School of Mathematical
and Physical Sciences

Department of Mathematics and Statistics

Preprint MPS-2012-05

27 February 2012

On Error Dynamics and Instability in Data Assimilation

by

Roland W.E. Potthast, Alexander J.F. Moodey, Amos S. Lawless and Peter Jan van Leeuwen

On Error Dynamics and Instability in Data Assimilation

Roland W.E. Potthast[1,2], Alexander J.F. Moodey[2], Amos S. Lawless[2], Peter Jan van Leeuwen[2]

[1] German Meteorological Service - Deutscher Wetterdienst Research and Development, Frankfurter Strasse 135, 63067 O enbach, Germany

[2] University of Reading, School of Mathematical and Physical Sciences,

state φ_k at time t_k into a new state φ_{k+1} at time t_{k+1}, where we assume that we consider time steps $t_0 < t_1 < t_2 < \ldots$. Let Y be our measurement space, which we assume to be fixed within this work. For the data assimilation task we are given measurements $f_k \in Y$ at time t_k. Also, we assume that we know some initial guess $\varphi_0 \in X$ at time t_0. The task of data assimilation is to successively calculate some analysis $\varphi_k^{(a)}$ at time t_k which is estimating the true system state $\varphi_k^{(true)}$ at time t_k[LLD06]. Usually, we are also interested in an estimate for the uncertainty of the estimate $\varphi^{(a)}$, or an estimate for the analysis error

$$e_k^{(a)} := \varphi_k^{(a)} - \varphi_k^{(true)}, \quad k \in \mathbb{N}_0. \tag{1.1}$$

Here, we are interested in Hilbert-space type error estimates for the finite-dimensional as well as the *infinite-dimensional* case, which provides a good model for high-dimensional systems. We call a data assimilation system stable, if $e_k^{(a)}$ remains bounded by some constant $C > 0$ for $k \to \infty$. If $e_k^{(a)} \to \infty$ for $k \to \infty$, we call it *unstable*. When we have an estimate

$$f_k \quad f_k^{(true)}$$

The nonlinear case is beyond the scope of this work. The state space X is assumed to be a Hilbert space with scalar product $h\,;\; i_X$ and the measurement space Y is a Hilbert space with scalar product $h\,;\; i_Y$. Often, we will drop the indices X or Y, respectively, since it is usually clear which scalar product is used.

At every time step t_k, $k\ 2$ N

where the difference in the last round bracket can be decomposed as the first difference in (3.3), i.e.

$$\epsilon^{(true)}_{k+1} - \epsilon^{(b)}_{k+1} = M_k \left(\epsilon^{(true)}_k - \epsilon^{(a)}_k\right) - \left(M_k - M^{(true)}_k\right) \epsilon^{(true)}_k. \tag{3.4}$$

This leads to

$$\begin{aligned} e^{(a)}_{k+1} = {} & \overbrace{(I - R\,H)}^{\text{reconstruction error}} \overbrace{\left(M_k e^{(a)}_k + \left(M_k - M^{(true)}_k\right) \epsilon^{(true)}_k \right)}^{\text{propagation of previous error + model error}} \\ & + \overbrace{R\, f^{(\;)}_{k+1}}^{\text{data error influence}} \\ & + \overbrace{R\left(H^{(true)} - H\right) \epsilon^{(true)}_{k+1}}^{\text{observation operator error}} \end{aligned} \tag{3.5}$$

We first study the simplified situation where we assume that our computational model is correct, i.e. $M^{(true)} = M$, TJ/FrI [(ie574 0 Td(f 15.946z -13.27 Td [(g -397.258 -34.658 -34.658

$$
\begin{aligned}
&= \quad {}^{k+1}e_0^{(a)} + \left(\sum_{\ell=1}^{k} {}^{\ell}\, e^{(\ )} \right) + e^{(\ )} \\
&= \quad {}^{k+1}e_0^{(a)} + \left(\sum_{\ell=0}^{k} {}^{\ell}\, e^{(\ )} \right); \qquad (3.11)
\end{aligned}
$$

which is the formula for k replaced by $k+1$ and the induction is shown. Then, in the case where *bounded by* > 0, $:= jj \quad jj$ *and* $:= jjR \quad jj$, *then the analysis error* $e_k^{(a)}$ *is estimated by*

$$e_k^{(a)} \quad {}^k \quad e_0^{(a)} \quad + \quad \left(\sum_{` = 0}^{k\,1} \quad ` \right) : \qquad (3.17)$$

If < 1, *we have*

$$e_k^{(a)} \quad {}^k \quad e_0^{(a)} \quad + \frac{1 \qquad {}^k}{1} \qquad (3.18)$$

such that

$$\limsup_{k\,!\,1} \quad e_k^{(a)} \qquad \frac{jjR \quad jj}{1} : \qquad (3.19)$$

We now come to the g1[(W)82(e)-298(e3a.1(8k))]TJe)-itt4 *such that*

$$\limsup_{k \to \infty} \| e_k^{(a)} \| \le \frac{\|R_\alpha\| \, + }{1 - \rho}. \tag{3.30}$$

4. Stabilizing Cycled Data Assimilation

We have described error estimates for the analysis error of cycled data assimilation in Theorems 3.3 and 3.4. One key assumption to keep the error bounded is the estimate $\rho = \|\Lambda\| < 1$. This condition means that $\|(I - R_\alpha H)M\| < 1$, i.e. the model dynamics is not increasing the error stronger than the regularized reconstruction can reduce it. Here, we investigate trace-class model operators M to show that we can obtain this condition for appropriately chosen regularization parameter α.

Lemma 4.1 *For a finite dimensional state space $X = \mathbb{R}^n$ and injective operators H, we can always make $N := I - R_\alpha H$ arbitrarily small in norm.*

Proof. This is due to the fact that $H^* H$ is self-adjoint and thus it has a complete basis $\varphi^{(1)}, \ldots, \varphi^{(n)}$ of eigenvectors with eigenvalues $\lambda_j > 0$ for $j = 1, \ldots, n$. We set

$$c := \min_{j=1,\ldots,n} |\lambda_j| > 0. \tag{4.1}$$

We can transform N into

$$N = I - R_\alpha H = I - (\alpha I + H^* H)^{-1} H^* H = \alpha (\alpha I + H^* H)^{-1} \tag{4.2}$$

and estimate

$$\|N\| \le \max_{j=1,\ldots,n} \frac{\alpha}{\alpha + \lambda_j} \le \frac{\alpha}{\alpha + c}. \tag{4.3}$$

Given c we can always choose $\alpha > 0$ sufficiently small such that the norm $\|N\|$ of N is arbitrarily small.

Remark. As a consequence of the previous lemma, given M we can choose $\alpha > 0$ such that $\|\Lambda\| \le \|N\| \|M\| < 1$.

For the case of infinite dimensions, which is much closer to realistic situations, we need to take more care, since in general the constant c in the above arguments is zero. Here, we will work out the case where M is a trace-class operator. We first collect notations and set-up our scene for further arguments.

Let $\{\psi_\ell : \ell \in \mathbb{N}\}$ be a complete orthonormal system in X. Then, any vector $\varphi \in X$ can be decomposed into its Fourier sum

$$\varphi = \sum_{\ell=1}^{\infty} \alpha_\ell \psi_\ell \tag{4.4}$$

with $\alpha_\ell := \langle \varphi, \psi_\ell \rangle$ for $\ell \in \mathbb{N}$ $\sum_{\ell=1}$ i f8r h $\sum_{\ell=1}$ ℓ := f0r h $\sum$

which yields (4.11). The second statement is a result of the order of the singular values $_1$ $_2$ $_3$ $:::$ > 0, such that we can choose su ciently small to make each term in (4.11) arbitrary small. Since there are only a nite number of such terms, we obtain the estimate (4.12).

Lemma 4.4 *The norm of the operator* Nj_{X_2} *is given by*

$$jjNj_{X_2}jj = 1 \tag{4.14}$$

for all n 2 N *and* > 0*.*

Proof. Since H is compact, we know that $_n$! 0 for n ! 1. This means that

$$\sup_{` = n+1;:::;1} \frac{}{+ \ {}^2_n} = 1 \tag{4.15}$$

for all n 2 N and all > 0, which implies the norm estimate (4.14).

We now [illegible] -412.022 -27.82 0 Td /(:

This means that given $\epsilon > 0$ there is an $n \in \mathbb{N}$ such that

$$\sum_{\ell=n+1} \sum_{j=1} |M_{j,\ell}|^2 = \sum_{\ell=n+1} |a_\ell|^2 < \epsilon^2. \tag{4.19}$$

This provides an estimate for $||M_2||$ for su ciently l 11.46ently l 11.46en2

(1.3). Here, we have shown that we need su ciently small to keep the error bounded over time. But the error bound will also depend on by the norm

$$jjR\ \ jj \quad \frac{1}{2\,p=}$$

of R , compare equation (4.20). The bound might in fact be rather large. Given there will be some optimal which leads to the best possible error bounds for a given set-up. However, the maximal for which $jj\ \ jj < 1$ is achieved might still cause severe numerical instabilities, such that there is the possibility that practically it is impossible to stabilise the scheme, depending on the particular set-up and operators under consideration.

5. Numerical Examples

Here, we present two numerical examples to demonstrate the theory presented in this work. Firstly we present a simple low-dimensional setup which con rms the theoretical results. Then, we investigate a more realistic system, the 2D Eady model [Ead49], which con rms the practical validity of the above results.

The numerics demonstrate that with a small regularization parameter we can achieve a stable cycled data assimilation scheme. For the simple low-dimensional system we use construct M to be a random $n \quad n$-matrix, using Matlab rand function, giving us a singular system by its SVD. To mimic a kind of trace-class operator, we then manipulated the singular values to decay su ciently strongly.

The interesting case is where some singular values are larger than one, leading to a system with growing modes. Further, we want the rest of h

$||e_k^{(a)}||_{l^2}$ over time t_k with respect to the l^2 norm. In Figure 3 we plot $||e_{10000}^{(a)}||_{l^2}$ as is varied. Here we observe for a fixed time t_{10000} that the analysis error is sensitively dependent on the regularization parameter we choose.

Figure 1. The l^2 norm of the analysis error as the scheme is cycled for the time index k, for a regularization parameter, $\quad = \frac{{}^2_{(o)}}{{}^2_{(b)}} \quad 1.2$.

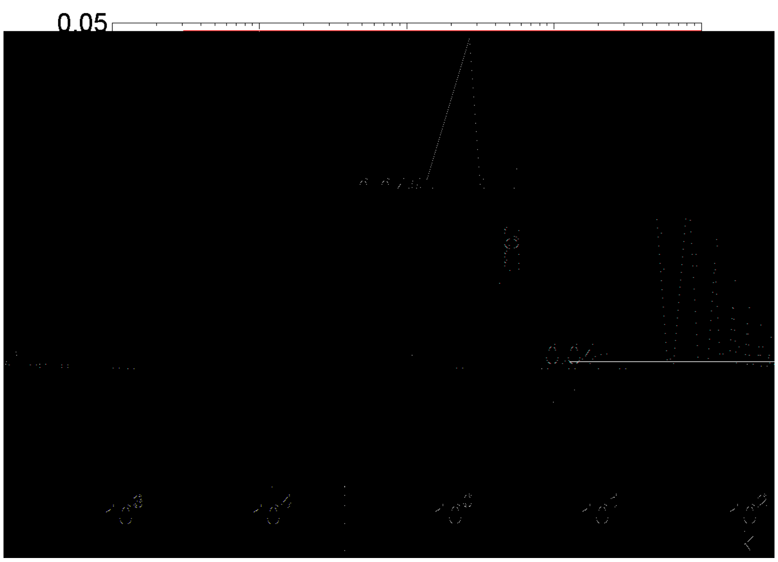

Figure 2. The l^2 norm of the analysis error as the scheme is cycled for the time index k, for a regularization parameter, $\quad 0.8$.

As a second example, we now consider a more realistic system where the model operator M arises from the discretization of a system of partial differential equations.

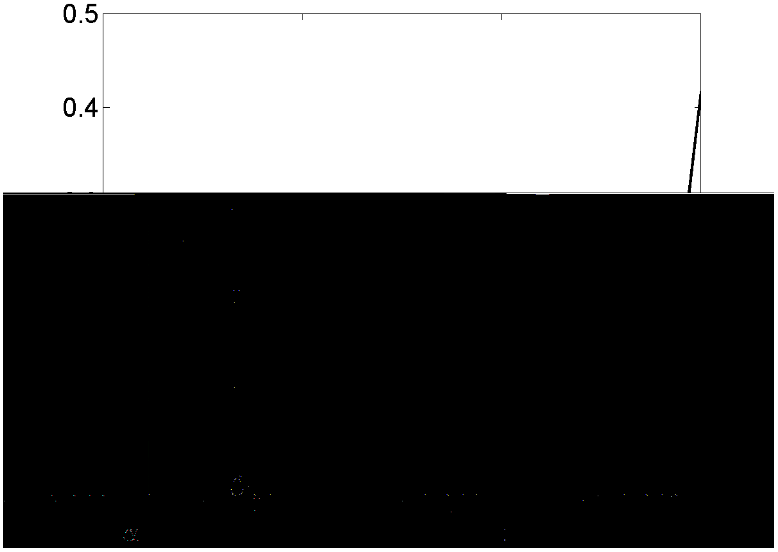

Figure 3. The l^2 norm of the analysis error as the scheme is cycled for a constant time index $k = 10000$, varying the regularization parameter, .

The system we consider is the two-dimensional Eady model, a simple model of atmospheric instability, see [Ead49] for a detailed introduction. The model is defined in the x z plane, with periodic boundary conditions in x and $z \in [-1/2, 1/2]$. The state vector consists of the nondimensional buoyancy b on the upper and lower boundaries and the nondimensional potential vorticity in the interior of the domain. For the current study we assume that the interior potential vorticity is zero and thus we only need to consider the dynamics on the boundaries. The buoyancy is advected along the boundaries forced by the nondimensional streamfunction according to the equation

$$\frac{\partial b}{\partial t} + z\frac{\partial b}{\partial x} = \frac{\partial}{\partial x} \qquad \text{on } z = \pm 1/2 \tag{5.1}$$

where the streamfunction satisfies

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} = 0 \qquad \text{in } z \in [-1/2, 1/2]; \tag{5.2}$$

with boundary conditions

$$\frac{\partial}{\partial z} = b \qquad \text{on } z = \pm 1/2. \tag{5.3}$$

The equations are discretized as described in [Joh03] and [JHN05] using 40 grid points in the horizontal, giving 80 degrees of freedom. The resulting discrete operator M has a maximum eigenvalue of 1.3066.

We simulate the consequence of compact observation operator H with a random 80×80 matrix with the last 5 singular values $_{76:80} = (10^{-6}, 10^{-8}, 10^{-10}, 10^{-12}, 10^{-14})$ respectively. Therefore, H is severely ill-conditioned with a condition number, $= 4.1367 \times 10^{15}$ with respect to the l^2 norm. We set up background and observation standard deviations as follows $_{(b)} = 0.25$ and $_{(o)} = 1$ respectively. Random normally distributed noise is added to the observations with mean, 0 and standard deviation, ${}_{(o)}$. Initially we choose $= \frac{{}^2_{(o)}}{{}^2_{(b)}}$ 16. In Figure 4 we observe that over time t_k the analysis error, $jje_k^{(a)}jj_{l^2}$ blows up with respect to the l^2 norm. Now in Figure 5 we choose a smaller regularization parameter, $= \frac{1}{0:09}$ $11{:}1$, in ating the background error variance, ${}^2_{(b)}$ from $0{:}25^2$ to $0{:}3^2$. Subsequently repeating with the same data, we observe a stable analysis error, $jje_k^{(a)}jj_{l^2}$ over time t_k with respect to the l^2 norm. In Figure 6 we plot $jje^{(a}$

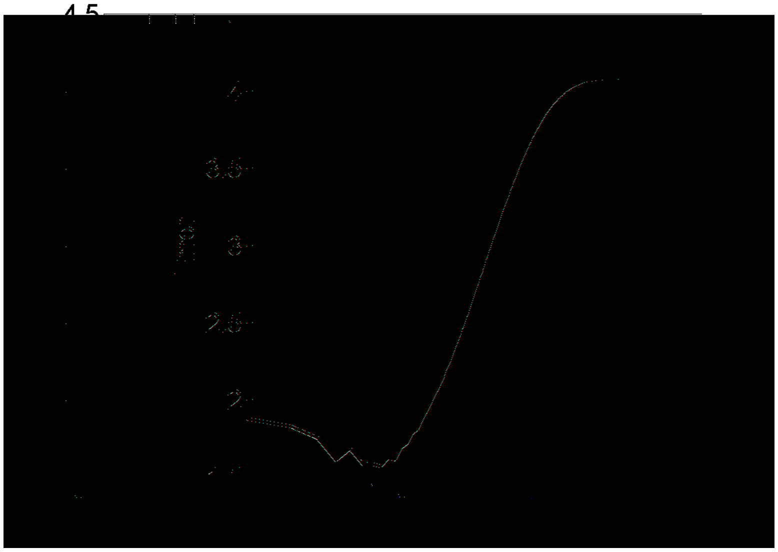

Figure 5. The l^2 norm of the analysis error as the scheme is cycled for the time index k, for a regularization parameter, 11.1.

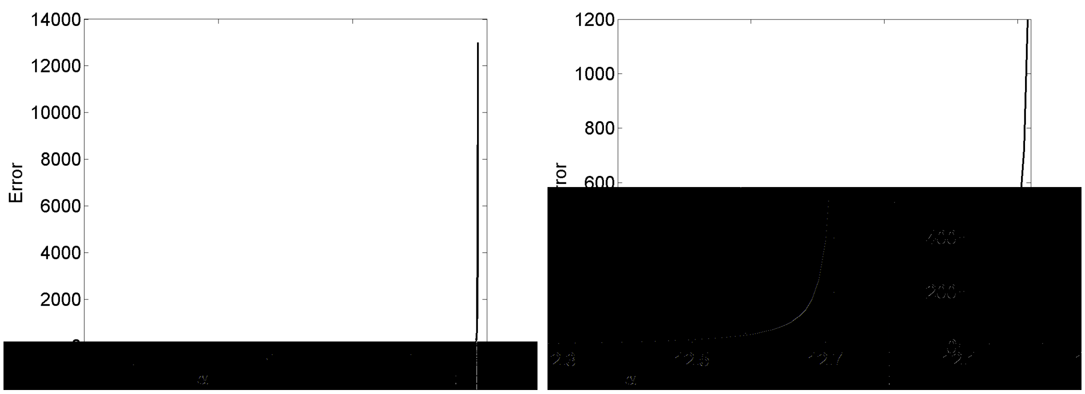

Figure 6. Left: The l^2 norm of the analysis error as the scheme is cycled for a constant time index $k = 10000$, varying the regularization parameter, . Right: Zoomed version.

too large, error which enters the system via the data in general will be amplified, such that the analysis error growth without limits in its long-term behaviour. This growth happens even for an arbitrarily small data error.

In a numerical part we studied simple examples and also applied the theory to a two-dimensional Eady model, a simple model of atmospheric instability. The numerical results confirm and demonstrate the general theory very well.

Acknowledgements. The authors would like to thank the following institutions for their financial support: Deutscher Wetterdienst (DWD), Engineering and Physical Sciences Research Council (EPSRC) and the National Centre for Earth Observation (NCEO, NERC). The Eady model used in Section 5 was developed by C. Johnson, now at the Met Office, and N.K. Nichols and B. J. Hoskins of the University of Reading. We are also grateful to N.K. Nichols for useful discussions during this work.

References

[AM79] B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Prentice Hall, 1979.

[BC85] S. Barnett and R.G. Cameron. *Introduction to mathematical control theory*. Oxford applied mathematics and computing science series. Clarendon Press, 1985.

[BH75] A.E. Bryson and Y.C. Ho. *Applied optimal control: optimization, estimation, and control*. Taylor & Francis, 1975.

[BLL+12] C.E.A. Brett, K.F. Lam, K.J.H. Law, D.S. McCormick, M.R. Scott, and A.M. Stuart. Stability of lters for the navier-stokes equation. Submitted, 2012.

[But69] A.G. Butkovskii. *Theory of optimal control of distributed parameter systems*. American Elsevier, 1969.

[CF03] M.J. Corless and A.E. Frazho. *Linear systems and control: an operator perspective*. Monographs and textbooks in pure and applied mathematics. Marcel Dekker, 2003.

[CP78] R.F. Curtain and A. J. Pritchard. *In nite dimensional linear systems theory*. Lecture notes in control and information sciences. Springer-Verlag, 1978.

[Cul12] M. J. P. Cullen. Analysis of cycled 4d-var with model error. Submitted, 2012.

[DP00] G.E. Dullerud and F.G. Paganini. *A course in robust control theory: a convex approach*. Texts in applied mathematics. Springer, 2000.

[Dul96] G.E. Dullerud. *Control of uncertain sampled-data systems*. Systems & control. Birkhäuser, 1996.

[Ead49] E. T. Eady. Long waves and cyclone waves. *Tellus*, 1:33{52, 1949.

[EHN00] H.W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*. Mathematics and its applications. Kluwer Academic Publishers, 2000.

[Jaz70] A. H. Jazwinski. *Stochastic processes and ltering theory*. Academic Press, 1970.

[JHN05] C. Johnson, B. J. Hoskins, and N. K. Nichols. A singular vector perspective of 4d-var: Filtering and interpolation. *Q. J. R. Meteorol. Soc.*, 131:1{19, 2005.

[Joh03] C. Johnson. *Information content of observations in variational data assimilation*. PhD thesis, Dept. of Meteorology, University of Reading, 2003.

[Kal60] R.E. Kalman. A new approach to linear ltering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, (82 (Series D)):35{45, 1960.

[Kna05] A.W. Knapp. *Advanced real analysis*. Cornerstones Series. Birkhäuser, 2005.

[KS04] J. Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems (Applied Mathematical Sciences)*. Springer, 1st edition, December 2004.

[Lio71] J.L. Lions. *Optimal control of systems described by partial di erential equations*. Springer, 1971.

[LLD06] J. M. Lewis, S. Lakshmivarahan, and S. Dhall. *Dynamic Data Assimilation: A Least Squares Approach*. Cambridge University Press, August 2006.

[Lue64] D. G. Luenberger. Observing the state of a linear system. *Military Electronics, IEEE Transactions on*, 8(2):74 {80, 1964.

[Lue66] D. G. Luenberger. Observers for multivariable systems. *Automatic Control, IEEE Transactions on*, 11(2):190 { 197, 1966.

[MP12] B. A. Marx and R. W. E. Potthast. On instabilites in data assimilation algorithms. 2012.

[O'R83] J. O'Reilly. *Observers for linear systems*. Mathematics in science and engineering. Academic Press, 1983.

[Rug96] W.J. Rugh. *Linear system theory*. Prentice-Hall information and system sciences series. Prentice Hall, 1996.

[Sim06] D. Simon. *Optimal state estimation: Kalman, H [in nity] and nonlinear approaches*. Wiley-Interscience, 2006.