ISSN: 0852-0682, EISSN: 2460-3945
Forum Geografi, Vol 30 (2) December 2016: 112-119

FORUM GEOGRAFI
Indonesian Journal of Spatial and Regional Analysis

Using 3D-Var Data Assimilation for Improving the Accuracy of Initial Condition of Weather Research and Forecasting (WRF) Model in Java Region (Case Study : 23 January 2015)

Novvria Sagita*[,1], Rini Hidayati[2], Rahmat Hidayat[2], Indra Gustari[1], and Fatkhuroyan[1]
[1)]Indonesian Agency for Meteorology, Climathology, and Geophysics (BMKG)
[2)]Department of Geophysics and Meteorology, Bogor Agricultural University (IPB)
*)Corresponding author (e-mail: novvria.sagita.ut@gmail.com)

Abstract. *Weather Research and Forecasting (WRF) is a numerical weather prediction model developed by various parties due to its open source, but the WRF has the disadvantage of low accuracy in weather prediction. One reason of low accuracy of model is inaccuracy initial condition model to the actual atmospheric conditions. Techniques to improve the initial condition model is the observation data assimilation. In this study, we used three-dimensional variational (3D-Var) to perform data assimilation of some observation data. Observational data used in data assimilation are observation data from basic stations, non-basic stations, radiosonde data, and The Binary Universal Form for the Representation of meteorological data (BUFR) data from the National Centers for Environmental Prediction (NCEP) , and aggregate observation data from all stations. The aim of this study compares the effect of data assimilation with different data observation on January 23, 2015 at 00.00 UTC for Java island region. The results showed that changes root mean square error (RMSE) of surface temperature from 2° C to 1.7° C - 2.4° C, dew point from 2.1° C to 1.9° C - 1.4° C, relative humidity from 16.1% to 3.5% - 14.5% after the data assimilation.*

Keywords: *WRF, initial condition, data assimilation, 3D-Var.*

Abstrak. *Weather Research and Forecasting (WRF) merupakan sebuah model prediksi cuaca numerik yang sedang dikembangkan oleh berbagai pihak karena bersifat terbuka (open source), tetapi WRF memiliki kekurangan berupa keakurasian prediksi cuaca yang kurang baik. Salah satu penyebabnya adalah ketidaksesuain syarat awal model (initial condition) terhadap kondisi atmosfer aktual. Salah satu metode untuk memperbaiki syarat awal model adalah dengan asimilasi data observasi menggunakan metode three dimensional varitional (3D-Var). Data observasi yang digunakan adalah data observasi permukaan dari stasiun utama, data observasi permukaan dari stasiun non-utama, data observasi rason, data observasi format BUFR dari National Centers for Environmental Prediction (NCEP), dan data gabungan dari data observasi dari seluruh stasiun. Tujuan penelitian ini untuk membandingkan pengaruh asimilasi data dengan data yang berbeda terhadap syarat awal model pada tanggal 23 Januari 2015 jam 00.00 UTC untuk wilayah pulau Jawa. Hasil penelitian menunjukkan perubahan nilai root mean square error (RMSE) syarat awal model pada estimasi suhu dari 2° C menjadi kisaran 1.7° C hingga 2.4° C, titik embun dari 2.1° C menjadi kisaran 1.4° C hingga 1.9° C dan kelembapan nisbi dari 16.1% menjadi 3.5% hingga 14.5% setelah asimilasi.*

Kata kunci: *WRF, syarat awal, asimilasi data , 3D-Var.*

1. Introduction

One of the problems of NWP system model is the inaccurate of initial conditions models. Initial condition of NWP system is one that is closer to the actual atmospheric conditions. One technique in forming the initial condition model that approaches the actual data of the atmosphere is by data assimilation. Data assimilation is the technique for combining observations data with an NWP product (the first guess or background forecast) and their respective error statistics to provide an improved estimate (the analysis) of the atmospheric state (Skamarock et al., 2005; Talagrand, 1997).

Indonesia Agency for Meteorology, Climatology and Geophysics (BMKG) difines the types of stations into the basic and the non-basic based on the obligation insending the data. The basic station has the obligation to send data to the international and non-basic station only has the obligation of data into a national system (BMKG, 2014). Data coming from the basic station is the one used in building global forecasting models. Utilization of observational data from non-basic station has not been made to establish a system of global NWP models. Data assimilation method is used to incoporate data observations from non-basic station into establishing the system of NWP models. Observation data itself is divided into two types i.e. surface and upper air data. Surface observation data are meteorological parameters up to a height of 10-meters, while the upper air observation data cover parameter up to a height of 10 kilometers (WMO, 2003).

The development of data assimilation method has improved weather prediction models. For instances, assimilation of observation data affects the predictions of planetary boundary layeras (PBL) heights as shown by Stauffer et al. (1991) who studied the effect of direct assimilation surface temperature by four-dimensional method of data assimilation (FDDA) resulting in reduced errors but the surface temperature can trigger high errors on the PBL height for changes that are not realistic for surface buoyancy flux. Alapaty et al. (2001) used assimilation method to increase the one-dimensional simulation of PBL. The method can significantly reduce the error model of PBL height. One method that has been known is the three-dimensional variational methods of data assimilation (3D-Var), which was introduced by Lorenc (1986). 3D-Var method can reduce overestimate precipitation WRF models in Tanzania (Athumani, 2012). Using the method of 3D-Var assimilation from Automatic Weather Station (AWS) data produces little improvement on the prediction of meteorological parameters (Junaedhi, IG et al., 2008; Dash, SK et al., 2013; Hou, T et al., 2013; Sahu, DK et al, 2014). 3D-Var method has a disadvantage that is not sensitive to the uncertainty of the vertical limit (Gao et al, 2004). Development of data assimilation methods produce another method called the four-dimensional variational methods of data assimilation (4D-Var), the technique of advanced data assimilation which takes into account the observation data every hour (Yang et al., 2009; Banister, R.N., 2007).

WRF is a numerical weather forecast system designed for atmospheric research to weather forecasting. WRF has been developed by the National Center for Atmospheric Research (NCAR), the National Centers for Environmental Prediction (NCEP), Forecast Systems Laboratory (FSL), Air Force Weather Agency (AFWA), Naval Research Laboratory, University of Oklahoma and the Federation Aviation Administration (FAA) since late 1990 (Skamarock et al., 2005). WRF allows researchers to generate atmospheric simulations based on real data observation. Part of WRF is used for the assimillation of the data referred to WRFDA (WRF Data Assimiation) (Skamarock et al., 2005).

This study will examine the difference impacts of the data assimilation using observation from the basic stations (bas), non-basic stations (non-bas), upper air data (rason), all observation data (all) and observation data from NCEP in the format of BUFR by comparing between the value of root mean square error (RMSE) of initial conditions without data assimilation and with data assimilation. This study also analyzes the spatial distribution of bias in the form of the difference of initial conditions between models with data assimilated initial condition and the model without data assimilation (control).

2. Research Method

We used WRF model data with input data from Global Forecasting System (GFS) with a resolution of 0.5º x 0.5º. Observation data is divided into a basic stations data, non-basic stations data, upper air observation data (rason) and observational data from NCEP in BUFR format.

NCEP data consists of a surface observation data, upper air observation data sent to the Global Telecommunication System (GTS), and data from satellites NESDIS (National Environmental Satellite Data and Information Service). This research will only be focused in java island region on January 23, 2015 at 00.00 UTC when. There 17 basic stations and 14 non-basic stations used in this study. Table 1 describes those stations. In addition, the only used rason data in Java is in Juanda meteorological stations.

The workflow of this study can be seen in Figure 1. This study uses a background (χ^b) and background error (B) WRF models. Background (χ^b) is the initial conditions of the model obtained from the simulation results in the WRF-GFS data input. Background error (B) is the statistical error of the model. Background errors are divided into: the global and regional background.

Background global error has been provided by the WRFDA. This study used the data assimilation with the background of global error that has been provided in the application WRFDA. γ^o is an observation data used to data assimilation and error value of observation shown by the parameter R.

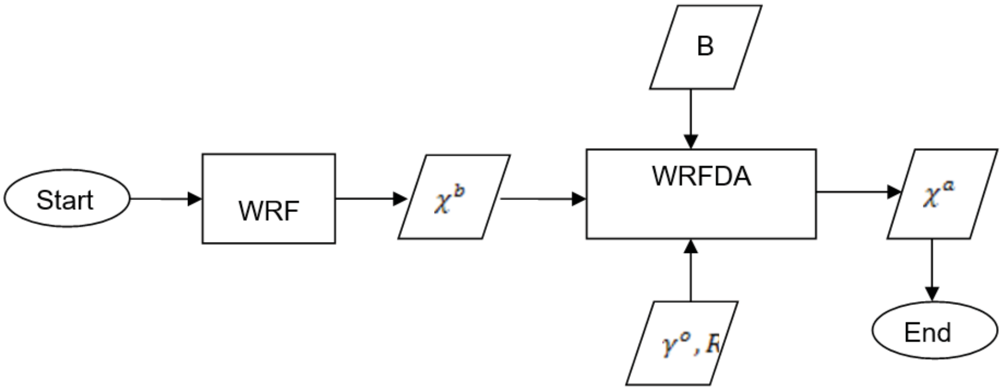

Figure 1. *Flowchart of data assimilation*

Table 1. Description of surface observation stations used in this study.

NO	Station	ID station	Latitude	Longitude	Elevation	Kind of Station
1	Serang	96737	-6.11185	106.11	50	Basic
2	Curug	96739	-6.287	106.57	46	Basic
3	Tj.Priok	96741	-6.10781	106.88	2	Basic
4	Kemayoran	96745	-6.15559	106.84	4	Basic
5	Cengkareng	96749	-6.12	106.65	8	Basic
6	Citeko	96751	-6.7	106.85	920	Basic
7	Jatiwangi	96791	-6.75	108.27	42	Basic
8	Tegal	96797	-6.85	109.15	3	Basic
9	Cilacap	96805	-7.73	109.02	6	Basic
10	Tj.Emas	96837	-6.98	110.38	3	Basic
11	A.yani	96839	-6.98	110.42	3	Basic
12	Bawean	96925	-5.85	112.63	3	Basic
13	Tj.Perak	96933	-7.13	112.46	3	Basic
14	Juanda	96935	-7.37	112.77	3	Basic
15	Tj.Perak 2	96937	-7.20937	112.74	3	Basic

NO	Station	ID Station	Latitude	Longitude	Elevation	Kind of Station
16	Kalianget	96973	-7.05	113.97	3	Basic
17	Banyuwangi	96987	-8.21	114.35	43	Basic
18	Pondok Betung	96733	-6.26107	106.75	26	Non-Basic
19	Tangerang	96735	-6.1	106.38	14	Non-Basic
20	Dramaga	96753	-6.5	106.75	207	Non-Basic
21	Bandung	96783	-6.88	107.6	829	Non-Basic
22	Banjarnegara	96807	-7.318	109.71	608	Non-Basic
23	Semarang	96835	-6.98	110.42	227	Non-Basic
24	Karangploso	96943	-7.90139	112.597	600	Non-Basic
25	Tretes	96945	-7.7	112.64	832	Non-Basic
26	Karangkates	96949	-8.15	112.45	325	Non-Basic
27	Sawahan	96975	-7.74	111.79	835	Non-Basic
28	Adi Sumarmo (AURI)	96845	-7.5	110.75	127	Non-Basic
29	Abdurahman Saleh (AURI)	96881	-7.917	112.7	523	Non-Basic
30	Atang Sanjaya (AURI)	96755	-6.5	106.75	163	Non-Basic
31	Tunggul Wulung	96805	-7.6167	109.05	21	Non-Basic

The intial condition (χ^b), the background error (B), and observation and the error (y^o, R) were used as input to the application WRFDA to assimilate the data and generate the new initial conditions (χ^a) after the model of assimilation. WRF configuration settings are performed in (Table 2) :

Table 2. WRF configuration

WRF configuration		
a.	Resolusi	9 km
b.	Time step	45 s
c.	Domain	one
d.	Grid	
	Utara-Selatan	120 grid
	Timur-Barat	120 grid
	Level vertikal	33 level
e.	Mikrofisik	Thompson scheme (Thompson, 2004)
f.	*Surface_Physics*	RUC
g.	Kumulus	Betts-Miller-Janjic scheme (Betts and Miller, 1986)
h.	Background error	Global (Barker et al., 2005)

The method used for the data assimilation is varitional three-dimensional analysis (3D-Var). Formula of 3D-Var is (Kalnay, 2003):

$$J\ x) = \frac{1}{2}\left\{\left[y^o - H(x)\right]^T R^{-1}\left[y^o - H(x\right] + \left(x - x^b\right)B^{-1}\left(x - x^b\right)\right\} \quad (1)$$

$$2J(x) = \left[y^o - H(x)\right]^T R^{-1}\left[y^o - H(x)\right] + \left(x - x^b\right)B^{-1}\left(x - x^b\right) \quad (2)$$

$$2J\ x = \left[y^o - H\ x\ \right]^T R^{-1}\left[y^o - H(x\right] + x - x^b\ B^{-1}\ x - x^b \quad (3)$$

$$2J(x = \left[y^o - H\ x\ \right]^T R^{-1}\left[y^o - H(x\right] + x - x^b\ B^{-1}\ x - x^b \quad (4)$$

$$2J\ x = x - x^b\ B^{-1}\ x - x^b + \left[\ y^o - H\ x\ - H\ x - x^b\ \right]^T R^{-1}\left[\ y^o - H\ x\ - H\ x - x^b\ \right] \quad (5)$$

Divide $(x\text{-}x^b)^T$to be :

$$\nabla J\left(x\right. = B^{-1}\ x - x^{b}\ + H^{T}R^{-1}H\ x - x^{b}\ - H^{T}R^{-1}\ y^{o} - H\ x \quad (6)$$

When $\nabla J(x) = 0$ so,

$$B^{-1} + H^{T}R^{-1}H\ x - x^{b}\ = H^{T}R^{-1}\ y^{o} - H\ x \quad x = \left(2-7\ + H^{T}R^{-1}H\ \right.^{-1} H^{T}R^{-1}\ y^{o} - H\ x \quad (7)$$

J (x) is a function that calculates the cost dissimilarity between models with observational data surface. The calculation of the cost function J (x) described in equations 2-1 where x and x^b is an analysis of data expected and the data background. H is the observation operator, R is the observation error covariance operator, B is operator background error, y is an observation data and y^o is observation data in the model grid.

3. Results and Discussions

This study analyzes the influence of data assimilation to the initial conditions (initial condition) models with several different observational data source. Analysis of the data assimilation impacts on the initial requirement model is an important step for checking the influence of assimilation of data for weather prediction. Analysis of the impact of the initial condition the model is done by comparing the value of RMSE between initial condition without data assimilation against observational data (TA-O) and the initial conditions of models with data assimilation (bas, non-bas, rason, sma, and BUFR) against observation data (AO).

RMSE value is used to identify influence data assimilation. Table 4 shows the changes of RMSE value for temperature from 2.0 °C (WA) to 1.7 °C (bass), 1.9 °C (non-bass), 2.4 °C (rason), 1.7 °C (all), and 1.7 °C (BUFR). For dewpoint temperature RMSE values have changed from 2.1 °C (WA) to 1.4 ° C (bass), 1.5 °C (non-bass), 1.9 °C (rason), 1.4 °C (all), and 1.5 °C (BUFR). The RMSE value for RH have changed from 16.1 % (WA) to 3.6 % (bas), 5.8 % (non-bas), 14.5 % (rason), 3.7 % (all), 3.5 % (BUFR). Almost all condition with data assimilation have RMSE value that are smaller than the initial condition without data assimilation. Data assimilation with rason generate increased RMSE for temperature parameters. This occurs because only one station used in data assimilation. While data assimilation with bass, all and BUFR have the smallest RMSE values for the third parameter (temperature, dewpoint, and relative humidity) weather among other data assimilation.This chapter contains the results of research.

Table 2. Estimation RMSE temperature, dew point temperature, and relative humidity (RH) model without assimilation (TA) and assimilation (A) versus observation.

Experiment	WA-O			A-O														
	RMSE without data assimilation (wa)			RMSE **bas** data assimilation			RMSE **non-bas** data assimilation			RMSE **ras** data assimilation			RMSE **all** data assimilation			RMSE **BUFR** data assimilation		
	Temperature (°C)	Dewpoint (°C)	Relative Humidity (RH) (%)	Temperature (°C)	Dewpoint (°C)	Relative Humidity (RH) (%)	Temperature (°C)	Dewpoint (°C)	Relative Humidity (RH) (%)	Temperature (°C)	Dewpoint (°C)	Relative Humidity (RH) (%)	Temperature (°C)	Dewpoint (°C)	Relative Humidity (RH) (%)	Temperature (°C)	Dewpoint (°C)	Relative Humidity (RH) (%)
RMSE	2.0	2.1	16.1	1.7	1.4	3.6	1.9	1.5	5.8	2.4	1.9	14.5	1.7	1.4	3.7	1.7	1.5	3.5

Figure 2 shows the estimation temperature bias at a height of 2 meters of the initial condition of the models with assimilation against initial condition of without assimilation. In general, all experiment of data assimilation resulted positive value of bias value for the eastern part of Java and negative bias for the region of western Java.

On the data assimilation BUFR provides a broad impact on the parameters indicated surface temperature of the surface temperature distribution bias initial condition of data assimilation models BUFR the initial condition of assimilation model without reaching areas outside Java.

4. Conclusion

The influence of data assimilation methods of 3D-Var to the initial conditions lead to a decrease of RMSE values of initial condition model for estimating surface weather parameter (surface temperature, dewpoint, and relative humidity). The best data assimilation used for decreasing RMSE value of the initial condition model are the ones that used data from the basic station, surface observation data from all stations, and data BUFR from NCEP. In the study area, the influences of data assimilation are also spatially visible for the surface temperature estimation which underestimated after data assimilation in the western part of Java and overestimated in the eastern part of Java.

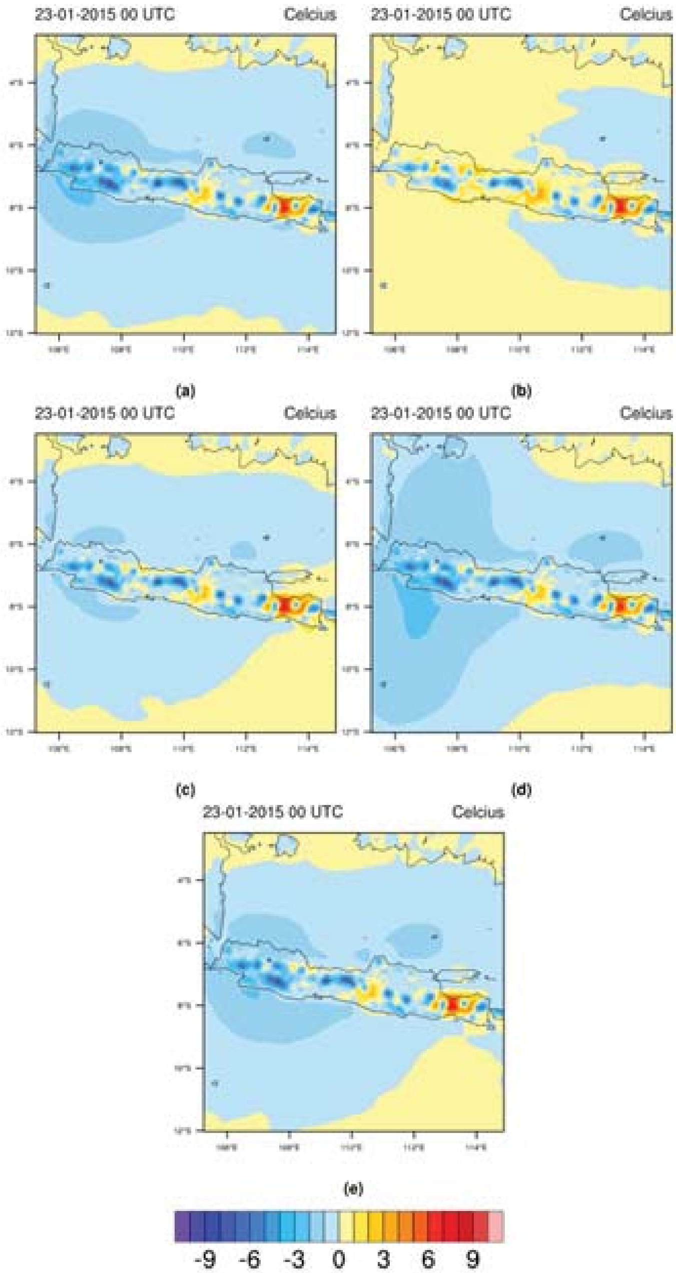

Figure 2. *Bias temperature estimation model of assimilation 2m initial conditions (a) bass, (b) rason, (c) non, (d) BUFR, and (e) all against without assimilation (control).*

5. References

Alapaty, K., Seaman, NL., Niyogi, DS., Hanna, A.F.(2001)*Assimilating Surface Data to Improve the Accuracy of Atmospheric Boundary Layer Simulations*. Journal Of Applied Meteorology.40, 2068-2082.

Yang, S.-C. et al., (2009)Comparison of Local Ensemble Transform Kalman Filter, 3DVAR, and

4DVAR in a Quasigeostrophic Model. *Monthly Weather Review*, 137, pp.693–709.

Bannister, R.N.,(2007)*Elementary 4-Dvar DARC Technical Report No.2*. Data Assimilation Research Centre, University of Reading, UK

Barker, D., Guo, Y.R., Lin, H.C.(2005)*Specification and Use of Forecast Error Covariances in WRF-Var*. NCAR, Taiwan.

Betts, A.K. and Miller, M.J., (1986) .*A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX, and Arctic Airmass data sets*. QQ. *J. R. Meteorol. Soc.*, 112, 693-710.

BMKG, (2014)Regulation of Chief Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG) No 12. Jakarta

WMO, (2003) Manual on the Global Observing System. World Meteorological Organization (WMO). Geneva.

Chin, H.N.S., Glascoe, L., Lundquist, J., Wharton, S.(2010)*Impact of WRF Physics and Grid Resolution on Low-Level Wind Prediction: Toward the Assesment of Climate Change Impact on Future Wind Power*.

Dash, S.K., Sahu, D.K., Sahu, S.C.(2013)*Impact of AWS observation in WRF-DVAR data assimilation system: a case study on abnormal warming condition in Odisha*. Net Hazard. 65, 767-798.

Gao, J.D. et al., (2004) *A three-dimensional variational data analysis method with recursive filter for Doppler radars*. Journal of Atmospheric and Oceanic Technology, 21(3), pp.457–469.

Hou, T., Kong, F., Chen, X. and Lei, H.,(2013)*Impact of 3DVAR data assimilation on the prediction of heavy rainfall over Southern China.Advances in Meteorology, 2013*, 1-17.

Junnaedhi, I.D.G.(2008)*Pengaruh asimilasi data dengan metode 3DVAR terhadap hasil prediksi cuaca numerik di Indonesia.* Tesis Institut Teknologi Bandung. Bandung. 73 hlm.

Kalnay, E.(2003)*Atmospheric modeling, data assimilation, and predictability*. New York. Cambridge University Press.

Kurniawan, R., et al.(2014)*Penggunaan Skema Konvektif Model Cuaca Wrf (Betts Miller Janjic, Kain Fritsch Dan Grell 3d Ensemble) (Studi Kasus: Surabaya Dan Jakarta)*. Jurnal meteorologi dan geofisika 15 (1), 25-36.

Lorenc, A. C. (1986) Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.* 112, 1177–1194.

Rajeevan, M., Kesarkar, A., Thampi, S.B., Rao, T.N., Radhakrishna, B. and Rajasekhar, M.(2010) *Sensitivity of WRF cloud microphysics to simulations of a severe thunderstorm event over Southeast India*. In *Annales geophysicae: atmospheres, hydrospheres and space sciences* 28(2), 603.

Sahu, D.K., Dash, S.K., Bhan, S.C.(2014) *Impact of Surface observations on simulation of rainfall over NCR Delhi using regional background error statistic in WRF-3DVAR model*. Meteorol Atmos Phys.125:17-42.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Wang, W. and Powers, J.G.(2005) *A description of the advanced research WRF version 2* (No. NCAR/TN-468+ STR). National Center For Atmospheric Research Boulder Co Mesoscale and Microscale Meteorology Div.

Stauffer, D.R., Seaman, N.L.,and Binkowski, F. S.(1991)*Use of four-dimensional data assimilation in a limited-area me- soscale model. Part II: Effects of data assimilation within the planetary boundary layer*.Mon. Wea. Rev., 119, 734–754.

Talagrand, O.(1997)*Assimilation of observation, an introduction*. J. Met.Soc. Japan Special Issue 75, 1B, 191-209.

Thompson, G., Rasmussen, R.M. & Manning, K.(2004)*Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part I: Description and Sensitivity Analysis*. Monthly Weather Review, 132(2), 519–542.

Yang, S.-C. et al., (2009)*Comparison of Local Ensemble Transform Kalman Filter, 3DVAR, and 4DVAR in a Quasigeostrophic Model*. Monthly Weather Review, 137, pp.693–709.