Stochastic Parametrization and Adjoint Methods:

Exploring a Toy Problem

MAX TROSTEL

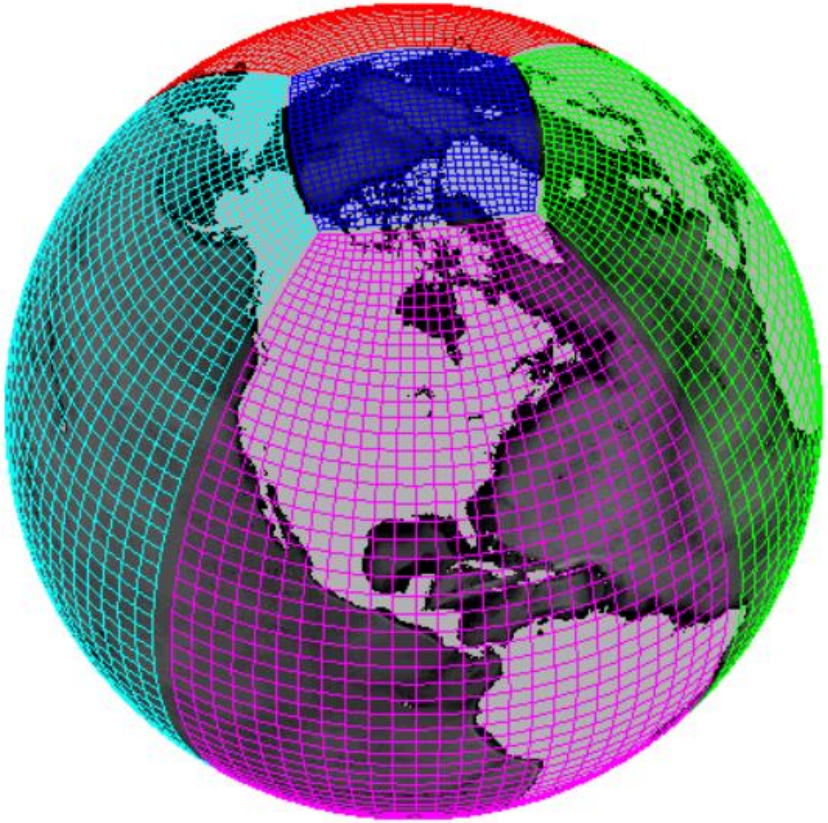
MARCH 22^{ND}, 2024

This talk

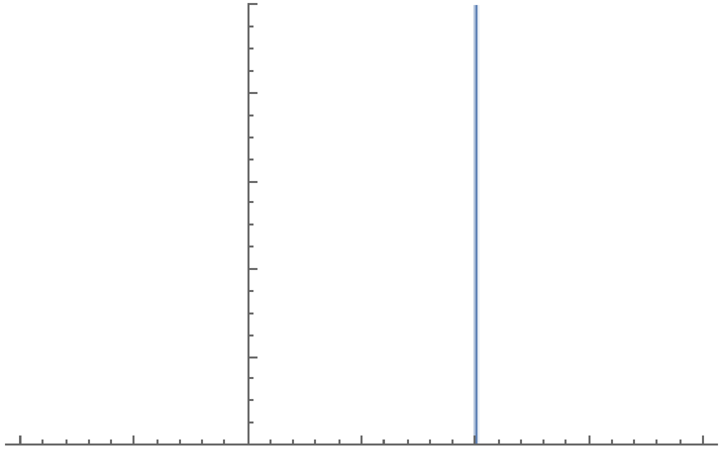
- Why stochastic parametrization?
- A toy problem: the Lorenz 96 2-scale model
- A look ahead

Subgrid-scale parametrization

- Approximate effect of the subgrid-scale on the grid-scale
- MITgcm packages:
  - gmredi: Gent-McWilliams/Redi Eddy Parameterization
  - Other mixing schemes: KPP, GGL90, KL10



Why stochastic parametrization?

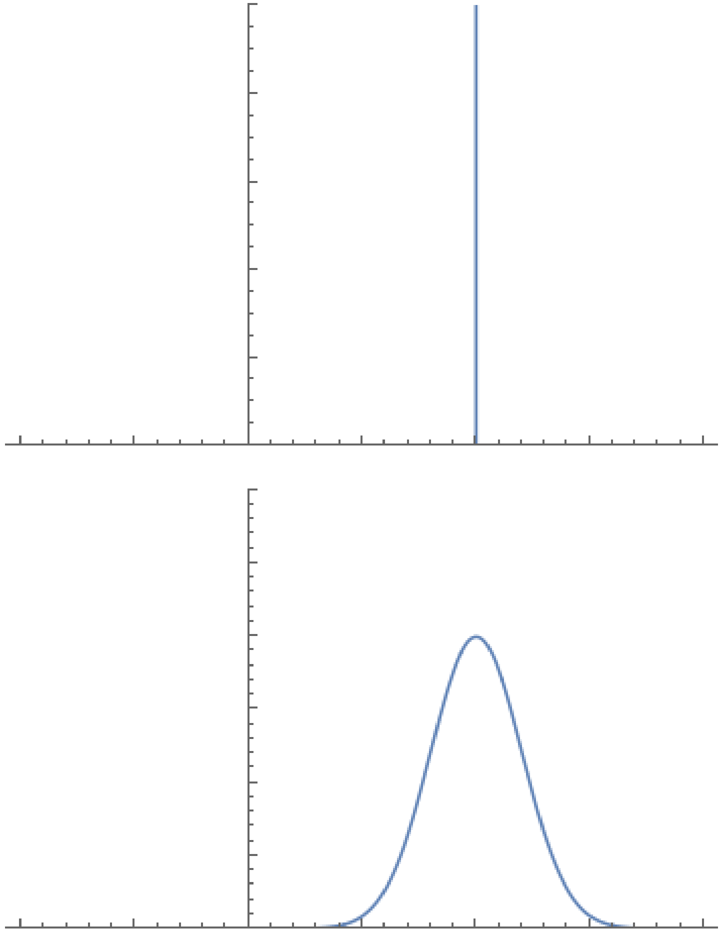


$$U_{\text{det}}\left(\text{params},\ \text{state}\right)$$

Why stochastic parametrization?

$$U_{\mathrm{det}}\left(\mathrm{params},\ \mathrm{state}\right)$$

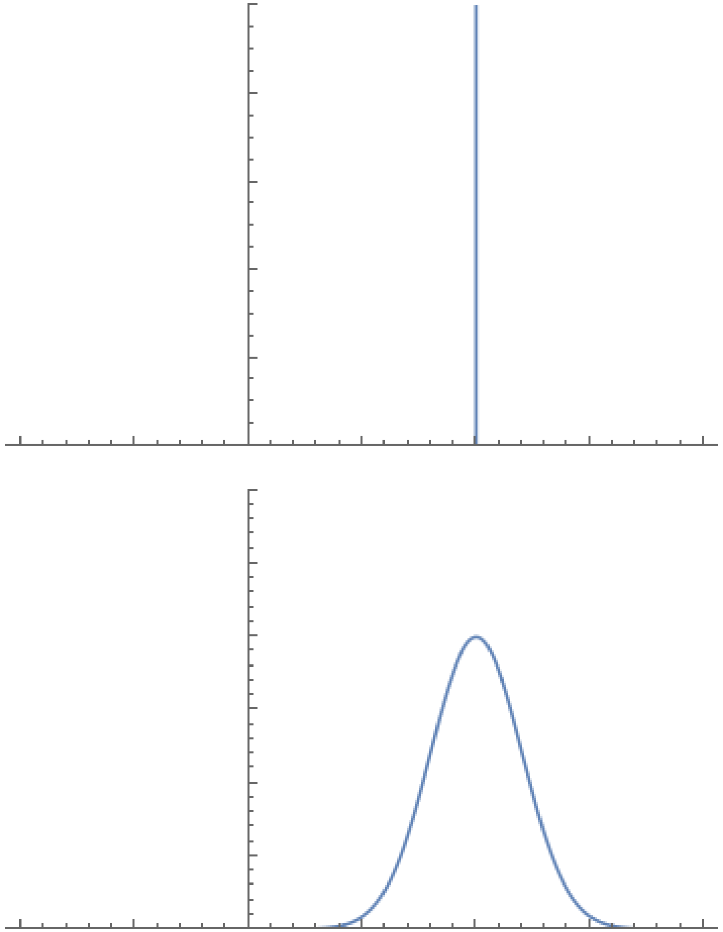
$$U_{\mathrm{stoch}}\left(\mathrm{params},\ \mathrm{state},\ \mathrm{stats}\right) = U_{\mathrm{det}}\ + e\left(t,\ \mathrm{stats}\right)$$



Why stochastic parametrization?

$$U_{\mathrm{det}}\left(\mathrm{params},\ \mathrm{state}\right)$$

$$U_{\mathrm{stoch}}\left(\mathrm{params},\ \mathrm{state},\ \mathrm{stats}\right) = U_{\mathrm{det}}\ + e\left(t,\ \mathrm{stats}\right)$$



- More accurate mean model state and *variability*
  - i.e. statistical consistency of data and model
- Realistic variability ⇒ better uncertainty estimates

Lorenz 96 2-scale model

$$\frac{dX_k}{dt} = -X_{k-1}\left(X_{k-2} - X_{k+1}\right) - X_k + F - \frac{hc}{b}\sum_{j=J(k-1)+1}^{kJ} Y_j\ ; \quad k = 1, ..., K$$

$$\frac{dY_j}{dt} = -cb\,Y_{j+1}\left(Y_{j+2} - Y_{j-1}\right) - c\,Y_j + \frac{hc}{b} X_{\mathrm{int}[(j-1)/J]+1}\ ; \quad j = 1, ..., JK$$

- *K*: number of *X* variables
- *J*: number of *Y* variables per *X* variable
- *h*: coupling (>0, sets degree of coupling)
- *F*: forcing (>0, sets degree of driving)
- *b*: spatial-scale ratio (>1 ⇒ large *X*, small *Y*)
- *c*: time-scale ratio (>1 ⇒ slow *X*, fast *Y*)

Lorenz 96 2-scale system

$$\frac{dX_k}{dt} = \boxed{-X_k + F} \qquad k = 1, ..., K$$

Self-damping/forcing

- *K*: number of *X* variables
- *J*: number of *Y* variables per *X* variable
- *h*: coupling (>0, sets degree of coupling)
- *F*: forcing (>0, sets degree of driving)
- *b*: spatial-scale ratio (>1 ⇒ large *X*, small *Y*)
- *c*: time-scale ratio (>1 ⇒ slow *X*, fast *Y*)

Lorenz 96 2-scale system

$$\frac{dX_k}{dt} = \boxed{-X_{k-1}\left(X_{k-2} - X_{k+1}\right)} \boxed{- X_k + F} \qquad k = 1, \ldots, K$$

- *K*: number of *X* variables
- *J*: number of *Y* variables per *X* variable
- *h*: coupling (>0, sets degree of coupling)
- *F*: forcing (>0, sets degree of driving)
- *b*: spatial-scale ratio (>1 ⇒ large *X*, small *Y*)
- *c*: time-scale ratio (>1 ⇒ slow *X*, fast *Y*)

Self-damping/forcing
Neighbor coupling

Lorenz 96 2-scale system

$$\frac{dX_k}{dt} = -X_{k-1}\left(X_{k-2} - X_{k+1}\right) - X_k + F - \frac{hc}{b}\sum_{j=J(k-1)+1}^{kJ} Y_j \; ; \qquad k = 1, \ldots, K$$

- *K*: number of *X* variables
- *J*: number of *Y* variables per *X* variable
- *h*: coupling (>0, sets degree of coupling)
- *F*: forcing (>0, sets degree of driving)
- *b*: spatial-scale ratio (>1 ⇒ large *X*, small *Y*)
- *c*: time-scale ratio (>1 ⇒ slow *X*, fast *Y*)

Self-damping/forcing
Neighbor coupling
Inter-scale coupling

Lorenz 96 2-scale system

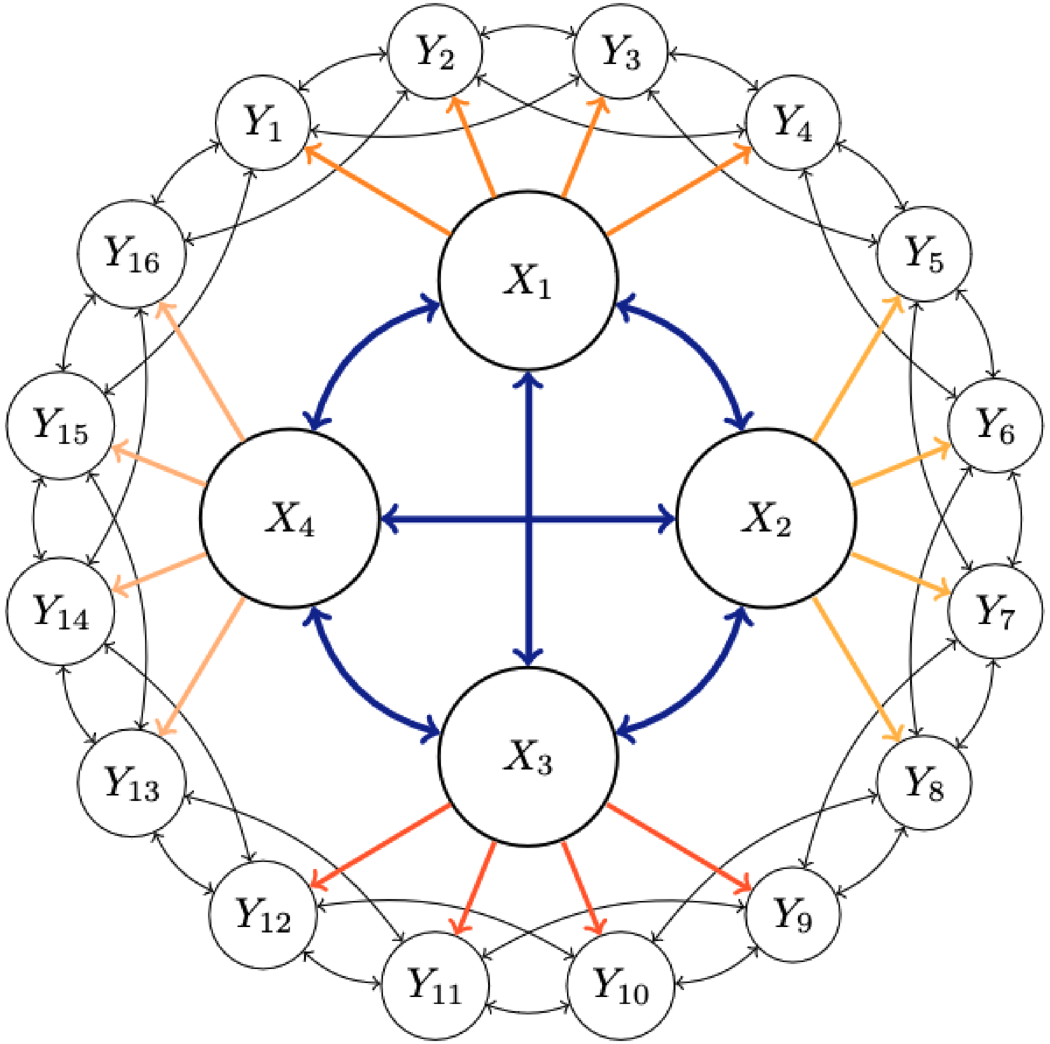
$$\frac{dX_k}{dt} = -X_{k-1}\left(X_{k-2} - X_{k+1}\right) - X_k + F - \frac{hc}{b} \sum_{j=J(k-1)+1}^{kJ} Y_j \; ; \qquad k = 1, \ldots, K$$

$$\frac{dY_j}{dt} = -cb\,Y_{j+1}\left(Y_{j+2} - Y_{j-1}\right) - c\,Y_j + \frac{hc}{b} X_{\mathrm{int}[(j-1)/J]+1} \; ; \qquad j = 1, \ldots, JK$$

- *K*: number of *X* variables
- *J*: number of *Y* variables per *X* variable
- *h*: coupling (>0, sets degree of coupling)
- *F*: forcing (>0, sets degree of driving)
- *b*: spatial-scale ratio (>1 ⇒ large *X*, small *Y*)
- *c*: time-scale ratio (>1 ⇒ slow *X*, fast *Y*)

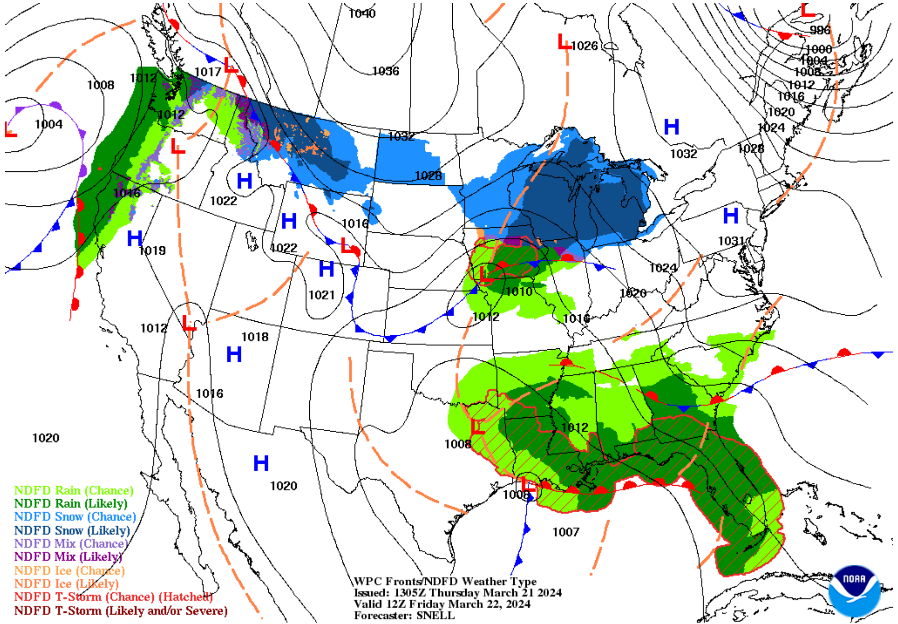
Self-damping/forcing
Neighbor coupling
Inter-scale coupling

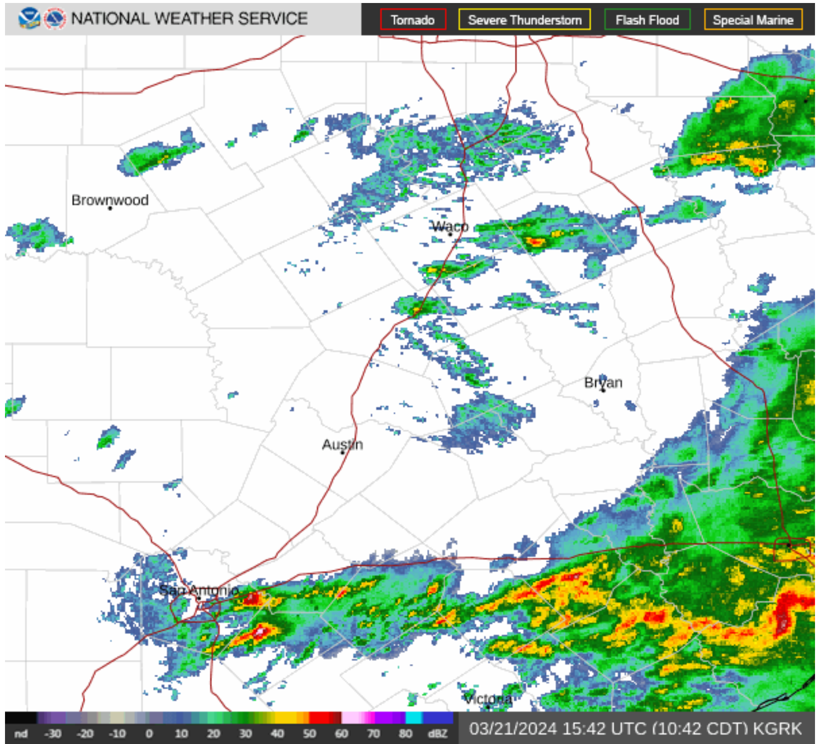
Lorenz 96 2-scale system



Lorenz 96 2-scale system

- Created as a toy weather model
  - *X*: large scale synoptic dynamics (e.g. low pressure system)
  - *Y*: convective events (e.g. local thunderstorms or showers)
  - Parameter choices yield chaotic *X* and *Y*

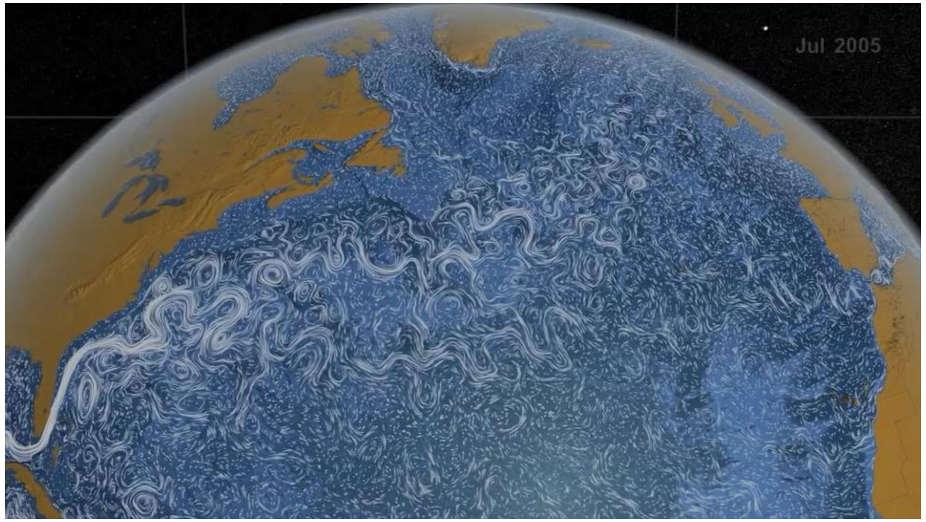




Lorenz 96 2-scale system

- Created as a toy weather model
  - *X*: large scale synoptic dynamics (e.g. low pressure system)
  - *Y*: convective events (e.g. local thunderstorms or showers)
  - Parameter choices yield chaotic *X* and *Y*
- Could it be a proxy for ocean processes as well?

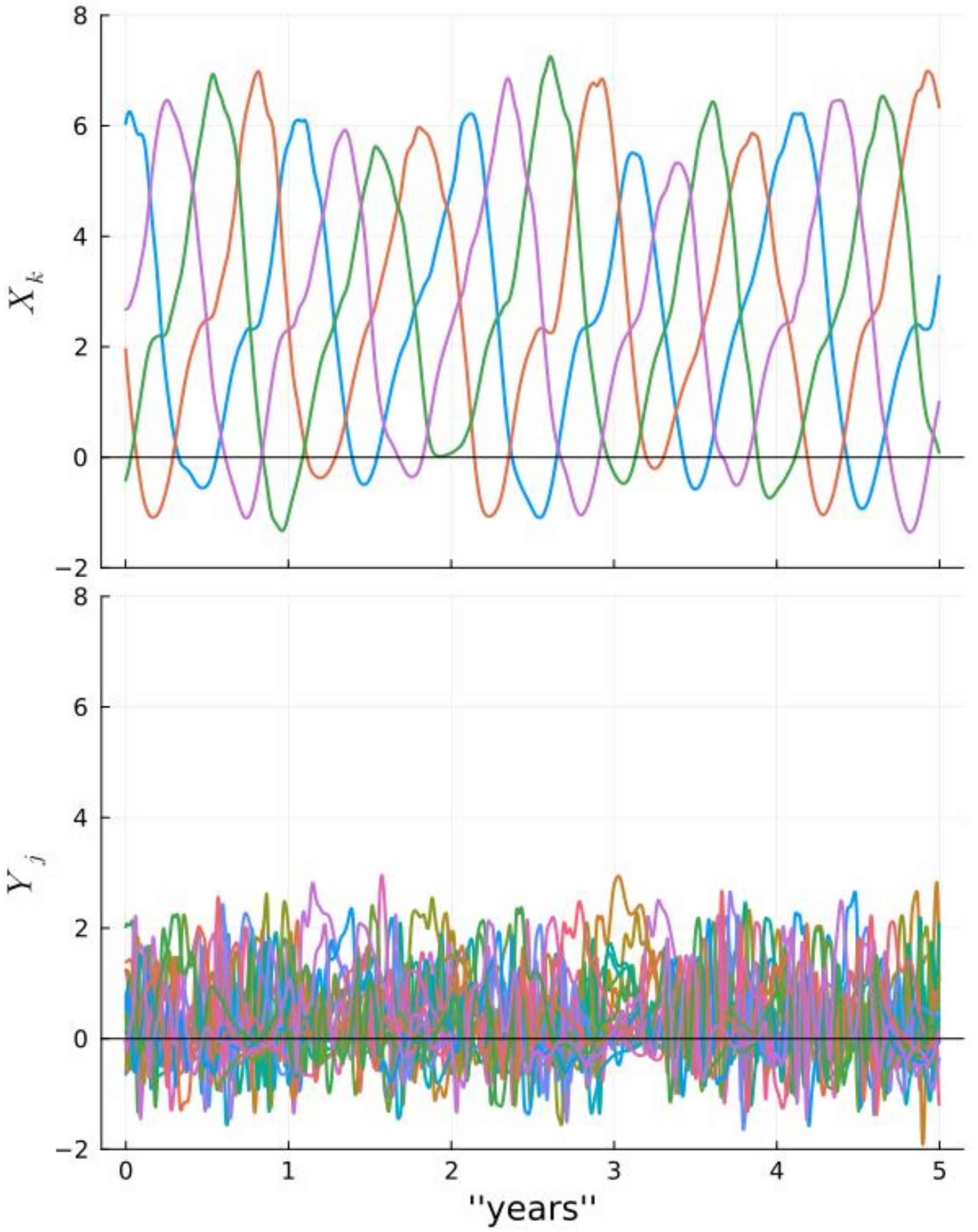
Lorenz 96 2-scale system



- Could it be a proxy for ocean processes as well?
  - *X*: large scale current and gyre dynamics
  - *Y*: ocean mesoscale eddies
  - Parameter choices yield **quasiperiodic** *X* and **chaotic** *Y*

- “Ocean” parameter choices:
  - $K = 4$ ⇒ 4 interacting currents
  - $J = 5$ ⇒ 5 eddy regions per current
  - $h = 2$ ⇒ Strong coupling
  - $F = 10$ ⇒ Enough to drive oscillations
  - $b = 5$ ⇒ Y amplitude 1/5 of X
  - $c = 5$ ⇒ Y frequency 5 times X

- "Ocean" parameter choices:
  - $K = 4$ ⇒ 4 interacting currents
  - $J = 5$ ⇒ 5 eddy regions per current
  - $h = 2$ ⇒ Strong coupling
  - $F = 10$ ⇒ Enough to drive oscillations
  - $b = 5$ ⇒ Y amplitude 1/5 of X
  - $c = 5$ ⇒ Y frequency 5 times X
- The result?
  - Quasiperiodic oscillations in X
  - Chaos in Y



Parametrization?

- An approximate *parametrized* model of L96
  - Explicitly model only the X_k
  - Parametrize the effect of the Y_j on the X_k
  - X_k drive Y_j dynamics, which in turn affect X_k

$$\frac{dX_k}{dt} = -X_{k-1}\left(X_{k-2} - X_{k+1}\right) - X_k + F - \frac{hc}{b}\sum_{j=J(k-1)+1}^{kJ} Y_j \;; \quad k = 1, ..., K$$

$$\Rightarrow \quad \frac{dX_k^*}{dt} = -X_{k-1}^*\left(X_{k-2}^* - X_{k+1}^*\right) - X_k^* + F + U_p\left(X_k^*\right) \;; \quad k = 1, ..., K$$

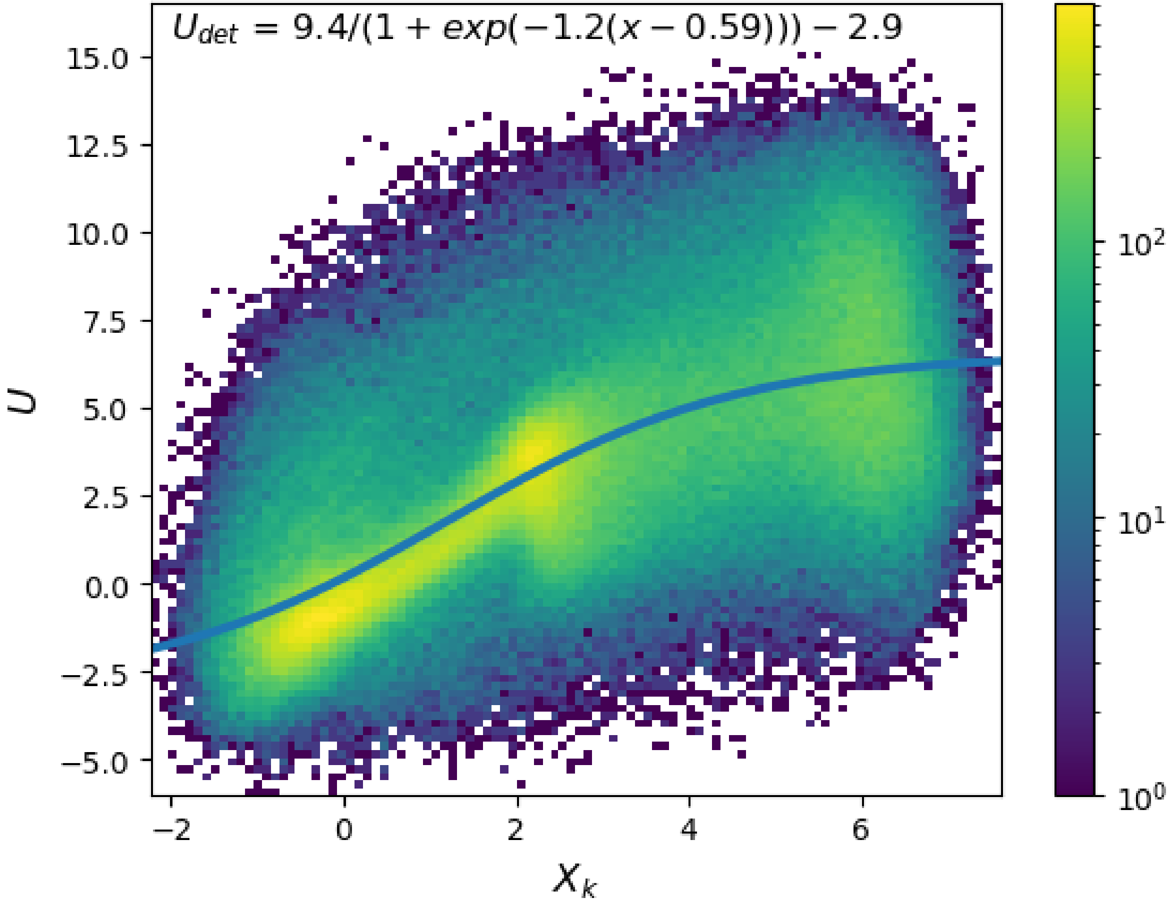
First guess

- Use full model “data”: least squares fit to a logistic curve

$$U_{\text{det}}\left(X^*\right) = \frac{L}{1 + e^{-k(X^* - X_0)}} + b$$

First guess

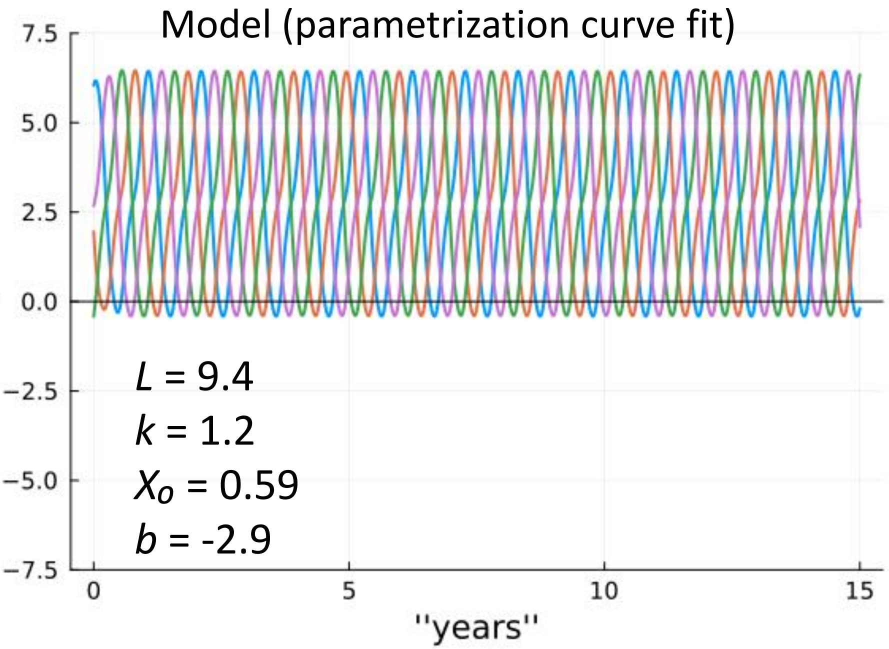
- Use full model "data": least squares fit to a logistic curve



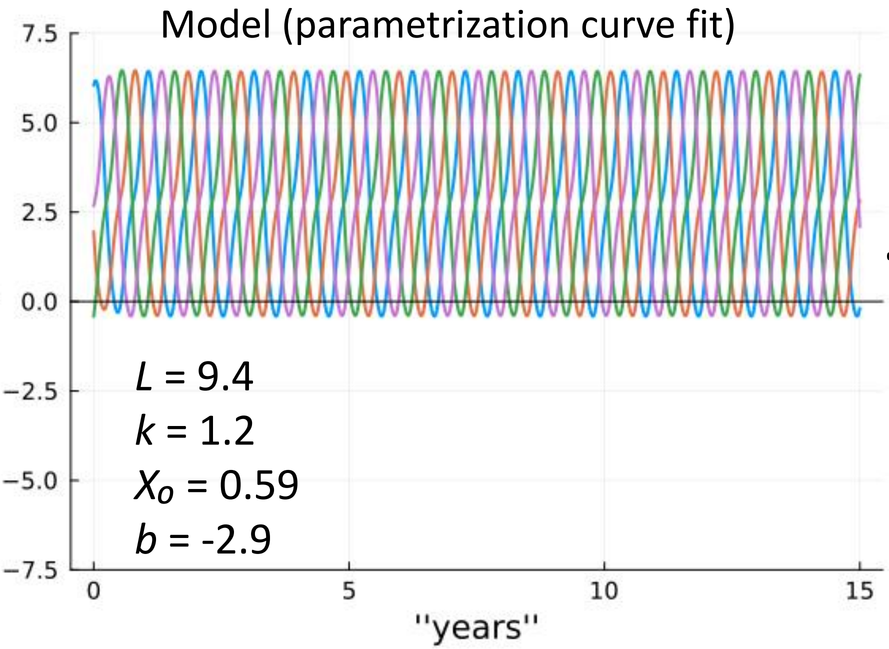
$$U_{\mathrm{det}}\left(X^*\right) = \frac{L}{1 + e^{-k(X^* - X_0)}} + b$$

- L = 9.4
- k = 1.2
- X_o = 0.59
- b = -2.9

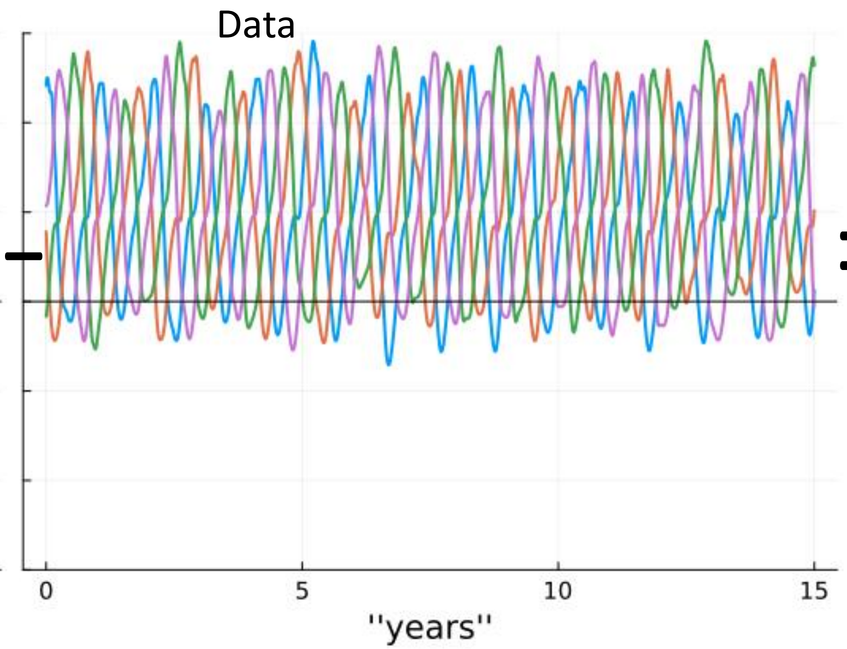
How does it do?



How does it do?

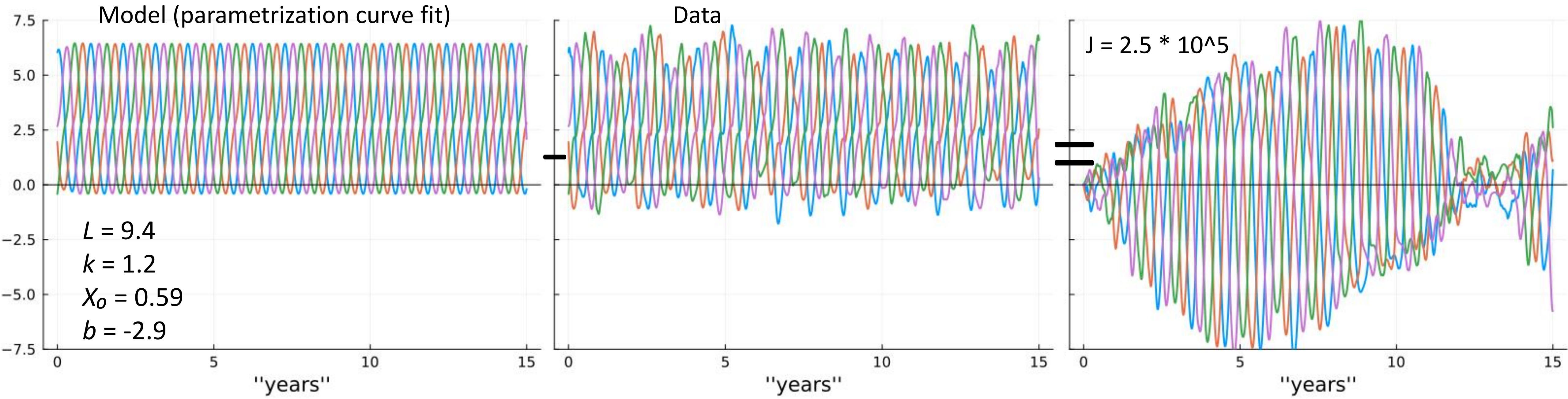


−

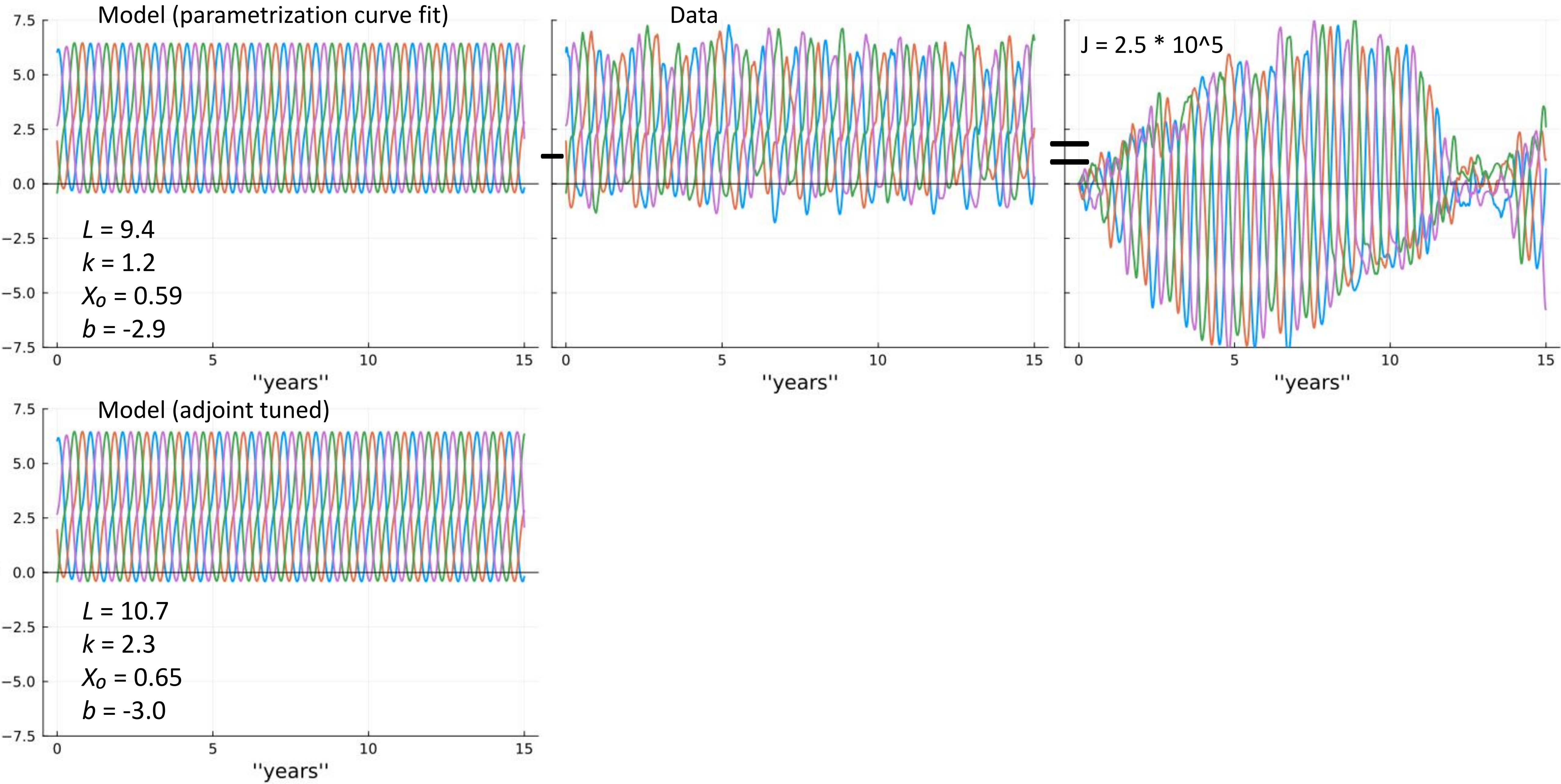


=

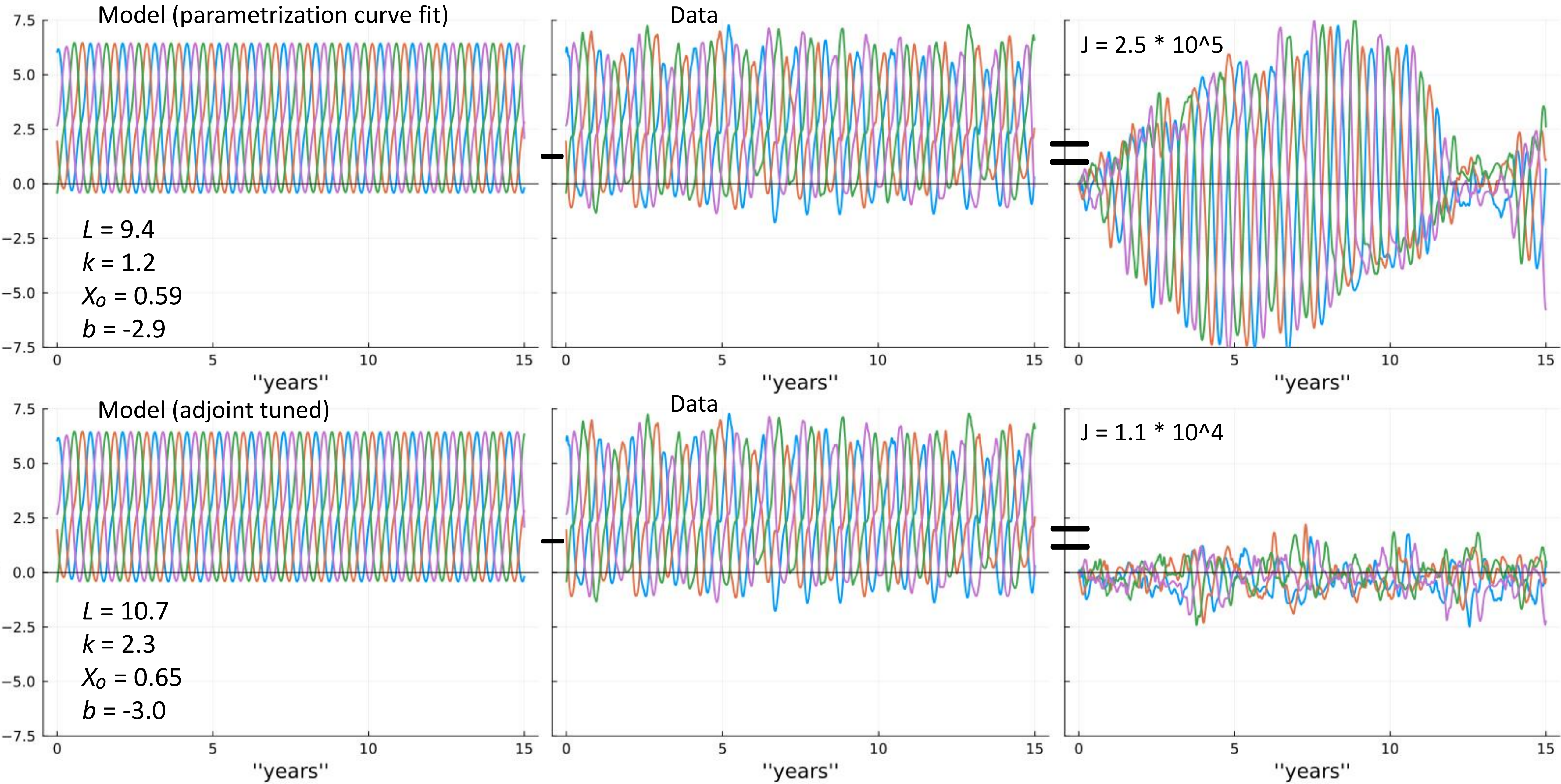
How does it do?



How does it do?

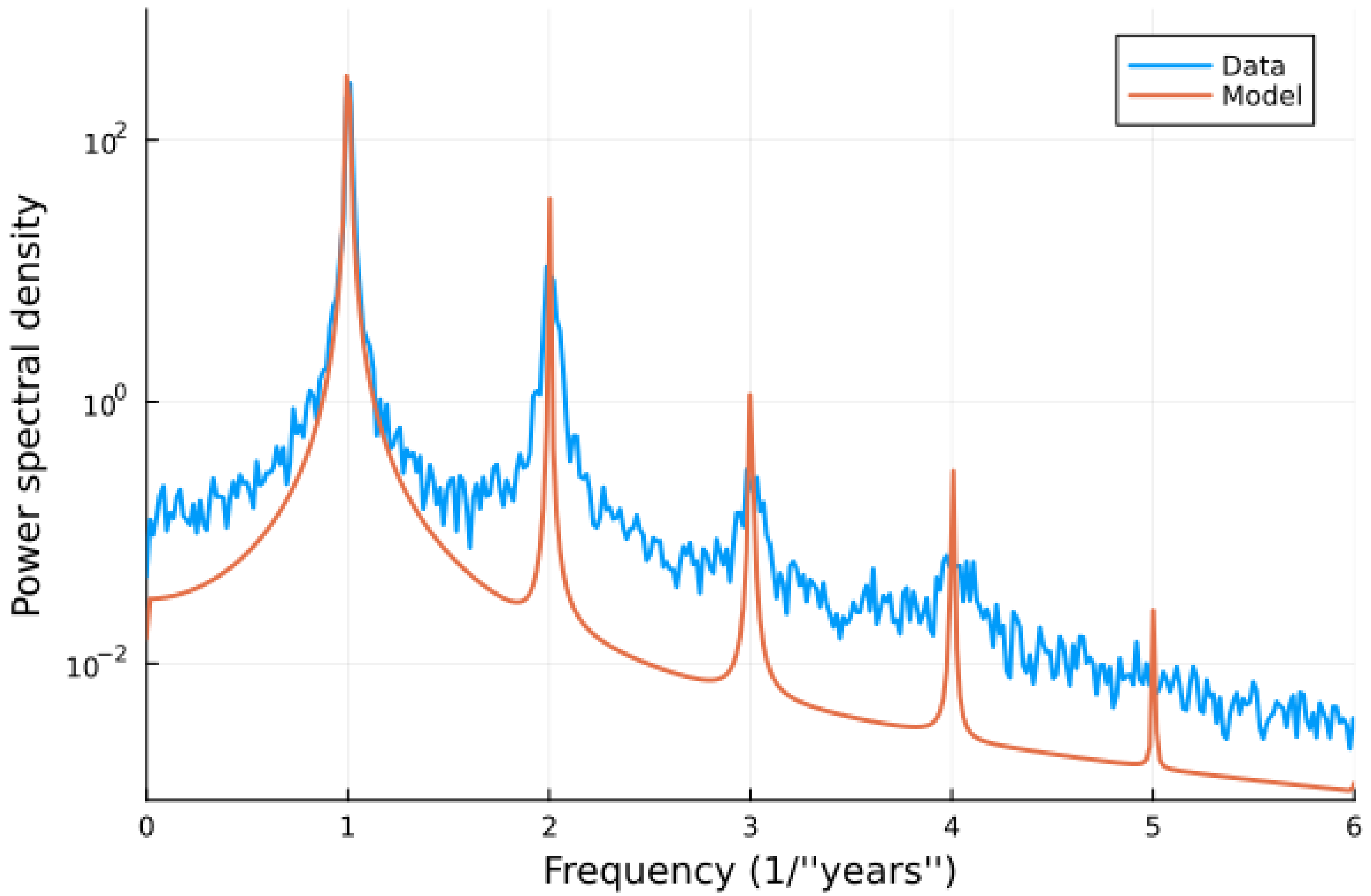


How does it do?



What about statistics?

What about statistics?



And stochastic parametrization?

$$\frac{dX_k^*}{dt} = -X_{k-1}^*\left(X_{k-2}^* - X_{k+1}^*\right) - X_k^* + F + U_p\left(X_k^*\right)\ ; \quad k = 1, ..., K$$

And stochastic parametrization?

$$\frac{dX_k^*}{dt} = -X_{k-1}^*\left(X_{k-2}^* - X_{k+1}^*\right) - X_k^* + F + U_p\left(X_k^*\right) \ ; \quad k = 1, ..., K$$

Deterministic: $$U_{\mathrm{det}}\left(X^*\right) = \frac{L}{1 + e^{-k(X^* - X_0)}} + b$$

And stochastic parametrization?

$$\frac{dX_k^*}{dt} = -X_{k-1}^*\left(X_{k-2}^* - X_{k+1}^*\right) - X_k^* + F + U_p\left(X_k^*\right) \;\; ; \qquad k = 1, ..., K$$

Deterministic: $$U_{\text{det}}\left(X^*\right) = \frac{L}{1 + e^{-k(X^* - X_0)}} + b$$

Stochastic: $$U_p\left(X^*\right) = U_{\text{det}}\left(X^*\right) + e(t)$$

$$e(t) = \phi\, e(t - \Delta t) + \sigma\sqrt{1 - \phi^2}\, z(t)$$

And stochastic parametrization?

$$\frac{dX_k^*}{dt} = -X_{k-1}^*\left(X_{k-2}^* - X_{k+1}^*\right) - X_k^* + F + U_p\left(X_k^*\right) \; ; \quad k = 1, ..., K$$

Deterministic: $$U_{\mathrm{det}}\left(X^*\right) = \frac{L}{1 + e^{-k(X^* - X_0)}} + b$$

Stochastic: $$U_p\left(X^*\right) = U_{\mathrm{det}}\left(X^*\right) + e(t)$$

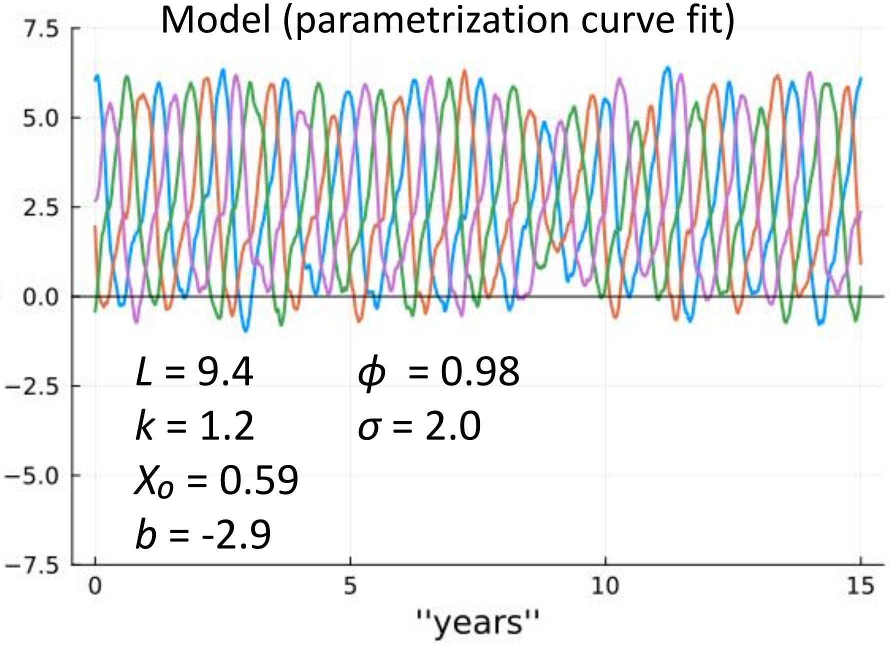
$$e(t) = \phi\, e(t - \Delta t) + \sigma\sqrt{1 - \phi^2}\, z(t)$$

lag-1 autocorrelation

SD of *U* residual

unit variance Gaussian white noise

How about stochastic parametrization?



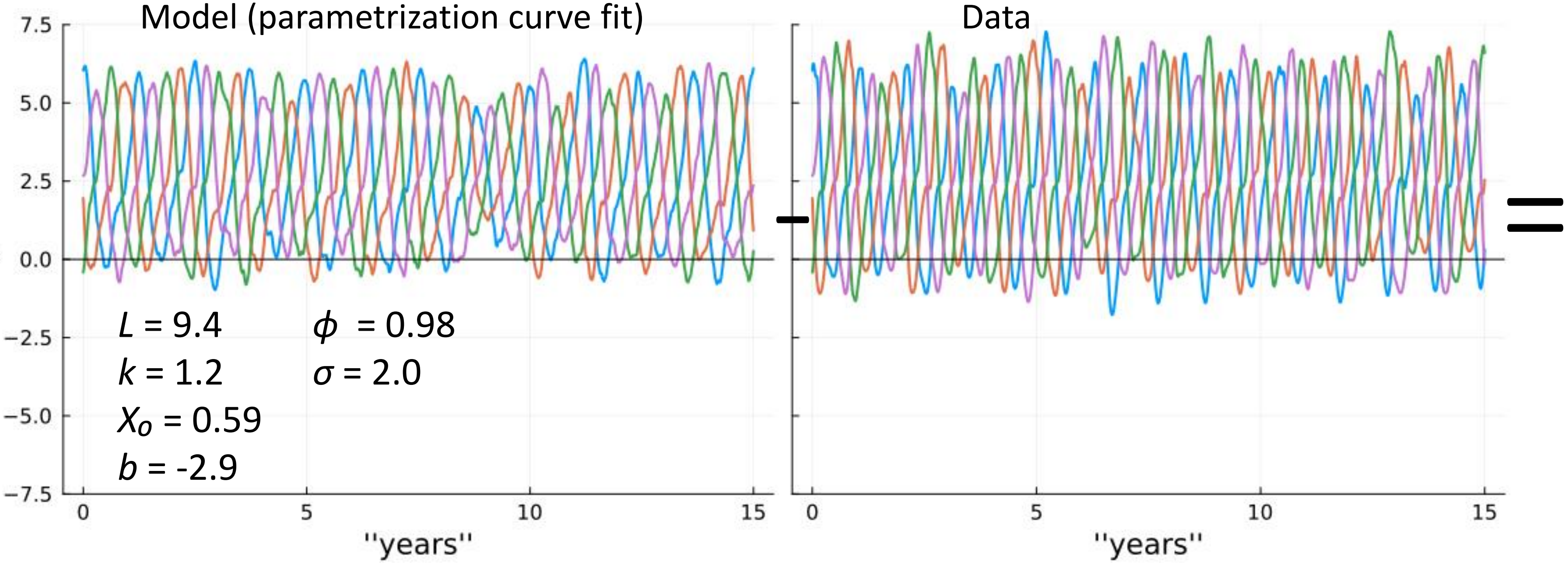
$L = 9.4$ $\phi = 0.98$

$k = 1.2$ $\sigma = 2.0$

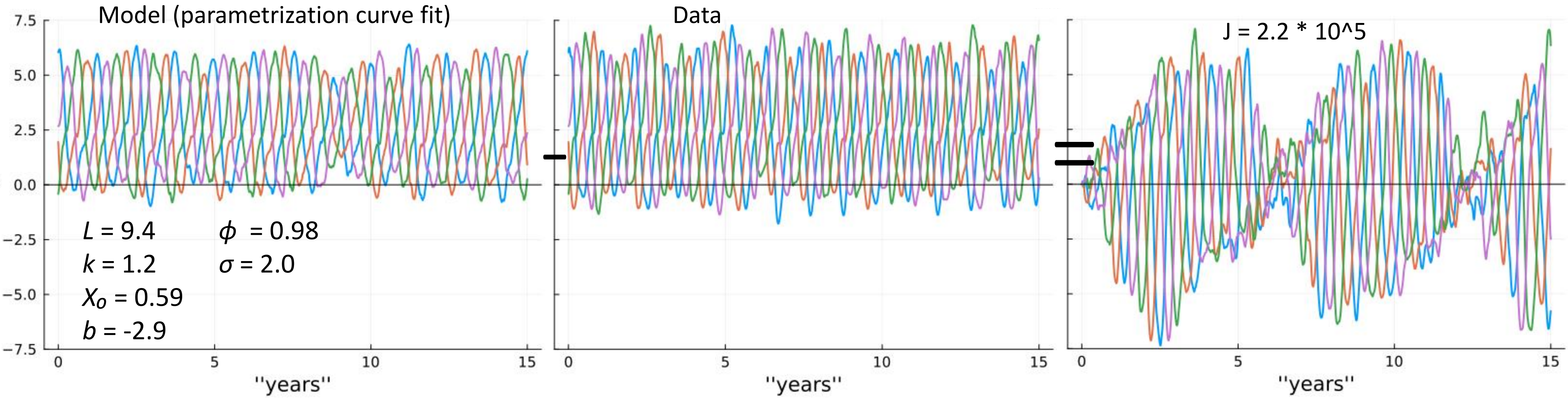
$X_o = 0.59$

$b = -2.9$

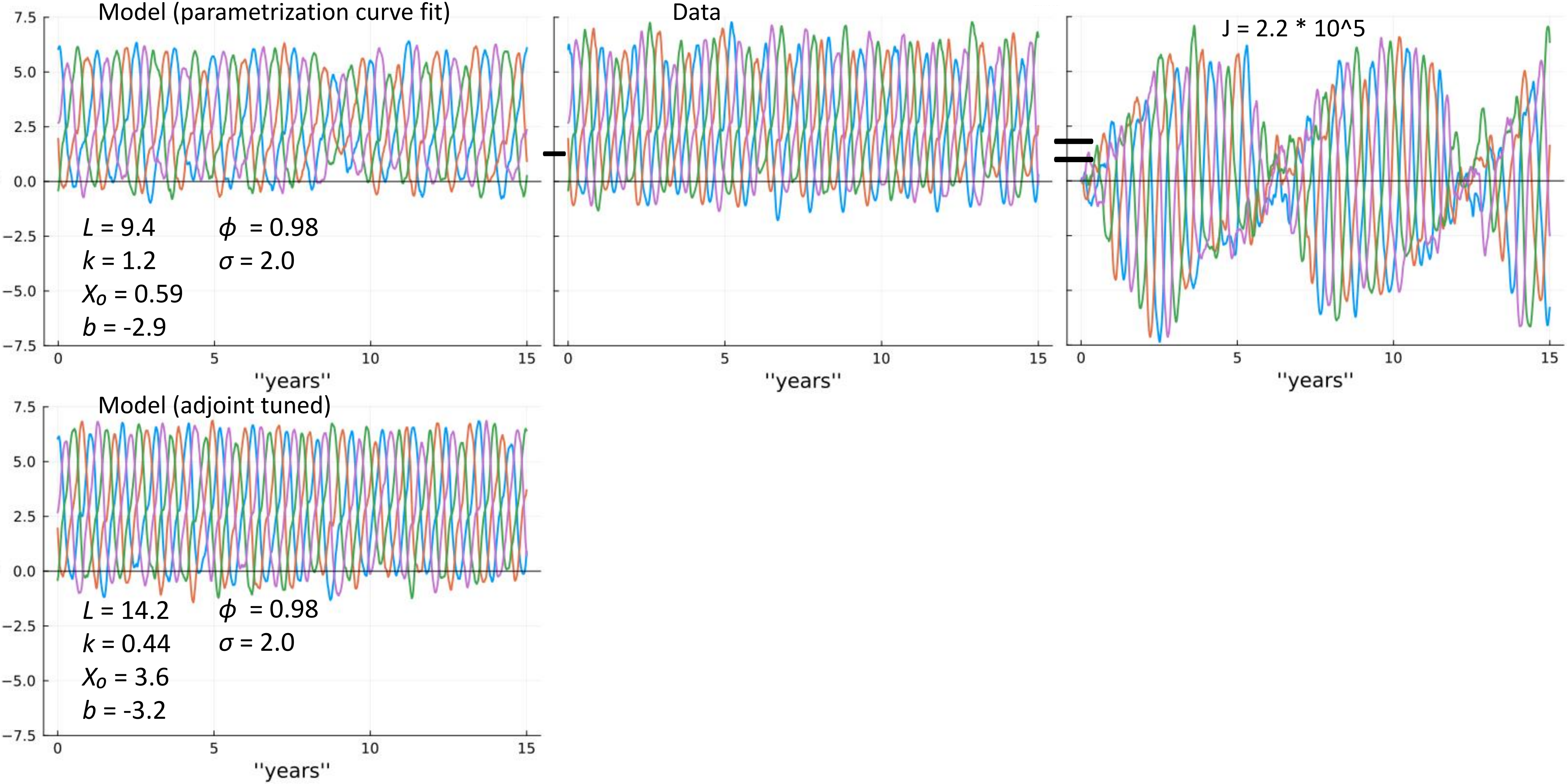
How about stochastic parametrization?



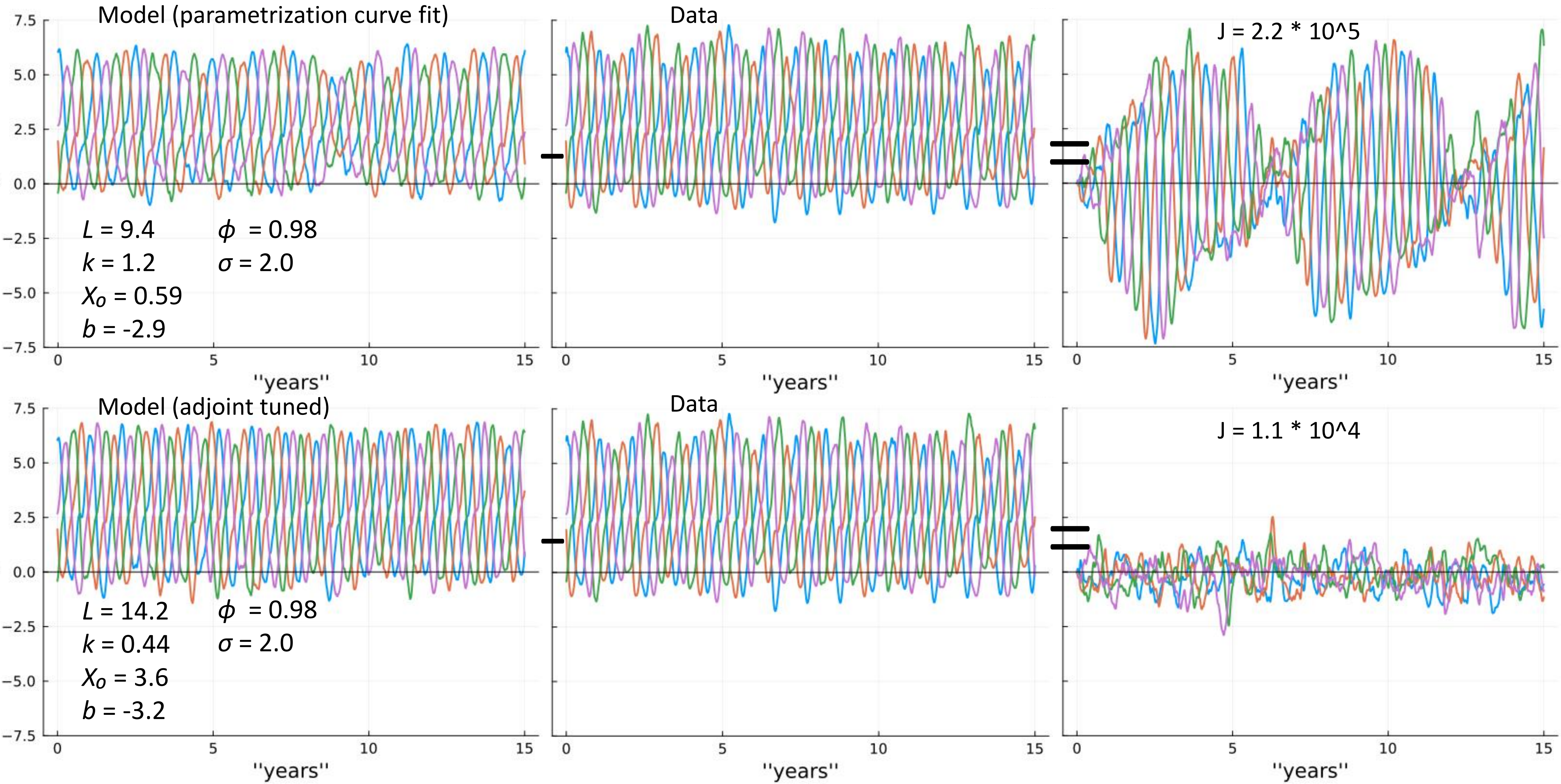
How about stochastic parametrization?



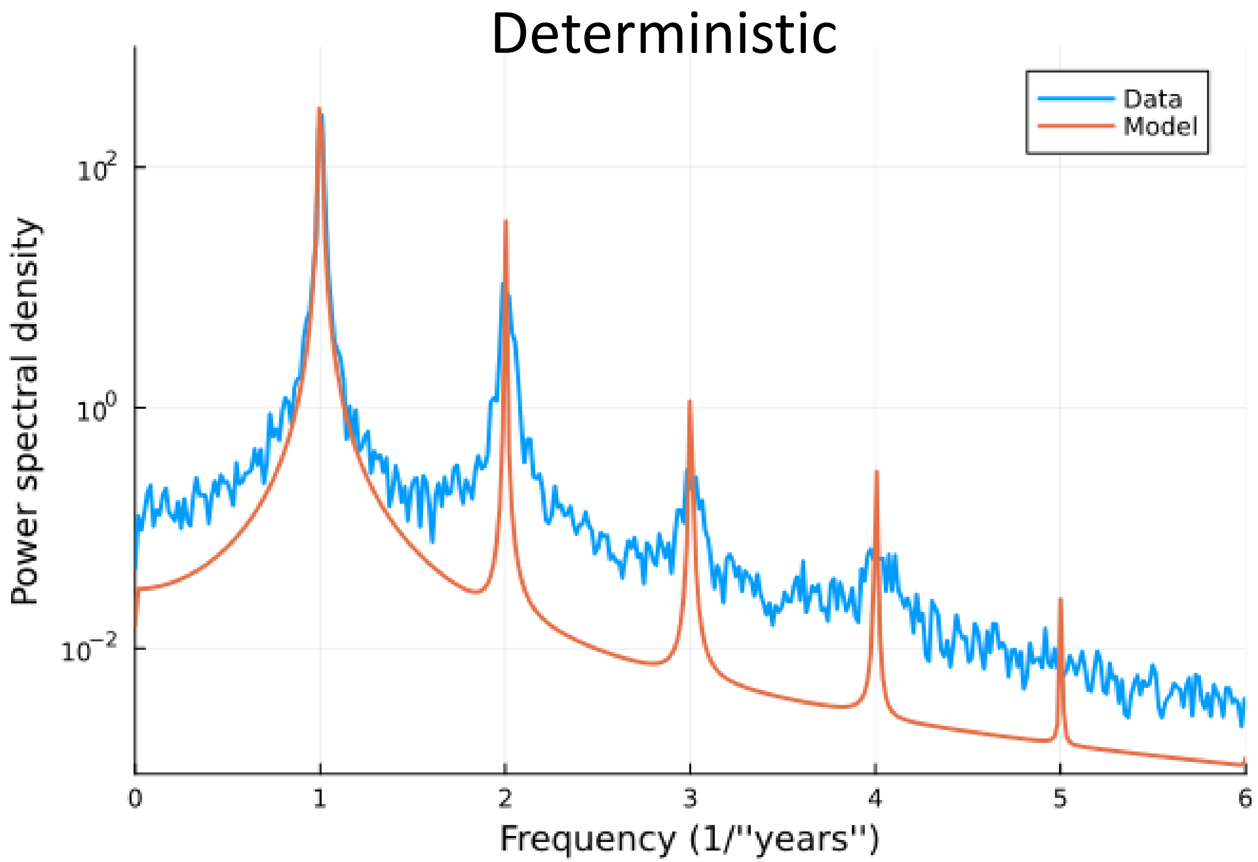
How about stochastic parametrization?



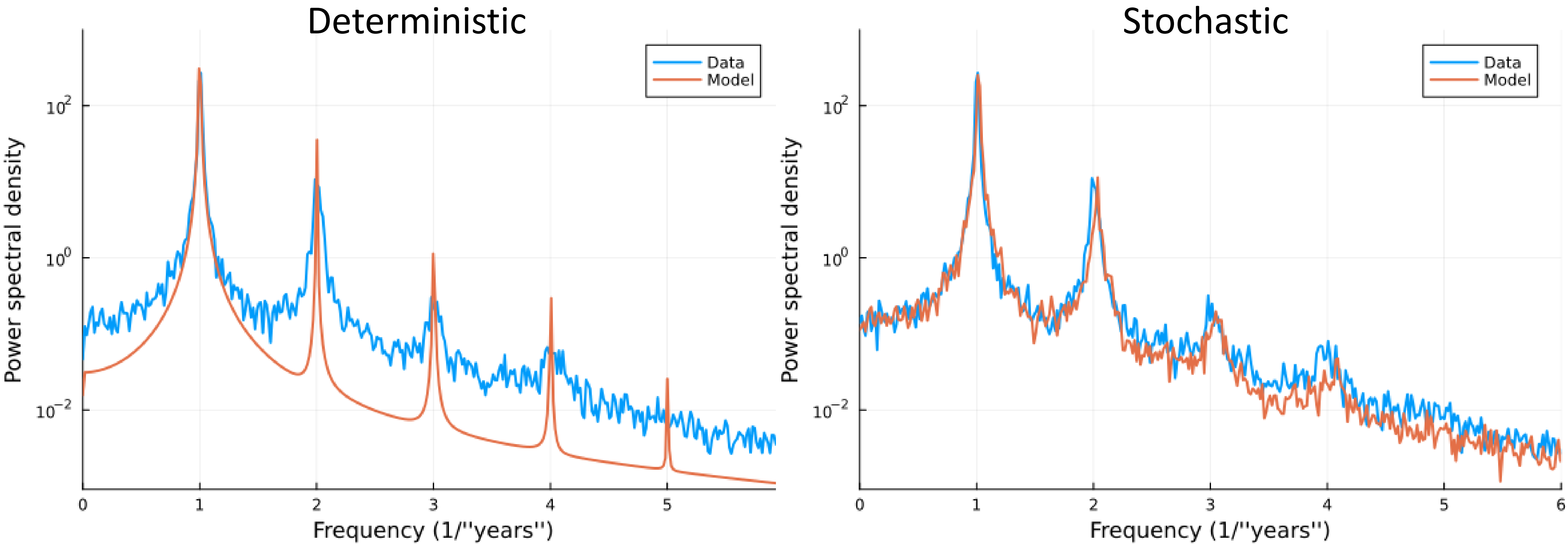
How about stochastic parametrization?



What about statistics?



What about statistics?



A look ahead

A look ahead

- Further explorations in L96
  - Uncertainty quantification
  - Stability
  - Forecast skill
  - PSD misfit in cost function?

A look ahead

- Further explorations in L96
  - Uncertainty quantification
  - Stability
  - Forecast skill
  - Adjoint tuned statistics? PSD misfit in cost function?
- Idealized MITgcm setups
  - Baroclinic gyre with eddy parametrization
  - soma (MPAS-Ocean ideal model with continental shelf)