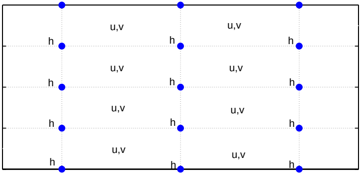
Model error and observation error

Tijana Janjić

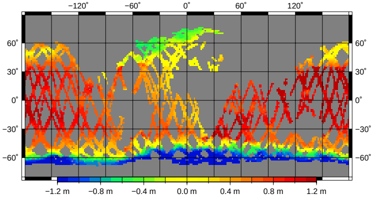
January 15, 2013

Data assimilation algorithm combine forecast and observations to produce the best analysis

$\mathbf{w}_k^f$



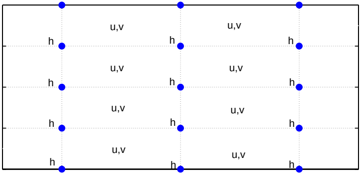
$\mathbf{y}_k^o$



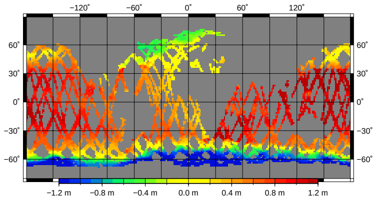
Analysis systems are dependent on appropriate statistics for observation and background errors.

Data assimilation algorithm combine forecast and observations to produce the best analysis

$\mathbf{w}_k^f$



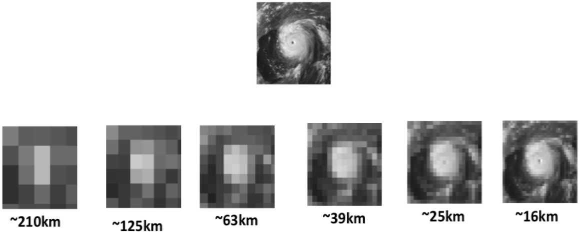
$\mathbf{y}_k^o$



Analysis systems are dependent on appropriate statistics for observation and background errors.

Our goal: Best analysis for a prediction.

One major contributor to the forecast uncertainty is the model error.



Model resolution (slide ECMWF)

Model error

Unfortunately model error statistics are not perfectly known and their determination remains a major challenge in assimilation systems.

Reasons behind the model error:

- accuracy of numerical schemes
- unrepresented subgrid scale processes
- inaccurate forcing and boundary conditions
- representation of orography as well as parametrisation uncertainty.

Model error

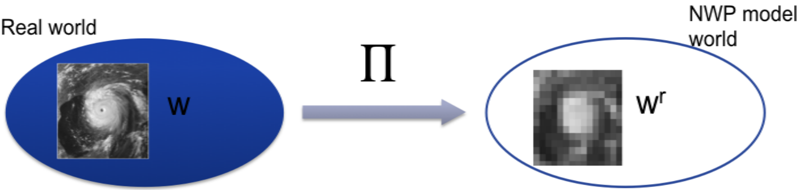
Unfortunately model error statistics are not perfectly known and their determination remains a major challenge in assimilation systems.

Reasons behind the model error:

- accuracy of numerical schemes
- unrepresented subgrid scale processes
- inaccurate forcing and boundary conditions
- representation of orography as well as parametrisation uncertainty.

Model error statistics produced by use of multiple physics packages,inclusion of stochastic kinetic energy backscatter scheme, parameter variations, as well as use of deterministic stochastic dynamical models (Berner et al. 2011).

Model Error

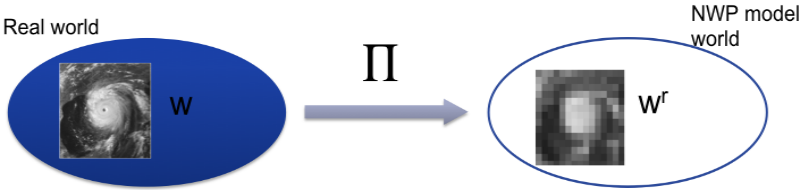


from time k to time k+1 atmosphere evolves without us knowing perfectly time propagator, F^c.

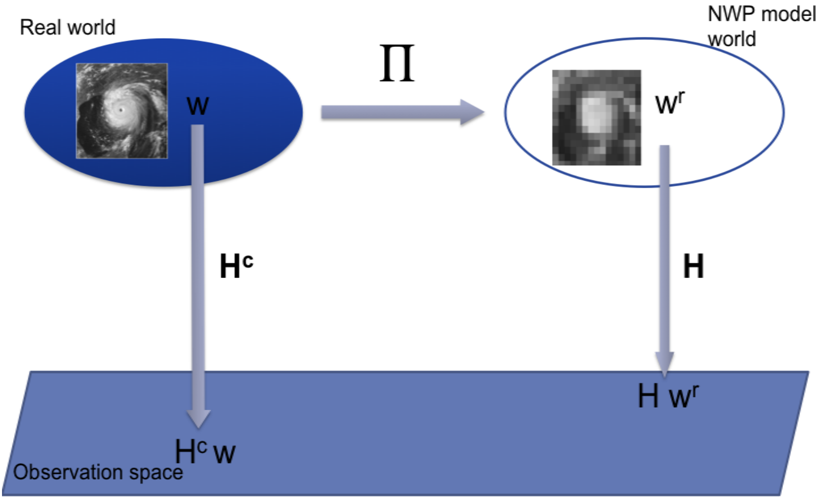
from time k to time k+1 numerical model, F, propagates w^r

Model error is the difference: $\Pi F^c(w) - F(w^r)$.

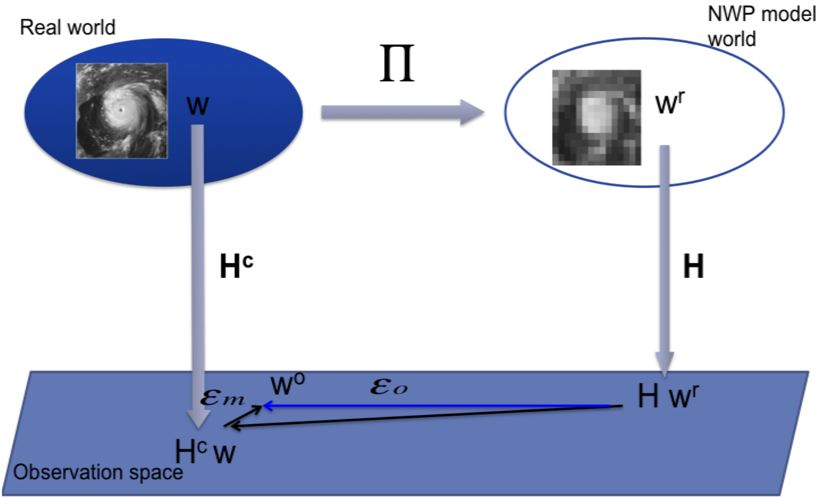
Observation Error



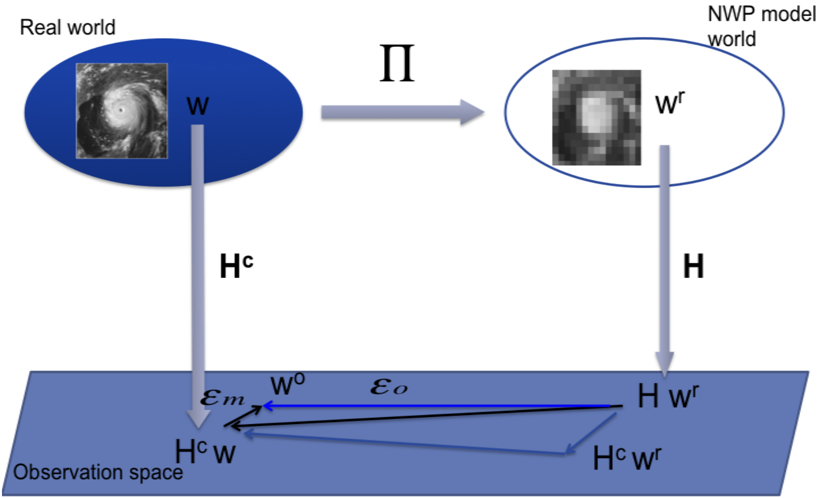
Observation Error



Observation Error



Observation Error



Atmospheric Data Assimilation

The state $\mathbf{w} \equiv \mathbf{w}(\mathbf{x}, t)$ of the atmosphere at time t_k:

$$\mathbf{w}(\mathbf{x}, t_k) \equiv \begin{bmatrix} w_1(\mathbf{x}, t_k) \\ \vdots \\ w_q(\mathbf{x}, t_k) \end{bmatrix}$$

where $w_i : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}, \forall i = 1, \ldots, q, \ w_i \in \mathcal{B}$.
($\mathcal{B}$ is a vector space of scalar valued, continuous functions.)

Atmospheric Data Assimilation

The state $\mathbf{w} \equiv \mathbf{w}(\mathbf{x}, t)$ of the atmosphere at time t_k:

$$\mathbf{w}(\mathbf{x}, t_k) \equiv \begin{bmatrix} w_1(\mathbf{x}, t_k) \\ \vdots \\ w_q(\mathbf{x}, t_k) \end{bmatrix}$$

where $w_i : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}, \forall i = 1, \ldots, q, \; w_i \in \mathcal{B}$.
($\mathcal{B}$ is a vector space of scalar valued, continuous functions.)

Discrete problem: Find an estimate of some projection $\mathbf{\Pi w}$ of $\mathbf{w}$ on the space of the dynamical model.

$$\mathbf{\Pi w}(\mathbf{x}, t_k) \equiv \begin{bmatrix} \Pi w_1(\mathbf{x}, t_k) \\ \vdots \\ \Pi w_q(\mathbf{x}, t_k) \end{bmatrix}$$

where $\Pi w_i : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}, \forall i = 1, \ldots, q$.
$\Pi w_i(\mathbf{x}, t_k) \in \mathcal{B}_N$, where $\mathcal{B}_N$ is an N-dimensional subspace of $\mathcal{B}$.

Observation error

ϵ_k^o consists of measurement error and representativeness error. It can be divided into three parts:

$$\epsilon_k^o = \epsilon_k' + \epsilon_k'' + \epsilon_k^m$$

where

$$\begin{aligned} \epsilon_k' &\equiv \mathbf{H}_k^c \mathbf{w}(\cdot, t_k) - \mathbf{H}_k^c \mathbf{\Pi} \mathbf{w}(\cdot, t_k) \\ &= \mathbf{H}_k^c (\mathbf{I} - \mathbf{\Pi}) \mathbf{w}(\cdot, t_k) \end{aligned}$$

ϵ_k' – will be called *error due to unresolved scales*.

$$\begin{aligned} \epsilon_k'' &\equiv \mathbf{H}_k^c \mathbf{\Pi} \mathbf{w}(\cdot, t_k) - \mathbf{H}_k \mathbf{\Pi} \mathbf{w}(\cdot, t_k) \\ &= [\mathbf{H}_k^c - \mathbf{H}_k] \mathbf{\Pi} \mathbf{w}(\cdot, t_k) \end{aligned}$$

ϵ_k'' – will be called *forward interpolation error*.

Representativeness error (Lorenc 1986; Daley 1993; Cohn 1997)

- representativeness error introduces spatial correlations in the observational error
- and it is state and time dependent (Janjic 2001, Janjic and Cohn 2006)

Representativeness error (Lorenc 1986; Daley 1993; Cohn 1997)

- representativeness error introduces spatial correlations in the observational error
- and it is state and time dependent (Janjic 2001, Janjic and Cohn 2006)
- difficult to estimate
- important for optimal use of observations, since it tell us how observations are to be provided to best adopt to model resolution
- for variable model resolutions needs to be scale adaptive
- it depends on the observation type

Representativeness error (Lorenc 1986; Daley 1993; Cohn 1997)

- representativeness error introduces spatial correlations in the observational error
- and it is state and time dependent (Janjic 2001, Janjic and Cohn 2006)
- difficult to estimate
- important for optimal use of observations, since it tell us how observations are to be provided to best adopt to model resolution
- for variable model resolutions needs to be scale adaptive
- it depends on the observation type
- Example: for 40×40 km Radiosonde/Dropsonde wind observation, observational error < 0.5 m/s and Assigned error: $2 - 3$ m/s

2 Estimation of $\mathbf{\Pi w}(\mathbf{x}, t)$

$$w(\mathbf{x}, t) = w^r(\mathbf{x}, t) + w^u(\mathbf{x}, t),$$

with

$$w^r(\mathbf{x}, t) \equiv \Pi w(\mathbf{x}, t)\ , \qquad w^u(\mathbf{x}, t) \equiv (I - \Pi)\, w(\mathbf{x}, t).$$

2 Estimation of $\mathbf{\Pi w}(\mathbf{x}, t)$

$$w(\mathbf{x}, t) = w^r(\mathbf{x}, t) + w^u(\mathbf{x}, t),$$

with

$$w^r(\mathbf{x}, t) \equiv \Pi w(\mathbf{x}, t) \ , \qquad w^u(\mathbf{x}, t) \equiv (I - \Pi)\, w(\mathbf{x}, t).$$

Dynamics:

$$w^r(\mathbf{x}, t_{k+1}) = F^r_{k+1,k} w^r(\mathbf{x}, t_k),$$

Suppose $w^u(\mathbf{x}, t_k)$ satisfies the dynamics

$$w^u(\mathbf{x}, t_{k+1}) = F^u_{k+1,k} w^u(\mathbf{x}, t_k) + F^{sl}_{k+1,k} w^r(\mathbf{x}, t_k).$$

Observations:
Assume $\mathbf{H}_k = \mathbf{H}_k^c|_{\mathcal{B}_N}$ (i.e., $\boldsymbol{\epsilon}_k'' = 0$),

$$\mathbf{w}_k^o = \mathbf{H}_k^c w^r(\cdot, t_k) + \boldsymbol{\epsilon}_k^o,$$

where $\boldsymbol{\epsilon}_k^o$ is

$$\begin{aligned} \boldsymbol{\epsilon}_k^o &= \mathbf{H}_k^c w^u(\cdot, t_k) + \boldsymbol{\epsilon}_k^m \\ &= \boldsymbol{\epsilon}_k' + \boldsymbol{\epsilon}_k^m. \end{aligned}$$

A Kalman filter method

Augmented vector [Janjic 2001, Janjic and Cohn 2006]

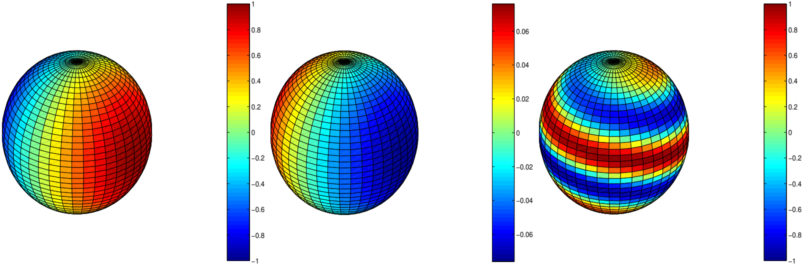
$$\mathbf{w}(\mathbf{x}, t_k) \equiv \begin{bmatrix} w^r(\mathbf{x}, t_k) \\ w^u(\mathbf{x}, t_k) \end{bmatrix},$$

$$\begin{bmatrix} w^r \\ w^u \end{bmatrix} (\mathbf{x}, t_{k+1}) = \begin{bmatrix} F^r_{k+1,k} & 0 \\ F^{ur}_{k+1,k} & F^u_{k+1,k} \end{bmatrix} \begin{bmatrix} w^r \\ w^u \end{bmatrix} (\mathbf{x}, t_k),$$

with observations given by

$$\mathbf{w}^o_k = \begin{bmatrix} \mathbf{H}^c_k & \mathbf{H}^c_k \end{bmatrix} \begin{bmatrix} w^r(\cdot, t_k) \\ w^u(\cdot, t_k) \end{bmatrix} + \boldsymbol{\epsilon}^m_k.$$

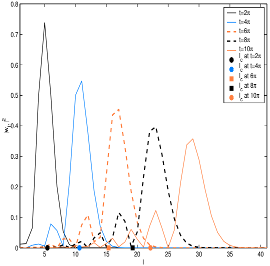
This way the correct equations can be derived to take into account unresolved scales.
We would require estimates of correlation between resolved and unresolved scales as well as an estimates of the unresolved covariance at observation points.



Initial state

Left: The state $w^r(\lambda, \varphi, t)$ which is being estimated.
Right: The full state $w(\lambda, \varphi, t)$ from which the observations are taken.

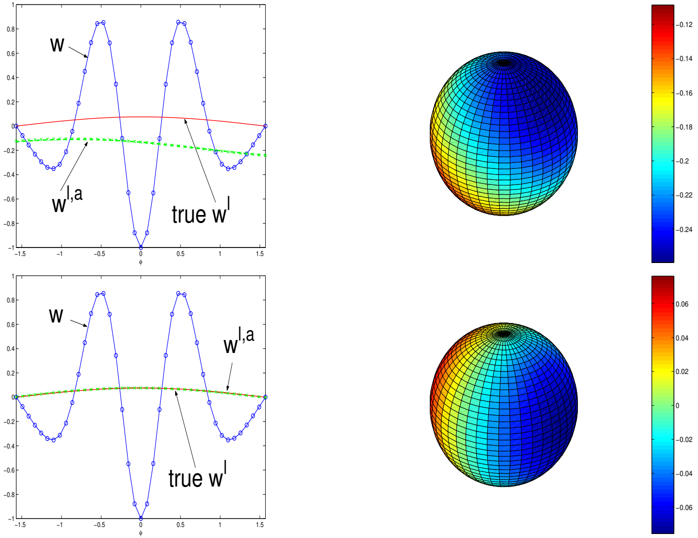
A method for estimation of time varying variance W^{uu} (Janjic 2001)



$$\sum_{l=l_c(t)-T_1}^{l_c(t)+T_2} \sum_{p=l_c(t)-T_1}^{l_c(t)+T_2} \sum_{m=-1}^{1} \sum_{q=-1}^{1} \langle \widehat{w}_{lm}(t)\widehat{w}_{pq}(t)\rangle \, Y_l^m(\lambda,\varphi) Y_p^q(\lambda,\varphi).$$

The characteristic wavenumber will be calculated using the definition

$$l_c(t)^2(l_c(t)+1)^2 \equiv \frac{\int_0^{2\pi}\int_{-\pi/2}^{\pi/2}\langle \Delta w^u(\lambda,\varphi,t)\Delta w^u(\lambda,\varphi,t)\rangle \cos\varphi \, d\varphi \, d\lambda}{\int_0^{2\pi}\int_{-\pi/2}^{\pi/2}\langle w^u(\lambda,\varphi,t) w^u(\lambda,\varphi,t)\rangle \cos\varphi \, d\varphi \, d\lambda}.$$



Cross section at $\lambda = \pi$ and estimate of $w^r(\lambda, \varphi, t)$.
Upper: $\mathbf{H}^c_{2k}[\mathbf{H}^c_{1k} W^{uu}_k(\cdot,\cdot)]^T$ approximated by zero.
Lower: $\mathbf{H}^c_{2k}[\mathbf{H}^c_{1k} W^{uu}_k(\cdot,\cdot)]^T$ approximated by adaptive method.

Conclusion

- Data assimilation algorithms require us to specify the statistical properties of the observation and model error.
- Both of these errors depend on the state of the atmosphere.
- Since we are searching for the best estimate for the scales that our model can represent,
- the unresolved scales are part of the model error as well as observation error.