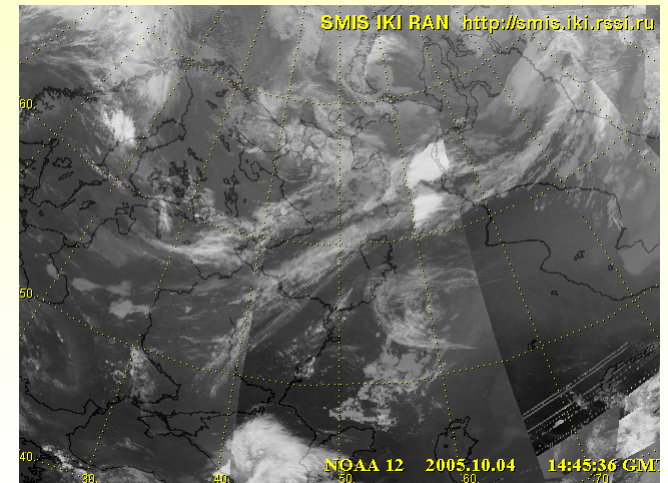
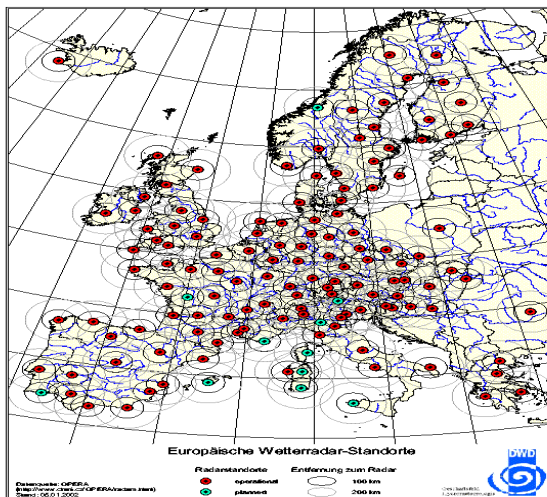


*Institute of Computational Technologies Russian Academy of  
Sciences, Siberian Branch, Ac. Lavrentjev Ave., 6,  
Novosibirsk, 630090, Russia*

[klimova@ict.nsc.ru](mailto:klimova@ict.nsc.ru)

*E. G. Klimova*

*Effective algorithm of data assimilation, based on  
the ensemble approach*



# *Data assimilation problem*

---

## **The known information:**

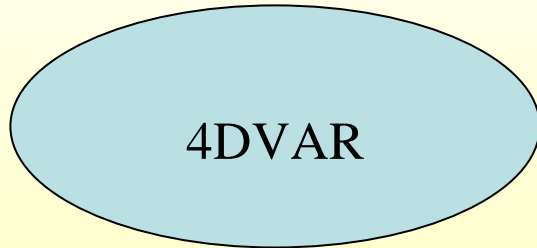
- observations;
- mathematical models;
- «the aprioristic information».

## **The basic problems:**

- necessity of work with the data of different types;
- many variables directly are not measured
- the estimation has huge dimension;
- forecast models are nonlinear.

# Data assimilation problem

---



ECMWF (1997)

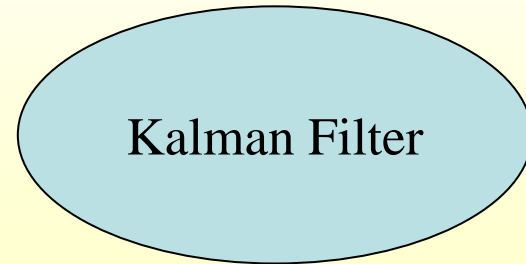
Meteo France (2000)

Met-Office (2004)

Japan (2005)

Canada (2005)

**4DVAR** → ***RRKF (reduced rank Kalman filter)***



Canada, Atmospheric Environment Service (H.Mitchel, P.L.Houtekmer):

***Ensemble Kalman Filter***

NSEP

***The local ensemble transform Kalman filter (LETKF)***

# The optimal filtration problem

---

Linearised atmospheric model:  $x_k^f = A_{k-1}x_{k-1}^a$ ,

$x_k^f$  - n-vector of forecast variables at time  $t_{k-1}$ ,

$x_{k-1}^a$  - n-vector of analysed variables at time  $t_{k-1}$ .

"True" state of the atmosphere:  $x_k^t = A_{k-1}x_{k-1}^t + \eta_{k-1}^t$ ,

$\eta_k^t$  - the rand vector of model "noise":

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

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

## *The optimal filtration problem*

---

The observations:  $y_k^0 = M_k x_k^t + \xi_k^0$ ,

$y_k^0$  - m-vector of observations at time  $t_k$ ,  $M_k$  -  $(n \times m)$  matrix,

$\xi_k^0$  - rand m-vector of observational errors:

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

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

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

The optimal filtration problem is the problem of minimization

$\min J$  on the observations at times  $t_k, (k = 0, \dots, K)$

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

# Kalman filter

---

$$\mathbf{x}_k^f = A_{k-1} \mathbf{x}_{k-1}^a;$$

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

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

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

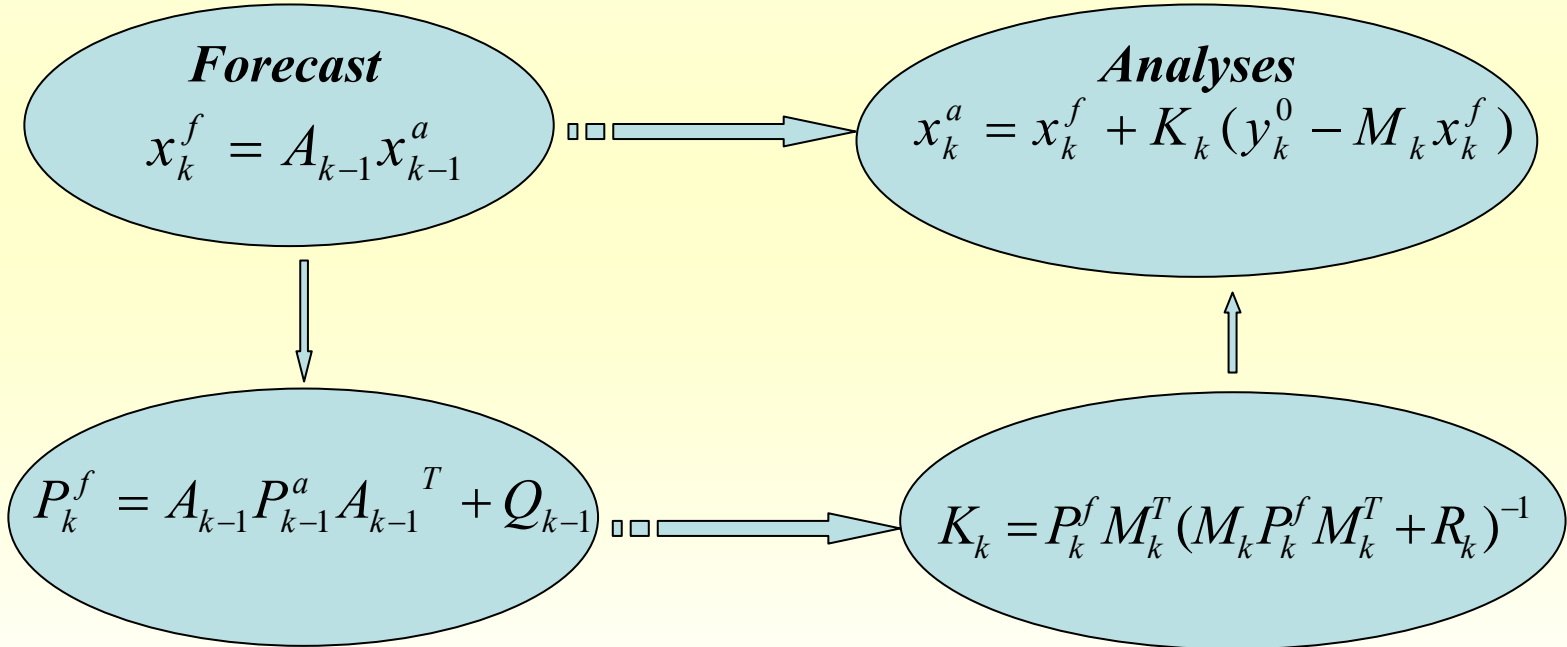
$$\mathbf{x}_k^a = \mathbf{x}_k^f + K_k (y_k^0 - M_k \mathbf{x}_k^f);$$

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

$$P_k^f = E(\mathbf{x}_k^f - \mathbf{x}_k^t)(\mathbf{x}_k^f - \mathbf{x}_k^t)^T; P_k^a = E(\mathbf{x}_k^a - \mathbf{x}_k^t)(\mathbf{x}_k^a - \mathbf{x}_k^t)^T.$$

# Kalman Filter

**Observations**  
 $y_k^0$



# Ensemble Kalman filter

**Ensembles of initial fields and forecasts:**

$$x_f^{0(i)} = \hat{x} + \Delta x_0^{(i)}, \quad i = 1, \dots, N$$

**Estimation of covariance matrix:**

$$x_{k+1}^{f(i)} = A_k(x_k^{a(i)}) + \xi_k^{(i)},$$

$$P_k^f M_k^T = \frac{1}{N-1} \sum_{i=1}^N (x_k^{f(i)} - \overline{x_k^{f(i)}})(M_k x_k^{f(i)} - \overline{M_k x_k^{f(i)}})^T,$$

$$M_k P_k^f M_k^T = \frac{1}{N-1} \sum_{i=1}^N (M_k x_k^{f(i)} - \overline{M_k x_k^{f(i)}})(M_k x_k^{f(i)} - \overline{M_k x_k^{f(i)}})^T,$$

**Ensemble of «analyses»:**

$$K_k = P_k^f M_k^T (M_k P_k^f M_k^T + R_k)^{-1},$$

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

$$y_k^{0(i)} = y_k^0 + r_k^{(i)}$$

$$\overline{x_{k+1}} = \frac{1}{N} \sum_{i=1}^N x_{k+1}^{(i)}$$

## Extended Kalman filter

---

Let's consider the forecast model  $\mathbf{x}_{k+1} = \mathbf{A}_k(\mathbf{x}_k)$

The algorithm of extended Kalman filter for nonlinear forecast model ( $k$  is the number of the time step) can be written in the following form:

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{A}_k(\mathbf{x}_k) + \xi_k + \mathbf{P}_{k+1} \mathbf{M}_{k+1}^T \mathbf{R}_{k+1}^{-1} (y_0 - \mathbf{M}_{k+1} \mathbf{A}_k(\mathbf{x}_k)), \\ \mathbf{P}_{k+1} &= \hat{\mathbf{P}}_{k+1} - \hat{\mathbf{P}}_{k+1} \mathbf{M}_{k+1}^T (\mathbf{M}_{k+1} \hat{\mathbf{P}}_{k+1} \mathbf{M}_{k+1}^T + \mathbf{R}_{k+1})^{-1} \mathbf{M}_{k+1} \hat{\mathbf{P}}_{k+1}, \\ \hat{\mathbf{P}}_{k+1} &= \tilde{\mathbf{A}}_k \hat{\mathbf{P}}_k \tilde{\mathbf{A}}_k^T + \mathbf{Q}_k,\end{aligned}$$

Where  $\mathbf{P}_{k+1}$  - covariance matrix of estimation error,

$\mathbf{x}_0 = \hat{\mathbf{x}}, \mathbf{P}_0 = \hat{\mathbf{P}}, \tilde{\mathbf{A}}_k$  - the operator, linearised about the previous time step,

$\mathbf{R}_{k+1}, \mathbf{Q}_k$  - covariance matrices of observations errors and model noise  $\xi_k$

# Ensemble $\pi$ -algorithm

---

*Ensemble of initial fields and forecasts:*

$$\mathbf{x}_0^{(i)} = \mathbf{x}_0 + \Delta \mathbf{x}_0^{(i)}, \quad i = 1, \dots, N$$

$$\mathbf{x}_{k+1}^{f(i)} = A_k(\mathbf{x}_k^{a(i)}) + \boldsymbol{\zeta}_k^{(i)} + P^{k+1} M^T R^{-1} (y_0 - M A_k(\mathbf{x}_k^{(i)}) - M \boldsymbol{\zeta}_k^{(i)}),$$

$$\overline{\mathbf{x}}_{k+1} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_{k+1}^{(i)}$$

## *Ensemble $\pi$ - algorithm*

---

$$1.) \quad f_k^{(i)} = A_k(x_k^{(i)}) + \xi_k^{(i)} - \overline{A_k(x_k^{(i)})}, \quad F = \{f_k^{(i)}\} \quad 2.) \quad \Pi^T = (C + 0,25I)^{\frac{1}{2}} - 0,25I$$

$$C = \frac{1}{N-1} F^T M^T R^{-1} M F$$

$$3.) \quad D^T = (I + \Pi^T)^{-1} F^T$$

$$D = \{dx^{(i)}, \quad i = 1, \dots, N\}$$

$$dx_{k+1}^{(i)} = x^{(i)}(t_{k+1}) - \overline{x^{(i)}(t_{k+1})}$$

$$4.) \quad (\Pi_2)_j^i = \frac{1}{N-1} dx_j^T M^T R^{-1} (y_0 - M A_k(x_k^{(i)}) - M \xi_k^{(i)})$$

$$X^{(k+1)T} = F_2^T + \Pi_2^T D^T$$

$$X^{(k+1)} = \{x_{k+1}^{(i)}, \quad i = 1, \dots, N\}$$

$$F_2^T = \{A_k(x_k^{(i)}) + \xi_k^{(i)}, \quad i = 1, \dots, N\}$$

$$5.) \quad \overline{x_{k+1}} = \frac{1}{N} \sum_{i=1}^N x_{k+1}^{(i)}$$

***$\Pi$ ,  $D$ ,  $F$ ,  $X$  – matrices of dimension  $N$   
(ensemble dimension)***

## *Ensemble $\pi$ - algorithm*

---

- The full description of the algorithm will be published in

*Klimova E.G. Data assimilation method, based on the ensemble  $\pi$  - algorithm application. - Meteorologiya I Gydrologiya, 2008, N9 in press (in Russian).*

- It is necessary to note, that the received formulas have much in common with very popular algorithm LETKF, but in too time there are essential distinctions.

*Hunt B.R. An efficient implementation of the local ensemble Kalman filter, 2006. In: [http:// arhiv.org/abs/physics/0511236](http://arhiv.org/abs/physics/0511236)*

## *Planning of adaptive observations*

---

*Accuracy of data assimilation algorithm based on Kalman filter is determined by a trace of a covariance matrix of estimation errors*

$$P_k = \frac{D^{(k)} D^{T(k)}}{N-1}.$$

*that matrix may be considered as covariance matrix of analysis errors*

$$P_k^a$$

*Matrix*

$$P_{k+1}^f = \frac{FF^T}{N-1}$$

*may be considered as covariance matrix of forecast errors*

*In Kalman filter.*

## *Planning of adaptive observations*

Let's consider columns of matrixes  $D^T$  and  $F^T$ , appropriate to ensemble of values in m-point of a grid  $d_m^T, f_m^T$  then, if to enter norm of a vector

$$\|x\|^2 = (x, x) = \sum_{i=1}^N x_i^2$$

$$\frac{1}{\|(I + \Pi^T)\|^2} \|f^T\|^2 \leq \|d^T\|^2 = (d^T, d^T) \leq \|(I + \Pi^T)^{-1}\|^2 \times \|f^T\|^2.$$

$$\|d^T\|^2 = (N - 1)p_{ii}^a, \quad \|f^T\|^2 = (N - 1)p_{ii}^f,$$

Where  $p_{ii}^a, p_{ii}^f$  - elements of matrixes  $P_k^a, P_{k+1}^f$ , accordingly. From an inequality follows, that change of an estimation error in m-th grid point after procedure of the analysis depends on  $\|(I + \Pi^T)\|$ .

## *Planning of adaptive observations*

Let's consider spectral radius as norm of a symmetric matrix . Eigen values of a matrix  $(I + \Pi^T)$  can be expressed from eigen numbers of a matrix  $C$  as follows:

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

Let's consider the sum of elements of matrix  $C$  in  $i$ -th row . In a case, when the operator of forecast model is linear,  $F=MD$  and

$$S_i = \frac{1}{N-1} \sum_{j=1}^N f_m^{(i)} f_m^{(j)} / r^2 = \frac{f_m^{(i)}}{(N-1)r^2} \sum_{j=1}^N f_m^{(j)}.$$

The sum represents average on ensemble value of the forecast error in  $m$ -th point of the grid, multiplied by  $N$ , and, in case of the linear operator of model, tends to zero if  $n \rightarrow \infty$  . Thus, matrix  $C$  has the eigen vector  $\{1, \dots, 1\}^T$  appropriate to zero eigen value.

## *Planning of adaptive observations*

For the entered norm of a matrix we shall have the following estimation:

$$\frac{1}{\{\rho(I + \Pi^T)\}^2} \leq \frac{p_{ii}^a}{p_{ii}^f} \leq 1,$$

as

$$\rho\{(I + \Pi^T)^{-1}\} = \frac{1}{\lambda_{\min}(I + \Pi^T)} = 1.$$

Hence, at realization of the analysis on the data in  $m$ -th point of a grid we shall have change of the forecast error covariance

$$\frac{1}{\lambda_{\max}(I + \Pi^T)} \leq \frac{p_{ii}^a}{p_{ii}^f} \leq 1.$$

From the eigen values of a matrix  $(I + \Pi^T)$  it is clear, that the maximal eigen values of this matrix will be at a maximum  $\lambda(C)$ .

## *Planning of adaptive observations*

Let's assume, that adaptive observation is given in some ***m-th*** pointt of a grid. The criterion of accuracy of the analysis in the given area will be  $\delta_i = \|(I + \Pi^T)\|$ , where  $\Pi^T$  it is estimated for this observation. If to calculate  $\delta_i$  for all grid point, it is possible to estimate areas with the greatest value  $\delta_i$ . The received areas will determine that region where accommodation of additional observations will give the greatest effect. At such analysis we assume, that the data used consistently, such procedure is possible if to assume, that errors of observations do not correlate.

Let's note, that in case it is supposed to have observations in a sequence of grid points  $\{m_1, \dots, m_l\}$ , and thus errors of observations do not correlate, the analysis can be carried out with the help of iterative procedure, using on one observation on each iteration. In this case

$$d^T = (I + \Pi_l^T)^{-1} \dots (I + \Pi_1^T)^{-1} f^T.$$

## *Planning of adaptive observations*

The important problem is the estimation of behaviour of errors on time after realization of the analysis on additional observation. The formula for an estimation of forecast error covariances in case of the linear operator of model is:

$$P^f = \frac{1}{N-1} MD(I + \Pi^T)^{-2} D^T M.$$

$\pi$  - algorithm has that property, that the matrix  $\Pi$  does not depend on a grid point, therefore the formula can be calculated consistently, all over again having calculated a matrix  $MD$ , and then an estimation of covariances.

On ensemble of forecast errors it is possible to estimate area  $D^{(f)}$ , in which errors achieve the greatest sizes, and area  $D^{(a)}$  in which it is the most expedient to place observation from the point of view of quality of the analysis. Then the region in which it is necessary to place observation for reduction of error growth can be estimated through

$$D^{(a)} \cap D^{(f)}.$$

## *Planning of adaptive observations*

In case of the linear operator of observations the trace of a matrix  $C$  looks like

$$\text{tr } C = \frac{1}{N-1} \sum_{i=1}^N (f_m^i)^2 / r^2.$$

In a case of linearised operator of model  $f^i = Md^i$  and, hence,

$$\text{tr } C = p_{mm}^f / r^2.$$

So,  $\lambda_{\max}(C)$  characterizes reduction of an error of the analysis at use of observation in the appropriate grid point. At the same time, the trace of a matrix  $C$  characterizes an error of the forecast. Thus, the estimation of areas for additional observations depends on accuracy of the information about the forecast errors which we have.

## *Numerical experiments*

*Numerical experiments were carried out for hemispheric barotropic quazi-geostrophical model :*

$$\frac{d(\Omega + l)}{dt} = 0,$$

$$\frac{d}{dt} = \frac{\partial}{\partial t} + u_g \frac{1}{a} \frac{\partial}{\partial \lambda} + v_g \frac{1}{a \cos(\varphi)} \frac{\partial}{\partial \theta},$$

$$\Omega = \frac{g}{l_0} \left\{ \frac{1}{a^2} \frac{\partial^2 h}{\partial \lambda^2} + \frac{1}{a \cos(\varphi)} \frac{\partial}{\partial \varphi} \frac{1}{a \cos(\varphi)} \frac{\partial h}{\partial \varphi} \right\} + l(\varphi),$$

$$u_g = -\frac{g}{l_0} \frac{1}{a \cos(\phi)} \frac{\partial h}{\partial \phi}, \quad v_g = \frac{g}{l_0} \frac{1}{a} \frac{\partial h}{\partial \lambda}$$

## *Numerical experiments*

*The equation was solved in area*

$$\{0 \leq \lambda \leq 2\pi, \varphi_1 \leq \varphi \leq \varphi_2\}, \varphi_1 = 25^\circ, \varphi_2 = 75^\circ$$

*The following finite- difference scheme was considered (the grid was latitude-longitudinal,  $\Delta\lambda = \Delta\varphi = 2,5^\circ$  ):*

$$\Omega_{i,j}^{n+1} = \Omega_{i,j}^{n-1} - \frac{\Delta t}{2a^2 \Delta\lambda \Delta\varphi \cos(\varphi)} \{ (z_{i,j+1}^n - z_{i,j-1}^n) \times (\Omega_{i+1,j}^n - \Omega_{i-1,j}^n) - (z_{i+1,j}^n - z_{i-1,j}^n) \times (\Omega_{i,j+1}^n - \Omega_{i,j-1}^n) \}, n > 1,$$

$$\Omega_{i,j}^{n+1} = \Omega_{i,j}^n - \frac{\Delta t}{4a^2 \Delta\lambda \Delta\varphi \cos(\varphi)} \{ (z_{i,j+1}^n - z_{i,j-1}^n) \times (\Omega_{i+1,j}^n - \Omega_{i-1,j}^n) - (z_{i+1,j}^n - z_{i-1,j}^n) \times (\Omega_{i,j+1}^n - \Omega_{i,j-1}^n) \}, n = 1.$$

*Time step  $\Delta t = 15$  min, the field of the objective analysis of 500 mb height was an initial value.*

## *Numerical experiments*

*We considered, that distribution of an errors at the initial moment of time is known. Initial errors were modeled in area*

$$D = \{(i - i_1)^2 + (j - j_1)^2 \leq 25\} \cup \{(i - i_2)^2 + (j - j_2)^2 \leq 25\}$$

$$(i_1, j_1) = (70, 20), (i_2, j_2) = (100, 15)$$

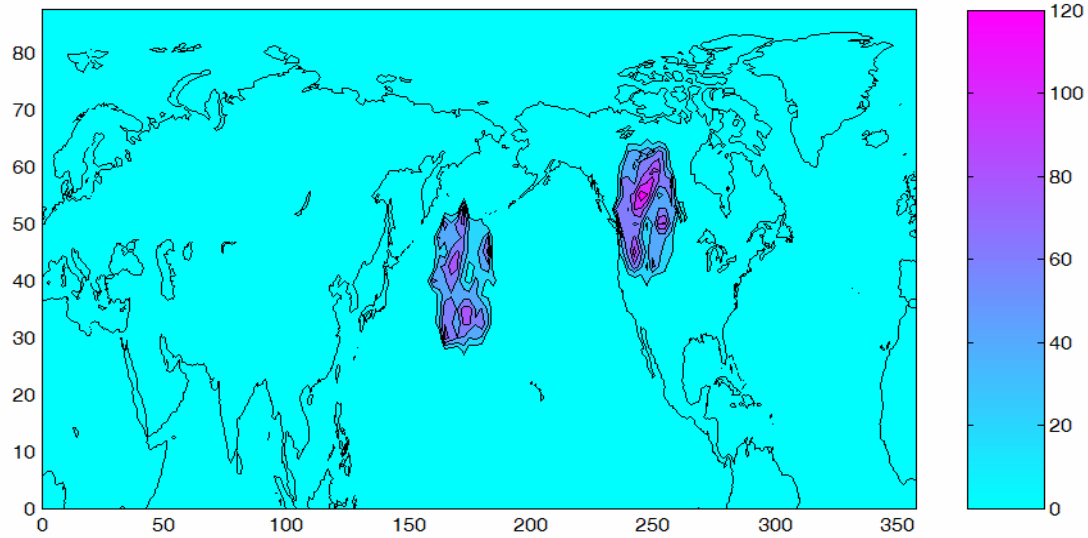
*Ensemble of random errors of an initial field*  $\{\delta_n, n = 1, \dots, N\}$

*was set so that*

$$\text{cov}(l_1, l_2) = \sigma_0^2 e^{-0,5(r_{12}/L)^2},$$

*where*  $r_{12}$  *- distance between points*  $l_1$  *and*  $l_2$ ,  $\sigma_0 = 20$  *m,*  $L = 600$  *km .*

## *Numerical experiments*



*RMS – error,  $t=0$  hour*

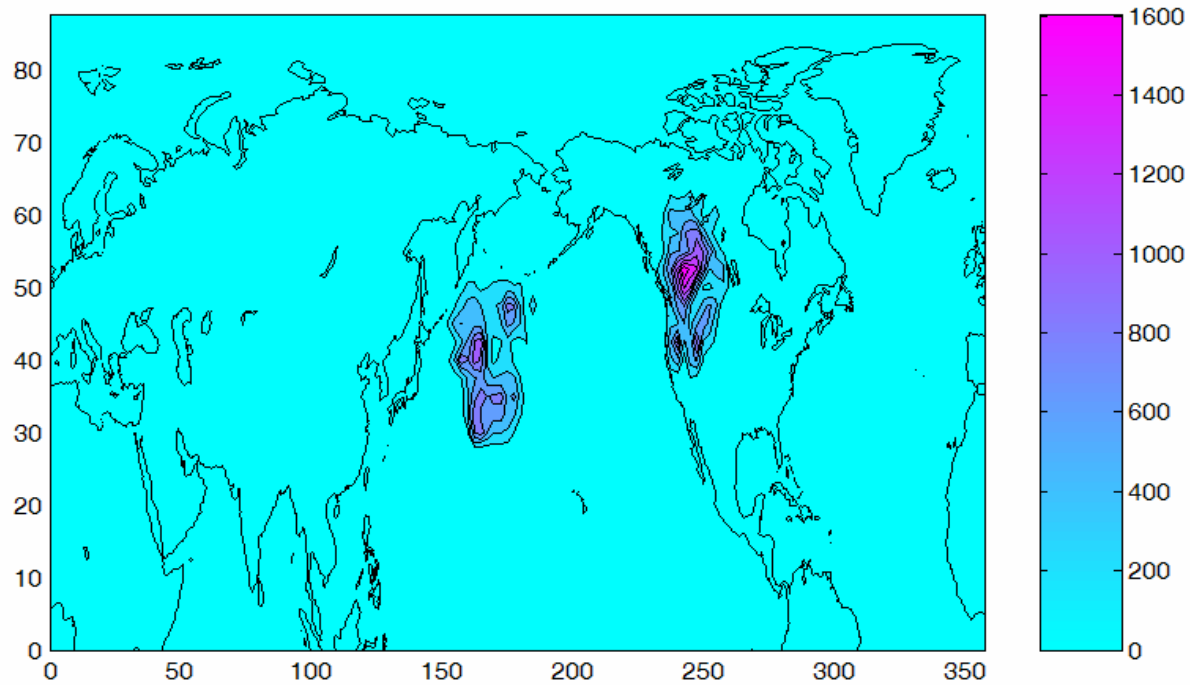
## *Numerical experiments*

*On 50 initial fields forecasts on 12 h were calculated and average value was*

$$\bar{z} = \frac{1}{N} \sum_{n=1}^N z^{(n)}$$

*For all grid points of considered area value  $S_{ij} = \rho(C_{ij})$ , was calculated under supposition, that elements of a matrix  $C_{ij}$  is in unit  $(i, j)$ .*

## *Numerical experiments*



*Fig.1*

$$S_{ij} = \rho(C_{ij}), \quad (\text{forecast on 12 hour}).$$

## Numerical experiments

In the following experiment the forecast from 12 h was calculated till 24 h. In linearised model the advection was carried out with the help of average on ensemble value of geostrophic wind. On fig. 2a, 2b function  $T_{ij} = \text{tr}C_{ij}$  for 12 and 24 hours are presented. As it was specified earlier, the trace of a matrix  $C_{ij}$  characterizes an error of the forecast in unit  $(i, j)$ .

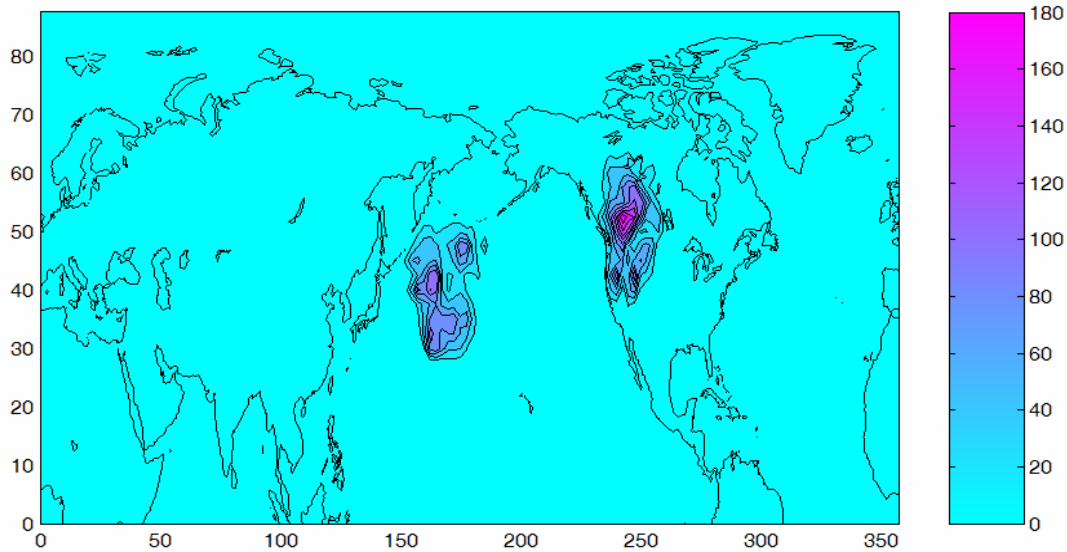
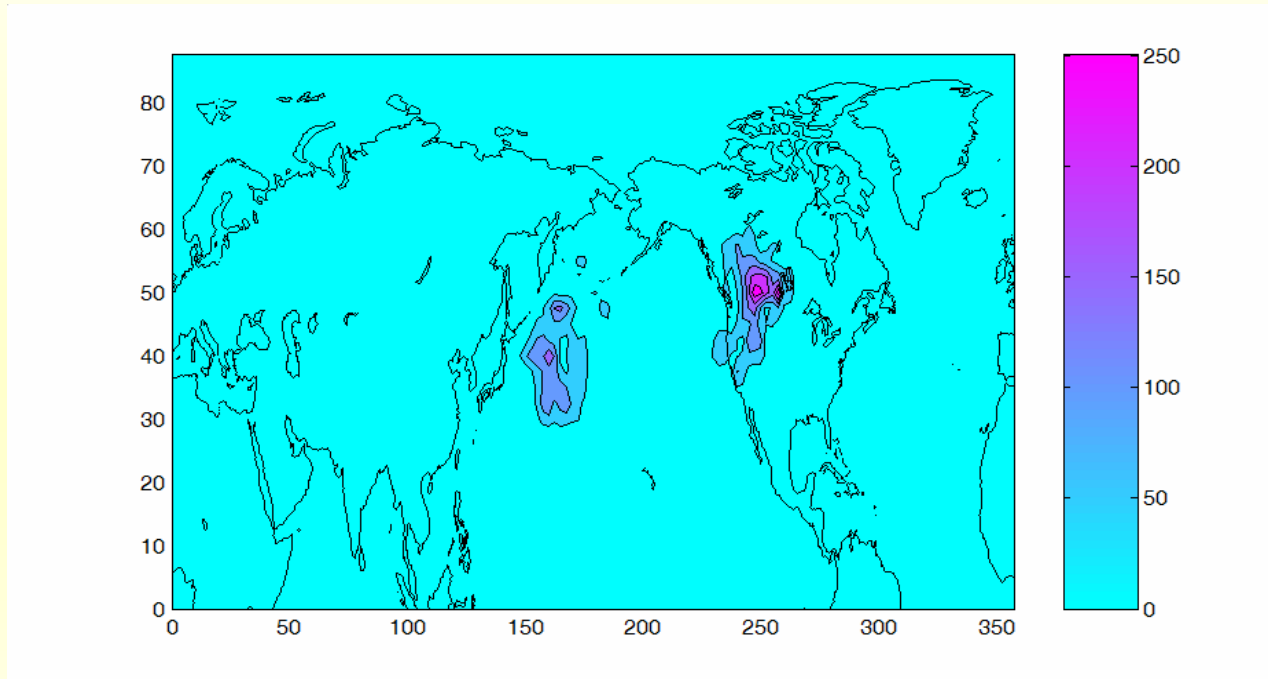


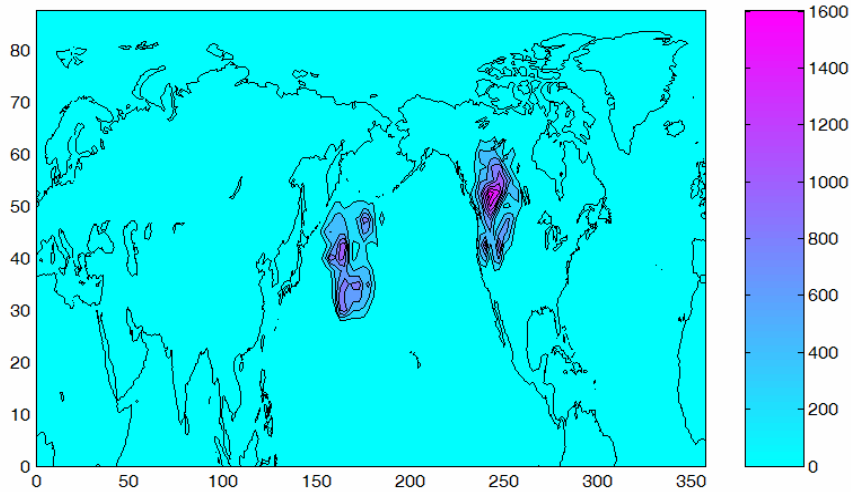
Fig. 2a

## *Numerical experiments*



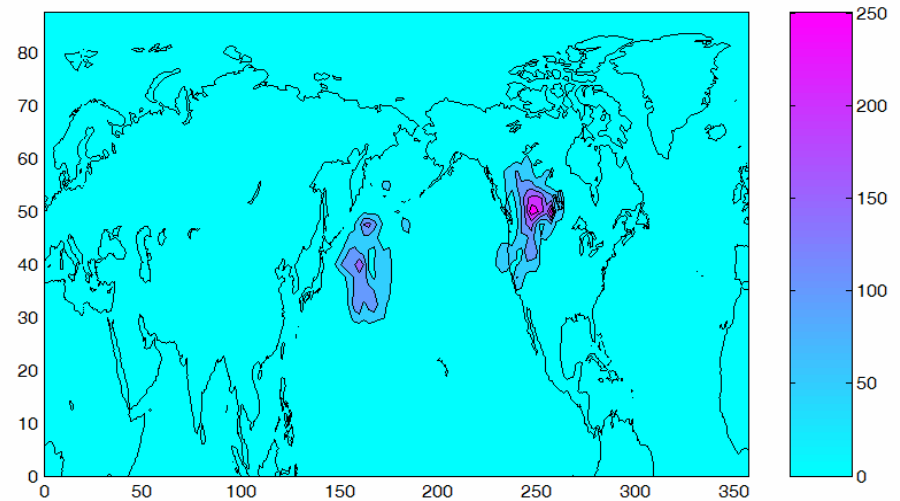
*Fig. 2b*

# Numerical experiments



$$T_{ij} = \text{tr} C_{ij} \quad t=24 \text{ h}$$

$$S_{ij} = \rho(C_{ij}) \quad t=12 \text{ h}$$



Thanks for attention!

