

Configuration d'un filtre de Kalman d'ensemble préopérationnel

Peter Houtekamer, Herschel Mitchell, Martin Charron, Gérard Pellerin,
Bjarne Hansen, Chantal Côté et Lubos Spacek

Situation:

Au CMC, nous utilisons la méthode de l'Interpolation Optimale (OI) pour générer les conditions initiales pour un ensemble de 16 prévisions de 10 jours. Nous voulons remplacer l'OI par une méthode moderne qui est capable de traiter en particulier les données de radiance.

Projet:

En 1996, nous avons commencé un projet pilote pour étudier la méthode du filtre de Kalman d'ensemble (EnKF). L'EnKF combine de façon cohérente l'assimilation 4-d des observations et la génération des conditions initiales pour la prévision d'ensemble.

État du projet:

Nous assimilons de vraies observations et nous obtenons des résultats intrigants.

Le long chemin vers des résultats acceptables

Il n'est pas évident de développer une toute nouvelle méthode d'assimilation et d'obtenir des résultats d'une qualité acceptable. Heureusement, nous avons pu profiter de l'expérience des gens qui travaillent au modèle et à des méthodes d'assimilation variationnelles.

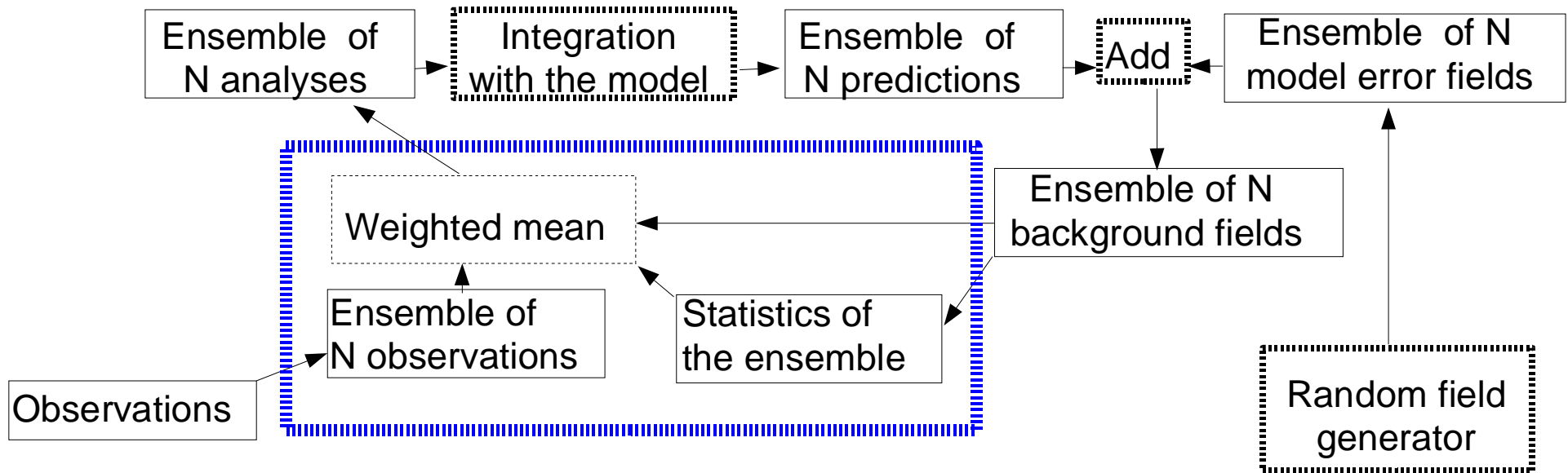
Remerciements:

- Luc Fillion (le développement initial de l'algorithme),
- Richard Ménard et Steve Cohn (opérateur de localisation),
- Gilles Verner et Pierre Koclas (traitement des observations),
- Josée Morneau (usage d'arcad),
- Stéphane Laroche et Jacques Halle (opérateurs d'interpolation),
- Mark Buehner (description statistique pour l'erreur du modèle),
- Michel Roch (modèle GEM et GEM-DM).
- L'équipe de Michel Valin (conseils informatiques).
- Dave Steenbergen et Louis Lefavre (pour la confiance).

Contenu de la présentation

- Introduction
- Environnement pour les expériences:
modèle
observations
- Produit de Hadamard
- Description de l'erreur du modèle
- Comparaison avec les statistiques des innovations
- La dynamique étrange des erreurs
- Comparaison avec le 3d-var
- Conclusions

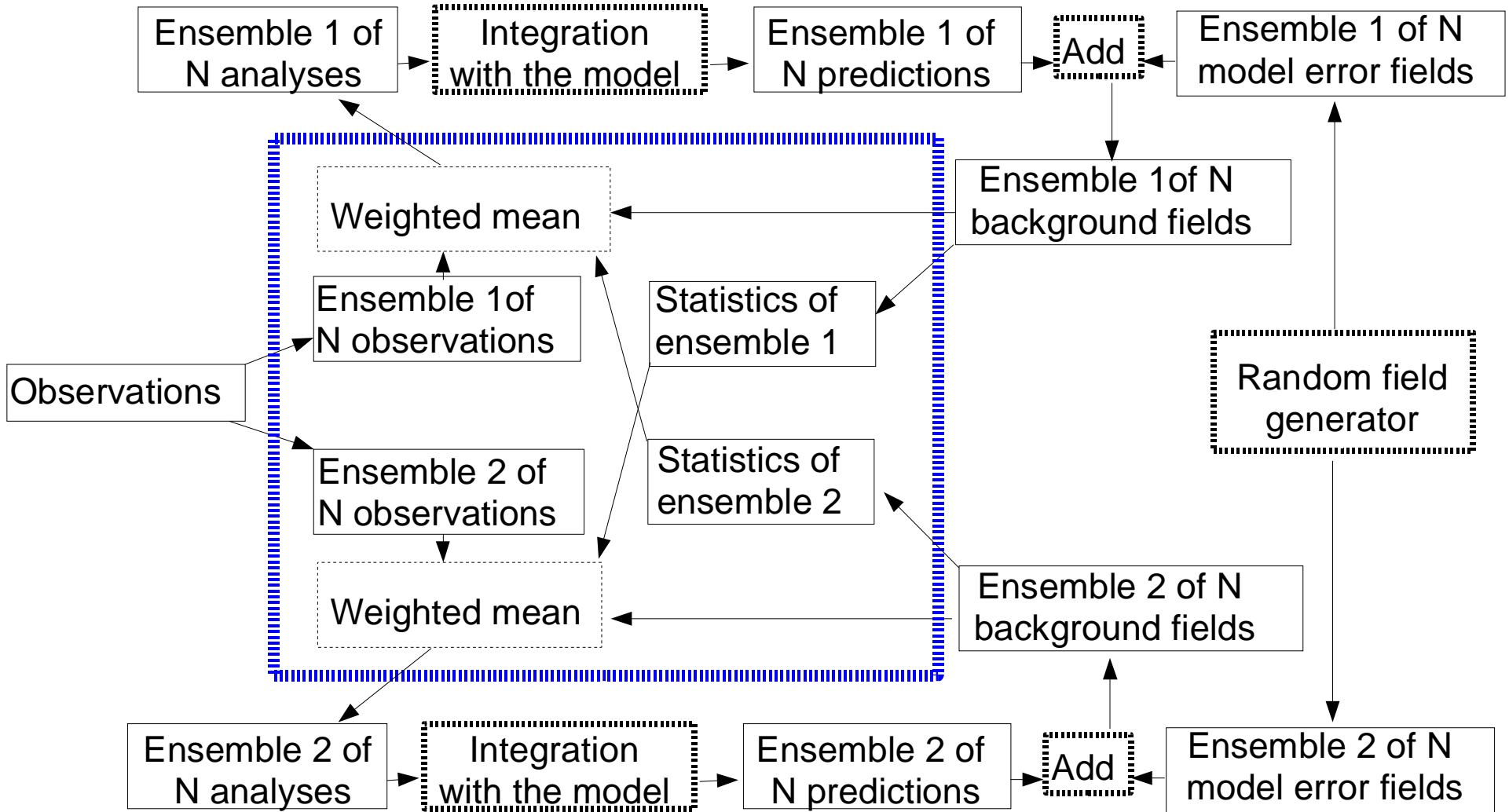
Single Ensemble Kalman Filter



Almost all other EnKF groups use a Single EnKF as depicted here. However the ensemble of background fields is used both to compute the weights and to test the quality of these very same weights. This may lead to an underdispersive ensemble and to filter divergence.

The challenge is to derive an algorithm that behaves well and that makes an efficient use of computer resources.

Double Ensemble Kalman Filter



Environnement: le modèle

Nous utilisons le modèle GEM-DM version 3.0.2. La configuration est identique au modèle GEM-DM 400 x 200 sauf pour:

la résolution horizontale réduite à 240 x 120,

le pas de temps de 60 minutes,

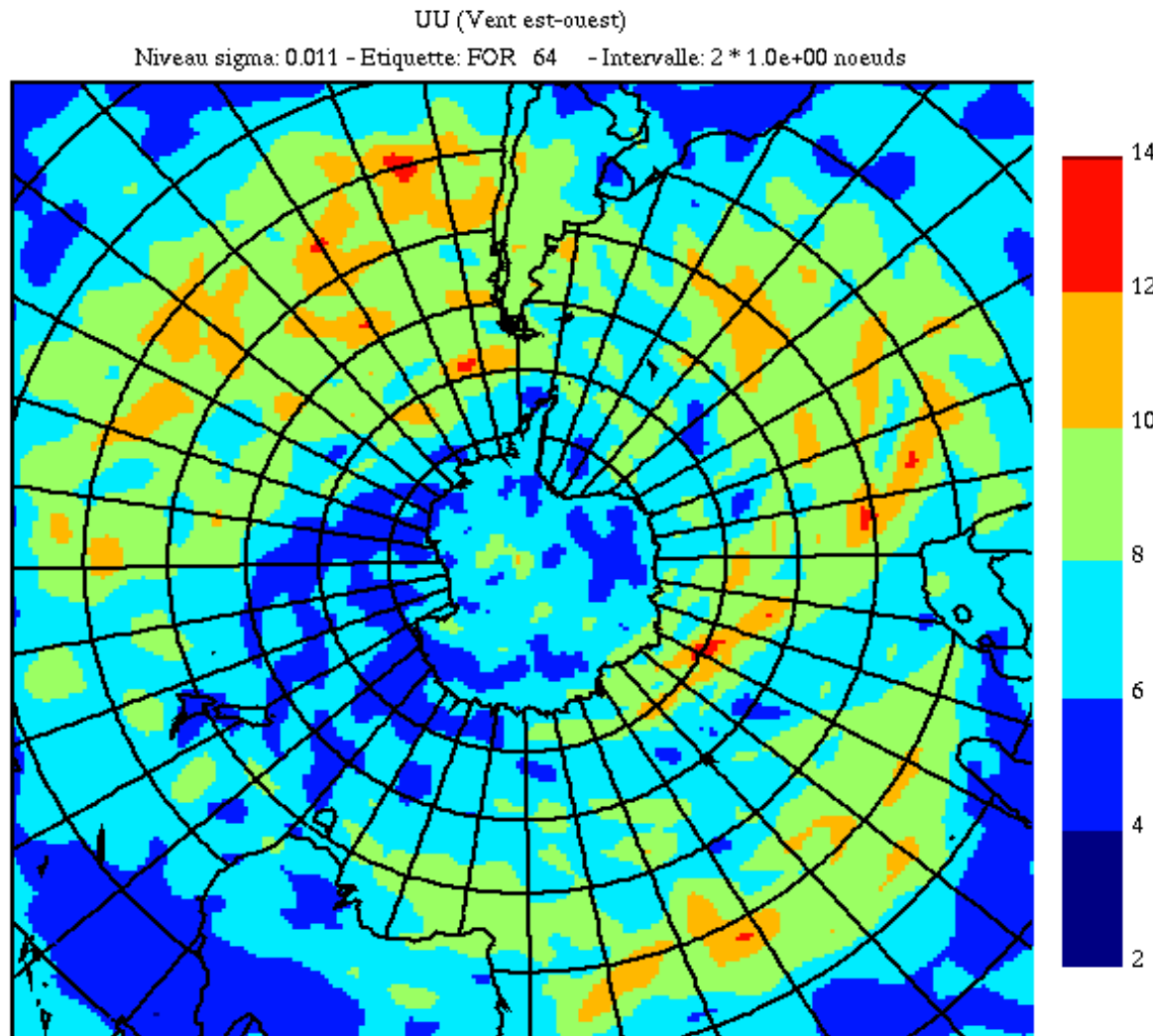
la diffusion horizontale légèrement augmentée,

les pôles du modèle aux pôles géographiques,

le traitement de la neige sur la glace.

Note: nous montrerons aussi quelques résultats avec une résolution à 144 x 72 et avec le modèle GEM v1.05.

Bon comportement du modèle GEM-DM au pôle sud



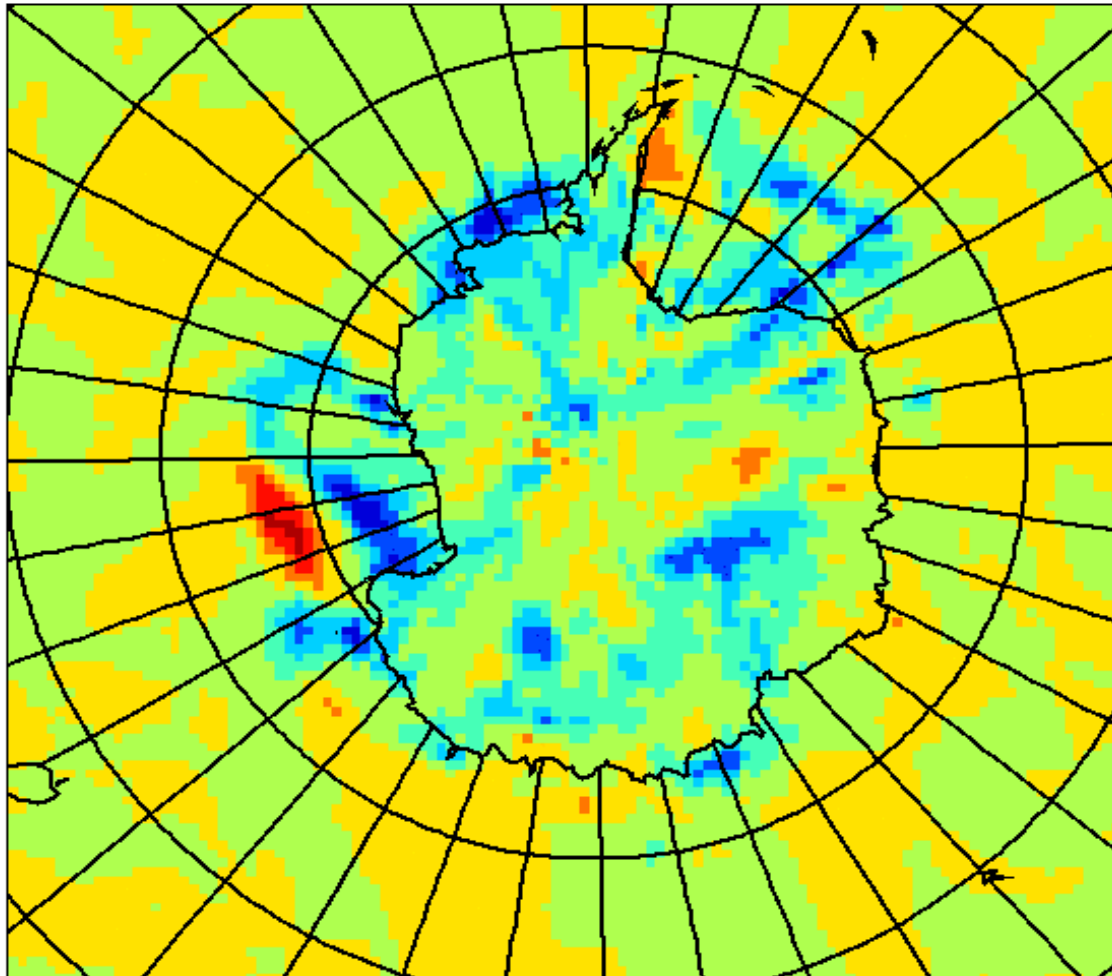
Prevision 06 heures valide 00:00Z le 22 mai 2002

Pour le vent zonale au niveau $\eta = 0.011$, nous regardons l'incertitude dans le champs d'essai (22 mai 2002 6 UT).

Les pôles du modèle sont aux pôles géographiques.

Nous pouvons donc éviter une étape d'interpolation entre l'analyse et le modèle de prévision.

Problème avec la température sur la glace marine.



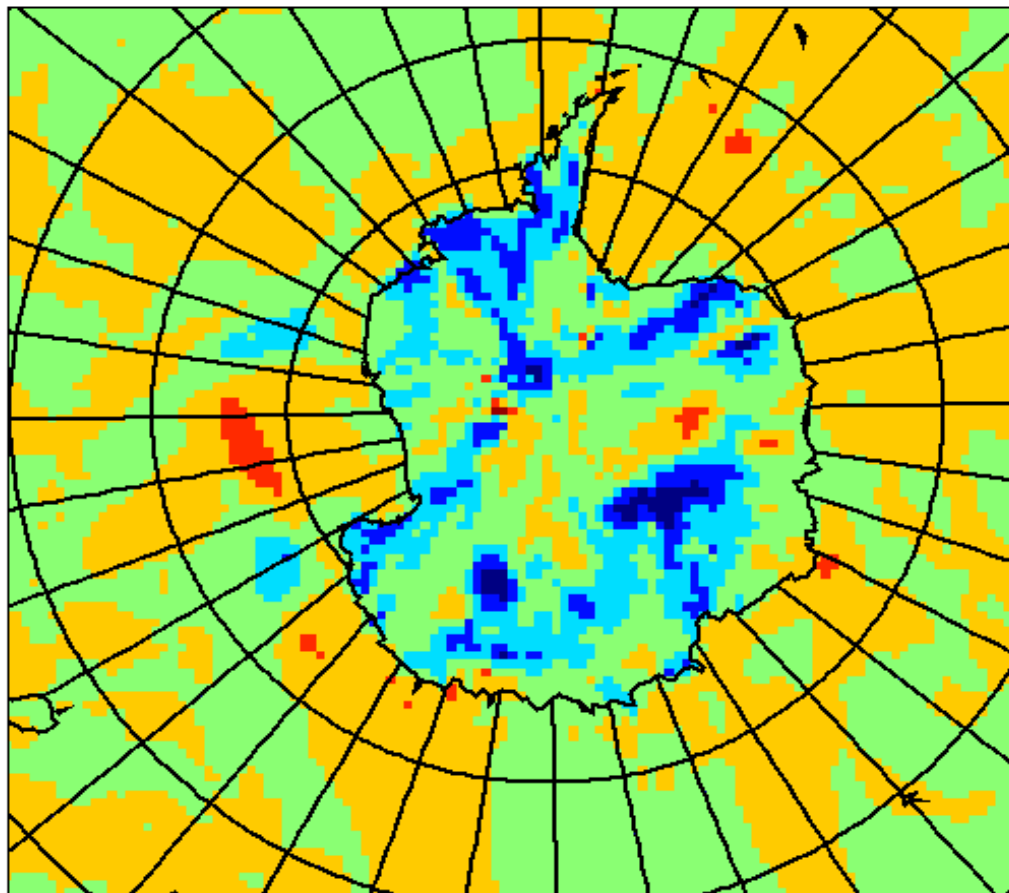
TT*P*12000*[6- 0]* 0*[V06:00Z 19mai2002-]*GDM_KAL008

Pour la température au niveau $\eta = 1.000$, nous regardons l'évolution pendant une intégration de 6 heures (19 mai 2002 0 - 6 UT).

Avec la version standard du modèle, nous notons un refroidissement très rapide de la température de surface sur la glace.

La neige sur glace n'a pas d'âge et "donc" très peu de conductivité.

Changement du traitement de la neige sur la glace marine.



TT*P*12000*[6- 0]* 0*[V06:00Z 19mai2002-]*GDM KAL008

Dorénavant, nous traitons la neige sur la glace comme la neige sur la terre (voir Gérard Pellerin pour les détails).

L'évolution rapide sur la glace a disparu. Il n'y a quasiment pas de changements (comparé avec l'expérience précédente) sur la terre.

Nous suivrons les développements.

Traitement des observations

Notre approche est d'utiliser toutes les observations acceptées par le 3d-var global opérationnel.

Nous prenons les observations du fichier postalt du système opérationnel. De cette façon, nous utilisons le “background check” et le qc-var opérationnel.

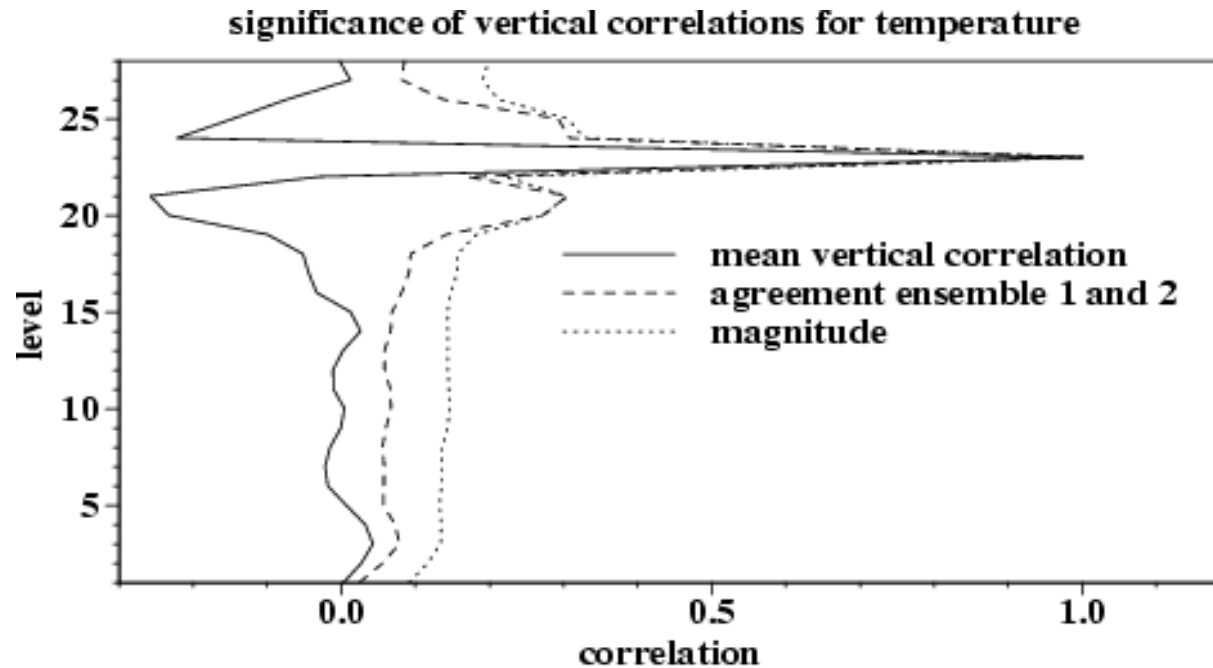
Les erreurs d'observation utilisées dans l'EnKF ont été prises du système 3d-var opérationnel .

On assimile:

- Les radiosondages: $u, v, T, q, p_{\text{surface}}$
- Avion: u, v, T
- Satellite: $u, v, \text{ TOVS 1b radiances}$
- Surface: T, p_{surface}

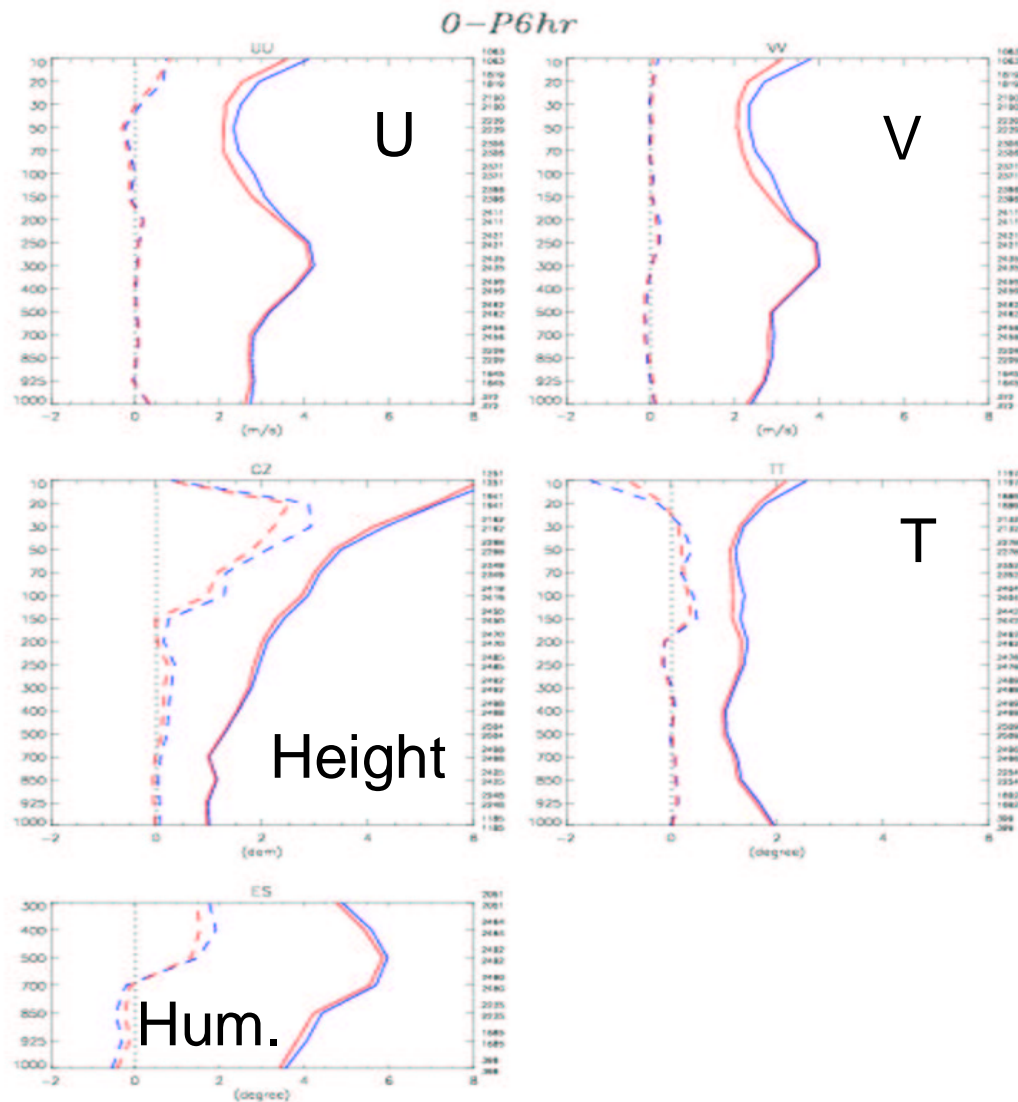
Les observations de surface du vent et de l'humidité ne sont pas encore assimilées.

Example of narrow vertical correlations in the upper levels



For a pair of 48 member ensembles the vertical correlations were computed with respect to level 23 (at about 100 hPa). The global mean correlations are negative at levels 22 and 24! Above level 26 and below level 19 the two 48 member ensembles tend to disagree. Estimated vertical correlations are at the noise level and should therefore be filtered. This is done with a Hadamard product in the vertical.

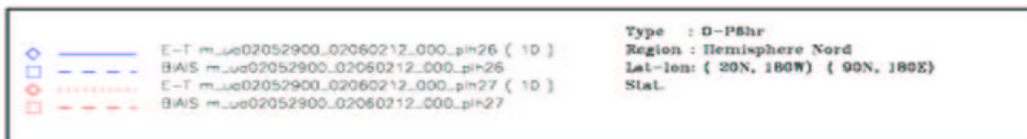
Impact of a vertical Hadamard product



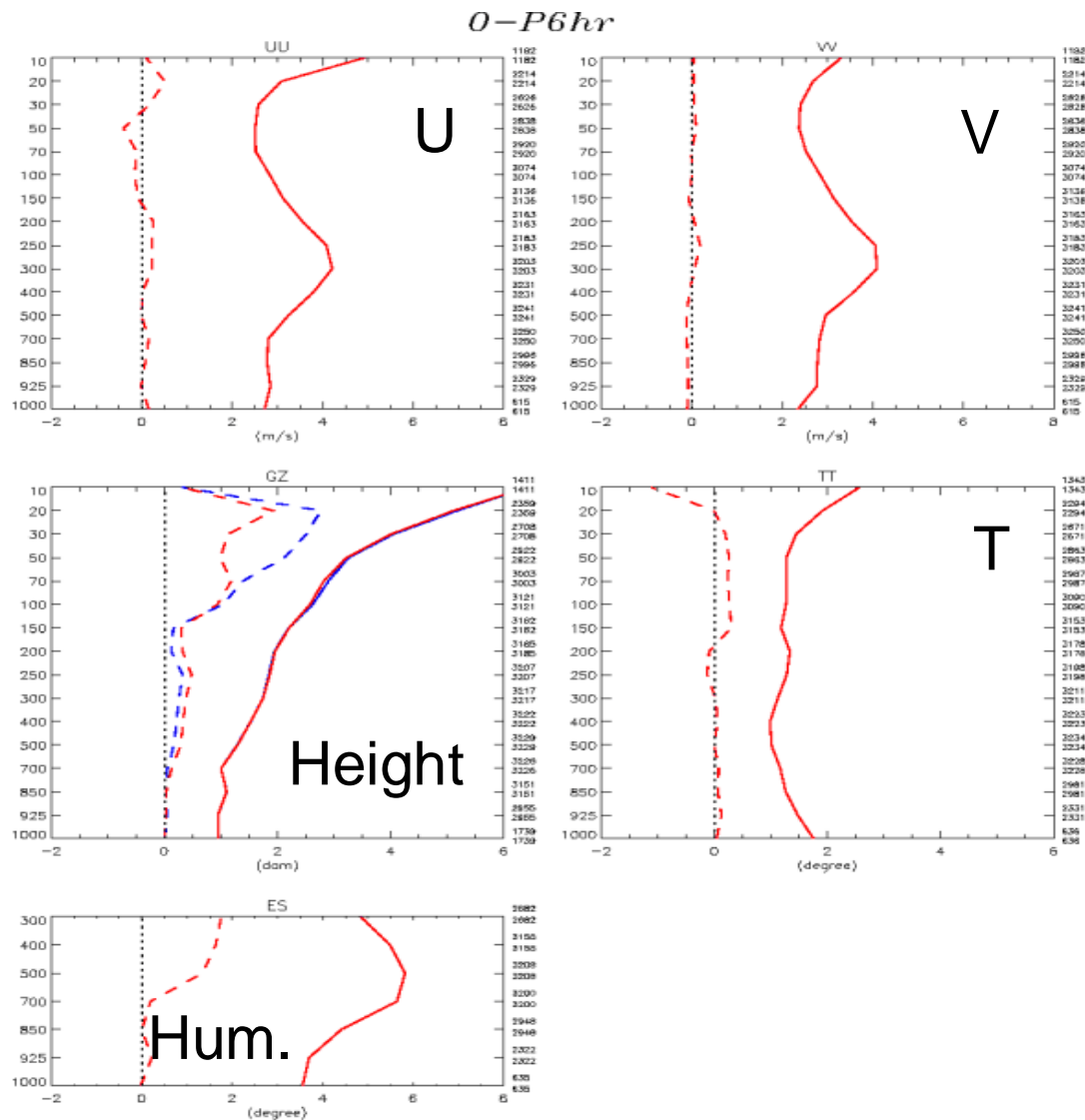
The reference experiment has no localization in the vertical.

It is compared with an experiment where correlations are forced to zero in 2 units of $\ln(\text{pressure})$.

The impact is very positive for in particular the upper levels where vertical correlations were seen to be very narrow.



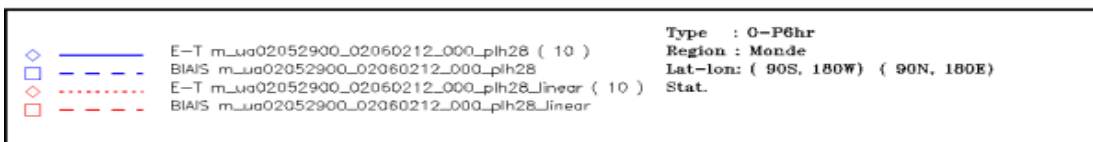
Interpolation linéaire pour la vérification avec arcad



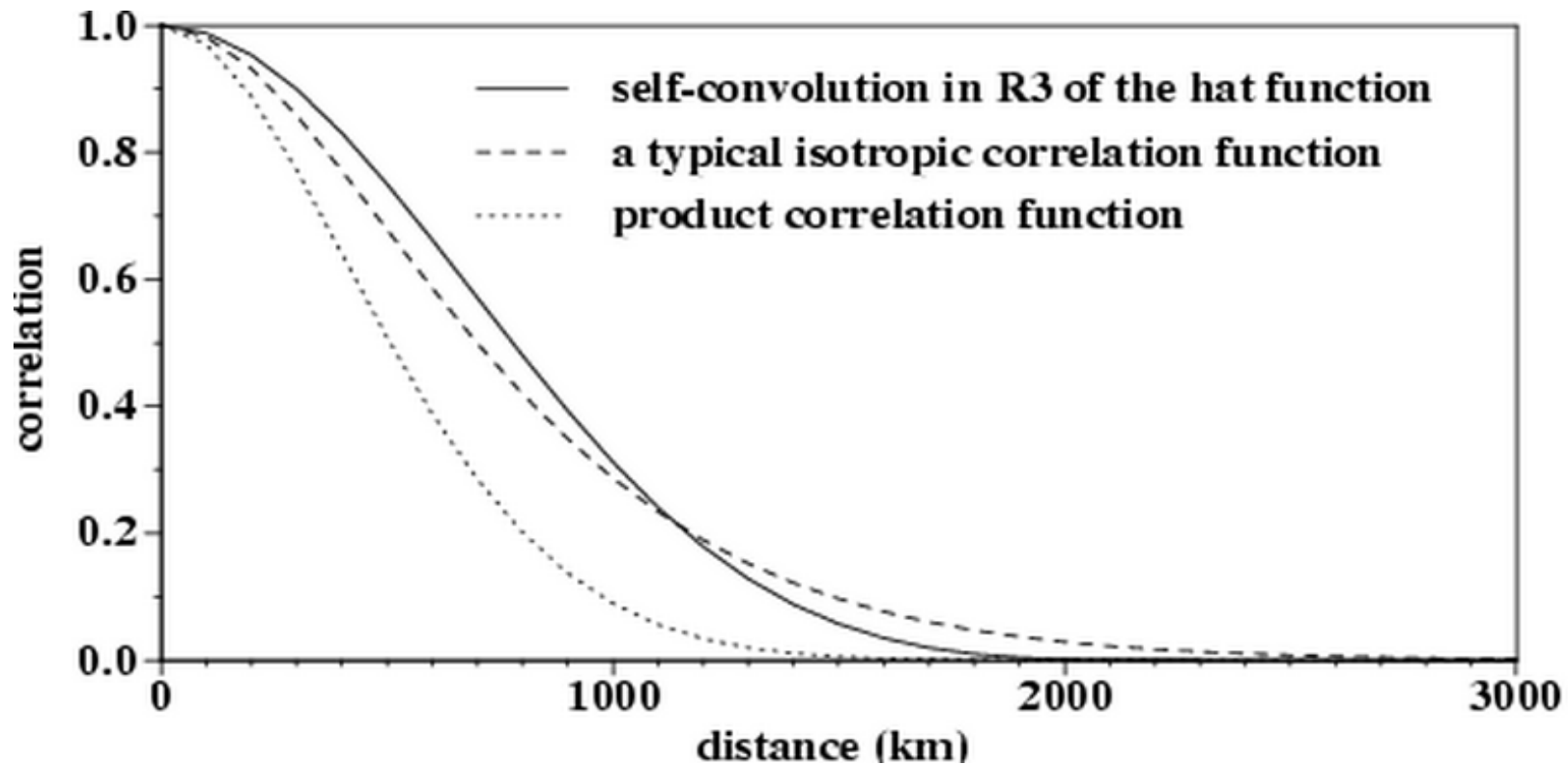
Le même cycle d'assimilation a été validé **une fois avec interpolation verticale cubique** et **une fois avec interpolation verticale linéaire** pour les hauteurs.

Le biais se regarde mieux après l'interpolation linéaire.

Nous concluons que la résolution verticale est insuffisante en haut de 200 hPa.

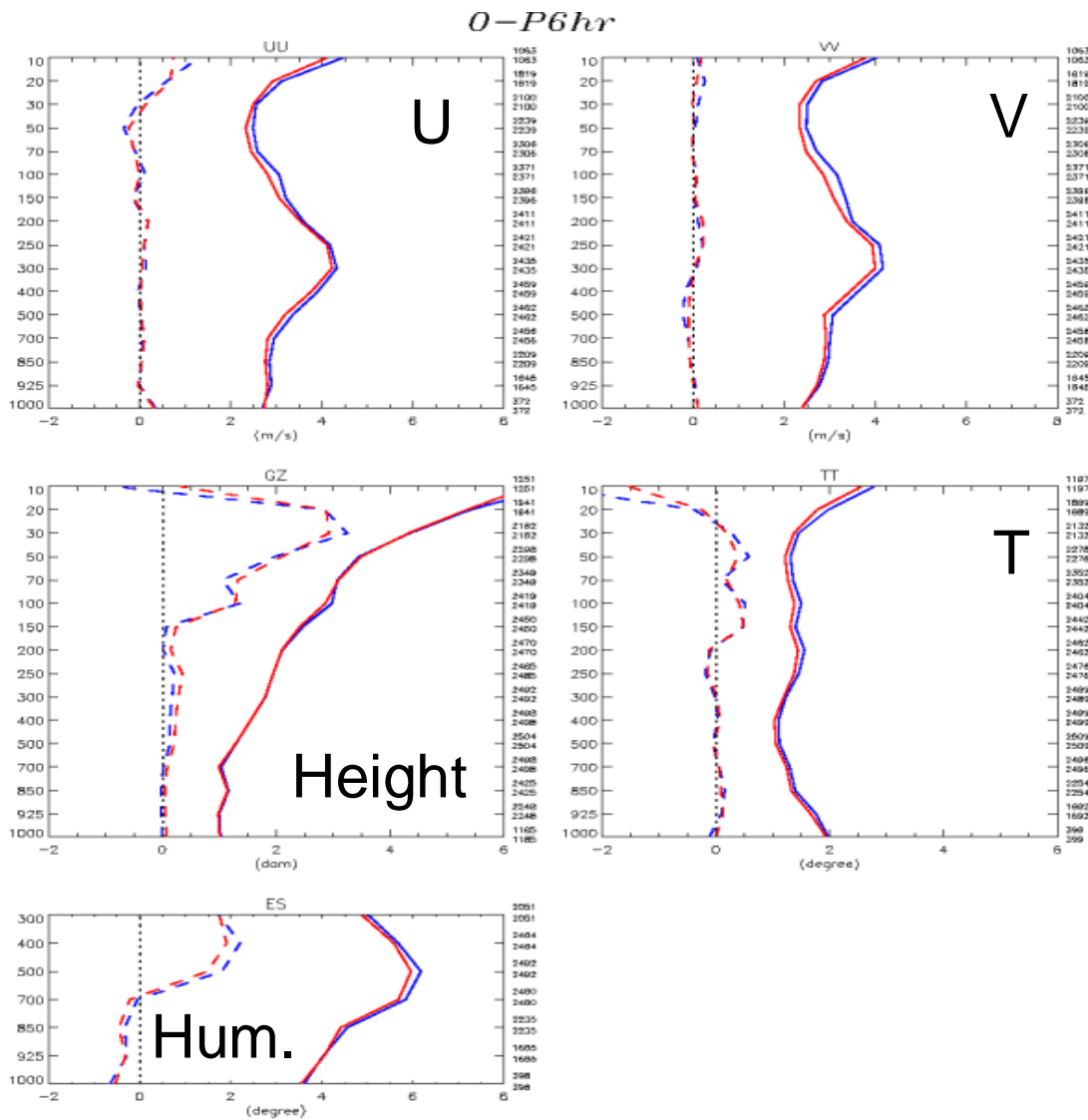


Horizontal Hadamard Product



To filter covariances at long distances we use a Hadamard product (which does a point-wise product of two matrices). Correlations are forced to zero at 2300 km!

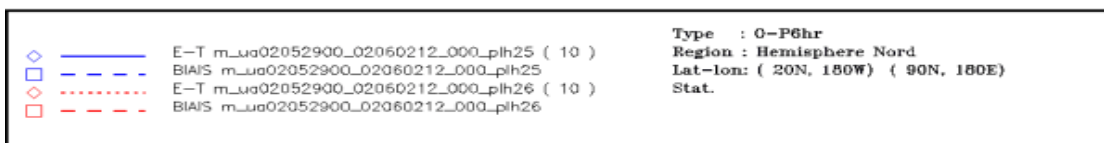
Impact of the horizontal Hadamard product



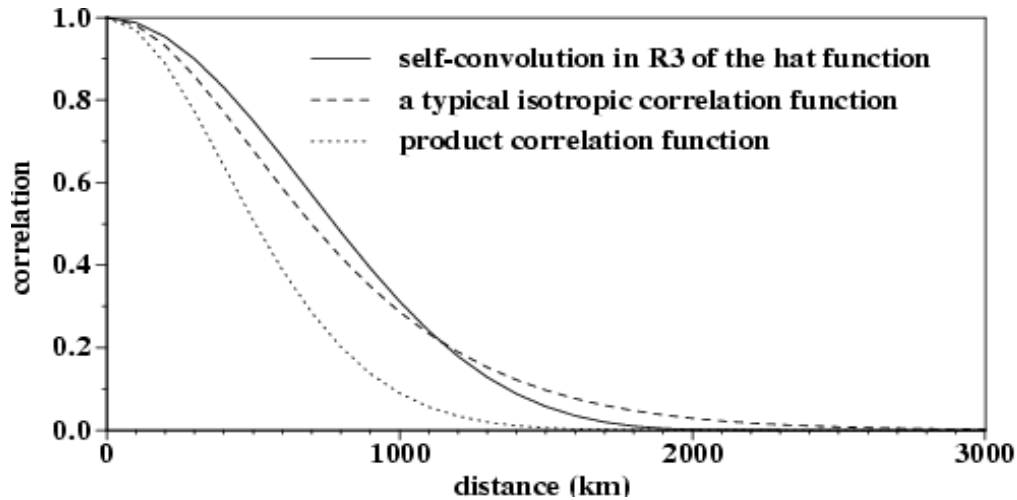
In the reference experiment correlations are forced to zero at a distance of 3400 km.

Smaller errors are obtained when forcing the errors to zero at a distance of 2300 km.

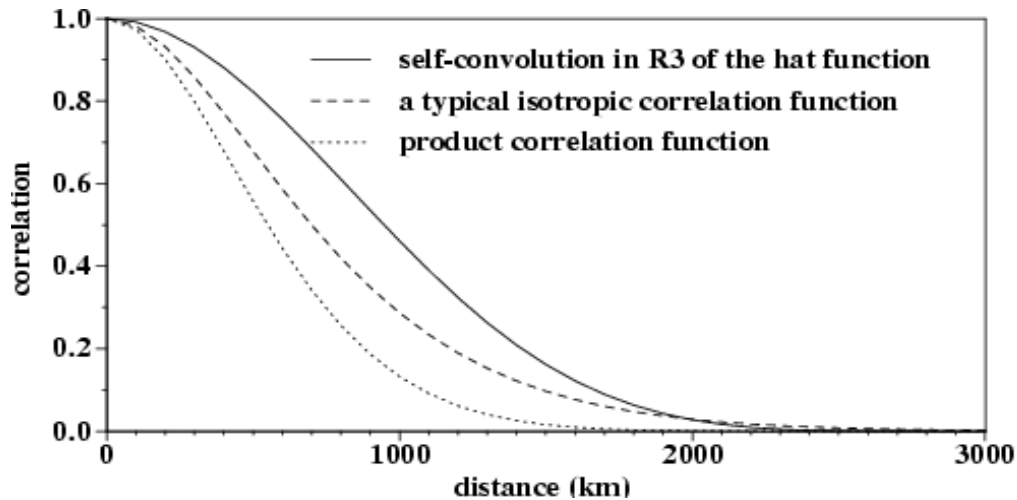
Based on these results one would use a strong localization with enforced zero-impact at 2300 km. Alternatively one could consider using a bigger ensemble.



Localisation des covariances

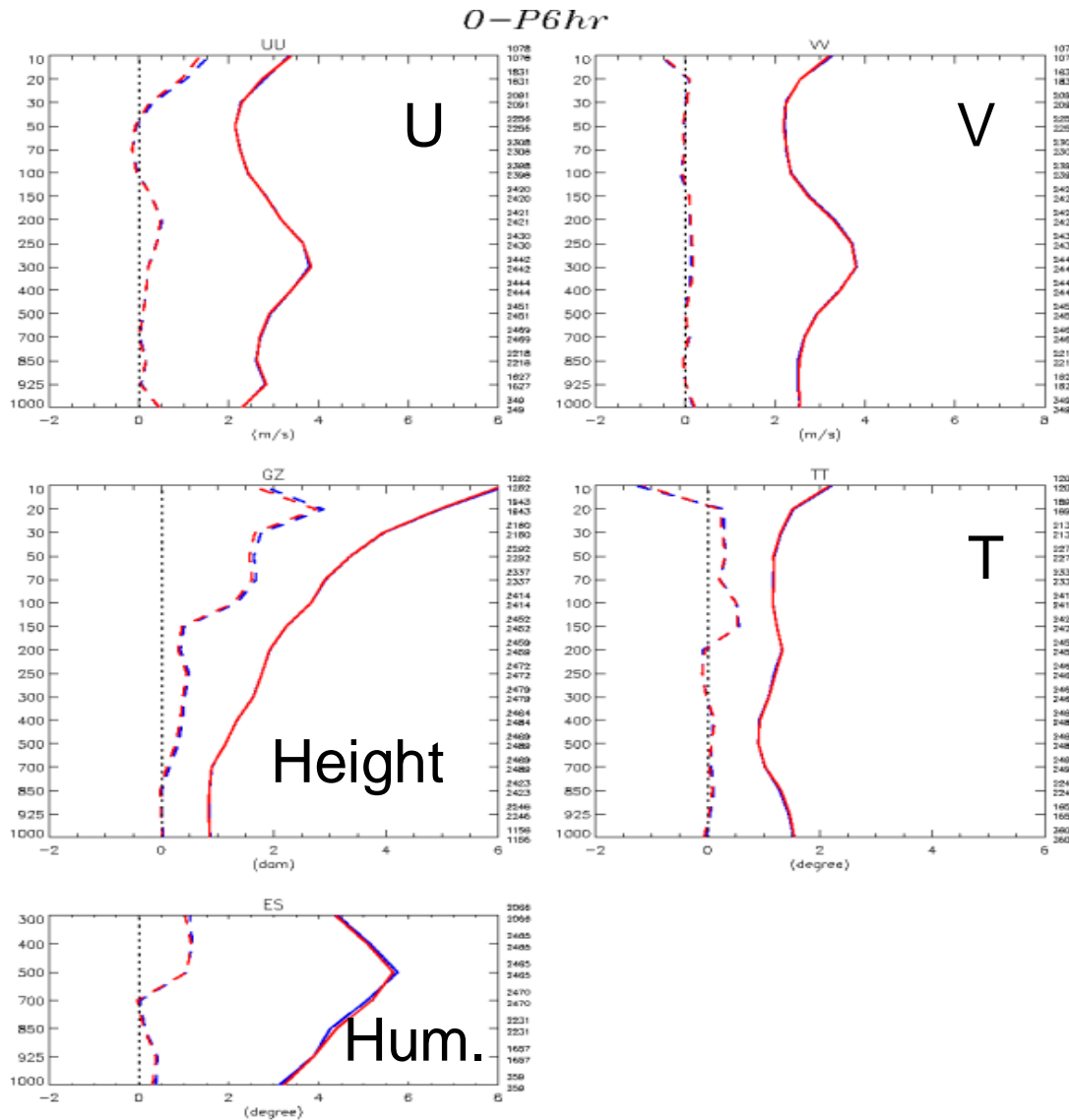


Localisation à 2300 km
permet d'utiliser un ensemble
de taille modeste (2 x 48).



Localisation à 2800 km
permet de mieux maintenir
l'équilibre des champs d'essai
(usage avec 2 x 64 membres).

Taille de l'ensemble

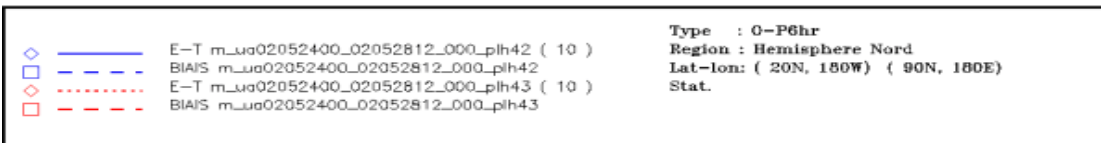


En bleu, nous avons une expérience avec 2 x 48 membres et localisation à 2300 km.

En rouge, nous avons superposé une expérience avec 2 x 64 membres et localisation à 2800 km.

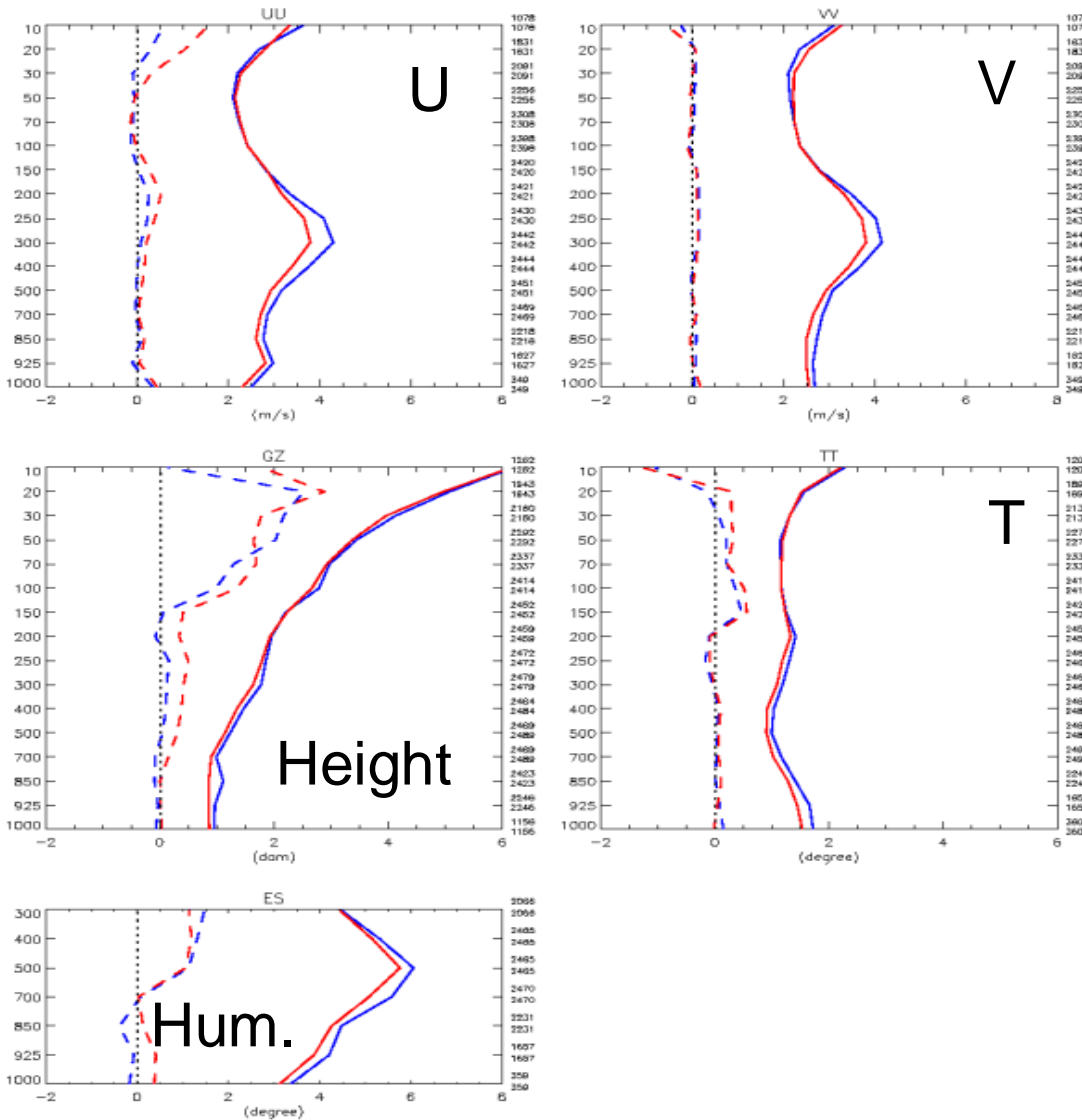
Nous voyons des résultats presque identiques (pour l'hémisphère nord).

Notre approximation du filtre de Kalman a convergé!



Résolution horizontale

0-P6hr

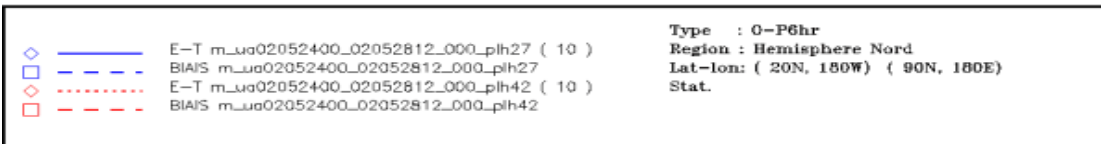


En bleu, nous voyons une expérience avec GEM v1.05 à 144 x 72.

En rouge, nous voyons une expérience avec GEM-DM à 240 x 120.

Nous constatons l'effet bénéfique de la résolution surtout en bas de 200 hPa.

Ayant le choix d'augmenter la taille de l'ensemble au delà de 2 x 48 ou d'augmenter la résolution au delà de 240 x 120 x 28, nous choisirons d'augmenter la résolution (horizontale ou verticale).



Simulation de l'erreur du modèle

Comme hypothèse de travail, nous présumons que les covariances de l'erreur du modèle sont similaires à celles de l'erreur de prévision du 3d-var.

$$\mathbf{P}^f(\mathbf{t} + 1) = \mathbf{M}\mathbf{P}^a\mathbf{M}^T + \mathbf{Q}$$

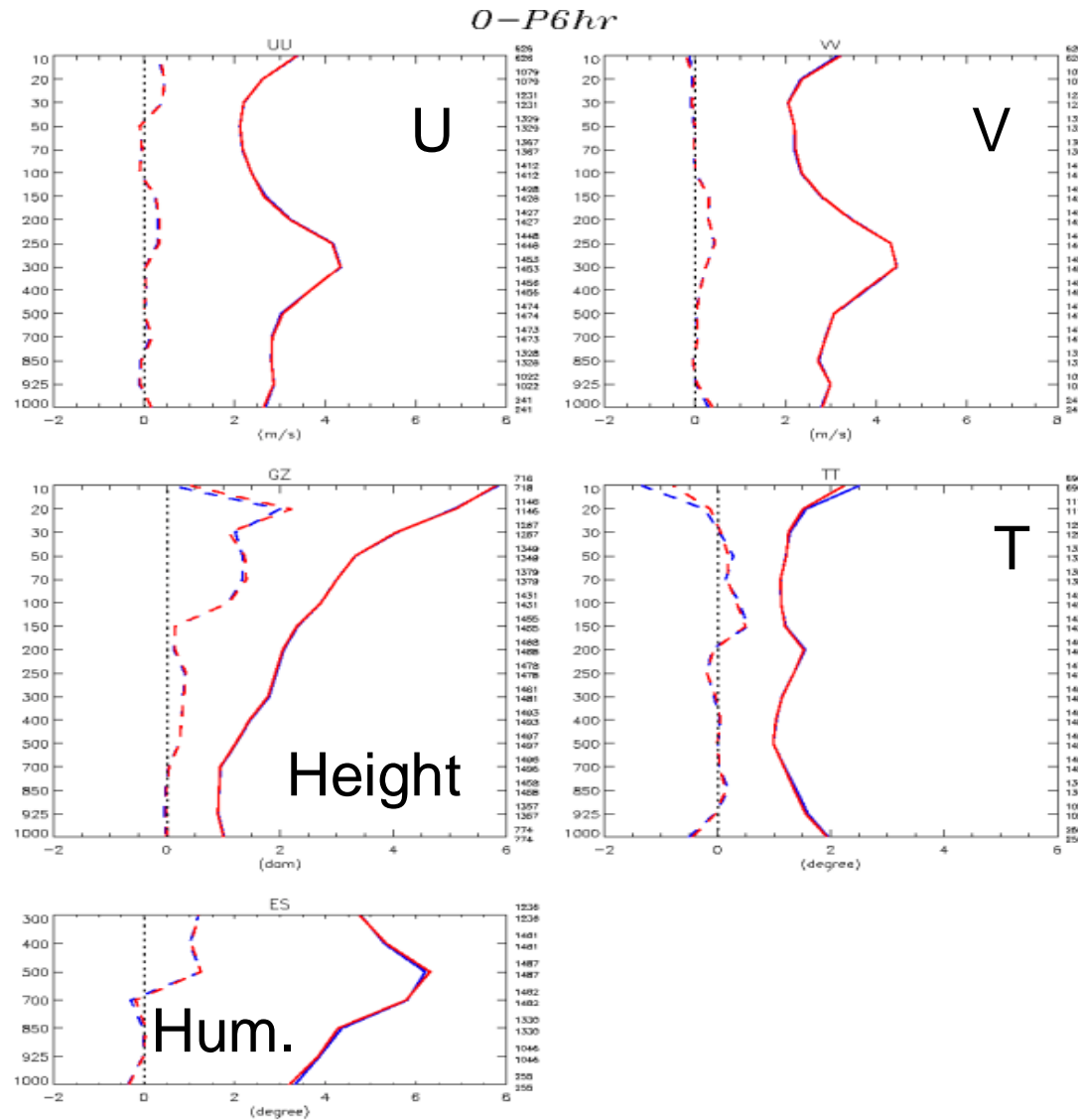
$$\mathbf{P}^f(\mathbf{t} + 1) = \mathbf{M}\mathbf{P}^a\mathbf{M}^T + 0.25\mathbf{P}_{3dvar}$$

Pour chaque membre, nous obtenons un champs aléatoire avec des statistiques isotropes.

Actuellement notre description de l'erreur du modèle a:

- Une composante balancée générée pour la fonction de courant Ψ . Après une transformation, nous obtenons une perturbation balancée pour les vents, la temperature et la pression de surface.
- Une composante non-balancée pour la température qui est importante près de la surface, près du toit et dans les tropiques.

Composante non-balancée pour la température



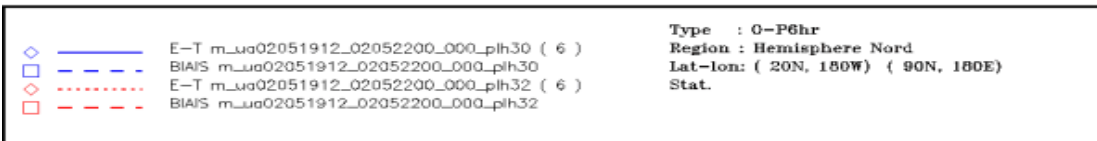
L'expérience de référence a une erreur du modèle pour Ψ uniquement.

La nouvelle expérience a aussi une composante non-balancée pour la température.

Les résultats sont meilleurs pour la température près du toit.

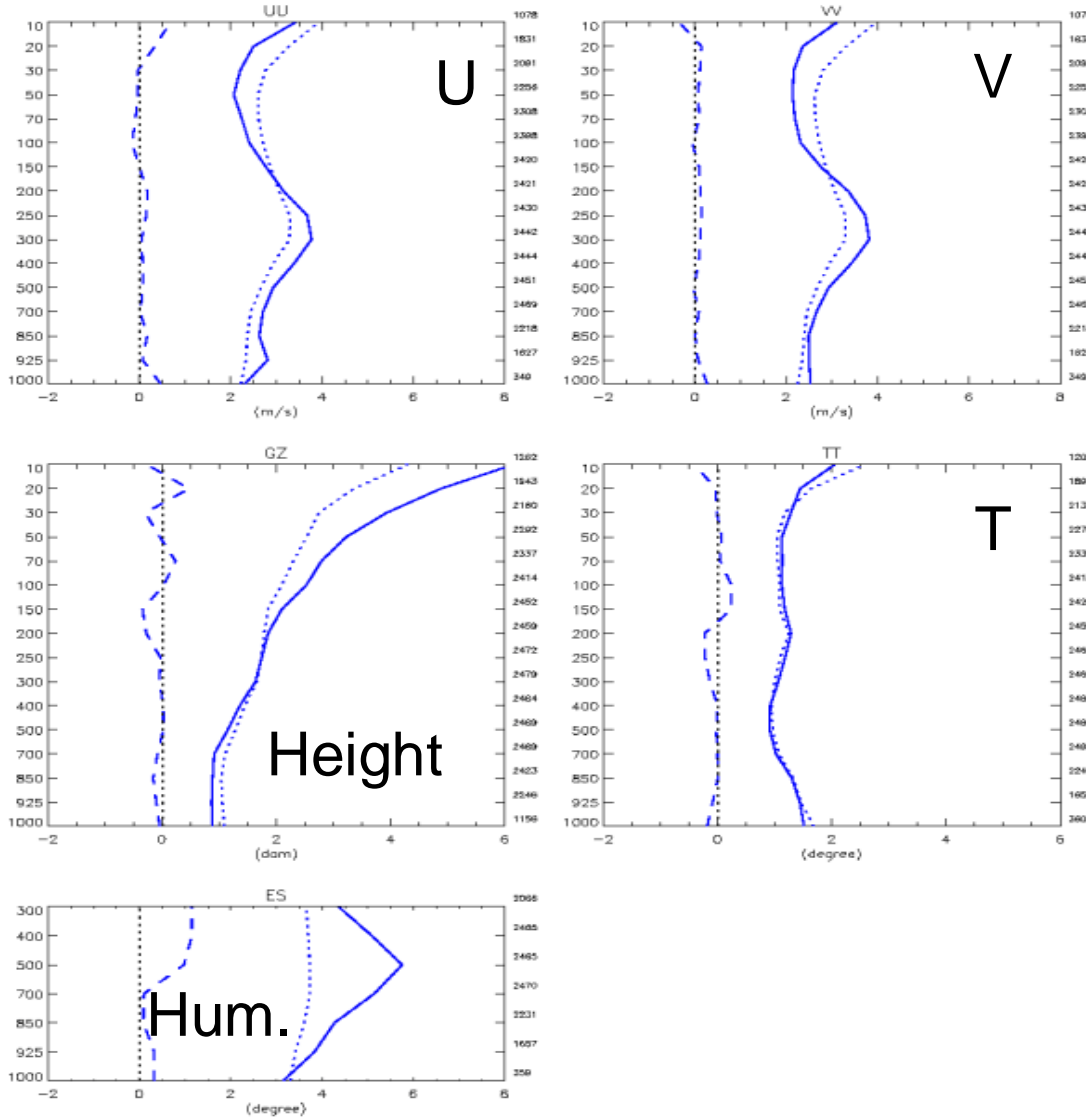
Nous considérons donc qu'il n'est peut-être pas nécessaire que l'erreur du modèle soit balancée.

Nous aurons donc plus de liberté pour simuler l'erreur du modèle de façon réaliste.



Quality of error statistics

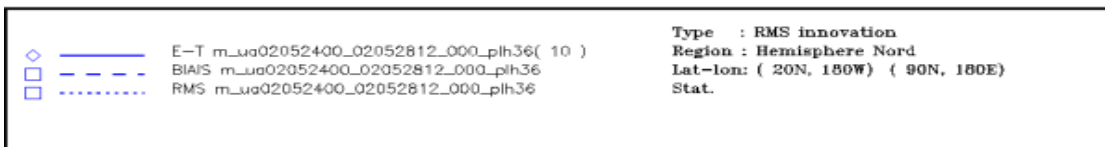
RMS innovation



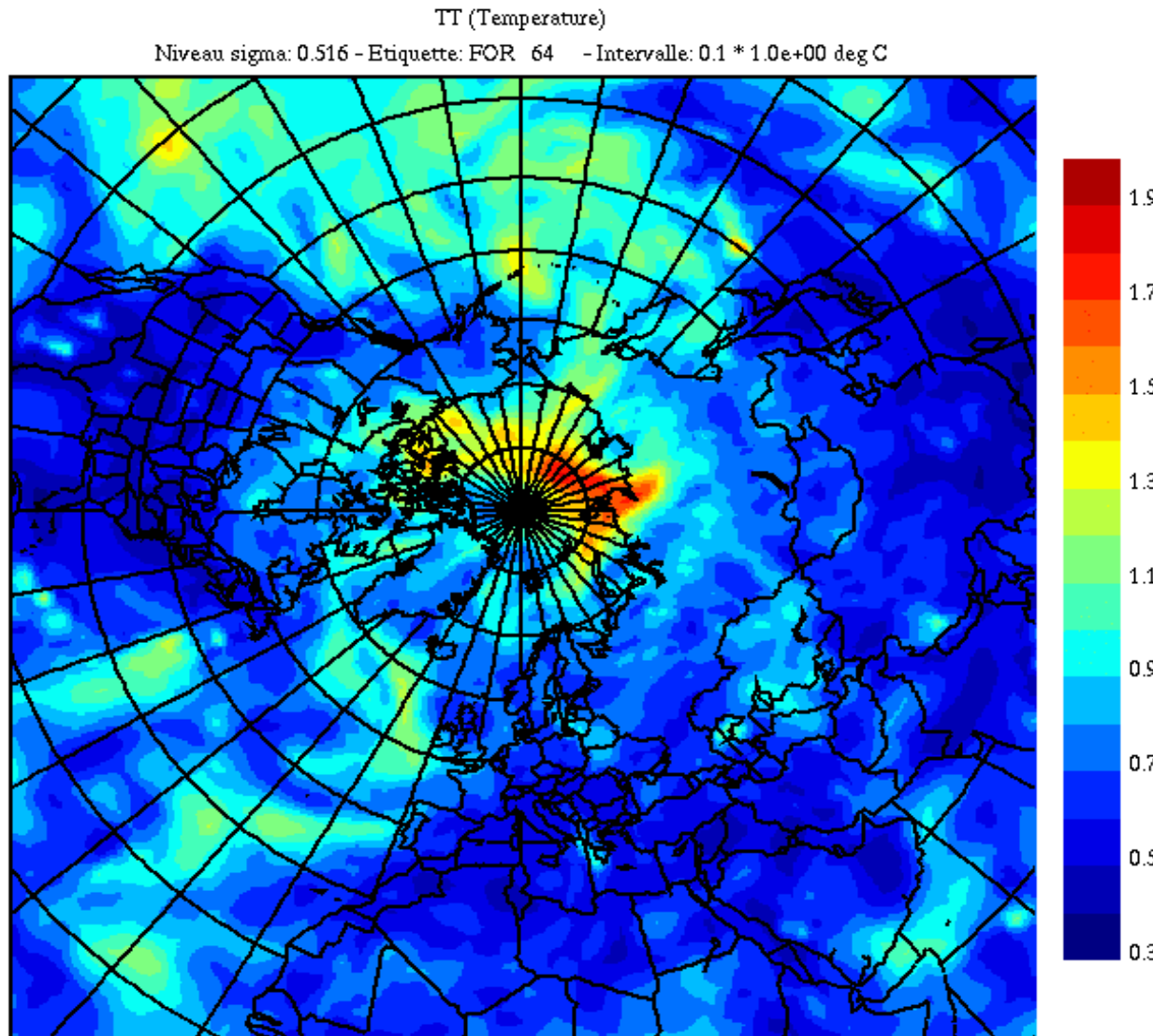
The solid line is the rms amplitude of the innovations for radiosonde observations.

The dashed line is the ensemble based prediction of the innovation amplitude. It is the root of the sum of observational variance and ensemble spread.

There is excellent agreement for the temperature. For winds the ensemble spread is too large near the model top. The spread is too small for humidity (need for some model error). For other areas the results are somewhat different (not shown).



L'écart-type dans un champs d'essai

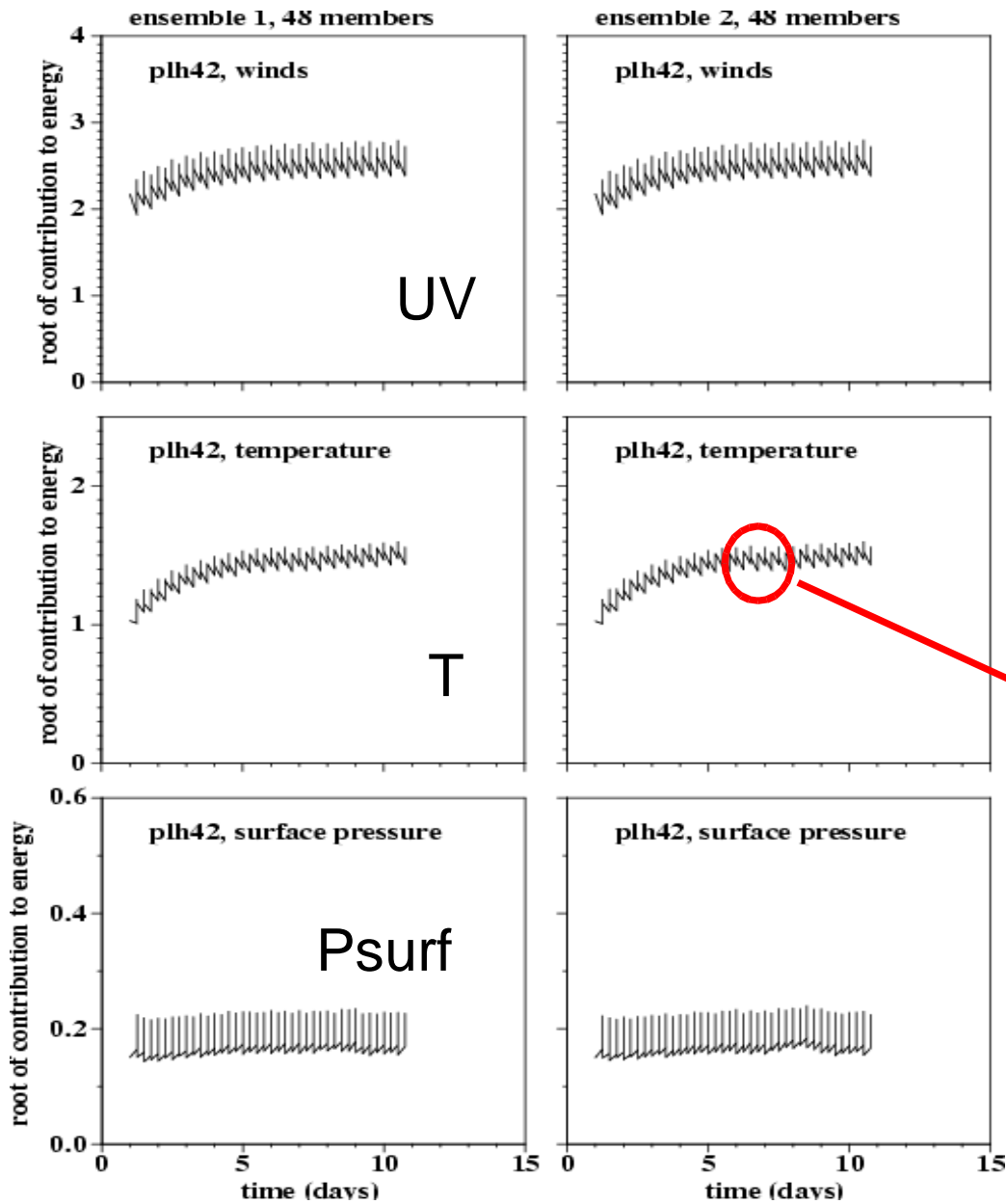


Pour la température au niveau $\eta = 0.516$, nous observons l'incertitude dans le champs d'essai (24 mai 2002 18 UT).

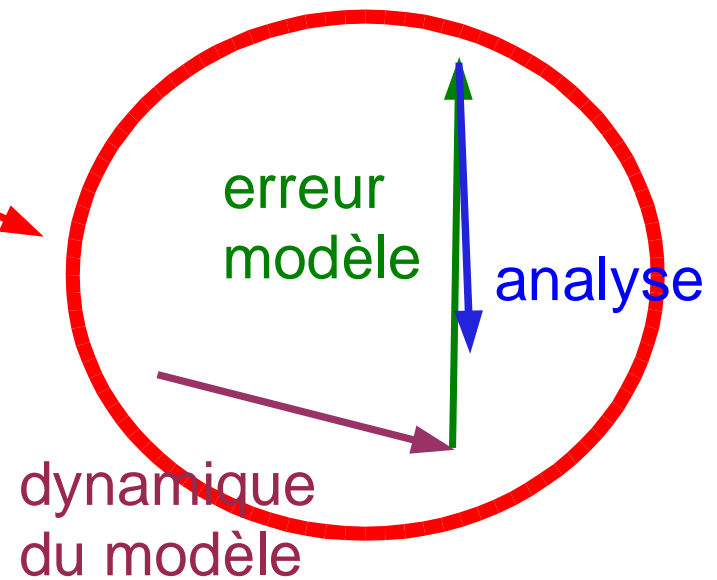
L'assimilation a réussi à réduire les erreurs sur les continents.

Il y a aussi des structures en apparence dynamique sur les océans.

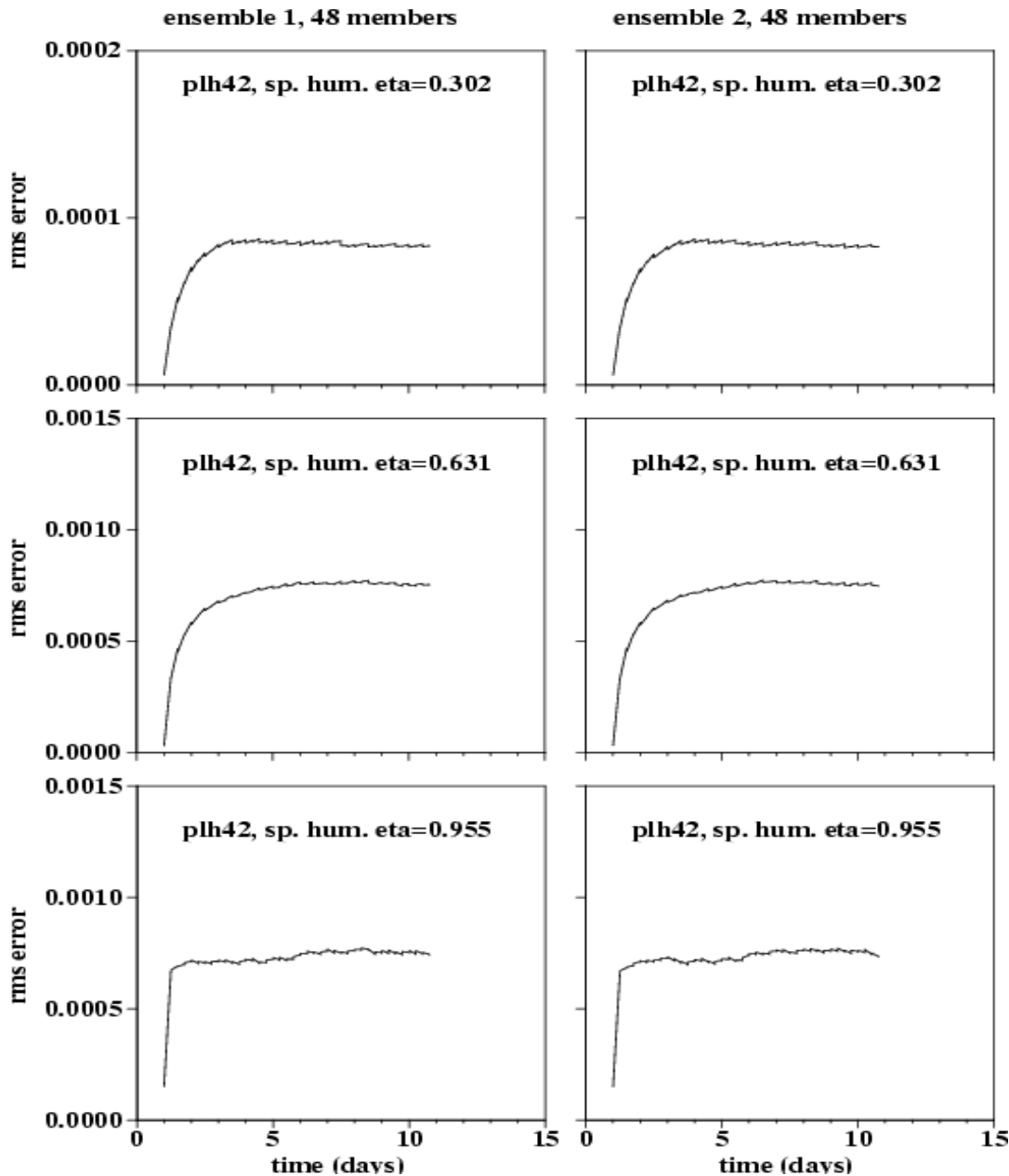
Dynamique des erreurs pendant le cycle d'assimilation



Il n'y a pas de croissance de l'erreur avec la dynamique!



Dynamique des erreurs pour l'humidité



Il y a très peu d'évolution pour l'humidité spécifique. L'analyse n'améliore pas le champs d'essai de façon systématique.

Avec l'EnKF, nous aurions pu espérer un effet positif des autres observations (u, v, T, radiances etc).

Comparaison avec le 3d-var

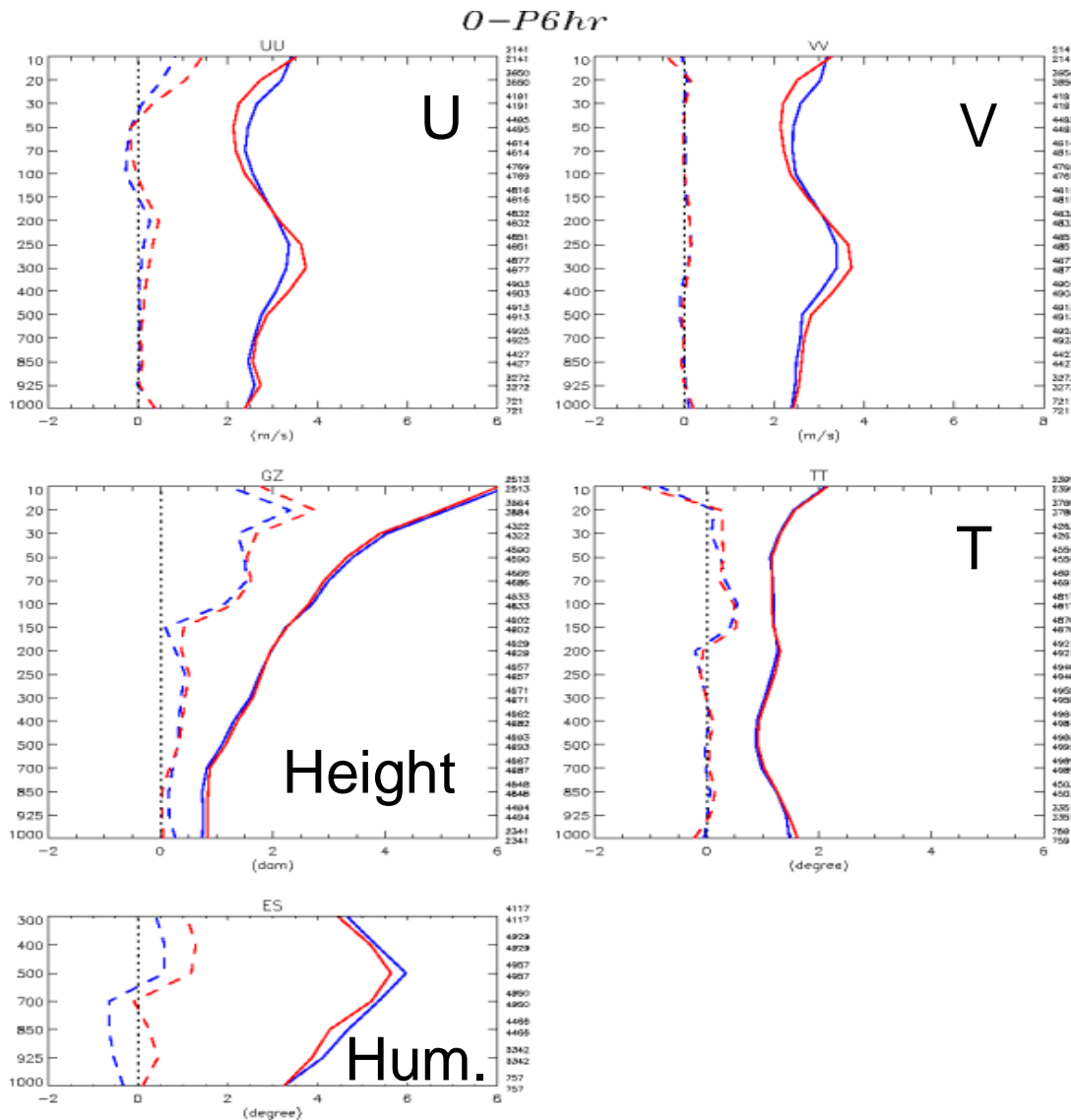
Avec l'EnKF, nous supplantons maintenant l'Interpolation Optimale (OI). C'est en partie grâce à un traitement moderne des observations (radiances TOVS-1b au lieu d'épaisseurs SATEM). Il est alors plus approprié de comparer l'EnKF avec un schéma moderne comme le 3d-var.

Pour l'EnKF, nous avons réalisé un cycle d'assimilation pour la période 2002051900 à 2002060212. On utilise 2 x 64 membres et le modèle GEM-DM à résolution 240 x 120. Les pôles du modèle sont aux pôles géographiques. Le filtre digital n'a pas été utilisé.

Pour le 3d-var, nous avons désarchivé les champs du `trials.glbpres2` et on a validé avec les mêmes observations avec `arcad`. Le modèle GEM avait une résolution de 400 x 200. Le 3d-var utilisait une grille de 240 x 120 pour "l'approche incrémentale".

Nous comparons les vérifications pour la période 2002052400 à 2002060212 (10 jours).

Comparaison avec le 3d-var (Hémisphère Nord)



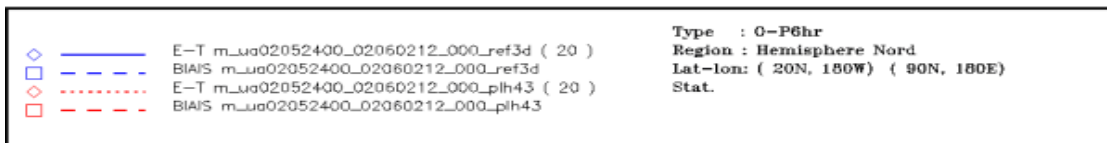
L'expérience en bleu vient des archives trial.glbpres2 (3d-var et GEM).

L'expérience en rouge est l'EnKF avec 2 x 64 membres à 240 x 120 avec GEM-DM.

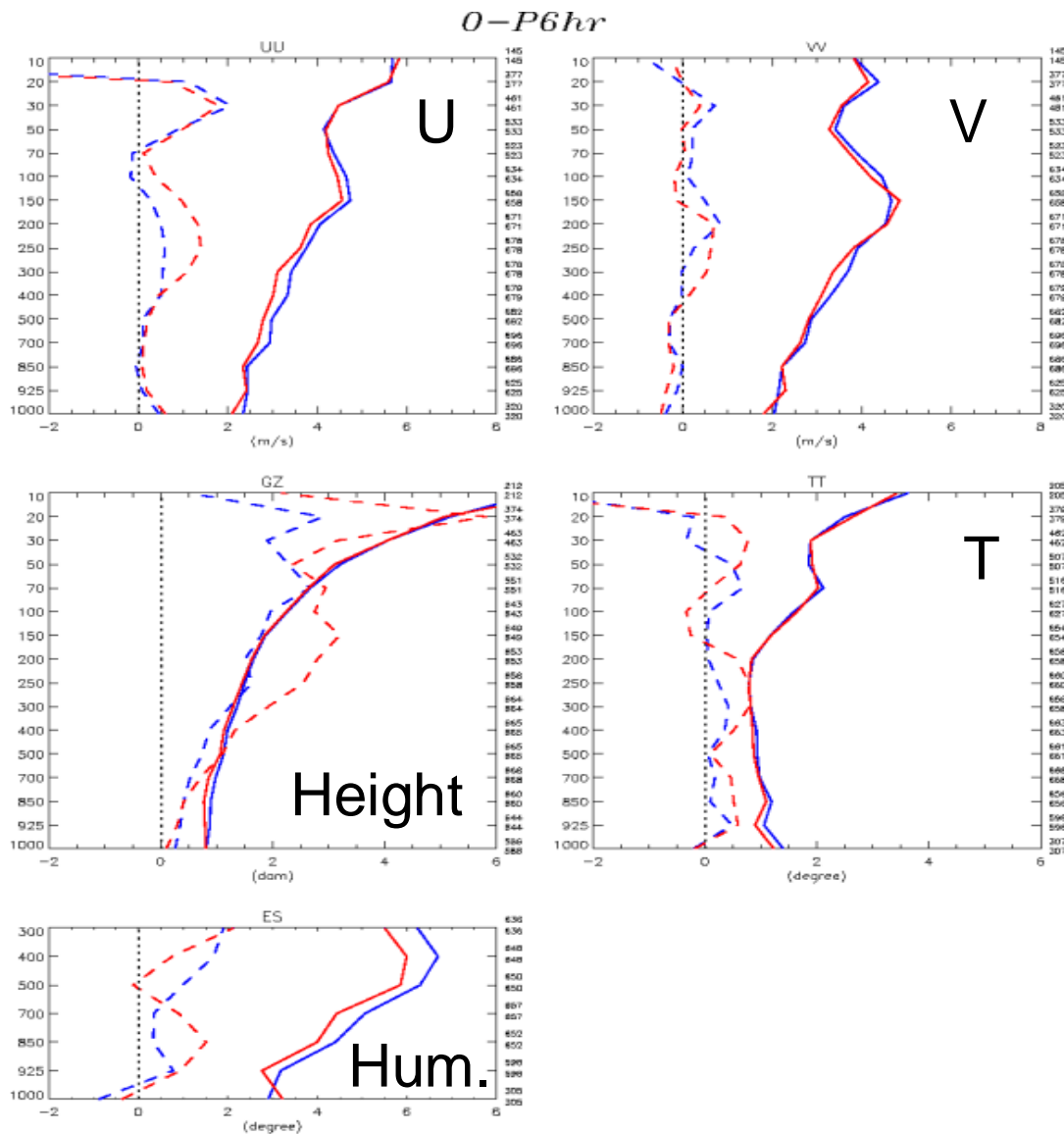
L'EnKF est meilleur pour les vents en haut de 200 hPa (les pôles ?) et pire pour les vents en bas de 200 hPa (la résolution?).

L'EnKF est pire pour les biais.

L'EnKF semble meilleur pour l'humidité.



Comparaison avec le 3d-var (Tropiques)



L'expérience en bleu vient des archives trial.glbpres2 (3d-var et GEM).

L'expérience en rouge est l'EnKF avec 2 x 64 membres à 240 x 120 avec GEM-DM.

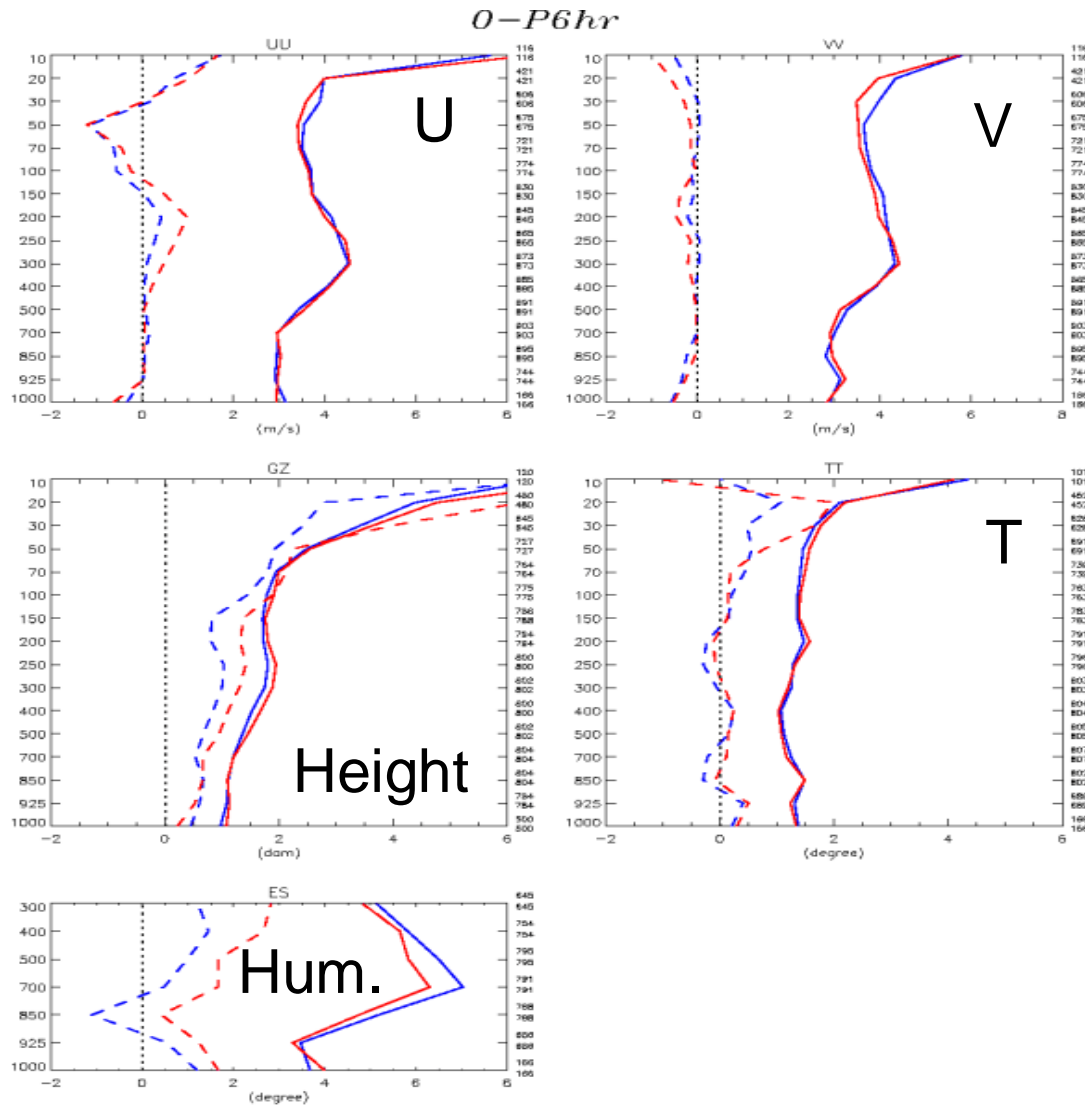
L'EnKF est meilleur pour les vents et l'humidité.

L'EnKF est pire pour les biais.

- ◇ — E-T m_u02052400_02060212_000_ref3d (20)
- - - BIAIS m_u02052400_02060212_000_ref3d
- ◇ - - E-T m_u02052400_02060212_000_plh43 (20)
- - - BIAIS m_u02052400_02060212_000_plh43

Type : 0-P6hr
 Region : Tropiques
 Lat-lon: (20S, 180W) (20N, 180E)
 Stat.

Comparaison avec le 3d-var (Hémisphère Sud)

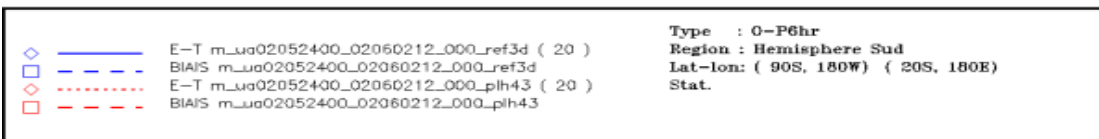


L'expérience en bleu vient des archives trial.glbpres2 (3d-var et GEM).

L'expérience en rouge est l'EnKF avec 2 x 64 membres à 240 x 120 avec GEM-DM.

L'EnKF est meilleur pour les vents et l'humidité et pire pour les hauteurs.

L'EnKF est pire pour les biais.



Conclusions

Le filtre de Kalman d'ensemble peut être utilisé pour l'assimilation des données atmosphériques.

La qualité de la moyenne d'ensemble est comparable à celle de l'analyse 3d-var à 400 x 200. La qualité est supérieure à celle de l'Interpolation Optimale (OI) utilisée pour le système de prévision d'ensemble de CMC.

Pourquoi est-ce qu'il y a si peu de différences entre les vérifications de l'EnKF et le 3d-var?

Hypothèses:

- 1) Toutes les observations d'une fenêtre de 6 heures sont assimilées au moment central (pour l'EnKF et le 3d-var).
- 2) La description de l'erreur du modèle ne dépend pas de la situation météorologique (et vient du 3d-var).
- 3) ???

Des projets futurs d'échange et de développement:

Publication scientifique.

Notre code est disponible (et documenté) pour des usagers à l'intérieur de SMC.

Établissement des standards de sorte que CMC et MRB peuvent utiliser les mêmes scripts pour l'EnKF.

Migration de la version MPI de RTTOV6 vers une version MPI de RTTOV7.

Pour plus d'information:

<http://iweb.cmc.ec.gc.ca/~armabha/enkf> et

<http://www.cmc.ec.gc.ca/rpn/enkf>