Predicting the cloud patterns of the Madden-Julian Oscillation through a low-order nonlinear stochastic model (Auxiliary Material)

N. Chen, A. J. Majda and D. Giannakis

Department of Mathematics and Center for Atmosphere Ocean Science, Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA.

1 Calibration of the Nonlinear Stochastic Models

Recall the low-order nonlinear stochastic model

$$\frac{du_1}{dt} = (-d_u\, u_1 + \gamma\,(v + v_f(t))\, u_1 - (a + \omega_u)\, u_2) + \sigma_u\, \dot{W}_{u_1}, \tag{1.1a}$$

$$\frac{du_2}{dt} = (-d_u\, u_2 + \gamma\,(v + v_f(t))\, u_2 + (a + \omega_u)\, u_1) + \sigma_u\, \dot{W}_{u_2}, \tag{1.1b}$$

$$\frac{dv}{dt} = (-d_v\, v - \gamma\,(u_1^2 + u_2^2)) + \sigma_v\, \dot{W}_v, \tag{1.1c}$$

$$\frac{d\omega_u}{dt} = (-d_\omega \omega_u + \hat{\omega}_u) + \sigma_\omega\, \dot{W}_\omega, \tag{1.1d}$$

where

$$v_f(t) = f_0 + f_t \sin(\omega_f t + \phi). \tag{1.2}$$

The optimal parameters in the stochastic model is determined by systematically minimizing the information distance (model error) of the signal equilibrium PDF of the stochastic model q compared with that of the actual data p [1, 2],

$$\mathcal{P}(p, q) = \int p \ln\left(\frac{p}{q}\right). \tag{1.3}$$

The sensitivity analysis is shown in Figure 1.1, which indicates the robustness of the low-order nonlinear stochastic model with respect to the parameters.

Prediction with random suboptimal parameters are shown in Figure 1.2 and 1.3, where the random suboptimal parameter are given by

$$\begin{aligned} &\sigma_u = \sigma_u^* + \mathcal{U}(-0.05, 0.05), \qquad d_u = d_u^* + \mathcal{U}(-0.35, 0.25), \qquad f_t = f_t^* + \mathcal{U}(0, 1) \\ &\gamma = \gamma^* + \mathcal{U}(0, 0.2), \qquad \sigma_v = \sigma_v^* + \mathcal{U}(-0.2, 0.2), \qquad d_v = d_v^* + \mathcal{U}(-0.35, 0.25), \end{aligned} \tag{1.4}$$

where the variables with stars are the optimal parameters and $\mathcal{U}(a, b)$ is the uniform distribution in the interval $[a, b]$.

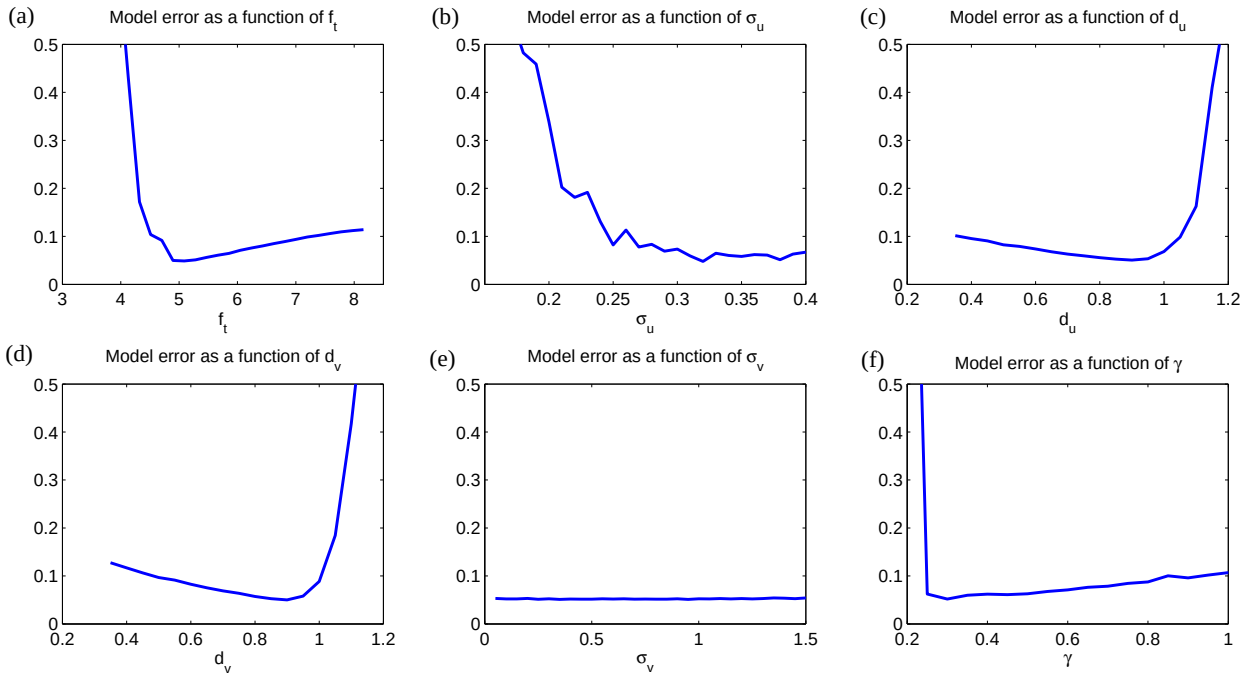

Figure 1.1: Sensitivity analysis of parameters $f_t, \sigma_u, d_u, d_v, \sigma_v$ and γ in the low-order nonlinear stochastic model (1.1).

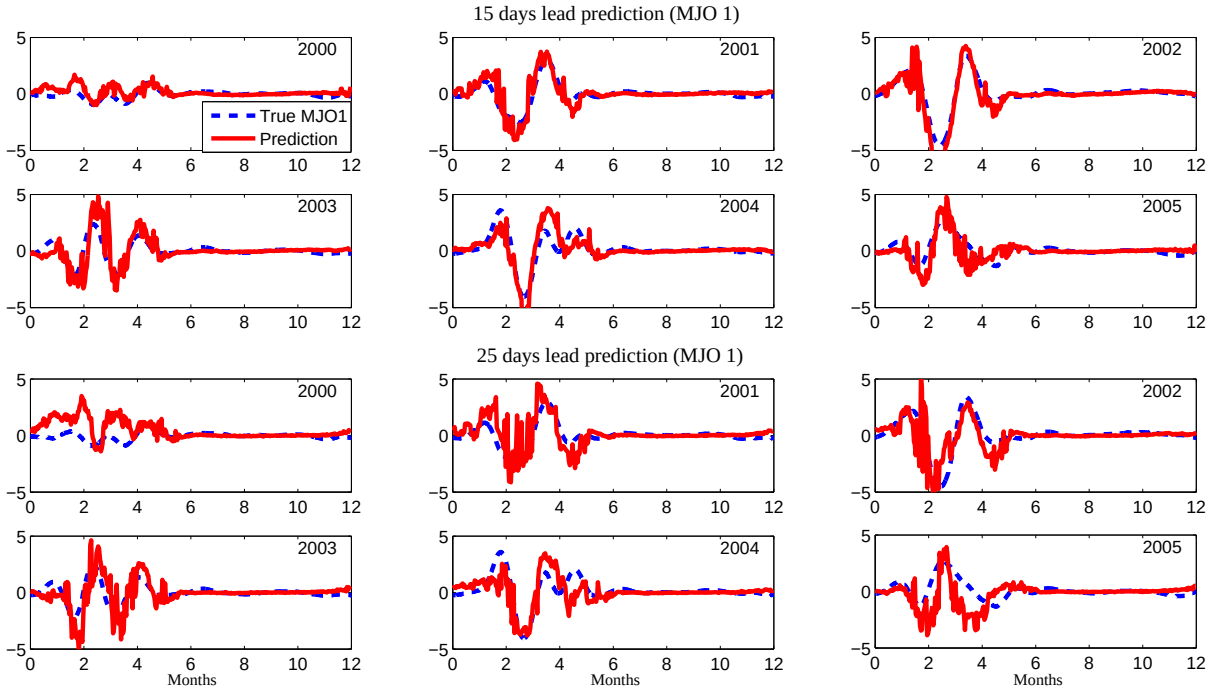

Figure 1.2: 15 and 25 days lead prediction skill with random suboptimal parameters.

The prediction with random suboptimal parameters has comparable prediction skill as that with optimal parameters, implying the robustness of the low-order nonlinear stochastic model in prediction.

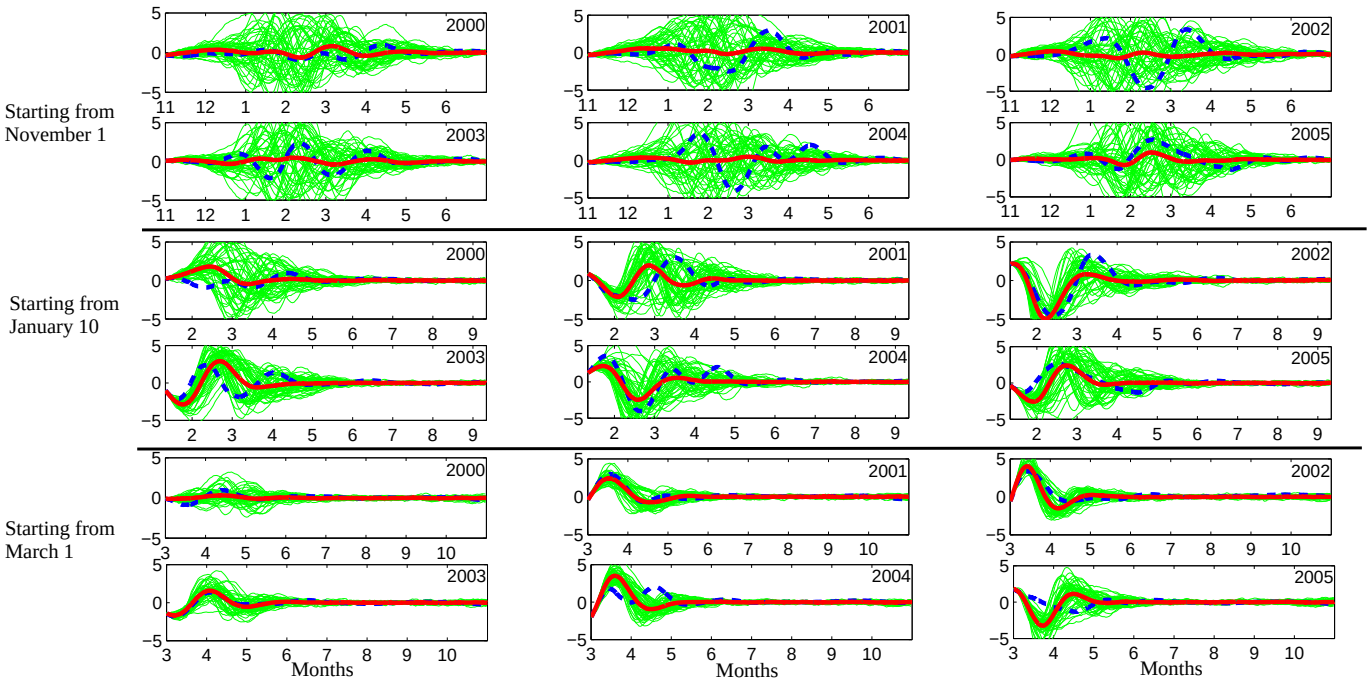

Figure 1.3: Prediction skill starting from November 1, January 10 and March 1 with random suboptimal parameters.

2 Details of data assimilation algorithm

In the prediction stage, the initial values of the hidden variables v and ω_u are determined via data assimilation algorithm.

One feature of the low-order nonlinear stochastic system (1.1) is that it is a conditional Gaussian system with respect to the observations u_1 and u_2, meaning that once u_1 and u_2 are given, the system is a linear equation with Gaussian statistics provided that the initial condition is Gaussian. For such type of system, the conditional distribution of v and ω_u given the observations u_1 and u_2, which is Gaussian, can be explicitly written down, which is contributed by Lipster and Shiryaev [3].

To run this data assimilation algorithm, we rewrite the low-order nonlinear stochastic model (1.1) in the following abstract form,

$$d\mathbf{U}_t = [\mathbf{A}_0(t,\mathbf{U}) + \mathbf{A}_1(t,\mathbf{U})\mathbf{\Gamma}_t]dt + \mathbf{B}_1(t,\mathbf{U})d\mathbf{W}_1(t), \tag{2.1a}$$

$$d\mathbf{\Gamma}_t = [\mathbf{a}_0(t,\mathbf{U}) + \mathbf{a}_1(t,\mathbf{U})\mathbf{\Gamma}_t]dt + \mathbf{b}_2(t,\mathbf{U})d\mathbf{W}_2(t), \tag{2.1b}$$

where $\mathbf{U}_t = (u_1, u_2)^T$ is a vector containing the observed variables while $\mathbf{\Gamma}_t = (v, \omega_u)^T$ represents for the unobserved processes. The matrices and vectors A_0, A_1, a_0, a_1, B_1 and b_2 for the low-order nonlinear stochastic are given as follows

$$\begin{aligned} A_0(t,\mathbf{U}) &= \begin{pmatrix} -d_u u_1 - a u_2 + \gamma v_f(t) \\ -d_u u_2 + a u_1 + \gamma v_f(t) \end{pmatrix}, \qquad A_1(t,\mathbf{U}) = \begin{pmatrix} \gamma u_1 & -u_2 \\ \gamma u_2 & u_1 \end{pmatrix}, \\ a_0(t,\mathbf{U}) &= \begin{pmatrix} -\gamma(u_1^2+u_2^2) \\ \hat{\omega}_u \end{pmatrix}, \qquad a_1(t,\mathbf{U}) = \begin{pmatrix} -d_v & \\ & -d_\omega \end{pmatrix}, \\ B_1(t,\mathbf{U}) &= \begin{pmatrix} \sigma_u & \\ & \sigma_u \end{pmatrix}, \qquad b_2(t,\mathbf{U}) = \begin{pmatrix} \sigma_v & \\ & \sigma_\omega \end{pmatrix}. \end{aligned} \tag{2.2}$$

The Theorem in [3] states that assuming the conditional initial conditions being Gaussian and some moments bound for the corresponding processes, the conditional mean μ_t and conditional covariance R_t of $\mathbf{\Gamma}$ based on the observed process $\mathbf{U}$ are expressed as follows

$$\begin{aligned} d\mu_t =& [\mathbf{a}_0(t,\mathbf{U}) + \mathbf{a}_1(t,\mathbf{U})\mu_t]dt + (R_t\mathbf{A}_1^*(t,\mathbf{U}))(\mathbf{B}_1\mathbf{B}_1^*)^{-1}(t,\mathbf{U})[d\mathbf{U}_t - (\mathbf{A}_0(t,\mathbf{U}) + \mathbf{A}_1(t,\mathbf{U})\mu_t)dt], \\ dR_t =& \left\{\mathbf{a}_1(t,\mathbf{U})R_t + R_t\mathbf{a}_1^*(t,\mathbf{U}) + (\mathbf{b}_2\mathbf{b}_2^*)(t,\mathbf{U}) - (R_t\mathbf{A}_1^*(t,\mathbf{U}))(\mathbf{B}_1\mathbf{B}_1^*)^{-1}(t,\mathbf{U})(R_t\mathbf{A}_1^*(t,\mathbf{U}))^*\right\}dt. \end{aligned} \tag{2.3}$$

The initial value of the hidden processes for this filtering process is taken as the Gaussian fit of the model equilibrium state. Therefore, with (2.3) the conditional Gaussian distribution of the hidden processes v and ω_u at each fixed time is determined.

As a remark, we point out that the filter (2.1) and (2.3) is an optimal filter if and only if the signal is generated from system (2.1). Yet, our observed signal is MJO indices while we utilize the low-order nonlinear stochastic model as the filter, and therefore our filter is merely a suboptimal filter.

In Figure 2.1 we show the posterior mean and variance of the recovery of the hidden variable v and ω_u in the historic period from 1983/09/03 to 1999/12/31. The cross-covariance of v and ω_u is negligible of order $O(10^{-18})$ and is not shown here. Note that when intermittency occurs, the posterior covariance is small, meaning that the recovered state has small uncertainty.

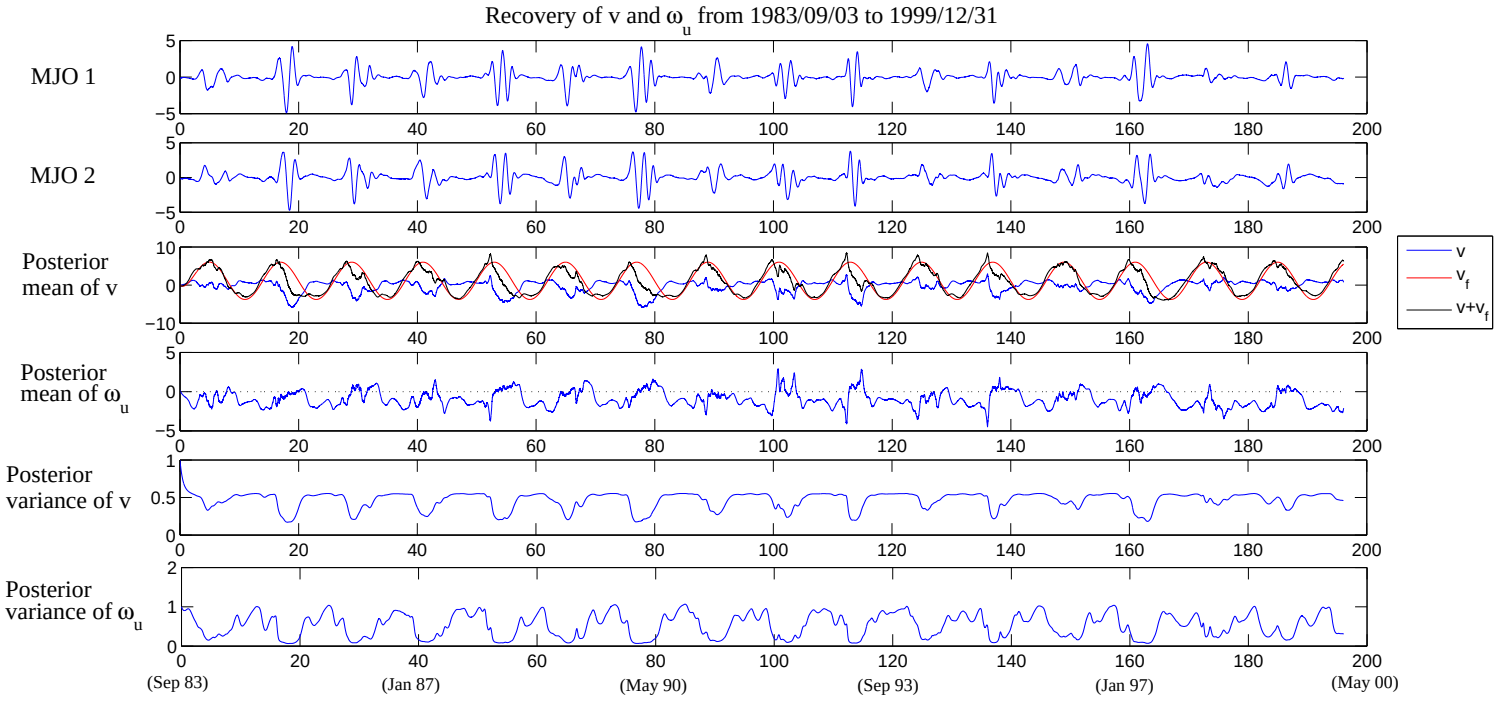

Figure 2.1: Recovery of posterior mean and variance of stochastic damping v and stochastic phase ω_u in (1.1) from 1983/09/03 to 1999/12/31 as a function of time compared with the observations of MJO indices. The cross-covariance of v and ω_u is negligible of order $O(10^{-18})$ and is not shown here.

3 Twin experiment

In this final section, we consider a twin experiment. That is, we assume the true signal is generated from the low-order nonlinear stochastic model (1.1). Thus, the data assimilation algorithm is based on a perfect filter. The signal generated by (1.1) is shown in Figure 3.1. In particular, the signal has the same length as MJO indices and contains weak and strong events in the prediction period. Figure 3.2 and 3.3 show the RMS error and bivariate correlation for 1-60 days lead prediction skill and the 15 and 25 days lead prediction curves. The MJO prediction skill in the paper is comparable to this internal prediction skill, suggesting the low-order nonlinear stochastic model has significant skill for determining the predicability limits of the large scale cloud patterns of the boreal winter MJO.

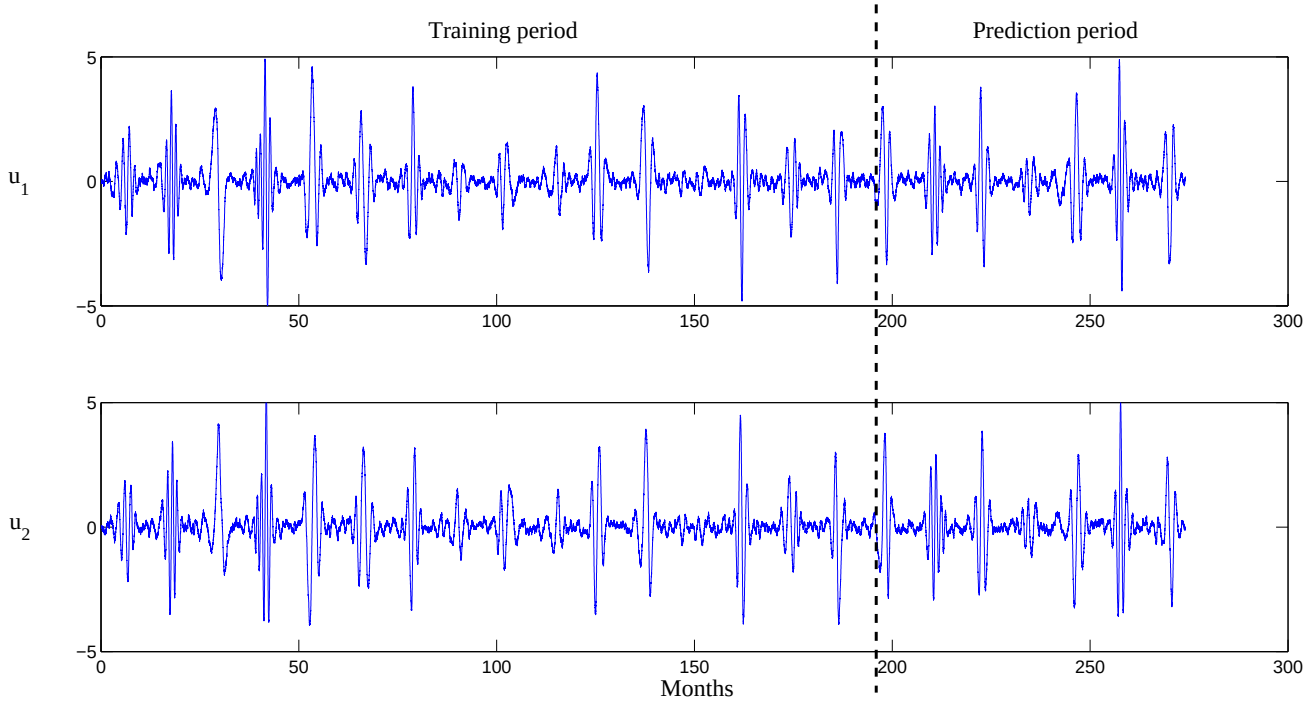

Figure 3.1: Twin experiment. The signal is generated from nonlinear model (1.1) that has the same length as the MJO data with the training period (1983/09/03–1999/12/31) and prediction period (2000/01/01–2006/06/30).

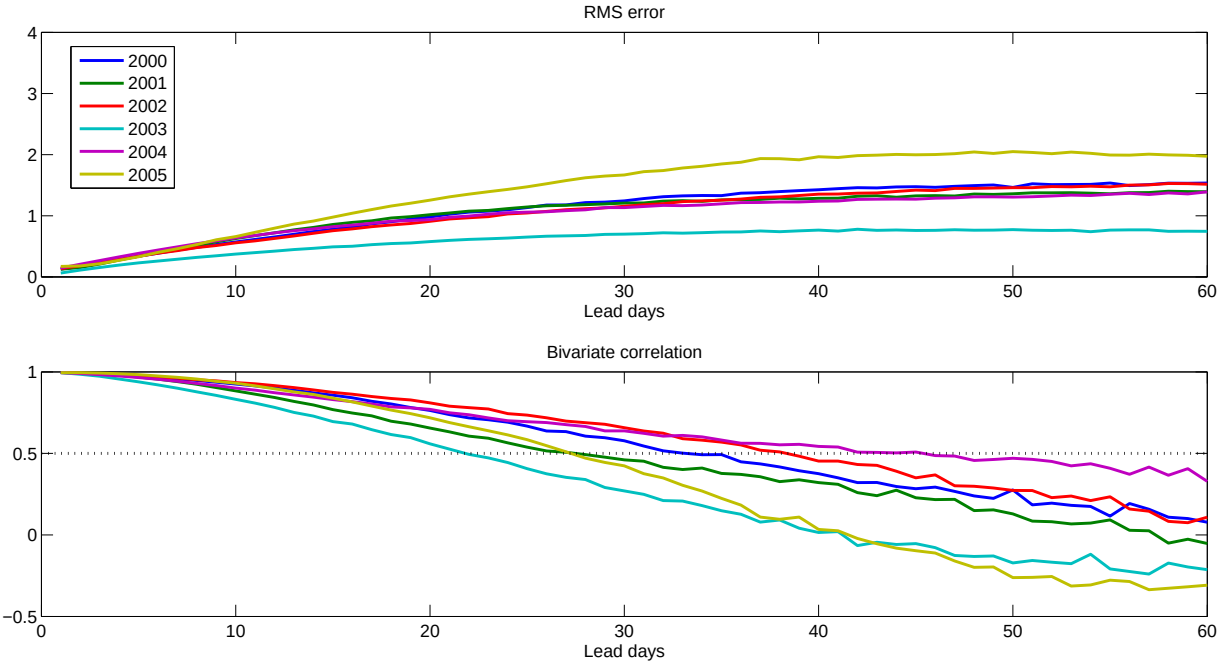

Figure 3.2: Twin experiment. Skill scores with RMSE (top) and bivariate correlation (bottom) for prediction in different years.

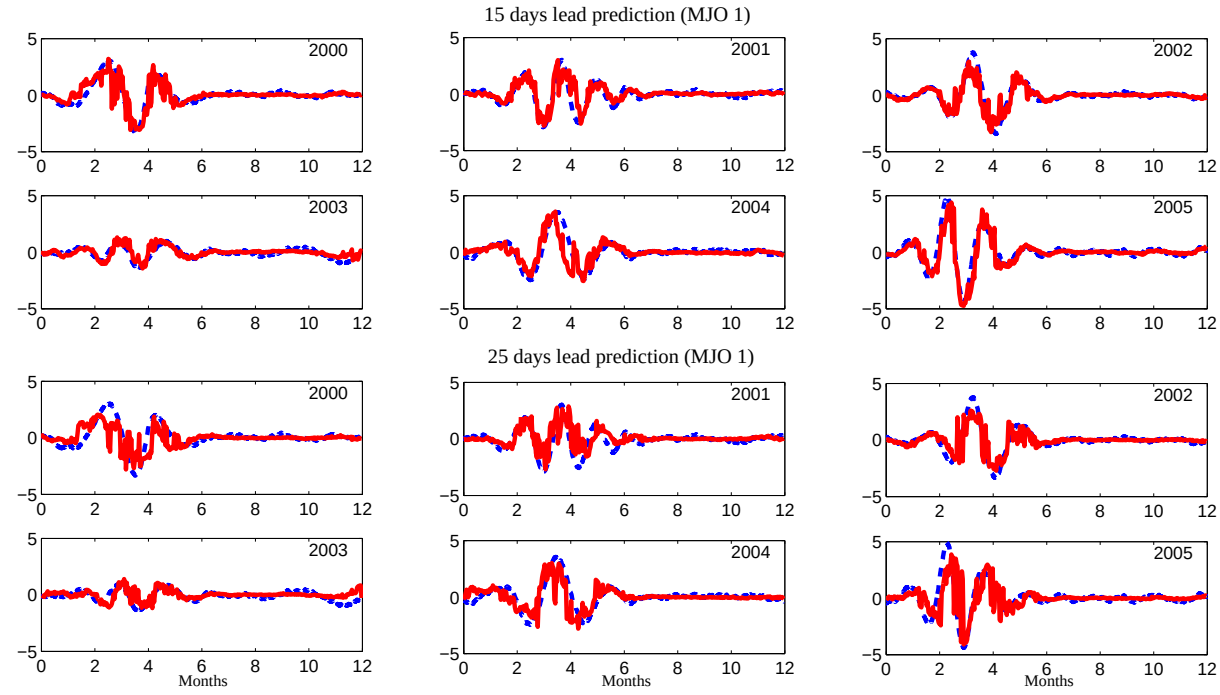

Figure 3.3: Twin experiment. Prediction skill of u_1 at a 15 (top) and 25 (bottom) days lead. The blue line shows the true signal and the red line shows the ensemble average of the predicted signal with 50 ensemble members.

References

[1] Majda, A. J., and B. Gershgorin (2010), Quantifying uncertainty in climate change science through empirical information theory, *Proceedings of the National Academy of Sciences*, *107*(34), 14,958–14,963.

[2] Majda, A. J., and B. Gershgorin (2011), Improving model fidelity and sensitivity for complex systems through empirical information theory, *Proceedings of the National Academy of Sciences*, *108*(25), 10,044–10,049.

[3] Liptser, R. S., and A. N. Shiryaev (2001), *Statistics of Random Processes II: II. Applications*, vol. 2, Springer.