Data Assimilation
- L04. Computational Complexity-

Shunji Kotsuki

Environmental Prediction Science Laboratory
Center for Environmental Remote Sensing (CEReS), Chiba University
(shunji.kotsuki@chiba-u.jp)

Data Assimilation
- L04. Computational Complexity -

Shunji Kotsuki

Environmental Prediction Science Laboratory
Center for Environmental Remote Sensing (CEReS), Chiba University
(shunji.kotsuki@chiba-u.jp)

Goal

- To understand computational complexity (CC)

- To understand CC of the LETKF

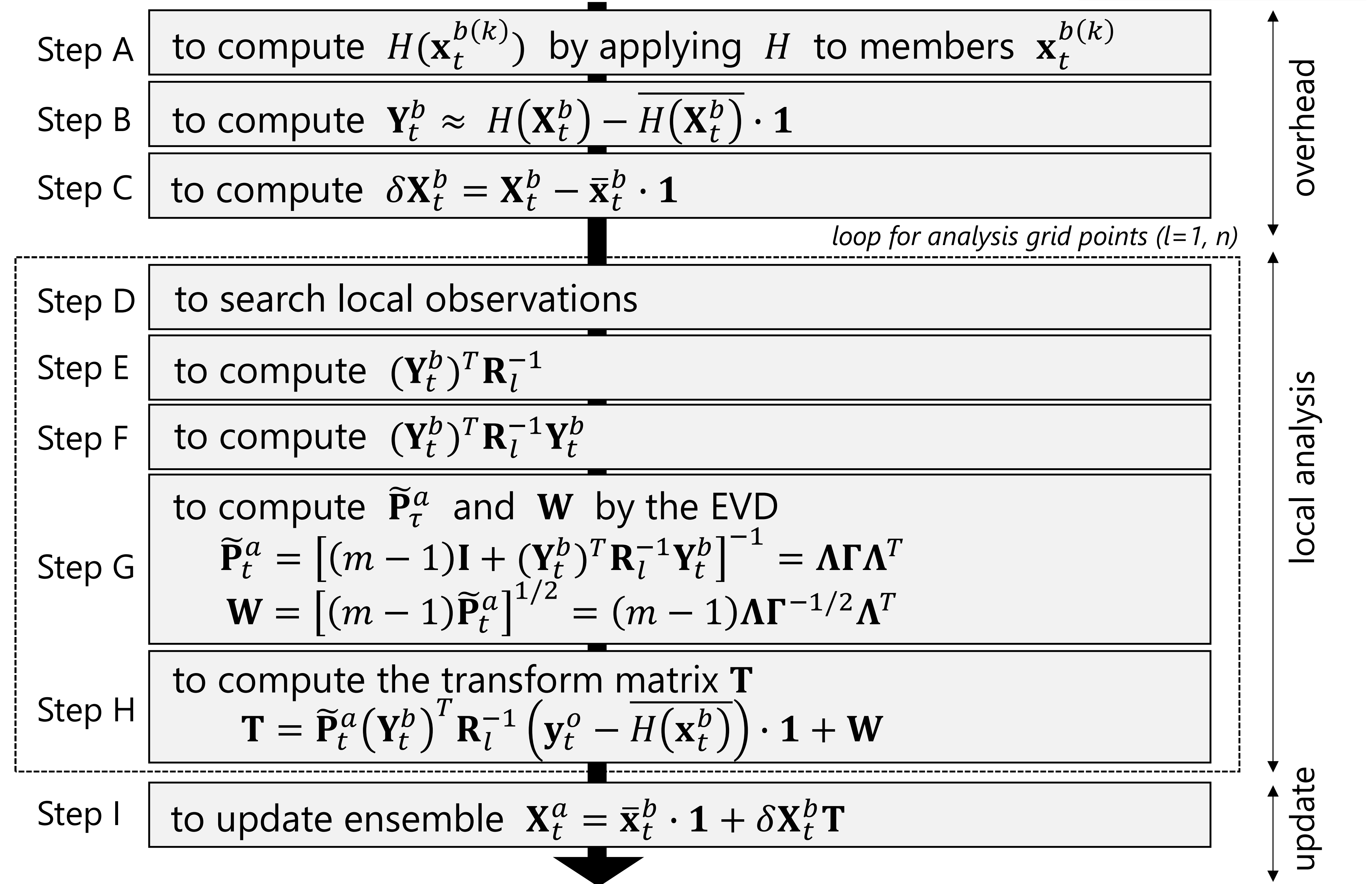
Computational Complexity

Note: Different notations

	Hunt et al. (2007)	This Slide
Ensemble size	k	m
# of model grid points	$m_{[g]}$	n
# of observations	$l_{[g]}$	p
# of local observations	l	p_L

NOTE: The notation of this slide is different from Hunt et al. (2007)
so as to be consistent with other slides on data assimilation

Miyoshi’s implementation



Step A: to compute $H(\mathbf{x}_t^{b(k)})$ by applying H to members $\mathbf{x}_t^{b(k)}$

Step B: to compute $\mathbf{Y}_t^b \approx H\left(\mathbf{X}_t^b\right) - \overline{H\left(\mathbf{X}_t^b\right)} \cdot \mathbf{1}$

Step C: to compute $\delta\mathbf{X}_t^b = \mathbf{X}_t^b - \bar{\mathbf{x}}_t^b \cdot \mathbf{1}$

loop for analysis grid points (l=1, n)

Step D: to search local observations

Step E: to compute $(\mathbf{Y}_t^b)^T\mathbf{R}_l^{-1}$

Step F: to compute $(\mathbf{Y}_t^b)^T\mathbf{R}_l^{-1}\mathbf{Y}_t^b$

Step G: to compute $\widetilde{\mathbf{P}}_\tau^a$ and $\mathbf{W}$ by the EVD

$$\widetilde{\mathbf{P}}_t^a = \left[(m-1)\mathbf{I} + (\mathbf{Y}_t^b)^T\mathbf{R}_l^{-1}\mathbf{Y}_t^b\right]^{-1} = \mathbf{\Lambda}\mathbf{\Gamma}\mathbf{\Lambda}^T$$

$$\mathbf{W} = \left[(m-1)\widetilde{\mathbf{P}}_t^a\right]^{1/2} = (m-1)\mathbf{\Lambda}\mathbf{\Gamma}^{-1/2}\mathbf{\Lambda}^T$$

Step H: to compute the transform matrix $\mathbf{T}$

$$\mathbf{T} = \widetilde{\mathbf{P}}_t^a\left(\mathbf{Y}_t^b\right)^T\mathbf{R}_l^{-1}\left(\mathbf{y}_t^o - \overline{H\left(\mathbf{x}_t^b\right)}\right)\cdot\mathbf{1} + \mathbf{W}$$

Step I: to update ensemble $\mathbf{X}_t^a = \bar{\mathbf{x}}_t^b \cdot \mathbf{1} + \delta\mathbf{X}_t^b\mathbf{T}$

Overhead

Step A | to compute $H(\mathbf{x}_t^{b(k)})$ by applying H to members $\mathbf{x}_t^{b(k)}$

m applications of H

Step B | to compute $\mathbf{Y}_t^b \approx H\left(\mathbf{X}_t^b\right) - \overline{H\left(\mathbf{X}_t^b\right)} \cdot \mathbf{1}$

$2mp$

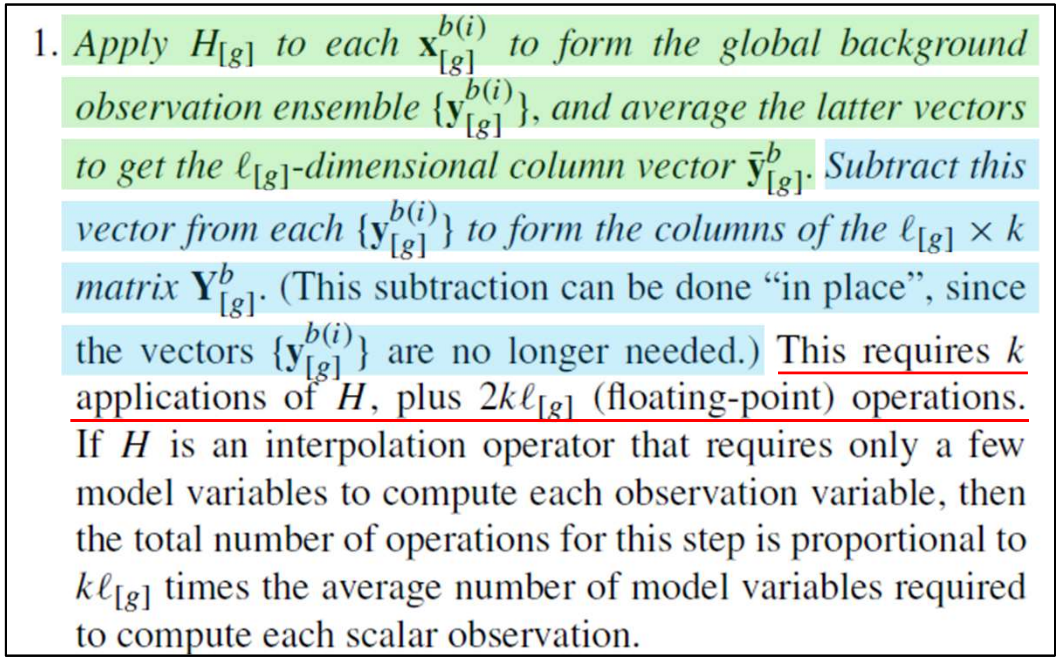
mp-times computations for $\overline{H\left(\mathbf{X}_t^b\right)}$ and $\mathbf{Y}_t^b$, respectively

Step C | to compute $\delta\mathbf{X}_t^b = \mathbf{X}_t^b - \bar{\mathbf{x}}_t^b \cdot \mathbf{1}$

$2mn$

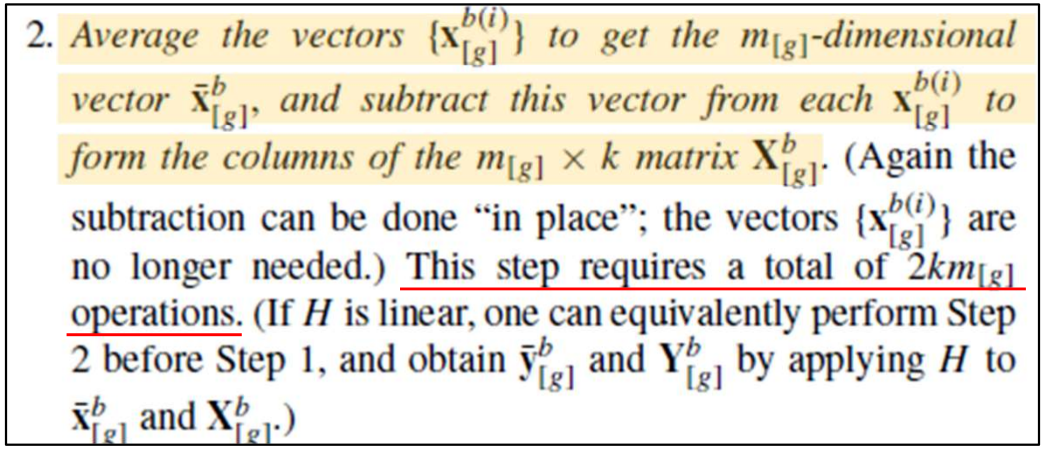
mn-times computations for $\bar{\mathbf{x}}_t^b$and $\delta\mathbf{X}_t^b$, respectively

Hunt et al. (2007) Step A Step B



1. *Apply $H_{[g]}$ to each $\mathbf{x}_{[g]}^{b(i)}$ to form the global background observation ensemble $\{\mathbf{y}_{[g]}^{b(i)}\}$, and average the latter vectors to get the $\ell_{[g]}$-dimensional column vector $\bar{\mathbf{y}}_{[g]}^{b}$. Subtract this vector from each $\{\mathbf{y}_{[g]}^{b(i)}\}$ to form the columns of the $\ell_{[g]} \times k$ matrix $\mathbf{Y}_{[g]}^{b}$.* (This subtraction can be done "in place", since the vectors $\{\mathbf{y}_{[g]}^{b(i)}\}$ are no longer needed.) This requires k applications of H, plus $2k\ell_{[g]}$ (floating-point) operations. If H is an interpolation operator that requires only a few model variables to compute each observation variable, then the total number of operations for this step is proportional to $k\ell_{[g]}$ times the average number of model variables required to compute each scalar observation.

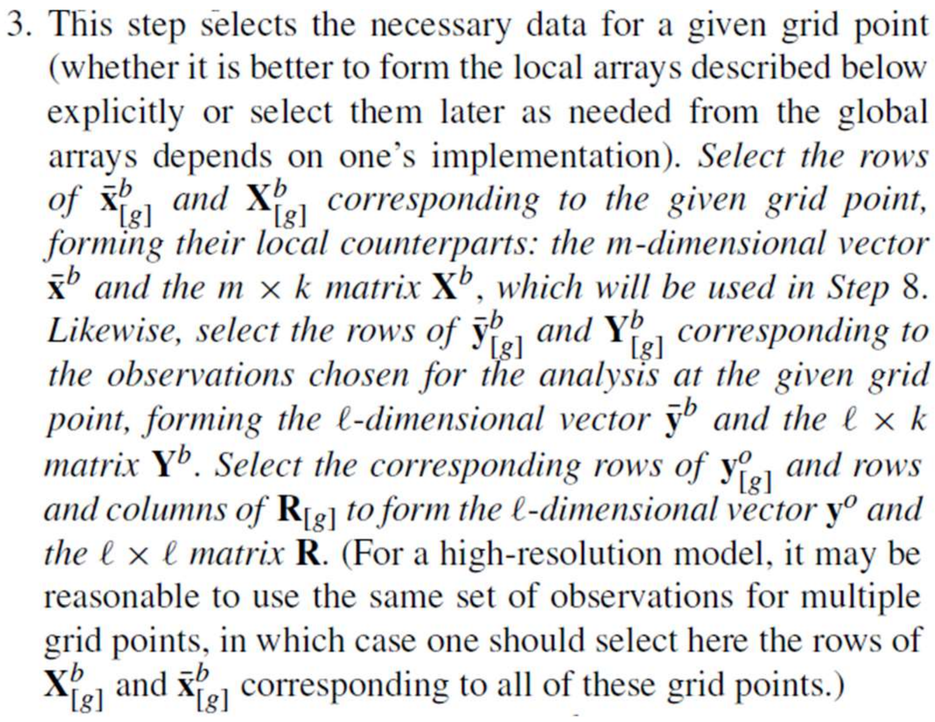
Step C



2. *Average the vectors $\{\mathbf{x}_{[g]}^{b(i)}\}$ to get the $m_{[g]}$-dimensional vector $\bar{\mathbf{x}}_{[g]}^{b}$, and subtract this vector from each $\mathbf{x}_{[g]}^{b(i)}$ to form the columns of the $m_{[g]} \times k$ matrix $\mathbf{X}_{[g]}^{b}$.* (Again the subtraction can be done "in place"; the vectors $\{\mathbf{x}_{[g]}^{b(i)}\}$ are no longer needed.) This step requires a total of $2km_{[g]}$ operations. (If H is linear, one can equivalently perform Step 2 before Step 1, and obtain $\bar{\mathbf{y}}_{[g]}^{b}$ and $\mathbf{Y}_{[g]}^{b}$ by applying H to $\bar{\mathbf{x}}_{[g]}^{b}$ and $\mathbf{X}_{[g]}^{b}$.)

Local Analysis (1)

Step D to search local observations

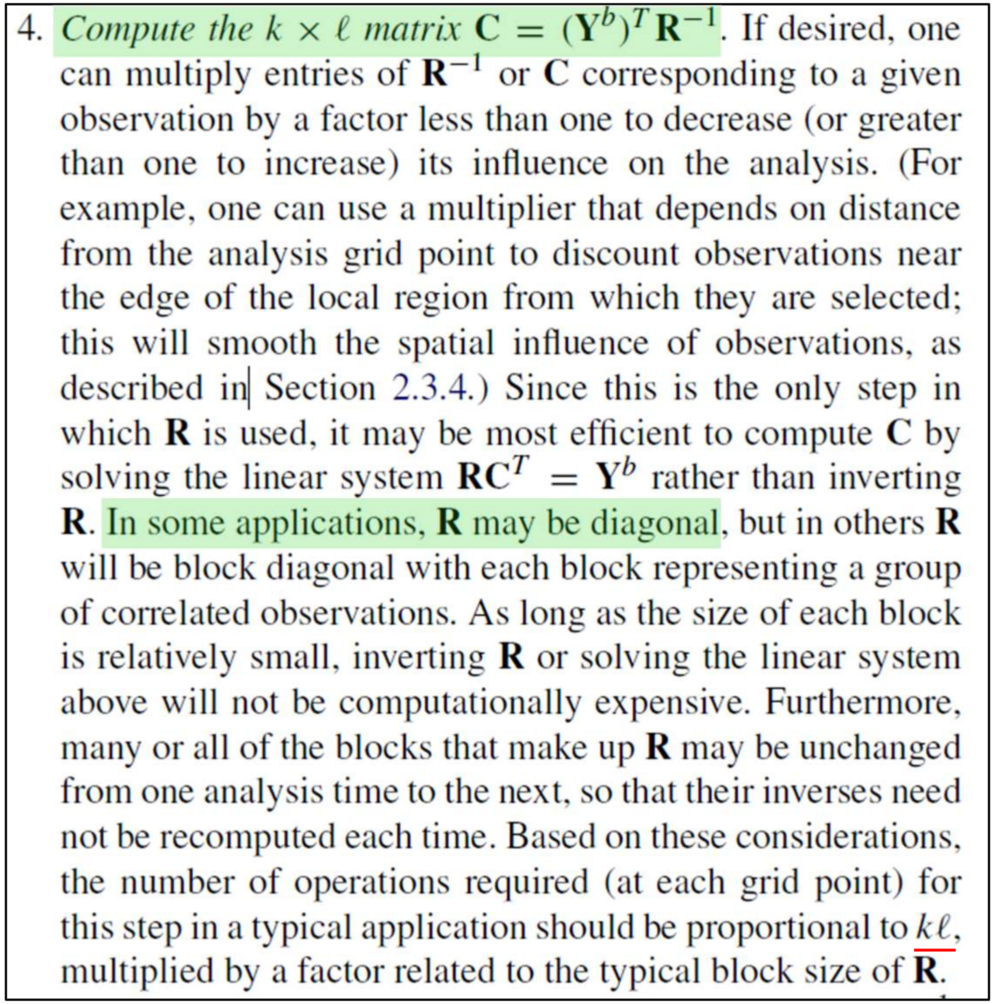


3. This step selects the necessary data for a given grid point (whether it is better to form the local arrays described below explicitly or select them later as needed from the global arrays depends on one's implementation). *Select the rows of $\bar{\mathbf{x}}^b_{[g]}$ and $\mathbf{X}^b_{[g]}$ corresponding to the given grid point, forming their local counterparts: the m-dimensional vector $\bar{\mathbf{x}}^b$ and the $m \times k$ matrix $\mathbf{X}^b$, which will be used in Step 8. Likewise, select the rows of $\bar{\mathbf{y}}^b_{[g]}$ and $\mathbf{Y}^b_{[g]}$ corresponding to the observations chosen for the analysis at the given grid point, forming the ℓ-dimensional vector $\bar{\mathbf{y}}^b$ and the $\ell \times k$ matrix $\mathbf{Y}^b$. Select the corresponding rows of $\mathbf{y}^o_{[g]}$ and rows and columns of $\mathbf{R}_{[g]}$ to form the ℓ-dimensional vector $\mathbf{y}^o$ and the $\ell \times \ell$ matrix* $\mathbf{R}$. (For a high-resolution model, it may be reasonable to use the same set of observations for multiple grid points, in which case one should select here the rows of $\mathbf{X}^b_{[g]}$ and $\bar{\mathbf{x}}^b_{[g]}$ corresponding to all of these grid points.)

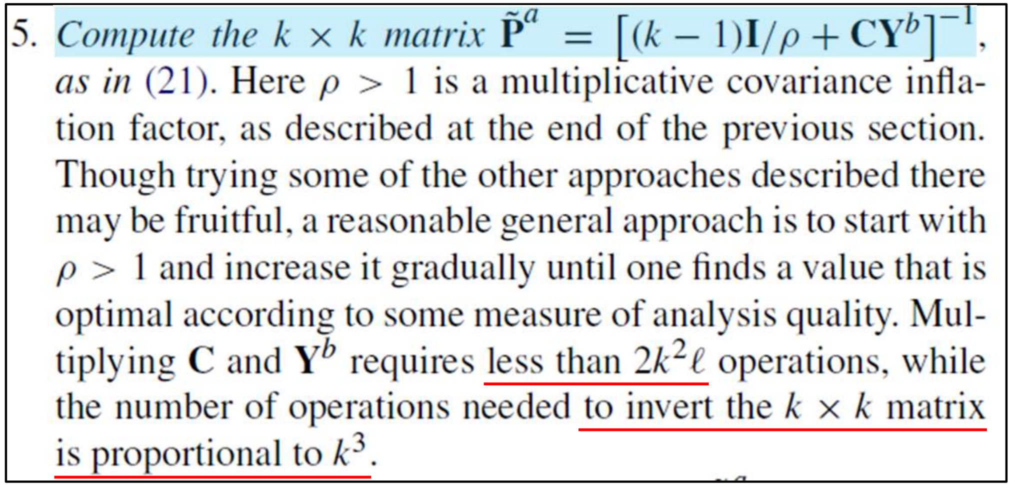
Local Analysis (2)

Step E | to compute $(\mathbf{Y}_t^b)^T\mathbf{R}_l^{-1}$ | mp_L

Step F | to compute $(\mathbf{Y}_t^b)^T\mathbf{R}_l^{-1}\mathbf{Y}_t^b$ | $\leq 2m^2p_L$



4. *Compute the $k \times \ell$ matrix* $\mathbf{C} = (\mathbf{Y}^b)^T\mathbf{R}^{-1}$. If desired, one can multiply entries of $\mathbf{R}^{-1}$ or $\mathbf{C}$ corresponding to a given observation by a factor less than one to decrease (or greater than one to increase) its influence on the analysis. (For example, one can use a multiplier that depends on distance from the analysis grid point to discount observations near the edge of the local region from which they are selected; this will smooth the spatial influence of observations, as described in Section 2.3.4.) Since this is the only step in which $\mathbf{R}$ is used, it may be most efficient to compute $\mathbf{C}$ by solving the linear system $\mathbf{R}\mathbf{C}^T = \mathbf{Y}^b$ rather than inverting $\mathbf{R}$. In some applications, $\mathbf{R}$ may be diagonal, but in others $\mathbf{R}$ will be block diagonal with each block representing a group of correlated observations. As long as the size of each block is relatively small, inverting $\mathbf{R}$ or solving the linear system above will not be computationally expensive. Furthermore, many or all of the blocks that make up $\mathbf{R}$ may be unchanged from one analysis time to the next, so that their inverses need not be recomputed each time. Based on these considerations, the number of operations required (at each grid point) for this step in a typical application should be proportional to $k\ell$, multiplied by a factor related to the typical block size of $\mathbf{R}$.



5. *Compute the $k \times k$ matrix* $\tilde{\mathbf{P}}^a = [(k-1)\mathbf{I}/\rho + \mathbf{C}\mathbf{Y}^b]^{-1}$, *as in* (21). Here $\rho > 1$ is a multiplicative covariance inflation factor, as described at the end of the previous section. Though trying some of the other approaches described there may be fruitful, a reasonable general approach is to start with $\rho > 1$ and increase it gradually until one finds a value that is optimal according to some measure of analysis quality. Multiplying $\mathbf{C}$ and $\mathbf{Y}^b$ requires less than $2k^2\ell$ operations, while the number of operations needed to invert the $k \times k$ matrix is proportional to k^3.

Step E
Step F

Local Analysis (2)

Step G

to compute $\widetilde{\mathbf{P}}_\tau^a$ and $\mathbf{W}$ by the EVD $O(m^3)$

$$\widetilde{\mathbf{P}}_t^a = \left[(m-1)\mathbf{I} + (\mathbf{Y}_t^b)^T \mathbf{R}_l^{-1} \mathbf{Y}_t^b\right]^{-1} = \mathbf{\Lambda\Gamma\Lambda}^T$$

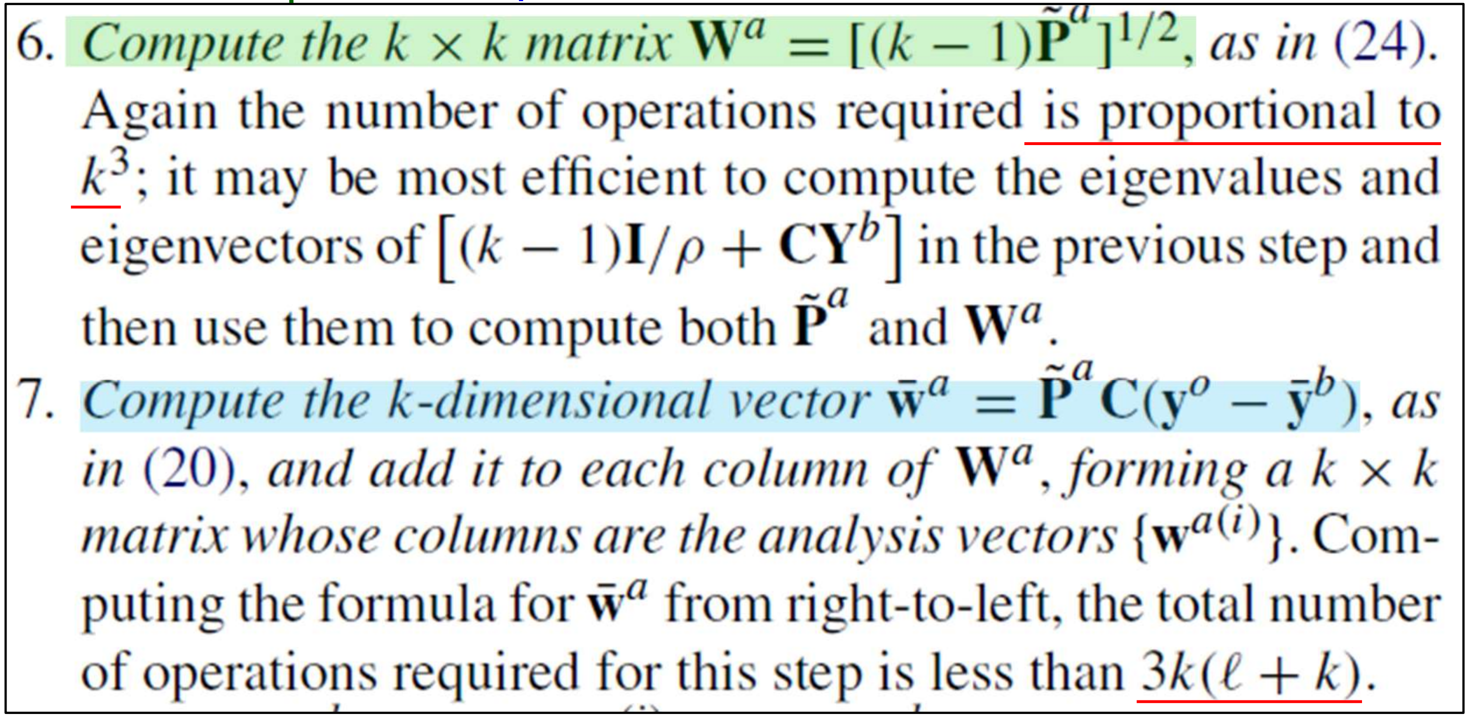
$$\mathbf{W} = \left[(m-1)\widetilde{\mathbf{P}}_t^a\right]^{1/2} = (m-1)\mathbf{\Lambda\Gamma}^{-1/2}\mathbf{\Lambda}^T \qquad O(m^3)$$

Step H

to compute the transform matrix $\mathbf{T}$

$$\mathbf{T} = \widetilde{\mathbf{P}}_t^a \left(\mathbf{Y}_t^b\right)^T \mathbf{R}_l^{-1} \left(\mathbf{y}_t^o - \overline{H\left(\mathbf{x}_t^b\right)}\right) \cdot \mathbf{1} + \mathbf{W} \qquad \leq 3m(m + p_L)$$

W of Step G Step H



6. *Compute the* $k \times k$ *matrix* $\mathbf{W}^a = [(k-1)\tilde{\mathbf{P}}^a]^{1/2}$, *as in* (24). Again the number of operations required is proportional to k^3; it may be most efficient to compute the eigenvalues and eigenvectors of $\left[(k-1)\mathbf{I}/\rho + \mathbf{C}\mathbf{Y}^b\right]$ in the previous step and then use them to compute both $\tilde{\mathbf{P}}^a$ and $\mathbf{W}^a$.
7. *Compute the* k*-dimensional vector* $\bar{\mathbf{w}}^a = \tilde{\mathbf{P}}^a \mathbf{C}(\mathbf{y}^o - \bar{\mathbf{y}}^b)$, *as in* (20), *and add it to each column of* $\mathbf{W}^a$, *forming a* $k \times k$ *matrix whose columns are the analysis vectors* $\{\mathbf{w}^{a(i)}\}$. Computing the formula for $\bar{\mathbf{w}}^a$ from right-to-left, the total number of operations required for this step is less than $3k(\ell + k)$.

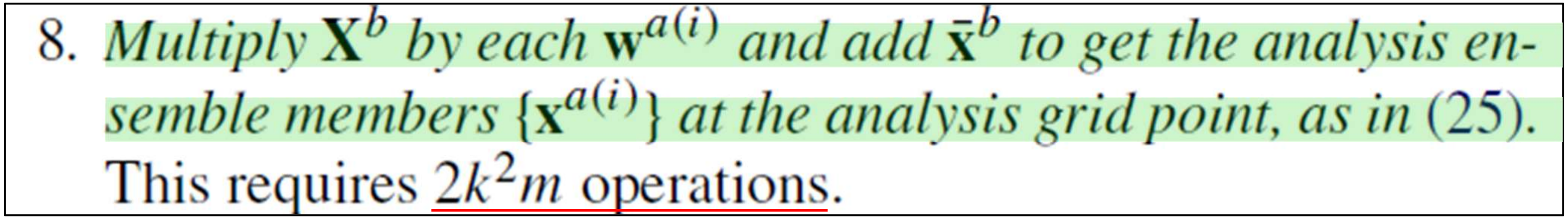
Analysis Update

Step I

to update ensemble $\mathbf{X}_t^a = \bar{\mathbf{x}}_t^b \cdot \mathbf{1} + \delta\mathbf{X}_t^b\mathbf{T}$

$2m^2n$

Step I



8. *Multiply $\mathbf{X}^b$ by each $\mathbf{w}^{a(i)}$ and add $\bar{\mathbf{x}}^b$ to get the analysis ensemble members $\{\mathbf{x}^{a(i)}\}$ at the analysis grid point, as in (25).* This requires $2k^2m$ operations.

Summary

Step	Computation	Hunt et al. (2007)	Kotsuki et al. (2022)
A	to compute $H(\mathbf{x}_t^{b(k)})$	m applications of H	m applications of H
B	to compute $\mathbf{Y}_t^b$	$2mp$	$2mp$
C	to compute $\delta\mathbf{X}_t^b$	$2mn$	$2mn$
D	to search local obs	problem dependent	problem dependent
E	to compute $(\mathbf{Y}_t^b)^T\mathbf{R}_l^{-1}$	mp_L	mp_L
F	to compute $(\mathbf{Y}_t^f)^T\mathbf{R}_l^{-1}\mathbf{Y}_t^f$	$\leq 2m^2p_L$	$\leq 2m^2p_L$
G	to compute $\widetilde{\mathbf{P}}_\tau^a$ and $\mathbf{W}$	$O(m^3)$	$O(m^3)$
H	to compute the trans. mtx. $\mathbf{T}$	$\leq 3m(m+p_L)$	$O\big(m(m+p_L)\big)$ (*1)
I	to update ensemble $\mathbf{X}_t^a$	$2m^2n$	$\leq 2m^2n$ (*2)

*1: to be updated

*2: based on the Strassen algorithm

Thank you for your attention!

Presented by Shunji Kotsuki
(shunji.kotsuki@chiba-u.jp)

Further information is available at
https://kotsuki-lab.com/

A World Beyond Predictions

Environmental
Prediction
Science
Laboratory