

# Discussion of “Feature Matching in Time Series Modeling” by Y. Xia and H. Tong

Edward L. Ionides

Xia and Tong have made a novel contribution to the debate on whether and how to carry out some sort of feature matching in preference to a statistically efficient alternative such as the maximum likelihood estimate (MLE). They show that an estimation criterion emphasizing long-term predictions has some advantages over the MLE on some misspecified time series models. However, emphasizing long-term predictions must lead to a down-weighting of higher-frequency information in the data. In particular, Xia and Tong’s catch-all approach does not typically share the statistical efficiency of MLE when the model fits the data adequately. Further, it is necessarily the case (whatever fitting method is used) that some scientific inferences one might wish to conclude from fitting a misspecified model are statistically invalid. Scientific interpretation of fitted parameter values and predictions using a model that is a statistically poor match to the data therefore requires considerable care. One seeks models that are simultaneously scientifically relevant and provide an adequate statistical description of the data, and then statistical efficiency becomes an important consideration for drawing scientific conclusions from limited data. Flexible modern inference methods facilitate the development and statistical analysis of such models. I will discuss these issues in the context of Xia and Tong’s analysis of Nicholson’s blowfly data. Similar considerations arise in their measles example, and have been investigated by He, Ionides and King (2010).

Xia and Tong’s  $APE(\leq 1)$  estimate is equivalent to the MLE only for a specific choice of stochastic model. From their equation (3.12), we see that  $APE(\leq 1)$  corresponds to the MLE for additive, Gaussian, constant-variance process noise with no measurement error. For Xia and Tong’s blowfly model, the log-likelihood at the  $APE(\leq 1)$  point estimate is  $-1568.5$  whereas the log-likelihood at the  $APE(\leq T)$  point estimate is  $-1569.5$ . A chi-squared approximation indicates that

a full likelihood-based analysis for this model should consider the  $APE(\leq 1)$  and  $APE(\leq T)$  point estimates to be both statistically plausible, since the difference of 1.0 log units is not large compared to typical values of  $1/2$  of a chi-squared random variable with five degrees of freedom. To check the extent to which either of these point estimates provides a reasonable statistical explanation of the data, I compared their goodness of fit with that of a simple phenomenological model. For oscillating populations, a log-ARMA model is an appropriate choice (He, Ionides and King, 2010). I fitted a stationary log-ARMA model to the 9th through 200th data points for which predictions are made by Xia and Tong’s model, in order to ensure that the resulting likelihood provides a fair comparison. A log-ARMA(2, 2) model gives a maximized log-likelihood of  $-1542.3$  based on estimating six parameters. Xia and Tong’s mechanistic model therefore explains the data considerably more poorly (e.g., judged by Akaike’s information criterion) than this simple black-box model. Is it possible to preserve the scientific interpretability of Xia and Tong’s model while also providing a statistically satisfactory explanation of the data? To address this question, I fitted a dynamic model adapted from Wood (2010) which has a similar structure to the model of Xia and Tong but differs by formulating the stochasticity in a scientifically motivated way. This alternative model is described in full in the Appendix below. I evaluated the likelihood by sequential Monte Carlo and computed the MLE by iterated filtering (Ionides, Bretó and King, 2006) implemented using the `pomp` package for R (King et al., 2010). Maximization over the six parameters led to a log-likelihood of  $-1465.4$ . Figure 1 shows that the skeleton of this alternative model matches the periodicity in the data, a measure of fit which Xia and Tong chose to emphasize in their Figure 8. The likelihood at the MLE also comfortably outperforms the log-ARMA(2, 2) benchmark so subsequent analysis can consider the model to be adequately specified, at least to a first approximation. Of course, the possibility of potential further advances in the model specification cannot be ruled out. Indeed, a careful and complete investigation would be expected

---

Edward L. Ionides is Associate Professor, Department of Statistics, University of Michigan, Ann Arbor, Michigan 48109, USA (e-mail: [ionides@umich.edu](mailto:ionides@umich.edu)).