Modelling of background error covariances for the analysis of clouds and precipitation

Thibaut Montmerle, Yann Michel and Benjamin Ménétrier

Météo-France/CNRM-GAME
42 av. G. Coriolis, 31057 Toulouse, France
thibaut.montmerle@meteo.fr

ABSTRACT

After a brief introduction to background error modelling, this paper describes several examples of covariances in clouds and precipitation computed from ensemble of forecast differences performed using non-hydrostatic Cloud Resolving Models (CRM). These kind of models allow realistic forecast of diabatic processes thanks to the explicit parametrization of most microphysical processes. The methodology is based on the use of geographical masks that allow to compute forecast errors statistics that are restricted to areas where rain or fog is predicted. The resulting covariances strongly differ from climatological statistics that are generally used in operations. They are also in good agreement with the physics processes that are active in those areas. In precipitating cases for instance, strong coupling of humidity, cloud and rain water contents with divergence has been found, as well as shorter horizontal correlation lengths. Overall, covariances in clouds and precipitation are characterized by different balance relationships between control variables, and by strong inhomogeneity and therefore flow dependency. Possible strategies to take these characteristics in current B matrix formulations are reviewed.

1 Introduction

A long-standing problem in numerical weather prediction (NWP) has been to provide adequate initial conditions for the quantitative prediction of clouds and precipitation, given available observations such as satellite microwave imagers, infrared radiances or Doppler radars. As pointed out by Errico *et al.*, 2007, the strong non-linearities of moist physical processes raise several issues to accommodate for these observations in assimilation schemes based on linear theory such as variational data assimilation (VAR), especially i) the linearization of observation operators, ii) the handling of non-Gaussian probability density functions (pdf) for hydrometeors errors, and iii) the modeling of appropriate background error covariances (so called **B** matrix). As a matter of fact, VAR needs a priori (or background) meteorological fields in order to provide information in non-observed areas and to provide realistic reference state needed in some nonlinear observation operators. As observations, the background state is prone to error, which are taken into account in VAR through the use of the **B** matrix.

As pointed out by Daley, 1991, **B** has a profound impact on the analysis, by weighting the importance of the a priori state, by smoothing and spreading information from observation points, and by imposing balance between the model control variables (CV). To estimate this matrix, two main practical difficulties occur. First, the "true" state needed to measure error against is unknown. Secondly, because of its size, **B** can be neither estimated at full rank nor stored explicitly. Covariances have thus to be modelled, typically by computing statistics on differences between forecasts that allow to mimic forecast errors (reviews of such methods can be found in Bannister, 2008. In order to retrieve climatological covariances, these statistics are usually performed for long time periods and moreover averaged over the whole computational domain, which explains why the signal of clouds and precipitation forecast errors is weak in the final results. As a consequence, current operational formulation of **B** are often inadequate in cloudy and rainy areas where couplings between errors of CV are misrepresented and where the use

of homogeneous and isotropic horizontal correlations are not appropriate. Lopez and Bauer, 2007 for instance pointed out that the lack of impact of precipitation observations on the dynamical and mass fields found in 4DVar analyses after the 1D+4DVar assimilation of radar data may be due to the absence of coupling relationships between those fields and humidity.

To address these issues, Auligné *et al.*, 2010 emphasize the need to relax the ergodicity assumption in clouds and precipitation by allowing for inhomogeneous background error modeling to represent the spatial variability of the background error matrix and by using of better balance relationships, based either on statistical regressions (Berre, 2000) or on more sophisticated diabatic non linear balance (Pagé *et al.*, 2007). By using geographical masks in the computation of **B**, recent studies have diagnosed background error covariances separately in precipitating and non-precipitating areas for regional models (Caron and Fillion, 2010) and for cloud resolving models (CRM) (Montmerle and Berre, 2010 - MB2010 hereafter, Michel *et al.*, 2010). The purpose of this paper is to review some of the main conclusions of the latter studies and to present some recent results obtained for fog based on the same kind of methodology. The very different behaviors that have been found may suggest to redesign **B** in those areas by using more adequate balance relationships between CV and by considering inhomogeneous flow-dependent covariances.

In section 2, we recall briefly the way **B** is modeled using control variable transforms (CVT) and decomposed as a product of a balance operator and of a spatial transform. In this context, covariances diagnosed for precipitating and non-precipitating areas are then discussed in section 3 and 4 for "traditional" CV (i.e wind, mass and humidity) and for cloud and rain water content respectively. Results for fog are also briefly presented in section 3, possible strategies for the specification of background error covariances modeling in VAR for clouds and precipitations being given in the final section.

2 Modelling of B in the CVT framework

Most operational centres are based on the CVT methodology, which consists in replacing the increment δx by a control variable χ in a VAR algorithm written in an incremental formulation (Courtier *et al.*, 1994). For a 3DVar, the total cost function writes:

$$J_\chi = \frac{1}{2}\chi^T\chi + \frac{1}{2}\left(\mathbf{HB^{1/2}}\chi - \mathbf{d}\right)^T \mathbf{R^{-1}}\left(\mathbf{HB^{1/2}}\chi - \mathbf{d}\right)$$

with

$$\delta\mathbf{x} = \mathbf{B^{1/2}}\chi \tag{1}$$

where **R** is the observation error covariance matrix, d the innovation vector $(\mathbf{y} - H[\mathbf{x_b}])$ that measures the difference between the observation y and its simulated counterpart computed by applying the non-linear observation operator H to the background $\mathbf{x_b}$, and where **H** is the linearized version of H. In the CVT formulation, the background term becomes trivial and the observation term represents the distance between the innovation vector d and the increment written in the observation space. Another advantage of the CVT is the improvement of the preconditioning.

It allows also to express univariate and multivariate aspects of **B** is a compact and efficient way. χ does not necessarily contain the same number of degrees of freedom than $\delta\mathbf{x}$. For instance, χ may include extra information that does not contribute directly to the analysis (e.g. VarBC Dee, 2005). It implies also that background errors of the control vector χ are uncorrelated and have unit variance. The challenge of background error covariance modeling is to capture in $\mathbf{B^{1/2}}$ the known important features of the background error covariance matrix. For instance, these features should be in agreement with atmospheric balance, non-separability of the flow, and take into account the effect of the observation

density and accuracy. The formalism of Derber and Bouttier, 1999 expresses $\mathbf{B}^{1/2}$ as a combination of two terms:

$$\mathbf{B}^{1/2} = \mathbf{K}\mathbf{B}_{\mathbf{u}}^{1/2}$$

Eq. (1) can be written:

$$\chi = \mathbf{B}^{-1/2}\delta\mathbf{x} = \mathbf{B}_{\mathbf{u}}^{-1/2}\mathbf{K}^{-1}\delta\mathbf{x}$$

$\mathbf{K}^{-1}$ is called the inverse of the balance operator (or parameter transform) and aims in taking increments of the model's variable and to output new less correlated parameters on the same grid using balance constraints, under the basic hypothesis that errors in geostrophically balanced parameters (associated with Rossby modes) are decoupled from errors in unbalanced parameters (associated with inertia-gravity modes). $\mathbf{B}_{\mathbf{u}}^{-1/2}$ is a block diagonal matrix called the spatial transform. It aims in projecting each parameter onto uncorrelated spatial modes, and then in dividing by the square root of the variance of each mode.

Several different representations of these operators have been developed at NWP centers. The ECMWF formulation (also used at Météo-France) designs the balance operators with either analytical or based on scale-dependent regression operators (Derber and Bouttier, 1999; Berre, 2000; Fisher, 2003). Spatial correlations are constructed through empirical orthogonal decomposition and spectral or wavelet diagonal assumptions. On the contrary, grid point formulations are used at NCEP and NCAR (Wu *et al.*, 2002; Barker *et al.*, 2004; Michel and Auligné, 2010) with physical space regression-based balances and with recursive filters (Purser *et al.*, 2003) to ensure spatial correlations. The Met-Office formulation (Ingleby, 2001; Lorenc *et al.*, 2003) is based on a grid point analytical balance but spectral (or wavelet) transforms for horizontal correlations. Control variables are also different. A review of these distinct formulations can be found in Bannister, 2008.

3 Diagnosis of error covariances in clouds and precipitation for traditional CV

3.1 Convective systems

In order to quantify the misrepresentation of clouds and precipitation in the current climatological **B** used in the operational NWP system at convective scale AROME (Seity *et al.*, 2010), MB2010 have diagnosed background error covariances for precipitating and non-precipitating areas. Computing such covariances from a CRM that uses an explicit microphysical scheme allows indeed to document covariances and balances in areas that are under-represented in samples of forecast differences used for climatological covariances computation. For that purpose, an AROME ensemble assimilation of several convective cases has been built, and statistics have been computed from forecast differences by considering separately rainy and no-rainy profiles. In the balance operator **K** used in this study, increments of specific humidity q is linked to other CV following Berre, 2000 multivariate formalism:

$$\delta q = \mathscr{Q}\mathscr{H}\delta\tilde{\zeta} + \mathscr{R}\delta\tilde{\eta}_u + \mathscr{S}(\delta\tilde{T}, \delta\tilde{P}_s)_u + \delta\tilde{q}_u \qquad (2)$$

where $\delta\tilde{\zeta}, \delta\tilde{\eta}, (\delta T, \delta P_s)$ are respectively forecast errors of vorticity, divergence, temperature and surface pressure; the subscript u stands for unbalanced (total minus balanced) fields; $\mathscr{Q}, \mathscr{R}$ and $\mathscr{S}$ are operators deduced from statistical regressions; $\mathscr{H}$ is a horizontal geostrophic balance operator that couples mass and vorticity. At Météo-France, Eq. 2 is used operationally at regional (ALADIN, (Fischer *et al.*, 2005)) and convective (AROME) scales.

Caron and Fillion, 2010 have shown that forecast errors drift away from linear geostrophic balance over precipitation areas and that this deviation is proportional to the intensity of precipitation. MB2010

have confirmed this result and have furthermore shown that precipitating areas are mainly characterized by: i) larger error standard deviations for divergence η and vorticity ζ, which denotes a more intense small scale dynamical activity, ii) twice shorter correlation lengths for q and T, iii) larger vertical auto-correlations in the mid troposphere reflecting the stronger vertical mixing within clouds, iv) very different contributions, in scale and in intensity, to the explained q error variances reflecting the presence of low level cold pool, low level convergence, latent heat release, and cloud top divergence.

To illustrate this last point, Fig. 1 shows the vertical distributions of variance ratios that explain q variance in rainy and non-rainy areas. The differences of behaviors are striking: in the rainy case, the total explained variance of q varies between 25% and 50% in the troposphere, and it is mainly controlled by unbalanced divergence $\delta\eta_u$; in contrast, in the non-precipitating case, the total explained variance of q varies only between 10% and 30% in the troposphere, and is mainly driven by the unbalanced mass field $(T,P_s)_u$. This strong coupling between moisture and convergence in precipitating areas is consistent with the analysis of mesoscale balance coupling in Pagé *et al.*, 2007, and is of great interest for the assimilation of e.g Doppler radar data in precipitation. Michel *et al.*, 2010 have also found different multivariate couplings and overall shorter horizontal (larger vertical) correlations with a different NWP model and a different **B** formulation, suggesting that the result may be valid for a large class of NWP models and weather situations at convective scales.

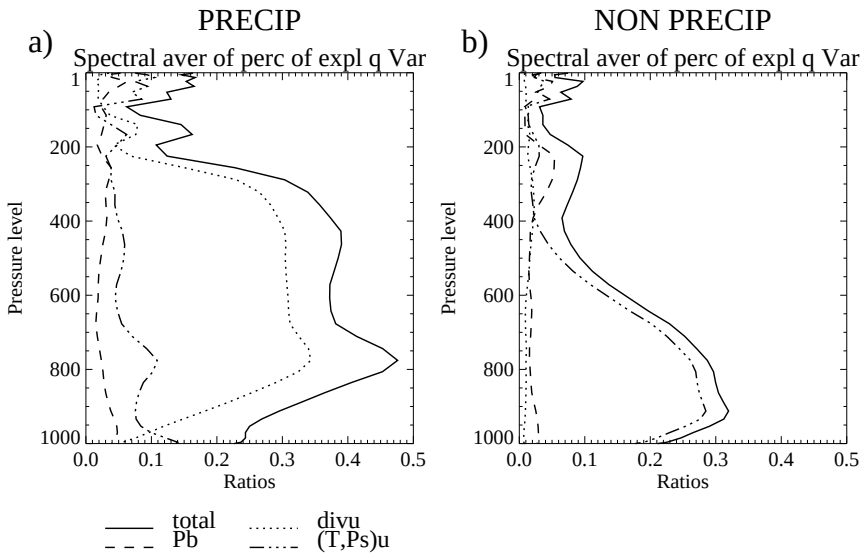

Figure 1: Explained variance ratios of specific humidity q error variances as a function of height, computed over (a) precipitating and (b) non-precipitating areas. P_b stands for the so-called balanced mass and *divu* for the unbalanced divergence (from Montmerle and Berre, 2010)

3.2 Fog

The same approach has been recently applied to compute specific background error covariances in fog. An ensemble assimilation based on AROME 3D-Var has been build for 18 cases of fog observed over France during the fall and the winter 2009, and statistics have been computed on differences of forecast separately in regions where fog was and was not predicted in both forecasts. To determine the best model diagnostic for fog, a long time period comparison between different parameters and the french CARIBOU fog and haze analysis (Guidard and Tzanos, 2007) was performed. The best predictor was found to be the nebulosity computed for model levels below 150 m.

As for convective cases, different behaviors were found for regions where fog is predicted. The mean vertical profile of T standard deviation exhibits a maximum at low levels, reflecting the vertical variation of the level of inversion (not shown). The presence of saturated conditions within fog explains why specific humidity q variance is strongly balanced and exhibits a very strong coupling (up to 75%) with the mass field $(T, P_s)_u$ and a moderate coupling with divergence in the lowest layers, whereas almost no coupling is found in clear air (Fig. 2).

Vertical auto-correlations also display strong differences for all CV consistent with the presence of a very stable layer of constant potential temperature near the ground. For example, Fig. 3 shows that T errors in the first 3 model levels ($<$150 m approximately) are totally uncorrelated from levels up above. In clear air or in the operational **B**, much more vertical mixing appends. This illustrates a well known problem of many operational models, where increments from ground-based observations propagate too much vertically in cases of very stable boundary layer and deteriorate the resulting forecasts (Lorenc, 2007). Using the covariances obtained for fog in the heterogeneous **B** matrix formulation should alleviate this problem (see section 5).

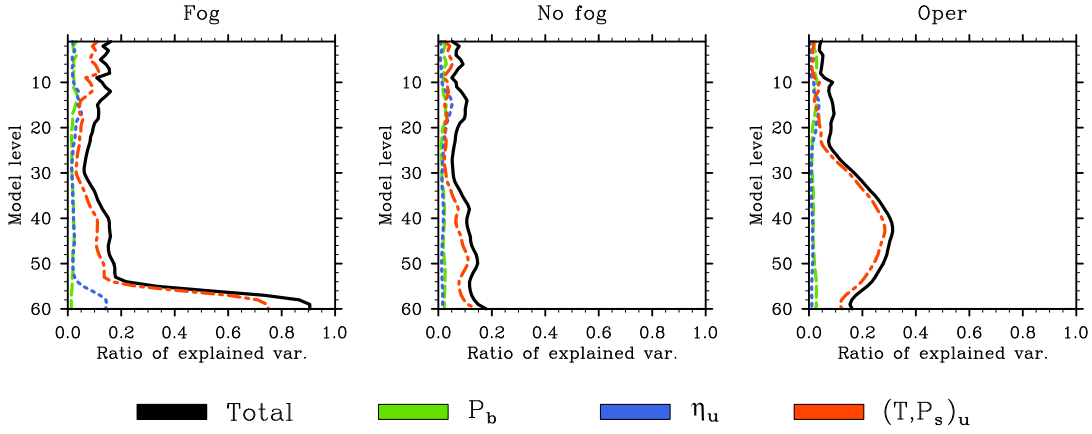

Figure 2: Same as Fig. 1 but for fog (left), for areas without fog (middle) and for the operational **B** used in AROME (right)

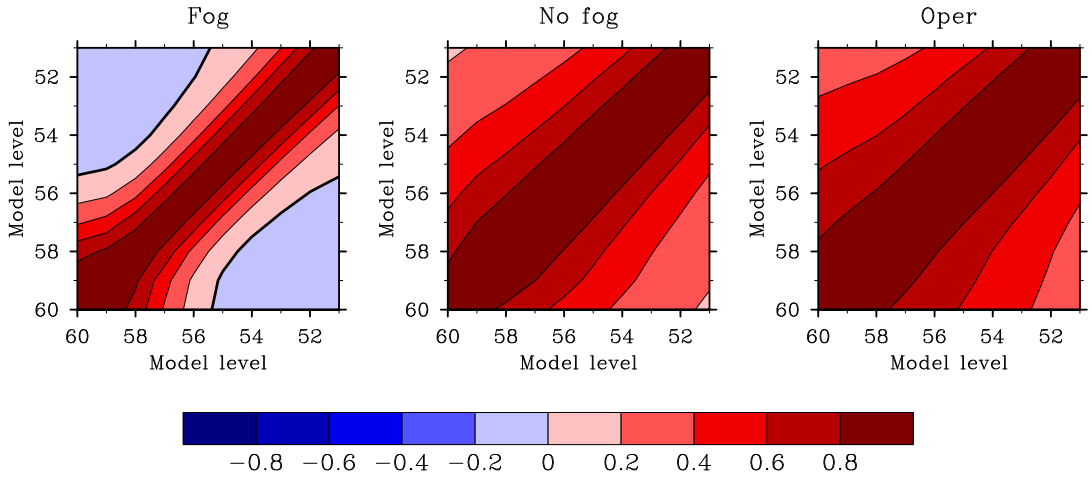

Figure 3: Vertical auto-correlations for T errors (zoom in the first 400m) for fog (left), for areas without fog (middle) and for the operational **B** used in AROME (right)

4 Diagnosis of error covariances in clouds and precipitation for cloud and rain water content

As mentioned in the introduction, the direct assimilation of cloud q_c and rain q_r water content raises several issues such as the linearization of strongly non-linear observation operators and the non-Gaussianity of the pdf of hydrometeor errors. To avoid those impediments, ECMWF use operationally a 1D+4DVar method derived from Marécal and Mahfouf, 2002 early work to assimilate the total water vapor column previously retrieved from rainfall rates observed by microwave imagers thanks to the use of a 1D-Var. Other studies have tried instead to retrieve, also using a 1D-Var, profiles of cloud and rain water contents from ATOVS, SSM/I or TRMM radiances. In those studies, vertical background error covariances where whether empirical (Chevallier *et al.*, 2002), or, for instance, obtained by perturbing the input of the model's moisture scheme (q and T, Moreau *et al.*, 2003). More recently, Amerault and Zou, 2006 have computed such vertical covariances for warm and cold hydrometeors, by performing differences on forecasts that were using different explicit moisture scheme. To directly assimilate these kind of profiles however requires to include hydrometeors in the CV and to compute full covariances of background errors for the hydrometeors that are considered (among other issues such as the observation operator).

Using the same methodology based on geographical masks than MB2010, Michel *et al.*, 2010 have used forecasts taken from an Ensemble Kalman Filter run at 3 km resolution for the prediction of a special case of convective storm to compute full covariances for q_c and q_r (Dowell and Stensrud, 2008). The balance operator used in WRF 3D-Var has been extended following a similar approach than Berre, 2000 for q, in order to obtain multivariate relationships between q_c and q_r and other CV. For traditional CV, very comparable results than those of MB2010 have been found. Even if hydrometeor errors are probably non-Gaussian distributed, Michel *et al.*, 2010 also show that reasonable and physically meaningful auto-covariances and statistical couplings with other variables (especially with divergence) could be obtain for these variables.

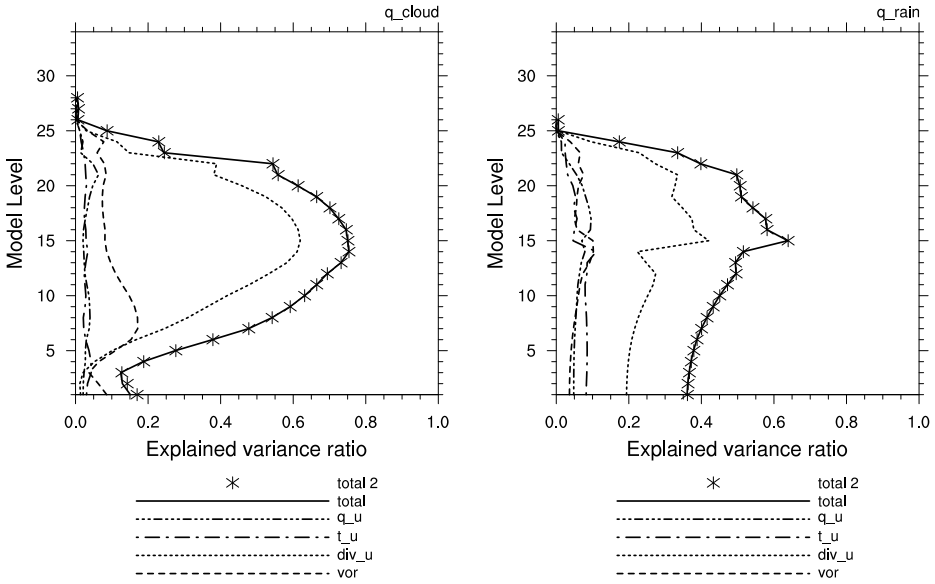

Figure 4: Explained variance ratio of (left) cloud content and (right) rain content as a function of model level, computed over the heavy precipitating areas (from Michel *et al.*, 2010)

Fig. 4 shows for instance that the prevalent coupling for δq_c and δq_r is, as for δq, with (unbalanced) divergence $\delta \eta_u$. This does not mean that coupling with temperature or humidity are small, as a significant part of temperature and humidity is explained by $\delta \eta_u$ (not shown). Covariances between δq_c

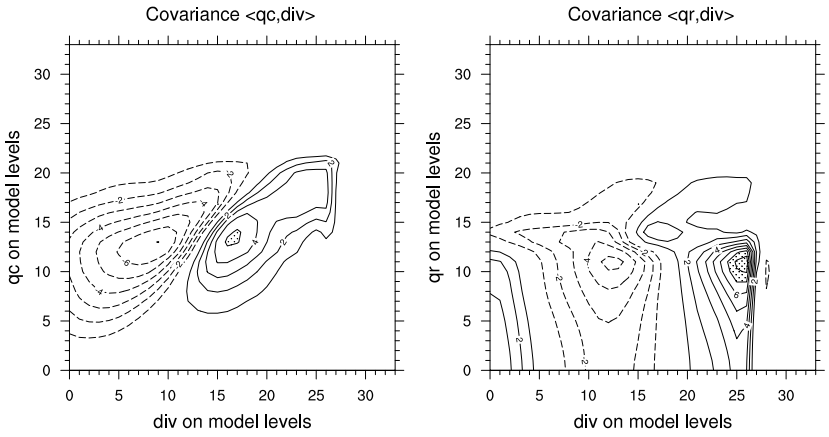

Figure 5: Vertical cross-covariances of background errors between (left) cloud content and divergence (contour interval $2.10^{-4} gkg^{-1}s^{-1}$) and (right) rain content and divergence (contour interval $10^{-4} gkg^{-1}s^{-1}$), computed over the heavy precipitating areas (from Michel *et al.*, 2010).

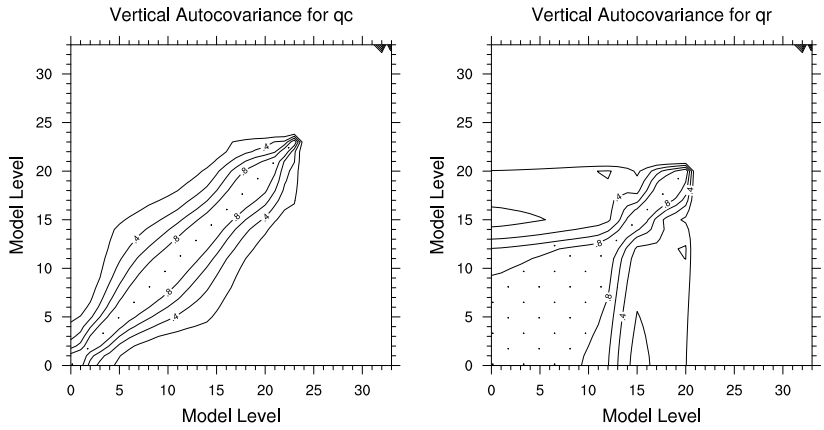

Figure 6: Average vertical auto-correlations for (left) cloud and (right) rain water content in heavy-precipitating areas as a function of model vertical level (contour interval 0.2). From Michel *et al.*, 2010.

and divergence $\delta\eta$ shares the same structures than covariances between δq and $\delta\eta$, also observed in MB2010, but translated vertically: a positive δq_c error is linked with low level convergence and higher level divergence in convective clouds (Fig. 5, left panel). Covariances between δq_r and $\delta\eta$ is more complex and displays structures that depend on level of free convection and freezing level height (Fig. 5, right panel). Finally, The structure of vertical auto-correlations for δq_c and δq_r shares some similarities with the results of Amerault and Zou, 2006, despite different models, methodologies and weather phenomenons. There is more vertical mixing of δq_c error within clouds because of the explicit convection, δq_r is vertically highly correlated under a certain level, which is the direct consequence of rain falling (Fig. 6).

5 How to get clouds- and precipitation-dependent statistics?

The previous results suggest that, compared with current operational formulation of **B**, the structures of forecast errors linked to clouds and precipitation are differently balanced, strongly inhomogeneous and flow dependent. Some leads to better consider those features in the CVT are proposed in this section.

5.1 Non-linear balances

Following Barker *et al.*, 2004, some flow-dependency can be obtain by modifying the geostrophic-like balance (that relates the balanced part of the wind increment with the balanced part of the mass increment, depending of the formulation of **K**) in order to take into account cyclostrophic terms that are important in regions of strong curvature. For that purpose, $\mathscr{H}$ has been recently replaced in some centres by a more sophisticated analytical Non-Linear Balance Equation (NLBE), linearized around the background (Fisher, 2003).

A similar approach allows to add an analytically balanced divergence from vorticity and temperature, according to the quasi-geostrophic (QG) omega and continuity equations linearized around the background (Fisher, 2003). For high resolution models, Pagé *et al.*, 2007 suggest to revisit this formulation with the introduction of diabatic forcing of balanced vertical motion. The inclusion of these additional terms in CVT formulation is technically difficult and has not been tried yet. It has however to be noted that large scale analytical balance relationships like the NLBE or the QG omega equation can be efficiently relaxed in clouds and precipitation that are characterized by smaller scale error structures when using ECMWF formulation that is based on scale-dependent regression coefficients. The proportional deviation from linear geostrophic balance with the intensity of precipitation, as shown by Caron and Fillion, 2010, could also be alleviated thanks to this scale dependency.

5.2 Ensemble flow-dependent B

The basic principle of ensemble variational data assimilation (EnVar) is to include partially information from a flow-dependent **B** matrix, computed from daily runs of an ensemble assimilation, into 3 or 4D-Var, without significant change of the existing setup of operational VAR systems. In the case where this **B** is used in the analysis, VAR systems share with the ensemble Kalman Filter (EnKF) the ability to represent flow-dependent error covariances. The difference lies in the fact that the EnKF uses a sample **B** matrix that is filtered afterwards (mainly through tapering, also called "localisation"), whereas the CVT formalism more or less directly filters the noise (by estimating **B** in a certain basis). Ensembles with perturbed observations allow to simulate the evolution of the model state errors both in the ensemble (Houtekamer *et al.*, 1996) and variational (Fisher, 2003) frameworks. Cycled perturbed analyses based, on one hand, on explicit observation perturbations (representative of observation errors), and on the other hand, on background perturbations which are either fully implicit (i.e integrated from the previous cycle) or partly explicit (to represent model error contribution) are used for that purpose.

Although very attractive, this method is computationally expensive (especially for CRM), as the representativeness of the forecast errors pdf depend on the number of members of the ensemble. A solution is to consider an ensemble with few members (eventually computed at lower resolution) and use optimized filtering techniques to reduce sampling error in the retrieval of **B**. Several approaches are currently tested in operational centres:

- Modulation of climatological covariances by spectrally filtered background error standard deviations σ_b (Raynaud *et al.*, 2009) and wavelet correlations (Fisher, 2003; Pannekoucke *et al.*, 2007) or anisotropic recursive filters (Sato *et al.*, 2009). At Météo-France, the global ARPEGE 4D-Var

already uses operationally these "σ_bs of the day" for all CV deduced from 6 perturbed global members with 4DVar (Berre, 2009).

- CVT in ensemble sub-space (En3DVar or En4DVar (Lorenc, 2003; Liu *et al.*, 2008; Buehner, 2008)) using localizations with Schur operators to reduce the analysis noise. Other spatialization methods, such as wavelets, diffusion operators or recursive filters, could be used for that purpose.

5.3 Heterogeneous covariances

The formalism of heterogeneous **B** can be found in MB2010. Following an initial idea by Courtier *et al.*, 1998, it consists in expressing in the CVT framework the analysis increment as a linear combination of two terms, each term corresponding to areas characterized by different meteorological behaviors (e.g rainy and non-rainy areas):

$$\delta \mathbf{x} = \mathbf{F}^{1/2}\mathbf{B}_1^{1/2}\chi_1 + \mathbf{G}^{1/2}\mathbf{B}_2^{2}\chi_2 \tag{3}$$

where the complementary **F** and **G** operators define the spatial locations where $\mathbf{B_1}$ and $\mathbf{B_2}$ are applied respectively. $\mathbf{B_1}$ and $\mathbf{B_2}$ can be retrieved by applying geographical masks on differences of forecast coming from an ensemble assimilation. As displayed in Fig. 7, such method allows to produce analysis increments that are differently balanced and spatially spread accordingly to the error covariances that are representative of the weather type in which observations are located. This figure displays for instance very different structures that reflect the strong differences of forecast errors in precipitating and non-precipitating areas discussed in section 3.1. Test on real cases are currently under progress for fog in the AROME 3D-Var framework, using the specific statistics of section 3.2 that allow to confine increments due to ground-based measurements in the first levels of the model. The technique does not however solve the maybe more fundamental problem of rain v.s. no rain that may occur because of displacement errors (Auligné *et al.*, 2010).

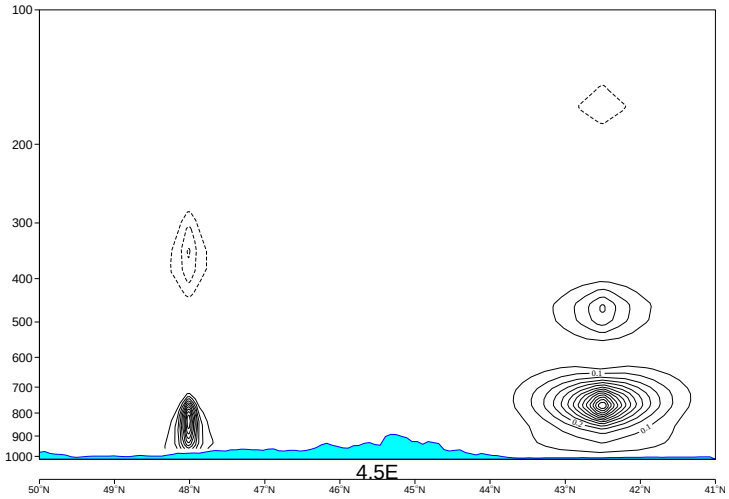

Figure 7: Vertical cross-sections of temperature increments (isocontours every 5.10^{-2} K), generated by four pseudo-observations located at 800 and 500 hPa and characterized by a -30% innovations of relative humidity. These increments have been obtained using the heterogeneous **B** matrix formulation and considering the northern half of the domain as rainy (from Montmerle and Berre, 2010).

6 Conclusion

This paper uses geographical masks to compute CRM background error statistics in specific areas. The masks depend on predicted cloud or rain water content. The covariances are computed for traditional CV in convection and in fog cases, but also for hydrometeor variables for a particular convective case. Overall, very different statistics have been obtained for those areas compared to their counterparts (i.e areas without precipitations or without fog) and to operational climatological covariances. In clouds and precipitation, the coupling between humidity and divergence seems predominant, the horizontal correlation lengths are shorter and the vertical correlations reflect the cloud vertical extension due to convection. Meaningful covariances have been also found for q_c and q_r: strong coupling with divergence, even shorter correlation lengths and structures that may depend on typical cloud characteristics such as the free convection or the freezing levels. Considering such multivariate relationships in the balance operator would allow to map q_c and q_r increments onto the full set of variables used by the nonlinear NWP model, even if a simpler linear observation operator is used.

These results suggest that operational formulations of **B** misrepresent forecast errors in clouds and precipitation, which implies that the assimilation of observations performed in those regions (e.g Doppler radars or profiles deduced from microwave imager measurements) may be far from optimal. This has two implications: localization can be made more restrictive in ensemble-based DA methods such as EnKF or EnVar, and VAR formulations at cloud scales should be able to represent these major inhomogeneities. To represent these inhomogeneities, nonlinear balance relationships could include diabatic processes, and covariances could incorporate some flow-dependency in their formulations, i) by modulating climatological values with covariances extracted and filtered from an ensemble assimilation, ii) by considering forecast error information from an ensemble in the CV after an appropriate filtering of the sampling noise, iii) by computing covariances specifically for clouds and precipitation and by using these covariances in the heterogeneous **B** framework.

References

Amerault, C. and X. Zou (2006). Comparison of model-produced and observed microwave radiances and estimation of background error covariances for hydrometeor variables within hurricanes. *Mon. Wea. Rev. 134*, 745?758.

Auligné, T., A. Lorenc, Y. Michel, T. Montmerle, A. Jones, M. Hu, and J. Dudhia (2010). Summary of the cloud analysis workshop, Sept 1-3, NCAR, Boulder, CO, USA. *Bulletin of Amer. Meteor. Soc.*.

Bannister, R. (2008). A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances. *Quart. J. Roy. Meteor. Soc. 134*, 1951–1970.

Bannister, R. (2008). A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics. *Quart. J. Roy. Meteor. Soc. 134*, 1971–1996.

Barker, D., W. Huang, Y. Guo, A. B. Q.N., and Xiao (2004). A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Mon. Wea. Rev. 132*, 897–914.

Berre, L. (2000). Estimation of synoptic and mesoscale forecast error cavariances in a limited area model. *Mon. Wea. Rev. 128*, 644–667.

Berre, L. (2009). Use of ensemble assimilation to represent flow-dependent B in 3d/4d-var. In *8th International Workshop on Adjoint Applications in Dynamic Meteorology 18-22 May 2009, Tannersville, PA, USA*.

Buehner, M. (2008). Towards an improved use of flow-dependent background error covariances in a variational data assimilation system. In *Banff, Canada. February 3-8*.

Caron, J.-F. and L. Fillion (2010). An examination of background error correlations between mass and rotational wind over precipitation regions. (early online realease). *Mon. Wea. Rev. 138*, 563–578.

Chevallier, F., P. Bauer, J.-F. Mahfouf, and J. Morcrette (2002). Variational retrieval of cloud liquid profile from atovs observations. *Quart. J. Roy. Meteor. Soc. 128*, 2511–2526.

Courtier, P., E. Andersson, W. Heckley, J. Pailleux, D. Vasiljevic, , A. Hollingsworth, M. Fisher, and F. Rabier (1998). The ECMWF implementation of three-dimensional variational assimilation (3D-var). I: Formulation. *Quart. J. Roy. Meteor. Soc. 124*, 1783–1807.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth (1994). A strategy for operational implementation of 4d-var using an incremental approach. *Quart. J. Roy. Meteor. Soc. 120*, 1367–1387.

Daley, R. (1991). *Atmospheric data analysis*. Cambridge University Press, pp. 460.

Dee, D. (2005). Bias and data assimilation. *Quart. J. Roy. Meteor. Soc. 131*, 3323–3343.

Derber, J. and F. Bouttier (1999). A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus 51A*, 195–221.

Dowell, D. and D. Stensrud (2008, October). Ensemble Forecasts of Severe Convective Storms. 24th AMS Conference on Severe Local Storms.

Errico, R., G. Ohring, P. Bauer, B. Ferrier, J.-F. Mahfouf, J. Turk, and F. Weng (2007). Assimilation of satellite clouds and precipitation observations in NWP models: Introduction to the JAS special Collection. *J. Atmos. Sci. 64*, 3737–3741.

Fischer, C., T. Montmerle, L. Berre, L. Auger, and S. Stefanescu (2005). An overview of the variational assimilation in the ALADIN/France numerical weather-prediction system wave–driven circulation of the mesosphere. *Quart. J. Roy. Meteor. Soc. 131*, 3477–3492.

Fisher, M. (2003). Background error modelling. In *Proceedings of the ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, pp. 45–63.

Guidard, V. and D. Tzanos (2007). Analysis of fog probability from a combination of satellite and ground observation data. *Pure and Applied Geophysics 164*, 1207–1220.

Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritchie, and H. L. Mitchell (1996). A system simulation approach to ensemble prediction. *Mon. Wea. Rev. 124*, 1225–1242.

Ingleby, N. (2001). The statistical structure of forecast errors and its representation in The Met. Office Global 3-D Variational Data Assimilation Scheme. *Quart. J. Roy. Meteor. Soc. 127*, 209–231.

Liu, C., Q. Xiao, and B. Wang (2008). An ensemble-based fourdimensional variational data assimilation scheme: Part I: Technical formulation and preliminary test. *Mon. Wea. Rev. 136*, 3363–3373.

Lopez, P. and P. Bauer (2007). 1d+4dvar assimilation of NCEP stage-IV radar and gauge hourly precipitation data at ECMWF. *Mon. Wea. Rev. 135*, 2506–2524.

Lorenc, A. (2003). The potential of the ensemble kalman filter for NWP - a comparison with 4d-var. *Quart. J. Roy. Meteor. Soc. 129*, 3183–3203.

Lorenc, A. (2007). Ideas for adding flow-dependence to the Met Office VAR system. In *ECMWF Workshop on Flow-dependent aspecyts of data assimilation, 11-13 June 2007*, pp. 1–10.

Lorenc, A., I. Roulstone, and A. White (2003). On the choice of control fields in VAR. In *Forecasting Research Tech. Report 419 (available from Met Office, Fitzroy Road, Exeter, Devon, EX8 3PB, UK*.

Marécal, V. and J.-F. Mahfouf (2002). Four-dimensional variational assimilation of total column water vapour in rainy areas. *Mon. Wea. Rev. 130*, 43–58.

Michel, Y. and T. Auligné (2010). Inhomogeneous Background Error Modeling and Estimation over Antarctica. *Mon. Wea. Rev. 138*(6), 2229–2252.

Michel, Y., T. Auligné, and T. Montmerle (2010). Diagnosis of convective scale background error covariances within and outside of precipitating areas (submitted). *Mon. Wea. Rev.*.

Montmerle, T. and L. Berre (2010). Diagnosis and formulation of heterogeneous background error covariances at mesoscale (in press). *Quart. J. Roy. Meteor. Soc.*.

Moreau, E., P. Bauer, and F. Chevallier (2003). Variational retrieval of rain profiles from spaceborne passive microwave radiance observations. *J. Geophys. Res. 108*, 4521.

Pagé, C., L. Fillion, and P. Zwack (2007). Diagnosing summertime mesoscale vertical motion: implications for atmospheric data assimilation. *Mon. Wea. Rev. 135*, 2076–2094.

Pannekoucke, O., L. Berre, and G. Desroziers (2007). Filtering properties of wavelets for the local background error correlations. *Quart. J. Roy. Meteor. Soc. 133*, 363–379.

Purser, R., W. Wu, D. Parrish, and N. Roberts (2003). Numerical Aspects of the Application of Recursive Filters to Variational Statistical Analysis. Part II: Spatially Inhomogeneous and Anisotropic General Covariances. *Mon. Wea. Rev. 131*, 1536–1548.

Raynaud, L., L. Berre, and G. Desroziers (2009). Objective filtering of ensemble-based background error variances. *Quart. J. Roy. Meteor. Soc. 135*, 1177–1199.

Sato, Y., M. D. Pondeca, R. Purser, and D. Parrish (2009). Ensemble-based background error covariance implementations using spatial recursive filters in NCEP grid-point statistical interpolation system. NCEP Office Note 459, Environmental Modeling Center, Camp Springs, Maryland.

Seity, Y., P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson (2010). The AROME-france convective scale operational model. *submitted to Mon. Wea. Rev.*.

Wu, W., R. Purser, and D. Parrish (2002). Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. *Mon. Wea. Rev. 130*, 2905–2916.