  Journal of Hydroinformatics | 05.1 | 2003

An adaptive neuro-fuzzy inference system for the post-calibration of weather radar rainfall estimation

Masoud Hessami, François Anctil and Alain A. Viau

ABSTRACT

An Adaptive Neuro-Fuzzy Inference System, based on a jack-knife approach, is proposed for the post-calibration of weather radar rainfall estimation exploiting available raingauge observations. The methodology relies on the construction of a fuzzy inference system with three inputs (radar *x* coordinate, *y* coordinate and rainfall estimation at raingauge locations) and one output (raingauge observations). Subtractive clustering is used to generate the initial fuzzy inference system. Artificial neural network learning provides a fast way to automatically generate additional fuzzy rules and membership functions for the fuzzy inference system. Fuzzy logic enhances the generalisation of the artificial neural network system. In order to demonstrate the steps of the radar rainfall post-calibration using the Adaptive Neuro-Fuzzy Inference System, CAPPIs of one-hour rainfall accumulation and corresponding raingauge observations have been selected. Results show that the proposed approach looks for a response that is a compromise between radar rainfall estimations and raingauge observations and does not necessarily consider the raingauge observations as ground truth. The algorithm is very fast and can be implemented for real time post-calibration. This algorithm makes use of all available data—raingauge observations are usually scarce—for training and checking the neuro-fuzzy inference system. It also provides a degree of reliability of the post-calibration.

Key words | jack-knife, neuro-fuzzy systems, post-calibration, rainfall, weather radar

Masoud Hessami
François Anctil (corresponding author)
Department of Civil Engineering,
Université Laval,
Sainte-Foy,
Quebec,
Canada G1K 7P4
Tel: +1 418-656-3653;
Fax: +1 418-656-2928;
E-mail: *francois.anctil@gci.ulaval.ca*

Alain A. Viau
Department of Geomatic Sciences,
Université Laval,
Sainte-Foy,
Quebec,
Canada G1K 7P4

INTRODUCTION

Accurate measurements of rainfall over time and space are critical for many hydrological and meteorological projects. The most usual tools to monitor rainfall events are raingauges and weather radar, whose usefulness is limited in different ways. Networks of raingauges provide accurate point estimates of rainfall, when appropriately set, but their usual low density considerably restricts the spatial resolution of the gathered information. The quality of raingauge data is also susceptible to some error sources, especially biological and mechanical fouling, and human and environmental interference (Steiner *et al.* 1999). Weather radar is much more efficient in providing the space–time evolution of a rainfall event, but the precision of their estimates is often plagued by many factors including ground clutter, bright band, anomalous propagation, beam blockage and attenuation (e.g. Zawadski 1984; Andrieu *et al.* 1997). The effectiveness of weather radar operation is strongly linked to a rigorous calibration (Serafin & Wilson 2000). The performance of radar rainfall estimation mainly depends on a proper choice of Z–R relationship (Anagnostou & Krajewski 1998), which may vary from event to event or even within a single storm, where Z is the radar reflectivity factor ($mm^6\ m^{-3}$) and R is the precipitation rate ($mm\ h^{-1}$). A recent experience of a proper choice of Z–R relationship returns to the work of Rongrui & Chandrasekar (1997) who have proposed a neural network based approach to determine a Z–R relationship.

Early on, Wilson (1970) has recognised the strengths and weaknesses of both observation systems and proposed to integrate them in order to enhance the space–time quality of the rainfall information. Since then, various methods have been proposed to achieve this. They can be classified into two main categories: deterministic and statistical. The deterministic approach involves the calibration of radar rainfall estimations against raingauge observations (Wilson 1970; Andrieu *et al*. 1997). The statistical approach includes multivariate analysis (Eddy 1979), Kalman filtering (French & Krajewski 1994; French *et al*. 1994) and cokriging (Krajewski 1987; Seo *et al*. 1990). Geostatistical approaches are known as the best methods for radar–raingauge data integration but they are usually inefficient in real time, especially when dealing with sampling rates of one hour or less, necessary for urban and small watershed applications. Such methods also rely on strong human expertise, which can lead to user-dependent results (Bollivier *et al*. 1997). The technique based on Kalman filtering needs known error estimates, which are not usually available (Huffman *et al*. 1995). Overall, these methods share a similar objective: to somehow perform a post-calibration of the radar estimation using raingauges as ground truth—recall that, on occasions, raingauges also depart from truth (Steiner *et al*. 1999).

Several authors have reported the usefulness of fuzzy methods for spatial data analysis. Bardossy *et al*. (1990*a, b*) have used the fuzzy set theory for variogram modelling. A simple application for an environmental impact analysis has been described by Anile *et al*. (1995), which shows the effectiveness of fuzzy arithmetic. Piotrowski *et al*. (1996) have introduced a fuzzy kriging interpolation for rationalisation of geological data. A kriging method based on a combination of Bayesian and fuzzy approaches has been presented by Bandemer & Gebhardt (2000). In this paper, we study the merging of radar rainfall estimations and raingauge observations based on the combination of fuzzy logic and artificial neural networks. The point of this approach is to map the input space (radar) to the output space (raingauges) through an Adaptive Neuro-Fuzzy Inference System in order to achieve a post-calibration of the weather radar rainfall estimation.

The remainder of the paper is organised as follows. A brief description of the selected rainfall event, along with instrumentation, is first presented. General concepts of fuzzy logic and the proposed adaptive neuro-fuzzy inference system are described next. Results of the post-calibration of the selected weather radar event, discussion and conclusion are finally presented in sequence.

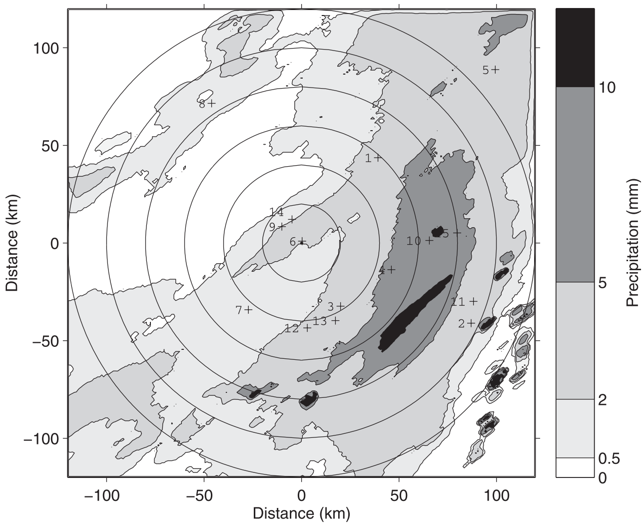

Figure 1 | CAPPI of the 8 am to 9 am rainfall accumulation: 17 June 1997. Crosses indicate raingauge locations.

DATA SELECTION

McGill Radar Weather Observatory has provided radar rainfall estimations for this study. The radar, located at the western tip of the island of Montreal, at Sainte-Anne-de-Bellevue, transmits in the S band (10 cm). It scans the atmosphere using a regular strategy. Data are collected at 24 elevation angles from 0.5° to 34.4° every 5 min. The reflectivity CAPPI (Constant Altitude Plan Position Indicator) is the radar image for displaying precipitation intensity. The CAPPIs used for this study are the one-hour rainfall accumulations from 5 am to 1 pm on 17 June 1997, with a resolution 1 km × 1 km, obtained from an altitude of 2 km (Figure 1). The precipitation melting layer (bright band) has been avoided by choosing an altitude of 2 km, but the image still suffers from some common problems, in particular ground clutter and anomalous propagation. Nevertheless, these imperfections do not affect the general data integration methodology presented here—the complete radar image correction is outside the scope of this paper. The corresponding hourly raingauges observations used for this study were collected by Environment Canada's network, supplemented by a private raingauge network (Figure 1).

FUZZY LOGIC

Fuzzy logic is based on the theory of fuzzy sets, which was first developed by Zadeh (1965). Fuzzy set theory allows for using partial membership, which differs from the traditional binary membership of Boolean set theory. Boolean set theory is two-valued in the sense that a member either belongs to a set (represented by 1) or does not belong to a set (represented by 0). In fact, fuzzy set theory is just the generalisation of Boolean set theory, which permits all real values ranging between 0 and 1. The degree of membership between 0 and 1 is defined by a membership function, a relation that defines how each point in the input space is mapped to a membership value between 0 and 1.

A precise mathematical description is not always necessary to optimise an operation. For example, most of human reasoning is based on imprecise knowledge. Fuzzy logic reasoning is devised in a similar fashion, by making use of approximate information and uncertainties for decision-making. Fuzzy set theory groups data without the need for clearly defined boundaries. It is mathematically designed to manage complex situations and imprecise data.

Fuzzy logic allows the mapping of an input space to an output space based on membership functions, fuzzy logic operations and parallel if–then rules: an overall process called fuzzy inference. The fuzzy inference system used in this paper is based on the Takagi–Sugeno–Kang method (Sugeno 1985). The main steps in the fuzzy inference process are: to fuzzify inputs by determining the membership functions, to determine if–then rules, to combine antecedent parts (if-parts) by fuzzy operators, to apply an implication method for each rule (combining if-parts and then-parts), to aggregate outputs (combining the outputs of each rule) to obtain a single fuzzy set, and to defuzzify the output membership function to provide a single output value.

ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM

As system complexity increases, reliable fuzzy rules and membership functions used to describe the system behaviour are difficult to determine. Neuro-adaptive learning techniques provide a method to integrate information from a data set (learn), in order to compute the membership function parameters that best allow the associated fuzzy inference system to track the given input/output data. This technique constructs a fuzzy inference system whose membership function parameters are adjusted through a learning process similar to that of neural networks. Because of their complementary nature, these two technologies can be integrated in a number of ways to overcome the drawbacks of each other. Neural network learning provides a good way to adjust the expert's knowledge, to automatically generate additional fuzzy rules and membership functions to meet certain specifications and to reduce design time and costs. On the other hand, fuzzy logic enhances the generalisation capability of artificial neural networks.

The first step in generating a Sugeno fuzzy inference system consists of clustering the input–output training data to identify a model that suits the data while minimising the number of rules. The number of clusters and their centres are estimated using a fast algorithm called subtractive clustering (Chiu 1994), which automatically detects clusters in the training data. However, cluster radii are unknown and must be selected manually. Small cluster radii result in more rules than large cluster radii. The optimum cluster radius is obtained by trial and error. The subtractive clustering algorithm is used for determining the initial rules of the adaptive neuro-fuzzy inference system. Further optimisation of these initial rules is performed by a neural network model. For the post-calibration of the weather radar estimations, the fuzzy inference system is built around three inputs, the raingauges x coordinate Xr and y coordinate Yr, and the radar rainfall estimation Zr, along with the raingauge observations Zp as output. For example, such a Sugeno model with two rules can be expressed as

IF *Xr is MF1x* **AND** *Yr is MF1y* **AND** *Zr is MF1z* **THEN** *Zo is MF1o*

IF *Xr is MF2x* **AND** *Yr is MF2y* **AND** *Zr is MF2z* **THEN** *Zo is MF2o*

where *Zo* is the rainfall predicted by the model; *MF1x*, *MF1y* and *MF1z* are the first Gaussian membership functions; *MF2x*, *MF2y* and *MF2z* are the second Gaussian membership functions of the input space; *MF1o* and *MF2o* are the first-order linear membership functions of output space:

$$Zo = PXr + QYr + RZr + S \quad (1)$$

where P, Q, R and S are all constants. Equation (1) provides the output of each rule, and the final output is the weighted average of each rule's output. Therefore, one can increase the complexity of the system by adding more rules to the system and approximate any non-linear mapping with desired accuracy based on the linear submodels. The parameters of this system are updated in an iterative way using hybrid learning. This hybrid procedure updates the antecedent and the consequent parameters by a backpropagation algorithm and least squares, respectively.

Over-fitting is one of the problems that may occur during the training of an adaptive neuro-fuzzy inference system. The usual approach to avoid over-fitting is to split the observations into two groups, where the first group is used for training and the second one is used separately to prevent deterioration of the generalisation quality of the model. However, when the overall database is small—for example, in this study, there are 15 raingauges to post-calibrate a 240 km × 240 km weather radar image—other ways to achieve the generalisation of the model must be sought. The following jack-knife approach has been selected to assure the generalisation performance of the adaptive neuro-fuzzy inference system. An observation is removed from the database. It is used as the checking data for the optimisation of a model based on the rest of the database, i.e. that the training process is stopped when the checking error (the checking observation minus the model realisation at this location) is minimised. This process is repeated until all observations have been removed once. Together, all these models form the adaptive neuro-fuzzy inference system sought. Let Zo_i be the output of the ith fuzzy inference system structure and the mean value of all outputs *Zm* is the result of the data fusion:

$$Zm = \frac{1}{N}\sum_{i=1}^{N} Zo_i \quad (2)$$

where N is the number of raingauges. A root mean square error can also be computed for each estimation:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Zo_i - Zm)^2}\,. \quad (3)$$

This approach allows the use of all available data for training and for checking the fuzzy inference system. Furthermore, *RMSE* provides an indication of the reliability of the estimation (convergence of all models).

RESULTS

In order to present step by step results, we have selected the event of 17 June 1997 from 8 am to 9 am. This event has been selected because it is typical of large systems rainfall and because it is well centred on the radar. For this event, the radar overestimates the raingauges in most locations (Figure 2), leading to a

$$RMSE_r = \sqrt{\frac{1}{15}\sum_{i=1}^{15}(Zp_i - Zr_i)^2} \quad (4)$$

of 1.14 mm. The 15 separate processes from variable fuzzification to defuzzification of the aggregate output have been repeatedly performed by removing one raingauge at a time. For example, the results when raingage 1 is removed are presented. For subtractive clustering of the remaining 14 raingauges, the cluster radii must be specified. An initial guess of 40 km for *Xr* and *Yr* and of 2 mm for *Zr* and *Zo* leads to a 5-rule fuzzy inference system. The *RMSE* of the training data is then 0 mm, which is excellent, but the checking error is 22 mm, which is huge considering that the checking value is 2 mm. These results are a good example of over-fitting. Consequently, the model has no generalisation capability, as revealed by its poor results when used to post-calibrate the weather radar rainfall estimation (Figure 3).

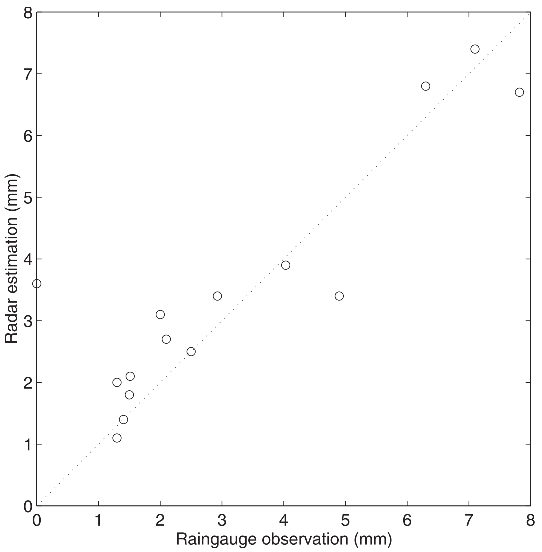

Figure 2 | Scatter plot of the raingauge observations and of the weather radar estimations: 17 June 1997 from 8 am to 9 am.

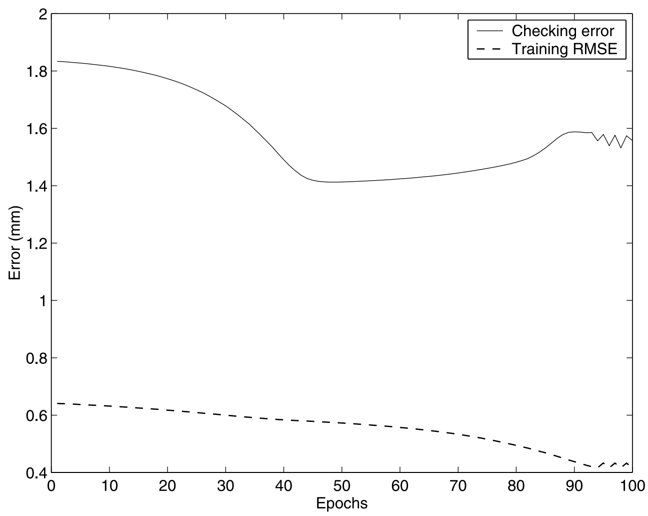

Figure 4 | Evolution of the training *RMSE* and of the checking error of the 2-rule fuzzy inference system, when raingauge 1 is left out, over 100 artificial neural network training epochs.

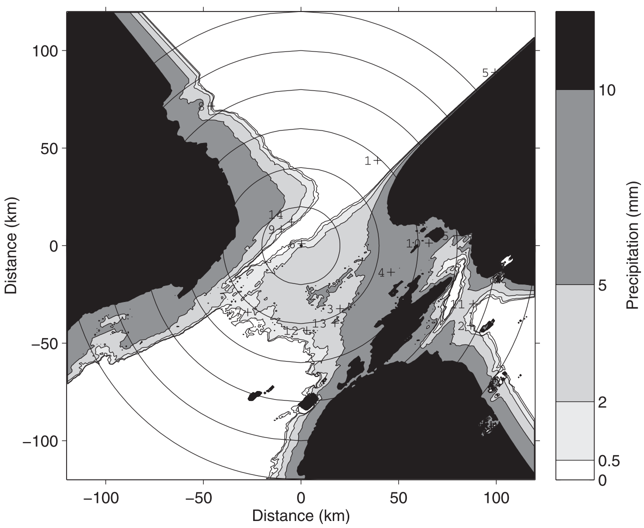

Figure 3 | CAPPI derived from the 5-rule fuzzy inference system, when raingauge 1 is left out: 17 June 1997 from 8 am to 9 am.

Over-fitting is controlled by reducing the number of fuzzy rules. This is achieved by the selection of larger cluster radii. A few trials rapidly lead to much improved models. For example, a 2-rule fuzzy inference system—*RMSE* of 0.64 mm and checking error of 1.83 mm—ensues from cluster radii of 50 km for *Xr* and *Yr* and of 2.5 mm for *Zr* and *Zo*. Such a model is significantly improved with respect to the checking observation while fitting the training observations reasonably well.

Once a satisfying fuzzy inference system is obtained, it is further optimised using an artificial neural network procedure. Figure 4 shows the evolution of the training *RMSE* and of the checking error over 100 artificial neural network training epochs, where the first epoch values are those of the fuzzy inference system. In this case, the smallest checking error occurs at epoch 48, for which parameters will be retained—training RMSE of 0.58 mm and training checking error of 1.41. Figure 5 shows the post-calibration results based on this model. One can see that it is a significant improvement over the 5-rule fuzzy inference system used to create Figure 3. The above 2-rule neuro-fuzzy inference system has a much better generalisation capability which is suitable for weather radar precipitation post-calibration.

The jack-knife procedure asks for the above modus operandi to be repeated employing alternatively one raingauge as the checking observation. The 15 training checking errors thus obtained are drawn in Figure 6. Most of them are close to zero. The worst one (raingauge 2) is associated with a raingauge observation of 0 mm and a radar estimation of 3.6 mm (see Figure 2).

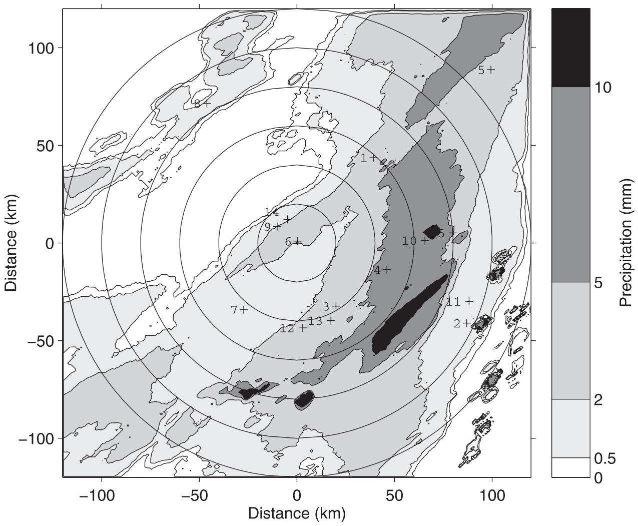

Figure 5 | CAPPI derived from the optimised 2-rule neuro-fuzzy inference system, when raingauge 1 is left out: 17 June 1997 from 8 am to 9 am.

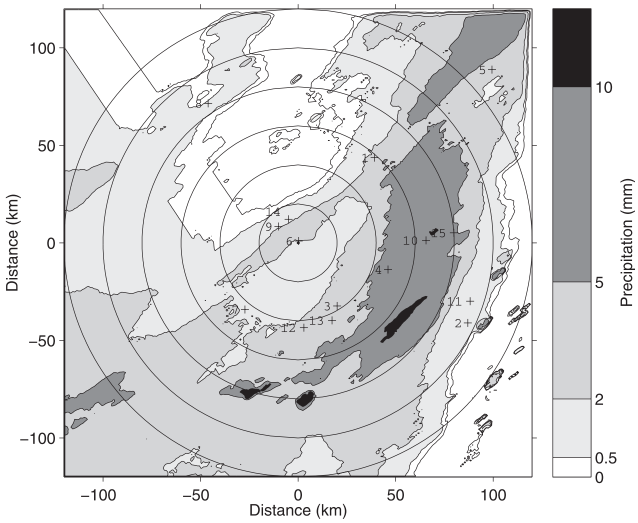

Figure 7 | CAPPI derived from the 2-rule adaptive neuro-fuzzy inference system: 17 June 1997 from 8 am to 9 am.

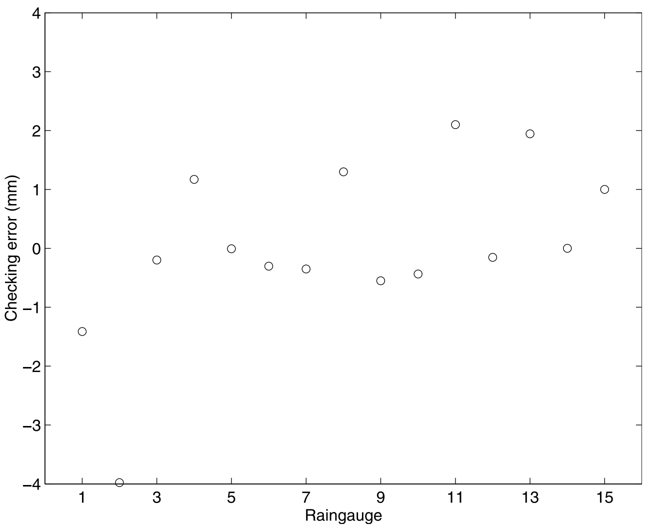

Figure 6 | Checking errors of the 2-rule adaptive neuro-fuzzy inference system: 17 June 1997 from 8 am to 9 am.

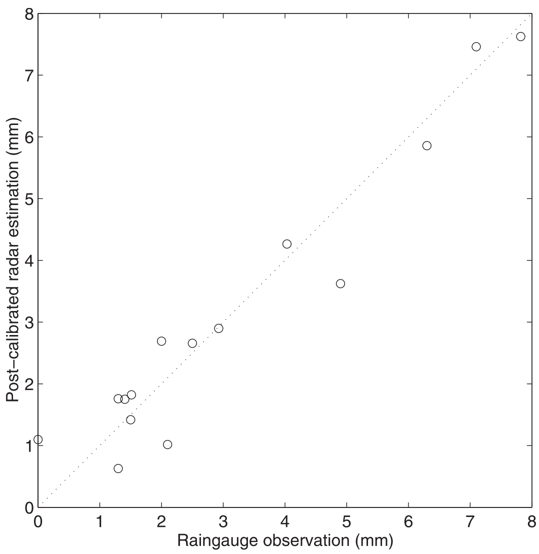

Figure 8 | Scatter plot of the raingauge observations and of the 2-rule adaptive neuro-fuzzy inference system: 17 June 1997 from 8 am to 9 am.

The post-calibrated CAPPI using the adaptive neuro-fuzzy inference system (Figure 7) is the combination (average) of the maps resulting from the 15 neuro-fuzzy inference systems (Equation (2)). One can see that it has retained most of the features of the original CAPPI (Figure 1). The good generalisation capability of the model allows a slow degradation of the performance when values away from the raingauges are post-calibrated. At the same time, raingauge observations are not forced during post-calibration, as shown by the scatter plot of the raingauge observations and the post-calibrated radar estimations (Figure 8). In fact, the adaptive neuro-fuzzy inference system leads to a compromise between both data sets–remember that the raingauge observations are also

Journal of Hydroinformatics | 05.1 | 2003

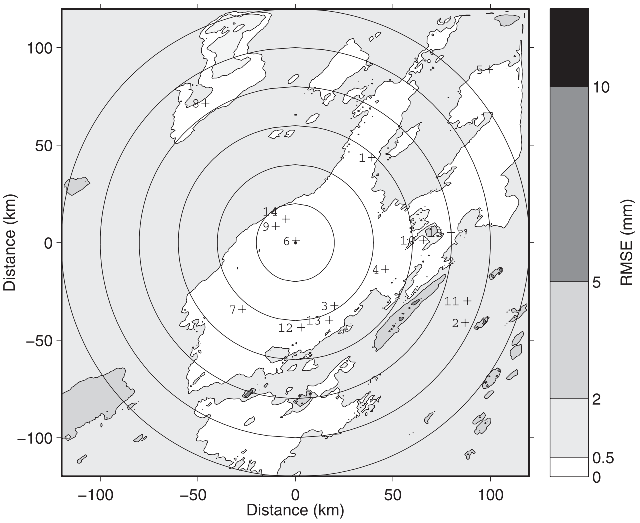

Figure 9 | *RMSE* derived from the 2-rule adaptive neuro-fuzzy inference system: 17 June 1997 from 8 am to 9 am.

Table 1 | Post-calibration results: 17 June 1997 from 5 am to 1 pm

	Radar vs raingauges		ANFIS vs raingauges	
Time	**Correlation coefficient**	**RMSE (mm)**	**Correlation coefficient**	**RMSE (mm)**
5:00 to 6:00	0.63	1.24	0.87	0.68
6:00 to 7:00	0.85	1.34	0.96	0.67
7:00 to 8:00	0.92	1.48	0.99	0.51
8:00 to 9:00	0.88	1.14	0.96	0.62
9:00 to 10:00	0.87	0.87	0.95	0.57
10:00 to 11:00	0.87	0.65	0.92	0.36
11:00 to 12:00	0.64	1.33	0.90	0.61
12:00 to 13:00	0.67	0.96	0.73	0.77

prone to errors. The *RMSE* error in Figure 8 is 0.62 mm, compared to 1.14 mm in Figure 2. Note that the post-calibrated value corresponding to raingauge 2 has been significantly adjusted.

The resulting *RMSE* map of the selected event is shown in Figure 9. It provides a degree of reliability for the estimation or, in other words, its repetitiveness from one inference model to the other. As expected, it can be seen that the errors increase with distance away from the raingauges. A large area where most raingauges are located shows *RMSE* values below 0.5 mm. Larger *RMSE* errors are associated with the most intense precipitations in the middle of the storm. It reflects the undeceiveness of the adaptive neuro-fuzzy inference system when large radar estimations are not supported by raingauge observations. Similar tests have been performed on other events with as satisfactory results.

Table 1 has summarised the results of the post-calibration of radar images for the complete event from 5 am to 1 pm of 17 June 1997.

DISCUSSION

Post-calibration of weather radar precipitation estimations based on an adaptive neuro-fuzzy inference system offers important advantages when compared to the other interpolation schemes for radar–raingauge data integration.

- The algorithm does not necessarily force the radar image to fit the raingauge measurements. The algorithm finds a response, which is a compromise between the radar rainfall estimations and the raingauge observations. This means that the algorithm does not necessarily consider raingauge data as truth.
- The training and the interpolation results can be obtained within just a few seconds using an ordinary personal computer which is incomparably faster than most interpolation methods, cokriging in particular. Therefore this algorithm would be very efficient for real-time post-calibration, especially when dealing with a sampling rate of one hour or less.
- Personalised rules for particular cases can be easily incorporated into such a system. For example, one may decide to force the post-calibrated CAPPI to zero whenever no precipitation is reported by the weather radar—although the capability of a weather radar to detect light rain diminishes with distance.

CONCLUSION

An adaptive neuro-fuzzy inference system, based on a jack-knife approach, has been presented and tested for the post-calibration of weather radar rainfall estimation. This methodology does not attempt to solve all the different flaws that typically hinder weather radar (viz. ground clutter, bright band, anomalous propagation, beam blockage and attenuation), but allows corrections compensating Z–R relationships, which are known to vary from event to event or even within a single storm.

The proposed adaptive neuro-fuzzy inference system offers the precision and learning capability of artificial neural networks combined with the advantages of fuzzy logic, namely flexibility and tolerance to imprecise data. Noisy information is easily incorporated into such a system, and rules are automatically extracted from a properly trained neuro-fuzzy inference system. This approach is flexible and works incomparably faster than geostatistical methods, making it suitable for real-time weather radar rainfall post-calibration objectives. The proposed approach allows the use of all the available data for training and checking the neuro-fuzzy inference system, and provides a degree of reliability of the post-calibration.

ACKNOWLEDGEMENTS

The first author acknowledges support by the Ministry of Science, Research and Technology of Iran. The second and third authors acknowledge funding obtained from the Natural Sciences and Engineering Research Council of Canada, the Ministère des Ressources naturelles du Québec, and the Fonds pour les chercheurs et l'aide à la recherche du Québec. Collaborations from Aldo Bellon and Isztar Zawadzki (J. S. Marshall Radar Observatory, McGill University) are also gratefully acknowledged.

REFERENCES

Anagnostou, E. N. & Krajewski, W. E. 1998 Calibration of the WSR-88D recipitation processing subsystem. *Weather Forecasting* **13**, 396–406.

Andrieu, H., Creutin, J. D., Delrieu, G. & Faure, D. 1997 Use of a weather radar for the hydrology of a mountainous area, Part I: Radar measurement interpretation. *J. Hydrol*. 193, 1–25.

Anile, A. M., Deodato, S. & Privitera, G. 1995 Implementing fuzzy arithmetic. *Fuzzy Set Syst.* **72**, 239–250.

Bandemer, H. & Gebhardt, A. 2000 Bayesian fuzzy kriging. *Fuzzy Set Syst.* **112**, 405–418.

Bardossy, A., Bogardi, I. & Kelly, W. E. 1990*a* Kriging with imprecise variograms I: Theory. *Math. Geol*. **22**(1), 63–79.

Bardossy, A., Bogardi, I. & Kelly, W. E. 1990*b* Kriging with imprecise variograms II: Application. *Math. Geol*. **22**(1), 81–93.

Bollivier, M., Dubois, G., Maignnan, M. & Kanevsky, M. 1997 Multilayer perceptron with local constraint as an emerging method in spatial data analysis. *Nucl. Instrum. Meth. Phys. Res.* **A389**, 226–229.

Chiu, S. 1994 Fuzzy model identification based on cluster estimation. *J. Intell. Fuzzy Syst.* **2**(3), 267–278.

Eddy, A. 1979 Objective analysis of convective scale rainfall using gauges and radar. *J. Hydrol*. **44**, 125–134.

French, M. N. & Krajewski, W. 1994 A model for real-time quantitative rainfall forecasting using remote sensing: 1 Formulation. *Wat. Res. Res.* **30**(4), 1075–1084.

French, M. N., Andrieu, H. & Krajewski, W. 1994 A model for real-time quantitative rainfall forecasting using remote sensing: 2 Case studies. *Wat. Res. Res.* **30**(4), 1085–1098.

Huffman, G., Adler, R., Rudolf, B., Schneider, U. & Keehn, P. 1995 Global precipitation estimates based on a technique for combining satellite-based estimates, rain gauge analysis and NWP model precipitation information *J. Clim.* **8**, 1284–1295.

Krajewski, W. F. 1987 Cokriging radar-rainfall and rain gage data. *J. Geophys. Res*. **92**(D8), 9571–9580.

Piotrowski, J. A., Bartels, F., Salski, A. & Schmidt, G. 1996 Geostatistical regionalization of glacial aquitard thickness in Northwestern Germany based on fuzzy kriging. *Math. Geol.* **28**(4), 437–452.

Rongrui, X. & Chandrasekar, V. 1997 Development of a neural network based algorithm for rainfall estimation from radar observations. *IEEE Trans. Geosci. Remote Sensing* **35**(1), 160–171.

Seo, D.-J., Krajewski, W. F. & Bowles, D. S. 1990 Stochastic interpolation of rainfall data from rain gauges and radar using cokriging: 1 Design of experiments. *Wat. Res. Res*. **26**(3), 469–477.

Serafin R. J. & Wilson, J. W. 2000 Operational weather radar in the United States: progress and opportunity. *Bull. Am. Meteorol. Soc*. **81**(3), 501–518.

Steiner, M., Smith, J. A., Burges, S. J., Alonso, C. V. & Darden, R. W. 1999 Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation. *Wat. Res. Res*. **35**(8), 2487–2503.

Sugeno, M. 1985 *Industrial Application of Fuzzy Control*. Elsevier, Amsterdam.

Wilson, J. W. 1970 Integration of radar and rain gauge data for improved rainfall measurement. *J. Appl. Meteorol*. **9**(3), 489–497.

Zadeh, L. A. 1965 Fuzzy sets. *Inform. Control* **8**, 338–353.

Zawadski, I. 1984. Factors affecting the precision of radar measurement of rain. *22nd Conf. on Radar Meteorol., Zurich, Preprints* pp. 251–256.