hastie and Tibshirani, 1987). This approach enables the user to explore the shapes of predictor responses along the gradient of the dependent variable (snow depth), enabling in turn an accurate fit of statistical models to highly non-linear relationships and the detection of abrupt changes in the responses of many natural processes (Lehmann *et al*., 2002). Thus, a GAM can be stated as

$$\mathbf{g(E(Y)) = PL} = \alpha + \mathbf{f_1(X_1)} + \mathbf{f_2(X_2)} + \mathbf{...} + \mathbf{f_n(X_n)} + \varepsilon \qquad (2)$$

where each predictor variable X_n is fitted using a smoothing function $f_n(Xn)$, α is a constant, and ε the remaining residuals. Consequently, a GAM involves the addition of different functions fitted to the independent variables in order to predict Y values.

- 4: Application of tree regression models to GAM residuals. This procedure enables the user to address potential interactions between predictors—an important issue that has received increasing attention (Austin, 2002). This novel approach for regression models (Maggini et al., 2006) consists of fitting a regression tree on residuals to enable the identification of significant interaction terms, thereby enhancing the predictive capacity of the initial model.

The agreement index (Willmott's D) was used to evaluate the four models. Willmott's D is a relative and bounded measure used to assess model accuracy: it retains mean information and does not amplify outliers (Willmott, 1982). The index also scales the magnitude of variables, thereby enabling comparisons between different experiments performed in other locations or at different times, independently of differences in the mean magnitude and range of the snowpack.

Model accuracy was assessed by comparing predictions with the observations used to conduct the models, as well as by cross-validation. Cross-validation first consists of splitting the data into a number of subsets (five in the present study) and omitting each subset in turn; the model is then fitted to the remaining cases, and the obtained equation is finally applied to the omitted subset to calculate its predicted value. A large bias obtained between model accuracy using all of the data and that quantified from an independent dataset (cross-validation) indicates that model is overfitted to the observations; consequently, the capability of the model to predict values for unsampled areas is questionable.

RESULTS

Effect of grid and sample size on model contribution by predictor variables

Figure 2 shows the mean contribution to the linear, tree, and GAM models (inter-grid size, inter-sample size and inter-replications average). It highlights the most influential variables in explaining snow-depth patterns in the catchment, revealing clear differences among the different regression-based methods. Altitude, curvature, and radiation are the most relevant contributors to the models, exceeding 10% in each of the three compared methods. Slope, northing, easting, and TPI are of secondary importance within the models, and CTI is the least influential variable.

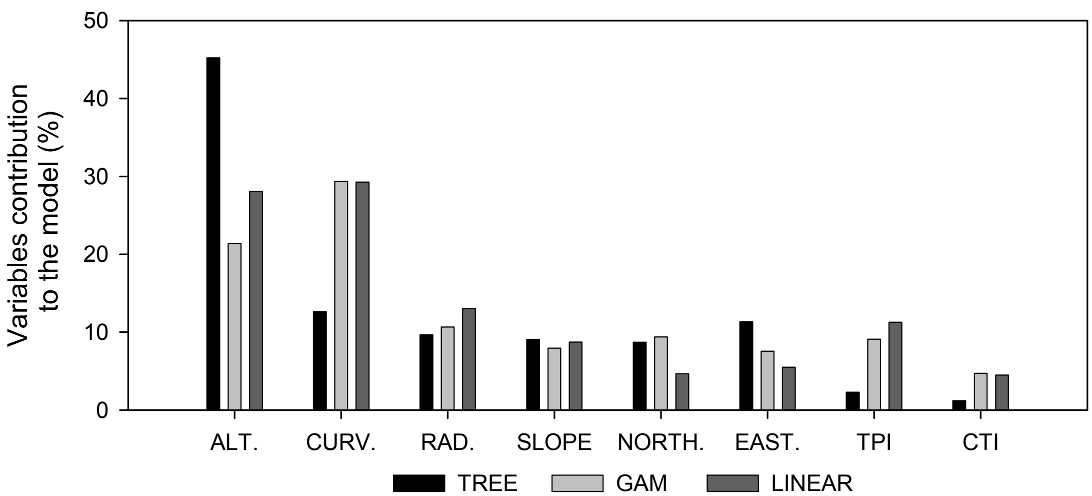

Figure 2. Mean variable contributions to the model (inter-grid size, inter-sample size, and inter-replication averages) of the different terrain characteristics considered as predictors.

Comparison of the three regression-based techniques reveals that tree models show important differences compared with the other two approaches. Altitude is of major importance in regression tree models, contributing on average around 45% to the models. Conversely, curvature, potential incoming radiation, slope, northing, and easting each make contributions of between 8.7 and 12.6% to regression tree models; TPI and CTI make only marginal contribution (2.3 and 1.2%, respectively).

The contributions of different terrain variables are somewhat similar between GAMs and linear models. In these two approaches, altitude and curvature are the most important predictors (each contributing around 20–30%); radiation and slope are also relevant factors (10%), while the contributions of the remaining variables are less than 4.5%.

Behind the averaged values shown in Figure 2, there exists a clear variability in the contribution of different terrain characteristics to the models when grid size and sample size are considered. Figure 3 shows the effect of sample size and grid size on the contributions of altitude, curvature, and radiation in different regression-based methods. The most important pattern observed in Figure 3 is the increasing contribution of altitude with decreasing spatial resolution of the DEM. With increasing grid resolution, curvature gains in importance in terms of its total contribution: this trend is particularly evident for tree models. The role of radiation tends to be relatively constant for different grid and sample sizes (with contributions of 10–20%), becoming slightly more important with increasing grid size.

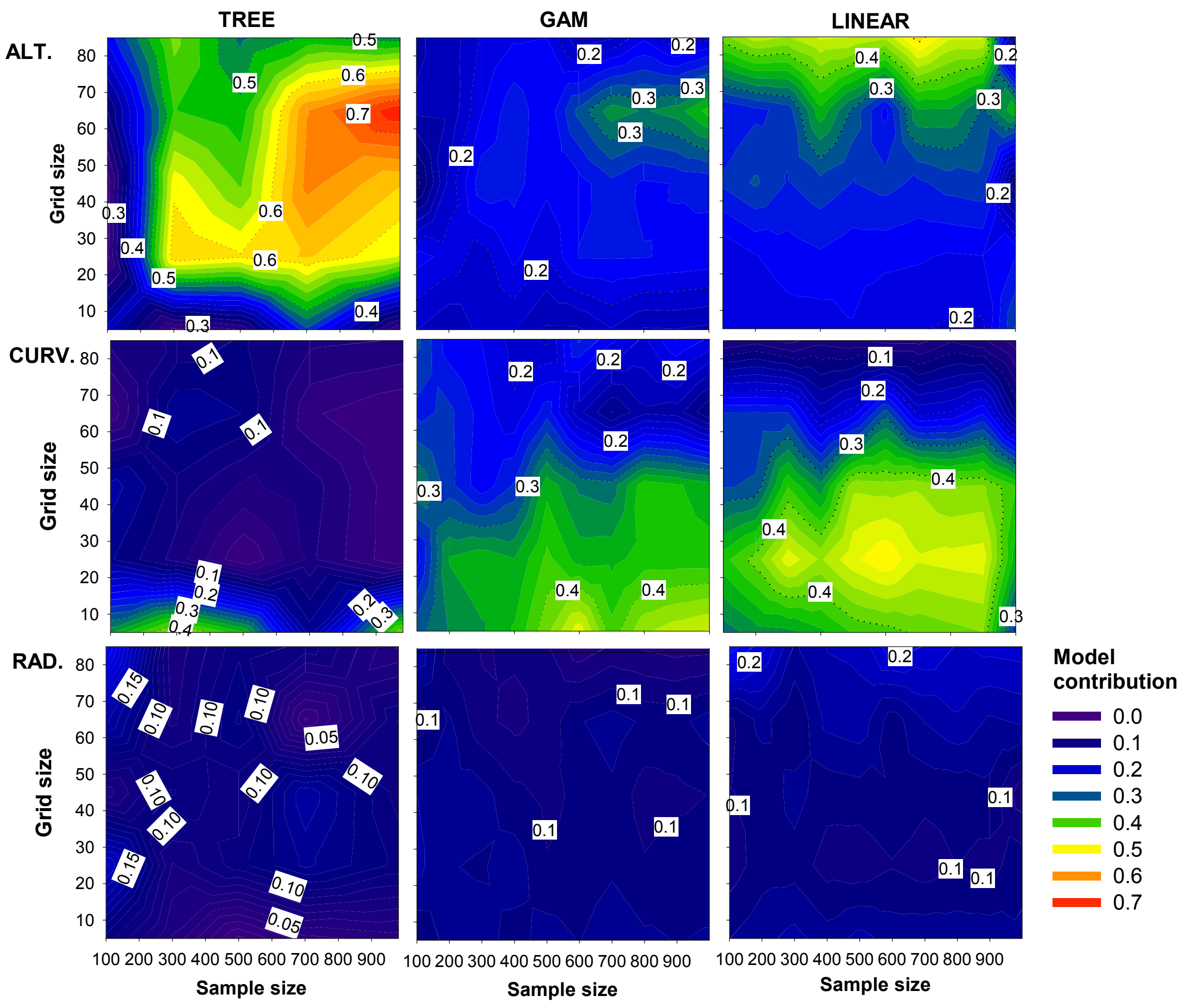

Figure 3. Effect of sample and grid size on contributions to the model of altitude, curvature, radiation, and slope in different regression-based methods.

Effect of grid and sample size on model accuracy

Figure 4 shows model accuracy calculated by cross-validation (average of Willmott's D of 10 replications) for the four regression-based methods under all possible combinations of grid and sample size. Also shown are Willmott's D values calculated using the same dataset as that used for modeling (dotted lines). Focusing on the results obtained by cross-validation, Figure 4 shows that application of the tree model to GAM residuals provides the best accuracy under all combinations of grid and sample size when an independent dataset is used for validation. For this method, D values range from 0.57 (sample size = 100; grid size = 95 m) to 0.82 (sample size = 991; grid size = 5 m). The accuracy of this method is more strongly related to grid size than sample size, as for a given grid size the accuracy remains relatively constant with increasing sample size above 200–300 cases.

The rest of the regression methods exhibit similar accuracy levels; however, there exist interesting differences associated with the number of cases and grid resolution. Regression trees results are strongly affected by sample size, as this method requires the most cases to perform reasonably accurate predictions; however, tree models are less sensitive to grid resolution than linear models and GAMs, providing better predictions for appropriate sample sizes (more than 500 data points). GAMs and linear models show similar accuracy under different grid and sample sizes, with GAMs being slightly more accurate than linear models. For the remaining methods, prediction quality remains relatively constant for a given grid size when sample size exceeds 300 cases; above this threshold, accuracy increases with DEM resolution. The best predictions (D > 0.75) obtained using GAMs and linear models are obtained with grid cells smaller than 25 m. Conversely, the use of cells larger than 65 × 65 m yields predictions that show a marked departure from observed values (D < 0.6).

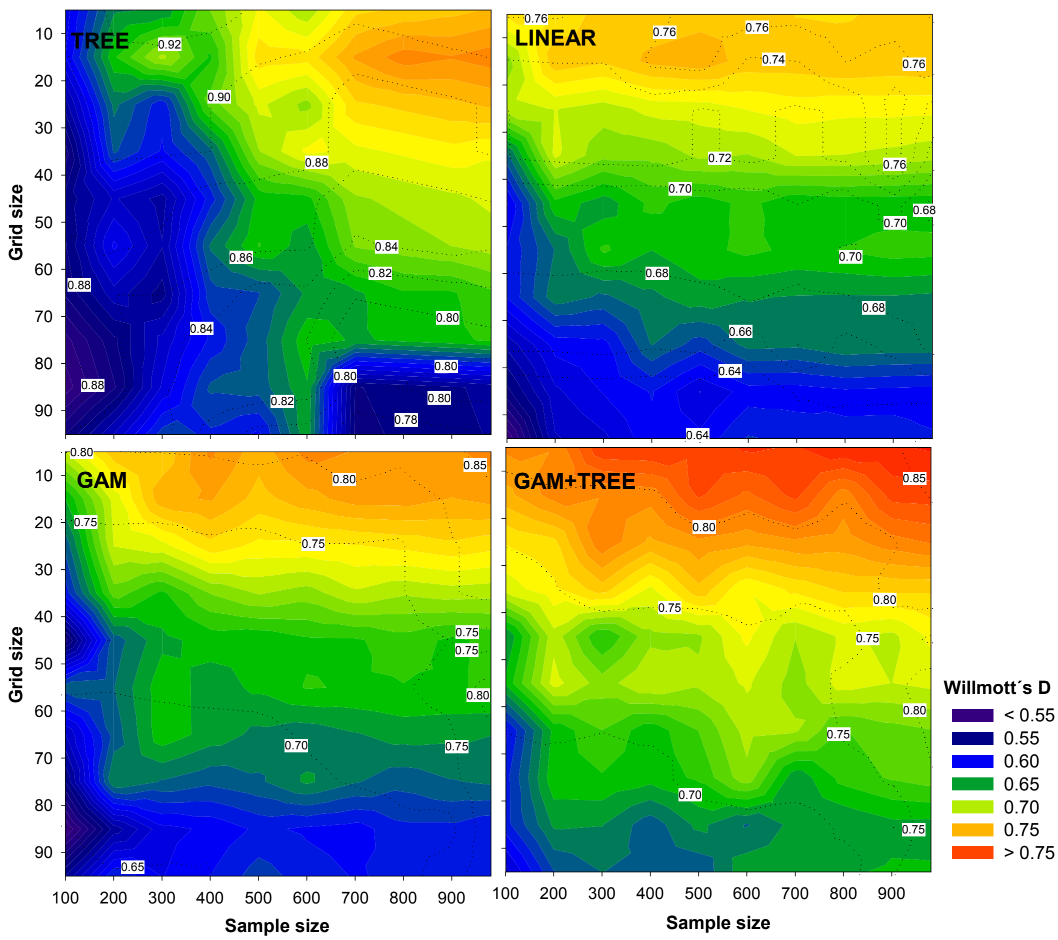

Figure 4. Effect of sample and grid size on model accuracy. Labelled dotted lines are Willmot´s D values calculated using the same dataset as that used for modelling. Colours are Willmott´s D values calculated by cross-validation.

Figure 4 also enables a comparison of model accuracies determined (i) using the same dataset as that used for modeling and (ii) calculated from an independent dataset (cross-validation). In most cases, Willmott's D value decreases noticeably when an independent dataset is used, suggesting that the models tend to overfit their predictions to observations. The accuracy determined from the same dataset as that used for modeling appears highly insensitive to sample size, even when 100 cases are used; grid size affects model accuracy, but to a much lesser degree than when an independent dataset is used.

The greatest differences between the two procedures employed in validating model quality are observed for tree models. For this method, values obtained using the same dataset as that for model evaluation are always above 0.74, even exceeding 0.9; in contrast, they never exceed 0.75, and are commonly often below 0.6, when an independent dataset is used.

GAMs are also moderately affected by overfitting, with a fall in Willmott's D values of between 5 and 10 units when an independent dataset is used. When the GAM residuals are corrected via a tree model, smaller differences are observed between the two types of validation. Thus, the accuracy scores obtained using the dataset employed for modeling are closely similar to those obtained by GAMs, but the prediction accuracies determined by cross-validation are superior. Linear models appear to be the most robust in relation to overfitting problems, with the two validation approaches yielding similar accuracy levels.

DISCUSSION AND CONCLUSIONS

The main results of this study are as follows:

1- On average, the most important variables in explaining snow distribution in the basin are altitude, curvature, and radiation. Although the contributions of the remaining predictors were relatively low, they still played a significant role in the employed models.

2- When the effects of grid and sample size are taken into account, grid size appears to be the most influential in determining the contributions of the different variables. In general, with a high-resolution DEM, curvature is the main predictor variable. With decreasing grid resolution, the contribution of curvature falls away, with altitude and solar radiation increasing in importance.

3- Grid and sample size are strong determinants of the accuracy of model predictions. Taken together, these factors may account for in excess of 0.3 Willmott's D units. Model accuracy is generally more sensitive to grid size than to sample size.

4- The application of tree models to GAM residuals provides the best accuracy scores. Although the other methods are less accurate, but still provide reasonable performance under various sample and grid sizes.

5- The four regression-based techniques show different thresholds in sample size above which the observed increases in model accuracy are very low. These thresholds should be considered when selecting the optimal number of required data points for an analysis.

6- In general, high accuracy scores are obtained with DEM grid cells smaller than 25 × 25 m. Grid sizes larger than 55 × 55 m appear to be inappropriate for modeling snowpack at the spatial scale considered in this study.

7- Model assessment must be performed using an independent dataset; otherwise, the quality of the results may be a consequence of overfitting to the used observations, and the potential error in unsampled areas would remain unknown.

The above conclusions agree with the results of previous studies regarding the importance of DEM resolution and/or quality in snow model outputs (Tang et al., 2001; Wechsler, 2007). In the present study, accuracy levels fell with reduced spatial resolution, as high-resolution DEMs (grid size of 5 × 5 m) are able to capture terrain features in detail, whereas those at lower resolutions do not capture the sharpness of the relief.

One of the main results of this work is the observed nonlinear response of model stability and accuracy to sample size. Between 200 and 400 observations seem to be sufficient to obtain accurate and robust models; the inclusion of additional cases generally results in only modest improvements, with an increase of less than 0.03 Willmott's D units. The detection of such thresholds, so-called data efficiency thresholds, has been described in several studies related to ecology (Peterson and Cohoon, 1999; David et al., 2002), but has remained barely addressed in snow studies.

Among the regression-based methods considered in the present study, the application of tree regression to GAM residuals has emerged as a promising tool for modeling snowpack distribution. The main strength of this approach is the combination of a highly flexible and nonlinear regression method (GAMs), which is valuable in snow modeling (Tyler et al., 2005; López-Moreno and Nogués-Bravo, 2006), with tree models, which are a powerful tool in considering interactions between variables (Breiman et al., 1984; Anderton et al., 2004; Molotch et al., 2005). Application of the three other regression models, using an appropriate DEM resolution and sample size, also yields satisfactory results.

The present results are not directly exportable to other geographical areas; moreover, they could vary with different previous climatic and/or nivological conditions to those of the snow sampling survey considered here. Thus, the aim of this work is to highlight the importance of grid and sample size in snow-model performance, as these factors are commonly selected in the absence of defined criteria or without the benefit of choice because of a lack of alternatives. The present results indicate that undertaking a number of slightly different analyses during the first surveys of a longer-term experiment may help to optimize the relation between the required amount of data (which is commonly difficult to collect in snow studies) and the quality of the results.

ACKNOWLEDGMENTS

The study was supported by the following projects: CGL 2004-04919-C02-01, CGL 2005-04508/BOS, and CGL2006-11619/HID financed by the Spanish Commission of Science and Technology and FEDER. J. Latron was the beneficiary of a research contract (Juan de La Cierva Programme) funded by the Spanish Ministry of Education and Science.

REFERENCES

Anderton SP, White SM, Alvera B. 2004. Evaluation of spatial variability in snow water equivalent for a high mountain catchment, *Hydrological Processes* **18** (3): 435-453.

Austin MP. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *EcologicalModelling* **157**: 101-118.

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

Cohen J, Cohen P, West SG, Aiken LS. 2003. *Applied multiple regression/correlation analysis for the behavioral sciences*. Lawrence Earlbaum Associates. Mahwah, NJ, USA. 537 pp.

David D, Stockwell RB, Townsend-Peterson A. 2002. Effects of sample size on accuracy of species distribution models. *Ecollogical Modelling* **148**: 2-13.

Gessler PE, Moore ID, McKenzie NJ, Ryan PJ 1995. Soil-landscape modeling and spatial prediction of soil attributes. *International Journal of GIS* **9 (4):** 421-432.

Jenness J. 2006. Topographic Position Index (tpi_jen.avx) extension for ArcView 3.x, v. 1.3a. Jenness Enterprises. Available at: http://www.jennessent.com/arcview/tpi.htm.

Jost G., Weiler M, Gluns DR, Alila Y. 2007. The influence of forest and topography on snow accumulation and melt at the watershed-scale. *Journal of Hydrology* **347**: 101-113.

Kienzle S. 2004. The effect of DEM Raster resolution on first order, second order and Compound Terrain derivatives. *Transactions in GIS* **8 (1):** 83-111.

López-Moreno JI, Nogués-Bravo D. 2006. Interpolating snow depth data: a comparison of methods. *Hydrological Processes* **20(10)**: 2217-232.

López-Moreno JI, Stähli M. 2008. Statistical analysis of the snowcover variability in a subalpine watershed: Assessing the role of topography and forest interactions. *Journal of Hydrology* **348 (3-4)**: 379-394**.**

Maggini R, Lehmann A, Zimmermann NE, Guisan A 2006. Improving generalized regression analysis for the spatial prediction of forest communities. *Journal of Biogeography* **33**: 1729-1749.
Molotch NP, Colee MT, Bales RC, Dozier J. 2005. Estimating the spatial distribution of snow water equivalent in an alpine basing using binary regression tree models: the impact of digital elevation data and independent variable selection. *Hydrological Processes* **19**: 1459-1479.
Peterson AT, Cohoon KP. 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. *Ecological Modelling* **117**: 159-164.
Pons X, Ninyerola X. 2008. Mapping a topographic global solar radiation model implemented in a GIS and refined with ground data. *International Journal of Climatology.*
Tang G, Hui Y, Strobl J, Liu W. 2001. The impact of resolution on the accuracy of hydrologic data derived from DEMs. *Journal of Geographical Sciences* **11 (4)**: 393-401.
Tyler A, Erickson T, Williams M. 2005. Persistence of topographic controls on the spatial distribution of snow in rugged mountain terrain, Colorado, United States. *Water Resources Research* **41:** W04014.
Wechsler S. 2006. Uncertainties associated with digital elevation models for hydrologic applications: a review. *Hydrology and Earth System Sciences* **11**: 1481-1500.
Willmott CT. 1982. Some comments on the evaluation of model performance. *Bulletin American Meteorological Society* **63** (11): 1309-1313.
Zhang W, Montgomery DR. 1994. Digital elevation model, gris size, landscape representation and hydrologic simulations. *Water Resources Research* **30**: 1019-1028.