Department of Earth and Atmospheric Sciences
Université du Québec à Montréal (UQAM)

Near-Surface Wind Data Assimilation using A Geo-Statistical Observation Operator – Results from OSE

Joël Bédard[1], Stéphane Laroche[2] and Pierre Gauthier[1]

1) Department of Earth and Atmospheric Sciences, UQAM, Canada
2) Data Assimilation and Satellite Meteorology Section, Environment Canada, Canada

2015 AMS WAF-NWP conference (Chicago)

Contact: *Joël Bédard*, Email: *bedard.joel@gmail.com*

Research Framework

Problem limiting near-surface wind data assimilation:

- Representativeness and systematic errors (Ingleby, 2014)
  - Mainly due to subgrid scale topographic effects
  - Representativeness errors limits the use of near-surface wind observations

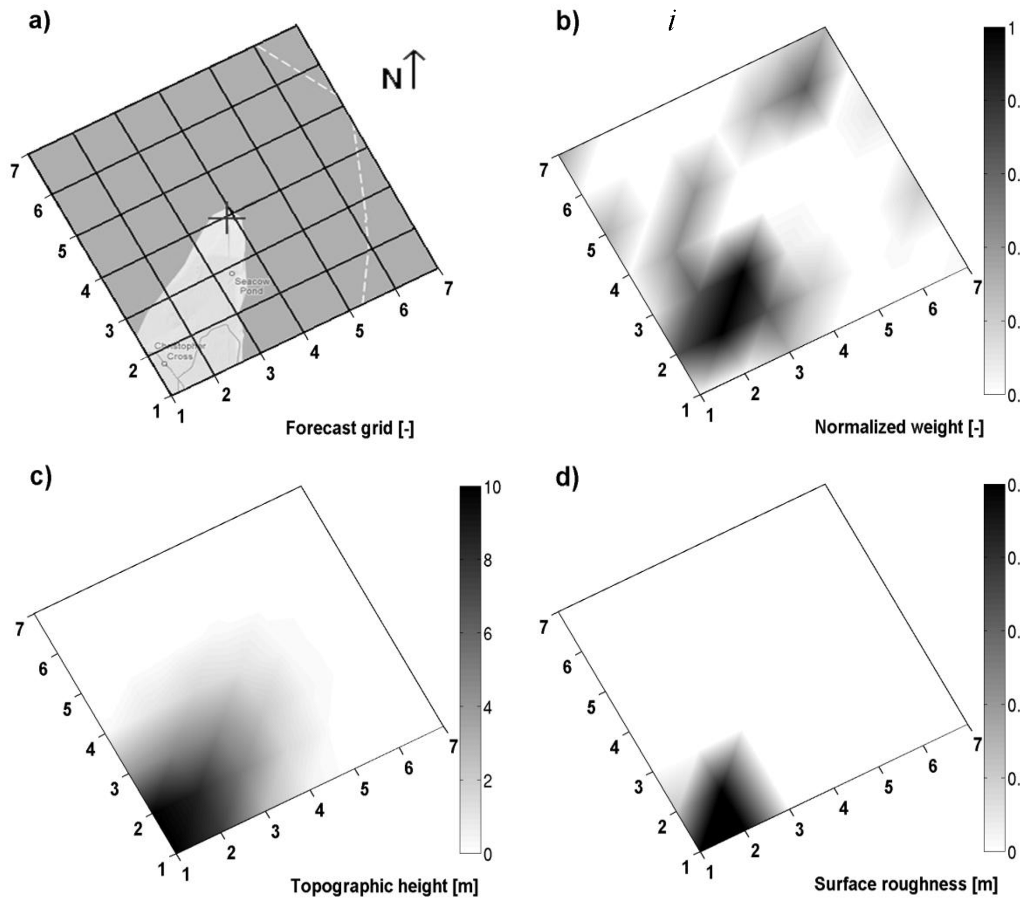
Development of a geo-statistical observation operator

- Reduces representativeness errors
- Eliminates biases
- Improves local forecasts in the very short-term
  - 3h to 6h lead times
- Increases the coherence between forecasts and analyses
  - Up to 48h

Observation Operator (Bédard et al. 2013)

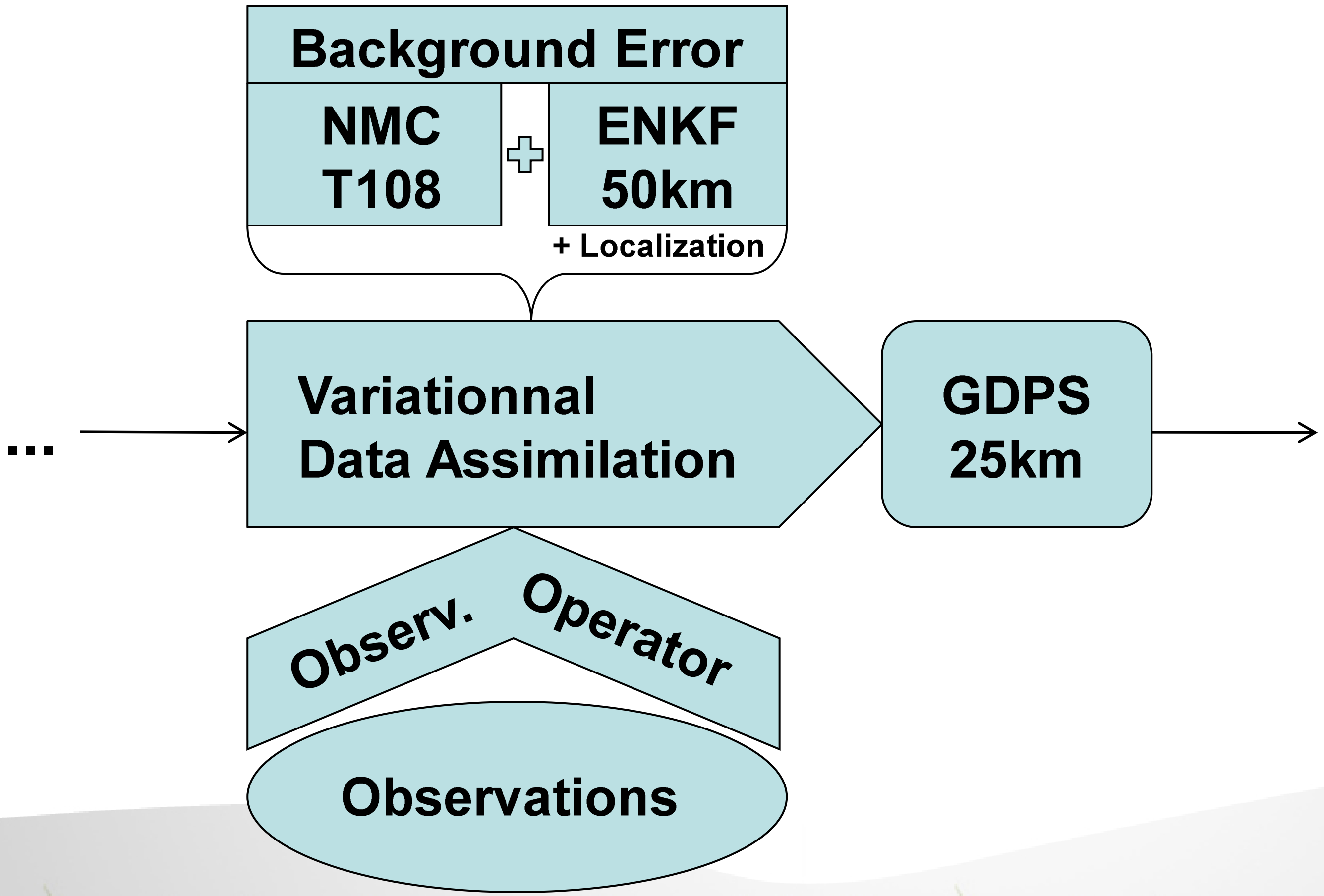
Interpolate model state variables to observation space

- Geophysic Model Output Statistics (GMOS) $$H(X) = \sum_{i} (A_i X_i) + C$$
  - Reduces representativeness errors by selecting the forecast point the most representative of the site
  - Corrects systematic errors due to sub-grid scale topographic effects



Data Assimilation System

Ensemble variational data assimilation system (4D-EnVar):



$X_a = X_b + K (Y - H(X_b))$

X_a: Analysis fields

X_b: Background fields

Y: Observations

K: Gain matrix

…

$K = B H^T (R + H B H^T)^{-1}$

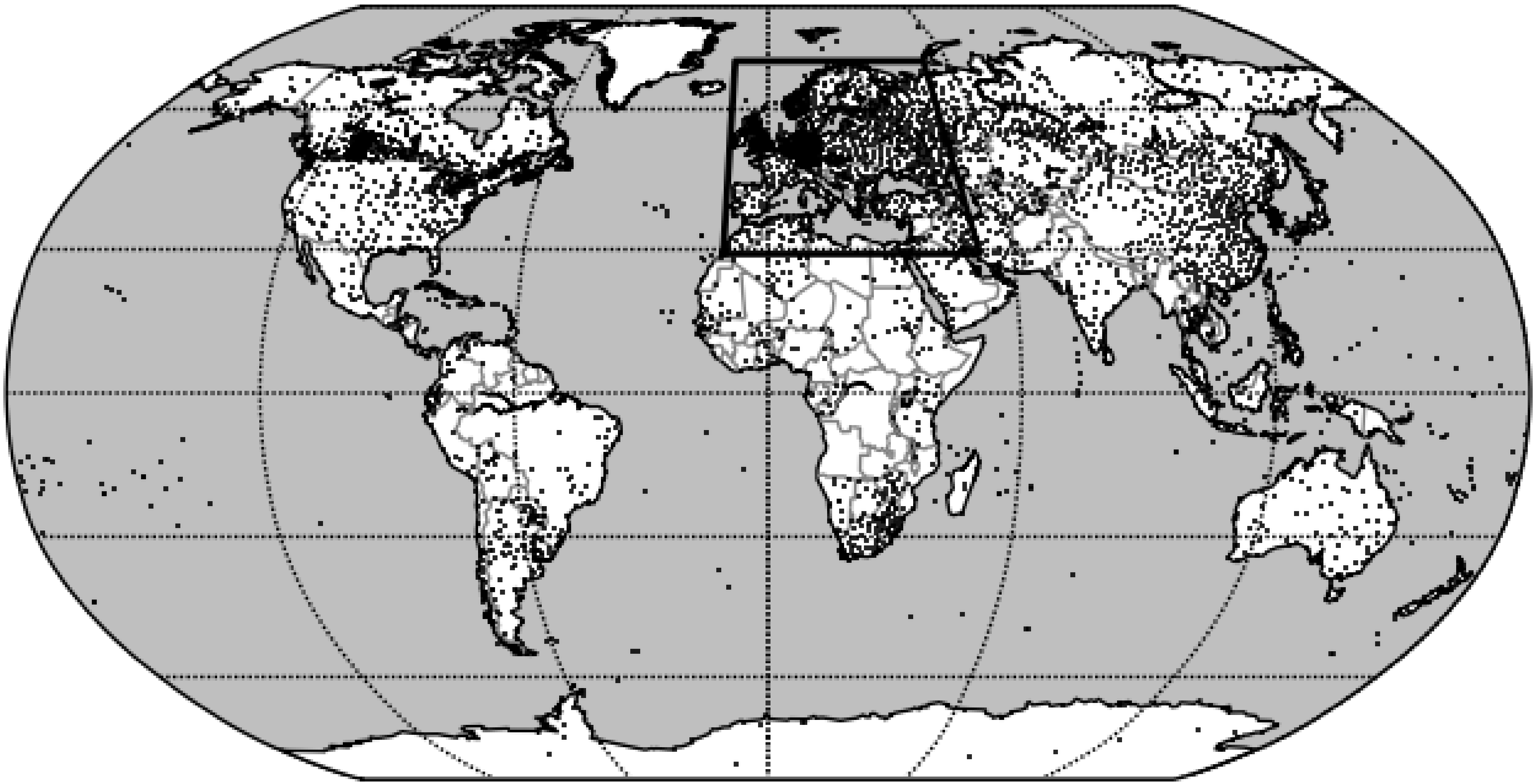
H: Observation operator

B: Background error covariances

R: Observation error covariances

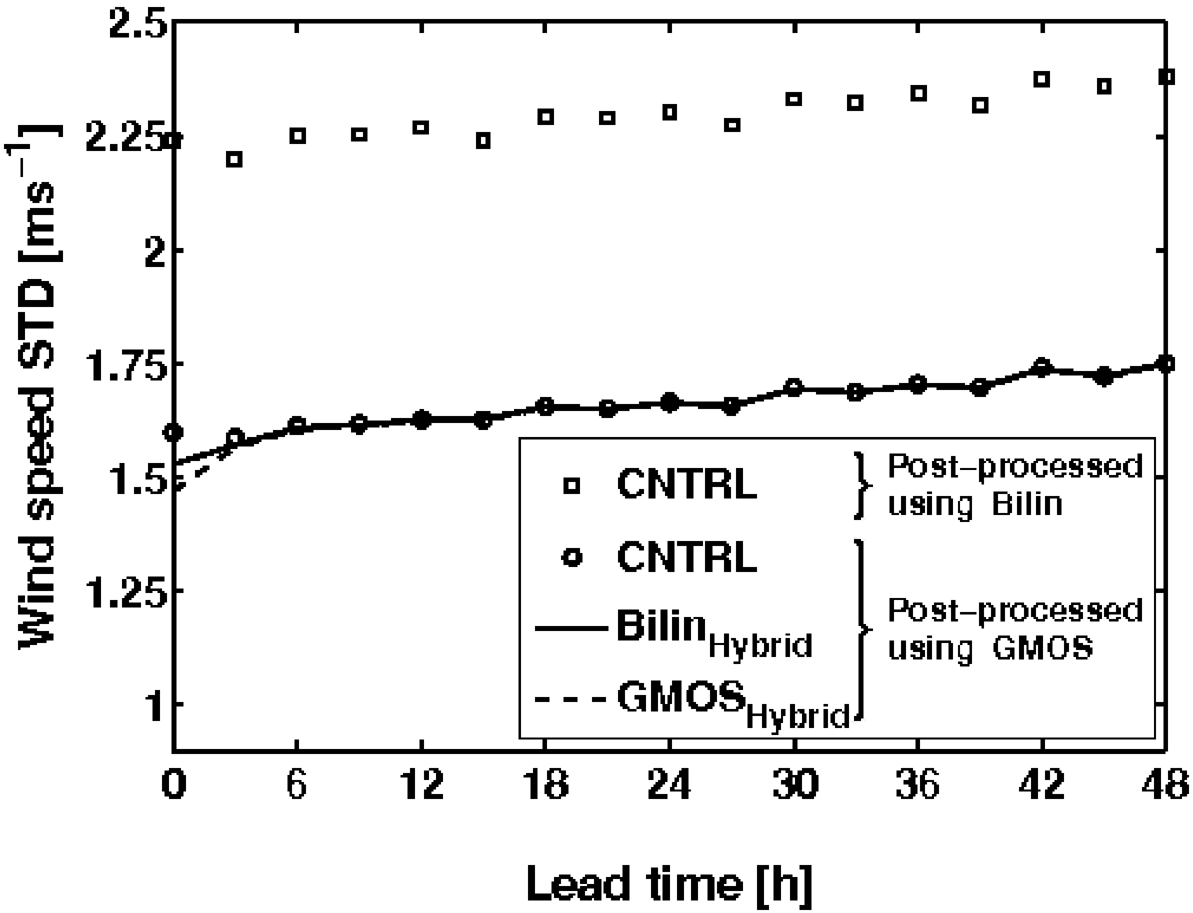
Observation Assimilated

Wind observations from 4942 SYNOP stations



Observation Impact

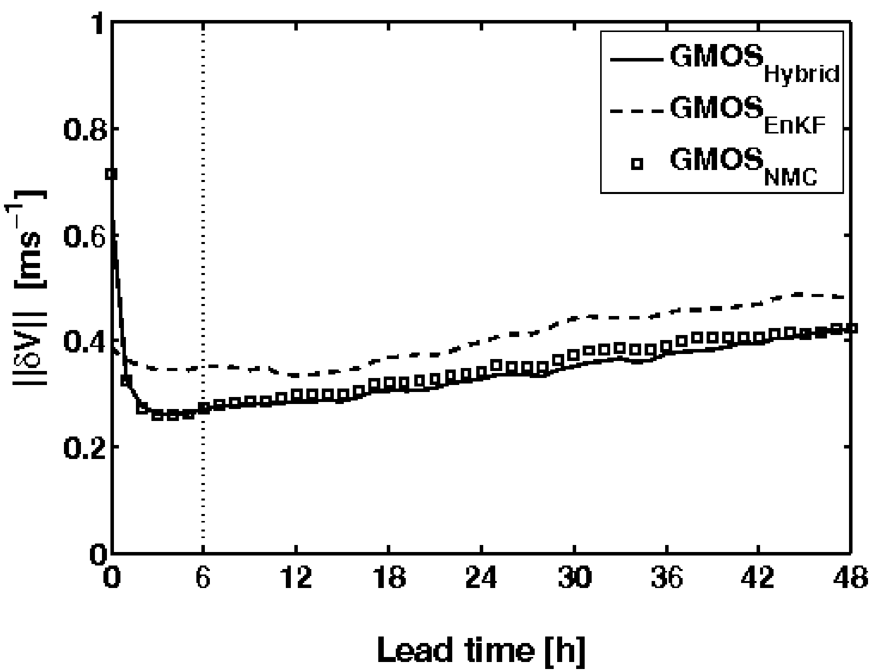
Forecast skill (evaluated against surface obs.) from the assimilation of SYNOP wind data vanishes quickly (< 6h)

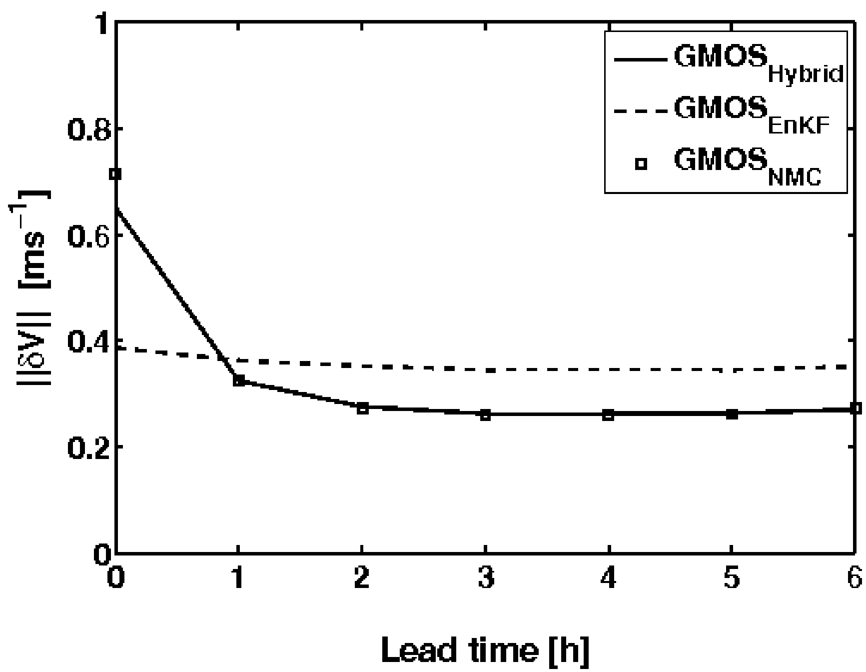


Forecast differences

Mean evolution of the forecast differences between the control (CNTRL) and the experiment (GMOS) runs

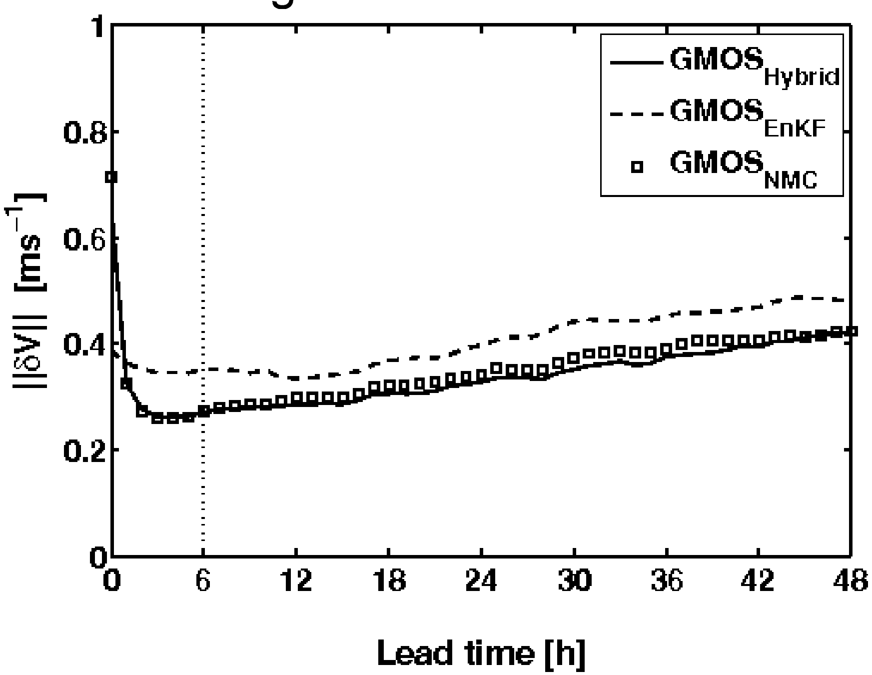
- CNTRL: no obs. are assimilated (the 6h background from the operational system is simply reinitialized to generate the following 48h forecasts)
- GMOS: using the same background as CNTRL, near-surface wind observations are assimilated using the geo-statistical observation operator

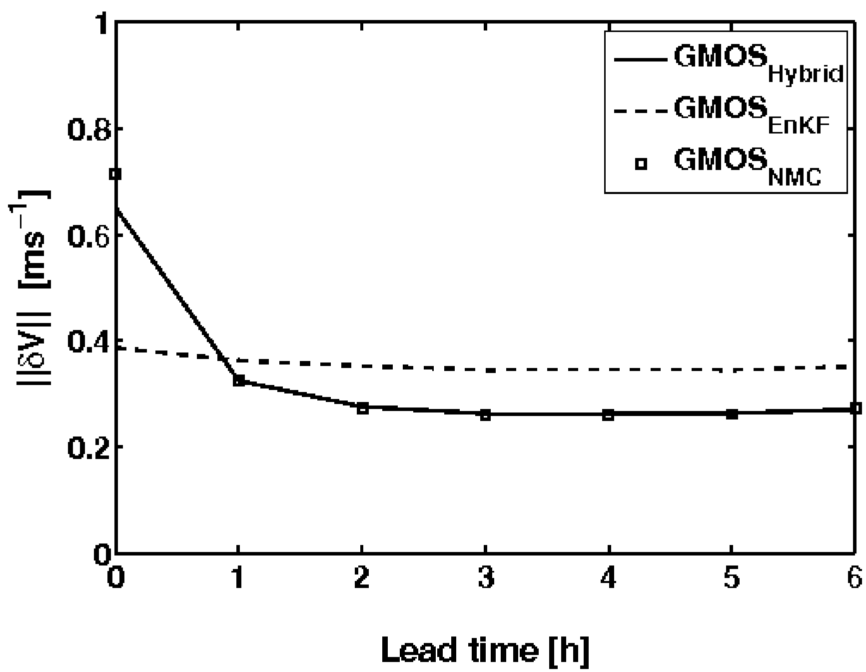




UQÀM CRSNG NSERC Hydro Québec Environment Canada

- **The NMC experiments (using homogeneous and isotropic background error covariances) produce the worst results**
  - rapid forecast difference reduction
  - most of the information from the observations is damped within 3h
- **The EnKF covariances produces better results**
  - most of the increment is propagated within the forecast
  - the forecasts only slightly converge before diverging due to perturbation growth within the NWP model.





Initial tendencies

Studying the momentum prognostic equation to assess the impact of the different background error statistics

$\mathbf{T}_{adv}$: advection;

$\mathbf{T}_{cori}$: Coriolis effect;

$\mathbf{T}_{v.d.}$: vertical diffusion;

$$\frac{\partial \mathbf{V}}{\partial t} = \mathbf{T}_{adv} + \mathbf{T}_{cori} + \mathbf{T}_{p.g.} + \mathbf{T}_{v.d.} + \mathbf{T}_{o.b.} + \mathbf{T}_{h.d.}$$

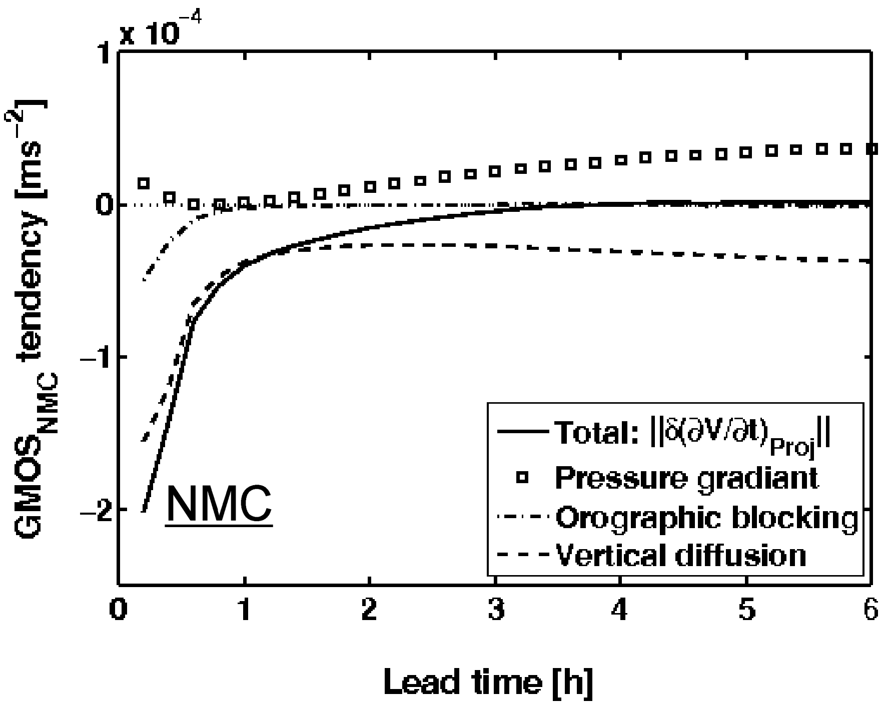
$\mathbf{T}_{h.d.}$: horizontal diffusion; $\mathbf{T}_{p.g.}$: pressure gradients; $\mathbf{T}_{o.b.}$: orographic blocking

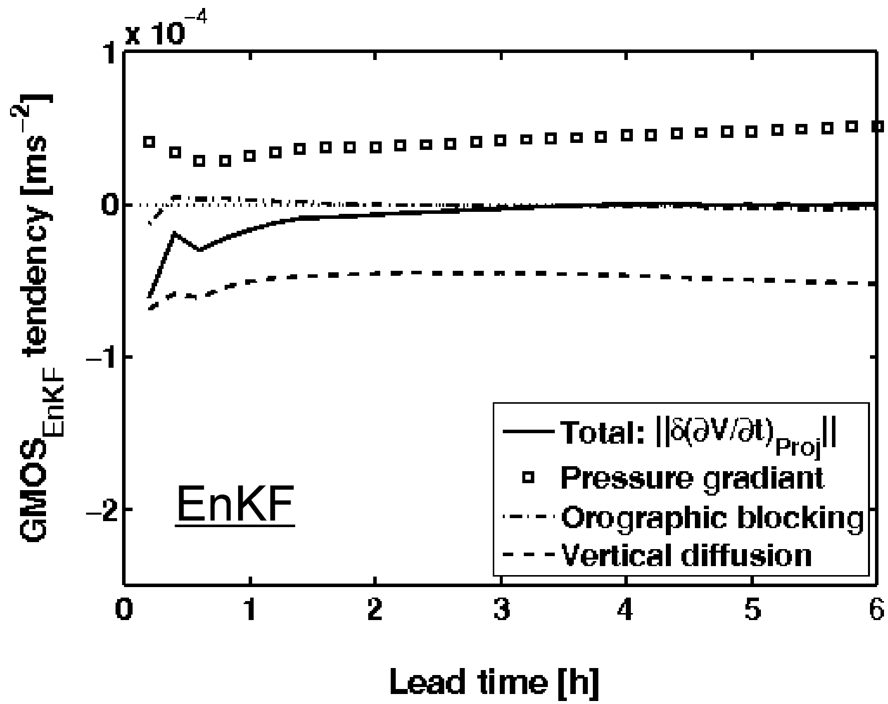
Systematic initial tendencies for the first time steps are quantified to assess the reduction of the differences

- The evolution of the forecast difference between CNTRL and the experiments ($\partial(\delta\mathbf{V})/\partial t$) can then be estimated by projecting $\delta(\partial\mathbf{V}/\partial t)$ on $\delta\mathbf{V}$ which approximates the evolution of the local analysis increments.

Initial tendencies

- NMC: error covariances between wind and mass fields near the surface are small
- NMC: vertical diffusion and orographic blocking damp the wind increments
- NMC: forecast difference between CNTRL and the $GMOS_{NMC}$ decreases rapidly
- EnKF: pressure gradient forces balance vertical diffusion and orographic blocking
- EnKF: the information from the obs. is propagated within the forecast ($\delta(\partial V/\partial t) \approx 0$)



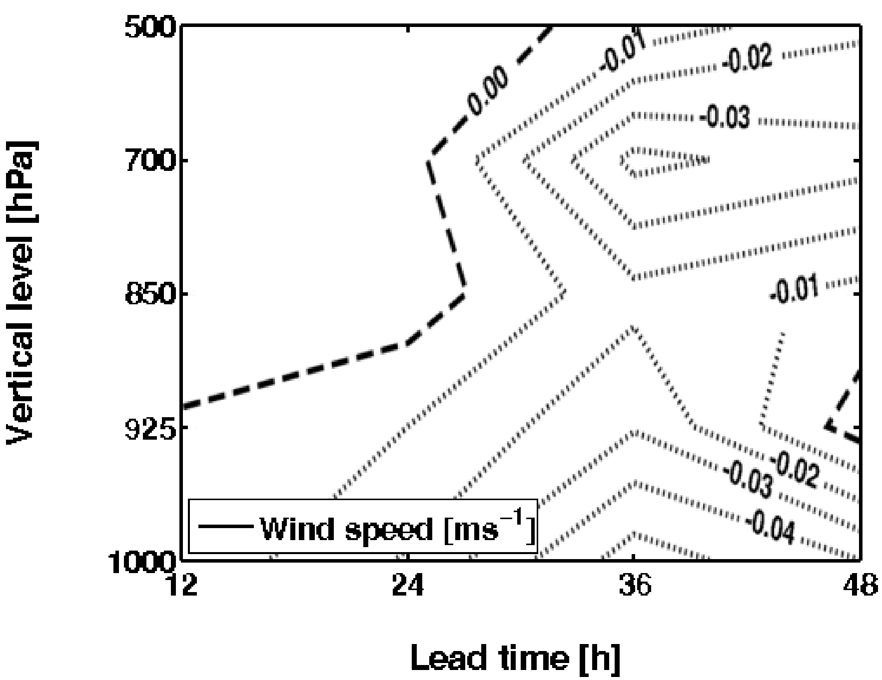


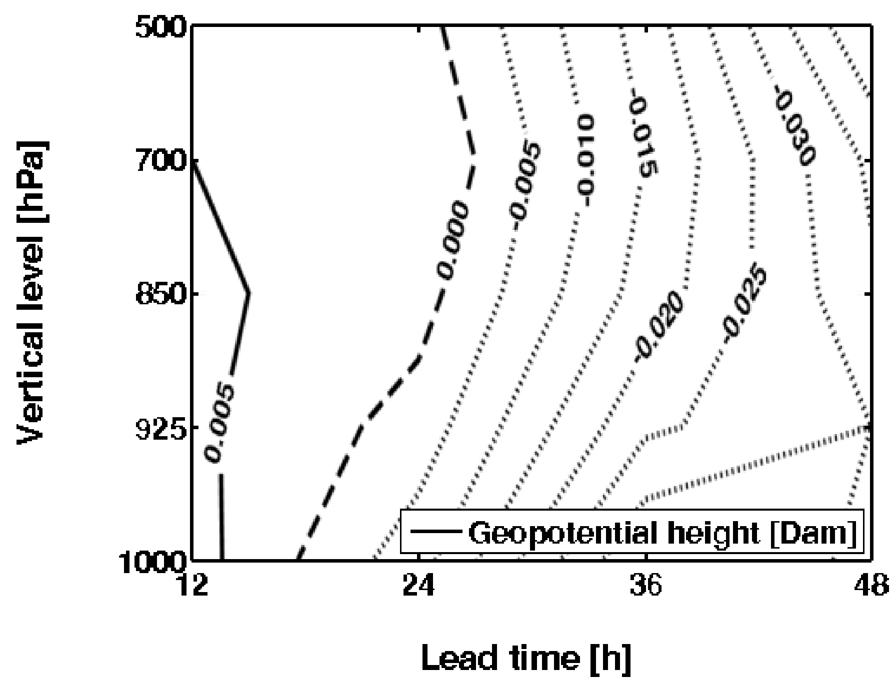
Observing System experiment (Bilin)

OSE evaluation against own analyses

- The scores for geopotential height (right) are generally negative at t > 24h
- The scores for wind speed (left) are also mostly negative

Forecasts and analyses from Bilin are not as consistent as those from the CNTRL run (the operational system at EC)



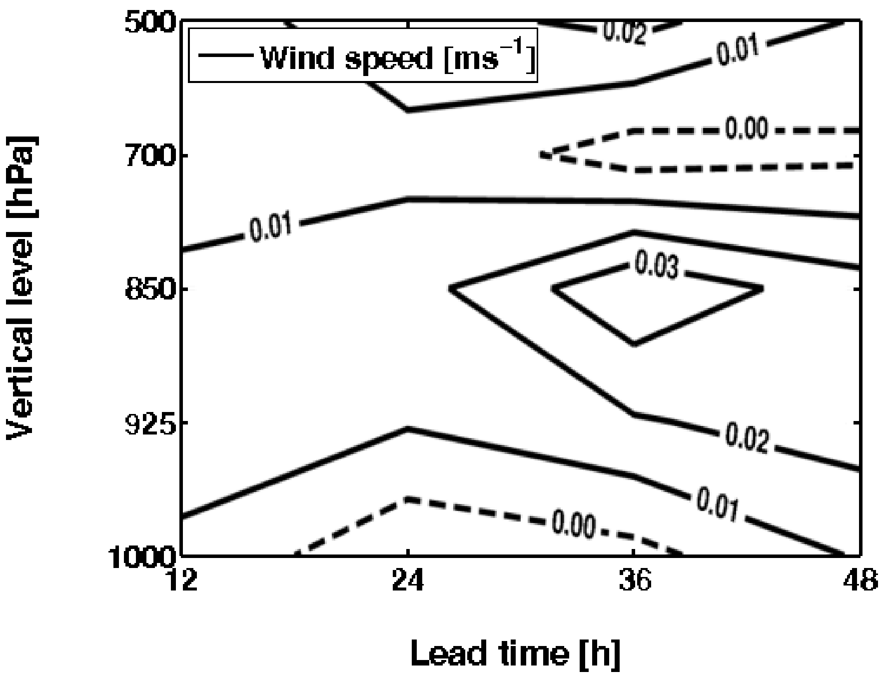


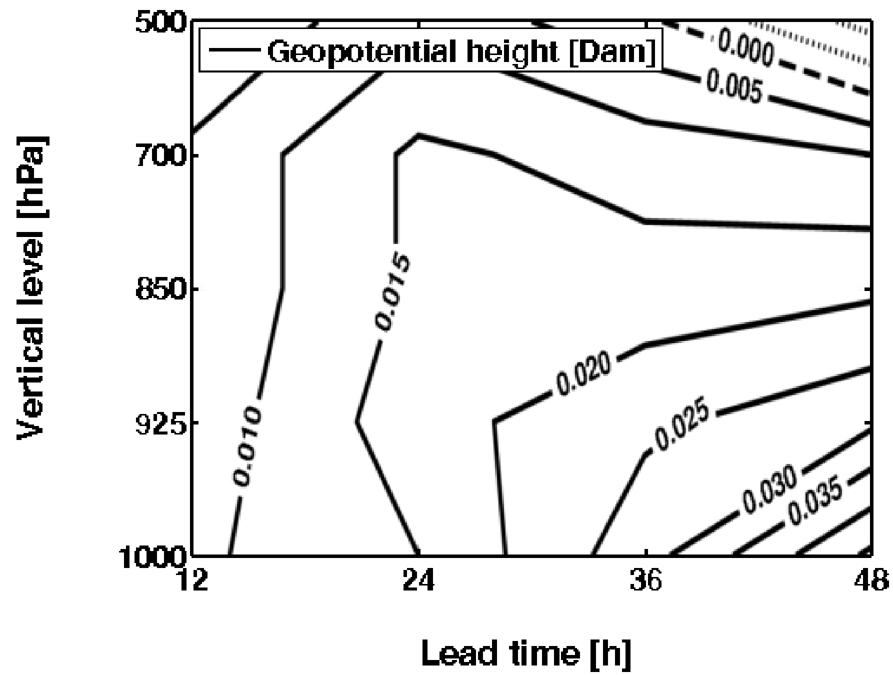
Observing System experiment (GMOS)

OSE evaluation against own analyses

- The wind speed (left) and geopotential height (right) scores are mostly positive
- Upper level gz scores (< 700 hPa) slowly fade out and become negative at 48h

Forecasts and analyses issued from the GMOS experiment are more in agreement than those from the CNTRL run

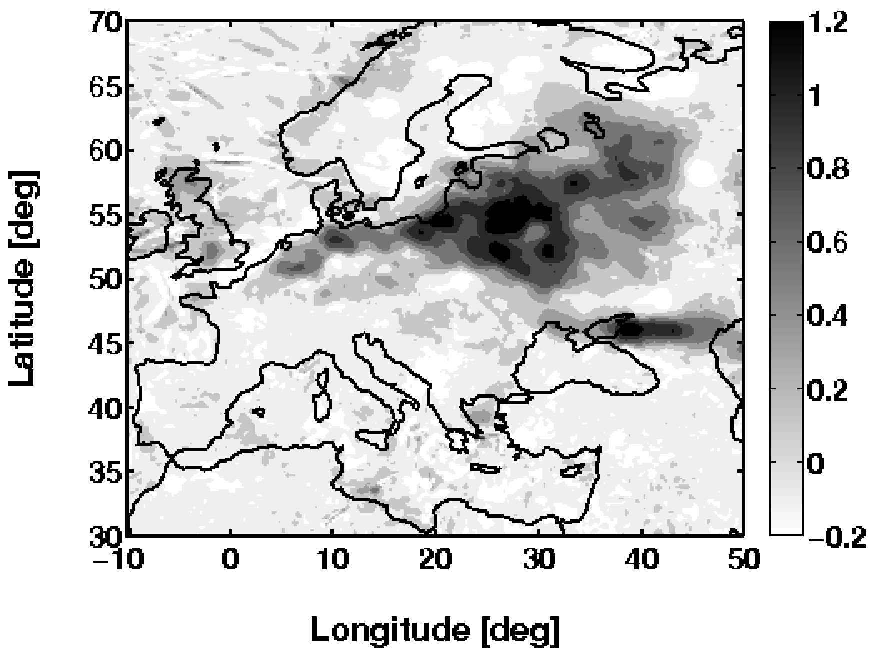


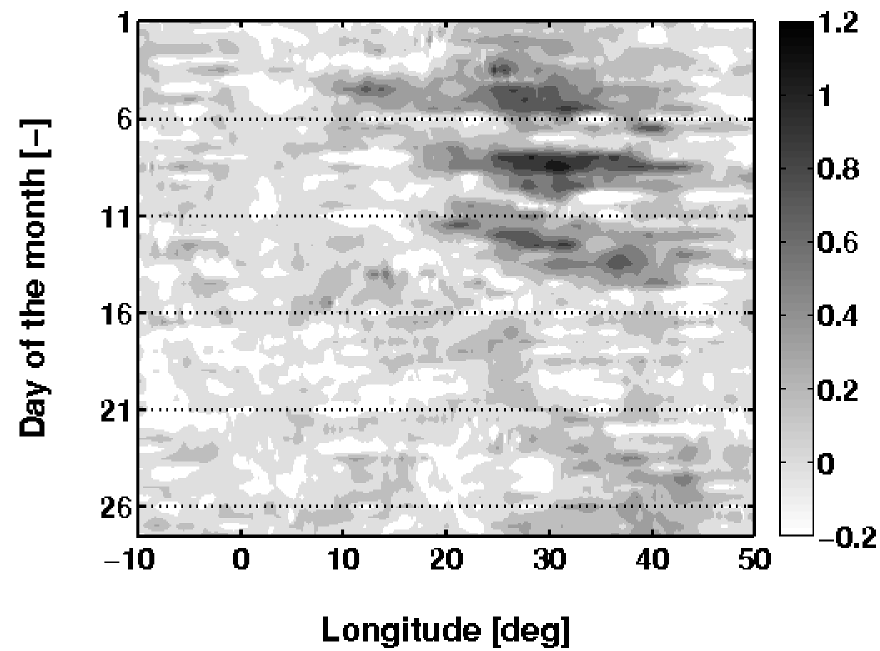


Results from OSE (GMOS vs Bilin)

Spatial and temporal distributions of verification scores

- Validation of near-surface winds (12h forecasts) for the february 2011 case
  - GMOS (Bilin) forecasts are more (less) coherent with the analyses for the area impacted by synoptic activity during the evaluation period (dark color)
  - Verification against radiosondes does not display GMOS improvements as northern and eastern Europe countries operate few radiosonde stations





Conclusion

The local forecast skill from the assimilation of SYNOP wind data vanishes quickly (< 6h)

- The spatiotemporal propagation of the information is strongly limited by the quality of the background error covariances
  - Homogeneous and isotropic error stats produce unbalanced increment
  - Flow dependent background error statistics modify both wind and mass fields in a coherent way (through multivariate correlations)
- Unless counterbalanced by proper pressure gradient forces, the atmospheric boundary layer parameterization schemes cause information from near-surface wind observation to be quickly damped

Results from OSEs

- By accounting for representativeness errors and biases, GMOS allows for a better use of near-surface wind observations and provides forecasts that are more coherent with analyses

PhD supervision

- **Pierre Gauthier (director)**,

  Département des sciences de la Terre et de l'atmosphère, UQAM
- **Stéphane Laroche (co-director)**,

  Meteorological Research Division, Environment Canada

References

- Bédard J, Yu W, Gagnon Y, Masson C. 2013. Development of a geophysic model output statistic module for improving short-term numerical wind predictions over complex sites. *Wind Energy*, **16** : 1131–1147.
- Bédard J, Laroche S, Gauthier P. 2015a. A geo-statistical observation operator for the assimilation of near-surface wind data. *Quart. J.R. Meteor. Soc.*, Accepted on April 22nd, 2015. DOI: 10.1002/qj.2569
- Bédard J, Laroche S, Gauthier P. 2015b. Near-Surface Wind Data Assimilation Using a Geo-Statistical Observation Operator. *Quart. J.R. Meteor. Soc.*, Submited on June 1st 2015.

Supplementary Material

Observation Operator

GMOS sensitivity to the size of the training period used

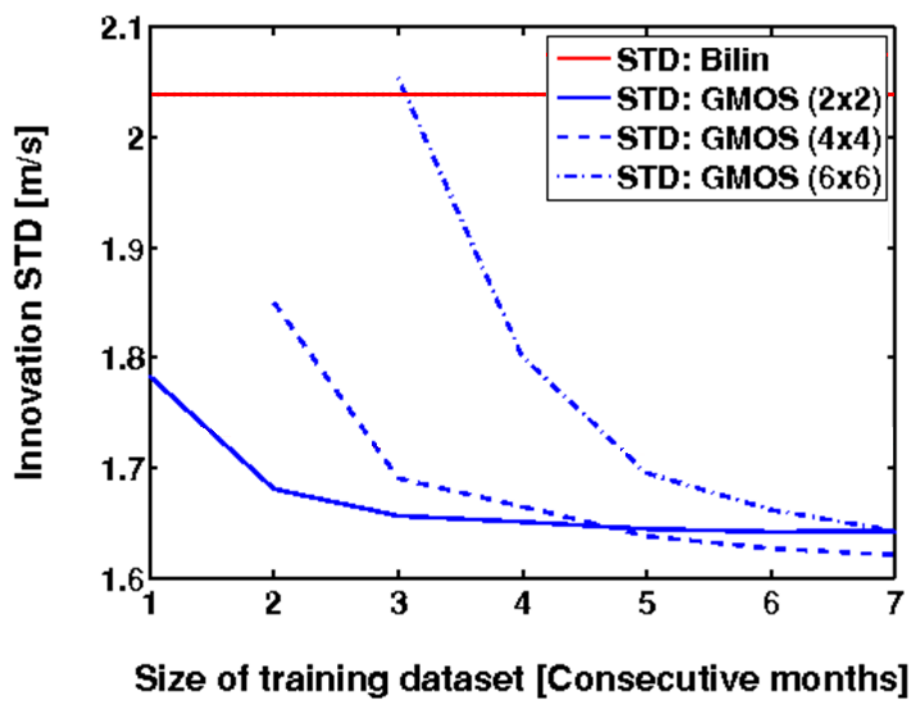
- GMOS using NxN grid-points need at least N month of training data
- Bias correction and representativeness error reduction (~0.4 m/s)

GMOS 2x2 grid-point architecture is used for comparison

- Same grid-points as bilinear
- Most robust GMOS scheme

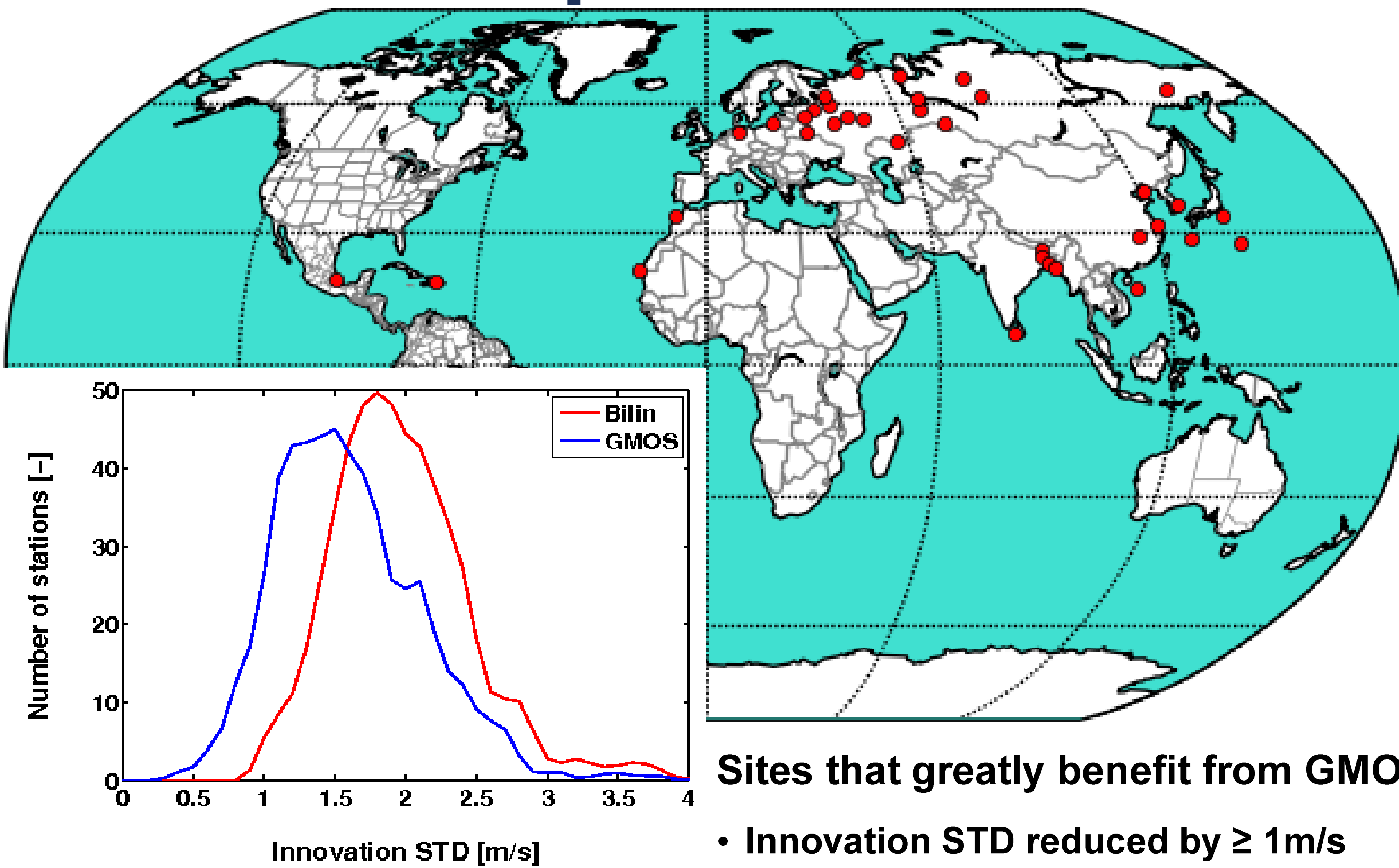
Validation methodology

- Evaluation using:
  - Surface wind observations from SYNOP stations
  - Radiosonde observations
  - Own analyses



UQÀM CRSNG NSERC Hydro Québec Environment Canada

Observation Operator



Sites that greatly benefit from GMOS

- **Innovation STD reduced by ≥ 1m/s**

Analysis Error Statistics

Analysis error based on independent observations

- Non-assimilated colocated radiosonde observations

$$\mathbf{y}_{Raob} - H_{Raob}(\mathbf{x}_a) = \mathbf{y}_{Raob} - H_{Raob}(\mathbf{x}_b + \mathbf{K}(\mathbf{y}_{Sfc} - H_{Sfc}(\mathbf{x}_b)))$$
$$= (\mathbf{y}_{Raob} - H_{Raob}(\mathbf{x}_b)) - (H_{Raob}\mathbf{K}(\mathbf{y}_{Sfc} - H_{Sfc}(\mathbf{x}_b)))$$
$$= \boldsymbol{\varepsilon}_{m_{Raob}} + \boldsymbol{\varepsilon}_{r_{Raob}} - H_{Raob}\boldsymbol{\varepsilon}_b - H_{Raob}\mathbf{K}(\boldsymbol{\varepsilon}_{m_{Sfc}} + \boldsymbol{\varepsilon}_{r_{Sfc}} - H_{Sfc}\boldsymbol{\varepsilon}_b)$$

Error metric based on analyse error variance

$$E(\sigma^2_{o_S}) = \left\langle \mathbf{y}_R - H_R(\mathbf{x}_b) - \left\langle \mathbf{y}_R - H_R(\mathbf{x}_b) \right\rangle^T \mathbf{y}_R - H_R(\mathbf{x}_b) - \left\langle \mathbf{y}_R - H_R(\mathbf{x}_b) \right\rangle \right\rangle$$

$$E\ \sigma^2_{o_S} = \sigma^2_{o_R} + \left(\left(\frac{\sigma_b^2}{\hat{\sigma}^2_{o_S} + \sigma_b^2} \right) - 1 \right)^2 \sigma_b^2 + \left(\frac{\sigma_b^2}{\hat{\sigma}^2_{o_S} + \sigma_b^2} \right)^2 \sigma^2_{o_S} - 2\left(\frac{\sigma_b^2}{\hat{\sigma}^2_{o_S} + \sigma_b^2} \right) r_r \sigma_{r_S} \sigma_{r_R}$$

Observation Error Statistics

Evaluation of the observation error variances

- The "optimal" value minimize the error metric defined previously as:

$$E\left(\sigma_{o_S}^2\right) = \sigma_{o_R}^2 + \left(\left(\frac{\sigma_b^2}{\hat{\sigma}_{o_S}^2 + \sigma_b^2}\right) - 1\right)^2 \sigma_b^2 + \left(\frac{\sigma_b^2}{\hat{\sigma}_{o_S}^2 + \sigma_b^2}\right)^2 \sigma_{o_S}^2 - 2\left(\frac{\sigma_b^2}{\hat{\sigma}_{o_S}^2 + \sigma_b^2}\right) r_r \sigma_{r_S} \sigma_{r_R}$$

Experimental assessment of the error variances

- From independent near-surface radiosonde observations, the innovation variance is used, along with σ_o to find a realistic σ_b value:

$$\left\langle \left(\left(\mathbf{y} - H(\mathbf{x}_b)\right) - \left\langle \mathbf{y} - H(\mathbf{x}_b) \right\rangle \right)^2 \right\rangle = \sigma_o^2 + \sigma_b^2$$

- Then, σ_b is used to assess the near-surface observation error components
  - Measurement error (σ_m)
  - Representativeness error (σ_r)

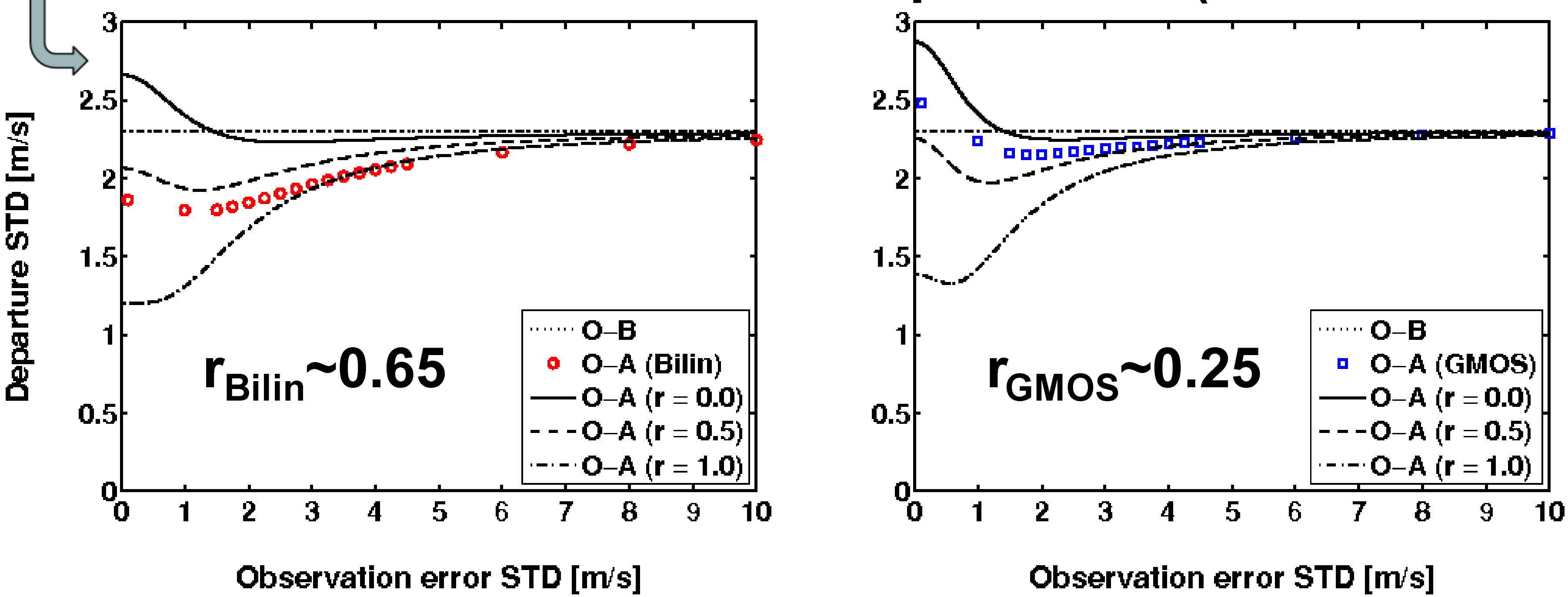
$$\sigma_o^2 = \sigma_m^2 + \sigma_r^2$$

Representativeness Error Correlation

Evaluation of the representativeness error correlation (r_r)

$$E\left[\sigma_{o_S}^2\right] = \sigma_{o_R}^2 + \left(\left(\frac{\sigma_b^2}{\hat{\sigma}_{o_S}^2 + \sigma_b^2}\right) - 1\right)^2 \sigma_b^2 + \left(\frac{\sigma_b^2}{\hat{\sigma}_{o_S}^2 + \sigma_b^2}\right)^2 \sigma_{o_S}^2 - 2\left(\frac{\sigma_b^2}{\hat{\sigma}_{o_S}^2 + \sigma_b^2}\right) r_r \sigma_{r_S} \sigma_{r_R}$$

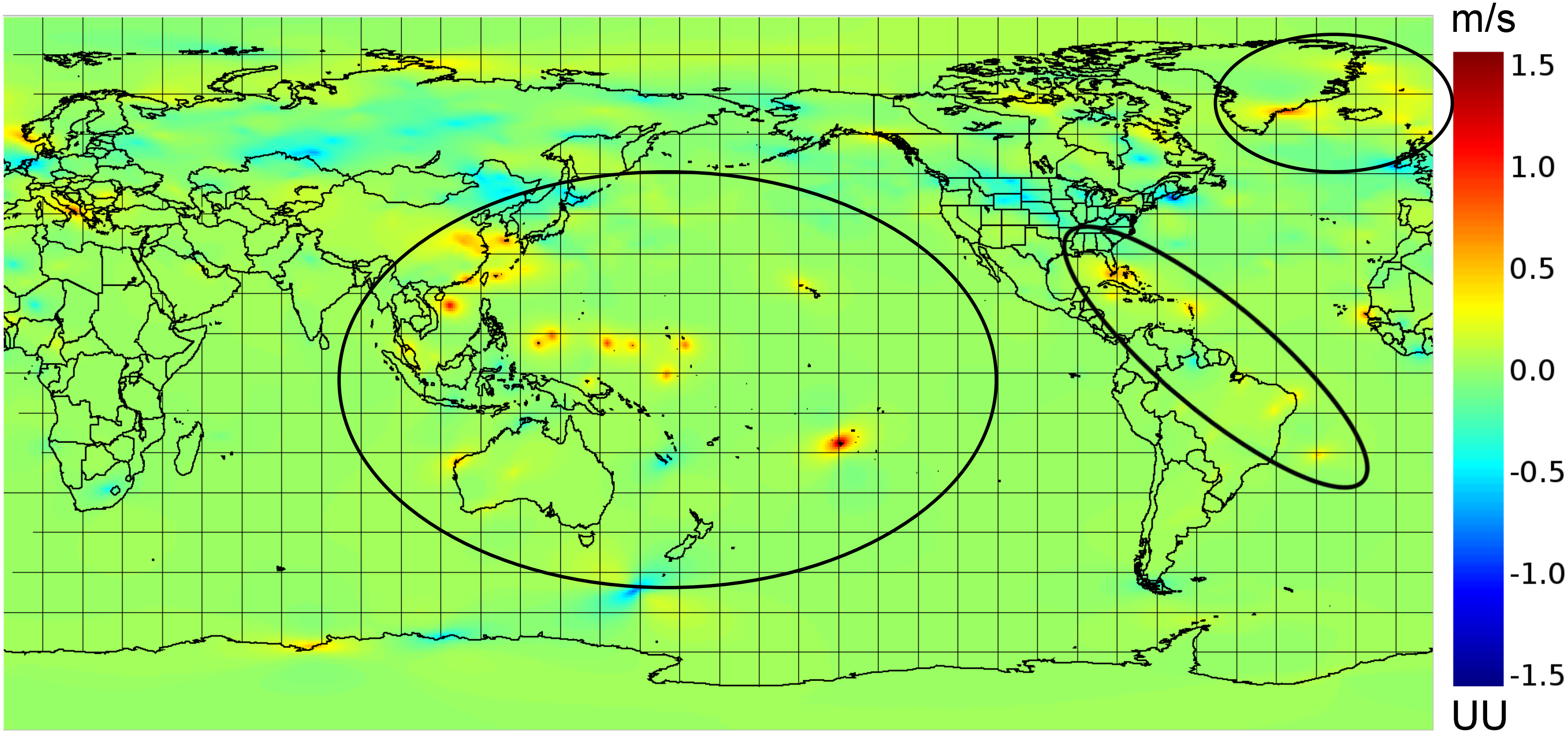
Results from 36 assimilation experiments (1 month each)



UQÀM CRSNG NSERC Hydro Québec Environment Canada

Efficient Bias Correction

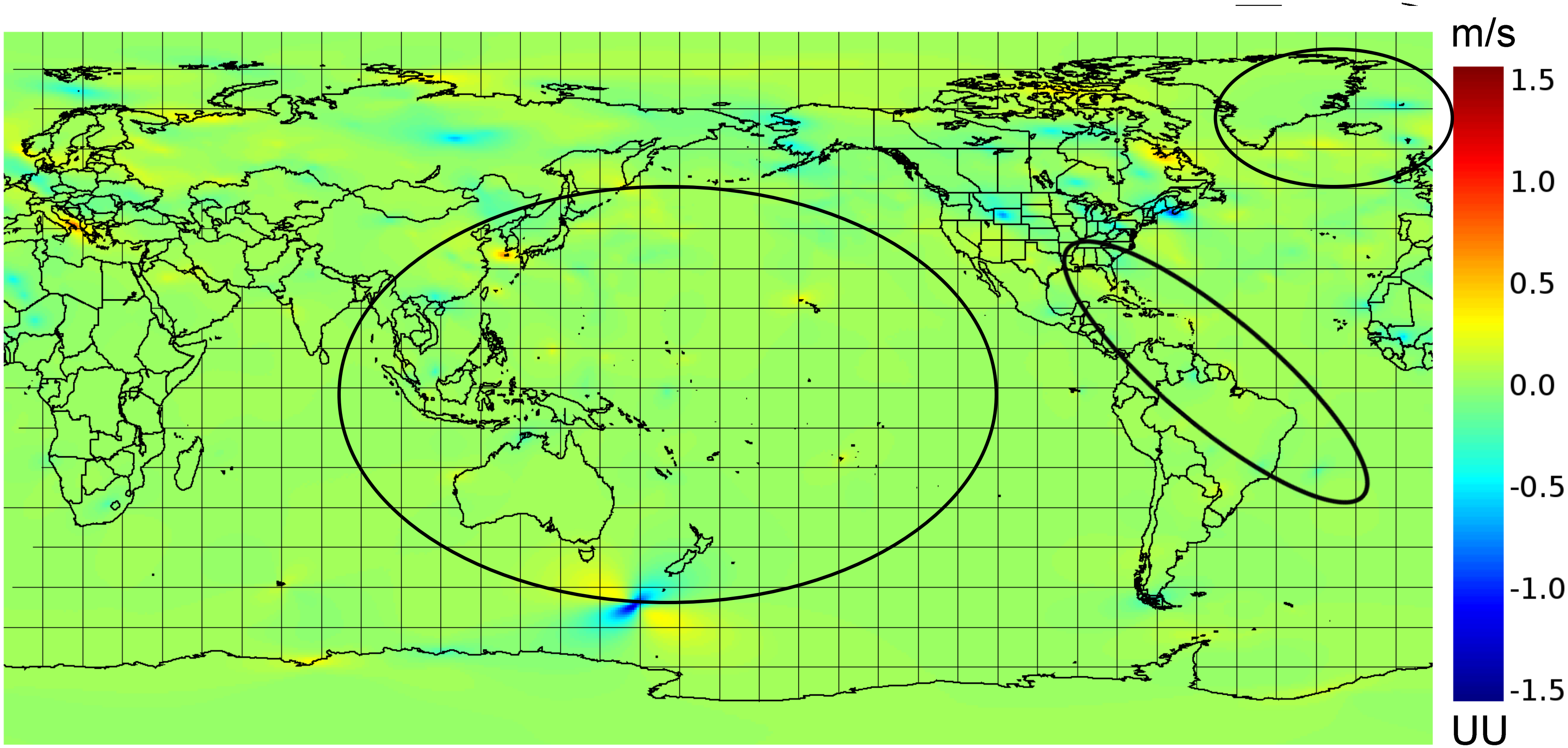
Large biases over complex terrain (e.g. coastal sites)



Mean analysis increments (February 2011) using Bilin operator

Efficient Bias Correction

GMOS reduces the large biases over coastal sites



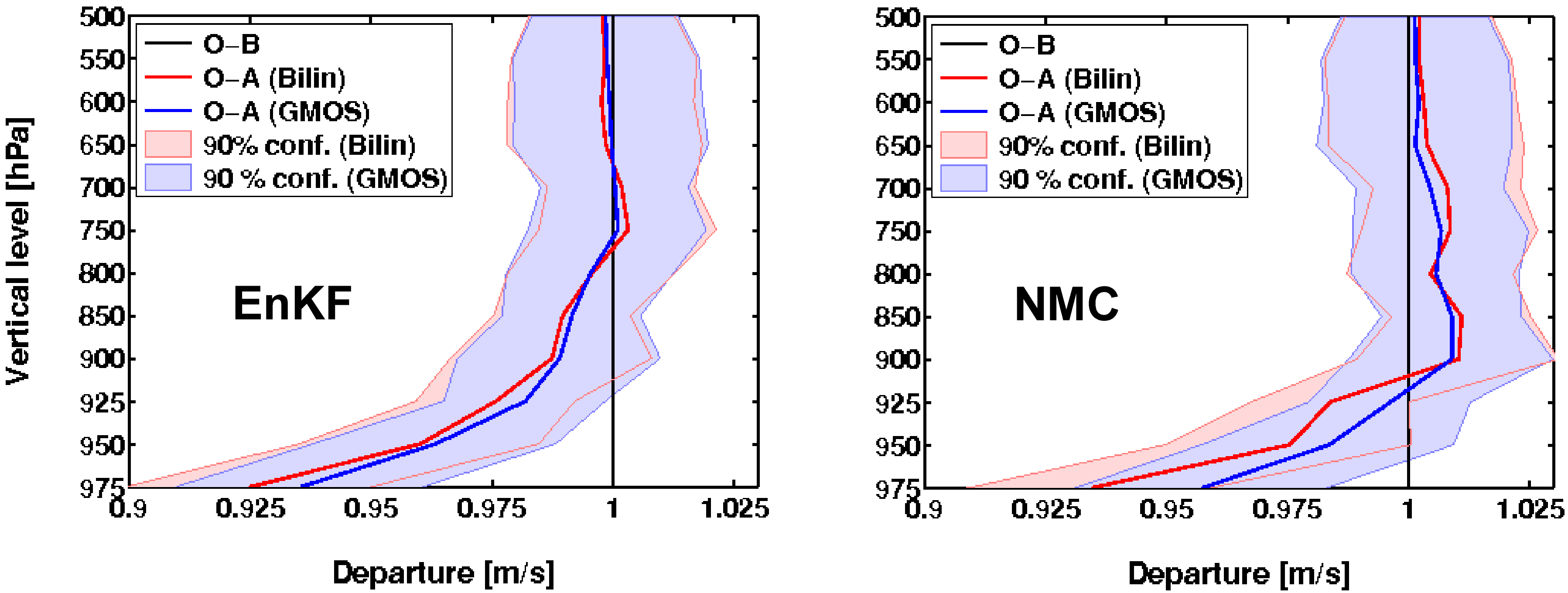
Mean analysis increments (February 2011) using GMOS operator

Impact Assessment

Vertical propagation of the information

- NMC: Propagation up to 650hPa and positive impact up to 900hPa
- EnKF: Propagation and significant positive impact up to p to 750hPa

Significant (5-10%) analysis improvement in the boundary layer



Impact Assessment

Vertical propagation of the information

- NMC: Propagation up to 650hPa and positive impact up to 900hPa
- EnKF: Propagation and significant positive impact up to p to 750hPa

Significant (5-10%) analysis improvement in the boundary layer