Available online at www.sciencedirect.com
SCIENCE DIRECT®
Remote Sensing of Environment 90 (2004) 76–85
Remote Sensing of Environment
www.elsevier.com/locate/rse

Artificial neural network-based techniques for the retrieval of SWE and snow depth from SSM/I data

M. Tedesco[a,*], J. Pulliainen[b], M. Takala[b], M. Hallikainen[b], P. Pampaloni[a]

[a] *Institute of Applied Physics 'Carrara', CNR, Florence, Italy*
[b] *Laboratory of Space Technology, HUT, Espoo, Finland*

Received 17 July 2003; received in revised form 27 November 2003; accepted 6 December 2003

Abstract

The retrieval of snow water equivalent (SWE) and snow depth is performed by inverting Special Sensor Microwave Imager (SSM/I) brightness temperatures at 19 and 37 GHz using artificial neural network ANN-based techniques. The SSM/I used data, which consist of Pathfinder Daily EASE-Grid brightness temperatures, were supplied by the National Snow and Ice Data Centre (NSIDC). They were gathered during the period of time included between the beginning of 1996 and the end of 1999 all over Finland. A ground snow data set based on observations of the Finnish Environment Institute (SYKE) and the Finnish Meteorological Institute (FMI) was used to estimate the performances of the technique. The ANN results were confronted with those obtained using the spectral polarization difference (SPD) algorithm, the HUT model-based iterative inversion and the Chang algorithm, by comparing the RMSE, the R^2, and the regression coefficients. In general, it was observed that the results obtained through ANN-based technique are better than, or comparable to, those obtained through other approaches, when trained with simulated data. Performances were very good when the ANN were trained with experimental data.

Keywords: Snow; Microwave remote sensing; Artificial neural networks

1. Introduction

A major purpose of space-borne passive remote sensing of snow is the retrieval of snow depth and snow water equivalent (SWE), which are fundamental for hydrological, meteorological, and climatological applications as well as for discharge forecasting for hydropower production. In general, snow parameters can be extracted from remote sensing data through the inversion of theoretical equations or by employing empirical algorithms based on spatial or temporal correlation analysis. Several algorithms for the retrieval of SWE, such as those by Aschbacher (1989), Chang et al. (1987), reviewed and updated for forested areas by Foster et al. (1997), Hallikainen and Jolma (1992), and Tait (1998), are discussed in the literature. In recent times, further studies were also conducted by Goita et al. (2003), Goodison and Walker (1995), and Mognard and Josberger (2002).

Empirical formulas for the retrieval of SWE and snow depth were proposed by Chang et al. (1987), Aschbacher (1989), and Hallikainen and Jolma (1992). The proposed formulas are linear and easy to invert. However, they have the disadvantage of being closely related to the local conditions and scene characteristics. Moreover, techniques based on the inversion of theoretical or semi-empirical emission models through numerical techniques were suggested by Pulliainen and Hallikainen (2001) and Davis et al. (1993). Inversion of the theoretical models is related to the inversion of non-linear integro-differential equations and may lead to the so-called 'ill-posed' problems. In the latter, the same set of brightness temperatures may correspond to different sets of snow parameters, and the stability of the solution, should there be any, is not guaranteed.

The main purpose of this study was to develop and test an inversion technique for the retrieval of snow water equivalent and dry snow depth based on artificial neural networks (ANN) by using 19- and 37-GHz Special Sensor Microwave Imager (SSM/I) measured brightness temperatures. A comparison between performances of ANN-based technique and literature techniques (i.e. SPD algorithm, HUT model itera-

* Corresponding author. Viale Italia, 199/b, 83100 Avellino, Italy. Tel.: +39-55-4235-299; fax: +39-55-410893.
E-mail address: M.Tedesco@ifac.cnr.it (M. Tedesco).

0034-4257/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2003.12.002

Fig. 1. Test sites locations selected for the ground truth data.

tive inversion) was consequently performed. The method was tested over Finland in the period going from 1997 to 1999, using ground measurements performed on different test sites located in northern, central, and southern Finland.

The ANN training sets were generated by using either simulated or measured brightness temperatures. Simulated values were obtained by using the HUT snow emission model (Pulliainen et al., 1999). Experimental data were obtained from ground measurements. Once trained, the ANN were interrogated by employing the SSM/I brightness temperatures as input. Retrieval performances were, subsequently, compared with those obtained using the spectral polarization difference (SPD) (Aschbacher, 1989), the HUT model-based iterative inversion (Pulliainen et al., 1999) and the Chang algorithm (Chang et al., 1987) and its modification for forested areas (Foster et al., 1997). The paper is divided as follows: in the following section SSM/I data and test sites characteristics are reported, together with ground-measured data; in Section 3, we examine inversion approaches taken from the literature and artificial neural networks approach; Section 4 contains the results. Section 5 regards a discussion of the results and the conclusions.

2. SSM/I data and test sites

The use of the SSM/I radiometer, flying on board of the Defense Meteorological Satellite Program (DMSP) series satellites, is essential for snow monitoring on a global scale. Indeed, its functionality is not limited by the presence of clouds and observations are available on a daily basis. The SSM/I is a seven-channel, four-frequency (19-, 22-, 37- and 85-GHz) microwave radiometric system.

All channels operate in dual vertical (V) and horizontal (H) polarizations, except for the one at 22 GHz, which is fixed at V polarization. The footprint ranges from 69×43 km^2 at 19 GHz to 12.5×12.5 km^2 at 85 GHz (Hollinger, 1989, 1990). The SSM/I data collected for this study over the selected test sites (see Fig. 1) consist of Pathfinder Daily EASE-Grid (Equal Area SSM/I Earth Grid) brightness temperatures, supplied by the National Snow and Ice Data Centre (NSIDC) (Armstrong et al., 1994–1999). Data covered the period from the beginning of 1996 to the end of 1999.

Table 1 shows the coordinates of the test sites and their land-cover distribution. The SWE data set is based on national operational snow observations of the Finnish Environment Institute (SYKE). The SWE data were acquired by an operator through a line of measurement that was typically about 2 km long. The test sites were chosen so that snow measurement lines were inside the chosen SSM/I pixel. The values were determined by sampling and weighing the snow. The samples were taken from a cross section of the undisturbed snow pack. The spatial coverage was estimated to be sufficient to describe the SWE accurately enough over the EASE-Grid cell. Measurements were performed twice a month; for the remaining days, the data were modeled using a computer program by SYKE. The snow depth measurements were available only for the test site #1 from the website www.wunderground.com. They were performed daily by an automatic station. In order to assure dry snow conditions, we employed data concerning only the days in which the maximum temperature was lower than -5 °C and with the same condition registered over the 2 previous

Table 1
Coordinates of the test sites and land-cover distribution

Test site	Latitude	Longitude	Forest (%)	Agricultural (%)	Water (%)	Other (%)
1	67.3015	26.5651	94.2	0.6	0.2	5
2	66.9257	29.2488	90	2.2	1.1	6.7
3	65.0423	28.8754	87.9	1.4	3.2	7.5
4	64.9802	28.3755	90.8	1.8	3.8	3.6
6	64.7057	28.0297	91.3	2.3	4.4	2
7	64.5579	28.6682	92.1	1.6	3.3	3
8	64.1955	29.4354	92.3	1.3	2.9	3.5
9	62.4177	25.0772	89.6	3.4	4.1	2.9
10	62.2126	25.2363	88.4	5	5.4	1.2
11	60.5613	25.0357	74.8	18.1	5.1	2
12	60.4107	24.2136	78.5	10.1	9.7	1.7

Table 2
Number of samples employed for each test site

Test site	No. of samples
1	184
2	188
3	157
4	158
6	158
7	158
8	113
9	56
10	60
11	32
12	32

days. The number of samples selected for each test site is shown in Table 2.

As an example of the data used in this study, Fig. 2 shows the temporal behaviour of 19-, 37- and 85-GHz SSM/I measured brightness temperatures, the maximum and minimum air temperatures, and the SWE for test site #1 (Tedesco, 2002). It is interesting to observe that, as expected (Macelloni et al., 2001), when the SWE increased, the difference in polarizations was very strong at 19 GHz, reduced at 37 GHz, and negligible at 85 GHz.

3. Inversion procedures

In this section, the inversion procedures proposed are briefly introduced and examined. Moreover, the ANN-based technique procedures developed are presented and tested with the data at our disposal.

3.1. The SPD algorithm

The SPD algorithm was first suggested by Aschbacher (1989) for the retrieval of SWE when no information on the land-cover categories in the area of interest is available. The algorithm, which is based on a combination of SSM/I channels (Eq. (1)), led to Eqs. (2) and (3) for snow depth (Eq. (2)) and snow water equivalent (Eq. (3)). The coefficients appearing in the equations are the following: $A_0 = 0.68$, $A_1 = -0.67$ and $B_0 = 2.20$, $B_1 = -7.11$ for all data and $A_0 = 0.72$, $A_1 = -1.24$ and $B_0 = 2.02$, $B_1 = -7.42$ if T_{max} is lower than 0 °C. To analyse the validity of the values of coefficients used in the SWE formula more in detail, they were recomputed by using the data available in this study. The values obtained are the following: $B_0 = 3.17$, $B_1 = -16.64$ (RMSE = 41.4 mm) when considering all available data; $B_0 = 2.66$, $B_1 = -6.78$ (RMSE = 33.9 mm) if only SWE values lower than 170 mm are considered; $B_0 = 2.56$, $B_1 = -11.97$ (RMSE = 29.7 mm) if data without springtime measurements (middle of February, corresponding to a SWE of about 100–120 mm as shown in Fig. 3) are taken into consideration. Note that the computed coefficients become similar to those proposed in Aschbacher (1989) if only SWE values lower than 170 mm or measurements until the end–middle February are considered.

3.2. The HUT snow emission model-based iterative inversion algorithm

The HUT snow emission model is a single-layer model assuming that scattering of microwave radiation inside the medium is mostly concentrated in the forward direction. In

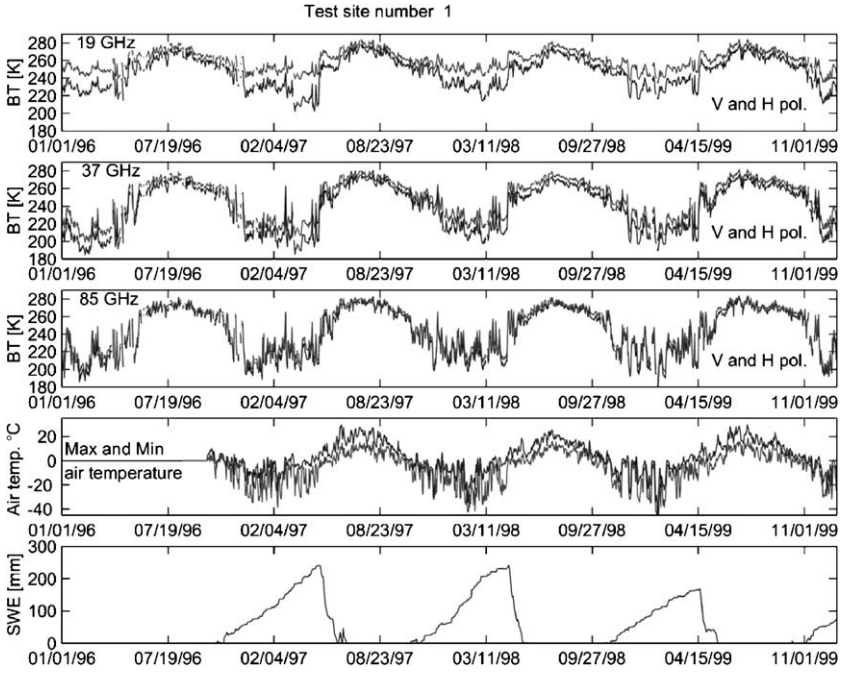

Fig. 2. Temporal behavior of 19-, 37- and 85-GHz SSM/I measured brightness temperatures, maximum and minimum air temperature and SWE for test site #1.

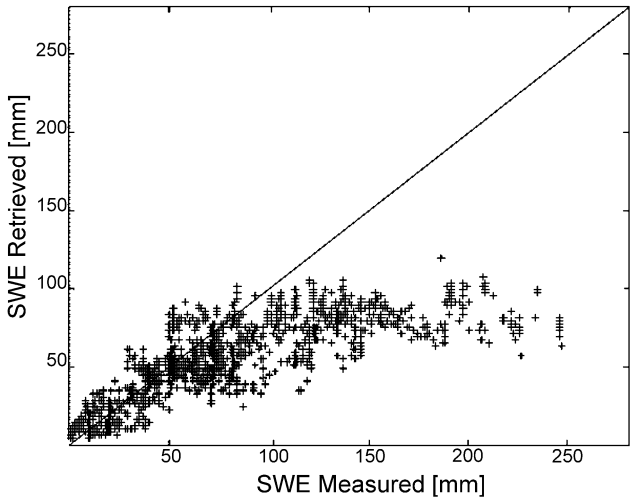

Fig. 3. Retrieved versus measured SWE with the SPD algorithm.

the model, the scattering coefficient is *weighted* by an empirical amplitude factor indicating the fraction of incident intensity scattered into the forward direction. The employed value of 0.96 was determined from multifrequency experimental data (Pulliainen et al., 1999). The brightness temperature is computed by solving the radiative transfer equation. As forest coverage is the main land-cover category for most test sites (Table 1), a boreal forest canopy model proposed by Kurvonen (1994) was used to take into account the influence of vegetation on the brightness temperature.

Atmospheric effects were disregarded and will be investigated in future works. However, in the case of methods based on experimental data, the atmospheric effects were directly taken into account by the data themselves.

The HUT iterative inversion algorithm is based on the minimization of a cost function containing measured and simulated values of the brightness temperature. Define $\boldsymbol{X}$ (Eq. (4)) as the vector containing the observed channel differences and $\boldsymbol{Y}$ (Eq. (5)) the corresponding model prediction. By assuming that the random error in brightness temperature predictions and the fluctuation of snow grain size are normally distributed, the conditional probability for SWE and D_0, given the observation $\boldsymbol{X}$, is supplied by Eq. (6), where $\langle D_0 \rangle$ is the expected value (a priori information) of snow grain size (diameter), σ is the standard deviation of the channel differences, and λ is the standard deviation of snow grain size. The cost function to be minimized for yielding the maximum likelihood estimate for SWE and D_0 is given in Eq. (7).

3.3. The Chang algorithm and its modification for forested areas

As already said, Chang et al. (1987) proposed an algorithm based on the combination of 18 and 36 GHz brightness temperatures at H polarization for the retrieval of snow depth (Eq. (8)). In this study, we used the equation proposed in Chang et al. (1987) by replacing the 18- and 36-GHz with the 19- and 37-GHz SSM/I channels. A revisited algorithm, developed from the original Chang algorithm, was proposed by Foster et al. (1997) in order to consider also the forest cover fraction. In the modified Chang algorithm, the original factor of 1.59 (see Eq. (8)) is replaced by the factor 0.78 for Eurasia, and the forest cover fraction is taken into account by dividing the a factor by $(1-f)$, where f is the forest cover fraction (Eq. (9)). A fixed value of density (0.27 g/cm^2) was employed to calculate the SWE from snow depth, as already suggested by Hallikainen and Jolma (1992).

3.4. Artificial neural networks technique

Multilayer perceptrons (MLP) may have one or more hidden layers of neurons between the input and output layers. They have a simple layer structure in which successive layers of neurons are fully interconnected, with connection weights controlling the strength of the connections. The input to each neuron in the next layer is the sum of all its incoming connection weights multiplied by their connecting input neural activation value. The trainable offset value associated with the neuron is added to the sum, and the result is fed into the function of the neuron (activation function). The latter can have many forms; the most common is the non-linear sigmoid function, which is the one used in this study. Other functions, such as the simple linear activation, threshold activation, and hyperbolic tangent activation can also be used. Activation functions used at the output of each neuron typically yield values in the $[-0.5, +0.5]$ range. Since the output units of the mapping network must generally produce an estimate of a parameter with an arbitrary range (i.e. not limited to $[-0.5, +0.5]$), the restriction on the output range must be removed. This can be accomplished by scaling the inputs and outputs of the network. In our case, the training phase of the ANN was based on the back-propagation (BP) learning rule to minimize the mean square error (MSE) between the desired target vectors and the actual output vectors. Training patterns were sequentially presented to the network, and the weights of each neuron were adjusted so that the approximation created by the neural network minimized the global error between the desired output and the added output created by the network. The trained neural network can be thought of as a type of non-linear, least mean square interpolation formula for the discrete set of data points in the training set: *multilayer feedforward networks are a class of universal approximators* (Hornik et al., 1989). The accuracy of the approximation depends mainly on training data. The training phase ends either when a fixed MSE error or a maximum number of iterations is reached, or when a stop is given by the *early stopping* technique (Caruana et al., 2000). In the latter, the training is interrupted when a desired mean square error on the training set is reached or when the error on a validation set increases, even when the error on

Table 3
Regression coefficients, R^2, and RMSE errors for nets locally trained over all test sites for the ANN, SPD, HUT iterative snow emission inversion, and Chang algorithm for the retrieval of SWE

Test site	ANN			SPD			Iterative			Chang		
	RMSE (mm)	R^2	r	RMSE (mm)	R^2	r	RMSE (mm)	R^2	r	RMSE (mm)	R^2	r
1	22.96	38.78	0.970	35.56	46.8	0.715	41.55	7.84	0.694	37.45	21.85	0.403
2	21.20	29.02	0.685	50.85	39.28	0.479	59.64	8.83	0.525	30.86	16.93	0.214
3	20.38	51.39	0.824	27.3	74.37	0.595	21.73	33.94	1.20	30.7	32.33	0.385
4	21.75	47.10	0.947	25.06	67.52	0.664	37.63	27.97	1.02	29.7	32.25	0.466
6	22.21	47.55	0.715	34.27	63.68	0.506	47.12	32.81	0.944	30.9	32.02	0.375
7	21.65	45.32	0.874	23.9	70.52	0.604	35.40	36.40	0.919	28.1	35.6	0.473
8	24.17	40.18	0.970	21.09	71.88	0.669	27.09	35.59	1.24	30.3	35.02	0.554
9	14.99	62.80	0.950	19.88	65.95	0.545	21.47	57.13	1.16	18.7	19.5	0.575
10	11.36	75.91	1.201	11.8	79.41	0.691	14.58	71.00	0.865	20.9	38.33	0.596
11	15.64	41.43	1.247	17.15	48.51	0.834	18.31	42.19	0.816	19.7	51.19	0.827
12	12.03	69.89	1.201	9.61	85.38	0.797	13.56	70.91	1.138	11.9	76.55	0.975
All	24.10	44.89(ave.)	0.855	25.13	65.07	0.591	30.72	38.54(ave.)	0.957(ave.)	26.29	35.59	0.531

the training set is still decreasing. This technique is very useful when dealing with ANN trained with values obtained from theoretical models. Indeed, in this case, it can be very useful to adopt a small subset of measured data to generate validation sets. In this way, the network is trained over a wide range of values (due to the possibility of generating the training set with the model) but is not *over-trained* for realistic values of the sought quantity (because of the early-stopping technique applied to the validation set made up of measured values).

In our study, two different approaches were used for the generation of the training sets. In the first one, brightness temperatures simulated by means of the HUT snow emission model were employed. The second approach made use of a subset of measured values. The aim of the two different approaches aimed at testing the capabilities of an artificial neural network to retrieve snow parameters when (1) trained simulated data, validated with measurements sampled in space or time, and (2) trained with the only measured brightness temperatures and the corresponding ground measured data.

The input layer of the ANN consisted of four neurons, made up of 19 and 37 GHz vertical and horizontal brightness temperatures, where the output was the snow parameter to be retrieved. One hidden layer was employed with a number of neurons depending case by case on the parameter to be retrieved. The learning rule employed a function updating weight and bias values according to Levenberg-Marquardt optimisation (Levenberg, 1944; Marquard, 1963). Several different networks were trained for each test site. In order to define a criterion for the best trained network, the following statistical parameters were used: root mean square error (RMSE, Eq. (10)), R^2 (Eq. (11)), and linear regression coefficient r (Eq. (12)), with y_{pred} as the predicted value, y_{true} the measured value, $\bar{y}_{\text{pred}}$, and $\bar{y}_{\text{true}}$, respectively, the mean of y_{pred} and y_{true} and N the number of samples. The one showing the lowest RMSE, the highest R^2, and linear regression coefficient r closest to 1 on the validation set was selected as best trained network. In those cases in which a validation set was not used (i.e. when the training set was made of a subset of experimental data), the net showing the lowest mean square error on the training set was selected.

4. Retrieval of dry snow parameters

The results obtained in regard to the retrieval of the snow parameters using the techniques discussed in Section 3 are reported here as follows, together with performances of the different techniques.

4.1. The SPD algorithm

Fig. 3 shows the measured versus SPD-retrieved SWE for all test sites (Eq. (3)). As expected (Armstrong et al., 1993; De Seve et al., 1997), the SPD algorithm worked well

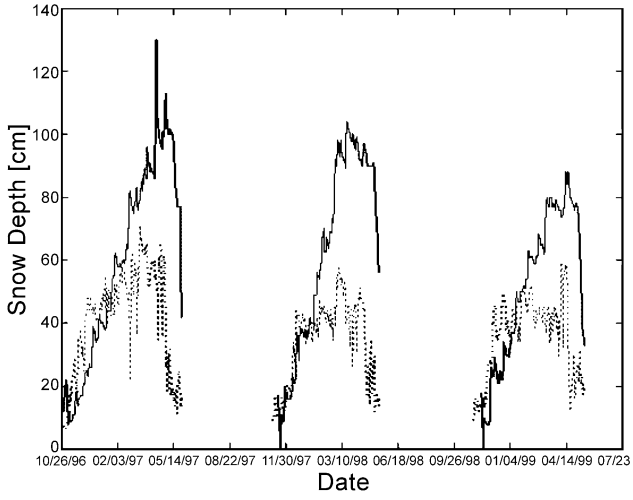

Fig. 4. Temporal behaviour of snow depth at Sodankylä (test site 1) measured (continuous line) and retrieved with the SPD algorithm (dashed line).

Table 4
'A priori' mean values of snow grain size and standard deviation of snow grain size used in the HUT iterative inversion technique for each test site

Test site	Mean	S.D.
1	1.7	1.6
2	1.6	1.5
3	1.4	1.3
4	1.4	1.3
6	1.3	1.2
7	1.4	1.3
8	1.3	1.2
9	1.3	1.2
10	1.5	1.4
11	1.6	1.5
12	1.4	1.3

for low SWE values (100 mm), whereas the error considerably increased for higher values. The RMSE, R^2, and linear regression coefficients were, respectively, RMSE = 32.27 mm, $R^2 = 57.32\%$, and $r = 0.577$. The results are summarized in Table 3. In addition, Fig. 4 shows the behaviour of snow depth at Sodankylä (test site #1) as measured (continuous line) and retrieved (dashed line) with the SPD algorithm (Eq. (2)). In this case, RMSE = 18.34 cm, $R^2 = 16.31\%$, and $r = 0.558$. The algorithm performances became worse when dealing with very deep snow.

4.2. The HUT snow emission model-based iterative inversion algorithm

The HUT snow emission model iterative inversion algorithm showed good results for low SWE values. The method has the advantage that the grain radius is not a fixed parameter, as in other inversion procedures, but it can slightly change. The fractional volume was fixed to $f = 0.3$. The expected values of snow depth and mean grain radius were given as inputs to the model as well as changes from the expected values, according to the normal distribution for finding the minimization of the function C (Eq. (7)). The average and standard deviation values employed for the grain size (diameter) were set for each test site and ranged from $\langle D_0 \rangle = 1.7$ mm and $\sigma = 1.6$ mm for the northern test sites to $\langle D_0 \rangle = 1.3$ mm and $\sigma = 1.2$ mm for the central and southern test sites. Table 4 reports the values of 'a priori' mean value of snow grain size and standard deviation of snow grain size for each test site. The results obtained with this method show a RMSE = 35.40 mm, $R^2 = 36.4\%$, and $r = 0.919$. The performances of the method can be improved by selecting different grain size values for each year or season. The results obtained are summarized in Table 3.

Table 5
Regression coefficients and RMSE obtained with the Chang et. al. algorithm (1987) and its modification (Foster et al., 1997) for different test sites and forest cover fraction

Test site	Forest (%)	Chang (1987)		Foster (1997)	
		r	RMSE	r	RMSE
11	74.8	0.82	19.7	1.6	38.4
12	78.5	0.9	11.9	2.05	27.16
3	87.9	0.38	30.7	1.56	122.5
10	88.4	0.59	20.9	2.5	82.3
9	89.6	0.57	18.7	2.68	88.2
2	90	0.21	30.86	1.07	154.3
4	90.8	0.46	29.7	2.47	158.4
6	91.3	0.37	30.9	2.07	172.2
7	92.1	0.47	28.1	2.95	174.9
8	92.3	0.55	30.3	3.55	193.3
1	94.2	0.40	37.45	3.48	322.8

Table 6
HUT snow emission model inputs range

Parameter	Min	Max
Snow depth (m)	0.05	1.45
Mean snow grain size (mm)	1.3	1.9
Snow temperature (C)	− 50	0
Snow fractional volume	0.15	0.35
Real part of ground permittivity	4	7
Imaginary part of ground permittivity	0	2
Forest coverage (m^3/ha)	0	200

4.3. The Chang algorithm and its modification for forested areas

The SWE was computed by setting the snow density to a fixed value of 0.27 g/cm^2, as already suggested by Hallikainen and Jolma (1992). Table 5 shows the regression coefficients and RMSE obtained by applying the Chang algorithm (Chang et al., 1987), along with its modification for forested areas in Eurasia (Foster et al., 1997). The results obtained showed that the original Chang algorithm (Chang et al., 1987) worked better for those areas where high cover fractions of forest are present. The modified algorithm (Foster et al., 1997) tended to get better as the forest cover fraction decreased. Further investigations will be necessary in order to improve the performances of the latter algorithm in very dense forested areas.

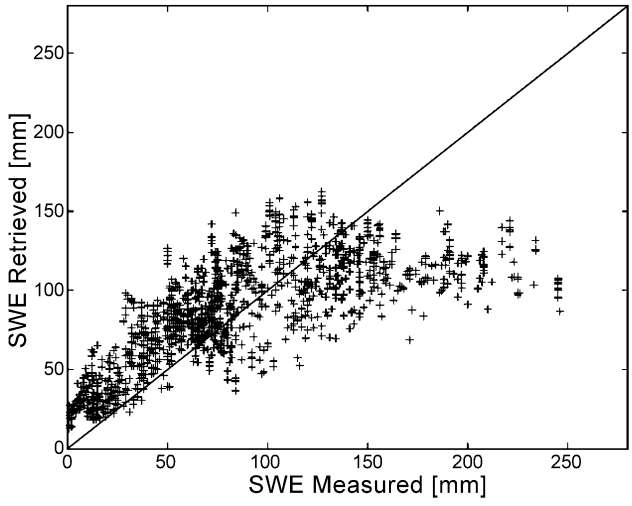

Fig. 5. Measured versus retrieved SWE with a single ANN trained over all test sites.

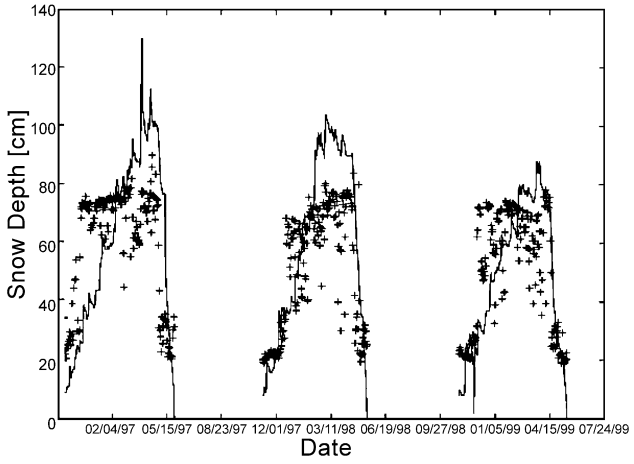

Fig. 6. Measured (line) and retrieved (crosses) snow depth with net trained with HUT model as a function of time for Sodankylä test site.

4.4. Artificial neural networks trained with the HUT snow emission model

In the case of ANN trained with the HUT model, the simulated brightness temperatures and the snow input parameters were used, respectively, as inputs and outputs of the network (i.e. snow parameters employed as input of the model become output parameters of the net, whereas brightness temperatures, being the output of the model, became the input of the ANN). In order to generate the training set, snow parameters such as snow depth, mean snow grain size, snow temperature, density, ground permittivity, and forest coverage ranged from a minimum to a maximum value (Table 6).

The retrieval of SWE was performed by training a single network for each test site (denoted as the *local* approach) or by training only one network for all test sites (denoted as the *generalized* approach). Note that, even if the selected training set was the same for all the test sites, the differently trained ANN showed different performances because their validation sets were different. In the case of the *local* approach, a subset consisted of one quarter of the measured data over the selected test site was employed as a validation set, whereas in the *generalized* case, the validation set was made up of one quarter of all selected ground data over all test sites. Once trained, the nets were interrogated over each single test site by employing both the *local* and the *general* approaches. The retrieval of SWE showed valid results for

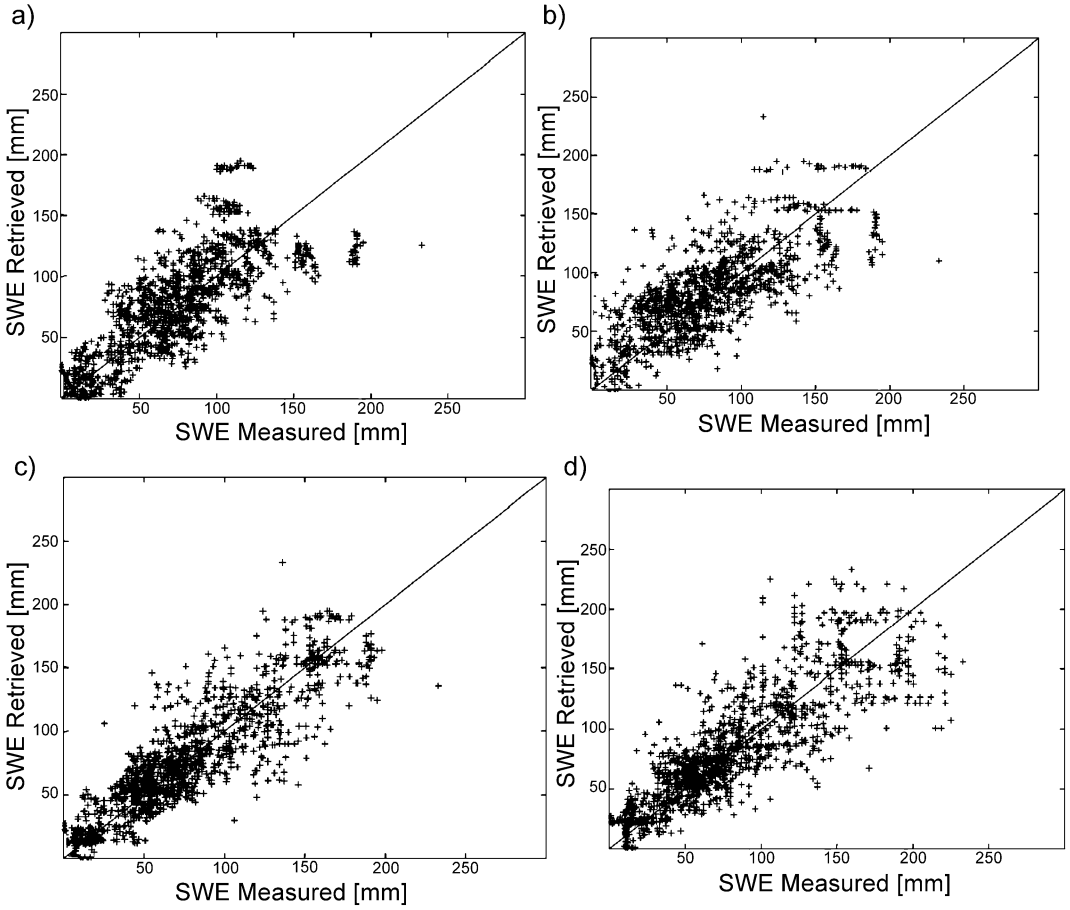

Fig. 7. Measured SWE versus the retrieved SWE employing, respectively, the nets ANN N/S (a), ANN N (b), ANN S (c), ANN N/C/S (d) over the test sites from #3 to #10.

values up to 150 mm, while, for higher values, results did not appear to be as accurate. This was also true in the case of the SPD algorithm (Fig. 5). Among the reasons for the increasing error with SWE, there was the use of the HUT snow emission model for generating the training set. As a matter of fact, being the latter model a single-layer model, it is unable to take into account the snow-pack stratification, strongly influencing microwave snow emission.

The procedure for retrieving the snow depth from the ANN trained with the HUT model simulated values was the same as the one adopted for the SWE. Fig. 6 shows measured (line) and retrieved (crosses) snow depth as a function of time. The statistical parameters were the following: $r = 0.85$, $R^2 = 47.48\%$, and RMSE = 14.64 cm. The results obtained are summarized in Table 3.

4.5. Artificial neural networks trained with experimental data

In the case of ANN trained with experimental data, the input set consisted of SSM/I measured brightness temperatures. The output set was made up of the snow parameters measured corresponding to the used radiometric data. Contrary to what was done in the case of ANN trained with the HUT model, in this case, the early stopping technique was not employed: the training phase stopped when a fixed mean square error was reached or a maximum number of iterations was performed. The ANN trained with the only experimental data can be seen as a sort of sigmoid interpolator whose weights are calculated during the training phase.

The retrieval of SWE was performed by generating the training sets with values of SWE at different locations in northern, southern and central Finland (test sites #1 and #2 for the north, test sites #11 and #12 for the south, and site #7 for the centre). In particular, test sites #1 and #2 were chosen as representative of northern data, test sites from #11 to #13 of southern data, and #7 of centre data. The nets trained with northern and southern data are denoted with ANN N/S, those with only north or south data, respectively, with ANN N and ANN S and those containing all data from the north, south and centre with ANN N/C/S. Test sites not employed for training were used to test the trained networks. Fig. 7 shows the measured SWE versus the retrieved SWE with the employment, respectively, of the ANN N (a), ANN N/C/S (b), ANN N/S (c), ANN S (d) over the test sites from #3 to #10 (with the exclusion of #7, which was used as a training set). Table 7 shows the RMSE, R^2, and r for the four different trained nets. The obtained results show that, in general, best performances were obtained with the ANN S with $R^2 = 80.44\%$ and RMSE = 19.53 mm.

Table 7
RMSE, R^2, and regression coefficients for the four different trained nets

Test sites 3–10	RMSE (mm)	R^2	r
ANN N/S	27.52	61.18	0.835
ANN N	29.65	54.95	0.821
ANN S	19.53	80.44	0.938
ANN N/C/S	24.02	74.15	0.914

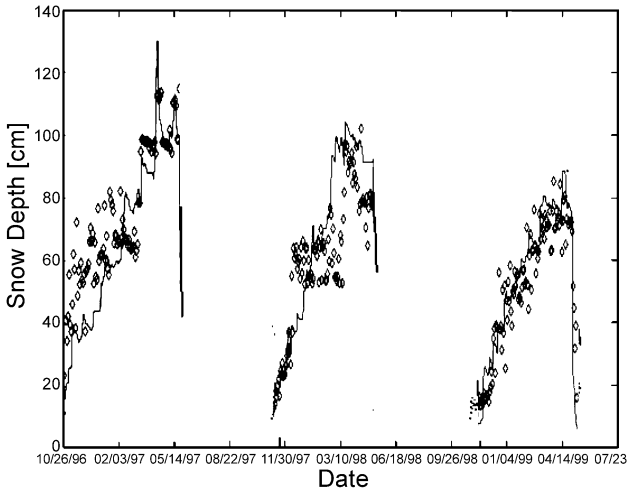

Fig. 8. Snow depth measured (line) and retrieved with the ANN trained with experimental data (diamonds) as a function of time for the test site #1.

The procedure followed for the retrieval of snow depth over test site #1 was the same as that one regarding the SWE. Fig. 8 shows the temporal trend of measured (line) and retrieved (diamonds) snow depth. The statistical coefficients are: RMSE = 14.28 cm, $R^2 = 72.98\%$, and $r = 0.975$. Even if the RMSE obtained in this case was very close to the one obtained with the ANN trained with the HUT model, we can say that performances undoubtedly improved when the R^2 and r coefficients were also compared.

5. Discussion and conclusions

Through this study, observations regarding artificial neural networks based techniques for retrieving snow water equivalent and snow depth from passive microwave spaceborne were presented and tested. Performances of ANN-based techniques were compared with those obtained using the SPD algorithm (Aschbacher, 1989), the iterative inversion of HUT snow emission model (Pulliainen et al., 1999), and linear regressions (Chang et al., 1987; Foster et al., 1997). Typically, the ANN trained with simulated data gave better or comparable results in comparison with the other examined approaches. The best performances were obtained with the ANN trained with experimental data. Table 8 shows a comparison of the performances of the different methods.

The retrieval of SWE and snow depth either with SPD algorithm or with the iterative inversion of HUT snow emission model gave good results when SWE values were lower than 170 mm, or when springtime measurements were disregarded. In particular, for highly stratified snow-packs, the error in snow depth retrieval increased (in the HUT model, the SWE was computed by considering the fractional volume as a fixed parameter), due to the fact that the model is a single-layer model. Also, the formula proposed

Table 8
Comparison of the results of the different approaches

Applied technique	RMSE	R^2	r
SWE			
SPD (all)	32.27 mm	57.32%	0.577
HUT iterative inversion (averaged)	30.72 mm	38.54%	0.957
ANN trained with model	24.1 mm	44.89%	0.85
ANN trained with experimental data	19.53 mm	80.44%	0.938
Depth			
SPD	18.34 cm	16.31%	0.558
ANN trained with model	14.64 cm	47.78%	0.854
Chang	18.97 cm	12.38%	0.502
ANN trained with experimental data (test site #1)	14.28 cm	72.98%	0.975

by Chang et al. (1987) for the retrieval of snow depth gave best results by taking into account only measurements performed until mid-February. The modification in the Chang algorithm for forested areas (Foster et al., 1997) showed best performances for those areas where the forest coverage was not very high.

Artificial neural networks trained with the HUT model were found to be an efficient tool both for the retrieval of SWE and the snow depth. The ANN trained with experimental data showed the best performances, for both the retrieved parameters, with RMSE and R^2 respectively equal to 19.53 mm and 80% in the case of SWE, and 14.28 cm and 72.98% in the case of snow depth. This result is encouraging for the practical application of these networks.

To summarize, the results obtained were very satisfactory showing that *artificial neural networks are able to retrieve the spatial or temporal variations of the unknown parameter from SSM/I data, especially if trained with data collected on test sites*. The technique developed in this study can be very useful when periodical ground measurements are collected in few stations only, and no information is available from any of the areas in between. Furthermore, the possibility to interrogate the trained network, without running the inversion of equations or formulas, makes this technique suitable for close to real-time applications.

Future studies will include application and testing in other areas, in order to validate the method in areas outside Finland. The operational and efficiency implications of training individual ANN for all sites, versus one ANN that can be applied regionally will also be taken into account. In addition, a preliminary operational algorithm which combines unsupervised classification and supervised retrieval technique is already under study.

Acknowledgements

This work was supported by the Centre for International Mobility (CIMO), the Laboratory of Space Technology of the Helsinki University of Technology, the Institute of Applied Physics "Carrara" of the Italian National Research Council, and the Italian Space Agency (ASI). It was accomplished within the framework of the ENVISNOW European Project EVG2-2001-00018.

Appendix A. Equations

$$\mathrm{SPD} = [\mathrm{Tb}(19V) - \mathrm{Tb}(37V)] + [\mathrm{Tb}(19V) - \mathrm{Tb}(19H)] \tag{1}$$

$$d_s = A_0 * \mathrm{SPD} - A_1 \ [\mathrm{cm}] \tag{2}$$

$$\mathrm{SWE} = B_0 * \mathrm{SPD} - B_1 \ [\mathrm{mm}] \tag{3}$$

$$\underline{X} = [T_{19V\ \mathrm{SSM/I}} - T_{37V\ \mathrm{SSM/I}} T_{19V\ \mathrm{SSM/I}} - T_{19H\ \mathrm{SSM/I}}]^T = [X_1 X_2]^T \tag{4}$$

$$\underline{Y} = [T_{19V\ \mathrm{Model}} - T_{37V\ \mathrm{Model}} T_{19V\ \mathrm{Model}} - T_{19H\ \mathrm{Model}}]^T = [Y_1 Y_2]^T \tag{5}$$

$$\rho(\mathrm{SWE}, D_0 \mid X) \propto \frac{1}{(2\pi)^{3/2}\sigma^2\lambda} \times \exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{2}\left(\frac{X_i - Y_i(\mathrm{SWE}, D_0)}{\sigma}\right)^2 + \left(\frac{d - < D_0 >^2}{\lambda}\right)\right]\right\} \tag{6}$$

$$C(\mathrm{SWE}, D_0) = \frac{1}{2\sigma^2}\{[Y_1(\mathrm{SWE}, D_0) - X_1]^2 + [Y_2(\mathrm{SWE}, D_0) - X_2]^2\} + \frac{1}{2\lambda^2}(D_0 - < D_0 >)^2 \tag{7}$$

$$d_s = 1.59 * (18H - 36H) \tag{8}$$

$$d_s = \frac{0.74 * (18H - 36H)}{1 - f} \tag{9}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(y(i)_{\mathrm{pred}} - \bar{y})^2} \tag{10}$$

$$R^2 = \frac{\sum_{i=1}^{N}(y_{\mathrm{pred}}(i) - \bar{y}_{\mathrm{pred}}) \sum_{i=1}^{N}(y_{\mathrm{true}}(i) - \bar{y}_{\mathrm{true}})}{\left(\sum_{i=1}^{N}(y_{\mathrm{true}}(i) - \bar{y}_{\mathrm{true}})\right)^2} \cdot 100 \tag{11}$$

$$r = \frac{\sum_{i=1}^{N}((y_{\mathrm{true}}(i) - \bar{y}_{\mathrm{true}})(y_{\mathrm{pred}}(i) - \bar{y}_{\mathrm{pred}}))}{\sum_{i=1}^{N}(y_{\mathrm{true}}(i) - \bar{y}_{\mathrm{true}})^2} \tag{12}$$

References

Armstrong, R., Chang, A., Rango, A., & Josberger, E. (1993). Snow depths and grain-size relationships with relevance for passive microwave studies. *Annals of Glaciology*, *17*, 171–176.

Armstrong, R. L., Knowles, K. W., Brodzik, M. J., & Hardman, M. A. (1994–1999). DMSP SSM/I Pathfinder daily EASE-Grid brightness temperatures. *Digital media and CDROM*. Boulder, CO: National Snow and Ice Data Center.

Aschbacher, J. (1989). Land surface studies and atmospheric effects by satellite microwave radiometry. *PhD thesis dissertation*, University of Innsbruck.

Caruana, R., Lawrence, S., & Giles, L. (2000). *Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping*. Denver, CO: Neural Information Processing Systems.

Chang, A. T. C., Foster, J. L., & Hall, D. K. (1987). Nimbus 7 SM derived global snow cover patterns. *Annals of Glaciology*, *9*, 39–44.

Davis, D. T., Chen, Z., Tsang, L., Hwang, J. N., & Chang, A. T. C. (1993). Retrieval of snow parameters by iterative inversion of a neural network. *IEEE Transactions on Geoscience and Remote Sensing*, *31*, 842–851.

De Seve, D., Bernier, M., Fortin, J. -P., & Walker, A. (1997). Preliminary analysis of snow microwave radiometry using the SSM/I passive-microwave data: The case of La Grande River watershed (Quebec). *Annals of Glaciology*, *25*, 353–361.

Foster, J. L., Chang, J. A., & Hall, D. (1997). Comparison of snow mass estimates from a prototype passive microwave snow algorithm, a revised algorithm and a snow depth climatology. *Remote Sensing of Environment*, *62*, 132–142.

Goita, K., Walker, A., & Goodison, B. (2003). Algorithm development for the estimation of snow water equivalent in the boreal forest using passive microwave data. *International Journal of Remote Sensing*, *24*(5), 1097–1102.

Goodison, B., & Walker, A. (1995). Canadian development and use of snow cover information from passive microwave satellite data. In B. Choudhury, Y. Kerr, E. Njoku, & P. Pampaloni (Eds.), *Passive microwave remote sensing of land–atmosphere interactions* (pp. 245–262). Utrecht, Netherlands: VSP.

Hallikainen, M., & Jolma, P. (1992). Comparison of algorithms for the retrieval of snow water equivalent from NIMBUS-7 SMMR data in Finland. *IEEE Transactions on Geoscience and Remote Sensing*, *30*, 124–131.

Hollinger, J. (1989). *DMSP special sensor microwave/imager calibration/validation report*, *vol. I*. Washington, DC: Naval Research Labs.

Hollinger, J. (1990). *DMSP special sensor microwave/imager calibration/validation report*, *vol. II*. Washington, DC: Naval Research Labs.

Hornik, K., Stincombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*, 359–366.

Kurvonen, L. (1994). Radiometer measurements of snow in Sodankyla 1991–93. *Report*, (16,. Espoo, Finland: Helsinki University of Technology, Laboratory of Space Technology.

Levenberg, K. (1944). Method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, *2*, 164–168.

Macelloni, G., Paloscia, S., Pampaloni, P., & Tedesco, M. (2001). Microwave emission from dry snow: A comparison of experimental and model results. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(12), 2649–2656.

Marquard, D. W. (1963). An algorithm for least squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, *11*, 431–441.

Mognard, N., & Josberger, E. (2002). Seasonal evolution of snowpack parameters, northern Great Plains. *Annals of Glaciology*, *34*, 15–23.

Pulliainen, J., & Hallikainen, M. (2001). Retrieval of regional snow water equivalent from space-borne passive microwave observations. *Remote Sensing of Environment*, *75*, 76–85.

Pulliainen, J. T., Grandell, J., & Hallikainen, M. (1999). HUT snow emission model and its applicability to snow water equivalent retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, *37*, 1378–1390.

Tait, A. (1998). Estimation of snow water equivalent using passive microwave radiation data. *Remote Sensing of Environment*, *64*, 286–291.

Tedesco, M. (2002). Dry snow mapping in finland employing space-borne passive microwave observations and artificial neural networks. *Report*, (*vol. 50*,. Espoo, Finland: Helsinki University of Technology, Laboratory of Space Technology.