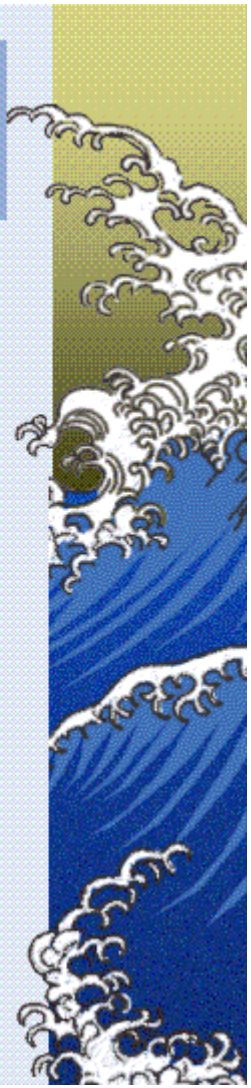
# The observations in the variational cost function : exercice

## Problematic of this exercice

We study in this exercice a simple 1-D situation : an analysis with two measurements done by a termometer in a room, knowing a background value of the temperature of the room.

We want to illustrate the importance of a good knowledge of the observation errors statistics : variances and correlations.

Since we don't carry horizontal errors correlations in the Arpege analysis, we show in this exercice how it is important to thin the data when they are correlated and at too high resolution.



# The observations in the variational cost function : exercice

Let **x** be the control variable of the minimisation problem

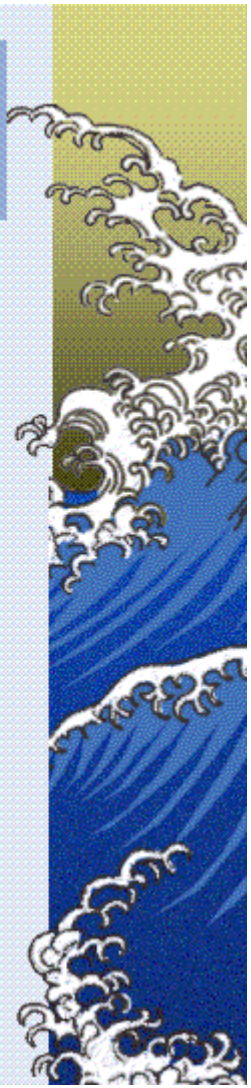
The available observations for a given time period are gathered into an observation vector **y**

To use it in an analysis, it is necessary to compare the observations with the state vector **x**

This is done through the use of a function from model state space to observation space called an observation operator, denoted by **H** generating values **H(x)**

**H** is a collection of interpolation operators from the model discretisation to the observation points, and conversions from model variables to the observed parameters

We call **y-H(x)** the vector of departures at the observation points. When calculated with background $\mathbf{x}_b$ it is called innovations, and with the analysis $\mathbf{x}_a$, analysis residuals



# The observations in the variational cost function : exercice

General hypotheses and notations :

$\mathbf{X_b}$ background model state

$\mathbf{X_a}$ analysis model state

**y** vector of observations

**H** observation operator

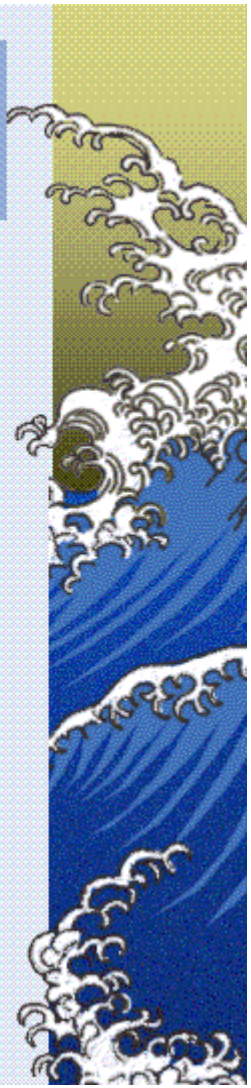
**B (or P)** covariance matrix of the background errors

**R** covariance matrix of observation errors

**B** and **R** positive definite matrices

Unbiased and uncorrelated errors

Optimal analysis : we look for an analysis state which is as close as possible to the true state in a rms sens ⇒ least-squares analysis equations



# The observations in the variational cost function : exercice

The optimal least-square estimator (or BLUE analysis) is defined by :

$$x_a = x_b + K(y - H[x_b])$$

$$K = BH^T(HBH^T+R)^{-1}$$

Where the linear operator K is called gain, or weight matrix, of the analysis

The BLUE analysis is equivalently obtained as a solution to the variational problem :

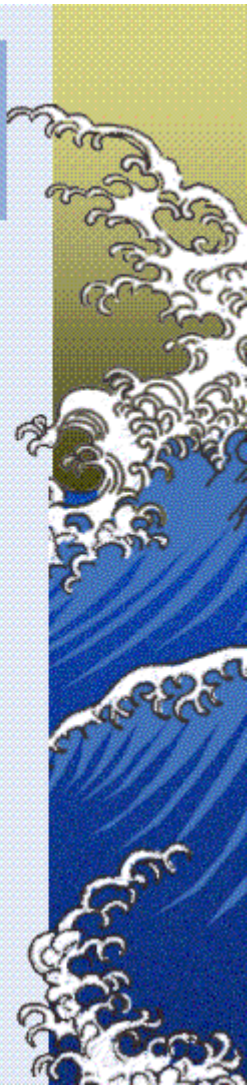
$$X_a = \min J$$

$$J(x) = (x-x_b)^T B^{-1}(x-x_b) + (y - H[x])^T R^{-1}(y - H[x])$$

$$= \quad J_b(x) \quad + \quad J_o(x)$$

Where J is called the cost function of the analysis, $J_b$ is the background term and $J_o$ is the observation term

Exercice n°1 : show this equivalence



# The observations in the variational cost function : exercice

The solution of the minimisation problem is given by $\nabla J(x_a)=0$

$$\nabla J(x_a) = 0 = 2B^{-1}(x_a-x_b) - 2H^TR^{-1}(y - H[x_a])$$

Since H is assumed to be linear, we have :

$$y - H(x_a) = y - H(x_a+ x_b- x_b) = y - H(x_b) - H(x_a-x_b)$$

then

$$0 = B^{-1}(x_a-x_b) - H^TR^{-1}(y - H[x_b]) - H^TR^{-1}H(x_a-x_b)$$

$$(x_a-x_b) = (B^{-1} + H^TR^{-1}H)^{-1} H^TR^{-1} (y - H[x_b])$$

With

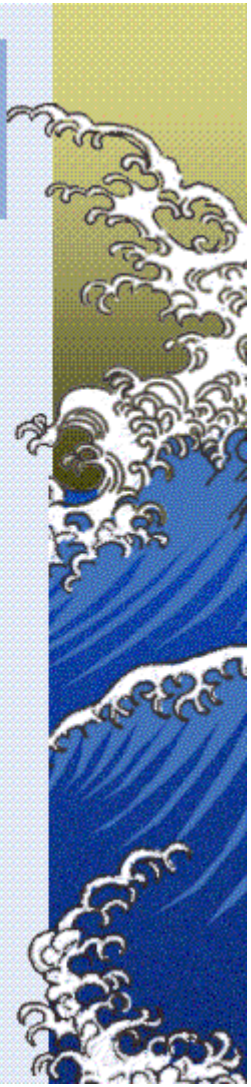
$$(B^{-1} + H^TR^{-1}H)^{-1} H^TR^{-1} = BH^T(HBH^T+R)^{-1}$$

We obtain :

$$x_a = x_b + K(y - H[x_b])$$

$$K = BH^T(HBH^T+R)^{-1}$$

Where the linear operator K is called gain, or weight matrix, of the analysis



# The observations in the variational cost function : exercice

We have one termometer in a room. We know the standard deviation of the error of this instrument : $\sigma$**o**.

We have done, starting from a previous analysis, a forecast of the temperature in this room. We know that the forecast error standard deviation is $\sigma$**b**.

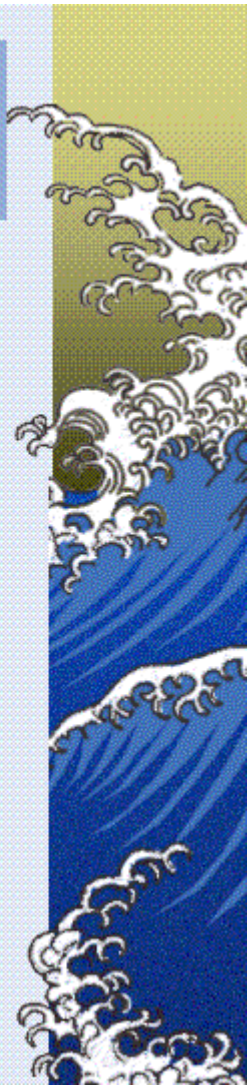
We do a measurement **to** of the temperature. We know the value **tb** of the forecast.

Exercice n°2 : compute the analysed temperature ta starting from the *optimal interpolation* formulation

Exercice n°3 : compute the analysed temperature ta starting from the *variational* formulation

Numerical application : $\sigma o = \sigma b = 1°K$

$tb = 20°K$ and $to = 24°K$



# The observations in the variational cost function : exercice

We have one termometer in a room. We know the standard deviation of the error of this instrument : $\sigma$**o**. We have done, starting from a previous analysis, a forecast of the temperature in this room. We know that the forecast error standard deviation is $\sigma$**b**.

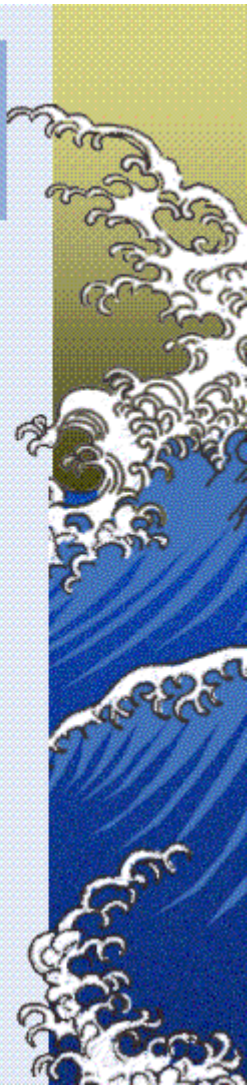
We do 2 measurements **t1 and t2** of the temperature. We know the value **tb** of the forecast. We considere these two measurements as being done by the same instrument, at the same time, and not correlated.

Exercice n°4 : compute the analysed temperature ta starting from the *optimal interpolation* formulation

Exercice n°5 : compute the analysed temperature ta starting from the *variational* formulation

Numerical application : $\sigma o = \sigma b = 1°K$

$tb = 20°K$ and $t1 = t2 = 24°K$



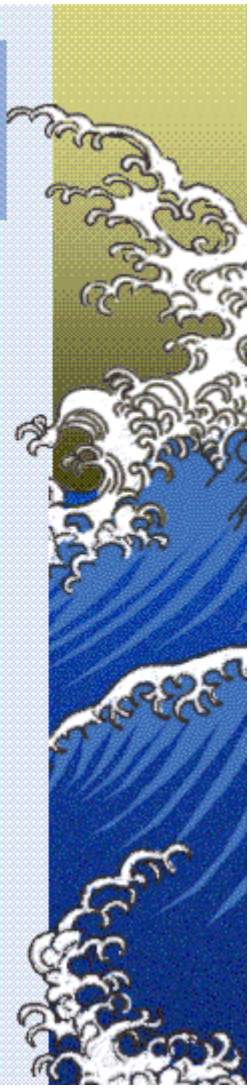
# The observations in the variational cost function : exercice

We have one termometer in a room. We know the standard deviation of the error of this instrument : $\sigma$**o**. We have done, starting from a previous analysis, a forecast of the temperature in this room. We know that the forecast error standard deviation is $\sigma$**b**.

We do 2 measurements **t1 and t2** of the temperature. We know the value **tb** of the forecast. We considere these two measurements as being done by the same instrument, at the same time, with an error correlation of **c**.

Exercice n°6 : compute the analysed temperature ta starting from the *optimal interpolation* formulation

Exercice n°7 : compute the analysed temperature ta starting from the *variational* formulation



# The observations in the variational cost function : exercice

Exercice n°8 : compare the results when $c \rightarrow 1$ and when $c \rightarrow 0$

Numerical application : $\sigma o = \sigma b = 1°K$

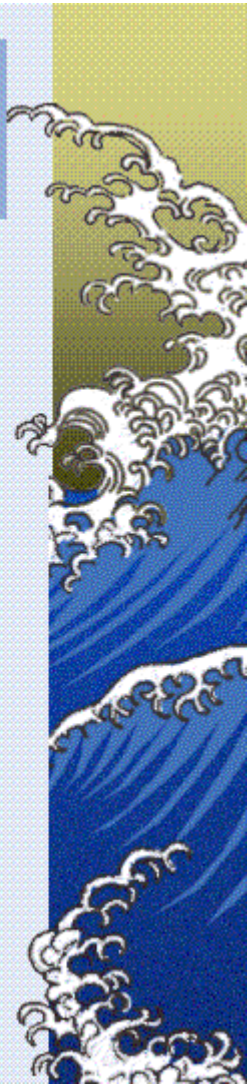
$tb = 20°K$ and $t1 = t2 = 24°K$

$c = 0$ and $c = 1$

Exercice n°9 : since the error covariance matrix of the analysis is :

$$A = (I - KH) P$$

compute $\sigma a$ using the results of the exercice 6.

Compare the results when $c \rightarrow 1$ and when $c \rightarrow 0$

Imagine the value of $\sigma a$ in the case of a big number (n) of non-correlated observations