A posteriori verification of analysis and assimilation algorithms

*Wafaa Sadiki ** *Claude Fischer***
&
*Gérald Desroziers***

* Direction de la Météorologie Nationale, CNRM, Maroc
** Météo-France, CNRM, Toulouse

Abstract :

☛ Improve the knowledge of the assimilation systems based on **variational approach** :

More Precisely

✔ It consists to develop a set of **diagnostics** allowing to validate ***a posteriori***, any analysis algorithm.

$\Downarrow$

• Estimate the accuracy of the **observation** and **background error statistics**, specified in entry of the assimilation system.

CONTENTS

I. ***A posteriori* diagnostics** in a variational scheme :

- "Jmin" Method
- Iterative fixed-point approach

II. Test in the Aladin 3D-Var system :

- **Tuning** of background information error parameters : **Lagged-NMC** statistics

I. *A posteriori* diagnostics in a variational scheme

☞ Incremental Formulation of 3d/4d-Var

Principle

✓ Seek the **increment** :

$$\partial \mathbf{x} = \mathbf{x}_{\mathbf{a}} - \mathbf{x}_{\mathbf{b}}$$

✓ which **minimizes** :

$$J(\partial x) = J_o(\partial x) + J_b(\partial x)$$

$$\mathbf{J(\partial x) = \frac{1}{2}(d - H\partial x)^T R^{-1}(d - H\partial x) + \frac{1}{2}\partial x^T B^{-1}\partial x}$$

✓ Where :

$$d = y - H(x_b)$$

Innovation vector

Talagrand (1997)

✓ **Complete** vector **information** denoted **"z"**, can be seen as :

$$\mathbf{z = (x_b \, . y)^T}$$

y : Proper **observations**, with dimension ***p***

$\mathbf{x_b}$: **Background** estimate, with dimension ***n***

☛ ***Statistical average of the terms in "J "***

➲ **Bennett (1993); Talagrand (1999)**

✍ The **expectation** of **"Jj"** at the **minimum** is :

$$\mathbf{E(J_j) = \frac{1}{2}\left[(m_j + Tr(\Gamma_j^T S_j^{-1} \Gamma_j P^a)\right]}$$

"Jj" stands for a term of **J** witch is the sum of mj elements :

$$J = \sum_j J_j \qquad\qquad m = \sum_j m_j$$

Γ_j Linearized observation operator

S_j Observation error covariance matrix with **mj**

P^a Estimation analysis error covariance matrix resulting from the analysis with **"z"**

☛ ***Application***

❶ ***"Jmin" approach***

➪ *Complete **observation** term :*

$$\Gamma_j = H \qquad S_j = R \qquad m_j = p$$

$$E(J_o) = \frac{1}{2}\left[(p - Tr(H^T R^{-1} H P^a)\right]$$

$$\boxed{E(J_o) = \frac{1}{2} Tr\left[I_p - HK\right]}$$

Where :

$$K = P^a H^T R^{-1}$$

➪ *Complete* ***background*** *term :*

$$\Gamma_j = I_n \qquad S_j = B \qquad m_j = n$$

$$\boxed{E(J_b) = \frac{1}{2}Tr(KH)}$$

➪ *Cost Function :*

Since :

$$Tr(HK) = Tr(KH)$$

It follows that :

$$\boxed{E(J_{min}) = \frac{p}{2}}$$

✍ If The information **error covariances** are properly **specified** in entry of the assimilation system.

➤➤ The **statistical average** of **"J"** at the **minimum** is simply **proportional** to the number of proper **observation**.

➥ Variational system is called "**coherent**"

*** Consequently**

✍ **E(Jmin) ≠ p/2**

➤➤ Indicate a **Bad specification** of the statistics on the information **errors**

This provides :

➥ **Simple *a posteriori*** diagnostic that requires no extra cost.

❷ Iterative fixed-point approach

✍ Use **statistical average** of the terms in **"J"** as **diagnostics tools**

However

Difficulty in the estimation of the error statistics, coming from the computation of :

$$\mathbf{Tr}(\Gamma_j^T S_j^{-1} \Gamma_j P^a)$$

➤➤ This method introduces a practical way for computation of the **trace** even when **"mj"** is **large**

For example

⇨ $mj = p$

$$E(J_o^{min}) = \frac{1}{2} Tr\left[I_p - HK \right]$$

➲ Girard 1989 (Monte-Carlo method)

Estimate **trace** by a **randomized** method, using a method proposed by *Girard* :

$$\mathbf{Rand}\left[\mathbf{Tr(HK)}\right] = \partial \mathbf{y}^{\mathbf{T}}\left[\mathbf{H}\partial \mathbf{x}^{\mathbf{a}}_{(\mathbf{y}+\partial \mathbf{y})} - \mathbf{H}\partial \mathbf{x}^{\mathbf{a}}_{\mathbf{y}}\right]$$

Where :

$\partial \mathbf{y}$ Vector of small perturbations with Gaussiane distribution

$\partial \mathbf{x}^{\mathbf{a}}_{(\mathbf{y}+\partial \mathbf{y})}$ Analysis increments obtained with **perturbed observations**

➪ **Similar** produce can be employed to determine

mj = n

$$E(J_b) = \frac{1}{2}Tr(KH)$$

➤➤ **Perturbing background** information :

$$\boxed{Rand[Tr(KH)] = -\partial x_b^T \left[H\partial x^a_{(x_b+\partial x_b)} - H\partial x^a_{(x_b)} \right]}$$

☛ ***Remark :***

Each **trace** evaluation ➤➤ **2 analyses steps,** with **unperturbed** and **perturbed** informations

Tuning of information error parameters: Desroziers (2000)

✍ **Optimize** the **weighting** parameters of background and observation terms.

◆ If **"Sb"** and **"So"** denote these 2 parameters

➤➤ Complete cost function can be re-written as :

$$\boxed{J = \frac{1}{S_o} J_o + \frac{1}{S_b} J_b}$$

◆ The idea is to find "So" and "Sb" such as :

➤➤ Values of **"1/Sb*Jb"** and **"1/So*Jo"** are close to their expected values

$\Downarrow$

1/2Tr(KH) & 1/2Tr(Ip-HK)

➡ **"Coherent"** variational system in the sense of the error statistics.

✓ Thus, we define **"so"** and **"sb"** such as :

$$S_o = s_o^2 = \frac{2E(J_o^{min})}{Rand\left[Tr(I_p - HK)\right]}$$

$$S_b = s_b^2 = \frac{2E(J_b^{min})}{Rand\left[Tr(KH)\right]}$$

✓ Form of the non-linear equations suggests the use of an **iterative fixed-point** method with the following sequence :

$$s_o^{k+1} = \sqrt{S_o(s_o^k, s_b^k)} \qquad s_b^{k+1} = \sqrt{S_b(s_o^k, s_b^k)}$$

-

II. Test in the Aladin 3D-Var system : Tuning of "lagged-NMC" statistics

Feature

♦ Use of a **mesoscale Jb** term in the **LAM** 3D-Var :

➤➤ Band-pass filter

$\Downarrow$

Not to reanalyze the **large scales** already analyzed by the global model and provided by the coupling data.

"Lagged NMC" method in Aladin

♦ **Reduce** the energy in the **large-scale spectrum**

❒ *Principle*

➤➤ Differences between two forecasts of various durations, 36h and 12h, for the same validating time ➥ (**F36 - F12**)

✓ **Identical LBC** for both F36 & F12

$\Downarrow$

$\boxed{\textbf{Same coupling data}}$

✓ **IC** of **F12** are provided by the **36h sequence**

✍ It was noted :

✔ Effect of **initialisation** by Digital Filters

✔ Different **IC**

✔ **Integration** of the model 24 hours later

➤➤ Build the structure functions

➪ ***Advantage % standard NMC method :***

♦ **"Reduce"** energy spatially in the **large scales**

⇓

➤➤ Pb of observation influence placed in the middle of the domain over the whole domain

♦ **"Keep"** the same structure functions in the **small scales**

⇓

➤➤ More mesoscale **representative analysis** increments

❐ *Limitation :*

As a consequence

♦ **Information** about the true level of the background error variance, compared to those of the observation, **is missing**.

❐ *Explanation :*

✓ Not use the information of the analyses and the **observations** from the H+24 lead time.

Problem

✓ **"Retrieve"** the missed information, to control the analysis increments

➤➤ **Tuning** the ratio **"Sb"** such as :

$$\boxed{J^{lagged} = J_o + \frac{1}{Sb} J_b}$$

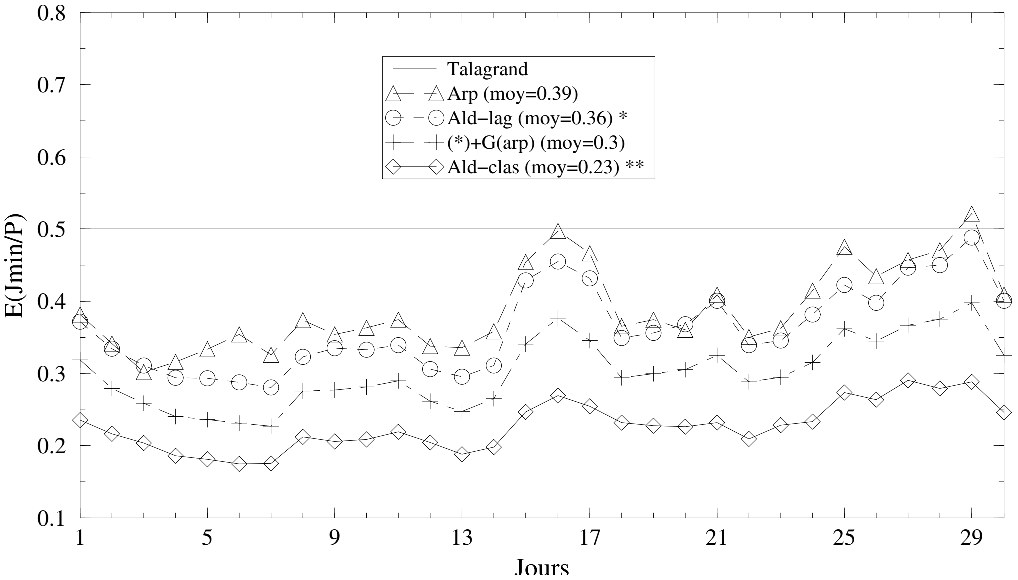
☞ Experiments

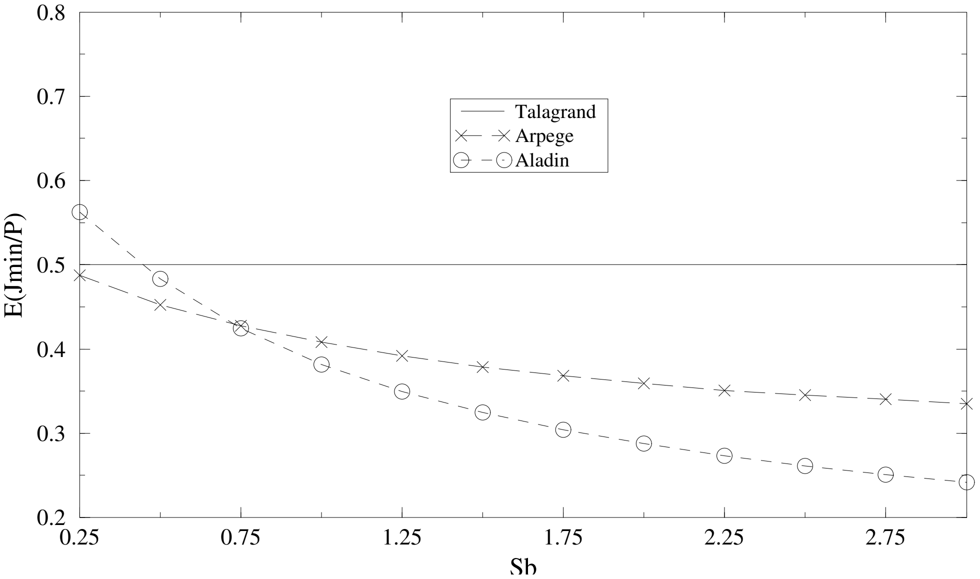
- Domain : Aladin-France
- Observation file : Without tovs data
- Period : 30 days (20 marsh / 18 April 2000), at 00 UTC

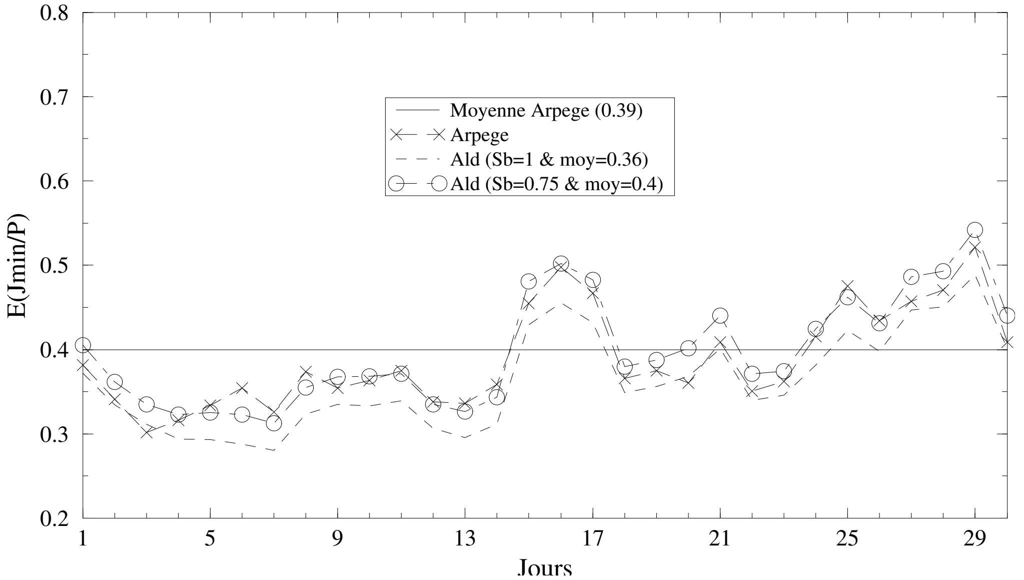
✓ *It's worth to notice :*

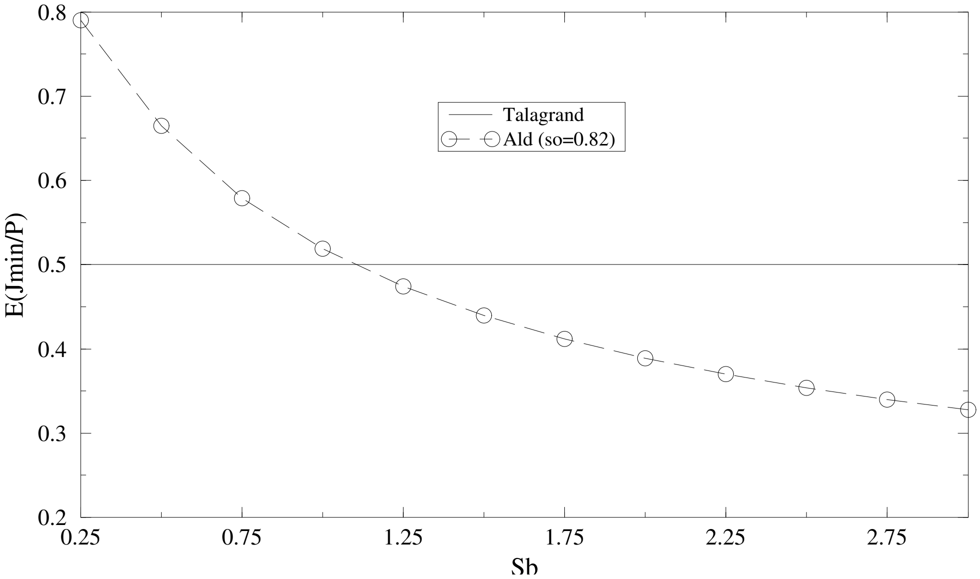
Scale / ***First guess***	**Large k* < 7-10 (*)**	**Medium**	**Small k* > 55 (*)**
06H Aladin forecast (P)	Coupling effect	Aladin model integration	
Interpolated Arpege analysis(A)	Arpege analysis (Post proccesing)		Mathematical signal

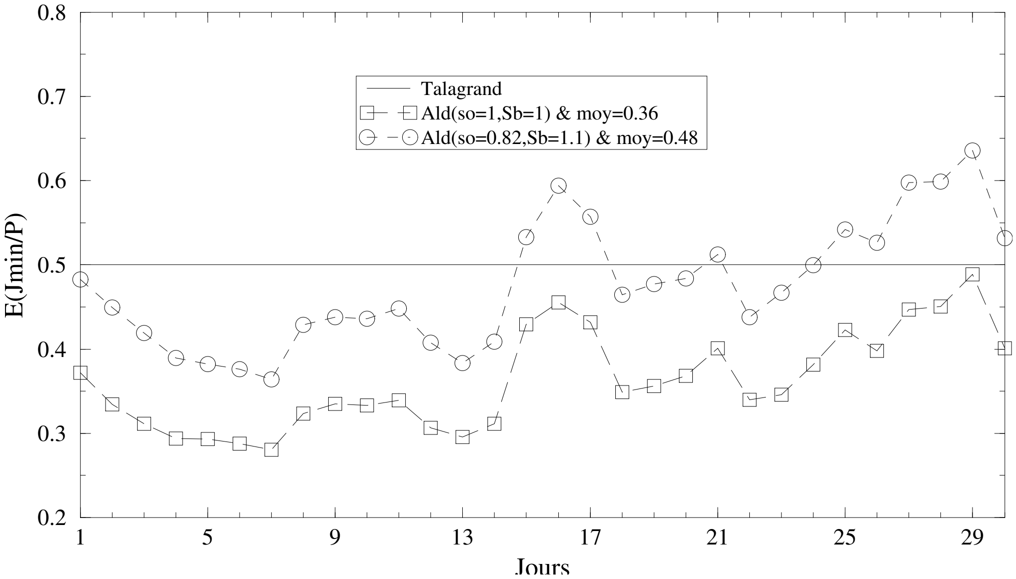
(*) : obtained values for Aladin-France


Talagrand
Arp (moy=0.39)
Ald−lag (moy=0.36) *
(*)+G(arp) (moy=0.3)
Ald−clas (moy=0.23) **
E(Jmin/P)
Jours


Talagrand
Arpege
Aladin
E(Jmin/P)
Sb


Moyenne Arpege (0.39)
Arpege
Ald (Sb=1 & moy=0.36)
Ald (Sb=0.75 & moy=0.4)
E(Jmin/P)
Jours


Talagrand
Ald (so=0.82)
E(Jmin/P)
Sb


Talagrand
Ald(so=1,Sb=1) & moy=0.36
Ald(so=0.82,Sb=1.1) & moy=0.48
E(Jmin/P)
Jours

III. Synthesis & perspectives

☛ ***A posteriori*** **diagnostics**

❶ *"Jmin" approach*

➲ ***Arpege :***

Aladin Domain ➢➢ Sb = **0.9**

&

E(Jmin)/p = **0.39**

➲ ***Aladin "lagged NMC" :***

➨First guess == **06h Aladin forecast**

Default ➤➤ Sb = **1**

&

E(Jmin)/p = **0.36** "on average"

Theoretical value ➤➤ $Sb \approx$ **0.5**

"Test situation"

Arpege Reference➤➤ $Sb \approx$ **0.75**

&

E(Jmin)/p = **0.4** "on average"

➨First guess = = **Interpolated Arpege analysis**

Default ➤➤ Sb = **1**

&

E(Jmin)/p = **0.3** "on average"

➥ Already relatively **correlated** to the observations

✍ **"Jmin"** diagnostic **sensitive** to the coherence between **innovation** statistics & **Jb**

⇓

➤➤ **Fast** & **effective** tool to test various data types to be specified in entry of 3d-Var

❷ *Iterative "fixed-point" approach*

- *12 April 2000 case at 00 UTC*

  ✔ (so=**0.82**, Sb=**1.1**) **&** E(Jmin)/p = **0.5**

- *Validation*

  ✔ E(Jmin)/p = **0.48** "on average"

  ✔ 2000032000 ➤➤ (so=**0.89**, Sb=**1.2**)

⇨ **SUMMARY**

☞ **Sb must be close to 1**

☞ How do the **"So < 1" result** can be interpret in **Aladin** ?

☛ PERSPECTIVE

✍ Realize tests with **blended fields**

✍ More **Stability** for the iterative fixed-point method :

♦ Estimate **Tr(KH)** / Tr(HK) by **perturbing background** information

$$\mathbf{Rand}\left[\mathbf{Tr(KH)}\right] = -\partial \mathbf{x}_b^{T}\left[\mathbf{H}\partial \mathbf{x}^{a}_{(\mathbf{x}_b+\partial \mathbf{x}_b)} - \mathbf{H}\partial \mathbf{x}^{a}_{\mathbf{x}_b}\right]$$

♦ Define **"So"** for each **observation type** / each **variable**

✍ **J = (1/So) Jo + (1/Sb) Jb**

Diagnostics	**Itération par point fixe**	**"Jmin"**
Situations/ Expériences	so= So Sb	E(Jmin)/p
12042000 ***à 00UTC*** ===== ***20032000*** ***à 00UTC***	1 1 0.75 1.1 0.86 1.15 0.79 1.03 . . -- 0.84 1.09 **0.82** **1.1** **0.82** **1.1** . . . . 0.82 1.07 0.83 1.14 -- . . 0.79 0.89 0.88 1.4 0.77 0.83 0.9 1.5 ==========================	➤➤ **0.5** ➤➤ avec ☛ **so=0.82 & Sb=1.1** ============ ☛ **so=0.89 & Sb=1.2**
17042000 ***à 00UTC***	1 1 0.91 1.82 0.88 1.03 1 1.96 . . . . 0.74 0.67 1.21 3.42 -- *exp1* exp2 *exp1* exp2 1 0.82 2 1.45 1 0.82 0.88 1.44 1 0.99 2.08 1.44 1 0.84 . . . . 1 2.44 1 0.77	➤➤ 0.5 avec so=1 & Sb=1 ➤➤ **0.57** ➤➤ avec ☛ **so=0.82 & Sb=1.44**