Neural Comput & Applic (1995)3:50–61

Neural Computing & Applications

Analysis of Radar Images for Rainfall Forecasting using Neural Networks

T. Denœux[1,2] and P. Rizand[3]

[1]Université de Technologie de Compiègne, URA CNRS Compiègne Cedex; [2]Lyonnaise des Eaux Dumez (LIAC); and [3]RHEA SA, Nanterre, France

This paper describes a new approach to the analysis of weather radar data for short-range rainfall forecasting based on a neural network model. This approach consists in extracting synthetic information from radar images using the approximation capabilities of multilayer neural networks. Each image in a sequence is approximated using a modified radial basis function network trained by a competitive mechanism. Prediction of the rain field evolution is performed by analysing and extrapolating the time series of weight values. This method has been compared to the conventional cross-correlation technique and the persistence method for three different rainfall events, showing significant improvement in 30 and 60 min ahead forecast accuracy.

Keywords: Radial basis function network; Competitive learning; Image processing; Meteorology; Weather radar; Rainfall forecasting

1. Introduction

The use of radar in meteorology dates back to World War II, when it was noticed that at wavelengths of 10 cm or less, reflections were occasionally obtained from meteorological targets. In the 1950s and 1960s, a better understanding of the physical phenomena involved in these reflections has been gained, allowing for the operational use of conventional radars for the detection and tracking of thunderstorms and cyclones [1,2]. With the availability of fast computers in the early 1970s, it later became possible to process large amounts of digitised data in real-time. Quantitative rainfall measurement became possible, as well as short-range forecasting, up to several hours ahead, of precipitation amounts. Weather radars are now routinely used in several countries for flood warning and water management in general. In parallel, work is going on to determine the accuracy of radar-derived precipitation data, and improve the forecasting performance [2–4].

In this paper, we present a new approach to the analysis of radar data for rainfall forecasting based on a neural network model. This approach consists in extracting synthetic information from radar images using the approximation capabilities of multilayer neural networks. Each image in a sequence is approximated using a modified radial basis function network trained by a competitive mechanism. Prediction of the rain field evolution is performed by analysing and extrapolating the time series of weight values.

The next section presents an overview of this approach, after a brief description of the principles of rainfall measurement and forecasting by radar. The competitive radial basis function network model is presented in Sect. 3, and its application to the approximation of radar images is described in Sect. 4. Finally, experimental results are given in Sect. 5.

2. Problem Description

2.1. Rainfall Measurement

For wavelengths greater than 3 cm, the scattering of electro-magnetic pulses by an assemblage of

Received for publication 14 September 1994.

Correspondence and offprint requests to: T. Denœux, Université de Technologie de Compiègne, U.R.A. C.N.R.S. 817 Heudiasyc, BP 649 F-60206 Compiègne cedex, France.

airborne liquid particles can be described by the following radar equation

$$P_r = C\frac{Z}{r^2} \tag{1}$$

where P_r is the average power in each echo pulse arriving at the receiver, C is a constant depending on the characteristics of the radar equipment, r is the range, and Z is the radar reflectivity factor. Z can be shown be depend on the 6th power of the diameters of precipitation particles. By making assumptions regarding the drop-size distribution and the fall speeds of different-sized drops, it is possible to relate the average value of Z to the rate of rainfall R using an equation of the form:

$$Z = aR^b \tag{2}$$

where a and b are constants. When Z is expressed in $\mathrm{mm}^6.\mathrm{m}^{-3}$ and R in $\mathrm{mm.h}^{-1}$, common values for a and b are $a = 200$ and $b = 1.6$. The resulting $Z - R$ relationship is known as the Marshall–Palmer law.

By rotating the radar antenna around a vertical axis with constant elevation, one obtains a 'snapshot' of the rain field over an area of several thousands km^2. Some examples of radar images corresponding to three different meteorological situations are shown in Fig. 1. These are 256 × 256 images with 16 reflectivity levels and a spatial resolution of 1 km. The situations of June 17 1991 and July 31 1991 are characterised by convective precipitation generated by a large number of small-scale structures (convective cells). The third situation (November 12 1991) is typical of stratiform precipitation caused by a large frontal rainband.

2.2. Rainfall Forecasting

The first radar-based rainfall forecasting systems relied on the statistical comparison of the rain fields in two successive images, using, for example, the cross-correlation function [5]. In this method, the cross-correlation coefficient is calculated for different displaced superpositions of the two images, and its maximum indicates the most probable displacement. The forecast is obtained by linear extrapolation of the displacement vector, applied to the whole image.

Although this simple technique performs quite well in frontal situations, characterised by large rainfall areas with uniform advection, it suffers from servere limitations when applied to other types of meteorological situations such as convective ones, where different rainfall areas may have different directions and speeds of motion. Since these situations are often those causing the most heavy rainfall, this drawback is particularly serious in hydrological applications. This remark has motivated research on more 'structured' approaches working with connected groups of pixels, or 'echoes', supposed to be associated with individual rainfall generating systems [3,6,7]. These approaches usually consist of four steps: (1) echo definition; (2) echo description by features; (3) echo matching; and (4) forecast by vector extrapolation. We have investigated [7] the use of a neural network for learning a matching rule whose function is to determine whether two echoes taken from successive images correspond to the same meteorological structure or not. The decision rule obtained has been found to yield a significant improvement over the state-of-the-art method introduced by Neumann [6], and has since been integrated into an operational system. However, this structured approach remains partly based on the three simplifying assumptions mentioned by Einfalt *et al.* [3]:

1. Uniform spatial behaviour of a rainfall structure (no split or merge).
2. Uniform temporal behaviour of a rainfall structure (no growth or decay).
3. Representation of a rainfall structure on the radar image by one echo.

Indeed, detailed analysis of the forecasting performance of the structured approach has revealed that futher improvements could be expected from the consideration of not only the *motion*, but also the *growth* and *decay* of rainfall areas [6]. Furthermore, the matching rules reported in [3,6,7] may fail in case of important deformation of rainfall areas from one time step to the other.

A previous attempt to apply connectionist algorithms to radar-based rainfall forecasting was reported recently [8]. In this approach, a three-layer backpropagation network receives as input the last image, and is trained to predict the rainfall field one hour ahead. The rainfall data used in this study were generated by a mathematical rainfall simulation model. The simulation domain was 100 × 100 km at a resolution of 4 km, yielding a grid of 25 × 25 points. The neural network was a three-layer backpropagation network with 625 input nodes, 625 output nodes and 15, 30, 45, 60 or 100 hidden nodes. The performance of the neural network was compared to the persistence method, in which the output is simply assumed to be identical to the input, and the cross-correlation method. According to the chosen performance criteria (comparison of mean intensity and areal coverage),

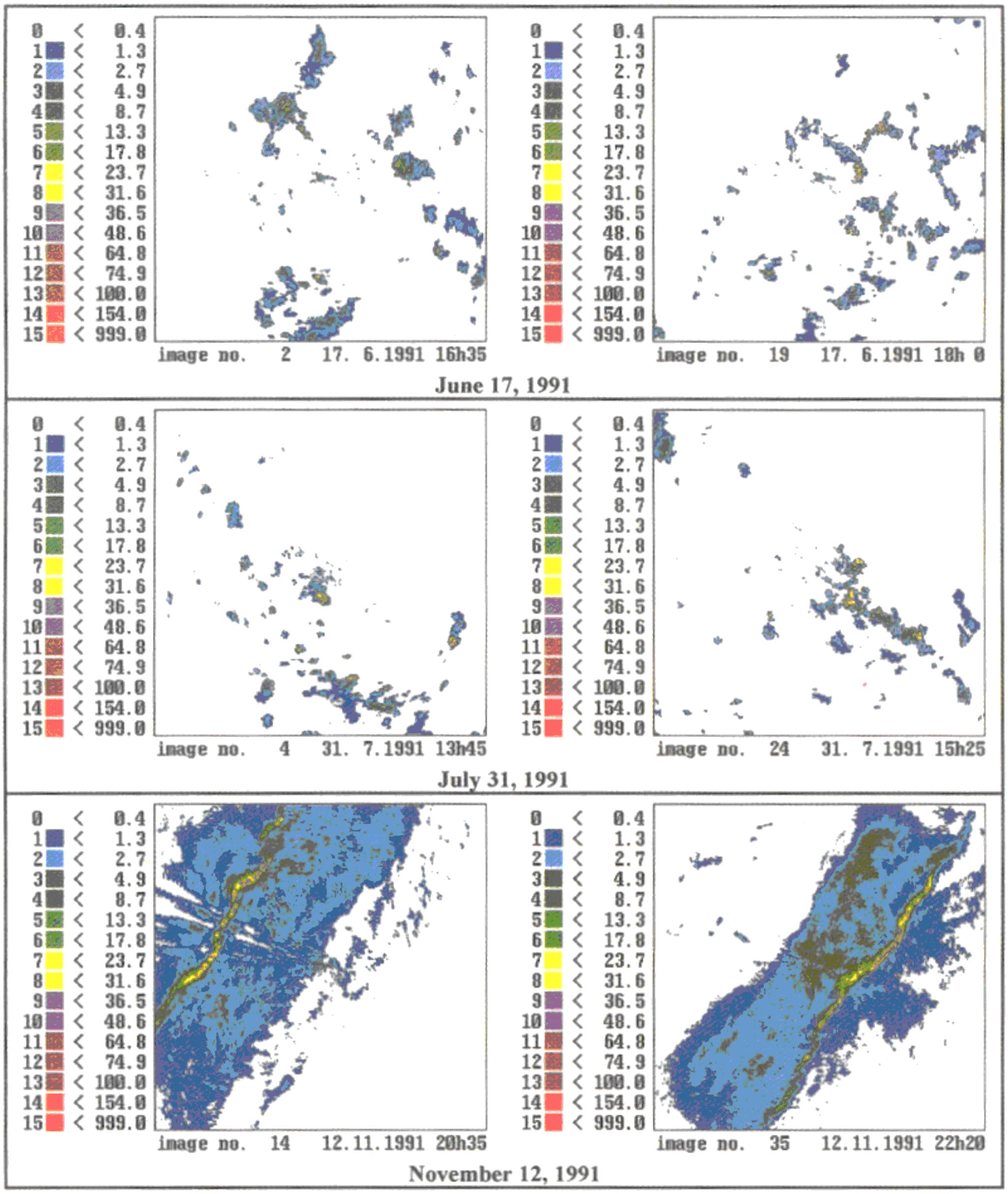

Fig. 1. Examples of radar images.

the study did not reveal any substantial improvement as compared to the reference methods. Moreover, this 'brute-force' approach was found to be very time-consuming (training times of up to 32 hours have been reported).

2.3. Proposed Approach

In this paper, we introduce a totally different approach intended to overcome the aforementioned limitations, making use of the approximation capabilities of artificial neural networks. The methodology is schematically depicted in Fig. 2.

The basic idea consists in coding a radar image using a small number of parameters. It is then expected that the analysis of the time series of parameter values corresponding to successive images will capture the main components of the rain field evolution, allowing for more accurate predictions than those resulting from simple translation of rainfall patterns.

To encode a radar image I_t at time t, of size $n \times n$, we propose to consider it as a function F_t: $\{1, \ldots, n\}^2 \rightarrow \Re$, such that for each $(i,j) \in \{1, \ldots, n\}^2$, $F_t(i,j)$ represents the intensity value of the pixel in image I_t at coordinates (i,j). Knowing the function F_t, it is possible to approximate it to any

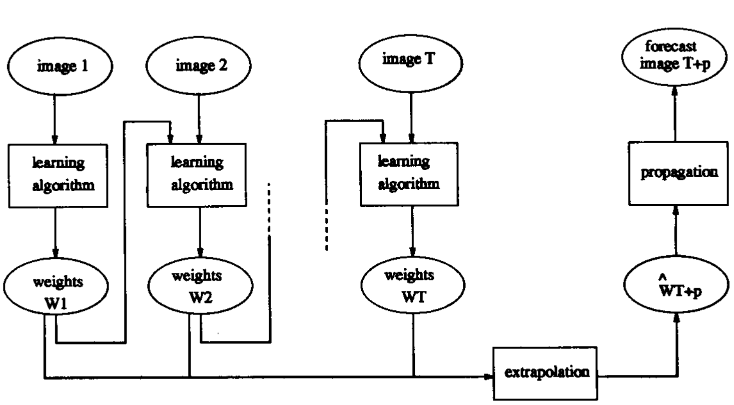

Fig. 2. The methodology.

accuracy using a multilayer feedforward neural network with two inputs, one output and a sufficient number of hidden units (Fig. 3). The associated weight vector $\mathbf{W}_t$ can be taken as a representation of the image I_t.

To preserve as much continuity as possible in the representations from one image to the next, the representation $\mathbf{W}_{t+1}$ of image I_{t+1} must be sufficiently 'close' to $\mathbf{W}_t$. This can be achieved by taking $\mathbf{W}_t$ as an initial guess for $\mathbf{W}_{t+1}$ before updating it by further training.

The neural network learning algorithm to be used for the approximation of F_t should meet the following requirements:

1. It should be fast enough to be usable in real-time.
2. It should be able to adapt its architecture to the complexity of the function F_t; this complexity may fluctuate due to the evolution of the meteorological situation (e.g. increase of the number of raincells due to convection intensification).
3. The class of functions implementable using a network architecture of 'moderate' size should contain sufficiently good approximations of the functions F_t corresponding to radar images, the shape of which is determined by the geometrical characteristics of rainfall structures.

As will be explained later, no existing learning algorithm has been found to be completely satisfactory according to all these criteria. For that reason, a new neural network paradigm, called the Competitive Gaussian Potential Function (CGPF) network has been developed for that application (see Sect. 3).

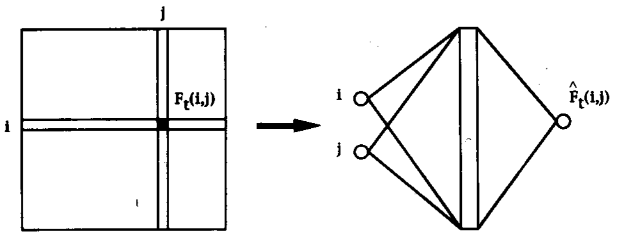

Fig. 3. Image encoding using a neural network.

As a result of the processing of a series of images I_t, $t = 1, \ldots, T$, a series of weight vectors $\mathbf{W}_t$, $t = 1, \ldots, T$ is available. However, due to the changes in the network architecture from one time step to the next, there is not necessarily a one-to-one correspondence between the components of successive weight vectors. For each component $W_t(i_k)$ of $\mathbf{W}_t$, one must consider the list $W_{T-q}(i_{k-q})$, ..., $W_{T-1}(i_{k-1})$ of its predecessors at previous time steps. This time series can be extrapolated to predict future values $\hat{\mathbf{W}}_{T+1}$, ... $\hat{\mathbf{W}}_{T+p}$ of the weights.

The learning algorithm and the prediction method will now be presented in greater detail.

3. Competitive Radial Basis Function Networks

3.1. Radial Basis Function Networks

Radial Basis Function (RBF) networks (also referred to as potential-function networks) are two-layer networks whose hidden layer is composed of units that respond to only a local region of the input space. The second layer performs a linear combination of the output signals from the hidden layer. The general form of the input-output mapping ϕ implemented by such a network is, for an input vector $\mathbf{x}$:

$$\phi(\mathbf{x}) = \sum_{i=1}^{M} w^i \, \psi(\mathbf{x},\mathbf{p}^i) + \theta \qquad (3)$$

where M is the number of hidden units, w^i the ith summation weight, $\mathbf{p}^i$ the parameter vector for unit

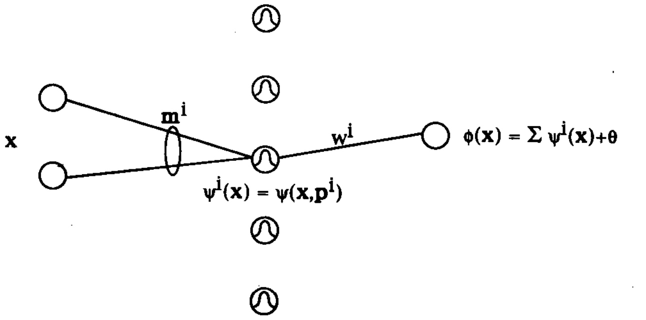
x
$\mathbf{m}^i$
w^i
$\phi(\mathbf{x}) = \Sigma\, \psi^i(\mathbf{x})+\theta$
$\psi^i(\mathbf{x}) = \psi(\mathbf{x},\mathbf{p}^i)$

Fig. 4. A radial basis function network.

i, ψ the activation function and θ a bias term (Fig. 4). In the most general case, $\phi(\mathbf{x})$ is a vector; however, it will be assumed to be a scalar in the following discussion.

Following Lee [9], we define a Gaussian potential function ψ^i by:

$$\psi^i(\mathbf{x}) = \psi(\mathbf{x},\mathbf{p}^i) \tag{4}$$

$$= \exp(-d_{\mathbf{K}^i}(\mathbf{x},\mathbf{m}^i)/2) \tag{5}$$

$$d_{\mathbf{K}^i}(\mathbf{x},\mathbf{m}^i) = (\mathbf{x}-\mathbf{m}^i)^t\mathbf{K}^i(\mathbf{x}-\mathbf{m}^i) \tag{6}$$

where $\mathbf{m}^i$ and $\mathbf{K}^i$ represent, respectively, the mean and shape matrix (defined as the inverse of the covariance matrix) of the potential function. The weight vector $\mathbf{m}^i$ can be interpreted as a prototype, or reference vector, in the input space. $d_{\mathbf{K}^i}(\mathbf{x},\mathbf{m}^i)$ is then the squared distance between the vectors $\mathbf{x}$ and $\mathbf{m}^i$, according the metric defined by $\mathbf{K}^i$.

The (j,k)th element of $\mathbf{K}^i$ can be expressed using the marginal standard deviations σ^i_j and σ^i_k and the correlation coefficient $h^i_{j,k}$:

$$k^i_{j,k} = \frac{h^i_{j,k}}{\sigma^i_j\, \sigma^i_k} \tag{7}$$

where $\sigma^i_j > 0$ and $h^i_{j,k} = 1$ if $j = k$, else $|h^i_{j,k}| \leq 1$. Note that $h^i_{j,k} = h^i_{k,j}$ for all i, j and k. In the case of a two-dimensional input space, we can therefore write $h^i_{1,2} = h^i_{2,1} = h^i$.

Gaussian potential-function networks meet relatively well the requirements listed in the former section:

1. Because the hidden unit weights receive a natural interpretation in terms of prototypes, they can easily be initialised better than randomly, resulting in faster training than, for example, back propagation networks with sigmoidal activation functions.
2. Relatively simple constructive algorithms are available for these networks [9, 10].
3. Because of the nature of rainfall-generating systems (essentially rainbands and convective cells), the rain field generally appears as the juxtaposition of more or less elongated 'bumps' that can be approximated using only a small number of Gaussian units.

3.2. Parameter Learning

For a fixed architecture, such a network can be trained in supervised mode by applying gradient-descent to an error function. Most commonly used is the quadratic error function, defined for the pth training pattern as:

$$E^p = \frac{1}{2}(t^p - \phi(\mathbf{x}^p))^2 \tag{8}$$

where t^p denotes the target value for the pth training pattern $\mathbf{x}^p$. The network weights are incrementally updated using the following learning rules [9]:

$$\Delta^p w^i = -\,\eta\frac{\partial E^p}{\partial w^i} = \eta(t^p - \phi(\mathbf{x}^p))\psi^i(\mathbf{x}^p) \tag{9}$$

$$\Delta^p m^i_j = -\eta\frac{\partial E^p}{\partial m^i_j} = \eta(t^p - \phi(\mathbf{x}^p))w^i\psi^i(\mathbf{x}^p) \sum_l k^i_{j,l}(x^p_l - m^i_l) \tag{10}$$

$$\Delta^p \sigma^i_j = -\eta\frac{\partial E^p}{\partial \sigma^i_j} = \eta(t^p - \phi(\mathbf{x}^p))w^i\psi^i(\mathbf{x}^p) \sum_l k^i_{j,l}\frac{(x^p_j - m^i_j)(x^p_l - m^i_l)}{\sigma^i_j} \tag{11}$$

$$\Delta^p \mathrm{h}^i_{j,k} = -\eta\frac{\partial \mathrm{E}^p}{\partial \mathrm{h}^i_{j,k}} = -\eta\frac{(\mathrm{x}^p_j - m^i_j)(\mathrm{x}^p_k - \mathrm{m}^i_k)}{\sigma^i_j\,\sigma^i_k} (t^p - \phi(\mathbf{x}^p))w^i\psi^i(\mathbf{x}^p) \tag{12}$$

$$\Delta^p \theta = -\eta\frac{\partial E^p}{\partial \theta} = -\eta(t^p - \phi(\mathbf{x}^p)) \tag{13}$$

where η is the learning step.

3.3. Architecture Adaptation

Platt [10] and Lee [9] have separately proposed similar algorithms for adapting the architecture of potential-function networks to the characteristics of the inputs. Basically, the method consists in combining gradient-descent with a mechanism for recruiting hidden units. Following an input presentation, a unit is recruited if the following conditions are satisfied:

1. The input pattern does not fall into the receive field of any hidden unit.

2. The output error for that pattern is greater than some threshold.

A new prototype is then created and initialised with the input vector. If the input falls within the accommodation boundary of an existing computational unit, or the error is small enough, then the network weights are simply updated using such learning rules as described in the previous section.

Initially, the receptive fields must be large enough to learn quickly a coarse representation of the input-output function. Their width is then gradually reduced for finer learning.

Although this learning technique offers good performance in terms of learning speed and compactness of representation, it presents some drawbacks, both in general terms, and specifically for our application:

- First, the parameter adaptation part of the learning rule is *global*, which means that all the weights are simultaneously updated to accommodate each input unit that does not trigger the allocation of a new unit. Since the representations formed by the network are essentially *local* (each unit responds to localised inputs), a possible alternative would be to use a *local* learning rule, which might result in faster training.
- Secondly, there is no mechanism for pruning the least relevant units; such a functionality is needed in our application, to account for the possible vanishing of rainfall-generating systems.
- Lastly, an additive output unit presents some drawbacks in our application, which can be shown by the following example. Let us consider an image at time t consisting of two echoes e_1 and e_2, as represented in Fig. 5, and let us assume that the predicted trajectories for these echoes converge at some point. From a physical point of view, it is not realistic to add up the rainfall intensities of the two echoes. A much better choice would be to consider at each point of the intersection the *maximum* of the contributions from e_1 and e_2.

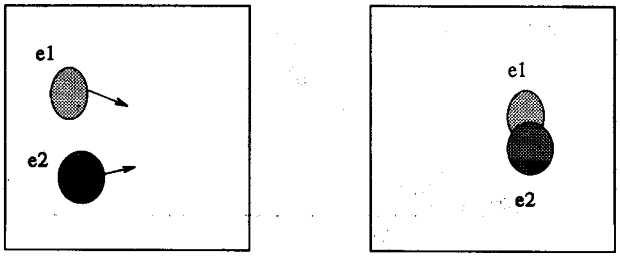

Fig. 5. Convergence of the trajectories of two echoes. The echoes e_1 and e_2 of the image at time t (left) merge into a single entity at time $t + \Delta t$ (right).

From these considerations, we have developed a new type of potential-function network, the Competitive Radial Basis Function (CRBF) network, that will now be described.

3.4. The Competitive Radial Basis Function Network

The CGPF network has the same architecture and hidden unit activation function as its non-competitive counterpart. The essential differences reside in its propagation equation and learning rule. Instead of Eq. 3, we now have:

$$\phi(\mathbf{x}) = \max_{i=1,\dots,M} w^i\psi^i(\mathbf{x}) + \theta^i \qquad (14)$$

where θ^i is a threshold for unit i and ψ^i is defined as before by Eqs. 4–6. Therefore, only the winning unit, defined as the hidden unit with the highest output signal for input $\mathbf{x}$, contributes to the network output. As in classical competitive learning, only this unit will be modified to accommodate the input. However, instead of simply moving the weights w^i in the direction of the input, as in the standard winner-take-all learning rule, all the parameters attached to unit i, i.e. w^i, $\mathbf{m}^i$, σ^i_k, $h^i_{j,k}$ and θ^i are updated using the rules described by Eqs. 9–13.

This simple adaptation mechanism can be combined with the same procedure for allocating units as described above. Moreover, we now have the possibility to prune a unit that has not been activated during a given number of presentations of the learning set.

We can now give a complete formulation of the learning algorithm. To simplify the notations, the input space will be assumed to be two-dimensional.

```
BEGIN
1. Initialise the first hidden unit with the first
   input vector $\mathbf{x}^1$:
   $M \leftarrow 1$; $\mathbf{m}^1 \leftarrow \mathbf{x}^1$; $\sigma^1_j \leftarrow \sigma_0$, $(j = 1,2)$
   $h^1 \leftarrow 0$; $\theta^1 \leftarrow 0$; $w^1 \leftarrow t^1$; $T \leftarrow 0$
   Initialise each individual learning rate at $\eta_0$
2. Repeat
     $it \leftarrow it + 1$
     For $p = 1$ to $P$/* number of training patterns
     */do
       compute $\phi(\mathbf{x}^p)$ according to Eq. 14
       $imax \leftarrow \arg\max_{i=1,\dots,M} w^i\psi^i(\mathbf{x}^p) + \theta^i$
         If $(d_{\mathbf{K}^{imax}}(\mathbf{x}^p, \mathbf{m}^{imax}) < \delta(it))$ or $(t^p -$
         $\phi(\mathbf{x}^p)) < \epsilon$ then
         update $w^{imax}$, $\mathbf{m}^{imax}$, $\sigma^{imax}_j$, $h^{imax}$ and $\theta^{imax}$
         using Eq. 9–13
       else
```

```
      create a new hidden unit
      $M \leftarrow M + 1$
      $\mathbf{m}^M \leftarrow \mathbf{x}^p$
      $\sigma_j^M \leftarrow \beta d_{\mathbf{K}}{}^{imax}(\mathbf{x}^p, m^{imax}), (j = 1,2)$
      $h^M \leftarrow 0$
      $\theta^M \leftarrow 0$
      $w^M \leftarrow t^p$
      endif
    end
    If a hidden unit has not fired during the
    last $L$ learning cycles, then
      delete it
      rename the indices accordingly
      $M \leftarrow M - 1$
    endif
    adapt the learning rates
  until convergence
END
```

σ_0, ϵ, L, β and η_0 are problem-dependent parameters. $\delta(it)$ is a time-decreasing parameter. In the simulations presented in the next section, $\delta(it)$ was made to decrease exponentially as a function of the number it of learning cycles:

$$\delta(it) = \max(\delta_0 \exp(-it/\tau), \delta_{min})$$

where δ_0, $\delta_{\min}$ and τ are constants.

Adaptation of the learning rates was performed using a modified version of the method described in [11].

4. Application

4.1. Learning Radar Images

Figures 6 to 8 show some typical results concerning the application of the algorithm described above to three radar images, corresponding to three different types of meteorological situations. The initial resolution of 256 × 256 pixels was lowered to 64 × 64 by averaging in order to reduce computing time. Reflectivity values were coded using 16 levels, as shown in Fig. 1, and converted to rainfall intensities using the Marshall–Palmer law.

Each image was smoothed using a median filter before applying the learning algorithm. The parameter values were defined as shown in Table 1. In our current implementation, the approximation of each image required about 1 min on a Sun Sparc 2 workstation.

As can be seen, the overall structure of the reflectivity patterns can be preserved using only 15 to 20 hidden units, i.e. approximately 100 to 150 parameters. The exact number of units and, correspondingly, the quality of the representation as measured by the mean squared error E obviously depend upon the tuning parameters of the learning algorithm. Our choice of parameter values has been found empirically to result in a good compromise between complexity on the one hand, and accuracy on the other hand. As a heuristic, it is interesting to compare E to the *discretisation error* E_d defined as:

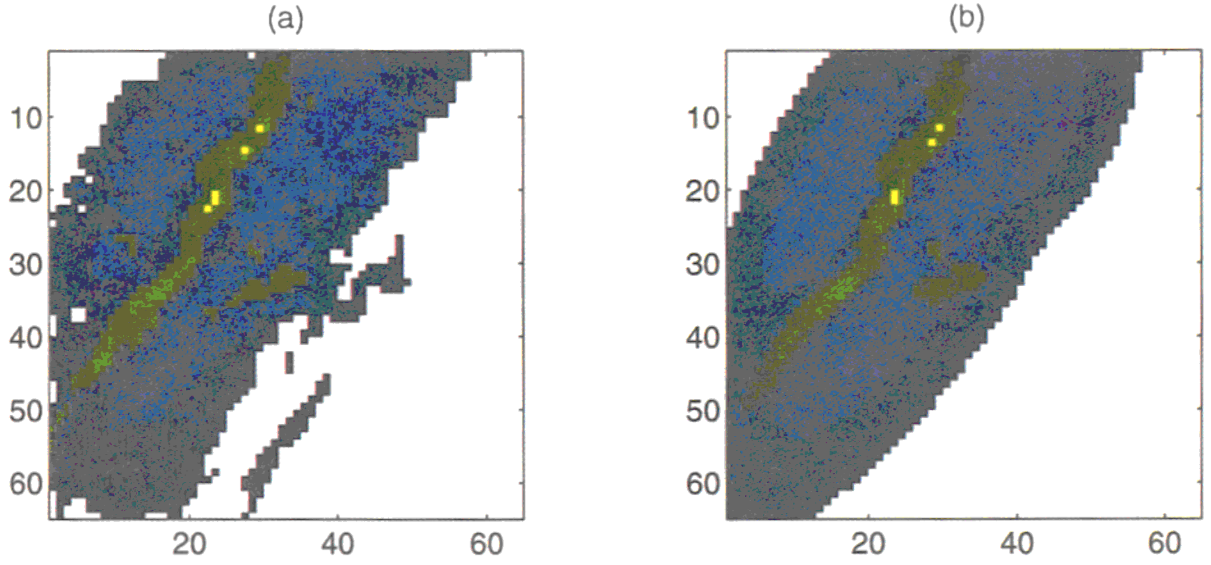

Fig. 6. Situation of November 12 1991. Original image after smoothing (a) and approximation (b) using 16 prototype units (E = 0.143; E_d = 0.39).

$$E_d = \sum_{i=1}^{n} \sum_{j=1}^{n} \left[\frac{R_{max}(N_{i,j}) - R_{min}(N_{i,j})}{2} \right]^2 \qquad (15)$$

where $N_{i,j}$ denotes the reflectivity level (between 0 and 15) of the pixel at coordinates (i,j) and $[R_{min}(N_{i,j}), R_{max}(N_{i,j})]$ represents the corresponding intensity interval. By convention, R_{max} (15) is taken as 207 mm/h.

Table 1. Parameter values.

σ_0	1	η_0	10^{-3}
ϵ	4	L	3
β	0.1	δ_0	40
δ_{min}	5	τ	30

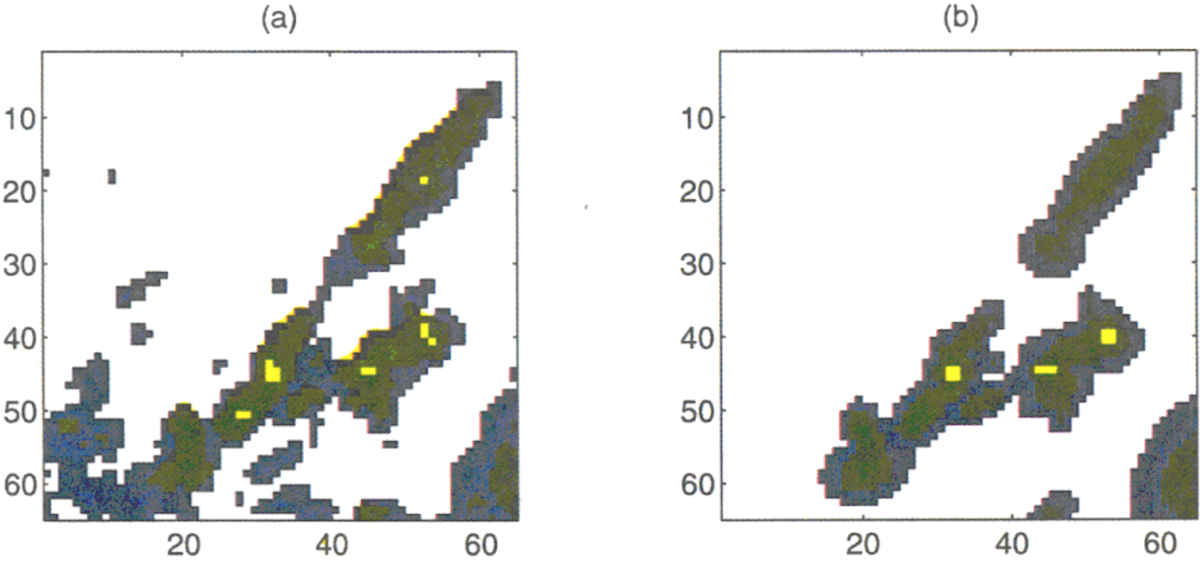

Fig. 7. Situation of August 7 1991. Original image after smoothing (a) and approximation (b) using 20 prototype units (E = 0.255; E_d = 0.408).

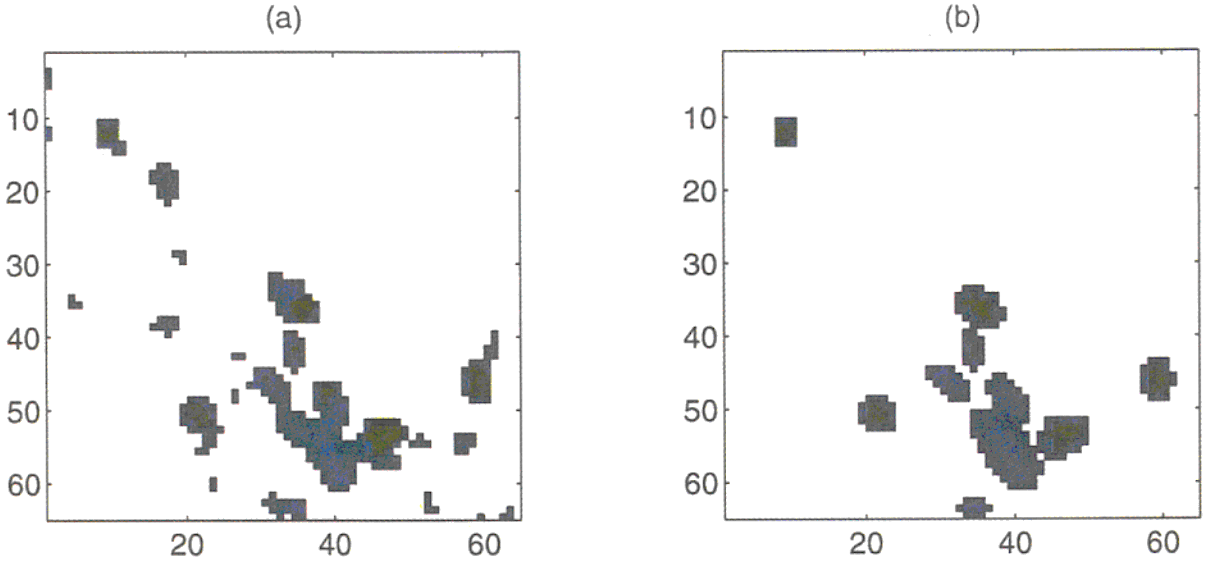

Fig. 8. Situation of July 31 1991. Original image after smoothing (left) and approximation (b) using 15 prototype units (E = 0.023; E_d = 0.083).

The discretisation error can be interpreted as an estimate of the mean squared error caused by the discretisation of reflectivity values, assuming the intensity to be uniformly distributed inside the intervals defined by R_{min} and R_{max}. It therefore provides a basis for estimating the error E needed to achieve a sufficiently good representation of the image. Empirically, good results have been obtained with $E \simeq 0.5E_d$.

4.2. Extrapolation of Connection Weights

By iteratively applying the learning procedure described above to successive images, taking the weight vector at time t as an initial guess for the weight vector at time $t + 1$, one obtains time series for each of the network weights $(m^i_j, \sigma^i_j, w^i, h^i, \theta^i), i = 1, \ldots, M; j = 1,2$. To extrapolate these series, some prior knowledge regarding the possible behaviour of each corresponding variable is necessary.

Assuming each hidden unit i to approximate the intensity distribution of one well-defined rainfall area, $\mathbf{m}^i$ can be interpreted as the centre of that area, whose shape (size, orientation, etc.) is described by the other parameters. According to that interpretation, one can expect to observe a quasi linear evolution for $\mathbf{m}^i$, corresponding to the generally linear advection of rainfall areas.

The evolution of the other network parameters such as σ^i_1, σ^i_2 and w^i is more difficult to predict. In convective situations, meteorological studies concerning the life-cycle of radar echoes [12] suggest that one could observe the same kind of 'bell-shaped' curves as observed for the size of rain cells.

As will be seen in the next section, experimental facts support these assumptions reasonably well, particularly concerning the linear displacement of the centres $\mathbf{m}^i$. This suggests that, given enough data, some extrapolation technique could be applied to the prediction of σ^i_1, σ^i_2 and w^i. However, in the present study, we have restricted ourselves to the linear extrapolation of the Gaussian potential centres $\mathbf{m}^i$ using a displacement vector computed by a moving average of the past observed displacements.

5. Experimental Results

5.1. The Data

The performances of our method and the cross-correlation method (hereafter referred to as NN and CC, respectively) have been compared for three typical meteorological situations respectively of frontal type (November 12 1991) and convective type (June 17 and July 31 1991). Sample images from these three situations are represented in Fig. 1.

The frontal situation is characterised by a narrow rainband embedded inside a larger area of stratiform precipitation, with uniform eastward advection and good stability of the shape of rainfall areas. In the two convective situations, much smaller echoes corresponding to convective cells are scattered all over the region covered by the radar. On June 17 1991 a heterogeneous advection with counter-clockwise rotation around some point near the centre of the image is clearly discernible on computer animations. The situation of July 31 1991 is characterised by a very slow advection with important deformation of reflectivity patterns and a tendency for rain cells to aggregate into an organised line of precipitation.

Figures 9 and 10 show the Gaussian potential centres and their predicted displacement for the next 30 and 60 minutes, respectively. As can be seen, both the uniform advection in the frontal situation and the heterogeneous advection of rain cells in the convective one are well captured by our procedure. The evolution of the network weights for three different hidden units in all three situations tested is represented in Figs 11 to 13. Not surprisingly, the weights are more stable during the frontal sitaution, whereas the stronger variability of

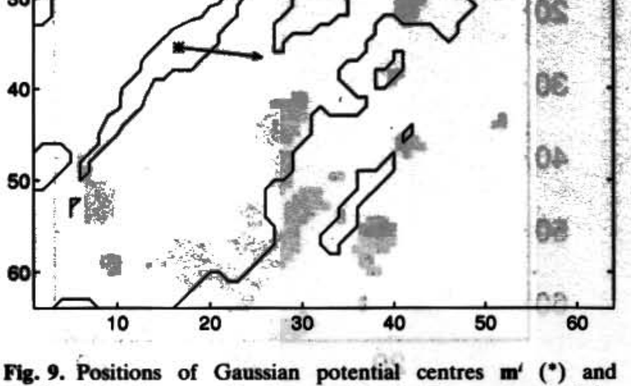

Fig. 9. Positions of Gaussian potential centres $\mathbf{m}^i$ (*) and predicted displacement for the next 30 min. The contours of the image correspond to the 1.3 mm/h and 4.9 mm/h levels (situation of November 12 1991).

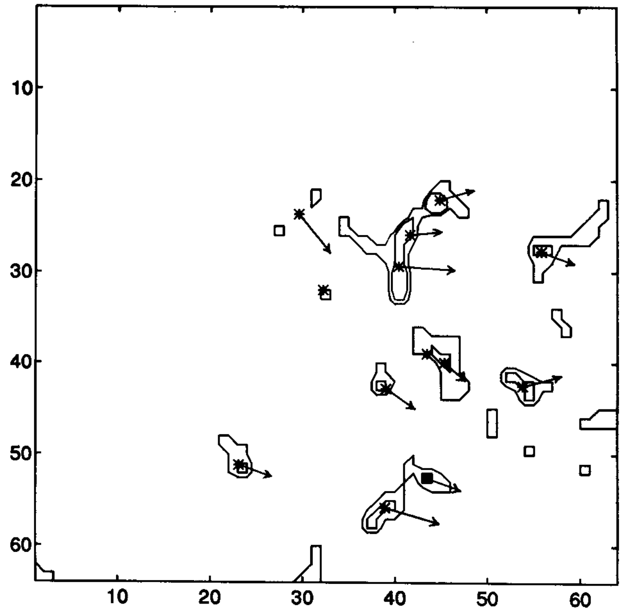

Fig. 10. Positions of Gaussian potential centres $\mathbf{m}^i$ (*) and predicted displacement for the next 60 min. The contours of the image correspond to the 1.3 mm/h and 2.7 mm/h levels (situation of June 17 1991).

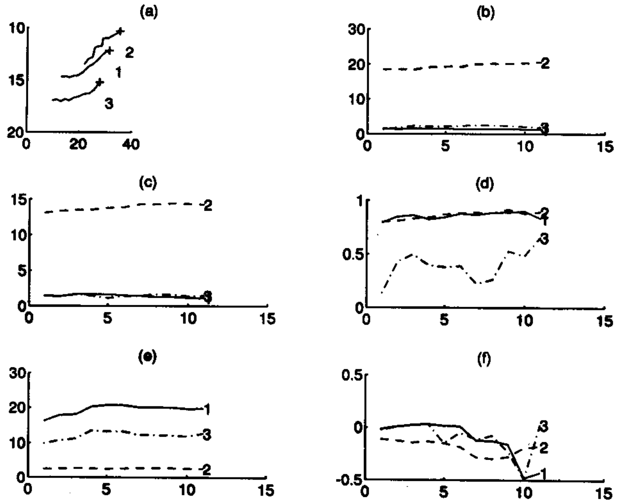

Fig. 11. Evolution of the weights for three hidden units (situation of November 12 1991). (a) Trajectories of Gaussian potential centres $\mathbf{m}^i$; (b) σ_1^i; (c) σ_2^i; (d) h^i; (e) w^i; (f) θ^i.

reflectivity patterns in the two convective situations results in greater weight variability.

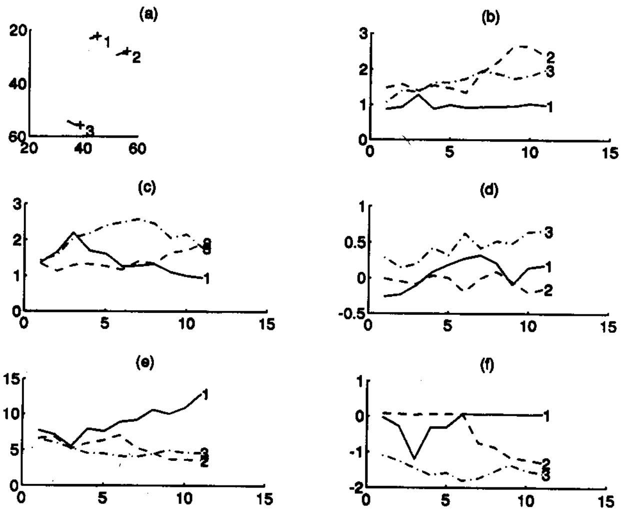

Fig. 12. Evolution of the weights for three hidden units (situation of June 17 1991). (a) Trajectories of Gaussian potential centres $\mathbf{m}^i$; (b) σ_1^i; (c) σ_2^i; (d) h^i; (e) w^i; (f) θ^i.

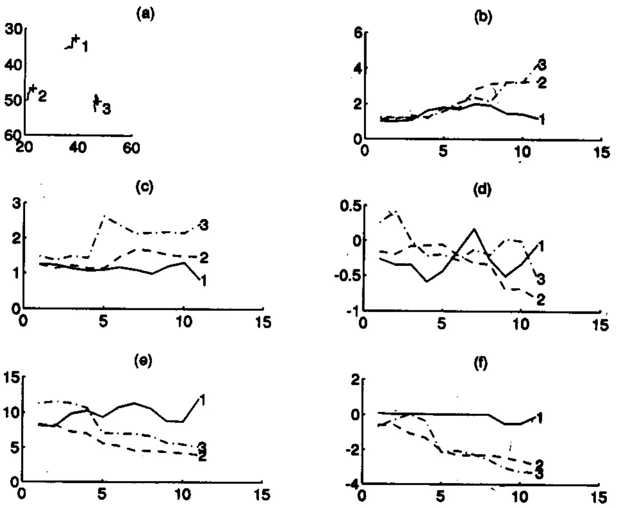

Fig. 13. Evolution of the weights for three hidden units (situation of July 31 1991). (a) Trajectories of Gaussian potential centres $\mathbf{m}^i$; (b) σ_1^i; (c) σ_2^i; (d) h^i; (e) w^i; (f) θ^i.

5.2. Results

Our forecast evaluation procedure has attempted to reflect as much as possible the use of quantitative rainfall forecasts in an operational context. The accumulated rainfall depth for the next 30 and 60 minutes was computed over the 4 × 4 km^2 areas corresponding to each of the 4096 pixels, using both forecasting methods (CC and NN). Predictions of rainfall depths were obtained by computing forecast images at 5 min intervals, and summing up the coresponding intensity values. In addition to CC and NN, we also considered for comparison the performance of the 'persistence' method (Pers.), which simply consists in taking the last observed image as the prediction for future time steps.

These predicted values were compared to estimates of the same quantities, computed in a similar manner using radar measurements performed in the same period. The sparseness of available rain-gauges did not permit the use of ground measurements for calibrating the radar data. However, the errors so introduced can reasonably be expected to affect only the absolute performance values, and not the relative values that only matter in this comparative study [4].

The mean forecasting errors for the three situations at five successive time steps are represented in Figs. 14 to 16. As expected, the CC and NN methods both outperform the trivial persistence method, and the overall best performance was obtained with the NN approach. The superiority of the NN method increases with range from 30 to 60 min, which suggests that it is actually due to a better estimation of the speed and direction of motion of rainfall patterns.

6. Conclusions

A new rainfall forecasting method based on the representation of radar images by radial basis function networks has been presented. For each incoming image, a neural network is trained so as to approximate the rain field, using the previous weights as an initial state. The learning algorithm is based on a competitive mechanism, combined with a procedure for adapting the network size to the complexity of the input distribution. Forecast images are obtained by extrapolating the time series of weight values.

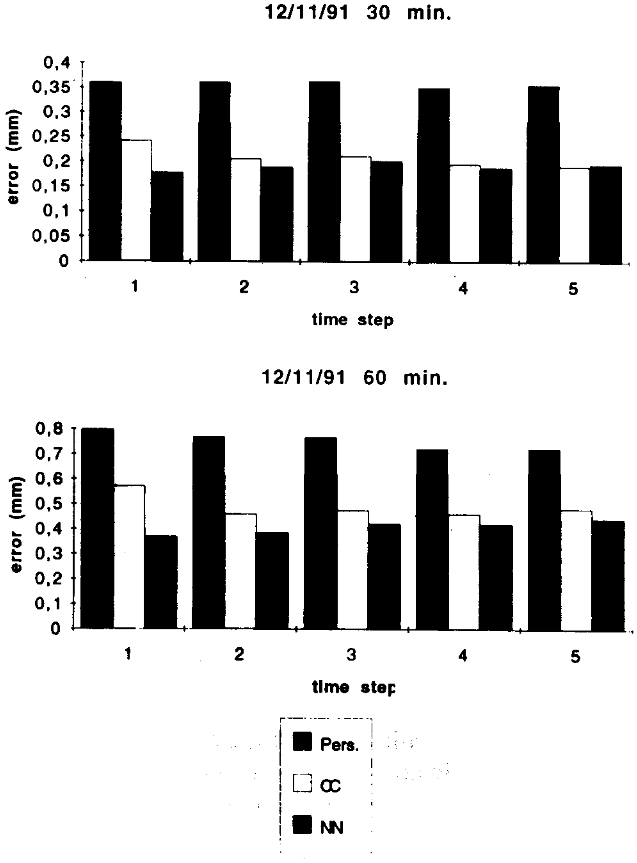

Fig. 14. Mean forecasting errors (situation of November 12 1991).

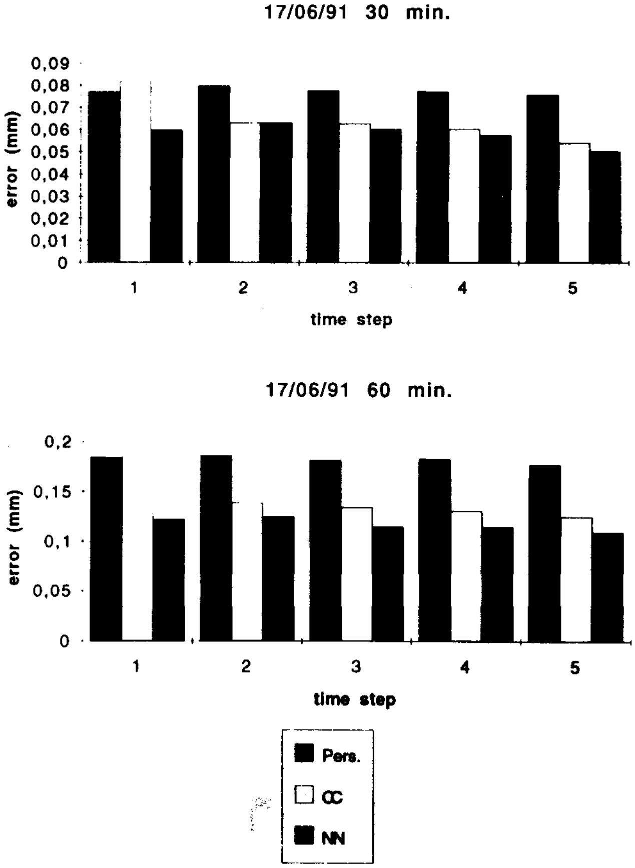

Fig. 15. Mean forecasting errors (situation of June 17 1991).

The ability of this method to predict accumulated rainfall volumes on small-size areas was evaluated for three rainfall events corresponding to different types of meteorological situations. For each of these events, significant gains in prediction accuracy were obtained, as compared to the standard cross-correlation method. Additional improvements are still expected from the analysis of a greater number of situations, that will allow both the experimentation of more sophisticated extrapolation techniques and the refinement of heuristics for tuning the method.

Although more experiments in various meteorological situations are still needed to complete the validation of this approach, the results obtained so far are considered as very encouraging. Our current work aims at carrying out further testing and refinement so as to convert the present prototype into a fully operational system.

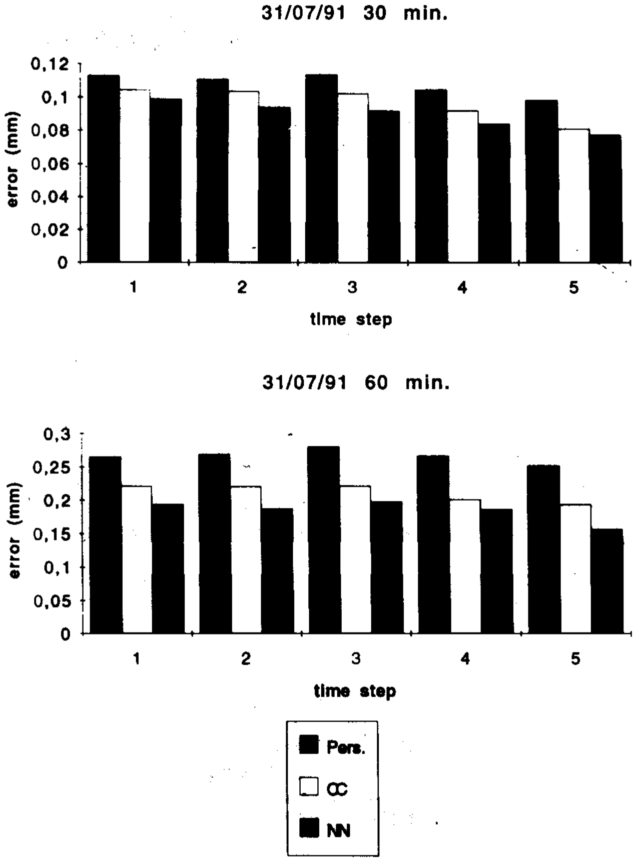

Fig. 16. Mean forecasting errors (situation of July 31 1991).

Acknowledgements. This work has been supported by EEC funded Esprit II project nr. 5433 (NEUFODI); partners: BIKIT, ARIAI, Elorduy Sancho y Cia, LABEIN, Lyonnaise des Eaux-Dumez, RHEA S.A.

References

1. Battan LJ. Radar Observation of the Atmosphere. University of Chicago Press, 1973
2. Clift GA. Use of radar in meteorology. Technical Report 181, World Meteorological Organization, 1985
3. Einfalt T, Denœux T, Jacquet G. A radar rainfall forecasting method designed for hydrological purposes. J Hydrology 1990; 114: 229–244
4. Denœux T, Einfalt T, Jacquet G. Determination in real time of the reliability of radar rainfall forecasts. J Hydrology 1991; 122: 353–371
5. Austin GL, Bellon A. The use of digital weather radar records for short-term precipitation forecasting. Quart J Roy Meteorological Soc 1974; 100: 658–664
6. Neumann A. Introduction d'outils de l'Intelligence Artificielle dans la prévision de pluie par radar (In French). PhD thesis, Ecole Nationale des Ponts et Chaussées, Paris, 1991
7. Ding X, Denœux T, Helloco F. Tracking rain cells in radar images using multilayer neural networks. In S Gielen and B Kappen, editors, Proc ICANN'93, pp 962–967. Springer-Verlag, London, 1993
8. French MN, Krajewski WF, Cuykendall RR. Rainfall forecasting in space and time using a neural network. J Hydrology 1992; 137: 1–31
9. Lee S. Supervised learning with gaussian potentials. In B Kosko, editor, Neural Networks for Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1992 pp 189–227
10. Platt JC. Learning by combining memorization and gradient descent. In RP Lippman, JE Moody, DS Touretzky, editors, Neural Information Processing 3, Morgan Kaufmann, San Mateo, CA, 1991, pp. 714–720
11. Silva FM, Almeida LB. Speeding up back-propagation. In R. Eckmiller, editor, Advanced Neural Computers, North-Holland, Amsterdam, 1990, 151–158
12. Lopez RE *et al.* Population characteristics, development processes and structure of radar echoes in south Florida. Monthly Weather Rev 1984; 112: 56–75