



Universidad de Buenos Aires

Facultad de Ciencias Exactas y Naturales

Departamento de Ciencias de la Atmósfera y los Océanos

Utilización de datos satelitales para la evaluación y mejora de los pronósticos numéricos en alta resolución a muy corto plazo

Tesis presentada para optar al título de Doctora de la Universidad de Buenos Aires en
el Área de Ciencias de la Atmósfera y los Océanos

Lic. Paola Corrales

Directora: **Vito Galligani**

Co-Director: **Juan Ruiz**

Consejera de estudios: **Celeste Saulo**

Lugar de trabajo: Centro de Investigaciones del Mar y la Atmósfera. CONICET-UBA

BUENOS AIRES, 2023

Acknowledgements

I want to thank a few people.

Preface

This is an example of a thesis setup to use the reed thesis document class (for LaTeX) and the R bookdown package, in general.

Índice

Introducción	1
0.1 Pronostico de eventos severos	1
0.2 Asimilación de datos como posible solución	2
0.3 Resultados previos de asimilación de distintas fuentes de observaciones	5
0.3.1 Asimilación de radianzas de satélites	6
0.4 Asimilación de datos en Sudamérica	9
0.5 Motivación y objetivos	10
Metodología {#metodologia}	12
0.6 El sistema de asimilación GSI	12
0.7 El modelo de transferencia radiativa CRTM	13
0.8 Caso de estudio	14
0.9 Configuración del ensamble	15
0.10 Verificación: obs y metricas	18
0.10.1 Conjunto de datos de validación	18
0.10.2 Métodos de verificación	18
0.11 Recursos computacionales	20
Capítulo 1: Asimilacion de observaciones de estaciones meteorológicas automáticas, vientos derivados de satélite y radianzas de satelites polares	21
1.1 Metodología	21
1.1.1 Observaciones asimiladas	21
1.1.2 Configuracion de los experimentos	27
1.2 Resultados	29
1.2.1 Consistencia del ensamble	29
1.2.2 Impacto de la asimilación de las distintas observaciones	32
1.2.3 Validación con observaciones independientes	41

1.3	Conclusiones	46
Capítulo 2: Graphics, References, and Labels	49	
2.1	Figures	49
2.2	Footnotes and Endnotes	52
2.3	Bibliographies	52
2.4	Anything else?	54
Conclusion	55	
Appendix A: The First Appendix	56	
Appendix B: The Second Appendix, for Fun	60	
References	61	

List of Tables

1	Generación de los 60 miembros del ensamble multifísica como combinación de parametrizaciones de Cumulus y PBL	17
1.1	Características de las observaciones asimiladas: El código de cada tipo de observación y su fuente, las variables disponibles, el error de observación y los umbrales de control de calidad utilizados.	23
1.2	Lista de los sensores disponibles cada plataforma, el número de canales aceptados para su asimilación y el porcentaje de observaciones asimiladas calculado sobre todas las observaciones de radiancias y todos los ciclos de asimilación.	25
1.3	Tipos de observaciones asimiladas en cada experimento.	28

List of Figures

1	Esquema de un ciclo de asimilación típico. El tiempo de las observaciones y el campo preliminar deberá coincidir.	2
2	Presión a nivel del mar (hPa, contornos negros), espesor 1000-500 hPa (contornos rojos discontinuos) y energía potencial convectiva disponible (sombreada) según ERA5 y temperatura de brillo del canal 13 de GOES-16 para a,d) 00 y b,e) 12 UTC, 22 de Nov y c,f) 00 UTC, 23 de Nov.	15
3	a) Dominio utilizado para las simulaciones (recuadro negro), dominio interior utilizado para la comparación entre experimentos (recuadro rojo), la región mostrada en b) (recuadro azul claro), y la ubicación de las Estaciones Meteorológicas Automáticas (EMA, cuadrados verdes) y las Estaciones Meteorológicas Convencionales (EMC, triángulos naranjas). b) Ubicación de los lanzamientos de radiosondeos durante RELAMPAGO. Los puntos verdes corresponden a los radiosondeos lanzados durante el IOP 7, los triángulos naranjas son radiosondeos lanzados durante el IOP 8, y los cuadrados morados son radiosondeos lanzados fuera de los periodos de medición intensiva. También se muestra la topografía en metros (sombreada).	17
1.1	Distribución espacial horizontal media de las observaciones disponibles en cada ciclo de asimilación para los experimentos a) CONV, b) AWS, c) SATWND y d) RAD calculados sobre cajas de $2,5^{\circ}$	26
1.2	a) Número de observaciones asimiladas por ciclo y b) número medio de observaciones asimiladas por ciclo dividido en capas verticales de 50 hPa de espesor para los experimentos CONV (cuadrados azules), AWS (puntos celestes), SATWND (triángulos naranjas) y RAD (diamantes rojos).	27

1.3	Diagrama de los ciclos de asimilación entre las 18 UTC del 20 de noviembre y las 12 UTC del 23 de noviembre más el periodo de spin up de 6 horas. La sección ampliada muestra la asimilación horaria que se realiza dentro de una ventana centrada de una hora y la incorporación de condiciones de borde del GFS cada 6 horas. Se muestran las dos misiones IOP de la campaña RELAMPAGO.	29
1.4	<i>sdRCRV</i> calculado para observaciones de superficie (de PREPBUFR y EMA) de a) viento, y b) temperatura promediados en cajas de 2.5° para el experimento RAD. Se usaron las observaciones agregadas de cada ciclo de asimilación horario paa todo el periodo del experimento.	31
1.5	Perfiles verticales de <i>meanRCRV</i> (línea punteada) y <i>sdRCRV</i> (línea sólida) para observaciones de a) temperatura y b) viento de sondeos y aviones, c) viento estimado por satélites, y d) temperatura de brillo para el experimento RAD. Se usaron las observaciones agregadas de cada ciclo de asimilación horario paa todo el periodo del experimento.	32
1.6	Diferencia entre la media del ensamble para la media del ensamble del análisis a) y d) AWS-CONV, b) y e) SATWND-AWS, y c) y f) RAD-SATWND para los perfiles verticales espacialmente promediados de la temperatura (a, b y c, en K) y la humedad específica (d, e y f en $g\ kg^{-1}$) calculados sobre el dominio interior (recuadro rojo en la Figura 3a) para cada ciclo de análisis.	34
1.7	Diferencia entre la media del ensamble para la media del ensamble del análisis a) y d) AWS-CONV, b) y e) SATWND-AWS, y c) y f) RAD-SATWND para los perfiles verticales espacialmente promediados del viento zonal (a, b y c, en K) y viento meridional (d, e y f en $g\ kg^{-1}$) calculados sobre el dominio interior (recuadro rojo en la Figura 3a) para cada ciclo de análisis. Los contornos negros corresponden al viento zonal y meridional para (a,d) AWS, (b,e) SATWND, y (c,f) RAD ya que son los experimentos tienen más observaciones asimiladas en cada panel.	36
1.8	Diferencia entre la media del ensamble del análisis de cada experimento y el ERA5 para los perfiles verticales espacialmente promediados de la temperatura del aire (K , a-d), la humedad específica ($g\ Kg^{-1}$, e-h) y el viento meridional ($m\ s^{-1}$, i-l) calculados sobre el dominio interior (recuadro rojo en la Figura 3a) para cada ciclo de análisis.	38

1.9	a) Agua precipitable (sombreada, $kg\ m^{-2}$) y viento medio del norte sobre los primeros 7 niveles sigma (desde la superficie hasta aproximadamente 800 hPa, contornos, $m\ s^{-1}$), b) Temperatura potencial media para la PBL (primeros 10 niveles sigma), y c) CAPE máxima y cortante del viento en ~0-6 km para 15 y 30 $m\ s^{-1}$ en cada experimento. Todos los campos corresponden a la media del ensamble del análisis para las 00 UTC del 22 de noviembre. Los contornos llenos de color gris corresponden a la topografía de más de 1500 metros sobre el nivel del mar.	40
1.10	Diagrama de Hövmoller de la media ajustada a la probabilidad de la precipitación acumulada horaria para cada banda de latitud estimada por IMERG (izquierda) y simulada (derecha) por cada experimento, promediada en un rango de longitudes entre $67^{circ}W$ y $54,5^{circ}W$. Los contornos se dibujan cada $0,5\ mm^{-1}$, comenzando en $0,5\ mm^{-1}$	42
1.11	FSS calculado sobre la precipitación acumulada a 1 hora en una ventana móvil de 6 horas para umbrales de 1 mm (a y c) y 25 mm (b y d), en escalas de 10 km (a y b) y 100 km (c y d), para el first-guess de los experimentos CONV (línea azul), AWS (línea celeste), SATWND (línea naranja) y RAD (línea roja).	43
1.12	Reflectividad máxima en la columna (COLMAX en dBZ), observada (fila superior) y pronósticada a 1 hora por los experimentos CONV (segunda fila) y RAD (tercera fila) a las 10 UTC (primera columna), 13 UTC (segunda columna), 16 UTC (tercera columna) y 19 UTC (cuarta columna) del 22 de noviembre de 2018. Los círculos negros en la primera fila muestran el rango de observación de cada radar.	45
1.13	RMSE (línea sólida) y Bias (línea discontinua) de a) la temperatura (K), b) la temperatura del punto de rocío (K), c) el viento u ($m\ s^{-1}$) y d) el viento v ($m\ s^{-1}$) calculados comparando el análisis de cada experimento con los sondeos de RELAMPAGO durante el IOP 7 y el IOP 8. La línea azul corresponde a CONV, la línea celeste a AWS, SATWND se representa con una línea naranja y RAD con una línea roja.	46
2.1	Reed logo	49
2.2	Mean Delays by Airline	51
2.3	Subdiv. graph	52
2.4	A Larger Figure, Flipped Upside Down	52

Abstract

The preface pretty much says it all.

Second paragraph of abstract starts here.

Dedication

You can have a dedication here if you wish.

Introducción

0.1 Pronostico de eventos severos

La simulación numérica de la atmósfera, es decir, la integración de las ecuaciones que rigen la evolución del sistema atmosférico es la base para la predicción del tiempo en diversas escalas temporales desde horas a semanas.

La predicción de fenómenos meteorológicos extremos es de particular importancia ya que pueden producir cuantiosas pérdidas humanas y materiales. En Argentina, una gran cantidad de estos fenómenos están asociados a la ocurrencia de convección profunda entre los que se cuentan tornados, ráfagas intensas, precipitaciones extremas en cortos períodos de tiempo, granizo de gran tamaño y actividad eléctrica. Es por tal motivo necesario avanzar en el conocimiento de estos fenómenos y en la capacidad de pronosticar la ocurrencia de los mismos.

El pronóstico meteorológico es un problema de condiciones iniciales y de borde. Si se cuenta con condiciones de borde apropiadas, es decir, una correcta representación de las características de la superficie terrestre y el tope de la atmósfera, la generación de pronósticos de calidad dependerá entonces, de la capacidad del modelo para representar los procesos atmosféricos y la exactitud de las condiciones iniciales usadas (Kalnay, 2002).

El pronóstico de los fenómenos severos es a su vez un desafío científico y tecnológico muy complejo debido a la predictibilidad limitada en la mesoescala y debido a la dificultad de conocer o diagnosticar el estado de la atmósfera en escalas espaciales pequeñas y tiempos cortos (por ejemplo de 1 a 10 km y del orden de los minutos).

Uno de los métodos que pueden utilizarse para el pronóstico de fenómenos meteorológicos severos es la utilización de modelos numéricos de la atmósfera que resuelvan explícitamente la convección profunda. Diversos estudios, han comprobado que estos modelos agregan valor al pronóstico a corto plazo y que en muchos casos proveen información sobre el modo de organización de las celdas convectivas y su intensidad (Aksøy et al., 2010; Stensrud et al., 2013). No obstante, la capacidad de los modelos

numéricos en anticipar la ubicación y tiempo de ocurrencia de eventos extremos asociados a convección es muy limitada si no se cuenta con una detallada información sobre el estado de la atmósfera en la escala de las tormentas en el momento en el que se inicializan los pronósticos numéricos (Clark et al., 2009).

0.2 Asimilación de datos como posible solución

Por otro lado es posible aplicar técnicas de asimilación de datos para generar una mejor estimación de las condiciones iniciales necesarias para integrar un modelo numérico. La asimilación de datos combina de manera optima un pronóstico numérico o campo preliminar en un tiempo t con las observaciones disponibles para ese mismo tiempo, generando un análisis. Esta combinación optima toma en cuenta el error asociado al modelo meteorológico (errores de pronóstico) y el error de las observaciones (que resulta de los errores instrumentales y los errores de representatividad) y su error será menor a los errores originales. Por esta razón el análisis es considerado desde un punto de vista estadístico y bajo ciertas hipótesis, como la *mejor aproximación* disponible del estado real de la atmósfera.

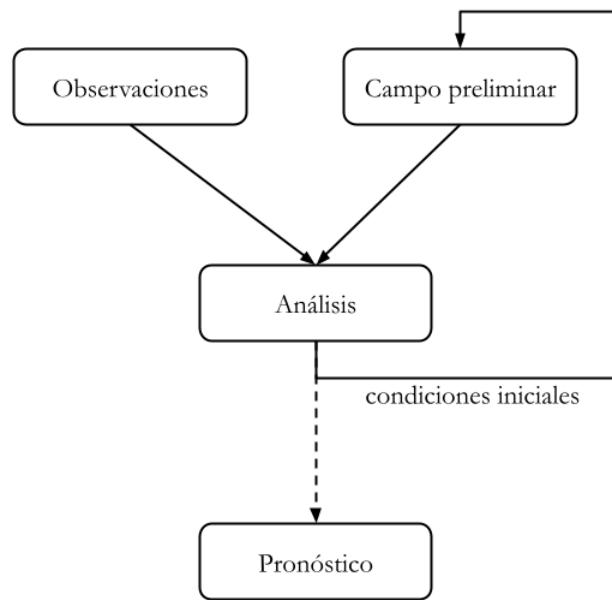


Figure 1: Esquema de un ciclo de asimilación típico. El tiempo de las observaciones y el campo preliminar deberá coincidir.

Un ciclo de asimilación de datos típico se muestra en la Figure 1. En un tiempo dado se comparan las observaciones disponibles con el campo preliminar para ese

mismo tiempo, generando así el análisis que se utilizará como condición inicial para un futuro pronóstico o campo preliminar. En el caso de modelos globales, típicamente cada ciclo de asimilación de 6 horas utiliza el campo preliminar previo, es decir el pronóstico a 6 horas inicializado a partir del análisis anterior y las observaciones disponibles para las 6 horas previas o en un periodo similar centrado en la hora del análisis. Para poder comparar y combinar el campo preliminar con las observaciones, este es interpolado a la ubicación de las observaciones. En determinados casos, por ejemplo cuando se trabaja con observaciones de satélite o radar, será necesario transformar las variables del modelo para obtener las variables observadas. En la siguiente ecuación H es el operador de las observaciones no lineal que se encarga de las interpolaciones y transformaciones necesarias sobre el campo preliminar x_b .

$$x_a = x_b + W[y_o - H(x_b)] \quad (1)$$

La diferencia entre las observaciones y_o y el campo preliminar se denomina innovación. El análisis x_a se obtiene aplicando las innovaciones al campo preliminar teniendo en cuenta un peso W que incluye información sobre los errores del pronóstico y de las observaciones.

Existen diferentes metodologías para obtener x_a . Los métodos variacionales, 3D-Var y 4D-Var, definen una función de costo que es proporcional a la distancia entre el análisis y el campo preliminar y el análisis y las observaciones simultáneamente. Esta función de costo J es minimizada para obtener el análisis.

$$J = \frac{1}{2}[y_o - H(x_a)]^T R^{-1} [y_o - H(x_a)] + (x_a - x_b)^T B^{-1} (x_a - x_b) \quad (2)$$

En la ecuación (2), el primer término corresponde a la distancia entre el campo del análisis y las observaciones, pesado por la covarianza de los errores de las observaciones R . El segundo término a la distancia entre el campo del análisis y el campo preliminar pesado por la covarianza de los errores del pronóstico B . Para el caso más simple, es decir, una variable de modelo y una observación, R y B son escalares. Para el caso multidimensional, serán matrices de covarianza de dimensión n (número de observaciones) que deben ser estimadas. Las ecuaciones (1) y (2) son equivalentes si H es un operador lineal y si se define a W como:

$$W = B H^T (H B H^T + R^{-1})^{-1} \quad (3)$$

El método 4D-Var extiende el uso del método 3D-Var para incluir la distancia a las

observaciones que pueden estar distribuidas temporalmente dentro de la ventana de asimilación en la misma función de peso. Sin embargo minimizar la función de costo requiere desarrollar el modelo tangente lineal y su adjunto, lo que puede ser costoso cuando se trabaja con modelos no lineales. Por esta misma razón, obtener la matriz B es un problema complejo y en general se la asume constante en el tiempo.

Los métodos secuenciales y en particular el filtro de Kalman extendido, actualizan el análisis a medida que las observaciones están disponibles. Este método tiene además la ventaja de actualizar la matriz B junto con el análisis. En este caso la matriz W toma el nombre de K o ganancia de Kalman que también actualiza en cada ciclo de asimilación t_i .

$$K_i = B(t_i)H^T(HB(t_i)H^T + R^{-1})^{-1} \quad (4)$$

La estimación de B utilizando el filtro de Kalman extendido es particularmente costosa en términos computacionales por lo que en la práctica se utiliza el filtro de Kalman por ensambles o EnKF. Un ensamble consiste en un conjunto de simulaciones ligeramente diferentes que se resuelven simultáneamente para incluir los posibles estados de la atmósfera y provee información dependiente de la dinámica durante la ventana de asimilación. A partir del ensamble, la matriz B se estima como:

$$B \approx \frac{1}{m-1} \sum_{k=1}^m (x_b^k - \bar{x}_b)(x_b^k - \bar{x}_b)^T \quad (5)$$

donde $k \in [1, m]$ el miembro k -ésimo del ensamble. Esta estimación será buena si el ensamble logra capturar los posibles estados futuros o en otras palabras el spread que acompaña los cambios en la incertidumbre de los pronósticos a lo largo de los ciclos de asimilación. Sin embargo, este método no es aplicable a menos que el tamaño del ensamble sea comparable a los grados de libertad de un modelo que resuelve 10^9 variables de estado, lo que resulta computacionalmente inviable.

El método Local Ensemble Transform Kalman Filter (LETKF) busca resolver los problemas anteriores resolviendo las ecuaciones previas en un espacio de dimensión reducida definido por los miembros del ensamble. Cómo en el resto de los métodos usando filtro de Kalman, se localiza o restringe el área de influencia de las observaciones a un determinado radio de localización reduciendo el costo computacional necesario. Además, calcula el análisis para cada punto de retícula uno a uno, incorporando todas las observaciones que puedan tener influencia en ese punto al mismo tiempo. De esta manera este método es hasta un orden de magnitud más rápido comparado con otros métodos desarrollados previamente (Whitaker et al., 2008).

Independientemente de la metodología aplicada, el modelo cumple un rol fundamental en la asimilación de datos ya que *transporta* información de regiones donde existe mucha información disponible (por ejemplo, los continentes) a regiones donde las observaciones son escasas (zonas oceánicas) manteniendo los balances físicos que rigen los procesos atmosféricos.

0.3 Resultados previos de asimilación de distintas fuentes de observaciones

Para que los métodos de asimilación de datos tengan éxito, deben utilizarse redes de observación con suficiente resolución temporal y espacial capaces de captar la variabilidad en las escalas que se quieren resolver, por ejemplo, la mesoescala.

Wheatley and Stensrud (2010) investigó el impacto de la asimilación de datos de presión de superficie en un sistema de asimilación de datos basado en conjuntos de mesoescala, pero encontró un impacto limitado en dos estudios de caso relacionados con sistemas convectivos de mesoescala. Ha and Snyder (2014) demostraron que la asimilación de la temperatura y la temperatura del punto de rocío de las redes de estaciones meteorológicas de superficie de alta resolución mejoraba sistemáticamente la estructura de la capa límite planetaria simulada y mejoraba la previsión de precipitaciones de corto alcance sobre los Estados Unidos. Chang et al. (2017), Bae and Min (2022) y Chen et al. (2016) informaron sobre los efectos beneficiosos de la asimilación de observaciones de estaciones meteorológicas de superficie en un sistema de asimilación de datos de alta resolución utilizando las metodologías de EnKF, 3D-Var y 4D-Var encontrando impactos positivos en el pronóstico de la temperatura y la humedad en la capa límite planetaria y en la localización de los sistemas de precipitación. Sobash and Stensrud (2015) demostró en un sistema de asimilación de datos de mesoescala que el impacto sobre la iniciación de la convección y el pronóstico de la precipitación de corto alcance es positivo si los datos se asimilan con frecuencia (en el orden de minutos, en lugar de en el orden de horas). Maejima et al. (2019) investigaron el impacto de la asimilación con frecuencia de 1 minutos de observaciones sintéticas en un caso de precipitación intensa, encontrando que la asimilación de observaciones de alta frecuencia y espacialmente espacialmente densas conducen a una mejor representación de la circulación de mesoescala aunque el número de observaciones proporcionadas por las estaciones de superficie es mucho menor que el proporcionado por los radares meteorológicos. Gasperoni et al. (2018) realizó un estudio de caso para evaluar el

impacto de la asimilación de las observaciones producidas por estaciones meteorológicas privadas que no se incorporan a los análisis operativos globales. Encontró un efecto positivo al asimilar estas observaciones sobre el inicio de la convección húmeda profunda a lo largo de una línea seca. Este resultado es especialmente importante para regiones con pocos datos, como el sur de Sudamérica, donde las redes operativas no son lo suficientemente densas como para captar los detalles de la mesoscala. En ese sentido, Dillon et al. (2021) utilizaron por primera vez observaciones de estaciones meteorológicas automáticas de redes privadas en el sur de Sudamérica, sin embargo, la contribución específica de este tipo de observaciones sobre esta región, no ha sido investigada hasta ahora.

Se ha investigado el impacto de otros tipos de observaciones de resolución espacial y temporal relativamente alta, como observaciones de satélites, en el contexto de la asimilación de datos de mesoscala. Estas observaciones incluyen radianzas y productos derivados como vientos derivados de satélite y perfiles de temperatura y humedad, Wu et al. (2014), Cherubini et al. (2006) y Sawada et al. (2019) observaron un impacto positivo de la asimilación de viento derivado de información satelital de alta frecuencia en un estudio de caso de un ciclón tropical utilizando un sistema de asimilación de datos basado en ensambles. Por otro lado, Gao et al. (2015) observaron un impacto positivo en pronósticos a corto plazo gracias a la asimilación de viento estimado a partir de las observaciones de satélites geoestacionarios sobre Estados Unidos.

0.3.1 Asimilación de radianzas de satélites

En esta sección se resumirán los avances en la asimilación de estas observaciones a nivel global y regional.

Los primeros satélites en proveer información meteorológica fueron desarrollados en las décadas de los 60 y 70. Estos estaban ubicados en órbitas polares, es decir, con cierta inclinación respecto del Ecuador, pasando cerca o sobre los polos. Incluían sensores infrarrojos y de microondas para monitorear la temperatura y humedad. Hacia finales de la década de los 70, Estados Unidos, Europa y Japón ya habían lanzando los primeros satélites geoestacionarios. Pocos años después este tipo de observaciones se incorporaban al Sistema de Observación Global (Global Observing System en inglés).

El primer conjunto de satélites compuesto por los sensores High-resolution Infrared Radiation Sounder (HIRS), Microwave Sounding Units (MSU) y Stratospheric Sounder Unit (SSU) o sistema TOVS (por su nombre en inglés, TIROS Operational Vertical

Sounder) podían cubrir el globo completo cada 12 hs. Si bien cada uno de estos sensores generaba información complementaria en la tropósfera y baja estratosfera, la resolución horizontal y vertical era limitada. En particular el primer HIRS, un sensor infrarrojo tenía una resolución horizontal de 40 km, mientras que actualmente HIRS4 tiene una resolución horizontal de 10 km. MSU, sensor sensible en las microondas, tenía una resolución de 110 km mientras que el sensor que lo reemplazó, Advanced Microwave Sounding Unit-A (AMSU-A) cuenta con una resolución de 50 km. En la vertical, el ancho la función de peso de los distintos canales ronda entre los 5 y 10 km y aún en los casos donde los canales se solapan, la resolución apenas alcanza los 3 km.

Las primeras pruebas de asimilación de observaciones de satélites fueron desarrolladas principalmente en Australia, motivadas particularmente por la escasez de observaciones en el hemisferio sur. Kelly et al. (1978) mostró una importante mejora en pronósticos a 24 horas de altura geopotencial entre 1000 y 200 hPa cuando se asimilaba de manera continua perfiles de temperatura derivados del satélite Nimbus-6, conocidos también como retrievals. A nivel global Ohring (1979) resume los avances de la década indicando los impactos son positivos aunque pequeños y que la mayor mejora se observa en los pronósticos en el hemisferio sur. Al mismo tiempo Ohring (1979) señala algunos de los posibles problemas asociados, por ejemplo la baja resolución vertical de los perfiles de temperatura y humedad y problemas en la generación de los mismos.

A principios de los 80 los centros de pronóstico mundiales continuaron estudiando la posibilidad de asimilar observaciones satelitales encontrando una disminución en el error de pronósticos a 6 horas principalmente en regiones donde hay poca disponibilidad de otras observaciones (Eyre et al., 2020). En particular el European Centre for Medium-Range Weather Forecasts (ECMWF) Seminar on Data Assimilation Systems and Observing System Experiments concluye que la asimilación de estas observaciones cumple un rol importante en el análisis de sistemas meteorológicos de larga escala en latitudes medias y altas, y en particular en el Hemisferio Sur. Sin embargo, hacia finales de los 80, los modelos de pronóstico habían mejorado sustancialmente haciendo que el potencial impacto de observaciones erróneas u observaciones asimiladas de manera incorrecta degradaran sustancialmente el pronóstico particularmente en el Hemisferio Norte. Andersson et al. (1991) mostró que los incrementos en el análisis presentaba patrones con importante sesgo cuando se asimilaba retrievals de TOVS.

Eyre et al. (2020) explica que la principal razón por la que los resultados obtenidos no fueran positivos era que se trataba a los retrievals como “sondeos de baja calidad” sin tener en cuenta las características particulares de las observaciones de satélite

como la resolución horizontal y vertical.

En la década de los 90, luego de que los centros de asimilación comenzaran a utilizar técnicas avanzadas de asimilación de datos como 3D-Var, se dieron las condiciones necesarias para asimilar radianzas de satélites de manera directa. Sin embargo, la correcta asimilación de estas observaciones depende de 3 factores, que las observaciones no tengan sesgo, que sus errores tengan una distribución Gaussiana y que el problema no sea afectado fuertemente por procesos no lineales (Eyre et al., 2022). Para asegurar estas condiciones fue necesario el desarrollo de técnicas de detección de nubes que permitan filtrar las regiones afectadas por nubosidad, principalmente para observaciones de sensores infrarrojos. Sin embargo, un avance clave en la asimilación de estas observaciones fue el desarrollo de modelos de transferencia radiativa que pudieran transformar el campo preliminar en radianzas comparables con las observaciones en tiempos razonables para ser usados de manera operativa. Finalmente, el desarrollo de métodos de corrección del bias de radianzas aplicados directamente en el proceso de asimilación fue determinante para la asimilación directa de este tipo de observaciones.

Junto al desarrollo de la asimilación de radianzas, también continuó el desarrollo de nuevos sensores, como la serie AMSU-A y AMSU-B y el sistema ATOVS (Advance TOVS) que cuenta con mayor cantidad de canales y por lo tanto una mayor resolución horizontal y vertical. Posteriormente el desarrollo de los sensores multiespectrales como IASI y AIRS permitieron obtener información con mayor resolución vertical al contar con más de 3000 canales en la región infrarroja del espectro electromagnético.

Una parte importante del desarrollo de la asimilación de datos en los últimos 20 años tiene que ver con el desarrollo de metodologías que tengan en cuenta la influencia de la superficie y la interacción entre las nubes y la energía electromagnética para los distintos canales infrarrojo y microondas. Inicialmente solo se asimilaron observaciones sobre agua y condiciones de cielos despejados. Sin embargo mejoras en los modelos de transferencia radiativa respecto del tratamiento de los distintos tipos de superficie y la representación y tratamiento de las nubes permiten en la actualidad incorporar observaciones que usualmente no podrían asimilarse.

Mientras que la asimilación directa de radianzas en modelos globales está establecida y estudiada [], las aplicaciones en modelos regionales, sin embargo, sigue siendo un desafío por la escasa cobertura de las observaciones debido a la órbita de los satélites polares, la corrección del sesgo y el tope de la atmósfera bajos usados en modelos regionales. Bao et al. (2015) estudió el impacto de la asimilación de datos de radiancia de microondas e infrarrojo en el pronóstico de temperatura y humedad en el oeste de EE.UU. y encontró una reducción del sesgo de la temperatura en niveles bajos y

medios como resultado de las observaciones de microondas, pero un efecto opuesto cuando se asimilaban radianzas en el infrarrojo. Más recientemente, Zhu et al. (2019) estudió el impacto de la asimilación frecuente de radiancias de satélites para un sistema regional y mostró una mejora para todas las variables, en particular para la humedad relativa en los niveles superiores. Wang and Randriamampianina (2021) estudiaron el impacto de la asimilación de observaciones en el infrarrojo en el Reanálisis Regional Europeo Copernicus de alta resolución e informaron que las observaciones de radiancia de satélite tuvieron un impacto neutro en los análisis de la altura geopotencial en la tropósfera baja, mientras que el impacto fue ligeramente negativo para la tropósfera superior y estratosfera. También observaron resultados similares para pronósticos a 3 hs inicializados a partir del análisis, pero un impacto positivo en las previsiones de mediano plazo (12 y 24 hs). Teniendo en cuenta los variados resultados, es necesario continuar estudiando la utilidad de asimilar las observaciones de radiancia en un sistema de asimilación de datos de área limitada sobre tierra.

0.4 Asimilación de datos en Sudamérica

La historia de la asimilación de datos en Sudamérica y en particular en Argentina es relativamente corta. A principios de la década del 90 Vera (1992) en su tesis doctoral desarrolló un Sistema de Asimilación de Datos Intermítente que utilizaba la interpolación óptima en un modelo cuasigeoestrófico en la región sur de Sudamérica. Algunos años después, en 1997, el Servicio Meteorológico Nacional se implementó un análisis utilizando el método de Cressman en un modelo de 10 niveles verticales (García Skabar, 1997).

Por otro lado el Centro de Pronóstico del Tiempo y Estudios Climáticos (CPTEC) de Brasil desarrolló un sistema de asimilación de datos global que utiliza el sistema Gridpoint Statistical Interpolation (GSI) en conjunto con su modelo global BAM y posteriormente aplicaciones regionales utilizando el modelo WRF en conjunto con el sistema de asimilación GSI. En particular, Goncalves de Goncalves et al. (2015) mostró experimentos realizados en el CPTEC usando el sistema de asimilación de datos regional para simulaciones de 12, 0 y 3 kilómetros durante un mes. Ferreira et al. (2017), Bauce Machado et al. (2017), Toshio Inouye et al. (2017) y Ferreira et al. (2020) también mostraron resultados positivos al aplicar asimilación de observaciones convencionales de superficie y altura y radar, en aplicaciones regionales sobre Brasil con resoluciones de entre 1 y 10 km.

En los últimos años, se documentaron importantes avances asociados a asimilación

de datos en Argentina. Por ejemplo Saucedo (n.d.) realizó un estudio teórico de asimilación de datos utilizando LETKF acomplado al modelo WRF donde estudio tecnicas para la representacion de diferentes fuentes de incertidumbre incluyendo los errores en las condiciones de borde y los errores de modelo. Posteriormente (???) avanzó en su tesis de doctorado en el desarrollo de un sistema de asimilación de datos reales y concluyó que la implementación de un emsable multifísica que considere los posibles errores del modelo y la inclusión de retrievals de temperatura y humedad en la asimilación tienen un impacto positivo en los análisis y pronósticos. [placeholder para incluir los trabajos de Paula]. Más recientemente, el Servicio Meteorológico Nacional (SMN) en conjunto con el Centro de Investigaciones del Mar y la Atmósfera desarrollaron y probaron el sistema de asimilación de actualización rápida LETKF-WRF de manera operativa durante la campaña de campo RELAMPAGO (Nesbitt et al., 2021). El sistemá incorporó observaciones convencionales, retrievals de satélites multiespectrales y viento derivado de observaciones satelitales y observaciones de radar de manera horaria y generó pronósticos a 36 hs cada 3 hs. Dillon et al. (2021) mostraron que el pronóstico inicializado a partir de los análisis muestra un rendimiento general similar al de los pronósticos inicializados a partir del sistema GFS, e incluso un impacto positivo en algunos casos. Actualmente el SMN está probando un sistema de asimilación similar al implementado en Dillon et al. (2021) para utlizarlo en la generación de pronósticos de manera operativa.

0.5 Motivación y objetivos

En base a los imporantes avances en la asimilación de datos en general y en las aplicaciones regionales en Argentina y Sudamérica, el objetivo principal de este trabajo es contribuir a la cuantificación y comparación del impacto de diversas fuentes de datos que aportan informacion en escalas espacio-temporales dentro de la mesoscala. Se utilizará el sistema WRF-GSI-LETFK para la generación de experimentos de asimilación de datos de actualización frecuente y basado en emsables. Mientras que el modelo WRF (por su nombre en inglés Weather Research and Forecasting, Skamarock et al. (2008)) es uno de los más utilizados y en constante avance, con importantes antecedentes en Argentina (por ejemplo Dillon et al. (2021)), el sistema de asmilación GSI (por su nombre en inglés Gridpoint Statistical Interpolation system, Shao et al. (2016)) y en particular su versión de LETKF, no ha sido probado en Argentina y es uno de los aportes originales de esta tesis.

Este trabajo estudiará el impacto en la asimilación de observaciones provenientes

de estaciones meteorológicas de alta resolución, observaciones de viento derivadas de satélite y radiancias satelitales polares y geoestacionarios en cielo claro en el contexto de los eventos de sistemas convectivos de mesoscala (SCM) debido a la importancia que cobran este tipo de eventos en la región. En particular se espera evaluar el aporte de cada una de las fuentes de datos en una región donde la red de observación convencional es bastante escasa y donde las contribuciones potenciales de sistemas de observación como redes de estaciones automáticas y observaciones de satélite son mayores.

El estudio de la asimilación de radiancias a nivel regional cobra aún mayor importancia en Sudamérica ya que no se conocen estudios realizados previamente. Por esta razón, en este trabajo se hará especial énfasis en la asimilación de radiancias y los controles de calidad necesarios para trabajar con estas observaciones. En primer lugar se estudiará el impacto de la asimilación de observaciones de satélites polares con sensores sensibles al espectro infrarrojo y microondas. Y en segundo lugar, se estudiará la implementación de la asimilación de observaciones del satélite geoestacionario GOES-16 y el impacto de asimilar observaciones en alta resolución espacial y temporal en un contexto regional.

Para alcanzar los objetivos de este trabajo, se realizaron distintos experimentos de asimilación de datos aplicados a un estudio de caso de un SCM que se desarrolló sobre el sur de Sudamérica durante el 22 y 23 de noviembre de 2018 durante el período de observación intensa de la campaña de campo RELAMPAGO.

Metodología {#metodologia}

0.6 El sistema de asimilación GSI

GSI por su nombre en inglés Gridpoint Statistical Interpolation, es un sistema de asimilación de datos de última generación, desarrollado inicialmente por el Environmental Modeling Center (EMC) del NCEP. Se diseñó como un sistema 3DVAR tradicional aplicado en el espacio de puntos de retícula de los modelos para facilitar la implementación de covarianzas anisotrópicas no homogéneas (Wu et al., 2002; Purser et al., 2003b, 2003a). Está diseñado para funcionar en varias plataformas computacionales, crear análisis para diferentes modelos numéricos de pronóstico, y seguir siendo lo suficientemente flexible como para poder manejar futuros desarrollos científicos, como el uso de nuevos tipos de observación, una mejor selección de datos y nuevas variables de estado (Kleist et al., 2009).

Este sistema 3DVAR sustituyó al sistema de análisis operativo regional de punto de retícula del NCEP por el Sistema de Predicción de Mesoscala de América del Norte (NAM) en 2006 y al sistema de análisis global Spectral Statistical Interpolation (SSI) usado para generar el Global Forecast System (GFS) en 2007 (Kleist et al., 2009). En los últimos años, GSI ha evolucionado para incluir varias técnicas de asimilación de datos para múltiples aplicaciones operativas, incluyendo 2DVAR (por ejemplo, el sistema Real-Time Mesoscale Analysis (RTMA); Pondeca et al. (2011)), la técnica híbrida EnVar (por ejemplo los sistemas de asimilación de datos para el GFS, el Rapid Refresh system (RAP), el NAM, el HWRF, etc.), y 4DVAR (por ejemplo, el sistema de asimilación de datos para el Sistema Goddard de Observación de la Tierra, versión 5 (GEOS-5) de la NASA; Zhu and Gelaro (2008)). GSI también incluye un enfoque híbrido 4D-EnVar que actualmente se utiliza para la generación del GFS.

Además del desarrollo de técnicas híbridas, GSI permite el uso de métodos de asimilación por ensambles. Para lograr esto, utiliza el mismo operador de las observaciones que los métodos variacionales para comparar el campo preliminar con las observaciones. De esta manera los exhaustivos controles de calidad desarrollados para los métodos

variacionales también son aplicados en la asimilación por ensambles. El código EnKF fue desarrollado por el Earth System Research Lab (ESRL) de la National Oceanic and Atmospheric Administration (NOAA) en colaboración con la comunidad científica. Contiene dos algoritmos distintos para calcular el incremento del análisis, el serial Ensemble Square Root Filter (EnSRF, Whitaker and Hamill (2002)) y el Local Ensemble Kalman Filter (LETKF, Hunt et al. (2007)) aportado por Yoichiro Ota de la Agencia Meteorológica Japonesa (JMA).

Para reducir el impacto de covarianzas espurias en el incremento que se aplica al análisis, los sistemas por ensamble aplican una localización a la matriz de covarianza K tanto en la dirección horizontal como vertical. Usa un polinomio de orden 5 para reducir el impacto de cada observación de manera gradual hasta llegar a una distancia límite a partir de la cual el impacto es cero. La escala de localización vertical se define como $-\log(P/P_{ref})$ y la escala horizontal usualmente se define en kilómetros. Estos parámetros son importantes a la hora de obtener un buen análisis y dependen de factores como el tamaño del ensamble y la resolución del modelo.

Otra característica importante del sistema es la implementación de un algoritmo de corrección de bias para las radiancias de satélites. La estimación hecha por el campo preliminar se compara las observaciones de radiancia para obtener la innovación. Esta innovación a su vez se utiliza para actualizar los coeficientes que permiten estimar el bias de las observaciones que se utilizó para obtener una innovación actualizada. Este proceso puede repetirse varias veces hasta que la innovación y los coeficientes de corrección de bias converjan. Este algoritmo se describe en mayor detalle en la Sección 1.1.1.

0.7 El modelo de transferencia radiativa CRTM

El Community Radiative Transfer Model (CRTM) es desarrollado por el Joint Center for Satellite Data Assimilation (JCSDA) y es usado para simular radiancias de sensores de microondas, infrarrojo y visible. Está compuesto por 4 modulos para el cálculo de la transmitancia, la emisión y reflección desde la superficie, la absorción por aerosoles y scattering y un modulo para resolver la transferencia radiativa.

El modelo forward del CRTM se utiliza para simular la radiancia medida por un satélite, que puede utilizarse para verificar la precisión y errores de la observación como así también los posibles cambios del sensor a largo plazo. El módulo de la matriz k se utiliza para calcular el Jacobiano (es decir, la derivada de la radiancia con respecto a parámetros geofísicos), que se utiliza para la asimilación de radiancias. El

uso de los módulos tangente-lineal y adjunto es equivalente al uso del módulo de la matriz k, y también se usa en algunas aplicaciones en la asimilación de la radiancia.

La transmitancia gaseosa describe la absorción generada por los gases en la atmósfera, de modo que se puede utilizar las observaciones en los sistemas de asimilación de datos y retrievals de temperatura del aire, la humedad y los gases traza como el CO₂, O₃, N₂O, CO y CH₄. El módulo de aerosoles es fundamental para identificar el tipo y la concentración de aerosoles para estudiar la calidad del aire. El módulo de nubes contiene las propiedades ópticas de seis tipos de nubes, proporcionando información del impacto radiativo generado por las nubes que es útil en la previsión meteorológica y los estudios climáticos. El modelo de superficie del CRTM incluye la emisividad/reflectividad estática de la superficie y se basa en un atlas que incluye varios tipos de superficie.

El CRTM es usado por GSI como operador de las observaciones de radiancia debido a que soportan más de 100 sensores distintos incluyendo todos los satélites meteorológicos. También permite la inclusión de nuevos sensores que solo requiere la generación de coeficientes espectrales y de transmitancia, componente clave en el desarrollo mejores pronósticos.

0.8 Caso de estudio

El caso de estudio utilizado en este trabajo corresponde a un SCM durante el 22 de noviembre de 2018. Previo al desarrollo de este SCM, el centro y norte de Argentina se encontraba inmerso en una masa de aire cálido y húmedo con altos valores de energía potencial disponible convectiva (CAPE), como lo muestra ERA 5 (Hersbach et al., 2018) en la Figura 2a. El 22 de noviembre de 2018 un frente frío cruzó el centro de Argentina (Figura 2b). Este frente frío desencadenó el desarrollo de células convectivas aisladas que rápidamente crecieron hasta convertirse en un SCM excepcionalmente grande (Figura 2d,e). Durante ese día, varias estaciones de superficie observaron actividad eléctrica, fuertes ráfagas de viento y lluvias intensas. Al norte de la región, el entorno cálido y húmedo contribuyó al desarrollo de convección aislada que finalmente creció y se fusionó con el SCM (Figura 2f). El SCM recorrió aproximadamente 2500 km de sur a norte, disipándose sobre Paraguay y el sur de Brasil después de 42 horas desde el inicio de su desarrollo.

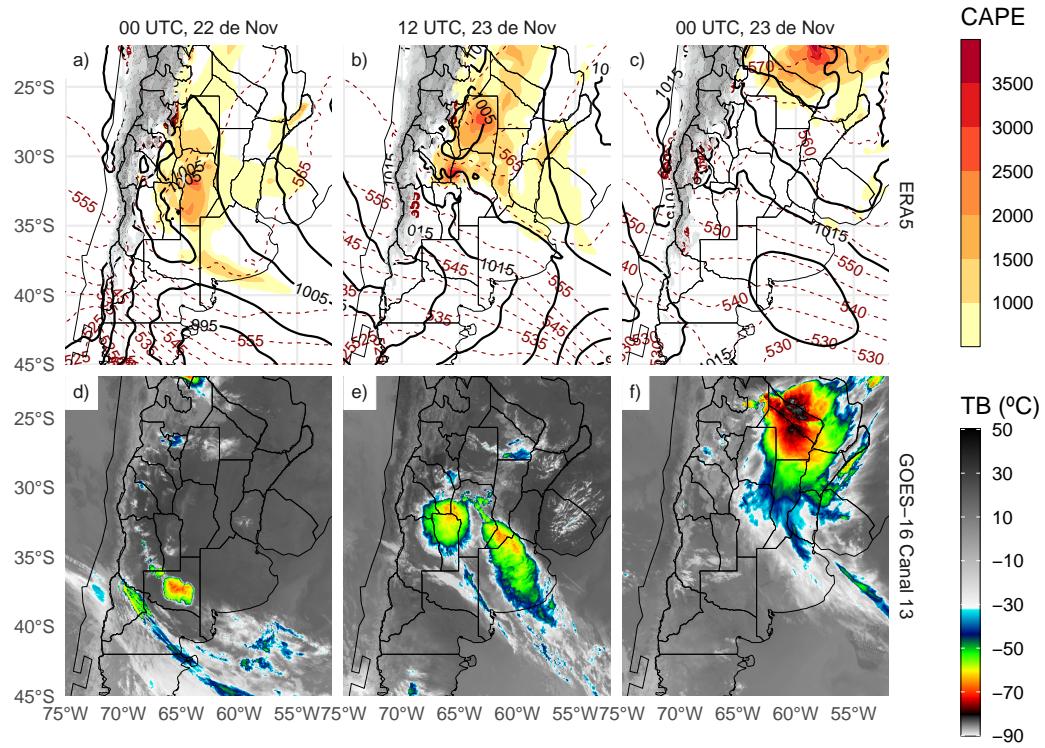


Figure 2: Presión a nivel del mar (hPa, contornos negros), espesor 1000-500 hPa (contornos rojos discontinuos) y energía potencial convectiva disponible (sombreada) según ERA5 y temperatura de brillo del canal 13 de GOES-16 para a,d) 00 y b,e) 12 UTC, 22 de Nov y c,f) 00 UTC, 23 de Nov.

0.9 Configuración del ensamble

Las simulaciones numéricas para el caso de estudio se realizan utilizando la versión 3.9.1 de modelo WRF (Skamarock et al. (2008)). Se utilizó una resolución horizontal de 10 km (150 x 200 puntos de retícula) y 37 niveles en la vertical con el tope del modelo en 50 hPa. Las condiciones iniciales y de contorno surgen del análisis del Global Forecast System (GFS) (resolución horizontal de 0,25° y resolución temporal de 6 horas; National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce (2015)). El dominio cubre la zona indicada en la Figura 3 para capturar el desarrollo del SCM durante el periodo simulado.

Los análisis se generaron utilizando la implementación LETKF (V1.3, Hunt et al. (2007)) que forma parte del sistema de análisis Gridpoint Statistical Interpolation (GSI V3.8; Shao et al. (2016)). Se utilizó un enfoque de actualización rápida con análisis horario y una ventana de asimilación centrada, lo que significa que se asimilaron

todas las observaciones dentro de ± 30 minutos del tiempo de análisis. Además, aas observaciones se asimilaron usando un enfoque 4D, es decir, comparándolas con el first guess más cercano que se genera en intervalos de 10 minutos. Para las observaciones de satelite, se utilizó el Community Radiative Transfer Model versión 2.3 (CRTM; Han et al. (2006)) como operador de observaciones para calcular las temperaturas de brillo simuladas por el modelo.

Utilizamos un conjunto de 60 miembros, cuya media al principio del ciclo de AD se inicializa utilizando el análisis deterministico del GFS al que se le suman perturbaciones aleatorias para generar el ensamble incial. Las perturbaciones se generaron como diferencias escaladas entre dos estados atmosféricos aleatorios obtenidos a partir de los datos del Reanálisis del Sistema de Predicción del Clima (CFSR) con una resolución horizontal de $0,5^\circ$ que tiene una con una evolución temporal suave (Necker et al., 2020; Maldonado et al., 2021). De este modo, preservamos el equilibrio casi hidrostático y geostrófico de las escalas mayores. Este método ayuda a evitar una subestimación del spread del ensamble (Ouaraini et al., 2015). Las perturbaciones también se aplicaron en los límites para mantener niveles adecuados de spread dentro del dominio del ensamble.

Además de las perturbaciones aleatorias en los límites laterales, se utilizó un esquema de multifísicas para representar mejor la incertidumbre en el modelo dentro del sistema de AD. Utilizamos 9 configuraciones diferentes que consisten en la combinación de 3 esquemas de convección húmeda (Kain-Fritsch (Kain, 2004), Grell-Freitas (Grell and Freitas, 2013), y Betts-Miller-Janjic (Janjić, 1994)) y 3 esquemas de capa límite planetaria (Esquema de la Universidad de Yonsei (Hong, Noh, et al., 2006), Esquema Mellor-Yamada-Janjic (Janjić, 1994), y Mellor-Yamada Nakanishi Niino (Nakanishi and Niino, 2009)). La distribución de estas parametrizaciones entre los 60 miembros del ensamble se muestra en la Tabla 1. Todos los miembros del conjunto utilizan las mismas parametrizaciones del modelo de superficie terrestre (Noah-MP, Chen and Dudhia (2001)), de microfísica (esquema de un solo momento de 6 clases del WRF (Hong, Kim, et al., 2006)) y de procesos radiativos (esquema de onda corta y onda larga del RRTMG (Iacono et al., 2008)). Para reducir el efecto de las correlaciones espurias en la estimación de las covarianzas de los errores de las observaciones, utilizamos un radio de localización horizontal de 180 km y un radio de localización vertical de 0,4 (en coordenadas de presión logarítmica) como en Dillon et al. (2021) para todos los tipos de observaciones. Se aplicó con un parámetro de inflación $\alpha = 0,9$ para mitigar el impacto de los errores de muestreo y para considerar los errores del modelo que no se tienen en cuenta en el enfoque mutifísica del ensamble (Whitaker and Hamill,

Table 1: Generación de los 60 miembros del ensamble multifísica como combinación de parametrizaciones de Cumulus y PBL

Cumulus	PBL		
	MYJ	MYNN2	YSU
BMJ	5, 14, 23, 32, 41, 50, 59	8, 17, 26, 35, 44, 53	2, 11, 20, 29, 38, 47, 56
GF	6, 15, 24, 33, 42, 51, 60	9, 18, 27, 36, 45, 54	3, 12, 21, 30, 39, 48, 57
KF	4, 13, 22, 31, 40, 49, 58	7, 16, 25, 34, 43, 52	1, 10, 19, 28, 37, 46, 55

2012).

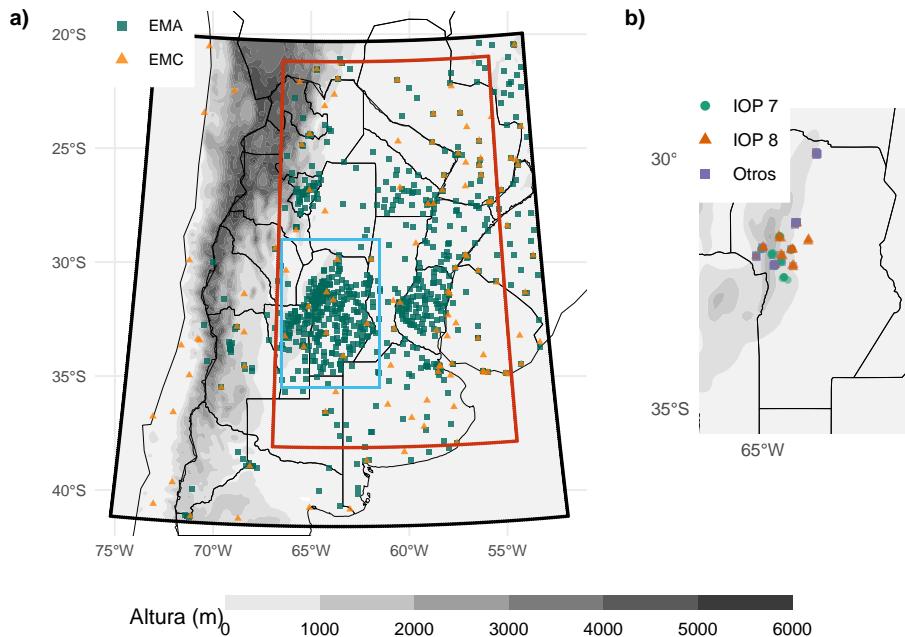


Figure 3: a) Dominio utilizado para las simulaciones (recuadro negro), dominio interior utilizado para la comparación entre experimentos (recuadro rojo), la región mostrada en b) (recuadro azul claro), y la ubicación de las Estaciones Meteorológicas Automáticas (EMA, cuadrados verdes) y las Estaciones Meteorológicas Convencionales (EMC, triángulos naranjas). b) Ubicación de los lanzamientos de radiosondeos durante RELAMPAGO. Los puntos verdes corresponden a los radiosondeos lanzados durante el IOP 7, los triángulos naranjas son radiosondeos lanzados durante el IOP 8, y los cuadrados morados son radiosondeos lanzados fuera de los períodos de medición intensiva. También se muestra la topografía en metros (sombreada).

0.10 Verificación: obs y metricas

0.10.1 Conjunto de datos de validación

Para evaluar el desempeño del sistema AD presentado en esta tesis, utilizamos los siguientes conjuntos de observaciones:

- **Datos horarios en niveles de presión de ERA5 de 1959 al presente (Hersbach et al., 2018):** Las variables de interés (temperatura del aire, humedad y viento) fueron interpoladas a la retícula del modelo para compararlas con el análisis de cada experimento.
- **Multi-Network Composite Highest Resolution Radiosonde Data (Earth Observing Laboratory, 2020):** radiosondeos en alta resolución lanzados desde varias ubicaciones durante el periodo de la campaña de campo de RELAMPAGO en junto con las radiosondeos operativos del SMN. Sólo se utilizaron para la validación los sondeos que no ingresaron en el sistema de asimilación. El periodo del experimento abarca las misiones IOP 7 y 8, durante las cuales se lanzaron 74 radiosondos en una pequeña zona cercana al centro del dominio experimental (Figura 3b).
- **IMERG Final Run (Huffman et al., 2018):** estimación de la precipitación a partir de datos de la constelación de satélites GPM (por su nombre en inglés Global Precipitation Measurement) con una resolución espacial de $0,01^{\circ}$ y una resolución temporal de 30 minutos para validar la habilidad de los pronósticos de 1 hora para representar la precipitación sobre el dominio.
- **Datos del Sistema Nacional de Radares Meteorológicos (SINARAME):** Se utilizaron observaciones de radar para realizar una evaluación cualitativa y visual de las características convectivas. Los datos provienen de 9 radares ubicados en el dominio y son provistos por la red de radares Doppler de doble polarización en banda C (de Elía et al., 2017) con una frecuencia temporal de 10 minutos. Para este trabajo se utilizó únicamente la máxima reflectividad de la columna (COLMAX) más cercana al momento de análisis.

0.10.2 Métodos de verificación

Se seleccionaron un conjunto de métricas para evaluar diferentes aspectos del análisis obtenido para cada experimento y los pronósticos inicializados a partir de ellos. Estas

métricas incluyen una validación de cómo se cuantifica la incertidumbre en el first guess y en el análisis, y cómo los diferentes experimentos se ajustan a un conjunto independiente de observaciones que no se asimilan.

Para evaluar la consistencia de la estimación de la incertidumbre en el first guess y en el análisis utilizamos la Reduced Centered Random Variable o RCRV (Candille et al. (2007)) que se define como:

$$RCRV = \frac{m - x_o}{\sqrt{\sigma_o^2 + \sigma^2}} \quad (6)$$

donde x_o es la observación asimilada y su error σ_o , m la media ensamble del análisis en el espacio de las observaciones, y la desviación estándar σ del ensamble. La media de $RCRV$ calculada sobre todas las realizaciones representa el sesgo de la media del conjunto con respecto a las observaciones normalizadas por la incertidumbre estimada:

$$meanRCRV = E[RCRV] \quad (7)$$

La desviación estándar de la $RCRV$ o $sdRCRV$ mide la concordancia de la dispersión del ensamble y el error de observación con respecto a la distancia entre la media del ensamble y las observaciones, y por lo tanto, la sobre o infradispersión sistemática del ensamble:

$$sdRCRV = \sqrt{\frac{1}{M-1} \sum_{i=1}^M (RCRV_i - meanRCRV)^2} \quad (8)$$

donde M es el tamaño del ensamble. Suponiendo que el error de observación fue estimado con precisión, un $sdRCRV > 1$ indica que el ensamble es infradispersivo (es decir, la distancia entre las observaciones y los pronósticos es mayor de lo esperado), y un $sdRCRV < 1$ indica que el conjunto es sobredispersivo (es decir, la distancia entre las observaciones y los pronósticos es menor de lo esperado). Un sistema consistente no tendrá sesgo ($meanRCRV = 0$) y una desviación estándar igual a 1 ($sdRCRV = 1$).

Para analizar el ajuste del first guess y el análisis a un conjunto de observaciones independientes, los radiosondeos de alta resolución de RELAMPAGO, se calculó la raíz del error cuadrático medio (RMSE) y el BIAS:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - O_i)^2} \quad (9)$$

$$BIAS = \frac{1}{N} \sum_{i=1}^N (X_i - O_i) \quad (10)$$

donde O y X representan las observaciones independientes y las simuladas respectivamente, y N es el tamaño de la muestra.

Para comparar la precipitación pronóstica a 1 hora con las estimaciones de precipitación de IMERG, calculamos el Fractions Skill Score (FSS, Roberts (2008)) para diferentes radios de influencia y umbrales de precipitación:

$$FSS = 1 - \frac{\sum_{i=1}^N (P_{xi} - P_{oi})^2}{\sum_{i=1}^N (P_{xi})^2 + \sum_{i=1}^N (P_{oi})^2} \quad (11)$$

donde P_{oi} es la proporción de puntos de reticula en el subdominio $i - th$ donde la que la precipitación acumulada observada es mayor que un umbral especificado. Siguiendo a Roberts et al. (2020), P_{xi} se calcula sobre el ensamble completo y cuantifica de la probabilidad de que la precipitación sea mayor al mismo umbral en cada punto de cuadrícula, y luego promediando sobre el subdominio $i - th$. El FSS se calculó a partir de la precipitación acumulada en ventanas móviles de 6 horas sumando los pronósticos de precipitación acumulada de 1 hora.

0.11 Recursos computacionales

Todos los experimentos corrieron en la supercomputadora Cheyenne del Centro Nacional de Investigación Atmosférica (NCAR) (Computational and Information Systems Laboratory, 2019). El posprocesamiento y análisis se realizó en [CIMA]. El análisis de datos se generó utilizando el lenguaje de programación R (R Core Team, 2020), utilizando los paquetes data.table (Dowle and Srinivasan, 2020) y metR (Campitelli, 2020), entre otros. Todos los gráficos se han realizado con ggplot2 (Wickham, 2009) y la versión final de la tesis se generó con knitr, rmarkdown (Xie, 2015; Allaire et al., 2019) y tesisdown [] .

Capítulo 1

Asimilacion de observaciones de estaciones meteorológicas automáticas, vientos derivados de satélite y radianzas de satelites polares

Este capítulo se analiza buscar contribuir a la cuantificación y comparación del impacto de asimilar observaciones de las estaciones meteorológicas automáticas de alta resolución, viento estimando por satélite y radiancias de cielo claro de satélites polaresa la hora de aplicar la asimilación de datos para mejorar los pronósticos numéricos de eventos severes) sobre Sudamérica, donde la red de observación convencional es bastante escasa y otras fuentes de información podrían potencialmente cubrir las áreas menos observadas.

1.1 Metodología

1.1.1 Observaciones asimiladas

Conventionales

Las observaciones convencionales (EMC) utilizadas forman parte del flujo de datos del Sistema Global de Asimilación de Datos (GDAS). Se asimilarán las observaciones convencionales incluidas en los archivos Binary Universal Form for Representation of

Meteorological Data (PREPBUFR) generados por NCEP. Los archivos PREPBUFR incluyen observaciones de superficie procedentes de 117 estaciones meteorológicas de superficie (EMS), barcos, y observaciones de altura procedentes de 13 sitios de lanzamiento de radiosondeos y aviones. Los triángulos naranjas de la Figura 3a indican la ubicación de las estaciones de superficie incluidas en este experimento. La frecuencia de estas observaciones varía entre 1 hora para las estaciones de superficie y 12/24 horas para las radiosondas. Las observaciones del viento en superficie sobre los océanos (ASCATW) proceden de los dispersómetros y también se incluyen en los archivos PREPBUFR.

La tabla 1.1 enumera todos los tipos de observación (presión en superficie, temperatura, humedad específica y viento) disponibles para cada fuente de observación, junto el error asociado para cada una definidos de acuerdo a la configuración por defecto de GSI. En algunos casos, el error varía con la altura y depende de la plataforma específica (avión y viento derivado del satélite). En cuanto al control de calidad aplicado a las observaciones convencionales, el operador de la observación realiza una primera comparación entre la innovación (la diferencia entre la observación y la observación simulada por el modelo para el campo preliminar) y un umbral predefinido que depende del error de observación (también incluido en la Tabla 1.1).

Red de Estaciones Meteorológicas Automáticas

También se asimilaron observaciones de 866 Estaciones Meteorológicas Automáticas (EMA) que forman parte de 17 redes de superficie públicas y privadas en la región. El conjunto de datos utilizado en este estudio fue obtenido del repositorio de datos de RELAMPAGO (Garcia et al., 2019). Estas estaciones se indican como cuadrados verdes en la Figura 3a. Tienen mayor cobertura espacial que las EMC y una frecuencia de muestreo de 10 minutos en la mayoría de los casos. Todas las estaciones miden temperatura, pero sólo 395 estaciones proporcionan humedad, 422 la presión y 605 el viento. Los errores de observación utilizados para asimilar estas observaciones son los mismos que para las observaciones de EMC (véase la tabla 1.1).

Vientos estimados por satélite

Las observaciones de viento derivadas de los satélites también se incluyen en los archivos PREPBUFR disponibles cada 6 hs, y consisten en estimaciones de GOES-16 (utilizando los canales visible, infrarrojo y de vapor de agua) y METEOSAT 8 y 11 (utilizando los canales visible y de vapor de agua). Debido al dominio elegido y la

Table 1.1: Características de las observaciones asimiladas: El código de cada tipo de observación y su fuente, las variables disponibles, el error de observación y los umbrales de control de calidad utilizados.

Código	Plataforma	Variable	Error	Umbral de error
CSWS AWS	Surface weather stations	Pressure	1-1.6 hPa*	3.6 hPa
		Temperature	1.5 K	7 K
		Specific humidity	20 %	8 gKg ⁻¹
		Wind	2.2 ms ⁻¹	6 ms ⁻¹
ADPUPA	Radiosondes	Pressure	1.1-1.2 hPa**	4 hPa
		Temperature	0.8-1.5 K*	8 K
		Specific humidity	20 %	8 gKg ⁻¹
		Wind	1.4-3 ms ⁻¹ *	8 ms ⁻¹
AIRCFT	Aircrafts	Temperature	1.47-2.5 K ⁺	7 K
		Wind	2.4-3.6 ms ⁻¹ +	6.5-7.5 ms ⁻¹ +
ASCATW	Advanced Scatterometers	Wind	1.5 ms ⁻¹	5 ms ⁻¹
SFCSHP	Ships and Buoys	Pressure	1.3 hPa	4 hPa
		Temperature	2.5 K	7 K
		Specific humidity	20 %	8 gKg ⁻¹
		Wind	2.5 ms ⁻¹	5 ms ⁻¹
SATWND	Satellite-derived winds	Wind	3.8-8 ms ⁻¹ +	1.3-2.5 ms ⁻¹ +

* El error de la observación varía con la altura.

** Observaciones por encima de 600 hPa son rechazadas.

+ El error de la observación depende del tipo de reporte.

cobertura de estos satélites, GOES-16 es la principal fuente de observaciones de vientos derivados de los satélites (99 % de las observaciones). Los errores de observación utilizados para asimilar estas observaciones siguen la configuración por defecto del GSI y se indican en la Tabla 1.1.

Radiancias de satélite

En este capítulo se utilizaron radiancias de satélites disponibles a través del flujo de datos del GDAS, que incluye observaciones en el espectro infrarrojos y microondas. Los sensores incluidos son el Advanced Microwave Sounding Unit - A (AMSU-A), Microwave Humidity Sounder (MHS), y 2 sensores multiespectrales; el Atmospheric Infrared Sounder (AIRS) y el Infrared Atmospheric Sounding Interferometer (IASI) ubicados en distintas plataformas (ver Tabla 1.2). Dado que el dominio regional se

encuentra en latitudes medias y que los satélite de interés están en órbitas polares, cada sensor escanea la zona sólo dos veces al día con una cobertura espacial que depende de la franja de cobertura satélite. Por esta razón, el número de observaciones de los satélites varía significativamente entre en cada ciclo de asimilación. En particular, los sensores multiespectrales proporcionaron entre 100 y 1000 observaciones por cada escaneo cada 12 horas, contribuyendo al 88 % de la cantidad total de radiancias asimiladas en los experimentos. La ubicación vertical de cada observación de radiancia se estimó como el nivel del modelo en el que se maximizaba su función de peso calculada por el CRTM. Los sensores multiespectrales tienen una buena cobertura vertical y son capaces de captar desde la baja troposfera hasta la baja estratosfera.

Los canales elegidos para la asimilación y sus errores asociados se definieron teniendo en cuenta el tope del modelo (50 hPa). El preprocesamiento de datos, que es un paso esencial en la asimilación de radiancias, se realizó dentro del sistema GSI para cada sensor específicamente. En primer lugar, se aplicó un thinning espacial de los datos utilizando una retícula de 60 km siguiendo a Singh et al. (2016), Jones et al. (2013) y Lin et al. (2017), donde las observaciones que se van a asimilar se eligen en función de su distancia a los puntos de la retícula del modelo, la calidad de la observación (basada en la información disponible sobre la calidad de los datos) y el número de canales disponibles (para el mismo píxel y sensor) que han pasado el control de calidad. Además, el algoritmo prefiere las observaciones sobre el mar a las de la tierra o nieve (Hu et al., 2018).

Luego del thinning, se aplica la corrección de bias. La corrección de bias (CB) tiene una componente dependiente de características termodinámicas del aire y otra dependiente del ángulo de escaneo (Zhu et al., 2014) y se calcula como una polinomio lineal de N predictores $p_i(x)$, con coeficientes asociados β_i . Por lo tanto, la temperatura de brillo corregida por bias (BT_{cb}) puede obtenerse como

$$BT_{cb} = BT + \sum_{i=0}^N \beta_i p_i(x) \quad (1.1)$$

GSI tiene un término de corrección de bias constante ($p_0 = 1$) mientras que los términos restantes y sus predictores son el contenido de agua líquida de las nubes (CLW), la tasa de cambio de temperatura con la presión, el cuadrado de la tasa de y la sensibilidad de la emisividad de la superficie. El bias dependiente del ángulo de escaneo se modela como un polinomio de 4º orden (Zhu et al., 2014).

En el sistema GSI, los coeficientes β_i se entrena utilizando un método de estimación variacional que genera los β_i que proporciona el mejor ajuste entre la simulación y las

Table 1.2: Lista de los sensores disponibles cada plataforma, el número de canales aceptados para su asimilación y el porcentaje de observaciones asimiladas calculado sobre todas las observaciones de radiancias y todos los ciclos de asimilación.

Sensor	Plataforma	Canales asimilados	Porcentaje sobre el total
AIRS	AQUA	52	31.63 %
	NOAA15	2	3.31 %
	NOAA18	2	4.45 %
AMSUA	METOP-A	2	2.08 %
	METOP-B	66	52.72 %
IASI	METOP-A	68	3.47 %
	NOAA19	2	0.68 %
	METOP-B	3	0.8 %
MHS	METOP-B	3	0.85 %

observaciones. Los coeficientes se inicializaron a las 18 UTC del 18 de noviembre de 2018 con los coeficientes generados para la misma hora por el sistema GFS. El sistema de asimilación se configuró para utilizar una varianza de error de los coeficientes constante de 0,01 para evitar grandes ajustes en los coeficientes estimados en cada momento.

En nuestros experimentos, sólo se utilizan observaciones de cielo claro. Para las radiancias de microondas, las observaciones potencialmente contaminadas por nubes se detectan utilizando los índices de dispersión y del Liquid Water Path (LWP) (Zhu et al., 2016; Weston et al., 2019). Para los canales infrarrojos, las observaciones contaminadas por nubes se detectan utilizando el perfil de transmitancia calculado dentro de los algoritmos CRTM. Además, GSI comprueba la diferencia entre las observaciones y la temperatura de brillo simulada en cada nivel para detectar los píxeles nublados. Por otro lado, el control de calidad de GSI aplicado a los sensores infrarrojos busca observaciones sobre agua con un gran ángulo cenital (más de 60°) para rechazar los canales cercanos al rango visible que pueden estar contaminados por reflexión. Para las observaciones en el infrarrojos y de microondas también realiza una chequeo de la emisividad para detectar observaciones contaminadas por efecto de la superficie.

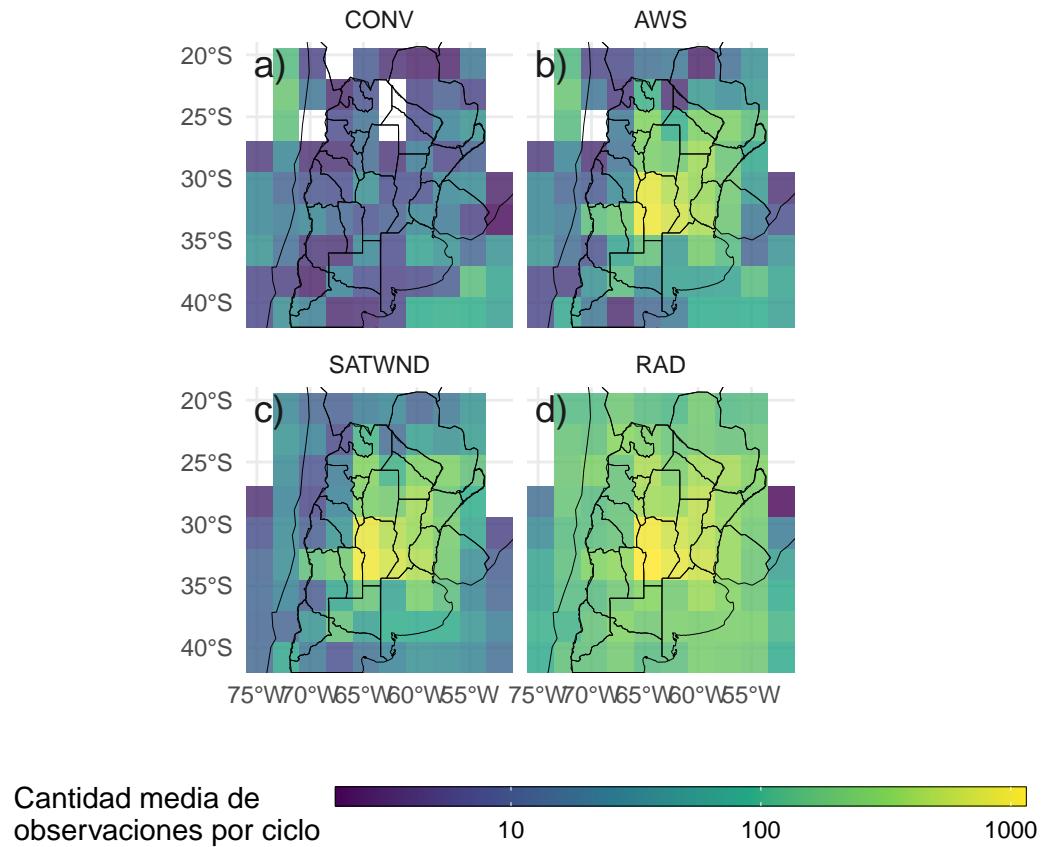


Figure 1.1: Distribución espacial horizontal media de las observaciones disponibles en cada ciclo de asimilación para los experimentos a) CONV, b) AWS, c) SATWND y d) RAD calculados sobre cajas de $2,5^{\circ} \text{circ}$.

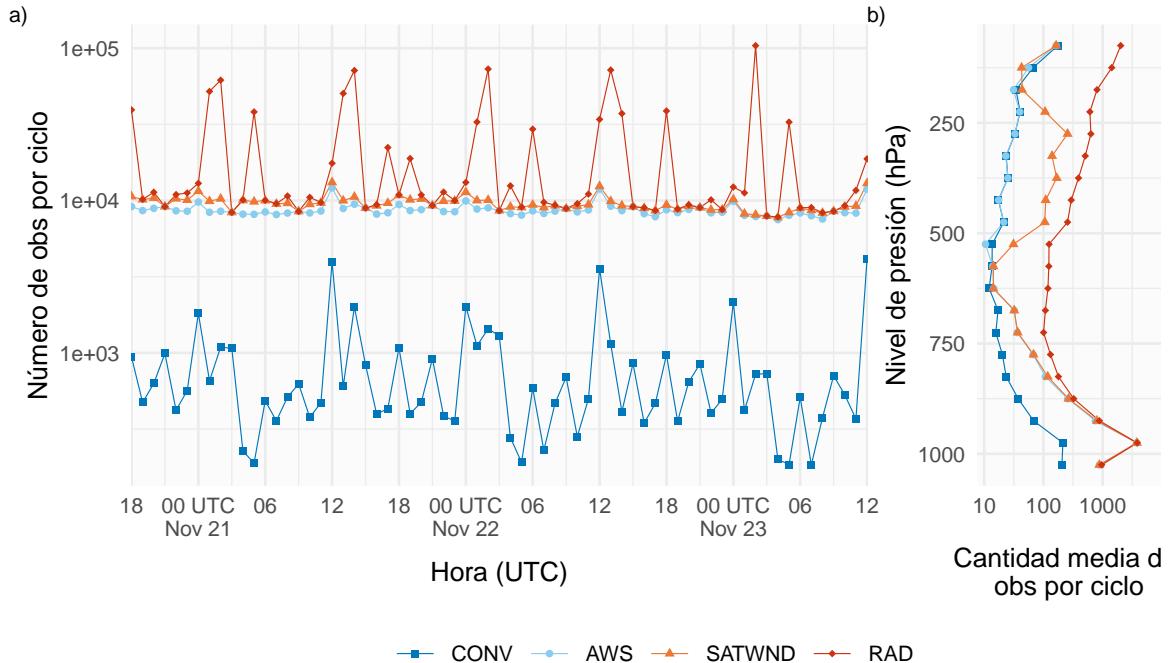


Figure 1.2: a) Número de observaciones asimiladas por ciclo y b) número medio de observaciones asimiladas por ciclo dividido en capas verticales de 50 hPa de espesor para los experimentos CONV (cuadrados azules), AWS (puntos celestes), SATWND (triángulos naranjas) y RAD (diamantes rojos).

1.1.2 Configuración de los experimentos

Para investigar el impacto de las diferentes observaciones en el análisis, se realizaron cuatro experimentos de asimilación de datos utilizando diferentes conjuntos de observaciones (Tabla 1.3). El experimento CONV utiliza únicamente las observaciones convencionales de PREPBUFR. En un segundo experimento, denominado AWS, se asimilan todas las observaciones incluidas en CONV más las observaciones de EMA con 10 minutos de frecuencia. En el tercer experimento, denominado SATWND, se asimilan las observaciones del experimento AWS junto con los vientos estimados por satélite. Por último, un cuarto experimento, denominado RAD, asimila todas las observaciones mencionadas previamente más las radiancias en cielo claro procedentes de los sensores a bordo de los satélites de órbita polar, como se describe en la sección 1.1.1. La distribución horizontal del promedio de observaciones asimiladas en cada ciclo de asimilación de cada experimento se muestra en la Figura 1.1. El mayor número de observaciones asimiladas sobre el centro y el este del dominio corresponde a las observaciones de EMA. En la Figura 1.2a se muestra el número de observaciones asim-

Table 1.3: Tipos de observaciones asimiladas en cada experimento.

Obs type	CONV	AWS	SATWNDRAD
Convencional (PREPBUFR)	x	x	x
Convencional (EMA)		x	x
Viento estimado por satélite		x	x
Radiancias			x

iladas a lo largo del tiempo. Los máximos locales a las 12 y a las 00 UTC encontrados principalmente en CONV corresponden a las observaciones obtenidas con los sondeos operativos. La fuerte variabilidad en el número de observaciones de radiancias por ciclo es también notable y depende de la cobertura de cada satélite. Los máximos a las 13-14 y 01-02 UTC en RAD corresponden a la contribución de los sensores multiespectrales. La distribución vertical del número medio de observaciones por ciclo (Figura 1.2b) muestra un máximo en niveles bajos debido a las observaciones de EMA. Los vientos estimados por satélite tienen un máximo en la troposfera alta (entre 500-250 hPa). Por encima de 850 hPa, la mayoría de las observaciones corresponden a radiancias.

Todos los experimentos de asimilación comienzan a las 18 UTC del 20 de noviembre de 2018 y continúan hasta las 12 UTC del 23 de noviembre (un total de 67 horas/ciclos de asimilación). El ensamble inicial de 60 miembros se genera según se explica en la sección ?? y todos los experimentos se inician con un periodo de spin up sin asimilación de observaciones entre las 12 UTC y las 18 UTC de Nov, 20 (Figura 1.3).

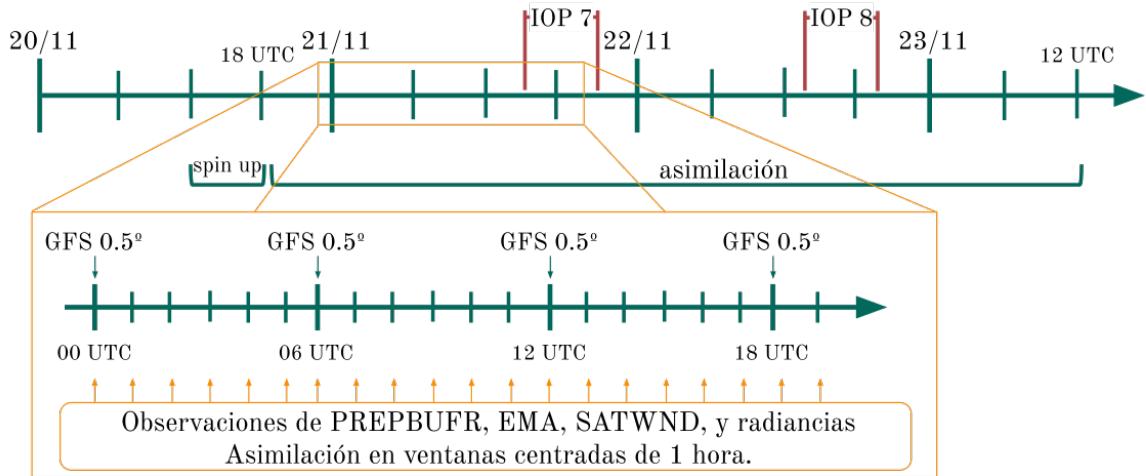


Figure 1.3: Diagrama de los ciclos de asimilación entre las 18 UTC del 20 de noviembre y las 12 UTC del 23 de noviembre más el período de spin up de 6 horas. La sección ampliada muestra la asimilación horaria que se realiza dentro de una ventana centrada de una hora y la incorporación de condiciones de borde del GFS cada 6 horas. Se muestran las dos misiones IOP de la campaña RELAMPAGO.

1.2 Resultados

1.2.1 Consistencia del ensamble

Para investigar la capacidad de la media del ensamblaje del first-guess para ajustarse a las observaciones teniendo en cuenta las incertidumbres del pronóstico y de las observaciones, se calcularon el *meanRCRV* y el *sdRCRV* para el experimento RAD. Como este experimento asimila todos los tipos de observaciones utilizados en este trabajo, es posible analizar la coherencia del ensamblaje comparándolo con cada tipo de observación. La figura 1.4 muestra el *sdRCRV* para las observaciones de superficie promediadas en una retícula de $2,5^\circ$. El *sdRCRV* para las observaciones del viento (Figura 1.4a) es cercano a 1, lo que sugiere una buena concordancia entre la dispersión del ensamblaje, el error de pronóstico y el error de observación. Para la temperatura (Figura 1.4b), los resultados son similares, salvo que para algunas zonas del oeste del dominio donde el *sdRCRV* puede llegar a ser de 4,5. Estos valores más altos de *sdRCRV* pueden estar asociados a errores sistemáticos derivados de las diferencias entre la topografía del modelo y al altura a la que están ubicadas las observaciones. La circulación de pequeña escala asociada al terreno complejo y podrían no estar bien resueltas por el modelo también podría contribuir a aumentar la distancia entre el

pronóstico y las observaciones. Estos aspectos no suelen ser captados por la dispersión del ensamble, a menos que se utilice un esquema de inflación dependiente del espacio bien ajustado, lo que conduce a mayores valores de $sdRCRV$.

La figura 1.5 muestra la media y el desvío estándar del RCRV para las observaciones de altura. Las figuras 1.5a-b muestran las estadísticas de RCRV para los sondeos (ADPUPA) y aviones (AIRCAR y AIRCFT). Tanto ADPUPA como AIRCFT muestran, en general, un buen acuerdo entre la dispersión del ensamble y el error de observación. Como las observaciones de sondeo y sus errores asociados son considerados de buena calidad, este resultado indica que el ensamble tiene una dispersión adecuada. AIRCAR presenta un perfil irregular con valores $sdRCRV$ que sugieren que el error de este tipo de observación está sobreestimado. ADPUPA y AIRCAR presentan un perfil de $meanRCRV$ cercano a cero en los niveles medios y altos. En niveles bajos, el perfil $meanRCRV$ es positivo, mostrando un bias cálido presente en el modelo, característica ya estudiada en Ruiz et al. (2010) y Dillon et al. (2021).

Las observaciones de vientos estimadas por satélites varían en número dependiendo del satélite y del nivel. En la Figura 1.5c sólo se incluye el $RCRV$ calculado con al menos 100 observaciones disponibles para cada sensor y nivel. En niveles bajos, donde no hay muchas observaciones disponibles, los perfiles de $meanRCRV$ y $sdRCRV$ muestran una mayor desviación del comportamiento esperado con un bias negativo, y una posible sobreestimación del error de observación. Las estimaciones del viento derivadas de los canales de vapor de agua son abundantes por encima de 500 hPa, donde su bias es cercano a cero. La única excepción son las observaciones de EUMETSAT que contribuyen muy poco en la región.

Los perfiles de $meanRCRV$ calculados a partir de las observaciones de radiancias (Figura 1.5d) no muestran casi ningún bias y lo mismo ocurre si se calcula el $meanRCRV$ sobre cada canal de cada sensor (no incluido en el trabajo). Esto indica que el algoritmo de corrección del bias funciona como se esperaba. Los valores de $sdRCRV$ son inferiores a 1 para todos los sensores, posiblemente debido a una sobreestimación de los errores de observación para reducir la influencia de las observaciones potencialmente erróneas.

En general, estos resultados indican que la dispersión del ensamble es coherente con el error de pronóstico a corto plazo y que los errores sistemáticos son relativamente pequeños para la mayoría de los tipos de observación utilizados en este trabajo. Además, estos resultados sugieren que el parámetro de inflación $\alpha = 0,9$ es adecuado para el sistema.

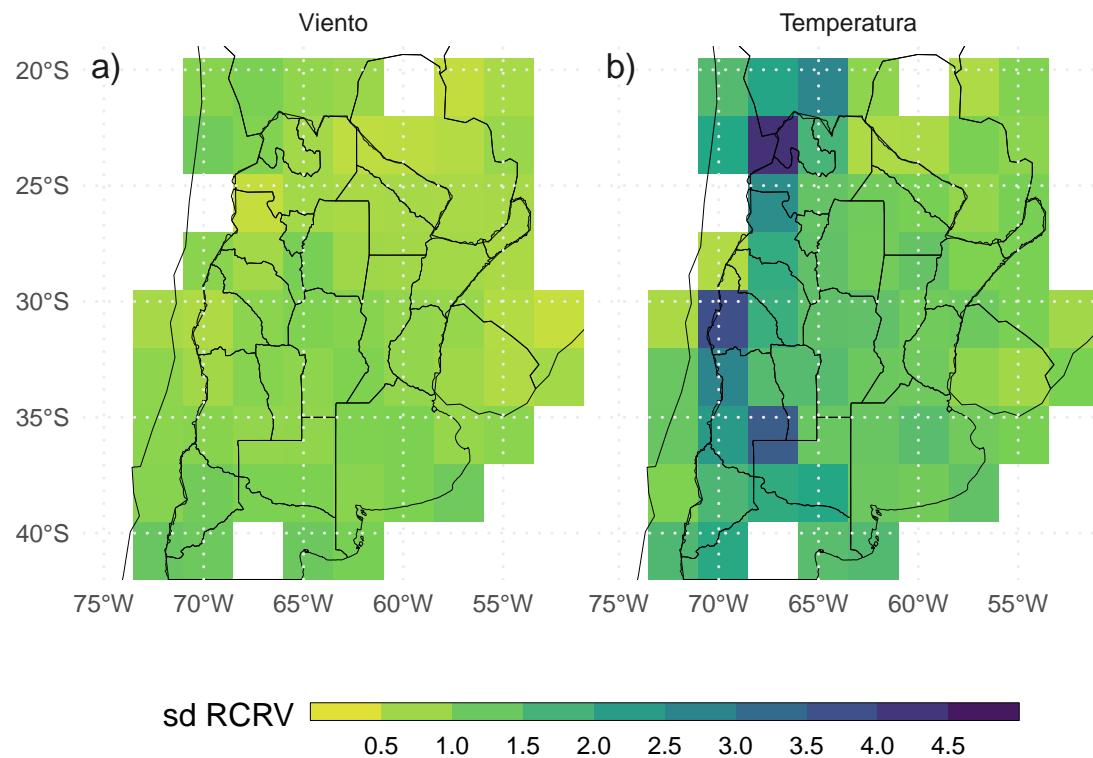


Figure 1.4: $sdRCRV$ calculado para observaciones de superficie (de PREPBUFR y EMA) de a) viento, y b) temperatura promediados en cajas de 2.5° para el experimento RAD. Se usaron las observaciones agregadas de cada ciclo de asimilación horario paa todo el periodo del experimento.

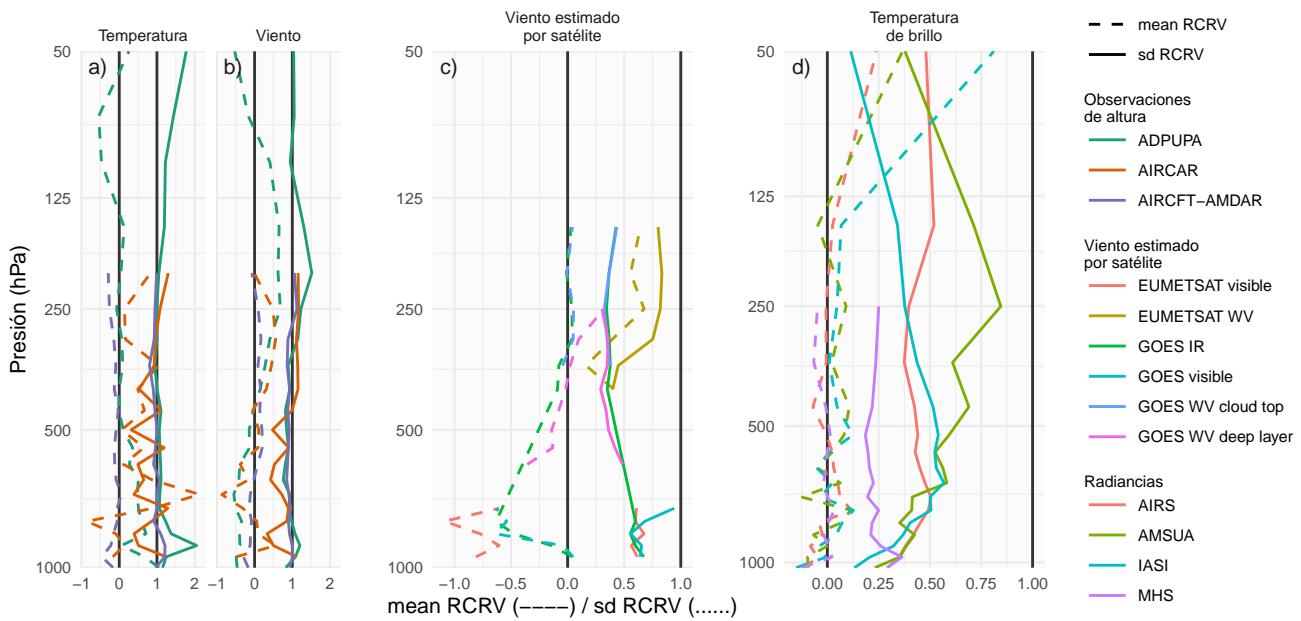


Figure 1.5: Perfiles verticales de *meanRCRV* (línea punteada) y *sdRCRV* (línea sólida) para observaciones de a) temperatura y b) viento de sondeos y aviones, c) viento estimado por satélites, y d) temperatura de brillo para el experimento RAD. Se usaron las observaciones agregadas de cada ciclo de asimilación horario para todo el periodo del experimento.

1.2.2 Impacto de la asimilación de las distintas observaciones

Esta sección presenta el impacto de la asimilación de diferentes tipos de observación sobre algunas variables que son particularmente relevantes para el desarrollo de convección húmeda profunda. El análisis se realiza sobre un dominio más pequeño (recuadro rojo en la Figura 3a) que corresponde a la región directamente afectada por el SCM. Las figuras 1.6a-c muestran la diferencia entre el perfil vertical de temperatura promediado espacialmente para el análisis de los experimentos. Al promediar las diferencias entre dos experimentos se puede aislar el impacto sistemático producido por los diferentes sistemas de observación sobre la situación analizada. Durante el primer día, la asimilación de las observaciones de EMA da como resultado el desarrollo de una PBL más fría. Este efecto de enfriamiento tiene un claro ciclo diurno, siendo más fuerte durante la noche (Figura 1.6a). Durante el segundo día del experimento, el impacto de las observaciones de EMA se extiende a la troposfera media y alta, coincidiendo con la fase madura del SCM. La diferencia positiva que se observa en AWS-CONV entre 500 y 200 hPa se produce por el desarrollo de convección más intensa en AWS en

comparación con CONV. Este es un buen ejemplo de cómo la información en niveles bajos de la atmósfera proporcionada por las estaciones meteorológicas de superficie puede extenderse rápidamente a la troposfera en presencia de convección húmeda profunda. Aunque la circulación media y alta puede tener un impacto importante en la organización y evolución del SCM sobre la región, los vientos estimados por satélite no tuvieron un impacto apreciable en la temperatura y humedad media (Figura 1.6b-e), posiblemente debido a alto valor del error de estas observaciones utilizados para la asimilación. Durante el primer día del experimento, la asimilación de las radiancias produce un efecto de calentamiento en la PBL que compensa parcialmente el efecto de enfriamiento de las observaciones de EMA (Figura 1.6c). No se encuentra un impacto sistemático claro por encima de la PBL durante este periodo para estas variables. Durante el segundo día, el impacto de las observaciones de radiancia se observa en toda la troposfera con una distribución que es similar al impacto encontrado en el experimento AWS pero con signo opuesto.

Comparando la representación de la humedad específica en los experimentos (Figuras 1.6d-f), el impacto de la asimilación de EMA, que tienen una resolución espacial y temporal mayor, es muy importante en niveles bajos (Figura 1.6d). La PBL en el experimento AWS es sistemáticamente más húmeda que en el experimento CONV, especialmente durante la noche. El aumento de la humedad en niveles bajos gracias a la asimilación de observaciones de una red de superficie más densa corrige el bias seco reportado previamente en el modelo WRF sobre la región (Casaretto et al., 2022, pp. @matsudo2021, @ruiz2010). El humedecimiento de la PBL es impulsado principalmente por la covarianza entre la temperatura y la humedad específica dentro de la PBL. En el experimento y sobre el centro del dominio, esta covarianza se mantiene negativa, generando un aumento de la humedad en niveles bajos a medida que las observaciones introducen correcciones negativas de temperatura. Como en el caso de la temperatura, el impacto sistemático de los vientos estimados por satélite sobre la humedad es pequeño (Figura 1.6e). La figura 1.6f muestra que las radiancias reducen la humedad en niveles medios y bajos durante el primer día del experimento. El efecto de secamiento se extiende a niveles medios y bajos durante el segundo día del experimento, coincidiendo con el desarrollo del SCM entre las 00 y las 12 UTC del 22 de noviembre.

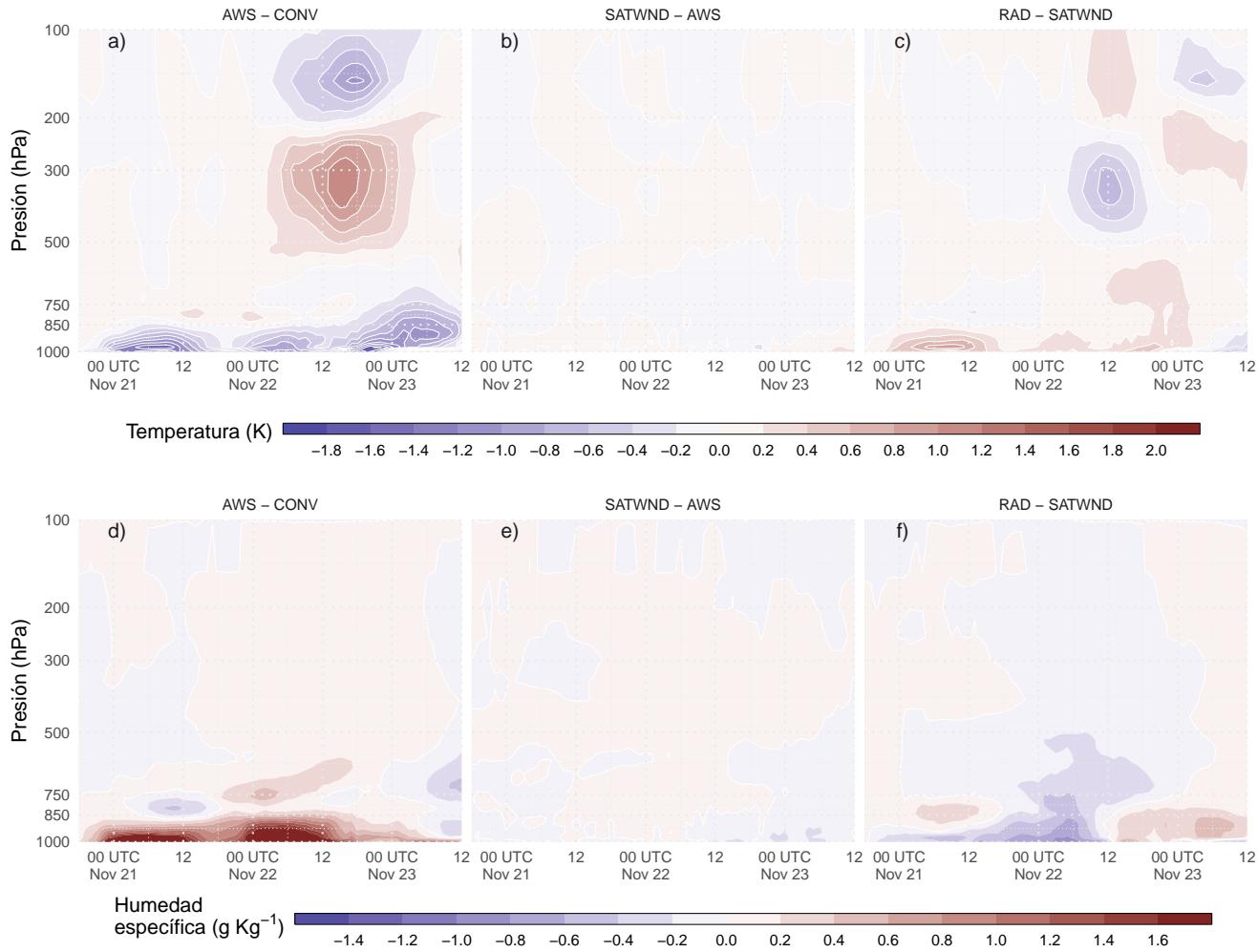


Figure 1.6: Diferencia entre la media del ensamble para la media del ensamble del análisis a) y d) AWS-CONV, b) y e) SATWND-AWS, y c) y f) RAD-SATWND para los perfiles verticales espacialmente promediados de la temperatura (a, b y c, en K) y la humedad específica (d, e y f en $g \text{ kg}^{-1}$) calculados sobre el dominio interior (recuadro rojo en la Figura 3a) para cada ciclo de análisis.

El impacto en las componentes del viento se muestran en la Figura 1.7, junto con la correspondiente componente del viento promediada en el experimento con el mayor número de observaciones asimiladas (por ejemplo, la Figura 1.7a muestra la diferencia de viento zonal entre AWS y CONV y el viento zonal para AWS). La asimilación de EMA produce un viento más del este y menos del norte en niveles bajos durante los dos primeros días de análisis (Figuras 1.7a,b). Existe un ciclo diurno asociado al impacto de las estaciones meteorológicas automáticas sobre el viento meridional (Figura 1.7d) con una mayor reducción del viento del norte durante las horas nocturnas.

Esto indica que las observaciones en superficie están reduciendo la intensidad del jet de capaz bajas presente en el entorno preconvectivo. Después de las 18 UTC del 22 de noviembre, se observa el efecto contrario cuando el SCM se desplaza por el dominio hacia el noreste. Tras el inicio de las celdas convectivas, el impacto sistemático en el campo de viento es mayor en los niveles medios y altos (Figuras 1.7d, f). Durante los días 22 y 23 de noviembre el impacto de asimilar EMA produce un aumento del viento del norte en los niveles superiores de la tropósfera. Esto podría ser una consecuencia de un SCM más intenso que produce un aumento del flujo de salida del lado polar de la tormenta. Aunque las observaciones de viento estimada por satélites producen el mayor impacto en niveles medios y altos, donde el número de observaciones es mayor, el impacto sistemático es en general menor que el producido por la asimilación de datos de EMA (Figuras 1.7b, e). La razón del pequeño impacto observado en SATWND podría estar asociada al valor del error de observación utilizado para la asimilación de estas observaciones.

La asimilación de las radiancias produce una reducción del viento del oeste con respecto a SATWND en niveles bajos y altos (Figura 1.7c). Para el viento meridional, estas observaciones producen un aumento 1 ms^{-1} en el flujo del norte en niveles bajos, opuesto al generado por la asimilación de las observaciones de EMA durante la noche, entre las 03 y las 12 UTC, previas al desarrollo del SCM (Figura 1.7f). En niveles superiores y durante los días 22 y 23 de noviembre, el impacto medio de la asimilación de radiancias es una disminución de la velocidad del viento. El campo de viento meridional a 200 hPa en diferentes momentos muestra que el flujo de salida del SCM es aún más intenso que en los otros experimentos, mientras que el viento sur por delante del SCM también aumenta produciendo una reducción media del viento del norte (Figura 1.7f).

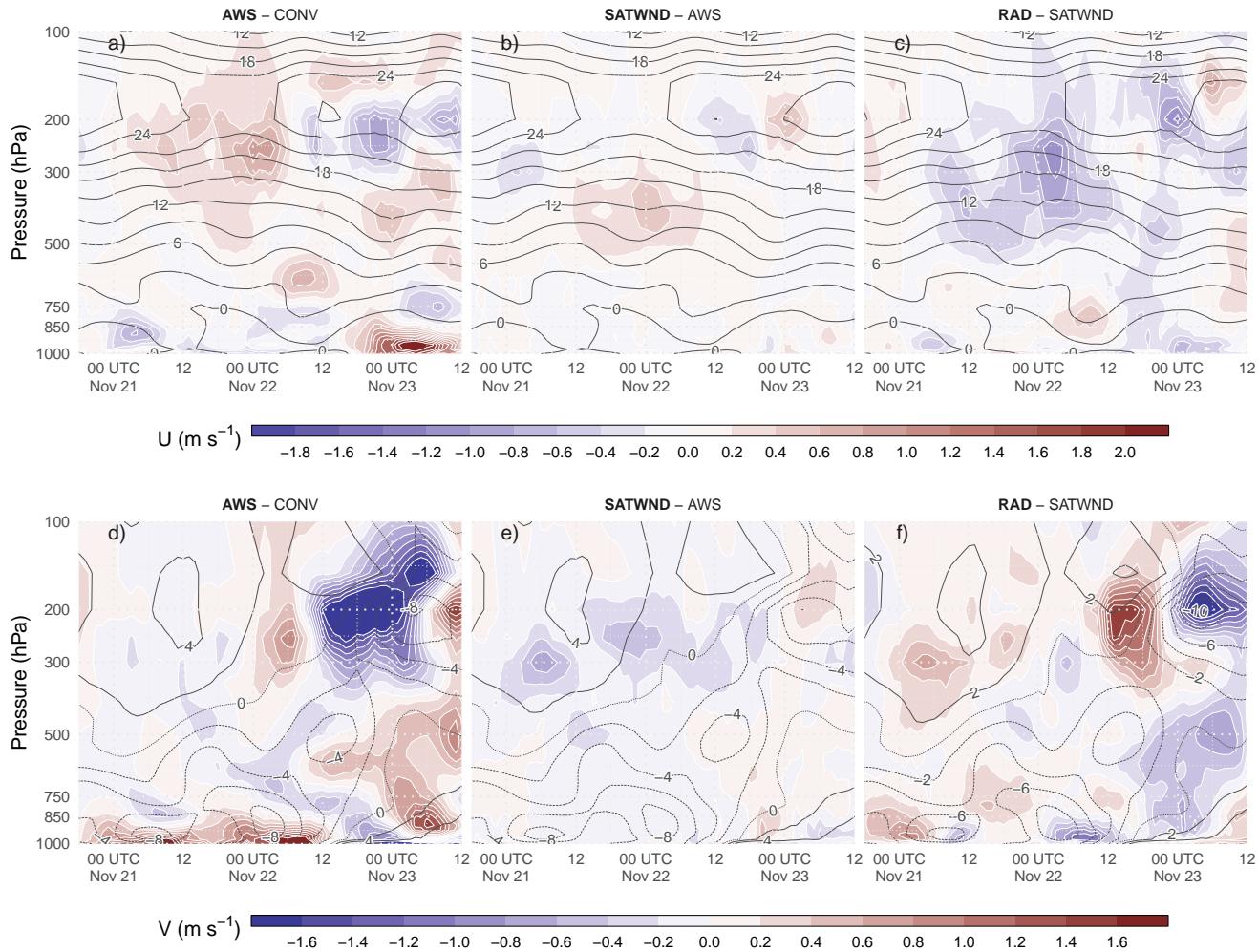


Figure 1.7: Diferencia entre la media del ensamble para la media del ensamble del análisis a) y d) AWS-CONV, b) y e) SATWND-AWS, y c) y f) RAD-SATWND para los perfiles verticales espacialmente promediados del viento zonal (a, b y c, en K) y viento meridional (d, e y f en $g \text{ kg}^{-1}$) calculados sobre el dominio interior (recuadro rojo en la Figura 3a) para cada ciclo de análisis. Los contornos negros corresponden al viento zonal y meridional para (a,d) AWS, (b,e) SATWND, y (c,f) RAD ya que son los experimentos tienen más observaciones asimiladas en cada panel.

La diferencia entre ERA5 (Hersbach et al., 2018) y la media del ensamble del análisis para cada experimento se compara en la Figura 1.8, apoyando los resultados observados en las Figuras 1.6 y 1.7. En concreto, la Figura 1.8a muestra un bias cálido en los niveles bajos (es decir, CONV es más cálido que ERA5) que disminuye en la Figura 1.8b cuando se asimilan las observaciones de EMA. En la misma dirección, la Figura 1.6a muestra una diferencia negativa entre AWS y CONV, lo que significa que las

observaciones de EMA están enfriando los niveles bajos de la atmósfera. Comparando ERA5-RAD (Figura 1.8d), hay un pequeño aumento del bias cálido, asociado al calentamiento producido por la asimilación de las observaciones de radiancia, como se muestra en la Figura 1.6c. Un efecto similar puede observarse en la humedad específica, las observaciones de EMA corrigen parcialmente el bias seco presente en la Figura 1.8e y la asimilación de las observaciones de radiancia reduce el impacto positivo de EMA. El impacto sobre las componentes del viento es menor, por lo que sólo se incluye el viento meridional en las figuras 1.8i-l, que muestran que las observaciones de radiancia son las principales responsables del impacto positivo observado en el análisis al reducir la distancia ERA5-RAD, en particular durante la fase madura del MCS. En general, los ajustes asociados a la asimilación de las observaciones de radiancia y de EMA conducen a un análisis más cercano a los reanálisis ERA5.

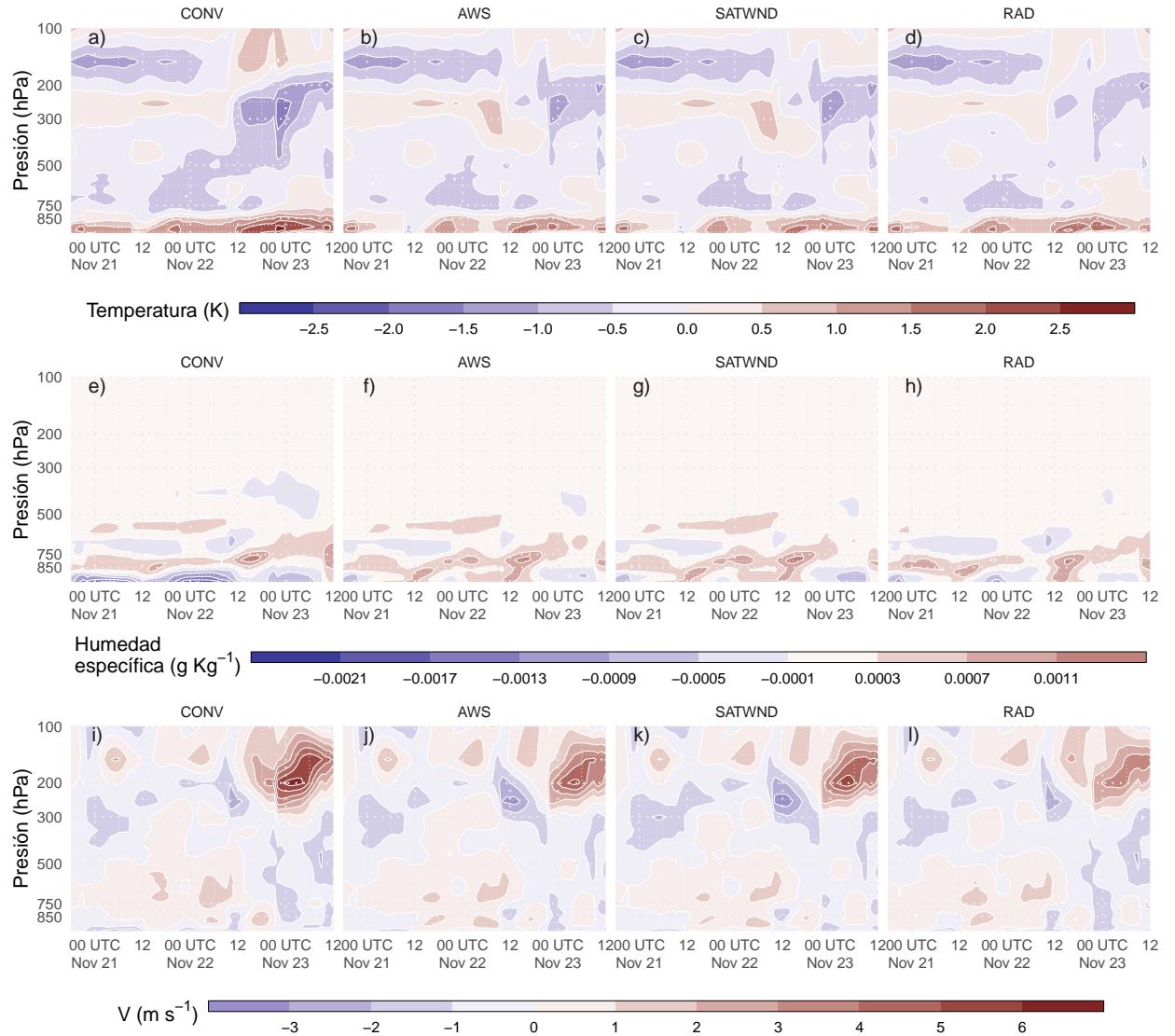


Figure 1.8: Diferencia entre la media del ensamble del análisis de cada experimento y el ERA5 para los perfiles verticales espacialmente promediados de la temperatura del aire (K, a–d), la humedad específica ($g\text{ Kg}^{-1}$, e–h) y el viento meridional ($m\text{ s}^{-1}$, i–l) calculados sobre el dominio interior (recuadro rojo en la Figura 3a) para cada ciclo de análisis.

Para investigar cómo los cambios en la PBL pueden modificar el entorno preconvectivo, se compara la distribución horizontal media de análisis del viento meridional del norte (para los primeros 7 niveles sigma), el agua precipitable, la temperatura en niveles bajos y el CAPE. A las 00 UTC del 22 de noviembre (después de 30 ciclos de asimilación) las primeras células convectivas se estaban desarrollando sobre la región sur del dominio a lo largo del frente frío. La figura 1.9a muestra el agua

precipitable (sombreada) y la componente de viento meridional en niveles bajos promediada verticalmente (contornos). La figura muestra que la región húmeda que se extiende en la zona norte del dominio se ve reforzada por la asimilación de las observaciones de superficie de EMA. El aumento de la humedad es particularmente intenso en el extremo sur de esta región, justo por delante del frente frío donde se produce la iniciación de la convección. Los experimentos AWS y SATWND son muy similares, con valores de agua precipitable superiores a 55 kg m^{-2} al norte de 30°S y una distribución vertical similar de la humedad específica (no mostrada). RAD tiene un contenido de agua precipitable menor que AWS y SATWND, pero mayor que CONV. La distribución de la humedad en niveles bajos en RAD parece ser el resultado de la combinación del efecto de humedecimiento de la asimilación de EMA -compensado parcialmente por la asimilación de las observaciones de radiancia- y un menor transporte meridional de humedad debido al flujo más débil del norte sobre el centro del dominio en comparación con CONV.

La distribución de la temperatura y la humedad en la PBL (Figura 1.9b) se asemeja a las características observadas en los perfiles de temperatura (Figura 1.6a-c) donde AWS produce una PBL más fría que CONV mientras que la PBL en RAD es más cálida que en SATWND. En promedio, la PBL en AWS y SATWND es más fría que en CONV, mientras que RAD muestra una PBL más cálida que AWS debido a la asimilación de radiacias. Una PBL más cálida aumenta la inestabilidad potencial y ayuda a generar un entorno adecuado para el desarrollo de la convección profunda. La figura ??c muestra la energía potencial convectiva disponible más inestable (MCAPE, sombreada) y la cortante del viento entre 0 a 6 km. Los valores de MCAPE en CONV no superan los 2000 J Kg^{-1} mientras que en el resto de los experimentos el MCAPE máximo supera los 4000 J Kg^{-1} . El MCAPE en el experimento RAD es menor en comparación con AWS o SATWND. Esto es consistente con una menor humedad en la PBL con respecto a estos experimentos pero puede ser parcialmente compensado por una PBL ligeramente más cálida en el experimento RAD. La cortante del viento es más intensa en AWS, SATWND y RAD, alcanzando valores superiores a 15 m s^{-1} en el extremo sur de la región con valores positivos de MCAPE. Además, en esta misma región, estos experimentos muestran valores de MCAPE mayores que CONV. Cortantes de viento superior a 15 m s^{-1} están asociadas al desarrollo de SCMs más intensos y organizados (Chen et al., 2015) y también a condiciones favorables para el desarrollo de supercélulas (Markowski and Richardson, 2010).

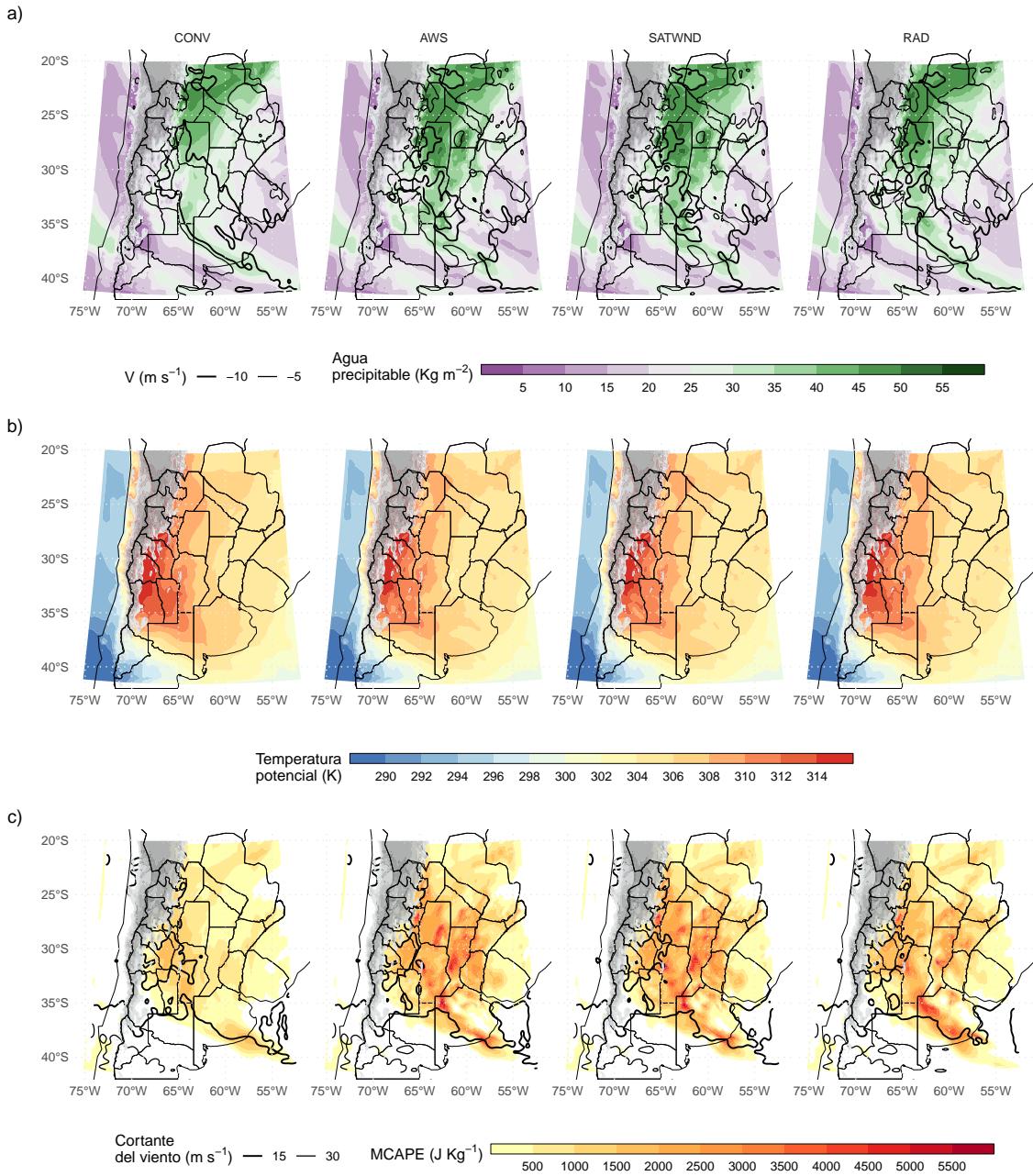


Figure 1.9: a) Agua precipitable (sombreada, kg m^{-2}) y viento medio del norte sobre los primeros 7 niveles sigma (desde la superficie hasta aproximadamente 800 hPa, contornos, m s^{-1}), b) Temperatura potencial media para la PBL (primeros 10 niveles sigma), y c) CAPE máximo y cortante del viento en $\sim 0\text{-}6$ km para 15 y 30 m s^{-1} en cada experimento. Todos los campos corresponden a la media del ensamble del análisis para las 00 UTC del 22 de noviembre. Los contornos rellenos de color gris corresponden a la topografía de más de 1500 metros sobre el nivel del mar.

1.2.3 Validación con observaciones independientes

En primer lugar, se analizó el impacto de la asimilación de diferentes tipos de observación en la representación del SCM y su precipitación asociada. La figura 1.10a muestra la precipitación acumulada horaria estimada por IMERG, y la media ajustada a la probabilidad (PM) (Clark, 2017) para la precipitación acumulada horaria del first-guess promediada entre 67°W y $54,5^{\circ}\text{W}$ en función del tiempo y la latitud en los diferentes experimentos. Las precipitaciones más intensas (más de 12 mmh^{-1}) comienzan durante la tarde del 22 de noviembre y continúan durante el 23 de noviembre posterior al periodo simulado (Figura 1.10a). En todos los experimentos, la precipitación acumulada en los pronósticos de corto plazo es subestimada. Esto es particularmente evidente en CONV (Figura 1.10b), donde el inicio de la convección se retrasa y ocurre más al norte con respecto al inicio observado. AWS, SATWND y RAD captan mejor el momento y la ubicación de la iniciación de la convección (Figuras 1.10c-e). AWS y RAD muestran una distribución más fragmentada en comparación con SATWND, posiblemente debido al desarrollo de una convección menos organizada durante el 22 de noviembre. Después de las 18 UTC del 22 de noviembre, RAD muestra mejoras en la tasa de precipitación y su distribución en comparación con los otros experimentos como resultado de un desarrollo de la convección más intenso.

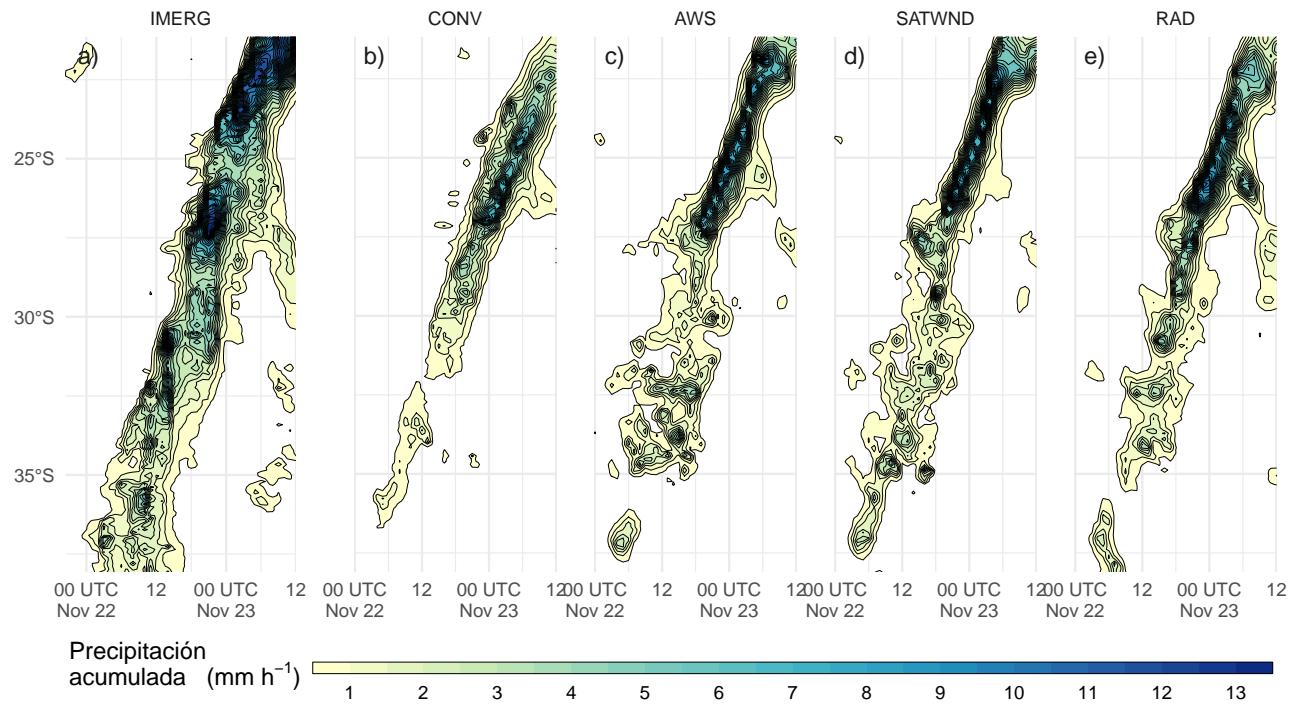


Figure 1.10: Diagrama de Hövmoller de la media ajustada a la probabilidad de la precipitación acumulada horaria para cada banda de latitud estimada por IMERG (izquierda) y simulada (derecha) por cada experimento, promediada en un rango de longitudes entre 67°W y $54,5^{\circ}\text{W}$. Los contornos se dibujan cada $0,5 \text{ mm}^{-1}$, comenzando en $0,5 \text{ mm}^{-1}$.

El FSS se calcula para cuantificar la coincidencia espacial entre la precipitación observada y la precipitación acumulada horaria simulada por el first-guess de los diferentes experimentos (Figura 1.11). Para cada umbral y escala espacial, se aplica la ecuación (11) en ventanas móviles de 6 horas a lo largo del periodo del experimento. Todos los experimentos muestran valores similares de FSS durante el inicio de la convección antes de las 06 UTC del 22 de noviembre, excepto RAD, que obtiene mejores resultados que el resto de los experimentos durante este periodo. Esto indica que las observaciones de radiancia tienen un impacto positivo en el análisis. El FSS de CONV es el más bajo en comparación con el resto de los experimentos y las diferencias son mayores durante la fase madura del SCM. AWS y SATWND muestran FSSs similares, lo que indica que la asimilación del viento estimado por satélites tiene poco impacto en la representación de la precipitación para este caso de estudio. La asimilación de radiancias conduce a una mejora general de la representación de la precipitación pronosticada a una hora, sobre todo para el umbral de 25 mm durante el

periodo de mayor precipitación durante 22 de noviembre (Figura 1.11b,d). La mejora también es importante en la fase de desarrollo del SCM (entre las 00 y las 12 UTC del 22 de noviembre y también para escalas espaciales superiores a 500 km, no mostradas).

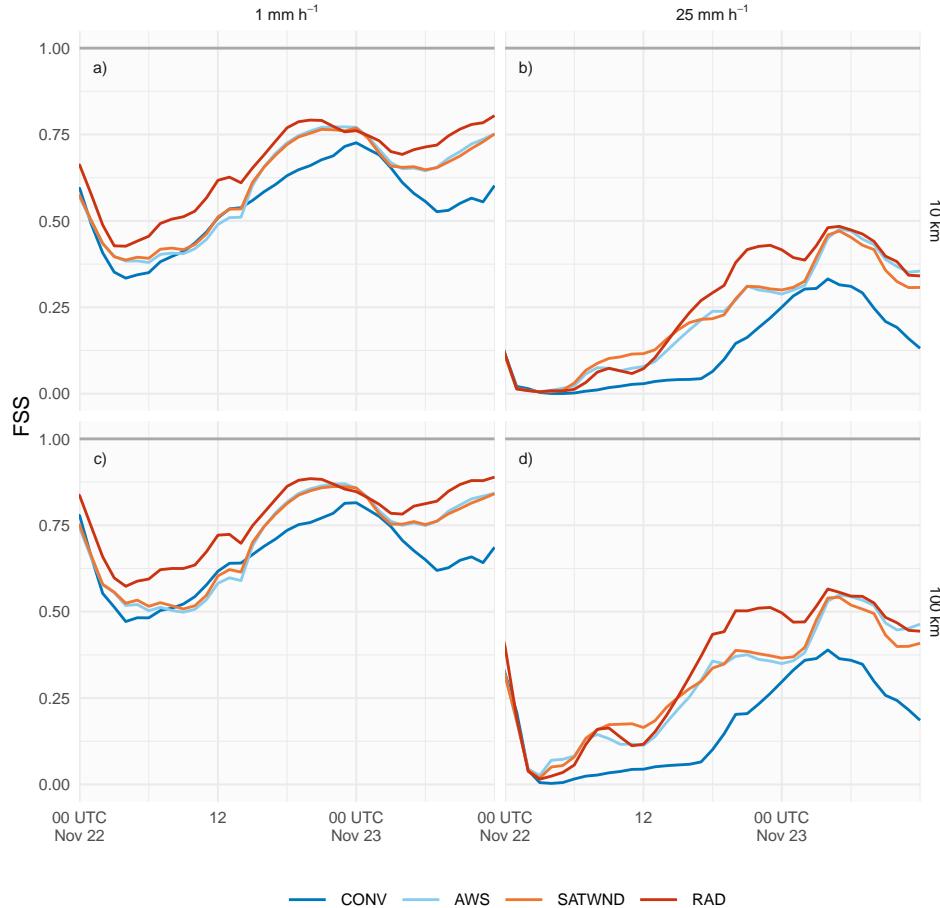


Figure 1.11: FSS calculado sobre la precipitación acumulada a 1 hora en una ventana móvil de 6 horas para umbrales de 1 mm (a y c) y 25 mm (b y d), en escalas de 10 km (a y b) y 100 km (c y d), para el first-guess de los experimentos CONV (línea azul), AWS (línea celeste), SATWND (línea naranja) y RAD (línea roja).

Para complementar el análisis, la Figura 1.12 muestra la reflectividad máxima observada en la columna vertical (COLMAX) y el COLMAX de la media del ensamble para los experimentos CONV y RAD en diferentes momentos entre las 10 y las 19 UTC del 22 de noviembre. Estos experimentos fueron elegidos porque representan el análisis con el mínimo (CONV) y el máximo (RAD) número de observaciones asimiladas. Además, son los experimentos con peor (CONV) y mejor (RAD) desempeño en cuanto a la habilidad para pronosticar la precipitación a una hora (Figura 1.11). En general,

ninguno de los pronósticos de corto alcance capta los detalles finos de la distribución de la reflectividad. Esto es esperable si se tiene en cuenta la resolución horizontal del modelo (10 km), que no es suficiente para representar adecuadamente la intensidad de la banda convectiva asociada al SCM. RAD representa mejor las características observadas del sistema mostrando un SCM más fuerte y organizado que CONV, sobre el centro del dominio a las 10 y 13 UTC (primera y segunda columnas en la Figura 1.12). Las celdas convectivas que se inician después de las 16 UTC a lo largo del frente cálido en la parte noreste del dominio son bien captadas por ambos experimentos, pero están mejor representadas en términos de intensidad en RAD. Además, CONV capta la ubicación del SCM, pero la convección parece estar menos organizada y ser mucho más débil que en RAD. Antes y después de los tiempos mostrados en la Figura 1.12, la concordancia entre la localización de las celdas convectivas observadas y las simuladas en los experimentos es bastante buena en las regiones donde se dispone de datos de radar, especialmente para RAD.

Finalmente, la Figura 1.13 muestra el RMSE y el bias calculados comparando los experimentos con los datos de radiosondeos durante las misiones de observación intensiva durante RELAMPAGO, IOP 7 del 15 al 21 UTC de noviembre (incluyendo 30 radiosondeos), e IOP 8 del 14 al 20 UTC de noviembre (incluyendo 22 radiosondeos).

El IOP 7 (Figuras 1.13a-d) proporciona una buena caracterización del entorno preconvectivo durante el primer día de nuestros experimentos. La zona donde se realizaron las observaciones se caracterizó por cielos mayoritariamente despejados y un flujo de norte en niveles bajos asociado a una advección cálida y húmeda. En general, los experimentos muestran un RMSE y un bias similares para todas las variables. Las observaciones de EMA lograron reducir el RMSE para la temperatura y la temperatura del punto de rocío y un bias seco en la PBL. Sin embargo, en esta región (Figura 3b) y para este periodo, los incrementos generados por EMAs en el análisis (Figura 1.7d) degradan el viento zonal entre 7 y 12 km aumentando el bias y el RMSE (Figura ??c).

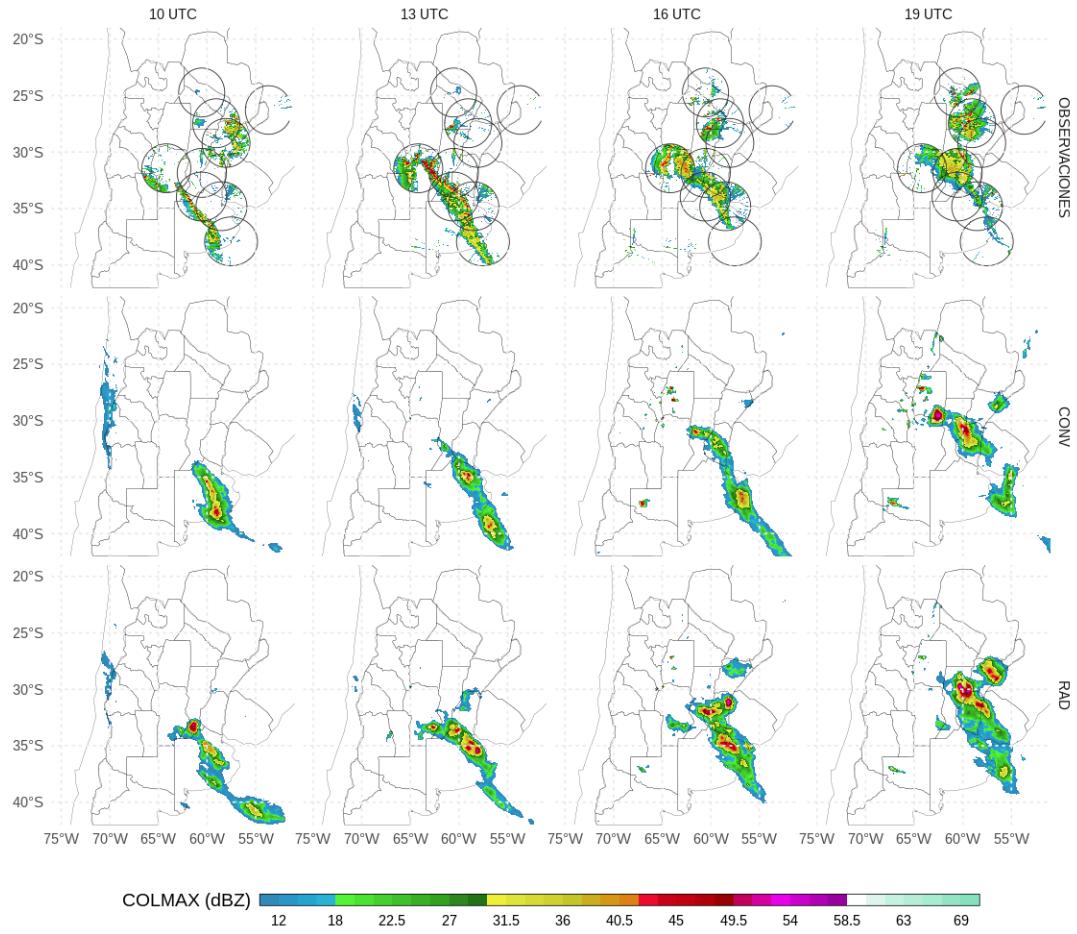


Figure 1.12: Reflectividad máxima en la columna (COLMAX en dBZ), observada (fila superior) y pronósticada a 1 hora por los experimentos CONV (segunda fila) y RAD (tercera fila) a las 10 UTC (primera columna), 13 UTC (segunda columna), 16 UTC (tercera columna) y 19 UTC (cuarta columna) del 22 de noviembre de 2018. Los círculos negros en la primera fila muestran el rango de observación de cada radar.

Durante el IOP 8 (Figuras 1.13e-h), la zona densamente observada se encontraba detrás del SCM, lo suficientemente alejada del sistema como para no estar directamente afectada por su circulación de mesoscala. Esta zona también estaba detrás del frente frío y se veía afectada por advección fría en niveles bajos. La asimilación de observaciones en AWS, SATWND y RAD reduce el bias de frío y el RMSE para la temperatura entre 5 y 12 km y el RMSE en la PBL en comparación con CONV (Figura 1.13e). La reducción del bias y del RMSE también es importante para la temperatura del punto de rocío (Figura 1.13f), siendo SATWND el que muestra el mayor impacto, seguido de AWS y RAD. El viento zonal está sobreestimado en los análisis y sólo RAD muestra una mejora con respecto a CONV en la troposfera ala

(Figura 1.13g). En niveles bajos el viento meridional (Figura 1.13g) presenta un bias negativo, indicando una subestimación del viento del sur detrás del frente frío principalmente en AWS, SATWND, y RAD. De hecho, los bias en niveles bajos en estos experimentos son mayores que en el experimento CONV, lo que indica un efecto negativo de las observaciones asimiladas (posiblemente asociado al efecto de EMA).

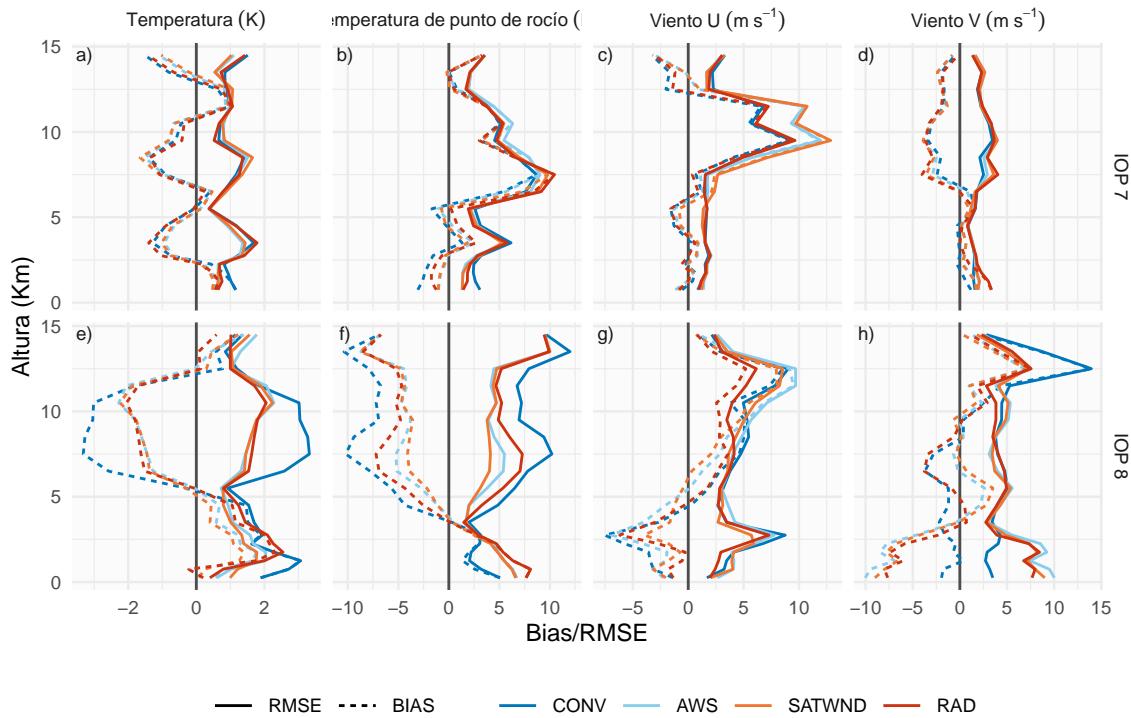


Figure 1.13: RMSE (línea sólida) y Bias (línea discontinua) de a) la temperatura (K), b) la temperatura del punto de rocío (K), c) el viento u ($m\ s^{-1}$) y d) el viento v ($m\ s^{-1}$) calculados comparando el análisis de cada experimento con los sondeos de RELAMPAGO durante el IOP 7 y el IOP 8. La línea azul corresponde a CONV, la línea celeste a AWS, SATWND se representa con una línea naranja y RAD con una línea roja.

1.3 Conclusiones

En este capítulo se investiga el impacto de los diferentes sistemas de observación en el desempeño de un sistema de asimilación de datos regional de mesoscala basado en ensambles. Este caso de estudio corresponde a un SCM masivo que se desarrolló sobre el sur de Sudamérica el 22 de noviembre de 2018 durante la campaña de campo de RELAMPAGO. En particular, se explora el impacto en la calidad del

análisis de la asimilación de las observaciones superficiales frecuentes y espacialmente densas provenientes de estaciones automáticas, los vientos estimados por satélite y las radiancias de cielo claro sensores a bordo de satélites polares.

En primer lugar, se evaluó la consistencia del ensamble para garantizar una buena concordancia entre la dispersión del ensamble y los errores de observación con respecto a la distancia entre la media del ensamble y las observaciones. Mientras que las desviaciones de las observaciones convencionales son coherentes con la dispersión del ensamble y los errores de observación asumidos, las desviaciones de las observaciones de viento estimado por satélite y radiancias son menores de lo esperado. Esto último podría ser el resultado de una sobreestimación de los errores de observación que se suele introducir para evitar el impacto perjudicial en el análisis de las observaciones de baja calidad.

En general todos los tipos de observación considerados (es decir, las estaciones meteorológicas automáticas, los vientos estimados por satélites y las radiancias de cielo claro procedentes de los satélites de órbita polar) mejoran la calidad del análisis y del pronóstico a corto plazo de precipitación cuando se comparan los resultados obtenidos al asimilar solo las observaciones de la red de observación convencional. En cuanto al análisis, las observaciones de las estaciones meteorológicas automáticas, que tienen una alta resolución espacial y temporal, produjeron impactos principalmente dentro de la PBL, pero que ocasionalmente se extienden por toda la troposfera durante los períodos en los que la convección húmeda es más intensa dentro del dominio. Estas observaciones también ayudaron a reducir el bias cálido y seco presente en el modelo, produciendo un análisis más cercano a ERA5. Durante el periodo preconvectivo, la asimilación de la temperatura de superficie, la temperatura del punto de rocío y el viento meridional mejoró el análisis en niveles bajos en comparación con los sondeos de la campaña RELAMPAGO. En particular, cuando se asimilan estas observaciones, el contenido de agua precipitable y la circulación meridional en niveles bajos condujeron a un aumento de la convección profunda y de la precipitación intensa acercando el análisis a lo observado.

También se encontraron resultados positivos al asimilar las observaciones de radiancia, que produjeron un mejor desarrollo de la convección y de su circulación de salida asociada, principalmente durante la fase madura del SCM, lo que condujo a un aumento de la precipitación acumulada en comparación con el caso en el que no se asimilan estas observaciones. Sin embargo, estas observaciones debilitaron el impacto de las observaciones de las estaciones meteorológicas automáticas dentro de la PBL, aumentando ligeramente el bias cálido y seco del modelo. Aunque esto debe estudiarse

más a fondo, podría estar relacionado con la asimilación de canales afectados por la superficie o con una corrección de bias de las observaciones subóptima. Comparando el experimento con sondeos independientes, la asimilación de radianes mejoró el viento en niveles medios y altos.

La asimilación del viento estimado por satélite no produjo un impacto notable en el análisis. Esto se debe posiblemente al número relativamente pequeño de observaciones en niveles bajos disponibles para este caso de estudio y gran error de observación asignado durante la asimilación. Sin embargo, se observan mejoras en la distribución de la precipitación acumulada en el pronóstico a 1 hora. Es necesario realizar un análisis más exhaustivo para comprender los mecanismos que subyacen al impacto de estas observaciones en los pronósticos de mayor alcance, que será tema de análisis en el capítulo 2.

Capítulo 2

Graphics, References, and Labels

2.1 Figures

If your thesis has a lot of figures, *R Markdown* might behave better for you than that other word processor. One perk is that it will automatically number the figures accordingly in each chapter. You'll also be able to create a label for each figure, add a caption, and then reference the figure in a way similar to what we saw with tables earlier. If you label your figures, you can move the figures around and *R Markdown* will automatically adjust the numbering for you. No need for you to remember! So that you don't have to get too far into LaTeX to do this, a couple **R** functions have been created for you to assist. You'll see their use below.

In the **R** chunk below, we will load in a picture stored as `reed.jpg` in our main directory. We then give it the caption of “Reed logo”, the label of “reedlogo”, and specify that this is a figure. Make note of the different **R** chunk options that are given in the R Markdown file (not shown in the knitted document).



Figure 2.1: Reed logo

Here is a reference to the Reed logo: Figure 2.1. Note the use of the `fig:` code here. By naming the **R** chunk that contains the figure, we can then reference that figure later as done in the first sentence here. We can also specify the caption for the

figure via the R chunk option `fig.cap`.

Below we will investigate how to save the output of an **R** plot and label it in a way similar to that done above. Recall the `flights` dataset from Chapter ???. (Note that we've shown a different way to reference a section or chapter here.) We will next explore a bar graph with the mean flight departure delays by airline from Portland for 2014.

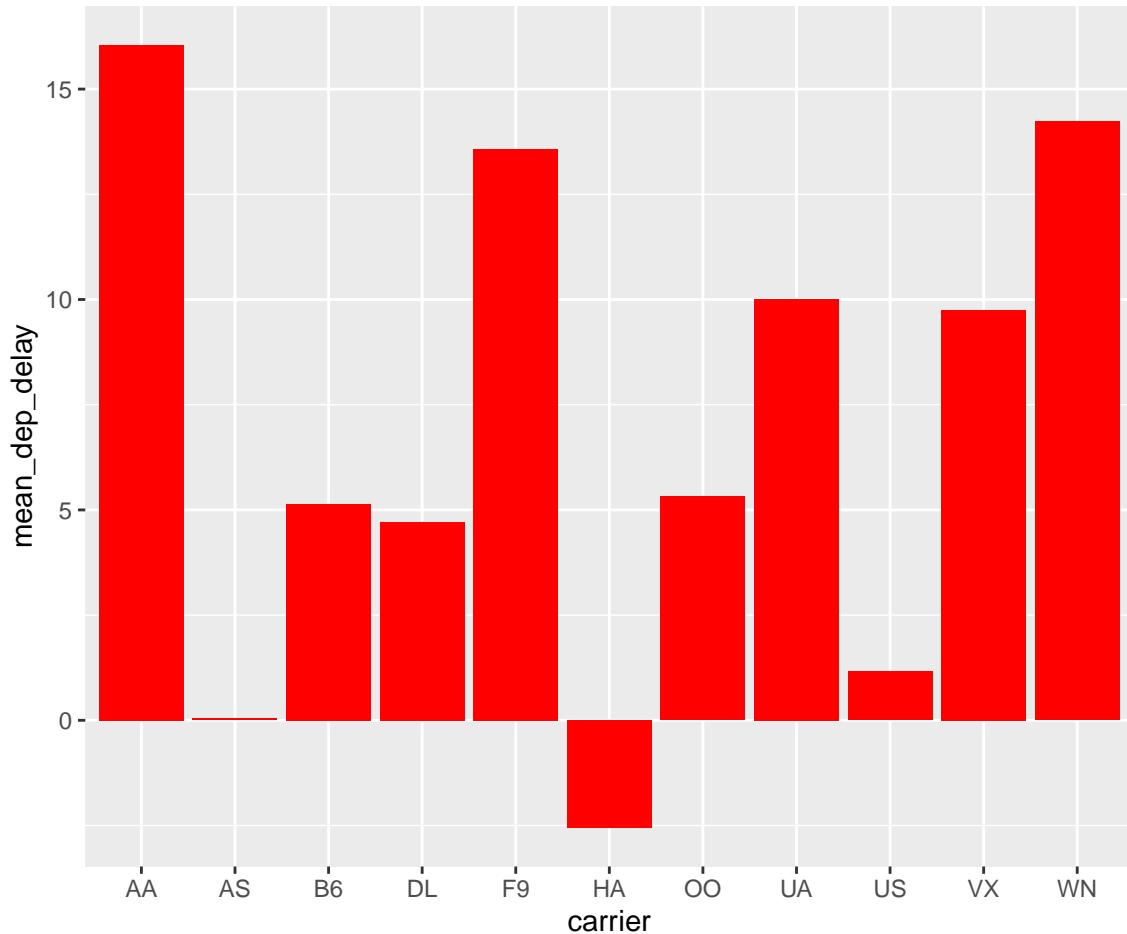


Figure 2.2: Mean Delays by Airline

Here is a reference to this image: Figure 2.2.

A table linking these carrier codes to airline names is available at <https://github.com/ismayc/pnwflights14/blob/master/data/airlines.csv>.

Next, we will explore the use of the `out.extra` chunk option, which can be used to shrink or expand an image loaded from a file by specifying "`scale=`". Here we use the mathematical graph stored in the "subdivision.pdf" file.

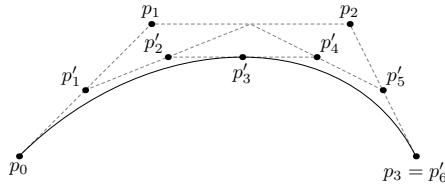


Figure 2.3: Subdiv. graph

Here is a reference to this image: Figure 2.3. Note that `echo=FALSE` is specified so that the **R** code is hidden in the document.

More Figure Stuff

Lastly, we will explore how to rotate and enlarge figures using the `out.extra` chunk option. (Currently this only works in the PDF version of the book.)

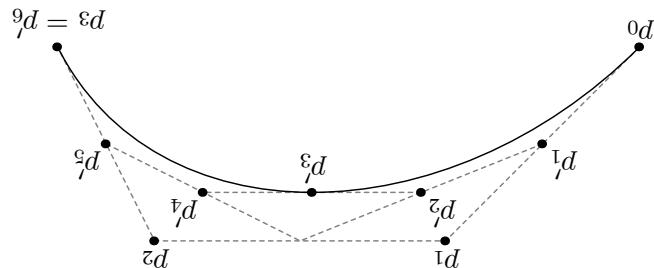


Figure 2.4: A Larger Figure, Flipped Upside Down

As another example, here is a reference: Figure 2.4.

2.2 Footnotes and Endnotes

You might want to footnote something.¹ The footnote will be in a smaller font and placed appropriately. Endnotes work in much the same way. More information can be found about both on the CUS site or feel free to reach out to `data@reed.edu`.

2.3 Bibliographies

Of course you will need to cite things, and you will probably accumulate an armful of sources. There are a variety of tools available for creating a bibliography database

¹footnote text

(stored with the .bib extension). In addition to BibTeX suggested below, you may want to consider using the free and easy-to-use tool called Zotero. The Reed librarians have created Zotero documentation at <https://libguides.reed.edu/citation/zotero>. In addition, a tutorial is available from Middlebury College at <https://sites.middlebury.edu/zoteromiddlebury/>.

R Markdown uses *pandoc* (<https://pandoc.org/>) to build its bibliographies. One nice caveat of this is that you won't have to do a second compile to load in references as standard LaTeX requires. To cite references in your thesis (after creating your bibliography database), place the reference name inside square brackets and precede it by the “at” symbol. For example, here's a reference to a book about worrying: (???). This Molina1994 entry appears in a file called `thesis.bib` in the `bib` folder. This bibliography database file was created by a program called BibTeX. You can call this file something else if you like (look at the YAML header in the main .Rmd file) and, by default, is placed in the `bib` folder.

For more information about BibTeX and bibliographies, see our CUS site (<https://web.reed.edu/cis/help/latex/index.html>)². There are three pages on this topic: *bibtex* (which talks about using BibTeX, at <https://web.reed.edu/cis/help/latex/bibtex.html>), *bibtexstyles* (about how to find and use the bibliography style that best suits your needs, at <https://web.reed.edu/cis/help/latex/bibtexstyles.html>) and *bibman* (which covers how to make and maintain a bibliography by hand, without BibTeX, at <https://web.reed.edu/cis/help/latex/bibman.html>). The last page will not be useful unless you have only a few sources.

If you look at the YAML header at the top of the main .Rmd file you can see that we can specify the style of the bibliography by referencing the appropriate csl file. You can download a variety of different style files at <https://www.zotero.org/styles>. Make sure to download the file into the csl folder.

Tips for Bibliographies

- Like with thesis formatting, the sooner you start compiling your bibliography for something as large as thesis, the better. Typing in source after source is mind-numbing enough; do you really want to do it for hours on end in late April? Think of it as procrastination.
- The cite key (a citation's label) needs to be unique from the other entries.

²(???)

- When you have more than one author or editor, you need to separate each author’s name by the word “and” e.g. `Author = {Noble, Sam and Youngberg, Jessica},`.
- Bibliographies made using BibTeX (whether manually or using a manager) accept LaTeX markup, so you can italicize and add symbols as necessary.
- To force capitalization in an article title or where all lowercase is generally used, bracket the capital letter in curly braces.
- You can add a Reed Thesis citation³ option. The best way to do this is to use the `phdthesis` type of citation, and use the optional “type” field to enter “Reed thesis” or “Undergraduate thesis.”

2.4 Anything else?

If you’d like to see examples of other things in this template, please contact the Data @ Reed team (email `data@reed.edu`) with your suggestions. We love to see people using *R Markdown* for their theses, and are happy to help.

³(???)

Conclusion

If we don't want Conclusion to have a chapter number next to it, we can add the `{-}` attribute.

More info

And here's some other random info: the first paragraph after a chapter title or section head *shouldn't be* indented, because indents are to tell the reader that you're starting a new paragraph. Since that's obvious after a chapter or section title, proper typesetting doesn't add an indent there.

Appendix A

The First Appendix

This first appendix includes all of the R chunks of code that were hidden throughout the document (using the `include = FALSE` chunk tag) to help with readability and/or setup.

In the main Rmd file

```
# This chunk ensures that the thesisdown package is
# installed and loaded. This thesisdown package includes
# the template files for the thesis.
if (!require(remotes)) {
  if (params$`Install needed packages for {thesisdown}`) {
    install.packages("remotes", repos = "https://cran.rstudio.com")
  } else {
    stop(
      paste('You need to run install.packages("remotes")',
            "first in the Console.'")
    )
  }
}
if (!require(thesisdown)) {
  if (params$`Install needed packages for {thesisdown}`) {
    remotes::install_github("ismayc/thesisdown")
  } else {
    stop(
      paste(
        "You need to run",
        "install.packages('remotes')",
        "first in the Console.'"
      )
    )
  }
}
```

```

    'remotes::install_github("ismayc/thesisdown")',
    "first in the Console."
)
)
}
}

library(thesisdown)
# Set how wide the R output will go
options(width = 70)

```

In Chapter 2:

```

# This chunk ensures that the thesisdown package is
# installed and loaded. This thesisdown package includes
# the template files for the thesis and also two functions
# used for labeling and referencing
if (!require(remotes)) {
  if (params$`Install needed packages for {thesisdown}`) {
    install.packages("remotes", repos = "https://cran.rstudio.com")
  } else {
    stop(
      paste(
        'You need to run install.packages("remotes")',
        "first in the Console."
      )
    )
  }
}

if (!require(dplyr)) {
  if (params$`Install needed packages for {thesisdown}`) {
    install.packages("dplyr", repos = "https://cran.rstudio.com")
  } else {
    stop(
      paste(
        'You need to run install.packages("dplyr")',
        "first in the Console."
      )
    )
  }
}

```

```
        )
    )
}

}

if (!require(ggplot2)) {
  if (params$`Install needed packages for {thesisdown}`) {
    install.packages("ggplot2", repos = "https://cran.rstudio.com")
  } else {
    stop(
      paste(
        'You need to run install.packages("ggplot2")',
        "first in the Console."
      )
    )
  }
}

if (!require(bookdown)) {
  if (params$`Install needed packages for {thesisdown}`) {
    install.packages("bookdown", repos = "https://cran.rstudio.com")
  } else {
    stop(
      paste(
        'You need to run install.packages("bookdown")',
        "first in the Console."
      )
    )
  }
}

if (!require(thesisdown)) {
  if (params$`Install needed packages for {thesisdown}`) {
    remotes::install_github("ismayc/thesisdown")
  } else {
    stop(
      paste(
        "You need to run",
        'remotes::install_github("ismayc/thesisdown")',
        )
    )
  }
}
```

```
    "first in the Console."
)
)
}
}

library(thesisdown)
library(dplyr)
library(ggplot2)
library(knitr)
flights <- read.csv("data/flights.csv", stringsAsFactors = FALSE)
```

Appendix B

The Second Appendix, for Fun

References

- Aksoy, A., Dowell, D.C., and Snyder, C., 2010. A Multicase Comparative Assessment of the Ensemble Kalman Filter for Assimilation of Radar Observations. Part II: Short-Range Ensemble Forecasts. *Monthly Weather Review*, 138, 4, 1273–1292.
- Allaire, J., Horner, J., Xie, Y., Marti, V., and Porte, N., 2019. Markdown: Render markdown with the c library 'sundown'.
- Andersson, E., Hollingsworth, A., Kelly, G., Lönnberg, P., Pailleux, J., and Zhang, Z., 1991. Global Observing System Experiments on Operational Statistical Retrievals of Satellite Sounding Data. *Monthly Weather Review*, 119, 8, 1851–1865.
- Bae, J.-H., and Min, K.-H., 2022. Forecast Characteristics of Radar Data Assimilation Based on the Scales of Precipitation Systems. *Remote Sensing*, 14, 3, 605.
- Bao, Y., Xu, J., Powell Jr., A.M., Shao, M., Min, J., and Pan, Y., 2015. Impacts of AMSU-A, MHS and IASI data assimilation on temperature and humidity forecasts with GSI-WRF over the western United States. *Atmospheric Measurement Techniques*, 8, 10, 4231–4242.
- Bauce Machado, V., gustavo de goncalves, luis, Vendrasco, E., Sinhori, N., Herdies, D., Sapucci, L., Levien, C., Quadro, M., Rodrigues, T., and Cardoso, C. and others, 2017. Investigating the impacts of convective scale hazardous weather events in Santa Catarina State through the CPTEC/INPE local data assimilation system. In. Presented at the Seventh International WMO Symposium on Data Assimilation.
- Campitelli, E., 2020, April. metR: Tools for Easier Analysis of Meteorological Fields.
- Candille, G., Côté, C., Houtekamer, P.L., and Pellerin, G., 2007. Verification of an Ensemble Prediction System against Observations. *Monthly Weather Review*, 135, 7, 2688–2699.
- Casaretto, G., Dillon, M.E., Salio, P., Skabar, Y.G., Nesbitt, S.W., Schumacher, R.S., García, C.M., and Catalini, C., 2022. High-Resolution NWP Forecast Precipitation Comparison over Complex Terrain of the Sierras de Córdoba during RELAMPAGO-CACTI. *Weather and Forecasting*, 37, 2, 241–266.
- Chang, W., Jacques, D., Fillion, L., and Baek, S.-J., 2017. Assimilation of Hourly

- Surface Observations with the Canadian High-Resolution Ensemble Kalman Filter. *Atmosphere-Ocean*, 55, 4-5, 247–263.
- Chen, F., and Dudhia, J., 2001. Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. *Monthly Weather Review*, 129, 4, 569–585.
- Chen, Q., Fan, J., Hagos, S., Gustafson, W.I., and Berg, L.K., 2015. Roles of wind shear at different vertical levels: Cloud system organization and properties. *Journal of Geophysical Research: Atmospheres*, 120, 13, 6551–6574.
- Chen, X., Zhao, K., Sun, J., Zhou, B., and Lee, W.-C., 2016. Assimilating surface observations in a four-dimensional variational Doppler radar data assimilation system to improve the analysis and forecast of a squall line case. *Advances in Atmospheric Sciences*, 33, 10, 1106–1119.
- Cherubini, T., Businger, S., Velden, C., and Ogasawara, R., 2006. The Impact of Satellite-Derived Atmospheric Motion Vectors on Mesoscale Forecasts over Hawaii. *Monthly Weather Review*, 134, 7, 2009–2020.
- Clark, A.J., 2017. Generation of Ensemble Mean Precipitation Forecasts from Convection-Allowing Ensembles. *Weather and Forecasting*, 32, 4, 1569–1583.
- Clark, A.J., Gallus, W.A., Xue, M., and Kong, F., 2009. A Comparison of Precipitation Forecast Skill between Small Convection-Allowing and Large Convection-Parameterizing Ensembles. *Weather and Forecasting*, 24, 4, 1121–1140.
- Computational and Information Systems Laboratory, 2019. Cheyenne: HPE/SGI ICE XA System (University Community Computing). National Center for Atmospheric Research Boulder, CO.
- de Elía, R., Vidal, L., and Lohigorry, P., 2017. El SMN y la red argentina de radares meteorológicos (<http://hdl.handle.net/20.500.12160/625>).
- Dillon, M.E., Maldonado, P., Corrales, P., Skabar, Y.G., Ruiz, J., Sacco, M., Cutraro, F., Mingari, L., Matsudo, C., and Vidal, L. and others, 2021. A rapid refresh ensemble based data assimilation and forecast system for the RELAMPAGO field campaign. *Atmospheric Research*, 105858.
- Dowle, M., and Srinivasan, A., 2020, July. Data.Table: Extension of 'data.frame'.
- Earth Observing Laboratory, U. -, 2020. Multi-network composite highest resolution radiosonde data. Version 1.3. UCAR/NCAR - earth observing laboratory.
- Eyre, J.R., English, S.J., and Forsythe, M., 2020. Assimilation of satellite data in numerical weather prediction. Part I: The early years. *Quarterly Journal of the Royal Meteorological Society*, 146, 726, 49–68.
- Eyre, J.R., Bell, W., Cotton, J., English, S.J., Forsythe, M., Healy, S.B., and Pavelin,

- E.G., 2022. Assimilation of satellite data in numerical weather prediction. Part II: Recent years. *Quarterly Journal of the Royal Meteorological Society*, 148, 743, 521–556.
- Ferreira, R.C., Herdies, D.L., Vendrasco, É.P., Beneti, C.A.A., and Biscaro, T.S., 2017. Impacto da Assimilação de Dados de Radar em Sistemas Convectivos de Mesoscala: Um Estudo de Caso. *Revista Brasileira de Meteorologia*, 32, 3, 447–458.
- Ferreira, R.C., Alves Júnior, M.P., Vendrasco, éder P., Aravéquia, J.A., Nolasco Junior, L.R., Biscaro, T.S., Ferreira, R.C., Alves Júnior, M.P., Vendrasco, éder P., and Aravéquia, J.A. and others, 2020. The Impact of Microphysics Parameterization on Precipitation Forecast Using Radar Data Assimilation. *Revista Brasileira de Meteorologia*, 35, 1, 123–134.
- Gao, F., Huang, X.-Y., Jacobs, N.A., and Wang, H., 2015. Assimilation of wind speed and direction observations: Results from real observation experiments. *Tellus A: Dynamic Meteorology and Oceanography*, 67, 1, 27132.
- Garcia, F., Ruiz, J., Salio, P., Bechis, H., and Nesbitt, S., 2019. Argentina mesonet data. Version 1.1. UCAR/NCAR - earth observing laboratory.
- García Skabar, Y., 1997. Análisis objetivo regional para inicializar un modelo de diez niveles en forma operativa. Tesis de licenciatura en ciencias de la atmósfera.
- Gasperoni, N.A., Wang, X., Brewster, K.A., and Carr, F.H., 2018. Assessing Impacts of the High-Frequency Assimilation of Surface Observations for the Forecast of Convection Initiation on 3 April 2014 within the Dallas–Fort Worth Test Bed. *Monthly Weather Review*, 146, 11, 3845–3872.
- Goncalves de Goncalves, L.G., Sapucci, L., Vendrasco, E., de Mattos, J.G., Ferreira, C., Khamis, E., and Cruz, N., 2015. A rapid update data assimilation cycle over South America using 3DVar and EnKF. In The 20th International TOVS Study Conference (ITSC-20).
- Grell, G.A., and Freitas, S.R., 2013. A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmospheric Chemistry and Physics Discussions*, 13, 9, 23845–23893.
- Ha, S.-Y., and Snyder, C., 2014. Influence of Surface Observations in Mesoscale Data Assimilation Using an Ensemble Kalman Filter. *Monthly Weather Review*, 142, 4, 1489–1508.
- Han, Y., Van Delst, P., Liu, Q., Weng, F., Yan, B., Treadon, R., and Derber, J., 2006. JCSDA Community Radiative Transfer Model (CRTM)—version 1 p. 40.
- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., and Rozum, I. and others, 2018. ERA5 hourly data

- on pressure levels from 1959 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), (Accessed on <08-08-2022>).
- Hong, S.-Y., Noh, Y., and Dudhia, J., 2006. A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. *Monthly Weather Review*, 134, 9, 2318–2341.
- Hong, S.-Y., Kim, J.-H., Lim, J.-o., and Dudhia, J., 2006. The WRF Single Moment 6-Class Microphysics Scheme (WSM6). *Journal of the Korean Meteorological Society*, 42, 129–151.
- Hu, M., Ge, G., Zhou, C., Stark, D., Shao, H., Newman, K., Beck, J., and Zhang, X., 2018. Grid-point Statistical Interpolation (GSI) User's Guide Version 3.7, Developmental Testbed Center, p. 149.
- Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E., Sorooshian, S., Tan, J., and Xie, P., 2018. NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG), National Aeronautics and Space Administration (NASA), p. 35.
- Hunt, B.R., Kostelich, E.J., and Szunyogh, I., 2007. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, 230, 1-2, 112–126.
- Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shephard, M.W., Clough, S.A., and Collins, W.D., 2008. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, 113, D13, D13103.
- Janjić, Z.I., 1994. The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes. *Monthly Weather Review*, 122, 5, 927–945.
- Jones, T.A., Otkin, J.A., Stensrud, D.J., and Knopfmeier, K., 2013. Assimilation of Satellite Infrared Radiances and Doppler Radar Observations during a Cool Season Observing System Simulation Experiment. *Monthly Weather Review*, 141, 10, 3273–3299.
- Kain, J.S., 2004. The Kain–Fritsch Convective Parameterization: An Update. *JOURNAL OF APPLIED METEOROLOGY*, 43, 12.
- Kalnay, E., 2002, November 6. Atmospheric Modeling, Data Assimilation and Predictability (<https://www.cambridge.org/highereducation/books/atmospheric-modeling-data-assimilation-and-predictability/C5FD207439132836E85027754CE9BC1A>).
- Kelly, G.a.M., Mills, G.A., and Smith, W.L., 1978. Impact of Nimbus-6 Temperature Soundings on Australian Region Forecasts. *Bulletin of the American Meteorological Society*, 59, 4, 393–406.

- Kleist, D.T., Parrish, D.F., Derber, J.C., Treadon, R., Wu, W.-S., and Lord, S., 2009. Introduction of the GSI into the NCEP Global Data Assimilation System. *Weather and Forecasting*, 24, 6, 1691–1705.
- Lin, H., Weygandt, S.S., Benjamin, S.G., and Hu, M., 2017. Satellite Radiance Data Assimilation within the Hourly Updated Rapid Refresh. *Weather and Forecasting*, 32, 4, 1273–1287.
- Maejima, Y., Miyoshi, T., Kunii, M., Seko, H., and Sato, K., 2019. Impact of Dense and Frequent Surface Observations on 1-Minute-Update Severe Rainstorm Prediction: A Simulation Study. *Journal of the Meteorological Society of Japan. Ser. II*, 97, 1, 253–273.
- Maldonado, P., Ruiz, J., and Saulo, C., 2021. Sensitivity to Initial and Boundary Perturbations in Convective-Scale Ensemble-Based Data Assimilation: Imperfect-Model OSSEs. *SOLA*, 17, 0, 96–102.
- Markowski, P., and Richardson, Y., 2010. Organization of Isolated Convection. In *Mesoscale Meteorology in Midlatitudes* pp. 201–244.
- Matsudo, C., Salles, M.A., and García Skabar, Y., 2021. Verificación de los pronósticos del esquema determinístico del modelo WRF para el año 2020.
- Nakanishi, M., and Niino, H., 2009. Development of an Improved Turbulence Closure Model for the Atmospheric Boundary Layer. *Journal of the Meteorological Society of Japan*, 87, 5, 895–912.
- National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, 2015. NCEP GFS 0.25 degree global forecast grids historical archive.
- Necker, T., Geiss, S., Weissmann, M., Ruiz, J., Miyoshi, T., and Lien, G., 2020. A convective-scale 1,000-member ensemble simulation and potential applications. *Quarterly Journal of the Royal Meteorological Society*, 146, 728, 1423–1442.
- Nesbitt, S.W., Salio, P.V., Ávila, E., Bitzer, P., Carey, L., Chandrasekar, V., Deierling, W., Dominguez, F., Dillon, M.E., and Garcia, C.M. and others, 2021. A storm safari in Subtropical South America: Proyecto RELAMPAGO. *Bulletin of the American Meteorological Society*, -1, aop, 1–64.
- Ohring, G., 1979. Impact of Satellite Temperature Sounding Data on Weather Forecasts. *Bulletin of the American Meteorological Society*, 60, 10, 1142–1147.
- Ouaraini, R.E., Berre, L., Fischer, C., and Sayouty, E.H., 2015. Sensitivity of regional ensemble data assimilation spread to perturbations of lateral boundary conditions. *Tellus A: Dynamic Meteorology and Oceanography*, 67, 1, 28502.
- Pondeca, M.S.F.V.D., Manikin, G.S., DiMego, G., Benjamin, S.G., Parrish, D.F., Purser,

- R.J., Wu, W.-S., Horel, J.D., Myrick, D.T., and Lin, Y. and others, 2011. The Real-Time Mesoscale Analysis at NOAA's National Centers for Environmental Prediction: Current Status and Development. *Weather and Forecasting*, 26, 5, 593–612.
- Purser, R.J., Wu, W.-S., Parrish, D.F., and Roberts, N.M., 2003a. Numerical Aspects of the Application of Recursive Filters to Variational Statistical Analysis. Part II: Spatially Inhomogeneous and Anisotropic General Covariances. *Monthly Weather Review*, 131, 8, 1536–1548.
- , 2003b. Numerical Aspects of the Application of Recursive Filters to Variational Statistical Analysis. Part I: Spatially Homogeneous and Isotropic Gaussian Covariances. *Monthly Weather Review*, 131, 8, 1524–1535.
- R Core Team, 2020. R: A language and environment for statistical computing, R Foundation for Statistical Computing,
- Roberts, B., Gallo, B.T., Jirak, I.L., Clark, A.J., Dowell, D.C., Wang, X., and Wang, Y., 2020. What Does a Convection-Allowing Ensemble of Opportunity Buy Us in Forecasting Thunderstorms? *Weather and Forecasting*, 35, 6, 2293–2316.
- Roberts, N., 2008. Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteorological Applications*, 15, 1, 163–169.
- Ruiz, J.J., Saulo, C., and Nogués-Paegle, J., 2010. WRF Model Sensitivity to Choice of Parameterization over South America: Validation against Surface Variables. *Monthly Weather Review*, 138, 8, 3342–3355.
- Saucedo, M.A., n.d. Estudio de los efectos de diferentes fuentes de error sobre la calidad de los análisis generados por un sistema de asimilación por filtros de Kalman, 279.
- Sawada, M., Ma, Z., Mehra, A., Tallapragada, V., Oyama, R., and Shimoji, K., 2019. Impacts of Assimilating High-Resolution Atmospheric Motion Vectors Derived from Himawari-8 on Tropical Cyclone Forecast in HWRF. *Monthly Weather Review*, 147, 10, 3721–3740.
- Shao, H., Derber, J., Huang, X.-Y., Hu, M., Newman, K., Stark, D., Lueken, M., Zhou, C., Nance, L., and Kuo, Y.-H. and others, 2016. Bridging Research to Operations Transitions: Status and Plans of Community GSI. *Bulletin of the American Meteorological Society*, 97, 8, 1427–1440.
- Singh, R., Ojha, S.P., Kishtawal, C.M., Pal, P.K., and Kiran Kumar, A.S., 2016. Impact of the assimilation of INSAT-3D radiances on short-range weather forecasts: Assimilation of INSAT-3D Radiances. *Quarterly Journal of the Royal Meteorological Society*, 142, 694, 120–131.
- Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.-Y., Wang, W., and Powers, J.G., 2008. A Description of the Advanced

- Research WRF Version 3 p. 125.
- Sobash, R.A., and Stensrud, D.J., 2015. Assimilating Surface Mesonet Observations with the EnKF to Improve Ensemble Forecasts of Convection Initiation on 29 May 2012. *Monthly Weather Review*, 143, 9, 3700–3725.
- Stensrud, D.J., Wicker, L.J., Xue, M., Dawson, D.T., Yussouf, N., Wheatley, D.M., Thompson, T.E., Snook, N.A., Smith, T.M., and Schenkman, A.D. and others, 2013. Progress and challenges with Warn-on-Forecast. *Atmospheric Research*, 123, 2–16.
- Toshio Inouye, R., Calvetti, L., Gonçalves, J., Maske, B., Neundorf, R., Beneti, C., Diniz, F., Vendrasco, E., Herdies, D., and gustavo de goncalves, luis, 2017. Impact of radar data assimilation on a severe storm study in brazil. In. Presented at the 97th American Meteorological Meeting Annual Meeting.
- Vera, C.S., 1992. Un sistema de asimilación de datos para la región extratropical de Sudamérica.
- Wang, Z.Q., and Randriamampianina, R., 2021. The Impact of Assimilating Satellite Radiance Observations in the Copernicus European Regional Reanalysis (CERRA). *Remote Sensing*, 13, 3, 3, 426.
- Weston, P., Geer, A., Bormann, N., and Bormann, N., 2019. Investigations into the assimilation of AMSU-A in the presence of cloud and precipitation.
- Wheatley, D.M., and Stensrud, D.J., 2010. The Impact of Assimilating Surface Pressure Observations on Severe Weather Events in a WRF Mesoscale Ensemble System. *Monthly Weather Review*, 138, 5, 1673–1694.
- Whitaker, J.S., and Hamill, T.M., 2002. Ensemble Data Assimilation without Perturbed Observations. *Monthly Weather Review*, 130, 7, 1913–1924.
- , 2012. Evaluating Methods to Account for System Errors in Ensemble Data Assimilation. *Monthly Weather Review*, 140, 9, 3078–3089.
- Whitaker, J.S., Hamill, T.M., Wei, X., Song, Y., and Toth, Z., 2008. Ensemble Data Assimilation with the NCEP Global Forecast System. *Monthly Weather Review*, 136, 2, 463–482.
- Wickham, H., 2009. Ggplot2: Elegant Graphics for Data Analysis, Springer-Verlag,
- Wu, T.-C., Liu, H., Majumdar, S.J., Velden, C.S., and Anderson, J.L., 2014. Influence of Assimilating Satellite-Derived Atmospheric Motion Vector Observations on Numerical Analyses and Forecasts of Tropical Cyclone Track and Intensity. *Monthly Weather Review*, 142, 1, 49–71.
- Wu, W.-S., Purser, R.J., and Parrish, D.F., 2002. Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. *Monthly Weather Review*, 130, 12, 2905–2916.

- Xie, Y., 2015. Dynamic documents with R and knitr, Second., Chapman and Hall/CRC,
- Zhu, K., Xue, M., Pan, Y., Hu, M., Benjamin, S.G., Weygandt, S.S., and Lin, H., 2019. The Impact of Satellite Radiance Data Assimilation within a Frequently Updated Regional Forecast System Using a GSI-based Ensemble Kalman Filter. *Advances in Atmospheric Sciences*, 36, 12, 1308–1326.
- Zhu, Y., and Gelaro, R., 2008. Observation Sensitivity Calculations Using the Adjoint of the Gridpoint Statistical Interpolation (GSI) Analysis System. *Monthly Weather Review*, 136, 1, 335–351.
- Zhu, Y., Derber, J., Collard, A., Dee, D., Treadon, R., Gayno, G., and Jung, J.A., 2014. Enhanced radiance bias correction in the National Centers for Environmental Prediction’s Gridpoint Statistical Interpolation data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 140, 682, 1479–1492.
- Zhu, Y., Liu, E., Mahajan, R., Thomas, C., Groff, D., Van Delst, P., Collard, A., Kleist, D., Treadon, R., and Derber, J.C., 2016. All-Sky Microwave Radiance Assimilation in NCEP’s GSI Analysis System. *Monthly Weather Review*, 144, 12, 4709–4735.