Radar based rainfall forecast for sewage systems control

H. Aspegren*, C. Bailly**, A. Mpé**, N. Bazzurro***, A. Morgavi***, E. Prem**** and N.E. Jensen†
* Malmö Water, S-205 80 Malmö, Sweden
** Lyonnaise des Eaux 59, avenue Emile Thiébaut 78110 Le Vésinet, France
*** AMGA SpA, Via S.S. Giacomo e Filippo, 7, 16122 Genova, Italy
**** Austrian Research Institute for Artificial Intelligence OFAI, Schottengase 3, A-1010 Vienna, Austria
†DHI Water and Environment, Agern Allé 1, DK-2970 Hørsholm, Denmark

Abstract There has been an increasing demand for accurate rainfall forecast in urban areas from the water industry. Current forecasting systems provided mainly by meteorological offices are based on large-scale prediction and are not well suited for this application. In order to devise a system especially designed for the dynamic management of a sewerage system the “RADAR” project was launched. The idea of this project was to provide a short-term small-scale prediction of rain based on radar images. The prediction methodology combines two methods. An extrapolation method based on a sophisticated cross correlation of images is optimised by a neural network technique. Three different application sites in Europe have been used to validate the system.
Keywords Radar; rain prediction; stormwater; drainage network

Introduction

Old sewerage systems are typically of the combined type, meaning that stormwater is mixed with wastewater. In case of rain, the hydraulic load on the treatment system increases dramatically and the sewerage system becomes a bottleneck in the transport of rainwater. This in turn leads to numerous adverse effects such as inundation, decreased treatment efficiency, pollution overflow, etc. If there had been knowledge about the flow in advance some of these adverse effects could have been attenuated or avoided. For this reason there has been an increasing demand for accurate rainfall forecasts in urban areas from the water utilities. Current forecasting systems provided mainly by meteorological offices are based on large-scale prediction, i.e. meteorological models based on worldwide evolution. These systems can provide a qualitative estimation of rainfall water on a limited area, such as a city but are very weak in the quantitative forecasts although this feature is essential for dynamic management of urban networks.

In order to devise a system especially designed for the dynamic management of a sewerage system the “RAD” project was launched in 1997. The idea of this project was to provide a short-term small-scale prediction of rain based on radar images. These predictions can be used as the entry point for the control and management of a drainage and sewage system. The system has been designed to meet the requirements of wastewater utilities and has been validated by three application sites across Europe.

The RADAR project

The RADAR project – ESPRIT 23475 – has been co-financed by the European Commission and was developed in-between the years 1997–1999 by a consortium of three water utilities, Malmö Water (Sweden), Lyonnaise des Eaux (France) and Azienda Mediterranea Gas e Acqua (Italy), the Danish Meteorological Institute DMI (Denmark), DHI Water and Environment (Denmark), the Austrian Research Institute for Artificial Intelligence OFAI (Austria) and the software and GIS company APIC (France). The

Water Science and Technology Vol 43 No 5 pp 79–86 © 2001 IWA Publishing and the authors

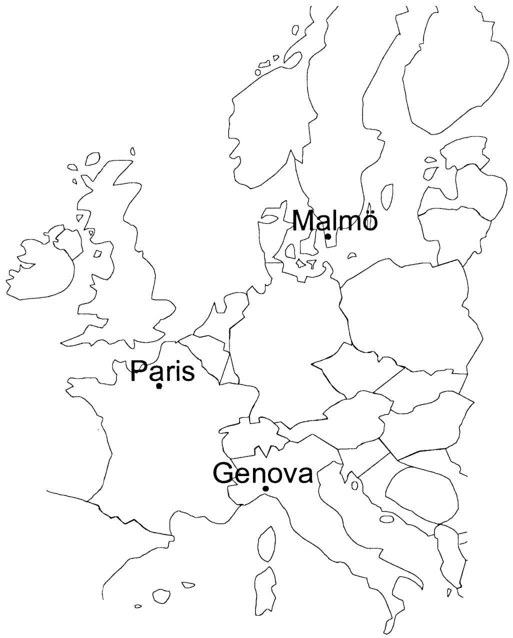

Figure 1 Test sites

knowledge and ability of al partners put together made it possible to design, develop and validate a relevant solution to the addressed problem.

The general approach to rainfall prediction in the RADAR project was designed with the following objectives in mind. The method should be optimised with respect to the needs of the water industry needs. Secondly, it should work robustly and continually with a system that is permanently on-line and does not need intervention by a human supervisor. Thirdly, the method should outperform classic prediction methods in this area (Andersson and Andersson, 1992). In order to evaluate the RADAR system, possible applications were also selected at the three test sites, Genoa, Malmö and Paris (Figure 1).

In Genoa the forecasting system is applied to the urban drainage system of the historical centre which consists of streams flowing in culverts conveying wastewater to a WWTP through a coastal trunk sewer and pumping stations. The system is used in Genoa in order to feed a quality/quantity model of the urban drainage system, which gives information regarding the flow rate amount and the pollution load related to the first foul flush in the urban drainage system.

In Malmö a prediction of the wastewater flow can be used to improve the overall strategy for in-line storage of stormwater and simultaneous equalisation of the diurnal dry weather flow in the Klagshamn wastewater system.

In Paris, the RADAR system will be used at a treatment plant especially designed for treatment of stormwater from the Orly airport. The dry-weather flow to this plant is rather small and a rainfall in the airport area will require a rapid reaction from the treatment plant operator. The forecast will be used to issue an alert if the rainfall will generate a flow that requires a mode shift at the treatment plant.

The RADAR software architecture

The basis for the RADAR system is time series of radar images collected in real time. These images come from a weather radar tuned to measure the reflectivity of raindrops in and below clouds. The radar provides a new image every 5 or 10 minutes. Different processes are then needed to achieve the prediction. The system is data driven. This means that a new processing cycle is launched by the arrival of new data. A task scheduler controls the arrival of information and the workflow of data through the different modules of the system.

In the pre-processing module each image is pre-processed to extract the reflectivity caused by raindrops from false reflections caused by buildings, hills or clutter and abnormal propagation. The various errors affecting the accuracy of derived radar precipitation come from the basic radar calibration, e.g. sensitivity and detection efficiency of radar hardware and the radar data, i.e. attenuation, clutter, anaprop, occultation, bright-band, vertical reflectivity profile variations, range, orographic growth etc.

The reflectivity measured by the radar is interpreted in terms of rain intensity. The general form of the relationship is given by the formula $Z = k\,R^n$ (Z = reflectivity, R = rainfall). The parameters (k, n) depend on the type of rain (drizzle, widespread rain or storm) and in order to get a more accurate evaluation of the intensity, rain gauge data are used to calibrate the constants on-line in a special calibration module.

In the prediction module the rain forecast is computed from a series of pre-processed images and results in a series of (future) images corresponding to a two hour window ahead. The predicted rain intensity on the watershed or on a specific location is then extracted from the images.

Depending on the requirements of the application site, the observed rain and the predicted rain are used to forecast the resulting flow to the treatment plant, to help human decision making or to automatically control pumping and gate devices.

The prediction method

An extensive investigation showed that certain scientifically promising methods, exhibit unsatisfactory results for a system targeted to a professional market (OFAI, 1998). For example none of the tested basis function techniques or the Gaussian mixture model were able to work robustly or to consistently deliver the promised results. The desired robustness together with a potential for an increase in performance may, however, be gained with a step-step approach. In the first step, a robust prediction is generated by a more conventional approach based on the extrapolation of rain pattern movements according to a vector field. In a second step, a neural network based model helps to adapt the system more tightly to a specific application site and thus capture peculiarities of an application, e.g. topology, local meteorology, etc.

There are several different ways of estimating the velocity:

1. *Global cross-correlation (GCC)*: In this classical method a movement vector is calculated from two consecutive images based on cross-correlation. The result is a velocity vector that is globally applied all over the image to extrapolate the movement into the future.
2. *Rain tracking (RT)*: As the movement in the image may be differnet for different image locations, or even rotary. An improvement to GCC is to divide the image into parts. Then a vector can be calculated for each of these parts. The desired vector field can then be obtained by using spatial interpolation techniques.
3. *Proximity cross-correlation (PCC)*: Another approach to overcome the limitations of GCC focuses on the patterns in the radar image rather than using a blind grid structure. The idea here is to search for a connected pattern in the image. These patterns can be thought of as rain cells. These cells are further put in rectangles. By means of cross-correlating the rectangles, a vector field is gained which is likely to be much more sensitive to the movement of single cells present in a radar measurement.

In order to select the best method an evaluation of the vector-based prediction methods had to be performed. Due to the large amount of data, it was decided to only test the 3×3

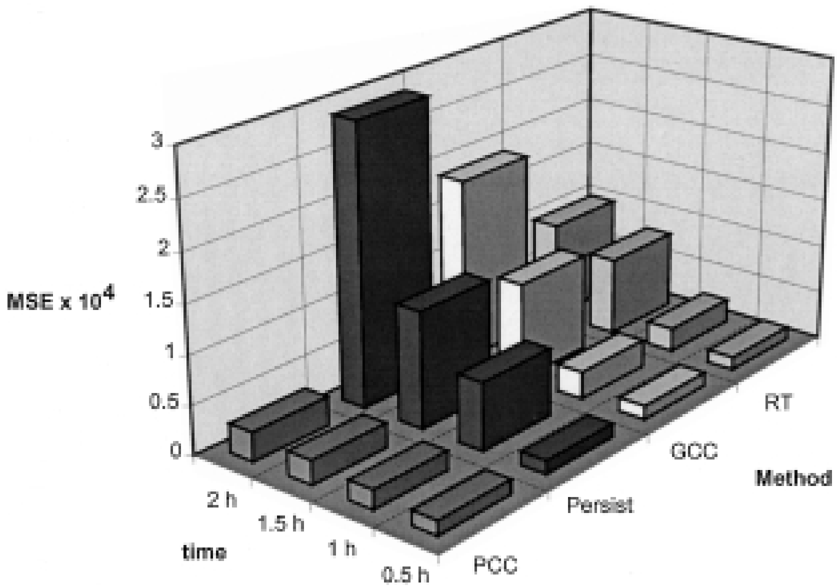

Figure 2 Mean squared error of the methods

version of rain-tracking method and to compare it with GCC, PCC, and for reasons of plausibility checking also to persistence (Persist). This method comprises a minimum criterion, which every more sophisticated prediction method should be able to beat. Here it is assumed that there is no movement in the image at all.

The results of the application of the three different methods to prediction are highly dependent on the weather type present in the image. In cases of extensive rain such as in typical frontal situations, all the methods (except persistence) seem to perform equally well. This is not surprising, since in these weather situations small errors in the vector field have no big consequences and the field itself will most likely always look very coherent. Indeed, in some cases the very simple cross-correlation does a good job. The situation is, however, different for thunderstorm situations. An example of a good result in the Malmö case is shown below. The results correspond to data from June 9, 1996.

The figure (Figure 2) shows the mean squared error (in 1000 mm/h) summed up to 30, 60, 90 and 120 min. In the example given here, PCC clearly outperforms all other techniques. However, the situation is not always as clear as in this example and sometimes a different prediction technique, especially RT, may be better than PCC.

Therefore, it was decided to first use PCC as the vector-based method, which is improved in a second trap by the neural network. The second step in the overall RADAR rain prediction consists in the application of a neural network to the prediction as generated in the first step. The network can be regarded as an error model that is trained to minimise the summed squared error of the real radar measurement and the radar value as predicted from the vector-based method. The neural network is trained to the specific characteristics of the radar at an application site. Being optimised with respect to single pixel values in the image, the network may capture geographical characteristics as well.

The network uses the forecasted images. The actual inputs to the network, however, concern only a subset of pixels covering the catchment area. Thus the system is capable of learning systematic deviations as they may be introduced through geographical peculiarities (e.g. mountains or winds).

The results showed the expected improvement of the summed squared error. The following figure (Figure 3) explains the results in more detail. The first picture shows a

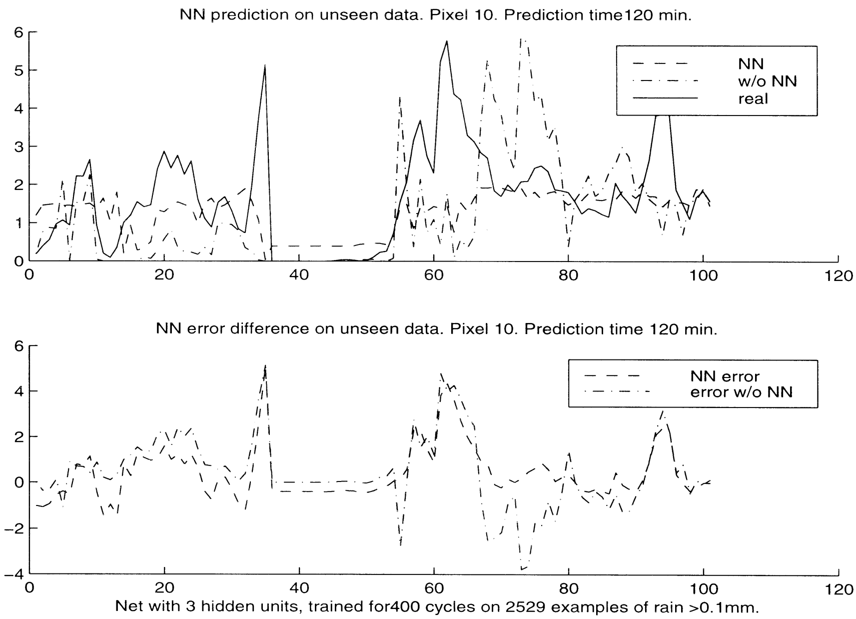

Figure 3 Example of the prediction with and without neural network (NN) on a sequence of 100 data points

randomly selected example of the generated prediction sequence, the real rainfall and the errors associated with it. The upper half of the image shows a sequence of 100 data points that exhibit significant rainfalls. The solid line represents the real rainfall as measured by the radar. The dashed line represents the prediction corrected by the neural network and the dash-dotted line represents the prediction as delivered by the proximity cross correlation algorithm. The lower half of the image shows the error of both systems.

As can be seen in the images, a clear assessment of these time-series is relatively difficult. In order to judge the results more easily, the sums of squared errors were computed for all prediction times (from 10 minutes to two hours in ten minute intervals). Most of the time, the method using a neural network is better than without the network. There are only a few exceptions. However, the overall trend is clear and it shows the desired reduction in error.

Validation

From a mathematical point of view, the RADAR prediction system proved to outperform other known methods. However, the usefulness of a rain prediction also depends on how this prediction is used by the sewage operators. The impact of a given rain (depth, duration) on a sewage network and on a treatment plant varies from one site to another. The benefits of a forecasting system depend on the storage capacity of the sewage system and on the identified operations to minimise the effect of the rain (closing valves, starting pumps etc.). The usefulness of a prediction system is consequently related to the accuracy of the prediction and to the effectiveness of the control strategy adopted by the sewage system operator.

There is a need for an indicator of the quality of prediction that can be interpreted in terms of operational benefits by the sewage operators. The mean square error and the Nash criterion (used for hydrologic models) are classical criteria. But they are not directly connected to the needs of the users. One feature of these criteria is that errors for low intensities have little importance compared to errors on high intensities. What is more problematic is that time shift may generate a great error. If the prediction is just shifted one time step from the reality, the prediction could still from a practical point of view be useful despite that the mean square error is high.

Criteria based on the occurrence of rain were therefore preferred. In this system a prediction is defined as correct if it forecasts rain when it rains (hit_yes) and no rain when it does not rain (hit_no). Furthermore the prediction is defined as incorrect if it forecasts rain when there is no rain (false_alarm) and no rain when it actually rains (miss). These criteria are summarised below:

hit frequency = (hit_no + hit_yes) / total number

false alarm frequency = false_alarm / (false_alarm + hit_yes)

miss frequency = miss / (miss + hit_yes)

The "hit frequency" is representative of the ability of the system to predict a rain of significant importance. It is not very different from a classical criterion. The "false alarm frequency" and the "miss frequency" are also of great interest for the sewage operation. They give an idea of how reliable the prediction is.

In the following analysis, a rain event is defined as a rain exceeding one of the three thresholds of 0.1, 2 or 5 mm/h. The hit frequency, false alarm frequency and miss frequency were calculated for the different methods. The hit frequency remains high (above 90%) in all the cases, which is due to the fact that the data set naturally contains a relatively large number of "no rain" cases. The false alarm frequency reaches the 50% rate after 40 minutes in the 2 mm case and after only 10 minutes in the 5 min situation. However, due to the very small number of heavy rains, only the 2 mm case is significant. This means that there is a relatively large chance to wrongly predict a medium rainfall after 40 minutes in the opposite case, for misses, the corresponding value is 50 minutes for the 2 mm rains and 20 minutes for the 5 mm rains. For rainfalls below a threshold of 0.1 mm, however, these frequencies remain below 50% up to a prediction of 2 hours.

A predictive accuracy of 50% false alarms after 40 minutes may not seem too good for the efforts involved. However, it is important to understand that the figures derive from a point-by-point comparison of true and predicted rainfall time series. This means that a rainfall of a given intensity which is predicted, e.g. 6 minutes earlier than it really happens, will count as a "miss" in this classification.

In order to relax this exact matching requirement of prediction and real rainfall time series, it is necessary to compute moving averages. In the following analysis the average of fixed-time intervals, i.e. 30 or 60 minutes, were calculated and treated as a rain event in the same sense as in the previous analysis. Practically this means that some flexibility is allowed for with respect to the timing of the prediction. The prediction can, e.g. by 10 minutes too early, but the moving average would still correctly predict a rain event. This is a very realistic quality criterion with respect to the application.

As a result, the false alarm and miss frequency both remain below 50% in the NN-PCC method up to a prediction time of 60 minutes for the 2 mm threshold. In the 5 mm case, the 50% miss rate is reached after 30–40 minutes, while the 50% false alarm rate lies at 20 minutes.

As a second result of this analysis, it turned out that the difference between the PCC and NN-PCC methods tends to disappear when looking at moving averages of different time intervals. This indicates that the neural network basically learned to perform exactly such a moving average of values. At a 20 minutes moving average time span, the RT3×3 and the PCC method both outperformed the NN-PCC; the RT3×3 and PCC methods are of comparable quality and show differences only on small and heavy rains.

Figure 4 Low-cost radar in Genoa

Low-cost radar

In Genoa, radar images were not readily available and instead a low-cost radar was developed by DEMI and DHI (Figure 4). The new low-cost radar is based on a commercial marine radar system using X-band at 9.41 GHz and 6 kW output power.

Tests with this first prototype have shown that the system is able to detect relatively weak precipitation (0.5 mm/h) at a distance of 25 km allowing a typical forecast time of 2 hours. The obtained images covers a smaller area than a conventional meteorological radar but the resolution is higher (up to 200 m), hence it is suitable for small catchment areas.

The system allows the collection of customisable information data, taking into account different boundary conditions in terms of catchment size and final purposes to be reached.

Conclusions

Some new methods for rain prediction have been designed, implemented and tested. The result is a system that provides a short term (up to two hours), quantitative prediction for one or several spots. It is aimed at helping drainage and sewage system control. The performance is promising. Three application sites must confirm the interest of this prediction for the water business.

A computer architecture and an embedded package "RADAR" has been designed to efficiently use the rain prediction methods in a plant or control room environment. It uses a real time link with a radar image provider and real time rain gauge measurements. The prediction can help a human operator to make a decision or can be used as an input to an automated control system.

The RADAR prediction system is used to optimise the operation of retention tanks, to optimise the management of the wastewater treatment plant and to minimise urban flood risks.

References

Andersson, T. and Andersson, B. (1992). Radar rain forecasts for sewerage systems. *2nd Int. Symp. On Hydrological Applications of Weather Radar*. Univ. Hannover, 7–10 September.

Bailly, C., Corbierre, V., Duchesne, H. and Labaquere (1996). A real-time decision aid for storm water treatment. *Proc. Int. Conf. on Hydroinformatics*, September 11–13, Zurich, Switzerland.

Denoeux, T. and Rizand, P. (1995). Analysis of radar images for rainfall forecasting using neural networks. *Neural Computing and Applications.*

Ding, X., Denoeux, T. and Helloco, F. (1993). Tracking Rain Cells in Radar Images Using Multilayer Neural Networks. *Proc. IJCANN 1993.*. S. Gielen and B. Kappan (eds). Springer-Verlag.

Nyberg, U., Andersson, B. and Aspegren, H. (1996). Real time control for minimizing effluent concentrations during storm water events. *Wat. Sci. Tech.*, **34**(3–4), 127–134.

OFAI (1998). *Identification of Selected Methods*, WP 400 task 430–D430.1. OFAI/RADAR/TN 001/05–98 v1.0, High Performance Rainfall Radar Image Processing for Sewage Systems Control, P. 23475.

Vejen, F., Allerup, P., Overgaard, S. and Illingworth, A.J. A review of radar correction and adjustment techniques. *R and D Note National Rivers Authority, UK.*