

ENVIROMIS-2018



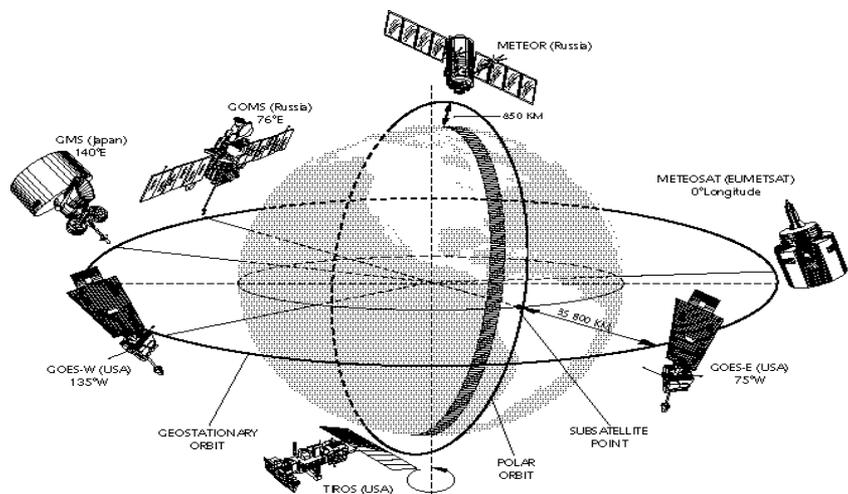
**Применение ансамблевого фильтра Калмана
в задаче усвоения данных об окружающей среде**

Климова Екатерина Георгиевна
ИИТ СО РАН

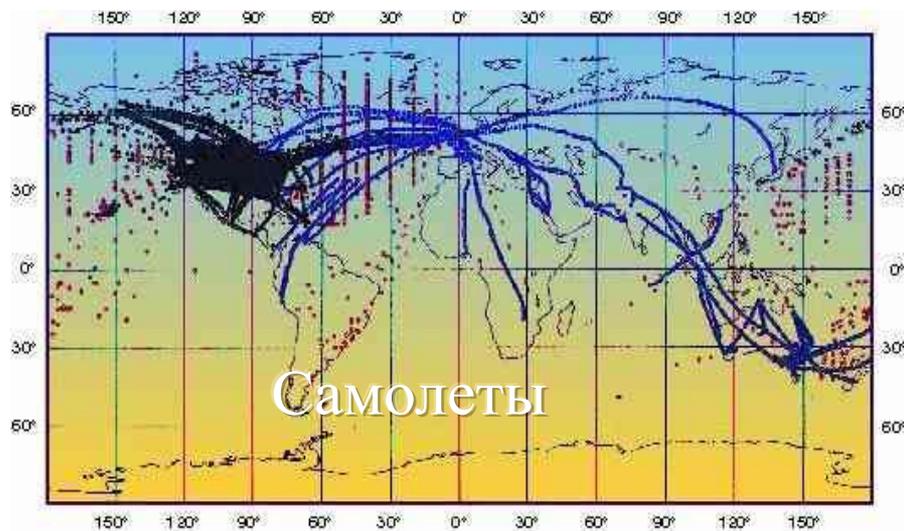
План доклада

1. Введение
2. Задача усвоения данных при моделировании процессов в окружающей среде
3. Ансамблевый фильтр Калмана
4. Задача оптимального сглаживания. Ансамблевое сглаживание (EnKS – ensemble Kalman smoother)
5. Оценка параметров в процедуре усвоения данных
6. Особенности задачи оценки потоков пассивных газовых составляющих в процедуре усвоения данных.
7. Заключительные выводы.

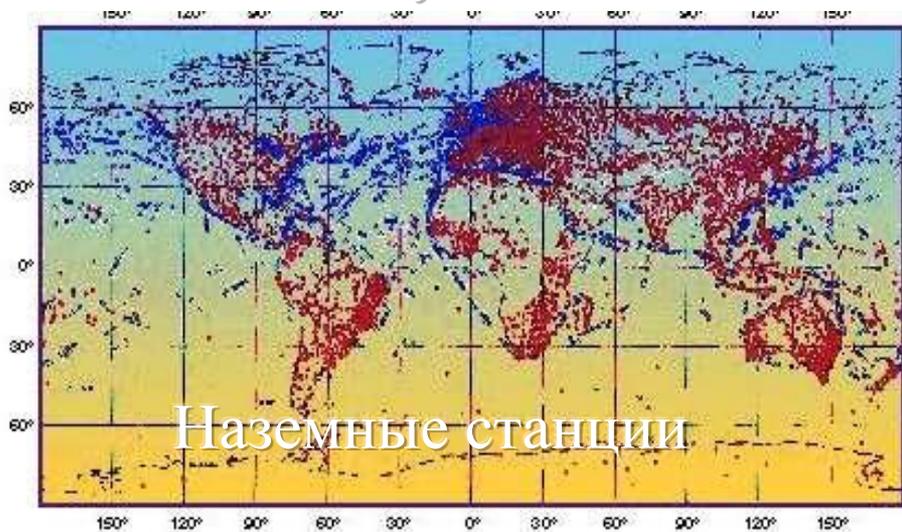
Наблюдательные системы



Спутники



Самолеты



Наземные станции



Аэрологические данные

Введение

- Под усвоением данных принято понимать задачу восстановления пространственно-временного распределения полей по данным наблюдений и математической модели, описывающей динамику полей по времени.
- Такие задачи в настоящее время рассматриваются для описания процессов в атмосфере, океане, а также распространения загрязняющих веществ в окружающей среде.
- Цель усвоения данных – как подготовка начальных полей для последующего прогноза, так и более общая – описание поведения по времени исследуемых полей, изучение климата, оценка параметров и т.д.

Введение

- Для решения задачи усвоения используются математические постановки задачи, при этом в мире в настоящее время развиваются два направления – вариационный (3DVAR-4DVAR) и динамико-стохастический (фильтр Калмана).
- Доклад посвящен динамико-стохастическому подходу к задаче усвоения.
- Рассматривается алгоритм усвоения данных, основанный на фильтре Калмана, в котором ковариационные матрицы вычисляются по ансамблю прогнозов: ансамблевый фильтр Калмана.

Математические постановки задачи оптимальной фильтрации и сглаживания

Линеаризованная модель атмосферы: $x_k^f = M_{k-1} x_{k-1}^a$,

x_k^f - n-вектор прогнозируемых значений в момент времени t_{k-1} ,

x_{k-1}^a - n-вектор проанализированных значений в момент времени t_{k-1} .

"Истинное" состояние атмосферы: $x_k^t = M_{k-1} x_{k-1}^t + \eta_{k-1}^t$,

η_k^t - случайный вектор "шумов" модели:

$$E\eta_k^t = 0;$$

$$E(\eta_k^t)(\eta_l^t)^T = Q_k \delta_{kl}$$

Математические постановки задачи оптимальной фильтрации и сглаживания

Данные наблюдений: $y_k^0 = H_k x_k^t + \xi_k^0$,

y_k^0 - m -вектор наблюдений в момент времени t_k , $H_k - (n \times m)$ матрица,

ξ_k^0 - случайный m -вектор ошибок наблюдений:

$$E \xi_k^0 = 0;$$

$$E(\xi_k^0)(\xi_l^0)^T = R_k \delta_{kl};$$

$$E(\xi_k^0)(\eta_l^t)^T = 0.$$

Задача оптимальной фильтрации состоит в поиске $\min J$ по данным наблюдений в моменты времени t_k , задача сглаживания — в момент $t_k, (k = 0, \dots, K)$

$$J = E(x_k^a - x_k^t)^T (x_k^a - x_k^t)$$

Алгоритм фильтра Калмана

$$x_k^f = M_{k-1} x_{k-1}^a;$$

$$P_k^f = M_{k-1} P_{k-1}^a M_{k-1}^T + Q_{k-1};$$

$$K_k = P_k^f H_k^T (H_k P_k^f H_k^T + R_k)^{-1};$$

$$P_k^a = (I - K_k H_k) P_k^f;$$

$$x_k^a = x_k^f + K_k (y_k^0 - H_k x_k^f);$$

$$k = 0, \dots, K.$$

$$P_k^f = E(x_k^f - x_k^t)(x_k^f - x_k^t)^T; P_k^a = E(x_k^a - x_k^t)(x_k^a - x_k^t)^T.$$

Ансамблевый фильтр Калмана

(*Ensemble Kalman filter – EnKF, Evensen, 1994; Mitchell, Houtekamer, 1998*)

Запишем нелинейную динамическую систему в виде уравнения процесса

$$\mathbf{x}_k^t = f(\mathbf{x}_{k-1}^t) + \boldsymbol{\eta}_{k-1}^t \quad (1)$$

и уравнения наблюдений

$$\mathbf{y}_k = h(\mathbf{x}_k^t) + \boldsymbol{\varepsilon}_k^t.$$

где $\boldsymbol{\eta}_{k-1}^t$ - вектор “шумов модели”, $\boldsymbol{\varepsilon}_k^t$ - вектор ошибок наблюдений, \mathbf{x}_k^t - вектор оцениваемых переменных в момент времени t_k , $\boldsymbol{\varepsilon}_k^t$ и $\boldsymbol{\eta}_{k-1}^t$ - Гауссовские случайные переменные:

$E[\boldsymbol{\varepsilon}_k^t (\boldsymbol{\varepsilon}_k^t)^T] = \mathbf{R}_k^t$, $E[\boldsymbol{\eta}_{k-1}^t (\boldsymbol{\eta}_{k-1}^t)^T] = \mathbf{Q}_{k-1}^t$. Будем считать \mathbf{x}_k^t “истинным” значением.

Ансамблевый фильтр Калмана

Ансамблевый фильтр Калмана состоит из ансамбля прогнозов $\{\mathbf{x}_k^{f,n}, n=1, \dots, N\}$

$$\mathbf{x}_k^{f,n} = f(\mathbf{x}_{k-1}^{a,n}) + \boldsymbol{\eta}_{k-1}^n \quad (2)$$

и ансамбля анализов $\{\mathbf{x}_k^{a,n}, n=1, \dots, N\}$

$$\mathbf{x}_k^{a,n} = \mathbf{x}_k^{f,n} + \mathbf{K}_k (\mathbf{y}_k^n + \boldsymbol{\varepsilon}_k^n - h(\mathbf{x}_k^{f,n})) . \quad (3)$$

Ансамбли (2) и (3) задают выборку значений “истины”, при этом среднее по выборке будет являться оптимальной оценкой, а отклонение от среднего – ансамблем ошибок анализа и прогноза, соответственно. Для осуществления ансамблевого варианта алгоритма фильтра Калмана требуется задание ансамбля ошибок наблюдений $\{\boldsymbol{\varepsilon}_k^n, n=1, \dots, N\}$ а также ансамбля ошибок прогноза $\{\mathbf{d}\mathbf{x}_k^{f,n} = \mathbf{x}_k^{f,n} - \overline{\mathbf{x}_k^{f,n}}, n=1, \dots, N\}$, где $\overline{\mathbf{x}_k^{f,n}} \cong \frac{1}{N} \sum_{n=1}^N \mathbf{x}_k^{f,n}$ и ансамбля шумов модели

$\{\boldsymbol{\eta}_{k-1}^n, n=1, \dots, N\} : E[\boldsymbol{\eta}_{k-1}^n (\boldsymbol{\eta}_{k-1}^n)^T] = \mathbf{Q}_k$. Матрица \mathbf{K}_k имеет вид

$$\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k)^{-1} ,$$

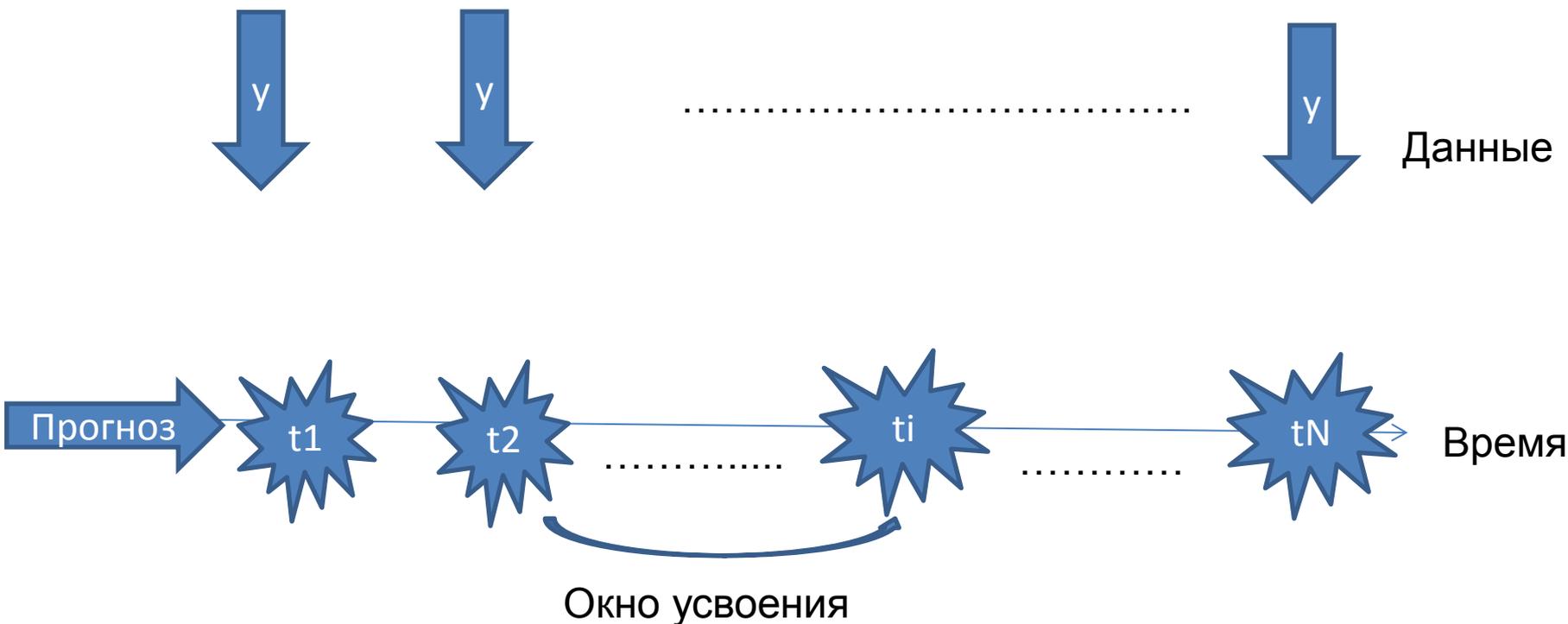
где \mathbf{P}_k^f и \mathbf{R}_k - матрицы, оцениваемые по ансамблю

$$\mathbf{P}_k^f \triangleq \frac{1}{N-1} \sum_{n=1}^N \mathbf{d}\mathbf{x}_k^{f,n} (\mathbf{d}\mathbf{x}_k^{f,n})^T , \mathbf{R}_k \triangleq \frac{1}{N-1} \sum_{n=1}^N \boldsymbol{\varepsilon}_k^n (\boldsymbol{\varepsilon}_k^n)^T ,$$

\mathbf{H}_k линеаризованный оператор $h(\mathbf{x}_k^{f,n})$ относительно $\overline{\mathbf{x}_k^{f,n}}$:

$$h(\mathbf{x}_k) \cong h(\overline{\mathbf{x}_k^{f,n}}) + \mathbf{H}_k \boldsymbol{\varepsilon}_k^f .$$

Практическая реализация алгоритма усвоения данных



- Оптимальная оценка по всем данным наблюдений в момент времени t_N – задача оптимальной фильтрации.
- Оптимальная оценка по всем данным наблюдений в момент времени t_i – задача сглаживания.

Задача оптимального сглаживания. Ансамблевое сглаживание (EnKS – ensemble Kalman smoother)

Данные наблюдений $Y = \{y_k, y_{k+1}, \dots, y_{k+l}\}$

Если

- ошибки наблюдений - Гауссовские случайные величины;
- ошибки наблюдений в разные моменты времени не коррелируют

то

алгоритм EnKS может осуществляться последовательно (циклически) по мере поступления данных, используя данные в момент времени t_j для оценки значений в момент времени t_k ($t_k < t_j$).

В конечный момент времени результаты EnKF и EnKS совпадают.

Evensen, G., P.J. van Leeuwen An ensemble Kalman smoother for nonlinear dynamics // Monthly Weather Review. 2000. V. 128. P. 1852-1867.

Evensen, G. Data assimilation. The ensemble Kalman filter. Berlin Heidelberg: Springer-Verlag, 2009. 307 p.

Задача оптимального сглаживания. Ансамблевое сглаживание (EnKS – ensemble Kalman smoother)

В EnKF оператор \mathbf{H} переводит значения в узлах сетки в точки наблюдений,
В EnKS – \mathbf{H} включает также интерполяцию по времени.

$$\mathbf{x}_k^{a,n} = \mathbf{x}_k^{f,n} + \sum_{n=1}^N \mathbf{d}\mathbf{x}_k^{f,n} (\mathbf{d}\mathbf{x}_k^{f,n})^T \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k)^{-1} (\mathbf{y}_k^n + \boldsymbol{\varepsilon}_k^n - h(\mathbf{x}_k^{f,n}))$$

$$\mathbf{P}_k^f \triangleq \frac{1}{N-1} \sum_{n=1}^N \mathbf{d}\mathbf{x}_k^{f,n} (\mathbf{d}\mathbf{x}_k^{f,n})^T, \mathbf{R}_k \triangleq \frac{1}{N-1} \sum_{n=1}^N \boldsymbol{\varepsilon}_k^n (\boldsymbol{\varepsilon}_k^n)^T$$

Оценка в момент
времени t_k по данным
 y_n -EnKF

$$\mathbf{x}_l^{a,n} = \mathbf{x}_{l-1}^{a,n} + \sum_{n=1}^N \mathbf{d}\mathbf{x}_{l-1}^{a,n} (\mathbf{d}\mathbf{x}_k^{f,n})^T \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k)^{-1} (\mathbf{y}_k^n + \boldsymbol{\varepsilon}_k^n - h(\mathbf{x}_k^{f,n}))$$

$$\mathbf{P}_k^f \triangleq \frac{1}{N-1} \sum_{n=1}^N \mathbf{d}\mathbf{x}_k^{f,n} (\mathbf{d}\mathbf{x}_k^{f,n})^T, \mathbf{R}_k \triangleq \frac{1}{N-1} \sum_{n=1}^N \boldsymbol{\varepsilon}_k^n (\boldsymbol{\varepsilon}_k^n)^T$$

Оценка в момент
времени t_l по
данным y_k -EnKS

Evensen, G., P.J. van Leeuwen An ensemble Kalman smoother for nonlinear dynamics // Monthly Weather Review. 2000. V. 128. P. 1852-1867.
Evensen, G. Data assimilation. The ensemble Kalman filter. Berlin Heidelberg: Springer-Verlag, 2009. 307 p.

Оценка параметров модели в процедуре усвоения данных

Рассмотрим уравнение процесса в виде

$$\mathbf{x}_k^t = f(\mathbf{x}_{k-1}^t, \mathbf{a}_{k-1}^t) + \boldsymbol{\eta}_{k-1}^t,$$

данные наблюдений

$$\mathbf{y}_k = h(\mathbf{x}_k^t, \mathbf{a}_k^t) + \boldsymbol{\varepsilon}_k^t,$$

где \mathbf{a}_k^t - вектор параметров. Будем считать, что параметр не меняется со временем: $\mathbf{a}_{k+1}^t = \mathbf{a}_k^t$.

Рассмотрим обобщенную задачу оценивания вектора $\mathbf{z} = [\mathbf{x}, \mathbf{a}]^T$. Опуская промежуточные выкладки, запишем сразу результат процедуры оценивания в общем виде:

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{P}_{xx} h_x^T (h_x \mathbf{P}_{xx} h_x^T + \mathbf{R})^{-1} [\mathbf{y}_0 - h(\mathbf{x}^f, \mathbf{a}^f)] + \mathbf{P}_{x\alpha} h_\alpha^T (h_x \mathbf{P}_{xx} h_x^T + \mathbf{R})^{-1} [\mathbf{y}_0 - h(\mathbf{x}^f, \mathbf{a}^f)],$$

$$\mathbf{a}^a = \mathbf{a}^f + \mathbf{P}_{\alpha x} h_x^T (h_x \mathbf{P}_{xx} h_x^T + \mathbf{R})^{-1} [\mathbf{y}_0 - h(\mathbf{x}^f, \mathbf{a}^f)] + \mathbf{P}_{\alpha\alpha} h_\alpha^T (h_x \mathbf{P}_{xx} h_x^T + \mathbf{R})^{-1} [\mathbf{y}_0 - h(\mathbf{x}^f, \mathbf{a}^f)].$$

В этих формулах $\mathbf{P}_{\alpha x}$ - кросс-ковариации ошибок \mathbf{x} и \mathbf{a} , $\mathbf{P}_{\alpha\alpha}$ - ковариационная матрица ошибок \mathbf{a} . h_x и h_α - линеаризованные операторы по \mathbf{x} и \mathbf{a} , соответственно. В случае, если h от \mathbf{a} не зависит, оценка \mathbf{x}^a проводится по той же формуле, что и в обычном фильтре Калмана.

Задача оптимального сглаживания. Ансамблевое сглаживание (EnKS – ensemble Kalman smoother). Оценка параметров.

$$\mathbf{a}_k^{a,n} = \mathbf{a}_k^{f,n} + \sum_{n=1}^N d\mathbf{a}_k^{f,n} (\mathbf{d}\mathbf{x}_k^{f,n})^T \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k)^{-1} (\mathbf{y}_k^n + \boldsymbol{\varepsilon}_k^n - h(\mathbf{x}_k^{f,n}))$$

$$\mathbf{P}_k^f \triangleq \frac{1}{N-1} \sum_{n=1}^N d\mathbf{x}_k^{f,n} (\mathbf{d}\mathbf{x}_k^{f,n})^T, \mathbf{R}_k \triangleq \frac{1}{N-1} \sum_{n=1}^N \boldsymbol{\varepsilon}_k^n (\boldsymbol{\varepsilon}_k^n)^T$$

Оценка в момент времени t_k по данным y_n -EnKF

$$\mathbf{a}_l^{a,n} = \mathbf{a}_{l-1}^{a,n} + \sum_{n=1}^N d\mathbf{a}_{l-1}^{a,n} (\mathbf{d}\mathbf{x}_k^{f,n})^T \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k)^{-1} (\mathbf{y}_k^n + \boldsymbol{\varepsilon}_k^n - h(\mathbf{x}_k^{f,n}))$$

$$\mathbf{P}_k^f \triangleq \frac{1}{N-1} \sum_{n=1}^N d\mathbf{x}_k^{f,n} (\mathbf{d}\mathbf{x}_k^{f,n})^T, \mathbf{R}_k \triangleq \frac{1}{N-1} \sum_{n=1}^N \boldsymbol{\varepsilon}_k^n (\boldsymbol{\varepsilon}_k^n)^T$$

Оценка в момент времени t_l по данным y_k -EnKS

Evensen, G., P.J. van Leeuwen An ensemble Kalman smoother for nonlinear dynamics // Monthly Weather Review. 2000. V. 128. P. 1852-1867.
Evensen, G. Data assimilation. The ensemble Kalman filter. Berlin Heidelberg: Springer-Verlag, 2009. 307 p.

Подходы к реализации ансамблевого фильтра Калмана

«Стохастический
подход» (EnKF)

«Истина» (true value)

$$x_{k+1}^t = f_k(x_k^t) + \eta_k^t$$

Ансамбль прогнозов

$$x_{k+1}^{f(i)} = f_k(x_k^{a(i)}) + \eta_k^{(i)}$$

Ансамбль анализов

$$x_k^{a(i)} = x_k^{f(i)} + K_k (y_k^{0(i)} - H_k x_k^{f(i)})$$

Ошибка оценки (skill)

$$\delta^a = \{x^t - \overline{x^{a(i)}}\}, \delta^f = \{x^t - \overline{x^{f(i)}}\}$$

Отклонение от среднего (spread)

$$X^a = \{x^{a(i)} - \overline{x^{a(i)}}\}, X^f = \{x^{f(i)} - \overline{x^{f(i)}}\}$$

«Детерминированный
подход» (ESRF, ETKF,
LETKF)

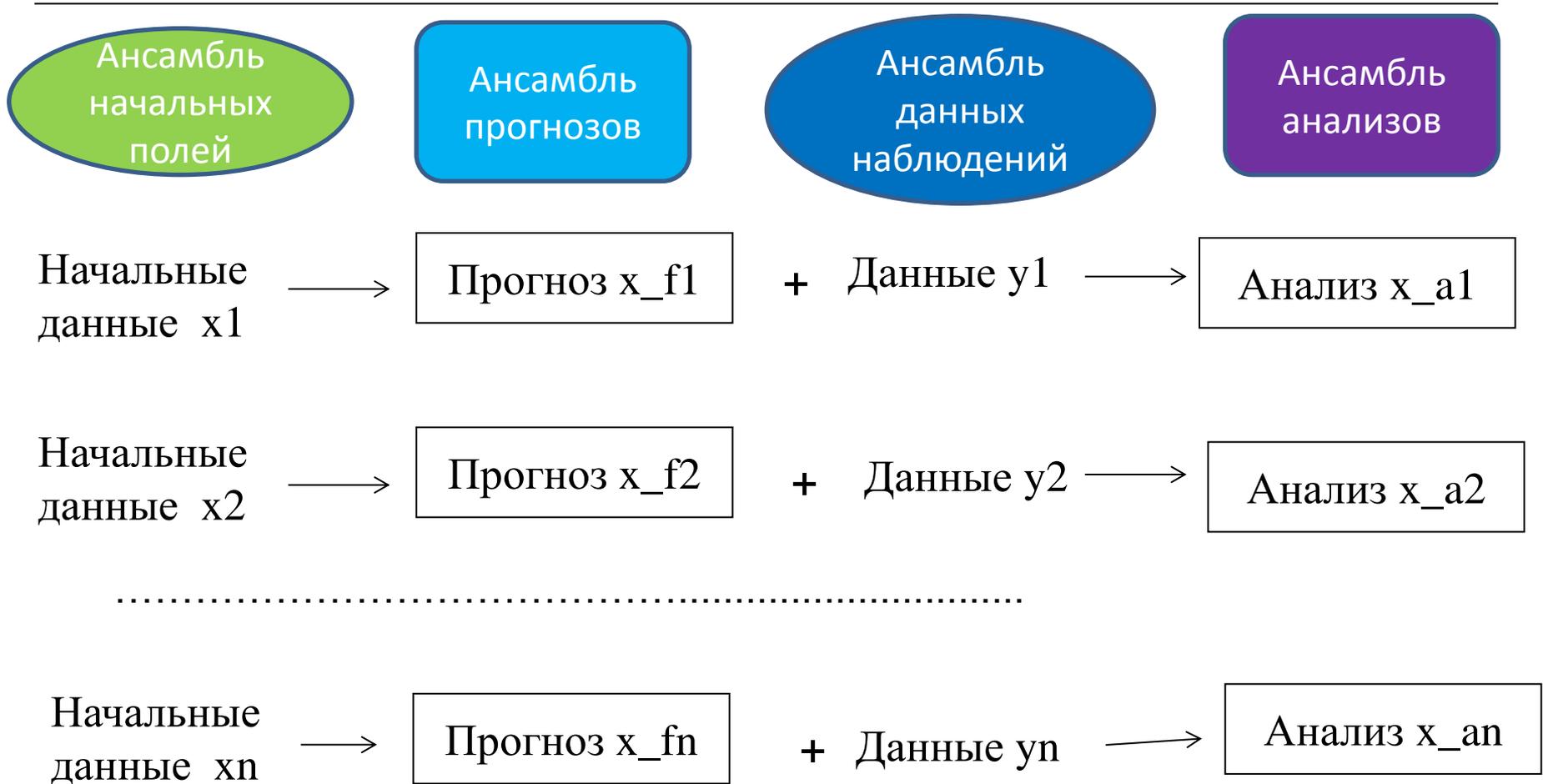
$$x^{a(i)} = \bar{x}^a + dx^{a(i)},$$

$$\overline{dx^{a(i)} dx^{a(i)T}} = P^a$$

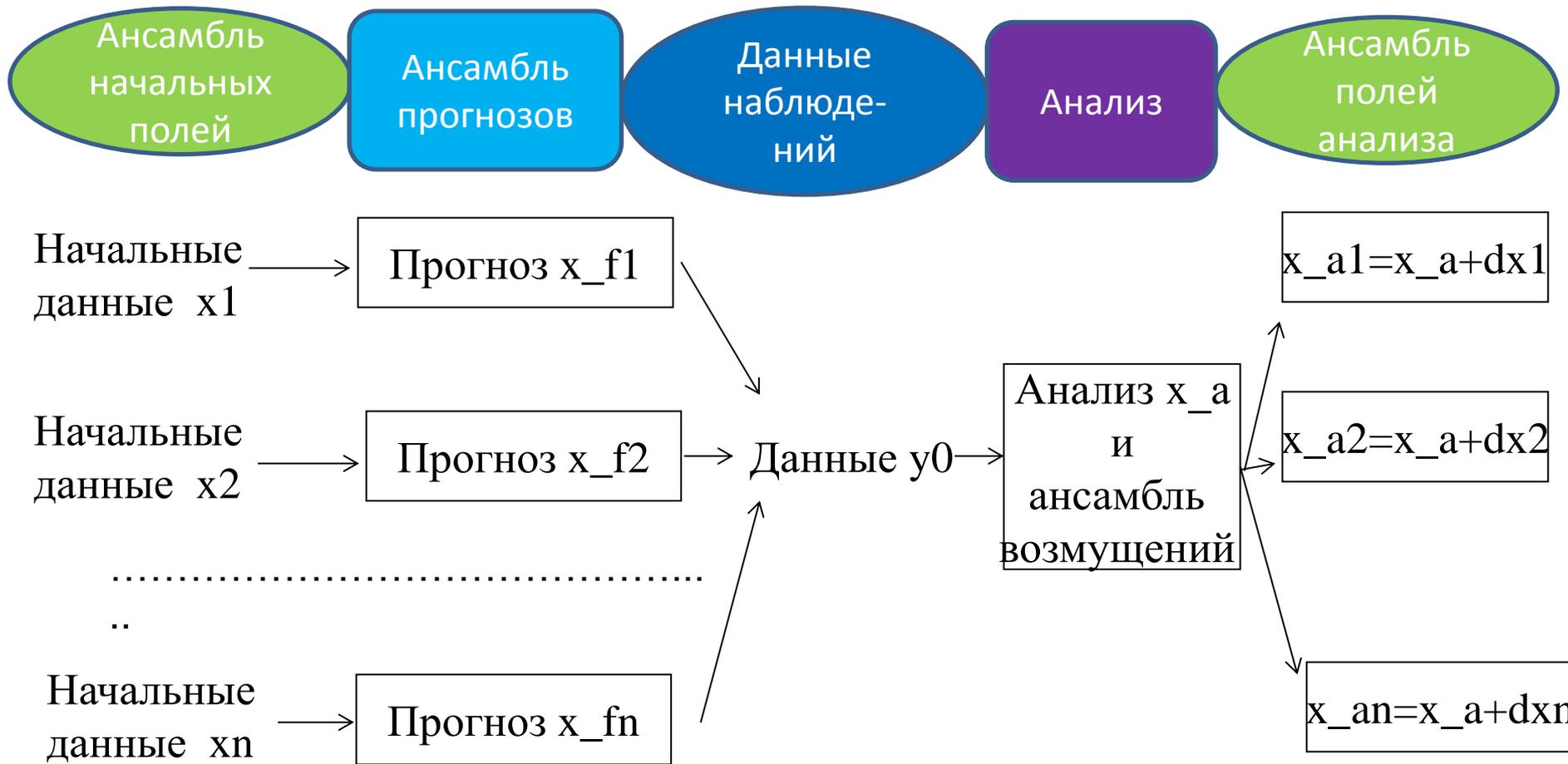
Трансформация
ансамбля прогнозов

$$dx_k^{a(i)} = A_k dx_k^{f(i)}$$

Ансамблевая схема усвоения (стохастическая)



Ансамблевая схема усвоения (детерминированная)



Ensemble Transform Kalman Filter - ETKF (Bishop et al, 2001)

- Ковариационная матрица P^f преобразуется в матрицу P^a с помощью «матрицы трансформации» T .
- Все операции осуществляются с матрицами размерности N (количество элементов ансамбля).

Local Ensemble Transform Kalman Filter- LETKF (Hunt et al, 2007)

- Объединение идей *Ensemble Transform Kalman Filter* и *Local Ensemble Kalman Filter* (анализ проводится локально в подобластях).
- Анализ может проводиться для каждого узла сетки независимо.
- Анализ проводится только для «среднего» значения.
- Вычисляется ансамбль анализов, соответствующий матрице ковариаций ошибок анализа.

Local Ensemble Transform Kalman Filter- LETKF (Hunt et al, 2007)

$$X^a = \{x^{a(i)} - \overline{x^{a(i)}}\}$$

- ансамбли анализов и прогнозов.

$$X^f = \{x^{f(i)} - \overline{x^{f(i)}}\}$$

Формулы LETKF:

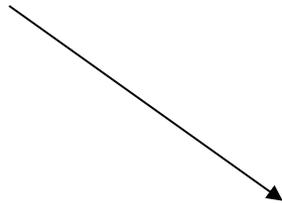
ρ - “inflation factor”.

$$\left\{ \begin{array}{l} 1) \tilde{P}^a = \left[(k-1)I / \rho + (Y^f)^T R^{-1} Y^f \right]^{-1}, Y^f = H X^f \\ 2) W^a = \left[(k-1) \tilde{P}^a \right]^{1/2} \\ 3) \bar{w}^a = \tilde{P}^a C (y_o - H \bar{x}^f), C = (Y^f)^T R^{-1} \\ 4) x^{a(i)} = \bar{x}^f + X^f w^{a(i)}, w^{a(i)} - i - \text{й столбец } W \\ 5) \bar{x}^a = \bar{x}^f + X^f \bar{w}^a \qquad 6) P^a = X^f \tilde{P}^a (X^f)^T \end{array} \right.$$

Подходы к реализации ансамблевого фильтра Калмана

«Стохастический
подход»:

EnKF

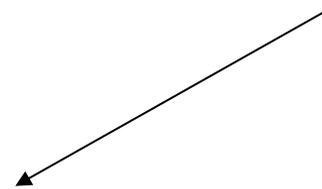


«Детерминированный
подход»:

ESRF

ETKF

LETKF



Ансамблевый пи-алгоритм:

Klimova E. A suboptimal data assimilation algorithm based on the ensemble Kalman filter. Quarterly Journal of the Royal Meteorological Society. 2012. DOI:10.1002/qj.1941.

Ансамблевый пи-алгоритм

Ансамблевый фильтр Калмана состоит из вычисления ансамбля прогнозов:

$$\mathbf{x}_{k+1}^{f,n} = f(\mathbf{x}_k^{f,n}) + \boldsymbol{\eta}_k^n$$

и ансамбля анализов:

$$\mathbf{x}_k^{a,n} = \mathbf{x}_k^{f,n} + \tilde{\mathbf{K}}_k (\mathbf{y}_k^{o,n} - h(\mathbf{x}_k^{f,n})),$$

где $\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k)^{-1}$. В этих формулах $\mathbf{x}_k^{f,n}$, $\mathbf{x}_k^{a,n}$ - ансамбли полей прогноза и анализа в момент времени t_k , f - оператор модели, h - оператор наблюдений,

$\mathbf{P}_k^f \triangleq \frac{1}{N-1} \sum_{n=1}^N \boldsymbol{\varepsilon}_k^{f,n} (\boldsymbol{\varepsilon}_k^{f,n})^T$, $\mathbf{R}_k \triangleq \frac{1}{N-1} \sum_{n=1}^N \boldsymbol{\varepsilon}_k^n (\boldsymbol{\varepsilon}_k^n)^T$, $\{\boldsymbol{\varepsilon}_k^{f,n} = \mathbf{x}_k^{f,n} - \overline{\mathbf{x}_k^{f,n}}, n=1, \dots, N\}$ - ансамбль

ошибок прогноза, $\overline{\mathbf{x}_k^{f,n}} \triangleq \frac{1}{N} \sum_{n=1}^N f(\mathbf{x}_{k-1}^n)$, $\mathbf{y}_k^{o,n} = \mathbf{y}_k - \boldsymbol{\varepsilon}_k^n$ - ансамбль наблюдений с

возмущениями (perturbed observations), $\boldsymbol{\varepsilon}_k^n$ - случайные ошибки наблюдений с нулевым математическим ожиданием и матрицей ковариаций \mathbf{R}_k , \mathbf{y}_k - вектор наблюдений в

момент времени t_k . В ансамблевом пи-алгоритме матрица \mathbf{K}_k имеет вид $\mathbf{K}_k = \mathbf{P}_k^a \mathbf{H}_k^T \mathbf{R}_k^{-1}$.

В классической теории фильтра Калмана эти записи шага анализа эквивалентны.

Ансамблевый пи-алгоритм

Наилучшей оценкой является среднее по ансамблю значение. Определим ошибку этой оценки на шаге анализа (отклонение от среднего значения) $\mathbf{dx}_k^n \triangleq \mathbf{x}_k^{a,n} - \overline{\mathbf{x}_k^{a,n}}$, где $\overline{\mathbf{x}_k^{a,n}} \triangleq \frac{1}{N} \sum_{n=1}^N \mathbf{x}_k^{a,n}$. Уравнение для этой ошибки имеет вид

$$\mathbf{dx}_k^n = (\mathbf{x}_k^{f,n} - \overline{\mathbf{x}_k^{f,n}}) - \mathbf{P}_k^a \mathbf{H}_k^T \mathbf{R}_k^{-1} (\boldsymbol{\varepsilon}_k^n + h(\mathbf{x}_k^{f,n}) - \overline{h(\mathbf{x}_k^{f,n})}).$$

Выразим матрицу \mathbf{P}_k^a через \mathbf{dx}_k^n :

$$\mathbf{P}_k^a \triangleq \frac{1}{N-1} \sum_{n=1}^N \mathbf{dx}_k^n (\mathbf{dx}_k^n)^T.$$

Систему уравнений относительно \mathbf{dx}_k^n можно записать в матричном виде

$$\mathbf{D}^T = \mathbf{F}^T - \mathbf{\Pi}^T \mathbf{D}^T,$$

где \mathbf{D} – матрица размерности $(L \times N)$, столбцами которой являются векторы $\{\mathbf{dx}_k^n, n = 1, \dots, N\}$, L – размерность этих векторов, соответствующая числу прогнозируемых переменных, $\mathbf{\Pi}$ – $(N \times N)$ - матрица

$$\mathbf{\Pi}^T = \frac{1}{N-1} \mathbf{D}^T \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{F} + \mathbf{E}), \mathbf{F} \text{ – матрица со столбцами } \{\mathbf{f}_k^n, n = 1, \dots, N\}:$$

$$\mathbf{f}_k^n = \mathbf{x}_k^{f,n} - \overline{\mathbf{x}_k^{f,n}}.$$

Ансамблевый пи-алгоритм

Матрица $\mathbf{\Pi}$ имеет вид:

$$\mathbf{\Pi}^T = (\mathbf{C} + 0,25\mathbf{I})^{\frac{1}{2}} - 0,5\mathbf{I},$$

$$\mathbf{C} = \frac{1}{N-1} \mathbf{F}^T \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{F} + \mathbf{E}) = \mathbf{C}_1 + \mathbf{C}_2.$$

В этих формулах \mathbf{E} - матрица, столбцы которой равны вектору $\boldsymbol{\varepsilon}_k^n$.

Для получения матрицы $\mathbf{\Pi}$ требуется вычисление корня из несимметричной матрицы.

После вычисления матрицы \mathbf{D} определяется ковариационная матрица \mathbf{R}_k^a и анализ для среднего по ансамблю значения:

$$\overline{\mathbf{x}_k^{a,n}} = \overline{\mathbf{x}_k^{f,n}} + \frac{1}{N-1} \mathbf{D} \mathbf{D}^T \mathbf{H}^T \mathbf{R}^{-1} \left[\mathbf{y}_k - \overline{h(\mathbf{x}_k^{f,n})} \right].$$

Ансамбль соответствующих анализов в матричном виде можно записать как

$$\mathbf{X}_k^a = \overline{\mathbf{x}_k^{a,n}} + \mathbf{D},$$

где \mathbf{X}_k^a - $(L \times N)$ матрица, столбцами которой являются векторы $\{\mathbf{x}_k^{a,n}, n = 1, \dots, N\}$.

Использование трансформаций в ансамблевом сглаживании (EnKS – ensemble Kalman smoother)

Данные наблюдений $Y = \{y_k, y_{k+1}, \dots, y_{k+l}\}$

Алгоритм EnKS,
оценка в момент
времени t_k :

$$\mathbf{x}_{tk}^{KS} = \left(\prod_{i=k+1}^{k+l} \mathbf{v}_i \right) \mathbf{x}_{tk}^{KF}, \quad \mathbf{X} = \{x_{tk}^n, n = 1, \dots, N\}$$

Использование трансформаций
ансамбля возмущений:

$$x^{a(tj)} = \bar{x}^{a(tj)} + dx^{a(tj)}, \quad dx^{a(tj)} = \mathbf{A}_j dx^{a(tj-1)}$$

При этом возможны детерминированный и стохастический варианты.

Evensen, G. Data assimilation. The ensemble Kalman filter. Berlin Heidelberg: Springer-Verlag, 2009. 307 p.

Подходы, основанные на предположении о стационарности временного ряда ошибок

- EnOI (*Evensen G.*) – «ансамблевая оптимальная интерполяция».
- Замена среднего по ансамблю средним по времени: делается предположение о стационарности временного ряда ошибок; ковариации ошибок оценивания (прогноза) оцениваются с помощью осреднения по времени.

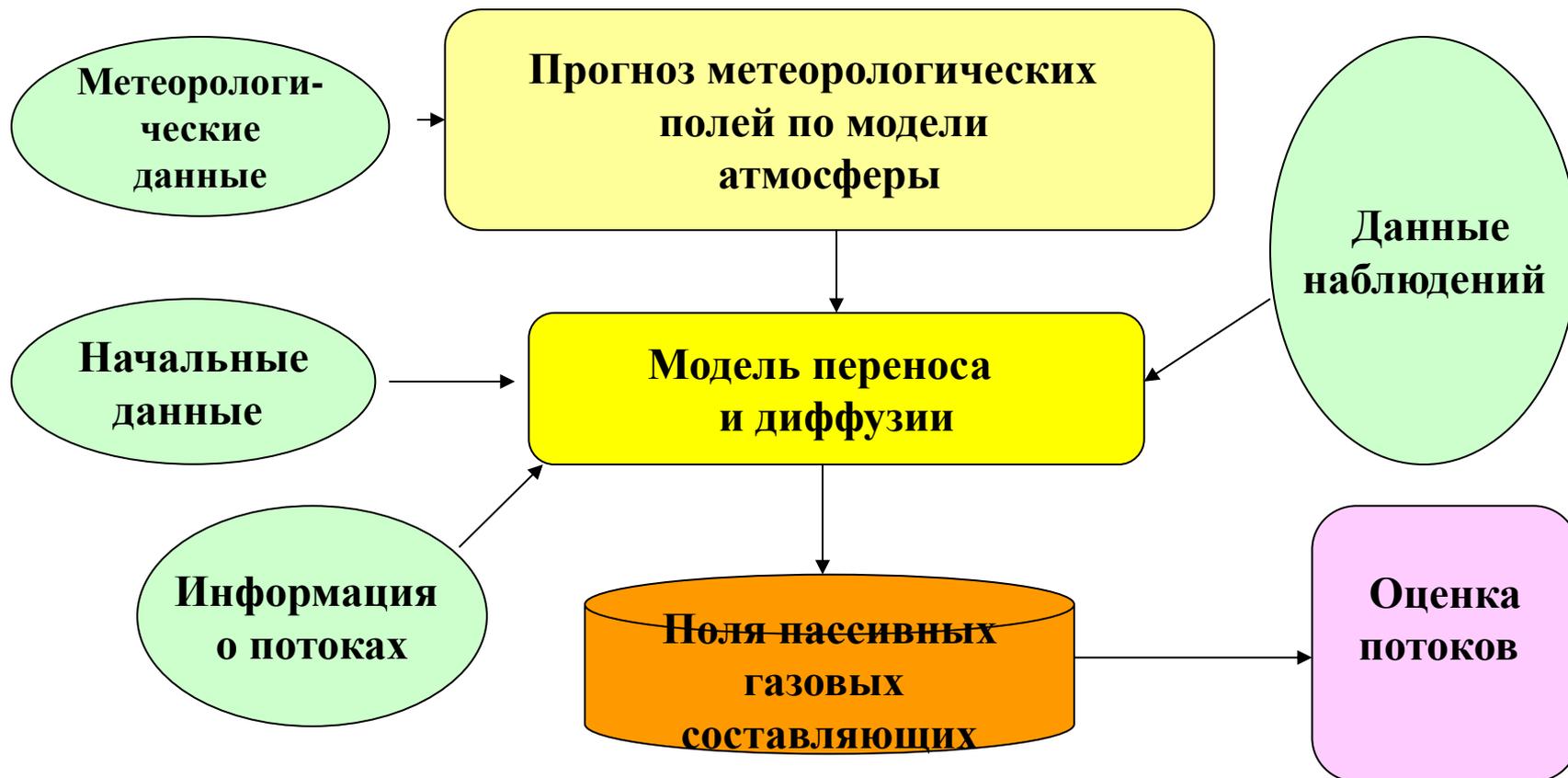
Evensen, G. Data assimilation. The ensemble Kalman filter. Berlin Heidelberg: Springer-Verlag, 2009. 307 p.

Климова Е.Г., Платов Г.А., Киланова Н.В. Разработка системы усвоения данных об окружающей среде на основе ансамблевого фильтра Калмана // Вычислительные технологии. 2014. Том 19, №3. С.27-37.

Усвоение данных наблюдений в задаче моделирования состояния окружающей среды

1. Моделирование распространения в атмосфере:
 - Пассивных газовых составляющих (“greenhouse”) (CO_2 , CH_4).
 - Химически активных веществ.
 - Аэрозолей.
2. Данные наблюдений.
3. Восстановление пространственно-временного распределения газовых составляющих с помощью системы усвоения данных.
4. Оценка областей, влияющих на распределение загрязняющих веществ в заданном регионе.
5. Оценка потоков газовых составляющих в атмосфере на основе системы усвоения данных наблюдений.

Усвоение данных о пассивных газовых составляющих



Современные направления в работах по оценке потоков парниковых газов по спутниковым данным

Цикл работ по оценке потоков CO₂

- [1] Feng L., P.I.Palmer, H.Bosch, and S.Dance Estimating surface CO₂ fluxes from space-borne CO₂ dry air mole fraction observations using an ensemble Kalman filter // Atmospheric chemistry and physics. 2009. V. 9. P. 2619-2633.
- [2] Feng L. et al. Evaluating a 3-D transport model of atmospheric CO₂ using ground-based, aircraft, and space-borne data // Atmospheric chemistry and physics. 2011. V. 11. P. 2789-2803.
- [3] Feng L. et al. Estimates of European uptake of CO₂ inferred from GOSAT X_{co2} retrievals: sensitivity to measurement bias inside and outside Europe // Atmospheric chemistry and physics. 2016. V. 16. P.1289-1302.
- [4] Feng L. et al. Consistent regional fluxes of CH₄ and CO₂ inferred from GOSAT proxy XCH₄:XCO₂ retrievals, 2010-2014 // Atmospheric chemistry and physics. 2017. V. 17. P. 4781-4797.

В серии работ группы авторов используется модель GEOS-Chem (3-D chemistry transport model) и спутниковые данные GOSAT. Кроме того, используются дополнительные наблюдения сети GLOBALVIEW, а также данные из других источников.

В этих работах рассматривается окно усвоения, состоящее из 12 циклов по 8 дней, и производится оценка средних за 8 дней значений потоков.

Современные направления в работах по оценке потоков парниковых газов по спутниковым данным

В этих работах проводится оценка значений средних по подобластям потоков \mathbf{x}^a по данным наблюдений \mathbf{y}_0 и прогнозу \mathbf{x}^f (поверхность Земли разбивается на квадраты равной площади (1000 км на 1000 км) по стандартной формуле фильтра Калмана:

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}[\mathbf{y}_0 - H(\mathbf{x}^f)],$$

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1}.$$

Для реализации ансамблевого фильтра Калмана задается ансамбль возмущений оцениваемого параметра:

$$\Delta \mathbf{x}^f = \frac{1}{\sqrt{N}} [\Delta x_1, \dots, \Delta x_N]^T,$$

Матрица \mathbf{P}^f оценивается по ансамблю

$$\mathbf{P}^f = \Delta \mathbf{x}^f (\Delta \mathbf{x}^f)^T,$$

$$\mathbf{K}_e = \Delta \mathbf{x}^f (\Delta \mathbf{y})^T [\Delta \mathbf{y} (\Delta \mathbf{y})^T + \mathbf{R}]^{-1},$$

$$\Delta \mathbf{y} = H(\mathbf{x}^f + \Delta \mathbf{x}^f) - H(\mathbf{x}^f),$$

где оператор H включает в себя прогноз по модели к моменту наблюдения, интерполяцию из узлов сетки в точки наблюдений, а также, в случае спутниковых данных, осреднение по вертикали с известными коэффициентами («average kernel»), N – число элементов ансамбля.

Современные направления в работах по оценке потоков парниковых газов по спутниковым данным

С точки зрения математической постановки задачи выделим следующие моменты:

1. производится оценка потоков без уточнения концентраций; решение задачи в такой постановке не будет оптимальным;
2. поскольку оператор H включает в себя математическую модель распространения примеси данные наблюдений являются асинхронными и данная задача является задачей сглаживания, а не фильтрации.

С точки зрения практической реализации для решения данной задачи требуется задание матрицы \mathbf{R} - матрицы ковариаций ошибок наблюдений и модели, начальные (климатические) значения потоков.

Современные направления в работах по оценке потоков парниковых газов по спутниковым данным

Tsuruta A. et al. Global methane emission estimation for 2010-2012 from Carbon tracker Europe-CH₄ v.10. // Geoscientific model development. 2017. V. 10. P. 1261-1287.

В данной работе, посвященной оценке эмиссии CH₄, рассматривается аналогичная постановка задачи и производится сравнительный анализ зависимости результатов от целого набора параметров. Как отмечается в этой работе, результат оценивания зависит от используемой модели, от параметризаций приземного и пограничного слоя в этих моделях, а также от методики усвоения данных.

The atmospheric chemistry transport model TM5

3-hourly ECMWF ERA-Interim meteorological fields with 25 vertical layers

GOSAT

Aircraft profiles

Surface CH₄ measurements

Современные направления в работах по оценке потоков парниковых газов по спутниковым данным

Kang J. et al. Estimating of surface carbon fluxes with an advanced data assimilation methodology // Journal of geophysical research. 2012. V.116. D24101, doi:10.1029/2012JD018259.

- Применение ансамблевого фильтра Калмана для оценки потоков CO_2 по данным AIRS.
- С математической точки зрения, главным отличием этой работы является то, что рассматривается процедура совместной оценки концентрации и потоков вместе с метеорологическими параметрами.
- Используется сравнительно небольшое окно усвоения (6 часов).
- В качестве метода используется алгоритм LETKF – детерминированный ансамблевый фильтр Калмана.
- Решение задачи совместного оценивания концентрации и эмиссии является оптимальным с математической точки зрения. Однако задача становится технически гораздо более сложно реализуемой

SPEEDY-C model

Тестирование системы усвоения с помощью модельных данных

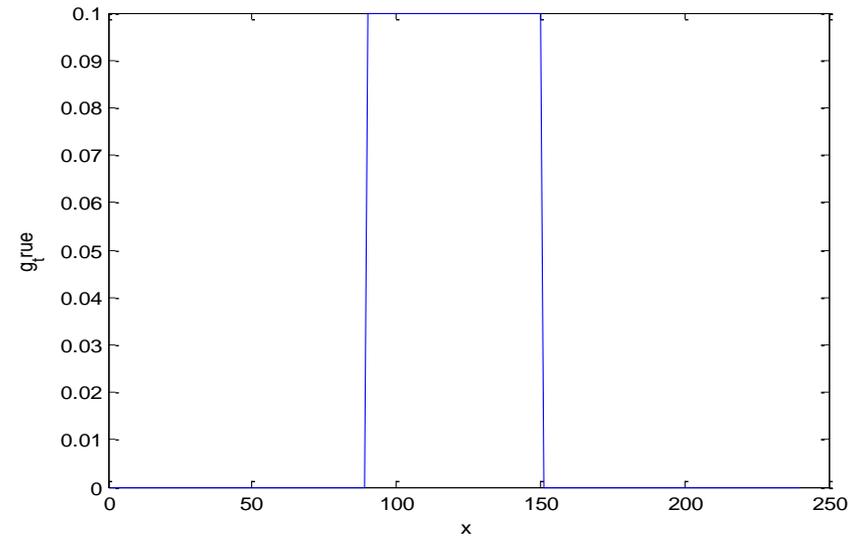
Тестовые расчеты с 1-мерной моделью переноса и диффузии

$$\Phi_{n+1} = \mathbf{F}_n \Phi_n + \mathbf{g}_n$$

$$\Delta x = \Delta t; u = 1$$

$$L = 240; N_{time} = 240$$

$$\text{inflation factor} = \begin{cases} 1.3, & \text{ensemble - pi} \\ 1.2, & \text{letkf} \end{cases}$$



Вариант 1 :

(первая серия экспериментов)

$$g_t(x, t) = g_0(x), g(x, 0) = 0$$

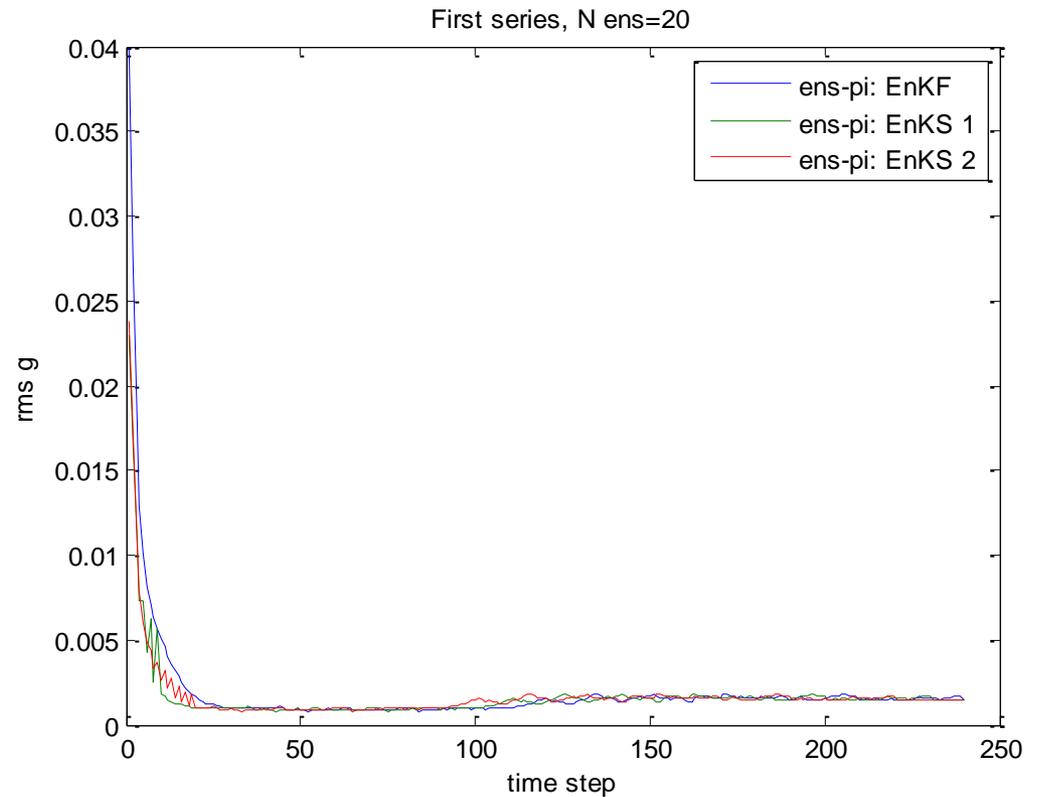
Вариант 2:

(вторая серия экспериментов)

$$g_t(x, t) = \begin{cases} g_0(x, t), & 0 \leq t \leq \bar{t} \\ 1.5 \times g_0(x, t), & t > \bar{t} \end{cases}, g(x, 0) = 0$$

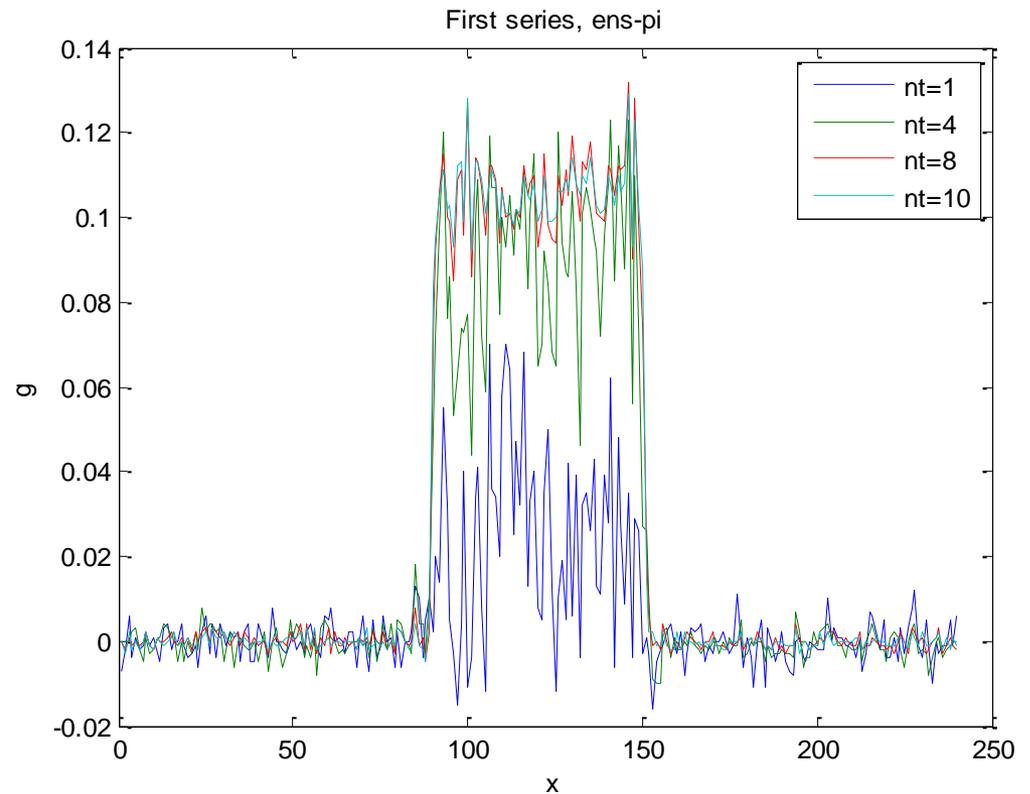
Тестовые расчеты с 1-мерной моделью переноса и диффузии

*Среднеквадратическая
ошибка оценки
параметра $g(x,t)$,
первая серия
экспериментов,
стохастические
алгоритмы EnKF, EnKS
(EnKS 1 – time lag=10,
EnKS 2 – time lag=20)*



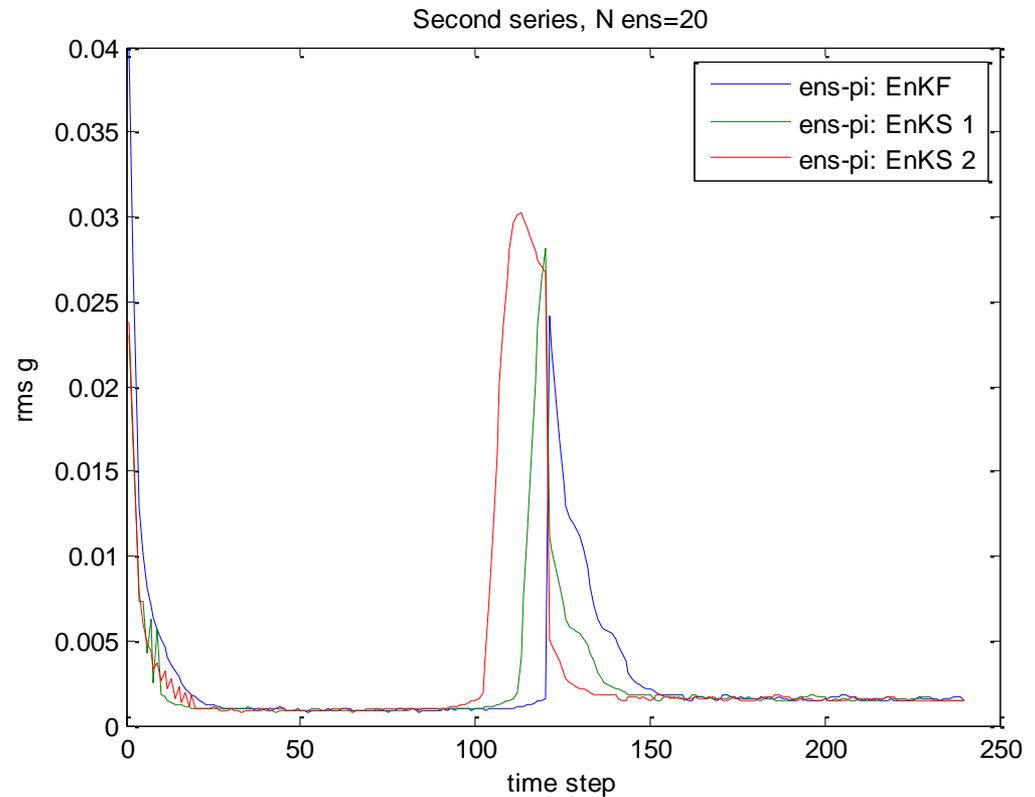
Тестовые расчеты с 1-мерной моделью переноса и диффузии

*Поведение по времени
оцениваемого
параметра $g(x,t)$,
первая серия
экспериментов,
 nt - номер шага по
времени*



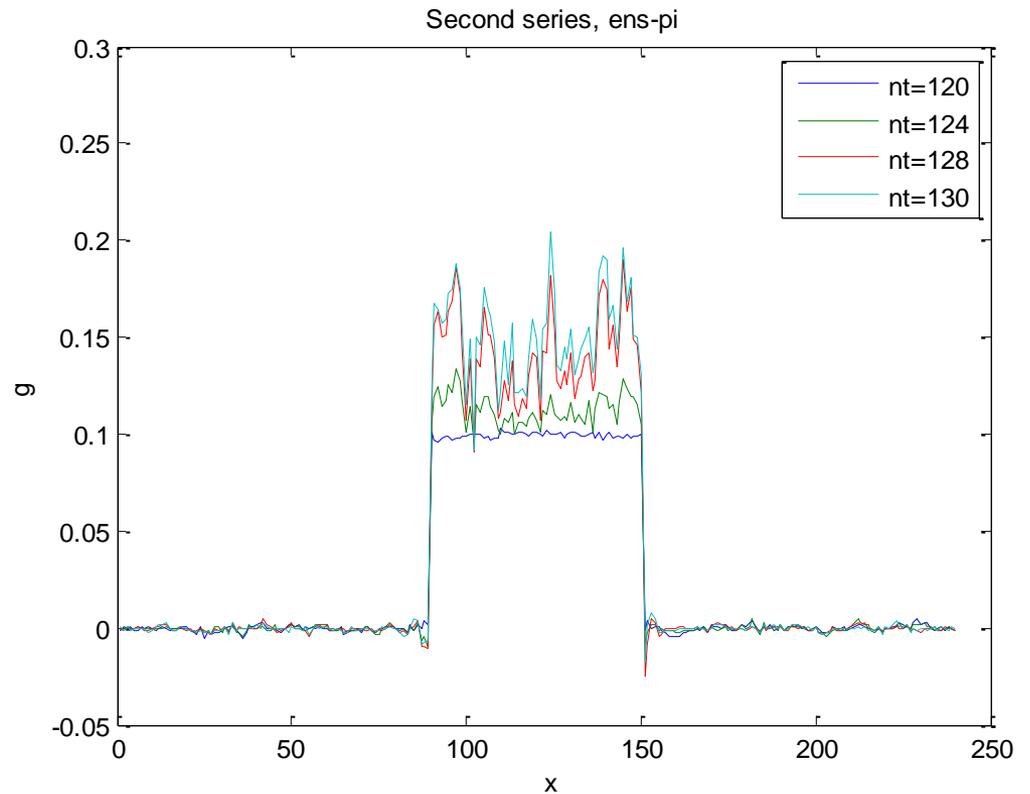
Тестовые расчеты с 1-мерной моделью переноса и диффузии

*Среднеквадратическая ошибка
оценки параметра $g(x,t)$,
вторая серия экспериментов,
стохастические алгоритмы EnKF,
EnKS (EnKS 1 – time lag=10,
EnKS 2 – time lag=20)*



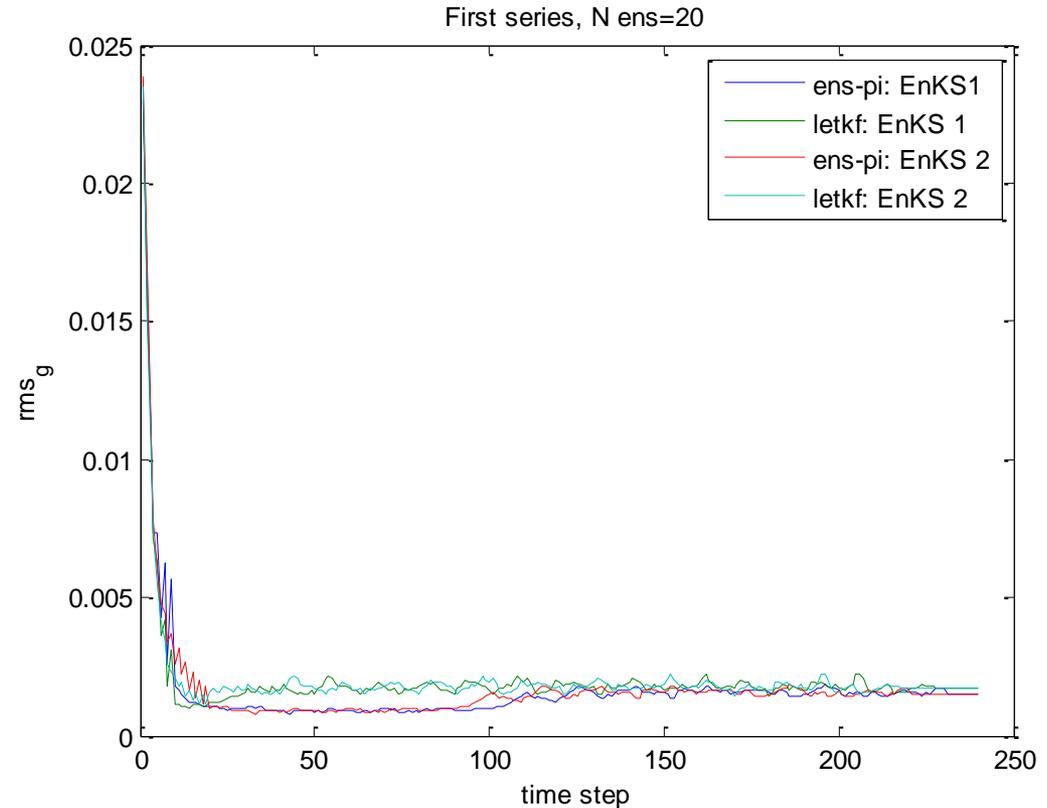
Тестовые расчеты с 1-мерной моделью переноса и диффузии

Поведение по времени оцениваемого параметра $g(x,t)$, вторая серия экспериментов, nt - номер шага по времени



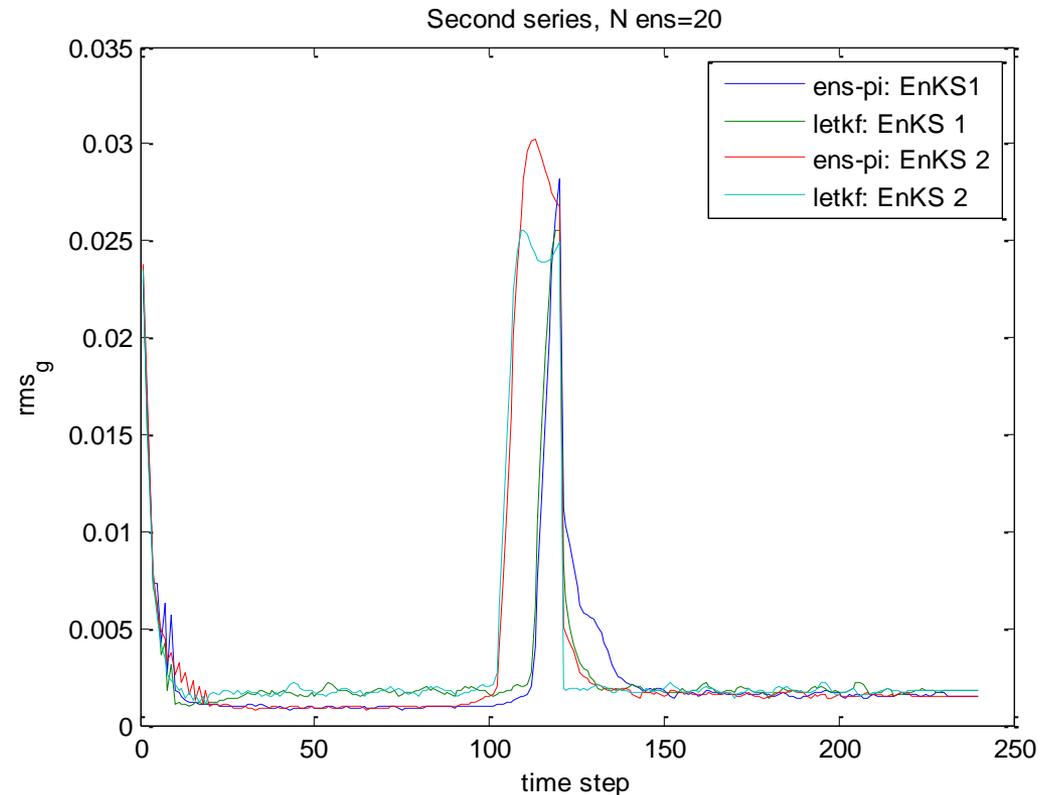
Тестовые расчеты с 1-мерной моделью переноса и диффузии

*Среднеквадратическая ошибка
оценки параметра $g(x,t)$
для детерминированного (letkf) и
стохастического (ens-pi)
алгоритмов EnKS (EnKS 1 – time
lag=10, EnKS 2 – time lag=20),
первая серия экспериментов*



Тестовые расчеты с 1-мерной моделью переноса и диффузии

*Среднеквадратическая ошибка
оценки параметра $g(x,t)$
для детерминированного (letkf) и
стохастического (ens-pi)
алгоритмов EnKS (EnKS 1 – time
lag=10, EnKS 2 – time lag=20),
вторая серия экспериментов*



Заключительные выводы

- Задача оценки состояния окружающей среды по данным наблюдений в настоящее время решается с помощью систем усвоения данных. При этом используются модели распространения примеси в атмосфере и метеорологические поля скорости ветра, температуры и т.д. для моделирования переноса и диффузии пассивных примесей в атмосфере.
- В качестве математической постановки задачи все чаще используется ансамблевый фильтр Калмана либо обобщение этого алгоритма – ансамблевое сглаживание.
- На конечный результат влияет большое количество факторов: используемые метеорологические данные, модель переноса и диффузии, параметризации приземного слоя атмосферы, методика усвоения данных, используемые данные наблюдений.
- Задача оценка потоков парниковых газов с поверхности Земли в настоящее время решается с помощью систем усвоения данных. При этом используются модели распространения примеси в атмосфере и метеорологические поля скорости ветра, температуры и т.д. для моделирования переноса и диффузии парниковых газов. Следует подчеркнуть, что реализация такой масштабной задачи требует объединения усилий как правило нескольких научных коллективов.

Спасибо за внимание!

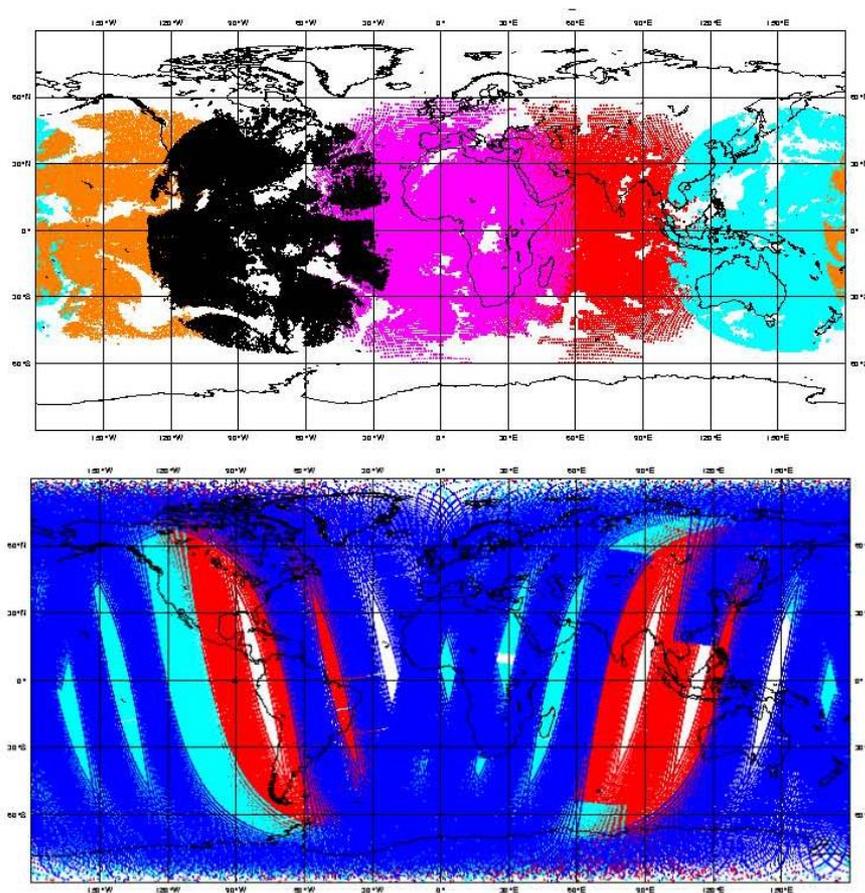


Figure 1: Typical data coverage provided by the Geostationary constellation (top): GOES-W/E (orange/black), Meteosat-7/5 (pink/red) and GMS-5/GOES-9 (cyan). Bottom plot displays the LEO constellation from the NOAA satellites (NOAA-15 in red, NOAA-16 in cyan, NOAA-17 in blue).