HAL
open science

Why preferring parametric forecasting to nonparametric methods?

F. Jabot

To cite this version:

F. Jabot. Why preferring parametric forecasting to nonparametric methods?. Journal of Theoretical Biology, 2015, 372, pp.205-210. 10.1016/j.jtbi.2014.07.038 . hal-01328031

HAL Id: hal-01328031

https://hal.science/hal-01328031v1

Submitted on 7 Jun 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Author-produced paper in Journal of Theoretical Biology, 2015, 372, 205-210. doi:10.1016/j.jtbi.2014.07.038,
Original paper available from http://www.sciencedirect.com

Why preferring parametric forecasting to nonparametric methods?

Franck Jabot

Laboratoire d'Ingénierie pour les Systèmes Complexes, IRSTEA, 9 avenue Blaise Pascal, CS 20085, 63178 Aubière, France. franck.jabot@irstea.fr

Abstract

A recent series of papers by Charles T. Perretti and collaborators have shown that nonparametric forecasting methods can outperform parametric methods in noisy nonlinear systems. Such a situation can arise because of two main reasons: the instability of parametric inference procedures in chaotic systems which can lead to biased parameter estimates, and the discrepancy between the real system dynamics and the modeled one, a problem that Perretti and collaborators call “the true model myth”. Should ecologists go on using the demanding parametric machinery when trying to forecast the dynamics of complex ecosystems? Or should they rely on the elegant nonparametric approach that appears so promising? It will be here argued that ecological forecasting based on parametric models presents two key comparative advantages over nonparametric approaches. First, the likelihood of parametric forecasting failure can be diagnosed thanks to simple Bayesian model checking procedures. Second, when parametric forecasting is diagnosed to be reliable, forecasting uncertainty can be estimated on virtual data generated with the fitted to data parametric model. In contrast, nonparametric techniques provide forecasts with unknown reliability. This argumentation is illustrated with the simple theta-logistic model that was previously used by Perretti and collaborators to make their point. It should convince ecologists to stick to standard parametric approaches, until methods have been developed to assess the reliability of nonparametric forecasting.

Keywords: chaotic dynamics, model checking, nonparametric forecasting, observation error, time series analysis.

Introduction

The societal demand for predictions in ecology poses a major challenge to the scientific community. It triggers the need to design efficient methodologies that can tackle the inherent complexity of ecosystems. Most predictive exercises in ecology so far have relied on parametric approaches, that is on the design of a model, on the estimation of model parameters that best fit available data, and on the use of the calibrated model to make predictions. Parametric methods encompass widely different approaches to predictions, from correlative approaches, in which the model is a statistical one (e.g., Austin 2002), to detailed individual-based modeling approaches (Grimm and Railsback 2005, Hartig et al. 2011, Jabot et al. 2013). In this general methodology, the generally accepted view is that by refining our understanding of the processes at stake in the dynamics of a system, we will improve our ability to make reliable predictions. This is one of the arguments advanced in favor of complex mechanistic modeling in ecology (Evans et al. 2013).

In contrast with this general approach, Perretti and colleagues have argued in a series of recent papers that model-free nonparametric approaches may be a better way to tackle ecosystem complexity (Perretti et al. 2013a,b,c). These nonparametric approaches rely on the idea that one should be able to forecast future dynamics by analogy with past dynamics. More precisely, a phase space representation of past dynamics is used, in which variables of interest are plotted as a function of each other, possibly with time delays (Kantz and Schreiber 2004). Forecasts are then built using the collection of previously observed dynamics, weighted by their proximity in the phase space (see Materials and Methods and Perretti et al. 2013a,b for additional details).

Perretti et al. (2013a,b) showed that such a nonparametric approach to forecasting could outperform parametric forecasts in some stochastic dynamical systems in chaotic regimes, even in situations in which the parametric forecast was based on the same model that was used to generate the data. This poor performance of parametric forecasting was due to two main reasons. The first reason is the well known instability of parametric inference in chaotic systems (Berliner

1991, Wood 2010). Hartig and Dormann (2013) recalled an alternative method to calibrate the model which is based on a segmentation of the time series to circumvent the sensibility to initial conditions encountered in chaotic regimes (Pisarenko and Sornette 2004). Armed with this trick, Hartig and Dormann (2013) have shown that the parametric forecast is more accurate than the nonparametric one in the case where the same model was used to generate the data and to ground the forecast.

The second reason for parametric forecasting failure was called by Peretti et al. (2013c) the "true model myth", and comes down to the very meaning of what a model is (Box 1976). Indeed, there is a necessary mismatch between reality and its simplified modeled representation. Perretti et al. (2013c) therefore repeated their analysis, using the theta-logistic model (see Methods) to generate a virtual dataset, and basing the parametric forecast on the simpler logistic model. In this setting, and using the inference method proposed by Hartig and Dormann (2013), they found that their nonparametric forecast was more accurate than the logistic based forecast.

As already argued by Hartig and Dormann (2013), parametric approaches offer the advantages of transferability and theoretical understanding which are lacking in nonparametric methods. They further enable to mobilize previous knowledge on the system studied through prior distributions in Bayesian analyses. The aim of this contribution is to demonstrate that parametric approaches are currently also preferable in a forecasting perspective, using the theta-logistic example of Perretti et al. (2013c). Two key comparative advantages of parametric approaches over nonparametric ones will be detailed. First, the likelihood of parametric forecasting failure can be diagnosed thanks to simple Bayesian model checking procedures. Indeed, the convergence of a parametric inference method does not ensure that the fitted model is a satisfactory representation of the data. Simple methods are available to check, with minor computing costs, the ability of the fitted model to recover major data characteristics (Rubin 1984, Gelman et al. 1996). In cases where the fitted model fails at such model checking tests, modifications of the model should be sought, before any reliable predictions can be made. Second, when parametric forecasting is

diagnosed to be reliable, forecasting uncertainty can be estimated thanks to virtual data generated with the fitted to data parametric model. This second advantage enables to complement a forecast with a detailed understanding of its inherent limitations. These two key advantages enable scientists to enrich their forecasts with a description of their reliability.

Methods

The theta-logistic model

In the following, the theta-logistic model with observation error is used to generate time series of population abundances. These time series represent the observed "reality" that we are trying to forecast. The deterministic theta-logistic model is defined by the following difference equation:

$$X(t+1) = X(t) \times r \times [\ 1 - (\ X(t) / K\ )^{\theta}\ ] \qquad (1)$$

where X(t) is the true population abundance at time t, r is the population growth rate, K is the carrying capacity and θ is a model parameter. When θ equals 1, the model is the well known logistic model of population dynamics, which can generate deterministic chaos (May 1974).

Equation (1) describes the true dynamics of the population, while observers only have access to noised estimations of population size N(t). Following Perretti et al. (2013b), the observation error is modeled here using:

$$N(t) = X(t) \times \exp[\ \varepsilon(t)\ ] \qquad (2)$$

where ε(t) are independent normal random variables with mean 0 and standard deviation σ_{obs}, leading to lognormal observation error.

Equations (1) and (2) are used to generate artificial data. Six values of θ were used {1; 1.1; 1.2; 1.5; 2; 4} with corresponding r values {3.7; 3.7; 3.5; 3; 2.5; 1.8} to ensure a chaotic regime and population survival. Three levels of noise σ_{obs} were used: 0.05, 0.1 and 0.2 (represented in Figure S1), leading to an increase of the temporal variance in observed population sizes N(t) of 3, 12 and 52% respectively, compared to the temporal variance of real population sizes X(t). This range of observation noise level is in line with previous theoretical studies (de Valpine and Hastings 2002) and with reported values estimated on the global population dynamics database (Knape and de Valpine 2012). For each triplet of values (θ; r; σ_{obs}), 50 time series of 100 time steps were generated, starting from an initial population size X(1) equal to 0.5, so that 900 time series were generated in total.

Parametric model fit

Each of these 900 time series was fit with the logistic model, which is the special case when θ equals 1 in equation 1. Only the 50 first steps of the time series were used for calibration, and the 50 remaining steps were used to assess the quality of the forecasts. Following Hartig and Dormann (2013), the time series were split into 10 subsets of 5 time steps each, and the X(t) values at the beginning of each subset were considered unknown. The rationale of this splitting procedure is to focus on short subset of the time-series so as to decrease the impact of the sensitivity to initial conditions which is specific to chaotic dynamics. The length of the splitting procedure was guided by Hartig and Dormann (2013) who have already shown that it was efficient for this parametric model. For other case studies, alternative splitting length might be more efficient. They could be compared by trial and error. Three MCMC chains were used with dispersed initial r values, 500,000 adaptation steps, 500,000 burn-in iterations and 1,000,000 posterior samples, of which 10,000 (1 every 100 steps) were retained to perform model checks and compute predictions. Prior distributions were uniforms on [2;4.5], [0.01;10] and [0.1;40] respectively for r, K and the observation precision equal to $1/\sigma_{obs}^2$. This last prior distribution was changed to uniform on [10;600] for cases of low and middle observation noise (σ_{obs} equal to 0.05 or 0.1 in the simulations). Convergence of the chains was checked with the multivariate potential scale reduction factor, that was required to be smaller than 1.2. Bayesian analysis was

performed with R version 3.0.1 and JAGS 3.4 (Plummer 2003). R and JAGS scripts are provided in the Appendix.

Model checking

The 10,000 parameter sets sampled in the MCMC were used to simulate 10,000 posterior predictive datasets with observation error of 50 time steps (for each of the 900 fitted time series). Note that this implies to have additionally recorded during the MCMC the fitted initial population size X(1). Five summary statistics were computed on each posterior predictive dataset and on the "real" fitted time series: the average observed population size $\overline{N}$, the standard deviation of observed population sizes through time $\hat{\sigma}_N$, and the autocorrelation coefficients R(τ) of observed population sizes with time lag τ equal to 1, 2 and 3 ($R(\tau) = \frac{1}{N-\tau}\sum_{t=\tau+1}^{N}\left[\left(N(t)-\overline{N}\right)\times\left(N(t-\tau)-\overline{N}\right)\right]/\hat{\sigma}_N^2$). For each summary statistic, a p-value was computed as the percentage of posterior predictive statistics below the value computed on the real dataset. The minimum p_{min} of these 5 p-values (or of 1 – p-value in cases it is lower than the p-value) is reported in the result section.

Posterior predictive p-values are not uniformly distributed between 0 and 1, but instead tend to be peaked around 0.5 (Robins et al. 2000, Gelman 2013). Interpreting them as standard p-values would thus lead to too liberal model checking procedures. A second difficulty arises here since five p-values are combined in p_{min}. Therefore, model checking statistics are not directly interpreted. Instead, each posterior predictive p-value is compared to the ones obtained on data sets generated with the model used for the fit. If the posterior predictive p-value obtained with the real data is lower than those obtained on data sets generated with the model, this will indicate a poor model fit quality of the real data.

We also tested the method proposed by Crespi and Boscardin (2009) to perform model checking on multivariate outputs. It was however found to be less efficient at discarding poor model fits than the above approach, in the present case study. Note also that Peretti et al. (2013c) mentioned that they performed some model-fit diagnostics when applying parametric approaches, but without further explanation. The fact that they were not able to detect model misfit is likely due to the use of too liberal model checking tests. Such too liberal tests are likely to stem from a direct interpretation of posterior predictive p-values or to the use of a set of summary statistics that is not sufficient to properly assess model fit quality. This should serve as a reminder that the power of model diagnostic procedures should be tested with simulated datasets.

Parametric Forecast error

Following Hartig and Dormann (2013), point estimates of the various parameters (the medians of the parameters sampled during the MCMC) were used to compute predicted future population sizes for up to the next 50 time steps. Concretely, for each parameter point estimates, 50 time series were computed starting at time step t (with t in [50;99]), and lasting during 100 – t steps. These time series were using as initial X(t) values, a maximum-likelihood estimator based on the last five observed population sizes N(t-4)…N(t), and using the model parameter point estimates. The mean square error between the simulated X(t) values and the reference X(t) values were then computed, and divided by the standard deviation of the reference time series to obtain standardized root mean square errors (SRMSE, see Hartig and Dormann 2013 for details, and the Appendix for the R scripts used). Constant forecasts equal to the population temporal mean would lead to a SRMSE equal to one. Therefore, forecasts with a SRMSE larger than one can be considered to have failed. Note that, contrary to what Perretti et al. (2013b,c) and Hartig and Dormann (2013) did, predictive time series were not launched from the true population sizes X(t), but from a maximum-likelihood estimator of the current population size. This choice ensures a more realistic setting where only past observed population sizes N(t) are known. Also note that it would have been possible to use as starting values for the validation time series the sampled MCMC values for the true population sizes X(t), but it would have implied to repeat the

parametric fitting for each starting value of the validation time series, which would have been here prohibitively costly computationally.

Overall forecasting efficiency was summarized in the statistics $\text{Area}_{\text{SRMSE}}$ defined as the overall improvement of the forecast compared to a forecast equal to the average population size (having a SRMSE equal to 1). $\text{Area}_{\text{SRMSE}}$ was computed as: $Area_{SRMSE} = \sum_{i=1}^{8} \max(1 - SRMSE(i), 0)$ where SRMSE(i) is the SRMSE of a forecast i time steps in the future. Negative values for 1-SRMSE(i) were not considered in this formula to focus on the short-term forecast efficiency of the parametric and nonparametric methods, since both will fail for long-term predictions in chaotic systems (Perretti et al. 2013b).

Nonparametric forecasting

Nonparametric forecasting was also implemented on these 900 time series using the S-Map model as in Perretti et al. (2013a,b). This consists in first, transforming the univariate time series N(t) into a time series of time-delayed vectors (N(t),N(t-1),...,N(t-E+1)), with E the embedding dimension. If n temporal points are available in the time-series, this produces n-E+1 vectors of dimension E. Each vector (N(t),N(t-1),...,N(t-E+1)) is associated to the following trajectory ((N(t+1),...,N(t+k)), observed in the k following time steps. The nonparametric forecast of the k next step consists in a weighted average of these following trajectories over the n-E+1-k vectors that have k following time steps available in the time series. The weight of a vector is given by $w(d) = \exp(-\theta d / \bar{d})$ where θ is a parameter of the S-Map model, d is the Euclidean distance between the current data and the focal vector whose weight is computed, and where $\bar{d}$ is the average of d over the various vectors. Embedding dimensions E equal to 2, 3 and 4, and θ values between 0.5 and 25 (with an increment of 0.5) were tested. The combination (E,θ) leading to the smallest mean square error on a forecasting interval of length 1 was determined by cross-validation as in Perretti et al. (2013b). Note again that, contrary to what Perretti et al. (2013a) did, observed population sizes N(t) in the past were used to ground the forecasts, instead of true

population sizes X(t), since this is the only piece of information available in real settings. R scripts are provided in the Appendix.

We repeated all the previous analyses including lognormal process errors in the model (with σ_{proc} equal to 0.005, and using 10 instead of 50 repetitions per parameter set). Results were qualitatively similar, although MCMC convergence was slower. These results are therefore not reported further in the following.

Results

The model checking procedure allowed to detect, in the majority of cases, the failure of the logistic model (θ=1) to satisfactorily fit the data in cases of low (σ_{obs}=0.05) and middle (σ_{obs}=0.1) noise level (Figure 1A,B). Indeed, 94% of the p_{min} values obtained on data generated with the theta-logistic (θ =4) were lower than the 5 percentile of p_{min} values obtained on data generated with the logistic model (θ =1), in the low noise case (Figure 1A). This number was still equal to 64% in the middle noise case (Figure 1B). In contrast, this model checking procedure failed in the large noise case (Figure 1C).

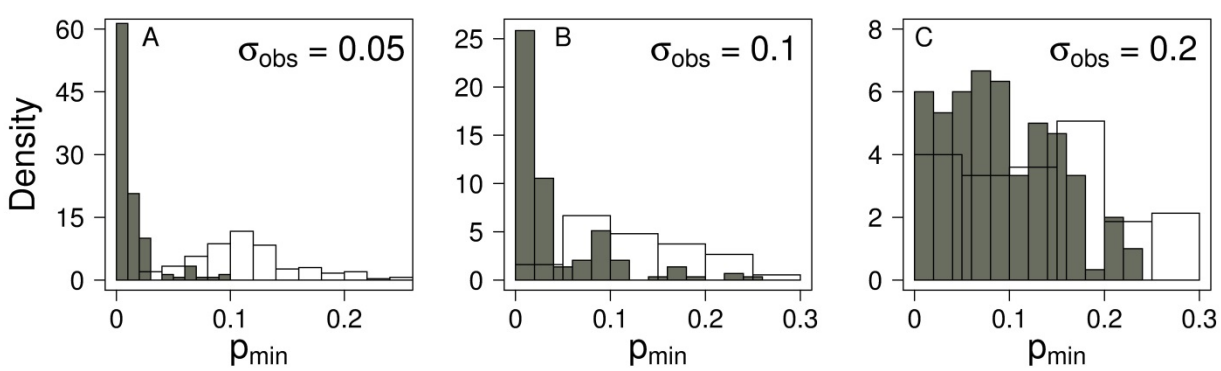

Figure 1. Histograms of model checking p_{min} values when the parametric fit is based on the "true" model in white (data and model fitted based on θ=1), or on a "wrong" model in grey (data generated with θ=4 while the model fitted uses θ=1). Panels A, B and C correspond respectively to cases of low, middle and large observation noise σ_{obs}.

The ability of the model checking procedure to detect model misfit was also influenced by the discrepancy between the model and the simulated data, as measured by the θ value used to simulate the data (Figure 2). In the low and middle noise cases (Figure 2A,B), model misfit could be detected for values of θ larger than or equal to 1.5, since p_{min} values were sharply smaller than those obtained with the reference model (with θ equal to 1). Model misfit could not be detected in the large noise case, irrespective of the value of θ (Figure 2C).

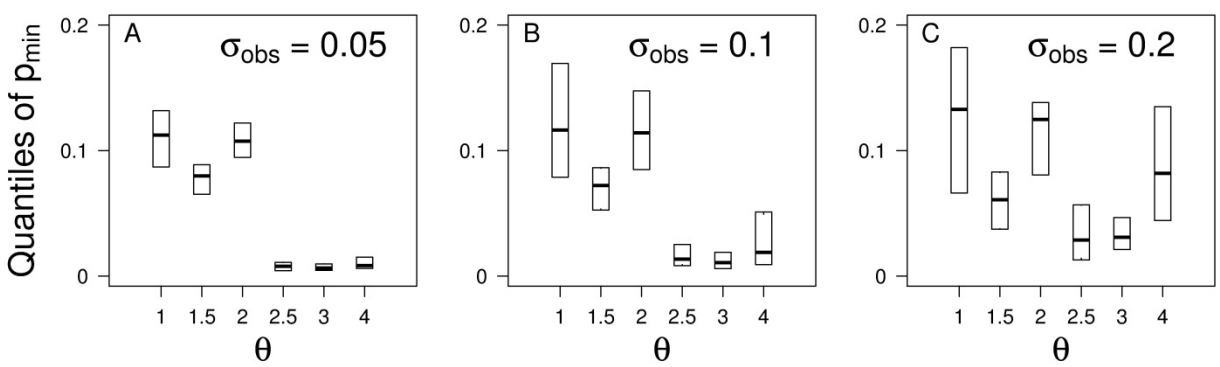

Figure 2. Quantile values of the model checking statistics p_{min} as a function of the discrepancy between the model fitted (θ=1) and data (simulated with θ). The boxes and middle lines correspond to the 25, 50 and 75 percentiles of the p_{min} distributions. Panels A, B and C correspond respectively to cases of low, middle and large observation noise σ_{obs}.

As the discrepancy between the model and the simulated data (measured by the θ value used to simulate the data) rose, forecast errors progressively increased for parametric forecasts, especially in the low noise scenario (Figure 3A, B and C). In contrast, nonparametric forecasts were only marginally affected by the value of θ used in the simulated datasets and were found to outperform parametric forecasts for large value of θ (Figure 3C and F), while the contrary was observed for smaller θ values (Figure 3A, D and E). SRMSE reached larger than one values for both approaches, indicating that they both performed worse than a simple average forecast in the long term. This result is not surprising given the notorious sensitivity to initial conditions inherent to chaotic systems, which prevents any reliable long-term forecast. No reliable predictions could be achieved for more than 5 time steps in the future in all cases. Parametric and nonparametric approaches were comparable, in all but one case (Figure 3C). For this low-noise case with strong model misfit, the nonparametric approach strongly outcompeted the parametric one. Notably, and contrary to what Perretti et al. (2013b,c) argued, in large noise cases, forecast efficiencies were poor for both parametric and nonparametric approaches (Figure 3D, E and F).

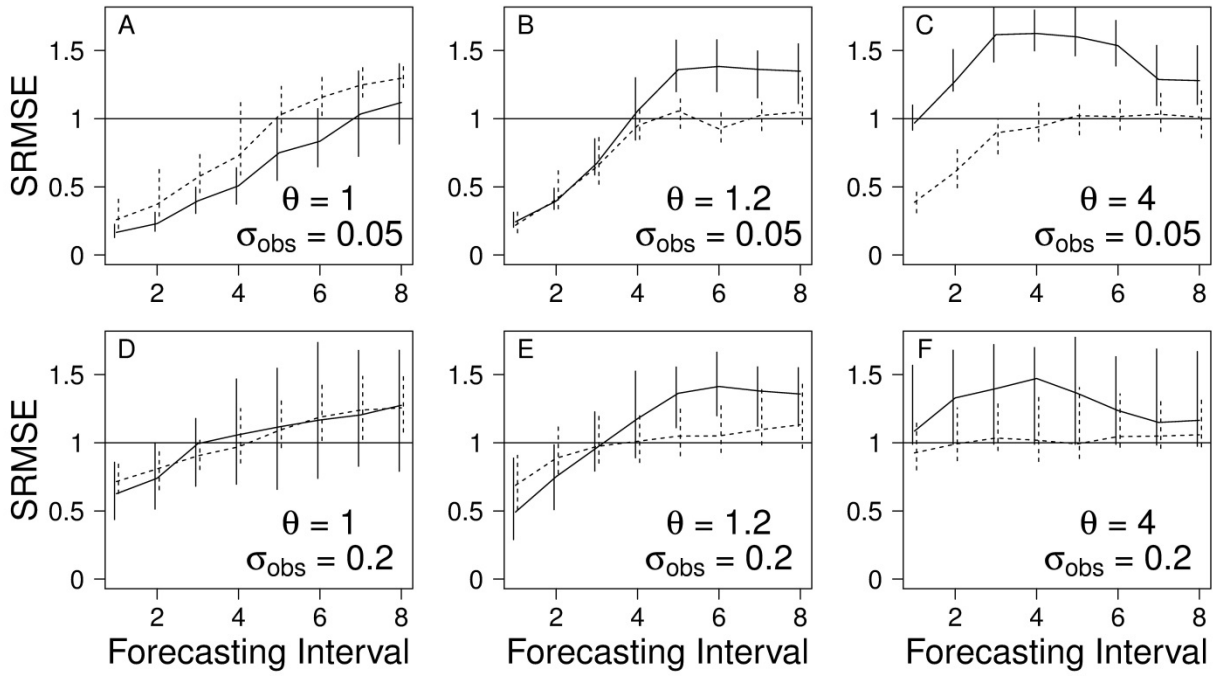

Figure 3. Forecast errors measured by the standardized root mean squared error (SRMSE) as a function of the number of time steps between current data and the predictions, using either the parametric (solid lines) or the nonparametric approach (dotted lines). Vertical bars stand for the 95% confidence interval of the SRMSE. Panels A, B, C (D, E, F) correspond to the low (large) noise cases. Panels A and D (respectively B and E, C and F) correspond to data generated with θ equal to 1, 1.2 and 4.

Finally, overall forecast efficiencies of the parametric approach were found to decrease with the level of observation noise and to be linked with model checking statistics (Figure 4). In both the low and middle noise cases, when model diagnostics would suggest that the model fit quality is acceptable (large p_{min} in Figure 2A, B obtained for θ values lower than 1.5), reasonably good parametric forecast accuracies were reached that were comparable to nonparametric ones (Figure 4A, B). For larger θ values, parametric forecasts were less efficient than nonparametric ones (Figure 4A, B). In the large noise case, forecasting accuracies were poor for both parametric and

nonparametric approaches (Figure 4C), and model checking procedures failed at detecting model misfit (Figure 2C). Importantly though, in this last case with large observation noise, forecasting quality was poor even when data were simulated with the model fitted (Figure 4C for a θ value equal to 1). This means that ecologists could in this case detect the lack of parametric forecast reliability, using simulated data with the fitted level of observation noise. In contrast, when forecasts performed with the nonparametric approach fail (Figure 4F), ecologists have currently no way to detect this failure.

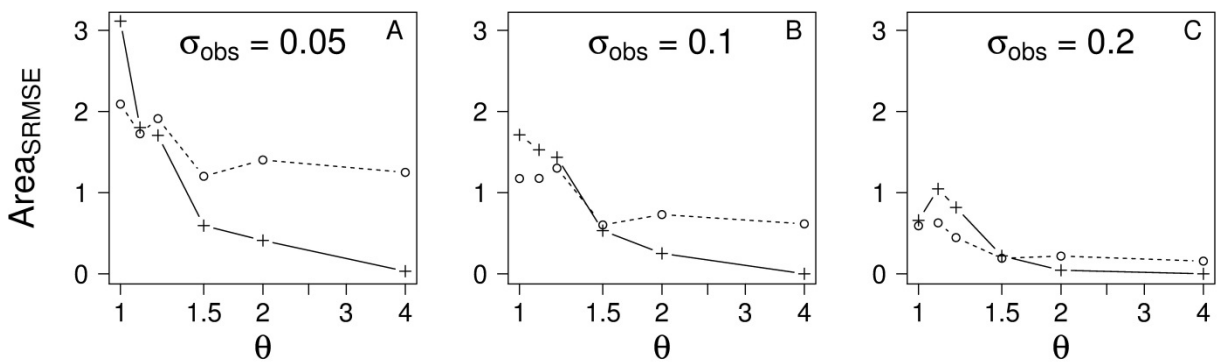

Figure 4. Forecasting efficiency measured by $Area_{SRMSE}$ (see methods) as a function of the discrepancy between the model fitted (θ=1) and data (simulated with θ). Panels A, B and C correspond respectively to cases of low, middle and large observation noise σ_{obs}. Crosses and plain lines stand for parametric forecasts. Circles and dashed lines stand for nonparametric forecasts.

Discussion

Why using parametric forecasting?

This study aimed at highlighting two key advantages of parametric approaches to forecasting. First, Bayesian model checking procedures are powerful tools to asses the misfit of a model

calibrated on real data. These tools enable to identify cases in which our modeled understanding of an ecological system is too crude to make reliable predictions. In such cases, modifications of the model should be sought, before any reliable predictions can be made. By re-examining the toy example of the theta-logistic model proposed by Perretti et al. (2013c), it was shown that a simple model checking procedure was sufficient, provided that observation noise was not too large, to distinguish cases in which model misfit was large and parametric forecasting poor from cases in which model misfit was not evidenced and parametric forecasting was reasonably good (Figure 2 and 4). For real case studies, it is difficult to advise the use of a particular threshold below which a model should be considered to fail at fitting the data. Such procedures enable to detect model misfit, when a posterior predictive p-value is clearly lower than the distribution of posterior predictive p-values obtained on data generated with the parametric model (e.g., Figure 2A and B, for θ values equal or larger than 1.5). In such cases, they will indicate particular patterns in the data that the model may not reproduce accurately, and therefore guide model improvements. Model checking procedures also confirm a good model fit to data, when the posterior predictive p-value is within the central range of the reference distribution. In less clear-cut cases, model checking still provides ecologists with a quantification of the model fit quality.

A second key advantage of parametric approaches is to provide ways to assess the uncertainty of the forecasts in cases in which no model misfit is evidenced. This can be performed using simulated datasets generated with the parametric model, as was done here: in cases in which observation noise was large, forecasting accuracy was found to be poor (irrespective of the approach used, Figure 4C). This last result differs from the ones of Perretti et al. (2013c). The discrepancy between their and the present results is due to the fact that forecasts were here grounded on observed population sizes (including observation noise) rather than true population sizes as they did. This choice of using observed population sizes is justified by the fact that only this piece of information is available in real settings.

The level of observation noise was found to be a key determinant of forecasting accuracy (Figure 4), recalling the common sense that good forecasts require good data with minimized observation error. In real case studies, one cannot know a priori whether the level of observation noise is sufficiently low to ground a reliable forecast. The benefit of parametric approaches is that they enable to assess the data quality with respect to a forecast objective. Indeed, such an assessment can be performed by looking at the forecast accuracy in simulated datasets based on the parametric model with fitted-to-data observation noises (Figure 4 for a θ value equal to 1). Still, users of parametric approaches will gain at collecting independent data on observation error, since this helps increase the reliability and accuracy of parametric inferences (Ives et al. 2003).

Proponents of the nonparametric approach argued that it provides better forecasts than parametric approaches in an array of noised and chaotic case studies (Perretti et al. 2013a,b). They further argued that their approach was easier to apply than computationally demanding parametric approaches. The results of this study do not contradict these two arguments. Nonparametric forecasts were indeed found to outperform parametric ones in a number of – but not all – cases with strongly reduced computational costs (Figure 4). However, no methods are yet available to assess the reliability of nonparametric forecasts which is likely to vary greatly (Figure 4). Not knowing when a forecast is likely to significantly fail is a major problem that should not be disregarded (May et al. 2008).

Circumventing the drawbacks of parametric forecasting

Parametric approaches do have real limitations. In particular, parametric inference can be challenging for complex models. A particular ad-hoc solution to this problem has been used in this study (Pisarenko and Sornette 2004). But many other ways to tackle parametric inference for complex models have been developed in recent years, notably methods based on computer-intensive simulations like approximate Bayesian computation (Beaumont 2010, Jabot and Chave 2011) or synthetic likelihoods (Wood 2010). These methodological innovations are further eased with dedicated statistical packages (e.g., Csilléry et al. 2012, Jabot et al. 2013) and should

drastically reduce limitations to parametric inference. Besides inference, model development has been eased in recent years with progress in techniques of sensitivity analysis (Lamboni et al. 2009) and more generally by the synthesis of good modeling practices (Grimm et al. 2014).

Valuing the advantages of nonparametric methods

The demonstration by Perretti et al. (2013a,b,c) of the great potential of nonparametric approaches suggests that a fruitful avenue for future research might be to combine the respective strengths of both parametric and nonparametric approaches. For instance, one could jointly use parametric and nonparametric approaches, using the first to assess forecast reliability, and the second to obtain more efficient forecasts in cases they are outcompeting parametric ones. Nonparametric embedding methodologies may also be useful for model development, by helping choosing the appropriate time-delays to use in process-based approaches. Another line of research might be to include an observation process in the nonparametric S-map methodology. This may enable to refine the recovery of the chaotic attractor in the embedded time series, and may enable to assess the forecast accuracy of the nonparametric approach. This last perspective should be seen as the key priority for proponents of the nonparametric approach, since the lack of methods for forecast reliability assessment is currently the Achilles' eel of the approach.

Acknowledgements

FJ was funded by the Irstea INDECO project DynIndic. I thank Florian Hartig, the editor and an anonymous reviewer for their insightful comments on a previous version of this manuscript.

References

Austin, M.P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecol. Modell. 157: 101-118.

Beaumont, M. A. 2010. Approximate Bayesian computation in evolution and ecology. Ann. Rev. Ecol. Evol. Sys. 41: 379-406.

Berliner, L.M. 1991. Likelihood and Bayesian prediction of chaotic systems. J. Am. Stat. Ass. 86: 938-952.

Box, G.E.P. 1976. Science and statistics. J. Am. Stat. Ass. 71: 791-799.

Crespi, C. M. and Boscardin, W.J. 2009. Bayesian model checking for multivariate outcome data. Comp. Stat. Dat. Anal. 53: 3765-3772.

Csilléry, K. et al. 2012. abc: an R package for approximate Bayesian computation (ABC). Meth. Ecol. Evol. 3: 475-479.

De Valpine, P. and Hastings, A. 2002. Fitting population models incorporating process noise and observation error. Ecol. Monog. 72: 57-76.

Evans, M.R. et al. 2013. Do simple models lead to generality in ecology? Trends Ecol. Evol. 28: 578-583.

Gelman, A. 2013. Two simple examples for understanding posterior p-values whose distributions are far from uniform. Elec. J. Stat. 7: 2595-2602.

Gelman, A. et al. 1996. Posterior predictive assessment of model fitness via realized discrepancies. Stat. Sin. 6: 733-807.

Grimm, V. and Railsback, S.F. 2005. Individual-based modeling and ecology. Princeton Univ. Press, Princeton, New Jersey, USA.

Grimm, V. et al. 2014. Towards better modelling and decision support: documenting model development, testing, and analysis using TRACE. Ecol. Modell. 280: 129-139.

Hartig, F. and Dormann, C.F. 2013. Does model-free forecasting really outperform the true model? Proc. Natl. Acad. Sci. USA 110: E3975-E3975.

Hartig, F. et al. 2011. Statistical inference for stochastic simulation models – theory and application. Ecol. Lett. 14: 816-827.

Ives, A.R. et al. 2003. Estimating community stability and ecological interactions from time-series data. Ecol. Monog. 73:301-330.

Jabot, F. and Chave, J. 2011. Analyzing tropical forest tree species abundance distributions using a nonneutral model and through approximate Bayesian inference. Am. Nat. 178: E37-E47.

Jabot, F. et al. 2013. EasyABC: performing efficient approximate Bayesian computation sampling schemes using R. Meth. Ecol. Evol. 4: 684-687.

Kantz, H. and Schreiber, T. 2004. Nonlinear time series analysis. Cambridge Univ. Press, Cambridge, UK.

Knape, J. and de Valpine, P. 2012. Are patterns of density dependence in the global population dynamics database driven by uncertainty about population abundance? Ecol. Lett. 15: 17-23.

Lamboni, M. et al. 2009. Multivariate global sensitivity analysis for dynamic crop models. Field Crops Res. 113: 312-320.

May, R.M. 1974. Biological populations with nonoverlapping generations: stable points, stable cycles, and chaos. Science 186: 645-647.

May, R. M. et al. 2008. Complex systems: ecology for bankers. Nature 451: 893-895.

Perretti, C.T. et al. 2013a. Nonparametric forecasting outperforms parametric methods for a simulated multispecies system. Ecology 94: 794-800.

Perretti, C.T. et al. 2013b. Model-free forecasting outperforms the correct mechanistic model for simulated and experimental data. Proc. Natl. Acad. Sci. USA 110: 5253-5257.

Perretti, C.T. et al. 2013c. Reply to Hartig and Dormann: the true model myth. Proc. Natl. Acad. Sci. USA 110: E3976-E3977.

Pisarenko, V.F. and Sornette, D. 2004. Statistical methods of parameter estimation for deterministically chaotic time series. Phys. Rev. E 69: 036122.

Plummer, M. 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. Proc. 3rd Inter. Workshop Distr. Stat. Comp.

Robins, J. M. et al. 2000. Asymptotic distribution of p values in composite null models. J. Am. Stat. Ass. 95: 1143-1156.

Rubin, D.B. 1984. Bayesianly justifiable and relevant frequency calculations for the applied statistician. Ann. Stat. 12: 1151-1172.

Wood, S.N. 2010. Statistical inference for noisy nonlinear ecological dynamic systems. Nature 466: 1102-1104.