The Cryosphere Discuss., 6, 2437–2475, 2012
www.the-cryosphere-discuss.net/6/2437/2012/
doi:10.5194/tcd-6-2437-2012

The Cryosphere Discussions

Snow cover thickness estimation by using radial basis function networks

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

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

Received: 7 May 2012 – Accepted: 19 June 2012 – Published: 16 July 2012

Correspondence to: V. Pedoia (valentina.pedoia@uninsubria.it)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

This work investigates learning and generalisation capabilities of radial basis function networks (RBFN) used to solve snow cover thickness estimation model as regression and classification. 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 in both regression and classification tasks. The snow cover thickness estimation by RBFN has been proved a valuable tool able to deal with several critical aspects arising from the specific experimental context.

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 are rooted in new techniques based on knowledge-based systems, neural networks, fuzzy logic and hybrids soft computing frameworks. These advances in geospatial computational methods open up new possibilities for the accurate estimation of complex environmental parameters (Belward et al., 2003).

In this work we focus our attention on snow cover thickness modelling which is an important scientific topic studied for different purposes such as snow avalanche risks, hydrological scope and permafrost distribution (Gong, 1996). Snow plays a significant role as an environmental and societal variable and, at the same time, it is also an important meteorological and climatological element.

The investigation of this problem can be considered a complex task for several reasons. One of the main challenges is related to the fact that snow cover thickness is strongly influenced by many climatic and topographic variables and for each of them it is not well defined the contribution factor for the evaluation of the snow height. In addition to the theoretical issues difficulties arise, also during the experiments; for example the processes of gathering, intersecting and clipping the available data to obtain an appropriate and coherent set of patterns dramatically reduced the numerosity of the initial data set.

Several works demonstrated that neural networks (NNs) modelling a variety of non-linear transfer functions, can successfully address the above critical aspects (Bishop, 1995). NNs are distribution free and do not require that data 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 comes from learning capabilities, robustness and ability to handle incomplete and imprecise information (Jain et al., 1996). They 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 approach can inductively learn from training data sets these explicit relationships without requiring prior knowledge. In the last 20 yr we 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 first raising the use of neural computing for the analysis of geological and/or geophysical data were created in the remote sensing community (Benediktsson et al., 1990; Gong, 1996). In these early works the potential of neural networks in handling spatial data coming from multiple sources is investigated and compared with conventional statistical and linear methods. Results obtained were encouraging confirming the superiority of neural models in dealing with data with any measurement scale and determining how much influence a source should have in the integrated analysis. The proven properties of high parallelism, robustness, ability to handle imprecise and fuzzy information outweigh the difficulties associated with the setting up of suitable internal parameters and complexity in performing the training stage (Skidmore et al., 1997). 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 (Baraldi et al., 2001;

Civco, 1993; Foody, 1995), landslides prediction (Binaghi et al., 2004; Lee et al., 2003; Guzzetti et al., 1999), forecasting of atmospheric events (Gardner and Dorling, 1998). NNs have been successfully applied to the analysis of climate variables enabling the construction of empirical models for the estimation of 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 deal properly with the complexity of these non linear 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 multi-layer 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, more and more accurate for increasing numbers of nodes in the hidden layers. However, benefits and limitations of MLP networks have become more and more visible and 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 globally optimum set of parameters. In recent years researches concerning different types of feedforward networks have been developed. Among the various kinds of promising networks are the so-called radial basis function networks (RBFNs). These neural feed forward 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 localised region of the input space (Moody and Darken, 1989). Several studies proved theoretically and experimentally that the 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 (Moody and Darken, 1989) demonstrated also that the RBF type networks learn faster than MLP networks. In geoscience literature there is a growing interest in studies that investigate the use of RBFNs to solve a variety of problems. Forecasting of daily streamflow is, for example, the topic discussed in (Moradkhani et al., 2004) and addressed by a RBFN integrating Self Organising principles. Dell'Acqua and Gamba (2003) use radial basis function NNs (RBFN) both to approximate the rain field and to forecast the parameters of this approximation in order to anticipate the movements and changes in geometric characteristics of significant meteorological structures. The study reports performances decisively better than the feed-forward network and the ordinary kriging (OK).

In the present work we investigate the performance of RBFNs used to estimate snow cover thickness in function of climate and topographic parameters. Solution is modelled in terms of both function regression and classification tasks. 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 neural model. Experiments have been extended reporting a deeper analysis of the results. A new task concerning snow cover mapping has been inserted including the description of spatialisation procedures and evaluation by comparison with snow mapping algorithm based on the normalized difference snow index (NDSI) derived from Landsat imagery (Crane, 1984). In this context the RBFNs specifically address a critical situation originated by weak pattern description and incompleteness among data due to error measurements.

2 Study area

The study area (Fig. 1a) includes the mountain sector of Lombardia Region, for a total area of about 8000 km^2, located in the Italian Central Alps. The elevation (Fig. 1b) range

TCD
6, 2437–2475, 2012

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 (MAAN) range between −1.05 and 13.56 °C with a mean annual lapse adiabatic rate of 0.51 °C/100 m and consequently the isotherm 0 °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 are mainly in snow form above 2400 m a.s.l. between October and May but the snow accumulation is much more variable than the precipitation regime because wind redistribution, avalanches and differential melting in function of the different nature of the surfaces. Snow cover distribution is crucial to the surviving of glaciers and permafrost areas that characterise the higher mountain landscape and ecosystems of this sector of Central Italian Alps.

3 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 as input data only precipitation and air temperature they require the definition of different physical thresholds. The literature shows that there is no a 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 others 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

TCD
6, 2437–2475, 2012

d. Daily precipitation
e. Cumulative rain over a given temporal interval T
f. Mean of measures in 1 within interval T
g. Mean of measures in 1 within interval T
h. Mean of measures in 1 within interval T

2. Geographic:

a. Elevation
b. Aspect
c. Slope

We assume that the value for T can be heuristically assessed by a trial and error procedure during experiments (see Sect. 6). Basing on the above listed climatic and geographic factors, we approach the snow cover thickness estimation as both 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 put in correspondence with specific sub-intervals of the overall range of variability of the snow cover thickness values. The meaning of these intervals is related to the final objective of the study. Proceeding then from the consideration that snow cover thickness is a key factor for the evaluation of Permafrost distribution we introduce the following classes:

Class A: absence of snow cover

Class B: 1–10 cm

Class C: 11–90 cm

Class D: greater than 90 cm

In Table 1 inputs and outputs for regression and classification tasks are summarised.

TCD
6, 2437–2475, 2012

4 Data set

The data set used was provided by Regione Lombardia[1] and consists in the climatic series recorded between 1987 and 2003 in 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 needed instruments; this operation has 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 by a total of 5476 observations heterogeneously distributed in 14 different locations (Fig. 1d). Table 2 lists the final data available for each station distinguished by months. Looking into the details we may observe that May, July and August are the most critical months among data that globally reaches the amount of an incompleteness of 46 %. Data related to input variables 1a, 1b, 1c listed in Sect. 3 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) which was freely download from the geoportal of Regione Lombardia and has spatial resolution of 20 m. Aspect and slope were computed using the dedicated tools (spatial analyst tools) of the ArcGIS 10 (ArcEditor).

5 Radial basis function networks

In our study, we model snow cover thickness estimation as a neural learning task according to which correlation between climatic/geographic factors and snow cover

[1] In the geoportal related to the Lombardy region are available a lot of geospatial data (vector and raster) http://www.cartografia.regione.lombardia.it/geoportale.

thickness is inferred by induction from supervised input-output pairs of data. We adopt a radial basis function networks in both regression and classification task.

5.1 Radial basis function networks

RBFNs are characterised 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 case of one output node, the global non linear function computed by the network can be expressed as a linear combination of M basis functions associated to each hidden layer neuron. In formula we have

$$f(\mathbf{x}) = \sum_{i}^{M} w_j h_j(\mathbf{x}) \tag{1}$$

where $\mathbf{x} = [x_1, \ldots, x_k]^T$ is the K-dimensional input vector, w_j are the weighting coefficients of the linear combination and $h_j(\mathbf{x})$ 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 $\mathbf{x}$ and the centre of each function $\mathbf{c}_j = [c_{1j}, \ldots, c_{kj}]$

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

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

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

TCD
6, 2437–2475, 2012

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

The model configuration 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, $j = 1,...,M$ of basis functions associated with first level processing units.

The second phase, having model parameters M, $\boldsymbol{c}_j, j = 1,...,M$, $r_j, j = 1,...,M$ known, computes $w_i, i = 1,...,M$ minimising the difference between predicted output and truth by Least Mean Squares, computed through the pseudo inverse. In formula

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

where

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

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

5.2 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 investigate the behaviour of RBFN when coping with multidimensional function estimation modelled 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 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 that take part in the regression task. The set Ω is composed by 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. 3, 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.

To model the multidimensional function estimation as classification we identify intervals of the function co-domain and we put them in correspondence with predefined classes. The underlying assumption is that precision required in regression task is arbitrary due to incompleteness and or inconsistencies among data. During training, input pattern vectors are put in correspondence with a predefined class labels, exemplifying a hard mapping at a lower granularity with respect to regression, with mutually exclusive classes. The network is configured with an output layer having a number of neurons equal to the number of classes. The formal definition can be easily reconducted to the regression task given in Eq. (5) considering the truth value y_i as a label that describe the belongingness to each of the defined classes.

6 Experiments

In our experiments different evaluation indexes have been adopted. The agreement between truth and classification results has been analysed by means of the confusion matrix and derived accuracy indexes (Congalton, 1991), overall accuracy (OA), producer accuracy (PA) and user accuracy (UA), capturing the percent agreement between truth and classification results, is complemented with the Cohen's kappa coefficient thought to be a more robust measure that takes into account the agreement occurring by chance (Cohen, 1960). The root mean square error (RMSE) index and its normalised version NRMSE are used in combination with the mean absolute error (MAE) to measure the magnitude of network mistakes. For completeness is given the formal definition of these indices:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum(\hat{y}_i - y_i)^2} \tag{6}$$

$$\mathrm{NRMSE} = \frac{\mathrm{RMSE}}{\mathrm{MAX}(y_i)} \times 100 \tag{7}$$

$$\mathrm{MAE} = \frac{1}{n}\sum|\hat{y}_i - y_i| \tag{8}$$

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_Ω) and test (TeS_Ω), respectively.

The radial basis function network configured for the tasks described above has been applied to solve the problem of estimating the snow cover thickness. In the experiments, attention has been given to the parameters calibration process. A sensitivity analysis has been conducted varying the input parameters described in Sect. 5.1. For the training phase focused on the 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 K-means algorithm imposes a small number of centroids to limit the computational complexity, the random choice strategy has been preferred.

6.1 Regression results

First of all, we present the results obtained using the RBFN performing regression task. The RBFN receives in input the vector of measurements derived from the set of features (input variables) described in Sect. 3. 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 the 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 best behavior setting 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_{Ω} and TeS_{Ω}, randomly generated from the overall dataset Ω and averaging the individual indexes obtained. After the value $M = 500$ the error indexes show a lower decrease. We choose than this value as a reference for an optimize balancing between computationl cost, training accuracy and generalisation power. In order to conduct a deeper analysis on how the snow cover thickness has been medelled, in Fig. 3 we plot the estimated values versus truth values which are sorted in ascending order. Figure 5a, b shows the mean weekly error of the modelled 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 not a significant relationship between the measured liquid precipitation and the errors although there is a general increase of the errors in late fall and spring when there are the bigger differences at different altitude because precipitation becomes solid (snow) at higher elevation and therefore it is not recorded only from the AWS. Nevertheless some peaks of errors (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 in modelling the snow cover thickness estimation task as a classification problem in an attempt to reduce the precision of the output values for the benefit of the significance.

6.2 Classification results

The four classes were chosen considering their importance for permafrost (perennially frozen ground) stability. Indeed, where the snow is absent (class A) during the winter ground suffers the extremely low air temperature, while class B (1–10 cm) has been distinguished by the class C (11–90 cm) because in the former, in addition to the more limited insulation of the thinner snow with respect to the air temperature, also a certain percentage of radiation can reach the ground surface, while above 90 cm (class D) the insulation of the snow can be considered almost total.

For this task the network architecture is structured as followed: 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 task. Also in this case the RBFN showed best behaviour setting temporal window dimension T at 45 and the number of centroids M equal to 500. Results obtained with this configuration are showed in Table 4. The RBFN performs classification with a good level of accuracy, showing an OA equal to 85.6 % and a K coefficient equal to 79.5 %. Examining the details, the crucial class A, representing the absence of snow cover, has the highest performance expressed in term of both PA and UA, with values 95.05 % and 92.47 %, respectively. This result can be reconducted to the fact that features values are very representative for discriminating between "there is snow" or "no 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, the experiments reveal good performances. Tables 5–10 slices global results showed in Table 4 for different level of elevation. Inevitably results of certain classes show low statistic as consequence of 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 considered elevation, as already seen in Table 4.

6.3 Snow cover mapping

In order to exploit the potentialities of RBFN in estimating snow cover distribution, we pose as further interest in our research, the production of snow cover maps allowing to obtain a synoptic view of the phenomenon under investigation. This task has been addressed by proceeding in the spatialisation of input climatic variables and then by using the RBF network to compute for each input pattern including climate and geographic input variables, the corresponding predicted snow cover values. The elevation has been derived from the Digital Elevation Model as variable related to the air temperature. With reference to a generic cell xy, steps were taken to homogenize in terms of elevation the known values. Homogeneization was obtained by performing linear regression between elevation and temperature values. Setting a reference evaluation value, each known temperature value was shifted in function of the angular coefficient of the linear dependence law. Subsequently spatialisation was performed by applying the inverse square distance method obtaining temperature values for each grid unit. These values have been finally modified reporting them at the original elevation. The spatialisation of precipitation patterns is a critical aspect requiring domain dependent complex analysis. In our study we decided to adopt 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 refrain then from conducting a more sophisticated analysis that could be arbitrary in our context. Nine weeks were chosen based on their relevance for the cryospheric development and for their variability. For these reasons a higher frequency of examined weeks was setted for the end of the spring (during the melting period). For each of the nine weeks a snow map have been generated having as attribute for each grid element the neural computed snow cover thickness.

TCD
6, 2437–2475, 2012

An indirect validation procedure has been 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, neural computed snow cover thickness values have been binarized setting euristically the threshold to 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). Comparison results have been organized in a confusion matrix in which classes considered are presence and absence of snow. Results obtained from NDSI-based snow algorithm play the role of reference data. Table 11 reports comparison results in terms of OA (Overall Accuracy), UA (User Accuracy) and PA (Producer Accuracy) indexes.

The RBFN maps show in general an overestimation of 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 the beginning autumn when the snow distribution suffered more intensively of the site conditions and it is extremely dishomogeneous, is more probable that snow remains 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 good in general at low elevation (<1600 m); at higher elevation, above 1900 m s.l.m., the accuracy decreases lower than 70 % during the melting season and at the beginning of autumn, with a general underestimation. A different result has been achieved between 1600 and 1900 where poor results have been obtained only during the winter core for 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 of the week 11 March 2003. The overlap of this map, hardened with 5 cm. Threshold, with the corresponding

TCD
6, 2437–2475, 2012

NDSI map is shown in Fig. 7. The experts examined the maps in the light of topographic features of the study areas, and judged the results satisfactory.

7 Conclusions

An detailed investigation on the use of RBFN for snow cover thickness estimation has been conducted. The RBFN model copes very well with several levels of complexity. One originates from the incompleteness among data principally due to different temporal lengths and large gaps in time series. Additional source of complexity is constituted by weak description of input patterns that include only a minimal set of topographic and climatic variables. The RBFN model shows high flexibility to move from regression and hard classification tasks which are usually complementary to understand complex environmental phenomena at different level of precision. As seen in our experimental context, we may consider the snow cover thickness estimation by RBFN a valuable tool that working on a minimal set of input variables could be used in a wide range of situations. Moreover the dynamic character of the implemented model allows to easily simulate different scenario setting the input variables to specific values of interest and obtaining the estimated snow cover thickness.

Acknowledgements. Work partially supported by the Italian MIUR project PRIN 2008 "Permafrost e piccoli ghiacciai alpini come elementi chiave della gestione delle risorse idriche in relazione al Cambiamento Climatico".

References

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. 2440

Baraldi, A., Binaghi, E., Blonda, P., Brivio, P., and Rampini, A.: Comparison of the muli-layer perceptron with neuro-fuzzy techniques in the estimation of cover class mixture in remotely sensed data, IEEE T. Geosci. Remote, 39, 994–1004, 2001. 2439

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 and Francis, 2003. 2438

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. 2439

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. 2440

Bishop, C.: Neural Networks for Pattern Recognition, Oxford University Press, USA, 1995. 2439

Civco, D.: Artificial neural networks for land-cover classification and mapping, Int. J. Geogr. Inf. Sci., 7, 173–186, 1993. 2440

Cohen, J.: A coefficient of agreement for nominal scales, Educ. Psychol. Meas., 20, 37, 1960. 2448

Congalton, R.: A review of assessing the accuracy of classifications of remotely sensed data, Remote Sens. Environ., 37, 35–46, 1991. 2447

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, 2003. 2440

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. 2441

Dozier, J.: Spectral signature of alpine snow cover from the landsat thematic mapper, Remote Sens. Environ., 28, 9–22, 1989. 2452

Crane, R. G. and Anderson, M. R.: Satellite discrimination of snow/cloud surfaceses, Int. J. Remote Sens., 5, 213–223, 1984. 2441

Foody, G.: Land cover classification by an artificial neural network with ancillary information, Int. J. Geogr. Inf. Sci., 9, 527–542, 1995. 2440

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. 2440

TCD
6, 2437–2475, 2012

Snow cover thickness estimation by using radial basis function networks

A. Guidali et al.

Gong, P.: Integrated analysis of spatial data from multiple sources: using evidential reasoning and artificial neural network techniques for geological mapping, Photogramm. Eng. Rem. S., 62, 513–523, 1996. 2438, 2439

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: Proceedings of the 11th international conference on Intelligent data engineering and automated learning: IDEAL'10, Paisley, UK, Springer-595 Verlag, 110–117, 2010. 2441

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. 2440

Hantel, M. and Hirtl-Wielke, L. M.: Sensitivity of alpine snow cover to european temperature, Int. J. Climatol., 27, 1265–1275, 2007. 2452

Hush, D. and Horne, B.: Progress in supervised neural networks, IEEE Signal Proc. Mag., 10, 8–39, 1993. 2441

Jain, A., Mao, J., and Mohiuddin, K.: Artificial neural networks: a tutorial, Computer, 29, 31–44, 1996. 2439

Jayawardena, A., Fernando, D., and Zhou, M.: Comparison of multilayer perceptron and radial basis function networks as tools for flood forecasting, IAHS P. Int. Assoc. Hydrol. Sci., 239, 173–182, 1997. 2440

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. 2442

Kay, P. A. and Kutiel, H.: Some remarks on climate maps of precipitation, Clim. Res., 4, 233–241, 1994. 2451

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. 2440

Lees, B. G.: Neural network applications in the geosciences: an introduction, Comput. Geosci., 22, 955–957, 1996. 2439

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. 2442

Moody, J. and Darken, C.: Fast learning in networks of locally-tuned processing units, Neural Comput., 1, 281–294, 1989. 2440, 2441, 2445, 2446

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. 2441

Park, J. and Sandberg, I. W.: Universal approximation using radial-basis-function networks, Neural Comput., 3, 246–257, 1991. 2441

Philip, N. S. and Joseph, K. B.: A neural network tool for analyzing trends in rainfall, Comput. Geosci., 29, 215–223, 2003. 2440

Skidmore, A., Turner, B., Brinkhof, W., and Knowles, E.: Performance of a neural network: mapping forests using GIS and remotely sensed data, Photogramm. Eng. Rem. S., 63, 501–514, 1997. 2439

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, 2003. 2439

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. 2442

TCD

6, 2437–2475, 2012

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

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

TCD
6, 2437–2475, 2012

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

		Months												
Station	m a.s.l.	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Tot
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

TCD
6, 2437–2475, 2012

Table 3. Regression results varying the number of centroids M, expressed in term 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

TCD

6, 2437–2475, 2012

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.

	Reference data					
Class data	A	B	C	D	Tot U	UA
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

TCD
6, 2437–2475, 2012

Snow cover thickness estimation by using radial basis function networks

A. Guidali et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀I | ▶I
◀ | ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

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.

	Reference data					
Class data	A	B	C	D	Tot U	UA
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

TCD
6, 2437–2475, 2012

Snow cover thickness estimation by using radial basis function networks

A. Guidali et al.

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.

	Reference data					
Class data	A	B	C	D	Tot U	UA
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

TCD
6, 2437–2475, 2012

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.

	Reference data					
Class data	A	B	C	D	Tot U	UA
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

TCD
6, 2437–2475, 2012

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.

	Reference data					
Class data	A	B	C	D	Tot U	UA
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

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

TCD

6, 2437–2475, 2012

Snow cover thickness estimation by using radial basis function networks

A. Guidali et al.

Table 9. Confusion matrix for the radial basis function network classifier evaluated on the test set TeS_{Ω} 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.

	Reference data					
Class data	A	B	C	D	Tot U	UA
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

TCD
6, 2437–2475, 2012

Table 10. Confusion matrix for the radial basis function network classifier evaluated on the test set TeS_{Ω} 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.

	Reference data					
Class data	A	B	C	D	Tot U	UA
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

TCD
6, 2437–2475, 2012

Snow cover thickness estimation by using radial basis function networks

A. Guidali et al.

Table 11. Summary of the RBFN results.

Data	OA	Presence of snow PA	Presence of snow UA	Absence of snow PA	Absence of snow UA
30 Apr 2002	76.03 %	88.20 %	64.94 %	68.25 %	89.66 %
28 May 2002	64.03 %	78.36 %	39.49 %	59.41 %	89.04 %
22 Oct 2002	62.94 %	69.81 %	20.72 %	62.11 %	93.55 %
3 Dec 2002	62.15 %	62.73 %	71.60 %	61.68 %	51.80 %
11 Feb 2003	83.31 %	87.20 %	84.18 %	78.79 %	82.61 %
25 Feb 2003	81.96 %	88.35 %	79.20 %	75.72 %	86.13 %
11 Mar 2003	85.29 %	84.72 %	89.91 %	86.64 %	80.15 %
1 Apr 2003	82.10 %	86.67 %	78.99 %	78.19 %	86.11 %
29 Apr 2003	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 %

TCD
6, 2437–2475, 2012

Table 12. Overall accuracy for 6 elevation range.

Data	Overall accuracy					
	0–1000	1000–1300	1300–1600	1600–1900	1900–2200	>2200
30 Apr 2002	99.95 %	94.69 %	99.82 %	82.73 %	56.13 %	68.14 %
28 May 2002	99.98 %	98.98 %	99.99 %	91.41 %	47.42 %	42.61 %
22 Oct 2002	99.30 %	96.25 %	99.14 %	91.94 %	58.70 %	37.15 %
3 Dec 2002	70.60 %	51.78 %	50.82 %	55.14 %	74.42 %	61.33 %
11 Feb 2003	99.67 %	91.98 %	93.60 %	67.60 %	60.77 %	87.44 %
25 Feb 2003	99.98 %	96.32 %	98.79 %	72.89 %	46.92 %	86.17 %
11 Mar 2003	99.93 %	97.40 %	94.43 %	73.35 %	69.26 %	86.51 %
1 Apr 2003	99.78 %	99.80 %	99.48 %	84.02 %	65.42 %	77.13 %
29 Apr 2003	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 %

TCD
6, 2437–2475, 2012

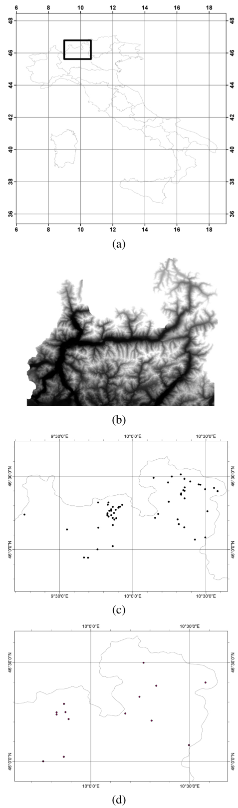


Fig. 1. In **(a)** the study area has been enclosed in a black rectangle. In **(b)** the Digital Elevation Model is depicted, related to the study area indicated in **(a)**. The locations of all stations of the initial data set equipped with different type of sensors are marked in **(c)**. The locations of the 14 stations equipped with air temperature sensor, snow thickness sensor and precipitation gauge, used for the final data set are represented in **(d)**.

TCD
6, 2437–2475, 2012

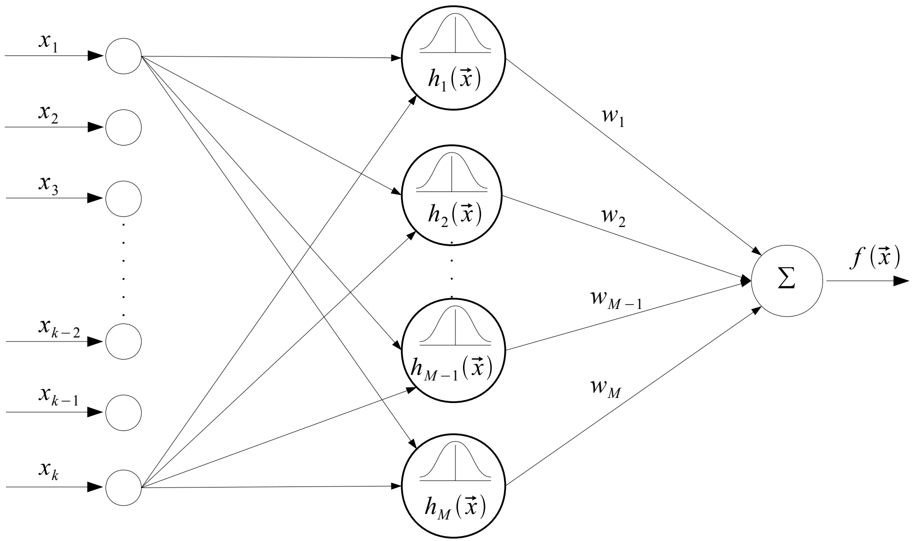


Fig. 2. Radial basis function network architecture for function approximation.

TCD
6, 2437–2475, 2012
Snow cover thickness estimation by using radial basis function networks
A. Guidali et al.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

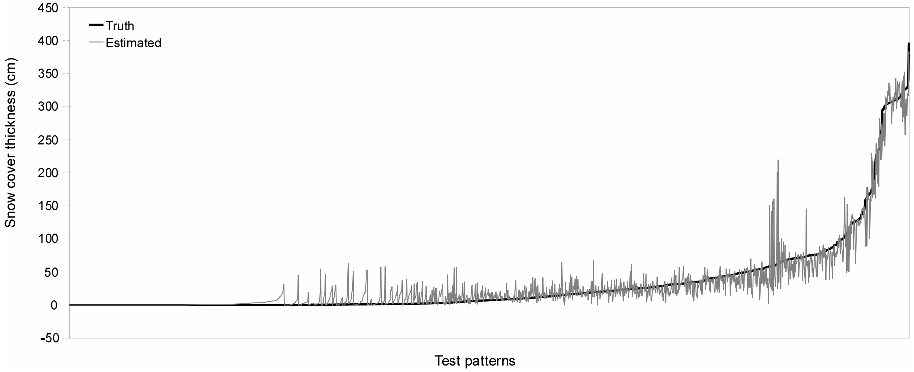


Fig. 3. Estimated snow cover thickness versus measured values in ascending order.

TCD
6, 2437–2475, 2012

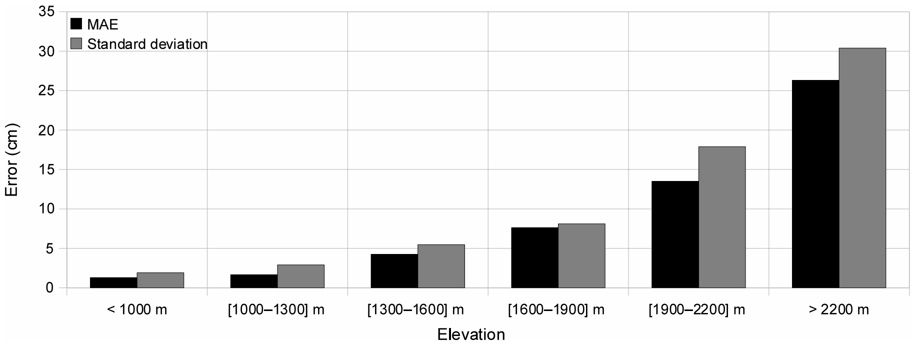
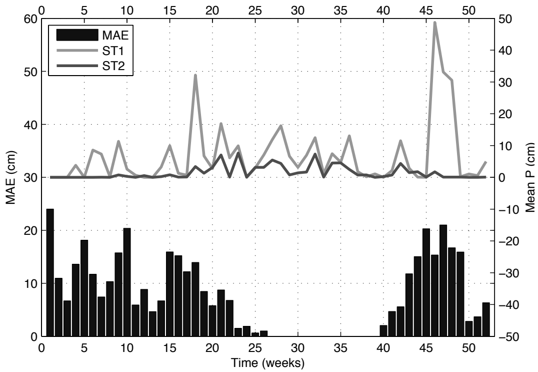
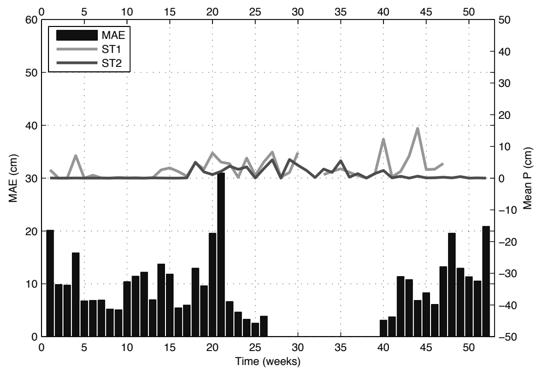


Fig. 4. Mean Absolute Error and standard deviation in regression task as a function of elevation ranges.

TCD
6, 2437–2475, 2012



(a)



(b)

Fig. 5. Mean Absolute Error (MAE) *n* 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.

TCD
6, 2437–2475, 2012

Snow cover thickness estimation by using radial basis function networks

A. Guidali et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀I | ▶I
◀ | ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

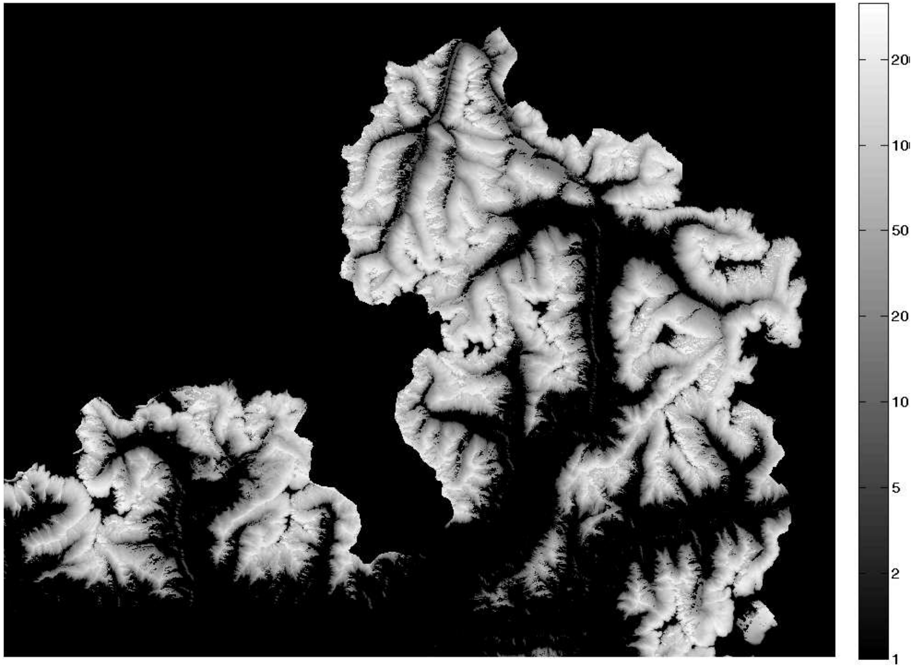


Fig. 6. Display on a logarithmic scale of the snow map produced by RBFN.

TCD

6, 2437–2475, 2012

Snow cover thickness estimation by using radial basis function networks

A. Guidali et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

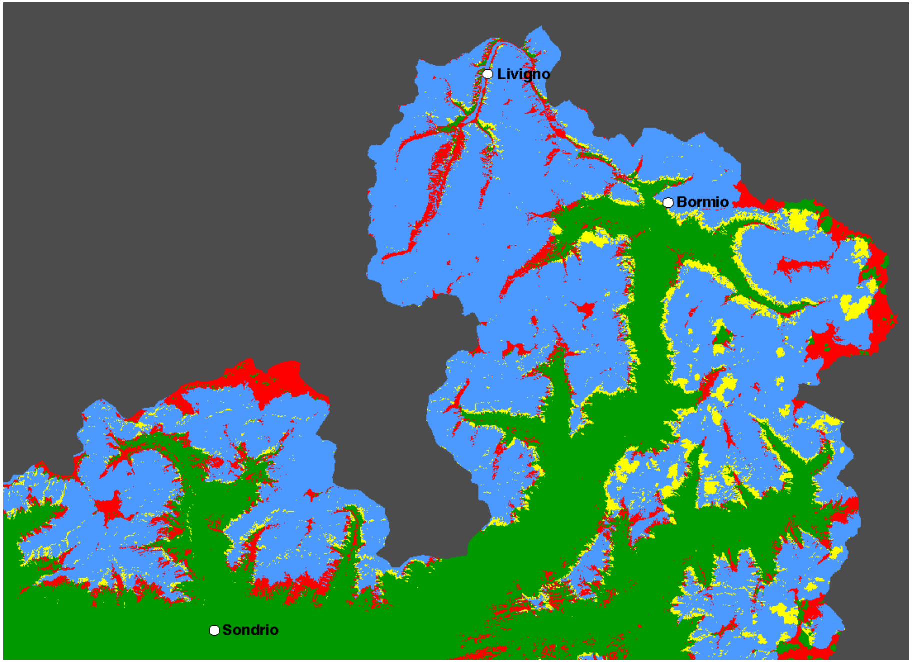


Fig. 7. Overlapping of maps produced by RBFN Network and NDSI. Considering as first result that produced by RBFN and as second that produced by NDSI, colors have the following meaning: blue = PS – PS; green = AS – AS; yellow = PS – AS; red = AS – PS, with PS = presence of snow and AS = absence of snow.

TCD
6, 2437–2475, 2012