Environmental Modelling & Software 59 (2014) 162–186

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Selection of smoothing parameter estimators for general regression neural networks – Applications to hydrological and water resources modelling

Xuyuan Li*, Aaron C. Zecchin, Holger R. Maier

School of Civil, Environmental and Mining Engineering, The University of Adelaide, Adelaide, South Australia 5005, Australia

ARTICLE INFO

Article history:
Received 23 December 2013
Received in revised form
5 May 2014
Accepted 6 May 2014
Available online xxx

Keywords:
General regression neural networks
Smoothing parameter estimators
Artificial neural networks
Multi-layer perceptrons
Extreme and average events
Hydrology and water resources

ABSTRACT

Multi-layer perceptron artificial neural networks are used extensively in hydrological and water resources modelling. However, a significant limitation with their application is that it is difficult to determine the optimal model structure. General regression neural networks (GRNNs) overcome this limitation, as their model structure is fixed. However, there has been limited investigation into the best way to estimate the parameters of GRNNs within water resources applications. In order to address this shortcoming, the performance of nine different estimation methods for the GRNN smoothing parameter is assessed in terms of accuracy and computational efficiency for a number of synthetic and measured data sets with distinct properties. Of these methods, five are based on bandwidth estimators used in kernel density estimation, and four are based on single and multivariable calibration strategies. In total, 5674 GRNN models are developed and preliminary guidelines for the selection of GRNN parameter estimation methods are provided and tested.

© 2014 Elsevier Ltd. All rights reserved.

Software availability

Software name: GRNNs
Developer: Xuyuan Li, Postgraduate Student, the University of Adelaide, School of Civil, Environmental & Mining Engineering, Adelaide, SA 5005, Australia
Phone: +61 8 8313 1575
Fax: +61 8 8303 4359
Email: xli@civeng.adelaide.edu.au
Hardware requirements: 64-bit AMD64, 64-bit Intel 64 or 32-bit ×86 processor-based workstation or server with one or more single core or multi-core microprocessors; all versions of Visual Studio 2012, 2010 and 2008 are supported except Visual Studio Express; 256 MB RAM
Software requirements: PGI Visual Fortran 2003 or later version
Language: English
Size: 4.74 MB
Availability: Free to download for research purposes from the following website:http://www.ecms.adelaide.edu.au/civeng/research/water/software/generalised-regression-neural-network/

* Corresponding author. Tel.: +61 8 8313 1575; fax: +61 8 8303 4359.
E-mail addresses: xli@civeng.adelaide.edu.au, xliadelaide@gmail.com, li-xuyuan@163.com (X. Li), aaron.zecchin@adelaide.edu.au (A.C. Zecchin), holger.maier@adelaide.edu.au (H.R. Maier).

http://dx.doi.org/10.1016/j.envsoft.2014.05.010
1364-8152/© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Over the last two decades, artificial neural networks (ANNs) have been used extensively in the field of hydrological and water resources modelling, and their popularity is still increasing (Maier et al., 2010; Abrahart et al., 2012; Wu et al., 2014). In the vast majority of these applications, multi-layer perceptrons (MLPs) have been used as the most common model architecture (Maier et al., 2010; Wu et al., 2014). While the use of MLPs has generally resulted in good model performance, their development is complicated by the fact that there are no rigorous methods for determining an appropriate model structure. Determination of the optimal number of hidden nodes is especially difficult, unless sophisticated Bayesian approaches are used (Kingston et al., 2008; Zhang et al., 2011), which are computationally demanding and require substantial technical expertise to implement. Therefore, the optimal model structure is generally determined by trial and error (Maier et al., 2010; Wu et al., 2014). This process usually involves a number of

steps, including (i) selection of a trial model structure, (ii) calibration of the model with the selected structure, and (iii) evaluation of the predictive performance of the calibrated model. These steps are repeated for models with different trial structures and the model structure that results in the best predictive performance is selected. Consequently, the model structure that is found to be optimal is a function of a number of factors, including:

(i) *The trial model structures selected for evaluation:* As the potential number of different model structures is generally large, the performance of a subset of all possible structures is usually evaluated. This can be achieved using different approaches, including ad-hoc, stepwise (e.g. constructive, pruning) or global approaches (Maier et al., 2010). Consequently, as different approaches generally result in the evaluation of different model structures, the structure obtained is a function of the adopted approach.

(ii) *The calibration method used:* The predictive performance of a model with a particular structure is a function of the quality of the calibration (training) process. Finding the combination of model parameters (connection weights) that gives the best predictive performance for a given network structure is complicated by the presence of a large number of local optima in the error surface (Kingston et al., 2005a). This is particularly the case if gradient-based calibration (training) methods are used (Maier and Dandy, 1999), such as the most commonly used back-propagation algorithm (Maier et al., 2010; Wu et al., 2014). In addition to the choice of calibration (training) methods, the parameters that control the searching behaviour of these methods (e.g. learning rate and momentum when the back-propagation algorithm is used) can also have a significant impact on the best predictive model performance obtained for a particular model structure (Maier and Dandy, 1998a,b). Consequently, unless the predictive performance that corresponds to the global optimum in the error surface can be identified for all models with different structures, it is not possible to identify which model structure results in the best predictive performance with certainty. As a result, the optimal model structure obtained is a function of the quality of the model calibration process.

(iii) *The calibration data used:* The available data are generally split into different subsets for calibration (training) and validation, which can be done using a number of different methods (see Maier et al., 2010). Consequently, which data points are included in the different subsets can vary, depending on which data division method is used (Bowden et al., 2002; May et al., 2010; Wu et al., 2012, 2013). This can also have an impact on which model structure is found to result in the best predictive performance. This is because different data points will result in different error surfaces during calibration, thereby potentially affecting calibration difficulty [see (ii)] and producing different global and local optima, which is likely to change which model structure results in the lowest error.

Given the factors described above, it is generally not possible to isolate the impact of model structure on the predictive performance of MLPs, making it difficult to know which model structure should be used. In addition, the trial-and-error process generally used to determine the optimal structure of MLPs is computationally expensive, as it necessitates the development of a potentially large number of models.

Although there are other alternative ANN based approaches, including Radial Basis Functions (RBFs) (Buhmann, 2003), Recurrent Neural Networks (RNNs) (Williams and Zipser, 1989) and Probabilistic Neural Networks (PNNs) (Specht, 1990), General regression neural networks (GRNNs) (Specht, 1991) provide an alternative ANN model structure that has been shown to perform well in a number of studies in water resources applications (e.g. Bowden et al., 2005b, 2006; Gibbs et al., 2006; Cigizoglu and Alp, 2006) and overcomes the shortcomings associated with MLPs discussed above, as the structure of GRNNs is fixed (Bowden et al., 2005a). This removes the ambiguity associated with determining which model structure is optimal. In addition, it increases the computational efficiency of the model development process, as there is no need to develop a number of models with different structures in order to determine which is optimal.

However, a potential issue with the application of GRNNs to hydrological and water resources problems is that there has been limited work on determining which smoothing parameter estimation methods should be adopted. As GRNNs are essentially a Nadaraya-Watson kernel regression method (Cai, 2001), parameter estimation only involves the determination of optimal values of one or more smoothing parameters, also known as kernel bandwidths. However, this is not a trivial issue, as illustrated by the vast amount of literature on kernel bandwidth estimation as applied to density estimation (e.g. Rudemo, 1982; Bowman, 1984; Scott and Terrell, 1987; Park and Marron, 1990; Hall et al., 1992; Wand and Jones, 1995). Overestimating the smoothing parameter can result in over-smoothing of the estimated density (i.e. kernel based probability density function (PDF)). In this case, the detailed local information (for instance the variation of daily rainfall in hydrological applications) will not be captured in the estimated density. In contrast, if values of the smoothing parameter are underestimated, the general trend of the estimated density (for instance the overall rainfall trend within a given time period) can be disturbed by localised features or noise.

Among the extensive literature on smoothing parameter (or kernel bandwidth) estimation in other areas of research, such as mathematics and statistics, there are a number of different approaches to obtaining optimal estimates of kernel density, which are based on assumptions about the form of the PDF and different fitness function types (i.e. the objective function on which the estimator is based). Consequently, their relative merits for determining the optimal values of the smoothing parameters for water resources GRNN models are likely to vary from case study to case study, depending on the distribution of the data and the modelling objective function used. However, the relationship between the performance of GRNNs with smoothing parameters obtained using different kernel density estimation methods and the properties of the water resources data used to develop them has not been considered previously, making it difficult to know which methods to use for particular case studies.

Therefore, the objectives of the current study are: (i) to compare the performance, in terms of both predictive accuracy and computational cost, of GRNN models for which smoothing parameters have been estimated using a range of methods, as well as that of a benchmark MLP model, for case studies with data that have varying degrees of normality, linearity and different modelling objectives (e.g. matching average or extreme events); and (ii) to develop and test empirical guidelines for the selection of the most appropriate methods for GRNN smoothing parameter estimation based on the properties of the available data (i.e. degree of normality and non-linearity) and the modelling objective.

The remainder of this paper is organised as follows. A brief introduction to GRNNs is provided in Section 2, followed by the

Download English Version:

https://daneshyari.com/en/article/6963800

Download Persian Version:

https://daneshyari.com/article/6963800

Daneshyari.com