RAPIDS: Early Warning System for Urban Flooding and Water Quality Hazards

Andrew P Duncan[1], **Albert S Chen**[1], **Edward C Keedwell**[1], **Slobodan Djordjević**[1] **and Dragan A Savić**[1]

Abstract. This paper describes the application of Artificial Neural Networks (ANNs) as Data Driven Models (DDMs) to predict urban flooding in real-time based on weather radar and/or raingauge rainfall data. A time-lagged ANN is configured for prediction of flooding at sewerage nodes and outfalls based on input parameters including rainfall. In the absence of observed flood data, a hydrodynamic simulator may be used to predict flooding surcharge levels at nodes of interest in sewer networks and thus provide the target data for training and testing the ANN. The model, once trained, acts as a rapid surrogate for the hydrodynamic simulator and can thus be used as part of an urban flooding Early Warning System (EWS). Predicted rainfall over the catchment is required as input, to extend prediction times to operationally useful levels. Both flood-level analogue and flood-severity classification schemes are implemented. An initial case study using Keighley, W Yorks, UK demonstrated proof-of-concept. Three further case studies for UK cities of different sizes explore issues of soil-moisture, early operation of pumps as flood-mitigation/prevention strategy and spatially variable rainfall. We investigate the use of ANNs for nowcasting of rainfall based on the relationship between radar data and recorded rainfall history; a feature extraction scheme is described. This would allow the two ANNs to be cascaded to predict flooding in real-time based on current weather radar Quantitative Precipitation Estimates (QPE). We also briefly describe the extension of this methodology to Bathing Water Quality (BWQ) prediction.

Keywords. ANN, early warning system, flood risk, machine learning, neural network, nowcasting, prediction, rainfall, urban flood.

1 INTRODUCTION

Recent studies [1], [2] have documented the increased frequency and likelihood of extreme precipitation events. In the UK, the existing installed base of combined drainage systems is huge. This means that a large proportion of urban rainfall runoff is immediately mixed with effluent, increasing the potential public health risks from urban flooding. Even flooding from separate storm sewers is in any case destructive and costly. An ageing network and increasing urbanisation further exacerbate these problems. Therefore models are required, which can provide predictions of location, severity and/or risk of flooding. In order to be operationally useful, these need to provide 2+ hour lead-time [3] and be able to operate rapidly in real-time.

Hydrodynamic simulators are used as standard to model the response of Urban Drainage Networks (UDNs) to rainfall events. However, especially for large UDNs, these can be slow and computationally expensive. A faster surrogate method is sought, which would permit modelling of very large networks in real-time, without unacceptable degradation of accuracy. However, if actual rainfall is used as input, the predictive ability of such models is limited by the Time of Concentration (ToC) for the sewer network, with the possibility of flooding at any node commencing from zero time onwards, following the start of precipitation. In practice, ToC would normally be of the order of minutes, rather than hours for all but the downstream sections of the very largest UDNs.

Therefore prediction of rainfall is a requirement to achieve the lead-times sought. Many papers have been written on rainfall nowcasting methods from radar rainfall images [3–11]. A novel machine-learning based approach to this is currently at an early stage of development within the Centre for Water Systems.

2 APPROACH USED ('RAPIDS')

As part of University of Exeter’s research under Work Package 3.6 of the Flood Risk Management Research Consortium Phase 2 (FRMRC2) [12] project , we developed the ‘RAdar Pluvial flooding Identification for Drainage System’ (RAPIDS) using ANN’s to predict flooding in sewer systems. This was described in our paper [13] and was further developed for an UKWIR-funded joint industry / University of Exeter project [14] in which three case studies were carried out for UDN's in South London, Portsmouth and Dorchester, with promising results.

The RAPIDS software (currently in MATLAB) includes two programs: RAPIDS1, which addresses the need for a faster surrogate for hydrodynamic simulators as well as classifier models for flood and other hydrological parameters, and RAPIDS2 (under development), which aims to provide nowcasting for rainfall over the catchment containing the modelled UDN. It is hoped to be able to demonstrate the cascading of these two systems to provide the required urban flood predictive model.

The RAPIDS1 program is based on a lagged-input, 2-layer, feedforward Artificial Neural Network (ANN), used to relate incoming rainfall data to the extent of flooding present at each node in the UDN. It has the same number of output neurons as sewerage nodes of interest – i.e. there is no requirement to model nodes identified from hydrodynamic modelling as never flooding, making an immediate computational saving. The ANN architecture is varied to establish an optimum. The supervised training regime uses either backpropagation of error quasi-Newton gradient-descent or NSGA-II [15] Evolutionary Algorithm method. A moving time-window approach is implemented whereby lagged time-series signals (e.g. rainfall intensity, cumulative rainfall, soil moisture, pump states, tidal levels etc) are provided in parallel over the time-window as inputs to the ANN. If no direct observation data is available for the UDN to be modelled, output target signals for training and evaluation of ANN model performance are provided from the flood-level, volume or flow hydrographs generated by

[1] Centre for Water Systems, University of Exeter, Harrison Building, North Park Road, Exeter EX4 4QF, UK. Email: {apd209, A.S.Chen, E.C.Keedwell, S.Djordjevic, D.Savic}@exeter.ac.uk

hydrodynamic simulator outputs for each sewerage node to be modelled. This only needs to be done for the training dataset of rainfall events. The trained ANN thus aims to generate the same hydrographs for new rainfall events as would the UDN itself, based on having learned and generalised the (non-linear) relationship between the provided input signals and observed or simulator-generated targets. Figure 1 illustrates the architecture of the RAPIDS1 system to predict sewer network outputs. The target signals selected are the flood levels at each sewerage node at a time-step that corresponds to the desired prediction lead-time (i.e. up to network ToC).

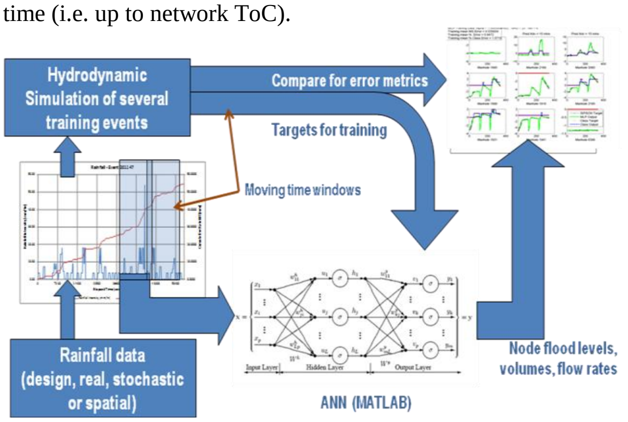

Figure 1. Architecture of RAPIDS1

Event profile data arrays of the input-signals are prepared for use as the time-series input to the ANN as illustrated in Figure 2. In line with best practice, all input data are normalised.

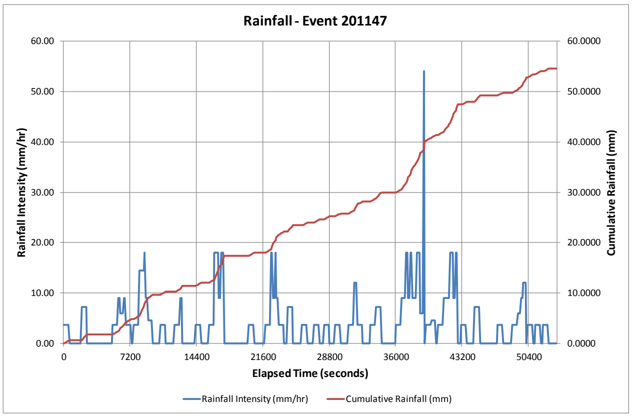

Figure 2. Selected ANN Input signals for a typical rainfall event

A selection of (historic) rainfall events is needed for the training dataset. These need to be representative of the envelope of likely intensities and rainfall totals for the future events to be modelled. If sufficient of these are not available, existing events can be augmented by factorally increasing rainfall intensity and modelling resulting target hydrographs using a hydrodynamic simulator.

Rainfall radar images are sourced from the UK Met Office NIMROD system [16], [17], which produces a composite 1km resolution Quantitative Precipitation Estimate (QPE) image covering the whole UK, every 5-minutes. A live RSS feed is available on request. Historic data images (from April 2004 to present) are available for download from [18]. Treatment of radar QPE images 1km pixel-by-pixel by an ANN is computationally prohibitive since, for example, for a 3-h prediction there would be 36-images, each with at least 3602-pixels (allowing for a maximum storm advection velocity of 60 km/h). This would potentially require ~5 × 10^6 neurons (at 1-neuron per pixel). Therefore features are extracted from the rain echoes in each time-step and associated with features from previous time-steps. These can then be applied to the inputs of an ANN as time-series signals. The feature extraction approach proposed is similar to Discrete Wavelet Transforms (DWT) using Haar wavelets [19], but using different sized grids depending on the proximity to the catchment being modelled. The mean rainfall for the whole area is evaluated; then residuals of mean rainfall over each sub-grid square are computed: see Figure 3. Standard deviations show that information is contained at all spatial scales [20].

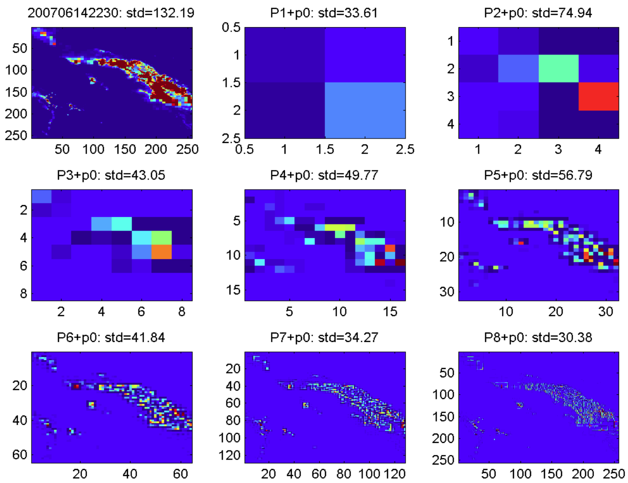

Figure 3. RAPIDS2: Rainfall Event 2007-06-14 – QPE snapshot at 22:30 showing original image (top left) and feature extraction of residuals at finer grid resolutions (128 to 1 km)

The extracted residuals from multiple images over the duration of each event become time-series signals, which can be applied as input signals to ANNs: see Figure 4.

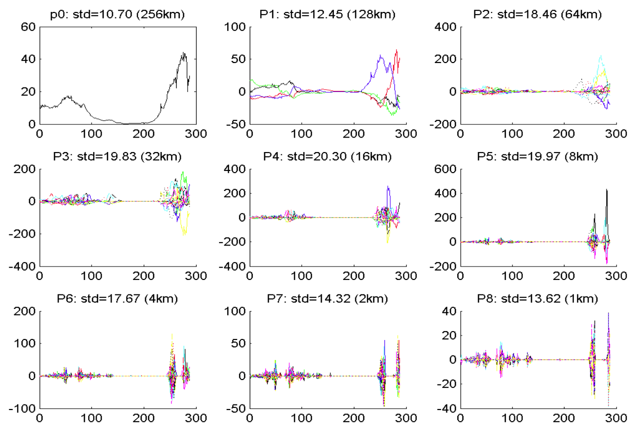

Figure 4. RAPIDS2: Rainfall Event 2007-06-14 – Time-series ANN input signals over 24-hours at spatial resolutions as shown; x-axis is radar image no.; y-axis is Δ rainfall intensity in mm/hr.

It is proposed to implement a similar time-windowed ANN framework as for RAPIDS1. Target rainfall for training and evaluating the ANN is derived from the rainfall intensities in grid squares covering the required catchment containing the UDN to be modelled, advected into the future by the required prediction period.

In summary, the proposed methodology is to cascade the two stages together (RAPIDS2 providing predicted rainfall, which can be applied to RAPIDS1 inputs) and thus provide flood predictions for each node of interest in the UDN, hopefully with operationally useful lead-times of 2+ hours.

3 CASE STUDIES

An initial "proof-of-concept" case study for RAPIDS1 was conducted as part of FRMRC2. An ANN with 123-outputs was used to model the Stockbridge sub-section of the combined rain/wastewater drainage system for the town of Keighley, West Yorkshire, containing 122 manholes and one combined sewer overflow (CSO). Design rainfall was used. The neural network gave a floating-point estimate of the level of flooding at each node. However, this level of accuracy is unlikely to be required for flood-warnings. Therefore a classification scheme to provide predictions of flood severity was implemented by post-processing ANN outputs. Results were reported in [13].

Under the UKWIR-funded joint-industry Real-time Machine Learning (RTM) project [14] the following 3 case studies were implemented, in a two-stage project to evaluate effectiveness in different sized catchments under different conditions; stage 1 used design rainfall and stage 2 used real rainfall:

Dorchester: small urban catchment (6km^2); evaluation of the significance of use of soil moisture as ANN input.

Portsmouth: medium urban catchment (30km^2); island location; tidal effects; need for pumping; evaluation of effectiveness of ANN models to provide early starting of pumps – as a flood-mitigation / prevention strategy.

Crossness (South London): large urban catchment (230km^2); evaluation of model effectiveness using spatially varying rainfall as ANN inputs.

In order to allow all partners to present results consistently, the MS Excel-based 'HydroMAT' model analysis tool was developed to provide automated assessment of ANN output using a number of metrics[2] including those recommended in [21]. Results below (Figures 7-9) were assessed using this tool.

4 RESULTS & DISCUSSIONS

Figure 5 shows average ANN training times of around 115 seconds for the 123-node network used in the FRMRC case-study. Fifteen 6-hour events (rainfall + runoff) were used for training. In comparison, hydrodynamic simulation for each took approx 240 seconds (total 3600 seconds). Once the ANN was trained, however, test run times were of the order of 0.1 seconds for each 6-hour event (Figure 6). Figures 7-9 illustrate the reporting of metrics provided by the HydroMAT tool; Figure 7 shows a typical spread of NSEC values over a 20-node sample for a single test rainfall event; Figure 8 compares ANN-generated hydrograph with the target hydrograph for a single node for a single test event; Figure 9 shows flood severity classification matrix for peak flood depths for a 20-node sample for a single event. This compares target classifications (rows A to C) with ANN-generated classifications (columns A to C). It also shows a colour-coded assessment of 3 'Accuracy bands'.

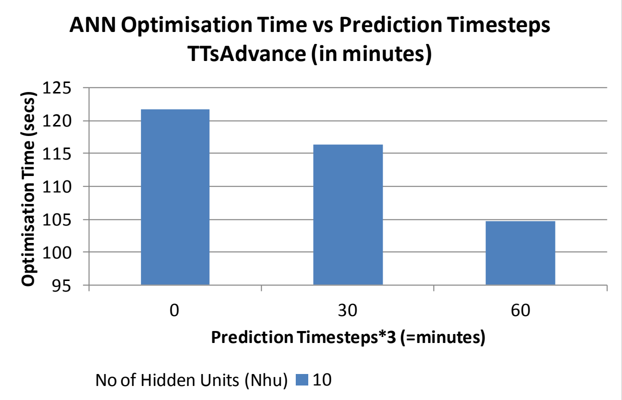

Figure 5. RAPIDS1 – typical 123-node ANN training times for FRMRC study.

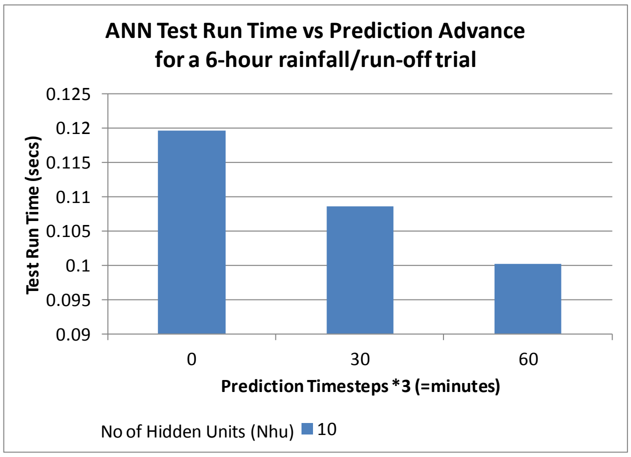

Figure 6. RAPIDS1 – typical 123-node ANN test times for FRMRC study.

In summary, results for UKWIR case studies demonstrated the following:

(Dorchester): Use of soil moisture levels (NAPI) as ANN input demonstrated a small improvement in model performance, but this was probably not sufficient to offset additional costs of data gathering, preparation and application to ANN model.

---

[2] Nash-Sutcliffe Efficiency Coefficient (NSEC); RMSE-Observations Standard Deviation Ratio (RSR) ; Percentage Bias (PBIAS) ; Total Volume Error (TVE) ; ANN Normalised Root Mean Square Deviation (NRMSD); % Samples in Limits - All Nodes; Amplitude Error of Hydrograph Peak ; Timing Error of Hydrograph Peak; R-Squared - All Nodes; Pearson Correlation Coefficient - All Nodes; ANN Output vs Target X-Y Plot (ATXY) - Single Node; ANN Output & Target Hydrographs - Single Node; Confusion Matrix for Peak Flood Depth Categories; Confusion Matrix for Flood Positives & Negatives; Confusion Matrix Accuracy Band summary analysis

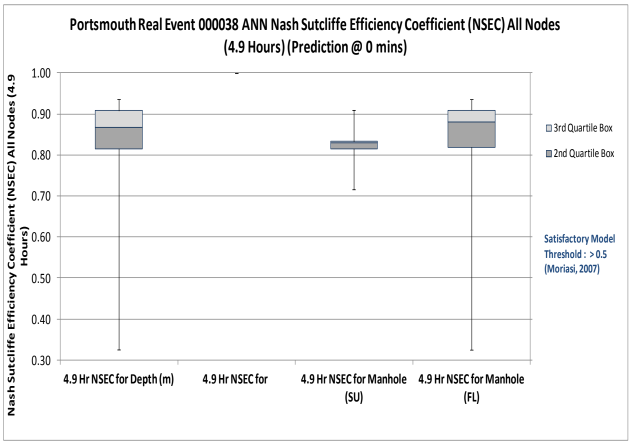

Figure 7. RAPIDS1 – typical spread of ANN output NSEC scores over 20-nodes for a single real rainfall event (Portsmouth case study)

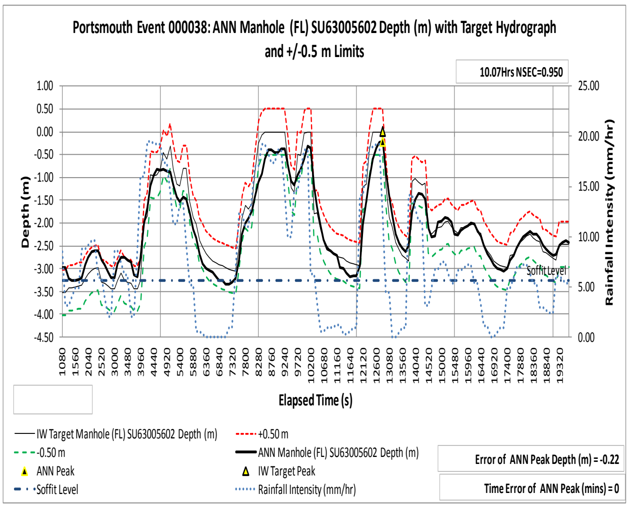

Figure 8. RAPIDS1 – typical target hydrograph and ANN response for a single manhole and rainfall event (Portsmouth catchment)

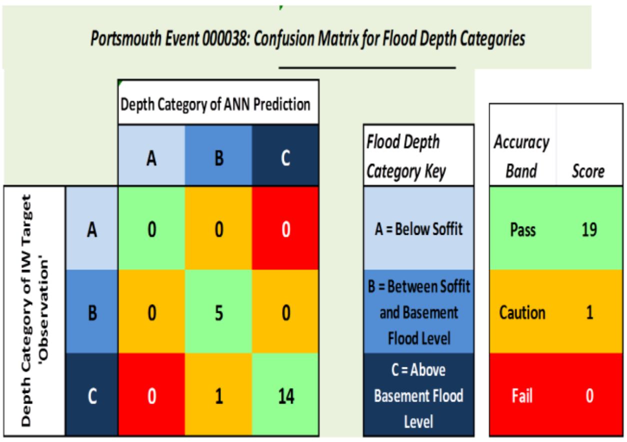

Figure 8. RAPIDS1 – typical classification matrix for three peak flood depth categories (A|B|C) at 20-sewer nodes, for a single rainfall event (Portsmouth catchment). Colour-coded accuracy bands for all nodes are also shown.

(Portsmouth): Use of ANN models were demonstrated successfully to prevent flooding in the 'Morass' area of Portsmouth, when used as a trigger for early initiation of pumping at the Eastney pumping station.
(Crossness): Results for the entire 230km^2 catchment using 23 raingauges as ANN input were poor. Spatial rainfall input worked best when applied to smaller areas (4-5 raingauges subcatchments). Further work is needed.

Work on RAPIDS2 rainfall nowcasting is at too early a stage to present results beyond those shown in Figures 3-4 for the proposed feature extraction approach; the methodology is still under development.

5 CONCLUSIONS & FUTURE WORK

Results for RAPIDS1 show that ANNs can provide a very significant speed improvement over conventional hydrodynamic simulators without excessive degradation in performance. They can moreover be used for flood severity classification. The RAPIDS1 method presents opportunities for automated generation of flood alarms / warnings right down to the individual sewer node, including potentially for networks of considerable size, without being computationally expensive.

However, flood prediction based on actual rainfall alone cannot provide operationally useful lead-times. Instead, prediction is limited in the worst case by the ToC for each node (typically <30 min). However, possibilities for stand-alone use of ANNs for rainfall nowcasting are being explored through a process of radar rainfall echo feature extraction and feature time-series prediction using ANNs (RAPIDS2). More work is needed to determine the value of this approach.

Extending prediction time to operationally useful values of 2+ hours could potentially be achieved by using Met Office rainfall prediction products in place of RAPIDS2.

Assuming that RAPIDS2 achieves satisfactory results, the possibility of cascading the two systems to provide flood-level prediction at manholes based on live radar rainfall images will be tested.

The RAPIDS1 package has been written to allow tailoring to other catchments and water-related EWS requirements to be readily achieved. At present a version of RAPIDS1 is being adapted to early warning of bathing water quality exceedances to comply with the EU directive [22], using a variety of ANN input parameters; principally antecedent rainfall over the catchment.

Acknowledgment. The research reported in this paper was conducted as part of the FRMRC2 and the UKWIR RTM projects, with support from the Engineering and Physical Sciences Research Council, the Department of Environment, Food and Rural Affairs/Environment Agency Joint Research Programme, UK Water Industry Research, Office of Public Works Dublin, and Northern Ireland Rivers Agency. Data were provided by the British Atmospheric Data Centre, Environment Agency, Halcrow, HR Wallingford, Met Office, Mouchel, Ordnance Survey, Richard Allitt Associates and Yorkshire Water. Our thanks go to all the above organisations for their support.

REFERENCES

[1] S.-K. Min, X. Zhang, F. W. Zwiers, and G. C. Hegerl, “Human contribution to more-intense precipitation extremes,” *Nature*, pp. 378–381, 2011.

[2] P. Pall, T. Aina, D. A. Stone, P. A. Stott, T. Nozawa, A. G. J. Hilberts, D. Lohmann, and M. R. Allen, “Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000,” *Nature*, pp. 382–386, 2011.

[3] T. Einfalt, K. Arnbjerg-Nielsen, C. Golz, N.-E. Jensen, M. Quirmbach, G. Vaes, and B. Vieux, “Towards a roadmap for use of radar rainfall data in urban drainage,” *Journal of Hydrology 299*, pp. 186–202, 2004.

[4] L. Li, W. Schmid, and J. Joss, “Nowcasting of Motion and Growth of Precipitation with Radar over a Complex Orography,” *Journal of Applied Meteorology*, vol. 34, no. 6, pp. 1286–1300, Jun. 1995.

[5] W. F. Krajewski and J. A. Smith, “Radar hydrology: rainfall estimation,” *Advances in Water Resources*, vol. 25, no. 8–12, pp. 1387–1394, Aug. 2002.

[6] N. E. Bowler, C. E. Pierce, and A. W. Seed, “STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP,” *Quarterly Journal of the Royal Meteorological Society*, vol. 132, no. 620, pp. 2127–2155, 2006.

[7] S. Achleitner, S. Fach, T. Einfalt, and W. Rauch, “Nowcasting of rainfall and of combined sewage flow in urban drainage systems,” *Water Science and Technology*, vol. 59, no. 6, p. 1145, 2009.

[8] A. N. A. Schellart, M. A. Rico-Ramirez, S. Liguori, and A. J. Saul, “QUANTITATIVE PRECIPITATION FORECASTING FOR A SMALL URBAN AREA: USE OF RADAR NOWCASTING,” in *8th INTERNATIONAL WORKSHOP on PRECIPITATION IN URBAN AREAS*, St Moritz, CH, 2009, pp. 22–26.

[9] P. Wang, A. Smeaton, S. Lao, E. O’Connor, Y. Ling, and N. O’Connor, “Short-Term Rainfall Nowcasting: Using Rainfall Radar Imaging,” *Eurographics ireland*, p. pp, 2009.

[10] S. THORNDAHL, T. Bøvith, M. R. Rasmussen, and R. S. Gill, “On comparing NWP and radar nowcast models for forecasting of urban runoff,” in *Proceedings of IAHS symposium held in Exeter, UK, April 2011*, Exeter, UK, 2011, vol. 351, pp. 620–625.

[11] S. Liguori, M. A. Rico-Ramirez, A. N. A. Schellart, and A. J. Saul, “Using probabilistic radar rainfall nowcasts and NWP forecasts for flow prediction in urban catchments,” *Atmospheric Research*, vol. 103, no. 0, pp. 80–95, Jan. 2012.

[12] “Flood Risk Management Research Consortium 2.” 2012-2009.

[13] A. Duncan, A. S. Chen, E. Keedwell, S. Djordjevic, and D. A. Savic, “Urban flood prediction in real-time from weather radar and rainfall data using artificial neural networks,” in *IAHS Red Book series no. 351*, Exeter, UK, 2011, vol. 351.

[14] R. Kellagher, “The Use of Artificial Neural Networks (ANNs) in Modelling Sewerage Systems for Management in Real Time: Volume 1 - UKWIR Main Report (12/SW/01/2).” UKWIR, 2012.

[15] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *Evolutionary Computation, IEEE Transactions on*, vol. 6, no. 2, pp. 182–197, 2002.

[16] B. W. Golding, “Nimrod: a system for generating automated very short range forecasts,” *Meteorological Applications*, vol. 5, no. 1, pp. 1–16, 1998.

[17] N. BADC, “Met Office - Rain radar products (NIMROD).” http://badc.nerc.ac.uk, 2011.

[18] N. BADC, “BADC UKMO Nimrod Data,” *British Atmospheric Data Centre*. 2013-2004.

[19] I. Daubechies, “Orthonormal bases of compactly supported wavelets,” *Communications on Pure and Applied Mathematics*, vol. 41, no. 7, pp. 909–996, 1988.

[20] I. Tchiguirinskaia, D. Schertzer, C. T. Hoang, and S. Lovejoy, “Multifractal study of three storms with different dynamics over the Paris region,” in *Proceedings of Weather radar and hydrology symposium, Exeter, UK*, 2011.

[21] D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith, “Model evaluation guidelines for systematic quantification of accuracy in watershed simulations,” *Transactions of the Asabe*, vol. 50, pp. 885–900, 2007.

[22] European Commission, “Revised Bathing Water Directive (2006/7/EC).” European Commission, 15-Feb-2006.