Meteorol. Appl. 11, 83–91 (2004) DOI:10.1017/S1350482704001173

Neural networks in satellite rainfall estimation

F. J. Tapiador[1], C. Kidd[1], K.-L. Hsu[2] & F. Marzano[3]
[1]*School of Geography, Earth and Environmental Sciences, University of Birmingham, UK*
[2]*Department of Hydrology and Water Resources, University of Arizona, USA*
[3]*Center of Excellence CETEMPS, University of L'Aquila, Italy*
Email: f.tapiador@bham.ac.uk

Neural networks (NNs) have been successfully used in the environmental sciences over the last two decades. However, only a few review papers have been published, most of which cover image processing, classification, prediction and geophysical retrieval in general, while neglecting rainfall estimation issues. This paper reviews, without aiming to be exhaustive, NN approaches to satellite rainfall estimation (SRE) by providing an overview of some of the methodologies proposed. A basic introduction to NNs is provided and the advantages of using NNs in SRE are explained, illustrating how NNs can be used to complement more computational-expensive methods to generate quick and accurate results in near real time. The role of the NNs in statistical-empirical algorithms is also reviewed. The last section aims to generate some discussion through comparing the empirical and deterministic algorithmic approaches and contrasting some of the apparent drawbacks of using NNs with a statistically based view of the satellite geophysical parameter estimation.

1. Introduction

Neural networks (NNs) aim to mimic the behaviour of brains. The brain cortex is nothing but a strongly interlaced processing system which takes the information provided by the sensory organs and generates an output (Rolls 1998): a flying mosquito responds to differences in air pressure (an input) by modifying the angle of attack of its wings (the output) after a quick processing of the data. In the same way, artificial neural networks simulate this processing mechanism by replacing the biological parts involved (that is, the sensory organs, neurotransmisors, neurons, etc.) for variables within a computer program. The same sort of decisions taken by the mosquito or a human pilot can be simulated by a NN-based system that can thus be implemented as an automatic control system (Rodin & Amin 1992).

NNs can also simulate more complex brain functions than simple control responses. These complex processing capabilities are the sorts of abilities required for satellite rainfall estimation (SRE hereafter), where partial or contradictory data can be present. However, the more complex the function becomes, the more computational power is required. A human brain contains about 10^{11} neurons connected by 10^{15} synapses. A typical multilayer perceptron (MLP) NN for SRE uses less than a hundred neurons, mainly due to processing constraints. Nevertheless, even with this scarce number of elements NNs can perform remarkable feats such as simulating radiative transfer equations for atmospheric correction (Göttsche & Olesen 2002). Considering that the NN performances are closely linked with the current state of the art of computing, improvements in the hardware can only benefit more complex reasoning, now that the theoretical foundations have been established. Moreover, since most of the NNs in SRE are presently implemented on serial computers, the developments in parallel computing with NNs will offer new possibilities in terms of training speed considerably shortening the time required to teach a net, improving real time systems and allowing more complex methods.

The most common approach to teach the NNs is through supervised learning. The idea is to train the NN using a sample of representative inputs and their expected outputs until the net is able to understand the link between them. The net learns how to compare its own outcomes with the outputs provided, adjusting the weight given to the synapses in order to minimise the differences between output and NNs estimation. Once the learning process has finished, the net can deal not only with the original training examples but also with new inputs not previously seen which would be correctly recognised (generalisation capability).

This procedure can be clarified with an example. Suppose we wish to generate a NN to simulate the GOES Precipitation Index (GPI) (Arkin 1979) using only infrared imagery. The required NN model is simple, as depicted in Figure 1.

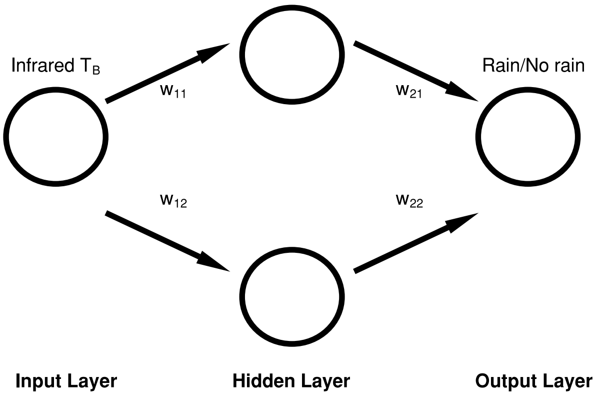

Figure 1. *NN model for the GPI simulation. w_{ij} represents the weights of the net.*

Table 1. *Training and testing dataset for the GPI NN simulation.*

Training dataset		Testing dataset	
Tb (K)	Rain (y/n)	Tb (K)	NN result (y/n)
200	y	10	y
210	y	209	y
234	y	233	y
235	y	235	y
236	n	236	n
240	n	245	n
250	n	252	n
300	n	910	n

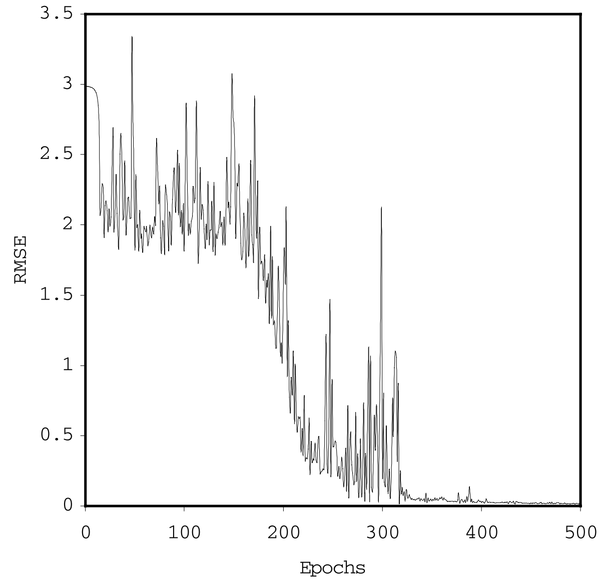

Figure 2. *Training graph of the NN of Figure 1 using the training dataset of Table 1.*

The first step is to train the net with the actual brightness temperature T_b and the expected result (0 mm/hr if the T_b is greater than 235 K and 3 mm/hr otherwise; or –1 for no rain and 1 for rain). After a few (several hundred) iterations, the error becomes small enough to conclude that the net is trained (Figure 2), meaning that the net is able to decide by itself which rainfall rate to assign.

If new data are presented, the net successfully recognises the cold values as rainfall (Table 1). This one-dimensional example is trivial but the more complex the input dataset becomes, the less the chance to deal with it in qualitative terms, and quantitative approaches in higher dimensions are required. NNs allow us to deal with these very complex problems in an elegant way, especially when we don't have a complete understanding of the physical processes involved.

Many papers have proclaimed the capabilities of NNs in remote sensing in recent years, both with regard to the classification and the retrieval of geophysical parameters. However, in spite of the achievements reported, NNs are still regarded as 'black boxes' capable of providing good results but lacking a rational explanation of the problem being solving. This paper also aims to clarify this statement, proving that NNs can no longer be considered as black boxes since much research in inversion and model selection has generated powerful methods to effectively investigate the net structure and behaviour. Even though the discussion that follows will be focused on satellite-based rainfall estimation, the methodologies outlined are of general applicability in remote sensing.

2. Potentiality of neural networks in satellite remote sensing

The non-parametric approach of the NN holds many advantages over other statistical procedures (Sarle 1994). As Hornik et al. (1989) have proved, multilayer perceptron NNs can approximate any measurable function up to an arbitrary degree of accuracy: as a semi-parametric regression estimator NNs can model a nonlinear function in a finite number of steps. This ability to extract nonlinear relationships is a very valuable feature in the remote sensing since many processes, such as rainfall, show this behaviour. Moreover, most of the statistical procedures of classification can be considered as particular cases of NNs. From a mathematical point of view, principal components analysis, maximum likelihood and a restricted form of the maximum entropy method are subsets of the more global NN set (Haykin 1999). Many applications have shown that NNs can improve classification accuracy by 10 to 30% compared with those conventional classification methods (Carpenter et al. 1997).

Additionally, NNs can deal with a great amount of data in an automatic and quick way without need for new calculations of the latest data: new information can also be easily integrated into the NN models without major modifications. Comparing this capacity with the important adaptations required in physically based models, NNs represent a great advantage in the operational field. NNs can deal even with noisy data – in fact some noise improves the convergence in the learning phase (Murray & Edwards 1994), which makes them a valuable tool in remote sensing retrieval where noise is almost always present.

However, NNs also present some problems. Hsieh & Tang (1998) have presented some drawbacks of using NNs in meteorological and climatologic data analysis, namely their nonlinear instability with short data records, the large spatial data fields, and the difficulties in interpreting the nonlinear NN results. Bearing their results in mind, the only drawback applicable to SRE could be in the need for statistically representative training examples. However, most of the physical models have been fine-tuned by using measurements at some stage (Kidd et al. 1998). These models are described with sets of equations that are constrained by physical laws, such as mass and energy balance. The equations are proposed upon some a priori knowledge of the system, and are then fine-tuned by measurements.

A common criticism directed at NNs is their apparent black-box nature. Nevertheless, much work has been developed to provide objective methods and procedures to investigate the NNs' behaviour and the meaning of the dynamics in the hidden layers. Both the dynamics of the learning phase and the model selection are well known and, in relation to the actual NN design, Anders and Korn (1999) used hypothesis testing, information criteria and cross-validation methods to find the more appropriate MLP structure. This model selection can then be guided by statistical modelling, providing a conscious selection both of the NN architecture and of its parameters (Murata et al. 1994). In addition to these MLP developments, other alternative methods such as the Cascade Correlation (CC) algorithm (Fahlman & Lebiere 1990) permit constructive creation of a NN by adding units from scratch until a final allowed error is achieved.

Another important argument against the black box claim is the possibility of studing the net once trained. A trained NN can be analysed further to find out what type of inputs can generate a given output. Inversion methods such as that presented by Jensen et al. (1999) can be applied to try to extract the assumptions the NN is making. Linden & Kindermann (1989), Linderman & Linden (1990) and Williams (1996) have also presented methods for the inversion of nets.

Qualitative approaches such as that presented in Hsieh & Tang (1998) might prove useful. Analysing the activation of the hidden units in the phase space can help in the interpretation of the net behaviour and also in choosing the appropriate number of neurons in the hidden layers, which is related to the degrees of freedom of the system. Spectral analysis can also facilitate the interpretation of the results, replacing the inputs by sinusoidal functions once the net has been trained and thus measuring the nonlinearity of the system. The presence and variation of such nonlinearities can then be related to unnoticed effects in the problem to be modelled, which can be exploited in conceptual models. In this sense, NNs can be used as a first guess when trying to physically model complex problems.

Perhaps even more important than inversion methods, NNs can be directly designed to solve not only the forward but also the inverse problem in remote sensing (Göttsche & Olesen 2002). However, this field is currently unexploited in SRE.

3. Overview of NN applications in satellite-based rainfall estimation

NNs have been successfully used in the environmental sciences (ASCE 2000; Govindaraju & Rao 2000; Maier & Dabdy 2000). However, only a few review papers have been published (Atkinson & Tatnall 1997; Gardner & Dorling 1998; Hsieh & Tang 1998; Krasnopolsky & Schiller 2003: Krasnopolsky & Chevallier 2003) mainly covering image processing, surface classification, series prediction and geophysical retrieval in general. Specific problems arising in SRE have only been indirectly addressed in these works.

The most common NN architecture used in remote sensing is the MLP NN. The GPI example corresponds with this kind of net. Other models such as Hopfield probabilistic nets (Hopfield 1982) or the Adaptative Resonance Theory NN models (ART, ARTMAP, Fuzzy ARTMAP and distributed ARTMAP; Grossberg 1969) are poorly represented in SRE. Only the self organizing maps (Kohonen 1997) or similar autorganizative models are sometimes used for classification purposes.

In order to clarify how NNs are used in SRE it is worthwhile differentiating two broad categories: one considering the kind of scientific paradigm in which NNs are inserted, and another considering the inputs required to generate the estimates.

3.1. Classification by inverse problem approaches

3.1.1. Physically-based methods

NNs have been used for physical inversion and forward procedures using radiative transfer equations (e.g. Li et al. 1997). The idea behind these methods is to utilise a NN as a quick method to simulate the numerical methods required to solve the non-analytically solvable equations. This overcomes the problems associated with the use of transfer functions such as the radiative transfer (RT) equations since it has been proven that the classical approach to the inversion problem is mathematically ill-posed (Parker 1994). This simulation of the transfer equations using NNs is possible since they are continuous mappings and a NN can approximate any continuous function with an arbitrary degree of accuracy just by increasing the number of neurons. In theory, NNs can map from the $[-1, 1]^n$ hypercube to the $(-1, 1)$ interval even using a single hidden layer (Kolmogorov 1963). This feature converts the NNs into universal approximators, which means that we can approximate almost any required function, including the Fourier transforms (Funahashi 1989). Many mathematical proofs have demonstrated these important results (Cybenko 1989 and Hornik 1993 among others).

There are two important reasons why we aim to simulate RT equations using NNs. First, when using classical inversion procedures, the high-quality measurements provided by satellites tend to be converted into lower-quality estimates – a methodology that amplifies the errors. Secondly, the use of NNs greatly improves the computational speed once trained. A clear demonstration of this, and of how to substitute RT models using NNs, is provided in Göttsche & Olesen (2001). In that work, comparison between the Moderate Resolution Transmittance Code 3 (MODTRAN-3) and an evolutionary NN demonstrated an improvement in the computational speed of $O(10^4)$ times. If any real time operation is intended, for example the use of Meteosat Second Generation data (15 minutes of temporal resolution) to derive rainfall rates, the improvement in the processing speed can be crucial (Levizzani et al. 2001).

Similar approaches to the RT simulations can be found in SRE. Krasnopolsky et al. (2000) compared a NN with a physically based retrieval algorithm (Wentz 1997). These authors simulated the columnar water vapour and columnar liquid water retrievals of the models using SSM/I measurements, closely reproducing its results, but improving the speed and thus allowing the use in forecast models such as the National Centers for Environmental Prediction (NCEP) numerical models. Cloud parameter retrieval models for inhomogeneous and fractional clouds have also been simulated (Faure et al. 2001) with reasonable accuracy.

3.1.2. Empirical or statistical methods

The empirical approach makes little or no assumption about the underlying relationships between the geophysical parameters measured and the estimated rainfall. The idea is to use the greatest possible number of available measurements to train the net against target rainfall measurements instead of model-derived values. Most of the operative procedures in rainfall estimation follow this approach. The PERSIANN system (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks) for example, demonstrates the NNs' capabilities when presenting a GOES-IR/TRMM-TMI working procedure to daily rainfall estimates at 1° resolution (Sorooshian et al. 2000).

In fact, most of the available empirical algorithms might be easily simulated using NNs. The simple mathematical operations involved in statistical rain mapping (Smith et al. 1998) and quasi-physical rain mapping algorithms of the second Precipitation Intercomparison Project (PIP-2) do not represent a difficulty for NN approaches. PIP-2's purely physical or profile-based algorithms could be more challenging, but the works mentioned in the last section have shown how almost all the equations involved in the current physical models can be accurately simulated through NNs.

3.2. Classification by satellite input data

3.2.1. Infrared-based methods

Rainfall estimates using only infrared imagery rely on the supposition that there is a relationship between cloud top temperature and surface rainfall. However, since the relationship between cloud top temperature and surface rainfall is highly biased for short-term estimates, pixel-based infrared (IR) methods work better when a large area and time-accumulated values are compared.

The major sources of IR data are the geostationary satellites. Using the moderate spatial resolution of these sensors (a few kilometres) and the high temporal sampling (a few hours) it is possible to train a NN with rain gauge data or other satellite rainfall estimates and generate pixel-resolution values. Some approaches aim to avoid the problem of indirectness of the measurement and take advantage of the scale of the rainfall processes using spatial and temporal adjustments. Thus, GOES data has been used to extract cloud features (Hsu et al. 2002) using a NN Self Organizing Map (SOM). These NNs permit the classification without supervision of a n-dimensional input vector into several classes (Kohonen 1997). By classifying the IR-derived cloud texture, geometry, brightness temperature and dynamics by NNs, it is possible to establish a relationship among these classes and measured rainfall distribution.

3.2.2. Passive microwave-based methods

The Special Sensor Microwave Imager (SSM/I) measurements (Hollinger et al. 1990) have been used as a primary source of passive microwave (PMW) information in SRE due to its large spatial coverage, the long-term aim of the program and the acceptable time sampling provided by the constellation. The advantages of using passive microwave information rather than infrared radiation rely on the directness of the observations, insofar as raindrops affect the upwelling Earth microwave emission. Thus, by matching the seven SSM/I channels at 19.35 (V&H), 22.23 (V), 37.0 (V&H) and 85.5 (V&H) GHz against a measured rainfall reference, a NN can be trained to extract a relationship between the 7-D vectors and the measured or estimated rainfall. For example, Mallet et al. (2002) used SSM/I to retrieve cloud Liquid Water Path (LWP) and Total Precipitable Water (TPW), finding the NN retrieval more regular than other algorithms when compared with radiosonde observations (RAOBs).

Target rainfall values for training can be derived from gauge, radar or other sensors. The NNs results obtained are closely linked with the accuracy of the data used: given a typically reported correlation between SSM/I rainfall and NN rainfall of 0.98 R^2 at 1 degree resolution, SSM/I vs. gauge correlation can yield a 0.75 R^2, with the NN vs. gauge not significantly better. The advantage in this case is being able to provide SSM/I-like information when there are no SSM/I data available (Bellerby et al. 2000).

Additionally, not only surface rainfall but also profiles can also be retrieved using NNs. Figures 3 to 5 illustrate this capability. The TRMM Microwave Imager (TMI) measures passive microwave measurements in nine channels. The rain estimates are derived using a Bayesian approach for previously calculated profiles using the Goddard Profiling (GPROF) algorithm (Kummerow et al. 1996). Training a NN with the TMI measurements as input and the GPROF profiles (2A12 product) as output, we can simulate the algorithm within a high degree of accuracy: Figure 3a shows how the NN is able to store almost perfectly all the profiles in the training database. However, once the NN has been trained, it maintains the ability to generalise: Figure 3b shows the correlation found when independent data are presented, while Figures 4 and 5 give a qualitative appraisal of the results.

3.2.3. Satellite data fusion methods

Fused methods aim to combine the temporal sampling of geostationary, high orbit satellites with the direct measurement of rainfall provided by low Earth orbit (LEO) passive microwave sensors. This synergic approach presents problems due to the different resolution of the sensors and the temporal sampling, but usually performs better than pure IR or PMW methods separately.

The above-mentioned approach of using textures and patterns instead of the brightness temperature of a single pixel is also used in the fused method to deal with the inherent uncertainty of pixel-based approaches. Thus, for example, GOES data can be used to train a net against TRMM PR measurements (Bellerby et al. 2000). These authors reported better results than the optimised GOES Precipitation index (UAGPI) when textural information and thresholds were used to train a feed-forward NN. Correlations up to 0.93 R^2 at 1° resolution were obtained in this work. Another example is the above-mentioned PERSIANN system. The same procedure has been applied to Meteosat satellites (Marzano et al. 2002). Small-scale, short-time estimation methods have also been proposed (Tapiador et al. 2004), showing how seemingly bad correlations at the instantaneous pixel-based scales can be improved, when spatial and temporal accumulations are performed as some researchers have pointed out (Turk et al. 2002).

4. Discussion

In spite of the good results obtained in applying NNs to SRE, much criticism has been aimed at their empirical methodology. It is also said that if the models are based on physical mechanisms, adjustments are possible to match different circumstances. The main argument is that the final goal of the research is to provide a set of equations that represent all the known physical effects, not just to generate good results. However, this is partially true: NNs are not a replacement for conceptual, physical models since those models are not solely operated for prediction purposes.

Conceptual models can also be criticised with their own arguments, especially when applied to SRE. First, most of the assumptions made when dealing with a large number of particles are also statistical constructs to simplify reality and to allow us to deal with problems in a mathematical way. Even for very detailed hypotheses on the shape and size of the drops, no deterministic model can deal with the huge variety that can be found in the clouds at global scale – or even at the local scale. Mie's or Rayleigh's assumptions are in fact useful simplifications which aim to make tractable some problems involving a large number of dissimilar particles. When non-trivial physical models are being developed the use of any kind of simplification is always involved, but these generalisations are considered to be rational assumptions, and we still believe that even with these simplifications we can provide an explanation of the very complex reality. We are willing to accept that, for some applications, raindrops can be considered as spherical-like since a statistical average of many almost-spherical drops will differ by an acceptable degree to a hypothetical perfect sphere or another easily parameterisable shape. We can refine this initial approach by including a few ensembles of different populations, but we can never take into account all the possibilities that actually occur in nature.

A second criticism against conceptual models is that the use of the a priori hypotheses required presents a major drawback: all the effects must be known prior to the estimation. If a model is built a prediction is made and if the results do not fit with empirical

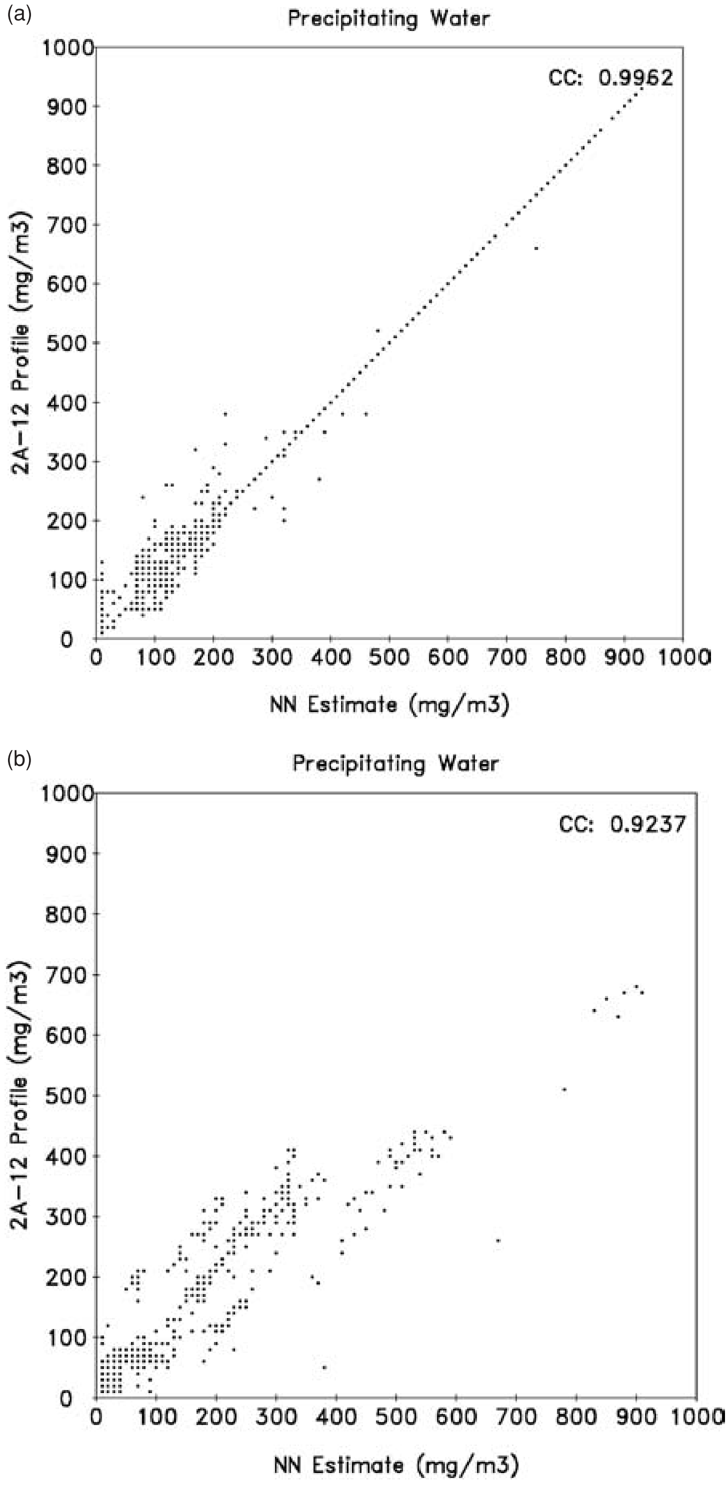

Figure 3. *Comparison between the TMI 2A12 land profile estimates and the NN estimate using the 19.35 GHz (V&H), 21.3 GHz, 37.0 GHz (V&H), and 85.5 GHz (V&H) TMI channels (1B11 product). In (a), the NN was trained with just 10 orbits of TMI 1B11/2A12 couples and then applied to one 1B11 overpass previously used for training. This demonstrates the storage capabilities of the NN. (b) represents an independent swath, which shows how the NN is able to generalise with data not seen before once it has been trained. The performances of the NN are linked both with the number of orbits used to train it and with the training epochs.*

data, new assumptions need to be included to tune the model. On the contrary, empirical methods do not need to know about all the effects involved. The sometimes used trial-and-error method of deterministic modelling is replaced in NNs by empirical testing that reflects the accuracy and precision of the model. If the sample used is representative, the results of the empirical methods can be extrapolated and new results can be obtained. Besides, conceptual models in SRE can benefit from NN results since, once the fit between the physical magnitudes measured by the satellite and the geophysical parameters has been established, there

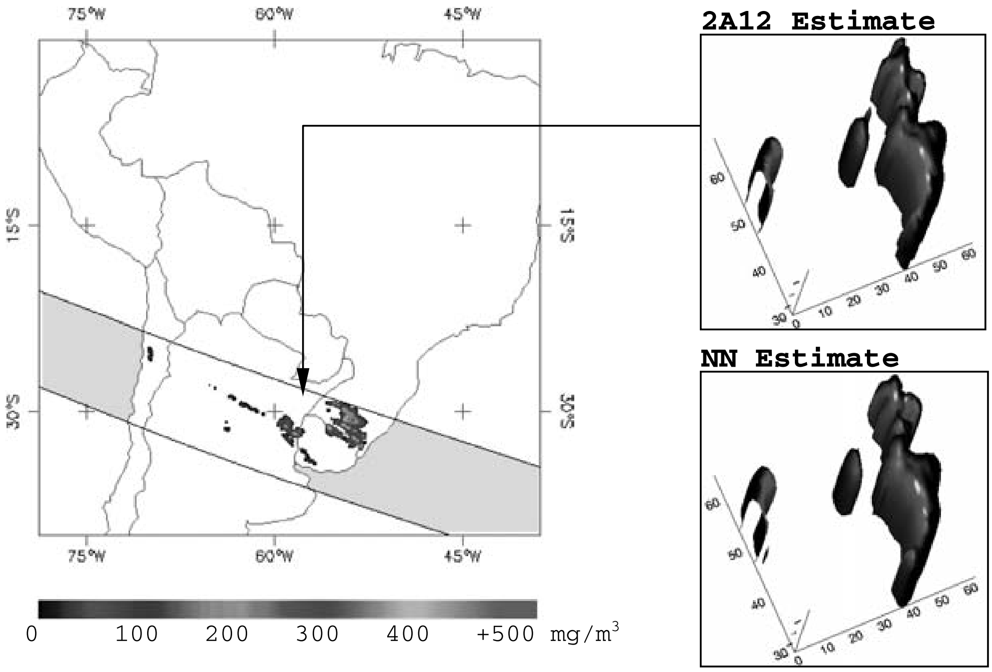

Figure 4. *Near surface rainfall retrieval over land (southern Brazil) for 1 July 1998, orbit 3394 (left) and comparison with the 3D structure provided by the 2A12 product and the NN simulation (right).*

are methods to investigate the net behaviour, as we have seen in section 2.

In the direct observation of nature – the first step of any scientific enquiry – the less a priori the hypotheses, the better will be the appraisal of the real behaviour of nature. In this sense, empirical procedures use only the available information to predict results. Much more importantly, this use of this available information is logically justified by the maximum entropy principle (Shore & Johnson 1980; Jaynes 1990).

If the final goal of science is to predict using an objective and repeatable procedure, there is no reason to prefer physically based algorithms over empirical ones, providing the results had the same precision and accuracy. Nevertheless, conceptual models are of course necessary. To gain an understanding of how the process is working, a physical- or conceptual-based model may provide a better insight, because they are based on our physical understanding of the process, no matter whether they are at molecular level or in a mean or statistical distribution. On many occasions, we

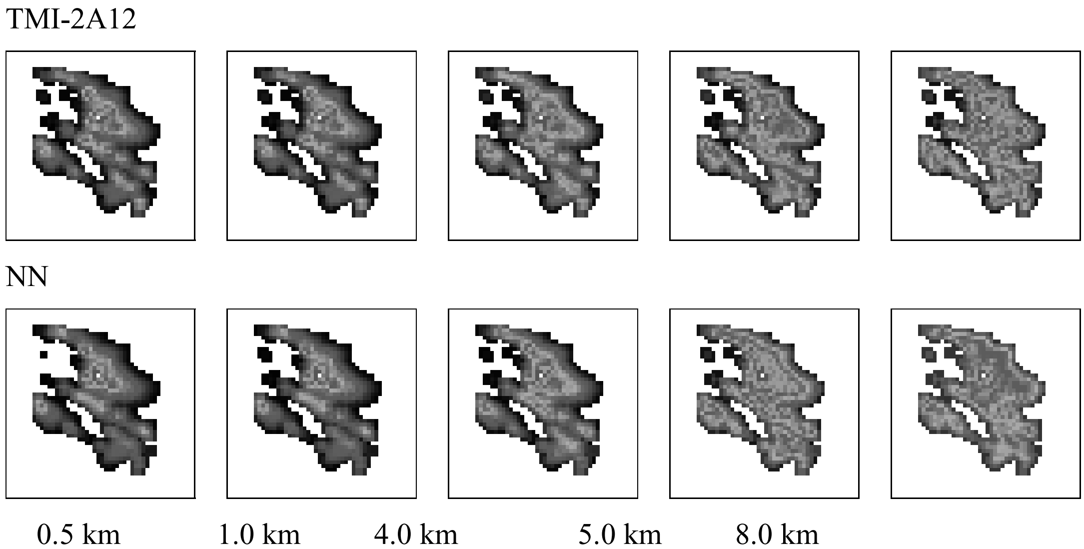

Figure 5. *A comparison of precipitable water maps for the land-NN corresponding to the rain area over southern Brazil in Figure 4 (1 July 1998; orbit 3394), showing different levels. The NN was trained against the two previous orbits. The top images show the 2A12 estimate while the images below are the NN estimates. The agreement is remarkable (colour-scale as Figure 4).*

are interested not only in the forecasting variables but also in the state variables or parameters in the system. NNs do not aim to replace conceptual models but to complement them, sometimes using their own results to improve the computational speed.

Rainfall estimation is a problem that contains a large number of unknown processes. It is unlikely that we can find an exact solution with insufficient information. The inverse problem of the radiative transfer process is an ill-posed scenario which yields many solutions for a given brightness temperature. In such circumstance, NNs models, which could learn from data and are able to generalise the model for rainfall estimation, are very useful.

5. Conclusions

NNs are universal approximators – any function currently used in physically based modelling can be approximated within an arbitrary degree of accuracy using a NN, which usually improves the computational speed and therefore allows near real time applications in SRE. Several recent works have used NNs to exploit these possibilities in terms of generalisation ability and learning by examples. Reported results improve other methods and present a promising initial development that can be improved through new theoretical work and new computational capabilities.

This paper has presented some of the developments in SRE using NNs, showing some examples of applications in each of the fields defined. Infrared, passive microwave and fused approaches have been described and the contribution of NNs has been explained.

Finally, NNs are not only powerful tools for defining relationships between parameters but also statistical means by which complex systems can be modelled. The satellite retrieval of geophysical parameters is, in fact, a measurement of averaged values over large inhomogeneous areas sometimes involving poorly known processes. In this scenario, the empirical approach of the NNs to SRE is not a drawback since when a large number of elements are involved and a certain degree of uncertainty is present, a probabilistic approach is always required. Physically based approaches play a key role in understanding the processes involved, but operational rainfall estimates using incomplete knowledge may require the use of empirical methodologies such as NNs.

Acknowledgments

This research was funded by the EURAINSAT project, a shared-cost project (contract EVG1-2000-00030), co-funded by the Research DG of the European Commission within the RTD Activities of a Generic Nature of the Environment and Sustainable Development sub-programme (5th Framework Programme).

References

Anders, U. & Korn, O. (1999) Model selection in neural networks. *Neural Networks* **12**: 309–323.

Arkin, P. A. (1979) The relationship between fractional coverage of high cloud and rainfall accumulations during GATE over the B-scale array. *Mon. Wea. Rev.* **107**: 1382–1387.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000) Artificial neural network in hydrology. II: Hydrologic applications. *J. Hydrologic Engineering, ASCE* **5**(2): 124–137.

Atkinson, P. M. & Tatnall, A. R. L. (1997) Neural networks in remote sensing – introduction. *Int. J. of Remote Sensing* **18**(4): 699–709.

Bellerby, T., Todd, M., Kniveton, D. & Kidd, C. (2000) Rainfall estimation from a combination of TRMM Precipitation Radar and GOES multispectral satellite imagery through the use of an artificial neural network. *J. Appl. Meteorol.* **39**: 2115–2128.

Carpenter, G. A., Gjaja, M. N., Gopal, S. & Woodcock, C. E. (1997) ART neural networks for remote sensing: vegetation classification from Landsat TM and terrain data. *IEEE Trans. on Geoscience and Remote Sensing* **35**: 308–325.

Cybenko, G. (1989) Approximation by superposition of a sigmoidal function. *Math. Control Signals Systems* **2**: 303–314.

Fahlman, S. E. & Lebiere, C. (1990) *The Cascade-Correlation Learning Algorithm*, Technical Report CMU-CS-90-100.

Faure, T., Isaka, H. & Guillemet, B. (2001) Neural network retrieval of cloud parameters of inhomogeneous and fractional clouds. Feasibility study. *Remote Sens. Environ.* **77**: 123–138.

Funahashi, K. (1989) On the approximate realization of continuous mappings by neural networks. *Neural Networks* **2**(3): 183–192.

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

Göttsche, F.-M. & Olesen, F. S. (2002) Evolution of neural networks for radiative transfer calculations in the terrestrial infrared. *Remote Sens. Environ.* **80**: 157–164.

Govindaraju, R. S. & Rao, A. R. (2000) *Artificial Neural Networks in Hydrology.* Dordrecht: Kluwer Academic Publishers, 329pp.

Grossberg, S. (1969) Some networks that can learn, remember and reproduce any number of complicated space-time patterns. *J. Mathematics and Mechanics* **19**: 53–91.

Haykin, S. S. (1999) *Neural Networks: A Comprehensive Foundation.* 2nd edn. Upper Saddle River, NJ: Prentice Hall.

Hollinger, J. P., Pierce, J. L. & Poe, G. A. (1990) SSM/I instrument evaluation. *IEEE Trans. on Geosci. and Remote Sensing* **28**: 781–789.

Hopfield, J. J. (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc. National Academy of Sciences* **79**: 2554–2558.

Hornik, K., Stinchcombe, M. & White, H. (1989) Multilayer neural networs are universal approximators. *Neural Networks* **2**: 359–366.

Hornik, K. (1993) Some new results on neural network approximation. *Neural Networks* **4**(2): 251–257.

Hsieh, W. W. & Tang, B. (1998) Applying neural network models to prediction and data analysis in meteorology and oceanography. *Bull. Am. Meteorol. Soc.* **79**: 1855–1870.

Hsu, K., Sorooshian, S., Hong, Y. & Gao, X. (2002) Cloud classification and rainfall estimation using GOES imagery. International Conference of Quantitative Precipitation Forecasting, Reading, UK, 2–6 September 2002.

Jameson, A. R. & Kostinski, A. B. (2002) What is a raindrop size distribution? *Bull. Am. Meteorol. Soc.* **82**: 1169–1177.

Jaynes, E. T. (1990) Probability theory as logic. In P. F. Fougère (ed.), *Maximum-Entropy and Bayesian Methods*, Dordrecht: Kluwer Academic Publishers.

Jensen, C. A., Reed, R. D., Marks II, R. J., El-Sharkawi, M. A., Jung, J.-B., Miyamoto, R. T., Anderson, G. M. & Eggen, C. J. (1999) Inversion of feedforward neural networks: algorithms and applications. *Proc. of the IEEE* **87**: 1536–1549.

Kidd, C., Kniveton, D. & Barrett, E. C. (1998) The advantages and disadvantages of statistical derived-empirically calibrated passive microwave algorithms for rainfall estimation. *J. Atmos. Sci.* **55**: 1576–1582.

Kohonen, T. (1997) *Self-Organizing Maps*, 2nd edn. Berlin: Springer.

Kolmogorov, A. N. (1963) On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition. *Am. Mathematical Soc. Trans.* **28**: 55–99.

Krasnopolsky, V. M., Gemmill, W. H. & Breaker, L. C. (2000) A neural network multiparameter algorithm for SSM/I ocean retrievals: comparisons and validations. *Remote Sens. Environ.* **73**: 133–142.

Krasnopolsky, V. M. & Chevallier, F. (2003) Some neural network applications in environmental sciences. Part II: Advancing computational efficiency of environmental numerical models. *Neural Networks* **16**: 335–348.

Krasnopolsky, V. M. & Schiller, H. (2003) Some neural network applications in environmental sciences. Part I: Forward and inverse problems in geophysical remote measurements. *Neural Networks* **16**: 321–334.

Kummerow, C., Olson, W. S. & Giglio, L. (1996) A simplified scheme for obtaining precipitation and vertical hydrometeor profiles from passive microwave sensors. *IEEE Trans. Geosci. and Remote Sensing* **34**: 1213–1232.

Levizzani, V., Schmetz, J., Lutz, H. J., Kerkmann, J., Alberoni, P. P. & Cervino, M. (2001) Precipitation estimations from geostationary orbit and prospects for METEOSAT second generation. *Meteorol. Appl.* **8**: 23–41.

Li, L., Vivekanandan, J., Chan, C. H. & Tsang, L. (1997) Microwave radiometric technique to retrieve vapor, liquid and ice: Part I. Development of a neural network-based inversion method. Part II: Joint studies of radiometer and radar winter clouds. *IEEE Trans. Geosci. and Remote Sensing* **35**: 224–247.

Linden, A. & Kindermann, J. (1989) Inversion of multilayer nets. *Proc. Int. Joint Conf. Neural Networks, vol II*, Washington, DC, 425–430.

Linderman, J. & Linden, A. (1990) Inversion of neural networks by gradient descent. *Parallel Comput.* 277–286.

Maier, H. R. & Dabdy, G. C. (2000) Neural networks for prediction and forecasting of water resources variables: a review of modeling issues and applications. *Environ. Modeling and Software* **15**: 101–124.

Mallet, C., Moreau, E., Casagrande, L. & Klapisz, C. (2002) Determination of integrated cloud liquid water path and total precipitable water from SSM/I data using a neural network algorithm. *Int. J. Remote Sensing* **23**(4): 661–674.

Marzano F. S., Fionda, E. & Ciotti, P. (2002) Inversion techniques for ground-based microwave radiometric retrieval of precipitation columnar contents and path attenuation. *Proc. IGARSS'02*, Toronto (Canada), 24–29 June.

Marzano, F. S., Palmacci, M., Cimini, D. & Turk, J. F. (2002) Statistical integration of satellite passive microwave and infrared data for high-temporal sampling retrieval of rainfall. *Proc. IGARSS'02*, Toronto (Canada), 24–29 June.

Murata, N., Yoshizawa, S. & Amari, S. (1994) A criterion for determining the number of parameters in an artificial neural network model. *IEEE Trans. Neural Networks* **5**: 865–872.

Murray, A. F. & Edwards, P. J. (1994) Enhanced (MLP) performance and fault tolerance resulting from synaptic weight noise during training. *IEEE Trans. Neural Networks* **5**(5): 792–802.

Parker, R. L. (1994) *Geophysical Inverse Theory*. Princeton: Princeton University Press.

Rodin, E. Y. & Amin, M. (1992) Maneuver prediction in air combat via artificial neural networks. *Computers and Mathematics With Applications* **24**(3): 95–112.

Rolls, E. T. (1998) *Neural Networks and the Brain*. Oxford: Oxford University Press.

Sarle, W. S. (1994) Neural networks and statistical models. *Proc. 19th Annual SAS Users Group International Conference*. Cary, NC, SAS Institute, 1538–1550.

Shore, J. E. & Johnson, R. W. (1980) Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inf. Theory*, **IT-26**: 26–39 and **IT-29**: 942–943.

Smith, E. A., Lamm, J., Adler, R., Alishouse, J., Aonashi, K. et al. (1998) Results of WetNet PIP-2 project. *J. Atmos. Sci.* **55**: 1483–1536.

Sorooshian, S., Hsu, K. L., Gao, X., Gupta, H. V., Imam, B. & Braithwaite, D. (2000) Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Am. Meteorol. Soc.* **81**: 2035–2046.

Tapiador, F. J., Kidd, C., Levizzani, V. & Marzano, F. S. (2004) A neural networks-based fusion technique to estimate half hourly rainfall estimates at 0.1° resolution from satellite passive microwave and infrared data. *J. Appl. Meteorol.* **43**, 309, pp. 98–116.

Turk, F. J., Ebert, E. E., Oh, H.-J. & Sohn, B.-J. (2002) Validation and applications of a realtime global precipitation analysis. *Proc. IGARSS, 24–28 June 2002*, Toronto, Canada.

Wentz, F. J. (1997) A well-calibrated ocean algorithm for special sensor microwave/imager. *J. Geophys. Res.* **102**: 8073–8718.

Williams, R. J. (1986) Inverting a connectionist network mapping by backpropagation of error. *Proc. 8th Annual Conf. Cognitive Science Society.* Hillsdale, NJ: Lawrence Erlbaum, 859–865.