The Cryosphere

Open Access

Snow cover thickness estimation using radial basis function networks

E. Binaghi, V. Pedoia, A. Guidali, and M. Guglielmin

Dipartimento di Scienze Teoriche e Applicate University of Insubria Via Mazzini 5, 21100 Varese, Italy

Correspondence to: E. Binaghi (elisabetta.binaghi@uninsubria.it)

Abstract. This paper reports an experimental study designed for the in-depth investigation of how the radial basis function network (RBFN) estimates snow cover thickness as a function of climate and topographic parameters. The estimation problem is modeled in terms of both function regression and classification, obtaining continuous and discrete thickness values, respectively. The model is based on a minimal set of climatic and topographic data collected from a limited number of stations located in the Italian Central Alps. Several experiments have been conceived and conducted adopting different evaluation indexes. A comparison analysis was also developed for a quantitative evaluation of the advantages of the RBFN method over to conventional widely used spatial interpolation techniques when dealing with critical situations originated by lack of data and limited *n*-homogeneously distributed instrumented sites. The RBFN model proved competitive behavior and a valuable tool in critical situations in which conventional techniques suffer from a lack of representative data.

1 Introduction

The increasing amount and quality of available geospatial, environmental data drive the need for new models with analytical recognition and predictive capabilities. These models, rooted in techniques such as knowledge-based systems, neural networks and soft computing open up new possibilities for the accurate estimation of complex environmental parameters (Belward et al., 2003). This work focuses on snow cover thickness modeling, an important scientific topic addressing snow avalanche risks, hydrological scope and permafrost distribution (Gong, 1996b). Snow is a significant environmental

and societal variable and it is also an important meteorological and climatological element. Estimating snow cover thickness is a complex task for several reasons. One of the main challenges is the fact that snow cover thickness is strongly influenced by many climatic and topographic variables whose to snow height is not well defined.

In addition to the theoretical issues, difficulties also arise during experiments. For example the processes of gathering, intersecting and clipping the available data to obtain an appropriate and coherent set of patterns, dramatically reduces the size of the initial data set. Several works demonstrated that neural networks (NNs) modeling a variety of nonlinear transfer functions can successfully address the above critical aspects (Bishop, 1995). NNs are distribution free and do not require data to conform to a fixed model, an aspect of great potential in the context of environmental studies which are based on the fusion of multiple heterogeneous data sets. The attractiveness of NNs also stems from learning capabilities, robustness and the ability to handle incomplete and imprecise information (Jain et al., 1996). NNs can provide practically accurate solutions for precisely or imprecisely formulated problems and for phenomena that are only understood through experimental data and field observations. Instead of assuming relationships between factors and output variables, the NN model can inductively learn these relationships from training data sets without requiring prior knowledge. The last 20 yr have seen a rapid growth in the use of NNs in geosciences and a variety of techniques have been investigated for different studies. Seminal papers proposing the use of neural computing for the analysis of geological and/or geophysical data have been published by the remote sensing community (Benediktsson et al., 1990; Gong, 1996b). These early works investigated the potential of NNs to handle

spatial data from multiple sources compared with conventional statistical and linear methods. Results were encouraging and confirmed the superiority of neural models in dealing with data with any measurement scale. The proven properties of high parallelism, robustness, the ability to handle imprecise and fuzzy information outweigh the difficulties associated with setting up suitable internal parameters and the complexity of the training stage. NNs have increasingly become practical tools for solving problems that more traditional systems have found intractable in several geoscience contexts (Lees, 1996; Tagliaferri et al., 2003) dealing with a variety of topics such as land cover mapping (Binaghi et al., 1997; Civco, 1993; Foody, 1995), landslide prediction (Binaghi et al., 2004; Lee et al., 2003; Guzzetti et al., 1999) or forecasting atmospheric events (Gardner and Dorling, 1998). NNs have been successfully applied to the analysis of climate variables enabling the construction of empirical models to estimate their temporal and spatial distributions. An adaptive basis function network for analysing trends in rainfall was proposed by Philip and Joseph (2003). Their study demonstrated experimentally that the periodicity of rainfall patterns may be understood using a neural model so that long-term predictions can be made. In Antoni et al. (2001) spatio-temporal distributions of climatic variables expressed at the level of monthly statistics are described as empirical functions of latitude, longitude, elevation and respective climatic time series obtained from a limited number of weather stations. To tackle the complexity of these nonlinear functional dependencies, a NN was used producing accurate results.

Among the many NN models available, the most used in geoscience and remote sensing studies has been the multilayer perceptron (MLP) coupled with the error back propagation (BP) algorithm. MLP networks are based on nonlinear sigmoid functions which give significant non-zero response in a wide region of the input space. Their approximations are smooth and continuous, and their accuracy increases with increasing numbers of nodes in the hidden layers. The benefits and limitations of MLP networks have become increasingly apparent and the results of comparative studies in diversified domains are now available (Corsini et al., 2003; Jayawardena et al., 1997). MLP is highly nonlinear in its parameters. The BP algorithm which uses the method of steepest descent does not guarantee convergence to a globally optimum set of parameters. In recent years research has developed on different types of feedforward networks. The various promising networks include the so-called radial basis function networks (RBFNs). These neural feedforward models are three-layer networks whose output nodes form a linear combination of the basis functions (usually of the Gaussian type) computed by the hidden layer nodes. Each node provides a significant non-zero response only when the input falls within a small localized region of the input space (Moody and Darken, 1989). Several studies proved theoretically and experimentally that RBFNs are capable of universal approximations and learning without local minima, thereby guaranteeing convergence to globally optimum parameters (Hush and Horne, 1993; Park and Sandberg, 1991). Moody and Darken (1989) also demonstrated that the RBF type networks learn faster than MLP networks.

Geoscience literature shows a growing interest in studies investigating the use of RBFNs to solve a variety of problems. Forecasting daily streamflow is, for example, a topic discussed in Moradkhani et al. (2004) and addressed by a RBFN integrating self-organizing principles. Dell'Acqua and Gamba (2003) used RBFNs both to approximate the rain field and to forecast the parameters of this approximation to anticipate the movements and changes in geometric characteristics of significant meteorological structures. The study reports decisively better performances than the MLP network and ordinary kriging. The present work investigates the performance of RBFNs in dealing with snow cover thickness estimation as a function of climate and topographic parameters. The RBFN model was chosen because of the small absolute number of sampling sites and by their non-homogeneous distribution in the study area. The estimation problem is modeled in terms of both function regression and classification, obtaining continuous and discrete thickness values respectively. The solutions investigated in this paper are an extension of those adopted in a previous work (Guidali et al., 2010) from which we inherit the choice of the neural model. Experiments have been extended reporting an in-depth analysis of the results. A new task concerning snow cover mapping has been inserted in which the neural model, trained on instrumented sites, is able to generalize the correlation between climate/geographic factors and precipitation measurements. The RBFN model acts as a spatial interpolation procedure estimating snow thickness values for non equipped sites. However, unlike the conventional spatial interpolation method, it does not consider true measurements values taken on a given date for a given map, but it induces approximate values (generalization stage) from a temporal series of instrumented values seen during the training stage (learning stage). Results obtained by the RBF model have been evaluated and compared with those produced by a snow mapping algorithm based on the Normalized Difference Snow Index (NDSI) derived from landsat imagery (Crane and Anderson, 1984). A comparison has been also developed using two of the most widely used spatial interpolation techniques (Tabios and Salas, 1985), the inverse distance weight (IDW) and Spline, to see whether the proposed RBFN method can be considered an alternative to conventional procedures in critical situations originated by lack of data and limited non-homogeneous instrumented sites.

2   Related work

The main contribution of this paper is the development of an RBFN model for snow thickness assessment. Therefore,

we surveyed the literature in the related field of neural models for snow parameter estimation. Machine learning methods have been widely used in hydrology and water resources (Gray and Male, 2004; Coulibaly et al., 2001; Agarwal et al., 2006). In particularly when applied to hydrologic time series modeling and forecasting the NNs have shown better performance than the classical techniques (Gong, 1996a). However, few studies involve the application of these methods in modeling snow parameters. In this context, NNs are especially applied to remote sensing data acquired from optical, infrared, and microwave sensors, allowing the estimation of snow thickness and snow water equivalent (SWE) parameters and the production of accurate arial measurements of snow cover extent. Tedesco et al. (2004) employed an MLP network for retrieval of snow depth and SWE from special sensor microwave imager (SSMI) data and compared the results with those obtained using the spectral polarization deference (SPD) algorithm, the Helsinki University of Technology (HUT) model-based iterative inversion and the Chang algorithm. They found that results obtained with the NN-based technique are better than or comparable to those obtained with other approaches. In Sun et al. (1997) classification techniques based on NNs were also successfully applied for deriving brightness temperature clusters related to different snow conditions from SSM/I data. In particular, a single-hidden-layer NN classifier was designed to learn the SSM/I signatures. A BP algorithm was applied and a winner-takes-all method was used to determine the snow condition. Results showed that the NN classifier was able to outline not only the snow extent but also the geographical distribution of snow conditions, thereby confirming the potential of using the nonlinear retrieval method for inferring land-surface snow conditions over varied terrain. Recently, a neural network approach was implemented to predicti changes in snow cover duration and distribution in the Black Forest mountain range of Germany (Sauter et al., 2010). On the basis of the International Panel on Climate Change (IPCC) A1B scenario, this work investigated the possible regional development of snow cover and snow duration in the Black Forest in southwest of Germany until 2050. In this application NNs are advantageous over other approaches as it is unnecessary to assume linear mixing of signals in a pixel. End members are not required and auxiliary information such as land cover can be easily incorporated. Once the neural network is trained, it is computationally efficient to produce snow fraction maps on both regional and global scales (Dobreva and Klein, 2011). Simpson and McIntire (2001) used a recurrent NN to differentiate between cloud, land, snow-covered and mixed pixels. Dobreva and Klein (2011) proposed a study in which a multilayer feed-forward NN trained through back propagation estimates fractional snow cover (FSC) using MODIS surface reflectance, NDSI, normalized difference vegetation index (NDVI) and land cover as inputs. The NN was trained and validated with higher spatial resolution FSC maps derived from Landsat Enhanced Thematic Mapper Plus (ETM+) bi-

nary snow cover maps. The literature reviewed highlights the following two main aspects: the most used NN models in snow studies have been MLP and recurrent networks; in almost all cases remote sensing plays a major role offering local, regional and global observations of snow, providing key information on snowpack processes. RBFNs have been used in few works. One of these by Xiao and Chandrasekar (1996) proposed a study in which an RBFN receives in input the whole radar observed reflectivity profile and produces snowfall estimations. The results of these early works are promising and suggest further studies and applications in snowpack processes.

Proceeding from these considerations, our study addresses the snow thickness estimation problem by measuring the ability of the RBFN to deal with difficulties arising from reduced data that do not include remote sensing data and limited instrumented sites. Under these conditions, the use of conventional statistical and parametric models is not recommended, being inapplicable and/or arbitrary (Binaghi et al., 1999).

Considering the different configurations with which neural models are considered in the reviewed studies (recurrent, recursive, feedforward) a key point investigated in our analysis is the design of a unique configuration of the RBF model for solving different tasks in snow thickness estimation.

3   Study area

Snow cover is essential to the survival of glaciers and permafrost areas characterizing the higher mountain landscape and ecosystems the of Central Italian Alps. The study area (Fig. 1a) includes the mountain sector of Lombardy Region, for a total area of about $8000\,\mathrm{km}^2$, located in the Central Italian Alps. The elevation (Fig. 1b) range varies between 186 m a.s.l. and 4025 m a.s.l. with an average of 818 m a.s.l. Mean annual air temperature (MAAT) ranges between -1.05 and $13.56\,°\mathrm{C}$ with a mean annual lapse adiabatic rate of $0.51\,°\mathrm{C}/100\,\mathrm{m}$ and consequently the isotherm $0\,°\mathrm{C}$ located at 2663 m a.s.l. The precipitation regime is extremely variable (ranging between 466 mm and 2254 mm) controlled mainly by the orographic systems. Precipitation is mainly in the form of snow above 2400 m a.s.l. between October and May but the snow accumulation is much more variable than the precipitation regime due to wind redistribution, avalanches and differential melting as a function of the different nature of the surfaces.
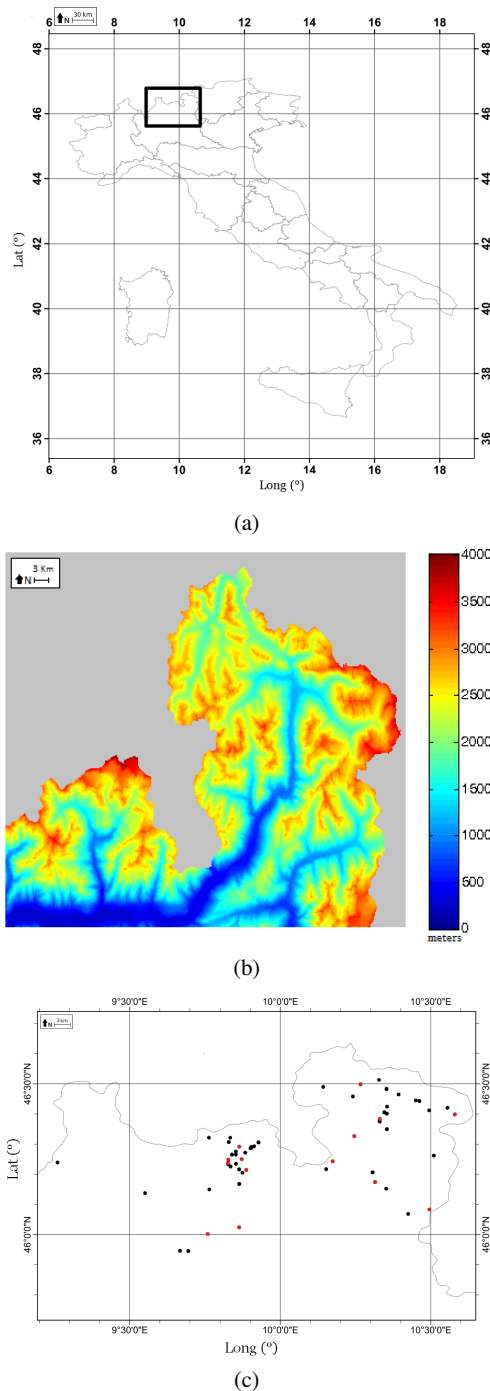
(a)

(b)

(c)

Fig. 1. (a) The study region enclosed in a black rectangle with latitudes (y-axes) and longitudes (x-axes) expressed. **(b)** The digital elevation model, related to the study area indicated in **(a)**. The locations of all stations of the initial data set equipped with different types of sensors are marked in figure **(c)**. The locations of the 14 stations equipped with air temperature sensor, snow thickness sensor and precipitation gauge, used for the final data set are colored in red.

4 Problem description

Several models and approaches were used to estimate the one-dimensional (z-direction) evolution of snow cover (e.g., Jordan, 1991; Melloh, 1999; Thorsen et al., 2010). When these models use only precipitation and air temperature as input data they require the definition of different physical thresholds. The literature shows that there is no universally accepted method for the evaluation of snow height that can be applied in every condition; often the choice of the most suitable method for the estimation of snow cover thickness (but also of other climatic data) depends on temporal resolution, spatial resolution, data quantity and also on the region of interest. Proceeding from these considerations we propose a model based on the following input variables:

1. Climatic:

   (a) Daily min temperature

   (b) Daily mean temperature

   (c) Daily max temperature

   (d) Daily precipitation

   (e) Cumulative rain over a given temporal interval T

   (f) Mean of measures in 1a (Daily min temperature) within interval T

   (g) Mean of measures in 1b (Daily mean temperature) within interval T

   (h) Mean of measures in 1c (Daily max temperature) within interval T.

2. Geographic:

   (a) Elevation

   (b) Aspect

   (c) Slope.

We assume that the value for T, the time interval in which the variables 1a–c are averaged for finding the variable h–f, can be heuristically assessed by a trial and error procedure during experiments (see Sect. 7). Based on the above listed climatic and geographic factors, we approach the snow cover thickness estimation as both a function regression and classification problem. In the first case, the output variable *snow cover thickness* assumes continuous values; in the second case it is modeled as a discrete variable whose values are labels of classes. Classes are made to correspond with specific sub-intervals of the overall range of variability of the snow cover thickness values. These intervals should be defined in the light of the object of the study. In our context, for the purpose of defining the classification scheme we used the following classes which are commonly used to evaluate permafrost distribution.

– Class A: absence of snow cover

– Class B: 1–10 cm

– Class C: 11–90 cm

– Class D: greater than 90 cm.

Table 1 summarizes inputs and outputs for regression and classification tasks.

5 Data set

The data set used was provided by Regione Lombardia[1] and consists of the climatic series recorded between 1987 and 2003 by 136 climatic sensors located in 64 different locations (Fig. 1c). Unfortunately the series has different temporal length and in many cases large gaps. Moreover the stations were equipped in different ways with different sensors. Consequently the derived data set was built using the spatial intersection of stations equipped with the instruments needed; this operation reduced the number of the stations to 16. In order to have comparable data, we decided to use as source of input data only the observations for a restricted time period (2002–2003). With the temporal intersection we obtained the final data set composed of 5476 observations heterogeneously distributed in 14 different locations (Fig. 1c in red). Table 2 lists the final data available for each station distinguished by months. Looking into the details, we observe that May, July and August are the most critical months in which 46 % of the data have gaps. Data related to input variables 1a, 1b, 1c listed in Sect. 4 are instrumental measurements drawn directly from the climatic database. Values of variables listed as 1e, 1f, 1g, 1h are obtained by applying cumulative and average procedures. Values of geographic variables are extracted from the digital elevation model (DEM) downloaded free of charge from the geoportal of Regione Lombardia and has a spatial resolution of 20 m. Aspect and slope were computed using the dedicated tools (Spatial Analyst Tools) of the ArcGIS 10 (ArcEditor).

6 Radial basis function networks

Our study modeled snow cover thickness estimation as a neural learning task according to which the correlation between climatic/geographic factors and snow cover thickness is inferred by induction from supervised input-output pairs of data. We adopt radial basis function networks in both regression and classification task.
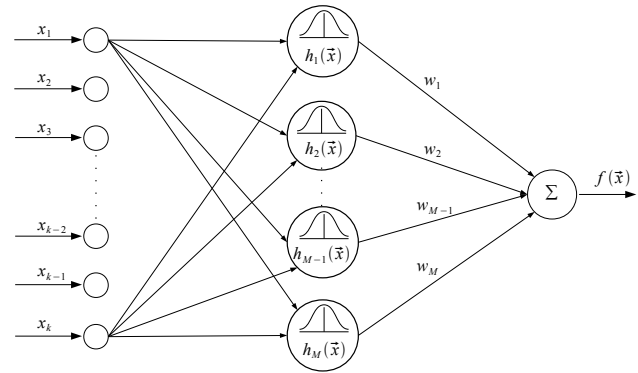
Fig. 2. Radial basis function network architecture for function approximation.

6.1 RBFN architecture and training algorithm

RBFNs are characterized by a very simple three layer architecture. The input layer propagates input values to a single hidden layer. In the output layer, each neuron receives a linear combination of the output of hidden neurons. In the case of one output node, the global nonlinear function computed by the network can be expressed as a linear combination of M basis functions associated with each hidden layer neuron. In the formula we have

$$f(\boldsymbol{x}) = \sum_i^M w_j h_j(\boldsymbol{x}), \tag{1}$$

where $\boldsymbol{x} = [x_1, .., x_k]^T$ is the K-dimensional input vector, w_j are the weighting coefficients of the linear combination and $h_j(\boldsymbol{x})$ represents the output of the Gaussian shaped basis function, with scale factor r_j, associated with the j-th neuron in the second layer. The response of j-th neuron decreases monotonically with the distance between the input vector $\boldsymbol{x}$ and the center of each function $\boldsymbol{c}_j = [c_{1j}, \ldots, c_{kj}]$:

$$h_j(\boldsymbol{x}) = \exp\left(-\frac{||\boldsymbol{x} - \boldsymbol{c}||^2}{r_j}\right). \tag{2}$$

During the training phase, the RBFN learns an approximation for the true input–output relationship based on a given training set of examples constituted by N input-output pairs $\{\boldsymbol{x}_i, y_i\}, i = 1, 2, \ldots, N$. Following Moody and Darken (1989), the training scheme is two-phased:

1. phase one is unsupervised and decides values for $\boldsymbol{c}_j, j = 1, \ldots, M$,

2. phase two solves a linear problem to find values for $w_i, i = 1, \ldots, M$.

Table 1. Input and output variables used in regression and classification tasks.

Features	Model	Output Setting
1a: Daily min temperature 1b: Daily mean temperature 1c: Daily max temperature 1d: Daily precipitation 1e: Cumulated rain on T 1f: Mean of measures in 1a on T 1g: Mean of measures in 1b on T 1h: Mean of measures in 1c on T 2a: Elevation 2b: Aspect 2c: Slope	Regression Classification	Real value [0cm–400cm] A: absence of snow cover B: 1cm–10cm C: 11cm–90cm D: greater than 90cm

Table 2. Amount of data in the final dataset. The division into months and altitude has been applied to emphasize the diversity and incompleteness of the data structure.

Station	m a.s.l.	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Tot
						Months								
Branzi	830	23	25	28	0	0	56	0	0	60	0	0	9	201
Grosio	1220	54	53	31	0	0	55	0	0	60	0	60	42	355
Val Torreggio	1350	17	28	35	30	0	60	0	0	60	31	30	50	341
Val Dorena	1575	46	22	7	0	0	0	0	0	0	4	29	31	139
Laghi di Chiesa	1596	59	56	62	60	2	59	0	0	60	62	59	61	540
Alpe Costa	1672	62	56	62	60	21	59	0	0	60	62	60	62	564
Piazzo Cavalli	1719	62	56	62	60	18	59	0	0	60	62	60	62	561
Monte Masuccio	1770	17	28	31	30	0	30	0	0	30	29	30	31	256
Carona	1955	31	27	30	30	1	30	0	0	30	40	60	52	331
Funivia Bernina	2014	62	56	62	59	32	0	62	62	7	62	60	62	586
Saviore dell'Adamello	2017	0	0	0	0	0	0	31	31	1	28	30	0	121
Cam Boer	2114	62	56	62	52	24	0	46	38	20	55	55	62	532
Monte Trela	2150	55	56	62	60	0	0	59	50	0	62	60	57	521
Isola Persa	2700	31	28	31	30	31	30	31	31	30	33	60	62	428
Tot		581	547	565	471	129	438	229	212	478	530	653	643	5476

The configuration of the model requires two user parameters:

1. the number M of first level local processing units and

2. the number p of the p-means heuristic (Moody and Darken, 1989), used to determine the scale factor r_j, $j = 1, \ldots, M$ of basis functions associated with first level processing units.

The second phase, having model parameters M, $c_j, j = 1, \ldots, M$, $r_j, j = 1, \ldots, M$ known, computes $w_i, i = 1, \ldots, M$ minimising the difference between predicted output and truth by least mean squares, computed through the pseudo inverse. In formula

$$w = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T y = \mathbf{H}^+ y, \qquad (3)$$

where

$$\mathbf{H} = \begin{pmatrix} h_1(\boldsymbol{x}_1) & h_2(\boldsymbol{x}_1) & \cdots & h_M(\boldsymbol{x}_1) \\ h_1(\boldsymbol{x}_2) & h_2(\boldsymbol{x}_2) & \cdots & h_M(\boldsymbol{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ h_1(\boldsymbol{x}_N) & h_2(\boldsymbol{x}_N) & \cdots & h_M(\boldsymbol{x}_N) \end{pmatrix} \qquad (4)$$

and $y = [y_1, ..y_N]$ is the vector of output data, $w = [w_1, ..w_M]^T$ are second level weights. The trained network is tested using a proper set of examples never seen during training.

6.2 Computation: regression and classification

This work is focused on the problem of learning an input-output mapping from a set of examples that can be regarded as an approximation of a multidimensional function. We investigated the behavior of RBFN when coping with

multidimensional function estimation modeled in the two different settings: regression and classification.

In the regression configuration, the RBFN learns from input-output pairs constituted as usual, by input patterns represented by a vector of measurements and output values representing numerical function values. The network is configured with a single output neuron.

We now formally define the components involved in the regression task. The set Ω is composed of all available data coupled with the relative truth value:

$$\Omega = \{(\boldsymbol{x}_i, y_i), i = 1, \ldots, N\}, \tag{5}$$

where $\boldsymbol{x}_i$ is a vector containing the input variables discussed in Sect. 4, y_i is the truth value related to $\boldsymbol{x}_i$ and N is the number of available data. The set Ω is then split into two partitions, namely the training set TrS_Ω and the test set TeS_Ω. If the regression task is arbitrary due to the poor reliability of the input data, the multidimensional function estimation can be conveniently modeled as a classification identifying intervals of the function codomain and making them correspond with a predefined class. The formal description of the model for classification can be easily obtained from the regression formalization, considering the value y_i in (5) as a label descrying membership in each of the defined classes. The network is configured with an output layer having a number of neurons equal to the number of classes. During training, input pattern vectors are made to correspond with predefined class labels, exemplifying a hard mapping at a lower granularity with respect to regression, with mutually exclusive classes.

7 Experiments

Our experiments adopted different evaluation indexes. The agreement between reference truth and classification results was analyzed by means of the confusion matrix and derived accuracy indexes (Congalton, 1991). A confusion matrix lists the values of the reference data in the columns and the values of the classified data in the rows. The main diagonal of the matrix lists the correctly classified pixels.

- Overall accuracy (OA): OA provides a general indication of the classifier's performance. The OA is equal to the ratio between the number of samples correctly classified, summing the values in the main diagonal of the confusion matrix, and the total number of observed samples.

- Producer accuracy (PA): PA measures the omission(exclusion) errors of the classifier. It is computed for each class. The PA for class i_{th} is equal to the ratio between the number of samples correctly classified in class i_{th} and the number of reference samples belonging to class i_{th} (column i_{th} total).

- User accuracy (UA): UA measures the commission(inclusion) errors of the classifier. It is computed for each class. The UA for class i_{th} is equal to the ratio between the number of samples correctly classified in class i_{th} and the total number of samples classified in class i_{th} (row i_{th} total).

The performance evaluation obtained using these indexes is complemented by the Cohen's kappa coefficient believed to be a more robust measure that takes into account the agreement occurring by chance (Cohen, 1960).

The kappa formal definition is

$$K = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)}, \tag{6}$$

where

$$\Pr(a) = \frac{\sum_{i=1}^{r} X_{i,i}}{N} \tag{7}$$

$$\Pr(e) = \frac{\sum_{i=1}^{r} X_{i,+} X_{+,i}}{N^2} \tag{8}$$

and

- r is the number of rows and columns in the confusion matrix;

- N is the total number of observations;

- $X_{i,j}$ is the observation in row i and column j;

- $X_{i,+}$ is the marginal total of the i-th row;

- $X_{+,i}$ is the marginal total of the i-th column.

Intuitively, $\Pr(a)$ is the proportion of judgments consistent among the classifier and the reference data and $\Pr(e)$ is the proportion of agreements that would be observed randomly classifying. In contrast to the overall accuracy, Cohen's kappa also takes non-diagonal elements into account. To measure the magnitude of network mistakes the root mean square error (RMSE) index and its normalized version NRMSE are used in combination with the mean absolute error (MAE). The formal definition of these index is given below:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum (\hat{y}_i - y_i)^2} \tag{9}$$

$$\text{NRMSE} = \frac{\text{RMSE}}{\text{MAX}(y_i)} \times 100 \tag{10}$$

$$\text{MAE} = \frac{1}{n} \sum |\hat{y}_i - y_i|, \tag{11}$$

where $\hat{y}_i$ is the estimated value. The overall data set composed of 5476 patterns was randomly split in the proportion of $\frac{2}{3}$, $\frac{1}{3}$ for training (TrS$_\Omega$) and test (TeS$_\Omega$), respectively. The radial basis function network configured for the tasks described above was applied to solve the problem of estimating snow cover thickness. The experiments focused on the parameters calibration process. A sensitivity analysis was conducted varying the input parameters described in Sect. 6.1. For the training phase focused on centroids identification, the K-means clustering algorithm was compared with a faster approach based on the random choice of M points in the input space. These two methods showed comparable performances. However, as the K-means algorithm imposes a small number of centroids to limit the computational complexity, the random choice strategy was preferred.

7.1 Regression results

First of all, we present the results obtained using the RBFN configured for the regression task. The RBFN receives in input the vector of measurements derived from the set of features (input variables) described in section 4. Concerning the network architecture, the input layer has 11 neurons equal to the number of features and the output layer has 1 neuron representing the predicted snow height value. Several configurations of the RBFN were considered varying the temporal window T used in the computation of features 1e, 1f, 1g and 1h, which assumed values ranging from 10 °C to 45 °C. For each window size, different RBFN configurations were considered distinguished by the different number M of basis functions which assumed values 100, 250, 500, 600, 750. The RBFN network showed the best behavior setting the temporal window dimension at 45. Table 3 shows the results obtained in this configuration, varying the neural internal parameter M. Results are expressed in terms of RMSE, NRMSE and MAE indexes and are obtained training and testing the network with five different pairs TrS$_\Omega$ and TeS$_\Omega$, randomly generated from the overall dataset Ω and averaging the individual indexes obtained. Above the value of $M = 500$ the error indexes significantly decrease for each incremental increase in M. After the value $M = 500$ the error indexes as a function of M show an asymptotic behavior. We than chose than this value as a reference for an optimize balancing between computational cost, training accuracy and generalization power. For an in-depth analysis of how the snow cover thickness was modeled, Fig. 3 plots the estimated values versus truth values which are sorted in ascending order. Figure 5a and b show the mean weekly error of the modeled snow cover with respect to the weekly mean of the liquid precipitation (rain) of two automatic weather stations (AWS) representative of the lower altitude (ST1) and the higher altitude (ST2) for 2002 and 2003, respectively. There is no significant relationship between the measured liquid precipitation and the errors although there is a general increase in errors in late Fall and Spring. At that time the greatest differences of snow height

Table 3. Regression results varying the number of centroids M, expressed in terms of RMSE, NRMSE and MAE.

# Centroids	RMSE	NRMSE	MAE
100	26.22	6.60 %	15.18
250	22.31	5.61 %	12.03
500	18.20	4.58 %	9.58
600	17.63	4.44 %	9.28
750	16.99	4.27 %	8.74

at different altitude occur because precipitation is in the form of snow only at higher elevation. Nevertheless, some error peaks (like the first week of 2002 and 2003 and the week 21 of 2003) are clearly not related to the data input. The RMSE values obtained indicate an acceptable mean disagreement between reference and predicted values. However, we have to consider that different intervals within the snow height range have different relevance in the environmental analysis, and errors computed on these intervals become unacceptable making arbitrary numerical predicted values. We proceeded to model snow cover thickness estimation as a classification task.

7.2 Classification results

The criteria with which we introduce classes subdividing snow height intervals, are derived from the idea of using them within a more general study concerning permafrost. The modeling of permafrost is outside the scope of the present paper which is focused on snow depth. Indeed, during the winter, the ground suffers from the extremely low air temperature in areas where the snow is absent (class A), while class B (1–10 cm) has been distinguished from C (11–90 cm) because a certain percentage of radiation in class B can reach the ground surface, in addition to the more limited insulation of the thinner snow with respect to the air temperature. Above 90 cm (class D) the insulation of the snow can be considered almost total.

For this task the network architecture is structured as follows: the input layer has 11 neurons equal to the number of features and the output layer has 4 neurons equal to the number of classes. Several configurations of the RBFN were considered varying parameters T and M as described for regression in the task. Also in this case, the RBFN showed the best behavior setting the temporal window dimension T at 45 and the number of centroids M equal to 500. Results obtained with this configuration are shown in Table 4. The RBFN performs classification with a good level of accuracy, showing an OA equal to 85.6 % and a Cohen's kappa coefficient equal to 79.5 %. In more detail, the crucial class A, representing the absence of snow cover, has the highest performance expressed in terms of both PA and UA, with values of 95.05 % and 92.47 %, respectively. This result can be explained by the fact that feature values are highly representative for

Table 4. Confusion matrix for the radial basis function network classifier evaluated on the overall test set TeS_Ω; Class A: absence of snow cover, Class B: 1–10 cm, Class C: 11–90 cm, Class D: greater than 90 cm.

Class data	Reference data				Tot U	UA
	A	B	C	D		
A	442	31	5	0	478	92.47 %
B	21	377	76	4	478	78.87 %
C	2	92	603	15	712	84.69 %
D	0	0	17	141	158	89.24 %
Tot P	465	500	701	160	–	–
PA	95.05 %	75.40 %	86.02 %	88.12 %	–	–

Total accuracy: 85.5969 % (1563 hit, 263 miss, 1826 total)
Total error: 14.4031 %
KAPPA value: 79.5523 %
KAPPA std.err: 0.0001

Table 5. Confusion matrix for the radial basis function network classifier evaluated on the test set TeS_Ω with elevation below 1000 m; Class A: absence of snow cover, Class B: 1–10 cm, Class C: 11–90 cm, Class D: greater than 90 cm.

Class data	Reference data				Tot U	UA
	A	B	C	D		
A	38	4	0	0	42	90.48 %
B	0	10	3	0	13	76.92 %
C	0	0	0	0	0	//
D	0	0	0	0	0	//
Tot P	38	14	3	0	–	–
PA	100.00 %	71.43 %	0 %	//	–	–

Total accuracy: 87.2727 % (48 hit, 7 miss, 55 total)
Total error: 12.7273 %
KAPPA value: 69.1259 %
KAPPA std.err: 0.0077

Table 6. Confusion matrix for the radial basis function network classifier evaluated on the test set TeS_Ω with elevation between 1000 and 1300 m; Class A: absence of snow cover, Class B: 1–10 cm, Class C: 11–90 cm, Class D: greater than 90 cm.

Class data	Reference data				Tot U	UA
	A	B	C	D		
A	39	0	0	0	39	100.00 %
B	2	67	3	0	72	93.06 %
C	0	2	4	0	6	66.67 %
D	0	0	0	0	0	//
Tot P	41	69	7	0	–	–
PA	95.12 %	97.10 %	57.14 %	//	–	–

Total accuracy: 94.0171 % (110 hit, 7 miss, 117 total)
Total error: 5.9829 %
KAPPA value: 88.4322 %
KAPPA std.err: 0.0017

Table 7. Confusion matrix for the radial basis function network classifier evaluated on the test set TeS_Ω with elevation between 1300 and 1600 m; Class A: absence of snow cover, Class B: 1–10 cm, Class C: 11–90 cm, Class D: greater than 90 cm.

Class data	Reference data				Tot U	UA
	A	B	C	D		
A	96	7	0	0	103	93.20 %
B	3	96	23	0	122	78.69 %
C	0	17	115	0	132	87.12 %
D	0	0	0	0	0	//
Tot P	99	120	138	0	–	–
PA	96.97 %	80.00 %	83.33 %	//	–	–

Total accuracy: 85.9944 % (307 hit, 50 miss, 357 total)
Total error: 14.0056 %
KAPPA value: 78.8497 %
KAPPA std.err: 0.0008

Table 8. Confusion matrix for the radial basis function network classifier evaluated on the test set TeS_Ω with elevation between 1600 and 1900 m; Class A: absence of snow cover, Class B: 1–10 cm, Class C: 11–90 cm, Class D: greater than 90 cm.

Class data	Reference data				Tot U	UA
	A	B	C	D		
A	114	8	1	0	123	92.68 %
B	14	95	13	0	122	77.87 %
C	0	27	183	0	210	87.14 %
D	0	0	0	0	0	//
Tot P	128	130	197	0	–	–
PA	89.06 %	73.08 %	92.89 %	//	–	–

Total accuracy: 86.1538 % (392 hit, 63 miss, 455 total)
Total error: 13.8462 %
KAPPA value: 78.6163 %
KAPPA std.err: 0.0006

discriminating between "presence of snow" or "absence of snow". The worst case is assigned to class B with PA equal to 75.40 % and UA equal to 78.87 %. This result can be correlated with the narrow range of snow cover thickness assigned to this class. Class D is the least represented, but since it is defined with a large interval, good performances were found in the experiments. Tables 5–10 show results obtained by partitioning the global results shown in Table 4 for different levels of elevation. Inevitably results of certain classes show low statistics due to the studied phenomenon. The second class is confirmed critical with the exception of the results in Table 6 related to elevation between 1000 and 1300 m. Misclassification errors are mostly committed between class B and class C at all different elevations considered, as already seen in Table 4.

Table 9. Confusion matrix for the radial basis function network classifier evaluated on the test set TeS$_\Omega$ with elevation between 1900 and 2200 m; Class A: absence of snow cover, Class B: 1–10 cm, Class C: 11–90 cm, Class D: greater than 90 cm.

Class data	Reference data				Tot U	UA
	A	B	C	D		
A	155	12	4	0	171	90.64 %
B	2	109	34	4	149	73.15 %
C	2	46	250	9	307	81.43 %
D	0	0	13	53	66	80.30 %
Tot P	159	167	301	66	–	–
PA	97.48 %	65.27 %	83.06 %	80.30 %	–	–

Total accuracy: 81.8182 % (567 hit, 126 miss, 693 total)
Total error: 18.1818 %
KAPPA value: 73.6529 %
KAPPA std.err: 0.0004

Table 10. Confusion matrix for the radial basis function network classifier evaluated on the test set TeS$_\Omega$ with elevation above 2200 m; Class A: absence of snow cover, Class B: 1–10 cm, Class C: 11–90 cm, Class D: greater than 90 cm.

Class data	Reference data				Tot U	UA
	A	B	C	D		
A	0	0	0	0	0	//
B	0	0	0	0	0	//
C	0	0	51	6	57	89.47 %
D	0	0	4	88	92	95.65 %
Tot P	0	0	55	94	–	–
PA	//	//	92.73 %	93.62 %	–	–

Total accuracy: 93.2886 % (139 hit, 10 miss, 149 total)
Total error: 6.7114 %
KAPPA value: 85.6978 %
KAPPA std.err: 0.0019

7.3 Snow cover mapping

In order to exploit the potential of RBFN in estimating snow cover distribution, addressed, the production of snow cover maps which offer a synoptic view of the phenomenon under investigation. This task was tackled by proceeding in the spatial interpolation of input climatic variables and then by using the RBF network to compute the corresponding predicted snow cover value for each input pattern including climate and geographic input variable. The spatial interpolation of temperatures is accomplished considering the variation of this parameter as a function of the elevation (Dodson and Marks, 1997). With reference to a generic cell xy, steps were taken to homogenize the known values in terms of elevation. Homogenization was obtained by performing linear regression between elevation and temperature values. Setting a reference evaluation value, each known temperature value was shifted as a function of the angular coefficient of the lin-
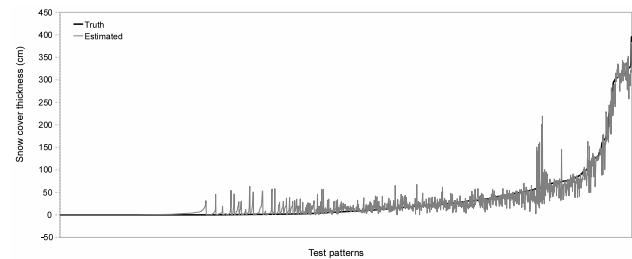

Fig. 3. Estimated snow cover thickness versus measured values in ascending order.

ear dependence law. Subsequently spatial interpolation was performed by applying the inverse square distance method obtaining temperature values for each grid unit. These values were finally modified reporting them at the original elevation. The spatial interpolation of precipitation patterns is a critical aspect requiring domain dependent complex analysis. Our study adopted a simple and easily controlled method based on Voronoi tessellation, implicitly assuming that the known values are representative of a given area around the point of measurement (Kay and Kutiel, 1994). We refrained from conducting a more sophisticated analysis that could be arbitrary in our context. Nine weeks were chosen based on their relevance for permafrost evolution (aggradation/degradation) and their variability. For these reasons, a higher frequency of examined weeks was chosen for the end of the spring (during the melting period). For each of the nine weeks, a snow map was generated with each grid element representing the neural-computed snow cover thickness.

An indirect validation procedure was accomplished comparing our results with those obtained by a snow mapping algorithm based on the normalized difference snow index (NDSI) derived from Landsat imagery. The NDSI-based snow algorithm together with its physical assumptions and derived decision rules for producing binary snow cover maps is described in (Dozier, 1989). To proceed in the comparison, neurally computed snow cover thickness values were binarized heuristically setting the threshold to a value of 5 cm. This value is a threshold commonly used to define the days with snow on the ground (snow duration) (Hantel and Hirtl-Wielke, 2007). Comparative results have been organized in a confusion matrix in which the classes considered are the presence and absence of snow. Results obtained from NDSI-based snow algorithm serve as of reference data. Table 11 reports comparative results in terms of OA (overall accuracy), UA (user accuracy) and PA (producer accuracy) indexes.

In general, the RBFN maps overestimate the presence of snow with respect to the NDSI maps because they reflect the network of AWS used for the training and the microclimatic conditions around them. This means that during the melting season and early autumn when the snow distribution suffered more intensively from the site conditions and it is extremely inhomogeneous, the snow remains more

Table 11. Summary of the RBFN results.

| Data	OA	Presence of Snow		Absence of Snow	
		PA	UA	PA	UA
30-April-02	76.03 %	88.20 %	64.94 %	68.25 %	89.66 %
28-May-02	64.03 %	78.36 %	39.49 %	59.41 %	89.04 %
22-Oct-02	62.94 %	69.81 %	20.72 %	62.11 %	93.55 %
3-Dec-02	62.15 %	62.73 %	71.60 %	61.68 %	51.80 %
11-Feb-03	83.31 %	87.20 %	84.18 %	78.79 %	82.61 %
25-Feb-03	81.96 %	88.35 %	79.20 %	75.72 %	86.13 %
11-March-03	85.29 %	84.72 %	89.91 %	86.64 %	80.15 %
1-April-03	82.10 %	86.67 %	78.99 %	78.19 %	86.11 %
29-April-03	79.99 %	88.70 %	70.61 %	74.28 %	90.42 %
Mean	75.31 %	81.64 %	66.63 %	71.68 %	83.27 %
Std	9.55 %	9.43 %	22.49 %	9.31 %	12.49 %
Max	85.29 %	88.70 %	89.91 %	86.64 %	93.55 %
Min	62.15 %	62.73 %	20.72 %	59.41 %	51.80 %
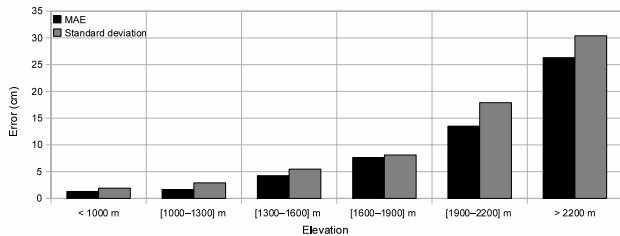
Fig. 4. Mean absolute error (MAE) and standard deviation in regression task as a function of elevation ranges.

probable on the flat point where AWS are located rather than in the other surrounding areas (generally except for northern exposed slopes). Table 12 shows OA values distinguished by different elevation ranges. Results obtained are generally good at low elevation (< 1600 m); at higher elevation, above 1900 m a.s.l., the accuracy decreases below 70 % during the melting season and at the beginning of Autumn, with a general underestimation. A different result was achieved between 1600 m and 1900 m, where poor results were obtained only during the winter core due to a problem of overestimation. The overall results obtained from the mapping procedure tallied in general with regression and classification results (see Figs. 3 and 4). Figure 6 shows the map produced by the RBF network when processing data for the week 11-03-2003. The overlap of this map, hardened with a 5 cm threshold, with the corresponding NDSI map is shown in Fig. 7. A snow scientist examined the maps in light of the topographic features of the study areas, and judged the results satisfactory.

7.4 Comparison analysis

A comparison analysis was developed for a quantitative evaluation of the advantages of the RBFN method with respect to conventional widely used spatial interpolation techniques when dealing with critical situations originated by
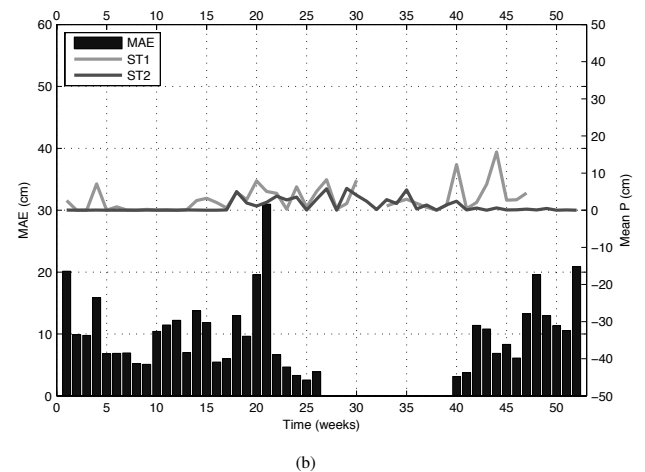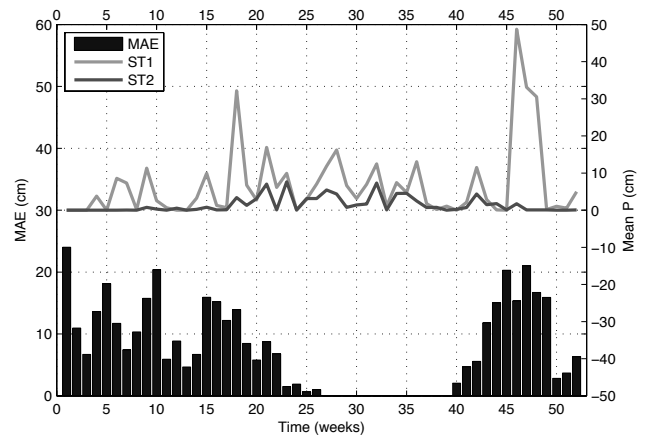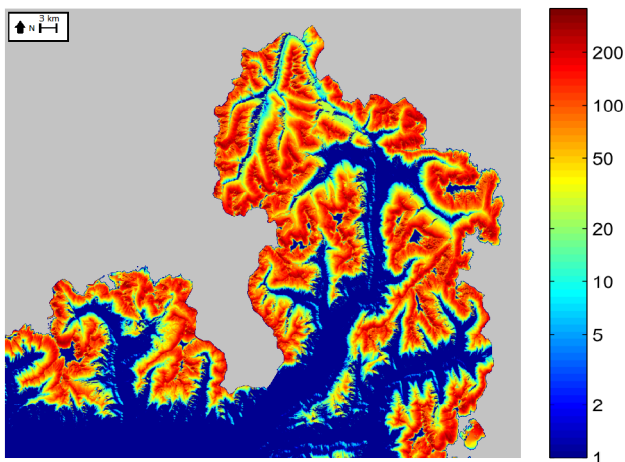

Fig. 5. Mean absolute error (MAE) in a regression task as a function of weeks during the years 2002 **(a)** and 2003 **(b)** compared with the mean weekly liquid precipitation measured close at the altitudinal limits of the data set: ST1 485 m a.s.l. and ST2 2150 m a.s.l.

lack of data and limited non-homogeneously distributed instrumented sites. Among the various techniques of spatial interpolation of hydrologic data, we adopted IDW and B-Spline that are widely diffused and of limited complexity (Tabios and Salas, 1985). We avoided the use of kriging techniques considering that the reduced sampling sites, the non-stationary character of the process under examination and the difficulty of configuring the variogram, can make the analysis arbitrary (Leuangthong et al., 2004). The analysis was conducted by comparing the two conventional techniques with the RBFN model. The binary maps produced by the NDSI decision rules served as of reference data. The IDW-based spatial interpolation was configured with a weighting power equal to 2 and the second method adopted a Cubic B-spline. We chose three weeks among the weeks in which the satellite images had a very low clouds coverage. In particular, we selected two weeks representing the full winter conditions and one spring week at the onset of snow melting
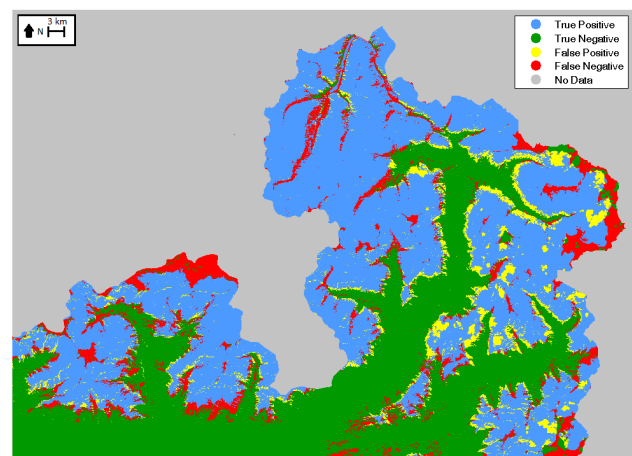
Table 12. Overall accuracy for 6 elevation ranges.

Date	Overall accuracy					
	0–1000	1000–1300	1300–1600	1600–1900	1900–2200	> 2200
30-April-02	99.95 %	94.69 %	99.82 %	82.73 %	56.13 %	68.14 %
28-May-02	99.98 %	98.98 %	99.99 %	91.41 %	47.42 %	42.61 %
22-Oct-02	99.30 %	96.25 %	99.14 %	91.94 %	58.70 %	37.15 %
3-Dec-02	70.60 %	51.78 %	50.82 %	55.14 %	74.42 %	61.33 %
11-Feb-03	99.67 %	91.98 %	93.60 %	67.60 %	60.77 %	87.44 %
25-Feb-03	99.98 %	96.32 %	98.79 %	72.89 %	46.92 %	86.17 %
11-March-03	99.93 %	97.40 %	94.43 %	73.35 %	69.26 %	86.51 %
1-April-03	99.78 %	99.80 %	99.48 %	84.02 %	65.42 %	77.13 %
29-April-03	99.97 %	98.31 %	99.92 %	91.96 %	61.96 %	71.93 %
Mean	96.57 %	91.72 %	92.89 %	79.01 %	60.11 %	68.71 %
Std	9.74 %	15.16 %	15.96 %	12.73 %	9.18 %	18.66 %
Max	99.98 %	99.80 %	99.99 %	91.96 %	74.42 %	87.44 %
Min	70.60 %	51.78 %	50.82 %	55.14 %	46.92 %	37.15 %

Fig. 6. Display on a logarithmic scale of the snow map produced by RBFN.

Fig. 7. Overlapping of maps produced by the RBFN Network and NDSI. Considering the RBFN map as the first result and the NDSI map as the second, colors have the following meaning: Blue = PS-PS; Green = AS-AS; Yellow = PS-AS; Red = AS-PS, with PS = Presence of Snow and AS = Absence of Snow

(29 April 2003). Both the full winter conditions and the early melting are crucial for several environmental issues, but mainly, for permafrost distribution because the snow-free areas in winter (above all at altitudes higher than 2200 m a.s.l. where air temperature are several degrees below 0 °C) are favorable for permafrost aggradation/conservation because the ground can be deeply frozen. On the other hand, the areas where snow melting is late-lying (and therefore later than the end of April, in this area) is also favorable for permafrost formation because positive air temperatures are prevented by the late snow cover.

Consistent with the procedure described in the previous section, maps produced by all the three methods were binarized setting the threshold value to 5 cm. Table 13 shows the OA values obtained by RBFN, IDW and Spline, distinguished by different elevation ranges consistently with the

analysis presented in the previous section. The OA values of the RBFN are always much higher than those obtained by the two conventional methods except between 1900 m and 2200 m a.s.l. in fully winter conditions when the spline algorithm performs slightly better. This exception is due to the higher number of weather stations available in this range of altitude. In fact, RBFN is particularly efficient, where the original dataset is poor as at altitude higher than 2200 m a.s.l. where the accuracy of RBFN is 72 % against values lower than 60 % of the two other methods. This result is particularly important because in all the mountain areas there are fewer weather stations at higher elevations where permafrost and glaciers are much more abundant.

Table 13. Comparison analysis among IDW, B-Spline and RBF: the results are presented in terms of overall accuracy for 6 elevation ranges.

Date	Methods	Overall accuracy						Mean	Var
		0–1000	1000–1300	1300–1600	1600–1900	1900–2200	> 2200		
11-Feb-03	RBF	99.67 %	91.98 %	93.60 %	67.60 %	60.77 %	87.44 %	*83.51 %*	*15.62 %*
	DW	7.46 %	12.90 %	15.45 %	29.50 %	59.86 %	85.58 %	*35.12 %*	*31.10 %*
	Spline	85.44 %	53.35 %	45.38 %	52.01 %	67.17 %	74.20 %	*62.92 %*	*15.30 %*
25-Feb-03	RBF	99.98 %	96.32 %	98.79 %	72.89 %	46.92 %	86.17 %	*83.51 %*	*20.63 %*
	IDW	7.47 %	9.68 %	11.07 %	17.95 %	41.86 %	80.64 %	*28.11 %*	*28.66 %*
	Spline	84.64 %	50.03 %	42.46 %	44.18 %	53.45 %	70.85 %	*57.60 %*	*16.68 %*
29-April-03	RBF	99.97 %	98.31 %	99.92 %	91.96 %	61.96 %	71.93 %	*87.34 %*	*16.38 %*
	IDW	60.41 %	35.13 %	30.97 %	27.23 %	33.62 %	58.42 %	*40.96 %*	*14.56 %*
	Spline	49.92 %	46.93 %	40.46 %	38.76 %	39.39 %	53.46 %	*44.82 %*	*6.17 %*

8   Conclusions

A method for snow cover thickness estimation is proposed within the context of a Permafrost studies in an alpine environment, based on the use of a radial basis function network capable of moving from regression and classification tasks which are usually complementary for understanding complex environmental phenomena at different levels of precision. As seen in the experimental context, the learning and approximation capabilities of the network allow the user to overcome the limitations of available data characterized by such critical aspects as the minimal set of climatic and topographic data and the reduced set of non-homogeneously distributed instrumented sites. The comparison analysis conducted between IDW, Spline interpolation techniques and the proposed model demonstrates the superiority of the neural implementation with respect to conventional deterministic approaches. The overall quantitative and qualitative evaluations of the experimental work substantiates the general claim that a variety of problems in hydrology and water resources studies benefit from the use of neural techniques, demonstrating in particular the performances of the RBFN model in the specific field of snow cover thickness estimation where its use is still little known. The results of our experiments suggest that RBFN approximation can provide a valuable solution for studies conducted in all mountain areas at high elevations where there are fewer weather stations and where permafrost and glaciers are more prevalent.

Edited by: J. L. Bamber

References

Agarwal, A., Mishra, S., Ram, S., and Singh, J.: Simulation of Runoff and Sediment Yield using Artificial Neural Networks, Biosyst. Eng., 94, 597–613, doi:10.1016/j.biosystemseng.2006.02.014, 2006.

Antoni, O., Krian, J., Marki, A., and Bukovec, D.: Spatio-temporal interpolation of climatic variables over large region of complex terrain using neural networks, Ecol. Model., 138, 255–263, 2001.

Belward, A., Binaghi, E., Lanzarone, G., and Tosi, G. (Eds.): Geospatial Knowledge Processing for Natural Resource Management, vol. 24 of Special Issue of International Journal of Remote Sensing, Taylor & Francis, 2003.

Benediktsson, J., Swain, P., and Ersoy, O.: Neural network approaches versus statistical methods in classification of multisource remote sensing data, IEEE T. Geosci. Remote, 28, 540–552, 1990.

Binaghi, E., Madella, P., Grazia Montesano, M., and Rampini, A.: Fuzzy contextual classification of multisource remote sensing images, IEEE T. Geosci. Remote, 35, 326–340, doi:10.1109/36.563272, 1997.

Binaghi, E., Brivio, P., Ghezzi, P., Rampini, A., and Zilioli, E.: Investigating the Behaviour of Neural and Fuzzy-Statistical Classifiers in Sub-Pixel Land Cover Estimation, Canad. J. Remote Sens., 25, 171–188, 1999.

Binaghi, E., Boschetti, M., Brivio, P., Gallo, I., Pergalani, F., and Rampini, A.: Prediction of displacements in unstable areas using a neural model, Nat. Hazards, 32, 135–154, 2004.

Bishop, C.: Neural networks for pattern recognition, Oxford University Press, USA, 1995.

Civco, D.: Artificial neural networks for land-cover classification and mapping, Int. J. Geogr. Inf. Sci., 7, 173–186, 1993.

Cohen, J.: A coefficient of agreement for nominal scales, Educ. Psychol. Meas., 20, 37–46, 1960.

Congalton, R.: A review of assessing the accuracy of classifications of remotely sensed data, Remote Sens. Environ., 37, 35–46, 1991.

Corsini, G., Diani, M., Grasso, R., De Martino, M., Mantero, P., and Serpico, S.: Radial Basis Function and Multilayer Perceptron neural networks for sea water optically active parameter estimation in case II waters: A comparison, Int. J. Remote Sens., 24, 3917–3931, http://www.informaworld.com/10.1080/0143116031000103781, 2003.

Coulibaly, P., Bobée, B., and Anctil, F.: Improving extreme hydrologic events forecasting using a new criterion for artificial neural network selection, Hydrol. Process., 15, 1533–1536, doi:10.1002/hyp.445, 2001.

Crane, R. G. and Anderson, M. R.: Satellite discrimination of snow/cloud surfaces, Int. J. Remote Sens., 5, 213–223, doi:10.1080/01431168408948799, 1984.

Dell'Acqua, F. and Gamba, P.: Pyramidal rain field decomposition using radial basis function neural networks for tracking and forecasting purposes, IEEE T.Geosci. Remote, 41, 853–862, 2003.

Dobreva, I. D. and Klein, A. G.: Fractional snow cover mapping through artificial neural network analysis of {MODIS} surface reflectance, Remote Sens. Environ., 115, 3355–3366, doi:10.1016/j.rse.2011.07.018, 2011.

Dodson, R. and Marks, D.: Daily air temperature interpolated at high spatial resolution over a large mountainous region, Climate Res., 8, 1–20, doi:10.3354/cr008001, 1997.

Dozier, J.: Spectral signature of alpine snow cover from the Landsat Thematic Mapper, Remote Sens. Environ., 28, 9–22, http://linkinghub.elsevier.com/retrieve/pii/0034425789901016, 1989.

Foody, G.: Land cover classification by an artificial neural network with ancillary information, Int. J. Geogr. Inf. Sci., 9, 527–542, 1995.

Gardner, M. and Dorling, S.: Artificial neural networks (the multilayer perceptron) – A review of applications in the atmospheric sciences, Atmos. Environ., 32, 2627–2636, 1998.

Gong, P.: Integrated analysis of spatial data from multiple sources: using evidential reasoning and artificial neural network techniques for geological mapping, Photogramm. Eng. Remote Sens., 62, 513–523, 1996a.

Gong, P.: Integrated analysis of spatial data from multiple sources: using evidential reasoning and artificial neural network techniques for geological mapping, Photogramm. Eng. Remote Sens., 62, 513–523, 1996b.

Gray, D. and Male, D.: Handbook of Snow: Principles, Processes, Management & Use, Blackburn Press, http://books.google.it/books?id=zhZFPgAACAAJ, 2004.

Guidali, A., Binaghi, E., Guglielmin, M., and Pascale, M.: Investigating the Behaviour of Radial Basis Function Networks in Regression and Classification of Geospatial Data, in: IDEAL'10, 110–117, 2010.

Guzzetti, F., Carrara, A., Cardinali, M., and Reichenbach, P.: Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy, Geomorphology, 31, 181–216, 1999.

Hantel, M. and Hirtl-Wielke, L.-M.: Sensitivity of Alpine snow cover to European temperature, Int. J. Climatol., 27, 1265–1275, doi:10.1002/joc.1472, 2007.

Hush, D. and Horne, B.: Progress in supervised neural networks, IEEE Signal Proc. Mag., 10, 8–39, 1993.

Jain, A., Mao, J., and Mohiuddin, K.: Artificial neural networks: A tutorial, Computer, 29, 31–44, 1996.

Jayawardena, A., Fernando, D., and Zhou, M.: Comparison of multilayer perceptron and radial basis function networks as tools for flood forecasting, IAHS-AISH P., 239, 173–182, 1997.

Jordan, R.: A one-dimensional temperature model for a snow cover: Technical documentation for SNTHERM. 89., CRREL Special Report 91-16, US Army Cold Regions Research and Engineering Laboratory, Hanover, NH, 1991.

Kay, P. A. and Kutiel, H.: Some remarks on climate maps of precipitation, Climate Res., 4, 233–241, 1994.

Lee, S., Ryu, J., Min, K., and Won, J.: Landslide susceptibility analysis using GIS and artificial neural network, Earth Surf. Proc. Land., 28, 1361–1376, 2003.

Lees, B. G.: Neural network applications in the geosciences: An introduction, Comput. Geosci., 22, 955–957, doi:10.1016/S0098-3004(96)00033-7, 1996.

Leuangthong, O., McLennan, J., and Deutsch, C.: Minimum Acceptance Criteria for Geostatistical Realizations, Nat. Resour. Res., 13, 131–141, doi:10.1023/B:NARR.0000046916.91703.bb, 2004.

Melloh, R.: A synopsis and comparison of selected snowmelt algorithms, CRREL Report 99-8, US Army Cold Regions Research and Engineering Laboratory. Hanover, NH, 1999.

Moody, J. and Darken, C.: Fast learning in networks of locally-tuned processing units, Neural Comput., 1, 281–294, 1989.

Moradkhani, H., Hsu, K., Gupta, H., and Sorooshian, S.: Improved streamflow forecasting using self-organizing radial basis function artificial neural networks, J. Hydrol., 295, 246–262, 2004.

Park, J. and Sandberg, I. W.: Universal approximation using radial-basis-function networks, Neural Comput., 3, 246–257, doi:10.1162/neco.1991.3.2.246,, 1991.

Philip, N. S. and Joseph, K. B.: A neural network tool for analyzing trends in rainfall, Comput. Geosci., 29, 215–223, 2003.

Sauter, T., Weitzenkamp, B., and Schneider, C.: Spatio-temporal prediction of snow cover in the Black Forest mountain range using remote sensing and a recurrent neural network, Int. J. Climatol., 30, 2330–2341, doi:10.1002/joc.2043, 2010.

Simpson, J. and McIntire, T.: A recurrent neural network classifier for improved retrievals of areal extent of snow cover, IEEE T. Geosci. Remote, 39, 2135–2147, doi:10.1109/36.957276, 2001.

Sun, C., Neale, C., McDonnell, J., and Cheng, H.-D.: Monitoring land-surface snow conditions from SSM/I data using an artificial neural network classifier, IEEE T. Geosci. Remote, 35, 801–809, doi:10.1109/36.602522, 1997.

Tabios, G. Q. and Salas, J. D.: A comparative analysis of techniques for spatial interpolation of precipitation1, J. Am. Water Resour. As., 21, 365–380, doi:10.1111/j.1752-1688.1985.tb00147.x, 1985.

Tagliaferri, R., Longo, G., D'Argenio, B., and Incoronato, A.: Introduction: Neural networks for analysis of complex scientific data: astronomy and geosciences, Neural Networks, 16, 295–295, doi:10.1016/S0893-6080(03)00012-1, 2003.

Tedesco, M., Pulliainen, J., Takala, M., Hallikainen, M., and Pampaloni, P.: Artificial neural network-based techniques for the retrieval of {SWE} and snow depth from SSM/I data, Remote Sens. Environ., 90, 76–85, doi:10.1016/j.rse.2003.12.002, 2004.

Thorsen, S., Roer, A., and Van Oijen, M.: Modelling the dynamics of snow cover, soil frost and surface ice in Norwegian grasslands, Polar Res., 29, 110–126, 2010.

Xiao, R. and Chandrasekar, V.: Multiparameter radar snowfall estimation using neural network techniques, in: Geoscience and Remote Sensing Symposium, 1996, IGARSS '96, "Remote Sensing for a Sustainable Future", International, 1, 566–568, doi:10.1109/IGARSS.1996.516405, 1996.