Journal of Theoretical Biology ■ (■■■■) ■■■–■■■

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

Letter to Editor

On estimating the reliability of ecological forecasts

ARTICLE INFO

Keywords:
Model checking
Nonparametric forecasting
Observation error
Time series analysis
Deterministic chaos
Nonlinear dynamics

ABSTRACT

Recent work has highlighted the utility of nonparametric forecasting methods for predicting ecological time series (Perretti et al., 2013. Proc. Natl. Acad. Sci. U.S.A. 110, 5253–5257). However, one topic that has received considerably less attention is the quantification of uncertainty in nonparametric forecasts. This important topic was brought to the forefront in the recent work by Jabot (2014. J. Theor. Biol.). Here, we add to this emerging discussion by reviewing the available methods for quantifying forecast uncertainty in nonparametric models. We conclude with a demonstration of one such method using the simulation model of Jabot (2014. J. Theor. Biol.). We find that nonparametric forecast error is accurately estimated with as few as 10 observations in the time series.

1. Introduction

Ecological prediction has been a topic of much recent interest (e.g., Clark et al., 2001; Glaser et al., 2013; Lindegren et al., 2010). The call for predictions is well founded, as there is a pressing need for anticipating how ecosystems will be impacted by a rapidly changing world. However, numerous methods exist for creating predictions, and an important open question is the relative efficacy of each method.

Methods for generating ecological predictions can be broadly divided into two categories: parametric and nonparametric. In the parametric approach, the modeler starts with a set (or sets) of plausible mechanisms that generate the dynamics and converts these into equations. Unknown parameters for these equations are calibrated to available data using increasingly complex methods, and predictions are made based on the calibrated model. Parametric models are often used for estimating the underlying 'true state' of the system, which is valid to the extent that the chosen model matches reality. Using such 'mechanistic' models with biologically interpretable parameters provides an intuitively obvious way of evaluating various 'what-if' scenarios.

Nonparametric approaches skip the step of converting mechanisms into equations, instead attempting to infer the dynamics directly from time series data. There are a wide variety of nonparametric models available, and they generally operate by estimating the relationships between variables using either a locally-weighted smoothing algorithm (e.g., Härdle and Vieu, 1992; Sugihara, 1994) or basis expansion (e.g., Ellner and Turchin, 1995; Hill et al., 1996). They can be extended to non-stationary time series after pre-processing with a detrending procedure such as first-differencing and a box–cox transformation (Box and Cox, 1964). Nonparametric models produce forecasts that are robust to model misspecification at the cost of reduced interpretability.

Jabot presents two arguments in favor of using parametric models over nonparametric approaches. The first is that parametric models permit 'model checking' and the second is that 'no methods are yet available to assess the reliability of nonparametric forecasts.' As we have argued elsewhere (Munch et al., 2005; Perretti et al., 2012; Thorson et al., 2013) we think that avoiding the need for model checking is a feature, not a pitfall, of nonparametric methods. Here we address Jabot's second argument.

Forecast error is measured as the discrepancy between predictions and observations. This differs from state-estimation error, in which one is concerned with estimating the 'true' state of a system given a set of observations. A 'reliable' forecast method is one that is able to accurately estimate its out-of-sample forecast error, where out-of-sample is defined as observations that are not used in the model-fitting procedure (Tashman, 2000). Fortunately, there are, in fact, a number of procedures available for estimating expected forecast error that are equally applicable to parametric and nonparametric models. The methods utilize partitioning and resampling procedures to estimate error, and unlike the parametric method proposed in Jabot (2014), they do not require the modeler to sift through a collection of mechanistic models in search of the best-fitting model.

Here, we provide a brief overview of some commonly used methods for estimating forecast error. Then, using the theta-logistic simulation model described in Jabot, we demonstrate the ability of one such method to accurately estimate expected forecast error from nonparametric forecasts.

2. A review of some nonparametric methods for estimating expected forecast error

The typical procedure for estimating forecast error is to divide the available data into a training set and a test set. The model is fit using the training set, and forecast error is evaluated on the test set.

http://dx.doi.org/10.1016/j.jtbi.2015.02.031
0022-5193/

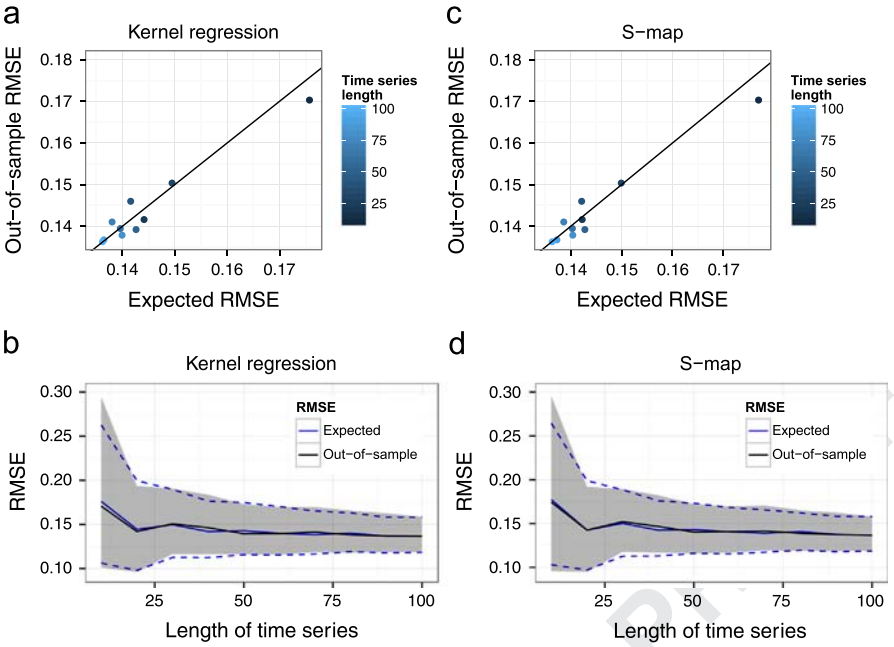

Fig. 1. Estimated nonparametric forecast error vs. out-of-sample forecast error. The median estimated forecast error closely matches the median out-of-sample forecast error for both the kernel regression and the S-map methods (a and c). The median, 10th and 90th percentiles are also accurately estimated by each method, and they become increasingly accurate with additional data (b and d).

The training set may be further divided into a learning set and validation set if a model selection step is required.

The simplest method of estimating forecast uncertainty is the *split-sample method* in which the available data is split into a single training set and test set. The prediction model is built using the training set, and expected forecast error is evaluated by predicting the test set (e.g., Ward et al., 2014). This method is computationally efficient, however, it may overestimate expected forecast error due to its incomplete usage of the available data (Molinaro et al., 2005).

Similar to the split-sample method, *k-fold cross-validation* divides the data into a training set and test set. However, instead of splitting the data once, it is partitioned into k equal sections, and each section is treated as a test set, with the remaining sections treated as the training set. In this procedure all portions of the training set are used to estimate forecast error which helps reduce the overestimation of expected error associated with the split-sample procedure. Leave-one-out cross-validation is obtained by setting k equal to the number of observations in the dataset, and is asymptotically equivalent to the Akaike Information Criterion for cross-sectional data (Stone, 1977). Similarly, leave-k-out is asymptotically equivalent to the Bayesian Information Criterion for linear models, where $k=n[1-1/(\log(n)-1)]$ and n is the number of observations (Shao, 1997). Simulation studies suggest that the optimal number of partitions for model selection is typically between five and ten (Kohavi, 1995).

Another variation on the k-fold scheme is *Monte Carlo cross-validation*. In this approach, the k-fold process is applied by randomly assigning points to each of the partitions and repeating this process many times. This has the advantage of generating a large number of simulated test sets, although the computational burden can become exceedingly large.

Finally, a bootstrap resampling method can be utilized to estimate expected forecast error. In one such scheme, known as the *0.632 bootstrap*, the observations are sampled from the original data with replacement to form a bootstrapped training set. The model is then fit to the bootstrapped training set, and expected forecast error is evaluated using a test set containing only data that was not included in the bootstrapped training set. This method is implemented in commonly used *R* packages such as *randomForests* and *e1071*. Although this bootstrap procedure has been shown to perform well for model selection, it will typically overestimate forecast error, as the proportion of unique observations in the bootstrapped training set will only be $\sim$0.632. Therefore the expected error is obtained by multiplying the test error by 0.632 and adding 0.368 of the training error (Efron and Gong, 1983; also see (Efron and Tibshirani, 1997).

3. A demonstration

Here, we use kernel regression and S-map methods to predict time series generated by the theta-logistic model used in Jabot (2014). Specifically, for the kernel regression we use a Gaussian kernel with a local-constant regression estimator (Nadaraya, 1964), with the bandwidth automatically selected using least-squares cross-validation (Racine and Li, 2004). For the S-map (Sugihara, 1994) we employ the same model structure as Jabot (2014), and fit it using leave-one-out cross-validation. We estimate expected forecast error using the simplest of the methods above in which the time series is divided evenly into a training set and a test set. Forecast error is quantified by root mean squared error (RMSE) rather than standardized RMSE (SRMSE) as was used in Jabot (2014), as SRMSE confounds the uncertainty in forecast error with the uncertainty in the variance of the observations.

Simulated time series were generated following the setup of Jabot (2014) using the theta-logistic model with parameters $\theta=1$, $r=3.7$, and $\sigma_{obs}=0.1$ (i.e., the logistic model with "medium" noise from Jabot (2014)). Kernel regression forecasts were generated using the *np* package (Hayfield and Racine, 2008) in *R* (see Appendix for code). To explore the relationship between the length of the time series and the accuracy of the expected prediction error, time series length varied from 10 to 100 observations by increments of 10. The expected prediction error was then compared to the true out-of-sample prediction error for a time series of length equal to the test set. For each time series length, 1000 replicates were performed to calculate the median, 10th, and 90th percentiles of the expected prediction error and out-of-sample prediction error.

In contrast to the discussion of Jabot (2014), we find that the uncertainty in the nonparametric prediction error is well estimated