

## Asimilación de Datos en un Modelo de Pronóstico Numérico del Tiempo en el Servicio Meteorológico Nacional.

*Lic. M. E. Dillon, Dra. Y. Garcia Skabar, Dra. E. A. Collini  
Colaboradores externos: Dra. E. Kalnay, Dr. J. J. Ruiz*

El concepto de **asimilación de datos** ha sido bien definido por Talagrand (1997) como “*usar toda la información disponible para determinar de la manera más precisa posible el estado del flujo atmosférico u oceánico*”. El estado de la atmósfera y/o el océano se estima combinando de forma óptima pronósticos numéricos de corto plazo y observaciones. A dicha estimación generalmente se la denomina “análisis”.

En particular, podemos pensar al **análisis** como una “imagen” aproximada del estado de la atmósfera en un tiempo determinado. Es útil en sí mismo como un diagnóstico claro y consistente de la atmósfera y también es utilizado como estado inicial de los modelos de pronóstico numérico. Demás está decir que cuanto mejor sea dicho análisis, es decir cuanto mejor estimemos las condiciones iniciales, mejor va a ser la calidad del pronóstico (*Warner y otros, 1989; Harms y otros, 1992; Kalnay, 2003*).

En la actualidad, los métodos de asimilación de datos más avanzados en nuestra disciplina son los filtros de Kalman por ensambles (EnKF, ETKF, LETKF) y el método variacional en 4 dimensiones (4DVAR). Estos métodos permiten obtener información de las observaciones, de modo de considerar cómo los errores de las variables en la atmósfera dependen del estado de la misma. Además, permiten extraer información de observaciones obtenidas en distintos tiempos y a partir de diferentes sensores, tanto remotos como in-situ.

¿Por qué sería importante tener un sistema de asimilación de datos operativo en el Servicio Meteorológico Nacional (SMN)? Por un lado, sería importante porque permitiría contar con análisis regionales de mayor resolución, tanto espacial como temporal, respecto de los actualmente disponibles. Por otro lado, esto permitiría generar pronósticos partiendo de condiciones iniciales más adecuadas. Además, abriría la posibilidad de incluir observaciones locales que no llegan a asimilarse en los modelos globales.

Teniendo como objetivo avanzar en el diseño de un sistema de asimilación de datos en la Argentina, con el fin de que luego se pueda implementar de manera operativa en el SMN, se contactó a la Dra. Eugenia Kalnay y al Dr. Juan Ruiz, quienes gentilmente no dudaron en brindar colaboración activa. Gracias a su experiencia, se decidió utilizar el método Local Ensemble Transform Kalman Filter (**LETKF**) desarrollado en la Universidad de Maryland (*Hunt y otros, 2007*). **La mayor ventaja de dicho método es producir un ensamble de análisis enriquecidos**, a diferencia de los métodos variacionales que calculan sólo uno. Esta capacidad es fundamental para poder generar luego pronósticos por ensambles, que permitan obtener información sobre la incertidumbre en los pronósticos a corto plazo. Además, es independiente del modelo con el que se lo utilice, resulta muy eficiente

computacionalmente, y posee parámetros que se pueden ajustar según las necesidades requeridas por cada aplicación particular. Entre éstos se destaca, por ejemplo, el radio de localización, que determina la distancia de influencia de una observación.

El modelo de pronóstico numérico elegido fue el **WRF-ARW** (*Michalakes y otros, 2005; Skamarock y otros, 2008*) debido a los antecedentes satisfactorios existentes no sólo a nivel global sino también en nuestra región (por ejemplo: *Weisman y otros, 2008; Jankov y otros, 2008; Blázquez, 2007; Saulo y otros, 2008; Ruiz y otros, 2009; Collini y otros, 2010; Ruiz y otros, 2010; Dillon y otros, 2010; Collini y otros, 2011; Dillon, 2012*).

En particular, el **sistema WRF-LETKF** fue implementado por Miyoshi y Kunii (2012) usando observaciones reales. Los resultados alentadores de dichos autores, sumado a los resultados obtenidos con otros modelos utilizando el método LETKF (*Miyoshi y otros, 2007; Yang y otros, 2009; Seko y otros, 2011; Miyoshi y Kunii, 2012*), nos impulsaron a implementar el sistema de asimilación WRF-LETKF en nuestra región. El código del mismo es libre y se encuentra disponible en <http://code.google.com/p/miyoshi/> .

Con el fin de diseñar un sistema consistente, acorde tanto a las necesidades como a las posibilidades del SMN, se planteó como primera instancia una serie de cuatro experimentos. Para los mismos, se seleccionó el período de los meses de Noviembre y Diciembre de 2012, en coincidencia con la fecha de la campaña CHUVA, llevada a cabo en Rio Grande do Sul, Brasil ( <http://chuvaproject.cptec.inpe.br/portal/noticia.ultimasSantamaria.logic> ), durante la cual ocurrieron varios casos de precipitaciones intensas en la región.

Se decidió abarcar un dominio que incluyera Argentina y los océanos subyacentes con una resolución horizontal de 40 km (Figura 1). Respecto a los análisis, se resolvió obtenerlos cada 3 horas, siendo 00, 03, 06, 09, 12, 15, 18 y 21 UTC los horarios seleccionados, imprimiendo así un alto valor agregado, ya que actualmente la frecuencia de análisis disponible en nuestra región es de seis horas. Además, la cantidad de miembros del ensamble fue de cuarenta.

Respecto a las observaciones asimiladas, se utilizaron las contenidas en los archivos denominados prepbufr del NCEP (National Centers for Environmental Prediction). Los mismos poseen observaciones provenientes de estaciones de superficie, radiosondeos, barcos y aviones, y estimaciones de viento de los sensores GOES (Geostationary Operational Environmental Satellite) y ASCAT (Advanced Scatterometer), en nuestra región. En particular, se aplicó una técnica de “superobing” (creación de superobservaciones), implementada por el Lic. Pablo Echevarría, a las estimaciones del ASCAT con el fin de disminuir su resolución espacial, y así mejorar la performance del sistema WRF-LETKF.

La ventana de asimilación escogida es de 3 horas y la cantidad de observaciones disponibles es muy variada durante los ocho ciclos de análisis diarios. En la Figura 2 está representado el promedio de la cantidad de observaciones asimiladas para cada variable en cada ciclo, considerando el período bimensual mencionado. De aquí se desprenden dos características importantes: por un lado la escasez de datos en los horarios secundarios,

principalmente a las 09 y a las 21 UTC; y por otro lado la falta de información del estado termodinámico de la atmósfera. Esto nos llevó a incluir en un experimento futuro la asimilación de los perfiles verticales de temperatura y humedad estimados de los AIRS (Atmospheric Infrared Sounder; *Aumann y otros, 2003*), los cuales impactaron positivamente en otras regiones (*Miyoshi y Kunii, 2012*). A modo de ejemplo, se muestra en la Figura 3 la distribución de los distintos tipos de observaciones asimiladas para el ciclo de análisis de las 12 UTC del 5 de diciembre de 2012.

Una de las maneras de evaluar el desempeño del método es calculando su error cuadrático medio (RMSE) y comparándolo con el error del modelo GFS (Global Forecasting System), el cual se puede considerar como una referencia para la escala sinóptica. En la Figura 4 se observa que el orden de magnitud del RMSE de las componentes U y V del análisis de ambos modelos coincide. Asimismo, se puede observar una mejora de los análisis del sistema de asimilación respecto al pronóstico de 3 horas, lo cual es esperable y demuestra el correcto desempeño del método. A pesar de que no se vislumbra una mejora radical de la calidad del análisis generado por el WRF-LETKF respecto al del GFS, la similitud de los errores permite estimar que los experimentos que se están desarrollando actualmente darán resultados más favorables, ya que los mismos se plantearon a partir de características a mejorar, como ser la inclusión de mayor cantidad de observaciones de temperatura y humedad, o la consideración del error del modelo. Asimismo se planea incluir en los experimentos la utilización del “*no cost smoother*”, método que mejora la calidad de los análisis con un bajo costo (*Kalnay and Yang, 2010*).

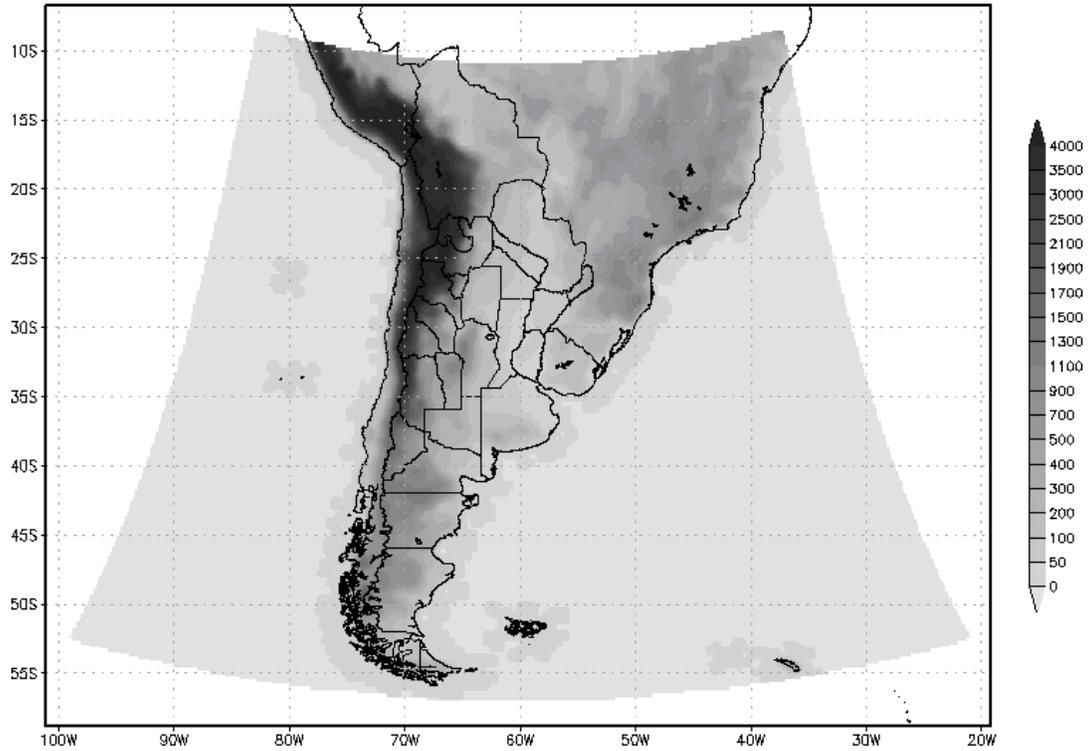
Actualmente, este sistema de análisis experimental corre en una máquina con 80 procesadores y demora aproximadamente 45 minutos en completar un ciclo de análisis. Con lo cual, la capacidad de cómputo necesaria para poder generar en tiempo real un ensamble de 40 análisis regionales no es sideral. Sin embargo, la mayor limitación recae en el almacenamiento: el tamaño de las salidas es muy grande y por lo tanto la necesidad de tener una alta capacidad de almacenamiento es imperiosa.

Además, dado que este desarrollo abre un abanico de posibilidades, como ser la generación de pronósticos por ensambles con el plazo que uno desee, el aumento de la resolución del modelo, o así mismo el aumento de la cantidad de miembros del ensamble, la necesidad de una máquina con mayor cantidad de procesadores también es muy importante.

Para concluir, **destacamos los siguientes puntos con el fin de representar tanto la potencialidad como la necesidad de la aplicación de este desarrollo:**

- Obtención de un ensamble de análisis regionales cada 3 horas.
- Obtención de un ensamble de pronósticos, por lo menos en los horarios principales de las 00 y las 12 UTC.

- Posibilidad de asimilar todo tipo de observaciones en un futuro, ya que el método LETKF es versátil en este sentido. Así se podrían asimilar observaciones que no están contenidas en los archivos que provee el NCEP.



*Figura 1. Topografía del dominio utilizado (m).*

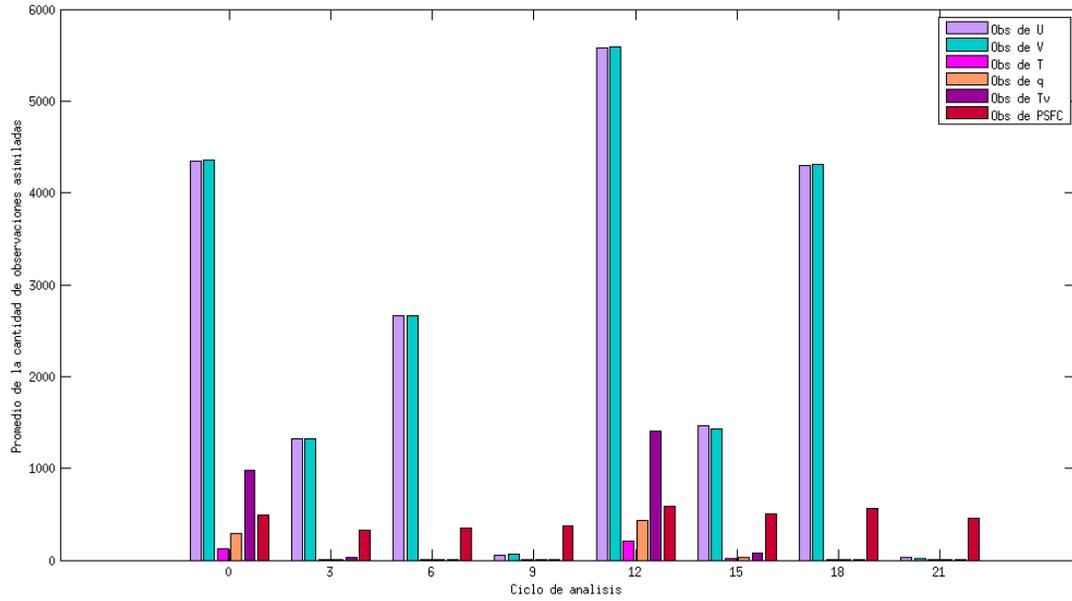


Figura 2. Promedio de la cantidad de observaciones asimiladas en cada ciclo de análisis para cada variable, considerando noviembre y diciembre de 2012. Las variables son: U (componente zonal del viento), V (componente meridional del viento), T (temperatura), q (humedad específica), Tv (temperatura virtual), PSFC (presión de superficie).

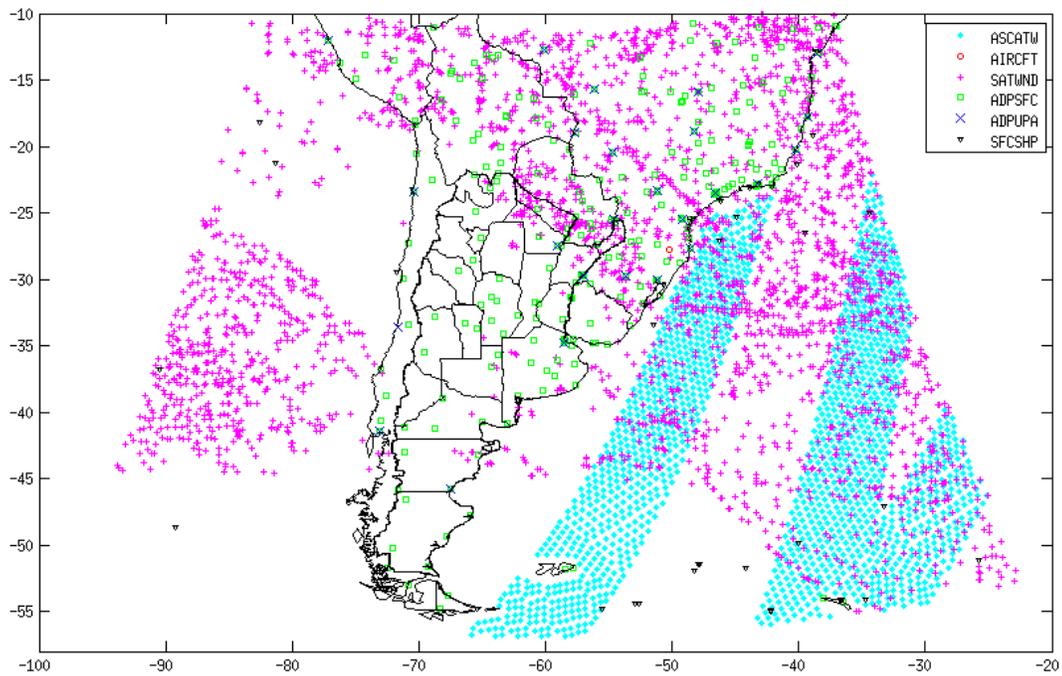
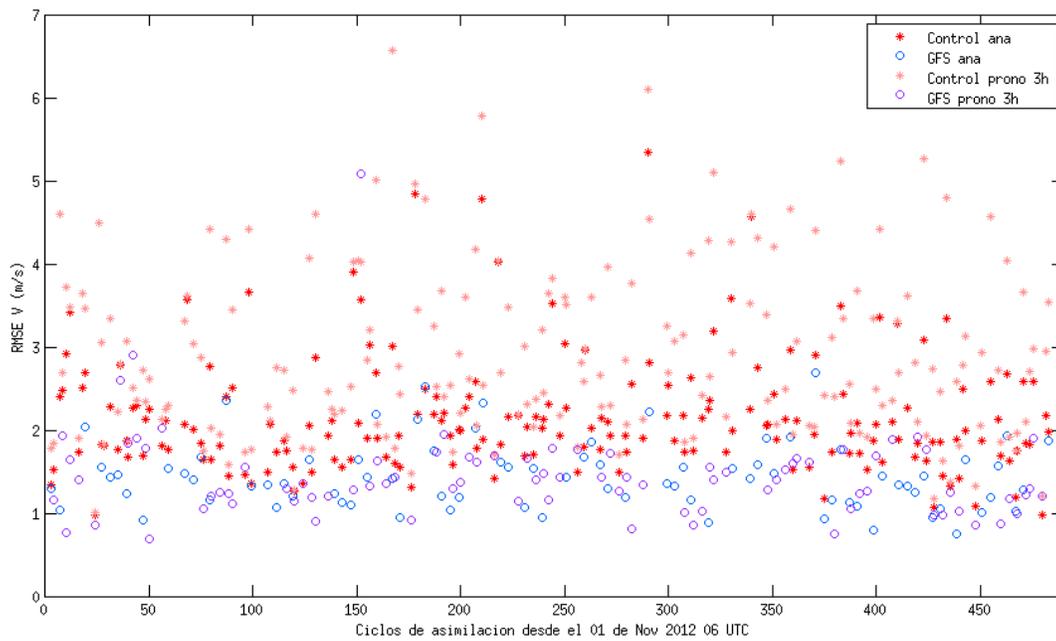
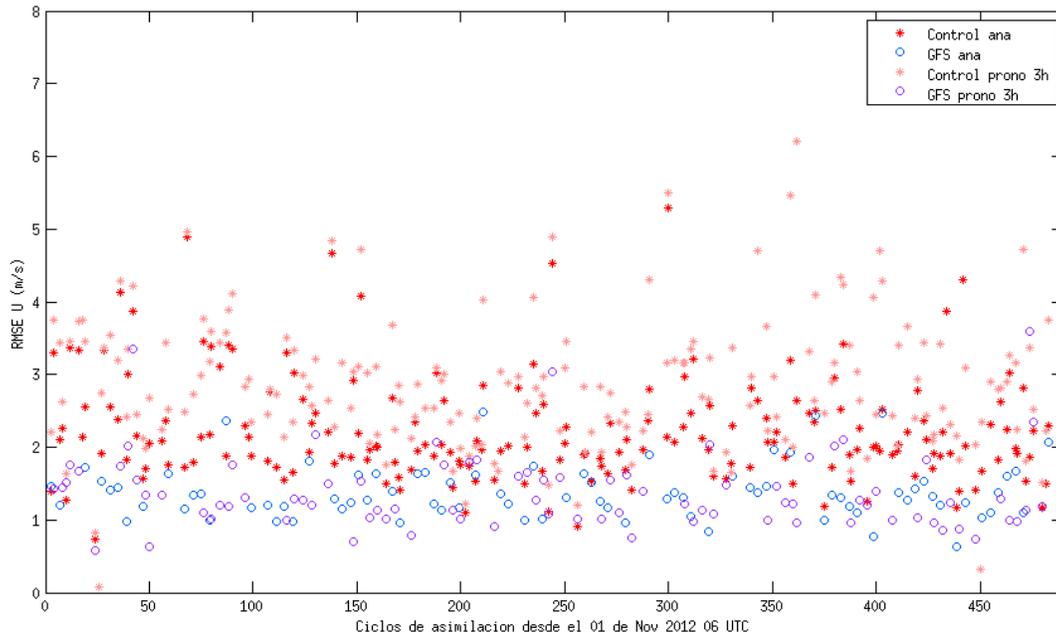


Figura 3. Observaciones asimiladas para el ciclo de análisis de las 12 UTC del 5 de diciembre de 2012. Los tipos de observaciones son: ASCATW (estimación del escaterómetro ASCAT), AIRCFT (mediciones de aviones), SATWND (estimación del satélite GOES), ADPSFC (estaciones de superficie), ADPUPA (radiosondeos), SFCSHP (mediciones de barcos).



*Figura 4. Comparación de los RMSE entre las observaciones y el análisis, y entre las observaciones y el pronóstico a 3hs, considerando el nivel de superficie, calculado para U (panel superior) y V (panel inferior); para los análisis y pronósticos generados por el WRF-LETKF (Control ana – Control prono 3h) y por el GFS (GFS ana – GFS prono 3h), para los ciclos de las 00, 03, 06, 09, 12, 15, 18 y 21 UTC, durante Noviembre y Diciembre de 2012.*

## Referencias.

Aumann, H. H., M. T. Chahine, C. Cautier, M. D. Goldberg, E. Kalnay, L. M. McMillin, H. Revercomb, P. W. Rosenkranz, W. L. Smith, D. H. Staelin, L. L. Strow, and J. Susskind, 2003: AIRS on the Aqua mission: Design, science objectives, data products, and processing systems. *IEEE Trans. Geosci. Remote Sens.*, 41, 253-264.

Blázquez J., 2007. Validación regional del modelo WRF bajo condiciones meteorológicas persistentes. Tesis de Licenciatura del Departamento de Ciencias de la Atmósfera y los Océanos, FCEyN, UBA

Collini E. A., L. Ferreira, M. E. Dillon, S. Osoro, G. Pujol, P. Martin, M. Suaya, A. Folch, 2011. Different Aspects of WRF-ARW applications at the National Meteorological Service – and Naval Hydrographic Service of Argentina. *WGNE Blue Book 2011*: <http://collaboration.cmc.ec.gc.ca/science/wgne/index.html>.

Collini E. A., M. E. Dillon, L. Ferreira, G. Pujol, 2010. Estudio de la sensibilidad del modelo WRF-ARW versión SMN empleando los campos de humedad de suelo provenientes de modelos globales y de sensores remotos. XXV Reunión Científica de la AAGG, Córdoba, Argentina; Resumen publicado en las Actas de la Reunión ISBN 978-987-25291-2-3.

Dillon M. E., 2012. Desempeño del pronóstico de modelos en alta resolución, en un área limitada durante una estación de verano. Tesis de Licenciatura del Departamento de Ciencias de la Atmósfera y los Océanos, FCEyN, UBA

Dillon M. E., M. Armanini, R. A. Valdivieso, M. Suaya, E. A. Collini, 2010. Verificación de temperaturas extremas de otoño del modelo WRF-ARW / SMN-SHN y su comparación con el modelo ETA / SMN en el territorio Argentino. XXV Reunión Científica de la AAGG, Córdoba, Argentina; Resumen publicado en las Actas de la Reunión ISBN 978-987-25291-2-3.

Harms, D. E., S. Raman, R. V. Madala, 1992. An Examination of Four-Dimensional Data- Assimilation Techniques for Numerical Weather Prediction. *Bull. Amer. Meteor. Soc.*, 73, No. 4, pp. 425–440.

Hunt, B. R., E. J. Kostelich and I. Szunyogh, 2007: Efficient Data Assimilation for Spatiotemporal Chaos: A Local Ensemble Transform Kalman Filter. *Physica D*, 230, 112-126.

Jankov I., W. Gallus, M. Segal, S. Koch, 2007. Influence of Initial Conditions on the WRF-ARW Model QPF Response to Physical Parameterization Changes. *Wea. and For.*, Vol 22, 501-519.

Kalnay E. and Yang S-C; 2010. Accelerating the spin-up of Ensemble Kalman Filtering. *QJRMS*

Kalnay E., 2003. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, pp 341, ISBN 0-521-79629-6.

Michalakes J., J. Dudhia, D. Gill, T. Henderson, J. Klemp, W. Skamarock, W. Wang, 2005. The Weather Research and Forecast Model: Software Architecture and Performance. *Proceedings of the Eleventh ECMWF Workshop on the Use of High Performance Computing in Meteorology*. Eds. Walter Zwiefelhofer and George Mozdzynski. World Scientific, pp 156 – 168.

Miyoshi T. and Kunii M., 2012. Using AIRS retrievals in the WRF-LETKF system to improve regional numerical weather prediction. *Tellus*, **64A**, 18408. DOI 10.3402/tellusa.v64i0.18408

Miyoshi T., S. Yamane, T. Enomoto, 2007. Localizing the error covariance by physical distances within a local ensemble transform Kalman filter (LETKF). *SOLA*, Vol. 3, 089 092, doi:10.2151/sola.2007 023

Ruiz J. J., C. Saulo, J. Nogués-Paegle, 2010. WRF Model Sensitivity to Choice of Parameterization over

South America: Validation against Surface Variables. *Mon. Wea. Rev.*, Vol. 138, 3342–3355. doi: 10.1175/2010MWR3358.1

Ruiz J. J., C. Saulo, E. Kalnay, 2009. Comparison of Methods Used to Generate Probabilistic Quantitative Precipitation Forecasts over South America. *Wea. Forecasting*, Vol. 24, 319–336. doi: 10.1175/2008WAF2007098.1

Saulo, C., S. Cardazzo, J. Ruiz, C. Campetella, A. Rolla, 2008. El sistema de pronóstico experimental del Centro de Investigaciones del Mar y la Atmósfera. *Meteorológica*, Vol. 33, 83 - 97.

Seko H., T. Miyoshi, Y. Shoji, K. Saito; 2011. Data assimilation experiments of precipitable water vapour using the LETKF system: intense rainfall event over Japan 28 July 2008. *Tellus*, 63A, 402-414.

Skamarock W., J. Klemp, J. Dudhia, D. Gill, D. Barker, M. Duda, X. Huang, W. Wang, J. Powers, 2008. A description of the Advanced Research WRF Version 3. NCAR Technical Note.

Talagrand O., 1997. Assimilation of observations: An introduction. *Journal of the Meteorological Society of Japan*, 75, 191 – 205.

Warner T., L. Key, A. Lario, 1989. Sensitivity of mesoscale model forecast skill to some initial data characteristics: data density, data position, analysis procedure and measurement error. *Mon. Wea. Rev.*, 117, No 6, 1281 – 1310.

Weisman M., C. Davis, W. Wang, K. Manning, J. Klemp, 2008. Experiences with 0-36h Explicit Convective Forecasts with the WRF-ARW Model. *Wea. And For.*, Vol 23, 407-437

Yang S-C, E. Kalnay, B. Hunt, N. E. Bowler; 2009. Weight interpolation for efficient data assimilation with the Local Ensemble Transform Kalman Filter. *QJR Met. Soc.*, 135, 251-262. DOI: 10.1002/qj.353