Hybrid Stochastic Forecasting Model for Management of Large Open Water Reservoir with Storage Function

Tomas Kozel 1, Milos Stary 1

1 University of Technology, Faculty of Civil Engineering, Institute of Landscape Water Management, Veveří 331/95 Brno, Czech Republic

kozel.t@fce.vutbr.cz

Abstract. The main advantage of stochastic forecasting is fan of possible value whose deterministic method of forecasting could not give us. Future development of random process is described better by stochastic then deterministic forecasting. Discharge in measurement profile could be categorized as random process. Content of article is construction and application of forecasting model for managed large open water reservoir with supply function. Model is based on neural networks (NS) and zone models, which forecasting values of average monthly flow from inputs values of average monthly flow, learned neural network and random numbers. Part of data was sorted to one moving zone. The zone is created around last measurement average monthly flow. Matrix of correlation was assembled only from data belonging to zone. The model was compiled for forecast of 1 to 12 month with using backward month flows (NS inputs) from 2 to 11 months for model construction. Data was got rid of asymmetry with help of Box-Cox rule (Box, Cox, 1964), value r was found by optimization. In next step were data transform to standard normal distribution. The data were with monthly step and forecast is not recurring. 90 years long real flow series was used for compile of the model. First 75 years were used for calibration of model (matrix input-output relationship), last 15 years were used only for validation. Outputs of model were compared with real flow series. For comparison between real flow series (100% successfully of forecast) and forecasts, was used application to management of artificially made reservoir. Course of water reservoir management using Genetic algorithm (GE) + real flow series was compared with Fuzzy model (Fuzzy) + forecast made by Moving zone model. During evaluation process was founding the best size of zone. Results show that the highest number of input did not give the best results and ideal size of zone is in interval from 25 to 35, when course of management was almost same for all numbers from interval. Resulted course of management was compared with course, which was obtained from using GE + real flow series. Comparing results showed that fuzzy model with forecasted values has been able to manage main malfunction and artificially disorders made by model were founded essential, after values of water volume during management were evaluated. Forecasting model in combination with fuzzy model provide very good results in management of water reservoir with storage function and can be recommended for this purpose.

1. Introduction

Main objective of researched was construction of stochastic forecasting model (generator of artificially flow line) for large open water reservoir with storage function. Advantage of stochastic forecasting model is fan of possibilities, which is able better describe future development of random process, than
deterministic model. Flow in measure profile could be classified as random process. For construction of stochastic forecasting model was used interface of program Matlab 2013. The main task for model was ability of forecasting drought month and month with average flow, which are critically important for management of large open water reservoir with storage function. The model works with principle of adaptability [4], what for model for each time step movement is forgotten previous calculations and calculation of model is started again, but with data for new time step.

2. Data

For construction of forecasting model was used 90 years long real flows series average monthly flows. The flow series was measured in measuring profile Bílovice, which is situated on the river Svitava (Czech Republic). Above the profile are not situated large open water reservoir, and therefore nature development of flow series is not affected by management of large open water reservoir. The first 75 years were used for construction of forecasting model. The last 15 years were used for validation of the model. Data were sorted according month, when they were measured. Every month has unique probabilistic distribution (elevation \( Q \)). Due to this reason each month data was transformed to standard normal distribution. Transformation was done in two steps. First step got rid of asymmetry (elevation \( Y \)) with using Box-Cox rule equation 1 [1]. Parameter \( r_j \) was calculated by grid method [5], where deviation of coefficient asymmetry from 0 was criteria of optimization.

\[
y_{i,j} = \frac{x_{i,j}^{r_j} - 1}{r_j},
\]

where \( y_{i,j} \) is transformed average monthly flow without asymmetry, \( x_{i,j} \) is average monthly flow, which will be transformed by equation, \( r_j \) is coefficient of transformation for chosen month, \( i \) is number of average monthly flow of particular month and \( j \) is number of month in year.

Data sets were transformed on standard normal distribution by appropriate transformation equations in next step (elevation \( Z \)).

3. Model

Model is hybrid based on zone model [2, 3, 7] and neural network model. In common zone model is problem with enough data in particular zone, because some zone has enough data but different zone is almost empty. For elimination of this problem could be data cutting the way, that each zone is contained same number of data. In zone are data which did not have to be as close as could be for different number of zone. This problem could be solving by one float (moving) zone.

Model has only one floating zone. All data in zone were transformed on elevation \( Z \). Zone is compiled around the last member of real flows series. Zone is constructed by sorting average monthly flows of months of last known member of real flow series and real flows from other months, which they occurred at same time series. Same time series is considered as 1-time step after last month with known real flow. During calibration of model was founded ideal size of zone.

After zone is compiled only data from zone are used for construction of matrix input – output relationship (training matrix), which is used for training of neural network (NS). Training matrix is consisted from targets, which are average monthly flow from forecasting months, and inputs, which they are sorted way that in the first column are average monthly flows from previous past month and in next one are average monthly flows from further past moved by one-time step. NS is learnt by backpropagation method. NS was learnt for every forecasting month (twelve NS). NS has three layers architecture. First hidden layers contain 15 hidden neurons and using sigmoid transformation function. Second hidden layers contain 8 hidden neurons and using sigmoid transformation function. Third layer is outputting layer. Third layer contain number of hidden neurons equals number of inputting monthly average flow to NS and using linear transformation function.

NS Model is given inputs series of flows transformed on elevation \( Z \). In the first step model forecast only one average monthly flow (output). For creation of fun possibilities is used principle of Monte Carlo method. To output are added different random numbers from standard normal
distribution $rnd_{i,j}$ due to this add is created fun of possibilities future flows. Outputs are transformed back to elevation $Q$, which is matching with distribution of month for whose forecast was calculated. Forecast for next month is created by repeating process. Only inputs are changed, because the last real flow is replaced by particular $Z_{i,j}^{-1}$ value of previous forecast flow. For other months is forecast made same algorithm only number of previous forecasted value $Z_{i,j-1}^{-1}$ to $Z_{i,j-k}$ in inputs series is increased as forecasted series is extended. From text above come out that this model is recurrent. Schema of model is showed in figure 1.

**Figure 1.** Hybrid model schema

### 4. Results and discussions

Model was compiled for purpose storage function of large open water reservoir management; therefore, it was made decision that forecasting model would be evaluation on reservoir. Reservoir is not situated in measured profile Bílovice nad Svitavou, due to this reason was made artificially water reservoir. The reservoir has one inflow $Q$, one outflow $O$ and volume on beginning of each time step $V_{\tau-1}$ (initial condition) for all steps of solution. Inflows $Q$ (boundary conditions) and outflows $O$ are described by series of inflows $Q \ (Q_1, Q_2, \ldots, Q_N)$ and managing outflows $O \ (O_1, O_2, \ldots, O_N)$, where $N$ is length of forecast and $\tau$ is time step of forecast. Behaviour of reservoir is described by equation of reservoir (2), therefore solution is provided with monthly step is used differentially shape of equation (2) with advance:

$$Q^\tau - O^\tau = \frac{V^\tau - V^{\tau-1}}{\Delta t}.$$  \ (2)
Management is provided on required (demand) outflow from reservoir $O_p$, which it is in choose scenario constant for all months. Outflow is in interval $(0, O_p)$, but values lower than $O_p$ are supposed to make error in supplement. Longer but shallower error in supplement is considered better than shorter but deeper error in supplement. Boundary conditions $Q$ are replaced by forecasted short series $Q_p$ of average monthly flows in main calculation. After each step is computed new ending step value of volume, which it is in next step starting volume $V_{\tau-1}$ (initial condition).

For evaluation was chosen as main criteria second square error of results given by forecasting model + fuzzy model and course of managing given by real flows series and Genetic algorithm. Genetic algorithm is robust and has ability to find in area of solution $\Omega$ ideal course of managing for criteria function $\pi[5, 9]$:

$$\pi = \sum_{\tau=1}^{N} (O_p - O^\tau)^2 \rightarrow \text{MIN}.$$  

(3)

Management is stochastic and using principle of adaptive. Only first member of calculated series of outflow is used for management. Method Monte Carlo is needed large number repeating calculation for accomplishment good result. In our scenario was used 500 repeating, which is been able to give good results for managing supply function of reservoir [6]. This number of repeating is huge time burden for calculation using Genetic algorithm (GA) as method optimization, therefore was used for management fuzzy model, which is able to give solid results for lesser computing time [6]. Fuzzy model used 4 month’s length forecast.

During evaluation process was searching ideal size of zone and number of output to forecasting model. In figure is on vertical axes time $T$ in months and on horizontal axes are average monthly managing flows using symbol $Q$ [m$^3$/s]. Results are showed in figure 2, that ideal size of zone is in interval $<25, 35>$. If model was used size of zone lower than 25 results of management started worsen. Same case was happened, if size of zone was bigger than 35. Results model using interval size of zone were really similar, therefore they were concluded same quality.

In figure 3 is showed comparing of the best results of forecasting model and fuzzy model with results given by Genetic algorithm and real flow series. Results are showed that the best results were achieved for probability of managing outflow 0.9.
The results in figure 3 are showed, that NS model + fuzzy made unnecessary artificially errors in supplement. In figure 4 is showed course of reservoir volume level for managing by NS and real flow series. From level volume course it can be see, that model was reacted right way, because forecasting
5. Conclusions

Results were showed, that one float zone is enough for compile matrix of input – output relationship and its approximation by NS, which it is able to approximated even strongly nonlinear relationship between months. Application of hybrid forecasting model on stochastic management open water reservoir with supply function showed that the best results were given by stochastic management model if size of float zone was in interval <25, 35>. If zone is contained smaller number of data, matrix of input – output relationship is too small and NS did not find relationship between data properly. If zone is contained too many data, true relationships are lost in useless data.

Stochastic management provided management of reservoir logically and it was able to manage main area of error in supplement. Results given by forecasting + fuzzy were good enough even in compare with real flow series and GA. Forecast of hybrid model is undervalued, which it leads to more aggressive management. Forecasting model in combination with fuzzy model is good tool for this purpose.

Acknowledgment(s)

This paper was supported by specific research project FAST-J-17-4044 Řízení zásobní funkce nádrží při uvažování jistot hydrologických vstupech s použitím metod umělé inteligence s podporou stochastických předpovědních modelů.

References

[1] Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations, Journal of the Royal Statistical Society, Series B, 26, 211-252.
[2] MENŠÍK, P.; STARÝ, M.; MARTON, D., Using Predictive Models of Mean Monthly Flows for Operative Outflows Control from Large Open Reservoirs, conference’s article Proceedings ITISE 2014, International work-conference on Time Series, ISBN 978-84-15814-97-9, Copicentro Granada S.L., Spain, Granada, 2014
[3] MARTON, D.; MENŠÍK, P.; STARÝ, M., Using Predictive Model for Strategic Control of Multi-Reservoir System Storage Capacity, article in Procedia Engineering, ISSN 1877-7058, Elsevier Science Publishers, Amsterdam, 2015
[4] NACHÁZEL, K., STARÝ, M., ZEZULÁK, J.: Využití metod umělé inteligence ve vodním hospodářství (Using method of artificially intelligence in water management), Academia, nakladatelství Akademie věd České republiky, Praha 2004, ISBN 80-200-0229-4-9.
[5] KOZEL, T.; STARÝ, M., Stochastic Management of the Open Large Water Reservoir with Storage Function with Using a Genetic Algorithm, conference’s article World multidisciplinary earth sciences symposium abstract collection, ISBN 978-80-270-0316-7, iposcience, Praha, 2016
[6] KOZEL, T.; STARÝ, M., STOCHASTIC MANAGEMENT OF THE OPEN LARGE WATER RESERVOIR WITH STORAGE FUNCTION WITH USING FUZZY LOGIC, conference’s article SGEM Conference Proceedingsc, ISSN 1314-2704, ISBN 978-619-7105-61-2, STEF92 Technology Ltd., 51 Alexander Malinov Blvd., 1712, Sofia, Bulgaria, 2016
[7] KOZEL, T.; STARÝ, M., ZONE STOCHASTIC FORECASTING MODEL FOR MANAGEMNT OF LARGE OPEN WATER RESERVOIR WITH STORAGE FUNCTION, conference’s article SGEM Conference Proceedingsc, ISSN 1314-2704, ISBN 978-619-7105-61-2, STEF92 Technology Ltd., 51 Alexander Malinov Blvd., 1712, Sofia, Bulgaria, 2016
[8] HAUPT, R. L. and HAUPT, S. E. (2003) Appendix II: MATLAB Code, in Practical Genetic Algorithms, Second Edition, John Wiley & Sons, Inc., Hoboken, NJ, USA. doi: 10.1002/0471671746.app2