Treatment, Estimation, and Issues with Representation Error Modelling

Elizabeth Satterfield
Naval Research Laboratory, Monterey, California, USA

Thanks to: Dan Hodyss, Dave Kuhl, Bill Campbell, Craig Bishop, Ben Ruston, Nancy Baker and Pat Pauley

4th Workshop on Assimilating Satellite Cloud and Precipitation Observations for NWP
ECMWF, Reading, UK
3-6 February 2020

Motivation

- Representation error arises from an incompatibility between coarse model grids (prior estimate **x**) and observations (**y°**) which observe a **higher resolution state**
- Includes contributions from errors in the **observation operator**
- The best state that the model can represent is a **smoothed version of reality**
- The state that we want to estimate is defined by the model: the pdf of a **discrete model state given the observation**, not the pdf of the **true state given the observations**

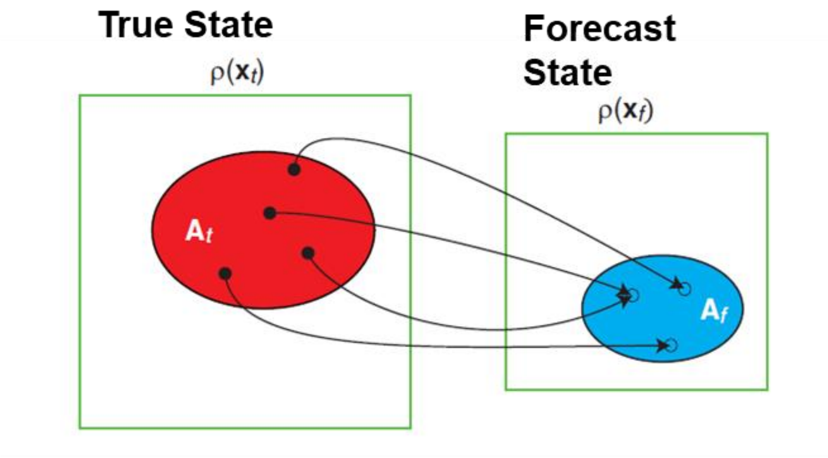


Image credit: Hodyss and Nichols, 2015

Outline of this Talk

1. **Definitions, Sources and Characteristics**
2. **Diagnosing observation error along with the focus on representation error**

   -Overview methods to diagnose observation error and discuss their shortcomings

   -Discuss the ability of these methods to estimate representation error
3. **Discuss current practice for handling representation error**
4. **Methods for incorporating a flow dependent model of representation error**
5. **Concluding Remarks/Discussion**

U.S. NAVAL RESEARCH LABORATORY

Components of Observation Error

- Observation error, broadly speaking, has two components, the **representation error** and the **instrument (or measurement) error**

$$\boxed{\boldsymbol{\varepsilon}^{o} = \boldsymbol{\varepsilon}^{i} + \boldsymbol{\varepsilon}^{r}}$$

- This means that even if the instrument error is small, the observation and the prior can still be quite different.

Sources of Representation Error (Janjić et al. 2018)

1. Error Due to Unresolved Scales and Processes: Discretization and Parameterizations

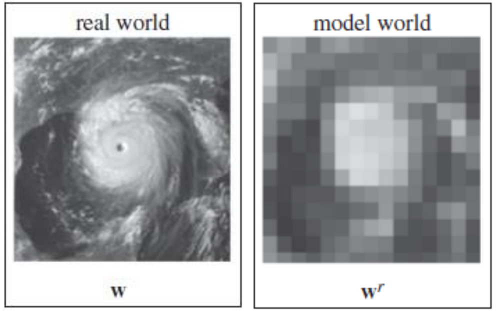
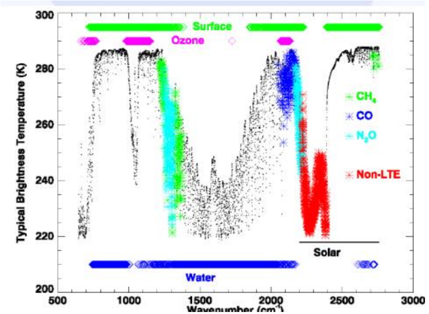


Image credit: Janjić et al 2018

2. Observation Operator: For non-local variables requires the use of complicated observation operators that introduce their own (inseparable) errors.



3. Preprocessing or QC: Imperfections in procedures that remove observations that cannot be adequately modeled or derive a state variable quantity from an observed variable.

Dominant source for cloudy assimilation

Dominant source for clear-sky assimilation

Defining the Components of Observation Error (Satterfield et al. 2017)

- If the NWP model were "perfect", the **instrument error** can be expressed as,

$$\boxed{\boldsymbol{\varepsilon}^{i} = \mathbf{y}^{o} - \mathbf{H}_{H}(\mathbf{x}^{t}{}_{H})}$$

- Where, $\mathbf{H}_H$ is the high-resolution observation operator that maps the high-resolution true state $\mathbf{x}^t$ to the observation location

- The **representation error** can be expressed by considering a (spectral truncation based) smoother $\mathbf{S}_{SC}$ and an imperfect low resolution observation operator.

$$\boxed{\boldsymbol{\varepsilon}^{r} = \mathbf{H}_{H}(\mathbf{x}^{t}{}_{H}) - \hat{\mathbf{H}}_{L}(\mathbf{S}_{SC}\mathbf{x}^{t}{}_{H})}$$

Characteristics of Representation Error

- **Modeling representation error** is not a straightforward problem
- Separation of representation error into **individual components** seldom possible
- It can also be difficult to distinguish between **model error** and **representation error**
- Representation error is **flow dependent** and a function of **model resolution**
- Key, often dominant, contributor to spatially and temporally **correlated observation errors**
- Needs to be considered for applications of **data assimilation**, **observation based verification**, as well as **observation inter-comparison**.

Methods to Recover Observation Error Statistics

1. **Estimate variances using innovation (observation-minus-background) data:**
   - Split innovation(o-b) statistics into observation and background error covariances (e.g. Rutherford, 1972; Höllingsworth and Lonnberg, 1986)
   - Iterative procedures based on assumptions of data assimilation (e.g optimality diagnostic of Desroziers and Ivanov, 2001; Desroziers et al., 2005)
   - Lag innovation statistics (e.g. Berry and Sayer 2013)

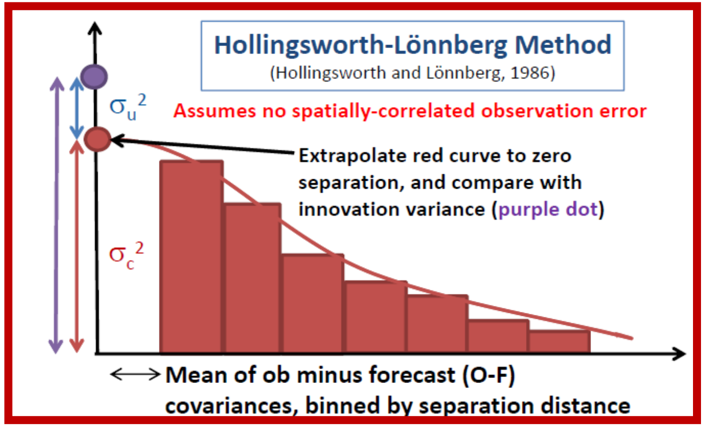
2. **Methods based on estimating a high resolution true state** using observations (e.g. Oke and Sakov, 2007) or model states (e.g. Etherton and Bishop, 2004)

3. Analysis and forecast **Sensitivity diagnostics**, such as those developed at NRL (Daescu and Langland, 2013)

4. **Ensemble Based Approaches** (e.g. Karspeck, 2016)

5. **Maximum likelihood** (e.g. Dee, 1995; Dee et al. 1999) or **Bayesian approaches** (e.g. Stroud and Bengtsson, 2007) to tune parameters of probability density functions (pdfs) using observed data.

Standard Diagnostics: Estimating Covariance Matrices



$$E[\mathbf{d}_b^o(\mathbf{d}_b^o)^T] = \mathbf{R} + \mathbf{HBH}^T$$

$$E[\mathbf{d}_b^a(\mathbf{d}_b^o)^T] = \mathbf{HBH}^T$$

$$E[\mathbf{d}_a^o(\mathbf{d}_b^o)^T] = \mathbf{R}$$

1. **Split Observation-minus-Background statistics into observation and background error covariances** (e.g. Höllingsworth and Lonnberg, 1986)
   - Requires a dense observing network
   - Dependent on the chosen correlation function
   - Assumes observation errors are uncorrelated or requires additional assumptions on background error statistics

2. **(Iterative) procedures based on assumptions of data assimilation,** Desroziers et al. (2006)
   - Easy to implement
   - Dependent on prescribed error covariance matrices
   - Iterative procedure may be required
   - Iteration on background and observation errors concurrently results in convergence in one step that may be in error

Can Innovation Based Diagnostics Estimate Representation Error?

- Hodyss and Satterfield (2017) showed that when the observation is at a higher resolution than the model state, innovation based methods will have contributions from representation error as well as errors from resolved scales

$$\mathbf{d}_L^a = \mathbf{y}^o - \mathbf{H}_L \mathbf{x}_L^a$$

- The Desroziers estimates results in

$$\left\langle \mathbf{d}_L^a (\mathbf{d}_L^b)^T \right\rangle = \mathbf{R}_D$$

$$\mathbf{R}_D = \mathbf{R}^i + \underbrace{\mathbf{H}_H \mathbf{P}_H \mathbf{H}_H^T - \mathbf{H}_L \mathbf{P}_L \mathbf{H}^T{}_L}_{\textit{Representation error + model error on resolved scales}}$$

Can deliver the correct covariance estimate if there is no model error on resolved scales

Accounting for Representation Error

- **Attempt to remove correlated (representation) error**
  - Super ob, averaging, thinning, quality control
  - Inflate to reduce effect of neglected correlated error

- **Account for representation error covariance**
  - Use correlated observation error covariance matrices (e.g. Borman, Collard, Bauer 2010; Weston et al. 2014; Waller et al. 2015; Campbell et al. 2017, Simonin et al. 2019)
  - Implementing spatially correlated errors can be a more challenging problem for parallelization (e.g. Simonin et al. 2019)
  - Using spatial correlations can improve smallest scales represented by the model (Rainwater et al. 2015)

- **Modify KF equations to attempt to account for the influence of the unresolved scales on resolved scales** (e.g. Janjić and Cohn, 2006, Janjić et al. 2018, Hodyss and Nichols, 2015)

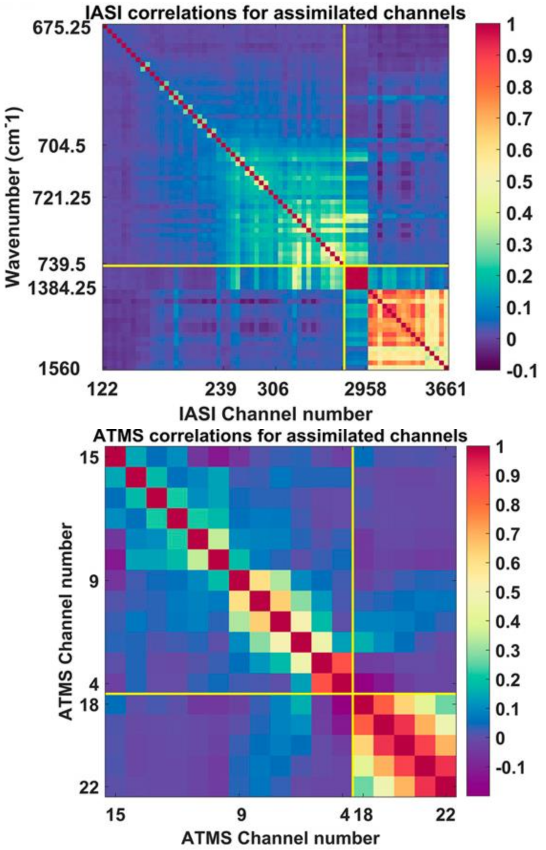


Image credit: Campbell et al. 2017

Flow Dependence Models of Representation Error

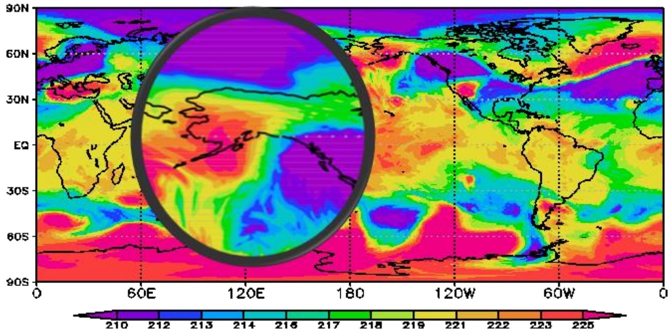
- Address spatial variations by calculating the observation error covariance **based on estimates of local turbulence** (e.g. Frehlich, 2006)

- Create a time dependent estimate of R using a **temporally smoothed Desrozier estimate** (e.g. Miyoshi et al., 2013; Waller et al. 2014)

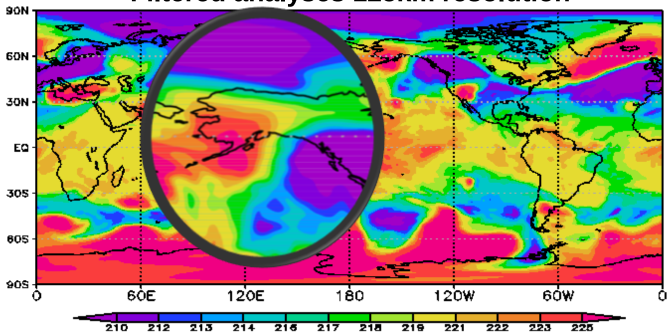
- Use an **ensemble-based estimator** (e.g. Karspeck, 2016)

- Satterfield et al. (2017) hypothesized that **ensemble variance could predict uncertainties associated with unresolved (smoothed) boundaries or gradients** associated with resolved scale features.

200 hPa temperature 1 Jan 2015
ECMWF analyses 32km resolution (TIGGEE)

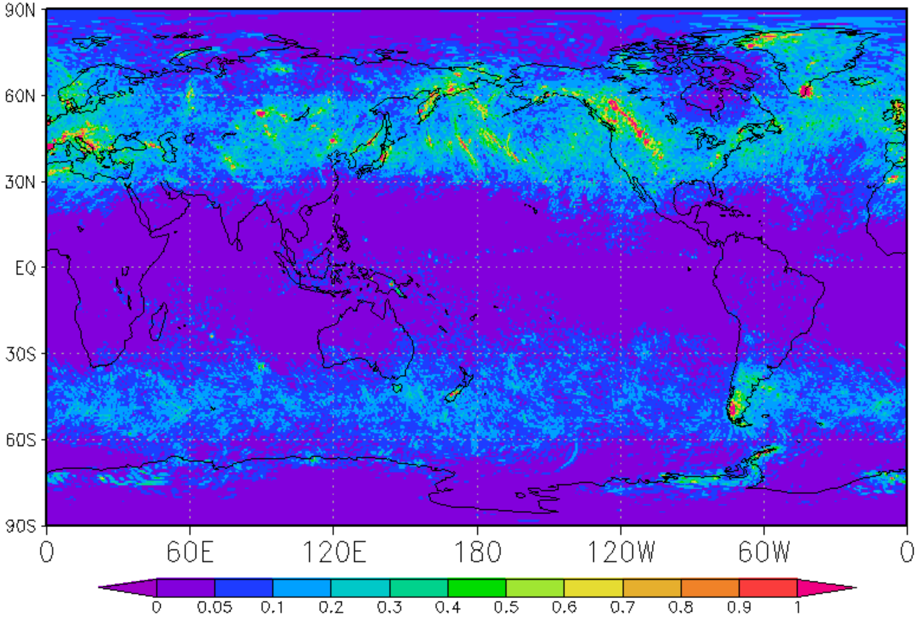


200 hPa temperature 1 Jan 2015
Filtered analyses 125km resolution

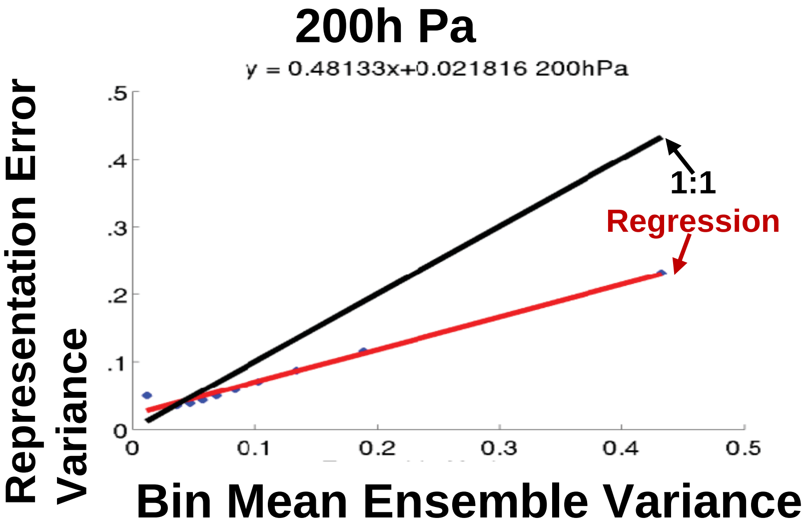


Can Ensemble Variance Predict Representation Error?

Error Variance due to Spectral Truncation for January 2015. Shown for temperature at 200hPa.



We bin ensemble variance and representation error (defined as high-minus-low resolution states) pairs, into equally populated bins as a function of ensemble variance.

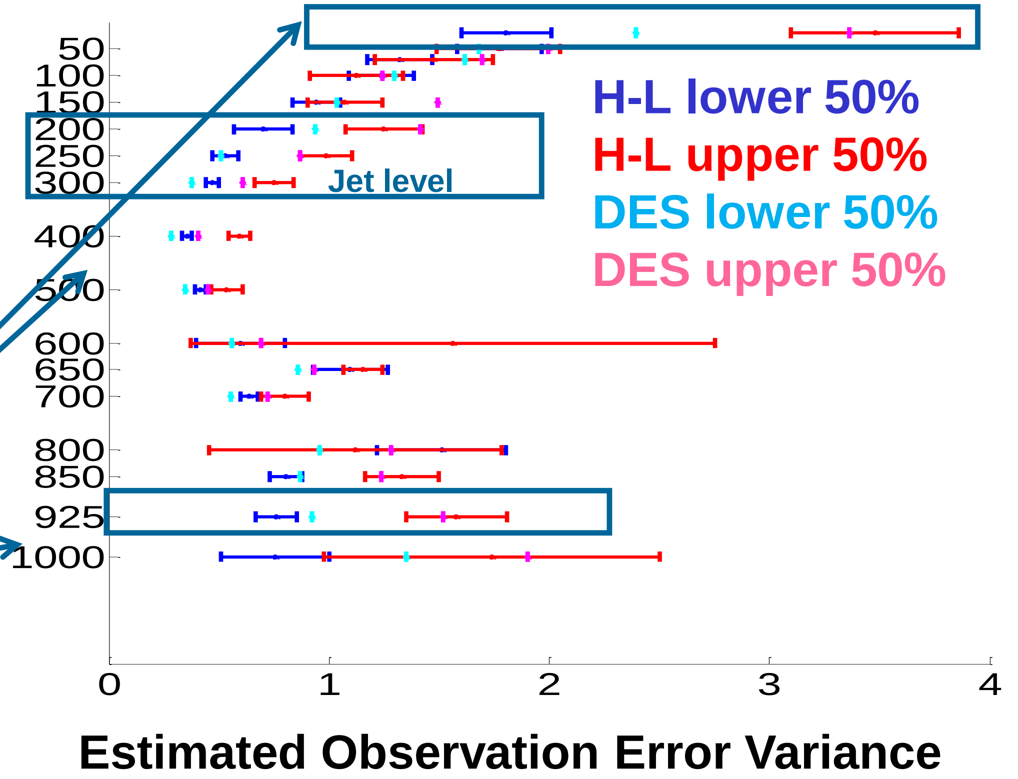


Do Desroziers and H-L estimates vary as a function of Ensemble Spread?

U.S. NAVAL RESEARCH LABORATORY

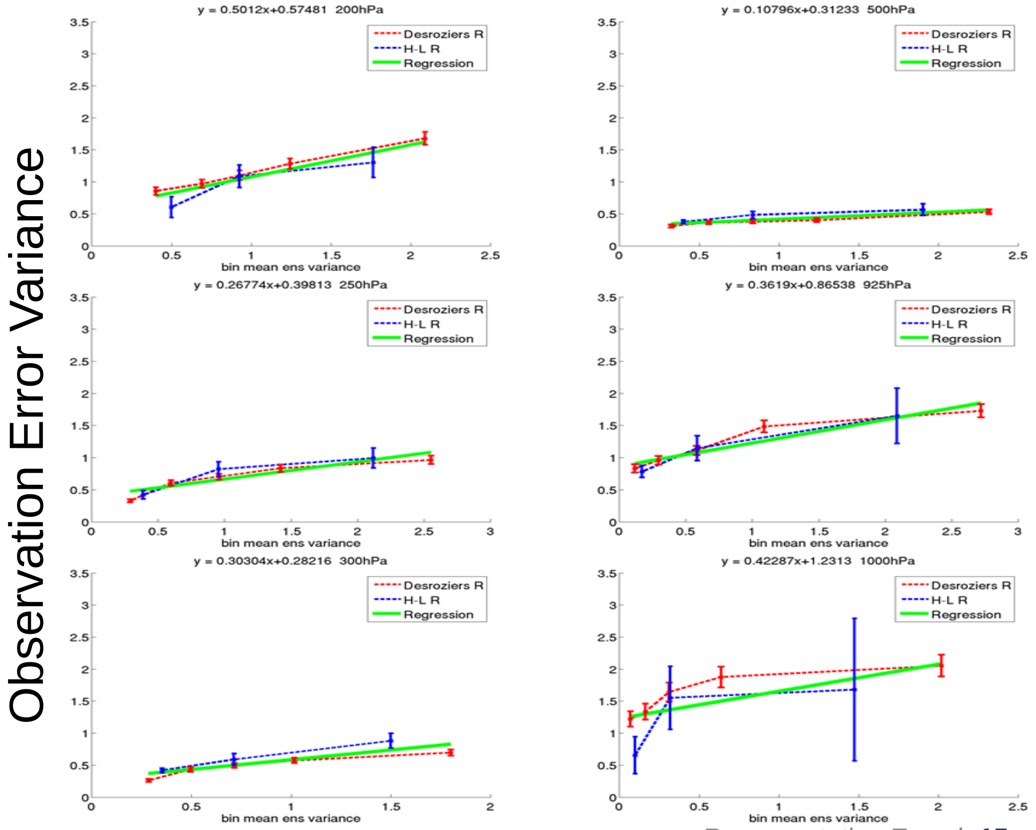
Using radiosonde temperature observations, we examined fluctuations in estimated observation error variances when the Desroziers and H-L methods are applied to subsets of innovations based on binning by ensemble variance.

Largest fluctuations agree with maps of variance due to spectral truncation, suggesting representation error



Accounting for Representation Error

- A regression based method would allow us to prescribe a static error and also allow for representation error to vary as a function of ensemble variance.
- Here we focus on those features that are mostly well represented by coarse model resolution and that have a high degree of spatiotemporal correlation.
- It remains to be determined how well a model like this would work for higher frequency fluctuations.



Discussion

- **Should we be using more quantitative techniques for diagnosing observation error?**
  - Current methods are imperfect, but still offer guidance. Likely more than one method needs to be used.
  - Using spatial correlations changes how observation information is filtered.
  - Capturing flow dependence is challenging, methods using temporal smoothers and ensemble based methods are encouraging.
  - How can existing methods for all sky that inflate observation error in the presence of cloud (e.g Geer, 2011; Minimide and Zhang 2017) be combined with innovation based methods?
  - How can we distinguish what is representation error vs. background error?

- **What are the implications of background error and observation error sharing characteristics and correlations?**
  - If model error and observation error are mutually correlated, the innovation relationship that is at the heart of most operationally used observation error estimation methods, no longer holds
  - Additional difficulty in distinguishing what is representation error and what is model error
  - What is the impact of ignoring the cross correlation in the gain? How much is gained from inflating observation error?

References

References

1. Berry, T., and T. Sauer, 2013: Adaptive ensemble Kalman filtering of non-linear systems. Tellus,Series A: Dynamic Meteorology and Oceanography, 65 (20331), 1–16.
2. Bormann N, Collard A, Bauer P (2010) Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. II: Application to AIRS and IASI data. Q J R Meteorol Soc 136:1051–1063
3. Campbell, W.F., E.A. Satterfield, B. Ruston, and N.L. Baker, 2017: Accounting for Correlated Observation Error in a Dual-Formulation 4D Variational Data Assimilation System. Mon. Wea. Rev., 145, 1019–1032, https://doi.org/10.1175/MWR-D-16-0240.1
4. Daescu, D.N. and Langland, R.H. (2013), Error covariance sensitivity and impact estimation with adjoint 4D-Var: theoretical aspects and first applications to NAVDAS-AR. Q.J.R. Meteorol. Soc., 139: 226-241. doi:10.1002/qj.1943
5. Dee, D. P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. Monthly Weather Review, 123 (4), 1128–1145.
6. Dee, D. P., G. Gaspari, C. Redder, L. Rukhovets, and A. M. da Silva, 1999: Maximum likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. Monthly Weather Review, 127 (1992), 1822–1834.
7. Desroziers, G., L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation, background and analysis-error statistics in observation space. Quart. J. Roy. Meteor. Soc.,131:3385–3396, 2005.
8. Desroziers, G., and S. Ivanov, 2001: Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. Quarterly Journal of the Royal Meteorological Society, 127 (574), 1433–1452.
9. Etherton, B.J. and C.H. Bishop, 2004: Resilience of Hybrid Ensemble/3DVAR Analysis Schemes to Model Error and Ensemble Covariance Error. Mon. Wea. Rev., 132, 1065–1080, https://doi.org/10.1175/1520-0493(2004)132<1065:ROHDAS>2.0.CO;2
10. Frehlich, R., 2006: Adaptive data assimilation including the effect of spatial variations in observation error. Quart. J. Roy. Meteor. Soc., 132, 1225–1257, doi:10.1256/qj.05.146.

References (continued)

1. Forget, G., and C. Wunsch, 2007: Estimated global hydrographic variability. J. Phys. Oceanogr., 563 37, 1997–2008.
2. Geer, A.J. and Bauer, P. (2011), Observation errors in all-sky data assimilation. Q.J.R. Meteorol. Soc., 137: 2024-2037. doi:10.1002/qj.830
3. Hodyss D, Nichols N (2015) The error of representation: Basic understanding. Tellus A 67:24 822–24 839
4. Hodyss, D and E. Satterfield. "The Treatment, Estimation, and Issues with Representation Error Modelling." Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. III). Park, Seon Ki, Xu, Liang (Eds.). Springer International Publishing, 2017. 177-194
5. Hollingsworth, A., and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. Tellus A. 38, 111-136.
6. Janjic T, Cohn SE (2006) Treatment of observation error due to unresolved scales in atmospheric data assimilation. Mon Weather Rev 134:2900–2915
7. Janjić, T, Bormann, N, Bocquet, M, Carton, JA, Cohn, SE, Dance, SL, Losa, SN, Nichols, NK, Potthast, R, Waller, JA, Weston, P. On the representation error in data assimilation, Q J R Meteorol Soc. 2018; 144: 1257– 1278. https://doi.org/10.1002/qj.3130
8. Karspeck, A.R., 2016: An Ensemble Approach for the Estimation of Observational Error Illustrated for a Nominal 1° Global Ocean Model. Mon. Wea. Rev., 144, 1713–1728, https://doi.org/10.1175/MWR-D-14-00336.
9. Liu Z-Q, Rabier F (2002) The interaction between model resolution, observation resolution and observation density in data assimilation: a one-dimensional study. Q J R Meteorol Soc 128:1367–1386
10. Minamide, M. and F. Zhang, 2017: Adaptive Observation Error Inflation for Assimilating All-Sky Satellite Radiance. Mon. Wea. Rev., 145, 1063–1081, https://doi.org/10.1175/MWR-D-16-0257.1
11. Miyoshi, T., E. Kalnay, and H. Li, 2013: Estimating and including observation-error correlations in data assimilation. Inverse Problems in Science and Engineering, 21 (3), 387–398.
12. Mitchell HL, Daley R (1997) Discretization error and signal/error correlation in atmospheric data assimilation (II). The effect of unresolved scales. Tellus A 49:54–73
13. Oke, P. R. and P. Sakov, 2008: Representation error of oceanic observations for data assimilation. J. Atmos. Oceanic Technol. 25, 1004-1017.
14. Satterfield, E., D. Hodyss, D.D. Kuhl, and C.H. Bishop, 2017: Investigating the Use of Ensemble Variance to Predict Observation Error of Representation. Mon. Wea. Rev., 145, 653–667, https://doi.org/10.1175/MWR-D-16-0299.1
15. Simonin, D, Waller, JA, Ballard, SP, Dance, SL, Nichols, NK. A pragmatic strategy for implementing spatially correlated observation errors in an operational system: An application to Doppler radial winds. Q J R Meteorol Soc. 2019; 145: 2772– 2790. https://doi.org/10.1002/qj.3592
16. Stroud, J. R., and T. Bengtsson, 2007: Sequential state and variance estimation within the ensemble Kalman filter. Monthly Weather Review, 135 (9), 3194–3208.
17. Rainwater S, Bishop CH, Campbell WF. 2015. The benefits of correlated observation errors for small scales. Quarterly Journal of the Royal Meteorological Society 141: 3439–3445, https://doi.org/10.1002/qj.2582.
18. Rutherford, I. D., 1972: Data assimilation by statistical interpolation of forecast error fields. Journal of the Atmospheric Sciences, 29 (5), 809–815.
19. Joanne A. Waller, Sarah L. Dance, Amos S. Lawless & Nancy K. Nichols (2014) Estimating correlated observation error statistics using an ensemble transform Kalman filter, Tellus A: Dynamic Meteorology and Oceanography, 66:1, DOI: 10.3402/tellusa.v66.23294
20. Waller JA, Dance SL, Nichols NK (2015) Theoretical insight into diagnosing observation error correlations using observation-minus-background and observation-minus-analysis statistics. Q J R Meteorol Soc. doi:10.1002/qj.2661
21. Weston P, Bell W, Eyre JR (2014) Accounting for correlated error in the assimilation of high-resolution sounder data. Q J R Meteorol Soc. doi:10.1002/qj.2306