UNIVERSITÉ DU QUÉBEC À MONTRÉAL
DEPARTMENT OF EARTH AND ATMOSPHERIC SCIENCES

NEAR-SURFACE WIND DATA ASSIMILATION USING A GEO-STATISTICAL OBSERVATION OPERATOR

THESIS
PRESENTED AS PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DOCTORAL PROGRAM IN
EARTH AND ATMOSPHERIC SCIENCES

BY
JOËL BÉDARD

FINAL SUBMISSION
MONTRÉAL, NOVEMBER 2015

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
DÉPARTEMENT DES SCIENCES DE LA TERRE ET DE L'ATMOSPHÈRE

ASSIMILATION DES VENTS DE SURFACE À L'AIDE D'UN OPERATEUR D'OBSERVATION GÉO-STATISTIQUE

THÈSE
PRÉSENTÉ COMME EXIGENCE PARTIELLE
AU PROGRAMME DE DOCTORAT EN
SCIENCES DE LA TERRE ET DE L'ATMOSPHÈRE

PAR
JOËL BÉDARD

DÉPOT FINAL
MONTRAL, NOVEMBRE 2015

ACKNOWLEDGEMENTS

I would like to thank my directors Prof. Pierre Gauthier (UQAM) and Dr. Stéphane Laroche (Environment Canada) for their advice and support in carrying out this work. More specifically, I am grateful to have had both give me the opportunity to work with them, as well as for their teachings and scientific supervision. Very special thanks to Dr. Laroche for his devotion to the project and to Prof. Gauthier for introducing me to different parties involved in data assimilation internationally. I also wish to thank Mr. Alain Forcione from IREQ and Dr. Wei Yu from Environment Canada who initiated this research project. Many thanks to Mr. Forcione for his support regarding the access to the tall anemometer tower observations from wind farms.

Thanks to Dr. Mark Buehner, Dr. Peter Houtekamer, and other colleagues from Environment Canada for their technical support and advices. Also, thanks to Mr. Michel Valin from the ESCER Center for many fruitful discussions. Moreover, thanks to the members of the present jury for their precious comments and their judicious evaluation.

On a personal note, very special thanks are reserved to Gabrielle, my family and my friends for being with me, for your constant support and your infinite patience.

This project is funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Hydro-Québec research institute (IREQ) and the Meteorological Research Division of Environment Canada. The author also acknowledges the contributions of the ESCER Centre (UQAM) in this research program.

Thank you very much!

TABLE OF CONTENT

LIST OF FIGURES VIII

LIST OF TABLES XII

RÉSUMÉ XIII

ABSTRACT XVI

INTRODUCTION 1

1.1. MOTIVATION 1

1.2. WIND POWER FORECASTING 2

1.3. OBJECTIVES AND METHODOLOGY 4

CHAPTER II 9

EXPERIMENTAL FRAMEWORK 9

2.1. DATA ASSIMILATION 9

2.1.1. Background error covariances 11

2.1.2. Observation error covariances 12

2.2. EXPERIMENTAL SETUP 13

2.2.1. Data assimilation system 13

2.2.2. Cycling experiments with the operational systems 15

2.2.3. Simplified non-cycling experiments 16

2.2.4. Near-surface wind observation operator 17

CHAPTER III

A GEO-STATISTICAL OBSERVATION OPERATOR
FOR THE ASSIMILATION OF NEAR-SURFACE WIND DATA 20

ABSTRACT 22

3.1. INTRODUCTION 23

3.2. DATA ASSIMILATION SYSTEM 26

3.2.1. Background error statistics 26

3.3. MEASUREMENT ERROR STATISTICS 29

3.4. OBSERVATION OPERATOR 31

3.4.1. Observation operator design 31

3.4.2. Training and statistical robustness of statistical operators 33

3.4.3. Temporal variability 35

3.4.4. Representativeness error correction 35

3.4.5. Observation error statistics 40

3.4.6. Tangent linear and adjoint models 41

3.5. OBSERVATION ERROR CORRELATIONS 43

3.6. EVALUATION OF THE IMPACT OF OBSERVATION OPERATORS 46

3.7. RESULTS FROM NON-CYCLING ASSIMILATION EXPERIMENTS 50

3.8. CONCLUSIONS 55

ACKNOWLEDGEMENTS 58

APPENDIX 59

REFERENCES 62

CHAPTER IV

NEAR-SURFACE WIND DATA ASSIMILATION:
TEMPORAL PROPAGATION OF THE ANALYSIS INCREMENT
AND MULTIVARIATE IMPACT ON FORECASTS 66

ABSTRACT 68

4.1. INTRODUCTION 69

4.2. THE DATA ASSIMILATION SYSTEM 72

4.2.1. Geo-statistical observation operator 73

4.2.2. Observation quality control 74

4.2.3. Observation error statistics 76

4.2.4. Evaluation dataset 78

4.3. SIMPLIFIED ASSIMILATION EXPERIMENTS 79

4.3.1. Evaluation against near-surface wind observations 81

4.3.2. Evaluation against upper-air observations 87

4.3.3. Impact of the background error covariance components 90

4.3.4. Constraining mass field using surface pressure observations 91

4.4. EXPERIMENTS WITH THE OPERATIONAL SYSTEM 93

4.4.1. Evaluation against near-surface wind observations 95

4.4.2. Upper-air evaluation 96

4.5. SUMMARY AND CONCLUSIONS 100

ACKNOWLEDGEMENTS 104

REFERENCES 105

CONCLUSION 108

5.1. ORIGINAL CONTRIBUTIONS 108

5.2. SUMMARY OF THE RESULTS 110

5.2.1. Observation operator 110

5.2.2. Background error statistics 110

5.2.3. Multivariate impact 111

5.2.4. Observation impact 112

5.3. LIMITATIONS 112

5.4. OUTLOOK 113

REFERENCES 116

LIST OF FIGURES

Figure	Caption	Page
2.1	Flow chart diagram of the non-cycling assimilation experiments.	16
3.1	Mean vertical profiles of the NMC and EnKF background error STD over land compared to the σ_b values obtained from radiosonde wind component innovations (February 2011) as well as the σ_o values used to assimilate operational radiosonde observations.	28
3.2	80 m wind measurement error STD from ~100 000 samples as a function of observed wind speed.	29
3.3	10 m wind components innovation (observed minus background states) bias and STD as a function of the training dataset length using Bilin, Bilin coupled with a MOS, and GMOS. The number of grid-points used for the GMOS operators is specified in parentheses (e.g., 2×2 when using the 4 nearest grid-points).	34
3.4	Frequency distribution of the wind components innovation STD for the 545 SYNOP stations using either Bilin or GMOS operator.	38
3.5	Spatial distribution of the sites benefitting from a significant observation error STD reduction ($\Delta\sigma_o \geq 1$ ms^{-1}) using the GMOS operator as opposed to Bilin.	39

3.6 Comparison between the wind error metric (lines) and experimental results (symbols) for different observation operators: Bilin (top left), Bilin with bias correction (top right), Bilin with both bias and amplitude corrections from MOS (bottom left), and GMOS (bottom right). Results are plotted for different specified values of the representativeness correlation coefficients ($r_r = 0.0$, 0.5, and 1.0). The analysis error variance is shown as a function of the prescribed observation error variance for each of the experiments. While the symbols represent the experimental results, the lines are the corresponding representation of the analysis error variance from (3.13) evaluated for three values of the representativeness error correlation. 49

3.7 Mean observation impact on wind components analysis (evaluated against radiosonde observations) for Bilin and GMOS operators. The dark (light) shaded area presents the 90 % confidence interval for the Bilin (GMOS) operator. From top to bottom, the panels show data assimilation results using static (top), dynamic (centre), or hybrid (bottom) background error statistics. 52

3.8 Mean zonal wind analysis increments for Bilin (top) and GMOS operators (bottom). 54

4.1 Spatial distribution of the 4942 SYNOP sites considered in this study (black dots). The rectangle refers to the area where the upper-air evaluation is performed. 76

4.2 Frequency distribution of the near-surface wind observation error statistics using the GMOS operator for the 4942 SYNOP stations considered in this study. 77

4.3 Spatial distribution of the 1487 SYNOP sites (left: black dots) located in the selected domain. This area is densely observed and also includes 124 radiosonde stations used for upper-air evaluation purposes (right: black stars). 79

4.4 Wind speed departure STD as a function of forecast lead time for different experiments (CNTRL, Bilin and GMOS). Note that the CRTL experiment is post-processed using both Bilin and GMOS in order to highlight the impact of the operator on post-processing. 82

4.5 Forecast differences ($\|\delta V\|$) between CRTL and the experiments (Bilin: Left panel and GMOS: Right panel) using different background error covariances (Hybrid, NMC and EnKF). Results are presented for the February 1st run launched at 0000 UTC. The upper panels show the results over the whole 48 h forecast period while only the first 6 h are depicted in the lower panels. 83

4.6 Contribution of the main terms of Eq. (4.3) on evolution of the forecast difference between CNTRL and the experiments (left: $GMOS_{NMC}$; right: $GMOS_{EnKF}$). Results for advection, Coriolis effect, horizontal diffusion are omitted as their respective influence is small compared to the pressure gradient, the vertical diffusion and orographic blocking. 86

4.7 24 h forecast departure bias from hybrid runs as evaluated against radiosonde observations over Europe for wind speed (left) and geopotential height (right). 88

4.8 Forecast departure STD from hybrid runs as evaluated against radiosonde observations over Europe for wind speed (left) and geopotential height (right). In each plot, results from 12 h, 24 h, 36 h and 48 h forecasts are shown from left to right respectively. 89

4.9 24 h forecast departure bias from GMOS runs as evaluated against radiosonde observation profiles over Europe for wind speed (left) and geopotential height (right). 90

4.10 24 h forecast departure bias from GMOS runs as evaluated against radiosonde observations over Europe for wind speed (left) and geopotential height (right). The CNTRL experiment is represented by circles, while assimilation experiments using only surface pressure observations (P0), only near-surface wind observations ($GMOS_{EnKF}$), or both surface pressure and near-surface wind observations (P0 + $GMOS_{EnKF}$) are shown by solid, dot and dash lines respectively. 92

4.11 Wind speed departure STD as a function of forecast lead time for different experiments ($CNTRL_{OSE}$, $Bilin_{OSE}$ and $GMOS_{OSE}$). Note that the GMOS operator is also used for post-processing in all experiments. 95

4.12 Mean 12 h forecast departure STD (against own analyses) differences (BilinoSE minus GMOSoSE) over Europe. Results for 10 m wind speed ([ms^{-1}]) are averaged over the February 2011 period. Positive (negative) values are represented by dark (light) colors. A positive value (dark gray) indicates that the GMOSoSE experiment is better than BilinoSE, while light gray color indicate neutral results. 98

4.13 Hovmöller diagram presenting the differences between BilinoSE and GMOSoSE 12 h forecast departure STD (against own analyses). Results for 10 m wind speed ([ms^{-1}]) are presented through February 2011 for different longitude bands over Europe. Positive (negative) values are represented by dark (light) colors. A positive value (dark gray) indicates that the GMOSoSE experiment is better than BilinoSE, while light gray color indicate neutral results. 100

LIST OF TABLES

Table	Caption	Page
3.1	Mean on-site wind component observation and forecast standard deviation using different observation operators: Bilin, MOS, and GMOS.	36
3.2	Configuration of the 60 analysis experiments performed. For each observation operators implemented, the different observation error value and background error matrix prescribed to the data assimilation system as well as the number of experiments are listed.	47
4.1	Configuration of the 9 simplified data assimilation performed. Each experiment is listed along with its own combination of near-surface wind observation operator, background and observation error statistics prescribed to the data assimilation system as well as assimilated observations. It is also specified if the experiments are cycled or not.	80
4.2	Configuration of the 3 cycling OSE performed. Each experiment is listed along with its own combination of near-surface wind observation operator, background and observation error statistics prescribed to the data assimilation system as well as assimilated observations. It is also specified if the experiments are cycled or not.	94

RÉSUMÉ

L'intégration de la production éolienne dans les réseaux électriques comporte d'importants défis reliés à la variabilité du vent. Afin de garantir la fiabilité du réseau, il faut se doter de réserves de puissance pouvant pallier les effets des variations de cette ressource. Pour des échéances allant jusqu'à 24 h, une connaissance précise de la production à venir (basée sur la prévision du vent et des paramètres météorologiques connexes) permet de limiter ces réserves au strict minimum, maximisant ainsi la valeur de la ressource éolienne. Les prévisions météorologiques de courte échéance sont imprécises et une partie de l'erreur peut être attribuée aux conditions initiales fournies par les analyses. Malgré la disponibilité d'observations de vent de surface au-dessus des continents, peu d'entre elles sont utilisées dans la production d'analyses en raison des interactions entre la topographie locale et l'écoulement à fine échelle non-résolu par les modèles météorologiques. La disparité d'échelles entre les observations et le modèle cause des erreurs de représentativité qui peuvent devenir significatives, particulièrement en terrain complexe.

L'objectif principal de ce projet est d'améliorer les prévisions troposphériques de courte échéance en assimilant les observations de vents continentaux dans le système d'assimilation de données ensembliste-variationnel d'Environnement Canada. Un nouvel opérateur d'observation basé sur les géo-statistiques (GMOS) a été développé dans le but de corriger les erreurs systématiques et de représentativité. Cet opérateur combine une méthode de correction statistique à une méthodologie utilisant plusieurs points de grille afin de prendre avantage des corrélations multi-échelles. GMOS est comparée à un opérateur conventionnel basé sur une interpolation bilinéaire (Bilin) dans le contexte d'analyses et de prévisions opérationnelles.

Des expériences d'assimilation ont été effectuées en assimilant les observations de vents de surface uniquement pour les stations où un radiosondage est également disponible. Ces expériences avaient pour but de comprendre l'impact de ces observations sur les analyses météorologiques et ainsi repérer les composantes du système d'assimilation nécessitant des améliorations afin de profiter du plein potentiel de ces observations. Les corrections apportées ont été systématiquement évaluées en comparant les analyses résultantes aux observations des radiosondages. Ces observations indépendantes n'étaient pas utilisées dans l'assimilation. Cette méthode de validation a permis d'estimer la variance d'erreur d'observation qui maximise l'impact des observations de vent de surface tout en générant une analyse

cohérente avec les observations de radiosondages. Les résultats des expériences d'assimilation (sur une période d'un mois) montrent qu'à cause de la structure verticale des covariances d'erreur de prévision, l'assimilation des vents de surface impacte principalement les basses couches de l'atmosphère. En sélectionnant statistiquement les points de grille du modèle qui sont les plus corrélés aux observations, GMOS permet de mieux représenter les phénomènes météorologiques *in situ*, ce qui élimine les biais, réduit les erreurs de représentativité et les corrélations d'erreurs d'observation.

Des résultats sont aussi présentés pour des expériences d'assimilation non cyclées utilisant cette fois les vents de surface de toutes les stations synoptiques globales (SYNOP). La vérification des analyses et des prévisions 48 h à l'aide d'observations de surface démontre que l'utilisation de GMOS permet d'améliorer les prévisions de vent locales. L'impact diminue rapidement et n'est valide que pour des échéances allant jusqu'à 6 h en raison de l'utilisation des statistiques d'erreur de prévision homogènes. Ces statistiques d'erreur génèrent des incréments d'analyse altérant principalement les vents et les paramétrages reliés au blocage orographique et à la couche limite atmosphérique atténuent l'incrément d'analyse. Toutefois, les statistiques d'erreurs dynamiques provenant de méthodes ensemblistes permettent une meilleure propagation de l'information à la verticale puisqu'elles varient en fonction de la stabilité atmosphérique. Elles conduisent aussi à des incréments mieux balancés à cause des corrélations plus accentuées entre les champs de vent et de masse. L'incrément peut alors se propager dans le temps: le gradient de pression généré supporte l'incrément de vent et contrebalance les forces diffusives des paramétrages du modèle. Les statistiques d'erreur de prévisions sont déterminantes pour la propagation de l'information dans les systèmes d'assimilation et de prévision.

En assimilant conjointement les observations de vent et de pression de surface, GMOS permet de réduire le biais des prévisions de vent jusqu'à 48 h. Les résultats de ces expériences suggèrent que la complémentarité entre les observations de vent et de masse, des statistiques d'erreur de prévisions dynamiques et l'opérateur GMOS sont des composantes essentielles à l'utilisation des vents de surface pour l'amélioration des conditions initiales des modèles de prévision numérique.

Enfin, les résultats d'expériences cyclées avec le système d'Environnement Canada utilisant toutes les observations assimilées opérationnellement, auxquelles on ajoute les observations de vents de surface, indiquent que les incréments générés par Bilin sont atténués par le système alors que ceux générés par GMOS persistent plus longtemps puisqu'ils sont mieux balancés. Ainsi, les prévisions issues des expériences employant GMOS (Bilin) sont plus (moins) cohérentes avec les analyses que les prévisions issues de l'expérience de contrôle n'assimilant pas les vents de surface au-dessus des continents.

Mots-clés: statistiques d'erreur d'observation et de prévision, paramétrages de la couche limite atmosphérique, énergie éolienne, assimilation de données de surface, analyse météorologique, prévision numérique du temps, validation contre analyses et observations, erreur de représentativité, opérateur d'observation, vents de surface.

ABSTRACT

The ability to predict hourly near-surface winds up to 24 h in the future plays a key role in the integration of wind power in the energy production of electrical facilities. Improvement in the accuracy of wind power predictions is essential in the overall energy management of a network comprising different energy sources and increases the value of wind energy. From numerical weather prediction assessments, large forecast errors appear to grow from an inadequate characterisation of the atmosphere in the initial conditions, defined by large-scale analyses, used by a numerical weather prediction (NWP) model. Although many observations describing the wind field in the lower troposphere are available, very few are assimilated over land because the sub-grid scale topographic interactions with the flow cannot entirely be reproduced by NWP models. Thus, representativeness errors can be significant, especially if the observing station is located on complex terrain or a coastal site.

This study aims at improving short-term tropospheric forecasts by assimilating near-surface wind observations over land in Environment Canada's ensemble-variational data assimilation system. A new geo-statistical observation operator (GMOS) has been developed to correct for systematic and representativeness errors. This method combines a statistical error correction with a multiple grid-point approach to correct representativeness error by taking advantage of the correlation between resolved and unresolved scales. It is tested and compared with the conventional bilinear interpolation scheme (Bilin) in the context of an operational NWP forecasting and assimilation suite.

Assimilation experiments are first performed in a simplified context assimilating only near-surface (10 m) wind observations over land to understand the impact of near-surface wind observations on the analysis and to point out aspects that need to be improved to make a better use of these observations. Due to the background-error covariances, the assimilation of near-surface wind observations impacts the lower part of the atmosphere. The resulting corrections have been evaluated by verifying the analyses against non-assimilated collocated radiosonde data. This assessment also made it possible to estimate the observation error variance to strike a balance between having an important impact at the surface and maintaining a good vertical fit to upper-air observations. Results from one month assimilation experiments show that, by its statistical correction and by using the forecast values at grid-points that are the most representative, GMOS better represents the meteorological phenomena onsite. It

eliminates biases and significantly reduces representativeness errors as well as observation error correlations, mainly over complex terrain. Overall, the analysis fit to non-assimilated collocated radiosonde observations is improved when assimilating wind observations from surface stations using both GMOS and Bilin.

Results from non-cycling forecast experiments are also presented. Only near-surface wind observations were assimilated to evaluate the spatio-temporal propagation of the information along with the multivariate upper-air impact in a controlled environment. The verification of the resulting analyses and subsequent 48 h forecasts against near-surface wind observations and independent radiosonde profiles show that very short-term wind predictions are significantly improved when using GMOS. However, the local impact decreases over time and is significant only for 6 h or shorter. When using static error covariances the mass field is not significantly altered and the boundary layer parameterization and orographic blocking schemes damp the poorly balanced increment. When using flow-dependent error statistics, the analysis increment modifies both wind and mass fields in a consistent way through multivariate covariances which result in increased temporal propagation of the information from the near-surface wind observations. Results also show that flow-dependent background error covariances from ensembles provide better vertical propagation of the information than static error statistics. The use of proper background error statistics is crucial to produce sustainable impacts on the atmosphere.

The assimilation of near-surface winds (with GMOS) in conjunction with observations of surface pressure provides significant upper-air improvements (as evaluated against radiosonde observations) in terms of wind speed bias up to 48 h lead time. This shows that the use of the GMOS operator along with flow dependent background error statistics allows taking advantage of the combined effect from both mass and wind observations.

Finally, results from cycling observing system experiments (assimilating near-surface wind data and all the other types of observations used operationally) indicates that Bilin does not provide a good model state comparison with the observations and its analysis increments are damped during subsequent forecasts. The increment produced when using GMOS is better balanced and the information persists in the system. Forecasts and analyses from GMOS (Bilin) are also slightly more (less) coherent than those from the control experiment (in which no wind observations are assimilated over land) as the information from the observations is (is not) propagated in time.

Keywords: observation and background error statistics, atmospheric boundary layer parameterization, wind energy, surface data assimilation, weather data analysis, numerical weather prediction, evaluation against analyses and observations, representativeness errors, observation operator, near-surface winds.

INTRODUCTION

1.1. Motivation

In the context of a changing climate and due to the many consequences of electricity generation, industries and governments are promoting and developing power generation from renewable sources such as wind power. This type of power generation has reduced impacts on the environment compared to more conventional power plants like nuclear, oil, natural gas and coal. In fact, in some European countries, up to 33 % of the power generation comes from wind power (Global Wind Energy Council, 2013). This clean energy is also increasingly being adopted and used in North America.

One of the major concerns regarding the integration of wind energy in electrical utilities is the variability of wind, and therefore, power. With the sustained growth of wind energy installed capacity for electricity generation, electrical system operators face the increasing challenge of balancing electrical grids, notably to maintain reliability and minimizing management and operation costs of other energy sources. To ensure the reliability of the electrical utility, system operators keep extra generating capacity (spinning reserves) that can be available within a short time interval. However, these spinning reserves, needed to cope with the variability of wind and power forecast uncertainty, translate into direct energy and financial losses.

For facilities, like Québec electrical utility, where wind will power up to 10 % of the peak electrical demand by the end of 2015 (Ministère des Ressources Naturelles et de

la Faune, 2006), hourly and daily high resolution wind power forecasts are necessary as systems operators need 24 to 48 h forecasts to plan the daily power generation. Hourly forecasts are then used to match the electrical generation with the predicted demand, either by using spinning reserves or by importing/exporting energy at the very last moment from/to neighbouring system operators. As these balancing solutions are expensive, reliable short-term wind power generation forecast is necessary for the technical and financial sustainability of large-scale wind energy integration, in both regulated and open markets.

1.2. Wind power forecasting

The reliability of the electrical utility is the main concern for system operators and to improve reliability requires the improvement of short-term forecasts (0 h to 48 h lead time). Landberg *et al.* (2003) and Giebel *et al.* (2011) provide complete reviews of wind power prediction models. The kinetic energy flux in the air is proportional to air density and to the cube of wind speed (through a perpendicular surface). Wind speed is thus the largest contributor to relative wind power output. Still, the local topographic (surface roughness, obstacles on the ground, etc.) impact on the wind characteristics (including wind shear) strongly depends on the wind direction (Bédard *et al.*, 2013). Air temperature, humidity and atmospheric pressure predictions influence the wind power forecasts through air density. Specific humidity also has a significant influence on icing events, which greatly influence the wind power production. As winds, temperature, humidity and atmospheric pressure contribute to wind power, the vast majority of short-term wind power forecast models are based on Numerical Weather Prediction (NWP) systems (Landberg *et al.*, 2003; Giebel *et al.*, 2011). The global and regional NWP models generally have grid-point spacings

ranging from 5 km to 25 km. As wind is closely related to fine-scale topographic effects not resolved by operational meso-scale models, representativeness errors can be significant. Thus, in many cases, forecasts uncertainties are related to the terrain complexity (Kariniotakis *et al.*, 2004).

To estimate the wind resources at the exact wind farm location, system operators generally run downscaling models. This operation can be done using a Limited Area Model (LAM) which resolves topography and surface roughness at a higher resolution (up to 1 km grid spacing) while employing the global or regional NWP as initial and lateral boundary conditions. Another solution is to use wind observations from the wind power plant along with historical forecasts to derive statistical regressions relating the forecasts to onsite measurements (so-called Model Output Statistics methods: MOS). Although the use of a LAM model generally provides improvements over complex terrain, the benefits almost disappear if MOS are used (Müller, 2011). In operational weather prediction, MOS are also routinely used to correct local biases and representativeness errors in NWP model forecasts for locations where observations are available.

In most wind farm applications, the end users (e.g., wind farm managers, transmission system operators, energy service suppliers and traders) are not the ones who run the NWP models; rather, they only use the forecasts. The end users experience directly the consequences of the forecast errors, such as the variation in the price of energy on the market, supply contracts, operating costs and security concerns. Therefore, end users benefit from an estimation of the error related to the power predictions. Such estimation can be provided by means of statistical approaches or using ensemble forecasts (Giebel *et al.*, 2011). Ensemble forecasts are also used to improve the operational robustness (reliability and accuracy) of the forecast systems (Nielsen *et al.*, 2007a) because a combination of an ensemble of

predictions generally outperforms the individual forecasts (Lange *et al.*, 2006; Nielsen *et al.*, 2007b). Ensembles of power forecast models can be generated by using NWP models from different national services, different physcial parameterizations used in the same model, and perturbed initial conditions (Nielsen *et al.*, 2004; Nielsen *et al.*, 2007a). End users generally combine the ensemble members by statistically weighting the NWP based on a combination of the different forecast properties (variance, kurtosis, skewness, etc.) and the weather conditions. Correlations developed using the turbine power curves are then used to relate the meteorological variables to electrical power. Wind power forecast models also generally take into account the wind farm layout to integrate the wake effect of a turbine on the aerodynamics of the whole wind farm into the final wind power forecast.

1.3. Objectives and methodology

To improve wind power predictions, a lot of effort has been invested in the last decades to improve the way NWP models are used (Giebel *et al.*, 2011). Forecasts from NWP models are generally skillfull in the short-term (up to 48 h) and persistence is an excellent local predictor for very short-term forecasts up to 3 – 6 h (Landberg and Watson, 1994; Liu, 2009). Thus, Nielsen *et al.* (2000) proposed to merge forecasts from both the persistence and the NWP model based on the autocorrelation of the wind speed from measured time series. Power prediction models, such as the Wind Power Prediction Tool (WPPT), now combine the forecasts from NWP models with those from nowcasting models which are generally based on time series analysis tools (e.g., neural networks). This type of approach locally

outperforms both individual models for all time horizons and it is now commonly used.

Although near-surface wind observations are used at the local scale, very few are used over land to generate the initial conditions provided to NWP models, mainly due to sub-grid scale topographic interactions with the flow. Indeed, near-surface observations tend to sample fine-scale structures which are not explicitly resolved by the NWP model (so-called sub-grid scale features), causing discrepancy between the characteristics of the measured and forecasted variables. While aircraft and radiosonde observations have the most positive impact on short-term forecasts, surface observations (temperature, humidity and pressure) strongly complement those upper-air observing systems (Benjamin *et al.*, 2010). Interest is thus growing in the use of near-surface wind observations over land in regional and global data assimilation systems to take advantage of their large scale correlations with the atmospheric flow.

To improve short term wind forecasts for different applications (e.g. wind energy, airport operations, road and construction site security, recreational activities, etc.), recent studies tried to assimilate near-surface wind observations over land. A research group at the National Center for Atmospheric Research (NCAR) performed simplified experiments to assess the near-surface wind data assimilation problem in a simplified context (Hacker and Snyder, 2005; Hacker and Rostkier-Edelstein, 2007; Hacker *et al.*, 2007; Rostkier-Edelstein and Hacker, 2010). Using a single column model, Rostkier-Edelstein and Hacker (2010) showed that improvements in the assimilation of surface observations can improve nowcasting capabilities. Furthermore, results from Hacker and Snyder (2005) indicate that near-surface observations (including winds) are effective at constraining the state of the Atmospheric Boundary Layer (ABL). As surface observations have smaller

correlation with the flow aloft compared to integrated variables such as surface pressure, their impact on analyses varies depending on the ABL coupling with the flow aloft (e.g., atmospheric stability). This may limit the use of stationary background error covariances. When using an Ensemble Kalman Filter (EnKF) to sample the flow dependent (in time and space) and multivariate background error covariances, the assimilation of near-surface observations impacted temperature, humidity and winds profiles (Hacker and Rostkier-Edelstein, 2007).

Other recent studies have focused on the impact of near-surface wind observations on forecasts by assimilating them in more realistic contexts. While Dong *et al.* (2010) showed that synthetic near-surface wind observations can provide significant forecast improvements for areas with poor low-level data coverage, many research groups working with real observations discarded them as representativeness errors and biases significantly affected the observation impact (Benjamin *et al.*, 2007; Ingleby, 2014). By performing observation targeting experiments to assess forecast sensitivity to wind observations, Zack *et al.* (2010) showed that near-surface wind forecasts (up to 3 h lead time) are sensitive to local low level wind initial conditions. Zack *et al.* (2011) then showed that near-surface local wind nowcasting capabilities can be improved by assimilating synthetic observations from tall anemometer towers (80 m). By assessing the observation impact on wind forecasts over Texas and Oklahoma using different assimilation systems, Ancell *et al.* (2015) showed that assimilating real near-surface wind observations improved wind nowcasting capabilities (up to 6 h). Improvements appeared to be more significant when using an assimilation system based on an EnKF rather than a 3D variational system (3DVAR), presumably because of the EnKF flow dependent error statistics.

To understand why near-surface wind observations have little or no impact on forecasts beyond 6 h, the present work focuses on the observation impact on analyses

and forecasts and the vertical and temporal propagation of the information in the forecast system. This study also aims at improving short-term tropospheric forecasts by assimilating near-surface wind observations over land. By assessing the influence of different components of the assimilation and prediction systems (namely the background error statistics, the observation operator and the boundary layer parameterization), it is intended to point out components of the system that need to be improved to make better use of these observations.

Unlike satellite data, which require complex observation operators, near-surface wind observations are more directly related to model state variables. The Canadian Meteorological Center (CMC) currently employs a near-surface wind observation operator that is a simple geometric bilinear horizontal interpolation of model state variables to the observation location. However, the observation operator used to compare forecasts locally with near-surface winds (for validation and operational forecast purposes) is generally based on Model Output Statistics (MOS). It is thus intended to use such a method in the assimilation system as well to increase the consistency between observations, analyses and forecasts in the ensemble-variational data assimilation system (4DEnVar) developed at Environment Canada (Buehner *et al.*, 2013; 2015).

While Chapter 2 describes the experimental setup, Chapter 3 presents a new observation operator based on a geo-statistical observation operator developed by Bédard *et al.* (2013). This operator is implemented in the 4DEnVar to reduce the representativeness errors for the assimilation of near-surface winds. By using only surface stations collocated with upper-air stations (545 stations), the statistical robustness of the operator as well as the systematic and representativeness error corrections are examined in a controlled environment by comparing the analysis against independent non-assimilated observations. A new approach based on

independent observations is also proposed to estimate observation error correlations. Data assimilation experiments over a one-month period are carried out where only wind data from surface stations are assimilated to carefully evaluate the observation impact on the analysis. The collocated radiosonde profiles are used to assess the vertical propagation of information when different background error statistics are used.

Chapter 4 presents results from both simplified Observing System Experiments (OSEs) carried out in non-cycling mode as well as fully cycled OSEs performed with the CMC operational system. All SYNOP stations distributed globally are considered (4942 stations) to take advantage of the full observation impact on the NWP system. At first, only wind data from surface stations are assimilated to allow for the evaluation of the spatio-temporal propagation of the information and multivariate impact in a controlled environment. The systematic model initial tendencies are used to assess influence of the boundary layer parameterizations on the temporal evolution of the analysis increment. Quality control issues related to near-surface wind observations are also examined. Realistic experiments are then performed to show the added value of near-surface wind observations in the presence of the other observations used operationally. Results from cycling OSEs using the Environment Canada operational 4DEnVar are presented. The analysis experiments are evaluated against surface observations, radiosonde profiles and analyses to assess the quality of the resulting analyses and subsequent 48 h forecasts.

CHAPTER II

EXPERIMENTAL FRAMEWORK

2.1. Data assimilation

Data assimilation is the process by which observations are combined with prior information on the atmospheric state (the so-called background state, which is generally based on a short-term NWP forecast) in order to produce the best estimate of the true state of the atmosphere at a given time (so-called analysis state). The analysis is generally used as initial conditions for the NWP model or as a reference for the evaluation of forecasts and climate simulations. Based on statistical estimation principles, the assimilation process relies on the observation and background error covariances (respectively $\mathbf{R}$ and $\mathbf{B}$) to optimally blend the observation vector ($\mathbf{y}$: containing all assimilated observations) and the operational background state ($\mathbf{x}_B$) to produce the analysis ($\mathbf{x}_A$) such as

$$\mathbf{x}_A = \mathbf{x}_B + \mathbf{K}\left(\mathbf{y} - H\left(\mathbf{x}_B\right)\right) \tag{2.1}$$

where H represents the full non-linear observation operator, $\mathbf{H}$ stands fot its Jacobian, a linearization of H around the basic state and $\mathbf{K}$ is the Kalman gain matrix defined as

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1} \tag{2.2}$$

Since the atmosphere relies on highly non-linear and multi-scale systems, a certain number of judicious assumptions must be made in order to operationally apply data assimilation methods, whether based on the variational (VAR) approaches (3D or 4D) or the Ensemble Kalman Filtering (EnKF) methods. A simple description of data assimilation for model initialization in atmospheric sciences can be found in Reichle (2008), while Bouttier and Courtier (1999) give a detailed course on data assimilation concepts and methods. A detailed implementation of the 4DVAR scheme previously used at the Meteorological Service of Canada can be found in Gauthier *et al.* (2007), while the operational EnKF (based on the algorithm found in Evensen, 1994) is described in Houtekamer *et al.* (2014). Also, the European Centre for Medium-Range Weather Forecasts (ECMWF) and Météo-France's implementations of the 4DVAR are found in Rabier *et al.* (2000), Mahfouf and Rabier (2000), Klinker *et al.* (2000) and Gauthier and Thépaut (2001).

While the EnKF generally processes small observation batches sequentially to compute the analysis following Eq. (2.1), 3DVar and 4DVar systems minimize the analysis error by means of a cost function J using all observations at once to compute the analysis by minimizing the functional

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_B)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_B) + \frac{1}{2}(H(\mathbf{x})-\mathbf{y})^T\mathbf{R}^{-1}(H(\mathbf{x})-\mathbf{y}) \tag{2.3}$$

The incremental method is commonly employed to minimize this cost function (Courtier *et al.*, 1994; Gauthier *et al.*, 2007). The innovation vector $(\mathbf{y}' = \mathbf{y} - H(\mathbf{x}_B))$ is first computed by comparing the observation and the background state using the observation operator (H). Then, the analysis increment ($\delta\mathbf{x} = \mathbf{x}_A - \mathbf{x}_B$) is obtained by minimizing

$$J_L(\delta\mathbf{x}) = \frac{1}{2}(\delta\mathbf{x})^T\mathbf{B}^{-1}(\delta\mathbf{x}) + \frac{1}{2}(\mathbf{H}(\delta\mathbf{x}) - \mathbf{y}')^T\mathbf{R}^{-1}(\mathbf{H}(\delta\mathbf{x}) - \mathbf{y}') \cdot \quad (2.4)$$

The analysis increment is then added to the background state to generate the analysis.

In all cases, when using the analysis as initial conditions for the NWP model, the forecast is also used as the background state to compute the subsequent analysis in time. Cycling the assimilation system allows the information from observations to propagate within the NWP system and contributes to subsequent analyses.

2.1.1. Background error covariances

While the 3DVAR scheme uses stationary background error covariances based on globally homogeneous correlations and simple balance relationships, 4DVAR and EnKF schemes employ three dimensional background error statistics that evolve in time (implicitly in 4DVAR and explicitly in EnKF). The 4DVar implicitly evolves the 3D error statistics using the Tangent Linear Model (TLM) approximation of the NWP model equations and its adjoint. Essentially, the adjoint model is the transposition of the TLM operator (Le Dimet and Talagrand, 1986). On the other hand, the EnKF is based on a Monte-Carlo approach: it uses the analysis error covariance to generate a set of perturbed initial conditions which are used as NWP model inputs. The different integrations of the model will tend to diverge due to error growth associated with atmospheric instabilities (non-linear error growth) and the spread of the ensemble is used to estimate the forecast error covariance (forecast uncertainty).

2.1.2. Observation error covariances

Comparing background state to observations implies different error terms as

$$\mathbf{y}-\mathbf{H}(\mathbf{x}_B)=(\mathbf{y}-\mathbf{y}_T)-\mathbf{H}(\mathbf{x}_B-\mathbf{x}_T)-(\mathbf{H}(\mathbf{x}_T)-\mathbf{y}_T)=\boldsymbol{\varepsilon}_M-\mathbf{H}(\boldsymbol{\varepsilon}_B)-\boldsymbol{\varepsilon}_R \quad , \quad (2.5)$$

where $\mathbf{y}_T$ and $\mathbf{x}_T$ are the atmospheric true state in the observation and model space respectively. Following Eq. (2.5), it is possible to distinguish between the measurement error ($\boldsymbol{\varepsilon}_M=\mathbf{y}-\mathbf{y}_T$), the background error ($\boldsymbol{\varepsilon}_B=\mathbf{x}_B-\mathbf{x}_T$) and the representativeness error ($\boldsymbol{\varepsilon}_R=\mathbf{H}(\mathbf{x}_T)-\mathbf{y}_T$).

While measurement and background errors are conceptually straightforward, the representativeness error is complex as it arises from the fact that the model state represent the average atmospheric state over a model grid-point (because of its limited horizontal resolution), while the observations represent the local atmospheric state. Observation operators ($\mathbf{H}$) cannot precisely reconstruct the local atmospheric states from the model average state and thus, representativeness is generated when comparing model state variables with observations. Again, the NWP model represents the average atmospheric state over a model grid-cell, not the local state. For this reason, the representativeness error is included in the observation error variance ($\boldsymbol{\sigma}_O^2=\boldsymbol{\sigma}_M^2+\boldsymbol{\sigma}_R^2$, assuming ε_M and ε_R are not correlated), meaning that the local observation is not representative of the average state the NWP model intends to compute.

Observation errors are also assumed to be uncorrelated in operational systems. This assumption appears reasonable for conventional observations as they are generally

collected using different instruments located in different geographical areas. For cases where the observation errors are correlated (e.g. satellite observations), increasing the observation error variance and/or spatially thinning the observations reduce the correlation impacts on the analysis (Liu and Rabier, 2002), and thus, this assumption allows the use of a diagonal observation error covariance matrix, which greatly simplifies the assimilation process.

2.2. Experimental setup

2.2.1. Data assimilation system

The 4D Ensemble-Variational data assimilation system (4DEnVar) of Environment Canada relies on 4D ensemble covariances to explicitly estimate the spatio-temporal background error statistics over a 6 h time window (Buehner *et al.*, 2013; 2015). The 4DEnVar uses the incremental approach to produce an analysis at the spatial resolution of the forecast model from an analysis increment computed at a lower resolution. The analysis increment is computed on a 800×400 grid, which corresponds to a horizontal grid spacing in latitude and longitude of 0.45°×0.45°. The Global Deterministic Prediction System (GDPS: Zadra *et al.*, 2013) uses the Global Environmental Multiscale (GEM) model (Côté *et al.*, 1998a,b; Charron *et al.*, 2012) configured on a global grid with a horizontal grid spacing in latitude and longitude of 0.35°×0.225° and 80 staggered vertical levels with the top at 0.1 hPa. The time step used to produce the forecasts is of 720 s, while the one used to produce the background state is of 450 s to match the temporal discretization of the ensemble covariances from the EnKF (Houtekamer *et al.*, 2014).

The background error statistics used in the 4DEnVar comprise a stationary homogeneous component and a flow-dependent component. While the static component is estimated using the NMC method (Parrish and Derber, 1992, Gauthier *et al.*, 1998; for details see Charron *et al.*, 2012), the flow dependent error covariances are estimated using an ensemble of 256 background states from the operational EnKF at Environment Canada (Houtekamer, 2014), which uses a model configuration adapted for ensemble forecasts (different than the EnVar model configuration, which is adapted for deterministic forecast). This hybrid formulation blends the two error covariance components using a 50-50% ratio below 40 hPa. Because the EnKF top is at 2 hPa and the top few model levels are strougly affected by a numerical "sponge" (i.e. an enhanced horizontal diffusion), only the static component is used above 10 hPa. Between 40 hPa to 10 hPa, the weighting (β) of the two components gradually changes from $\beta_{EnKF} = 0.5$ and $\beta_{NMC} = 0.5$ to $\beta_{EnKF} = 0$ and $\beta_{NMC} = 1$ (Buehner *et al.*, 2013).

In the 4DEnVar, the cost function used to compute the analysis increment is preconditioned using $\boldsymbol{\xi} = \mathbf{B}^{-1/2}\delta\mathbf{x}$, such as

$$J_L(\boldsymbol{\xi}) = \frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\xi} + \frac{1}{2}(\mathbf{H}(\delta\mathbf{x}(\boldsymbol{\xi})) - \mathbf{y}')^T\mathbf{R}^{-1}(\mathbf{H}(\delta\mathbf{x}(\boldsymbol{\xi})) - \mathbf{y}') . \tag{2.6}$$

As explained in Buehner *et al.* (2013), the control vector $\boldsymbol{\xi}$ is made up of two components ξ_{EnKF} and ξ_{NMC} associated respectively with the ensemble ($\mathbf{B}_{EnKF}$) and static ($\mathbf{B}_{NMC}$) background-error covariances. They are combined to obtain the analysis increment $\delta\mathbf{x}$ as

$$\delta\mathbf{x} = \beta_{NMC}^{1/2}\mathbf{B}_{NMC}^{1/2}\xi_{NMC} + \beta_{EnKF}^{1/2}\mathbf{B}_{EnKF}^{1/2}\xi_{EnKF} . \tag{2.7}$$

Only observations passing a gross error and background check are assimilated (a variational quality control is also included as part of the minimization procedure) and the analysis increment is added to the background state using a 4D Incremental Analysis Update scheme (4D-IAU, Buehner *et al.*, 2015) to initialize the forecasts. The 4D-IAU distributes the time varying increments across the assimilation window to perform a smooth transition through time. This method also allows recycling key physical variables which are not directly analyzed (condensate mixing ratio, turbulent kinetic energy, turbulence regime, mixing length, friction velocity and ABL height, etc.) to reduce the forecast model spin up.

2.2.2. Cycling experiments with the operational systems

Performing experiments using the full operational system permits the evaluation of the impact of near-surface wind observations in a realistic context to show the value of such observations for short-term forecasts. In this case, the control run is simply based on the operational version of the 4DEnVar at Environment Canada while the experimental runs assimilate near-surface wind observations in addition to all operationally assimilated observations. This system is relatively sophisticated and it currently assimilates ~14 million observations per day from many different observing systems, which include radiosondes, aircrafts, land stations, ships, buoys, scatterometers, atmospheric motion vectors, satellite based radio occultation, microwave and infrared satellite sounders/imagers. Results from such experiments are presented in the second part of Chapter 4 (Section 4.4) to examine the full observation impact in the presence of all other assimilated observations.

2.2.3. Simplified non-cycling experiments

Throughout Chapter 3 and the first part of Chapter 4 (Section 4.3), a simplified approach is employed to assess observation impact on the assimilation system in a controlled environment. Rather than cycling the complete system, simplified non-cycling experiments are carried out using the 6 h forecasts from the operational system as background fields. Figure 2.1 presents the flow chart diagram of such simplified experiments.

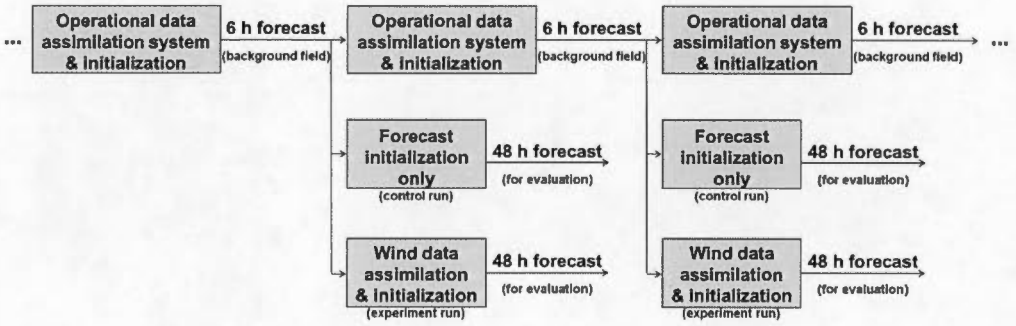

Figure 2.1: Flow chart diagram of the non-cycling assimilation experiments.

For the experimental runs, the background field from the operational 4DEnVar (not assimilating near-surface wind observations over land) is used to assimilate the near-surface wind observations over land. In this case, the analysis is computed by adding the increment to the background state at the middle of the assimilation window. The analysis is used along with the land surface analysis (e.g. soil moisture, surface temperature, etc.) from the operational system as initial conditions to initialize the forecast system (without 4D-IAU or any filtering).

The control is identical to the experiments, except that it does not assimilate the wind data. For cases where the experiment assimilates only the wind data, the control run assimilates no observations. Thus, the 6 h background field (from the operational system) is simply used to reinitialize the model and to simulate an analysis field to be used in the same way as the analysis from the experiments (to generate the 48 h control forecasts).

By performing non-cycled assimilation experiments, the cumulative impact from the observations is not captured. Still, by evaluating the quality of the resulting analyses against non-assimilated radiosonde observations for an impartial evaluation, the results can be directly attributed to near-surface wind observations because the analyses are not influenced by any other observation type or by background state differences.

2.2.4. Near-surface wind observation operator

To assimilate near-surface wind observations over land, a novel observation operator is developed in Chapter 3 based on the Geophysic Model Output Statistics (GMOS) proposed by Bédard *et al.* (2013). Although NWP model resolution can be relatively high (up to 2.5 km), they do not have a sufficiently refined grid to properly represent the meteorological phenomena over complex or coastal sites. To cope with such representation error, GMOS is based on multiple linear regressions relating the surrounding NWP grid-points to the observation site and exploit the correlations between the observing sites and the surrounding forecast points.

GMOS differs from other Model Output Statistics (MOS) that are widely used by meteorological centres as it implicitly takes into account the surrounding geophysical parameters (through its geo-statistical weights) such as surface roughness, terrain height, etc. From Bédard *et al.* (2013), GMOS can provide a more reliable representation of the site than common operator such as MOS or simple bilinear interpolation methods and the topographic signature of the forecast error (uneven distribution of the forecast error related to the surface characteristics) due to misrepresentation issues is significantly reduced.

Bédard *et al.* (2013) studied the influence of different predictor on near-surface winds over land (wind speed and direction, time of day, atmospheric stability, etc.). Results showed that the local wind components, the geophysical characteristics of the observing sites (implicitly accounted for by the multi-point linear regression) as well as wind direction (not wind speed) are the properties most significantly affecting GMOS impact on latitudinal and longitudinal wind forecasts. Atmospheric stability is also important for the vertical interpolation of the wind field, but it is already accounted for through the surface layer parameterization of the NWP model (to compute the 10 m wind fields) and thus, GMOS does not need to take this into account explicitly.

To tackle the anisotropic nature of the representativeness error, Bédard *et al.* (2013) developed a refined version of GMOS that considers the wind direction as a predictor. Due to the discontinuous nature of this operator and the relatively numerous statistical coefficients to train (requiring a large training dataset), a simplified version of the operator is implemented in the data assimilation system. Therefore, to assimilate near-surface wind over land, the GMOS observation operator is based on the multi-linear regression algorithm using the 10 m latitudinal (longitudinal) winds from the surrounding forecast grid-points to forecast the

latitudinal (longitudinal) wind at the observation site (located 10 m above ground level).

Based on Bédard *et al.*, (2013), the least mean squared error minimization algorithm is used to obtain the regression coefficients that minimize the observed minus forecast root mean squared value. Different regression coefficients are computed for each site because representativeness errors are site dependent.

CHAPTER III.

A GEO-STATISTICAL OBSERVATION OPERATOR FOR THE ASSIMILATION OF NEAR-SURFACE WIND DATA

This chapter presents a paper to appear in the *Quarterly Journal of the Royal Meteorological Society* (accepted on April 22, 2015):

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.*, **141** : 2857–2868.

It introduces a new geo-statistical observation operator (GMOS) developed to reduce biases and representativeness errors for the assimilation of near-surface winds over land (10 minutes averaged zonal and latitudinal winds). A new approach to estimate observation error correlations based on independent collocated observations is also presented. Only wind data from surface synoptic stations collocated with radiosonde observations (545 stations) are used. By evaluating the analysis against non-assimilated collocated observations, the impact of wind observations on the analyses and the influence of the background error statistics on the vertical propagation of the information are evaluated. The impact of the bias and representativeness error corrections on the analyses is also assessed.

A GEO-STATISTICAL OBSERVATION OPERATOR FOR THE ASSIMILATION OF NEAR-SURFACE WIND DATA

Joël Bédard [1],[a], Stéphane Laroche[b], and Pierre Gauthier[a]

[a] ESCER Centre, *Department of Earth and Atmospheric Sciences*
Université du Québec à Montréal (UQAM), Montréal (Québec), Canada

[b] *Data Assimilation and Satellite Meteorology Section*
Environment Canada, Dorval (Québec), Canada

(accepted version)

---

[1] Corresponding author: Joël Bédard, ESCER Centre
Department of Earth and Atmospheric Science
Université du Québec à Montréal
P.O. Box 8888, Downtown Station
Montréal (Québec) CANADA H3C 3P8
Email: bedard.joel@gmail.com

Abstract

Although many near-surface wind observations are available, very few are assimilated over land mainly due to sub-grid scale topographic interactions with the flow. The main objectives of this study are to understand the impact of near-surface wind observations on the analysis and to point out aspects that need to be improved to make a better use of these observations. A geo-statistical observation operator has been developed to correct for systematic and representativeness errors. Assimilation experiments are performed in a simplified context assimilating only near-surface wind observations over land in the ensemble-variational data assimilation system developed at Environment Canada. Due to the background-error covariances, the assimilation of near-surface wind observations impacts the lower part of the atmosphere. The resulting correction has been evaluated by verifying the analyses against non-assimilated collocated radiosonde data. This assessment also made it possible to estimate the observation error variance to strike a balance between having an important impact at the surface and maintaining a good vertical fit to upper air observations. Results from one month of assimilation experiments show that the geo-statistical operator eliminates biases and significantly reduces representativeness errors as well as observation error correlations in the analysis, mainly over complex terrain. Results also show that flow-dependent background error covariances from ensembles provide better vertical information propagation than static error statistics. Overall, the analysis fit to non-assimilated collocated radiosonde observations is improved when assimilating wind observations from surface stations.

Keywords: Observation error statistics, Representativeness error, Error correlations, Atmospheric boundary layer, Evaluation against collocated radiosonde observations

3.1. Introduction

Detailed evaluations of Numerical Weather Prediction (NWP) systems indicate that errors in initial conditions and atmospheric boundary layer (ABL) parameterizations appear to be the main limitations for short-range near-surface wind prediction capabilities. Bédard *et al.*, (2013) showed that, when compared with persistence, NWP models offer poor short-range near-surface wind forecasts (up to 6 h). Zack *et al.* (2010; 2011) showed that such predictions are sensitive to local initial conditions. Rostkier-Edelstein and Hacker (2010) suggests that the assimilation of near-surface observations can significantly improve nowcasting predictions, more than enhanced forecasting models. Thus, interest is growing in the improvement of NWP modeling by means of more accurate initial conditions defined by large-scale analyses. Although many observations describing the wind field in the lower troposphere are available from the global observing system, very few are assimilated over land mainly due to sub-grid scale topographic interactions with the flow. Until recently, near-surface wind observations over land were not used. However, with the increasing vertical and horizontal resolution of NWP models, finer topographic features are now resolved such that the assimilation of near-surface wind observations can be revisited.

A number of recent studies show that the assimilation of near-surface observations, including 10 m winds over land, can be beneficial for short-range forecasts in the lower troposphere (Hacker and Snyder, 2005; Dong *et al.*, 2010, Benjamin *et al.*, 2010, Ingleby, 2014). However, most of these studies exclude winds over complex terrain as representativeness errors can be significant. From its near-surface data assimilation assessment, Ingleby (2014) showed that biases and representativeness errors limit the global impact of near-surface wind observations. Winds from small islands, sub-grid scale headlands and tropical lands are still excluded from the UK

Met Office data assimilation system. Similarly, the Rapid Update Cycle (RUC) system uses narrow quality control parameters to prevent degrading the near-surface wind analysis due to representativeness errors (Benjamin *et al.*, 2007). Hence, representativeness and systematic errors in the analysis need to be addressed before assimilating near-surface wind observation over land.

Efforts are currently made to assess the representativeness error of surface and near-surface variables for air quality and wind energy predictions (Deng and Stull, 2005; Henne *et al.*, 2010; Koohkan and Bocquet, 2012; Bédard *et al.*, 2013). To cope with such misrepresentation issues, Koohkan and Bocquet (2012) showed that it is possible to include a station representativeness coefficient in the observation operator to assimilate carbon monoxide observations in the lower troposphere. A Geophysical Model Output Statistics (GMOS: Bédard *et al.*, 2013) operator was also developed for short-range near-surface wind forecast applications. This method combines a statistical error correction to a multiple grid-point approach. In the absence of collocated observations, GMOS cannot rely on bias correction schemes as used for satellite observations (Auligné *et al.*, 2007; Lea *et al.*, 2008; Dee and Uppala, 2009). It is difficult to separate the observation error bias from that of the background error and GMOS corrects for representativeness and background biases altogether. By its statistical error correction and by implicitly taking into account the geographical parameters of the surrounding grid-points, GMOS reduces representativeness errors for wind forecasts over complex sites (Bédard *et al.*, 2013).

Fundamentally, this geo-statistical operator only partially corrects representativeness errors. GMOS takes advantage of the correlation between resolved scales and unresolved scales to correct the stationary component of the representativeness error. However, Janjic and Cohn (2006) describe the representativeness error as flow-dependent and anisotropic. Even when using GMOS, the non-stationary component

of the representativeness error remains. To tackle the anisotropic nature of the representativeness error, Bédard *et al.* (2013) developed a refined version of GMOS that considers the wind direction as a predictor. Due to the discontinuous nature of this operator and the relatively numerous statistical coefficients to train (requiring a large training dataset), it is difficult to implement such an operator in a data assimilation system. The PBL regime is also important for the vertical interpolation of the wind field. As this is implicitly performed through the surface layer parameterization of the NWP model, GMOS does not need to integrate such property explicitly. The GMOS operator can be useful to compare model state variables with near-surface wind observations and reduce the representativeness errors that alter this comparison. The present work focuses on the reduction of the stationary and isotropic components of the representativeness error associated with local geographical characteristics. As Model Output Statistics (MOS) methods are typically used for forecast post-processing, it is appealing to use such statistical approaches in the data assimilation system to increase the consistency between the observations, analyses and forecasts.

This paper presents a novel observation operator to reduce the representativeness errors for the assimilation of near-surface winds based on the GMOS developed by Bédard *et al.* (2013). This operator has been implemented in the ensemble variational data assimilation system (4DEnVar) developed at Environment Canada (Buehner *et al.*, 2013). The objective of this work is to understand the observation impact on the analysis and point out aspects of the data assimilation system that need to be improved to make a better use of these observations. By using only surface stations that are collocated with upper-air stations, the statistical robustness of the operator and the actual representativeness error and bias corrections are examined in a controlled environment. A new approach to estimate observation error correlations based on independent observations is proposed. Data assimilation experiments over a

one-month period (February 2011) are carried out. Only wind data from surface stations are assimilated to carefully evaluate the observation impact on the analysis. The results presented can be directly attributed to near-surface wind observations as they are not influenced by any other observation type. The collocated radiosonde profiles, which are not assimilated, are used to assess the vertical propagation of information when different background error statistics are used.

3.2. Data assimilation system

Environment Canada 4DEnVar is a state-of-the-art variational data assimilation system which uses the 4D background error statistics derived from Ensemble Kalman Filtering (EnKF) techniques (Buehner *et al.*, 2013). The background error statistics used in the 4DEnVar comprise a stationary homogeneous (static) component and a flow-dependent (dynamic) component from the Environment Canada operational EnKF (Houtekamer *et al.*, 2014). While the static component is estimated using the NMC method (Parrish and Derber, 1992), the EnKF is using 256 members and the analysis increment is computed on a 800×400 grid (~50 km).

3.2.1. Background error statistics

To verify if the background error variances (σ_b^2) used in the data assimilation system are consistent with the observation error variances (σ_o^2), the innovation standard deviation (STD) is computed using radiosonde observations and 6 h background states provided by a reference EnVar assimilation cycle for February 2011:

$$\left\langle \left(\left(\mathbf{y} - H(\mathbf{x}_b) \right) - \left\langle \mathbf{y} - H(\mathbf{x}_b) \right\rangle \right)^2 \right\rangle = \sigma_o^2 + \sigma_b^2 , \tag{3.1}$$

where observations ($\mathbf{y}$) are compared to the background state ($\mathbf{x}_b$) using a bilinear interpolation (Bilin, hereafter) as observation operator H. Here, $\langle ... \rangle$ stands for the statistical average. This test depends on the observation error statistics being used. The radiosonde observation error is assumed to be properly tuned and implicitly contains representativeness errors. Reference σ_b^2 values (globally averaged) are obtained from the innovation variance based on radiosonde observations. The error variances from both the NMC and EnKF methods are compared to these reference values. Figure 3.1 presents the mean vertical structure of background and observation errors for wind components over land. Since the focus of this study is in the lower troposphere, only results from 1000 hPa to 500 hPa are presented. Hereafter, the results presented are an average of both zonal and meridional wind components as their error statistics are very similar.

In comparison with the σ_b value computed from the innovation variance, Figure 3.1 shows that the mean flow-dependent error has similar amplitude, whereas those from the static component are overestimated by ~40 %. Their vertical structures are similar on average, except near the surface where the EnKF background error statistics underestimates the wind error variance. This may be attributed to the fact that the surface analyses (e.g. surface roughness, albedo, soil moisture, etc.) are currently not perturbed in the EnKF (Lavaysse *et al.*, 2013). While the mean near-surface wind error in the static (flow-dependent) error covariance matrix is of 1.61 ms^{-1} (0.66 ms^{-1}), the innovation variance indicates that a value of 1.06 ms^{-1} is more realistic on average, assuming that the radiosonde observation error variance is accurate.

From the flow-dependent characteristics of the EnKF and from the statistical robustness of the NMC method, Buehner (2005) showed that 4DEnVar benefits from using both static and flow-dependent error covariance matrices. The hybrid formulation of the 4DEnVar blends the two error covariance components using a 50-50% ratio below 40 hPa. To allow for a detailed assessment of the vertical propagation of the information (Section 3.7), assimilation experiments are performed using each of the background error components independently as well as using the combination of both error covariances (hybrid formulation).

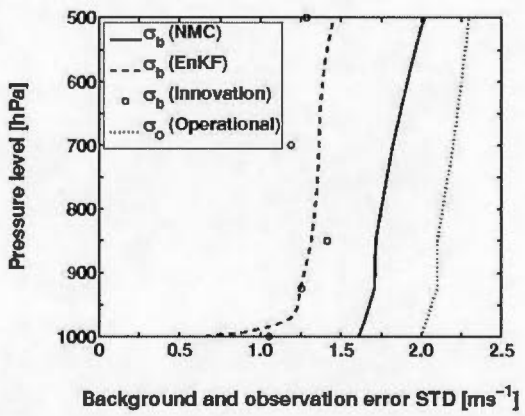

Figure 3.1: Mean vertical profiles of the NMC and EnKF background error STD over land compared to the σ_b values obtained from radiosonde wind component innovations (February 2011) as well as the σ_o values used to assimilate operational radiosonde observations.

3.3. Measurement error statistics

Data from a dozen tall meteorological towers (80 m), provided by Hydro-Québec, are used to assess the near-surface wind measurement error. These masts are heavily instrumented and present redundancy in terms of instruments to provide reliable measurements for wind energy assessments. Although the International Energy Agency standard for wind measurements is different than the WMO standard, the difference between vertically and horizontally collocated instruments from 80 m towers is used as it allows for an approximation of the measurement errors. Figure 3.2 presents the measurement differences as a function of the wind speed for direct (speed and direction) and indirect (zonal or meridional components) wind observations from similar instruments located at the same height on the same mast. The indirect observation error is represented by the mean error from both wind components.

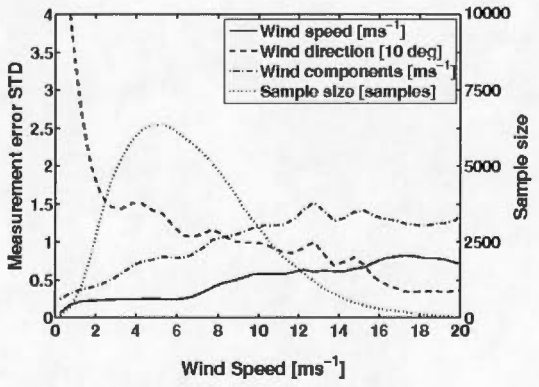

Figure 3.2: 80 m wind measurement error STD from ~100 000 samples as a function of observed wind speed.

While anemometer wind speed measurement errors are relatively proportional to the wind speed, the wind direction errors have a non-linear behaviour. When the wind speed drops below 2 ms^{-1}, the wind direction error significantly increases partly because wind direction variability increases. Comparing the horizontal wind components from these instruments, the resulting error characteristics are very similar to the wind speed error. Although it seems counter-intuitive, the measurement error for the horizontal wind component is larger than the wind speed error because the wind component error is function of the wind speed and direction errors modulated by trigonometric functions. As shown in Figure 3.2, the direction error is significant and outweighs the wind speed uncertainty when computing the wind components. The wind component measurements are still reasonably accurate owing to the precision of the instruments gathering wind observations. Nonetheless, the relative error in the horizontal wind components remains large for very light wind conditions (e.g. an error of ~0.4 ms^{-1} for a wind speed of ≤ 1 ms^{-1}). To circumvent this problem, Huang *et al.* (2013) proposed to assimilate near-surface wind observations in terms of wind speed and direction, rather than horizontal wind components. However, wind speed and direction are respectively positive definite and periodic variables. For simplicity and for consistency with Environment Canada observation processing scheme, horizontal wind components are used in this study. To preserve a good signal to noise ratio, only wind observations at wind speed higher than 1 ms^{-1} are considered. The wind speed frequency distribution (dotted line in Figure 3.2) shows that very light wind events are not frequent at 80 m above ground. An evaluation of the 10 m SYNOP observations confirms that the overall quantity of information brought to the system is not significantly affected from rejecting these light wind observations. The resulting averaged near-surface wind component measurement error (σ_m) is 0.94 ms^{-1}.

A similar experiment performed using collocated radiosonde observations was documented in a report from the World Meteorological Organization (Nash *et al.*, 2011). These horizontal wind retrievals are performed by directly measuring the radiosonde displacements using the global positioning system (GPS) technology. The measurement error for this type of observation is significantly smaller. However, observations following the release of the instrument are of poorer quality than upper level ones. From Nash *et al.* (2011), Chapter 11, results indicate that near-surface radiosonde wind measurement error is ~0.5 ms^{-1} on average.

3.4. Observation operator

3.4.1. Observation operator design

Bédard *et al.* (2013) designed a GMOS operator to reduce representativeness errors for short-range near-surface wind forecasts over complex sites. Similar to Deng and Stull (2005), the idea behind this geo-statistical operator is that, although they may be located near each other, the nearest grid-point may not properly represent the observing station, especially if the latter is located on complex terrain or a coastal site. Indeed, coastal grid-points often present aggregated onshore and offshore characteristics which are generally not representative of the onshore observing site. The NWP model considers topographic changes between grid points as steady slopes and cannot accurately represent sharp topographic details. NWP models also cannot reproduce sub-grid scale surface roughness changes. Consequently, the geographical parameters of the nearest grid-point can significantly differ from the observing site characteristics. On the other hand, amongst the surrounding grid-points, there may be one or many that are more representative of the observing site. Thus, Bédard *et al.*

(2013) replaced the underlying bilinear horizontal interpolation scheme used in MOS by a new set of geo-statistical weights relating the surrounding NWP grid-points to the observation site. By exploiting the correlations between the observing sites and their surrounding forecast points, GMOS provides a more reliable representation of the site. The MOS operator (H_{MOS}) is formulated as

$$H_{MOS}(\mathbf{x}) = A \cdot L(\mathbf{x}) + C, \tag{3.2}$$

where $\mathbf{x}$ is either the zonal or meridional modeled wind component, L is the bilinear interpolation operator (Bilin), A is an amplitude coefficient and C is the systematic error correction coefficient. The GMOS multi-point linear regression (H_{GMOS}) is defined as

$$H_{GMOS}(\mathbf{x}) = \sum_i (A_i x_i) + C, \tag{3.3}$$

where A_i are amplitude coefficients, C is the systematic error correction coefficient and the subscript i is the index of the surrounding grid-points. As in MOS, the systematic error coefficient corrects for biases while the amplitude coefficient adjusts the forecast temporal variability to best fit observations. The surrounding grid-points used by the GMOS operator are distributed around the observing site in an N×N square pattern. Using more grid-points increases the probability of considering a point with geographical characteristics similar to the site and thus further reduces the representativeness error. However, it also increases the number of statistical parameters to train and thus, a larger training dataset is needed.

Based on Bédard *et al.*, (2013), the least mean squared error minimization algorithm is used to obtain the regression coefficients A and C that minimize the innovation (observed minus forecast residual) root mean squared value. Different regression coefficients are computed for the zonal and meridional wind components. This allows predicting the local zonal (meridional) wind component using the modeled zonal (meridional) winds from the surrounding N×N grid-points. The wind component observations from 545 SYNOP stations collocated with radiosonde stations are used every 6 hours. These observations are compared with corresponding short-range (0-12 hours) wind forecasts updated twice a day from Environment Canada global deterministic prediction system (Global Environmental Multiscale model at 0.35°×0.23° latitude/longitude resolution (~25 km): Côté *et al.*, 1998a; Charron *et al.*, 2012). While the time period used to train the observation operators is different from that used in the assimilation experiments (November 2012 to June 2013 vs. February 2011), the forecasts are generated using the same model configuration.

3.4.2. Training and statistical robustness of statistical operators

Eight months (November 2012 to June 2013) of near-surface wind forecasts and observations are used to assess the statistical robustness of different GMOS architectures (varying the number of grid-points used). The diminishing return achieved from increasing the length of the training period is evaluated using the innovation bias and STD calculated on an independent dataset. For each case, at least 1 month of data is kept for validation. This evaluation was repetitively performed using observations from different seasons. Results are insensitive to the season used for both training and validation. Figure 3.3 shows the average diminishing return achieved when using a longer training period.

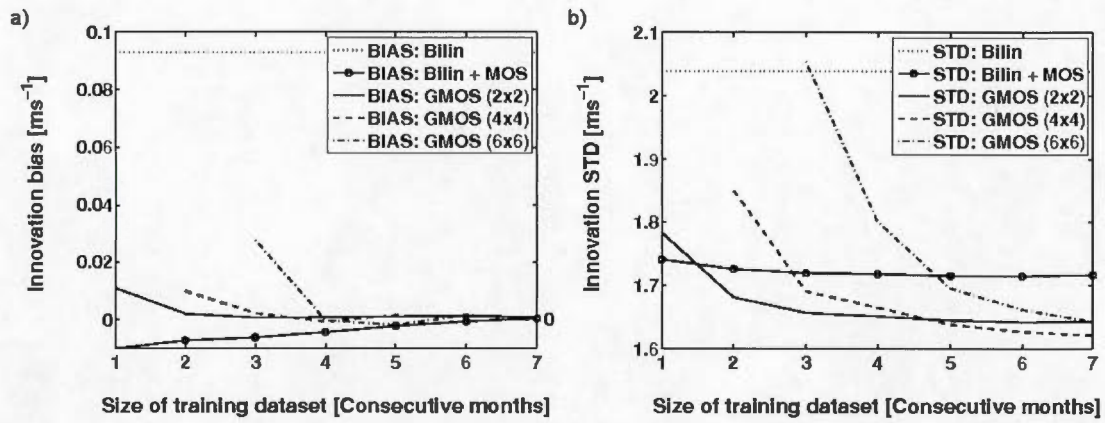

Figure 3.3: 10 m wind components innovation (observed minus background states) bias and STD as a function of the training dataset length using Bilin, Bilin coupled with a MOS, and GMOS. The number of grid-points used for the GMOS operators is specified in parentheses (e.g. 2×2 when using the 4 nearest grid-points).

Figure 3.3 indicates that both MOS and GMOS significantly reduce the innovation STD with respect to Bilin. In general, a GMOS approach using more grid-points performs slightly better than the simpler ones. However, the added complexity requires more training data to fit the additional coefficients. When applied to an independent dataset, the solutions from the geo-statistical operator using N×N grid-points generally diverge when using less than N/2 months of training data and converge towards a robust solution when using at least N months of data. The introduction of a statistical amplitude error correction (MOS) in the observation operator does reduce the innovation STD (~0.3 ms^{-1}). When properly trained (training dataset $\geq$ N months), the multi grid-point approach (GMOS) further reduces this criteria by ~0.1 ms^{-1}. Both statistical operators (MOS and GMOS) also significantly reduce the bias (Figure 3.3, a). However, the bias obtained with MOS only becomes comparable with GMOS when a sufficiently large training period is used ($\geq$5 months). This is due to GMOS ability to select the most representative grid-point. For

a proper comparison between the various operators, the GMOS architecture using 2×2 grid-points is employed hereafter. In this case the MOS and GMOS operators use the same forecast grid-points as Bilin. This is also the simplest and most robust GMOS scheme.

3.4.3. Temporal variability

It is known that simple averaging methods such as Bilin or MOS reduce the forecast variability (smoothing effect). Table 3.1 compares the mean on-site observation and forecast standard deviation (so-called temporal variability) for the different operators to show that GMOS does not have this limitation. Results confirm that both Bilin and MOS smooth the forecasts and increase the difference between forecast and observed temporal variability. On the other hand, GMOS decreases this difference. By using amplitude coefficients, the sum of the weights attributed to the surrounding grid-points is not always equal to 1 when using GMOS. This operator is thus able to use the most representative forecast points and modulate their signal to adjust the forecast temporal variability towards the natural on-site observation variability.

3.4.4. Representativeness error correction

To assess the error reduction achieved (Figure 3.3) when using these different observation operators, the innovation **d** is used. It is defined as

$$\mathbf{d} = \mathbf{y} - H(\mathbf{x}_b) = (\mathbf{y} - \mathbf{y}_t) - (H(\mathbf{x}_b - \mathbf{x}_t) - (\mathbf{y}_t - H(\mathbf{x}_t)) = \boldsymbol{\varepsilon}_m + \boldsymbol{\varepsilon}_r - H(\boldsymbol{\varepsilon}_b) \tag{3.4}$$

where observations ($\mathbf{y}$) are compared to the background state ($\mathbf{x}_b$) using an observation operator H. The subscript t denotes the true state in either the observation or model space ($\mathbf{y}_t$ and $\mathbf{x}_t$ respectively). Expressing the innovation ($\mathbf{d}$) in terms of measurement, representativeness and background errors ($\boldsymbol{\varepsilon}_m = \mathbf{y} - \mathbf{y}_t$, $\boldsymbol{\varepsilon}_r = \mathbf{y}_t - H(\mathbf{x}_t)$ and $\boldsymbol{\varepsilon}_b = \mathbf{x}_b - \mathbf{x}_t$ respectively), the innovation variance is statistically decomposed in terms of measurement, representativeness and background error variances. Therefore,

$$\begin{aligned}\left\langle (\mathbf{d} - \langle \mathbf{d} \rangle)^T (\mathbf{d} - \langle \mathbf{d} \rangle) \right\rangle &= tr\left(\left\langle (\mathbf{d} - \langle \mathbf{d} \rangle)(\mathbf{d} - \langle \mathbf{d} \rangle)^T \right\rangle \right) \\ &= tr\left(\left\langle \boldsymbol{\varepsilon}'_m \boldsymbol{\varepsilon}'^T_m \right\rangle \right) + tr\left(\left\langle \boldsymbol{\varepsilon}'_r \boldsymbol{\varepsilon}'^T_r \right\rangle \right) + tr\left(H \left\langle \boldsymbol{\varepsilon}'_b \boldsymbol{\varepsilon}'^T_b \right\rangle H^T \right),\end{aligned} \tag{3.5}$$

where $\boldsymbol{\varepsilon}'$ is the unbiased component of the error. This expression is equivalent to (3.1) as the observation error is generally considered to include both the measurement and representativeness errors ($\left\langle \boldsymbol{\varepsilon}'_o \boldsymbol{\varepsilon}'^T_o \right\rangle = \left\langle \boldsymbol{\varepsilon}'_m \boldsymbol{\varepsilon}'^T_m \right\rangle + \left\langle \boldsymbol{\varepsilon}'_r \boldsymbol{\varepsilon}'^T_r \right\rangle$ for uncorrelated measurement and representativeness errors).

Table 3.1: Mean on-site wind component observation and forecast standard deviation using different observation operators: Bilin, MOS, and GMOS.

Dataset	Variance (ms^{-1})	Difference (ms^{-1})
Observation	3.78	-
Bilin	3.28	0.50
MOS	3.19	0.59
GMOS	3.98	-0.20

The same observations and background fields are used when computing the innovations for each observation operator and the measurement error variance remains the same. However, the error variance of the interpolated background values is influenced by the observation operator used through $H\left\langle \boldsymbol{\varepsilon}_b \boldsymbol{\varepsilon}_b^T \right\rangle H^T = H\mathbf{B}H^T$, where $\mathbf{B}$ is the background error covariance matrix. Considering that GMOS amplitude coefficients are used to adjust the forecast temporal variability to best fit the observations, the sum of weights attributed to the surrounding grid-points is not always equal to 1 (as opposed to Bilin). When computing $H\mathbf{B}H^T$, the background error variance in observation space is 42% smaller on average for the GMOS than for Bilin. A global influence coefficient (α) is therefore introduced to account for the observation operator influence on $H\mathbf{B}H^T$ such that $tr(H\mathbf{B}H^T) = \sum \alpha^2 \sigma_b^2$.

Using $tr(H\mathbf{B}H^T) = \sum \alpha^2 \sigma_b^2$ in (3.5), the representativeness error statistics are obtained by subtracting the measurement and background error variances from the innovation variance. From Figure 3.3, the innovation STD is reduced from 2.04 to 1.65 ms^{-1}, which corresponds to a variance reduction of 1.46 m^2s^{-2}. Following Eq. (3.5), this significant variance reduction is attributed to a reduction of the background error variance in observation space and a reduction of the representativeness error variance by the GMOS operator (0.47 m^2s^{-2} and 0.99 m^2s^{-2} respectively). These results indicate that the GMOS operator provides significantly lower representativeness errors compared to Bilin (globally, σ_r = 1.21 and 1.42 ms^{-1} respectively).

To better understand the error correction provided by GMOS, innovations are computed for each of the 545 observation sites. Figure 3.4 presents the frequency distribution of the innovation STD for 0.1 ms^{-1} bin intervals. When comparing the frequency distribution of both Bilin and GMOS operators, the improvement using the

geo-statistical operator is systematic (~0.4 ms^{-1}). Sites that featured the highest innovation STD when using Bilin are also the most improved by the GMOS operator (observation error reduced by ≥ 1 ms^{-1}). These specific sites (36 stations in total) are plotted on a World map in Figure 3.5.

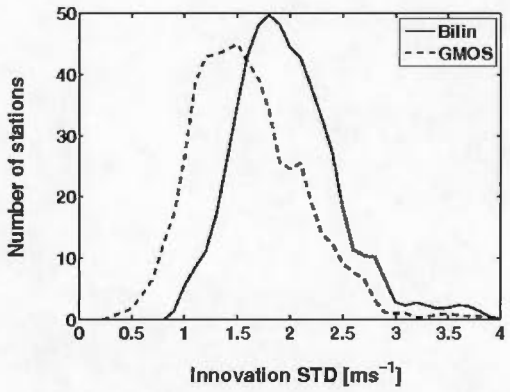

Figure 3.4: Frequency distribution of the wind components innovation STD for the 545 SYNOP stations using either Bilin or GMOS operator.

From Figure 3.5, it is clear that the GMOS operator has a significant impact on coastal sites, small islands and shoreline of major rivers. Also, a number of airports, small farming/desert villages (some sites in Russia, Germany and China) and relatively dense urban environments (some sites in Russia, Bangladesh and Kazakstan) are significantly impacted by the use of GMOS. These sites are located in complex environments having inhomogeneous surface roughness that the NWP model cannot reproduce accurately. Sites on shorelines also generally suffer from high representativeness errors as coastal NWP grid-points use aggregated surface fluxes and roughness, thus combining land and sea flow characteristics.

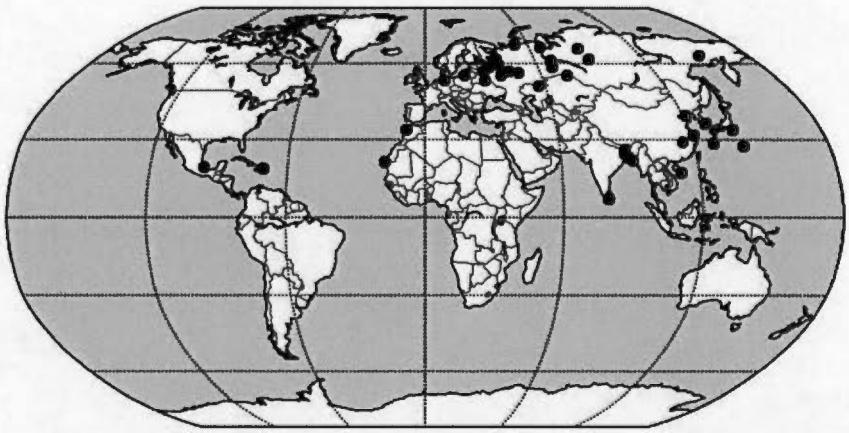

Figure 3.5: Spatial distribution of the sites benefitting from a significant observation error STD reduction ($\Delta\sigma_o \geq 1$ ms^{-1}) using the GMOS operator as opposed to Bilin.

When decreasing the observation error reduction threshold from 1.0 ms^{-1} to 0.5 ms^{-1} (not shown) the number of impacted sites increases significantly (from 36 to 139 sites out of 545) and includes complex mountainous areas (e.g.: Himalayas, Rocky Mountains, Alps, Ural, Drakensberg, Carpathian Mountains). On the other hand, differences between modeled and observing site topographic height, as well as model height variability around the forecast site do not appear to have direct correlations with the error correction (not shown). It is hard to attribute GMOS improvements to a specific geographic parameter as its benefits are site dependent. But, it is clear that the model cannot reproduce abrupt topographic changes because of its limited horizontal resolution (0.35°×0.23° here). In most cases, GMOS helps cope with such misrepresentation by giving more weight to the most representative grid-points of a given site. Hence, it appears that GMOS is mainly correcting for representativeness errors related to surface roughness, coastal effects and terrain complexity, but the improvements are not directly related to the topographic height itself.

3.4.5. Observation error statistics

The method proposed by Desroziers *et al.* (2005) can be used to assess the error covariance matrices used in data assimilation systems. This method needs to be applied iteratively (which is computationally expensive) and the outputs of the method significantly vary depending on the set of observation types assimilated in the system. To be properly applied, the Desroziers method needs to be performed while assimilating all observation types, which is beyond the scope of this study. Therefore, the evaluation of the observation error statistics is simply based on the innovation variances following Eq. (3.1), which is relatively straightforward.

This methodology can provide site dependent error variances, which would be helpful to better represent the local representativeness error related to the characteristics of the observation site. The establishment of site dependent error statistics is beyond the scope of the present work and is examined in future work using coherent site dependent observation operators and error variances. Thus, a global observation error value is computed for each observation operator. It should be noted that the error decomposition presented in Section 3.4.4 is only used to examine the representativeness error correction, not to establish σ_o.

As mentioned previously, both homogeneous and flow-dependent background error components from the operational systems (subscript *op*: $\sigma^2_{b_{op}}$) are sub-optimal near the surface. To be consistent with these error statistics, the observation error variances evaluated experimentally (subscript *ex*: $\sigma^2_{o_{ex}}$) are scaled as follows:

$$\sigma^2_{o_{op}} = \sigma^2_{o_{ex}} \frac{\sigma^2_{b_{op}}}{\sigma^2_{b_{ex}}} . \qquad (3.6)$$

Assuming the radiosonde observation error variance is correctly prescribed, the resulting SYNOP wind observation error statistics ($\sigma_{o_{op}}$) for Bilin and GMOS operators are respectively 2.59 and 2.33 ms^{-1} when using the static background error covariances; 1.06 and 0.96 ms^{-1} when using the flow-dependent covariances; and 1.98 and 1.79 ms^{-1} when using the hybrid covariances.

3.4.6. Tangent linear and adjoint models

GMOS relies on linear regressions and the derivation of its tangent linear and adjoint models is straightforward. In an incremental data assimilation system such as the 4DEnVar, the analysis increment is computed at a lower horizontal resolution with respect to the resolution of the background state used to calculate the innovation vector. The basic analysis cost function $J(\mathbf{x})$ is

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}(H(\mathbf{x})-\mathbf{y})^T\mathbf{R}^{-1}(H(\mathbf{x})-\mathbf{y}), \tag{3.7}$$

and in its incremental form, $J(\mathbf{x})$ becomes $J_L(\delta\mathbf{x})$ following

$$J_L(\delta\mathbf{x}) \approx \frac{1}{2}(\delta\mathbf{x})^T\mathbf{B}^{-1}(\delta\mathbf{x}) + \frac{1}{2}(\mathbf{H}'(\delta\mathbf{x})-\mathbf{y}')^T\mathbf{R}^{-1}(\mathbf{H}'(\delta\mathbf{x})-\mathbf{y}'), \tag{3.8}$$

where $\mathbf{x}_b$ stands for the background state, $\delta\mathbf{x} = (\mathbf{x} - \mathbf{x}_b)$ is the analysis increment, $\mathbf{y}' = \mathbf{y} - H(\mathbf{x}_b)$ represents the innovation vector, while the terms $\mathbf{B}$ and $\mathbf{R}$ represent respectively the background and observation error covariances matrices. In the

incremental form, the observation operator H is replaced by its Jacobian linearized around the background state, denoted by $\mathbf{H'}$, for the interpolation of the analysis increment (Gauthier *et al.*, 2007). Since the GMOS coefficients are computed from historical high-resolution innovations, they cannot be directly used to interpolate the low-resolution increments to the observation site. Therefore, in the minimization process, the low-resolution increments are locally interpolated on the high resolution grid using Bilin and then interpolated to the observation site using the GMOS operator such as

$$\mathbf{H'}(\delta\mathbf{x}) = \frac{\partial H}{\partial \mathbf{x}}(L(\delta\mathbf{x})) \quad . \tag{3.9}$$

By proceeding this way, the analysis cost function is consistent with the fact that, once computed, the low-resolution analysis increment ($\delta\mathbf{x}_a$) is interpolated to the high-resolution grid using Bilin (L) and added to the background field to generate the high resolution analysis such as $\mathbf{x}_a = \mathbf{x}_b + L(\delta\mathbf{x})$. The minimization of the low-resolution cost function and the high resolution analysis are mathematically consistent. This coherence also prevails in the application contexts where GMOS is used as a post-processing module (e.g. forecasts for wind energy purposes). With $\frac{\partial H}{\partial \mathbf{x}}(L(\delta\mathbf{x})) = \sum_i A_i\left(L_i(\delta\mathbf{x}_i)\right)$ and $\delta\mathbf{x} = \mathbf{B}^{1/2}\xi$, the implemented cost function is thus

$$J_L(\xi) = \frac{1}{2}\xi^T\xi + \frac{1}{2}\left(\sum_i \mathbf{A}_i\left(L_i(\mathbf{B}^{1/2}\xi_i)\right) - \mathbf{y}'\right)^T\mathbf{R}^{-1}\left(\sum_i \mathbf{A}_i\left(L_i(\mathbf{B}^{1/2}\xi_i)\right) - \mathbf{y}'\right) \quad , \tag{3.10}$$

where A are the GMOS amplitude coefficients for either the zonal or meridional wind component at the i^{th} surrounding grid-point.

Eq. (3.10) shows that, through the minimization process, the geo-statistical observation operator allows observations to influence consistently the innovation and the propagation of the information. By constraining the analysis increment using GMOS, observations mainly affect the analysis grid-points that are the most representative of the observation site. Still, due to the lower resolution increments, this heterogeneous impact is smoothed.

3.5. Observation error correlations

Several methodologies based on innovation and analysis residual stastistics have been developed to assess the observation error statistics (e.g. Hollingsworth and Lönnberg, 1986; Desroziers *et al.*, 2005). These studies assume observation errors are uncorrelated. Here we propose a new approach in which non-assimilated collocated observations are used to assess the representativeness error correlations associated with different observation operators.

To do this, only near-surface wind observations from the SYNOP stations ($\mathbf{y}_S$) are assimilated. The resulting analysis is $\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}\left(\mathbf{y}_S - H_S\left(\mathbf{x}_b\right)\right)$, where H_S is the SYNOP wind observation operator and $\mathbf{K} = \mathbf{B}\mathbf{H}_S^T(\hat{\mathbf{R}}_S + \mathbf{H}_S\mathbf{B}\mathbf{H}_S^T)^{-1}$ represents the gain matrix associated with the assimilation of near-surface data. Taking $\mathbf{H}_S$ and $\mathbf{B}$ as given, the analysis error is estimated by comparing the analysis to non-assimilated collocated radiosonde observations ($\mathbf{y}_R$) for an independent evaluation. Although $\mathbf{y}_S$ and $\mathbf{y}_R$ are horizontally and temporally collocated, they differ as winds are measured with different instruments at different heights. The analysis quality is estimated from the analysis residuals, or analysis departures with respect to the radiosonde data

($\mathbf{a}_R = \mathbf{y}_R - H_R(\mathbf{x}_a)$, where H_R is the radiosonde wind observation operator). The analysis error variance defines the error metric $Err(\hat{\sigma}^2_{o_S}, \mathrm{r}_\mathrm{r})$ used to assess the representativeness error correlations associated with different observation operators. $Err(\hat{\sigma}^2_{o_S}, \mathrm{r}_\mathrm{r})$ is a function of the surface observation error variance prescribed in the data assimilation system $\left(\hat{\sigma}^2_{o_S}\right)$ and the representativeness error correlations (r_r) such as

$$Err(\hat{\sigma}^2_{o_S}, \mathrm{r}_\mathrm{r}) = \left\langle \left(\mathbf{a}_R - \langle \mathbf{a}_R \rangle\right)^T \left(\mathbf{a}_R - \langle \mathbf{a}_R \rangle\right) \right\rangle = tr\left(\left(\mathbf{a}_R - \langle \mathbf{a}_R \rangle\right)\left(\mathbf{a}_R - \langle \mathbf{a}_R \rangle\right)^T \right). \tag{3.11}$$

Since the analysis departures can be represented using information concerning the background error ($\boldsymbol{\varepsilon}_b = \mathbf{x}_b - \mathbf{x}_t$) and the observation error ($\boldsymbol{\varepsilon}_o = \mathbf{y} - \mathbf{y}_t$), this yields

$$\begin{aligned} \mathbf{a}_R &= \mathbf{y}_R - H_R(\mathbf{x}_b + \mathbf{K}(\mathbf{y}_S - H_S(\mathbf{x}_b))) \\ &= (\mathbf{y}_R - H_R(\mathbf{x}_b)) - (H_R\mathbf{K}(\mathbf{y}_S - H_S(\mathbf{x}_b))) \\ &= \left(\boldsymbol{\varepsilon}_{o_R} - H_R\boldsymbol{\varepsilon}_b\right) - H_R\mathbf{K}(\boldsymbol{\varepsilon}_{o_S} - H_S\boldsymbol{\varepsilon}_b) \end{aligned}, \tag{3.12}$$

The observation error variances prescribed in the data assimilation system ($\hat{\mathbf{R}}_S$) may not be optimal: they may differ from the real observation error variances (estimated in Section 3.4.5) defined as $\mathbf{R}_S = \left\langle \boldsymbol{\varepsilon}'_{o_S} \boldsymbol{\varepsilon}'^T_{o_S} \right\rangle$, where $\boldsymbol{\varepsilon}'_{o_S}$ is the unbiased component of the error. This is why they are both accounted for independently. Moreover, as the surface and radiosonde measurement (nearest to the surface) are geographically very close, they are impacted by the same representativeness error associated with sub-grid scale variability. They may be correlated and this needs to be taken into account. It is shown in the Appendix that

$$Err\left(\hat{\sigma}_{o_S}^2, r_r\right) = \sum \begin{pmatrix} \sigma_{o_R}^2 + \left(\left(\frac{\alpha^2\sigma_b^2}{\hat{\sigma}_{o_S}^2 + \alpha^2\sigma_b^2}\right) - 1\right)^2 \sigma_b^2 + \left(\frac{\alpha\sigma_b^2}{\hat{\sigma}_{o_S}^2 + \alpha^2\sigma_b^2}\right)^2 \sigma_{o_S}^2 \\ -2\left(\frac{\alpha\sigma_b^2}{\hat{\sigma}_{o_S}^2 + \alpha^2\sigma_b^2}\right) r_r \sigma_{r_S} \sigma_{r_R} \end{pmatrix} \tag{3.13}$$

where r_r is the representativeness error correlation parameter and α is the coefficient introduced to account for the influence of the observation operator (**H**) on the background error (**B**) through the term **HBH**^{T}. Setting the derivative of $Err(\hat{\sigma}_{o_S}^2, r_r)$ with respect to $\hat{\sigma}_{o_S}^2$ to zero yields

$$\hat{\sigma}_{o_S}^2 = \alpha\sigma_b^2 \frac{\left(\sigma_{o_S}^2 - \alpha r_r \sigma_{r_S} \sigma_{r_R}\right)}{\left(\alpha\sigma_b^2 + r_r \sigma_{r_S} \sigma_{r_R}\right)} = \alpha\sigma_b^2 \frac{\left(\sigma_{m_S}^2 + \sigma_{r_S}^2 - \alpha r_r \sigma_{r_S} \sigma_{r_R}\right)}{\left(\alpha\sigma_b^2 + r_r \sigma_{r_S} \sigma_{r_R}\right)} \tag{3.14}$$

The representativeness error correlation (r_r) must be assessed with caution as it does affect the optimality of the observation error variance prescribed ($\hat{\sigma}_{o_S}^2$) to the system. This means that, without considering r_r explicitly, both the analysis error and the optimal observation error cannot be calculated using (3.14). On the other hand, using σ_b, σ_m and σ_r as estimated in Sections 3.2, 3.3 and 3.4, it is possible to assess the representativeness error correlations associated with different observation operators by fitting $Err(\hat{\sigma}_{o_S}^2, r_r)$ to experimental results.

3.6. Evaluation of the impact of observation operators

The analysis error metric $Err(\hat{\sigma}^2_{o_s}, r_r)$ defined in Section 3.5 is compared with experimental results to evaluate the impact of the various observation operators on analyses and assess the representativeness error correlations. For each set of experiments, the prescribed observation errors value ($\hat{\sigma}^2_{o_s}$) is varied (from 0.1 ms^{-1} to 10 ms^{-1}) to make a fit consistent with Eq. (3.13). Although experiments are performed using different background error statistics (static, dynamic, and hybrid), only the results for the static background error covariance matrix are presented to evaluate the observation operators in a simplified context. The resulting 60 analysis experiments performed are listed in Table 3.2. As this study focuses on Bilin and GMOS, these operators present 19 experiments each, whereas Bilin + bias correction and Bilin + MOS only have 11.

The same operational background fields are used for all experiments to study the impact of near-surface wind observations on the analysis in a simplified context. The fact that the analyses are not cycled allows performing experiments in a controlled environment, but does omit the cumulative impact from the observations. Each analysis experiment is conducted over a one month period (February 2011) assimilating only near-surface wind observations from the 545 SYNOP stations collocated with radiosondes. Analyses are produced twice daily considering only observations in the middle of the 6 h assimilation window. To ensure a good quality control, gross error check, background check and variational quality control are systematically applied on all datasets. The background check is always performed using Bilin to ensure that all experiments benefit from the same observations and to avoid feedback between the quality control and the statistical correction performed (from the statistical coefficient training: Auligné and McNally, 2007).

Table 3.2: Configuration of the 60 analysis experiments performed. For each observation operators implemented, the different observation error value and background error matrix prescribed to the data assimilation system as well as the number of experiments are listed.

Operator	**Number of experiments**	**Observation errors values** ($\hat{\sigma}^2_{o_S}$ **ms^{-1}**)	**Background error**
Bilin	19	0.1, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5, 2.75, 3.0, 3.25, 3.5, 3.75, 4.0, 4.25, 4.5, 6.0, 8.0, 10.0	Static
Bilin + bias correction	11	0.1, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 6.0, 8.0, 10.0	Static
Bilin + MOS	11	0.1, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 6.0, 8.0, 10.0	Static
GMOS	19	0.1, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5, 2.75, 3.0, 3.25, 3.5, 3.75, 4.0, 4.25, 4.5, 6.0, 8.0, 10.0	Static

The resulting analyses are verified against collocated radiosonde data for an independent validation. Results are compared to the estimated analysis error from Eq. (3.13). Figure 3.6 presents the analysis error STD from both the analysis experiments and the analytical derivation. As mentioned in Section 3.5, the analytical results depend on the correlation of representativeness error between winds from SYNOP and radiosonde stations. Figure 3.6 presents results for different representativeness error correlation coefficients (r_r = 0.0, 0.5 and 1.0) and also gives the r_r value that provide the best fit with experimental results.

Figure 3.6 shows that when using a high observation error variance value ($\hat{\sigma}_o^2$), the observations have a relatively low impact and the analysis remains close to the background state. At the point where the optimal σ_o is reached (as defined in Section 3.4), the impact becomes maximal and the independent analysis departure is minimized. Also, when the observation error is further reduced, the analyses start degrading. Overall, the near-surface observations improve the local analyses when using the optimal value for $\hat{\sigma}_o$, as validated against non-assimilated radiosonde observations. Still, the analysis departure appears to be smaller when using Bilin because the verification is not against observation with independent observation error. As mentioned in Section 3.5, it is misleading to evaluate the analysis error without considering r_r explicitly. To perform a fair evaluation of the different observation operators, only the representativeness error correlation coefficients (r_r) are compared.

Figure 3.6 shows that the representativeness errors from both observation types are highly correlated when using Bilin (r_r = 0.64). Because the near-surface radiosonde and SYNOP observations are nearly collocated, it is clear that they can both suffer from similar representativeness error associated with local sub-grid scale effects (e.g. coastal effects). Still, the site dependent bias correction slightly reduces the representativeness error correlation (r_r = 0.58). On the other hand, when using GMOS, Figure 3.6 shows that the representativeness error correlations are significantly reduced to r_r = 0.22. Similar results are obtained when coupling a MOS to Bilin (r_r = 0.23). It was shown previously that the multi-point scheme further reduces the representativeness error (Figure 3.3), but as the correlation between the representativeness errors is already small when using MOS, the additional improvement using GMOS is slight in Figure 3.6. Overall, the statistical error correction from either MOS or GMOS improves the average fit to the observations. Their representativeness error reduction also attenuates the near-surface wind

observation error correlation with the representativeness errors from radiosonde observations.

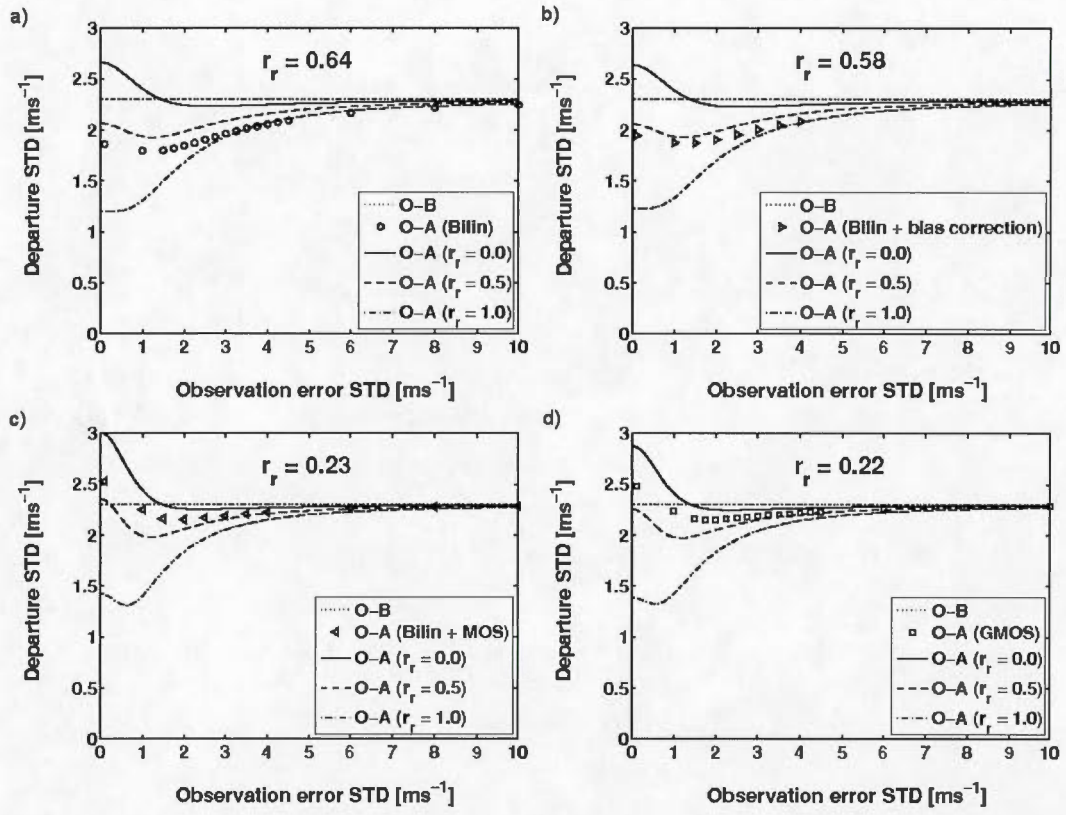

Figure 3.6: Comparison between the wind error metric (lines) and experimental results (symbols) for different observation operators: Bilin (a), Bilin with bias correction (b), Bilin with both bias and amplitude corrections from MOS (a), and GMOS (d). Results are plotted for different specified values of the representativeness correlation coefficients (r_r = 0.0, 0.5, and 1.0). The analysis error variance is shown as a function of the prescribed observation error variance for each of the experiments. While the symbols represent the experimental results, the lines are the corresponding representation of the analysis error variance from (3.13) evaluated for three values of the representativeness error correlation.

3.7. Results from non-cycling assimilation experiments

After performing the assimilation experiments with the observation error variance estimated in Section 3.4, the analyses are compared to collocated radiosonde observations. The departure STD is computed at different levels to evaluate the vertical impact from near-surface wind observations. Stations were first divided into two groups (inland / coastal), but as differences are not significant, the results are presented for all the stations altogether. Figure 3.7 shows vertical profiles of the analysis departure STD for cases using static, dynamic and hybrid background error statistics (using their optimal σ_o). The analysis departure STD with respect to radiosonde observations is normalized by the innovation STD to condense the information. Results can thus be represented for all vertical levels in the same figure. The 90% confidence intervals are also plotted to fully appreciate the improvement brought by the assimilation of near-surface wind observations.

The confidence intervals (shaded areas) in Figure 3.7 show that results are highly variable, which makes it difficult to affirm that the 2 curves are significantly different. Subsequently, observation operators are not compared in this section. Only the vertical propagation is examined. When comparing the analysis to the collocated radiosonde observations, the static background error statistics propagate the near-surface wind information up to 600 hPa and the impact is significantly positive in the ABL. However, the impact appears to be slightly negative between 600 and 900 hPa. As the static vertical correlation length does not account for the atmospheric stability, the analysis is degraded above the ABL where the correlations should vanish due to the decoupling with the free atmosphere. The flow-dependent error covariances provide a positive impact for all vertical levels from the surface up to 750 hPa. Again, this improvement is significant in the ABL. By accounting for atmospheric stability, the flow-dependent error statistics limit the vertical propagation of the information

for stably stratified cases. As a result, the vertical propagation is slightly shallower (750 hPa vs. 600 hPa), but the impact is strictly positive. Also, by its flow dependency, the error statistics from ensembles give a more appropriate weight to both the background and observations, thus generally enhancing the analyses by approximately 1 % (except near the surface where the spread of the ensemble is lacking). For the hybrid formulation, the degradation from the static component above 900 hPa is compensated by the positive impact from the flow-dependent error covariance matrix. As expected, the hybrid configuration provides a positive impact from the surface up to 900 hPa, but does not yet seem as good as the flow-dependent configuration above 925 hPa.

From the experiments using the static error covariance matrix, it is found that the near-surface cross-correlations between variables are very small as the impact from wind observations on temperature and humidity is negligible (not shown). The inter-variable correlations are larger in the flow-dependent error statistics due to the dynamic coupling between wind and temperature variables. Still, the improvement brought by the assimilation of near-surface wind observations on temperature fields is small (≤ 2 % up to 750 hPa) and does not extend beyond the 90 % confidence interval (not shown). Similarly, the correlations between wind and humidity fields are somewhat negligible. The hybrid formulation of the background error covariances presents similar behaviours, but with smaller amplitudes (not shown) due to the blending between flow-dependent and static error components.

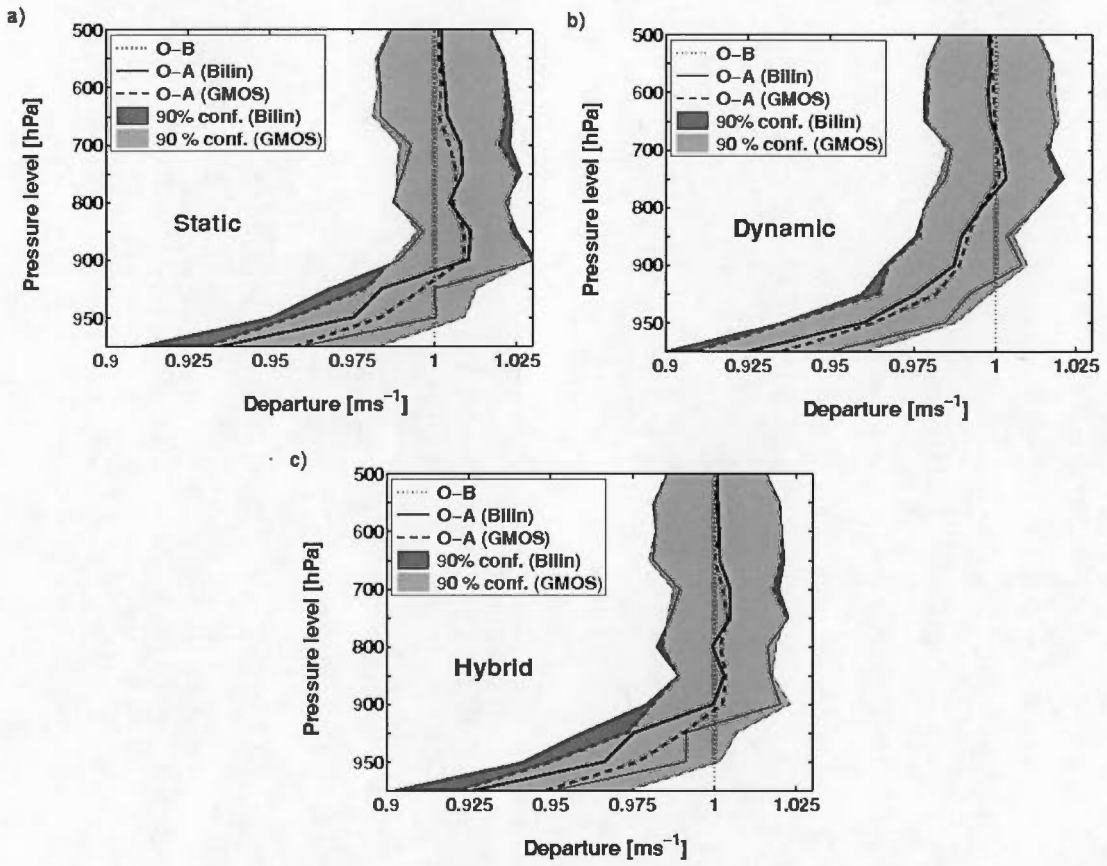

Figure 3.7: Mean observation impact on wind components analysis (evaluated against radiosonde observations) for Bilin and GMOS operators. The dark (light) shaded areas present the 90 % confidence interval for the Bilin (GMOS) operator. From top to bottom, the panels show data assimilation results using static (a), dynamic (b), or hybrid (c) background error statistics.

Given the fact that the vertical propagation of the information in the analysis is provided by the background error covariances, the use of different methods to generate these statistics produces results that differ. Results show that flow-dependent background errors have some ability to reconstruct the vertical structure of the

atmospheric state using near-surface observations whereas static covariances has more limited vertical propagation capabilities. A fair number of studies draw similar conclusions for the assimilation of surface pressure using either EnKF (Whitaker *et al.*, 2004; Anderson *et al.*, 2005; Compo *et al.*, 2006) or 3D-Var (Bengtsson *et al.*, 2004) data assimilation systems. This difference shows that both the validation method using independent observations and the observation operator performances depend on the background error used. Still, in most cases, results show that the vertical correction from the data assimilation experiments concurs with non-assimilated radiosonde observations.

To assess how the observation operators affect the system biases, monthly mean analysis increments are calculated for Bilin and GMOS operators. Figure 3.8 presents systematic zonal wind analysis increments for both Bilin and GMOS operators. To properly understand what these results imply, the systematic analysis increments ($\langle \delta \mathbf{x} \rangle = \langle \mathbf{x}_a - \mathbf{x}_b \rangle$) are decomposed as $\langle \delta \mathbf{x} \rangle = \langle \mathbf{K}(\mathbf{y} - H(\mathbf{x}_b)) \rangle$, where $\mathbf{x}_a$, $\mathbf{x}_b$ represent the analysis and the background state, $\mathbf{y}$ represents the assimilated SYNOP observations, H is the near-surface wind observation operator, $\langle ... \rangle$ stands for the temporal statistical average (monthly mean), and $\mathbf{K}$ represents the gain matrix. Following Eq. (3.4), it is expressed as a function of measurement, representativeness and background biases ($\boldsymbol{\varepsilon}_m$ $\boldsymbol{\varepsilon}_r$ and $\boldsymbol{\varepsilon}_b$ respectively) such as $\langle \delta \mathbf{x} \rangle = \mathbf{K}(\langle \boldsymbol{\varepsilon}_m \rangle + \langle \boldsymbol{\varepsilon}_r \rangle - \langle H(\boldsymbol{\varepsilon}_b) \rangle)$. Systematic analysis increments can be interpreted as a model space representation of the differences between systematic observation and background errors.

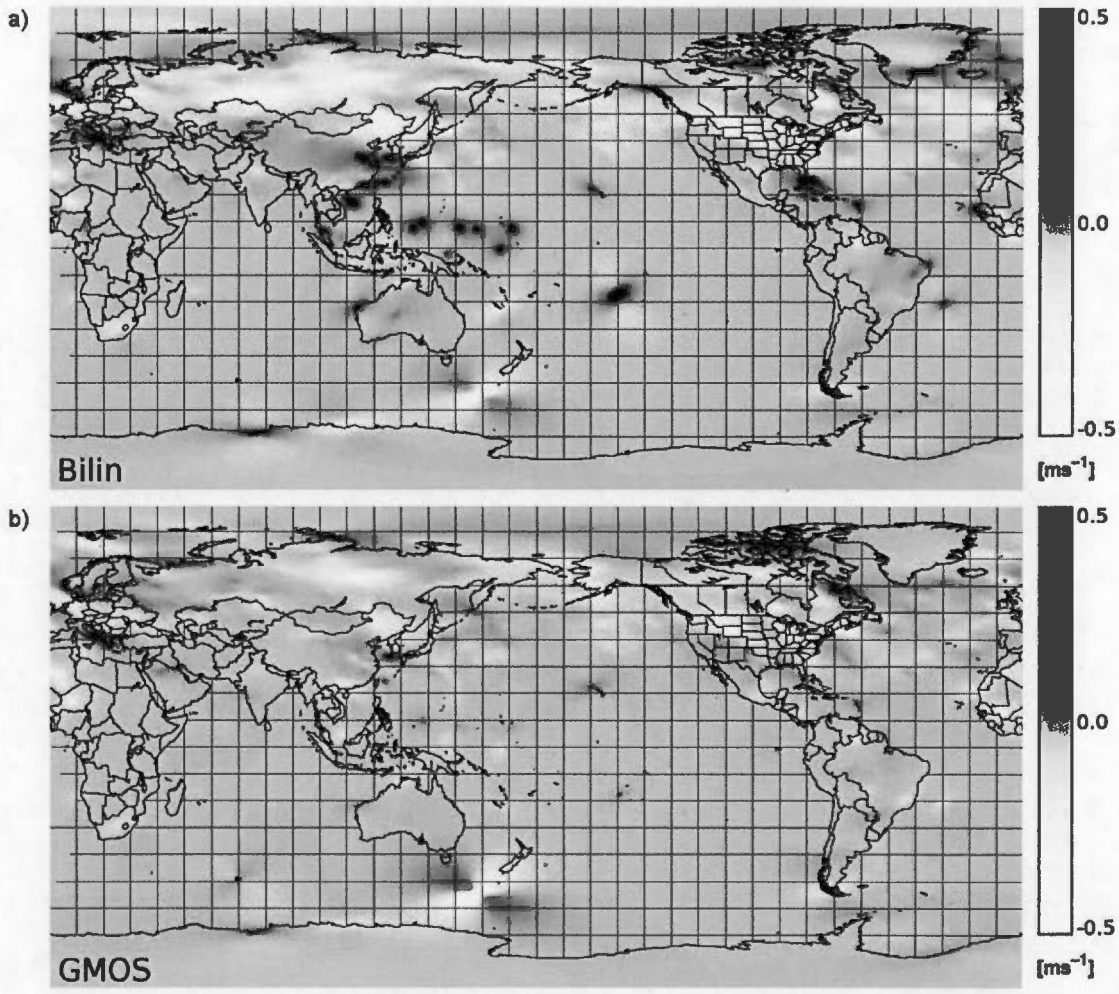

Figure 3.8: Mean zonal wind analysis increments for Bilin (a) and GMOS operators (b).

Figure 3.8 shows that when using Bilin (a), observations located in coastal tropical sites (e.g. in the western Pacific, Caribbean, etc.) cause large systematic analysis increments (dark shaded areas) as noted in Ingleby (2014). However when using the GMOS operators (b), these systematic increments are significantly reduced for most sites. This indicates that the data assimilation scheme (through GMOS) accounts for

systematic differences between local observations and corresponding large scale background states (especially for coastal tropical sites). The bias correction provided by the GMOS operator allows an observation not having the same bias as the NWP model to be more effectively assimilated.

The mean analysis increments STD are also smaller when using GMOS rather than Bilin (0.67 ms^{-1} vs. 0.79 ms^{-1}). This indicates that, although the innovation is smaller, the data assimilation system is still capable of efficiently using the information brought by the observations. The fact that GMOS can produce a good quality analysis while producing relatively smaller increments than Bilin indicates that GMOS is less prone to generate imbalances in the resulting analysis because this operator provides a better correspondence between background fields and observations.

3.8. Conclusions

This study aims at improving the lower tropospheric analyses by making a better use of near-surface wind observations over land. More specifically, the objective of this study is to: understand the observation impact on the analysis, point out aspects of the data assimilation system that need to be improved, and increase the consistency between the analyses and forecasts. Previous works have shown that the representativeness and systematic errors associated with the observation operator are important limitations in this field. To cope with such misrepresentation issues, a geo-statistical observation operator is introduced based on GMOS, including bias and representativeness error corrections. A new observation error correlation evaluation method based on independent observations is also proposed.

An evaluation based on observations innovations shows that the GMOS operator has the ability to correct for representativeness errors and background errors in the observation space. These improvements are related to surface roughness and coastal effects and allow GMOS to better represent the meteorological phenomena onsite. While being consistent with forecast post-processing modules, it is shown that GMOS provides a significant reduction of the innovation bias and STD (~0.1 and ~0.4 ms^{-1}). Results also show that such statistical methods are sensitive to the amount of training data used. In this case, when using at least N months of data, the training algorithm provides robust statistical coefficients to the GMOS architecture using N×N grid-points.

Results from non-cycling data assimilation experiments show that the background state and observations are generally more consistent when using the GMOS operator rather than Bilin. By using the most representative forecast points, the GMOS operator pushes the forecast variability towards the natural on-site variability rather than smoothing the forecast. Results also indicate that, by removing spurious signal from representativeness errors, GMOS has the ability to diminish the observation error variances and prevent analysis biases. When evaluated against radiosonde observations, GMOS and Bilin operators produce relatively good quality near-surface wind analysis (in terms of STD). However, the fact that GMOS accounts for biases, while producing relatively smaller increments than Bilin, indicates that GMOS is less prone to generate strong perturbations in the resulting analysis and allows an observation not on the model attractor to be more effectively assimilated.

The analysis skill in the lower troposphere, measured against non-assimilated radiosonde observations, is increased by 0 to 10 % when assimilating SYNOP wind observations. When assessing the vertical correction against non-assimilated radiosonde observations, it is found that flow-dependent background error

covariances from the EnKF allow for a better vertical propagation of the information in the vertical than static (NMC) error statistics. Also, by its flow-dependency, the error statistics from the EnKF give a more appropriate weight to both the background and observation fields, thus generally enhancing the analyses by ~1 %. Using these error statistics, the vertical correction from data assimilation experiments extends from the surface up to 750 hPa (900 hPa using the hybrid error covariances) and generally concurs with the radiosonde observations.

Results illustrate the importance of the representativeness errors for near-surface wind observations. The value of such observations in a global data assimilation system depends on a proper treatment of these errors. The implementation of a geo-statistical observation operator that corrects for representativeness error by giving more weight to the forecast point that best reproduces the natural on-site variability (as measured by the observations) is a major step in this direction. The analysis improvements provided by SYNOP wind observations show the value of near-surface wind measurements. Such observations appear to be a complementary source of information to provide NWP models with improved initial conditions. The simplified context of this paper (using only SYNOP stations collocated with radiosonde observations) does not allow evaluating the observation impact on forecasts. 545 stations distributed globally are not enough to make significant differences on forecasts. This is the scope of a subsequent paper where ~5000 global SYNOP stations are used in observing system experiments to evaluate the observation impact on forecast. By evaluating the temporal and horizontal propagation of the information, this study will show that improvements can be achieved in observation dense environment (e.g. Europe). Quality control issues related to near-surface wind observations will also be assessed in this subsequent study.

Acknowledgements

The authors acknowledge the contributions of Alain Forcione from Hydro-Québec research institute (IREQ) for his support concerning the access to the tall anemometer tower observations. The authors wish to thank Mark Buehner (Environment Canada), Peter Houtekamer (Environment Canada) and Michel Valin (UQAM's ESCER Centre and Environment Canada) for their suggestions and comments. Also, thanks to Josh Hacker and two anonymous reviewers for providing insightful comments. This project is funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), Hydro-Québec and Environment Canada.

Appendix

When decomposing the errors in mean and perturbation parts ($\boldsymbol{\varepsilon} = \langle \boldsymbol{\varepsilon} \rangle + \boldsymbol{\varepsilon}'$), $Err(\hat{\sigma}^2_{o_S}, \mathrm{r})$ is no longer bias dependent as

$$(\mathbf{a}_R - \langle \mathbf{a}_R \rangle) = \left(\boldsymbol{\varepsilon}'_{o_R} - H_R \boldsymbol{\varepsilon}'_b - H_R \mathbf{K} (\boldsymbol{\varepsilon}'_{o_S} - H_S \boldsymbol{\varepsilon}'_b) \right). \tag{3.15}$$

Combining equations (3.11) and (3.15), it is possible to distinguish the respective contribution of the different error sources to the error metric $Err(\hat{\sigma}^2_{o_S}, \mathrm{r})$. Using the definitions $\mathbf{R}_S = \langle \boldsymbol{\varepsilon}'_{o_S} \boldsymbol{\varepsilon}'^T_{o_S} \rangle$ and $\mathbf{R}_R = \langle \boldsymbol{\varepsilon}'_{o_R} \boldsymbol{\varepsilon}'^T_{o_R} \rangle$ and assuming the background and observation error to be uncorrelated ($\langle \boldsymbol{\varepsilon}'^T_{o_R} \boldsymbol{\varepsilon}'_b \rangle = 0$ and $\langle \boldsymbol{\varepsilon}'^T_{o_S} \boldsymbol{\varepsilon}'_b \rangle = 0$), then

$$Err(\hat{\sigma}^2_{o_S}, \mathrm{r}) = tr \begin{pmatrix} \mathbf{R}_R + \mathbf{H}_R \mathbf{B} \mathbf{H}_R^T + \mathbf{H}_R \mathbf{K} \mathbf{R}_S \mathbf{K}^T \mathbf{H}_R^T + \mathbf{H}_R \mathbf{K} \mathbf{H}_S \mathbf{B} \mathbf{H}_S^T \mathbf{K}^T \mathbf{H}_R^T \\ - \langle \boldsymbol{\varepsilon}'_{o_R} \boldsymbol{\varepsilon}'^T_{o_S} \rangle \mathbf{K}^T \mathbf{H}_R^T - \mathbf{H}_R \mathbf{K} \langle \boldsymbol{\varepsilon}'_{o_S} \boldsymbol{\varepsilon}'^T_{o_R} \rangle - \mathbf{H}_R \mathbf{B} \mathbf{H}_S^T \mathbf{K}^T \mathbf{H}_R^T - \mathbf{H}_R \mathbf{K} \mathbf{H}_S \mathbf{B} \mathbf{H}_R^T \end{pmatrix}. \tag{3.16}$$

It is assumed that the observation error for radiosonde data ($\boldsymbol{\varepsilon}_{o_R}$) and surface data ($\boldsymbol{\varepsilon}_{o_S}$) are impacted by a similar representativeness error ($\boldsymbol{\varepsilon}_r$) associated with sub-grid scale variability near the ABL. As Liu and Rabier (2002) showed that it is important to account for the correlation of representativeness errors from nearby observations, $\langle \boldsymbol{\varepsilon}'_{o_R} \boldsymbol{\varepsilon}'^T_{o_S} \rangle$ is considered having a non-null value. Thus, as measurement errors ($\boldsymbol{\varepsilon}_m$) are uncorrelated for surface and radiosondes observations ($\boldsymbol{\varepsilon}_{o_S}$ and $\boldsymbol{\varepsilon}_{o_R}$ respectively), $\langle \boldsymbol{\varepsilon}'_{o_R} \boldsymbol{\varepsilon}'^T_{o_S} \rangle$ becomes a function of the representativeness errors ($\boldsymbol{\varepsilon}_r$) as

$$\left\langle \boldsymbol{\varepsilon}'_{o_R} \boldsymbol{\varepsilon}'^{T}_{o_S} \right\rangle = \left\langle \left(\boldsymbol{\varepsilon}'_{m_R} + \boldsymbol{\varepsilon}'_{r_R} \right) \left(\boldsymbol{\varepsilon}'_{m_S} + \boldsymbol{\varepsilon}'_{r_S} \right)^T \right\rangle = \left\langle \boldsymbol{\varepsilon}'_{r_R} \boldsymbol{\varepsilon}'^{T}_{r_S} \right\rangle . \tag{3.17}$$

Assuming that a bilinear interpolation has no influence on the background error (e.g. when assimilating radiosonde observations), general background error variances (σ_b^2) are imposed such as

$$tr(\mathbf{H}_R \mathbf{B} \mathbf{H}_R^T) = \sum \sigma_b^2 . \tag{3.18}$$

This hypothesis appears reasonable as the wind components are directly related to model variables. However, considering that the background error variance is influenced by the amplitude correction from the statistical observation operators used to assimilate SYNOP observations (namely GMOS), the background error variance in observation space becomes

$$tr(\mathbf{H}_S \mathbf{B} \mathbf{H}_S^T) = \sum \alpha^2 \sigma_b^2 . \tag{3.19}$$

where α is a coefficient introduced to account for the influence of the observation operator on $\mathbf{HBH}^T$. From Section 3.4.3, α = 0.76 (1.0) when using the GMOS (Bilin) operator. Given the proximity of the surface and radiosonde data location, the vertical background error correlation (r_v) is approximately equal to 1 in the 4DEnVar covariance matrices and $tr(\mathbf{H}_R \mathbf{B} \mathbf{H}_S^T) = \sum r_v \sigma_{b_R} \alpha \sigma_{b_S} \approx \sum r_v \alpha \sigma_b^2 \approx \sum \alpha \sigma_b^2$. Hence, the expression $\mathbf{H}_R \mathbf{K}$ becomes

$$tr(\mathbf{H}_R \mathbf{K}) = tr\left(\mathbf{H}_R \mathbf{B} \mathbf{H}_S^T \left(\hat{\mathbf{R}}_S + \mathbf{H}_S \mathbf{B} \mathbf{H}_S^T \right)^{-1} \right) = \sum \alpha \sigma_b^2 \left(\hat{\sigma}_{o_S}^2 + \alpha^2 \sigma_b^2 \right)^{-1} . \tag{3.20}$$

Assuming that all results evaluated at the radiosonde location can be bundled together as samples of the same situation, (3.16) can be simplified to

$$Err\left(\hat{\sigma}_{o_S}^2, \mathrm{r}_\mathrm{r}\right) = \sum \left(\begin{array}{l} \sigma_{o_R}^2 + \left(\left(\dfrac{\alpha^2 \sigma_b^2}{\hat{\sigma}_{o_S}^2 + \alpha^2 \sigma_b^2} \right) - 1 \right)^2 \sigma_b^2 + \left(\dfrac{\alpha \sigma_b^2}{\hat{\sigma}_{o_S}^2 + \alpha^2 \sigma_b^2} \right)^2 \sigma_{o_S}^2 \\ -2 \left(\dfrac{\alpha \sigma_b^2}{\hat{\sigma}_{o_S}^2 + \alpha^2 \sigma_b^2} \right) \left\langle \boldsymbol{\varepsilon}'_{r_S} \boldsymbol{\varepsilon}'^{T}_{r_R} \right\rangle \end{array} \right) \tag{3.21}$$

The representativeness error covariance can then be rewritten as a function of the representativeness correlation parameter (r_r) and both representativeness error variances. Applying $r_r = \left\langle \boldsymbol{\varepsilon}'_{r_S} \boldsymbol{\varepsilon}'^{T}_{r_R} \right\rangle / \left(\sigma_{r_S} \sigma_{r_R} \right)$, the error metric becomes

$$Err\left(\hat{\sigma}_{o_S}^2, \mathrm{r}_\mathrm{r}\right) = \sum \left(\begin{array}{l} \sigma_{o_R}^2 + \left(\left(\dfrac{\alpha^2 \sigma_b^2}{\hat{\sigma}_{o_S}^2 + \alpha^2 \sigma_b^2} \right) - 1 \right)^2 \sigma_b^2 + \left(\dfrac{\alpha \sigma_b^2}{\hat{\sigma}_{o_S}^2 + \alpha^2 \sigma_b^2} \right)^2 \sigma_{o_S}^2 \\ -2 \left(\dfrac{\alpha \sigma_b^2}{\hat{\sigma}_{o_S}^2 + \alpha^2 \sigma_b^2} \right) r_r \sigma_{r_S} \sigma_{r_R} \end{array} \right) \tag{3.22}$$

.

References

Anderson JL, Wyman B, Zhang S, Hoar T. 2005. Assimilation of surface pressure observations using an ensemble filter in an idealized global atmospheric prediction system. *J. Atmos. Sci.*, **62** : 2925–2938.

Auligné T, McNally AP. 2007. Interaction between bias correction and quality control. *Quart. J.R. Meteor. Soc.*, **133** : 643–653.

Auligné T, McNally AP, Dee DP. 2007. Adaptive bias correction for satellite data in a numerical weather prediction system. *Quart. J.R. Meteor. Soc.*, **133** : 631–642.

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.

Bengtsson L, Hodges KI, Hagemann S. 2004. Sensitivity of the ERA-40 reanalysis to the observing system: Determination of the global atmospheric circulation from reduced observations. *Tellus A*, **56** : 456–471.

Benjamin SG, Moninger WR, Sahm SR, Smith TL. 2007. Mesonet wind quality monitoring allowing assimilation in the RUC and other NCEP models. Preprints, *22nd Conf. on Weather Analysis and Forecasting/18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Meteor. Soc., P1.33.

Benjamin SG, Jamison BD, Moninger WR, Sahm SR, Schwartz B, Schlatter TW. 2010. Relative short-range forecast impact from aircraft, profiler, radiosonde, VAD, GPS-PW, METAR, and mesonet observations via the RUC hourly assimilation cycle. *Mon. Wea. Rev.*, **138** : 1319–1343.

Buehner M. 2005. Ensemble-derived stationary and flow-dependent background error covariances: evaluation in a quasi-operational setting for NWP. *Quart. J.R. Meteor. Soc.*, **131** : 1013–1044.

Buehner M, Morneau J, Charette C. 2013. Four-dimensional ensemble-variational data assimilation for global deterministic weather prediction. *Nonlin. Processes Geophys.*, **20** : 669–682.

Charron M, Polavarapu S, Buehner M, Vaillancourt PA, Charette C, Roch M, Morneau J, Garand L, Aparicio JM, MacPherson S, Pellerin S, St-James J, Heilliette S. 2012. The stratospheric extension of the canadian global deterministic medium-range weather forecasting system and its impact on tropospheric forecasts. *Mon. Wea. Rev.*, **140** : 1924 –1944.

Côté J, Gravel S, Méthot A, Patoine A, Roch M, Staniforth A. 1998a. The operational CMC/MRB global environmental multiscale (GEM) model: Part I – design considerations and formulation. *Mon. Wea. Rev.*, **126** : 1373–1395.

Dee DP, Uppala S. 2009. Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quart. J.R. Meteor. Soc.*, **135** : 1830–1841.

Deng X, Stull R. 2005. A mesoscale analysis method for surface potential temperature in mountainous and coastal terrain. *Mon. Wea. Rev.*, **133** : 389 – 408.

Desroziers G, Berre L, Chapnik B, Poli P. 2005. Diagnosis of observation, background and analysis-error statistics in observation space. *Quart. J.R. Meteor. Soc.*, **131** : 3385–3396.

Dong J, Xue M, Droegemeier K. 2010. The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteorology and Atmospheric Physics*, **112** : 41–61.

Gauthier P, Tanguay M, Laroche S, Pellerin S, Morneau J. 2007. Extension of 3DVAR to 4DVAR: implementation of 4DVAR at the Meteorological Service of Canada. *Mon. Wea. Rev.*, **135** : 2339–2354.

Hacker JP, Snyder C. 2005. Ensemble Kalman filter assimilation of fixed screen-height observations in a parameterized PBL. *Mon. Wea. Rev.*, **133** : 3260–3275.

Henne S, Brunner D, Folini D, Solberg S, Klausen J, Buchmann B. 2010. Assessment of parameters describing representativeness of air quality in-situ measurement sites. Atmospheric Chemistry and Physics, **10** : 3561–3581.

Houtekamer PL, Deng X, Mitchell HL, Baek SJ, Gagnon N. 2014. Higher resolution in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **142** : 1143–1162.

Huang X, Gao F, Jacobs N, Wang H. 2013. Assimilation of wind speed and direction observations: a new formulation and results from idealised experiments. *Tellus A*, **65**, 19936, http://dx.doi.org/10.3402/tellusa.v65i0.19936.

Ingleby B. 2014. Global assimilation of air temperature, humidity, wind and pressure from surface stations. *Quart. J.R. Meteor. Soc.*, **141** : 504–517.

Janjic T, Cohn SE. 2006. Treatment of observation error due to unresolved scales in atmospheric data assimilation. *Mon. Wea. Rev.*, **134** : 2900–2915.

Koohkan MR, Bocquet M. 2012. Accounting for representativeness errors in the inversion of atmospheric constituent emissions: application to the retrieval of regional carbon monoxide fluxes. Tellus B, 64 : doi:10.3402/tellusb.v64i0.19047.

Lavaysse C, Carrera M, Bélair S, Gagnon N, Frenette R, Charron M, Yau MK. 2013. Impact of Surface Parameter Uncertainties within the Canadian Regional Ensemble Prediction System. *Mon. Wea. Rev.*, **141** : 1506–1526.

Lea DJ, Drecourt JP, Haines K, Martin MJ. 2008. Ocean altimeter assimilation with observational- and model-bias correction. *Quart. J.R. Meteor. Soc.*, **134** : 1761–1774.

Liu ZQ, Rabier F. 2002. The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study. *Quart. J.R. Meteor. Soc.*, **128** : 1367–1386.

Nash J, Oakley T, Vömel H, Wei LI. 2011. WMO Intercomparison of high quality radiosonde systems, Yangjiang, China, 12 July - 3 August 2010. Instruments and Observing Methods report No. 107, World Meteorological Organization, 248 p. Available from https://www.wmo.int/pages/prog/www/IMOP/publications-IOM-series.html.

Parrish DF, Derber JC. 1992. The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120** : 1747–1763.

Rostkier-Edelstein D, Hacker JP. 2010. The roles of surface-observation ensemble assimilation and model complexity for nowcasting of PBL profiles: a factor separation analysis. *Weather and Forecasting*, **25** : 1670–1690.

Whitaker JS, Compo GP, Wei X, Hamill TM. 2004. Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.*, **132** : 1190–1200.

Zack J, Natenberg EJ, Young S, Knowe GV, Waight K, Manobianco J, Kamath C. 2010. Application of ensemble sensitivity analysis to observation targeting for short term wind speed forecasting in the Washington - Oregon region`. Technical Report, Lawrence Livermore National Laboratory. 65pp. Available from http://computation.llnl.gov/casc/StarSapphire/pubs/LLNL-TR-458086.pdf.

Zack J, Natenberg EJ, Knowe GV, Waight K, Manobianco J, Hanley D, Kamath C. 2011. Observing system simulation experiments (OSSE) for the Mid-Columbia basin. Technical Report, Lawrence Livermore National Laboratory. 17pp. Available from https://e-reports-ext.llnl.gov/pdf/515298.pdf.

CHAPTER IV.

NEAR-SURFACE WIND DATA ASSIMILATION: TEMPORAL PROPAGATION OF THE ANALYSIS INCREMENT AND MULTIVARIATE IMPACT ON FORECASTS

This chapter presents a paper in preparation:

Bédard J, Laroche S, Gauthier P., 2015b. Near-surface wind data assimilation: temporal propagation of the analysis increment and multivariate impact on forecasts. In preparation.

This paper focuses on the multivariate observation impact on forecasts and evaluates the temporal propagation of the information in the forecast system. The influence of different components of the assimilation and prediction systems (namely the background error statistics and boundary layer parameterization) is also examined to point out aspects that need to be improved to make a better use of near-surface wind observations. The full observation impact from 4942 surface synoptic stations is evaluated through observing system experiments with the operational system to show the value of near-surface wind measurements for short-term tropospheric forecasts.

NEAR-SURFACE WIND DATA ASSIMILATION: TEMPORAL PROPAGATION OF THE ANALYSIS INCREMENT AND MULTIVARIATE IMPACT ON FORECASTS

Joël Bédard [(1),a], Stéphane Laroche[b], and Pierre Gauthier[a]

[a] ESCER Centre, *Department of Earth and Atmospheric Sciences*
Université du Québec à Montréal (UQAM), Montréal (Québec), Canada

[b] *Data Assimilation and Satellite Meteorology Section*
Environment Canada, Dorval (Québec), Canada

(revised version)

---

[1] Corresponding author: Joël Bédard, ESCER Centre
Department of Earth and Atmospheric Science
Université du Québec à Montréal
P.O. Box 8888, Downtown Station
Montréal (Québec) CANADA H3C 3P8
Email: bedard.joel@gmail.com

Abstract

This study revisits the assimilation of near-surface wind observations over land to assess its multivariate impact as well as the temporal evolution of the analysis increments within the model. Aiming to improve nowcasting and short-term tropospheric wind forecasts, a geo-statistical observation operator (GMOS) is also introduced and compared with the conventional bilinear interpolation scheme (Bilin). Without cycling, results from assimilating only wind data from 4942 SYNOP stations show that very short-term near-surface wind forecasts are improved when using GMOS. The local impact decreases over time and is only significant for a forecast lead time of 6 h or shorter. Initial model tendencies indicate that, when using flow-dependent error statistics, most of the analysis increment is propagated as it modifies both wind and mass fields in a coherent way through multivariate covariances. When using static error covariances the mass field is not significantly altered and the boundary layer parameterization and orographic blocking schemes damp the poorly balanced increment locally. The use of proper background error statistics is crucial to produce sustainable impacts on the atmospheric forecasts. Also, the increment produced when using GMOS is in better agreement with the model stated and the information persists longer in the system. The use of near-surface winds (with GMOS) in a more realistic context where the mass field is constrained with surface pressure observations provides upper-air improvements in terms of wind speed bias. Results from cycling observing system experiments (assimilating near-surface wind data and all the other types of observations used operationally) are generally neutral. Encouragingly, for specific cases, forecasts and analyses from GMOS (Bilin) were found to be more (less) coherent than those from the control experiment (in which no wind observations are assimilated over land) as the information from the observation is (is not) propagated in time.

Keywords: Initial systematic tendencies, Site dependent observation error statistics, Background error covariances, Atmospheric boundary layer parameterization, Upper-air verification, Evaluation against own analyses.

4.1. Introduction

In a continuous effort to improve short term wind forecasts for different applications (e.g. wind energy, airport operations, road and construction site security, recreational activities, etc.), recent studies were made trying to assimilate near-surface wind observations over land in relatively simplified context (Hacker and Snyder, 2005; Hacker and Rostkier-Edelstein, 2007; Hacker *et al.*, 2007; Dong *et al.*, 2010; Rostkier-Edelstein and Hacker, 2010; Zack *et al.*, 2010; Zack *et al.*, 2011; Ancell *et al.*, 2015). Using a single column model, Rostkier-Edelstein and Hacker (2010) showed that improvements in the assimilation of surface observations can meaningfully improve nowcasting capabilities. Similarly, by performing observation targeting experiments to assess forecast sensitivity to wind observations, Zack *et al.* (2010) showed that near-surface wind forecasts (up to 3 h lead time) are sensitive to local low level wind initial conditions. Zack *et al.* (2011) then showed that near-surface local wind nowcasting capabilities can be improved by assimilating synthetic observations from tall anemometer towers. By assessing the observation impact on wind forecasts over Texas and Oklahoma using different assimilation systems, Ancell *et al.* (2015) showed that assimilating real near-surface wind observations improved wind nowcasting capabilities (up to 6 h). Improvements appeared to be more significant when using an assimilation system based on an EnKF rather than a 3D variational system (3DVAR). As surface observations have smaller correlation with the flow aloft compared to integrated variables such as surface pressure, their impact on analyses varies depending on the Atmospheric Boundary Layer (ABL) coupling with the flow aloft (e.g., atmospheric stability) and this may limit the use of stationary background error covariances which are based on globally homogeneous correlations and simple balance relationships. When using an Ensemble Kalman Filter (EnKF) to sample the flow dependent (in time and space) and multivariate background error covariances, the assimilation of near-surface observations improved

temperature, humidity and winds profiles and the information propagates through the cycling of the data assimilation system (Hacker and Rostkier-Edelstein, 2007). Still, results showed to be sensitive to the way flow dependent error statistics are handled by the assimilation system (Hacker *et al.*, 2007).

While results indicate that near-surface observations (including winds) are effective at constraining the state of the ABL (Hacker and Snyder, 2005) and while synthetic near-surface wind observations can provide significant forecast improvements for areas with poor low-level data coverage (Dong *et al.*, 2010), assimilating near-surface wind observations over land in a realistic context remains a challenging task because such observations tend to sample fine scales which are not explicitly resolved by Numerical Weather Prediction (NWP) models. Although many near-surface wind observations (over land) are available from the global observing system, they had not been used in most data assimilation systems until recently due to the discrepancy between the characteristics of the measured and forecasted variables. The model misrepresentation of surface characteristics generates so-called representativeness errors. Ingleby (2014) showed that biases and representativeness errors limit the global influence of near-surface wind observations on operational forecasting systems. Winds from small islands, sub-grid scale headlands and over land throughout the tropics are still excluded from the UK Met Office data assimilation system, while other operational systems simply blacklist all wind observations from land stations (e.g., Environment Canada, hereafter EC). Similarly, as presented by Benjamin *et al.* (2007, 2010), the NCEP Rapid Update Cycle (RUC) uses strict quality control checks to prevent degrading the near-surface wind analyses due to representativeness errors.

Bédard *et al.* (2015a) addressed representativeness and systematic error issues by developing a geo-statistical observation operator based on a multiple grid-point

approach called GMOS (Geophysical Model Output Statistics: Bédard *et al.*, 2013). They examined the impact of near-surface wind observations (over land) on analyses using 454 surface synoptic (SYNOP) stations collocated with radiosonde observations. The GMOS operator has been tested and compared with the results obtained from a conventional bilinear interpolation scheme used in most forward operators for *in situ* observations (hereafter, Bilin). As local representativeness errors associated with local sub-grid scale effects are accounted for, collocated observation error correlations are significantly reduced when using GMOS operator with respect to Bilin. Also, the bias correction provided by GMOS allows the observations to be more effectively assimilated. By attributing higher weights to the most representative grid-points and by taking into account the natural on-site variability, GMOS better represents meteorological phenomena locally. By extension, if the surrounding grid points are only weakly representative, GMOS will give a reduced weight to the observation and its impact on analyses will be reduced. By making background states and observations generally more consistent, GMOS produces relatively smaller innovations and analysis increments than bilinear interpolation, and it is less prone to generating strong perturbations in the resulting analysis. Yet, considering the simplified context of this previous study, the impact from such perturbations on forecasts and the value of near-surface wind observations in a forecasting system could not be evaluated. This is the scope of the present paper where near-surface wind observations from 4942 global SYNOP stations are assimilated in the ensemble variational data assimilation system (4DEnVar) developed at EC (Buehner *et al.*, 2013; 2015).

The objective of this work is to understand the multivariate impact and the temporal propagation of the local analysis increment from assimilating near-surface wind observations over land using the GMOS or Bilin operator in EC's global NWP system. Simplified data assimilation experiments over a one-month period (February

2011) were first carried out in a non-cycling mode, assimilating only wind data from surface stations, to assess the influence of the background error statistics and the different components of the NWP model in a controlled environment. The systematic model initial tendencies are used to assess the influence of the boundary layer parameterizations and the background error covariances on the temporal evolution of the analysis increment. Quality control issues related to near-surface wind observations are also examined. The second part of this paper focuses on results from realistic observing system experiments (OSE) to assess the observation impact on short-term tropospheric forecasts in the context of a data assimilation cycle.

Section 4.2 introduces the characteristics of the assimilation system, Section 4.3 presents the simplified experiments and associated results, and the results from the realistic OSEs are presented in Section 4.4. A summary and conclusions are presented along with an outlook in Section 4.5.

4.2. The data assimilation system

The operational 4DEnVar developed at EC is a state-of-the-art variational data assimilation system (Buehner *et al.*, 2013; 2015) that uses 4D ensemble-based flow dependent background error statistics to produce 4D analysis increments. The background error statistics comprise a stationary homogeneous (static) and a flow dependent (dynamic) component to produce spatiotemporal analysis increments. The static covariance component is estimated using the NMC method (Parrish and Derber, 1992; Gauthier *et al.*, 1998) and the flow dependent component uses 256 ensemble members from EC's operational Ensemble Kalman Filter (EnKF) (Houtekamer *et al.*, 2014). In the operational configuration, the analysis increment is computed on a

800×400 grid (~50km grid spacing). The two background error components are blended equally below 40 hPa and gradually weighted towards the static error statistics above. However, the system can be run using only the static or the EnKF background error components. This feature is convenient to test the impact of the different background-error components on the propagation of the information.

4.2.1. Geo-statistical observation operator

By combining a multiple grid-point approach with statistical error corrections, GMOS takes advantage of the correlation between resolved scales and unresolved scales to correct the stationary and isotropic components of the systematic and representativeness error associated with local geographical characteristics (e.g., surface roughness or coastal effects). Similar to Deng and Stull (2005), the idea behind this operator is that the nearest grid points may not properly represent an observing station, especially if the station is located on complex terrain or coastal sites. On the other hand, amongst the surrounding grid points, there are generally one or several grid-points that are more representative of the observing site. Thus, GMOS uses a set of geo-statistical weights relating the most representative NWP grid-points to the observation site.

The GMOS multi-point linear regression (H_{GMOS}) is formulated as

$$H_{GMOS}(x) = \sum_i (A_i x_i) + C \quad , \tag{4.1}$$

where x is either the latitudinal or longitudinal modeled wind component, A_i, the amplitude coefficients, and C the systematic error correction coefficient. The subscript i is the index of the 4 closest grid-points to the observing site in a 2×2 square pattern. The systematic error coefficient corrects for biases while the amplitude coefficient adjusts the forecast variability to best fit observations. More details can be found in Bédard *et al.* (2015a).

The observations from global surface synoptic (SYNOP) stations are used every 3 hours in this study. To compute the statistical coefficients (A_i and C), these observations are compared with corresponding short-range (0-12 hours) wind forecasts produced twice a day by EC's global deterministic prediction system (Global Environmental Multiscale model at 0.35°×0.23° latitude / longitude resolution (~25 km): Côté *et al.*, 1998a; Charron *et al.*, 2012, Zadra *et al.*, 2014). While the time period used to train the observation operators is different than for the assimilation experiments, the forecasts are generated using the same model configuration.

4.2.2. Observation quality control

Some of the stations do not provide data for some periods of time, while many others only report irregularly. The quality control of the observations systematically applies a gross error check and a background check, while a variational quality control is present at the assimilation stage. These tests are also performed on the training dataset as GMOS relies on the availability of a relatively large amount of good quality data to train its statistical coefficients (Bédard *et al.*, 2013, 2015a).

A GMOS operator using a 2×2 grid-point stencil requires a minimum of 2 months of training data. From the 8 months training dataset available (November 2012 to June 2013), 6 months are used to compute the statistical coefficients and 2 months are kept to evaluate the statistical robustness of the GMOS operator at each site. Sites where GMOS does not perform well with the verifying dataset are considered to have too few observations to generate statistically robust coefficients. These stations, along with those not benefiting from a minimum of 2 months of data are blacklisted.

The background check is sensitive to the observation operator, and the set of statistical coefficients computed can be slightly biased by the operator used. The statistical coefficients are therefore obtained by an iterative process. The GMOS coefficients are used within the background check to filter out erroneous observations and then, the coefficients are recomputed. The coefficients rapidly converge towards stable values for most stations (e.g. after 2 or 3 iterations). A maximum of 10 iterations is allowed to accommodate stations presenting fewer observations. Still, a few stations do not converge because the number of good observations is too small. In such case, a slight change in the coefficients may alter significantly the relative quantity of good data (from the few available). These stations are also blacklisted as they cannot provide robust coefficients to the geo-statistical operator. Although each experiment has its own quality control, the Bilin experiments use the same observing stations as the GMOS experiment to ensure that all experiments benefit from the same amount of observations to perform a fair evaluation. The resulting 4942 SYNOP stations are shown in Figure 4.1.

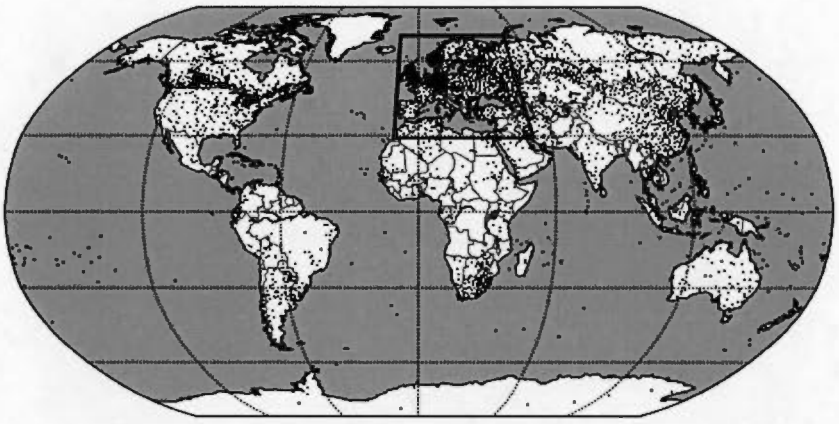

Figure 4.1: Spatial distribution of the 4942 SYNOP sites considered in this study (black dots). The rectangle refers to the area where the upper-air evaluation is performed.

As opposed to other studies where near-surface wind observation can be blacklisted based on a difference between modeled and observed station height, all stations that passed the quality control are used and thus, all surface wind observations having wind speed above 1 ms^{-1} are assimilated (wind vane have a poor accuracy at wind speed below 1 ms^{-1}: Bédard *et al.*, 2015a), provided that they pass the background check.

4.2.3. Observation error statistics

GMOS reduces the representativeness part of the observation error, which also affects other collocated observations such as low level radiosonde data (Bédard *et al.*, 2015a). In addition, GMOS removes the observation error bias. As representativeness

errors are related to the characteristics of the observation site, the error statistics computed from local innovation variances are site dependent. The observation error variances (σ_o^2) are computed (for each sites) using the optimal background error variances (for near-surface winds over land) to be used in the assimilation system (σ_b^2: as computed in Bédard *et al.*, 2015a) following

$$\left\langle \left(\left(y - H(\mathbf{x}_b) \right) - \left\langle y - H(\mathbf{x}_b) \right\rangle \right)^2 \right\rangle = \sigma_o^2 + \sigma_b^2 \quad , \tag{4.2}$$

where observations ($\mathbf{y}$) are compared with the background state ($\mathbf{x}_b$) using an observation operator H. Here, $\langle ... \rangle$ stands for the statistical average. Figure 4.2 shows the frequency distribution (0.05 ms^{-1} bin intervals) of the site dependent error statistics (using GMOS) for the 4942 sites considered in this study.

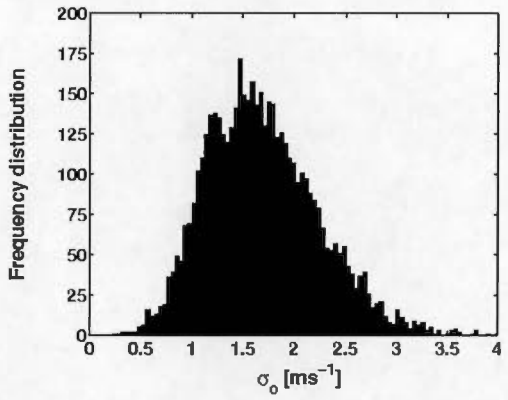

Figure 4.2: Frequency distribution of the near-surface wind observation error statistics using the GMOS operator for the 4942 SYNOP stations considered in this study.

When using GMOS rather than Bilin, Bédard *et al.* (2015a) showed that it is possible to diminish the global observation error statistics (σ_o) prescribed in the assimilation system from 1.98 ms^{-1} to 1.79 ms^{-1} on average. For GMOS cases, the site dependent error statistics described above are employed to be consistent with the fact that GMOS considers local representativeness errors. Bilin cases use a homogeneous value of 1.98 ms^{-1} to be consistent with results from previous studies (Benjamin *et al.*, 2007; 2010; Ingleby, 2014).

4.2.4. Evaluation dataset

Observations and analyses are used to evaluate the forecast departure bias and standard deviation (STD) and the impact of near-surface wind observations. SYNOP stations are used to evaluate local near-surface wind forecasts. Radiosonde profiles and analyses are used to diagnose both the mass (humidity, temperature and geopotential height) and wind fields at various vertical levels. Throughout this study, only scores for wind speed and geopotential height are presented because: 1) bias scores for temperature and humidity are neutral, and 2) the STD for temperature and humidity are similar to those for geopotential height. While the near-surface wind evaluation is performed globally, upper-air diagnostics are only shown over Europe and neighboring countries where the SYNOP station density and their impact are the most significant (see Figure 4.1). Figure 4.3 shows the 1487 SYNOP and 124 radiosonde stations located in this area, which covers latitudes ranging from 30°N to 70°N and longitudes ranging from 10°W to 50°E.

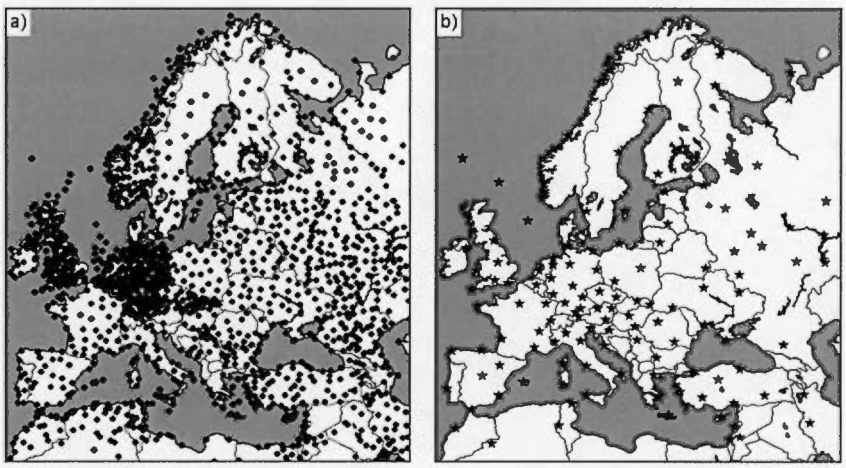

Figure 4.3: Spatial distribution of the 1487 SYNOP sites (a: black dots) located in the selected domain. This area is densely observed and also includes 124 radiosonde stations used for upper-air evaluation purposes (b: black stars).

4.3. Simplified assimilation experiments

The temporal propagation of the multivariate impact of the analyses on the forecasts are evaluated through data assimilation experiments performed in a simplified and controlled environment in which only wind data from surface stations are assimilated. These non-cycling experiments use the same background fields provided by a full assimilation cycle that did not assimilate the near-surface wind data (for an objective comparison). The different experiments are described in Table 4.1.

Table 4.1: Configuration of the 7 simplified data assimilation performed. Each experiment is listed along with its own combination of near-surface wind observation operator, background and observation error statistics prescribed to the data assimilation system as well as assimilated observations. It is also specified if the experiments are cycled or not.

Experiment name	**Cycling**	**Sfc. wind observation operator (over land)**	**Sfc. wind observation errors (over land)**	**Background errors**	**Observations considered**
Control experiment					
CNTRL	No	N/A	N/A	N/A	None
Bilin experiments					
Bilin$_{Hybrid}$	No	Bilin	Homogeneous	Hybrid	SYNOP: winds
GMOS experiments					
GMOS$_{Hybrid}$	No	GMOS	Site dependent	Hybrid	SYNOP: winds
GMOS$_{EnKF}$	No	GMOS	Site dependent	EnKF	SYNOP: winds
GMOS$_{NMC}$	No	GMOS	Site dependent	NMC	SYNOP: winds
Surface pressure experiments					
P0$_{EnKF}$	No	N/A	N/A	EnKF	SYNOP: pressure
P0 + GMOS$_{EnKF}$	No	GMOS	Site dependent	EnKF	SYNOP: pressure and winds

Two observation operators (Bilin and GMOS) were used to assimilate the near-surface wind observations. The experiments are evaluated against the control run (CNTRL) which is a simple model integration from the background field. Most experiments assimilated only wind observations from SYNOP stations over land. These globally distributed stations report 10 m wind speed and direction every 3 h. The hybrid formulation of 4DEnVar was used first. Then, to test the impact of multivariate covariances from the different background error components, other experiments were performed using either the NMC or the EnKF covariances. Additional experiments (P0 and P0 + $GMOS_{EnKF}$) were performed to compare the impact of assimilating near-surface wind together with surface pressure observations. Each experiment was conducted over a period of one-month (February 2011). The analyses and subsequent 48 h forecasts were produced twice daily at 0000 and 1200 UTC. The resulting forecasts are verified against the surface observations themselves and radiosonde profiles. Also, EC's digital filter was turned off to avoid filtering the analysis increment.

4.3.1. Evaluation against near-surface wind observations

GMOS is typically used for near-surface wind forecast post-processing (e.g., wind power forecasting). To show the improvement in the post-processing brought by GMOS, Figure 4.4 presents results from the CNTRL experiment when post-processed using both Bilin and GMOS (square and circle symbols respectively). The solid and dash lines present the forecast results for the $Bilin_{Hybrid}$ and $GMOS_{Hybrid}$ experiments, respectively. It shows that GMOS produces significantly lower wind departure STD compared to Bilin and is better suited to compare model states with near-surface wind observations from the same stations as those used for assimilation purposes. For this

reason and for consistency with the forecast post-processing tools, GMOS is used to evaluate the surface wind forecasts against SYNOP observations.

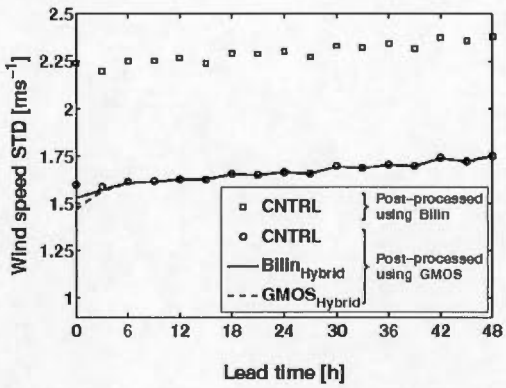

Figure 4.4: Wind speed departure STD as a function of forecast lead time for different experiments (CNTRL, Bilin and GMOS). Note that the CRTL experiment is post-processed using both Bilin and GMOS in order to highlight the impact of the operator on post-processing.

Figure 4.4 clearly shows that the assimilation of near-surface wind observations is only beneficial for forecasts during the first 6 h, which is consistent with the results of Ancell *et al.* (2015). While the Bilin experiment improves the near-surface wind speed by 0.07 ms^{-1} (at 0 h) over the CNTRL experiment, GMOS improvements are of 0.13 ms^{-1}.

To better understand why the forecast skill from the assimilation of near-surface wind observation vanishes so quickly, the mean evolution of the forecast differences between CNTRL and the experiments ($\|\delta\mathbf{V}(t)\|$) where $\delta\mathbf{V} = \begin{bmatrix}\delta u & \delta v\end{bmatrix}^T$, is shown in Figure 4.5 using different observation operators (Bilin and GMOS) and background error statistics (Hybrid, NMC and EnKF) for the month of February 2011.

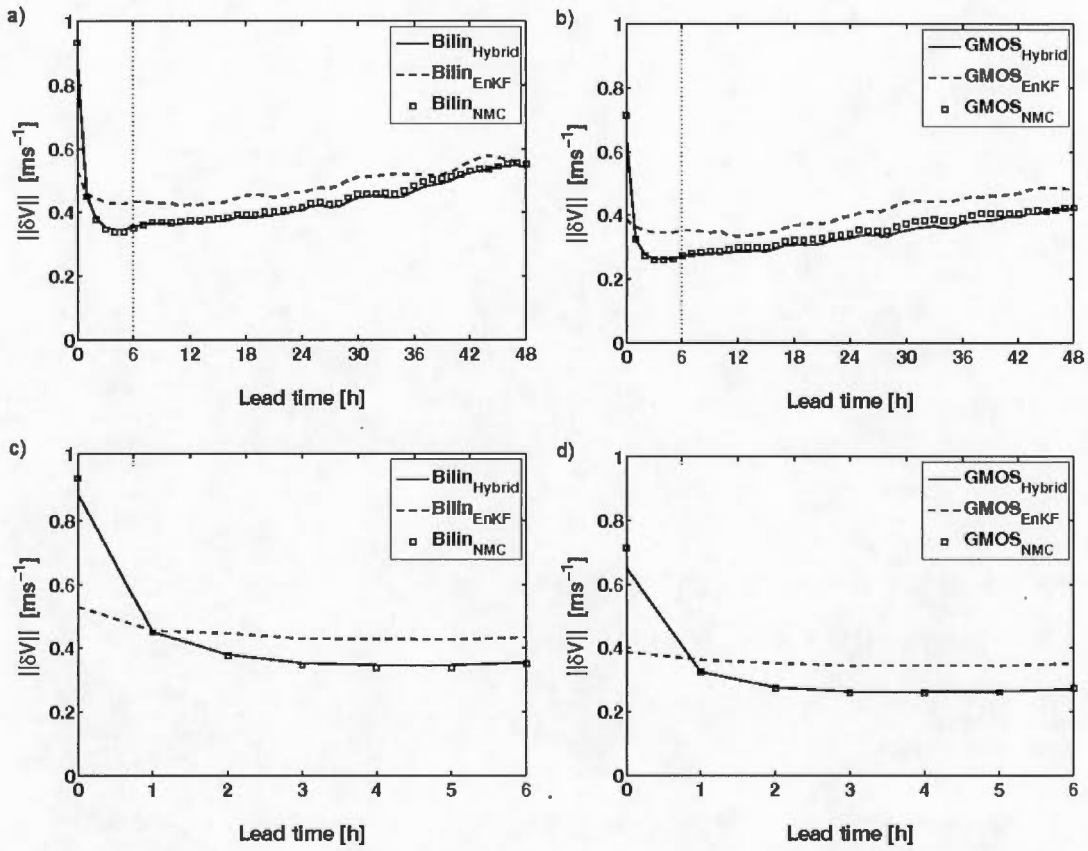

Figure 4.5: Forecast differences (||δV||) between CRTL and the experiments (a: Bilin and b: GMOS) using different background error covariances (Hybrid, NMC and EnKF). Results are presented for the February 1st run launched at 0000 UTC. The upper panels show the results over the whole 48 h forecast period while only the first 6 h are depicted in the lower panels.

It shows that GMOS (b) generates smaller analysis increments than Bilin (a). This can be observed for the different experiments with static, flow-dependent or hybrid background error covariances. For the EnKF cases, a greater fraction of the increment persists when using GMOS compared to Bilin (e.g., the increment initial magnitude is reduced by 6 % (14 %) during the first hour when using GMOS (Bilin)). The fact that

information persists longer in the system suggest that the increment produced when using GMOS are in better agreement with the model state. The NMC experiments produce the worst results in which the rapid forecast difference reduction (so-called forecast convergence) confirms that most of the information from the observations is damped during the first 6 h of model integration. Using the EnKF covariances produces better results as most of the increment is propagated within the forecast. In this case, the forecasts only slightly converge before diverging due to perturbation growth within the NWP model. By construction, the results for the hybrid case are intermediate, but are relatively similar to NMC as the static variances are larger than the flow dependent ones from the EnKF.

To understand the impact of the different background error statistics on the forecasts convergence, the momentum prognostic equation

$$\frac{\partial \mathbf{V}}{\partial t} = \mathbf{T}_{adv} + \mathbf{T}_{cori} + \mathbf{T}_{p.g.} + \mathbf{T}_{v.d.} + \mathbf{T}_{o.b.} + \mathbf{T}_{h.d.} \tag{4.3}$$

is studied. The terms on the right hand side of Eq. (4.3) are the tendencies from advection ($\mathbf{T}_{adv}$), Coriolis effect ($\mathbf{T}_{cori}$), pressure gradients ($\mathbf{T}_{p.g.}$), vertical diffusion ($\mathbf{T}_{v.d.}$), orographic blocking ($\mathbf{T}_{o.b.}$) and horizontal diffusion ($\mathbf{T}_{h.d.}$) terms. Similar to Rodwell and Palmer (2007), the systematic initial tendencies in the first forecast time steps are quantified to assess the reduction of the forecast differences. As defined previously, the CNTRL run is simply a time integration from the background field and the experimental run is a model integration from the analysis. It is defined as $\mathbf{V} = \mathbf{V}_{CNTRL} + \delta\mathbf{V}$ at $t = 0$, where $\delta\mathbf{V}$ is the wind analysis increment such that

$$\frac{\partial \mathbf{V}}{\partial t} = \frac{\partial (\mathbf{V}_{CNTRL} + \delta\mathbf{V})}{\partial t}. \tag{4.4}$$

Considering that the analysis increment is small, Eq. (4.4) can be expressed as

$$\frac{\partial(\delta \mathbf{V})}{\partial t} \approx \frac{\partial \mathbf{V}}{\partial t} - \frac{\partial \mathbf{V}_{CNTRL}}{\partial t} = \delta\left(\frac{\partial \mathbf{V}}{\partial t}\right). \tag{4.5}$$

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. This can explain the convergence of the forecasts during the first 6 h of model integration. Each component of Eq. (4.3) is computed at every station and every model time step for the first 6 h forecast of each experiment. The local differences between the CNTRL and experiment tendencies are calculated and projected on $\delta\mathbf{V}$ from the corresponding time step following

$$\delta\left(\frac{\partial \mathbf{V}}{\partial t}\right)_{Proj} = \left(\frac{\delta\mathbf{V}^T \delta\left(\frac{\partial \mathbf{V}}{\partial t}\right)}{\delta\mathbf{V}^T \delta\mathbf{V}}\right)\delta\mathbf{V}, \tag{4.6}$$

where $\delta(\partial\mathbf{V}/\partial t) = \left[\delta(\partial u/\partial t) \quad \delta(\partial v/\partial t)\right]^T$. The norm of the projection is calculated and the results are averaged over the 4942 stations. Figure 4.6 shows the monthly mean influence of each component of the tendency on $\|\delta\mathbf{V}\|$ for the $GMOS_{NMC}$ and $GMOS_{EnKF}$ experiments. The advection, Coriolis effect and horizontal diffusion terms are negligible compared to the other terms. Only the total tendency as well as pressure gradient, vertical diffusion and orographic blocking tendencies are presented in Figure 4.6.

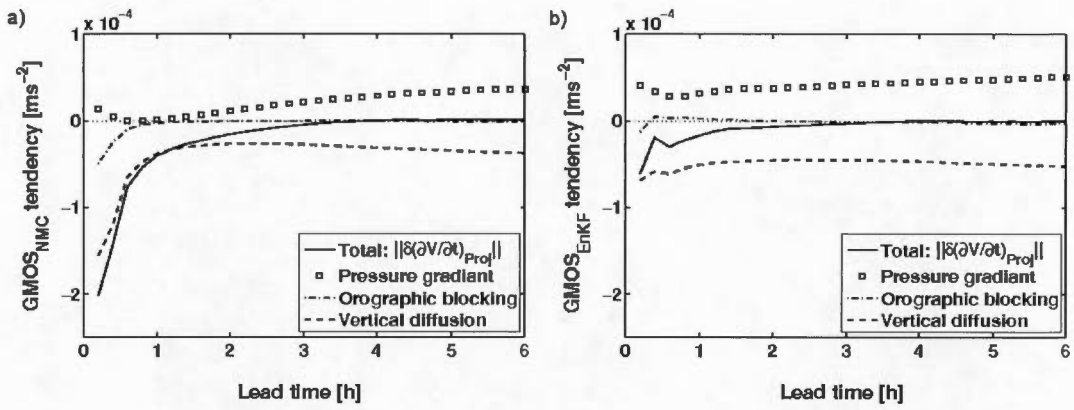

Figure 4.6: Contribution of the main terms of Eq. (4.3) on evolution of the forecast difference between CNTRL and the experiments (a: $GMOS_{NMC}$; b: $GMOS_{EnKF}$). Results for advection, Coriolis effect, horizontal diffusion are omitted as their influence is small compared to the pressure gradient, the vertical diffusion and orographic blocking.

When using the static background error covariances, Figure 4.6 (a) shows that the near-surface wind observations have limited influence on the pressure gradient force (square symbols). The NMC error covariances between wind and mass fields near the surface are small. In this case, the vertical diffusion and orographic blocking damp the surface wind increments (dash and dash-dotted lines respectively). It explains why the forecast difference between CNTRL and the $GMOS_{NMC}$ experiment decreases sharply over time. On the other hand, the EnKF error statistics provide more balanced correlations between mass and near-surface wind fields. From Figure 4.6 (b), the vertical diffusion (dashed line) and orographic blocking (dashed-dotted line) are balanced with the pressure gradient force (square symbol). The flow dependent background error statistics of the EnKF modify both wind and mass fields in a coherent way (through its multivariate correlations) and most of the information extracted from the observations is propagated within the forecast (solid line: $\delta(\partial \mathbf{V}/\partial t)$

≈ 0). Figure 4.6 shows that, unless counterbalanced by proper pressure gradient forces, the atmospheric boundary layer parameterization schemes (i.e., vertical diffusion and orographic blocking) cause the forecast skill from the assimilation of near-surface wind observation to be quickly damped. The use of flow-dependent background error statistics providing coherent multivariate correlations is thus crucial to produce sustainable analysis increments in the lower troposphere.

4.3.2. Evaluation against upper-air observations

Independent radiosonde profiles from 124 stations (see Fig.1) are used to diagnose the forecast error of both wind speed and mass fields at different levels. As mentioned earlier, the upper-air diagnostics are performed over Europe because this area possesses the highest SYNOP station density. Figures 4.7 and 4.8 present vertical profiles of the observed minus forecast departure bias and standard deviation (STD) for the February 2011 period. Figure 4.7 only presents the 24 h lead time biases because results from the 12 h, 36 h and 48 h lead times are similar. On the other hand, Figure 4.8 displays the forecast departure STD at all lead times to fully appreciate the temporal evolution of the information brought in by near-surface wind observations. Since the focus of this study is the lower troposphere, only results from the surface up to 500 hPa are presented.

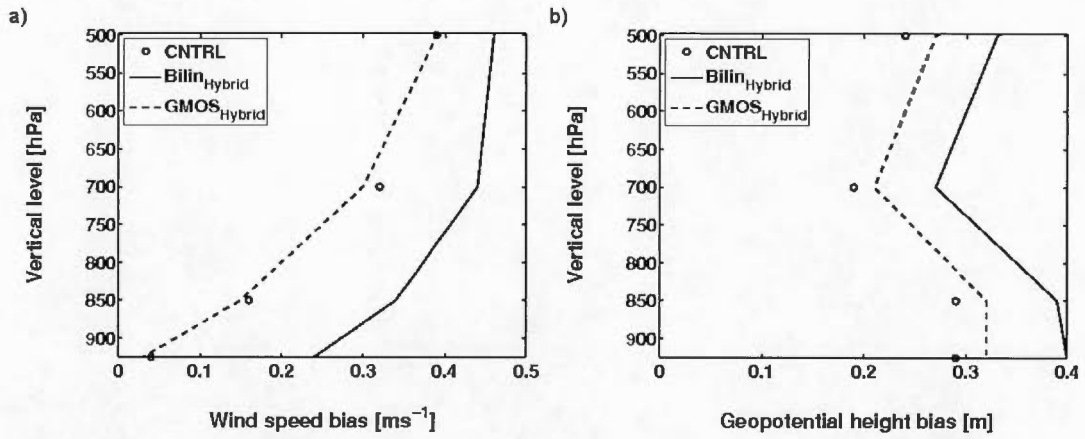

Figure 4.7: 24 h forecast departure bias from hybrid runs as evaluated against radiosonde observations over Europe for wind speed (a) and geopotential height (b).

From Figure 4.7, it is clear that in the Bilin experiment, there is a significant degradation in wind speed and geopotential height biases. On the other hand, GMOS slightly improves wind speed biases while slightly degrading the geopotential height biases. The STD scores from Figure 4.8 also show that the GMOS experiment generally performs better than Bilin. While GMOS results for geopotential height are generally neutral, Bilin significantly degrades those at all vertical levels. GMOS slightly improves 48 h wind speed forecasts at most levels, while Bilin has a neutral effect on this variable.

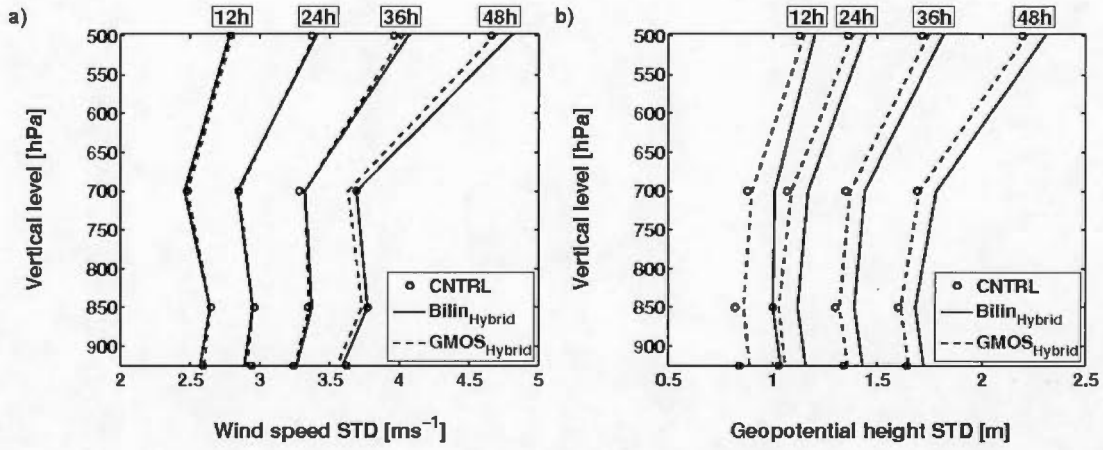

Figure 4.8: Forecast departure STD from hybrid runs as evaluated against radiosonde observations over Europe for wind speed (a) and geopotential height (b). In each plot, results from 12 h, 24 h, 36 h and 48 h forecasts are shown from left to right respectively.

The main difference between the two operators resides in the fact that GMOS slightly improves 48 h wind speed STD whereas Bilin significantly degrades the geopotential height STD at all lead times. In terms of bias, Bilin causes large wind speed and geopotential height degradation, while GMOS only allows some slight degradation in the latter. This degradation may be due to the fact that only near-surface wind observations are assimilated and there are no observations to provide direct information regarding the surface pressure field. Also, the cross-correlation between near-surface wind and pressure fields in the background error covariance matrix may be suboptimal. An evaluation of these issues is presented in the next two sub-sections.

4.3.3. Impact of the background error covariance components

In the experiments using hybrid error covariance statistics, it was found that the assimilation of near-surface winds over land is detrimental to the surface pressure biases. To assess the multivariate impact of the background error covariances used, experiments were performed using the static error covariance matrix (NMC), the flow-dependent error statistics from the EnKF and the hybrid formulation. Figure 4.9 presents the upper-air forecast departure bias for these three experiments. Upper-air forecast departure STD are not shown as results are neutral.

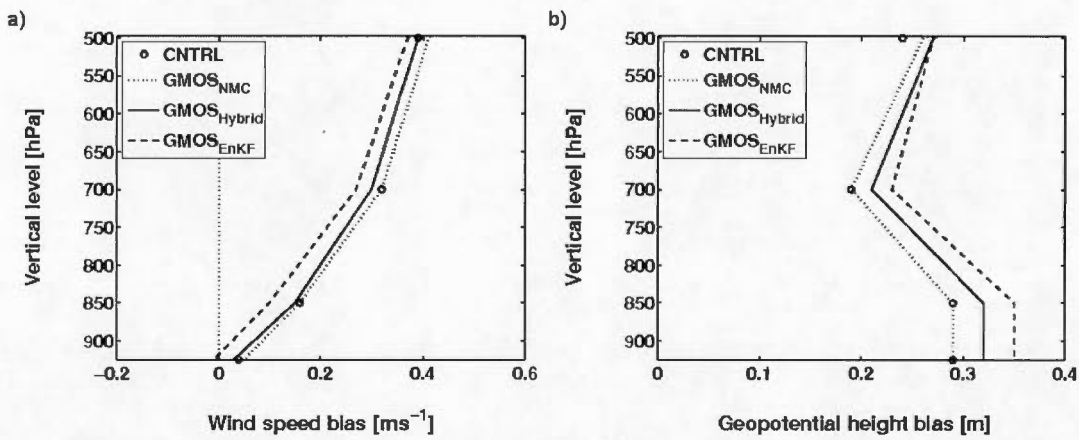

Figure 4.9: 24 h forecast departure bias from GMOS runs as evaluated against radiosonde observation profiles over Europe for wind speed (a) and geopotential height (b).

Figure 4.9 shows that the static background error covariances have a neutral impact on the biases of both wind speed and geopotential height. In contrast, the EnKF experiment has a significant multivariate impact on the bias scores: it has positive (negative) impact on the wind speed (geopotential height) bias. Again by construction, results for the hybrid case are intermediate.

The results show that the impact of near-surface wind observations depend on the quality of the background error statistics used. It seems that homogeneous and isotropic background error statistics (NMC) lack multivariate correlations near the surface and are not suited for the assimilation of near-surface winds observations. The geopotential height increment for the EnKF experiment is induced by the estimated flow-dependent cross-covariance which creates the bias degradation observed. Kepert (2009) discusses this issue and suggest that imbalances could be attributed to localization or other weaknesses of the EnKF.

4.3.4. Constraining mass field using surface pressure observations

Only wind observations were assimilated in the previous experiments which are not representative of the operational context where both the mass and wind fields are constrained by different types of observations. The surface geopotential height bias degradation observed previously may be due to the fact that there are no observations to constrain the mass field in these experiments. To assess if the use of near-surface winds in a more realistic context (where the mass field is constrained with surface pressure observations) can limit the geopotential height bias, conventional surface pressure observations are assimilated along with the near-surface winds. Figure 4.10 presents upper-air forecast departure bias from these experiments using flow dependent background error covariances.

Surface pressure observations have small positive impact on wind and mass forecast STD (not shown). However, the assimilation of surface pressure observations alone has a neutral effect on biases. When combining surface pressure and near-surface wind observations, the geopotential height bias degradation observed in Sections

4.3.2 and 4.3.3 is avoided by constraining the mass field using surface pressure observations. The two experiments assimilating wind observations (experiments $GMOS_{EnKF}$ and P0 + $GMOS_{EnKF}$) significantly reduce the wind speed bias at all vertical levels. Overall, the P0 and the P0 + $GMOS_{EnKF}$ experiments lead to similar results, except that the latter decreases the wind speed bias.

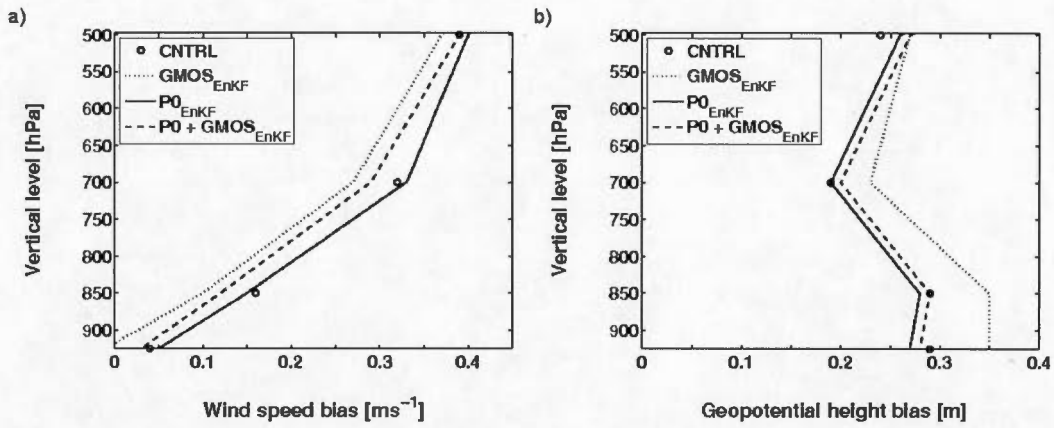

Figure 4.10: 24 h forecast departure bias from GMOS runs as evaluated against radiosonde observations over Europe for wind speed (a) and geopotential height (b). The CNTRL experiment is represented by circles, while assimilation experiments using only surface pressure observations ($P0_{EnKF}$), only near-surface wind observations ($GMOS_{EnKF}$), or both surface pressure and near-surface wind observations (P0 + $GMOS_{EnKF}$) are shown by solid, dot and dash lines respectively.

The geopotential height bias degradation observed previously is significantly reduced when observations are used to constrain the mass field in the analysis. While benefitting from the general positive impact (on forecast STD) of surface pressure observations, the use of complementary near-surface wind observations improves wind bias. As surface pressure is an integrated variable, it contains information about

the atmosphere in its full depth while wind observations near the surface are more representative of the boundary layer flow. Such local observations do not have the same improvement potential as surface pressure observations to reconstruct the full atmospheric state. Still, adding near-surface winds to conventional surface pressure observations provide improvements in terms of wind speed bias. Near-surface wind observations could certainly be used in a combined approach to obtain more information on the atmospheric state when only few data are available (e.g. the early twentieth century reanalysis initiative by Compo *et al.*, 2006).

4.4. Experiments with the operational system

To evaluate the full observation impact from SYNOP wind data, cycling OSEs were performed. Cycling the system allows the information to propagate within the NWP system and contribute to subsequent analyses. Again, the two observation operators (Bilin and GMOS) were used to assimilate the near-surface wind observations. The experiments with EC's operational global deterministic prediction system (including 4DEnVar) are listed in Table 4.2.

In these experiments, all observations assimilated in the global deterministic prediction system are used, which include those from radiosondes, aircraft, wind profilers, land stations, ships, buoys, scatterometers, atmospheric motion vectors, satellite based radio occultation, microwave and infrared satellite sounders/imagers. Wind observations from the 4942 SYNOP stations over land are also assimilated in GMOSOSE and BilinOSE experiments. Although Bédard *et al.* (2015a) showed that the use of flow dependent error statistics is beneficial for the assimilation of near-surface wind data, the assimilation system generally benefits from using this hybrid

representation of the background error covariances (Buehner, 2005). Therefore the operational 4DEnVar configuration employing hybrid covariances was selected. In total, 3 OSEs were performed. Each experiment is cycled over a five-week period (January 24th – February 28th 2011). The experiments are evaluated over the February 2011 period only. All three experiments were initialized using EC's 4D Incremental Analysis Update scheme (4DIAU: Buehner *et al.*, 2015). The 48 h forecasts were initialized twice a day (at 0000 UTC and 1200 UTC) and are systematically verified against surface stations, radiosonde profiles and own analyses.

Table 4.2: Configuration of the 3 cycling OSE performed. Each experiment is listed along with its own combination of near-surface wind observation operator, background and observation error statistics prescribed to the data assimilation system as well as assimilated observations. It is also specified if the experiments are cycled or not.

Experiment name	**Cycling**	**Sfc. wind observation operator (over land)**	**Sfc. wind observation errors (over land)**	**Background errors**	**Obs. assimilated**
CNTRLOSE	Yes	N/A	N/A	Hybrid	All
BilinOSE	Yes	Bilin	Homogeneous	Hybrid	All + SYNOP winds
GMOSOSE	Yes	GMOS	Site dependent	Hybrid	All + SYNOP winds

4.4.1. Evaluation against near-surface wind observations

Near-surface wind observations from the same 4942 SYNOP stations as those used for assimilation purposes were used to diagnose the observation impact on local short-term wind forecasts. Figure 4.11 presents the wind speed forecast STD as a function of forecast lead time for the February 2011 period (post-processed using GMOS). It shows clearly that the assimilation of near-surface wind observations is beneficial for very short-term local wind predictions. However, this impact decreases in time and is only significant up to 6 h because the vertical diffusion and orographic blocking schemes damp the surface wind increments during the first hours of the model integration. Still, the GMOS$_{OSE}$ (Bilin$_{OSE}$) experiment improves the fit of near-surface wind analyses to the observations by 0.16 ms^{-1} (0.10 ms^{-1}) over the CNTRL$_{OSE}$ experiment. These results show that using GMOS in the assimilation increases the consistency between observations, analyses and forecasts.

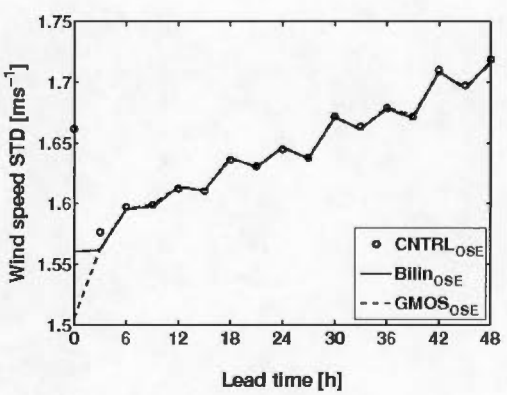

Figure 4.11: Wind speed departure STD as a function of forecast lead time for different experiments (CNTRL$_{OSE}$, Bilin$_{OSE}$ and GMOS$_{OSE}$). Note that the GMOS operator is also used for post-processing in all experiments.

As expected, considering that the background states and analyses carry information from all assimilated observations, such improvements are small compared to the improvement obtained when using GMOS as a post-processing module in Figure 4.4 (see also Bédard *et al.*, 2013). In the OSE context, the fact that near-surface wind observations can further improve low-level wind analyses and very short-term forecasts (in the presence of all operationally assimilated observations) is encouraging.

4.4.2. Upper-air evaluation

The forecast of both wind and mass fields is evaluated at different levels over Europe and adjacent countries. The 124 radiosonde stations presented in Figure 4.3 are used to compute the upper air scores against observations. Analyses generated using the same weather prediction system as each of the forecast experiment (so-called own analyses) are also used to assess the forecast. The score against own analyses are computed at a 1.5° horizontal resolution as recommended by the WMO standard. Unlike the forecast verifications against observations, those against their own analyses cannot be used to diagnose how good short range forecasts are because the forecast errors have a strong imprint of the analysis error. Nevertheless, by evaluating the forecasts using analyses from the same NWP system, it is possible to assess how consistent they are with each other. Following this, cases where the analysis increment is propagated in time shall provide forecasts that are more coherent with analyses (as the information persist in the system) than cases where the increments are diffused by the ABL parameterizations. Higher consistency in itself is desirable (information is propagated rather than diffused), but does not indicate that the forecasts are more accurate.

The scores against upper-air observations from both experiments (GMOSoSE and BilinoSE) indicate that near-surface wind observations have a neutral impact on short-term tropospheric forecasts (not shown). Again, the forecast evaluation against own analyses shows mostly neutral results (not shown): the GMOSoSE (BilinoSE) forecasts and analyses are slightly more (less) consistent than those from the CNTRLoSE experiment, but the differences are not statistically significant.

The weather over the area of interest (Figure 4.3) in February 2011 is examined to detail the verification scores. A careful inspection of the meteorological conditions indicates that three low pressure systems moved across the area during the first two weeks of February. More specifically, the depressions developed over the Norwegian Sea and moved over Scandinavia and the Baltic Sea before hitting Russia. Then a large anticyclone formed over Russia and northern Europe and remained quasi-stationary during the second half of February.

As opposed to the verification scores against radiosondes, the verification scores against analyses described previously were calculated for the whole area of interest without putting emphasis on more densely observed regions. It also covers regions which are not well sampled by radiosonde observations. The spatial and temporal distributions of the verification scores against analyses are thus studied to assess the difference between GMOSoSE and BilinoSE experiments. The geographical distribution of the score differences is shown in Figure 4.12. Results for 12 h forecasts of 10 m wind speed are presented for the February 2011 period.

It shows that GMOSoSE and BilinoSE scores are similar over Western Europe (light gray shading) because this region is densely observed by radiosonde stations (see Figure 4.3) and it is characterized by low synoptic activity during the evaluation period. This is consistent with the neutral upper-air evaluation as the latter is biased

towards the most densely observed region (e.g. Western Europe). This figure also indicates that GMOSose (Bilinose) forecasts are more (less) coherent with the analyses for the area impacted by synoptic activity during the evaluation period (dark gray shading). As opposed to the evaluation against own analyses, the verification against radiosonde observations does not display GMOSose improvements as northern and eastern Europe countries operate few radiosonde stations (see Figure 4.3). It suggests that the positive impact from near-surface wind observations is only significant over less densely observed areas. On the other hand, these countries possess many SYNOP stations which provide useful information to the data assimilation system.

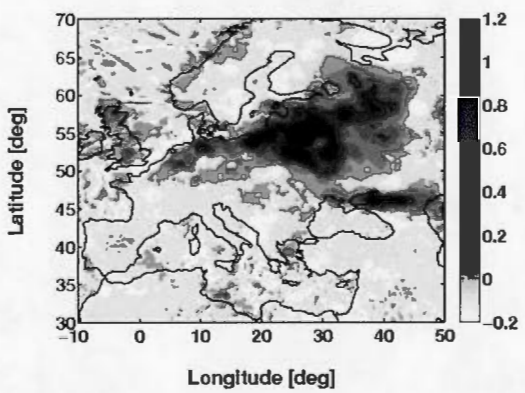

Figure 4.12: Mean 12 h forecast departure STD (against own analyses) differences ($Bilin_{OSE}$ minus $GMOS_{OSE}$) over Europe. Results for 10 m wind speed ([ms^{-1}]) are averaged over the February 2011 period. Positive (negative) values are represented by dark (light) colors. A positive value (dark gray) indicates that the $GMOS_{OSE}$ experiment is better than $Bilin_{OSE}$, while light gray color indicate neutral results.

A Hovmöller diagram is presented in Figure 4.13 to assess the forecasts coherence with the analyses as a function of longitude and time (averaged over latitudes 30°N to

70°N). It presents the differences between the Bilinose and GMOSose 12 h forecast departure STD (against own analyses). The Hovmöller diagram (Figure 4.13) shows that results are generally neutral except during the first two weeks of February between longitudes 20°E and 40°E where three dark bands oriented from upper-left to lower-right depict positive impacts moving from west to east. The SYNOP wind observations impact coincides with the depressions described previously. Overall, by assimilating near-surface wind data associated with synoptic weather elements, GMOSose produces forecasts that verify better against own analyses. Providing that the analysis increments are small, the changes brought to the assimilation system by GMOS are propagated by the NWP model and allows near-surface wind observations over land to have an impact on forecasts downstream.

Bilin operator provides a poor comparison of the model state with the observations and generates strong perturbations in the analyses. As these perturbations are not in agreement with the model state, the increments are diffused and the consistency between forecasts and analyses is slightly reduced. In contrast, the GMOSose increments are in better agreement with the model state. The resulting forecasts are in better agreement with analyses produced using SYNOP wind observations (along with all observations operationally assimilated) and the information propagates in time. It suggests that GMOS has the capacity to integrate persisting information from near-surface wind observations in the NWP system and thus, it provides forecasts that are more coherent with analyses than both CNTRL and Bilinose experiments. Although these results are encouraging, the improvements are slights as the operational system already assimilate a large number of observations (14 million per day).

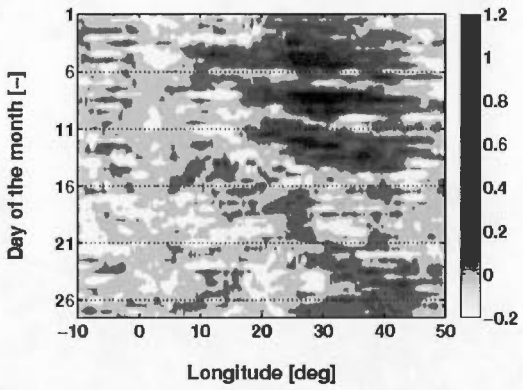

Figure 4.13: Hovmöller diagram presenting the differences between $Bilin_{OSE}$ and $GMOS_{OSE}$ 12 h forecast departure STD (against own analyses). Results for 10 m wind speed ([ms^{-1}]) are presented through February 2011 for different longitude bands over Europe. Positive (negative) values are represented by dark (light) colors. A positive value (dark gray) indicates that the $GMOS_{OSE}$ experiment is better than $Bilin_{OSE}$, while light gray color indicate neutral results.

4.5. Summary and conclusions

The objective of this study is to understand the multivariate observation impact on analyses and forecasts and assess the spatio-temporal propagation of the information in the NWP system. The influence of different components of the assimilation and prediction systems has been examined to identify and understand the mechanisms of the assimilation system that need to be improved. This study also aimed at improving short-term tropospheric forecasts and wind nowcasting capabilities by assimilating near-surface wind observations over land using an improved observation operator (GMOS) in EC's 4DEnVar.

Results show that using GMOS in the assimilation can increase the consistency between observations, analyses and forecasts. The forecast verification against near-surface wind observations show that very short-term wind predictions are significantly improved when using GMOS. As a consequence, this shows that near-surface wind data provide more useful information for very short-term wind forecasts in the lower troposphere.

Results also reveal that the homogeneous and isotropic error statistics produces unbalanced increments which are quickly damped by the vertical diffusion and orographic blocking schemes. By generating pressure gradient forces that balance the diffusion from the ABL parameterization schemes, flow-dependent error statistics allow the observation information to propagate further within the forecast. When using the hybrid formulation, the local observation impact decreases over time and is only significant for 6 h lead time. The spatiotemporal propagation of the information is strongly limited by the quality of the background error covariances. The use of flow dependent background error statistics providing coherent multivariate correlations is thus critical to produce sustainable impacts on the atmosphere by means of correlations between wind and mass fields.

Results from experiments assimilating only near-surface winds show that GMOS improves wind speed bias and STD, while Bilin degrades wind speed bias and geopotential height STD for all lead times. Both experiments degrade geopotential height bias. The use of near-surface winds in a more realistic context where the mass field is constrained with surface pressure observations limits this detrimental effect. Using near-surface winds along with conventional observations (e.g. surface pressure) provide additional improvements in terms of wind speed bias.

Finally, full cycling OSEs assimilating global wind data from SYNOP stations, along with all operationally assimilated observations, were produced to assess the impact of near-surface wind observations. The evaluation against upper-air observations suggests that such observations have a neutral impact on short-term tropospheric forecasts. On the other hand, the evaluation against their own analyses shows that the forecasts and analyses issued from the GMOS (Bilin) experiment can be more (less) consistent than those from the control experiment for 12 h forecast lead times. When presenting the results in a Hövmuller diagram as a function of space and time, the scores show that the small positive impacts are associated with the synoptic activity that occurred during the evaluation period. A detailed analysis of a particular case indicates that assimilating near-surface wind observations with the GMOS (Bilin) operator slightly improves (degrades) the forecasts for three low pressure systems propagating from the North Atlantic towards Eastern Europe in February 2011. Results indicate that the GMOS experiment allows for a better use of near-surface wind observations: it provides forecasts that are more coherent with analyses and the analysis increments are propagated throughout those major synoptic weather elements. As the impacted region possesses few radiosonde stations, this improvement is not captured in the evaluation against upper-air observations. It suggests that the impact of near-surface wind observations may be more significant over less densely observed areas (e.g. areas with few radiosonde stations).

Considering the decoupling of the flow at the top of the atmospheric boundary layer (ABL), near-surface wind observations have a limited influence in the vertical (Bédard *et al.*, 2015a). Also, a large quantity of observations is already assimilated in the system (14 million observations per day). These may be a factor limiting the observation impact on the 3D full NWP model, and improvements on short-term tropospheric forecasts are modest (especially over already well observed areas). However, because of their relatively low cost and their use for monitoring, safety and

climatological needs, near-surface wind observations are abundant (~30 000 stations over the globe). By sampling low level flow, they can provide useful information for local very short-term wind forecasts.

The assimilation of near-surface wind observations represents a real challenge. Consistent with previous work, results show that the representativeness and systematic errors associated with the observation operator are important limitations in this field. The value of such observations for nowcasting applications and for short-term tropospheric forecasts depends on a proper treatment of these errors. The geo-statistical observation operator corrects for biases and partly for representativeness error by giving more weight to the forecast point that best reproduces the natural on-site measurements. It makes a better use of near-surface wind observations and produces analyses that are more coherent with both the observations and forecasts. It also allows for modest forecast improvements. Although representativeness error decreases with NWP model increased resolution, results from Bédard *et al.* (2013) indicate that GMOS still provides significant forecast improvements for a model with 2.5 km grid spacing. GMOS is likely to be useful in the future as representativeness error is expected to remain significant until the local surface characteristics are adequately represented in NWP models.

Yet, EC's hybrid 4DEnVar is limited by the background error covariances it uses. It is mostly influenced by static error statistics because their variances are larger than the flow-dependent ones from the EnKF. With the objective of making a better use of near-surface wind observations over land and improving their impact on short-term tropospheric forecasts, future work must focus on the improvement of background error statistics. With the improvement of flow dependent background error covariances and the reduction of the static error component in the 4DEnVar, the impact of near-surface wind observations may propagate further in time and space.

Acknowledgements

The authors acknowledge the contributions of Jean Côté (UQAM's ESCER Centre and EC), Claude Girard (EC), Paul Vaillancourt (EC) and Ayrton Zadra (EC) for their suggestions on computing the momentum equation components. The authors wish to thank Mark Buehner (EC) for his suggestions and comments. This project is funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), Hydro-Québec research institute (IREQ) and EC Atmospheric Science and Technology Directorate.

References

Ancell B, Kashawlic E, Schroeder J. 2015. Evaluation of wind forecasts and observation impacts from variational and ensemble data assimilation for wind energy applications. *Mon. Wea. Rev.*, doi:10.1175/MWR-D-15-0001.1, in press.

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.*, **141** : 2857–2868.

Benjamin SG, Moninger WR, Sahm SR, Smith TL. 2007. Mesonet wind quality monitoring allowing assimilation in the RUC and other NCEP models. Preprints, *22nd Conf. on Weather Analysis and Forecasting/18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Meteor. Soc., P1.33.

Benjamin SG, Jamison BD, Moninger WR, Sahm SR, Schwartz B, Schlatter TW. 2010. Relative short-range forecast impact from aircraft, profiler, radiosonde, VAD, GPS-PW, METAR, and mesonet observations via the RUC hourly assimilation cycle. *Mon. Wea. Rev.*, **138** : 1319–1343.

Buehner M. 2005. Ensemble-derived stationary and flow-dependent background error covariances: evaluation in a quasi-operational setting for NWP. *Quart. J.R. Meteor. Soc.*, **131** : 1013–1044.

Buehner M, Morneau J, Charette C. 2013. Four-Dimensional Ensemble-Variational Data Assimilation for Global Deterministic Weather Prediction. *Nonlin. Processes Geophys.*, **20**, 669–682.

Buehner M, McTaggart-Cowan R, Beaulne A, Charette C, Garand L, Heilliette S, Lapalme E, Laroche S, Macpherson SR, Morneau J, Zadra A. 2015. Implementation of deterministic weather forecasting systems based on ensemble-variational data assimilation at Environment Canada. Part I: the global system. *Mon. Wea. Rev.*, doi:10.1175/MWR-D-14-00354.1, in press.

Charron M, Polavarapu S, Buehner M, Vaillancourt PA, Charette C, Roch M, Morneau J, Garand L, Aparicio JM, MacPherson S, Pellerin S, St-James J, Heilliette S. 2012. The stratospheric extension of the canadian global deterministic medium-range weather forecasting system and its impact on tropospheric forecasts. *Mon. Wea. Rev.*, **140** : 1924 –1944.

Compo GP, Whitaker JS, Sardeshmukh PD. 2006. Feasibility of a 100-year reanalysis using only surface pressure data. *Bull. of the Amer. Meteor. Soc.*, **87** : 175 – 190.

Côté J, Gravel S, Méthot A, Patoine A, Roch M, Staniforth A. 1998a. The operational CMC/MRB global environmental multiscale (GEM) model: Part I – design considerations and formulation. *Mon. Wea. Rev.*, **126** : 1373–1395.

Deng X, Stull R. 2005. A mesoscale analysis method for surface potential temperature in mountainous and coastal terrain. *Mon. Wea. Rev.*, **133** : 389 – 408.

Gauthier P, Buehner M, Fillion L. 1998. Background-error statistics modelling in a 3D variational data assimilation scheme: estimation and impact on the analyses. Proceedings of the *ECMWF Workshop on the diagnostics of assimilation systems*, 2-4 November 1998, Reading, U.K., 131–145.

Hacker JP, Snyder C. 2005. Ensemble Kalman filter assimilation of fixed screen-height observations in a parameterized PBL. *Mon. Wea. Rev.*, **133** : 3260–3275.

Hacker JP, Anderson JL, Pagowski M. 2007. Improved Vertical Covariance Estimates for Ensemble-Filter Assimilation of Near-Surface Observations. *Mon. Wea. Rev.*, **135** : 1021–1036.

Hacker JP, Rostkier-Edelstein D. 2007. PBL state estimation with surface observations, a column model, and an ensemble filter. *Mon. Wea. Rev.*, **135**, 2958-2972

Houtekamer PL, Deng X, Mitchell HL, Baek SJ, Gagnon N. 2014. Higher resolution in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **142** : 1143–1162.

Ingleby B. 2014. Global assimilation of air temperature, humidity, wind and pressure from surface stations. *Quart. J.R. Meteor. Soc.*, **141** : 504–517.

Kepert, JD. 2009. Covariance localisation and balance in an Ensemble Kalman Filter. *Quart. J.R. Meteor. Soc.*, **135** : 1157–1176.

Parrish DF, Derber JC. 1992. The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120** : 1747–1763.

Rostkier-Edelstein D, Hacker JP. 2010. The roles of surface-observation ensemble assimilation and model complexity for nowcasting of PBL profiles: a factor separation analysis. *Weather and Forecasting*, **25** : 1670–1690.

Rodwell MJ, Palmer TN. 2007. Using numerical weather prediction to assess climate models. *Quart. J.R. Meteor. Soc.*, **133** : 129–146.

Zack J, Natenberg EJ, Young S, Knowe GV, Waight K, Manobianco J, Kamath C. 2010. Application of ensemble sensitivity analysis to observation targeting for short term wind speed forecasting in the Washington - Oregon region`. Technical Report, Lawrence Livermore National Laboratory. 65pp. Available from http://computation.llnl.gov/casc/StarSapphire/pubs/LLNL-TR-458086.pdf.

Zack J, Natenberg EJ, Knowe GV, Waight K, Manobianco J, Hanley D, Kamath C. 2011. Observing system simulation experiments (OSSE) for the Mid-Columbia basin. Technical Report, Lawrence Livermore National Laboratory. 17pp. Available from https://e-reports-ext.llnl.gov/pdf/515298.pdf.

Zadra A, Antonopoulos S, Archambault B, Beaulne A, Bois N, Buehner M, Giguère A, Marcoux J, Petrucci F, Poulin L, Reszka M, Robinson T, St-James J, Rahill A. 2014. Improvements to the Global Deterministic Prediction system (GDPS) (from version 2.2.2 to 3.0.0), and related changes to the Regional Deterministic Prediction System (RDPS) (from version 3.0.0 to 3.1.0). Technical Note, Canadian Meteorological Centre, 88p. Available from http://collaboration.cmc.ec.gc.ca/cmc/CMOI/product_guide/docs/lib/op_systems/doc_opchanges/technote_gdps300_20130213_e.pdf.

CONCLUSION

This study aimed at improving short-term tropospheric forecasts and wind nowcasting capabilities by making a better use of near-surface wind observations over land within a numerical weather prediction system. More precisely, the objective of the study was to develop an improved observation operator for the assimilation of near-surface wind observations over land, to understand the multivariate observation impact on analyses and forecasts and assess the spatio-temporal propagation of the information in the system. The influence of different components of the assimilation and prediction systems has been examined to identify and understand the mechanisms of the assimilation system that need to be improved.

5.1. Original contributions

The assimilation of near-surface wind observations over land presents a real challenge. Previous studies have shown that the representativeness and systematic errors associated with the observation operator are important limitations. This study addresses an important deficiency with the observation operator and a new approach has been introduced to address misrepresentation issues. The improved operator is based on a geo-statistical interpolation scheme (GMOS), including bias and representativeness error corrections. The proposed approach produces analyses which lead to forecasts consistent with surface observations as they rely on MOS methods. Near-surface wind observations over land with high representativeness error were

previously excluded from most operational NWP systems and GMOS now makes it possible to use them without discarding sites on complex terrain or coastal sites.

In Chapter 3, the changes brought to the assimilation were assessed in a controlled environment by evaluating the quality of the resulting analyses (based only on near-surface wind data) using non-assimilated radiosonde observations. Historically, this type of evaluation is performed using assimilated observations. In this case, the use of non-assimilated observations allows for a more rigorous evaluation. While conventional approaches used to estimate observation errors assumes the errors are uncorrelated (e.g., Hollingsworth and Lönnberg, 1986; Desroziers *et al.*, 2005), the proposed method uses the non-assimilated collocated observations to estimate experimentally the observation error correlations associated with different observation operators. This revealed that, through representativeness error, there are error correlations between different types of collocated observations. It was shown that GMOS reduces significantly this error correlation on top of correcting for the systematic and representativeness errors.

The evaluation of the temporal evolution of the analysis increment is presented in Chapter 4. Following Rodwell and Palmer (2007), the physical tendencies were used to understand why near-surface wind observations had little or no impact on forecast beyond 6 h. By quantifying the systematic initial tendencies for the first forecast time steps, it was shown that balance is needed between vertical diffusion, orographic blocking and the pressure gradient for the increment to have more impact on forecasts. Finally, results from experiments were presented to assess the impact of different background error covariance matrices on the analyses and subsequent forecasts.

5.2. Summary of the results

5.2.1. Observation operator

The results from this study show that, while being consistent with forecasts MOS post-processing, GMOS can correct for representativeness errors associated with surface roughness and coastal effects. As a result, GMOS provides a significant reduction of the innovation bias and STD. Also, by using the most representative forecast points, GMOS brings the forecast variability closer to the natural on-site variability (rather than smoothing the forecast) and the background state is generally more consistent with observations when using the GMOS operator rather than Bilin. By removing spurious signals from systematic and representativeness errors, GMOS diminishes the observation error variances and prevents biases in the analyses. The fact that GMOS accounts for biases, while producing relatively smaller increments than Bilin, indicates that GMOS is less prone to generate strong perturbations in the resulting analysis and allows an observation not displaying the same dynamical behaviours as the model (e.g., due to different topographic characteristics) to be more effectively assimilated. Overall, the analysis skill in the lower troposphere is increased when assimilating SYNOP wind observations.

5.2.2. Background error statistics

Although it appears that the decoupling of the flow at the top of the ABL limits the vertical extent of the analysis increment, the spatiotemporal propagation of the information is also strongly limited by the background error covariances. The results from simplified OSEs presented in Chapter 4 show that the static error statistics,

based on homogeneous and isotropic error correlations, produce unbalanced increments which are quickly damped by the vertical diffusion and orographic blocking schemes. On the other hand, the flow-dependent error statistics lead to a better consistency between mass and wind increments through improvements in the multivariate correlations which can vary spatially and temporally. By generating pressure gradient forces that balance the surface drag from the ABL parameterization schemes, the flow-dependent error statistics allow the influence of observations to propagate further within the forecast. Also, the flow-dependent background error covariances from the EnKF also allow for a better vertical propagation of the information in the vertical as it varies depending on the atmospheric stability.

5.2.3. Multivariate impact

Results from experiments assimilating only near-surface winds show that GMOS decreases the wind speed bias and error variance while Bilin degrades the wind speed bias and geopotential height error variance for all lead times. Both experiments also degrade geopotential height bias when using the flow dependent error statistics, but the use of near-surface winds in a more realistic context where the mass field is constrained with surface pressure observations reduces this detrimental effect. While wind observations have a smaller impact on forecasts than surface pressure observations, results show that mass and wind observations provide complementary information that is beneficial to the NWP system. Using near-surface winds along with conventional observations (e.g., surface pressure) provides significant additional improvements in terms of wind speed bias.

5.2.4. Observation impact

Fully cycled OSEs assimilating global wind data from SYNOP stations, along with all operationally assimilated observations, were produced to assess the impact of near-surface wind observations over land. The evaluation against upper-air observations suggests that such observations have a neutral impact on short-term tropospheric forecasts, but the evaluation against their own analyses shows that the GMOS experiment provides forecasts that are more coherent with analyses than those from the control and Bilin experiments for lead times up to 48 h. Results indicate that the positive impacts are associated with the major synoptic activity that occurred during the evaluation period. The improvements were not captured in the evaluation against upper-air observations as the impacted region possesses few radiosonde stations. It also suggests that the impact of near-surface wind observations is more significant over less densely observed areas (e.g. areas with few radiosonde stations).

5.3. Limitations

Being a statistical operator, GMOS is sensitive to the amount of training data used. It was empirically determined that the algorithm requires at least N months of data to provide robust statistical coefficients to an operator using N×N grid-points. Because the operator needs a moderate amount of data from the surrounding grid-points (not just interpolated values), the training needs to be repeated in conjunction with major model changes (e.g., increased horizontal resolution). Also, the assimilation of date from new observing stations would be delayed in an operational system in order to gather the proper amount of training data. Autoregressive methods can be implemented to cope with such issues by training the coefficients online rather than

offline. The fact that the selected GMOS architecture (2×2 grid-points) only need 2 months of data makes it compatible with such schemes. Excluding observations may be a good alternative in the short-term (e.g., Benjamin *et al.*, 2010; Ingleby, 2014), but in the long-term, making a better use of the observations already available will certainly be beneficial.

Although GMOS makes better use of near-surface wind observations by correcting for biases and representativeness errors, it was shown that using it to assimilate near-surface wind observations in the operational system only allows for modest forecast improvements. The observation impact is limited by the background error covariances used. Because the static error variances are larger than the flow-dependent ones from the EnKF, EC's hybrid 4DEnVar is mostly influenced by the static error statistics and thus, the boundary layer parameterization scheme diffuses the increments from near-surface wind observations and the local observation impact rapidly decreases over time. It was shown that the use of background error statistics which vary spatially and temporally while providing consistent multivariate correlations is critical to produce sustainable impacts on the atmosphere. This is one of the factors to consider in order to make a better use of near-surface wind observations and improving their impact on short-term tropospheric forecasts and flow-dependent background error statistics should be improved to reduce the reliance on static error covariances in the future.

5.4. Outlook

As wind energy is increasingly being adopted and used worldwide, the need for improved wind power forecasts based on NWP rises. In an effort to improve forecasts

in the lower troposphere for lead times ranging from 0 h to 48 h, it was shown here that local wind predictions can be significantly improved in the short-term (0 h to 6 h) by assimilating near-surface winds using GMOS. On the other hand, improvements on operational tropospheric forecasts up to 48 h are modestly positive on average because near-surface wind data have a relatively local influence. Results suggest that the use of near-surface wind observations is mainly beneficial for forecasts associated with low pressure systems. These synoptic weather elements are generally associated with high impact weather (e.g., fronts) and such improvements (along with improved nowcasting capabilities) are likely to benefit wind turbine operators and electrical system operators for decision making.

In this study, SYNOP observations were selected because they are abundant all over the globe. Tall anemometer towers (~80 m) from wind power plants are growing in number and could readily be available for data assimilation purposes. As opposed to 10 m observations from METAR or SYNOP stations, measurements from tall anemometer towers are located high above ground level. Although representativeness errors associated with local geographical characteristics are expected to remain significant at 80 m, the observations are more likely to be representative of the flow above the atmospheric boundary layer (especially for stable atmospheric conditions). Thus, it is possible that such observations could have more impact in the vertical. Also, as the effect from the vertical diffusion scheme decreases with height, it is expected that the information from such observations will propagate further in time. Although tall anemometer towers are yet few in number (their actual global impact is probably not significant), they could eventually provide useful near-surface wind observations to improve local nowcasting and short-term forecast capabilities associated with low pressure systems. Thus, it appears worthwhile to evaluate their impact on analyses and forecasts in the future. It would allow a quantitative evaluation of the benefits from assimilating tall anemometer wind observations (in

terms of the value added to existing observations) and how this translates on power forecasts for wind turbines located downstream of the observation and on forecasts related to high impact weather.

REFERENCES

Ancell B, Kashawlic E, Schroeder J. 2015. Evaluation of wind forecasts and observation impacts from variational and ensemble data assimilation for wind energy applications. *Mon. Wea. Rev.*, doi:10.1175/MWR-D-15-0001.1, in press.

Anderson JL, Wyman B, Zhang S, Hoar T. 2005. Assimilation of surface pressure observations using an ensemble filter in an idealized global atmospheric prediction system. *J. Atmos. Sci.*, **62** : 2925–2938.

Auligné T, McNally AP. 2007. Interaction between bias correction and quality control. *Quart. J.R. Meteor. Soc.*, **133** : 643–653.

Auligné T, McNally AP, Dee DP. 2007. Adaptive bias correction for satellite data in a numerical weather prediction system. *Quart. J.R. Meteor. Soc.*, **133** : 631–642.

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.*, **141** : 2857–2868.

Bédard J, Laroche S, Gauthier P. 2015b. Near-Surface wind data assimilation: temporal propagation of the analysis increment and multivariate impact on forecasts. In preparation.

Bengtsson L, Hodges KI, Hagemann S. 2004. Sensitivity of the ERA-40 reanalysis to the observing system: Determination of the global atmospheric circulation from reduced observations. *Tellus A*, **56** : 456–471.

Benjamin SG, Moninger WR, Sahm SR, Smith TL. 2007. Mesonet wind quality monitoring allowing assimilation in the RUC and other NCEP models. Preprints, *22nd Conf. on Weather Analysis and Forecasting/18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Meteor. Soc., P1.33.

Benjamin SG, Jamison BD, Moninger WR, Sahm SR, Schwartz B, Schlatter TW. 2010. Relative short-range forecast impact from aircraft, profiler, radiosonde, VAD, GPS-PW, METAR, and mesonet observations via the RUC hourly assimilation cycle. *Mon. Wea. Rev.*, **138** : 1319–1343.

Buehner M. 2005. Ensemble-derived stationary and flow-dependent background error covariances: evaluation in a quasi-operational setting for NWP. *Quart. J.R. Meteor. Soc.*, **131** : 1013–1044.

Buehner M, Morneau J, Charette C. 2013. Four-dimensional ensemble-variational data assimilation for global deterministic weather prediction. *Nonlin. Processes Geophys.*, **20** : 669–682.

Buehner M, McTaggart-Cowan R, Beaulne A, Charette C, Garand L, Heilliette S, Lapalme E, Laroche S, Macpherson SR, Morneau J, Zadra A. 2015. Implementation of deterministic weather forecasting systems based on ensemble-variational data assimilation at Environment Canada. Part I: the global system. *Mon. Wea. Rev.*, doi:10.1175/MWR-D-14-00354.1, in press.

Charron M, Polavarapu S, Buehner M, Vaillancourt PA, Charette C, Roch M, Morneau J, Garand L, Aparicio JM, MacPherson S, Pellerin S, St-James J, Heilliette S. 2012. The stratospheric extension of the canadian global deterministic medium-range weather forecasting system and its impact on tropospheric forecasts. *Mon. Wea. Rev.*, **140** : 1924 –1944.

Compo GP, Whitaker JS, Sardeshmukh PD. 2006. Feasibility of a 100-year reanalysis using only surface pressure data. *Bull. of the Amer. Meteor. Soc.*, **87** : 175 – 190.

Côté J, Gravel S, Méthot A, Patoine A, Roch M, Staniforth A. 1998a. The operational CMC/MRB global environmental multiscale (GEM) model: Part I – design considerations and formulation. *Mon. Wea. Rev.*, **126** : 1373–1395.

Côté J, Gravel S, Méthot A, Patoine A, Roch M, Staniforth A. 1998b. The operational CMC/MRB global environmental multiscale (GEM) model: Part II – Results. *Mon. Wea. Rev.*, **126** : 1397–1418.

Courtier P, Thépaut JN, Hollingsworth A. 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J.R. Meteor. Soc.*, **120**, 1367–1387

Dee DP, Uppala S. 2009. Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quart. J.R. Meteor. Soc.*, **135** : 1830–1841.

Deng X, Stull R. 2005. A mesoscale analysis method for surface potential temperature in mountainous and coastal terrain. *Mon. Wea. Rev.*, **133** : 389 – 408.

Desroziers G, Berre L, Chapnik B, Poli P. 2005. Diagnosis of observation, background and analysis-error statistics in observation space. *Quart. J.R. Meteor. Soc.*, **131** : 3385–3396.

Dong J, Xue M, Droegemeier K. 2010. The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteorology and Atmospheric Physics*, **112** : 41–61.

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99** (C5), 10143–10162.

Gauthier P, Buehner M, Fillion L. 1998. Background-error statistics modelling in a 3D variational data assimilation scheme: estimation and impact on the analyses. Proceedings of the *ECMWF Workshop on the diagnostics of assimilation systems*, 2-4 November 1998, Reading, U.K., 131–145.

Gauthier P. Thépaut JN. 2001. Impact of the digital filter as a weak constraint in the preoperational 4DVAR assimilation system of Météo-France. *Mon. Wea. Rev.*, **129** : 2089-2102.

Gauthier P, Tanguay M, Laroche S, Pellerin S, Morneau J. 2007. Extension of 3DVAR to 4DVAR: implementation of 4DVAR at the Meteorological Service of Canada. *Mon. Wea. Rev.*, **135** : 2339–2354.

Giebel G, Brownsword R, Kariniotakis G, Denhard M, Draxl C. 2011. The State-Of-The-Art in Short-Term Forecasting of Wind Power - A Literature Overview, Deliverable Report of the Anemos Project, D-1.2. 109 p. Available from http://www.anemos-plus.eu/index.php?option=com_content&view=article&id=91&Itemid=69.

Global Wind Energy Council. 2013. Global Wind Report 2013. Online, 80 p. Available from www.gwec.net/wp-content/uploads/2014/04/GWEC-Global-Wind-Report_9-April-2014.pdf.

Hacker JP, Snyder C. 2005. Ensemble Kalman filter assimilation of fixed screen-height observations in a parameterized PBL. *Mon. Wea. Rev.*, **133** : 3260–3275.

Hacker JP, Anderson JL, Pagowski M. 2007. Improved Vertical Covariance Estimates for Ensemble-Filter Assimilation of Near-Surface Observations. *Mon. Wea. Rev.*, **135** : 1021–1036.

Hacker JP, Rostkier-Edelstein D. 2007. PBL state estimation with surface observations, a column model, and an ensemble filter. *Mon. Wea. Rev.*, **135**, 2958-2972

Henne S, Brunner D, Folini D, Solberg S, Klausen J, Buchmann B. 2010. Assessment of parameters describing representativeness of air quality in-situ measurement sites. Atmospheric Chemistry and Physics, **10** : 3561–3581.

Hollingsworth A, Lonnberg P. 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus A*, **38**: 111–136.

Houtekamer PL, Deng X, Mitchell HL, Baek SJ, Gagnon N. 2014. Higher resolution in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **142** : 1143–1162.

Huang X, Gao F, Jacobs N, Wang H. 2013. Assimilation of wind speed and direction observations: a new formulation and results from idealised experiments. *Tellus A*, **65**, 19936, http://dx.doi.org/10.3402/tellusa.v65i0.19936.

Ingleby B. 2014. Global assimilation of air temperature, humidity, wind and pressure from surface stations. *Quart. J.R. Meteor. Soc.*, **141** : 504–517.

Janjic T, Cohn SE. 2006. Treatment of observation error due to unresolved scales in atmospheric data assimilation. *Mon. Wea. Rev.*, **134** : 2900–2915.

Kariniotakis G, Martí I, Casas D, Pinson P, Nielsen TS, Madsen H, Giebel G, Usaola J, Sanchez I, Palomares AM, Brownsword R, Focken U, Lange M, Tambke J, Louka P, Kallos G, Lac C, Descombes G, Sideratos G. What performance can be expected by short-term wind power prediction models depending on site characteristics? Paper for the European Wind Energy Conference and Exhibition, London (UK), 22-25 November 2004.

Kepert, JD. 2009. Covariance localisation and balance in an Ensemble Kalman Filter. *Quart. J.R. Meteor. Soc.*, **135** : 1157–1176.

Klinker E. Rabier F. Kelly G. Mahfouf JF. 2000. The ECMWF operational implementation of four dimensional variational assimilation. Part III: Experimental results and diagnostics with operational configuration. *Quart. J.R. Meteor. Soc.*, **126** : 1191–1215.

Koohkan MR, Bocquet M. 2012. Accounting for representativeness errors in the inversion of atmospheric constituent emissions: application to the retrieval of regional carbon monoxide fluxes. *Tellus B*, 64 : doi:10.3402/tellusb.v64i0.19047.

Landberg L, Giebel G, Nielsen HA, Nielsen TS, Madsen H. 2003. Short-Term Prediction - An Overview, *Wind Energy*, **6** : 273–280.

Landberg L, Watson SJ. 1994. Short-term predictions of local wind conditions. *Boundary-Layer Meteorology*, **70** : 171–195.

Lange M, Focken U, Meyer R, Denhardt M, Ernst B, Berster F. 2006. Optimal combination of different numerical weather models for improved wind power predictions. In: *International Workshop on Large-Scale Integration of Wind Power and Transmission Networks for Offshore Wind Farms*, Delft, 2006.

Lavaysse C, Carrera M, Bélair S, Gagnon N, Frenette R, Charron M, Yau MK. 2013. Impact of Surface Parameter Uncertainties within the Canadian Regional Ensemble Prediction System. *Mon. Wea. Rev.*, **141** : 1506–1526.

Lea DJ, Drecourt JP, Haines K, Martin MJ. 2008. Ocean altimeter assimilation with observational- and model-bias correction. *Quart. J.R. Meteor. Soc.*, **134** : 1761–1774.

Le Dimet, F, Talagrand O. 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, *Tellus A*. **38** : 97–110.

Liu ZQ, Rabier F. 2002. The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study. *Quart. J.R. Meteor. Soc.*, **128** : 1367–1386.

Liu H. 2009. Wind speed forecasting for wind energy application, PhDThesis, York University, Toronto, Canada, 240 p.

Mahfouf JF. Rabier F. 2000. The ECMWF operational implementation of four dimensional variational assimilation. Part II: Experimental results with improved physics. *Quart. J.R. Meteor. Soc.*, **126** : 1171–1190.

Ministère des Ressources Naturelles et de la Faune. 2006. Québec energy strategy 2006-2015. Québec, 16p.

Müller MD. 2011. Effects of model resolution and statistical postprocessing on shelter temperature and wind forecasts. *J. Appl. Meteor. Climatol.*, **50** :1627–1636.

Nash J, Oakley T, Vömel H, Wei LI. 2011. WMO Intercomparison of high quality radiosonde systems, Yangjiang, China, 12 July - 3 August 2010. Instruments and Observing Methods. Report No. 107, World Meteorological Organization, 248 p. Available from https://www.wmo.int/pages/prog/www/IMOP/publications-IOM-series.html.

Nielsen TS, Madsen H, Christensen HS. 2000. WPPT - a tool for wind power prediction. In Proceedings of the Wind Power for the 21st Century Conference, Kassel, 2000.

Nielsen HA, Nielsen TS, Madsen H, Badger J, Giebel G, Landberg L, Sattler K, Feddersen H. 2004. Comparison of Ensemble Forecasts with the Measurements from the Meteorological Mast at Risø National Laboratory. Project report of the Anemos project, PSO (FU 2101). 45 p.

Nielsen HA, Pinson P, Christiansen LE, Nielsen TS, Madsen H, Badger J, Giebel G, Ravn HF. 2007a. Improvement and Automation of Tools for Short-Term Wind Power Forecasting. In: European Wind Energy Conference, Milan (Italy), 2007.

Nielsen HA, Nielsen TS, Madsen H, San Isidro Pindado MJ, Martì I. 2007b. Optimal Combination of Wind Power Forecasts, *Wind Energy*, **10** : 471-482.

Parrish DF, Derber JC. 1992. The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120** : 1747–1763.

Rabier F., Järvinen H., Klinker E., Mahfouf J.F. and Simmons A., 2000. The ECMWF operational implementation of four dimensional variational assimilation. Part I: experimental results with simplified physics. *Quart. J.R. Meteor. Soc.*, **126** : 1143–1170.

Rodwell MJ, Palmer TN. 2007. Using numerical weather prediction to assess climate models. *Quart. J.R. Meteor. Soc.*, **133** : 129–146.

Rostkier-Edelstein D, Hacker JP. 2010. The roles of surface-observation ensemble assimilation and model complexity for nowcasting of PBL profiles: a factor separation analysis. *Weather and Forecasting*, **25** : 1670–1690.

Whitaker JS, Compo GP, Wei X, Hamill TM. 2004. Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.*, **132** : 1190–1200.

Zack J, Natenberg EJ, Young S, Knowe GV, Waight K, Manobianco J, Kamath C. 2010. Application of ensemble sensitivity analysis to observation targeting for short term wind speed forecasting in the Washington - Oregon region`. Technical Report, Lawrence Livermore National Laboratory. 65pp. Available from http://computation.llnl.gov/casc/StarSapphire/pubs/LLNL-TR-458086.pdf.

Zack J, Natenberg EJ, Knowe GV, Waight K, Manobianco J, Hanley D, Kamath C. 2011. Observing system simulation experiments (OSSE) for the Mid-Columbia basin. Technical Report, Lawrence Livermore National Laboratory. 17pp. Available from https://e-reports-ext.llnl.gov/pdf/515298.pdf.

Zadra A, Antonopoulos S, Archambault B, Beaulne A, Bois N, Buehner M, Giguère A, Marcoux J, Petrucci F, Poulin L, Reszka M, Robinson T, St-James J, Rahill A. 2014. Improvements to the Global Deterministic Prediction system (GDPS) (from version 2.2.2 to 3.0.0), and related changes to the Regional Deterministic Prediction System (RDPS) (from version 3.0.0 to 3.1.0). Technical Note, Canadian Meteorological Centre, 88p. Available from http://collaboration.cmc.ec.gc.ca/cmc/CMOI/product_guide/docs/lib/op_systems/doc_opchanges/technote_gdps300_20130213_e.pdf.