Available online at www.sciencedirect.com
ScienceDirect
Neural Networks 20 (2007) 444–453
Neural Networks
www.elsevier.com/locate/neunet
2007 Special Issue

Nonlinear analog predictor analysis: A coupled neural network/analog model for climate downscaling

Alex J. Cannon*

Meteorological Service of Canada, Environment Canada, 201-401 Burrard Street, Vancouver, BC, V6C 3S5, Canada

Abstract

Synoptic downscaling models are used in climatology to predict values of weather elements at one or more stations based on values of synoptic-scale atmospheric circulation variables. This paper presents a hybrid method for climate prediction and downscaling that couples an analog, i.e., *k*-nearest neighbor, model to an artificial neural network (ANN) model. In the proposed method, which is based on nonlinear principal predictor analysis (NLPPA), the analog model is embedded inside an ANN, forming its output layer. Nonlinear analog predictor analysis (NLAPA) is a flexible model that maintains the ability of the analog model to preserve inter-variable relationships and model non-normal and conditional variables (such as precipitation), while taking advantage of NLPPA's ability to define an optimal set of analog predictors that maximize predictive performance. Performance on both synthetic and real-world hydroclimatological benchmark tasks indicates that the NLAPA model is capable of outperforming other forms of analog models commonly used in synoptic downscaling.
Crown Copyright © 2007 Published by Elsevier Ltd. All rights reserved.

Keywords: Analog; *k*-nearest neighbor; Neural network; Downscaling; Climatology; Multivariate; Forecasting; Hydroclimatology

1. Introduction

Synoptic downscaling models are used to predict local-scale climate variables (e.g., station temperatures or precipitation amounts) from synoptic-scale atmospheric circulation data (e.g., sea-level pressure or geopotential height fields), often as a means of creating climate scenarios at high temporal and spatial resolution, e.g., for assessing the impacts of climate change on a watershed, crop, or ecosystem (Wilby & Wigley, 1997; Xu, 1999). Synoptic downscaling techniques may be physically based, for example using limited-area numerical weather prediction models, or they may be statistically based, for example using linear or nonlinear regression models, stochastic weather generators, or analog models. Statistical models are attractive because they require few computational resources and are capable of performing as well as more complicated physically based approaches (Hellstrom, Chen, Achberger, & Raisanen, 2001; Murphy, 1999, 2000; Wood, Leung, Sridhar, & Lettenmaier, 2004).

Linear regression models offer a simple and direct way to link the synoptic-scale forcing with the local climatic response, and, as a result, they are commonly used in statistical climate downscaling. Most regression-based downscaling models are developed for a variable, e.g., precipitation or temperature, at an observing site. If many variables are required at many sites, different regression equations are usually developed for each separately. Using this approach, the model developer has little control over the consistency of outputs between sites and variables, perhaps only through the specification of a common set of synoptic-scale predictors. Maintaining realistic relationships between sites and variables is particularly important in hydroclimatological applications, as streamflow depends strongly on the spatial distribution of precipitation in a watershed and on interactions between precipitation and temperature that determine whether precipitation falls as rain or snow.

Recent studies have addressed multi-site downscaling by moving away from classical multiple regression models to multivariate linear models, e.g., canonical correlation analysis, CCA, or maximum covariance analysis, MCA (Uvo et al., 2001). Others have extended multivariate linear models by adopting cost functions that force the covariance matrix of

* Tel.: +1 604 664 9245; fax: +1 604 664 9004.
E-mail address: alex.cannon@ec.gc.ca.

0893-6080/$ - see front matter Crown Copyright © 2007 Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2007.04.002

the model predictions to match observations, e.g., expanded downscaling (Bürger, 1996, 2002). While these methods may be better suited to spatial downscaling than standard multiple regression models, they may fail when nonlinear relationships are present, for example when trying to predict daily precipitation amounts (Yuval & Hsieh, 2002).

Flexible nonlinear models, such as artificial neural networks (ANNs), may perform better than the classical linear approaches when nonlinear relationships are present. Comparisons between multivariate ANNs (i.e., ANNs with multiple outputs) and ANNs with a single output have also demonstrated the potential of the multivariate approach, due in part to the sharing of weights between outputs of the multivariate neural network (Caruana, 1997). It is conceivable that inter-site correlations would be modeled more accurately using this approach. However, constraints similar to the one used in expanded downscaling would likely still be needed to ensure close correspondence between modeled and observed spatial relationships. Modeling a variable like precipitation can also be a challenge, as the amount of precipitation is conditional on the occurrence of precipitation. Occurrences and amounts thus need to be modeled sequentially, which adds an additional layer of complexity to the modeling process.

Instead, analog (i.e., k-nearest neighbor) models have often been used for downscaling tasks, particularly for the prediction of precipitation at multiple sites (Zorita & von Storch, 1999). The analog model is nonlinear, and, due to the sampling process, is capable of preserving spatial correlations between sites and handling conditional variables like precipitation. Analog models are, however, prone to the 'curse of dimensionality' (Hastie, Tibshirani, & Friedman, 2001), which means that they may perform poorly when the dimensionality of the dataset is high and the number of cases is low. To mitigate this problem, predictor scaling algorithms have been developed that give more weight to relevant variables than irrelevant or redundant variables, thereby reducing the overall size of the predictor space (Fraedrich & Ruckert, 1998; Mehotra & Sharma, 2006). Alternatively, dimensionality reduction can be undertaken by linear models that create a smaller set of analog predictors from the original predictors. For example, principal component analysis (PCA) (Zorita & von Storch, 1999), CCA (Fernandez & Saenz, 2003), and multiple linear regression (Wilby, Tomlinson, & Dawson, 2003) have been used to reduce the dimensionality of the predictors, with outputs from the linear models acting as inputs to the analog model. Results have shown this to be a promising avenue for multi-site downscaling research.

Given the improved performance of multivariate ANNs relative to multivariate linear models, and the recent work into nonlinear versions of classical multivariate linear models, such as nonlinear PCA (NLPCA) and nonlinear CCA (NLCCA) (Hsieh, 2004), a logical next step would be to combine an ANN with an analog model in a manner similar to what has been done previously for linear models. In the current study, an approach that couples an ANN to an analog model is developed. The coupled model, referred to as nonlinear analog predictor analysis (NLAPA), is based on nonlinear principal predictor analysis (NLPPA) (Cannon, 2006), which is a nonlinear multivariate model that is closely related to both NLPCA and NLCCA. Rather than creating ANN and analog models separately, the analog model in NLAPA is instead coupled to an ANN in such a way that the ANN portion of the model is used to define the predictors for the analog model. The end result is a flexible model that maintains the ability of the analog model to preserve inter-variable relationships and to model non-normal and conditional variables like precipitation, while taking advantage of NLPPA's ability to define an optimal set of analog predictors that results in improved predictive performance.

The remainder of the paper is organized as follows. First, Section 2 describes the analog model and the NLPPA model; this is followed by the development of the NLAPA model. Next, Section 3 describes the synthetic benchmark tasks and compares the performance of the NLAPA model with standard analog models. Section 4 applies the NLAPA model to two real-world downscaling tasks, both focused on multivariate predictand datasets that are of relevance in hydroclimatology. The first involves downscaling of temperature, rainfall, and snowfall at a station near Whistler, British Columbia (BC), Canada; the second involves downscaling of precipitation at multiple sites in coastal BC. Finally, results are discussed in Section 5 along with conclusions and recommendations for future research directions.

2. Method

2.1. Analog model

The analog model is a data-driven nonparametric algorithm that relates a set of analog predictors (or model inputs) to a set of predictands (or model outputs) by assuming that cases that appear "close" to one another in predictor space share similar predictand values.

Given a set of L analog predictors observed at time t, $\mathbf{p}(t) = [p_1(t), \ldots, p_L(t)]$, the analog model predicts values of the K predictands at time t, $\hat{\mathbf{y}}(t) = [\widehat{y}_1(t), \ldots, \widehat{y}_K(t)]$ (where the hat denotes the predicted value of a variable), by searching through N historical observations of the analog predictors, $\mathbf{p}_l^{(h)} = [p_l^{(h)}(1), \ldots, p_l^{(h)}(N)]^T (l = 1 \ldots L)$, finding the time u with analog predictors $\mathbf{p}^{(h)}(u)$ that are closest to $\mathbf{p}(t)$ in terms of weighted Euclidean distance D,

$$D = \sqrt{\sum_{l=1}^{L} s_l [p_l(t) - p_l^{(h)}(u)]^2}, \tag{1}$$

and then taking the historical predictand values at time u, $\mathbf{y}^{(h)}(u)$, as the predicted values at time t, $\hat{\mathbf{y}}(t)$. The scaling factor $\mathbf{s}$ is included to allow different weights to be assigned to each analog predictor. Together, this series of operations will be referred to as the analog operator A.

If predictions are based on a single historical analog, A can preserve spatial correlations between multiple sites. The analog operator can be extended to use more than one analog or neighbor by taking predictions to be the weighted mean or

Download English Version:

https://daneshyari.com/en/article/404828

Download Persian Version:

https://daneshyari.com/article/404828

Daneshyari.com