

Artificial Neural Network Methods Applied to Forecasting River Levels

Fernando Dornelles*, Joel Avruch Goldenfum*, Olavo C. Pedrollo*

fds_eng@yahoo.com.br; j.goldenfum@gmail.com; pedrollo@gmail.com

Recebido: 17/10/11 - revisado: 26/06/12 - aceito: 07/06/13

ABSTRACT

The use of data-driven models may be an important alternative in several scientific fields, especially when the available data do not allow utilizing physical hydrologic models because these data must be measured in the basin. . This paper explores important aspects of ANN use: initial training conditions, performance assessment, partitioning of the strong seasonal component in short-term samples and ranking results by a weighted score. Sequential partitioning of the sample was shown to be adequate for cases where the data series has a strong seasonal component and short time response. The non-exceeded error was associated with its frequency, giving a measure of performance that is easily understood and which does not depend on the long familiarity required by traditional methods to evaluate results. A weighted score calculated from several indices removed the difficulty of how to reconcile several statistical measures of performance. The need for repeated artificial neural network training using random starting conditions is established, and the ideal number of repetitions to ensure good training was investigated. A straightforward approach to visualization of forecasting errors is presented, and a pseudo-extrapolation region at the domain extremes is identified. The methods were explored using the Quaraí river basin, whose specific characteristics include a rapid response to precipitation events. It therefore provides a good test of artificial neural network methods, including the use of rainfall forecasts which, to be combined with existing data resources, required novel methodological approaches.

Keywords: *Neural networks. Quantitative rainfall. Forecasting. Water-level prediction.*

INTRODUCTION

The hydroinformatics is being increased since end of 1980's due the fast evolution of computer process capacity (ABBOT, 2008), in especial the data-driven modelling and computational intelligence, that have proven their applicability to various water-related problems: modelling, short-term forecasting, data classification, reservoir optimisation, building flood severity maps based on aerial or satellite photos, etc. Data-driven models would be useful in solving a practical problem or modelling a particular system or process if a considerable amount of data describing this problem is available;

there are no considerable changes to the modelled system during the period covered by the model. Such models are especially effective if it is difficult to build knowledge-driven simulation models (e.g. due to lack of understanding of the underlying processes), or the available models are not adequate enough (SOLOMATINE; SEE; ABRAHART, 2008).

Nestes casos, as redes neurais artificiais progressivas são candidatas naturais, por serem identificadas como aproximadores universais de funções (HORNIK; STINCHCOMBE; WHITE, 1989).

Focusing on the mathematical methods for the multi-layer neural networks empirical modelling this paper deals with important aspects of this technique, such as the problem of convergence, the identification of optimum architecture, sample partitioning, and indices of model performance. In most published work, these topics either receive little.

*IPH - Instituto de Pesquisas Hidráulicas – UFRGS

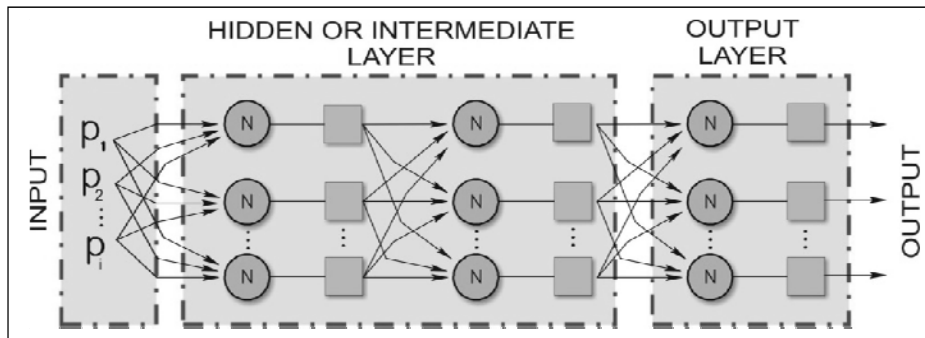


Figure 1 – Example of neural network structure

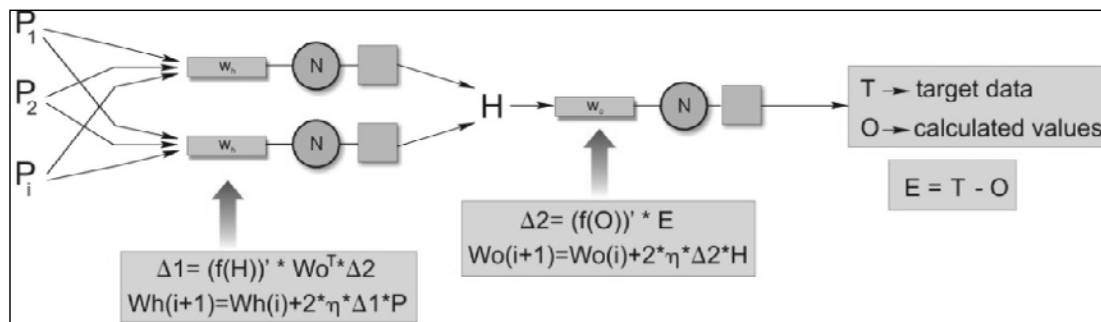


Figure 2 – Training scheme for training by back-propagation

Difficulties in the use of neural networks arise particularly where there is strong seasonality and where data are limited: a situation commonly encountered when modelling hydrological phenomena. The objective of the paper is therefore to establish whether the mathematical and strategic resources available for use can contribute, when correctly used, to the solution of some of the problems that occur where artificial neural networks are used to forecast river levels. To illustrate the application of mathematical methods for neural network modeling, a small hydrologic basin was chosen for which forecasting is inherently difficult and for which data series are short.

On this way the aim here is not compare results of different models, the application case was employed just to show the effects of the techniques proposed.

ARTIFICIAL NEURAL NETWORKS

Neural networks are computational models inspired by the way in which biological neurons function, and consist of processing layers with each

layer having several neurons. Data are received by the neurons in the first layer (the entry layer) and produce output signals which in turn stimulate neurons in the next layer, until the final layer (the output layer) is reached, as illustrated in figure 1.

The principal characteristics of artificial neural networks are that they learn and generalize. They learn in terms of their capacity to extract information from a sample of observations of input vectors and their corresponding outputs; they generalize through their capacity to respond to situations not previously encountered during the learning process.

The most widely-used training method for a multi-layer network is by error back-propagation (RUMELHART; HINTON; WILLIAMS, 1986). The weights of each neuron are updated by using what is termed the “delta rule” separately for each neuron:

$$w_{k+1} = w_k - \eta \nabla (E(w_k)) \tag{1}$$

where η is the magnitude of the learning step, taken with a negative sign; w are the neuron synaptic weights, and E is the quadratic error of the network output.

The procedure used in what follows is described by Kovacs (2002). The inputs to each neuron are the outputs from the preceding layer, and its errors, when the neuron is not an output neuron, are found from the product of the weights in the following layer with the derivative of the activation function, and from the errors of that layer (Figure 2).

Convergence of this method can be speeded up by using a variable learning step which is increased (decreased) when the error in the previous iteration is smaller (greater) than in the current iteration. An inertial term, momentum, can also be introduced which reduces oscillation towards convergence by effectively maintaining the most recently observed trend.

When training a neural network it is possible for overfit to occur; this is when a model is fitted to noise terms in the training sample, generating unsatisfactory results when the fitted network is applied to new data not used for training. Giustolisi and Laucelli (2005) observed that this problem always produces negative effects on generalization, which are worse if the noise is non-Gaussian. The problem can be avoided by using a cross-validation sample which is used to respond in parallel whilst the network is being trained. In the case of the training sample the mean error always becomes smaller as the number of iteration increases, but for the cross-validation sample there will be a number of iterations for which the error is a minimum (Figure 3).

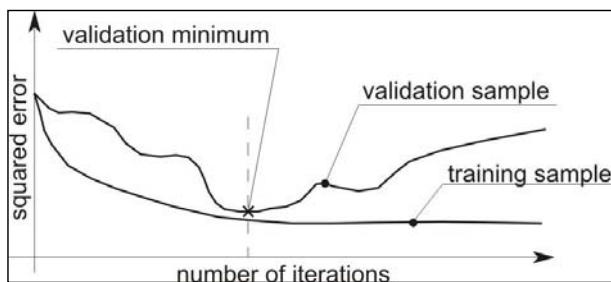


Figure 3 – Quadratic error function for training and validation samples

In the case of supervised training with both input and output data available, the cross-validation method divides the total available sample into three parts: one part for training, one for cross-validation, and one for verification.

Neural networks have been used successfully for hydrological forecasting (ACHELA; JAYAWAR-

DENA, 1998; BRAVO et al., 2009; TOKAR; MARKUS, 2000), but many methodological aspects remain to be explored, such as the initial training conditions, error evaluation and partitioning of the strong seasonal component when the available records are short. In this paper, these aspects are explored using as an example the forecasting of river water-levels in a drainage basin with short hydrological records.

APPLICATION

Characteristics of the system

The drainage basin lies upstream of the cities Quaraí and Artigas, in Brazil and Uruguay respectively, and has an area of roughly 4500 km². Its soil is shallow (~50cm), giving a high mean runoff coefficient (~ 0.46) for a rural basin. Its time of concentration is about 28h, and the time between peak rainfall and peak discharge is about 12h. These characteristics are such that 90% of the total annual runoff from the basin occurs in 30% of the time (PPGICBRQ, 2005).

Characteristics of the system

The region for which water-level forecasts are required has a telemetered water-level gauge monitored by the DNH (Dirección Nacional de Hidrología del Uruguay), giving measurements of river level that are continuously updated. A hydrometeorological forecast, downloaded from the internet, is issued twice daily (at 0h and 12h) by the Brazilian agency CPTEC (Centro de Previsão do Tempo e Estudos Climáticos). This quantitative forecast is generated for a network of points on a 40km grid and with a 5-day time horizon. Four grid-points were selected that lie within the basin and a forecast of mean daily rainfall was calculated using the areas of Thiessen polygons as weights.

The raw data were organized to obtain the samples used by the neural network, and consisted of three inputs and one output variable. The inputs to the neural network used to explore the proposed method were the river levels one and two days prior to the day on which the forecast is issued, together with the sum of rainfall forecasts for the basin, up to the day for which the water level is forecast. The period for which rainfall forecasts were available began on 5 January 2005 and ended 5 January 2007.

Sample partition

Each of the five sample series (one for each forecast horizon) was divided into three parts: one for training, another for validation and the third for verification. The partitioning that is normally used, in which the sample is divided into three separate sub-samples that are each contiguous in time, gives rise to problems (LACHTERMACHER; FULLER, 1994) as, due to the possible existence of anomalous periods, the samples will be not homogeneous, leading to an inadequate neural network training. So that an alternative procedure was sought, in which the separation was effected by taking four samples, the first and third for training, the second for validation and the fourth for verification, in the scheme shown in figure 4.

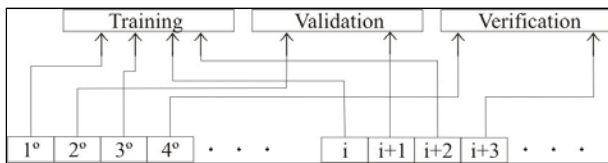


Figure 4 – Sequential partitioning of samples

The reason for this form of partitioning was to obtain samples that were more representative of the whole data-set. The usual method of partitioning samples would have led to very heterogeneous sub-samples, because of the pronounced seasonality in basin hydrological regime. Another method for sample partition was used by Dawson et al. (2006), where the three parts was obtained with a random selection that produced reasonable samples with different catchment types and sizes in each sub-set. To ensure that the training sample is as representative as possible, it is larger than those for cross-validation and verification, so that the proportions of records allocated to training, cross-validation and verification were 50%, 25% and 25%, respectively.

Characteristics of the neural networks used

Neural networks can have various configurations according to the combination of their characteristics. A three-layer configuration was used here, since by virtue of the Kolmogorov-Nielsen theorem of 1957, cited by Kovacs (2002), this is adequate for approximating functions by means of neural networks.

In an initial trial, both the activation functions and the training parameters (step-size, number of iterations, stopping criteria, minimum error and

decrease in its slope) were fixed, with only the number of neurons in the intermediate level being allowed to vary by trial and error. The activation functions were TANSIG (sigmoid with response values between +1 and -1) for the input layer, and POSLIN (linear with response values equal to or greater than zero) for the output layer (DEMUTH; BEALE, 2004). The function used for the output avoids the occurrence of negative values without imposing an upper limit, so that forecasts in the verification sample can have values greater than those found in the training sample (Figure 5).

These functions take values only over their permitted ranges, so that data samples must be scaled so that they too lie within these ranges. The stopping criteria were that the mean square error should be less than 0.001 and that the number of iterations should be less than 20,000. This limit was adopted by observing the decrease in mean square error, which is almost irrelevant beyond this number of iterations. Another stopping criterion was to set the value of least gradient at 1×10^{-10} , which would indicate non-converging trend. The value of the training step was set at 0.02.

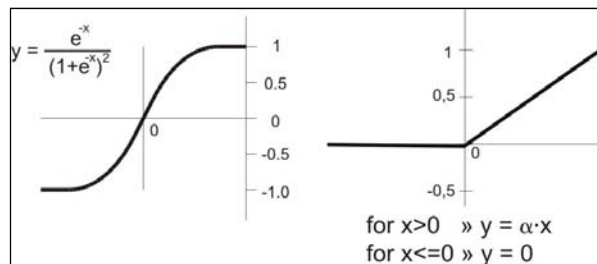


Figure 5 – Graphs of the activation functions TANSIG (left) and POSLIN (right)

A number of inflections due to increase in error calculated in the validation sample was permitted, the maximum being set at 50. This value was chosen because oscillations were observed shortly after training began, which interrupt the convergence process, still giving very large mean square errors.

Criteria for evaluating performance

Results for evaluation of comparative performance used the same verification sample for both models analyzed in this paper. This sample had 122 values from which the following performance measures were calculated: error mean square (EMS), absolute mean error (AME), absolute stand-

ard error (ASE), coefficient of linear correlation (R), coefficient of persistence (CP) and indices made of absolute error non-exceeded for frequencies 50%, 75% and 90%. The efficiency coefficient is the proportion of variation in the observed variable that is explained by the model, and the “non-exceeded error” for a given frequency is the numerical value which is not exceeded by a specified percentage of model errors, giving a simple intuitive measure of the quality of model predictions (PEDROLLO, 2005). Symbolically,

$$EMS = \frac{\sum (P_c - P_o)^2}{n} \quad (2)$$

$$AME = \frac{\sum |P_c - P_o|}{n} \quad (3)$$

$$ASE = \sqrt{\frac{\sum (P_c - P_o)^2}{n}} \quad (4)$$

$$R = \frac{\sum (P_o - \bar{P}_o) \cdot (P_c - \bar{P}_c)}{\sqrt{\sum (P_o - \bar{P}_o)^2 \cdot \sum (P_c - \bar{P}_c)^2}} \quad (5)$$

$$CP = 1 - \frac{\sum (P_{o_i} - P_{c_i})^2}{\sum (P_{o_i} - P_{o_{i-1}})^2} \quad (6)$$

where PC , PO and n are the forecast level, observed level and number of values in the sample respectively. The “non-exceeded error” is the p-quantile of the absolute errors |Po-Pc| for a given frequency p.

The use of a single measure of fit to evaluate performance of a neural network invariably favours one single characteristic, such as the magnitude of extreme events, or the magnitude of a mean value over time. Taking several such indices together gives a better idea of network performance, but analyzing them singly can be time-consuming. The approach used in this paper is to combine a number of performance indices into a single weighted value; by judicious selection of weights, greater emphasis can be allocated to (for example) extreme values, without discarding all other aspects of model fit.

Definition of the smallest number of training repetitions

There is no way to be sure that good initial values of the synaptic weights (i.e., weights given to the links between inputs and activation functions) have been chosen, and usually they have been ini-

tialized with small and random values (ASCE TASK COMMITTEE, 2000), so that distinct paths were always traced out on the errors surface at each new initiation, thus repeatedly ending up at a different local minimum. It is desirable that training be repeated with the best results being selected, but there is little guidance about how many repetitions should be used; this aspect was therefore analyzed in this study with a view to ascertaining the number of repetitions needed for which no further improvement in results occurs when this number is exceeded. Antcil et al. (2006) evaluated performance variability, which depends of the random initial values of the neural network weights, by repeating the complete neural network training 50 times and then using the median result as a reference performance, but they did not show the behaviour of other training runs. The median may also not be the most appropriate measure.

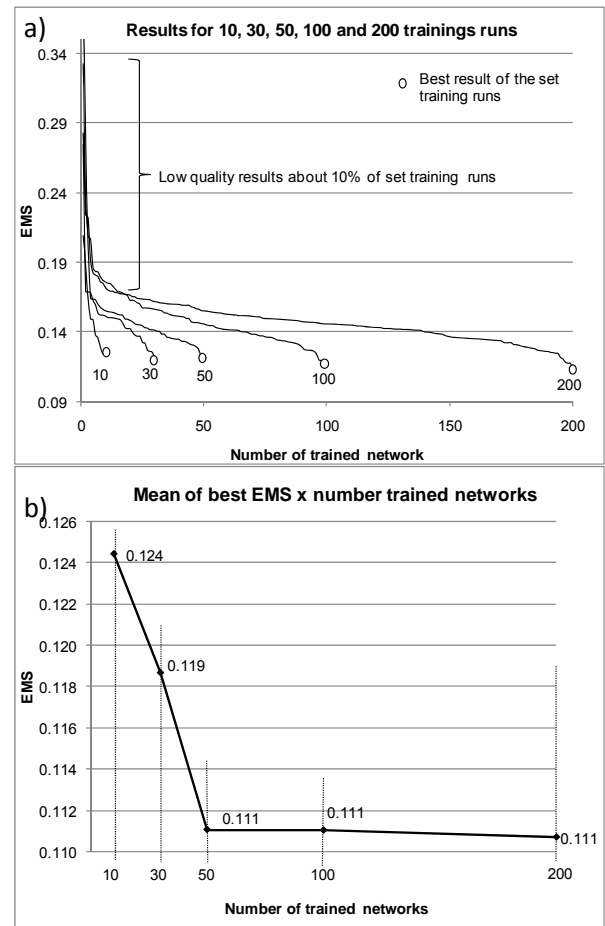


Figure 6 – Behaviour of the index as a function of the number of training runs

To find a more robust measure of the influence of number of training runs, the present study used ten simulations for each initialization set of 10, 30, 50, 100 and 200, resulting in the mean values of EMS shown in figure 6(b). Values were also obtained for other measures of fit (AME, ASE, R, CP and percentage non-exceeded), giving 100, 300, 500, 1000 and 2000 trained networks for the 10, 30, 50, 100 and 200 initialization sets respectively.

Combination of indices of model performance

An obvious first approach to judging which training run gave the best result would be to choose the run giving the best indices; however it is rare for one training run to give the best values for all indices of performance, and a run that is best for some indices may show mediocre values for others. Further, the contradictory indices results may indicate different neural networks or neuron synaptic weights implying on a dilemmas problem to choice the best model.

To deal of the dilemmas of ranking the results performance and to limit the time required both to select the best neural network architecture and to compare results from different models, a combined index was calculated by weighting the statistics obtained, with the weights selected so as to place more importance on those indices which penalize errors in peak levels (Table 1).

The combined index goal is to give a more robust statistic than analyse just the set of indexes, once, frequently, the indexes indicate discrepant conclusions, although near with each other.

Table 1 – Weights applied to indices of neural network performance

Error mean square (EMS)	3
Absolute mean error (AME)	2
Absolute standard error (ASE)	1
Coefficient of linear correlation (R)	2
Coefficient of persistence (CP)	3
90% non-exceeded error	3
75% non-exceeded error	2
50% non-exceeded error	1

The scaled value $V_{i,j}$, for each index is

$$V_{i,j} = w_i \cdot \left(\frac{|\text{best}(I_i) - I_{i,j}|}{\text{best}(I_i) - \text{worst}(I_i)} \right) \quad (7)$$

where w_i is the weight given to the i -th index; I is the vector containing the results of the i -th indices that are to be weighted; j is the number of the training run.

The score NP_j for each training run is then

$$NP_j = \frac{\sum V_{i,j}}{\sum w_i} \quad (8)$$

These indexes were developed placing bigger importance to peak level errors; consequently, they will be appropriate only for applications where the estimation of the peak levels and their recurrences is the main goal. If the main objective is different, other indexes can be necessary. For example, when applying the methodology to a flood warning system, the main variables would be alert and flooding levels and their respective times of occurrence. In this case, the above presented indexes are not the most appropriate; to penalize a specific water level in the judging which training run gave the best result, it is necessary to create an exclusive index and to attribute a high weight to the selected water level target.

RESULTS AND ANALYSES

Table 2 shows the best mean indices, for 10, 30, 50, 100 and 200 training runs. These were obtained from the ten simulations made for each number of training runs, with a one-day-ahead forecast horizon.

Based on the test results, the minimum of 50 training runs was used for each of neural network architecture, since there was little evidence of any useful increase in precision beyond that number. The best training run from the sets of 50 was therefore identified, and having chosen the best architecture for the forecasting model, the selected network was applied to the verification sample which had played no part in either training or cross-validation of the selection stages, and so was totally independent of them. Table 3 shows the results obtained from the verification sample using the best model selected for each forecast horizon.

The performance indices can be interpreted individually. Whilst, for example, the coefficient of linear correlation R for one-day-ahead forecasts was 0.899, the 90% non-exceeded error was 0.48m. Based on this value of R , a decision-maker familiar with the model results could decide whether this result is favourable or not, taking account of basin size and other factors. But even without any familia-

Table 2 – Best index for set of 10, 30, 50, 100 and 200 training runs

N° runs	training	EMS (m ²)	AME (m)	ASE (m)	R	CP	90% non-exceeded error (m)	non-75% exceeded error (m)	non-50% exceeded error (m)
10		0.124	0.170	0.352	0.951	0.579	0.401	0.181	0.073
30		0.119	0.164	0.344	0.953	0.597	0.374	0.171	0.068
50		0.111	0.160	0.333	0.956	0.623	0.369	0.166	0.069
100		0.111	0.159	0.333	0.956	0.623	0.359	0.164	0.067
200		0.111	0.159	0.332	0.957	0.625	0.359	0.164	0.068

Table 3 – Results from the verification sample – neural network with forecasts of rainfall

Horizon	EMS (m ²)	AME (m)	ASE (m)	R	CP	90% non-exceeded error (m)	non-75% exceeded error (m)	non-50% exceeded error (m)	Score
Day 1	0.255	0.227	0.505	0.899	0.361	0.480	0.204	0.079	0.824
Day 2	0.436	0.345	0.660	0.812	0.414	0.795	0.333	0.157	0.494
Day 3	0.546	0.422	0.739	0.757	0.442	0.946	0.486	0.207	0.275
Day 4	0.545	0.425	0.738	0.759	0.503	0.991	0.515	0.219	0.324
Day 5	0.637	0.477	0.798	0.735	0.475	1.073	0.541	0.276	0.142

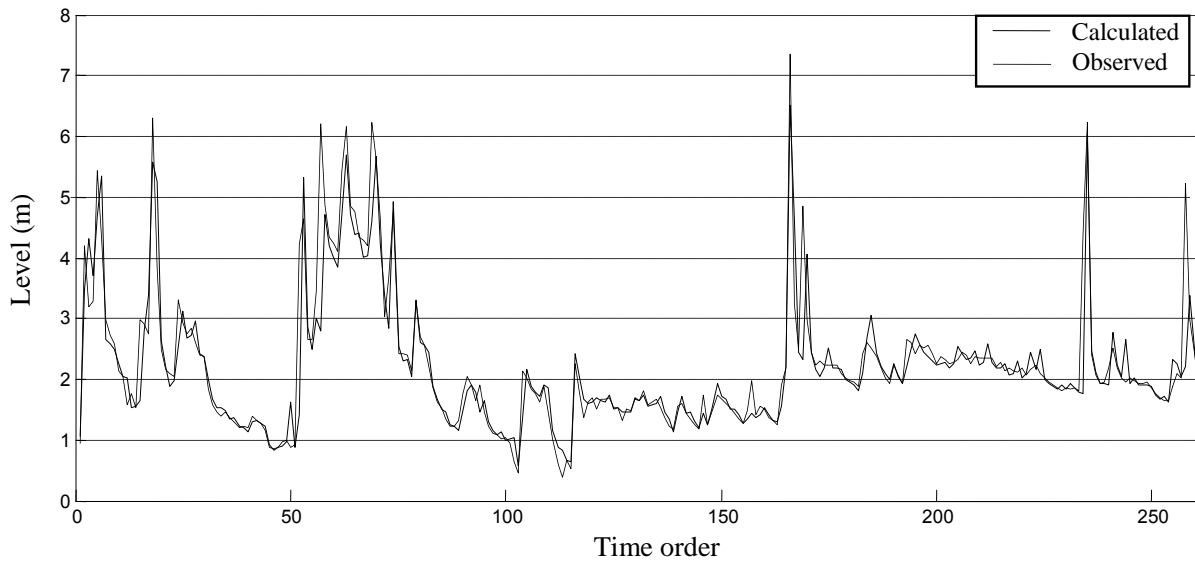


Figure 7 – Observed water-levels and one-day-ahead forecasts

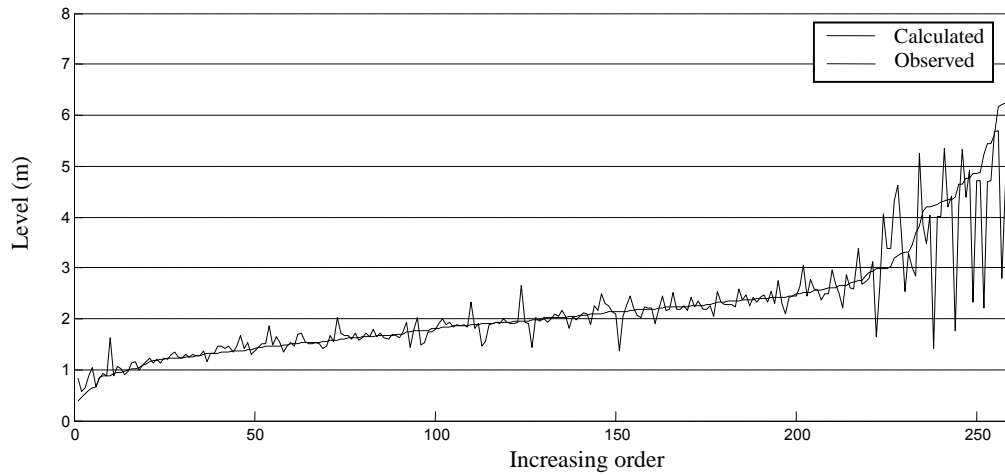


Figure 8 – One-day-ahead forecasts and corresponding observations, in increasing order of magnitude

ity with index usage, a decision-maker could consider whether a model in which the error in forecasts will not be greater than 48cm for 90% of the time is adequate for his/her needs.

However, a better comparison may be obtained by using the weighted score in order to reduce ambiguities. The decline in performance as a function of increasing forecast horizon shown by the indices given above is reversed between Days 3 and 4. The quality of the rainfall forecasts when they are placed in rank order of days is found to be 1, 2, 5, 4 and 3, which suggests some hypotheses that explain this result. The water-level forecast uses forecast values of accumulated rainfall, so that the drop in rainfall forecast accuracy on Day 3 exerts a greater influence than when the forecast horizon is 4 days.

Figure 7 shows a graph of observed and predicted water-levels for a one-day forecast horizon.

A look at values plotted in time order is a useful first step for showing whether the model has identified a false fit.

One way to evaluate the errors in water-level forecasts for events of greater magnitude is to graph the forecast level against the magnitude of the level that the model should have shown, i.e., the observed level. This is shown in figure 8.

CONCLUSIONS

The research reported in the paper was undertaken to explore modelling techniques, which were applied to a small drainage basin for which

forecasting water-levels is very difficult, since short hydrological records limit the flexibility in modeling approach. Although the final numerical results were not particularly encouraging for an water-level forecasting operational use in the specific drainage basin of the river Quaraí, the purpose of the work was to explore and develop methodology without compare with results in other basins, that can be an issue for future papers to prove the effectiveness of the techniques here proposed. The particular characteristics of the Quaraí basin include a high runoff coefficient, a concentration time of about 28h, and a time between peak rainfall and peak runoff of 12h, showing that the basin response to rainfall is rapid. Thus it presents a challenging test for artificial neural network modelling even with the correct use of existing methods, whilst some novel developments were also needed.

Sequential partitioning of the data record was found to be convenient because the data sequences were strongly seasonal and fairly short. This sequential method ensured that the three sample groups, for training, validation and verification, maintained the same statistical characteristics, which is a convenient basis for training and testing neural networks. The statistical equivalence of the samples was also important for determining the band of random variation during cross-validation.

The use of non-exceeded error values for different frequencies was found to be particularly useful for evaluating model results, giving direct measures of error magnitudes for frequencies 50, 75 and 90%. Such measures provide a very convenient description of forecast qualities because they are easy to understand even by users who have not seen

them before. This is in contrast to more traditional measures of model fit for which a certain experience is needed before their use comes naturally. The non-exceeded errors also agree with the other indices used which show that forecast performance falls off as the forecast horizon increases.

The use of a weighted score of several indices was shown to be useful in that it removed the dilemma of which among several different indices should be used for comparing the performance of alternative models.

The problem of convergence to local minima when training a neural network was investigated using repeated training runs and analysis of their results. It was established that about 10% of training runs reach minimum values that gave much poorer model fits than the best training run. It was also found that the best performance measure from 50 runs was roughly equal to that given by 100 and 200 runs, independent of network architecture and the training algorithm used.

A firm conclusion from the study, therefore, is that a number of training runs are needed with randomized initial conditions, since it was shown that results from a single training run, used in much published work, are unreliable.

It would be very desirable to have a general recommendation concerning the number of training runs needed, and the results given here are a step towards achieving this end, although further confirmation is required by tests with data from other case studies. If such research shows that no general recommendation is possible, more complex guidelines will need to be developed.

Presenting the results of model forecasts, in the usual way, by plotting observed and forecast water-levels in time sequence may obscure model errors, in particular for events where water levels are high. The alternative of showing results in order of increasing water levels gave a very clear picture of the poor fit obtained at high water levels.

With these tools, a decision-maker can look at a graph similar to those shown and decide whether the errors that will occur, particularly at higher water levels, are such that his/her objectives will fail to be met, even though statistical indices look reasonably satisfactory, in their bands of variation.

It can be further concluded from this test-case that neural networks are unable to extrapolate outside the domains observed, where functions are to be approximated. This is already accepted by many, but the present work confirms that neural networks have difficulty even in approximating regions that occur only infrequently within the do-

main of observation. Or, expressed another way, there is a region of pseudo-extrapolation near domain extremes which needs much further study.

ACKNOWLEDGEMENTS

The authors would like to thank CNPq for its award of a study grant; DNH (Dirección Nacional de Hidrología del Uruguay) for giving access to water-level data; and CPTEC (Centro de Previsão do Tempo e Estudos Climáticos) for providing meteorological forecasts.

REFERENCES

- ABBOTT, M. B. Some future prospects in hydroinformatics in hydroinformatics. In: ABBOTT, M. B. *Computational intelligence and technological developments in water applications*. Berlin: Springer, 2008. chapt. 1, p. 3-16. (Water Science and Technology Library, 68).
- ACHELA, F. D. K.; JAYAWARDENA, A. W. Runoff forecasting using RBF networks with OLS algorithm. *Journal of Hydrology Engineering*, p. v. 3, n. 3, 203-209, 1998.
- ANCTIL, F.; LAUZON, N.; ANDRÉSSIAN, V.; OUDIN, L.; PERRIN, C. Improvement of rainfall-runoff forecasts through mean areal rainfall optimization. *Journal of Hydrology*, v. 328, n. 3-4, p. 717-725, 2006.
- ASCE TASK COMMITTEE. Artificial neural networks in Hydrology I. *Journal of Hydrologic Engineering*, v. 5, n. 2, p. 115-123, 2000.
- BRAVO, J. M.; PAZ, A. R.; COLLISCHONN, W.; UVO, C. B.; PEDROLLO, O. C.; CHOU, S. Incorporating forecasts of rainfall in two hydrologic models used for medium-range streamflow forecasting. *Journal of Hydrology Engineering*, v. 14, n. 5, p. 435-445, 2009.
- DAWSON, C. W.; ABRAHART, R. J.; SHAMSELDIN, A. Y.; WILBY, R. L. Flood estimation at ungauged sites using artificial neural networks. *Journal of Hydrology*, v. 319, n. 1-4, p. 391-409, 2006.

DEMUTH, H.; BEALE, M. *Neural network toolbox: user's guide version 4.0, for use with MATLAB*. Natick, MA: Mathworks, 2004. 1v.

GIUSTOLISI, O.; LAUCELLI, D. Improving generalization of artificial neural networks in rainfall-runoff modeling. *Hydrological Sciences Journal*, v. 50, n. 3, p. 439-457, 2005.

HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. *Neural Networks*, v. 2, n. 5, p. 359-366, 1989.

KOVACS, Z. L. *Redes neurais artificiais: fundamentos e aplicações*. 3. ed., rev. São Paulo: Livraria da Física, 2002.

LACHTERMACHER, G.; FULLER, J. D. Backpropagation in hydrological time series forecasting. In: HIPEL, K. W.; MCLEOD, A. I.; PANU, U. S.; SINGH, V. P. (Ed.). *Stochastic and statistical methods in hydrology and environmental engineering*. Dordrecht: Kluwer, 1994. 229-242. v. 3.

PEDROLLO, O. C. Forecasting River Levels with Neural Networks Application for Rosário do Sul – RS (in Portuguese). In: ÁGUASUL - SIMPÓSIO DE RECURSOS HÍDRICOS DO SUL, 1.; SIMPÓSIO DE ÁGUAS DA AUGM, 1., 2005, Santa Maria. *Anais...* Santa Maria: [s.n.], 2005. v. 1.

PPGICBRQ. Pilot Project on Integrated Flood Management in the Basin of the River Quarai. Final report (in Portuguese). DNH (Dirección Nacional de Hidrografía del Uruguay) and IPH (Instituto de Pesquisas Hidráulicas, Brasil). 2005.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning Internal Representations by Error Propagation. In: RUMELHART, D. E.; MCCLELLAND, J. L. *Parallel distributed processing: explorations in the microstructure of cognition*. Cambridge, MA: MIT Press, 1986. v. 1.

SOLOMATINE, D. L. M.; SEE, L. M.; ABRAHART, R. J. Data-driven modelling: concepts, approaches and experiences. In: ABRAHART, R.J.; SEE, L. M.; SOLOMATINE, D. P. *Practical in hydroinformatics: computational intelligence and technological developments in water applications*. Berlin: Springer, 2008. chapt. 2. (Water science and technology library; 68).

TOKAR, A. S.; MARKUS, S. Precipitation-runoff modeling using artificial neural network and conceptual models. *Journal of Hydrologic Engineering*, v. 5, n. 2, p. 156-161, 2000.

RESUMO

A utilização de modelos orientados pelos dados (“data-driven models”) pode ser uma alternativa importante, principalmente quando não se dispõe de dados que permitam a utilização de modelos de base física, providos de parâmetros estabelecidos em função das propriedades medidas no sistema. O presente trabalho explora aspectos importantes na utilização das redes neurais artificiais: Condições iniciais aleatórias do treinamento, a avaliação do desempenho, o particionamento dos dados para amostras pequenas com forte sazonalidade e o ordenamento dos resultados por meio de um índice ponderador de diversas estatísticas. Uma técnica de particionamento sequencial mostrou-se adequada para casos em que a série de dados apresenta forte sazonalidade e rápida resposta temporal. Os quantis das amostras dos erros, utilizados como índices de não-excedência associados à frequência de ocorrência fornecem uma estatística de desempenho de fácil percepção, cujo significado, em termos absolutos, permite interpretação direta, independentemente da experiência prévia, como acontece com os tradicionais métodos de avaliação de desempenho de resultados. Um índice ponderado calculado com base em vários índices de desempenho removeu a dificuldade de como conciliar a contradição entre diversas estatísticas de medição de desempenho. A necessidade de repetir o treinamento da rede neural artificial usando condições iniciais aleatórias é confirmada, e foi investigado o número ideal de repetições necessárias para garantir um bom treinamento. Uma visualização dos erros em função do nível d'água em ordem crescente é apresentada, e uma região de pseudo-extrapolação para os valores extremos é identificada. Os métodos foram explorados em uma aplicação para a bacia do rio Quarai, que apresenta uma rápida resposta para eventos de precipitação. As dificuldades resultantes da rapidez das respostas, por um lado, limita o desempenho que é possível alcançar, porém, por outro, constitui uma oportunidade para avaliar as metodologias aplicadas, incluindo o uso de previsões de precipitação, que, combinada com os dados de monitoramento existentes, acabam por requerer uma nova metodologia de abordagem.
Palavras-chave: Redes neurais. Quantidade de água. Previsões. Nível de água.