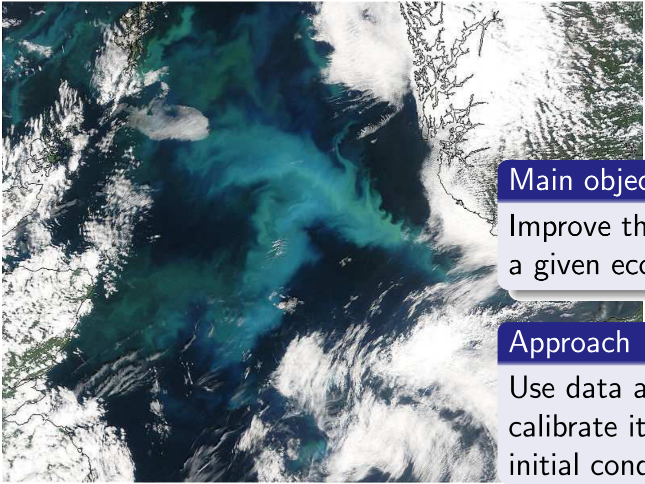
Model-reduced 4D-Var data assimilation in application to 1D ecosystem model

Joanna S. Pelc[2,3], Ehouarn Simon[1], Laurent Bertino[1], Ghada El Serafy[2,3], and Arnold W. Heemink[2]

Nansen Environmental and Remote Sensing Center, Bergen, Norway[1]
Delft University of Technology, Delft, The Netherlands[2]
Deltares, Delft, The Netherlands[3]

Outline

Introduction



Main objective

Improve the predictions of a given ecological model

Approach

Use data assimilation to calibrate its parameters & initial conditions

1D Ecological Model

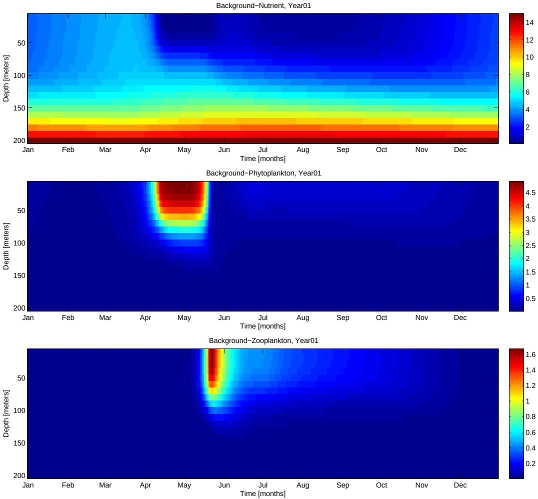
State variables

1. Nutrients (N)
2. Phytoplankton (P)
3. Herbivorous zooplankton (H)

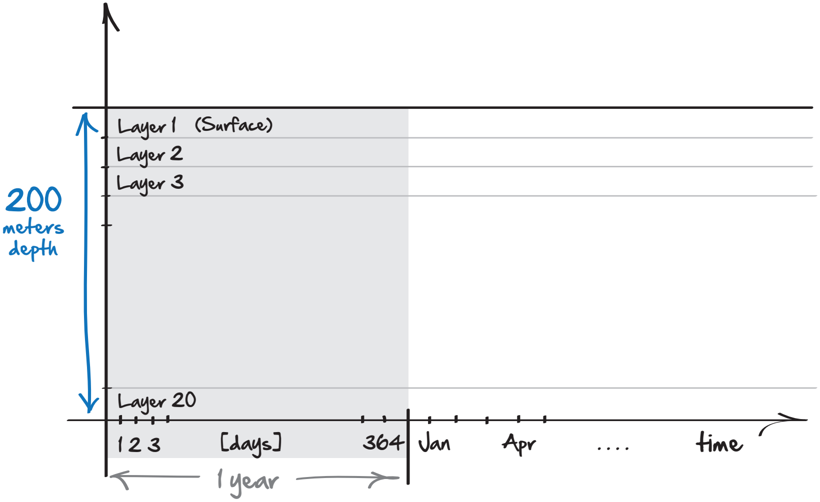
Estimated parameters

f grazing efficiency
g loss to carnivores
r plant metabolic loss

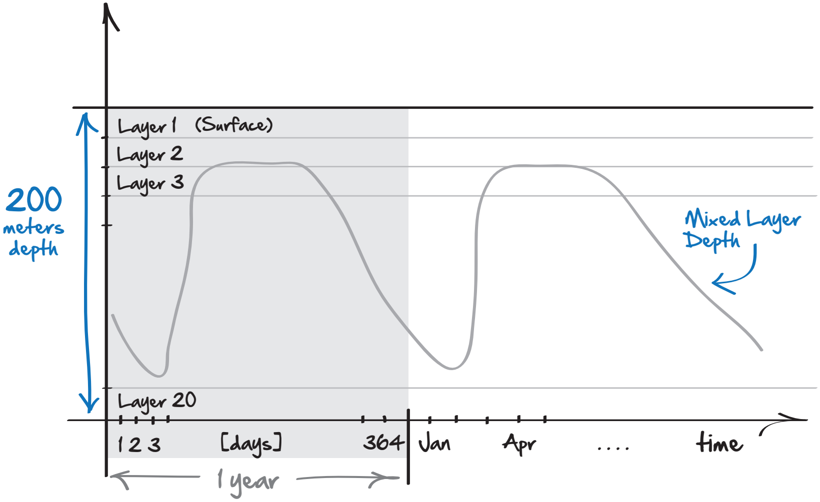
Evans and Parslow (1985)
Eknes and Evensen (2002)



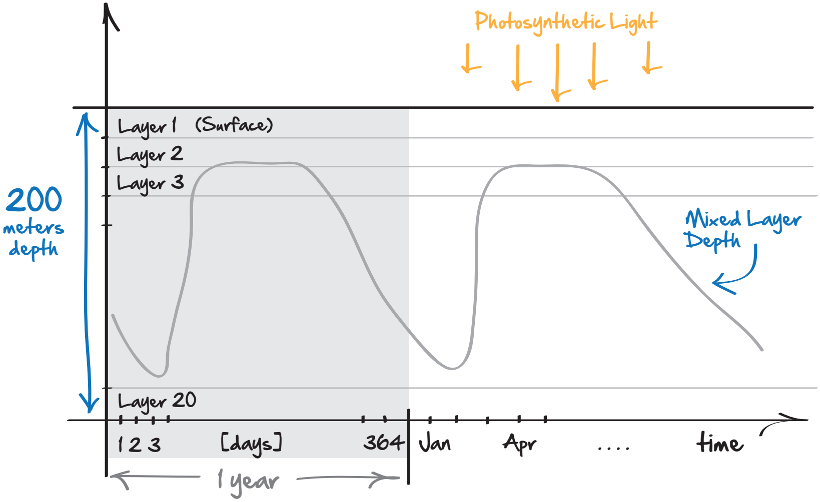
1D Ecological Model



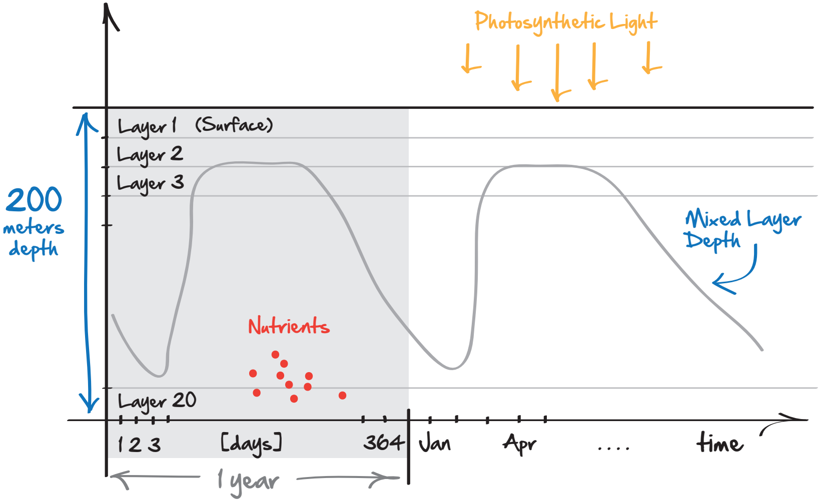
1D Ecological Model



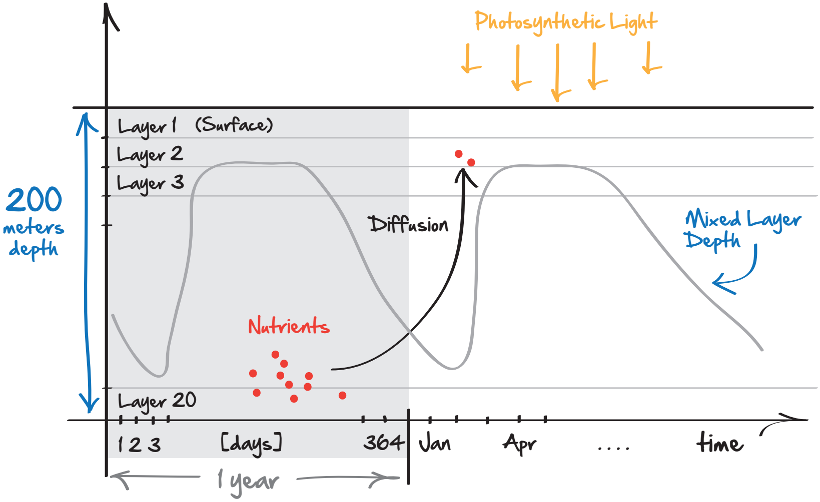
1D Ecological Model



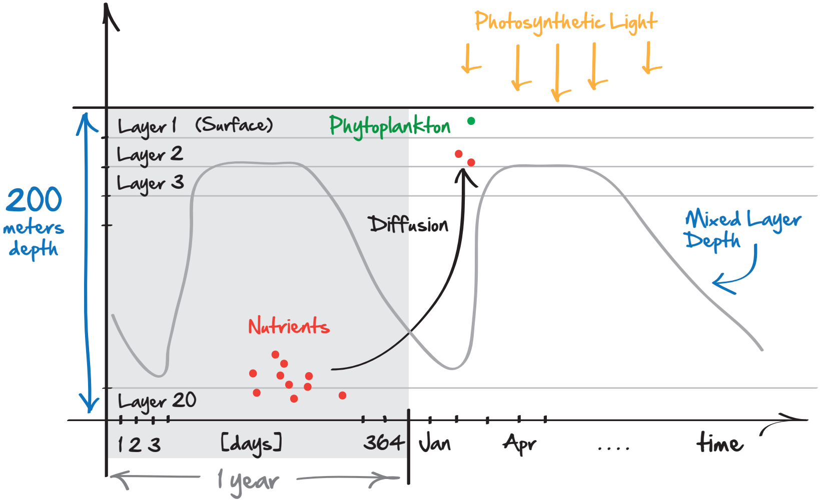
1D Ecological Model



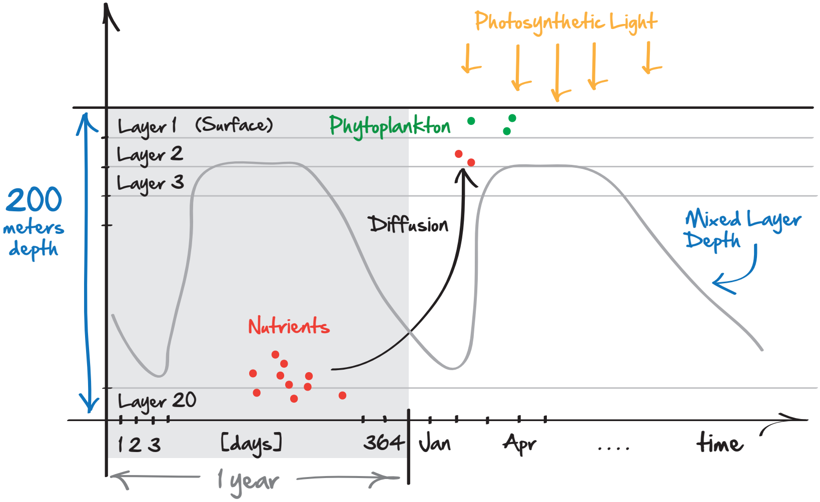
1D Ecological Model



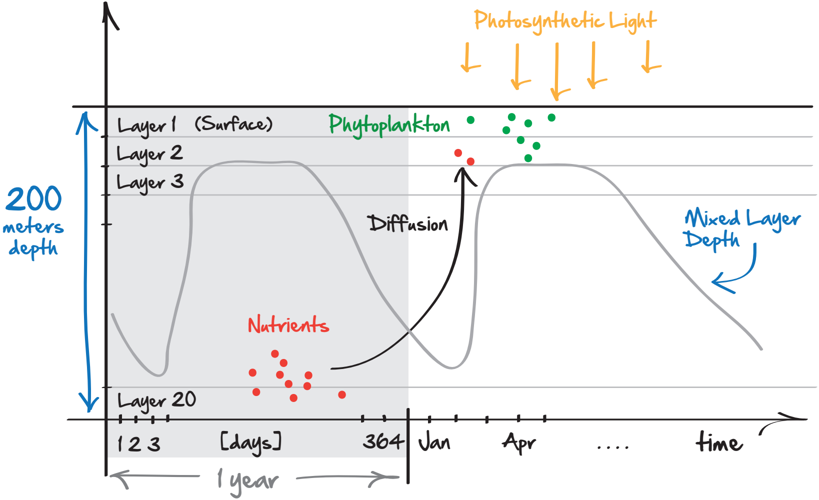
1D Ecological Model



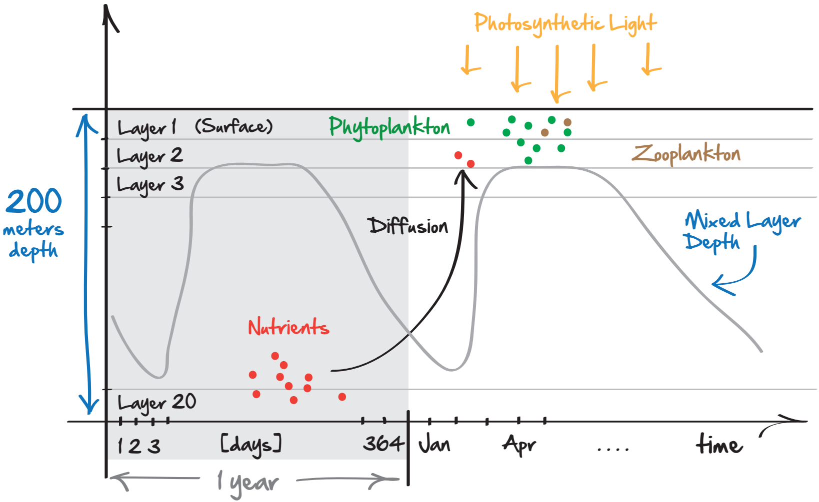
1D Ecological Model



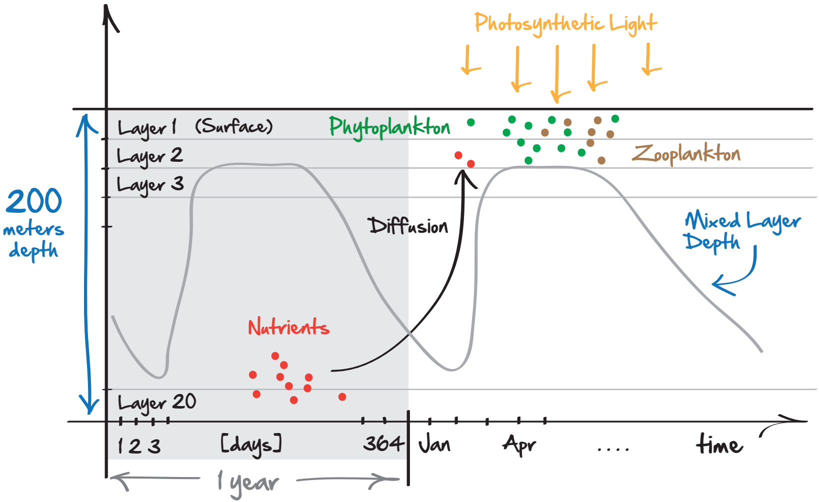
1D Ecological Model



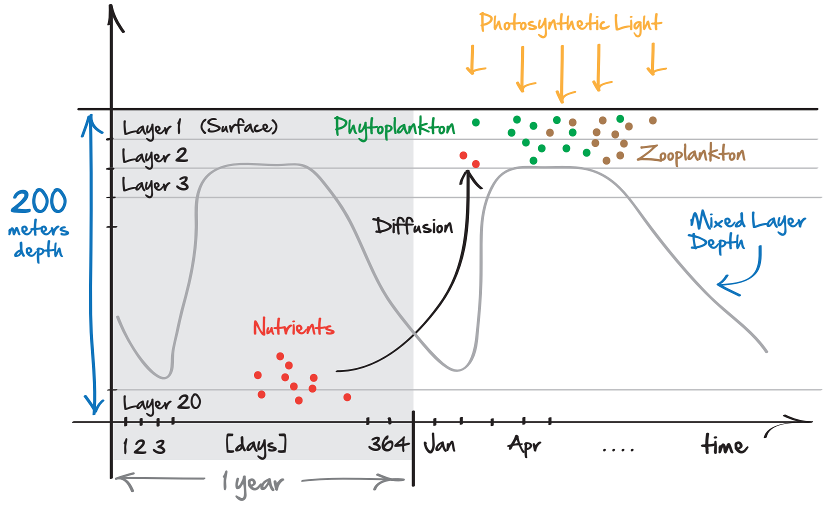
1D Ecological Model



1D Ecological Model



1D Ecological Model



4D-Var: 4D Variational Data Assimilation

Cost function in 4D-Var

$$J(\alpha) = \sum_{i=1}^{n} \underbrace{(\mathcal{H}_i(x_i) - y_i)^T}_{Distances} R_i^{-1} \underbrace{(\mathcal{H}_i(x_i) - y_i)}_{Distances} + \underbrace{(\alpha - \alpha_b)^T B^{-1} (\alpha - \alpha_b)}_{Background \quad term}$$

minimization with constraints:
$x_i = \mathcal{M}_i(x_{i-1}, \alpha)$

B background error covariance matrix
R_i observation error covariance matrix

To minimize J over α, we need: $\nabla_\alpha J(\alpha)$

Needed to get $\nabla_\alpha J(\alpha)$

- exact derivatives of the model
- approximate the derivatives of the model with finite differences

4D-Var: 4D Variational Data Assimilation

Cost function in 4D-Var

$$J(\alpha) = \sum_{i=1}^{n} \underbrace{(\mathcal{H}_i(x_i) - y_i)}_{Distances}^T R_i^{-1} \underbrace{(\mathcal{H}_i(x_i) - y_i)}_{Distances} + \underbrace{(\alpha - \alpha_b)^T B^{-1} (\alpha - \alpha_b)}_{Background \quad term}$$

minimization with constraints:
$x_i = \mathcal{M}_i(x_{i-1}, \alpha)$

B background error covariance matrix
R_i observation error covariance matrix

To minimize J over α, we need: $\nabla_\alpha J(\alpha)$

Needed to get $\nabla_\alpha J(\alpha)$

- **Complicated**
- approximate the derivatives of the model with finite differences

4D-Var: 4D Variational Data Assimilation

Cost function in 4D-Var

$$J(\alpha) = \sum_{i=1}^{n} \underbrace{(\mathcal{H}_i(x_i) - y_i)^T}_{Distances} R_i^{-1} \underbrace{(\mathcal{H}_i(x_i) - y_i)}_{Distances} + \underbrace{(\alpha - \alpha_b)^T B^{-1} (\alpha - \alpha_b)}_{Background\ \ term}$$

minimization with constraints:
$x_i = \mathcal{M}_i(x_{i-1}, \alpha)$

B background error covariance matrix
R_i observation error covariance matrix

To minimize J over α, we need: $\nabla_\alpha J(\alpha)$

Needed to get $\nabla_\alpha J(\alpha)$

- **Complicated**
- **Time consuming**

4D-Var: 4D Variational Data Assimilation

Cost function in 4D-Var

$$J(\alpha) = \sum_{i=1}^{n} \underbrace{(\mathcal{H}_i(x_i) - y_i)^T}_{Distances} R_i^{-1} \underbrace{(\mathcal{H}_i(x_i) - y_i)}_{Distances} + \underbrace{(\alpha - \alpha_b)^T B^{-1} (\alpha - \alpha_b)}_{Background \quad term}$$

minimization with constraints:
$x_i = \mathcal{M}_i(x_{i-1}, \alpha)$

B background error covariance matrix
R_i observation error covariance matrix

To minimize J over α, we need:
$\nabla_\alpha J(\alpha)$

Solution

POD Model-reduced 4D-Var
(Vermeulen and Heemink, MWR, 2006)

POD Model Reduced 4D-Var

Incremental cost function in 4D-Var

$$J(\delta\alpha) = \sum_{i=1}^{n} (\mathbf{H}_i(\delta x_i, \delta\alpha) + d_i)^T R_i^{-1} (\mathbf{H}_i(\delta x_i, \delta\alpha) + d_i) + \delta\alpha^T B^{-1} \delta\alpha$$

minimization with constraints

$$\delta x_i = \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial x_{i-1}} \delta x_{i-1} + \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial \alpha} \delta\alpha$$

POD Model Reduced 4D-Var

Incremental cost function in 4D-Var

$$J(\delta\alpha) = \sum_{i=1}^{n} (\mathbf{H}_i(\delta x_i, \delta\alpha) + d_i)^T R_i^{-1} (\mathbf{H}_i(\delta x_i, \delta\alpha) + d_i) + \delta\alpha^T B^{-1} \delta\alpha$$

minimization with constraints

$$\delta x_i = \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial x_{i-1}} \delta x_{i-1} + \boxed{\frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial \alpha}} \delta\alpha$$

With a small number of parameters, the finite differences method is feasible

POD Model Reduced 4D-Var

Incremental cost function in 4D-Var

$$J(\delta\alpha) = \sum_{i=1}^{n} (\mathbf{H}_i(\delta x_i, \delta\alpha) + d_i)^T R_i^{-1} (\mathbf{H}_i(\delta x_i, \delta\alpha) + d_i) + \delta\alpha^T B^{-1} \delta\alpha$$

minimization with constraints

$$\delta x_i = \boxed{\frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial x_{i-1}}} \delta x_{i-1} + \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial \alpha} \delta\alpha$$

If the size of the state x is huge, the finite differences would be too expensive

POD Model Reduced 4D-Var

Incremental equation

$$\delta x_i = \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial x_{i-1}} \delta x_{i-1} + \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial \alpha} \delta \alpha$$

Project the increments into smaller subspace

$$P^T \delta x_i = P^T \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial x_{i-1}} \delta x_{i-1} + P^T \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial \alpha} \delta \alpha$$

POD Model Reduced 4D-Var

Incremental equation

$$\delta x_i = \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial x_{i-1}} \delta x_{i-1} + \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial \alpha} \delta \alpha$$

Project the increments into smaller subspace

$$P^T \delta x_i = P^T \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial x_{i-1}} P\, P^T \delta x_{i-1} + P^T \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial \alpha} \delta \alpha$$

POD Model Reduced 4D-Var

Incremental equation

$$\delta x_i = \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial x_{i-1}} \delta x_{i-1} + \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial \alpha} \delta \alpha$$

Project the increments into smaller subspace

$$\underbrace{P^T \delta x_i}_{\delta z_i} = P^T \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial x_{i-1}} P \underbrace{P^T \delta x_{i-1}}_{\delta z_{i-1}} + P^T \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial \alpha} \delta \alpha$$

POD Model Reduced 4D-Var

Incremental equation

$$\delta x_i = \frac{\partial \mathcal{M}_i(x^b_{i-1}, \alpha^b)}{\partial x_{i-1}} \delta x_{i-1} + \frac{\partial \mathcal{M}_i(x^b_{i-1}, \alpha^b)}{\partial \alpha} \delta \alpha$$

Project the increments into smaller subspace

$$\underbrace{P^T \delta x_i}_{\delta z_i} = P^T \frac{\partial \mathcal{M}_i(x^b_{i-1}, \alpha^b)}{\partial x_{i-1}} P \underbrace{P^T \delta x_{i-1}}_{\delta z_{i-1}} + P^T \frac{\partial \mathcal{M}_i(x^b_{i-1}, \alpha^b)}{\partial \alpha} \delta \alpha$$

$\delta z_i = P^T \delta x_i$ increment of the state in the reduced space

POD Model Reduced 4D-Var

Incremental equation

$$\delta x_i = \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial x_{i-1}} \delta x_{i-1} + \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial \alpha} \delta \alpha$$

Project the increments into smaller subspace

$$\underbrace{P^T \delta x_i}_{\delta z_i} = P^T \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial x_{i-1}} P \underbrace{P^T \delta x_{i-1}}_{\delta z_{i-1}} + P^T \frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial \alpha} \delta \alpha$$

$\delta z_i = P^T \delta x_i$ increment of the state in the reduced space

$$\frac{\partial \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\partial x_{i-1}} p \simeq \frac{\mathcal{M}_i(x_{i-1}^b + \epsilon p, \alpha^b) - \mathcal{M}_i(x_{i-1}^b, \alpha^b)}{\epsilon}$$

directional derivative approximation

How to get matrix P?

We want to project into a smaller subspace, such that the most important dynamics of the system are kept

How to get matrix P?

We want to project into a smaller subspace, such that the most important dynamics of the system are kept

STEP 1: Generate ensemble of perturbed model simulations

$$\alpha^b + \Delta\alpha_1 \rightarrow x_1^{\Delta 1}, \quad x_2^{\Delta 1}, \quad \dots \quad x_n^{\Delta 1}$$
$$\alpha^b + \Delta\alpha_2 \rightarrow x_1^{\Delta 2}, \quad x_2^{\Delta 2}, \quad \dots \quad x_n^{\Delta 2}$$

How to get matrix P?

We want to project into a smaller subspace, such that the most important dynamics of the system are kept

Step 1: Generate ensemble of perturbed model simulations

$$\begin{array}{lcl} \alpha^b + \Delta\alpha_1 & \rightarrow & x_1^{\Delta 1}, \; x_2^{\Delta 1}, \; \dots \; x_n^{\Delta 1} \\ \alpha^b + \Delta\alpha_2 & \rightarrow & x_1^{\Delta 2}, \; x_2^{\Delta 2}, \; \dots \; x_n^{\Delta 2} \end{array}$$

Step 2: Create a covariance matrix

$$\mathbf{C_X} = \mathbf{\Delta X}\,\mathbf{\Delta X}^{\mathbf{T}}/(\mathbf{n}-\mathbf{1})$$

How to get matrix P?

We want to project into a smaller subspace, such that the most important dynamics of the system are kept

STEP 1: Generate ensemble of perturbed model simulations

$$\begin{array}{lcllll} \alpha^b + \Delta\alpha_1 & \rightarrow & x_1^{\Delta 1}, & x_2^{\Delta 1}, & \ldots & x_n^{\Delta 1} \\ \alpha^b + \Delta\alpha_2 & \rightarrow & x_1^{\Delta 2}, & x_2^{\Delta 2}, & \ldots & x_n^{\Delta 2} \end{array}$$

STEP 2: Create a covariance matrix

$$\mathbf{C_X} = \mathbf{\Delta X}\ \mathbf{\Delta X}^{\mathbf{T}} / (\mathbf{n} - \mathbf{1})$$

STEP 3: Decompose $\mathbf{C_X}$ with eigenvalue decomposition

$$\mathbf{\Delta X}\ \mathbf{\Delta X}^{\mathbf{T}} / (\mathbf{n} - \mathbf{1}) = \mathbf{P D P}^{\mathbf{T}}$$

P - eigenvectors, **D** - diagonal matrix with eigenvalues

How to get matrix P?

Step 1: Generate ensemble of perturbed model simulations

$$\alpha^b + \Delta\alpha_1 \;\rightarrow\; x_1^{\Delta 1},\; x_2^{\Delta 1},\; \ldots\; x_n^{\Delta 1}$$
$$\alpha^b + \Delta\alpha_2 \;\rightarrow\; x_1^{\Delta 2},\; x_2^{\Delta 2},\; \ldots\; x_n^{\Delta 2}$$

Step 2: Create a covariance matrix

$$\mathbf{C_X} = \mathbf{\Delta X}\ \mathbf{\Delta X}^{\mathbf{T}}/(\mathbf{n}-\mathbf{1})$$

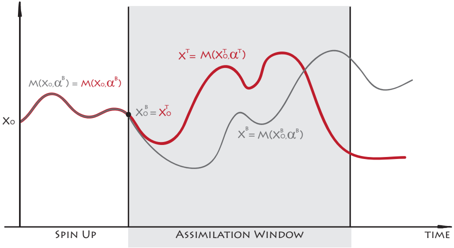
Step 3: Decompose $\mathbf{C_X}$ with eigenvalue decomposition

$$\mathbf{\Delta X}\ \mathbf{\Delta X}^{\mathbf{T}}/(\mathbf{n}-\mathbf{1}) = \mathbf{P}\mathbf{D}\mathbf{P}^{\mathbf{T}}$$

P - eigenvectors, **D** - diagonal matrix with eigenvalues

$$\mathbf{P}^{\mathbf{T}}\mathbf{\Delta X}\ (\mathbf{P}^{\mathbf{T}}\mathbf{\Delta X})^{\mathbf{T}}/(\mathbf{n}-\mathbf{1}) = \mathbf{\Delta Z}\ \mathbf{\Delta Z}^{\mathbf{T}}/(\mathbf{n}-\mathbf{1}) = \mathbf{D}$$

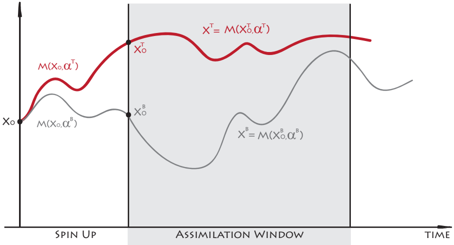
Twin Experiment Setup



Perfect Initial Condition

The spin up starts with

- prior par. for prior sol.
- prior par. for true sol.



Perturbed Initial Condition

The spin up starts with

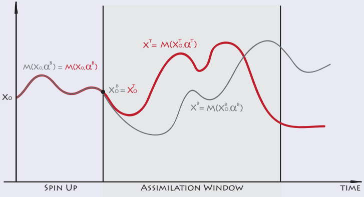
- prior par. for prior sol.
- true par. for true sol.

Experiment 1 - Estimate: Parameters

Experiment Setup - **Estimate Parameters**

Perfect Initial Condition

Spin up time	6 years
Assimilation time	4 years



Parameters Setup

	f	**g**	**r**
Prior	0.50	0.07	0.07
Truth	**0.90**	**0.11**	**0.11**
	80%	57%	57 %

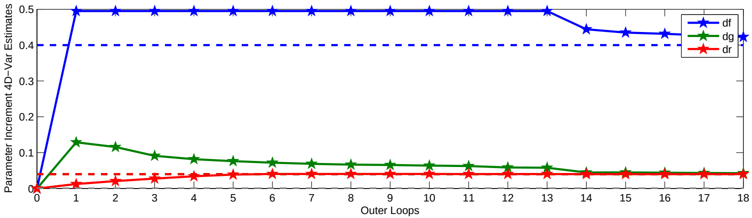
Observation Setup

Observe

What	surface Phyto
When	every 4 days
Error	30 %

Experiment 1 - Estimate: Parameters

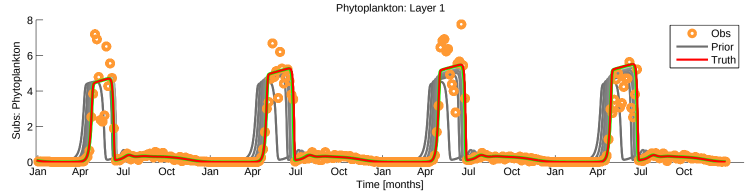
	Parameter Estimation			Cost Function	
	f	**g**	**r**		
Prior	0.5	0.07	0.07	Prior	3.91e+6
Truth	0.9	0.11	0.11	$\frac{1}{2}$Obs	364
4DVar	**0.9223**	**0.1127**	**0.1099**	4DVar	**340.45**

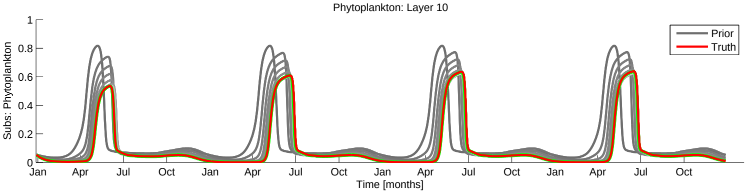


Parameter Increments shown within outer loops

Experiment 1 - Estimate: Parameters

Phytoplankton within time shown at the surface layer and at the 10th layer of the water column.





Grey - Outer Loops, **Green - 4DVar**

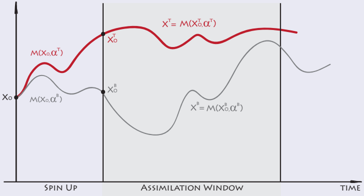
Red - Truth, **Orange - Observations**

Experiment 2 - Parameters & Initial Condition

Experiment Setup - **Estimate Parameters & Initial Condition**

Perturbed Initial Condition

Spin up time	455 days
Assimilation time	5 years



Parameters Setup

	f	g	r
Prior	0.50	0.07	0.07
Truth	**0.90**	**0.11**	**0.11**
	80%	57%	57 %

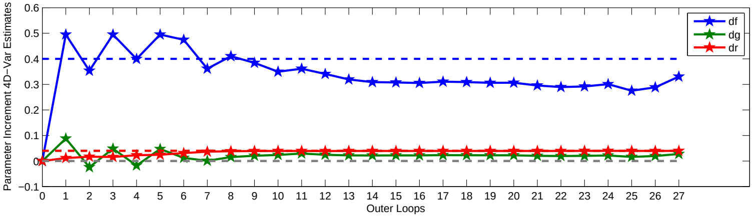
Observation Setup

Observe

What	surface Phyto
When	every 4 days
Error	30 %

Experiment 2 - Parameters & Initial Condition

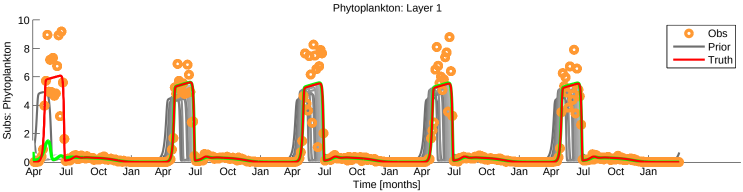
	Parameter Estimation			Cost Function	
	f	**g**	**r**		
Prior	0.5	0.07	0.07	Prior	5.49e+6
Truth	0.9	0.11	0.11	$\frac{1}{2}$Obs	433
4DVar	**0.8304**	**0.0976**	**0.1098**	4DVar	**611.02**

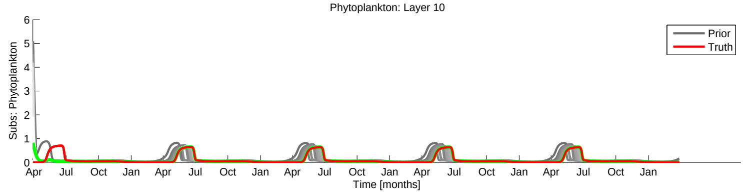


Parameter Increments shown within outer loops

Experiment 2 - Parameters & Initial Condition

Phytoplankton within time shown at the surface layer and at the 10th layer of the water column.



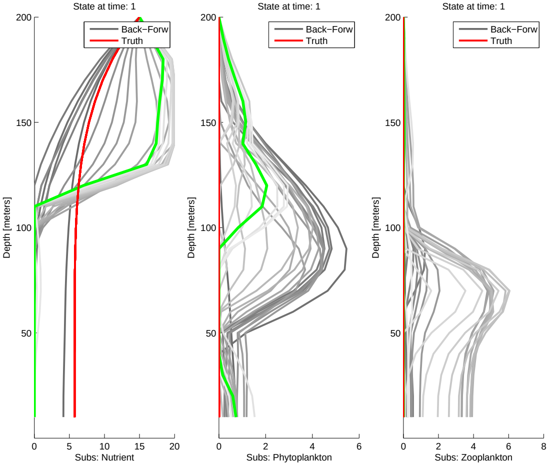


Grey - Outer Loops, **Green - 4DVar**

Red - Truth, **Orange - Observations**

Experiment 2 - Parameters & Initial Condition

Initial Condition estimation

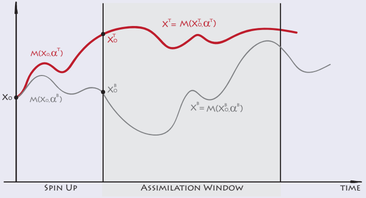


Grey - Outer Loops
Green - 4DVar
Red - Truth

Experiment 3 - Unrealistic Setup

Experiment Setup - **Estimate Parameters & Initial Condition**

Perturbed Initial Condition	
Spin up time	730 days
Assimilation time	5 years



Parameters Setup

	f	**g**	**r**
Prior	0.50	0.07	0.07
Truth	**0.5444**	**0.0766**	**0.0766**
	9%	9%	9 %

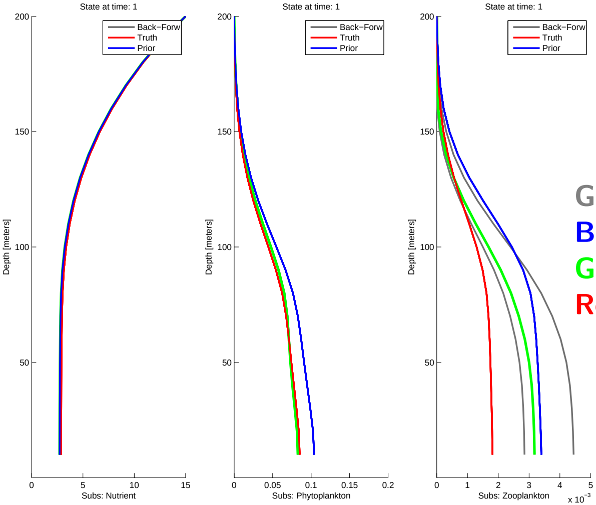
Observation Setup

Observe

What	surf Phyto
When	every 4 days
Error	**1 %**

Experiment 3 - Unrealistic Setup

Initial Condition estimation



Grey - Outer Loops
Blue - Prior
Green - 4DVar
Red - Truth

1. Outline
2. 1D Ecological Model
3. Model Reduced 4D-Var
4. Results
5. **Conclusions**

Conclusions

Model Reduced 4D-Var for 1D Ecosystem Model

- the parameter estimation - works fine when a good quality initial condition is available
- the combined parameter and initial condition estimation - needs an improvement

Conclusions

Model Reduced 4D-Var for 1D Ecosystem Model

- the parameter estimation - works fine when a good quality initial condition is available
- the combined parameter and initial condition estimation - needs an improvement

- Model Reduced 4D-Var is a potential tool in ecosystem applications

Conclusions

Model Reduced 4D-Var for 1D Ecosystem Model

- the parameter estimation - works fine when a good quality initial condition is available
- the combined parameter and initial condition estimation - needs an improvement

- Model Reduced 4D-Var is a potential tool in ecosystem applications

Future Work

- improve the combined parameter and initial condition approach
- apply the method to a large ecosystem model

Thank you for your attention!

Questions?