Category and Session Number:
NL2 Nonlinear Analysis, Assimilation and Predictability of Geophysical Data

Preferred Mode of Presentation: Oral

Double embedding for local prediction, nonlinear modelling, and simulation

MICHAEL SMALL

Electronic and Information Engineering
Hong Kong Polytechnic University

Many time series, including geophysical data, exhibit dynamics on wildely disparate time scales. We describe an iterated modelling procedure that captures these various dynamics and illustrate the application of this method to daily temperature data with local models that capture both the daily and seasonal variation in the data.

Let $\{x_t\}$ denote a scalar time series and $x_t^{(v)}$ denote the vector observation obtain via the time delay embedding $v = (v_1, v_2, \ldots, v_m)$ such that $x_t^{(v)} := (x_{t-v_1}, \ldots, x_{t-v_m})$. Suppose that the embedding v describes the long term dynamics (possibly, v may include lags to capture seasonal fluctuations in a time series of daily temperature readings), and suppose that $u = (u_1, u_2, \ldots, u_n)$ captures the short term dynamics (we suppose that u includes lags concerned with the immediate past). Then we build two predictive models, $f(x_t^{(v)}) = x_{t+1}+e_{t+1}$, and $g(f(x_t^{(v)}), x_t^{(u)}) = x_{t+1}+\epsilon_{t+1}$ where both $\sum e_{t+1}^2$ and $\sum \epsilon_{t+1}^2$ are somehow minimal. For simplicity, we choose to construct models $f(\cdot)$ and $g(\cdot)$ using the method of analogues. Namely $f(x_t^{(v)}) = \sum_{(s_1,s_2,\ldots,s_d)} x_{s_i+1}$ where s_i for $(i = 1, 2, \ldots, d)$ are such that $x_{s_i}^{(v)}$ are the n nearest neighbours ($\|x_{s_i}^{(v)} - x_t^{(v)}\|$ is smallest) of $x_t^{(v)}$ and $s_i < t$ for all i. The second model $g(\cdot)$ is constructed similarly. This is exactly equivalent to taking an average of the successors of those points in the past history that are most like the current state. The critical parameter here is the embedding strategy v (or u), as this is the metric which we apply to decide which past observations most closely resemble the current state.

From various computational simulations we observe that $\sum {e_{t+1}}^2 > \sum \epsilon_{t+1}^2$ and therefore, we find that this *double embedding* strategy improves short term predictability. Furthermore, and less trivially, we see that in general, the simulations (i.e. iterated prediction) of this composite model exhibit dynamics more similar to those of the underlying deterministic system, or the observed true data.