Nonlin. Processes Geophys. Discuss., 2, C510–C525, 2015
www.nonlin-processes-geophys-discuss.net/2/C510/2015/

Nonlinear Processes in Geophysics
Discussions
Open Access

Interactive comment on “Diagnosing non-Gaussianity of forecast and analysis errors in a convective scale model” *by* R. Legrand et al.

R. Legrand et al.

raphael.legrand@meteo.fr

Received and published: 23 October 2015

Dear Referee,

Thank you for those questions and comments. Please find below, our answers and the associated changes added to the manuscript.

Best regards,

Raphaël Legrand, Yann Michel and Thibaut Montmerle

(Reviewer comments are written in black, and authors answers are in blue)

Comments from Referee→(1) There is some confusion in the paper about the roles of

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

linearity and Gaussianity in assimilation. The abstract reads Two common derivations respectively lead to the Kalman filter and to variational approaches. They rely on either assumptions of linearity or assumptions of Gaussianity of the probability density functions of both observation and background errors . Maybe I am mistaken on the authors' intentions, but these sentences mean in effect that the hypotheses of linearity (leading to Kalman filter) and Gaussianity (leading to variational assimilation) are mutually exclusive. They are not. Both Kalman filter and variational assimilation are based on the same linear assumptions (and both are empirically extended to weakly non-linear situations). Under these linear assumptions, they are only two different algorithms that solve the same problem. In addition, they both achieve Bayesian estimation in the case when the errors affecting the data are Gaussian.

More precisely P. 1063, ll. 10-11. ..., up to now operational Numerical Weather Prediction (NWP) has relied on assimilation schemes that are Gaussian .... The authors do not say which assimilation schemes they have in mind, but I presume they mean schemes of the general 'Kalman' form

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_b) \tag{1}$$

where $\mathbf{x}_b$ and $\mathbf{x}_a$ are respectively the background and the analysis, $\mathbf{y}$ is the observation, $\mathbf{H}$ the corresponding (linear) observation operator, the difference $\mathbf{d} \equiv \mathbf{y} - \mathbf{H}\mathbf{x}_b$ being the innovation vector. $\mathbf{K}$ is the gain matrix which, in the context of least variance estimation, is defined as $\mathbf{K} \equiv \mathbf{C}_{zd}\mathbf{C}_{dd}^{-1}$, where $\mathbf{C}_{zd}$ is the cross-covariance matrix of the background error $\mathbf{z} \equiv \mathbf{x} - \mathbf{x}_b$ with the innovation, and is $\mathbf{C}_{dd}$ is the covariance matrix of the innovation itself. I stress there is nothing necessarily 'Gaussian' in Eq. (1) above. That equation can be obtained as defining the Best Linear Unbiased Estimator (BLUE) of $\mathbf{x}$ from $\mathbf{x}_b$ and $\mathbf{y}$, independently of any Gaussian hypothesis. It can also be obtained, also independently of any Gaussian hypothesis, on a principle of maximum entropy. Linearity, on the other hand, is always necessary. Gaussianity is only a 'plus' which, if

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

it comes in addition to linearity, ensures Bayesianity of the estimation.

The authors write (p. 1065, ll. 2-3, efforts [to] be made to improve linear assumptions ... Well, if Gaussianity is obtained at the expense of linearity, this may result in a degradation of the accuracy of the final estimate.

P. 1064, ll. 7-8. It [the 4D-Var algorithm] solves for the most probable state [...] by minimizing a non-quadratic cost-function .... If there are non-linearities and the cost-function is non-quadratic, it is very unlikely that minimizing it will lead to the most probable state. Actually, that is guaranteed only in the linear and Gaussian case.

Please revise all parts of the paper relative to the basic principles of assimilation and to the questions of linearity, Gaussianity and Bayesianity. It must be clear in particular that, among the hypotheses to be made for Kalman filtering and variational assimilation, linearity must come before Gaussianity.
Author's response→ There are two common derivations of the assimilation problem in the literature. The first one actually matches your derivation. It derives the Kalman gain as the best linear unbiased estimate, in the sense of a minimum variance estimate. Then it is possible to derive variational data assimilation as a minimization problem solving for this Kalman gain. The Gaussian assumption then is not necessary but ensure Bayesianity of the assimilation.

The second derivation takes maybe an other step: it starts from the derivation of a maximum likelihood problem using Bayes rules. This is the approach presented in (e.g.) Lorenc (1986), Bannister (2008), Bocquet et al. (2010), and Fisher et al. (2011). The Gaussian hypothesis is done next, leading to a minimization problem of a non-

NPGD
2, C510–C525, 2015

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

linear cost function $\mathcal{J}$ that is given for instance for a 3D-Var as:

$$2\mathcal{J} = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + (\mathbf{y} - \mathcal{H}(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{y} - \mathcal{H}(\mathbf{x})) \quad (2)$$

with **B** and **R** the covariance matrix of background and observations errors. Non-linearities arise from the observation operators $\mathcal{H}$ (in addition to the direct model operator for the 4D-Var). As mentioned by Bocquet et al. (2010) about the minimization of $\mathcal{J}$, instead of using stochastic optimization methods which are intractable for NWP applications, a remedy is to use a succession of quadratic optimization problem with simplified and linearised operators. Fisher et al. (2011) explain the several linearisation as a way to resolve the minimisation problem with "a range of efficient methods". So, with this second derivation, the Gaussianity (or correction of Gaussianity) is seen as the only tractable choice and appears very soon in the $\mathcal{J}$ designing. Then, linearisation is seen as an additional technical assumption leading to better efficiency in the minimization process.

The equivalence between those two possible derivations may be obtain using some kind of EnKF approaches.

Author's changes in manuscript→In the abstract: "In numerical weather prediction, the problem of estimating initial conditions with a variational approach is usually based on a Bayesian framework associated with a Gaussianity assumption of the probability density functions of both observations and background errors. In practice, Gaussianity of errors is tied to linearity, in the sense that a nonlinear model will yield non-Gaussian probability density functions. In this context, standard methods relying on Gaussian assumption may perform poorly.", and at the beginning of the introduction: "In data assimilation, the analysis step may be seen as finding a maximum likelihood of the probability density functions (PDF) of the state x given the available observations y and a background state (usually a short range forecast). Usual Bayesian formulation yields (Kalnay, 2003)".

NPGD

2, C510–C525, 2015

Comments from Referee→(2) The significance of the D'Agostino test, and the interpretation to be given to the results it produces, must be clarified.

I mention first that formulæ (2) and (3) for the skewness G_3 and the kurtosis G_4 are not exact. The denominator in the expression for the variance should be N_s -1, and similar corrections are to be made in the expressions for the third- and fourth-order moments. Author's response→ We don't think that there is a mistake here. According to (e.g.) Thode (2002) p45-46, G_3 and G_4 are using sample moments given as

$$m_k = \frac{1}{N_s} \sum_{i=1}^{N_s} (\mathbf{x}_i - \bar{\mathbf{x}})^k \tag{3}$$

Such definition of the second sample moment is used to compute G_3 and G_4, not the sample (unbiased) variance.
Author's changes in manuscript→We agree with this point.

Comments from Referee→More importantly, the fundamental purpose of the test is the following. For a given ensemble size N_s and exact Gaussianity, by how much can one expect G_3 and G_4 to deviate from their Gaussian values 0 and 3 ?
Author's response→ According to Kendall and Stuart (1977) in case of exact normality G_3 is a zero mean random variable with a variance of

$$\sigma^2(G_3) = \frac{6(N_s - 2)}{(N_s + 1)(N_s + 3)} \tag{4}$$

Asymptotically (N_s large enough), G_3 tends to be normally distributed with a zero mean and a variance of $6/N_s$. In our case for $N_s = 90$, G_3 is not exactly normally distributed and, with a bilateral testing at level 95%, the normality hypothesis is accepted when $-0.494 < G_3 < 0.494$ (Table B5, Thode, 2002).

NPGD
2, C510–C525, 2015

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

As regards G_4, in case of normality G_4 is asymptotically (N_s large enough) normally distributed with a mean of 3 and a variance of $24/N_s$. For finite ensemble size, mean and variance are given as

$$E[G_4] = \frac{3(N_s - 1)}{(N_s + 1)} \tag{5}$$

$$\sigma^2(G_4) = \frac{24N_s(N_s - 2)(N_s - 3)}{(N_s + 1)^2(N_s + 3)(N_s + 5)} \tag{6}$$

In our case for $N_s = 90$, G_4 is not exactly normally distributed and, with a bilateral testing at level 95%, the normality hypothesis is accepted when $2.24 < G_4 < 4.09$ (Table B6, Thode, 2002).

Those bounds are not mentioned in the text since, G_3 and G_4 are not used in this study. As justified below, only $f_3(G_3)$ and $f_4(G_4)$ are shown and analysed.
Author's changes in manuscript→None

Comments from Referee→The authors define transformed skewness and kurtosis $f_3(G_3)$ and $f_4(G_4)$ through formulæ whose significance is obscure (and which would be in my opinion more appropriately put in an appendix than in the main text of the paper). The transformed $f_3(G_3)$ and $f_4(G_4)$ are said to be standard Gaussian (i.e. with expectation 0 and variance 1) if the original variable is Gaussian. For which values of N_s is that statement true (it cannot be for any N_s, in view for instance of a term N_s in several of the formulæ leading to the definition of $f_4(G_4)$)?
Author's response→ If the original variable is Gaussian, the normality of the transformed skewness and kurtosis is valid respectively for any values of $N_s > 8$ and $N_s > 20$ (resp. p48 and p52, Thode, 2002). Thus, the $N_s - 3$ coefficient that we found in P definition is actually coming from the variance of G_4, which is used in Q as a normalization coefficient.

NPGD

2, C510–C525, 2015

Author's changes in manuscript→In section 2.1: "For a Gaussian PDF and N_s higher than 20 (Thode, 2002), $f_3(G_3)$ and $f_4(G_4)$ could be both assumed to follow a Gaussian law with a zero mean and a unity variance.". Fig.1 has been changed to be in accordance with this threshold.

Comments from Referee→The next step is to test the Gaussianity of the transformed $f_3(G_3)$ and $f_4(G_4)$. But what is then the interest of making the test on $f_3(G_3)$ and $f_4(G_4)$ rather than on the raw G_3 and G_4 ? Is it that a possible non-Gaussianity will show up more clearly on the former ? Is so, say it clearly. In any case, explain.
Author's response→We see three main reasons of using $f_3(G_3)$ and $f_4(G_4)$ instead of G_3 and G_4. The first reason is that for $N_s = o(10^3)$, the asymptotic behaviour of G_3 and G_4 is not reached (D'Agostino, 1970; Anscombe and Glynn, 1983). So in order to simplify hypothesis testing, a transformation is needed to transform them as normal random variables. The second reason is that their values weakly depend on N_s. This make possible to compare several studies using different ensemble sizes. The third reason is that $f_3(G_3)$ and $f_4(G_4)$ are both normally distributed. So the role of each of them in a possible deviation from Gaussianity could be compared, and they could be used to build an omnibus test of normality as K^2.
Author's changes in manuscript→None

Comments from Referee→The authors then introduce the parameter K^2 of which they write (p. 1068, ll. 9-10) that it follows an approximate χ^2 distribution with two degrees of freedom. Well, if $f_3(G_3)$ and $f_4(G_4)$ are independent standard Gaussians, K^2 will follow an exact χ^2 distribution with two degrees of freedom (with expectation 2 and variance 4). Is it because G_3 and G_4 are not independent in the first place that the distribution cannot be expected to be an exact χ^2 ?
Author's response→You are right, K^2 is not distributed with an exact χ^2 since G_3 and G_4 are uncorrelated but not independent (p54, Thode, 2002). An other reason is

NPGD
2, C510–C525, 2015

that normality behaviour of G_3 and G_4 is only asymptotic. For those two reasons it is possible to correct critical values of the K^2 test (chapter *Moment* $(\sqrt{b_1}, b_2)$ *techniques*, D'Agostino and Stephens, 1986). For instance, with $N_s = 100$ the critical value is $K^2 = 6.271$ and not 5.991 as for an exact χ^2 distribution.
Author's changes in manuscript→None since it is already mentioned that K^2 is only "approximately" following a χ^2 distribution.

Comments from Referee→It is not clear how the values obtained for $f_3(G_3)$, $f_4(G_4)$ and **K**^{2} must be interpreted. The authors write (p. 1069, last sentence) describing the values of K^2 has the advantage to prevent the results from depending on the chosen confidence level. Which confidence level are you referring to ? A level similar to the one given (p. 1068, l. 11) for N_s = 100 ?
Author's response→ When using hypothesis testing, conclusion of the test is always associated with a confidence level α (usually $1 - \alpha = 95\%$). Critical values X_c of the test are defined according to this level as (for a right-unilateral testing)

$$P(X > X_c) = \alpha \tag{7}$$

Instead of K^2, we could have shown binary result giving "this point is Gaussian or not". But since the critical value is depending on the confidence level of the test, results would have been different when using different confidence level. Moreover we want to see where the NG is the largest, and see structures. That's why we choose to show raw values of K^2.
Author's changes in manuscript→None

Comments from Referee→But that does not say how to interpret the values obtained for K^2 . One could expect that a χ^2 mean value of 2 for K^2 , with a variance of 4, could be interpreted as proof of Gaussianity. And you mention a value of 2.7 (p. 1072, l. 11) as indicating Gaussianity. But Fig. 3a shows values, at all levels and for all

NPGD
2, C510–C525, 2015

variables except q, which are about 4, which seems to indicate significant deviation from Gaussianity. Nevertheless, you write in the conclusion (p. 1076, l. 14) Deviation from Gaussianity for U, V , and T only appears in the boundary layer . All that is confusing.
Author's response→ As it is stated in the text, with unilateral testing at level 95%, the Gaussian hypothesis is rejected for $K^2 > 6.271$, so K^2 value around 4 are small enough to accept the Gaussian hypothesis of the sample tested. Moreover, discussion on K^2 values allow us to compare quantitatively Gaussian behaviour between variables.
Author's changes in manuscript→None

Comments from Referee→A similar remark applies to the parameters $f_3(G_3)$ and $f_4(G_4)$, of which it is not clearly said (except for the large values of $f_3(G_3)$) how they must be interpreted. For instance, how the fact that the values of $f_4(G_4)$ are positive in Fig. 3c must be interpreted ($f_4(G_4)$ clearly does not have the standard Gaussian distribution to be expected if the basic variables are Gaussian)?
Author's response→ Positive values of $f_4(G_4)$ means that distribution tails are heavier than Gaussian distribution, and also a bigger modal peak. Negative values of $f_4(G_4)$ show lighter tails and smaller modal peak.

Despite negative values appear in Fig.5, you are right noticing that $f_4(G_4)$ are in a large part positive. But it doesn't mean that $f_4(G_4)$ does not follow a standard Gaussian distribution. Indeed this conclusion needs the spatial ergodicity assumption which is not straightforward to us (since $f_3(G_3)$ and $f_4(G_4)$ distribution may depend on meteorological situation i.e spatial inhomogeneity). In this study we would simply test the normality of $f_4(G_4)$ with an hypothesis testing (see if $f_4(G_4)$ is larger than a critical value). An other way to test the normality of $f_4(G_4)$, would be to look at the distribution of an ensemble of $f_4(G_4)$ for each grid point. But this is costly since it

NPGD
2, C510–C525, 2015

Interactive Comment

needs an ensemble of ensembles.
Author's changes in manuscript→Added in section 2.1: "While positive (negative) values of $f_3(G_3)$ point out distributions with a median smaller (higher) than the mean and with a longer right (left) tail, positive (negative) values of $f_4(G_4)$ mean that distribution tails are heavier (lighter) than Gaussian distribution's, with also a bigger (smaller) modal peak."

Comments from Referee→All those aspects must be clarified. In particular, explain in what it is better to use the parameters $f_3(G_3)$ and $f_4(G_4)$ (and K^2) rather than the raw diagnostics G_3 and G_4 . And explain better how the values found for $f_3(G_3)$, $f_4(G_4)$ and K^2 must be interpreted (see also comment 4 below).
Author's response→ We hope that previous answers are clarifying the use and interpretation of $f_3(G_3)$, $f_4(G_4)$, and K^2.
Author's changes in manuscript→None

Comments from Referee→(3) Subsection 4.2.1 and associated Fig. 9. You present diagnostics for control variables, and particularly vorticity and divergence and for a 3-hour forecast. You have shown previously that, for other variables, the analysis ensembles are more Gaussian than the forecast ensembles. I suggest you also present diagnostics for the analysed control variables.
Author's response→ To be consistent with the Fig.7, diagnostics of NG for vorticity ζ and total divergence η have been computed before and after the assimilation step of the 4^{th} November 2011 at 03:00. Results are shown below in Fig.1 of this comment (caption: "Vertical profiles of Vorticity ζ and total Divergence η before ("background") and after ("analysis") the assimilation process. Results are computed from the ensemble of 90 background and analysis states valid the 4^{th} November 2011 at 03:00."). While NG for levels higher than 900hPa are almost unchanged, the averaged K^2 of ζ and η is systematically lower for the analysis than the background state in

NPGD
2, C510–C525, 2015

the boundary layer. However the order of magnitude of the decrease is much smaller than for T and q, thus the dynamical variables ζ and η remain by far much more non-Gaussian than T and q.(Those conclusions have been added in the text).
Author's changes in manuscript→added in section 4.2.1: "As for Fig.7, diagnostics of NG for vorticity ζ and total divergence η have been computed before and after the assimilation step (not shown). While NG of levels higher than 900hPa are almost unchanged, the averaged K^2 of ζ and η is systematically lower for the analysis than the background state in the boundary layer. However the order of magnitude of the decrease is much smaller than for T and q, and the dynamical variables ζ and η remain by far much more non-Gaussian.

Comments from Referee→(4) Subsection 4.2.1. You write on the basis if Fig.9 that the vorticity, unlike the wind components, is strongly non-Gaussian. This is what comparison of Figs 3 and 9 may suggest, but the vorticity is a linear function of the wind components, and cannot be as such be less Gaussian than those components. This requires clarification.
Author's response→This remark is similar to that made by reviewer 1 (and 2) in his second point. In order to study relative impact on NG of heteroscedasticity and spatial derivatives, NG diagnostics have been computed for the temperature T, which is a nearly Gaussian field (cf Fig. 3a), for the temperature normalized by its standard deviation $\frac{T}{\sigma_T}$, and for their respective first-order spatial derivatives ($\frac{\partial T}{\partial x}$ and $\frac{\partial \frac{T}{\sigma_T}}{\partial x}$). Results are shown and explained in the answer to reviewer 1. A comment has also been added in the manuscript.
Author's changes in manuscript→Same as for reviewer 1.

Comments from Referee→(5) Concerning also Fig. 9, you write that the unbalanced divergence η_u , like vorticity, is strongly non-Gaussian, while the variables T_u and q_u display much more Gaussian profiles. Well, according to the caption of Fig. 9, it is

T_u which, in addition to vorticity, shows large values of K^2 , while η_u shows smaller values. Is there an error in the caption, or what ?
Author's response→There was indeed an error in the caption. T was inverted with η_u. This has been corrected.
Author's changes in manuscript→Caption of Fig.9 corrected.

Comments from Referee→And, speaking of vorticity, you use the Greek letter ξ (pronounced xi) to denote it. The usual notation is ζ (pronounced zeta). I suggest you follow the established practice.
Author's response→We agree with this point.
Author's changes in manuscript→ This has been corrected in the manuscript.

Comments from Referee→(6). P. 1073, ll. 9-11. For q, NG is mainly found in "cloudy" areas, [...] with two peaks around 900 and 700 hPa. According to Fig. 6a, there is a much more marked peak in the layer 100-300 hPa.
Author's response→ As noted in section 3.2, largest NG for q in high troposphere appear where q is almost non-existent. As it is stated in the text, those large values of NG have then to be taken with caution.
Author's changes in manuscript→None

Comments from Referee→(7) Abstract, ll. 18-19, The mass control variables used in our data assimilation, namely vorticity and divergence. Well, vorticity and divergence are not mass variables (check for other possible similar mistakes elsewhere in the paper)
Author's response→We agree with this point.
Author's changes in manuscript→This has been corrected in the manuscript as "dynamical control variables".

NPGD

2, C510–C525, 2015

Interactive Comment

Comments from Referee→(8) P. 1066, ll. 3-4, Positive (negative) values are associated with a mode of the PDF smaller (larger) than its mean. This statement may not be true of the mode of the distribution (which can be arbitrarily modified with infinitesimal change to the distribution), but is true of its median.
Author's response→We agree with this point.
Author's changes in manuscript→ This has been corrected in the manuscript.

Comments from Referee→(9) And there are erroneous statements concerning the relationship between skewness and tails pp. 1071, l. 13, and 1072, l. 1.
Author's response→We agree with this point.
Author's changes in manuscript→ This has been corrected in the manuscript.

Comments from Referee→(10) P. 1068, l. 11, what is unilateral testing ?
Author's response→ When testing an hypothesis (e.g.) H_0 : "$X_{observed}$ is sampled from a Gaussian random variable", with a confidence level α the test is right-tailed unilateral if H_0 is rejected when $P(X > X_{observed}) < \alpha$.
Author's changes in manuscript→ The "right-tailed" adjective has been added in the manuscript.

Comments from Referee→(11) P. 1073 and 1076, ll. 11 and 19, forecast terms ranges
Author's response→We agree with this point.
Author's changes in manuscript→ This has been changed

Comments from Referee→(12) P. 1064, l. 15, Laroche and Pierre, 1998. Do you mean Laroche and Gauthier ?
Author's response→We agree with this point.
Author's changes in manuscript→ This has been changed

Comments from Referee→(13) P. 1064, l. 20 (and elsewhere). The proper spelling is Järvinen (with a diaeresis)
Author's response→We agree with this point.
Author's changes in manuscript→ This has been changed

Comments from Referee→(14) P. 1077, ll. 7-8, ... does not include model error, neither in the analysis nor in the forecast steps (what you write is analogous to writing in French Je n'ai pas vu personne)
Author's response→We agree with this point.
Author's changes in manuscript→ This has been changed

Comments from Referee→(15) P. 1076, l. 3, *below* the tropopause
Author's response→We agree with this point.
Author's changes in manuscript→ This has been changed

References

Anscombe, Francis J. and Glynn, William J.: Distribution of the kurtosis statistic b2 for normal samples, Biometrika, 70 (1), 227–234, 1983.

Bannister, Ross N.:A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances, Quarterly Journal of the Royal Meteorological Society, 134 (637),1951–1970,2008.

Bocquet, M., Pires, C. A., and Wu, L.: Beyond Gaussian statistical modeling in geophysical data assimilation, Mon. Weather Rev., 138, 2997–3023, 2010.

D'Agostino, R. B.: Transformation to normality of the null distribution of G1, Biometrika, 57, 679–681, 1970.

D'Agostino, Ralph B. and Stephens, Michael A.: Goodness-of-fit techniques, Marcel Dekker, 1986.
Fisher, M. and Tremolet, Y. and Auvinen, H. and Tan, D. and Poli, P.:Weak-constraint and long-window 4D-Var, ECMWF Techn. Rep, 655, 2011.
Holm, E., Andersson, E., Beljaars, A., Lopez, P., Mahfouf, J.-F., Simmons, A., and Thépaut, J.-N.: Assimilation and modelling of the hydrological cycle: ECMWF's status and plans, European Centre for Medium-Range Weather Forecasts (ECMWF), Technical Memorandum, 2002.
Kalnay, E.: Atmospheric modeling, data assimilation, and predictability, Cambridge University Press, 2003.
Kendall, M. and Stuart, A.: The advanced theory of statistics. Vol. 1: Distribution theory, 4th ed., 1977.
Lawson, W. G. and Hansen, J. A.: Alignment error models and ensemble-based data assimilation, Mon. Weather Rev., 133, 1687–1709, 2005.
Lorenc, A. C.: Analysis methods for numerical weather prediction, Q. J. Roy. Meteorol. Soc., 112, 1177–1194, 1986.
Ravela, S., Emanuel, K., and McLaughlin, D.: Data assimilation by field alignment, Physica D, 230, 127–145, 2007.
Thode, Henry C.:Testing for normality, volume 164, Marcel Dekker, CRC Press, 2002.
Trémolet, Yannick: Incremental 4D-Var convergence study, Tellus A, 59 (5), 706–718, 2007.

Interactive comment on Nonlin. Processes Geophys. Discuss., 2, 1061, 2015.

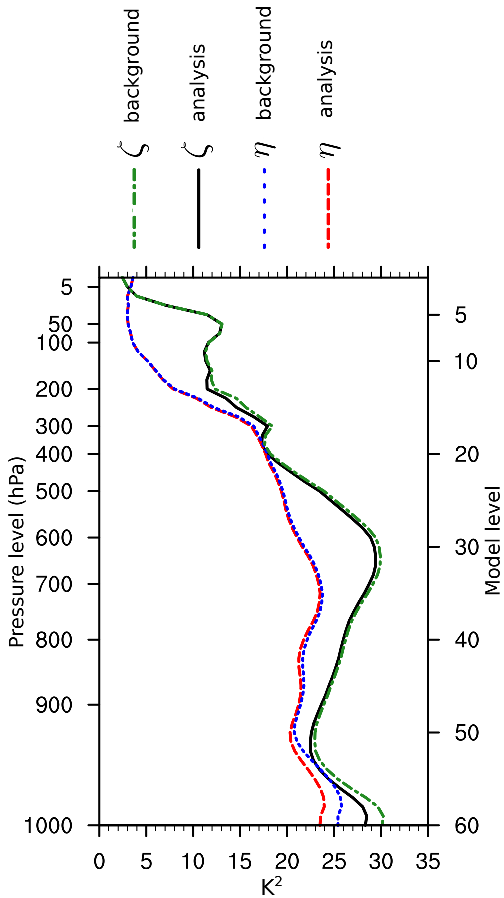

Fig. 1. Vertical profiles of Vorticity and total Divergence before ("background") and after ("analysis") the assimilation process (see text).