Data-driven forecasting without a model and with a partially known model

Tyrus Berry

The Pennsylvania State University

June 11, 2015

Collaborators: John Harlim (PSU) and Dimitrios Giannakis (Courant)
Support: Office of Naval Research

Goal of nonparametric forecasting

(Loading Video...)

Forecasting densities and observables

- Consider an uncertain initial state $x(0)$ with dynamics:

$$dx = a(x)\,dt + b(x)\,dW_t$$

- Probability $p(x,t)$ solves *Fokker-Planck* PDE, $p_t = \mathcal{L}^* p$

$$\mathcal{L}^* f = -\nabla \cdot (a(x)f(x)) + \sum_{i,j,k} \frac{\partial^2}{\partial x_i \partial x_j}\left(\frac{b_{ik}(x)b_{jk}(x)}{2} f(x)\right)$$

- Solution is a semi-group: $p(x,t) = e^{t\mathcal{L}^*} p(x,0)$
- Adjoint semi-group evolves functionals $\mathbb{E}[f(x(t))] = e^{t\mathcal{L}} f(x(0))$
- Solving is hard, data gives a shortcut!

The Shift Map

- Given data samples $x_i = x(t_i)$ with $\tau = t_{i+1} - t_i$
- Define the *shift map* of a function by $Sf(x_i) = f(x_{i+1})$
- Unbiased estimate $\mathbb{E}[Sf(x_i)] = \mathbb{E}[f(x_{i+1})] = e^{\tau\mathcal{L}} f(x_i)$
- Using the Itô lemma we can show:

$$Sf(x_i) = e^{\tau\mathcal{L}} f(x_i) + \int_{t_i}^{t_{i+1}} \nabla f^\top b \, dW_s + \int_{t_i}^{t_{i+1}} Bf \, ds$$

- Need to minimize the stochastic integrand $\nabla f^\top b$

Forecasting with the Shift Map

- Project onto a basis $\{\varphi_j(x)\}$ of smooth functions:

$$\begin{array}{ccc} p(x,t) & \xrightarrow{\text{Nonparametric Forecast}} & p(x,t+\tau) \\ \Big\downarrow \langle p,\varphi_j \rangle & & \Big\uparrow \sum_j c_j \varphi_j p_{\mathrm{eq}} \\ \vec{c}(t) & \xrightarrow{A_{lj} \equiv \mathbb{E}[\langle \varphi_j, S\varphi_l \rangle_{p_{\mathrm{eq}}}]} & \vec{c}(t+\tau) = A\vec{c}(t). \end{array}$$

- Easy numerical approximations for $c_l(t)$ and A_{lj} (Monte-Carlo)
- Estimate of A_{lj} has error $\mathcal{O}(||\nabla \varphi_l||_{p_{\mathrm{eq}}} \sqrt{\tau/N})$

Choosing a basis

- The minimizers of $||\nabla\varphi_l||_{p_{\mathrm{eq}}}$ are a generalized Fourier basis
- Let $\Delta_{p_{\mathrm{eq}}} = \Delta + \frac{\nabla p_{\mathrm{eq}}}{p_{\mathrm{eq}}} \cdot \nabla$ be the Laplacian weighted by p_{eq}
- The eigenfunctions $\Delta_{p_{\mathrm{eq}}}\varphi_j = \lambda_j\varphi_j$ minimize $||\nabla\varphi_j||_{p_{\mathrm{eq}}} = \lambda_j$
- These are the eigenfunctions produced by Diffusion Maps ($\alpha = 1/2$).
- Use a graph Laplacian $L_{p_{\mathrm{eq}}}$ approximates the operator $\Delta_{p_{\mathrm{eq}}}$
- Eigenvectors of $L_{p_{\mathrm{eq}}}$ approximate φ_j minimizing $||\nabla\varphi_j||_{p_{\mathrm{eq}}}$

The Alogrithm

- Given data $\{x_i\}$ use Diffusion Maps to get basis $\{\varphi_j(x_i)\}$
- Use Shift Map to get the matrix $A \approx \frac{1}{N}\sum_{i=1}^{N} \varphi_j(x_i)\varphi_l(x_{i+1})$
- Now forecast any density by:

$$\begin{array}{ccc} p(x,t) & \xrightarrow{\text{Nonparametric Forecast}} & p(x,t+\tau) \\ \Big\downarrow \langle p,\varphi_j\rangle & & \Big\uparrow \sum_j c_j\varphi_j p_{\mathrm{eq}} \\ \vec{c}(t) & \xrightarrow{A_{lj}\equiv\mathbb{E}[\langle\varphi_j,S\varphi_l\rangle_{p_{\mathrm{eq}}}]} & \vec{c}(t+\tau) = A\vec{c}(t). \end{array}$$

Nonparametric forecast on a torus

- Stochastic dynamics on a torus $(\theta, \phi) \in [0, 2\pi)^2$

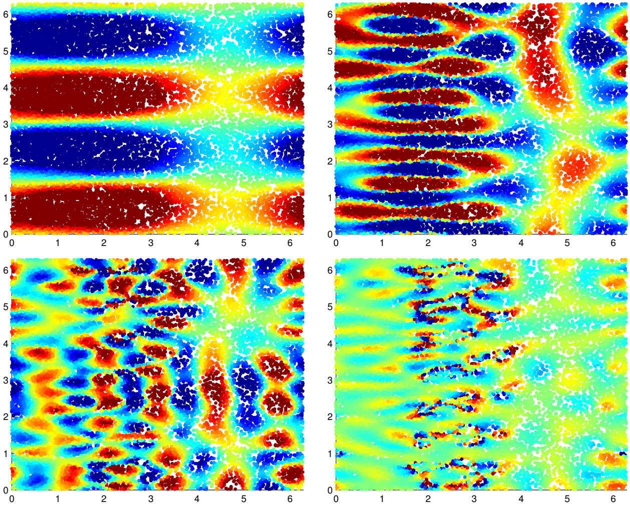
$$d(\theta, \phi)^\top = a(\theta, \phi)\, dt + b(\theta, \phi)\, dW_t$$

- Drift and diffusion coefficients,

$$a(\theta, \phi) = \begin{pmatrix} \frac{1}{2} + \frac{1}{8}\cos(\theta)\cos(2\phi) + \frac{1}{2}\cos(\theta + \pi/2) \\ 10 + \frac{1}{2}\cos(\theta + \phi/2) + \cos(\theta + \pi/2) \end{pmatrix},$$

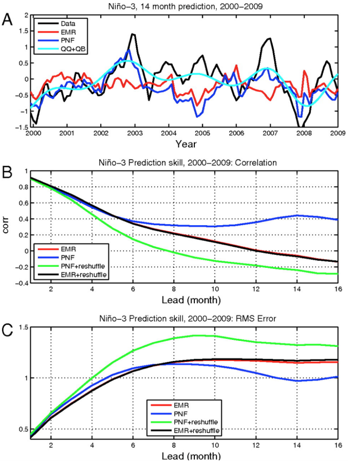
$$b(\theta, \phi) = \begin{pmatrix} \frac{1}{4} + \frac{1}{4}\sin(\theta) & \frac{1}{4}\cos(\theta + \phi) \\ \frac{1}{4}\cos(\theta + \phi) & \frac{1}{40} + \frac{1}{40}\sin(\phi)\cos(\theta) \end{pmatrix}.$$

Eigenfunctions on the torus

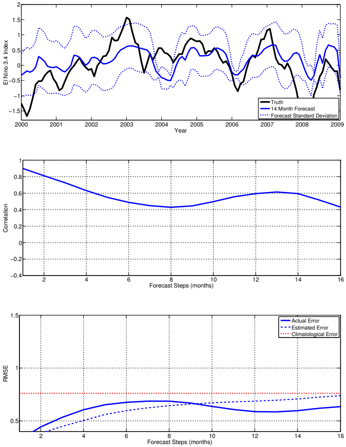


(Loading Video...)

Forecasting the El Nino Index



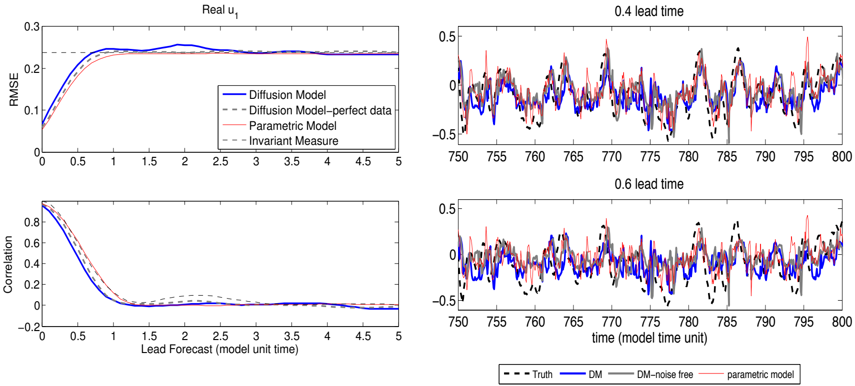
Chekrouna, Kondrashov, and Ghil, PNAS 2011,108,no.29



Diffusion Forecast

Forecasting a Fourier mode of TBH

Diffusion Forecast compared to parametric model of A.Majda,J.Harlim,
Physics constrained nonlinear regression models for time series, Nonlinearity 26 (2013).



Model Error and the Curse of Dimensionality

- Nonparametric model *interpolates* from the training data
- Data required grows exponentially in the dimension of the manifold
- Assume we have a partially known model $\dot{x} = f(x, \theta)$
- Dynamics $d\theta = a(\theta)\,dt + b(\theta)\,dW_t$ are unknown
- Idea: build a nonparametric model for $p(\theta, t)$
- Challenge: make the nonparametric model play nicely with $f(x, \theta)$

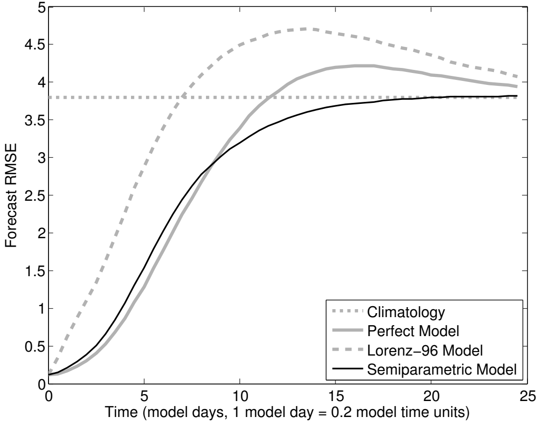
Semiparametric Model

- Assume we have a partially known model $\dot{x} = f(x, \theta)$
- Dynamics $d\theta = a(\theta)\, dt + b(\theta)\, dW_t$ are unknown
- Build a nonparametric model for $p(\theta, t)$
- Sample $\theta^k(t) \sim p(\theta, t)$ to use with ensemble forecast (x^k, θ^k)

$$
\begin{array}{ccc}
(x^k(t), \theta^k(t)) & \xrightarrow{\dot{x}=f(x,\theta)} & (x^k(t+\tau), \theta^k(t+\tau)) \\
\uparrow \theta^k(t) & & \uparrow \theta^k(t+\tau) \\
p(\theta, t) & \overset{\text{Nonparametric Forecast}}{- - - - - - \longrightarrow} & p(\theta, t+\tau)
\end{array}
$$

Example: 40-dimensional Lorenz-96 system

$$\dot{x}_i = f(x_i, \theta) = \theta x_{i-1}x_{i+1} - x_{i-1}x_{i-2} - x_i + 8$$



Additional challenges of semiparametric modeling

- Need a training data set for θ
- Need initial condition $p(\theta, t)$ for nonparametric forecast
- We developed semiparametric filtering to address these
- Still require that evolution of θ does not depend on x

For more information: http://personal.psu.edu/thb11/

Building the basis

- Coifman and Lafon, *Diffusion maps.*
- Berry and Harlim, *Variable Bandwidth Diffusion Kernels.*
- Berry and Sauer, *Local Kernels and the Geometric Structure of Data.*

Nonparametric forecast

- Berry, Giannakis, and Harlim, *Nonparametric forecasting of low-dimensional dynamical systems.*
- Berry and Harlim, *Forecasting Turbulent Modes with Nonparametric Diffusion Models.*

Semiparametric forecast

- Berry and Harlim, *Semiparametric forecasting and filtering: correcting low-dimensional model error in parametric models.*