ABSTRACT

Title of Dissertation:	IMPROVEMENTS IN THE ASSIMILATION OF DOPPLER RADIAL WINDS AT NCEP IN REGIONAL AND GLOBAL FORECAST SYSTEMS
	Donald E. Lippi, Doctor of Philosophy, 2023
Dissertation directed by:	Dr. Daryl T. Kleist, Department of Atmospheric and Oceanic Science

Doppler radial winds have been an underutilized observation in U.S. operational forecast systems. This has typically been owing to limitations in formulation of the observation operator, the amount of data thinning via super-obbing, or simple exclusion from assimilation in global modeling systems. In this work we explore some of the more feasible aspects of radial wind assimilation that could more readily be applied to the operational systems with the main goal to improve the use of radial winds in the operational forecast systems used by NOAA. It will be shown that improvements could be made to systems not only operating at the convective scale, but also that global systems could benefit from the assimilation of radial winds. Experiments featuring results from the regional NAM version 4 forecast system along with the GFS version 15 and 16 will be shown. The GFS experiments explore the potential radial wind impact via hypothetical observing networks tested with observing

system simulation experiments. We further extend this study to a real-data case with a land falling tropical cyclone event and novel, rapidly-updated version of the GFS.

This work is a first step toward improving the use of radial wind observations and tests their use for the very first time in the GFS. This effort demonstrates potential for radial wind assimilation in the GFS, a potentially important observation type as we increase update cadence and spatial resolution.

IMPROVEMENTS IN THE ASSIMILATION OF DOPPLER RADIAL WINDS AT NCEP IN REGIONAL AND GLOBAL FORECAST SYSTEMS

by

Donald E. Lippi

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Dr. Daryl T. Kleist (Chair)
Dr. Jacob R. Carley
Prof. Kayo Ide
Prof. Jonathan Poterjoy
Prof. Brian Hunt (Dean's Representative)

Foreword

Foreword by Dr. Laura Slivinski, Research Scientist at the Cooperative Institute for Research in Environmental Sciences at CU Boulder and the NOAA Physical Sciences Laboratory

Aspects of the work discussed in this dissertation (specifically, Chapter 4, "Overlapping Windows in a Global Hourly Data Assimilation System and a Radial Wind Assimilation Experiment") have been published in Slivinski et al (2022), on which Donald Lippi was second author (of seven total). He made substantial contributions to that work, including the development of the code to implement the Global Rapid Refresh System within NCEP's Global Data Assimilation System, running experiments to test the system, analyzing subsequent results, and actively participating in weekly meetings to discuss progress and goals of the larger project (supported by NOAA Grant NA19OAR0220182). He contributed significantly to the writing of the text and produced several of the figures in that paper. Therefore, that work deserves to be included in this dissertation.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to Dr. Daryl Kleist and Dr. Jacob Carley for their guidance, feedback, and encouragement not just in regard to this work but also in helping me develop into my career. This work could not have been possible without both of these brilliant scientists; I am incredibly lucky to have both of them as mentors. I would also like to thank Professors Kayo Ide, Jonathan Poterjoy, and Brian Hunt for providing useful discussions and feedback to this dissertation and their willingness to participate on my defense committee. I am incredibly grateful to the EMC, I.M. System Group (IMSG), and Lynker Technologies for providing the opportunity and funding for this work as well as to maintain a healthy work-life balance. I would like to give a special thank you to Dr. Laura Slivinski and Dr. Jeffrey Whitaker for an incredibly useful collaboration and their willingness to allow me to include our work on the overlapping windows as an additional topic to this dissertation. I am also grateful to other colleagues at NOAA EMC, especially Shun Liu, for their useful discussions.

A huge thank you to my best friend and wife, Jessica Lippi. Thank you for all your support throughout this journey. I truly couldn't have done it without you. Last but certainly not the least, I want to give a special thank you to my parents, Donald and Teresa Lippi; both of you have provided a tremendous amount of support and encouragement.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

3D	Three-Dimension
3DEnVar	Three-Dimension Ensemble-Variational
3DVAR	Three-Dimension Variational
4D	Four-Dimension
4DEnVar	Four-Dimension Ensemble Variational
4DIAU	Four-Dimension Incremental Analysis Update
4DVAR	Four-Dimension Variational
ADT	Advanced Dvorak Technique
AMV	Atmospheric Motion Vector
ARPS	Advanced Regional Prediction System
ATL	Atlantic Ocean (verification domain)
BCMSE	Bias Corrected Mean Square Error
BCRMSE	Bias Corrected Root Mean Square Error
BEC	Background Error Covariance Matrix
BUFR	Binary Universal Form for Reading Meteorological Data
CASA	Collaborative Adaptive Sensing of the Atmosphere
CC	Correlation Coefficient
CONUS	Contiguous United States
DA	Data Assimilation
dBZ	Decibels of Z (of radar reflectivity factor)
DTC	Developmental Testbed Center
Dual-Pol	Dual Polarization
ECMWF	European Center for Medium-Range Weather Forecasts
EKF	Extended Kalman Filter
EMC	Environmental Modeling Center
En4DVar	Ensemble 4DVar
EnKF	Ensemble Kalman Filter
EnSRF	Ensemble Square Root Filter
ESRL	Earth System Research Laboratory
ETKF	Ensemble Transform Kalman Filter
ETS	Equitable Threat Score
EUR	European (verification domain)
FAA	Federal Aviation Administration
FAT	Feature Alignment Technique
FBIAS	Frequency Bias
FGAT	First Guess at Appropriate Time
FSS	Fraction Skill Score
FV3	Finite Volume Model in the Cubed Sphere Grid

GDAS	Global Data Assimilation System
GFDL	Geophysical Fluid Dynamics Laboratory
GFS	Global Forecast System
GLOBAL	OSSE experiment using the global radar network defined in the WRD
GOES	Geostationary Operational Environmental Satellite
GPS	Global Positioning System
GRR	Global Rapid Refresh
GSD	Global Systems Division
GSI	Gridpoint Statistical Interpolation System
HRRR	High-Resolution Rapid Refresh
HWRF	Hurricane Weather and Research Forecast System
HYPO	OSSE experiment using the hypothetical radar network
IAU	Incremental Analysis Update
IMSG	I.M. Systems Group
JCSDA	Joint Center for Satellite Data Assimilation
JDOP	Joint Doppler Operational Project
JMA	Japanese Meteorological Administration
KDP	specific Differential Phase
KF	Kalman Filter
LEKF	Local Extended Kalman Filter
LETKF	Local Extended Transform Kalman Filter
LLJ	Low Level Jet
MET	Model Evaluation Tools
MRMS	Multi-Radar/Multi-Sensor System
NAM	North American Mesoscale Forecast System
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NEXRAD	Next Generation Radar Network
NGGPS	Next Generation Global Prediction System
NHM	Northern Hemisphere (verification domain)
NMC	National Modeling Center
NMMB	Non-hydrostatic Multiscale Model on the Arakawa B grid
NOAA	National Oceanic and Atmospheric Administration
NODA	OSSE No Data Assimilation experiment
NOXP	NOAA-X-POL
NSSL	National Severe Storm Laboratory
NWP	Numerical Weather Prediction
NWS	National Weather Service
OAR	Oceanic and Atmospheric Research
OFCM	Office of the Federal Coordinator for Meteorological Services and Supporting Research

OI	Optimal Interpolation
OPERA	Operational Program for the Exchange of Weather Radar Information
OSSE	Observing System Simulation Experiment
OW	Overlapping Windows
PDF	Probability Distribution Function
PRF	Pulse Repetition Frequency
PSD	Physical Sciences Laboratory
RAP	Rapid Refresh
RMS	Root Mean Square (error)
RMSD	Root Mean Square Difference
RRFS	Rapid Refresh Forecast System
RTPP	Relaxation to Prior Perturbation
RTPS	Relaxation to Prior Spread
RUC	Rapid Update Cycle
SHM	Southern Hemisphere (verification domain)
TC	Tropical Cyclone
TDWR	Tail Doppler Weather Radar
TLNMC	Tangent Linear Normal Mode Constrain
TRP	Tropics (verification domain)
VAD	Velocity Azimuth Display
VAR	Variational
VCP	Volume Coverage Pattern
WMO	World Meteorological Organization
WoF	Warn-on-Forecast
WRD	World Radar Database
WRF	Weather and Research Forecast System
WSR-57	Weather Surveillance Radar - 1957
WSR-74	Weather Surveillance Radar - 1974
WSR-88D	Weather Surveillance Radar - 1988 Doppler
ZDR	Differential Reflectivity

Chapter 1: Introduction

A Weather-Ready Nation (https://www.weather.gov/wrn/) is the vision of the National Weather Service and a continual effort is being made to improve the weather information available to the public. It is well known that a free-running model will accumulate errors until it is no longer useful due to the atmosphere's chaotic nature and limited predictability (Lorenz 1969). Therefore, the ability to produce and disseminate useful weather information to the public is not only fundamentally dependent on having an accurate forecast model, but it also requires providing that model with accurate initial and boundary conditions. Data assimilation (DA) is the methodology of combining observations with a numerical weather prediction (NWP) model and their respective uncertainties to obtain the most accurate state of the atmosphere to use as initial conditions for generating a useful weather forecast.

Skillful convective-scale forecasts depend on having high accuracy at large scales and the frequent assimilation of convective-scale observations such as radial wind (among other types) to correct the locations of front, storms, tropical cyclones, etc. Therefore, it is critical to improve the use the use of radial wind observations in both regional and global systems. This dissertation takes on one small aspect of the continuing effort to improve our nation's weather-readiness by improving the use and assimilation of Doppler radial wind observations in both regional and global operational forecast systems.

1.1 Motivation

The National Centers for Environmental Prediction (NCEP) has had real-time radar observations since May 2005 and has been assimilating radial winds operationally since 2006 in the North American Mesoscale (NAM) forecast system. The radial wind assimilation scheme, however, has remained mostly unchanged since its original implementation despite improvements in regional modeling, increased model resolutions, data assimilation techniques, and radar data quality control (Liu et al. 2016).

Radial wind observations have been notoriously difficult to assimilate skillfully (Fabry and Meunier 2020), especially in operational systems. For example, the assimilation of radial wind was assessed in the High-Resolution Rapid Refresh (HRRR) and was found to provide little to no benefit (DTC 2018). Generally, the studies that have found the most benefit from assimilating radial wind observations have stemmed from convective-scale systems in the literature using sub-hourly update cadences, high resolution modeling at 3-km grid spacing or less, and over very small domain sizes (e.g., Gao and Stensrud 2014; Jones et al. 2015). While some operational systems currently use 3-km grid-spacing, they do not update at sub-hourly cadences. Sub-hourly cadences of operational systems would likely be predicated on increased computing power and more scalable algorithms than currently employed in order to keep up in real time. The main goal of this dissertation then is to understand how the assimilation of radial wind data could lead to improvements in operational regional and global systems.

Table 1 summarizes characteristics of various systems that have previously tested the assimilation of radial wind observations or are considered for radial wind assimilation in this dissertation. These characteristics include the grid-spacing of the model, update frequency, domain size, the assimilation of radial wind observations, and the relative impact of the radial wind DA on the respective experiments. Note that these systems have much smaller domains relative to operational domains such as the NAM CONUS nest (1827 × 1467 grid points at 3-km grid spacing).

Radial winds are an underutilized observation type in operational systems and there are limitations to the current assimilation routines that are used in operations. For example, the forward operator has historically not included vertical velocity. Furthermore, there are clear benefits in studies in the literature which have many of the components missing in the operational assimilation routines. It is conceivable then that radial wind assimilation could have a larger impact in the operational systems than what has previously been found. Therefore, the focus of the first part of this study (Chapter 2) is to improve the use of radial wind observations in the operational, convective-scale regional systems at NCEP. This work includes updating the observation operator to include vertical velocity and optimizing the super-observation parameters by expanding upon the work of Lippi et al. (2016) to improve the operational analysis and subsequent forecast of convective-storms.

Table 1. Summary table of systems and studies that have tested radar DA with corresponding inner-most domain size and grid spacing, update frequency, use of reflectivity and/or radial wind, and the subjective impact of assimilating the data in each corresponding system.

Study/System	**(Inner Domain) Grid-spacing**	**Update Frequency**	**(Inner) Domain Size (grid points)**	**Radial wind DA**	**Impact**
(Gao and Stensrud 2014)	1-km	5-min	57 × 57	Yes	Encouraging
(Jones et al. 2015)	3-km	15-min/ 5-min	170 × 170	Yes	Encouraging
(Jones et al. 2016; Wheatley et al. 2015)	3-km	15min	170 × 170	Yes	Encouraging
(Lawson et al. 2018)	3-km	15-min	250 × 250	Yes	Encouraging
HRRR	3-km	1-hr	1799 × 1059	Yes	Small
NAM (CONUS nest)	3-km	1-hr	1827 × 1467	Yes	Small
GFS	C768 (~13-km)	6-hr	768 × 768 per face	No	N/A

Global modeling systems, on the otherhand, do not assimilate any radial wind data and thus the value of their assimilation is completely unknown in global systems (Sun 2005b). Over the last several years the operational Global Forecast System (GFS) has undergone many upgrades that justify exploring the impacts of assimilating radial wind observations in a global system. Furthermore, over 800 Doppler radars are registered in the World Meteorological Organization's world radar database (WRD) [Available online: https://wrd.mgm.gov.tr/Home/Wrd]; however, no Doppler radial winds are assimilated within NCEP's GFS. Therefore, the main goal of the second part

of this work (Chapter 3) will be to understand if it is possible to effectively assimilate radial wind observation in the global system. Owing to the lack of available radar observations from the global radar network, an observing system simulation experiment (OSSE) was designed to test this concept.

The final part of this work (Chapter 4) combines and builds upon aspects of Chapters 2 and 3. Chapter 4 discusses the development of a global, hourly updating forecast system, the Global Rapid Refresh (GRR), in a near-operational configuration (Slivinski et al. 2022) and also preliminary testing assimilating radial wind observations during a landfalling tropical cyclone event. There are several motivating factors for developing an hourly updating global system. One to highlight is that more frequent updates hypothetically enables better use of high frequency, low-latency observations such as radial wind observations (among other types) by using them closer to their receipt time and thus closer to the end of the assimilation window. Observations at the end of the assimilation window have been shown to have a greater impact on the analysis (McNally 2019).

There are other benefit to having an hourly updating global system as well. First, the GFS currently provides updated analyses and forecast fields on a 6-hourly basis whereas higher resolution, regional systems update on a 1-hourly basis. Regional systems such as the Rapid Refresh (RAP; Benjamin et al. 2016) and High Resolution Rapid Refresh (HRRR; Alexander et al. 2020; Dowell et al. 2022; James et al. 2022) ultimately depend on a global system to provide initial and lateral boundary conditions. The current GFS is reinitialized every 6-hours, but is responsible for providing hourly lateral boundary conditions to the RAP. One motivator for an hourly updating global

system is that more frequent updates should provide more accurate initial and lateral boundary conditions to the RAP and other downstream systems. Second, high frequency observations are becoming much more common aside from radar observations. For example, full disk observations from geostationary satellites such as GOES-16/17[1] and Himawari-8/9[2] are available every 10 minutes, but it is likely that they are not being used as effectively as possible in a 6-hourly updating system. Kim and Kim (2019) found that observation impacts were greater when shorter background forecasts were used. It is concievable then that more frequent assimilation and the assimilation of observations within 1 hour of their receipt time could make better use of the observations. A 4DVar system such as that employeed by the European Center for Medium-Range Weather Forecasts (ECMWF) which uses continuous DA (Lean et al. 2021) could also be implemented which would also use observations as they become available; however, the challenges of the 4DVar approach likely stem from the use of the linear model. Third, 6-hourly analysis update interval in the presentday GDAS may not be frequent enough to best constrain the position of hurricanes and other convective storms. Longer update cadences, requiring longer background forecasts, will inherintly include larger error in the background fields compared to shorter update cadences. For the TC problem, the current 6-hourly assimilation cadence often struggles to accurately constrain the position of hurricanes which can lead to double vorticies and other spin-up/spin-down issues (Chen and Snyder 2007; Lin et al. 2018). To combat large position errors that can result in double vorticies and other spin-up/spin-down issues, the current

[1] https://www.ospo.noaa.gov/data/messages/2019/MSG0861424.html
[2] http://www.jma.go.jp/jma/jma-eng/satellite/news/himawari89/himawari89_leaflet.pdf

operational Hurricane Weather Research and Forecasting model (HWRF) uses "vortex initialization" which includes a "vortex relocation" scheme to correct the background fields by adjusting the hurricane position prior to assimilating observations (Liu et al. 2020b). Another motivator for an hourly updating global systems is that more frequent updates should better constrain such systems through shorter and theoretically more accurate background forecasts with smaller displacement errors between each analysis update. One of the main challenges of implementing an hourly updating, global system is the handling of data latency. This chapter discusses two potential configurations for handling data latency: catch-up cycles and asymmetric, overlapping windows.

This dissertation is motivated by the successes of the assimilation of radial winds in previous idealized work (3-km grid spacing or less with sub-hourly updates), and yet often unclear impacts in the current operational systems (3-km grid spacing with 1-hourly updates). We explore some of the more feasible aspects of the idealized configurations of radial wind assimilation that could more readily be applied to the operational systems. We first investigate the impacts and sensitivity of radial wind assimilation in an operational, convection-permitting system with specific attention to the formulation of the observation operator and super-observation pre-processing. Second, we seek to extend radial wind assimilation to a global system to investigate the impacts and sensitivities of radial wind assimilation within the idealized context of an OSSE. And third, we formulate an hourly updating version of the GFS and address issues of data latency and investigate the impacts of assimilating radial wind observations in a real-data experiment of a land-falling tropical cyclone. The main goal

of this dissertation is to understand how the assimilation of radial wind observations could lead to improvements in the operational regional and global systems.

1.2 Background

Weather radar networks, such as the United States' Weather Surveillance Radar 1988 Doppler (WSR-88D) Next Generation Weather Radar (NEXRAD; Crum and Alberty 1993) network, are an important aspect of the weather forecasting and warning infrastructure of nations around the globe. The NEXRAD network consists of 160 irregularly spaced radars, each able to provide a single volume scan in about 5-min. Each volume scan may include elevation angles ranging from 0.5 to 19.5-degrees, with azimuthal and along beam range sampling resolutions of 1-deg and 250-m respectively. As one of the only observing networks capable of observing storm-scale motions[3], the NEXRAD network plays a crucial role in the severe weather warning infrastructure of the United States (e.g., improved tornado warning lead times with NEXRAD; Bieringer and Ray 1996). These in-storm observations coupled with recent and future technological and scientific advancements in the operational forecasting community are already playing a crucial role in integrating frequently-updating, convective-scale, model-based probabilistic forecasts into the warning decision-making process (Lawson et al. 2018; Stensrud et al. 2009; Wheatley et al. 2015). This Warn-on-Forecast paradigm is opposed to the warning strategies focusing primarily on Doppler radar and other observation-based detection methods. Ultimately, this will improve the warning lead times of a variety of hazardous weather phenomena including severe

[3] For example, Doppler radar is an important tool for detecting tornado vortex signatures (or TVS) which are patterns in the radial wind field that indicates an area of intense concentrated rotation.

thunderstorms, flash floods, and tornadoes to further enhance NOAA's strategic mission goal designed to reduce the loss of life, injury, and economic costs due to these types of events.

The prediction of a particular scale of atmospheric motion not only requires a model with sufficient resolution and an accurate representation of the major processes at those scales, but also requires that model to be initialized by assimilating observations that can describe that scale (Sun 2005b). One such observational data set for convective scales are Doppler radar radial winds[4].

1.3 History of Radar

Whiton et al. (1998a, 1998b) provide a detailed history of radar technology from its earliest beginnings in 1935 when the British physicist Sir Robert Watson-Watt began investigating a method of detecting aircraft with the use of radio waves (i.e., Radio Detection and Ranging—RADAR). The concept of using radar to observe the weather, however, was not conceived until World War II when radar operators discovered that the weather could cause radar echoes and potentially mask enemy targets. Due to the need for secrecy surrounding the war, these important findings were not disclosed until the war ended in 1945. Special interest governmental agencies prioritized these findings and eventually established the U.S.'s first weather radar network in the 1960s which consisted of 66 WSR-57 radars. By the 1970s, the U.S. radar network was expanded to help better serve the operational forecasting needs by adding 62 WSR-74 radars to the existing WSR-57 network. Neither radar model,

[4] Radial wind is the wind motion directly toward or away from a Doppler radar as measured along the radar beam path.

however, had Doppler capabilities and thus were not able to report wind speed and along-beam direction, which are crucial to detecting rotation in thunderstorms.

Crum and Alberty (1993) provide a detailed overview of the NEXRAD program which began with the formation of the Joint Doppler Operational Project (JDOP) in 1976 at the National Severe Storm Laboratory (NSSL) in Norman, Oklahoma. The mission of the JDOP was to investigate the use of a Doppler radar to identify severe thunderstorms and tornadic features in real-time (JDOP 1979) which eventually led to the decision to replace all non-Doppler meteorological radars currently employed by the NWS with the WSR-88D. The implementation of WSR-88Ds into production began in 1992 with the final radar installed in 1996. Fig. 1 shows the completed NEXRAD sites as of 2010. NCEP first assimilated radial winds during the 1996 summer Olympics in Atlanta, Georgia; however, it was not until 2006 that radial winds were assimilated operationally in the 12-km North American Mesoscale (NAM[5]) forecast system.

[5] https://www.emc.ncep.noaa.gov/impdoc/NAM/WRFinNAM/Update_WRF-NMM_replacing_Eta_in_NAM2.pdf

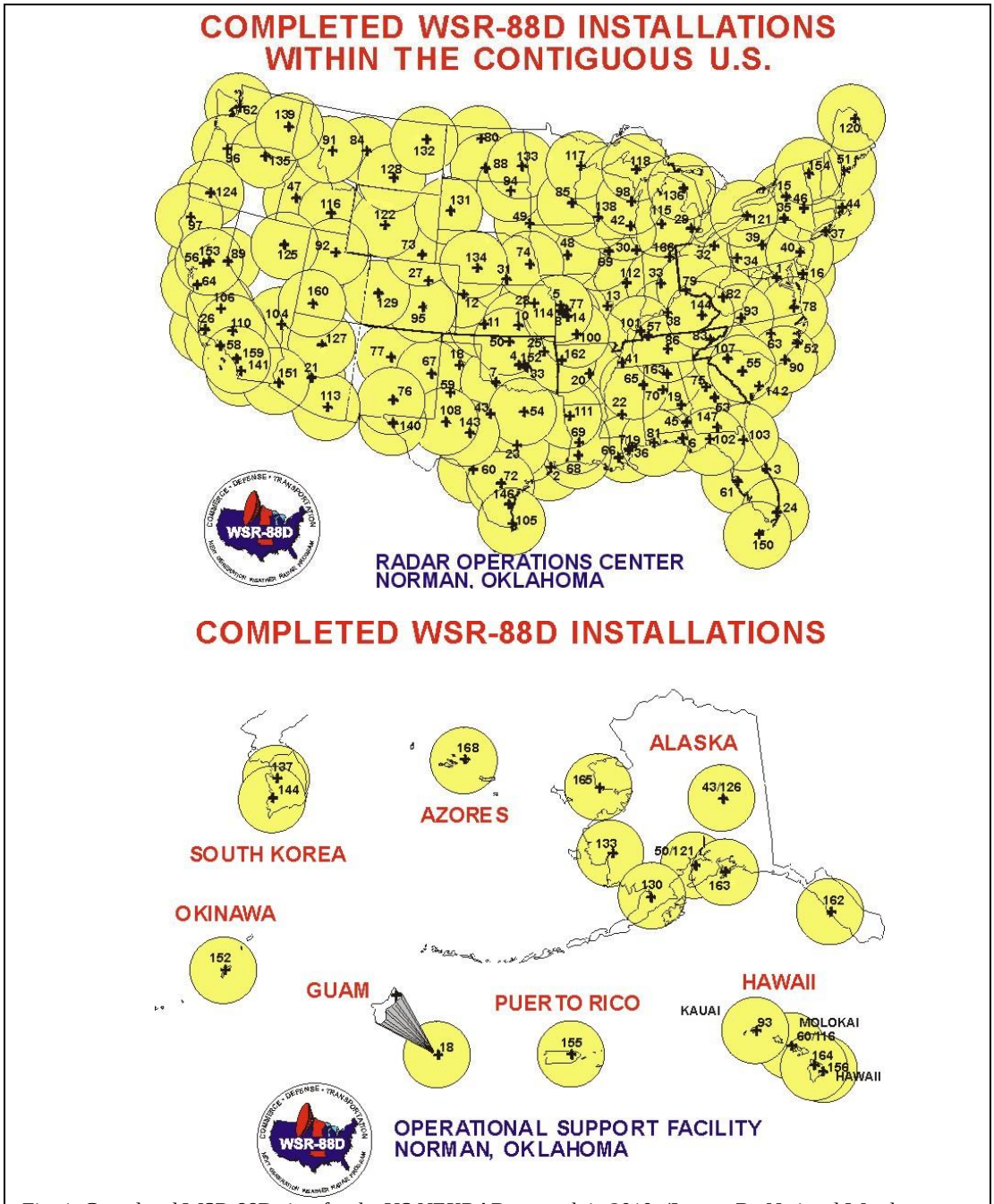

Fig. 1. Completed WSR-88D sites for the US NEXRAD network in 2010. (Image: By National Weather Service NOAA/Public domain)

1.4 Fundamentals of Doppler Weather Radar

1.4.1 The Basics

Before diving deeper into this radar focused thesis, it is important to understand the basic operating principles of Doppler weather radar. There are many types of radar, but here, we will discuss mainly the WSR-88D radars in operations in the U.S. today. These weather radars can be further classified as *cloud pulse radars*, meaning that they emit short pulses of electromagnetic energy at regular intervals into the atmosphere and then wait and "listen" for a return signal scattered back by precipitation sized scatterers. This process of transmitting and listening repeats several hundred to over one thousand times per second. The number of pulses per second (Hz) is called the pulse repetition frequency (PRF). The PRF of a radar can be tuned for observing the phenomenon of interest and can range from 318Hz to 1304Hz for *short pulses* and 318-452Hz for *long pulses* (Heiss et al. 1990). The PRF is directly linked to the *Doppler dilemma* which will be discussed later in this section.

Each pulse travels at the speed of light (approximately $3x10^8$ ms^{-1}) until it hits a target, the target then scatters (Rayleigh scattering) that energy in all directions and only a fraction of the original amount of energy is scattered directly back toward the radar. The amount of energy returned or "reflected back" (i.e., reflectivity) can then be related to the object's intensity which is a function of the object's size, shape, state, and concentration. The radial distance, r, of each intensity measurement can also be computed using the speed of light, c, and half the time, T, it took for the signal to return to the radar since the signal must travel from the radar to the target and back,

$$r = \frac{cT}{2} \tag{1}$$

There are two types of scattering: Rayleigh and Mie scattering. In the Rayleigh region, the backscattering cross-sectional area of a sphere is proportional to the sixth power of the diameter. Rayleigh scattering occurs when the target diameter, D, is much smaller than the wavelength of the transmitted radio wave ($D << \lambda$). "Much smaller" is typically interpreted to mean $\frac{D}{\lambda} \leq \frac{1}{10}$ but some specify $\frac{D}{\lambda} \leq \frac{1}{16}$ for this condition to apply (Rinehart 2004). The WSR-88D's wavelength is approximately 10.7 cm, so Rayleigh scattering occurs with targets whose diameters are ≤ 0.7 cm. Raindrops usually do not exceed 0.7 cm (Marshall and Palmer 1948; Yakubu et al. 2016). Therefore, it is generally safe to assume all liquid drops are Rayleigh scatterers; however, it is important to know if the targets are liquid, ice, or something else. The backscattering cross-sectional area is dependent on the material, the temperature, and the wavelength of the radar so proper adjustments to the calculation will be required otherwise there could be significant errors in the reflectivity measurements. Battan (1973) provides a few tables for the appropriate values for the parameter related to the complex index of refraction for water and ice at various temperatures and radar wavelengths.

1.4.2 Volume Coverage Patterns

Weather radars continuously scan the atmosphere by completing Volume Coverage Patterns (VCPs). Rotating about a vertical axis, the radar antenna scans all directions surrounding the radar site one elevation angle at a time. Once a full rotation is completed, the elevation angle of the antenna is incremented to the next angle in the

current operating VCP. Sometimes the antenna will revert to a lower elevation mid volume scan for better updates of the lower levels. Rinehart (2004) puts things into a great perspective to fully appreciate how incredible weather radar is,

> "...Even though the antenna may be scanning at a speed of from 10°/s to as much as 70°/s (up to 10 RPM or faster), the speed of light is *much* faster. It only takes 2ms (milliseconds) for a radar signal to travel about 300km and back. If the antenna is rotating at 10 RPM, the antenna will move less than a tenth of a degree during this time. So, from the human perspective, the signal went out and back with the antenna virtually stationary. The radar repeats its transmission/listening cycle several hundred to a couple thousand times each second. So, we mere mortals watch the antenna moving smoothly and fairly rapidly, but we often don't appreciate the fact that the signal is going out and back very long distances at speeds that are almost incomprehensible to the human mind...".

There are two main operating states of the WSR-88Ds. Those states include clear-air mode and precipitation mode. For each of these states, there are multiple VCPs which can be used and is chosen in real-time to provide the most effective coverage of the local weather phenomena. The VCPs can be broken down into the following categories with the naming conventions according to the Federal Meteorological Handbook No. 11 WSR-88D Meteorological Observations Part C from the OFCM (2017) (see acronym table):

1. **Deep convection group**: two digits beginning with 1; VCPs 11 and 12
2. **Shallow convection group**: two digits beginning with 2; VCP 21
3. **Multiple PRF dealiasing algorithm and Sachidanananda-Zirnic group**: three digits beginning with 1 or 2; VCPs 121, 211, 212, and 221
4. **Clear air group**: two digits beginning with 3; VCPs 31 and 32.

Each VCP has its own strengths and weaknesses and for more information regarding those and the intricacies of each, the reader is referred to the NEXRAD technical documentation of VCPs OFCM (2017).

1.4.3 Operating Bands

The WSR-88Ds of the NWS can also be classified as S-band radars meaning they operate with a frequency range of 2-4GHz. The WSR-88Ds specifically operate at 2.7-3.0 GHz or with a wavelength of approximately 10.7 cm (Heiss et al. 1990). Because of their wavelength and frequency, the WSR-88D S-band radars do not often suffer from signal attenuation but can happen under large hail cores or in extremely heavy precipitation. Because attenuation is minimal in the WSR-88Ds, this makes them useful for near and far range weather observations. The drawback to this band is that it requires a large antenna dish and a large amount of energy to power it.

Attenuation is the reduction in energy via scattering and absorption of a radar pulse as it travels through a medium. It occurs when large or a high concentration of objects scatter most of the energy back to the radar, leaving little to travel down radial to distant storms. The amount of attenuation depends on the type and concentration of the medium that the signal is traveling through as well as the wavelength of the radar signal itself. Smaller wavelengths (i.e., higher frequencies) attenuate at a faster rate than longer wavelengths because higher frequencies are more easily scattered and absorbed by clouds, rain, hail, etc. The backscattering cross-sectional area of the target is proportional to the inverse of the wavelength to the 4^{th} power meaning radars with smaller wavelengths are more sensitive and thus can detect smaller particles. It also means that smaller wavelengths experience more attenuation. Attenuation can also

cause down-radial storms to appear weaker by underestimating the return signals because of the unexpected decrease in energy of the original pulse or even block the beam entirely causing down-radial storms to go completely unobserved.

In contrast to the WSR-88D S-band radars, there are also many other operating bands for radars. For example, the C-band radar which operates on a wavelength of 4-8 cm and a frequency of 4-8 GHz. At this wavelength, the antenna dish does not need to be as large and requires significantly less energy to operate than the S-band radar. Because C-band are more easily attenuated by precipitation than the S-band radar, they are mostly useful for close range weather observation and mostly used by the FAA for the Terminal Doppler Weather Radar (TDWR) network and local TV stations. There are other operating bands including X-band. These radars can be smaller yet, however, they are also more easily attenuated than C-band radars but can detect much smaller particles. This makes X-band radars optimal for field research applications where mobile radar platforms are particularly useful, such as studying tornadogenesis (e.g., Bluestein et al. 2019). NSSL maintains a mobile X-band radar known as NOAA-X-POL (NOXP) used for field research. There are also radars with even higher frequencies which include K-band and mm or W-band.

1.4.4 Doppler Velocity

At the heart of this thesis is Doppler radial velocity. Most of what has been discussed thus far has been related to the intensity of the returned signal, which is proportional to the well-known quantity of equivalent radar reflectivity factor. Doppler velocity is a measurement of the phase shift of the returned signal over multiple pulses. To understand this measurement, it is key to first understand the Doppler effect. The

Doppler effect is the phenomenon that you may have experienced as a car passes you. The sound waves from the approaching car are compressed, having a higher frequency, and have a higher pitched sound. This is also known as a positive phase shift. Once the car has passed and is moving away from you, the sound waves are stretched, having a lower frequency, and producing a sound with a lower pitch. The same phenomenon happens with all waves, including electromagnetic radiation, which is why traffic police can determine how fast you are driving.

Weather radar measures the change in wave phase or the fractional offset of a full wavelength and determines whether the target is moving toward or away from the radar by whether the phase shift is positive or negative, respectfully. The larger the phase shift, the higher the target's radial velocity. To measure this phase shift, we need to know the distance to the target in radians, which can be computed from the distance measured in the number of wavelengths. Suppose the target is a distance of r from the radar using Eq. (1), then the total distance, r_{total}, that the radar beam has traveled is given by,

$$r_{total} = 2r \tag{2}$$

This distance can then be converted into the number of wavelengths, $r_{wavelengths}$, the radar is from the target by dividing by the wavelength,

$$r_{wavelengths} = \frac{2r}{\lambda} \tag{3}$$

and then using the fact that $1\ \text{wavelength} = 2\pi\ \text{radians}$, we can convert (2) into distance measured in radians, $r_{radians}$,

$$r_{radians} = \frac{4\pi r}{\lambda}. \tag{4}$$

Having the distance in radians is useful because then it is possible to measure the phase shift of the electromagnetic wave by measuring from some reference point (i.e., one for cosine and zero for sine) along the wave pattern. By using this reference point, the phase shift will never be more than $\pm\pi$ radians.

To measure the phase shift, we must take the time derivative of the returned radar phase. Suppose the radar transmits an initial phase of ϕ_0, then the phase of the return signal will be

$$\phi = \phi_0 + \frac{4\pi r}{\lambda}. \tag{5}$$

The phase shift from pulse to pulse is then given by the time derivative of Eq. (5),

$$\frac{d\phi}{dt} = \frac{4\pi}{\lambda}\frac{dr}{dt} \tag{6}$$

where $\frac{dr}{dt}$ is equal to the velocity of the target. Equation (6) represents the angular frequency. Regular frequency is the measure of the number of full oscillations per second. Angular frequency $\frac{d\phi}{dt}$ is the measure of number of radians per second. Thus if, for example, there is a wave that oscillates one full oscillation per second, that wave would have a frequency of $1\ \text{oscillation}\ s^{-1}$ (Hz). The angular frequency of the same wave would have an angular frequency of $2\pi\ \text{radians}\ s^{-1}$. Thus, we can define angular frequency as a factor of 2π larger than the frequency

$$\frac{d\phi}{dt} = 2\pi f_d \tag{7}$$

where f_d is the frequency shift of the returned signal. Then combining (6) and (7) we can write the radial velocity, V_r, as

$$V_r = \frac{f_d\lambda}{2}. \tag{8}$$

With this level of understanding of radial velocity, it is now possible to discuss the physical limitations of this measurement.

There are limitations in the velocities and ranges that a radar can unambiguously analyze. In other words, the radar could have an incorrect interpretation of the actual velocity and range of a target. As a simple example, let us say that the maximum unambiguous range is 150km (a PRF of 1000Hz) and there is a storm at 175km. The speed of light is extraordinarily fast, but it is not instantaneous; therefore, it requires some amount of time to travel 150km. In fact, it would take roughly 1ms to travel out and back. Therefore, in this example, the radar is emitting one pulse every 1ms; however, the return signal from the storm at 175km took 1.16ms to return and is now overlapping with the subsequent pulse which was emitted only 0.16ms earlier. The radar assumes this return signal is from this second pulse and now, incorrectly, interprets this signal as a weakened storm at 25km since it only took 0.16ms to return. The storm is weakened because of the loss of energy over the actual distance the pulse has traveled compared to what the radar interpreted it to have traveled. Essentially, overlapping return signals from subsequent pulses can have multiple interpretations on the range, in other words the range of the storm is ambiguous beyond a particular range given a specified PRF. There are also ambiguities that can arise in the velocity which will be discussed next.

1.4.4.1 Maximum Unambiguous Velocity

Recall that the phase shift should never meet or exceed $\pm\pi$ radians. A phase shift of $+\pi$ radians is indistinguishable from a phase shift of $-\pi$ radians because both represent the same phase shift of 180° which is actually no phase change, and the radar

would assume the target is stationary. If the amount of time between pulse transmissions is long (i.e., low PRF) this greatly reduces the range of velocities that would produce a phase shift of less than $\pm\pi$ radians. Thus, only slower velocities would be able to be resolved unambiguously.

The Shannon-Nyquist sampling theorem states that to measure a frequency, f_d, it is necessary to sample at a frequency of at least $2f_d$; therefore, $PRF = 2f_d$. Using this relationship, and Eq. (8), we can find the maximum unambiguous velocity, V_{max}, for any given combination of PRF and wavelength

$$V_{max} = \frac{\pm PRF \cdot \lambda}{4} \tag{9}$$

Conceptually, this equation means that longer wavelengths and higher PRFs yield a wider range of velocities that would cause a phase change of less than $\pm\pi$ radians. Such a configuration would be advantageous for sampling weather phenomena with strong winds and wind shear (i.e., severe storms, land falling tropical storms, etc.). A higher PRF would provide more frequent updates between pulses reducing the phase change for a given velocity. However, as PRF increases the range is reduced.

1.4.4.2 Maximum Unambiguous Range

If the amount of time between pulse transmissions is short (i.e., high PRF) then there is less time for the radar to "listen" for return signals. This greatly reduces the distances that a radar can unambiguously determine because the beam has less time to travel to a target and back. The equation that relates the maximum unambiguous range, r_{max}, and the PRF is

$$r_{max} = \frac{c}{2PRF} \tag{10}$$

which is similar to Eq. (1) since $T = \frac{1}{PRF}$. This relationship shows that maximum range is inversely proportional to the PRF. As PRF increases, the maximum range decreases.

1.4.4.3 Doppler Dilemma

The maximum unambiguous velocity and range define the constraints of the Doppler Dilemma. Simply, a low PRF having longer time between pulse transmissions allows the signal to travel to more distant objects but reduces the ability to distinguish velocities. On the other hand, a high PRF having shorter time between pulses improves the ability of the radar to determine velocity but does not allow the signal to travel as far. Using Eq. (9) and (10) the Doppler Dilemma can be mathematically explained as

$$V_{max} r_{max} = \frac{c\lambda}{8} \tag{11}$$

For a given radar, the right-hand side of Eq. (11) is constant, thus if we want our radar to be able to unambiguously resolve a high V_{max} we must compromise for a low r_{max} and vice versa. Having a large V_{max} is important because it is crucial to the NWS mission to be able to detect and warn against the threat of an eminent tornado via a tornadic vortex signature from radial winds as well as other sources of strong and damaging winds that pose a threat to the public. Furthermore, longer ranges are only helpful to a certain point. At some distance the height above ground level becomes so large that it is not particularly helpful to have a long r_{max}.

While there are many different VCPs used depending on the current local weather patterns, the most common operational VCPs generally follow a similar surveillance scheme for PRFs. PRFs have a number code that ranges from 1-8 and each of which is described in Table 2 which is a reproduction of table 5-10 of OFCM (2017).

PRF 1 is used for the lowest scan angles of ≤ 2°; PRF 2 is used for angles ≤ 4°, and PRF 3 is used for angles < 7°. Beyond 7°, contiguous Doppler without range unfolding (CDX) is used. At these higher scan angles, a high PRF is used because the likelihood of range folding is small for a beam that is within reasonable height above ground level. For example, at an elevation angle of 8.0°, the radar beam will be approximately 17 km (56 kft) above ground level, well above the troposphere, at 117 km from the radar which is the minimum R_{max} of all the PRFs used. While the minimum R_{max} for each PRF 1, 2, and 3 is 233 km, the radar routine algorithms in our data assimilation system only use an R_{max} of 100 km to avoid signal quality issues.

Table 2. Table 5-10 of OFCM (2017) showing the PRF identification number and corresponding characteristics including the Pulse Range Frequency (PRF), maximum unambiguous range (R_{max}), and maximum unambiguous velocity (V_{max})

PRF No.	**PRF (s^{-1})**	**R_{max} (km)**	**V_{max} (ms^{-1})**
1	322	467	8
2	446	336	11
3	644	233	16
4	857	176	22
5	1014	148	26
6	1095	137	28
7	1181	128	30
8	1282	117	33

1.4.5 Refractivity

Refraction is another important aspect of radar meteorology. Refraction is the bending of light as it travels at an angle from one substance into another substance with different refractive indices. According to Snell's law, the bending of light is caused primarily by two factors: the change in refractive index from substance to substance

which causes a change in the speed of light through the materials and the angle of incidence that the ray is traveling. If the light enters the new substance at a 90° angle, the light will still slow down, but it will not change direction. The refractive index of the troposphere has been found to depend on atmospheric pressure, temperature, and water vapor (i.e., density). It is not the refractive index of the material the ray is traveling through but the *change* in refractive index. Therefore, if there was no atmosphere or if the atmosphere was perfectly uniform everywhere, radar rays would travel in straight lines.

Let us first consider straight line optics. If the radar beam was oriented perfectly parallel to the earth's surface at a given location and ignoring any effects of refraction, the beam would appear to an observer on the earth to bend upward as the beam travels farther and farther from the radar location. This is the result of the curvature of the earth. We can write curvature as $\frac{\delta\theta}{\delta S}$ where $\delta\theta$ is the change in angle experienced with change in distance, δS. Considering a circle of radius R. Traveling a distance around the circle would give $\delta S=2\pi R$ which is just the circumference and the angular change in traveling around that circle would be $\delta\theta=2\pi$ radians. Therefore, curvature can just be written as

$$C=\frac{\delta\theta}{\delta S}=\frac{2\pi}{2\pi R}=\frac{1}{R} \tag{12}$$

and is just the reciprocal of the radius of that circle.

Now, for a radar ray traveling through a non-uniform atmosphere the curvature of the radar ray will change based on the changes in the refractive index with height $\frac{\delta n}{\delta H}$.

Battan (1973) gives the following equation to describe the apparent curvature of the radar beam due to the curvature of the earth and due to atmospheric refraction.

$$\frac{\delta\theta}{\delta S}=\frac{1}{R}+\frac{\delta n}{\delta H} \tag{13}$$

It is common to use Eq. (13) to estimate an effective earth curvature, R' using standard atmospheric conditions of $\frac{\delta n}{\delta H}=\frac{-39\cdot10^{-6}}{\text{km}}$, thus we have

$$\frac{1}{R'}=\frac{1}{R}+\frac{\delta n}{\delta H}=\frac{1}{6374\text{km}}+\frac{-39\cdot10^{-6}}{\text{km}}=\frac{1.179\cdot10^{-4}}{\text{km}} \tag{14}$$

or an effective earth radius R' = 8483 km ≅ 1.33R ≅ 4/3 R. Hence, this estimation of an effective earth radius under standard atmospheric conditions is known as the "4/3rds Rule".

The magnitude and direction of beam bending is the result of the combined effects of earth's curvature and atmospheric refraction with increasing distance from the radar. Earth's curvature always results in the apparent bending of the beam upward. Since atmospheric density is variable, the effects of atmospheric refraction can thus bend the beam in either direction, but most typically will bend the beam downward. Under standard atmospheric conditions, the net beam bending is upward due to earth's curvature being greater than the downward bending due to standard atmospheric refraction. The height of the radar beam above radar level can be calculated using the assumption of *standard refraction* and is given by the equation from Rinehart (2004) as

$$H=\sqrt{r^2+\left(\frac{4}{3}R_e\right)^2+2r\left(\frac{4}{3}R_e\right)\sin\phi}-\frac{4}{3}R_e \tag{15}$$

Where H is the height of the beam above radar level in km, r is the range (distance) to the target in km, ϕ is the elevation angle, and R_e is the radius of the earth. Fig. 2 uses the so-called 4/3rds Rule to show the height at various ranges for various elevation angles.

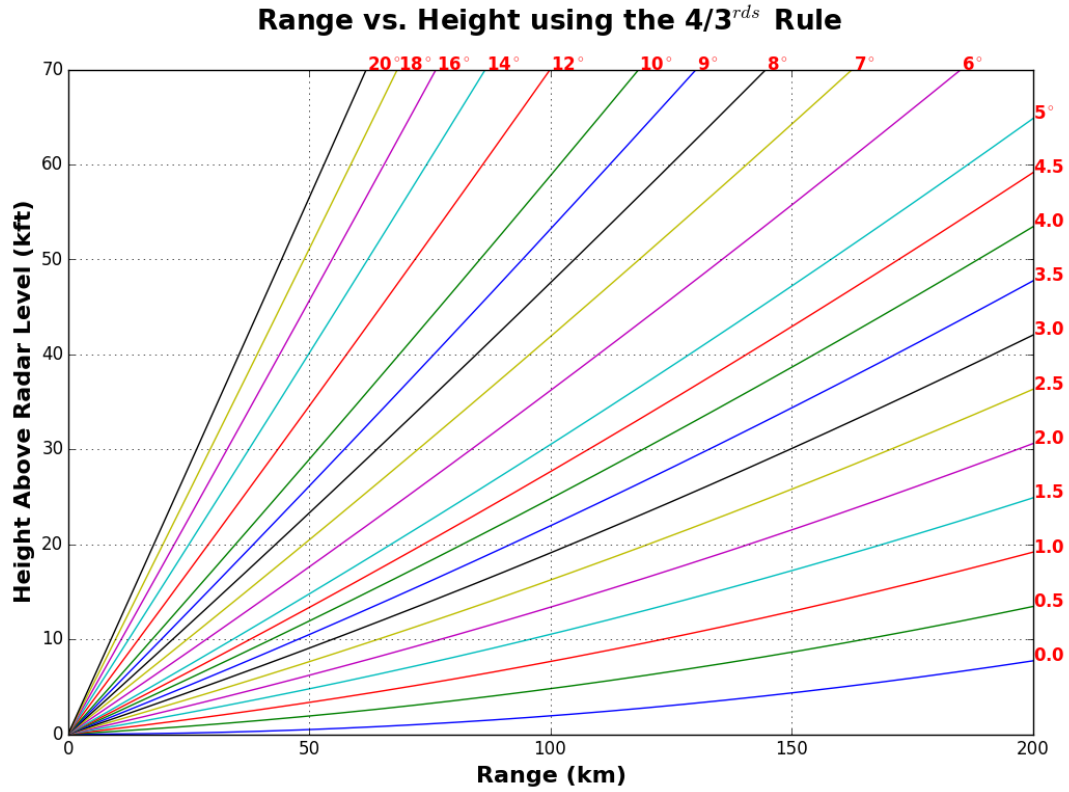

Fig. 2. Range (km) vs height (kft) of various radar scan elevation angles using the 4/3rd approximation.

The atmosphere does not always conform to standard-conditions, so there are times in which this rule is invalid, and the amount of beam bending can be greater or less than the amount of bending under standard conditions, especially in the boundary layer. The case where the beam bends more than the standard rate is called *super refraction*. This can cause low angled radar beams to sense more ground clutter since the beam is bent downward at a higher rate. Conversely, there is also *subrefraction*

which is the case where the amount of bending is less. Subrefraction poses a challenge because low elevation angles may not be able to identify storms at low levels.

The 4/3rds Rule is the methodology used in practice, and while it is generally a reasonable approximation, it is an approximation and will not always be valid especially in the boundary layer thus is mostly a potential issue for the lower scan angles. The alternative to using this approximation would be to know the vertical profile of refractive index. There have been studies which seek to estimate the low level refractivity profile using radars and point targets (e.g., Feng et al. 2016). In the future, the 4/3rds rule approximation could be modified by a correction factor for improving the height estimation of radar observations, but this is beyond the scope of this thesis.

1.5 A Review of Radar Radial Wind Data Assimilation

This section presents a review of data assimilation methods and their use primarily as it relates to the assimilation of Doppler radar radial wind observations. This section is organized by beginning with a brief history of how radar observations were first used in retrieval methods for estimating fields that are not directly observed by radar. This will lead into a discussion of the basic principles of data assimilation as well as discussion about different data assimilation methods such as variational approaches, ensemble approaches, and then close with hybrid methods that combine the advantages of both variational and ensemble methods. Following each subsection on the various DA methods is a discussion on how each method has been used as it pertains to the assimilation of radial wind observations.

1.5.1 Single and Multiple Doppler Retrieval Methods

The U.S. WSR-88D network (Fig. 3) resulted in a great observational advancement for not only the detection of storms but also for storm-scale DA and NWP. With the implementation of the radar network in the early 1990's, radar DA has received a great deal of attention because Doppler radar is one of the only instruments capable of sampling the four-dimensional, storm-scale environment needed to initialize convection-permitting models. The most used observations from Doppler radar are radar reflectivity and radial winds, both of which are not model prognostic variables. However, reflectivity and radial winds can be related to other prognostic model variables. For example, one may attempt to indirectly assimilate radar observed reflectivity by diagnosing the latent heating that occurs because of precipitation processes, or radial winds could be used to retrieve an estimate of the three-dimensional wind components.

There are other observations that come from Doppler radar and those include products resulting from the upgrade to Dual-Polarization (Dual-Pol) capabilities which provide additional information about the shapes of the targets in a radar volume. Such types of observations include differential reflectivity (ZDR), correlation coefficient (CC), specific differential phase (KDP), as well as many other derived products which include hydrometeor classification and precipitation estimation. The assimilation of Dual-Pol products is beyond the scope of this project and might be better suited for future considerations.

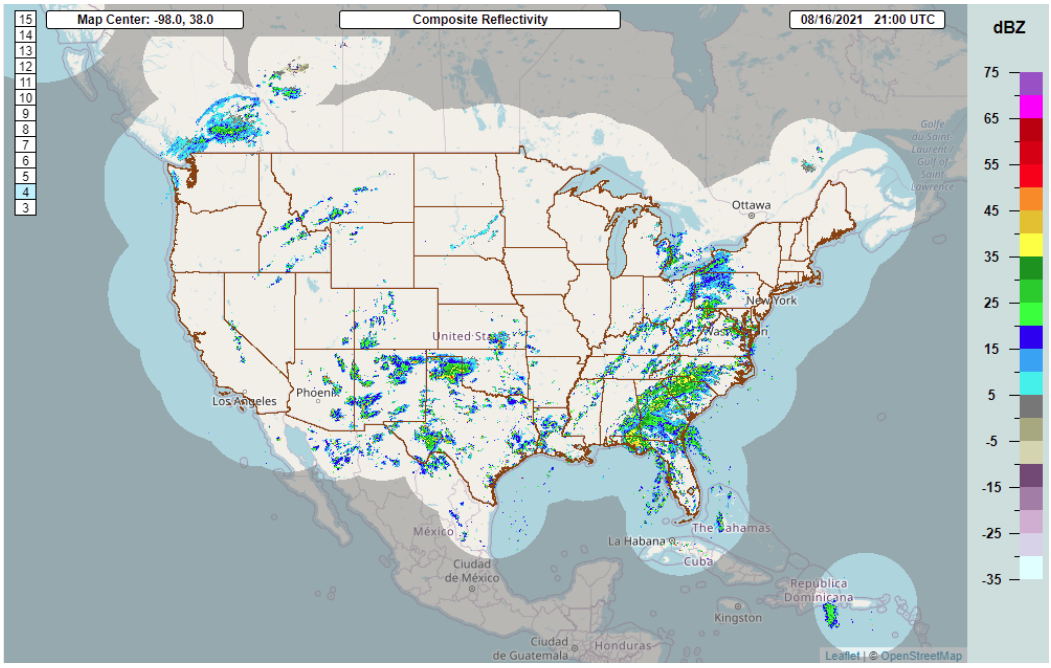

Fig. 3. Example of radar observational coverage capabilities. Composite reflectivity (dBZ) from land falling Hurricane Fred August 2021. (Image: Captured from the MRMS Operational Product Viewer https://mrms.nssl.noaa.gov/)

Historically, the initialization of modeled convective storms began as researchers demonstrated, by using Dual-Doppler observations, that the wind field could be derived through the use of the continuity equation (Armijo 1969) and thermodynamic fields through the use of the equations of motion (Gal-Chen 1978; Hane and Scott 1978). Dual-Doppler radar is simply where two or more radars are positioned such that they have overlapping coverage, providing a better determined problem. This type of analysis is typically exclusive to research applications such as with a mobile radar since the spacing of the NEXRAD network is too large. Following these dual-Doppler retrieval methods, Lin et al. (1993) developed a method to initialized a convective model with the derived 3D fields of velocity, rainwater, temperature, and pressure. They were able to initialize the tornadic storm from the 20 May 1977, Del City, Oklahoma case with good agreement.

Since the US NEXRAD network is not spaced geometrically such that multiple- or even dual-Doppler analyses are possible without an external aid, such as a mobile radar, single-Doppler analyses have also been developed (e.g., Weygandt et al. 2002). Weygandt et al. (2002) developed a single-Doppler retrieval method using the wind retrieval technique of Shapiro et al. (1995) to estimate the three-dimensional wind field from a time series of single-Doppler observations of reflectivity and radial wind. The thermodynamic retrieval of Gal-Chen (1978) is then used to estimate temperature and pressure from the three-dimensional wind fields. While this technique proves efficient and inexpensive, it requires multiple steps, and the quality of the thermodynamic retrieval depends greatly on the quality of the retrieved three-dimensional wind in the first step. Therefore, the use of such retrieval methods and the underlying assumptions and associated errors hinders the optimal analysis that could be achieved (Hu et al. 2006). Instead, the direct assimilation of radar observations has become common practice.

1.5.2 Bayesian Principles

Numerical weather prediction (NWP) is the foundation on which nearly all operational forecasts are made. NWP is a process in which the governing equations of the atmosphere are discretized onto a three-dimensional grid and integrated forward in time to predict future states of the atmosphere.

NWP is an initial and boundary value problem, and here our focus is on the initial values, meaning that we must know the current state of the atmosphere to predict a future state of the atmosphere. Through a process known as data assimilation, an estimate of the initial conditions is created to represent the atmosphere's current state

using Bayes Theorem. Because NWP is an initial value problem, if the model starts with a poor estimate of the initial state of the atmosphere, the resulting forecast will also, quite likely, be poor. Initial conditions are a very important aspect of NWP because even small errors at the start can quickly grow through the simulation, resulting in a forecast dominated by large error at extended time periods (Lorenz 1963). It is important to have the most accurate initial conditions as possible to have the greatest chances of accurately predicting future states of the atmosphere.

One of the most important formulas in all of probability theory and one that is central to data assimilation is Bayes theorem. Lorenc (1986) demonstrated how Bayesian principles could be extended to numerical weather prediction for determining the best analysis of the state. Bayes theorem is most generally written as

$$\mathrm{p}(\mathbf{x}|\mathbf{y}^o) = \frac{\mathrm{p}(\mathbf{y}^o|\mathbf{x})\mathrm{p}(\mathbf{x})}{\mathrm{p}(\mathbf{y}^o)} \tag{16}$$

where, in data assimilation, $\mathbf{x}$ is the state and $\mathbf{y}^o$ are the observations. Bayes theorem provides a theoretical framework for data assimilation and is a way to estimate the *posterior* probability distribution function (PDF), $\mathrm{p}(\mathbf{x}|\mathbf{y}^o)$, or the conditional probability of event $\mathbf{x}$ to occur given that event $\mathbf{y}^o$ has occurred. The right-hand side of the equation is product of three probabilities resulting in the posterior probability. The three probabilities on the right hand side are: $\mathrm{p}(\mathbf{y}^o|\mathbf{x})$, which is the conditional probability that $\mathbf{y}^o$ will occur given that $\mathbf{x}$ has occurred; $\mathrm{p}(\mathbf{x})$, which is the *prior* or the independent probability that $\mathbf{x}$ will occur; and finally, $\mathrm{p}(\mathbf{y}^o)$, which is the independent probability that $\mathbf{y}^o$ will occur and is unity since the observations are given. Let us now briefly describe what each of these variables mean in the data assimilation problem (e.g., Lorenc 1986)

- $\mathrm{p}(\mathbf{x}|\mathbf{y}^o)$ is the *posterior* PDF or the updated PDF of the state after the analysis.
- $\mathrm{p}(\mathbf{y}^o|\mathbf{x}) \propto \mathrm{L}(\mathbf{x}|\mathbf{y}^o)$ is the likelihood or the PDF of the observations conditioned on the state.
- $\mathrm{p}(\mathbf{x}) = \mathrm{p}(\mathbf{x}^b|\mathbf{x})$ is the *prior* PDF which contains our knowledge about the state before new observations. This is given by the background forecast, $\mathbf{x}^b$.
- $\mathrm{p}(\mathbf{y}^o)$ is the marginal PDF of the observations, $\mathbf{y}^o$, and it does not depend on $\mathbf{x}$. In Bayes' theorem, this is regarded as a normalizing factor such that the area under the posterior PDF is unity. Therefore, only the terms in the numerator are required to compute the posterior.

Therefore, we can now rewrite Bayes' theorem as

$$\mathrm{p}(\mathbf{x}|\mathbf{y}^o) \propto \mathrm{p}(\mathbf{y}^o|\mathbf{x})\mathrm{p}(\mathbf{x}) \tag{17}$$

Solving the Bayesian problem is difficult in the absence of any assumptions about the shapes for the prior and likelihood of the observations. Therefore, it is common practice to assume that these take the form of Gaussian distributions; although, there are DA methods that do not rely on this assumption (e.g., Particle Filters; Poterjoy et al. 2019). The Gaussian assumption is generally reasonable and greatly simplifies the analysis problem due to certain properties of Gaussian distributions. For example, a product of two Gaussians is a Gaussian, Gaussians remain Gaussian under linear transformation, the mean and covariance are computable, and matrices are invertible. The general form of the single-dimensional Gaussian probability density function is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{18}$$

where $f(x)$ is the PDF, σ is the standard deviation, μ is the mean or expectation of the distribution. Now we need to translate the general form into a multidimensional problem. To aid us, the general form of the normal PDF is given by

$$f(x_1, ..., x_N) = \frac{1}{\sqrt{(2\pi)^N |\mathbf{\Sigma}|}} e^{-\frac{1}{2}[(\mathbf{x}-\boldsymbol{\mu})^T (\mathbf{\Sigma})^{-1} (\mathbf{x}-\boldsymbol{\mu})]} \tag{19}$$

where $\mathbf{x}$ is a N-dimensional vector. We still have the leading constant, $\frac{1}{\sqrt{(2\pi)^N}}$, to ensure that the equation integrates to unity. In the multidimensional problem the variance, $\boldsymbol{\sigma}^2$, is given by the diagonal of the covariance matrix $\mathbf{\Sigma}$. Therefore, using (19), we can write each of the terms in Eq. (17) in the multidimensional problem as

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{N/2} |\mathbf{B}|^{1/2}} e^{-\frac{1}{2}\left[(\mathbf{x}-\mathbf{x}^b)^T (\mathbf{B})^{-1} (\mathbf{x}-\mathbf{x}^b)\right]} \tag{20}$$

$$p(\mathbf{y}^o|\mathbf{x}) = \frac{1}{(2\pi)^{L/2} |\mathbf{R}|^{1/2}} e^{-\frac{1}{2}[(\mathbf{y}^o-\mathrm{H}\mathbf{x})^T (\mathbf{R})^{-1} (\mathbf{y}^o-\mathrm{H}\mathbf{x})]} \tag{21}$$

In Eq. (20) and (21) the general covariance matrix $\mathbf{\Sigma}$ is replaced by the background error covariance matrix of the model $\mathbf{B}$ (N×N matrix) and the observation error covariance matrix $\mathbf{R}$ (L×L matrix) respectively. Furthermore, $\boldsymbol{\mu}$, the mean or known state, turns into the background forecast $\mathbf{x}^b$ and the observations $\mathbf{y}^o$ respectively. Finally, the observation operator H is also introduced here, which operates on the model fields to obtain a model equivalent value in observation space. Therefore, the conditional probability in equation 17 follows as

$$p(\mathbf{x}|\mathbf{y}^o) \propto p(\mathbf{x}) p(\mathbf{y}^o|\mathbf{x}) = \frac{e^{-\frac{1}{2}\left[(\mathbf{x}-\mathbf{x}^b)^T (\mathbf{B})^{-1} (\mathbf{x}-\mathbf{x}^b) + (\mathbf{y}^o-\mathrm{H}\mathbf{x})^T (\mathbf{R})^{-1} (\mathbf{y}^o-\mathrm{H}\mathbf{x})\right]}}{2\pi^{\frac{N}{2}} |\mathbf{B}|^{\frac{1}{2}} 2\pi^{\frac{L}{2}} |\mathbf{R}|^{\frac{1}{2}}} \tag{22}$$

These Bayesian principles will form the basis of the data assimilation methods discussed in the next few sections. For example, in the maximum likelihood scenario,

the likelihood is maximized when the exponential in Eq. (22) is minimized. Conveniently, the term in the exponential becomes our cost function for variational methods described in the next section.

1.5.3 Variational Methods (Maximum Likelihood)

Variational methods, or maximum likelihood methods, seek *to find the most likely state of* **x** *given the two independent measurements* $\mathbf{x}^b$ *and* $\mathbf{y}^o$, *which is the state that maximizes the joint probability*, and it does so by minimizing a cost function which can be defined using the Bayesian principles from the previous section. This section discusses three- and four-dimensional variational methods. Each subsection will begin with the derivation of the analysis equation for each method and then will conclude with a discussion of previous work as it relates to radial wind assimilation using variational methods.

1.5.3.1 Three-Dimensional Variational (3DVar) Method

Three-dimensional variational (3DVar; e.g., Lorenc 1986) data assimilation, as its name suggests, is a 3D data assimilation method meaning that there is no time component to the assimilation. For example, all observations are assumed to be valid at the analysis time which are assimilated simultaneously in combination with the prior state valid at the analysis time to estimate the best state at the analysis time. The joint probability of the likelihood of the observations and the probability of the state before new observations [i.e., the product of the right-hand side of Eq. (22)] is the product of two Gaussian probabilities. Maximizing the likelihood attains the most likely state **x** which is equivalent to minimizing the term in the exponential of Eq. (22); therefore, a

cost function can be defined as (following the notation of Ide et al. (1997) wherever possible from here on)

$$\begin{aligned} J(\mathbf{x}) &= J_B + J_O \\ &= \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{y}^o - H(\mathbf{x}))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}^o - H(\mathbf{x})) \end{aligned} \tag{23}$$

Equation (23) is the cost function for the 3DVar algorithm. In general, the cost function, $J(\mathbf{x})$, is the sum of at least two terms (sometimes more). The first term, J_B, measures the departure of the analysis or control variable, $\mathbf{x}$, from the background, $\mathbf{x}^b$, and is weighted by the inverse of the background error covariance matrix, $\mathbf{B}$ (referred to from here on as the BEC), which is typically chosen to be static (i.e., never changes from day-to-day or season-to-season since, as a choice, only one BEC is generated and used year round) and typically estimated from climatology (Bannister 2008). The BEC is estimated since it is too large to explicitly compute and store in computer memory[6]. One method for estimating the BEC statistics is the NMC method (Parrish and Derber 1992). The review paper Bannister (2008) provides additional information about the NMC method and other methods for estimating the BEC. The NMC method uses the difference between two forecasts valid at the same time but at different leads (e.g., a 48-hour and 24-hour forecast valid at the same time) and averaging those over many cases to find the expected value of the forecast error. The BEC is a crucial part of variational methods because it is an estimate of the error correlations and cross-correlations of model analysis variables. It not only provides information about the

[6] A rough estimation of the size of the static B given GFSv16 at C384 would require a matrix on the order of 10^9 x 10^9 and assuming only single precision that would require about 4,000 PB (Petabytes) of memory to store the matrix. Inversion of a k x k matrix has a time complexity of $O(k^3)$ which means that if a 100 x 100 matrix would take 1 second to invert, then a 10^9 x 10^9 matrix would take 10^{18} seconds to invert or roughly 31.7billion years.

uncertainty of the first guess and how an observation should inform the analysis at the observation location, but also how an observation should impact the analysis at other locations as well as how an observation should affect other relatable variables. Generating new BECs is beyond the scope of this work but is an important aspect of the assimilation system that will be considered in the future. More discussion on the choice of method for estimating BEC will be discussed in future work (section 5.3).

The second term, J_O, measures the departure of the analysis at the observation location using the non-linear interpolation operator, *H*, from the observations, $\mathbf{y}^o$, and is weighted by the inverse of the observation error covariance matrix, **R**. The observation error covariance matrix is also important since it provides information about the uncertainty of the observations.

The minimum of the cost function $J(\mathbf{x})$ is attained when

$$\nabla_x J(\mathbf{x}) = 0$$

$$\nabla_x^2 J(\mathbf{x}) = \text{positive definite}$$

To find an exact solution, we can expand the second term, J_O, in Eq. (23) under the assumption that analysis is close to the truth and is also close to the observations. We can then linearize *H* (**H**) around the background such that our innovation vector becomes

$$\mathbf{d} = \mathbf{y}^o - H(\mathbf{x}) = \mathbf{y}^o - H(\mathbf{x}^b + (\mathbf{x} - \mathbf{x}^b)) \approx (\mathbf{y}^o - H(\mathbf{x}^b)) - \mathbf{H}(\mathbf{x} - \mathbf{x}^b) \tag{24}$$

Substituting Eq. (24) into (23) we get

$$\begin{aligned} 2J(\mathbf{x}) = {} & (\mathbf{x} - \mathbf{x}^b)^{\mathrm{T}} \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b) \\ & + [(\mathbf{y}^o - H(\mathbf{x}^b)) - \mathbf{H}(\mathbf{x} - \mathbf{x}^b)]^{\mathrm{T}} \mathbf{R}^{-1} [(\mathbf{y}^o - H(\mathbf{x}^b)) - \mathbf{H}(\mathbf{x} - \mathbf{x}^b)] \end{aligned} \tag{25}$$

and then expanding the second term further we get

$$2J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b)$$

$$+ (\mathbf{x} - \mathbf{x}^b)^T \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} (\mathbf{x} - \mathbf{x}^b) - \left(\mathbf{y}^o - H(\mathbf{x}^b)\right)^T \mathbf{R}^{-1} \mathbf{H} (\mathbf{x} - \mathbf{x}^b) \quad (26)$$

$$- (\mathbf{x} - \mathbf{x}^b)^T \mathbf{H}^T \mathbf{R}^{-1} \left(\mathbf{y}^o - H(\mathbf{x}^b)\right) + \left(\mathbf{y}^o - H(\mathbf{x}^b)\right)^T \mathbf{R}^{-1} \left(\mathbf{y}^o - H(\mathbf{x}^b)\right)$$

The gradient of the cost function can then be written as

$$\nabla_x J(\mathbf{x}) = \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b) + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} (\mathbf{x} - \mathbf{x}^b) - \mathbf{H}^T \mathbf{R}^{-1} \left(\mathbf{y}^o - H(\mathbf{x}^b)\right) \quad (27)$$

and setting $\nabla_x J(\mathbf{x}) = 0$ ensures that the solution is a minimum then we have the solution

$$\mathbf{x} = \mathbf{x}^b + (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} \left(\mathbf{y}^o - H(\mathbf{x}^b)\right) \quad (28)$$

This can also be written in incremental form as

$$\delta\mathbf{x} = (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} \delta\mathbf{y}^o$$

analysis increment, $\delta\mathbf{x}$, and the innovation vector, $\delta\mathbf{y}^o$, are defined as

$$\delta\mathbf{x} = \mathbf{x} - \mathbf{x}^b$$

$$\delta\mathbf{y}^o = \mathbf{y}^o - H(\mathbf{x}^b)$$

For easier implementation and simplification of the cost function [Eq. (23)] can be written and solved in the incremental form (Courtier et al. 1994) so that its 1st and 2nd derivatives are written in terms of $\delta\mathbf{x}$ and is minimized with respect to $\delta\mathbf{x}$. Thus, the incremental form of the cost function [Eq. (23)] is

$$J(\delta\mathbf{x}) = \frac{1}{2} (\delta\mathbf{x})^T \mathbf{B}^{-1} (\delta\mathbf{x}) + \frac{1}{2} (\delta\mathbf{y}^o - \mathbf{H}\delta\mathbf{x})^T \mathbf{R}^{-1} (\delta\mathbf{y}^o - \mathbf{H}\delta\mathbf{x}) \quad (29)$$

The first guess to minimize this equation (i.e., $\nabla J(\delta\mathbf{x}) = 0$) is to prescribe the analysis as the background ($\mathbf{x} = \mathbf{x}^b$). The analysis $\mathbf{x}$ using 3DVar can then be found by an iterative numerical solver such as the conjugate gradient method (Derber and Rosati

1989; Navon and Legler 1987); however, it is also typical to first precondition the cost function by using the BEC to improve convergence (e.g., Kleist et al. 2009a). For more detailed documentation on this the reader is referred to p. 17 of the GSI's Advanced User's Guide (available online: https://dtcenter.ucar.edu/com-GSI/users/docs/users_guide/AdvancedGSIUserGuide_v3.4.0.0.pdf)

In the 3DVar algorithm, observations are collected within the full window and assumed to be valid at the nominal analysis time which is usually in the center of that window. This can potentially lead to inaccurate innovations. This leads to an important variation of the 3DVar algorithm: 3DVar FGAT which stands for first guess at appropriate time. This specific application of the 3DVar algorithm utilizes multiple background fields and splits the observations into time bins according to their valid times so that the corresponding background at the appropriate time is used for temporal interpolation in the $H(\mathbf{x})$ calculation (e.g., Bannister 2008). The innovation vector, however, is constant in time. The 3DVar FGAT algorithm is distinct from the 4DVar algorithm discussed next.

For more information about 3DVar, 3DVar FGAT, other data assimilation methods, the reader is directed to the review paper Bannister (2017) which summarizes many different data assimilation methods.

1.5.3.2 The Four-Dimensional Variational (4DVar) Method

The four-dimensional variational method (4DVar; e.g., Thépaut et al. 1993a; Thépaut et al. 1993b) is a powerful and dynamic extension of 3DVar which incorporates a linearized model to include a time component to the assimilation, relaxing the temporal constraints of 3DVar. While 3DVar produces the best estimate

of the state for a discreet time, 4DVar does the same but for a given trajectory over which we have observations. The 4DVar BEC is typically chosen to be the same BEC used in the 3DVar method; however, the 4DVar method is able to extract information from the model dynamics and thus be able to implicitly include flow-dependency (i.e., errors of the day) within the assimilation (Thépaut et al. 1993a) . Incremental 4DVar involves solving for the optimal solution, $\delta\mathbf{x}_0$, at the beginning of a time window where k=0, obtained by minimizing the cost function

$$\mathrm{J}(\mathbf{x}_0) = \frac{1}{2}(\delta\mathbf{x}_0)^{\mathrm{T}}\mathbf{B}^{-1}(\delta\mathbf{x}_0) + \frac{1}{2}\sum_{k=0}^{K}\left(\delta\mathbf{y}_k^o - \mathbf{H}\delta\mathbf{x}_k\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(\delta\mathbf{y}_k^o - \mathbf{H}\delta\mathbf{x}_k\right) \tag{30}$$

Equation (30) is the incremental form of the 4DVar cost function and follows from the 3DVar cost function in incremental form [Eq. (23)]. The 4DVar cost function, $\mathrm{J}(\mathbf{x}_0)$, is the sum of the same two (sometimes more) terms such as in the 3DVar equation, but the optimal solution is solved for at time k=0 and the observations are used at their valid times. Therefore, $\mathbf{x}_0$ is the forecast state at the beginning of the time window, K in the total size of the time window, k is the time window index. The 4DVar method is distinct from the 3DVar FGAT method because in 4DVar, the innovation is propagated by the linearized model to solve for the 4D state. We can define the analysis increment at time k=0, the analysis increment at time k, and the innovation at time k respectfully as

$$\delta\mathbf{x}_0 = \mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}}$$

$$\delta\mathbf{x}_k = \mathbf{M}_k\delta\mathbf{x}_0$$

$$\delta\mathbf{y}_k^o = \mathbf{y}_{\mathrm{k}}^{\mathrm{o}} - \mathbf{H}\mathbf{M}_{\mathrm{k}}\mathbf{x}_0^{\mathrm{b}}$$

where $\mathbf{M}$ is the linearized model.

Similarly to 3DVar, the 4DVar cost function is minimized when the first derivative is zero and the second derivative is positive definite using a minimization technique such as the conjugate gradient method (Derber and Rosati 1989; Navon and Legler 1987) .

1.5.3.3 Discussion of Variational Methods and Their Use for Radial Wind DA

Variational methods such as 3DVar and 4DVar are some of the major building blocks for the more advanced data assimilation algorithms used today for operational NWP. Extensive research has been done to assimilate radial winds into NWP models using the 3DVar and 4DVar methods with varying levels of success. While this thesis does not solely focus on convective-scale radar DA, the vast majority, if not all, previous literature on radar DA is in the context of improving convective-scale forecasts. For example, Sun and Crook (1997) developed a 4DVar adjoint dynamic retrieval technique and demonstrated that the 3D wind and thermodynamic fields could be obtained from dual- and single-Doppler observations. Sun and Crook (1997, 1998) showed that this variational Doppler radar analysis system was able to retrieve the three-dimensional wind, thermodynamic, and microphysical fields from a Florida air mass storm within reasonable approximation to a dual-Doppler analysis. This system was eventually converted from a single-Doppler parameter retrieval research tool into a real-time low-level wind and temperature analysis system (Sun and Crook 2001). Sun and Crook (2001) tested the feasibility of initializing an NWP cloud model with single Doppler observations and predicting the evolution of the thunderstorms and verified against the dual Doppler analyses with success. Sun (2005a) developed a variational Doppler radar analysis system that produce real time boundary layer wind

and temperature analyses using WSR-88D radar data with success. Sun and Wang (2013) showed that during a squall line event, 4DVar produced higher precipitation skill scores compared to 3DVar.

One advantage of variational methods is that constraints can be imposed directly into the cost function (Xie et al. 2002). For example, Ge et al. (2012) incorporated a diagnostic pressure equation into a storm-scale 3DVar system in the form of a weak constraint [e.g., adding another term (J_C) term in the cost function of Eq. (23)] in addition to a mass continuity equation constraint. They found that both constraints yielded the best results. These constraints are meant to maintain dynamically consistent optimized solutions in the data assimilation system, which has been shown to greatly improve the analysis and subsequent forecasts of deep convective storms.

Variational methods are powerful, robust, and well-developed methods for determining an analysis via Bayesian principles. Even today, as much more sophisticated methods exist (discussed in a later section, e.g., hybrid methods), 3DVar is still used in research applications, especially in early system testing and development. This is because 3DVar has a relatively low computational expense compared to other methods. Therefore, while 3DVar for most operational NWP applications is outdated, there still resides some useful applications of 3DVar and is still an important aspect of NCEP's data assimilation system.

On the other hand, the use of a pure 4DVar system, while it has been shown to provide improved analyses when compared to 3DVar, would not make the best use of resources as the preferred hybrid ensemble-variational methods (discussed later) are

computationally less resource intensive, do not require the development and maintenance of a linearized dynamical model, and still produce better analyses. Before discussing hybrid methods (or a blend of variational and ensemble methods), it is advantageous to first cover ensemble methods. Therefore, in the following section, the discussion will begin with the minimum variance approach of the Kalman Filter which will eventually lead to its ensemble approximation.

1.5.4 Sequential Methods (Minimum Variance)

Minimum variance seeks *to find the most optimal weight matrix that minimizes the analysis error covariance matrix.* One might think of this in the simplest terms as least squares linear regression. In the least squares estimation, all the errors are the same so one would fit a line to the observations that minimizes the sum of the square of the residuals (or distances of each observation to the line). Generally, not all observations will have the same error variances and thus the observations with lower error variances should be given more weight. Thus, in minimum variance approach of the multiple linear regression problem we attempt to find the optimal weight matrix which is the *best linear unbiased estimator,* and we can do this using the Gaussian assumptions that we made using a Bayesian approach.

This section first discusses a practical implementation of the original Kalman filter: optimal interpolation. This will then lead into Kalman filtering and simplifications of Kalman filtering that are more commonly found in weather forecasting. Each section will begin with the derivation of the analysis equations for the various methods and then this section will conclude with a discussion of previous work as it relates to radial wind assimilation using minimum variance methods.

1.5.4.1 Kalman Filters (KF)

The Kalman Filter (KF) was first formulated by Kalman (1960) and is formally very similar to optimal interpolation (OI; Eliassen 1954; Gandin 1963). The goal of OI and the KF is to obtain the best possible estimate of the state, $\mathbf{x}$, by combining the background equal to the true value the state, $\mathbf{x}^t$, plus the background error, $\boldsymbol{\varepsilon}^b$

$$\mathbf{x}^b = \mathbf{x}^t + \boldsymbol{\varepsilon}^b \tag{31}$$

with observations equal to the true value of $\mathbf{x}$ interpolated to observation space by the observation operator, H, plus the observation error, $\boldsymbol{\varepsilon}^o$

$$\mathbf{y}^o = H(\mathbf{x}^t) + \boldsymbol{\varepsilon}^o \tag{32}$$

as a linear combination such that the analysis is equal to the background plus the innovation weighted by the optimal weight matrix

$$\mathbf{x} = \mathbf{x}^b + \mathbf{W}(\mathbf{y}^o - H(\mathbf{x}^b)). \tag{33}$$

The innovation vector, $\mathbf{d}$, is the difference between the observation and the background mapped to the observational point via the nonlinear forward operator H and is defined as

$$\mathbf{d} = \mathbf{y}^o - H(\mathbf{x}^b) \tag{34}$$

The weight matrix, $\mathbf{W}$, is usually called the *gain matrix* $\mathbf{K}$, especially in the Kalman filter literature. Now, we attempt to find the optimal weight matrix which is the best linear unbiased estimator using a Bayesian approach.

Recall in Eq. (22) that the posterior is simply the product of two Gaussian probability distribution functions: the prior times the likelihood of the observations. The analysis in Eq. (33) is also simply the true state, $\mathbf{x}^t$, plus the analysis error, $\boldsymbol{\varepsilon}$

$$\mathbf{x} = \mathbf{x}^t + \boldsymbol{\varepsilon} \tag{35}$$

The analysis error variance is what we want to minimize and is *the essence of Kalman Filtering*.

Now we define the error covariance matrices for the analysis, $\mathbf{P}^a$, the background, $\mathbf{B}$, and the observations, $\mathbf{R}$, respectively by multiplying the respective error vectors by their transpose and then averaging over many cases (or over an ensemble) to find the expected value denoted by the "$E[\quad]$"

$$\mathbf{P}^a = E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T] = E[(\mathbf{x} - \mathbf{x}^t)(\mathbf{x} - \mathbf{x}^t)^T] \tag{36}$$

$$\mathbf{P}^b = \mathbf{B} = E\left[\boldsymbol{\varepsilon}^b \boldsymbol{\varepsilon}^{bT}\right] \tag{37}$$

$$\mathbf{P}^o = \mathbf{R} = E\left[\boldsymbol{\varepsilon}^o \boldsymbol{\varepsilon}^{oT}\right] \tag{38}$$

The analysis error, $\boldsymbol{\varepsilon}$, is the difference between the analysis state and the true state. Then, plugging in Eq. (33), we can write the analyses variance as the sum of the background error variance and the optimally weighted innovation vector

$$\begin{aligned} \boldsymbol{\varepsilon} &= \mathbf{x} - \mathbf{x}^t \\ &= (\mathbf{x}^b - \mathbf{x}^t) + \mathbf{W}\mathbf{d} \\ &= \boldsymbol{\varepsilon}^b + \mathbf{W}(\boldsymbol{\varepsilon}^o - \mathbf{H}\boldsymbol{\varepsilon}^b) \end{aligned} \tag{39}$$

Therefore, the analysis error covariance, $\mathbf{P}^a$, can be written as

$$\mathbf{P}^a = E\left[[\boldsymbol{\varepsilon}^b + \mathbf{W}(\boldsymbol{\varepsilon}^o - \mathbf{H}\boldsymbol{\varepsilon}^b)][\boldsymbol{\varepsilon}^b + \mathbf{W}(\boldsymbol{\varepsilon}^o - \mathbf{H}\boldsymbol{\varepsilon}^b)]^T\right]$$

$$= \mathbf{B} - \mathbf{B}\mathbf{H}^T\mathbf{W}^T - \mathbf{W}\mathbf{H}\mathbf{B} + \mathbf{W}[\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T]\mathbf{W}^T.$$

The optimal solution that minimizes the total variance of the analysis (i.e., $E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T]$) can be found by differentiating the total analysis variance (by taking the trace of $\mathbf{P}^a$), with respect to the weight matrix $\mathbf{W}$ and then setting the derivative to zero

$$Tr(\mathbf{P}^a) = Tr(\mathbf{B}) - Tr(\mathbf{B}\mathbf{H}^T\mathbf{W}^T) - Tr(\mathbf{W}\mathbf{H}\mathbf{B}) + Tr(\mathbf{W}[\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T]\mathbf{W}^T)$$

$$= -2Tr(\mathbf{B}\mathbf{H}^T\mathbf{W}^T) + Tr(\mathbf{W}[\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T]\mathbf{W}^T) \quad (40)$$

$$= \mathbf{W}\,[\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R}] - \mathbf{B}\mathbf{H}^{\mathrm{T}} = 0.$$

From Eq. (40) we can obtain the *optimal weight matrix*

$$\mathbf{W} = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}}+\mathbf{R})^{-1} = \mathbf{K}. \quad (41)$$

Finally, substituting the Kalman Gain into Eq. (33) we arrive at the KF analysis equation

$$\mathbf{x} = \mathbf{x}^b + \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}}+\mathbf{R})^{-1}\mathbf{d} \quad (42)$$

The KF algorithm determines the analysis, $\mathbf{x}$, by weighting the innovation vector, $\mathbf{d}$, by the optimal weight matrix $\mathbf{W}$. Note that in the Kalman Filter, the error covariance is advanced by using the model itself instead of using a static BEC such as is done in OI. Notice that the result for the analysis equation in 3DVar Eq. (28) looks very similar to Eq. (42) obtained here for KF/OI. It has been shown that 3DVar and KF/OI are equivalent (Kalnay 2002 p. 171); however, they use two distinct methods to solve for the analysis. There are many variations of the Kalman Filter; therefore, the remainder of this section will discuss only a few of the variations that are relevant to this thesis.

1.5.4.2 Extended Kalman Filter (EKF)

The Extended Kalman Filter (EKF; Einicke and White 1999) is the nonlinear counterpart of the Kalman Filter, and it is often regarded as the "gold standard" of data assimilation methods. In the EKF, the forecast error covariance is obtained by linearizing the model about the nonlinear trajectory of the model. It has been shown that under certain conditions, the solution for the four-dimensional variational (4DVar)

and EKF (i.e., the nonlinear counterpart of the KF; Einicke and White 1999; Kalnay 2002 p. 177-180) are equivalent. The equivalence of the EKF and 4DVar hold, as noted by Daley (1991 p. 384) when the forecast model is linear and perfect, the observation operator is linear, and the BEC at t_0 is the BEC used in the 4DVar equation. This equivalence between EKF and 4DVar has also been shown by Lorenc (1986). The problem with the EKF is that it is too expensive since the linearized model has a rank *n* (degrees of freedom in the model) and updating the error covariance is equivalent to performing O(*n*) model integrations (Kalnay 2002; p. 180). It is for this reason that many simplifications have been made.

1.5.4.3 Ensemble Kalman Filter (EnKF)

The EnKF method (Evensen 1994; Houtekamer and Mitchell 1998) is an example of a simplified, ensemble representation of the KF method which relies on an ensemble to approximate the BEC from an ensemble or a collection of individual forecasts. Because the BECs are estimated from an ensemble forecast and changes from day-to-day, the BEC from the EnKF will contain the errors of the day often referred to as *flow-dependent errors*. The flow dependent errors provided explicitly by the ensemble is distinctly different and stronger than the flow dependency induced by the model in 4DVar.

When formulating an ensemble data analysis system (such as the EnKF), an ensemble forecast framework is first developed with *M* members.

Ensemble $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_M)$

Mean $\bar{\mathbf{x}} = \frac{1}{M}\sum_{m=1}^{M} \mathbf{x}_m$

Perturbations $\quad \widehat{\mathbf{X}} = (\mathbf{x}_1 - \bar{\mathbf{x}}, \ldots, \mathbf{x}_M - \bar{\mathbf{x}})$

Covariance $\quad \widehat{\mathbf{B}} = \frac{1}{M-1}\widehat{\mathbf{X}}\widehat{\mathbf{X}}^{\mathrm{T}}$

From the ensemble, the covariance matrix can be estimated using the ensemble mean and the ensemble perturbations. One must also perturb the observations, $\widehat{\mathbf{Y}}$ being the vector of perturbed observations, with random noise N(0, $\mathbf{R}^o$). In the perturbed observation EnKF, this is necessary for the ensemble to main proper spread**.** The Kalman gain matrix can then be estimated similarly as in KF/OI by the following

$$\mathbf{K}^{EnKF} = \frac{\frac{1}{M-1}\widehat{\mathbf{X}}\widehat{\mathbf{Y}}^{\mathrm{T}}}{\frac{1}{M-1}\widehat{\mathbf{Y}}\widehat{\mathbf{Y}}^{\mathrm{T}} + \mathbf{R}^o} \approx \mathbf{K}^{KF} = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R}^o)^{-1} \tag{43}$$

The equivalence holds if the observation operator (**H**) is linear. Likewise, the analysis for each member can be estimated using a similar formulation as was done for KF with a slight modification. Recall that for KF the analysis can be computed as

$$\mathbf{x} = \mathbf{x}^b + \mathbf{K}^{KF}\mathbf{d}$$

then the analysis for the EnKF is computed as

$$\mathbf{x}_m = \mathbf{x}_m^b + \mathbf{K}^{EnKF}\left(\mathbf{y}_m^o - \mathbf{H}\mathbf{x}_m^b\right) \tag{44}$$

where $\mathbf{y}_m^o = \mathbf{y}^o + \mathbf{e}_m^o$ and $\mathbf{e}_m^o$ is random error drawn from $\mathrm{N}(0,\mathbf{R}^o)$, a normal distribution of mean zero and a standard deviation of $\mathbf{R}^o$. The analysis error, $\mathbf{P}^a$, is given by

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B} \tag{45}$$

This formulation of the EnKF is known as Perturbed Observation (PO) EnKF and referred to as *stochastic* EnKF (Evensen 1994; Houtekamer and Mitchell 1998). In the PO formulation, Burgers et al. (1998) showed that it is necessary to add noise to the

observations in order to prevent the analysis covariance from being underestimated leading to filter divergence.

Alternatives to the stochastic EnKF algorithm have been developed which do not require perturbing observations. The alternative is a *deterministic* formulation of the EnKF which are a class of EnKF known as ensemble square root filters (EnSRF; e.g., Whitaker and Hamill 2002). Whitaker and Hamill (2002) have shown that the EnSRF is more accurate than the stochastic EnKF for the same ensemble size. In the EnSRF formulation of Whitaker and Hamill (2002), which removes the necessity to perturb the observations, the analysis covariance is given by

$$\mathbf{P}^a = \left(\mathbf{I} - \widetilde{\mathbf{K}}\mathbf{H}\right)\mathbf{B}\left(\mathbf{I} - \widetilde{\mathbf{K}}\mathbf{H}\right)^{\mathrm{T}} \tag{46}$$

and, for a single observation, has the solution of

$$\widetilde{\mathbf{K}} = \left[1 + \sqrt{\frac{\mathbf{R}^o}{\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R}^o}}\right]^{-1} \mathbf{K} \tag{47}$$

so, the corresponding analysis perturbation ensemble update is

$$\mathbf{Z}^a = \left(\mathbf{I} - \widetilde{\mathbf{K}}\mathbf{H}\right)\mathbf{Z}^b \tag{48}$$

where $\mathbf{Z}$ is the *square root* of the respective covariance matrix. Thus, the analysis and forecast error covariance matrices can be represented as

$$\begin{aligned} \mathbf{P}^a &= \mathbf{Z}^a \mathbf{Z}^{a\,\mathrm{T}} \\ \mathbf{B} &= \mathbf{Z}^b \mathbf{Z}^{b\,\mathrm{T}} \end{aligned} \tag{49}$$

respectively.

There have been many different implementations of deterministic EnKF which are all part of a sub-class of EnKF known as EnSRF and are all formally equivalent

and differ only in how they are implemented (Tippett et al. 2003). Within the classification of EnSRF, there is the Local Ensemble Transform Kalman Filter (LETKF; Hunt et al. 2007) among others. The LETKF, similarly to the Local Ensemble Kalman Filter (LEKF; Ott et al. 2002; Ott et al. 2004), is a localized version of the ensemble transform Kalman filter (ETKF; Bishop et al. 2001). The ETKF seeks to *transform* the background ensemble perturbations

$$\widehat{\mathbf{X}}^{b} = (\mathbf{x}_1{}^{b} - \bar{\mathbf{x}}^{b}, \ldots, \mathbf{x}_M{}^{b} - \bar{\mathbf{x}}^{b}) \tag{50}$$

into analysis perturbations, $\widehat{\mathbf{X}}^{a}$, by multiplying by a transformation matrix, $\mathbf{T}$

$$\widehat{\mathbf{X}}^{a} = \widehat{\mathbf{X}}^{b}\mathbf{T} \tag{51}$$

The transformation matrix is derived in (Bishop et al. 2001; Wang and Bishop 2003; Wang et al. 2004) and is given by

$$\mathbf{T} = \mathbf{C}(\mathbf{\Gamma} + \mathbf{I})^{\frac{1}{2}} \tag{52}$$

where Bishop et al. (2001) shows that $\mathbf{C}$ and $\mathbf{\Gamma}$ are the orthonormal eigenvector matrix and diagonal matrix containing the eigenvalues of $\widehat{\mathbf{X}}^{b\,\mathrm{T}}\widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{H}}\widehat{\mathbf{X}}^{b}$ respectively (where $\widetilde{\mathbf{H}}$ is the normalized observation operator) and $\mathbf{I}$ is the identity matrix. In other words, the ETKF works in the subspace of the ensemble perturbations and then seeks to minimize error covariance of the amplifying modes or those modes which are the most unstable and lead to the greatest forecast errors. The ETKF was later made more efficient via the LETKF (Harlim and Hunt 2005; Hunt et al. 2007) and is formally equivalent to performing the ETKF on localized domains by only assimilating the observations within a given radius. Each local analysis involves much less data than the full global analysis to compute. The global ETKF requires the ensemble to contain the global

uncertainty of the model, whereas the LETKF and LEKF only require the ensemble to contain the uncertainty in the local region, this localization effectively lowers the dimensionality of the data assimilation problem. As noted by Harlim and Hunt (2005), each of the local analyses can be computed independently and in parallel thereby making this scheme much more efficient.

The relative advantages and limitations between EnKF and 4DVar are thoroughly discussed in (Gustafsson 2007; Kalnay et al. 2007a; Kalnay et al. 2007b; Lorenc 2003). Both EnKF and 4DVar methods try to incorporate some degree of flow dependence. The EnKF does so explicitly via the ensemble perturbations while 4DVar is able to include the flow-dependent errors implicitly through the model propagation (Thépaut et al. 1993a; Thépaut et al. 1993b). The EnKF is much simpler to implement, maintain, and run and often has comparable skill to 4DVar (e.g., Caya et al. 2005). 4DVar struggles under highly non-linear processes such as at the convective scales while the main disadvantage of a pure EnKF method is its rank deficiency of the BEC which is usually measured by a relatively small ensemble size compared to the number of degrees of freedom in the model. The effects of rank deficiency such as spurious long distant correlations can be reduced by localization and by ensemble inflation which otherwise would lead to filter divergence (or the ensemble having too low of variance thus being overconfident and completely ignoring the observations).

Due to the rank-deficiency of the error covariance in the EnKF, it is crucial to apply localization techniques to remove long range spurious correlations. Covariance localization is performed by using a Schur (element wise) product of the ensemble-based covariance estimates weighted by a distant-dependent correlation function that

varies from 1.0 at the observation location to 0.0 at some pre-specified distance (Hamill et al. 2001; Houtekamer and Mitchell 2001). By applying localization, the effective ensemble size is also increased (Hamill et al. 2001; Oke et al. 2007; Ott et al. 2004). In addition to covariance localization, there is also domain localization such as through the use of a local ensemble Kalman filter (Ott et al. 2004).

1.5.4.4 Discussion of Ensemble Methods and Their Use for Radial Wind DA

Ensemble based methods have also been successfully utilized in radar DA. Once again, we find that previous work has almost exclusively focused on convective-scale data assimilation with an analysis grid of 1-2km (Aksoy et al. 2009; Caya et al. 2005; Dawson et al. 2012; Dong et al. 2011; Dowell and Wicker 2009; Dowell et al. 2011; Dowell et al. 2004; Snyder and Zhang 2003; Tong and Xue 2005). For example, Snyder and Zhang (2003) made one of the first attempts to assimilate Doppler-radar observations using an EnKF DA method and was performed within the context of convective-scale NWP. In this study, they tested the EnKF using simulated observations of radial velocity from an isolated supercell thunderstorm in a perfect model scenario. The observations were simulated at 5-min intervals at each grid point where the rainwater met a particular threshold with random noise added to simulate imperfect observations. With 5-min EnKF analysis updates, the simulation of the thunderstorm was able to closely resemble that of the reference state after the assimilation of only a few volume scans.

Using the same setup from Snyder and Zhang (2003), Caya et al. (2005) compared 4DVar against the EnKF in the context of radar DA. They found that 4DVar and EnKF to produce comparatively accurate analyses, although dissimilarities may be

noted depending on the number of assimilation cycles and the assimilation window length. 4DVar generally performs better than the EnKF within the first few cycles with then EnKF outperforming 4DVar at later cycles.

There is a wealth of other studies, aside from those already listed, using an ensemble-based methods for radar DA (e.g., Jones et al. 2015; Snook et al. 2012; Tong et al. 2020; Tong and Xue 2008). The use of ensemble methods for radar DA have become relatively mature for convection-allowing applications.

Both variational and ensemble methods have their respective strengths and weaknesses; therefore, hybrid methods (next section) have been suggested by Hamill and Snyder (2000) as they overcome the challenges that exist in both the ensemble and variational methods.

1.5.5 Hybrid methods (3DEnVar, 4DEnVar, & hybrid gain)

To overcome the challenges of ensemble methods and the traditional variational methods such as 3DVar and 4DVar, Hamill and Snyder (2000) proposed the use of hybrid methods. Hybrid methods are a class of algorithms that solve the data assimilation problem by blending a static BEC with an ensemble BEC and have been shown to outperform variational methods and have generally comparative performance to that of ensemble-based methods.

Hybrid methods are featured in this dissertation. Hybrid 3DEnVar was used in Chapter 2, hybrid 4DEnVar in Chapter 3, and hybrid gain in Chapter 4. The choice of which method to use was dependent entirely upon what was used within the current system at the time.

1.5.5.1 Hybrid Three-Dimensional Ensemble Variational (3DEnVar)

The first hybrid method discussed here is the hybrid three-dimensional ensemble variational (3DEnVar) method. Hybrid 3DEnVar is a variational method which determines the analysis (i.e., the most likely state given the prior forecast and observations) by the direct minimization of a cost function. The hybrid 3DEnVar cost function may be written in incremental form as (see Wang (2010)):

$$J(\delta\mathbf{x_f}, \boldsymbol{\alpha}) =$$

$$\beta_f \frac{1}{2}(\delta\mathbf{x_f})^T\mathbf{B_f}^{-1}(\delta\mathbf{x_f}) + \beta_e \frac{1}{2}\sum_{n=1}^{N}(\boldsymbol{\alpha}^n)^T\mathbf{L}^{-1}(\boldsymbol{\alpha}^n) + \frac{1}{2}(\mathbf{d} - \mathbf{H}\delta\mathbf{x_t})^T\mathbf{R}^{-1}(\mathbf{d} - \mathbf{H}\delta\mathbf{x_t}) \quad (53)$$

The incremental form is used for simplicity and the analysis increments, and the innovation are defined as follows

$$\delta\mathbf{x_f} = \mathbf{x} - \mathbf{x}^b$$

$$\delta\mathbf{x_t} = \delta\mathbf{x_f} + \sum_{n=1}^{N}[\boldsymbol{\alpha}^n \circ \delta\mathbf{x}_e^n]$$

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

The analysis increment derived from the static error covariance, $\delta\mathbf{x_f}$, measures the departure of the optimal state or analysis, $\mathbf{x}$, from the previous forecast or background, $\mathbf{x}^b$, and is weighted by the inverse of the static (fixed) background error covariance matrix, $\mathbf{B_f}$. The innovation or observation residual, $\mathbf{d}$, measures the departure of the observations, $\mathbf{y}^o$, and the background state by means of the non-linear, H and linear observation operators $\mathbf{H}$, and is weighted by the inverse of the observation error covariance, $\mathbf{R}$. The flow-dependent, ensemble estimated covariances are included in the cost function through the extended control variable method (Lorenc 2003; Wang

2010). The extended control variable, $\boldsymbol{\alpha}^n$, represents the ensemble weights locally at each grid point, which are used to construct the analysis increments through linear combinations of the ensemble perturbations, $\mathbf{x}_e^n$ using a Schur product where N is the number of ensemble members. The matrix, **L**, denotes the error covariance for the alpha control variable and is specified to be of unit amplitude through which covariance localization can be imposed. The total analysis increment, $\delta\mathbf{x_t}$, is a linear combination of the analysis increment derived from the static error covariance and that which is derived from the ensemble perturbation prescribed by the alpha control variable. The tuning parameters, β_f^{-1} and β_e^{-1}, are used to control the weight given to the static or ensemble contributions respectively (e.g., β_f^{-1}=0.25 gives 25% of the weight to the static and 75% to the ensemble error covariance). These tuning parameters are sometimes assumed to sum to one, but that is not a strictly necessary constraint since neither the modeled covariance or the ensemble estimated covariance are perfect estimations of the true covariance (Bishop and Satterfield 2013).

1.5.5.2 Hybrid Four-Dimensional Ensemble Variational (4DEnVar)

Like 4DVar as an extension of 3DVar, hybrid 3DEnVar can be extended to the hybrid four-dimensional ensemble variational (4DEnVar). However, unlike 4DVar and hybrid En4DVar (Bannister 2017), hybrid 4DEnVar no longer requires a linear model and its adjoint to include flow-dependency within the analysis. Non-linear ensemble perturbations, $\boldsymbol{\mathcal{M}}_k$, valid at multiple times, k, are used to replace the linearized model. This makes hybrid 4DEnVar not only easier to maintain, but also less expensive to run relative to 4DVar. The hybrid 4DEnVar cost function can be written as follows

$$J(\delta\mathbf{x_f}, \boldsymbol{\alpha}) = \beta_f \frac{1}{2}(\delta\mathbf{x}_0)^T\mathbf{B}^{-1}(\delta\mathbf{x}_0) + \beta_e \frac{1}{2}\sum_{n=1}^{N}(\boldsymbol{\alpha}^n)^T\mathbf{L}^{-1}(\boldsymbol{\alpha}^n)$$
$$+\frac{1}{2}\sum_{k=1}^{K}(\mathbf{d}_k - \mathbf{H}_k\delta\mathbf{x}_k)^T\mathbf{R}_k^{-1}(\mathbf{d}_k - \mathbf{H}_k\delta\mathbf{x}_k) \quad (54)$$

where

$$\delta\mathbf{x}_0 = \delta\mathbf{x_f} + \sum_{n=1}^{N}[\boldsymbol{\alpha}^n \circ \delta\mathbf{x}_e^n]$$

$$\delta\mathbf{x}_k = \mathcal{M}_k\delta\mathbf{x}_0$$

$$\delta\mathbf{x}_k = \mathcal{M}_k\delta\mathbf{x_f} + \sum_{n=1}^{N}\left[\boldsymbol{\alpha^n} \circ (\delta\mathbf{x}_e)_k^n\right]$$

Recall that 4DVar solves for the 4D increments at the beginning of the window and propagates the innovation through the time window using the linearized model and its adjoint to find the best analysis across multiple states within the time window. The BEC in 4DVar is also time invariant meaning the flow dependency that is implicitly introduced comes from the propagation through the time interval by the linearized model. This makes the flow dependency weaker than from say explicitly including the flow dependency from an ensemble directly such as in hybrid EnVar methods. Compared to 4DVar, the hybrid 4DEnVar method is fundamentally different in how the 4D increments are determined because the background states from the nonlinear model are provided directly by the ensemble rather than through a linear model. In a sense hybrid 4DEnVar is like 3DVar FGAT in that multiple backgrounds are provided throughout an analysis window, but hybrid 4DEnVar is computing a 4D increment and 3DVar FGAT computes a 3D increment (i.e., not varying in time).

There are advantages and disadvantages to both En4DVar and 4DEnVar aside from the relative computational costs. The En4DVar, having "4DVar" in the name suggests that it uses the same tangent-linear machinery such as in 4DVar to implicitly include flow-dependency by propagating the full-rank BEC but is still restricted to using the static BEC at t=0 (Bannister 2017; Poterjoy and Zhang 2015). This, however, also allows for easier inclusion of time-dependent localization since localization can be applied at t=0 and then implicitly propagated by the tangent-linear model. On the other hand, 4DEnVar is unable to evolve the static BEC in time and including time-dependent localization within the 4DEnVar system is non-trivial and is typically not accounted for within the system (Poterjoy and Zhang 2015). Poterjoy and Zhang (2015) showed that En4DVar produces superior results relative to 4DEnVar when localization is included and otherwise produce similar results when no localization is used.

1.5.5.3 Hybrid Gain

Up until now, the hybrid methods that have been discussed have been focused on supplementing an operational variational system with dynamic error covariance information from an ensemble. Penny (2014) introduced a different approach which instead supplements an operational EnKF (or other ensemble method) with information from a 3DVar analysis. This type of hybrid method falls into a class of methods called hybrid gain. Essentially, hybrid gain methods combine the gain matrices of the ensemble and variational methods, rather than linearly combining the respective BECs.

Recall that in the EnKF, the Kalman gain matrix is the weighting applied to the observational innovation. Penny (2014) then defines a general hybrid gain matrix as

$$\mathbf{K}^{Hybrid} = \beta_1 \mathbf{K} + \beta_2 \mathbf{K}^{B}$$

$$= \mathbf{K} + \alpha \mathbf{K}^{B}(\mathbf{I} - \mathbf{H}\mathbf{K})$$

where $\mathbf{K}^{B}$ is just the optimal weight matrix using the static BEC,

$$\mathbf{K}^{B} = \mathbf{K}^{KF} = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}$$

$\mathbf{K}$ is the optimal weight from the ensemble part and where the beta parameters (β_1 and β_2) can be defined according to a specific configuration and are the weightings for the EnKF increment and 3DVar increment respectively. In Penny (2014), the Hybrid/Mean configuration defines $\beta_1 = 1 - \alpha$ and $\beta_2 = \alpha$ giving the analysis as

$$\mathbf{x} = \bar{\mathbf{x}}^{b} + \mathbf{K}^{Hybrid}(\mathbf{y}^{o} - \mathbf{H}\bar{\mathbf{x}}^{b})$$

$$= \bar{\mathbf{x}}^{b} + \mathbf{K}^{Hybrid}\mathbf{d}$$

where $\bar{\mathbf{x}}^{b}$ is simply the ensemble mean background. However, β_1 and β_2 need not necessarily sum to 1. The hybrid gain method has the advantage of easy implementation and has a relatively cheaper computational cost compared to hybrid variational methods with comparable performance (Whitaker 2020).

1.5.5.4 Discussion of Hybrid Methods and Their Use for Radial Wind DA

Some studies have shown that the use of a hybrid DA system are effective for convective-scale forecasting (Carley 2012; Gao and Stensrud 2014; Johnson et al. 2015; Li et al. 2012). Li et al. (2012) compared tropical cyclone forecasts initialized from 3DVar and hybrid 3DEnVar analyses with radial wind observations. All experiments assimilating radar data showed improved intensity analyses and forecast compared to experiments without; however, the experiment using hybrid 3DEnVar produced better fits to the observed radial winds and better precipitation forecasts. Carley (2012) used hybrid 3DEnVar for radar data assimilation and testing the impacts on the short term prediction of storms and compared the performance relative to the

3DVar approach and found that the hybrid approach led to better analyses. Gao and Stensrud (2014) also studied storm-scale radar data assimilation by assimilating simulated radar observations in an observing system simulation experiment comparing the 3DVar, hybrid 3DEnVar, and EnKF algorithms and found that the hybrid 3DEnVar and EnKF methods both performed relatively similar and were both better than 3DVar.

Lippi et al. (2016), tested the assimilation of radial winds using a hybrid 3DEnVar algorithm in the NAMv4 system with the goal of improving the assimilation of the data by updating the radial wind routines in the GSI. The results of which prompted Lippi et al. (2019) (i.e., Chapter 2 of this dissertation) where the radial wind assimilation in the GSI was further improved.

1.6 Filters vs. Smoothers

Each of the previous ensemble and variational methods discussed can be further classified as a filter or a smoother. A filter considers only current and past observations to create an analysis at the current time. A smoother considers all past and future observations and when new observations become available an analysis is computed for all previous times up until the current analysis time thereby creating an analysis of the model trajectory. Some of the most common filters used in atmospheric DA include the EnKF, 3DVar, 3DEnVar, 4DEnVar, and PF and the most common smoothers include 4DVar and the ensemble Kalman smoother (EnKS; Evensen and van Leeuwen 2000; van Leeuwen and Evensen 1996).

Kurosawa and Poterjoy (2021) explored the relative performance of ensemble and variational filters and smoothers in the context of the 40-variable Lorenz model (Lorenz 1996). The methods that were compared include EnKF-MDA (Multiple Data

Assimilation; Emerick and Reynolds 2012), EnKS-MDA, En4DVar, 4DEnVar, and the local PF and using the conventional EnKF and EnKS as a baseline. They found that each method has advantages under specific conditions which they identified to be a function of sampling error, nonlinearity in observation operators, and observation density. In summary, they found that smoothers tend to perform well under weakly nonlinear regimes (i.e., where there are dense observations) and where sampling error is the dominant source of bias. Filters, on the other hand, tend to perform better where nonlinearity in the observation operators and model dynamics play a larger role in the application.

It would be worthwhile to test different algorithms in future studies to help guide future data assimilation decisions in the operational framework; however, it is outside the scope of this work to determine the best algorithm to use since we are seeking to improve the use of radial winds within the systems that have already been implemented in operations. One other challenge to choosing the method of assimilation in operations is that it is an optimization problem between fast, efficient algorithms in order to disseminate analyses and forecast in a timely manner and employing the most accurate algorithms therefore the relative costs associated with each method is ultimately at the core of the decision.

1.7 Project Overview and Objectives

This dissertation is organized as follows:

Chapter 2 [published in the *Monthly Weather Review* as Lippi et al. (2019)] describes improvements made to the assimilation of Doppler radial winds for convection-

permitting forecasts of a heavy precipitation event. The major objectives of this chapter are to

i. Investigate the benefit of extending the observation operator to include vertical velocity,
ii. Add vertical velocity as an analysis control variable, and
iii. Investigate the benefit of tuning the super-observation parameters.

Chapter 3 describes and observing system simulation experiment in which the use of radial wind observations is tested through various network configurations. The major objectives of this chapter are to

i. Build upon Lippi et al. (2019) by formally exploring the impacts of radial wind observations in the GFS by assessing their impacts from a network that is accessible today (i.e., the NEXRAD network),
ii. Then extend that to the potential scenario of having access to all radial wind observations worldwide (i.e., the GLOBAL network),
iii. Test a purely hypothetical scenario where the worldwide network was designed with highly uniform spatial coverage (i.e., hypothetical network; HYPO for short), and
iv. Consider the effects of upstream DA on the downstream forecast.

Chapter 4 describes the development of an hourly global data assimilation system and builds upon Chapter 2 and 3 by testing the assimilation of radial wind within this new system. The major objectives of this chapter are to

i. Develop and test possible configurations in an hourly updating global data assimilation system including an overlapping window approach and an approach using catch-up cycles,

ii. Determine the best approach to handle imbalances caused by the initial conditions; the 4-dimensional Incremental Analysis Update (4DIAU) and an approach without IAU are examined, and

iii. Test the assimilation of radial winds in the GRR for a land-falling Tropical Cyclone case.

Chapter 5 will conclude this dissertation. This discussion will include a summary of the work presented in this dissertation and a thought for future work. This section will also include a summary of some unsuccessful efforts.

Chapter 2: Improvements to the Assimilation of Doppler Radial Winds for Convection-Permitting Forecasts of a Heavy Rain Event (Lippi et al. 2019)

This chapter is a slightly modified version of "Improvements to the assimilation of Doppler radial winds for convection-permitting forecast of a heavy rain event" published in the Monthly Weather Review as (Lippi et al. 2019) and has been reproduced here with the permission of the copyright holder obtained on 4/14/2022.

Lippi, D. E., J. R. Carley, and D. T. Kleist, 2019: Improvements to the Assimilation of Doppler Radial Winds for Convection-Permitting Forecasts of a Heavy Rain Event. *Monthly Weather Review*, **147,** 3609-3632, doi:10.1175/MWR-D-18-0411.1.

2.1 Introduction and Background

Only four national centers assimilate Doppler radar radial winds in operational, convective-scale systems: Météo-France, the Met Office, the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Prediction (NCEP), and the Japanese Meteorological Agency (Gustafsson 2018). NCEP first assimilated radial winds during the 1996 summer Olympics in Atlanta, Georgia; however, it was not until 2006 that radial winds were assimilated operationally in the 12-km NAM system [http://www.emc.ncep.noaa.gov/NAM/clog.php]. Operational numerical weather prediction (NWP) models have since progressed to convection-permitting resolutions (i.e., where convection is not parameterized) as a result of increased computer power and scientific advancement. Operational data assimilation systems, on the other hand, have not yet progressed to best leverage radial winds for convective-scale NWP systems. The assimilation of radial winds from the United

States WSR-88D network (Crum and Alberty 1993) has received a great deal of focus over the past few decades in the research community since Doppler radar is one of the only networks of instruments capable of sampling the storm-scale environment. Many studies (e.g., Gao and Stensrud 2014; Gao et al. 2004; Johnson et al. 2015; Xiao et al. 2005) have shown methods for assimilating radial winds that can benefit the analysis and forecast of convection-permitting models—most of these configurations are not yet suited for operational application owing to computational expense, e.g., compute time, nodes required, disk space, etc. In this study, the focus is on the assimilation of the radial wind observations. The exploration of the direct assimilation of radar reflectivity along with radial winds is beyond the scope of this manuscript and is the subject of complementary, collaborative efforts (e.g., Duda et al. 2019). However, reflectivity is used indirectly in all experiments in this study by applying radar-derived latent heating tendencies in the forward part of the digital filter initialization (Gustafsson et al. 2018; Peckham et al. 2015; Rogers et al. 2017).

Radial wind observations (Liu et al. 2016) from scan angles greater than 5° have historically been ignored for assimilation in operations because data coverage reduces rapidly as the elevation angle increase (Fig. 4). Additionally, model horizontal grid spacing was sufficiently coarse that one could ignore the analysis of non-hydrostatic motions when the radial wind assimilation algorithm was originally implemented at NCEP. NCEP's radial wind observation operator did not account for vertical motion which necessitated the restriction of a maximum allowable elevation angle in order to exclude observations with high scan elevation angles from contaminating the analysis, as the contribution from vertical motions would potentially become non-negligible at

such angles. Each of the other centers that assimilate radial wind observations also only consider the horizontal wind components in their corresponding observation operator (Ishikawa and Koizumi 2006; Lindskog et al. 2004; Liu et al. 2005; Simonin et al. 2014). By limiting the number of scan elevations (Fig. 4) in the assimilation process, a considerable amount of potentially useful information is discarded. Moreover, ignoring vertical velocity in the vicinity of strongly non-hydrostatic flows, i.e., where vertical velocity can be significant (5-50 m s^{-1}), is potentially problematic. Therefore, the GSI's (Wu et al. 2002) radial wind observation operator is extended to include vertical velocity. In future related studies, the inclusion of vertical velocity in the assimilation system will have potentially important implications for storm-scale models by helping to establish some of the non-hydrostatic dynamics associated with deep convective storms, thus leading to an improved forecast of these storms. However, due to current limitations of the model used in this study, the analysis of vertical velocity does not directly feed back into the forecast but acts primarily as a sink term in the observation operator. This allows the analysis of vertical velocity to have an indirect impact on the resulting forecast which has been shown through testing.

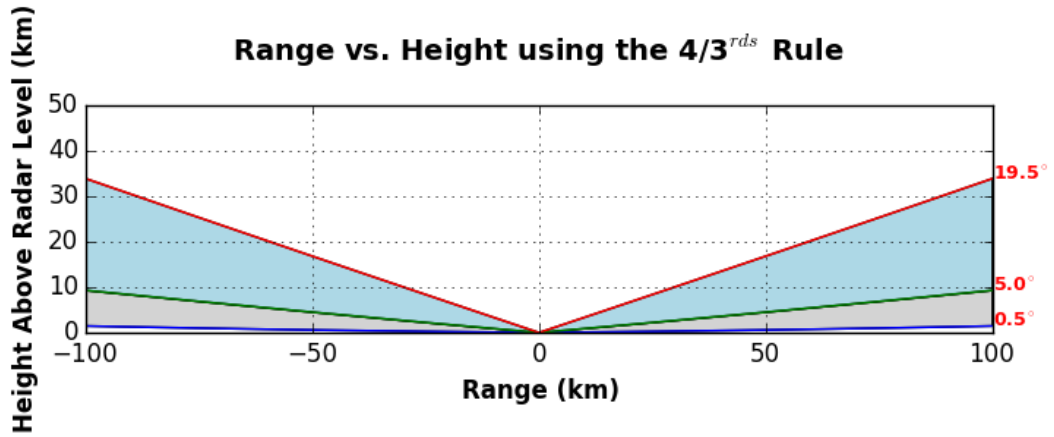

Fig. 4. Radar beam height estimation in km above radar level as a function of range and elevation angle using the 4/3rds rule for standard atmospheric refraction. The grey shaded area shows the elevation angles assimilated prior to this work with the maximum elevation angle restricted to 5°. The blue shaded area corresponds to the max elevation angle of WSR-88Ds from the available volume coverage patterns and can be considered the volume that could potentially be assimilated when accounting for vertical velocity in the observation operator.

Another aspect to consider for improving the radial wind assimilation is reducing the errors that arise from the disparity between observation and model resolution, i.e. errors of representativeness (Janjić et al. 2018). Prior to assimilation, the radial wind observations are on local polar radar coordinates having dimensions approximately 1° azimuth × 250-m gates with a temporal frequency of 5-10 minutes. The spatial and temporal resolution of this data exceeds the resolution of most operational analysis systems (e.g., a 9-km analysis grid). To account for some degree of representativeness error, it is standard practice to combine the radial wind observations to create so-called "super-observations" (e.g., Lindskog et al. 2004; Rihan et al. 2008; Simonin et al. 2014; Wheatley et al. 2015). Super-observation methods vary in technique, but all have the same purpose and naturally involve some degree of smoothing of the information content and underlying variance in the observation field. The super-observation technique used in the GSI follows the method of (Alpert and Kumar 2007) using tunable parameters that have not been modified in operations since

the advent of convection-permitting NWP. The super-observation parameters require revisiting as operational NWP and data assimilation moves toward resolutions where some of these mesoscale convective circulations are resolved. Refining these parameters will help retain storm-scale information content that, historically, would have been necessarily smoothed.

This study investigates the impacts and sensitivity of radial wind assimilation on a convection-permitting forecast system with specific attention to the formulation of the observation operator and super-observation pre-processing. The modifications to the radial wind assimilation scheme in the present study are considered first steps toward advancing the use of these observations at the convective-scale in an operational setting.

In Section 2.2, the model and data assimilation systems are described along with a description of the enhancements to the existing radial wind observation operator. The results are presented in section 2.3. This work is concluded in section 2.4 with a summary and a discussion of future work.

2.2 Methods

2.2.1 Model and Data Assimilation Configuration

Retrospective forecasts were produced using the North American Mesoscale version 4 (NAMv4) forecast system (Rogers et al. 2017). The NAMv4 runs hourly data assimilation cycles and was reconfigured to issue 36-hour forecast four times per day at 0000, 0600, 1200, and 1800 UTC over the 12-km parent and 3-km CONUS (Fig. 5) domains and 18-hour forecasts at each intermediate hour (e.g., 0100, 0200, 0300, 0400,

and 0500 UTC etc.) over the 3-km CONUS nest. Gustafsson et al. (2018) and Djalalova et al. (2016) provide additional information about the data assimilation cycling and other model characteristics of the NAMv4 CONUS nest forecast system. The review paper Gustafsson et al. (2018) also provides configurations for other operational centers supporting convective-scale data assimilation and NWP. The configurations used by the parent and CONUS nest domains are summarized in Table 3.

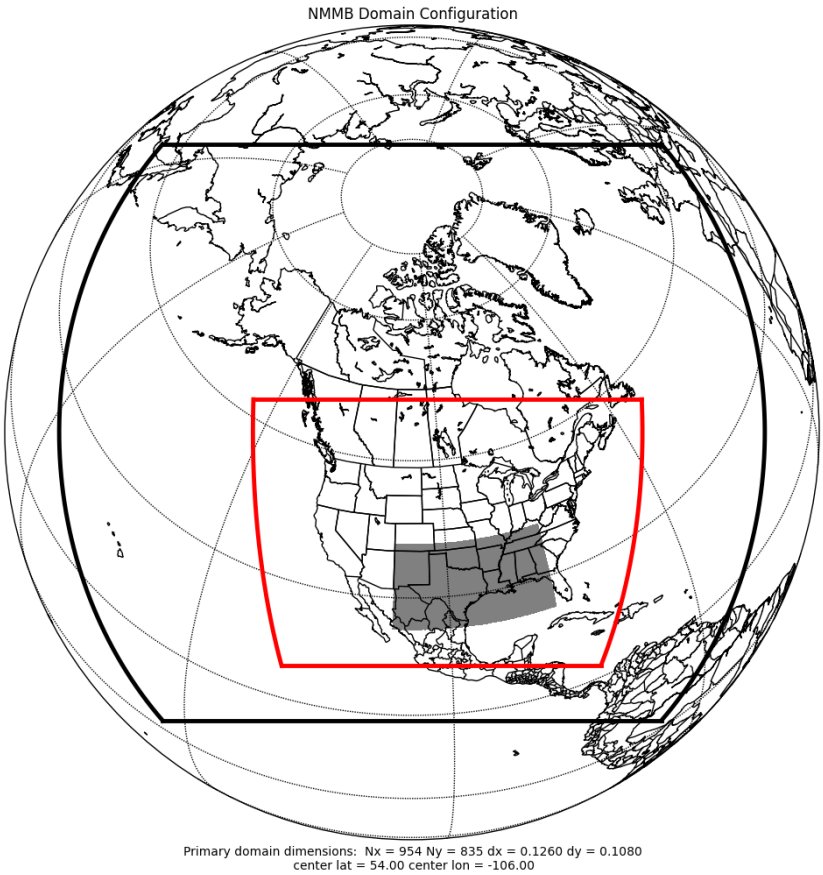

Fig. 5. Computational and verification domains. The outer (black) and inner (red) domains are the 12-km parent and 3-km CONUS nest computational domains, respectively. The objective verification statistics were calculated over the inner (gray shaded) domain located over the South Central (SC) U.S. using 5-km grid spacing.

Table 3. Summary of the configuration and physical parameterizations used by the two NMMB domains.

Domain	Grid space	Radiation (LW/SW)	Microphys	Turbulence	Sfc. lyr.	Land sfc.	Gravity Wave Drag	Cumulus
Parent	12-km	RRTMG (Mlawer et al. 1997; Iacono et al. 2008)	Ferrier-Aligo (Aligo et al. 2018)	MYJ (Janjić 2001)	MYJ (Janjić 2001)	Noah (Ek et al. 2003)	On (Alpert 2004)	BMJ (Janjić 1994)
CONUS nest	3-km	RRTMG (Mlawer et al. 1997; Iacono et al. 2008)	Ferrier-Aligo (Aligo et al. 2018)	MYJ (Janjić 2001)	MYJ (Janjić 2001)	Noah (Ek et al. 2003)	None	None

Data assimilation is performed using the GSI (Wu et al. 2002) system. The GSI is a variational analysis system formulated in model grid space and is used in many operational applications at NCEP, such as the NAMv4 forecast system (Rogers et al. 2017), the GFS (Kleist et al. 2009b), the Rapid and High Resolution Rapid Refresh systems (Hu et al. 2017), and the Real-Time Mesoscale Analysis system (De Pondeca et al. 2011). In this study, the GSI is configured using a hybrid three-dimensional ensemble-variational algorithm (3DEnVar; Kleist and Ide 2015b; Lorenc 2013; Wang et al. 2013; Wu et al. 2017) where the static covariance is implicitly blended with an ensemble covariance through the extended control variable method (Lorenc 2003; Wang 2010) and uses a T574 (~35-km grid spacing) 80-member ensemble from the GFS to provide the ensemble background error covariance (Wu et al. 2017). The weighting between the static and the flow-dependent, ensemble background error covariance is set to 25% and 75% respectively, which is consistent with the settings used in the operational NAMv4's 3-km CONUS nest data assimilation scheme. The univariate portion of the static background error covariance is modeled using recursive filters (Purser et al. 2003a, 2003b) and statistical balances are employed to account for the cross-covariances between control variables (e.g., streamfunction with balanced

temperature, velocity potential, and surface pressure). In regional applications, recursive filters are also used to impose localization of the alpha control variable as part of the ensemble-based background error covariance (Purser et al. 2003a, 2003b; Wu et al. 2017). The static background error covariance statistics were estimated using the so-called NMC method (Parrish and Derber 1992), which averages the differences between lagged forecast pairs (e.g., 24- and 48-hour forecasts) valid at the same time. The static background error covariance is identical to that used in the operational NAMv4 system, formulated to be isotropic and vary vertically and latitudinally.

The comparatively coarse nature of the static background error covariance and ensembles leveraged in the hybrid 3DEnVar system limits the fine scale nature of the analysis increments that can be prescribed. Therefore, the data assimilation is performed on a grid that is three times coarser than the model grid, i.e., a 9-km analysis grid. This practice is consistent with the operational configuration of the NAMv4 3-km CONUS nest system and allows for efficient use of computational resources without changing the characteristics of the analysis; however, this aspect will be revisited in future work when a fine-scale convection permitting ensemble is included in the EnVar algorithm.

2.2.2 Extending the Radial Wind Operator for Vertical Velocity

Data assimilation requires the comparison of model state variables to observations. As is the case for many observation types, radial wind is not a direct model prognostic variable; therefore, a relationship which transforms the model state, $\mathbf{x}$, into an observation equivalent, $\mathbf{y}$, must be used. This transformation occurs with the

use of the observation operator, H. The general relationship between the model state variables and observations can be written as:

$$\mathbf{y} = \mathrm{H}(\mathbf{x}) + \boldsymbol{\varepsilon}^{\mathbf{o}} \tag{55}$$

where $\boldsymbol{\varepsilon}^{\mathbf{o}}$ is the observational error. Errors associated with radial wind observations can be attributed to four main sources: instrument error, pre-processing errors, errors of representativeness, and errors introduced by the observation operator (Waller et al. 2016). Errors introduced by the observation operator are the result of omissions and approximations of the equation. Therefore, the prescribed observation uncertainty should also be updated when modifying the observation operator; however, this will be left for future work.

2.2.2.1 The Radial Wind Observation Operator

The original GSI radial wind operator took the form:

$$\mathbf{V_r}(\theta,\alpha) = u\cos(\theta)\cos(\alpha) + v\sin(\theta)\cos(\alpha) \tag{56}$$

where $\mathbf{V_r}$ is the radial wind observation, u and v are the model horizontal wind components, θ is 90° minus the azimuth angle of the radar, and α is the elevation (or tilt) angle of the radar. The formulation here uses the azimuthal directions based on the unit circle rather than the Cardinal Directions. The effects of Earth's curvature and atmospheric refraction are accounted for via the four-thirds approximation (e.g., Ge et al. 2010). Wind rotation is also included to accommodate relevant coordinate conversions. These corrections are included in the azimuth and elevation angles of Eq. (56).

Historically, the radial wind operator has not accounted for the influence of vertical motion, owing to the hydrostatic scales of motion to which it was applied (e.g.,

12-km). Due to this restriction, radial wind observations associated with scan elevation angles greater than 5° were routinely discarded in order to reduce the impact of vertical motion on the radial wind field. Using the simplified observation operator [Eq. (56)] will produce a bias error in the model equivalent value when high elevation angles (e.g., Fig. 4) are used considering the vertical component of the wind could make a non-trivial contribution in the observed radial wind especially in the presence of strongly non-hydrostatic motions. However, the sensitivity to this assumption will be a function of several factors such as the model resolution.

To extend the radial wind observation operator to include the vertical component of the wind, Eq. (56) was modified to include a third term:

$$\mathbf{V_r}(\theta,\alpha) = u\cos(\theta)\cos(\alpha) + v\sin(\theta)\cos(\alpha) + w\sin(\alpha) \qquad (57)$$

where w represents vertical velocity.

The difference between these two observation operators is shown in Fig. 6. When w is zero, there is no difference between the observation operators. When w is positive (negative), Eq. (57) underestimates (overestimates) the radial wind. The magnitude of error is strongly dependent on the vertical velocity as well as the elevation angle. The maximum possible error for an elevation angle more practically used in an operational scenario, less than 20°, with an assumed vertical velocity of 10 m s^{-1}, is +/- 3 m s^{-1}. Such large vertical motions are not unusual for deep convection. This error would be much greater (less) in regions of very strong (weak) vertical motion and/or steeper (less steep) elevation angles.

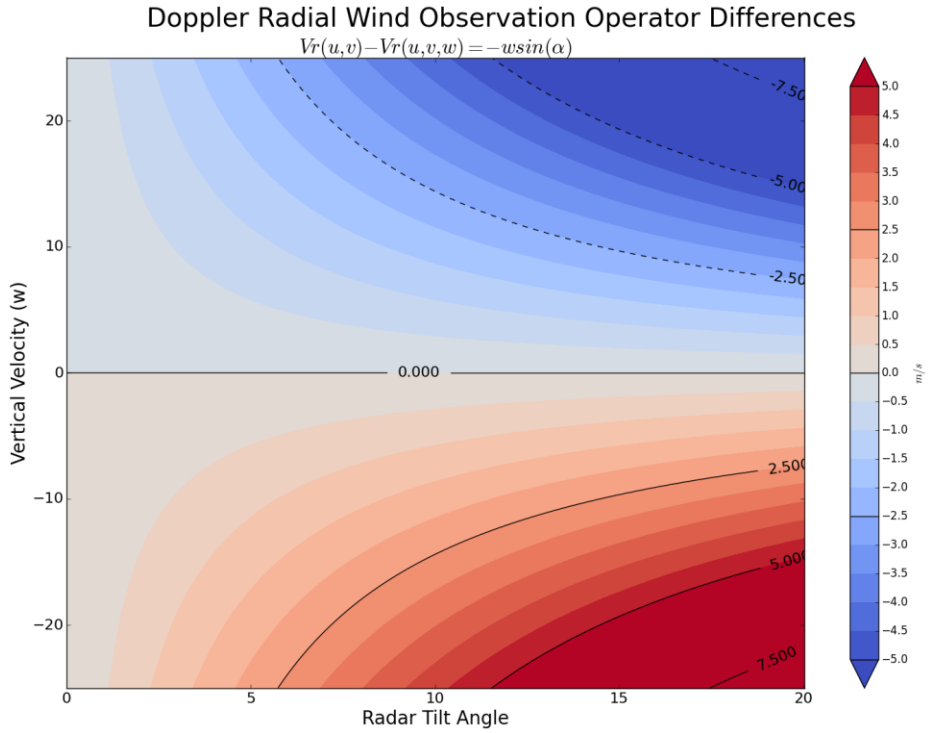

Fig. 6. The difference [Eq.(57) – Eq.(58)] between the original observation operator [Eq. (57)] and the observation operator with a term for vertical velocity [Eq.(58)].

Large raindrops, around 5 mm in diameter, have terminal velocities (w_t) of about 10 m s^{-1} (Spilhaus 1948) leading to an error from Eq. (57) of $w_t \sin(\alpha)$, or about 1.7 m s^{-1} in the most extreme cases for elevation angles of 10°—the highest elevation angle considered in the experiments. Eq. (57) more accurately simulates the radial winds as compared to Eq. (56), although still contains simplifications that should be considered in the future (e.g., hydrometeor terminal velocity and beam broadening; Fabry 2010). Nonetheless, this change results in a more physically consistent forward model, reduces representativeness error, and allows for relaxing the maximum radar elevation angle to beyond 5 degrees, ultimately allowing a larger volume of observations to be assimilated.

2.2.2.2 Vertical Velocity Analysis Control Variable

The variational analysis (e.g., Hybrid 3DEnVar) is determined based on the minimization of a cost function. The cost function is measured, in part, by the combination of the departures of the observations and background from a select set of analysis control variables. The control variables may be any set of model or model-related variables. In the case of the GSI, the control variables are chosen with considerations towards Gaussianity and multivariate balance. For example, the control variables for the NAMv4 CONUS nest applications are stream function, unbalanced velocity potential, unbalanced temperature, unbalanced surface pressure, and normalized specific humidity. During the minimization of the cost function, increments to the control variables are made and then added to the initial background forecast. Only increments to variables related to the control variables (directly or indirectly through multivariate relationships) can be affected in the resulting analysis. Vertical velocity was added as a control variable to support the analysis of non-hydrostatic motion, which follows from the work above to improve the radial wind observation operator.

The experiments here utilize an 80-member ensemble from the GDAS running at T574 (~35-km). The ensemble-based analysis increment is constructed by utilizing the extended control variable method (Lorenc 2003) used in the GSI hybrid 3DEnVar cost function (Wu et al. 2017). The ensemble perturbation wind variables are u and v and are not converted into balanced and unbalanced parts. Ensemble perturbations of w are not used at this time, owing to the fact that the perturbations are from a global, hydrostatic ensemble.

The model used in this study, the Nonhydrostatic Multiscale Model on the B-grid (NMMB; Janjić and Gall 2012), *diagnoses* vertical velocity via the non-hydrostatic continuity equation with the knowledge of the hydrostatic and non-hydrostatic pressures and temperature. Therefore, while the model is fully non-hydrostatic, it does not feature a prognostic term specifically for vertical velocity (Janjić and Gall 2012). This presents some difficulty in terms of providing an analyzed vertical motion field to the model; therefore, the associated analysis of vertical velocity in this study serves primarily to reduce error present in the observation operator and therefore improve representativeness. However, limited testing has shown that the analysis of vertical velocity can indirectly impact the forecast.

2.2.2.3 Single Observation Tests

A useful tool for diagnosing several aspects of a data assimilation system is the single observation test. Here, we seek to verify the successful implementation of the radial wind operator using synthetic[7] radial wind observations using a 3DVar configured GSI. Two single observation tests were used to isolate and test the vertical and horizontal components of the observation operator by assigning the elevation angle to two extremes: 90° and 0°, respectively. A 90° elevation angle corresponds to the radar looking straight up, and the 0° elevation angle corresponds to the radar looking tangentially to the Earth's surface at the observation location. For the 90° elevation test, the observation was placed about 1-km above ground (corresponding to approximately model level 18; between 800 and 900-hPa) directly over the Fort Hood, Texas radar

[7] Synthetic observations may refer to observations that cannot be physically possible but are useful in testing the functionality of the data assimilation system.

(KGRK). The observation for the 0° elevation test was placed about 20-km directly to the east of KGRK about 200-m above ground level (approximately model level 4; between 1000 and 900-hPa). Although it is not geometrically possible for a radial wind observation with an altitude above ground level to have a corresponding 0° elevation angle, synthetic observations are used to facilitate testing by relaxing such rule. This also applies to the 90° test since elevation angles in the real world do not exceed 20° for WSR-88Ds.

An azimuth angle must be defined for the 0° elevation test. The assignment of the azimuth angle will not affect the results of the 90° elevation test since only the vertical term in Eq. (57) will remain. For demonstration purposes, we would like to achieve a positive increment in only the zonal direction for the 0° elevation test (and vertical direction for the 90° elevation test); therefore a 0° azimuth, or eastward pointing radar was chosen. To obtain a positive analysis increment with an eastward (upward) pointing radar, a positive innovation is needed. This is because the innovations are calculated in observation space, and outbound radial winds are positive, which is a westerly (upward) wind for the 0° (90°) elevation test. For both single observation tests, each observation was assigned an innovation of +1.0 m s^{-1} with an observation error of 1.0 m s^{-1}. The statistical balance relationships associated with the static background error covariance were also set to zero to avoid multivariate increments to isolate the impact from the single observation tests (e.g., increments in temperature when only a *u* observation was given). Finally, the assigned winds were assumed earth relative, and thus the winds were not rotated from latitude-longitude coordinates. This ensures the observation is exactly as prescribed.

From these single observation tests, one can verify that the newly implemented observation operator correctly maps the model state into an observed equivalent quantity so that an analysis increment can be made to the correct wind component according to the given observation specifications (e.g., elevation and azimuth angles). A comparison of the two experiments confirms the successful implementation of the radial wind operator. As expected, the 90° elevation test shows an analysis increment for only vertical velocity (Fig. 7b) depicting an isotropic analysis increment, as expected in a 3DVar analysis. This test used a globally constant value for the vertical velocity background error covariance with a value of 0.7 m^2 s^{-2} and a horizontal and vertical scale of influence of 27-km and 10 grid units respectively. In contrast, the 0° elevation test has isotropic increments only in the zonal wind component (Fig. 8a). The length scales for the horizontal wind components, originating from the operational NAMv4 forecast system, are much larger than for vertical velocity hence the comparatively broad increments.

The background error covariance statistics were not recalculated to include vertical velocity for this study since the vertical motion field is a diagnostic quantity in the forecast model[8], and therefore acts primarily as a sink term. The specifications used for the background error covariance in the single observation tests are used in the main set of experiments presented in this study. Some sensitivity experiments were also done with alternate length scales and those results will be discussed briefly as well.

[8] Recall that the model in this study diagnoses vertical velocity via the non-hydrostatic continuity equation using hydrostatic and non-hydrostatic pressures and temperature. It does not feature a prognostic term for vertical velocity, instead the associated nonhydrostatic prognostic term is nonhydrostatic pressure.

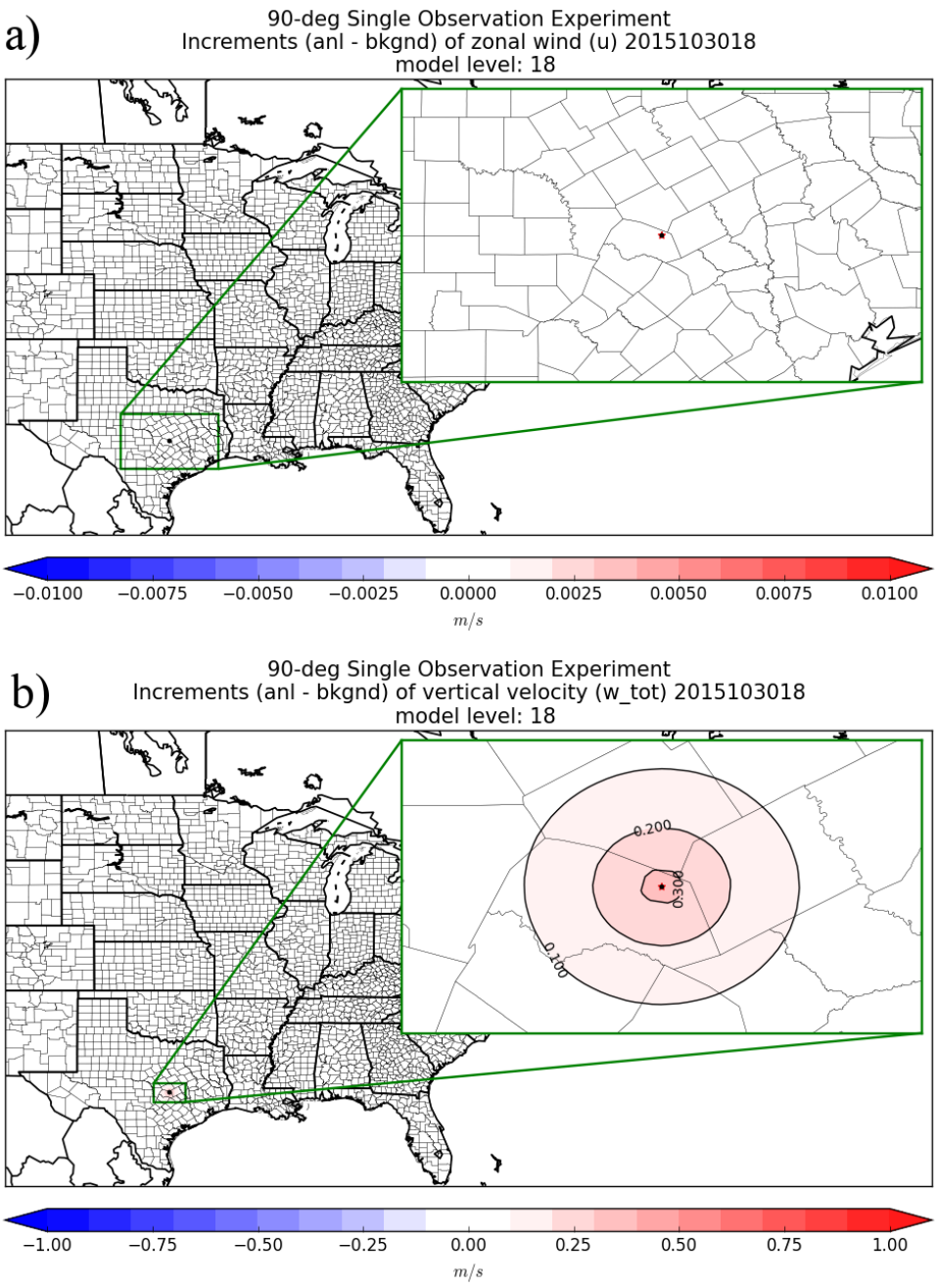

Fig. 7. Results from the 90° tilt single observation test (SOT). The pseudo-observation (black dot) was placed 1-km above the Fort Hood, TX radar KGRK (red star) with an innovation of +1 m s^{-1} and an observation error of 1 m s^{-1}. 3DVar analysis increments are shown for (a) zonal wind *u* and (b) vertical velocity *w*.

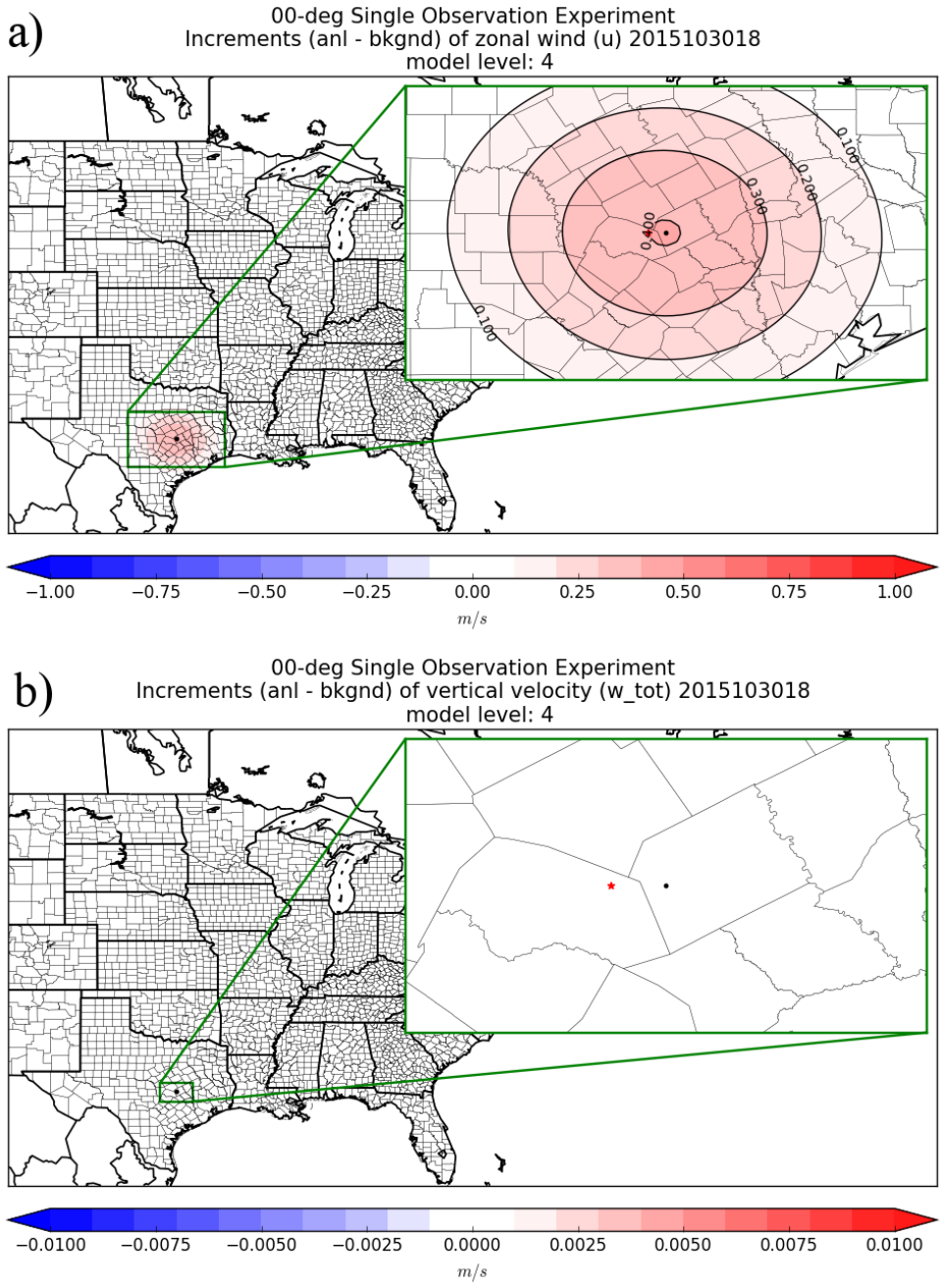

Fig. 8. Results from the 0° tilt single observation test (SOT). The observation (black dot) was placed 20-km directly to the east of the Fort Hood, TX radar KGRK (red star) about 200-m above ground level with an innovation of +1 m s^{-1} and an observation error of 1 m s^{-1}. 3DVar analysis increments are shown for (a) zonal wind *u* and (b) vertical velocity *w*.

2.2.3 Creating Super-Observations of Radial Wind Data

Prior to assimilation, the radial wind observations (Liu et al. 2016) undergo an aggregation and smoothing procedure to generate super-observations. The super-observation technique is employed to remove representativeness errors and reduce the volume of observations, which are often redundant and come at the price of additional computational expense (Alpert and Kumar 2007). The super-observation procedure involves a spatial and temporal averaging of the radial wind observations within a volume defined in radar coordinates via the parameters noted in Table 4.

Table 4. List of the super-observation parameters and their default and experimental values.

Configuration	**Azimuth range (degrees)** $\Delta\boldsymbol{\theta}$	**Elevation angle range (degrees)** $\Delta\boldsymbol{\alpha}$	**Gate spacing (meters)** $\Delta\boldsymbol{r}$	**One-half time range (hours)** $\Delta\boldsymbol{t}$	**Max elevation angle (degrees)** $\propto_{\boldsymbol{max}}$	**Minimum number of samples** $\boldsymbol{N}$
Default	5	0.25	5000	+/- 0.500	5	50
Experimental	3	0.25	3000	+/- 0.125	10	10

There are seven super-observation parameters: del azimuth, $\Delta\theta$; del elevation angle, $\Delta\alpha$; del range, Δr; del time, Δt; maximum elevation angle, α_{max}; minimum number, N; and maximum range R_{max}. $\Delta\theta$, $\Delta\alpha$, and Δr controls the width, height, and length of the super-observation box azimuthally, in elevation, and along the beam respectively. Δt is the one-half time window for which to include observations. N is the minimum number of observations that must exist within the super-observation box defined in space and time in order for a super-observation to be calculated. Finally, α_{max} and R_{max} are the maximum allowable radar elevation angle and maximum range from the radar, i.e., an observation is discarded if it exceeds either of these values.

Together, these parameters define a 4-dimensional box in radar observation space and time that forms the bounds for which the observations are processed into super-observations. The effects of the super-observation parameters can be captured by visualizing the radar observations before (Fig. 9a) and after (Fig. 9c, d) the radial winds have undergone super-observation pre-processing. The corresponding radar reflectivity observations are shown in (Fig. 9b) for reference. Through visual comparison, one can see that Fig. 9d features more detail than Fig. 9c when compared to the unprocessed, raw data depicted in Fig. 9a.

The maximum possible number of super-observations per radar at a single time can be estimated using the following equation:

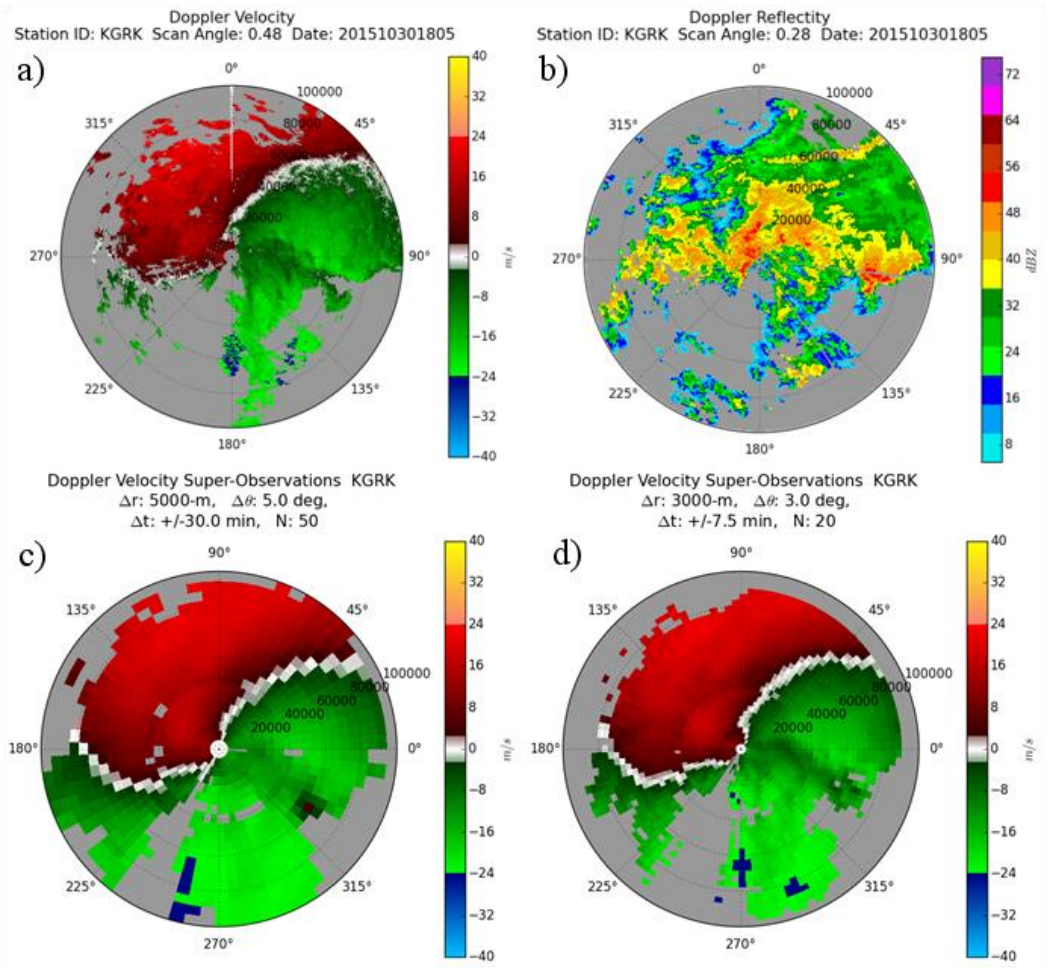

Fig. 9. (a) Level-II radial wind observations before (c) and after super-observation processing using the default super-observation settings (d) and experimental super-observation settings for the radial winds. The associated radar reflectivity for ~1800 UTC 30 October 2015 is shown in (b) for reference.

$$\text{Maximum possible number of super-observations} = \frac{360^\circ}{\Delta\theta} \times \frac{R_{max}}{\Delta r} \times \frac{\alpha_{max}}{\Delta\alpha}.$$

For the default super-observation settings, there could be up to 28,800 super-observations per radar and up to 158,400 with the modified experimental values. The actual number of super-observations will generally be much lower because observations will exist in only a fraction of the total radar volume. For example, the super-observation counts that correspond to the Fort Hood, Texas radar (KGRK) for the convective period 1800 UTC on 30 October 2015 shown in Fig. 9a are 8,000 and

25,000 super-observations for the default and modified super-observation settings respectively.

In addition to visualizing how modifying the super-observation parameters will change the spatial and temporal geometry of the super-observation box, it is also important to understand how this will affect the processing of the underlying data. A rough estimation of the maximum number of radial wind observations (M_{max}) that can be contained within a single super-observation box can be obtained by:

$$M_{max} = \frac{\Delta\theta}{1^\circ} \times \frac{\Delta r}{250\text{m}} \times 120 \frac{\Delta t}{5\text{min}}$$

assuming the radial wind observation spatial resolution of 1° x 250-m, a temporal resolution of 5-min, and only one scan per elevation angle. The maximum number of observations per super-observation box can only be estimated due to a few factors:

1. The actual beam width may differ slightly from the assumed 1°. This could increase or decrease the number of observations.
2. During severe weather events, when more aggressive volume coverage patterns may be used, there could be more than one scan at the same tilt. This will increase the number of observations.
3. The 0.25° super-observation box height ($\Delta\propto$) may include data from two elevations, possibly increasing the number of observations.
4. When the radar mode changes, scan times may not be 5-min, which could increase or decrease the number of observations.

The estimated number of observations per super-observation box are 1200 and 108 for the specified parameters in Table 4 for the default and modified configurations respectively.

Fig. 10a, b depicts the distributions of the number of observations per super-observation box for a single time and radar that corresponds with Fig. 9c, d respectively. The default configuration shows a bimodal distribution with a peak near zero and a peak around 650-700 observations but shows that few of the super-observations were generated at or near the maximum possible number of radial winds in a super-observation box (M_{max}). The features represented by radial winds evolve at sub-hourly intervals which may partially explain the large difference in the actual and estimated number of observations per super-observation box. The default time window for assimilating radial winds is one hour (+/- 30 minutes), but there is no guarantee that those features will exist in or totally encompass a super-observation box during that entire period. For example, a super-observation box may be located near a storm boundary and may include observations from only a few minutes. Observation quality control may also play a role by discarding observations of poor quality.

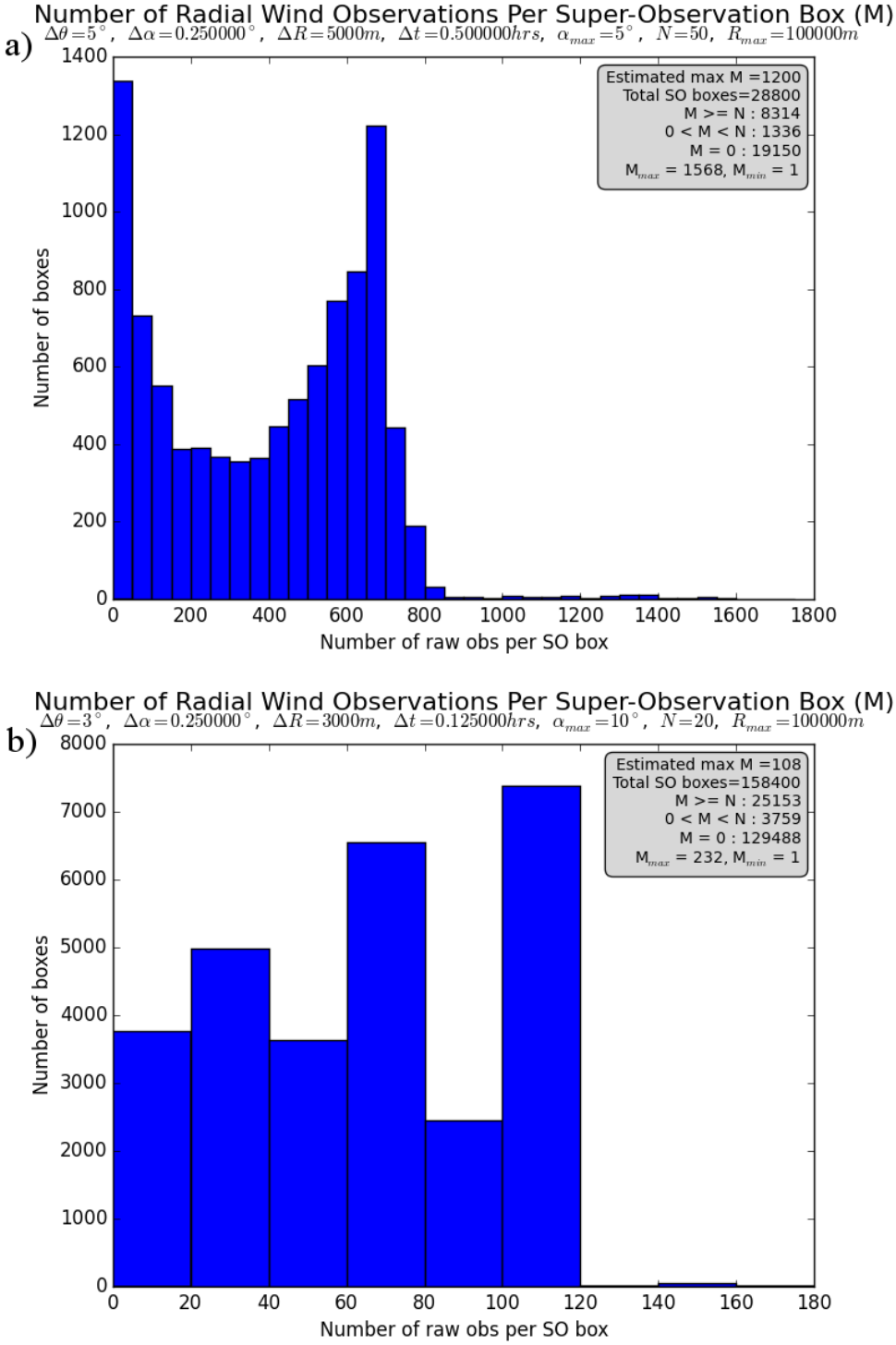

Fig. 10. Distribution of the number of radial wind observation per super-observation box (M) for a single time and radar that corresponds with for Fig. 9c,d histogram a, b respectively. Additional information appears in the annotation box which includes the estimated maximum value for M, the total number of super-observations boxes for one radar, the number of super-observation that would be assimilated (M ≥ N), the count of super-observations that do not meet the requirement of N but have at least one Level-II observation, the count of super-observation boxes with 0 radial wind observations, and finally the maximum and minimum values for M for this case.

The large time window for aggregation has the potential to smooth out convective-scale features of interest as they evolve on relatively short time scales. By reducing the time window of observations to 15 minutes (+/- 7.5 minutes), the super-observations should be more representative of the analysis time, and we see a peak in Fig. 10b at 100-120 observations corresponding to the estimated max number of observations of 108 for the given settings.

2.2.4 Experimental Design

To test the impact of adding vertical velocity to the radial wind observation operator and the associated sensitivities with super-observation parameters refined to retain more detail, a control simulation and four experiments were initialized at 0000 UTC 30 October 2015 on the 3-km CONUS nest grid (Table 2). The modified super-observation parameter values were chosen based on limited, preliminary testing that showed improved wind analyses balanced with considerations towards the data assimilation system memory limitations.[9]

1. **control**: configured to match the operational NAMv4 with the exception of refined radial wind quality control discussed in Lippi et al. (2016) which eliminated the quality check against Velocity Azimuth Display (VAD) winds. The VAD winds are radar-specific wind profiles derived from the radial wind field and are based on a linearized wind model (Holleman et al. 2005). The procedure of the VAD quality control is to compare the radial wind super-

[9] There are specific run time and disk space limitations that must be considered for implementing a system in operations. In this example we must exercise caution assimilating raw, unprocessed observations which can exceed the available memory on compute nodes. Thinning observations and creating super-observations helps eliminate this issue while also addressing other scientific concerns, such as representativeness.

observations to the VAD derived winds. Only observations that agree with the VAD winds within a predetermined tolerance are accepted. This has the impact of retaining only those winds that align with the larger-scale flow derived from the VAD wind profile and discards the winds representative of the convective-scale. Since this study focuses on the assimilation of radial wind observations associated with a convective system the VAD quality control step was removed in this work (Lippi et al., 2016).

2. **w_incl**: as in control but includes *w* in the observation operator and as an analysis variable.
3. **w_so_elev5**: as in w_incl but uses refined super-observation (so) parameters to retain more detail and keeps the maximum elevation angle at 5° (elev5).
4. **w_so_elev10**: as in w_so_elev5 but raises the maximum elevation angle to 10° (elev10).
5. **so_elev10**: as in w_so_elev10 but *w* is <u>not</u> used in the observation operator.

Table 5. List of NAMv4 radial wind data assimilation experiments and the configurations for the super-observation settings as well as if vertical velocity is used in the observation operator and as a control variable. Each super-observation parameter ($\Delta\theta$, $\Delta\propto$, Δr, Δt, $\propto_{max,}$ and N) has a distinct impact on the resulting super-observation box size in time and space. $\Delta\theta$, $\Delta\propto$, and Δr control the width, height, and length of the super-observation box azimuthally, in elevation, and along the beam respectively. Δt is the time window for which to include observations. N is the minimum number of observations that must exist within the super-observation box defined in space and time in order for a super-observation to be calculated. Finally, $\propto_{max}$ is the maximum allowable radar elevation angle, i.e. an observation is discarded if it exceeds this value.

Exp	**w**	**Azimuth range (degrees)** $\Delta\boldsymbol{\theta}$	**Elevation angle range (degrees)** $\Delta\propto$	**Gate spacing (meters)** $\Delta\boldsymbol{r}$	**One-half time range (hours)** $\Delta\boldsymbol{t}$	**Minimum number of samples** $\boldsymbol{N}$	**Max elevation angle (degrees)** $\propto_{\boldsymbol{max}}$
control	No	5	0.25	5000	+/- 0.500	50	5
w_incl	**Yes**	5	0.25	5000	+/- 0.500	50	5
w_so_elev5	Yes	**3**	0.25	**3000**	**+/- 0.125**	**20**	5
w_so_elev10	Yes	3	0.25	3000	+/- 0.125	20	**10**
so_elev10	**No**	3	0.25	3000	+/- 0.125	20	10

2.2.5 Verification

Frequency bias (FBIAS) and the fraction skill score (FSS) were used to objectively evaluate forecasts of accumulated precipitation. The verification was computed on a common, 5-km grid that encompassed the area of focus for the case study (Fig. 5; gray shaded). A description for each of the statistical methods follows.

FBIAS is used to determine if the forecast is too "dry" or too "wet" and is calculated in the following manner

$$\text{FBIAS} = \frac{hits + false\ alarms}{hits + misses} \tag{58}$$

Stated more simply, FBIAS is the ratio of "yes" forecasts to "yes" observations. FBIAS greater than 1 (less than 1) indicates more (fewer) points predicting a given threshold compared to observations, i.e., a wet bias (dry bias).

		Observed: Yes	Observed: No
Forecast	Yes	*hit*	*false alarm*
	No	*miss*	*correct rejection*

Fig. 11. 2x2 contingency table which shows the relationship between events for a dichotomous verification situation (Wilks 2011)

The FSS is a neighborhood verification approach which relaxes the requirement for forecast and observed events to match exactly at the grid scale. Instead, the fractional coverage of predicted and observed grid-point events above a specified threshold are compared over a range of increasing spatial windows (Roberts and Lean 2008). Thus, FSS reveals how well the forecast resembles the observations at a given spatial scale. The FSS is computed in the following manner

$$\mathrm{FSS} = 1 - \frac{\frac{1}{N}\sum_N (P_f - P_o)^2}{\frac{1}{N}\left[\sum_N P_f^2 + \sum_N P_o^2\right]} = 1 - \frac{\mathrm{FBS}}{\mathrm{FBS}_{worst}} \tag{59}$$

where N is the number of grid-points contained within the neighborhood area; P_f is the fractional coverage of forecast events that exceed a pre-determined threshold; and P_o is the fractional coverage of observed events that exceed the threshold. The fractional values (P_f and P_o) are what is computed at the increasingly large spatial windows. The FSS is thus the ratio between the fractional Brier score (FBS; Roberts 2005; related to the difference between the fractional coverage of predicted and observed events that exceed the threshold) and the worst possible FBS (FBSworst; Roberts 2005; related to the summation of the total number of forecast and observed fractions that exceed the

threshold). The FSS ranges from 0 to 1 where 0 would be a complete mismatch and 1 would be a perfect match of forecast events to observed events.

All statistics were aggregated in 3-hourly periods for accumulated precipitation out to 18-hours for the 0900, 1200, 1500, and 1800 UTC forecast cycles on 30 October 2015; the 3-hourly cycles are used to accommodate the 3-hour buckets of the precipitation observations. A bootstrapping technique, using 2000 replications, was used to test for statistical significance at the 90% confidence level.

2.2.6 Case Study Overview

Retrospective forecasts during the convectively active period of 30-31 October 2015 in the southern plains were made to evaluate the impact of the modifications to the radial wind observation operator as well as the modifications to the super-observation parameters. This case was chosen because it consists of a fairly diverse set of forcing mechanisms including an upper-level short-wave trough, cold front, interaction between the low-level jet (LLJ) and a warm front, and a pre-frontal confluence zone. Additionally, this case exhibited heavy precipitation, flooding, damaging winds, and several tornadoes. More than one foot of rain fell in less than 24-hours between Austin and San Antonio, Texas (Fig. 12), which caused extensive flooding throughout that region. Another line of heavy precipitation occurred later in the period near Houston, Texas. Several tornadoes rated between EF-0 and EF-2 were reported in these two regions (Fig. 13).

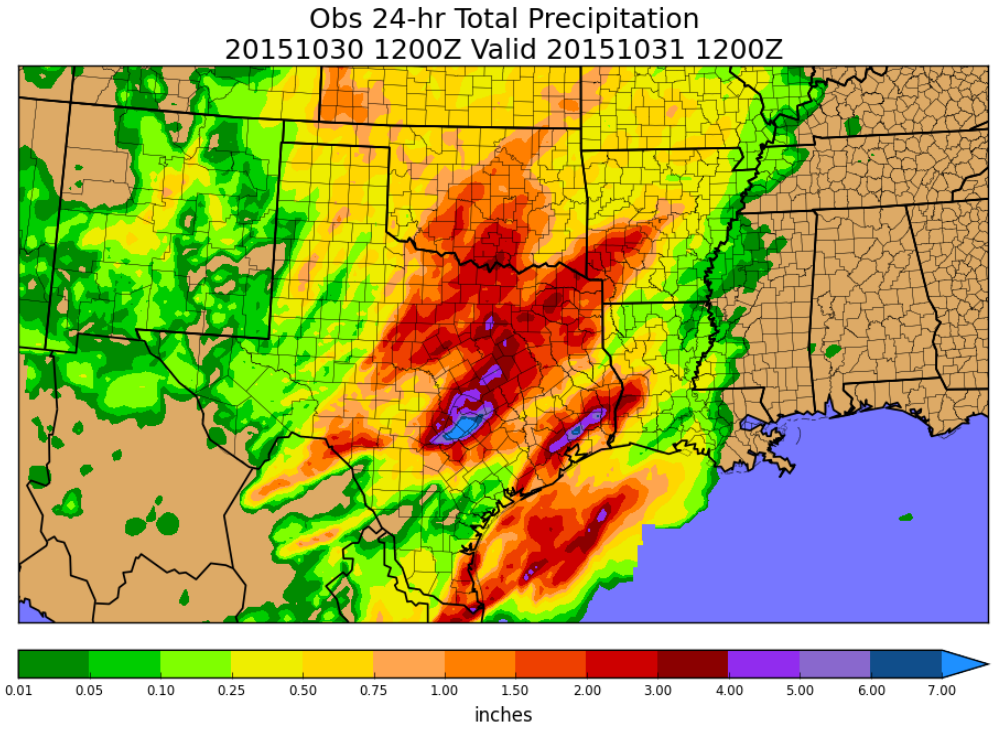

Fig. 12. NCEP stage IV observed total 24-hour precipitation (inches) valid 31 October 2015 1200Z (Lin and Mitchell 2005).

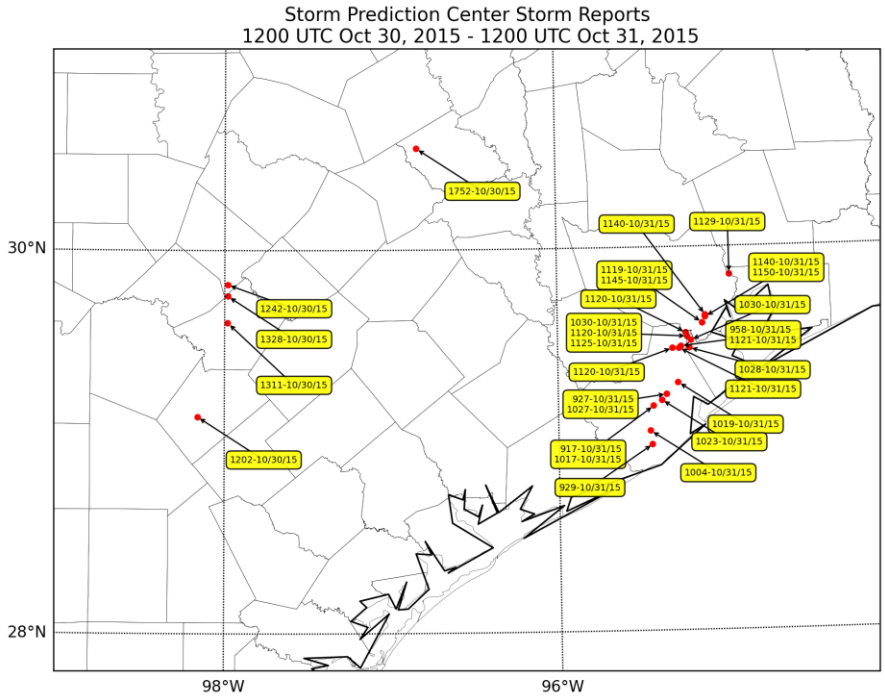

Fig. 13. Tornado reports for 1200 UTC 30 October 2015 to 1200 UTC 31 October 2015 from the Storm Prediction Center from the Austin, Texas and Houston Texas areas. All times are in UTC. [Data available online at: http://www.spc.noaa.gov/exper/archive/event.php?date=20151030]

2.3 Results

2.3.1 Data Assimilation System Fit-To-Observations

The analysis impact of the modified observation operator and refined super-observation processing can be diagnosed by comparing the observation-minus-background (OmB) and observation-minus-analysis (OmA) residuals between experiments. Three statistics from the fit-to-observations will be analyzed: observation counts, root mean square (RMS) innovation, and bias.

There are two factors that affect the observation counts: event-specific observation availability and experimental-specific factors which includes the pre-processing of the observations via tunable super-observation parameters. The increase in observation counts from 0900 UTC (Fig. 14a) through 1800 UTC (Fig. 14d) is the result of storms developing within the CONUS domain, and thus increasing the number of scattering particles for the radar(s). Additionally, super-observation processing over smaller spatial and temporal box sizes results in a larger number of total super-observations. For example, the control and w_incl assimilate the same number of super-observations (e.g., Fig. 14d red line vs. black line) since there is no change to the super-observation parameters. There are more super-observations in w_so_elev5 due to the reduction in super-observation box sizes, thus allowing for more super-observation boxes (e.g., Fig. 10). By raising the elevation angle from 5° to 10° in experiments w_so_elev10 and so_elev10, there is an additional increase in the observation counts. The two *elev_10* experiments (e.g., Fig. 14d blue line vs. purple line) assimilate the same number of super-observations since the only difference between those experiments is related to the observation operator.

A direct comparison of RMS and bias values cannot be made between experiments with differing super-observations parameters due to sensitivities related to changing the underlying characteristics of the super-observations themselves, e.g., differing sample sizes as noted in the prior paragraph. For example, changing the degree of super-observation processing, by changing the spatial box size and time window length, will change the statistics within the super-observation box (e.g., variance). This effect is then compounded from the use of data assimilation cycling and complicates the ability to interpret the results between experiments using differing super-observation parameter settings. Therefore, only two comparisons will be made, both of which only consider the effects of including vertical velocity in the observation operator: control vs. w_incl and so_elev10 vs. w_so_elev10.

There is an overall neutral impact on the OmB/OmA RMS (Fig. 15) and bias (Fig. 16) as a result of extending the forward operator to include *w*. The OmB RMS and bias for experiment w_so_elev10 (w_incl) have little to no differences compared to the counterpart experiment so_elev10 (the control). Based on this limited data set, there is no consistent improvement or degradation in the RMS and bias of the fit-to-observations for differences as a result of the improved observation operator. In general, the OmB RMS/bias values are larger than their corresponding OmA RMS/bias for each experiment, which is expected.

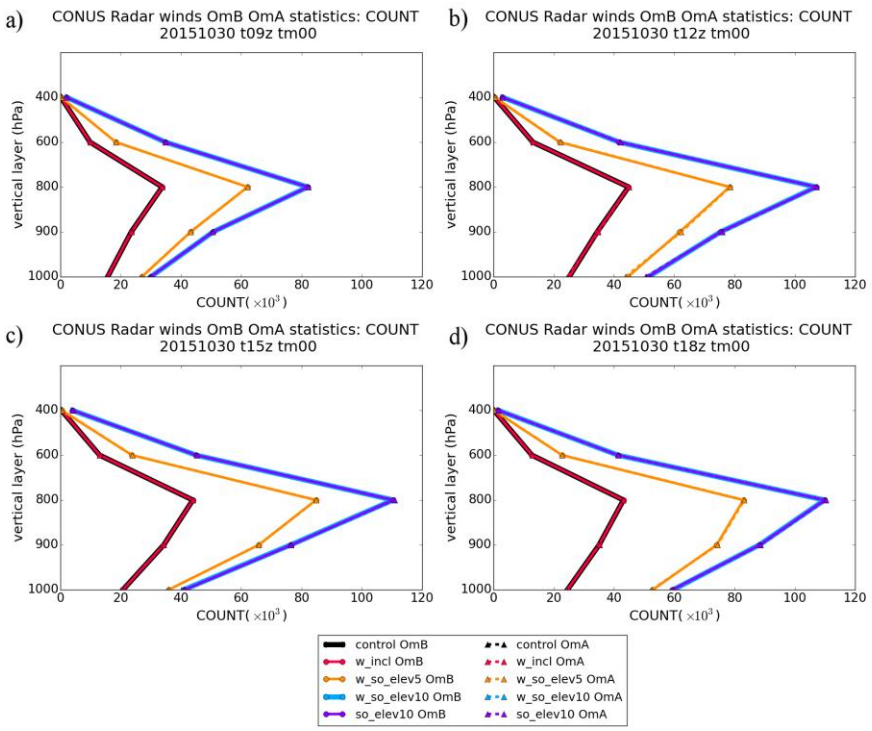

Fig. 14. Counts from the observation minus background (OmB; circles/solid line) and observation minus analysis (OmA; triangles/dashed lines) statistics files over entire CONUS nest domain for the forecast cycles a) 0900 UTC b) 1200 UTC c) 1500 UTC d) 1800 UTC 30 October 2015. The individual circle/triangle markers along the x-axis denote the aggregate values for the total CONUS atmosphere. Note: the OmB and OmA observation counts are nearly identical and therefore the lines overlap sufficiently to be indistinguishable.

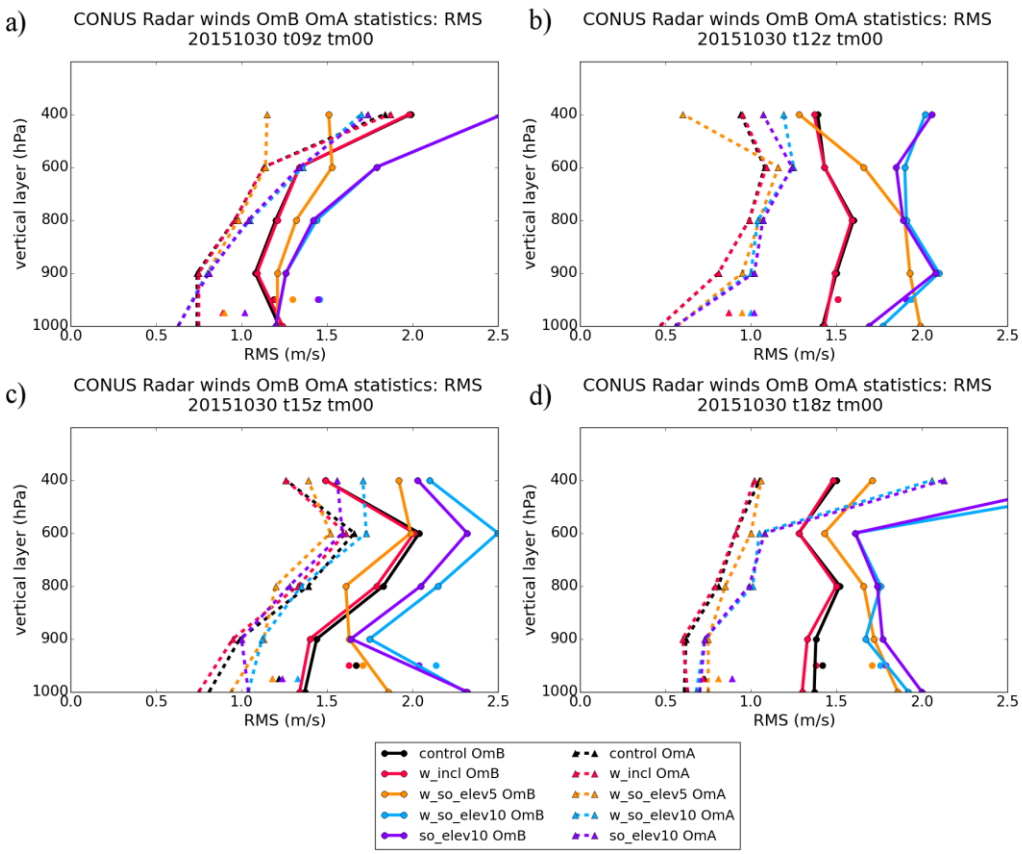

Fig. 15. Root mean square (RMS) error from the GSI observation minus background (OmB; circles/solid line) and observation minus analysis (OmA; triangles/dashed lines) statistics files over entire CONUS nest domain for the forecast cycles a) 0000 UTC b) 0600 UTC c) 1200 UTC d) 1800 UTC 30 October 2015. The individual circle/triangle markers along the x-axis denote the aggregate values for the total CONUS atmosphere.

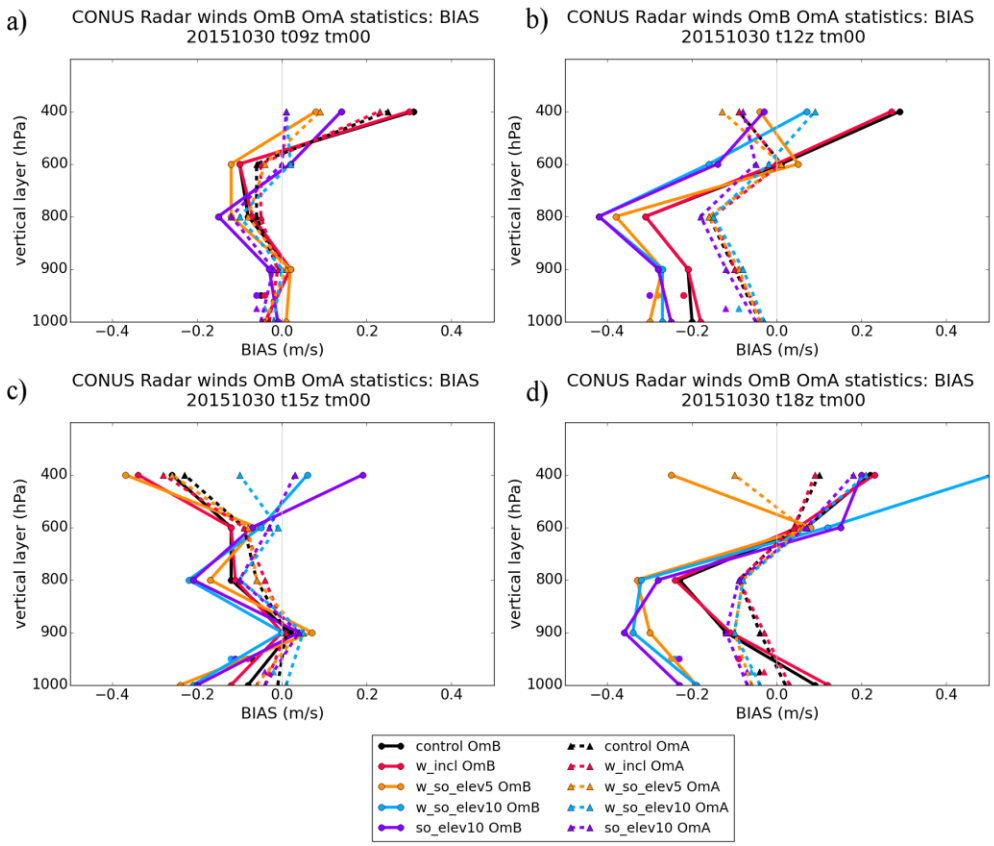

Fig. 16. Bias from the GSI observation minus background (OmB; circles/solid line) and observation minus analysis (OmA; triangles/dashed lines) statistics files over entire CONUS nest domain for the forecast cycles a) 0000 UTC b) 0600 UTC c) 1200 UTC d) 1800 UTC 30 October 2015. The individual circle/triangle markers along the x-axis denote the aggregate values for the total CONUS atmosphere.

2.3.2 Forecast Assessment

To assess the impacts on the resulting forecasts, we begin with a qualitative comparison of forecasts of accumulated precipitation from the 1200 UTC initialization time on 30 October 2015 (Fig. 17). This is followed by an assessment of FSS and FBIAS scores from forecasts initialized at 0900, 1200, 1500, and 1800 UTC on 30 October 2015 (Fig. 18 and Fig. 19). Each experiment (refer to Table 2) was tested for statistical significance at the 90% level using bootstrap confidence intervals constructed from 2000 replications on the difference curves computed relative to the control simulation. Statistical significance is found where the confidence intervals do

not encompass zero. Statistics are considered only over the approximate region that experienced heavy rainfall, which is a broad area covering the southern CONUS (Fig. 5; gray shaded). Further, FSS statistics are summarized across several thresholds and box sizes via the use of a so-called scorecard (Fig. 20).

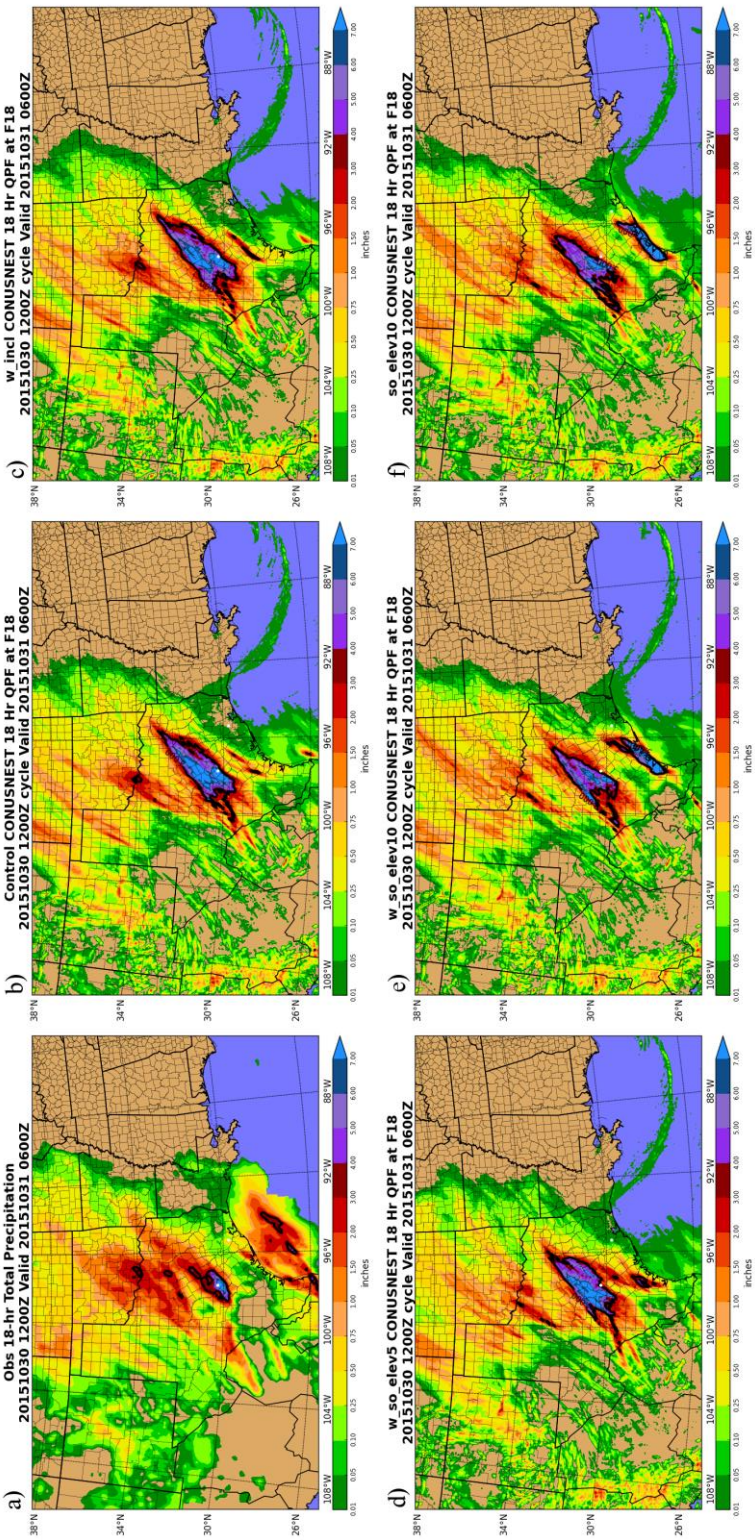

Fig. 17. 18-hour accumulated precipitation for the period 1200 UTC 30 October 2015 through 0600 UTC 31 October 2015. The observations are displayed in (a) and the experiments are (b) control, (c) w_incl, (d) w_so_elev5, (e) w_so_elev10, and (f) so_elev10. The white dot (30.27 N, 97.74 W) denotes the location of Austin, Texas and the white dot (29.76 N, 95.37 W) denotes the location of Houston, Texas. The black contour denotes the area of 3 in.

The control (Fig. 17b) simulated a much larger area of very heavy precipitation located near central Texas as compared to observations (Fig. 17a) and compared to the experiments with modified super-observation parameters (Fig. 17d, e, and f respectively). The changes in the coverage of the heavy precipitation can be noted by the 3 in. contour (Fig. 17 bold black line). The overprediction of the precipitation is reduced in this area when the super-observation parameters were adjusted, suggesting some sensitivity to the super-observation processing. There appears to be qualitatively little-to-no impact made when vertical velocity was included in the observation operator when comparing against the control (experiment w_incl; Fig. 17b). The FSS (Fig. 18) and FBIAS (Fig. 19a) difference scores support this mostly neutral change as the control (black) and w_incl (red) lines are consistently very similar among all metrics and thresholds, with the exception of FBIAS degradation at 0.10 and 0.25 in. and an improvement at 1.00 in. The scorecard (Fig. 20) shows that the difference between w_incl and the control is slightly degraded for the lower thresholds (e.g., 0.01, 0.05, and 0.10 in.) at all box sizes and slightly improved for the higher thresholds (e.g., 0.50 and 1.00 in.).

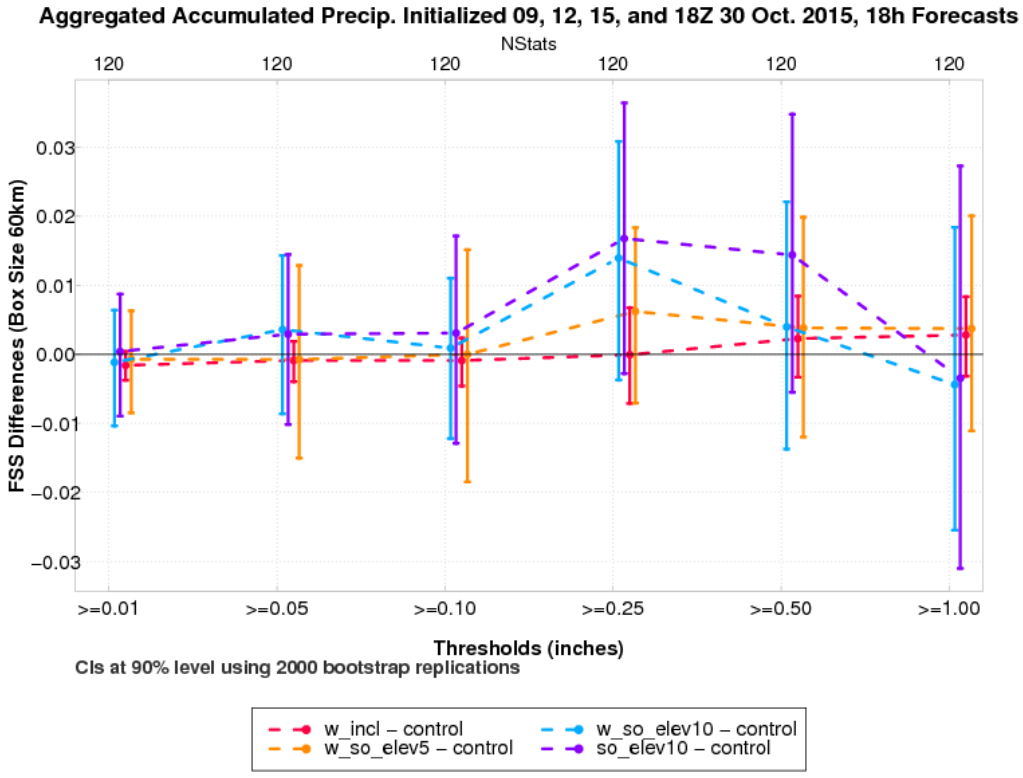

Fig. 18. Fractions skill score (FSS) differences from control for 3-hour precipitation forecasts out to 18-hours aggregated from 3-hourly forecasts initialized beginning at 0900, 1200, 1500, and 1800 UTC 30 October 2015.

There was a slightly more distinct impact in w_so_elev5 (Fig. 17d), as compared to the control and the w_incl experiments. In experiment w_so_elev5, the spatial coverage of the wet bias of the precipitation bullseye near Austin, Texas was reduced. While the wet bias in this region was reduced, the placement of precipitation was slightly eastward of the observed location. The w_so_elev5 FSS (Fig. 18) difference scores show similar but larger impacts to that of experiment w_incl. Slight degradation (improvement) is found for the lower (higher) thresholds. For FBIAS (Fig. 19a) there is degradation at thresholds 0.01 and 0.05 in., i.e., increasing the wet bias of control (Fig. 19b), and a trend toward improvements at thresholds 0.50 and 1.00 in. The scorecard (Fig. 20) shows w_so_elev5 to have a similar pattern to w_incl, but there

is less degradation at the low thresholds and stronger improvements at higher thresholds. There is an overall slight improvement for all box sizes and all thresholds as compared to w_incl.

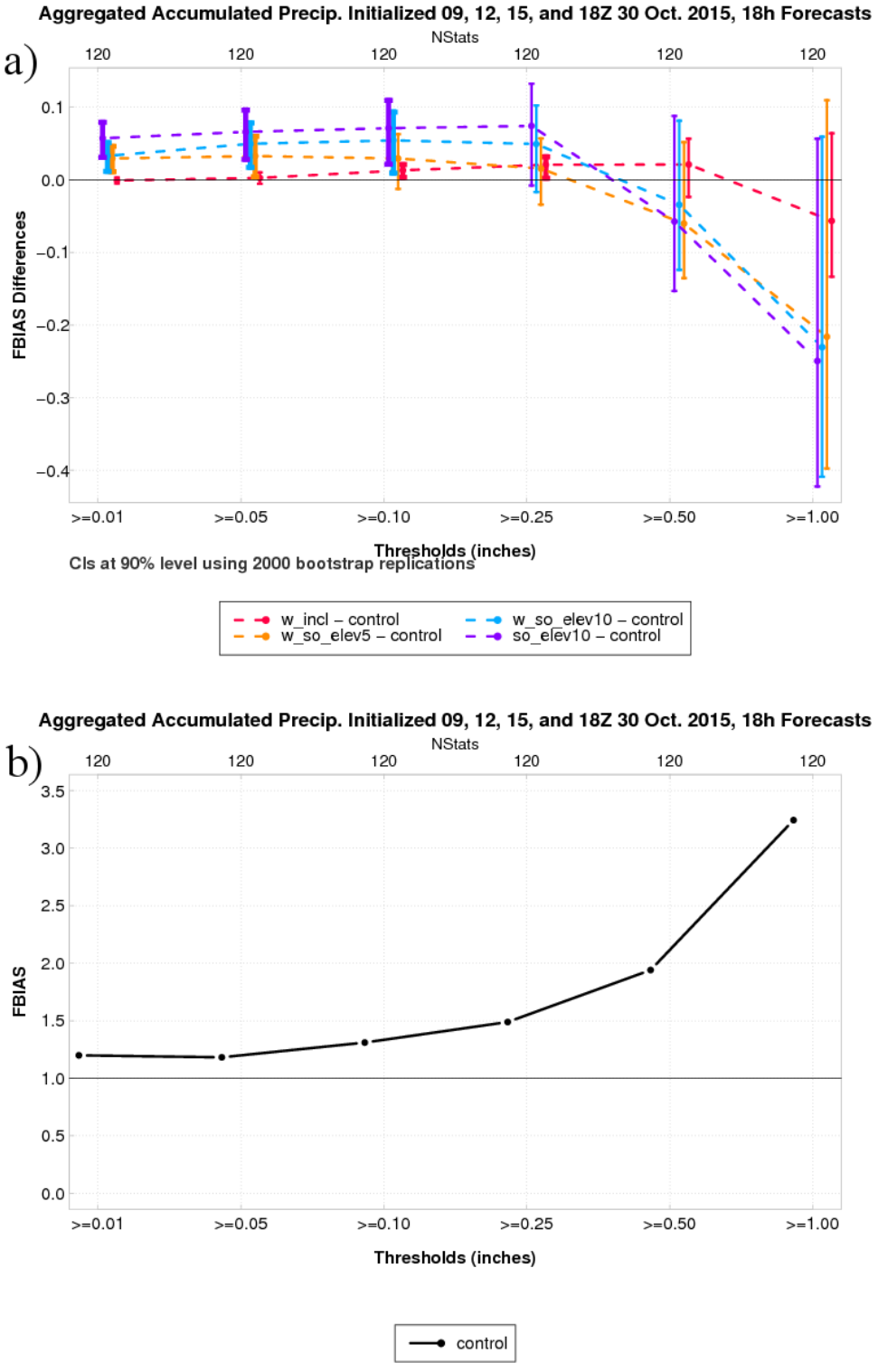

Fig. 19. Frequency bias (FBIAS) difference from control (a) and the control FBIAS values (b) for 3-hour precipitation forecasts out to 18-hours aggregated from 3-hourly forecasts initialized beginning at 0900, 1200, 1500, and 1800 UTC 30 October 2015.

The w_so_elev10 and so_elev10 experiments (Fig. 17e and f respectively), having similar forecasts, further reduce spatial coverage of the precipitation bullseye as compared to w_so_elev5. The *elev10* experiments appear to have the subjectively better overall forecasts as the broad wet bias in the central part of Texas appears improved and more closely reflects the extent and magnitude seen in observations. While both experiments still miss the heavy rain in southeast Texas extending over the Gulf, they do reflect heavy precipitation in the region associated with rotating storms that occurred later in the period (discussed in Section 3.3.4), unlike the control and w_incl. The FSS statistics (Fig. 18) confirm the positive forecast impact at 0.25 and 0.50 in. but slight degradation at 1.00 in. The FBIAS statistics (Fig. 19a) show a reduction in wet bias for the 0.50 and 1.00 in. thresholds. From this set of experiments, there is not a clear signal that suggests that including vertical velocity provides a statistically significant advantage. The largest improvements may be attributed to the enhanced super-observation settings with raised elevation angles, in the margins of statistical significance.

The experiments with modified super-observation parameters tended to have the greatest positive impact on the forecast. The extension of vertical velocity to the observation operator was found to have mostly neutral impacts in combination with the modified super-observation parameters (Fig. 20). The three super-observation experiments exhibited heavy precipitation located near Houston, Texas (Fig. 17d, e, and f) with increasing magnitudes when higher elevation angles were included. This is an important distinction because there was a marked increase in simulated storm rotation correlating to the locations of tornado reports.

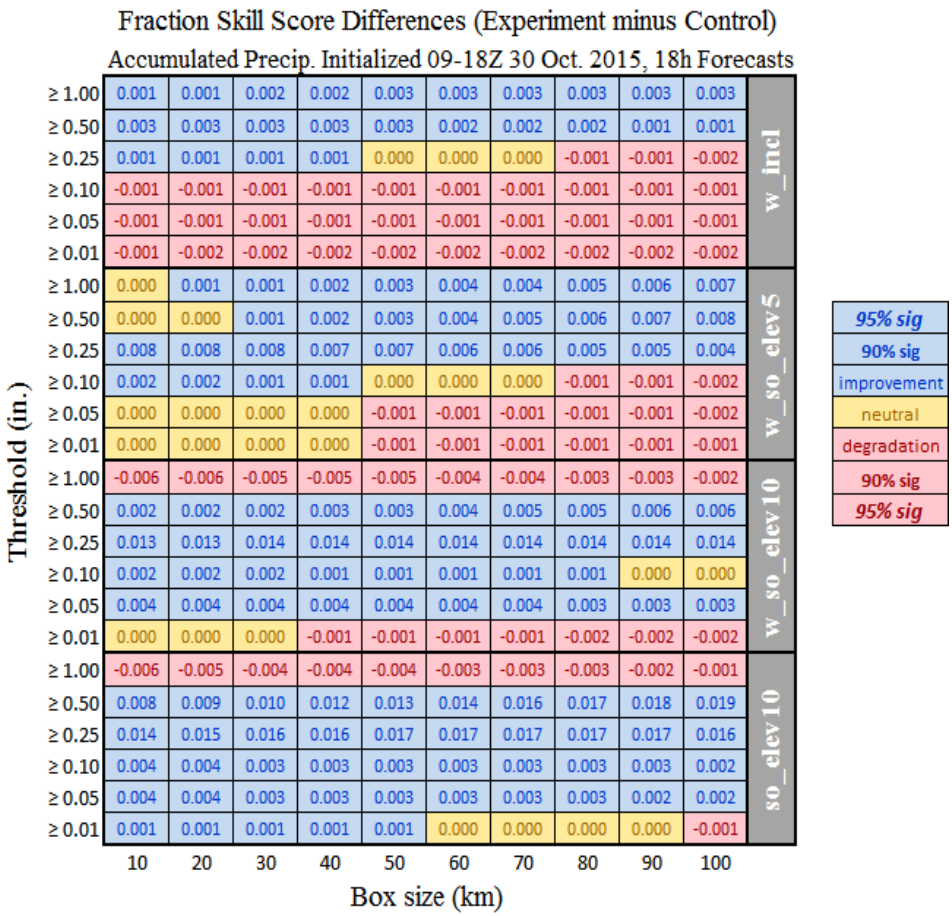

Fig. 20. Fraction skill score (FSS) scorecard for the differences between the control and each experiment for various accumulated precipitation thresholds and box sizes. Improvements and degradations are shown, rounded to three decimal places, with the corresponding statistical significance at the 90% and 95% confidence intervals.

2.3.3 Houston, Texas Region Precipitation

Here we discuss the aspects of the experimental configurations that led to the differences in precipitation forecasts along the Texas Gulf Coast and into Houston, Texas and use tornado reports as a proxy for storms with rotating updrafts.

Fig. 13 notes the time and date of each tornado report for the 1200 UTC 30 October 2015 to 1159 UTC 31 October forecast period; the earliest tornado report in the Houston, Texas region occurred after 0900 UTC on 31 October 2015. According to observed radar reflectivity, the storms associated with the tornado reports in the

Houston area initiated just off the coast of Port O'Connor, Texas around 0300 UTC 31 October 2015. The storms traveled northeastward into the Houston area where several tornadoes occurred. This discussion will focus on the 1200 UTC initialized forecast performance with a focus on the updraft helicity forecasts near Houston, Texas.

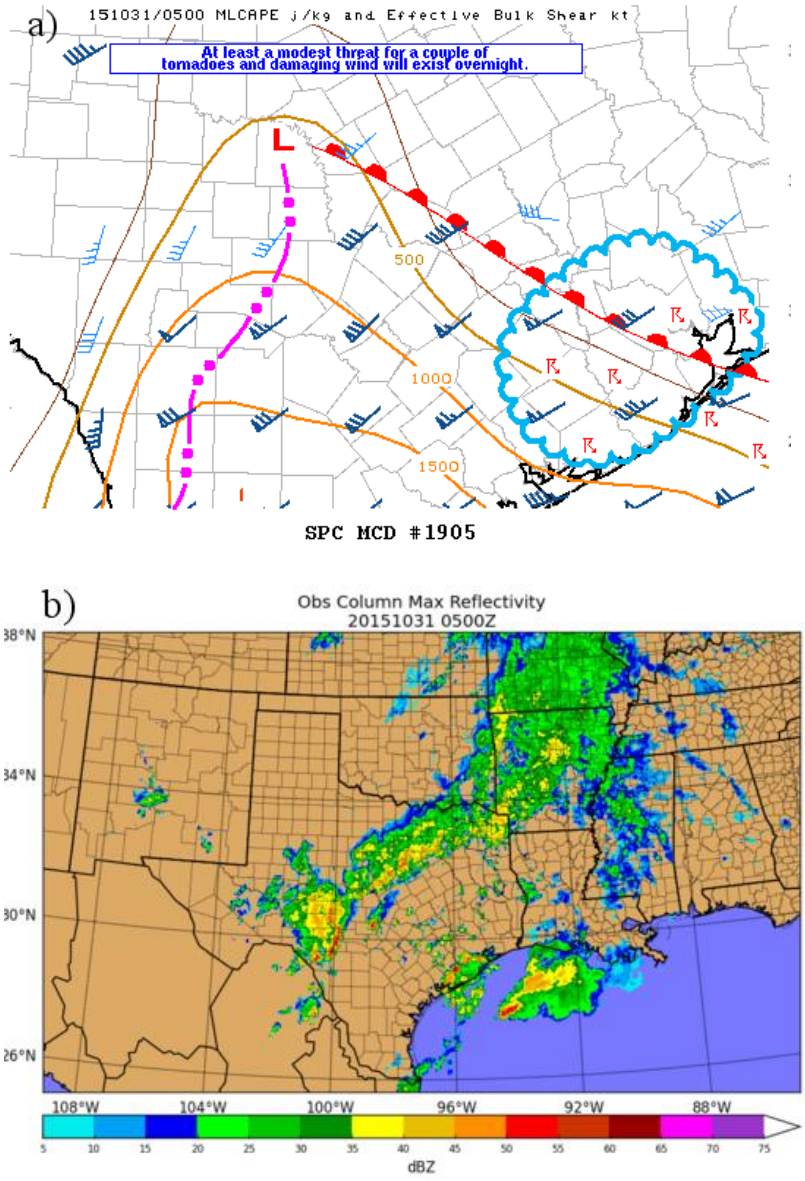

Fig. 21. Storm prediction center mesoscale analysis for 0500 UTC 31 October 2015[Available online at: http://www.spc.noaa.gov/exper/archive/event.php?date=20151030] accompanied by observed reflectivity valid at the same time.

The 1200 UTC initialized super-observation experiments, especially so_elev10 and w_so_elev10 all predicted heavy precipitation along the Texas Gulf Coast beginning near Port O'Connor, Texas and extending into the Houston, Texas region. This is a major difference between these experiments and the non-super-observation experiments as the non-super-observation experiments were *not* able to capture the initiation and life cycle of the storms associated with the warm frontal boundary (Fig. 21a) and low-level jet. As it was noted before, the storms in this region initiated around 0300 UTC 31 October 2015 while the super-observation experiments tended to have a much earlier initiation of several hours (e.g., 6+ hours) prior to verification. The *elev10* experiments clearly predicted strong, rotating storms (Fig. 22c, d, e, and f) to occur in the location that was being monitored by the Storm Prediction Center (Fig. 21a) and more closely corresponds to the locations of the observed tornado reports (Fig. 13) in the Houston, Texas area as well as with observed radar reflectivity (Fig. 22h). The control run predicted storms from 1900 UTC to 0000 UTC on the 31st (Fig. 22a, b) that were associated with interaction with the warm front. Those storms were less organized and less persistent as compared to the modified super-observation parameter experiments.

The forecasts are clearly sensitive to the super-observation parameters. The super-observation settings were modified to average observations over a smaller spatial box and temporal window; therefore, there were fewer observations per super-observation box and thus a lesser degree of smoothing of the observations. By smoothing over smaller spatial and temporal windows, there was a greater density of total super-observations assimilated where each was more representative of the raw

form of the observations occurring at the time of assimilation. The demonstrated sensitivity indicates that the operational super-observation settings (i.e., those used in the control) smoothes the observations to a greater degree than necessary and discards potentially useful information. Averaging over a long time period may also have the undesirable impact of dampening significant convective-scale motions, e.g., situations when the observations change rapidly at sub-hourly time intervals. The modified configuration used within this study, however, may not necessarily be the optimal configuration for these parameters but identifies sensitivities that result in potential forecast improvement thus forming a foundation for continued development toward improving the assimilation of these data.

2.3.4 Forecast Sensitivity to Background Error Covariance Length Scales

The three experiments that feature vertical velocity in the observation operator: w_incl, w_so_elev5, and w_so_elev10 were also run with the same global constant background error covariance 0.7 m^2 s^{-2} but different horizontal and vertical length scale: 100-km and 1-grid units respectively. These values correspond to default values that may be used by the GSI. The forecast impacts were also assessed for this new set of experiments using the 18-hour total accumulated precipitation plot from the 1200 UTC cycle and the corresponding FSS stats summarized by a scorecard for multiple thresholds and box sizes (not shown). The experiments that featured vertical velocity in the observation operator and as an analysis control variable were rerun to demonstrate that there is a forecast sensitivity to these parameters despite that the fact that vertical velocity is a diagnostic variable in the NMMB. Future research should consider generating new background statistics when considering analyzing vertical

velocity and that the results of this study are not conclusive to the importance or lack thereof of including vertical velocity. There was consistency in results that modifying the super-observation parameters improved the results.

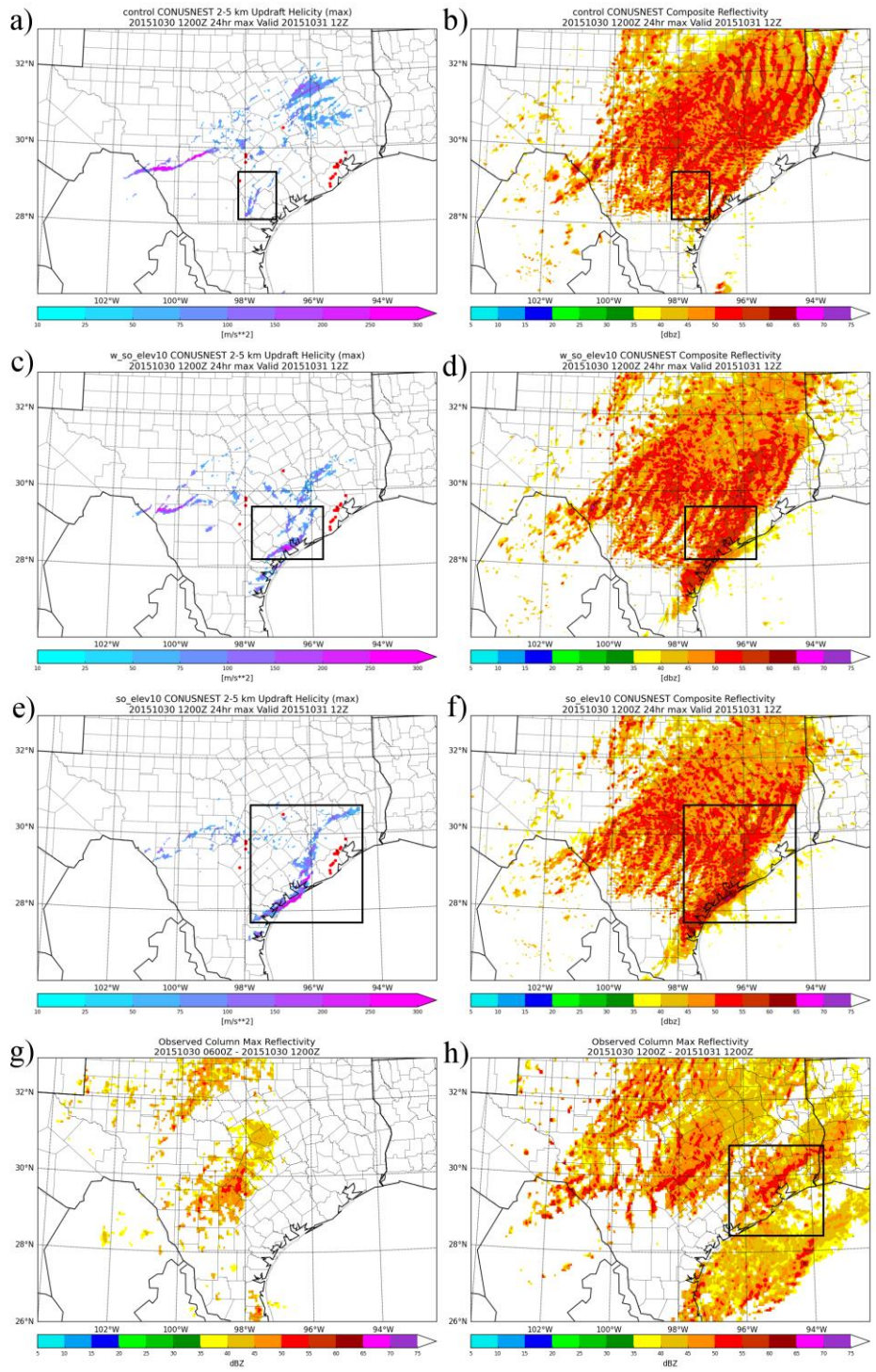

Fig. 22. The 1200 UTC initialized, 24-hour maximum updraft helicity (UH: masked below 50 m/s^2) and radar reflectivity (masked below 35 dBZ) valid at 1200 UTC 31 October 2015 for (a,b) control, (c,d) w_so_elev10, and (e,f) so_elev10 respectively. For reference, the max radar reflectivity for the (g) convectively active regions during the 6 hour data assimilation period (0600 UTC to 1200 UTC), and (h) the 24-hour forecast period are also shown. The updraft helicity that is boxed corresponds to the storms that developed at the intersection of the LLJ and warm front boundary. The updraft helicity that is boxed corresponds to the storms that developed at the intersection of the LLJ and warm front boundary (Fig. 20a) and the storms which produced tornadoes (tornado reports: red dots on (a), (c), and (e)).

2.4 Summary and Conclusions

The radial wind observation operator in the GSI was extended to include vertical velocity and testing associated with the refinement of the radial wind super-observation processing was performed in a single case study. To evaluate the impacts of this extension, experimental configurations using an hourly configuration of the NAMv4 data assimilation and model framework was used beginning at 0000 UTC 30 October and ran through 0000 UTC 31 October 2015. The data assimilation consists of a hybrid 3DEnVar data assimilation system which uses the hydrostatic GDAS ensemble Kalman filter (~35-km) for the ensemble contribution. The NAMv4 was run using a 12-km parent domain and 3-km CONUS nest domain with the verification run on a 5-km grid.

The experiments started with the control which mimics the operational configuration with relaxed quality control parameters for the radial wind observations. Incremental modifications to the experimental setup followed by adding vertical velocity to the forward operator, adjusting some of the super-observation parameters, and finally, adjusting the maximum elevation angle to 10°. The experiments were evaluated using objective analyses of accumulated precipitation using categorical and neighborhood verification metrics aggregated for 18-hour forecasts over the 0900, 1200, 1500, and 1800 UTC 30 October 2015 forecast cycles to accommodate the 3-hour precipitation observation buckets. A constant value of 0.7 $m^2 s^{-2}$ was used for the background error covariance for vertical velocity and a horizontal and vertical scale of influence of 27-km and 10-grid units respectively in the main set of experiments. These

values were chosen based on preliminary testing and tuning; however, future studies should generate a new background error covariance which includes vertical velocity.

The fit-to-observations were used to assess the impact to the data assimilation system between experiments with differences only related to the radial wind observation operator. For example, the RMS innovation from the GSI observation minus background and analysis for the control was compared with w_incl and experiment so_elev10 was compared with w_so_elev10. The RMS scores showed a neutral impact on the analysis for including vertical velocity in the observation operator.

Quantitative results for the *elev10* experiments showed slight improvements in the FSS of box size 60-km for accumulated precipitation at the 0.25 and 0.50 in. thresholds (Fig. 18) and in frequency bias at the 0.50 and 1.00 in. thresholds (Fig. 19). There was a slight degradation in the lower FBIAS thresholds. A summary of FSS statistics across a variety of thresholds and scales were also presented in a scorecard (Fig. 20). Results from the scorecard demonstrated that the strongest sensitivity was to the super-observation parameters where those experiments mostly demonstrated some (slight) improvement relative to the control. No results showed statistical significance.

Because the model is not directly forced by the analysis of vertical velocity, the results from these experiments can be attributed to two factors: 1) reducing the error associated with the original Doppler radar radial wind forward operator via the inclusion of vertical velocity and 2) refining the super-observation parameters and adjusting the max elevation angle. The most impactful changes resulting in forecast

improvement were associated with the refined super-observation procedure in conjunction with the adjustment to the max elevation angle.

The *elev10* experiments (w_so_elev10 and so_elev10) predicted a band of high precipitation accumulations along the Texas Gulf Coast that was not present in the control or w_incl experiment. This band of heavy precipitation corresponds to a simulated storm(s) with rotating updrafts. These model simulated storms matched closer to the location of observed tornado reports and hence storms with rotating updrafts but still displaced ~100-km to the west.

There is support that the *elev10* experiments showed the greatest improvements to the forecast of accumulated precipitation (Fig. 20) and demonstrated a more accurate forecast of updraft helicity swaths (Fig. 22) corresponding to tornado reports (Fig. 13). These findings suggest that convective-scale model forecasts are sensitive to the methods and settings for assimilating radial winds.

Testing in an ideal framework revealed that errors of up to 1 m/s could be expected for elevation angles around 10 degrees and vertical motions of 10 m/s (Fig. 6). However, the results of this study appear contradictory; the impact of adding vertical velocity to the forward operator was effectively negligible. This result is likely due to several factors which limits the generality of the findings. Such limiting factors likely include: the lack of a convective-scale ensemble in the hybrid 3DEnVar algorithm, the treatment of vertical velocity as a sink term in the analysis, limitations to a single case, and frequency of assimilation of radial wind observations.

A second set of experiments that featured vertical velocity in the observation operator and as a control variable were run using the same background error covariance

but a horizontal and vertical scale of influence of 100-km and 1 grid units respectively. This additional set of experiments showed mixed results and that the tuning of the background error covariance statistics plays a crucial role in obtaining positive forecast impacts even while using a non-hydrostatic model without a fully prognostic term for vertical velocity.

The results of this work are to be considered as first steps towards refining and improving the use of these abundant, convective-scale data. There are additional areas of convective-scale data assimilation to be explored in an operational context, such as the use of a convective-scale ensemble and a more impactful use of the analysis of vertical velocity instead of limiting it to a sink term in the observation operator. These experiments also only featured hourly data assimilation updates while trying to characterize features that evolve on sub-hourly timescales. Finally, future work toward expanding these tests to additional cases will likely yield more conclusive results.

Chapter 3: Impacts of Doppler Radial Wind Assimilation in the GFS with a Global Observing System Simulation Experiment (Lippi et al. 2022 in prep.)

3.1 Introduction and Background

More than 800 Doppler radars are registered in the World Meteorological Organization's world radar database [Available online: https://wrd.mgm.gov.tr/Home/Wrd]; however, no radial wind data is assimilated within the National Centers for Environmental Prediction's (NCEP) Global Forecast System (GFS). The assimilation of radial winds has been proven to be beneficial for improving the skill of regional, convection-permitting NWP model forecasts (Gao and Stensrud 2014; Gao et al. 2004; Johnson et al. 2015; Lippi et al. 2019; Xiao et al. 2005), but the value of their assimilation is not well known for global modeling systems (Sun 2005b). Radial winds, which are highly dense and frequent 3D observations, are not typically assimilated in global models (e.g., the NCEP GFS; Kleist et al. 2009b) due to the large mismatch between the spatiotemporal resolution of the observations and that of global modeling systems. Until recently, it has been impractical to run global systems at sufficiently high resolution and with a fine enough update frequency to make the best use of radial winds. Therefore, the need to assimilate such observations has not extended beyond those limited area systems for which NWP centers employ high resolution data assimilation. Furthermore, assimilation of radial winds has been limited by the lack of a service to collect and distribute the global dataset of radar observations. Unlike most other observing platforms, radial wind observations across the globe are

not typically available for assimilation beyond the region from which they originate with the exception of a few regional exchanges. For example, radar information is coordinated among the weather services of European countries via the Operational Program on the Exchange of Weather Radar Information (OPERA; Saltikoff et al. 2019). The U.S. also receives Canadian radars as well as radars from Puerto Rico, but ultimately, there is no global exchange of this data.

Table 6. Recent changes to the GFS

Year	GFS Version	Dynamic Core	Horizontal Resolution (control/ensemble)	Number of Vertical Levels	DA
2015	GFSv13	Global Spectral Model (GSM)	T1534 (13km)/ T574 (35km)	64	Hybrid 3DEnVar
2016	GFSv14	‘’	‘’	‘’	Hybrid 4DEnVar
2019	GFSv15	Finite-Volume Cubed Sphere (FV3)	C768 (13km)/ C384 (25km)	‘’	‘’

The operational GFS has recently undergone many upgrades that justify exploring the impacts of assimilating radial wind observations. Some of the major changes that are discussed are also summarized in Table 6. In 2015, the horizontal resolution was increased to T1534 (~13-km) and T574 (~35-km) for the deterministic GFS and ensemble Kalman Filter members, respectively (NWS 2014). This is a similar grid-spacing to the 12-km North American Mesoscale (NAM) forecast system, which first began assimilating radial winds in 2006. In 2016 the hybrid 3DEnVar data assimilation scheme was upgraded to a hybrid 4DEnVar algorithm; adding a temporal component to the analysis (Kleist and Ide 2015a, 2015b; Lorenc 2003; Wang and Lei 2014; Wu et al. 2017). In 2019, the GFS was upgraded to version 15 (Kleist et al. 2018; NWS 2019) which included replacing the global spectral model with the non-hydrostatic Finite-Volume Cubed-Sphere (FV3) dynamical core (Harris et al. 2020;

Harris and Lin 2013; Lin 2004; Putman and Lin 2007). The GFSv15 upgrade also included the introduction of a more sophisticated microphysics parameterization which expanded the number of prognostic cloud species from one (Zhao and Carr 1997) to five (Chen and Lin 2013; Zhou et al. 2019). During this implementation the deterministic GFS was maintained at roughly the same resolution but on the cubed sphere grid at C768 (~13-km), with the ensemble resolution upgraded to C384 (~25-km).

These improvements in physics, ensemble resolution, and data assimilation methods merit an evaluation of including higher-resolution sources of observational data. Historically, high spatiotemporal resolution observations in global NWP has long since been the primary domain of satellite radiances. However, while such observations are dense in time and space, they do not match the capabilities of radar which is able to observe fine scale structures of storms.

A major challenge that this work faces is the lack of routinely available Doppler radial wind observations from the global network of radars; therefore, an observing system simulation experiment (OSSE) methodology is adopted to assess the potential impact of assimilating these observations in a controlled manner. In an OSSE, prospective observations can be simulated from a nature run. For example, using the geographical information in the world radar database, it is possible to create a simulated version of the global radar network to test the potential usefulness of assimilating all radar observations in the GFS.

Keeping in mind the limitations of an OSSE, the goal of this study is to lay the groundwork for assimilating radial wind observations within the GFS by first assessing

the impact of assimilating radial wind observations from a network that is accessible today (i.e., the NEXRAD network). This framework is then extended to the potential scenario of having access to all radial wind observations worldwide (i.e., the GLOBAL network). Finally, a purely hypothetical scenario where the worldwide network was designed with highly uniform spatial coverage (i.e., hypothetical network; HYPO for short). This final experiment is designed to provide an upper limit on the impact, useful for assessing the relative values of the real networks. In each of these experiments, only radial wind observations are assimilated to simplify and isolate the impact of such observations.

In Section 3.2, the model and data assimilation systems are described along with a description of the OSSE design. The results are presented in Section 3.3 and conclusions with a short summary and discussion of future work is presented in Section 3.4.

3.2. Methods

3.2.1 Model and Data Assimilation Configuration

The GFSv15 was used as the basis for this study, allowing for some modifications to conserve both computational resources and improve model stability. For computational resources, the maximum forecast length was reduced from the operational 384-hour to 168-hour (1-week) forecast. For stability, the model time step was reduced (from 225s to 90s) to overcome imbalances associated with the assimilation of highly dense simulated radial wind observations. The same time step was used for all experiments; the longer time step (225s) was used only in the Nature

run. No other modifications were made relative to the operational configuration of the GFS. The system maintained the standard 6-hourly data assimilation cycle using the hybrid 4DEnVar framework issuing forecasts four times per day at the usual synoptic times of 0000, 0600, 1200, and 1800 UTC respectively. Each experiment was cycled over the 1-week period.

The data assimilation experiments begin 6 hours into the nature run (after the initial spin up period described in Section 3.2.2.3) and are performed using the Global Data Assimilation System (GDAS) cycling algorithm and the Gridpoint Statistical Interpolation system (GSI; Kleist et al. 2009b; Wu et al. 2002). The GSI is a variational analysis system formulated in model grid space and is used in many operational applications at NCEP, such as the North American Mesoscale Forecast System (Gustafsson et al. 2018; Rogers et al. 2017; Wu et al. 2017), the Global Forecast System (Kleist et al. 2009b), the Rapid and High-Resolution Rapid Refresh systems (Benjamin et al. 2016; Dowell et al. 2022; James et al. 2022), and the Real-Time Mesoscale Analysis system (De Pondeca et al. 2011).

The GSI is configured to be consistent with the GFSv15 which uses hybrid 4DEnVar (Kleist and Ide 2015a; Lorenc 2013; Wang et al. 2013; Wu et al. 2017) where the static covariance is implicitly blended with an ensemble covariance through the extended control variable method (Lorenc 2003; Wang 2010). An 80-member ensemble with dual resolution capabilities was used (C768 and C384 for the deterministic and ensemble resolution respectively) and updated from a serial square root filter form of the EnKF (Whitaker and Hamill 2002). The weighting between the static and the flow-dependent ensemble background error covariance (BEC) is set to

be 12.5% and 87.5% respectively. The static BEC is identical to that used in the operational GFSv15. The tangent linear normal mode constraint (TLNMC; Kleist et al. 2009a), implemented to help improve the balance of the initial conditions, is enabled during the GSI minimization for each of the data assimilation experiments. All the previously mentioned settings and parameters are consistent with the GSI configuration for GFSv15.

3.2.2 Observing System Simulation Experiment (OSSEs)

Because radial wind observations from the existing global network of radars is not currently available for this study, an OSSE methodology was used. Furthermore, the OSSE methodology also allows the possibility to test an idealized global network.

3.2.2.2 OSSE Design

The OSSE was designed to investigate the potential impact of assimilating radial wind observations from three configurations: the U.S. NEXRAD network, a global radar network based upon the known worldwide distribution of radars, and a hypothetical global radar network (Fig. 23).

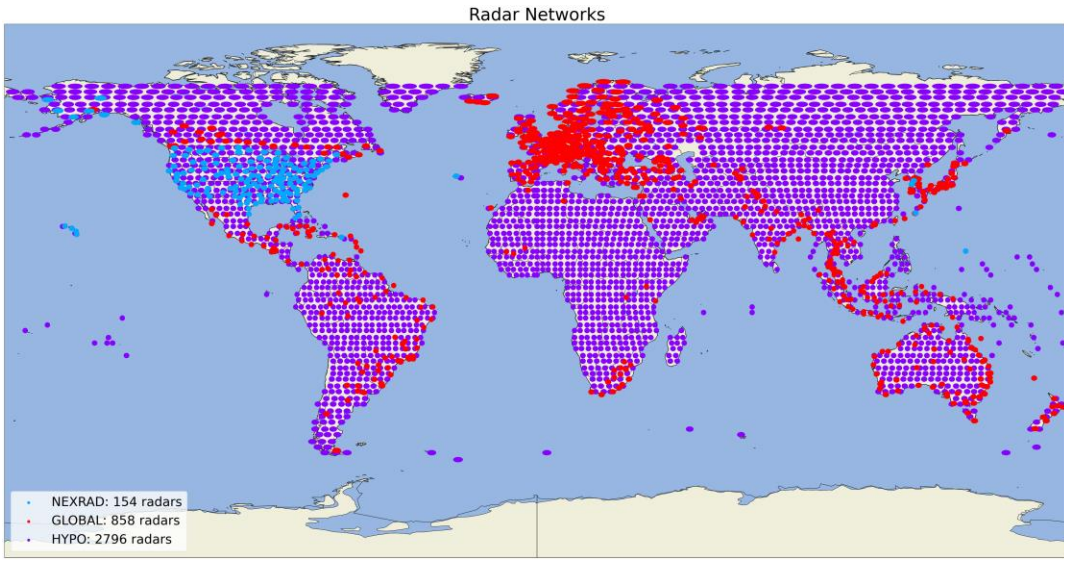

Fig. 23. Radar networks with 100-km radius circles including HYPO (purple), GLOBAL (red), and NEXRAD (blue). Note that the GLOBAL also includes all the NEXRAD radars.

For an OSSE (Errico and Privé 2018), a reference state or nature run is first generated by making a climatologically sound, free-running simulation using a reliable atmospheric model; the nature run is considered as a proxy truth against which subsequent assimilation experiments will be verified (e.g., Tong and Xue 2005; Xue et al. 2001). Simulated, imperfect observations from the observing systems are then generated from the nature run using the radial wind observation operator from the assimilation and adding random Gaussian, unbiased errors (e.g., Gao et al. 2004; Miller and Sun 2003; Xu and Gong 2003). These observations are then used in the data assimilation experiments where the impact of the simulated observations on the analyses and forecasts are assessed.

For simplicity, the same model is used for both the nature run and the experiments. While there are certain limitations to consider while using this identical twin approach, it is not uncommon practice (e.g., Gao et al. 2004; Tong and Xue 2005;

Xue et al. 2006; Xue et al. 2001) and can be useful if results are interpreted with the correct perspective. Therefore, one should consider the results from such an idealized design as a “best-case scenario” and a demonstration of potential plausibility.

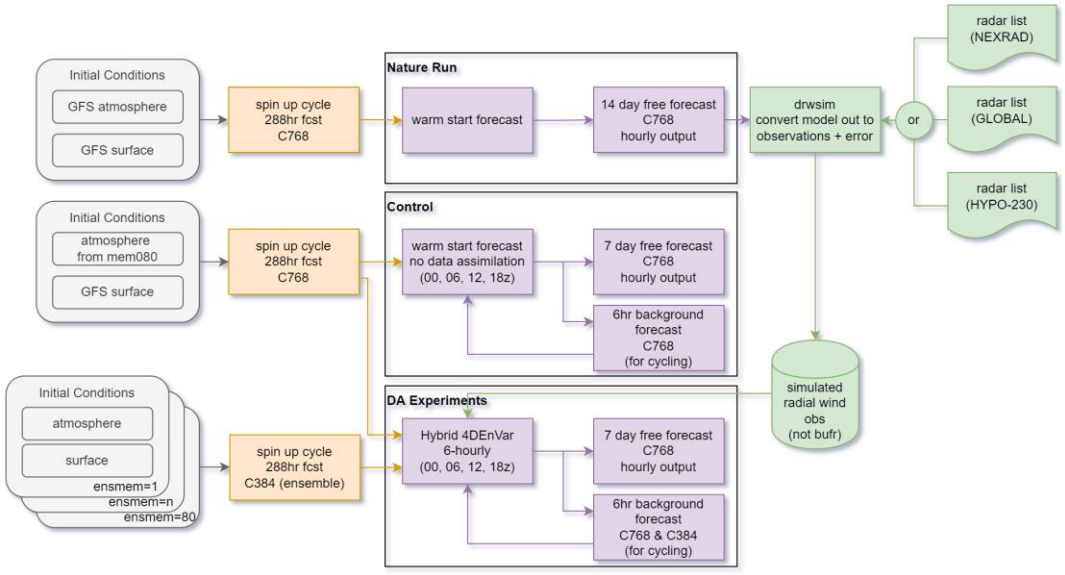

Fig. 24. Diagram describing the OSSE setup and the spin-up of initial conditions for the Nature and experiments. The spin-up cycle is highlighted in orange and represent the initial conditions after running the control and ensemble state for 288hrs.The forecast and data assimilation steps are shown in purple. The observation simulator and where those observations are used are highlighted in green.

3.2.2.3 Initial Conditions and Model Spin-Up

The model-spin up process is shown in the OSSE flowchart (Fig. 24). The perturbed initial conditions used for the control and all other experiments were generated by replacing the atmospheric state (lower boundary states were not replaced) with that of an arbitrary ensemble member from the operational GFS EnKF, in this case, the 80th member was used. The initial conditions for the nature run used the GFS control member atmospheric state.

Each set of initial conditions were then used to initialize a model free-forecast to determine the point at which the root mean squared difference (RMSD) of

temperature, specific humidity, and horizontal wind at 500-hPa (Fig. 25) between the two forecasts (the nature and perturbed initial conditions) saturates, i.e. when the RMSDs no longer continue to grow. This saturation criteria identifies the point at which the model has tracked to its climatology so that both simulations are in a balanced state. The RMSD of 500-hPa temperature, specific humidity, and horizontal wind saturate at approximately 288-hours for both C384 and C768 model resolutions. The 288-hour model state was then used as initial conditions for the OSSE experiments.

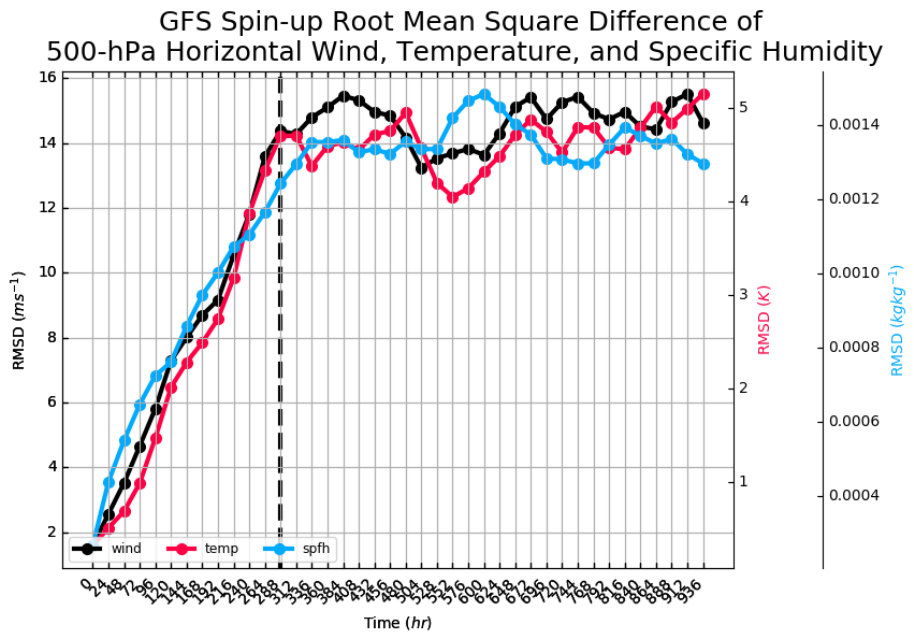

Fig. 25. Root mean square difference (RMSD) for 500-hPa wind (black), temperature (red), and specific humidity (blue) for model resolution C384 for the period starting at 0000 UTC on 11 September 2018. Similar results were found with C768 resolution (not shown). Hour 288 is highlighted by the vertical dashed line.

3.2.2.4 The Nature Run

The nature run is a 336-hour, deterministic, free-running simulation (starting after the 288-hour spin up period; Fig. 24) of the GFSv15, absent of any data assimilation, run at C768 (13-km) and 64 vertical levels with hourly output for the 1-

week period. Reproducing actual weather events is not the goal in this study as might be the case in some OSSE based studies (e.g., Ge et al. 2012). The initial spin up cycle was initialized with an actual GFS analysis as a convenient approach to get a realistic initial model state.

3.2.2.5 Simulation of Doppler Radial Wind Observations

A Doppler radial wind observation simulator[10] was developed based on the GSI radial wind observation operator, and is represented in Fig. 24 via the green box labeled "drwsim". The two main inputs are the nature run forecasts and a radar list containing location data for each radar network. For simplicity, all radars are considered to operate similarly to the radars of the NEXRAD network (e.g., S-band, 250-m gates, 1° azimuth; OFCM 2017 p. 2-22 V#99) but also imposed a maximum observation range (R_{max}) of 100-km which is consistent with convention for radial wind pre-processing for assimilation purposes (Alpert and Kumar 2007; Lippi et al. 2019). We chose 100-km under the assumption that all radars operate in a scan mode that uses a higher pulse repetition frequency, which yields a larger Nyquist velocity, yet a lower maximum unambiguous range. Thus, only data within the 100-km range is used to match real-world application of this data which is meant to avoid the decay of signal quality near the edge of a radial beam.

To obtain realistic simulated radial wind observations, observations were created only in locations where radar reflectivity exceeded a minimum threshold of 10

[10] https://github.com/delippi/drwsim

dBZ. This imposes a requirement that scatters of sufficient size be present. If using the Marshall Palmer Z-R relationship (Marshall and Palmer 1948)

$$Z=200R^{8/5} \tag{60}$$

where Z is the reflectivity factor ($mm^6\ m^{-1}$) and R is the rain rate ($mm\ h^{-1}$) and the conversion for Z to obtain the logarithmic reflectivity L_z (dBZ) is,

$$L_z=10log_{10}(Z), \tag{61}$$

then the relationship between dBZ and rain rate can be estimated as,

$$R=\left(\frac{10^{\frac{L_z}{10}}}{200}\right)^{\frac{5}{8}} \tag{62}$$

A reflectivity value of 10 dBZ approximately corresponds to a light mist with a rain rate of <0.01 in. h^{-1}. Furthermore, studies such as Tong and Xue (2005), Xue et al. (2006), and Liu et al. (2020a) have also used a 10 dBZ minimum threshold for detecting regions of precipitation for simulating radar observations. Therefore, in this study, a threshold of 10 dBZ is used as a proxy for determining regions with precipitation.

Radars have different scanning patterns (Volume Coverage Pattern; OFCM 2017) depending on the weather regime. The VCP 212 (OFCM 2017) is the typical radar scan pattern when precipitating regions exist. In our study, we assume all radars are operating in this mode, have Doppler capabilities, and operate at the S-band frequency such that signal attenuation can be ignored. The scan elevation angles of VCP 212 include 0.5, 0.9, 1.3, 1.8, 2.4, 3.1, 4.0, 5.1, 6.4, 8.0, 10.0, 12.5, 15.6, and 19.5 degrees. Clear air techniques for ground-based radar are not considered in this study.

The radial wind observations are simulated using the following equation:

$$\mathbf{V_r}(\theta,\alpha) = u\cos(\theta)\cos(\alpha) + v\sin(\theta)\cos(\alpha) + w\sin(\alpha) + R(0, 1) \quad (63)$$

where $\mathbf{V_r}$ is the radial wind observation, u and v are the nature run horizontal wind components, w represents the vertical wind, θ is 90° minus the azimuth angle of the radar, and α is the elevation (or tilt) angle of the radar. The formulation here uses the azimuthal directions based on the unit circle rather than the Cardinal Directions (i.e., 0° is East). The last term in Eq. (63), $R(0, 1)$, represents random, uncorrelated error drawn from a Gaussian distribution with a mean of 0 m s^{-1} and standard deviation of 1 m s^{-1}, which follows prior OSSE studies (Gao et al. 2004; Liu et al. 2020a; Snyder and Zhang 2003; Tong and Xue 2005; Xue et al. 2006). These are errors which may arise from representation errors (coarse model vs. high resolution observations and an imperfect observation operator) and measurement errors. In the case of this OSSE study, the radial wind observations are simulated from the nature run, which is an identical model and thus there are no errors of representation or of the observation operator; however, instrument measurement errors, are included by adding Gaussian, unbiased errors to the observations. For simplicity, observation errors are assumed uncorrelated.

Radial wind observations were then simulated using the settings found in Table 7. Simulated observations are constrained to the spatiotemporal resolution of the nature run; therefore, each simulated observation is thinned based on a prescribed three-dimensional box. The parameters include: azimuthal width ($\Delta\theta$), the elevation angle width ($\Delta\alpha$), and the gate spacing length (Δr) which describe the width, height, and length of the box respectively. The limiting parameters are the maximum elevation

angle (α_{max}) and maximum range (R_{max}) which put a maximum threshold for the elevation angle and maximum distance from the radar, respectively.

Table 7. List of parameters used for simulating the radial observations. Azimuth range, radial range, min/max elevation angle, and max observations range are as described in Lippi et al. (2019) (i.e., chapter 2 of this thesis) and as used in the GSI. The remaining parameters are specific to the observation simulator.

Azimuth range (degrees) $\Delta\theta$	**Elevation angle width (degrees)** $\Delta\propto$	**Gate spacing (meters)** Δr	**Min/Max elevation angle (degrees)** $\alpha_{max}/\alpha_{min}$	**Max observation range (meters)** R_{max}	**Ob error mean (m s^{-1})** μ	**Ob error standard deviation (m s^{-1})** σ	**Minimum dBz threshold (dBz)** dBz_{min}	**VCP**
1	0.5	13,000	0.5/20	100,000	0.0	1.0	10	212

The horizontal 2D spatial geometry for these observations was specified to be 1° by 13-km. The 13-km radial range was the finest resolution afforded by the nature run. The 1° azimuthal width is consistent with the configuration of the radial wind products of the NEXRAD network.

3.2.3 Experimental Design

To test the impact of assimilating Doppler radar radial winds, three DA experiments (see Fig. 24) were designed in addition to a control simulation (see Fig. 24) that does not assimilate any observations (NODA). Each of these is explained below and summarized in Table 8:

1. **NEXRAD**: A data assimilation experiment to assess the impact of assimilating radial wind observations from the only network that is accessible to the GFS today, the NEXRAD radar network (Fig. 23; blue).
2. **GLOBAL**: A data assimilation experiment to assess the impact of assimilating radial wind observations from the global network of radars (Fig. 23; red) as if

they were accessible to the GFS, the global network of currently deployed radars. This also includes the NEXRAD network over the U.S.

3. **HYPO**: A data assimilation experiment to assess the impact of assimilating radial wind observations from a hypothetical, optimally designed global network of radars (Fig. 23; purple) based on homogeneous spacing matching that of the average spacing for the NEXRAD network, which is 230-km (Huber and Trapp 2009). This experiment, while highly idealized, is designed to contextualize the results from the NEXRAD and GLOBAL experiments by providing an estimate for an upper limit on the impact of radial wind assimilation.

Table 8. List of FV3GFS radial wind OSSE experiments and their configurations.

Exp	ICs	Res.	Length (hours)	Cycled (hours)	DA	Ensemble	Time step (seconds) (C768/C384)
Nature	GFS	C768	336	N/A	N/A	N/A	225
NODA	Perturbed GFS	C768	168	6	None	N/A	90
NEXRAD	NODA	C768	168	6	Hybrid 4DEnVar	80 (C384)	90/150
GLOBAL	NODA	C768	168	6	Hybrid 4DEnVar	80 (C384)	90/150
HYPO	NODA	C768	168	6	Hybrid 4DEnVar	80 (C384)	90/150

3.2.4 Verification

Verification is performed using the Model Evaluation Tools (MET) verification software (Newman et al. 2022). The impacts of assimilating radial wind observations from the various network configurations are objectively assessed by calculating the fraction skill score (FSS) and frequency bias (FBIAS) of forecast quantitative precipitation as well as the bias corrected root mean square error (BCRMSE) and mean

error (bias) of geopotential height, temperature, horizontal wind, and surface pressure forecasts. The verification statistics were computed on a common, 0.125° grid over the following six verification domains: a) Northern Hemisphere (NHM), b) Tropics (TRP), c) Southern Hemisphere (SHM), d) Contiguous U.S. (CONUS), e) Atlantic Ocean (ATL), and f) Europe (EUR) (Fig. 26). All variables and metrics were evaluated via scorecards (not shown) for each verification domain; the major results are summarized into series plots. A bootstrapping technique, using 2000 replications with replacement, was used to test for statistical significance at the 95% confidence interval for series plots.

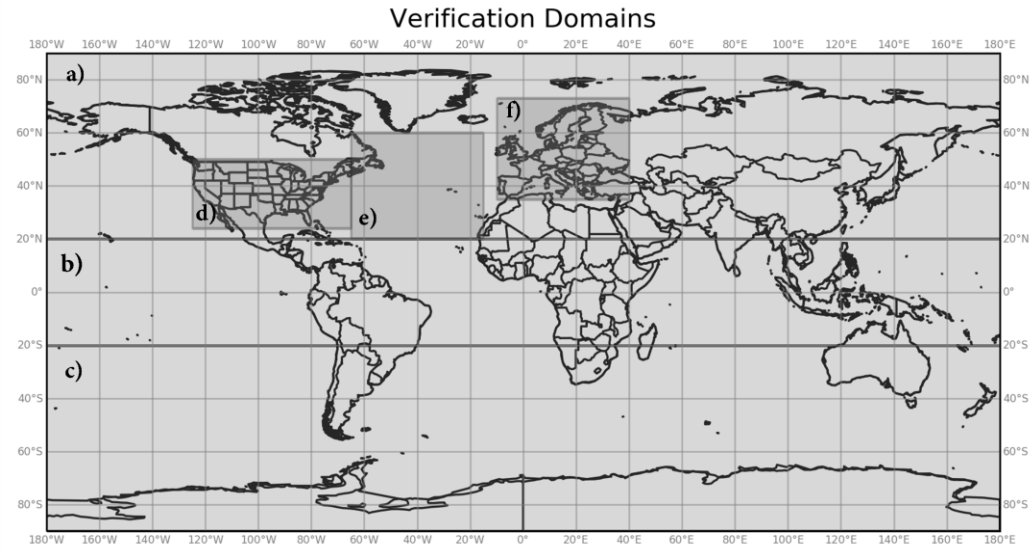

Fig. 26. Verification domains denoted by the gray shaded areas: a) Northern Hemisphere (NHM), b) Tropics (TRP), c) Southern Hemisphere (SHM), d) Continuous U.S. (CONUS), e) Atlantic Ocean (ATL), and f) Europe (EUR).

FBIAS is a categorical metric and is computed as the ratio of "yes" forecasts to "yes" observations. FBIAS greater than 1 indicates more points predicting a given threshold compared to observations (a wet bias). FBIAS less than 1 indicates fewer

points predicting a given threshold compared to observations (a dry bias). FBIAS is calculated in the following manner

$$\text{FBIAS} = \frac{hits + false\ alarms}{hits + misses} \tag{64}$$

		Observed	
		Yes	No
Forecast	Yes	*hit*	*false alarm*
	No	*miss*	*correct rejection*

Fig. 27. 2x2 contingency table which shows the relationship between events for a dichotomous verification situation (Wilks 2011)

where each of the terms in Eq. (64) are described in Fig. 27. A *hit* corresponds to a scenario where the forecast correctly matches the observations. A *miss* corresponds to a "no" forecast event occurred where there was an observed event. A *false alarm* corresponds to an event that was forecast but did not occur in reality. A *correct rejection* corresponds to an event that was correctly forecast to not occur.

The FSS is a neighborhood verification approach which relaxes the requirement for forecast and observed events to match exactly at the grid scale. Instead, the fractional coverage of predicted and observed grid-point events above a specified threshold are compared over a range of increasingly large box widths (Roberts and Lean 2008). Thus, FSS reveals how well the forecast resembles the observations at a given spatial scale. The FSS is computed in the following manner

$$\text{FSS} = 1 - \frac{\frac{1}{N}\sum_N (P_f - P_o)^2}{\frac{1}{N}\left[\sum_N P_f^2 + \sum_N P_o^2\right]} = 1 - \frac{\text{FBS}}{\text{FBS}_{worst}} \tag{65}$$

where N is the number of grid-points contained within the neighborhood area; P_f is the fractional coverage of forecast events that exceed a pre-determined threshold; and P_o is the fractional coverage of observed events that exceed the threshold. The fractional values (P_f and P_o) are what is computed at the increasingly large box widths. The FSS ranges from 0 to 1 where 0 would be a complete mismatch and 1 would be a perfect match of forecast events to observed events. FSS at a constant box width can also be used to verify the forecast similarly to the Equitable Threat Score (ETS or Gilbert's Skill Score; Wilks 2011).

Bias corrected root mean squared error (BCRMSE) is the standard deviation of the forecast errors that is not accounted for by the bias (mean error) and is computed in the following manner in the MET verification software by taking the square root of the bias corrected mean squared error (BCMSE)

$$\text{BCRMSE} = \sqrt{BCMSE} = \sqrt{RMSE^2 - bias^2} = \overline{(f-t)^2} - \overline{f-t}^2 \qquad (66)$$

where f is the experimental forecast, t is the truth from the nature run and the over bars denote an arithmetic mean.

3.3 Results

To determine how well each experiment performed in this OSSE, we consider the FSS and FBIAS of forecast precipitation as well as the BCRMSE and bias of 500-hPa geopotential heights. Other fields including wind, temperature, and surface pressure were also verified and generally showed similar results, so, for simplicity, those results are not shown.

3.3.1 FSS and FBIAS of Precipitation

The FSS of precipitation over increasingly longer lead times (i.e., days 1-7 in Fig. 28 as indicated by the vertical black lines) shows that, in this OSSE, the benefits of assimilating radial winds can persist for several days especially for the experiments with networks with broader coverage (e.g., ~5 days over the Northern Hemisphere for the HYPO and GLOBAL experiments). Since the impacts are strongest and clearest at shorter lead times, the day 1 forecast (i.e., average of the 1-24-hour forecast) results will be the focus of the remainder of this study.

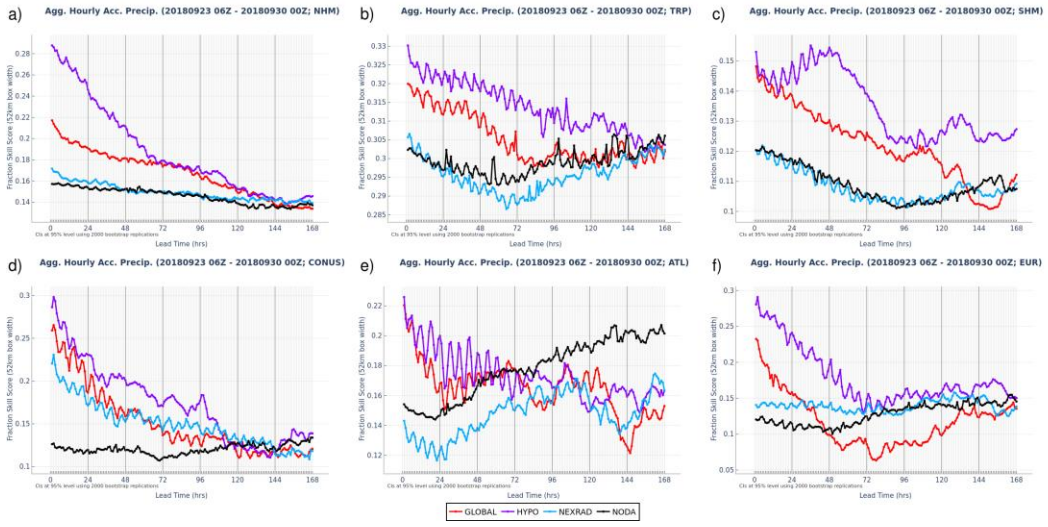

Fig. 28. FSS at 52km box width of the hourly accumulated precipitation as a function of forecast lead time over the 5 verification domains (Fig. 26). The vertical black lines indicate 24-hour periods.

The FSS difference (relative to NODA) of day 1 aggregated hourly accumulated precipitation at 52-km box width over each of the verification domains are shown in Fig. 29 and the corresponding FBIAS in Fig. 30. Fig. 29 reveals that the assimilation of only NEXRAD is modestly helpful in the Northern Hemisphere especially over the CONUS and Europe; however, it is not necessarily universally helpful such as over the Atlantic Ocean (Fig. 29e). This suggests that network design and spatial uniformity

have downstream impacts – a topic which is explored more completely in Section 3.3.3. Furthermore, the NEXRAD experiment was the only experiment that did *not* produce statistically significant improvements over the Southern Hemisphere and in the Tropics which is consistent with the fact that there are no NEXRAD radars in those regions nor are there radars upstream from those regions. It also shows that the GLOBAL experiment has significant improvements over the NEXRAD for all domains. Therefore, future implementation of radial wind assimilation in a global system might be met with greater success if the full existing global network of radars is used instead of those just over the U.S. Finally, the GLOBAL often underperforms compared to the HYPO experiment by as much as a 24% difference at short forecast leads over the Northern Hemisphere (Fig. 29a). The HYPO was expected to perform better since the network was designed with greater spatial uniformity and coverage and therefore shows that the existing global network of radars could also be further enhanced providing additional benefit.

The FBIAS of day 1 (1-24-hour) precipitation is shown in Fig. 30. In the context of an identical twin OSSE (although a reduced time step was used), it would be expected that there is no bias in the FBIAS; however, this is not the case. This discrepancy is thought to be due to initial imbalances within the assimilation which manifest more prominently in the shorter forecast leads (e.g., day 1 FBIAS scores). Recall that the time step was reduced for the data assimilation experiments due to imbalances in the model. Reducing the timestep did not remove those initial imbalances but allowed the model to better adjust to those imbalances without causing numerical instability in the model. Over longer forecast leads and larger verification domains the

FBIAS is relatively unbiased (not shown). At the shorter forecast leads, the lower FBIAS thresholds generally have good agreement with the truth.

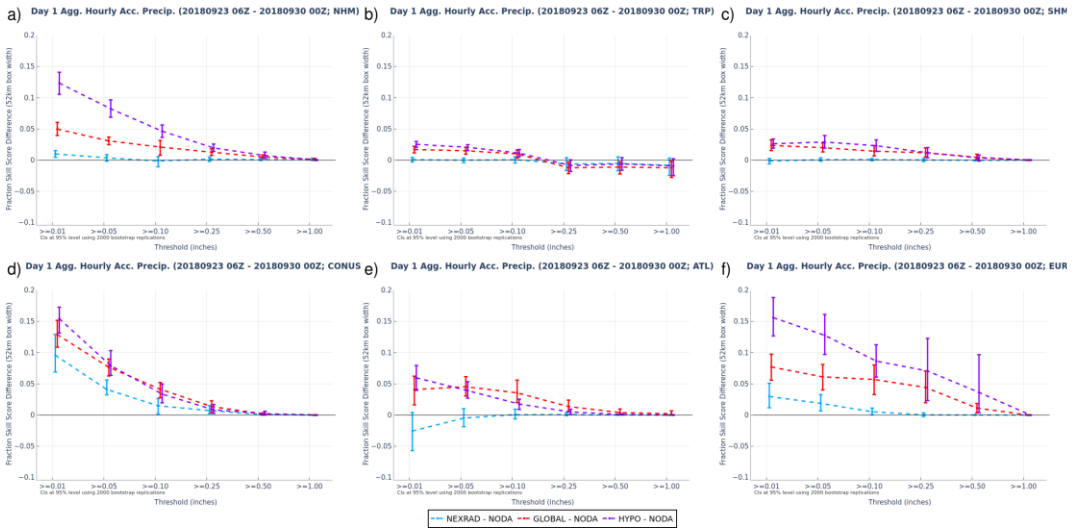

Fig. 29. FSS difference (relative to NODA) of Day 1 aggregated hourly accumulated precipitation at 52-km box sizes over the 1-week cycled period over the 5 verification domains (Fig. 26). FSS difference scores greater than zero indicate the respective experiment performed better than NODA.

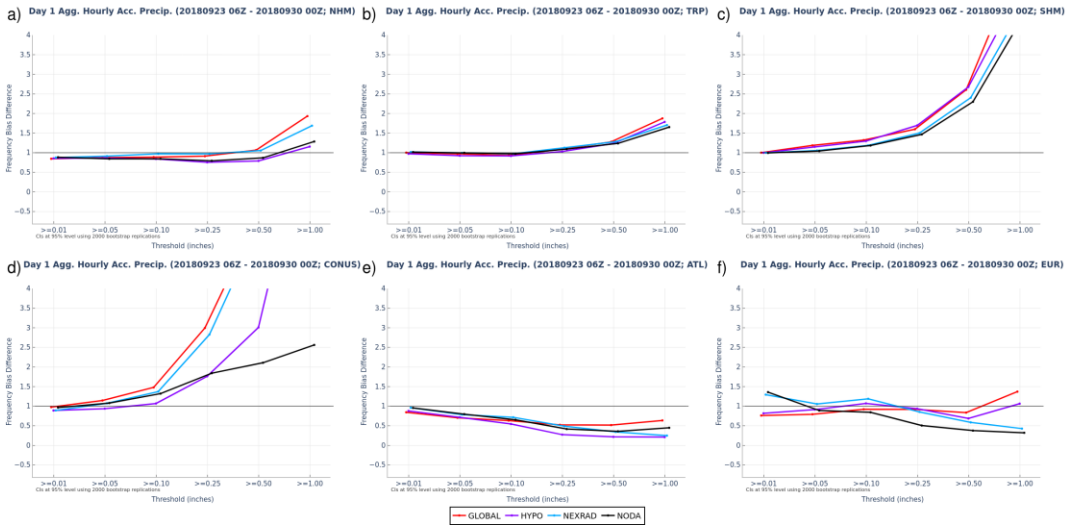

Fig. 30. FBIAS of Day 1 aggregated hourly accumulated precipitation over the 1-week cycled period over the 5 verification domains (Fig. 26).

3.3.2 BCRMSE of 500hPa Geopotential Height

To determine the potential impacts of assimilating observations from the different radar networks, the day 1 500-hPa geopotential height BCRMSE (Fig. 31) were compared as a function of cycle time (i.e., forecast initialization time).

Over the Northern Hemisphere (Fig. 31a), the NEXRAD experiment has little change relative to the NODA while the GLOBAL and HYPO experiments have a significant reduction in BCRMSE in 500-hPa geopotential height, especially during the later cycles. The improvements for the GLOBAL and HYPO experiments can be attributed to their larger area of observational coverage relative to NEXRAD. Fig. 32 shows the observation counts (total, Northern Hemisphere, and Southern Hemisphere) for each of the different experiments as a function of assimilation cycle. On average, the NEXRAD, GLOBAL, and HYPO experiments assimilate roughly 1×10^6, 3.5×10^6, and 15×10^6 observations per assimilation cycle respectively with most of those observations coming from the Northern Hemisphere. To reiterate, it is not only that the GLOBAL and HYPO have more observations, but that those networks have greater spatial coverage.

Over the Tropics (Fig. 31b) there are slight degradations among all experiments (~0-50m). The NEXRAD experiment shows the least amount of change here which is consistent with the fact that most NEXRAD radars reside north of the Tropical verification domain and thus should not have much impact on the forecast there. Both GLOBAL and HYPO experiments show larger BCRMSE in the Tropical verification domain than the NODA. This results is currently not well understood, but it is suspected that the limited sampling of the convectively environment of the tropics might be

triggering area of convection due to mass convergence/divergence. It would be important to further investigate the impacts of radial wind assimilation in the tropics using real-data experiments.

Over the Southern Hemisphere (Fig. 31c), there is virtually no change for the NEXRAD experiment relative to the control in BCRMSE 500-hPa geopotential height. This is consistent since there are no NEXRAD radar sites in the Southern Hemisphere. There are slight improvements for the GLOBAL and HYPO in the second half of the period; note the observation counts (Fig. 32; dotted lines) for the Southern Hemisphere for each experiment. The relatively smaller impact and lower observation counts in the Southern Hemisphere, relative to what was found in the Northern Hemisphere, is due to the smaller percentage of landmass available for ground-based observing networks (see Fig. 23).

Over the CONUS (Fig. 31d), all experiments have similar performance with small improvements (less than about 50 m height differences) prior to 26 September 2018 0600 UTC and larger improvements begin after this period (100-150 m height differences). It is expected that each data assimilation experiment performed with similar skill over the CONUS since the GLOBAL and NEXRAD coverage is identical, and the HYPO network was based on the average spacing of the NEXRAD network but are evenly spaced.

Over the Atlantic Ocean domain (Fig. 31e), there is improvement in the earlier cycles among all experiments but the NEXRAD and GLOBAL exhibit an unexpected, anomalous spike in forecast error peaking around 29 Sept 2018 0600 UTC (as indicated by the vertical line); a brief discussion will follow on this point in section 3.3.3.

Finally, over Europe, all experiments show some level of improvement. During the early cycles there are comparatively smaller differences (0-25 m height differences) between the NODA and each experiment. It takes about 5-6 assimilation cycles for improvements to appear in the NEXRAD experiment whereas improvements are immediate and large (~150 m height differences) in the GLOBAL and HYPO experiments. The improvement in the NEXRAD experiment is evidence of upstream DA impacting the time lagged downstream forecast, hence why we observe a time lag of 5-6 cycles before improvements are detected.

The positive downstream impacts over Europe and the negative impacts over the north Atlantic could be further investigated by running additional, radial-wind only experiments over a longer period to give more robust results as the results from this study are based on a limited sample size.

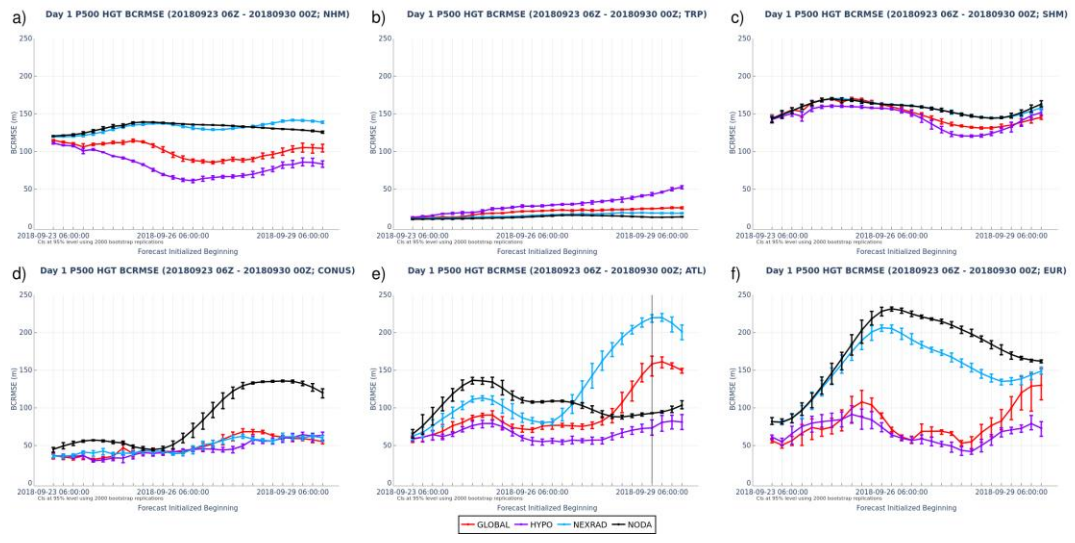

Fig. 31. BCRMSE of day 1 500 hPa geopotential height forecast by forecast initialization time over the 5 verification domains (Fig. 26).

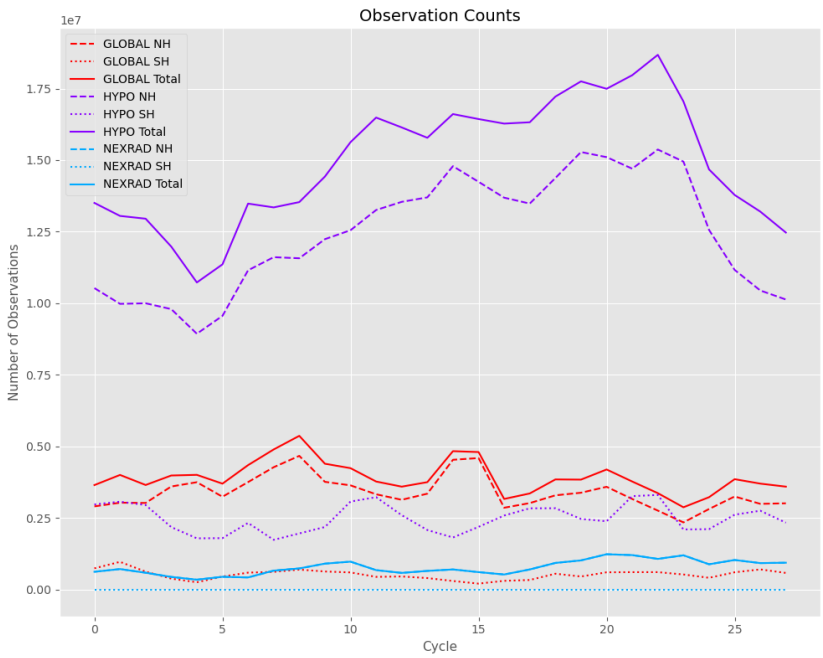

Fig. 32. Observation counts from the NEXRAD (blue), GLOBAL (red), and HYPO (purple) experiments. The counts for each experiment are also given for the Northern Hemisphere (dashed), Southern Hemisphere (dotted), and total (solid) regions.

3.3.3 Error Growth Over the Atlantic

The day 1 BCRMSE for the 500-hPa geopotential heights are plotted (Fig. 31e) as a function of forecast initialization time for each of the different experiments over the Atlantic Ocean. During the first 18 cycles the NEXRAD experiment, as expected, performed better than the control, but then by 29 September 2018 0600 UTC the NEXRAD 500-hPa geopotential height BCRMSE sharply rises to more than 200 m (more than twice that of the control; ~100m). Similarly, in the GLOBAL experiment, the 500-hPa geopotential height BCRMSE also sharply rises to more than 150 m which peaks roughly at the same time to that of the NEXRAD, but the upward trend in error does not occur until 4-5 cycles (24-30hrs) later. The 6-hr geopotential height forecast from the 0600 UTC cycle on 29 September 2018 (the 25th DA cycle; vertical line in

Fig. 31e) is shown comparing the NODA (Fig. 33a), NEXRAD (Fig. 33b), GLOBAL (Fig. 33c), and the HYPO (Fig. 33d) experiments all against the NATURE. The NEXRAD and GLOBAL experiments generate a deep upper-level low positioned over the northeast corner of the Atlantic verification domain and noted by the large negative height differences relative to NATURE (blue shading, Fig. 33b, c). Neither the NODA nor the HYPO experiments featured this large error growth.

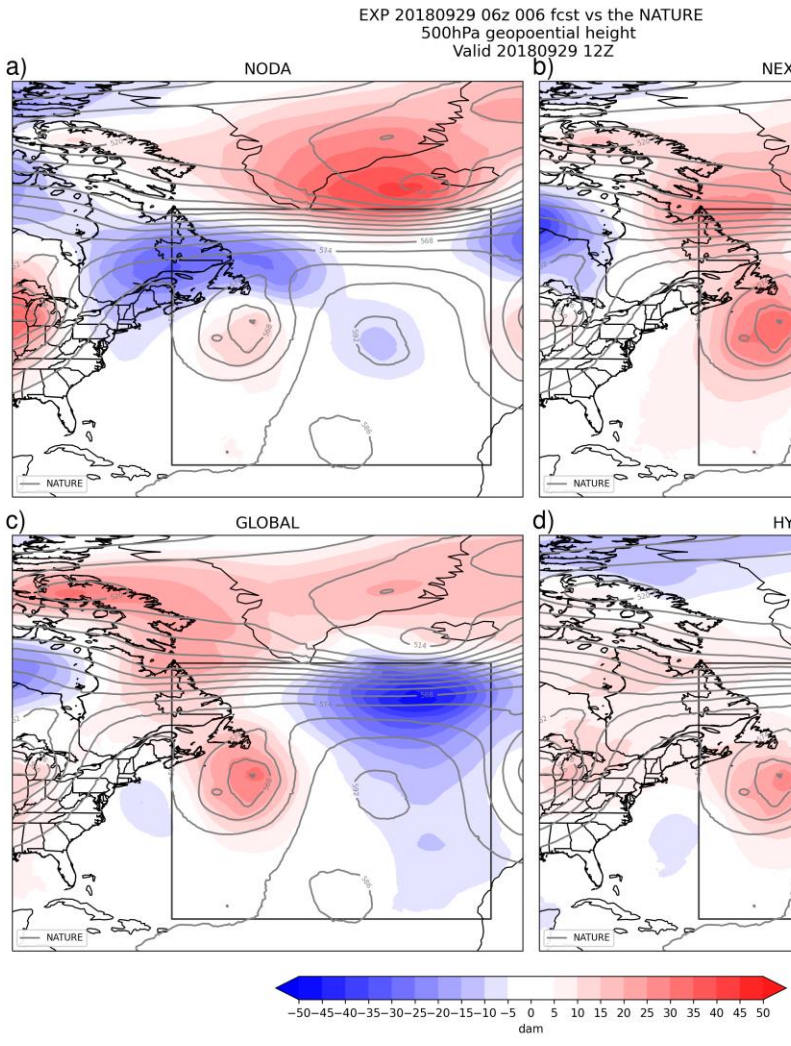

Fig. 33. Comparison of a 6-hr, 500-hPa geopotential height forecast for (a) NODA, (b) NEXRAD, (c) GLOBAL, and (d) HYPO vs the NATURE (gray contours) from the 25th DA cycle denoted by the vertical line in Fig. 31e. Geopotential height differences are denoted by the red/blue shading where blue indicates where the respective experiment has lower geopotential heights relative to the NATURE. The Atlantic Ocean verification domain is indicated by the black square.

We hypothesize that the NEXRAD and GLOBAL experiments have this large error growth for the following reasons: 1) the discrete nature of the radar network creates the opportunity for partial sampling to occur, especially along coastlines, and thus the possibility of aliasing of the atmospheric state when these observations are assimilated and 2) they lack the observational coverage that the HYPO network provides (even though HYPO also suffers from the first point).

Aliasing occurs when only a portion of a signal is projected onto a pattern that is not fully observed. This issue of aliasing has also been noted in a similar study (Djalalova et al. 2016) and is analogous to what we have found here. Djalalova et al. (2016) tested the effectiveness of assimilating winds from a network of onshore wind profiling radars (WPRs) located in the Northeastern U.S. In this study, the authors found that assimilating observations from these coastal WPRs generally improved downstream forecasts but had one case that showed clear degradation. Upon further analysis the authors found the degradation was due to aliasing. The similar network configuration and flow regime in our case lends us to hypothesize that we see a similar phenomenon.

Benjamin et al. (2004) also noted that aliasing is a possibility when using any discrete observational network. Therefore, boundaries in the radar (or any) observing network due to natural (i.e., coasts) or cultural (i.e., undeveloped regions with no explicit need for radar observations) can produce areas that are potentially prone to aliasing. What we have found is very similar to that found in Djalalova et al. (2016), having a surface low pressure near the Atlantic Coast of the U.S. with a land confined observing network (of wind). The HYPO network, which also has discrete boundaries,

likely avoided the large error over the Atlantic because it has better spatial coverage (e.g., Fig. 23). For this specific case, radars located over Nova Scotia and over Greenland were crucial to the success of the HYPO experiment.

The 10-m wind (black wind barbs) 6-hour forecasts from the NEXRAD (a) and HYPO (b) experiments are compared against the NATURE (c) at the same valid time (Fig. 34). It is at the time shown in which a low initially formed off the mid-Atlantic Coast which led to the large degradation over the Atlantic verification domain. Several variables are shown in each plot and include: the 10-m wind speed (black wind barbs), surface pressure, reflectivity (< 10dBZ), the 100-km radar range rings, and lowest level wind increments (red barbs) with the magnitude also shaded in gray. Superimposed on the NEXRAD and NATURE experiments are the corresponding storm tracks for either scenario (gray line). They HYPO experiment did not exhibit a low forming off the coast at this time.

At the time shown in Fig. 34, the main low in the NATURE is located east of Nova Scotia with a secondary low beginning to develop off the mid-Atlantic coast (as indicated by the gray storm track line; the red dot identifies the center location of the low at the corresponding valid time). The NEXRAD experiment produced low level wind increments of about 10 m s^{-1} in magnitude of offshore flow in Fig. 34a enhancing the cyclonic flow and aiding in the development of this surface low pressure. The low that develops in the NEXRAD tracks northeastward and northward following the gray line whereas the low in the NATURE tracks more eastward during the same period. Since, in the NATURE, the low tracks out of range of the radars, there are no radars to constrain the currently formed low in the NEXRAD experiment. On the other hand, the

HYPO experiment does not generate strong offshore increments, nor does it generate a low off the mid-Atlantic coast. The HYPO is able observe the main low in the NATURE as is traverses Nova Scotia noticeable by the large increments in Fig. 34b.

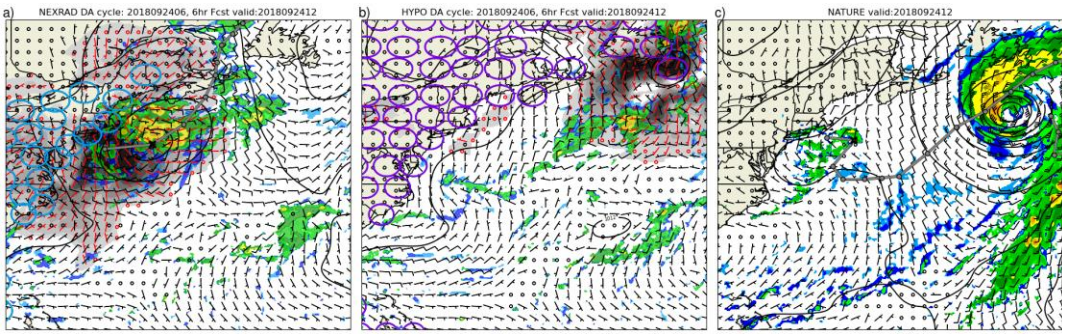

Fig. 34. A comparison of the 6-hour 10m wind forecast (black barbs) for the NEXRAD (a) and HYPO (b) against the NATURE (c) at a time when the error growth begins in the NEXRAD experiment. The 100-km radar rings for the NEXRAD and HYPO networks are both shown. In addition, the DA experiments have the wind increments (red wind barbs) with the magnitude of the wind speed increment shaded in gray. Storm tracks (gray lines) are superimposed on the NEXRAD and NATURE experiment noting the location of the storm forming off the mid-Atlantic Coast. The current storm location is noted by a black x.

Ultimately, spatial design of the observing network is important and brings into view the deficiencies in our data assimilation. For example, a larger ensemble could be one element in our data assimilation system that may allow for better utilization of observations of partially sampled features. However, in real-data studies, the full suite of observations will be available and would likely be less susceptible to such issues but is an important issue to document. Furthermore, we suspect that these issues will not be fully mitigated by the use of the full suite of observations, and hence, it was a useful exercise to perform an OSSE experiment testing the assimilation of only radial wind data which might have been obscured by other observations.

3.4 Summary and Conclusion and Future Work

This work is motivated by several upgrades over recent years that warrant the consideration for using radial winds within the global paradigm. Some of these improvements include the upgrades in DA methods, replacing dynamic cores, and increasing model resolution. While the data from the global network of radars does currently exist, it is not currently accessible for use by any individual system. Therefore, an OSSE was used to evaluate the potential impacts of assimilating Doppler radial winds from the global network of radars (as opposed to only the NEXRAD network) in the GFS.

Six verification domains were created, including the Northern Hemisphere, Tropics, Southern Hemisphere, CONUS, Atlantic Ocean, and Europe to test the impacts of assimilating radial winds. The Atlantic Ocean (for all experiments) and European (for only the NEXRAD experiment) were intended to test the impacts of upstream DA on the downstream forecast. The forecasts were verified, and their relative performance was evaluated using Fraction Skill Score (FSS) and frequency bias (FBIAS) of precipitation as well as the bias corrected root mean squared error (BCRMSE).

Three data assimilation experiments were designed each using a different radar network. The HYPO experiment used a hypothetical radar network and was used as an upper limit for assessing the real radar networks. Two real radar networks were tested: the NEXRAD which is currently the only network available to the U.S. for operational use and the full GLOBAL network of radars registered within the World Radar

Database (https://wrd.mgm.gov.tr/Home/Wrd). A control experiment was also created which did not assimilate any observations (NODA).

The HYPO experiment performed the best and generally set a clear upper limit to the benefits that could be potentially gained from assimilating radial wind observations. On a global average, in terms of day 1 accumulated precipitation FSS scores, the NEXRAD performed about 1% better than NODA, the GLOBAL experiment performed 13% better than NEXRAD, and the HYPO experiment performed 13% better than the GLOBAL experiment. These relative performances will change depending on the verification region. There were clear improvements for the NEXRAD data especially over the CONUS but also over the European verification domain. While there was not much difference between the NEXRAD and GLOBAL experiments over the CONUS verification domain, there is evidence that upstream DA impacts the downstream forecast. In both cases of the real radar networks, there was downstream degradation of the forecast over the Atlantic Ocean verification domain; however, the severity of this degradation is not expected when using the full suite of observations. Additional testing would be needed to assess the relative magnitude of which radar locations (outside the NEXRAD network) would make the greatest impacts on the downstream forecast over the CONUS. For U.S. landfalling tropical cyclone forecasts, the Caribbean radar network would be likely most useful.

Based on the results of this study, there is support for testing Doppler radar radial wind assimilation in the GFS in real-data experiments. Future, real-data experiments will include testing the assimilation of radial winds using the most recent version of the GFS (i.e., GFSv16; NWS 2021) and using the hourly GFS system

workflow (Slivinski et al. 2022). It is hypothesized that the hourly GFS could make better use of observations, especially high spatiotemporal observations like radial winds in global NWP, by updating with observations more frequently such as is done in convective-scale NWP models. It is hypothesized that an hourly-updating global system would be able to better constrain rapidly evolving systems such as hurricanes and convective storms and make better use of high temporal observations including observations such as radial winds. More on this project and initial tests using radial winds will follow in Chapter 4.

Chapter 4: Overlapping Windows in a Global Hourly Data Assimilation System (Slivinski et al. 2022) and a Radial Wind Assimilation Experiment

This chapter is based on the "Overlapping Windows in a Global Hourly Data Assimilation System" published in the Monthly Weather Review as a co-author as (Slivinski et al. 2022) and many of the ideas and motivations therein have been at least partially reproduced here with the permission of the copyright holder and the lead authors. Modifications were also made to adapt this work to fit within the radial wind theme of this dissertation.

Slivinski, L. C., D. E. Lippi, J. S. Whitaker, G. Ge, J. R. Carley, C. R. Alexander, and G. P. Compo, 2022: Overlapping Windows in a Global Hourly Data Assimilation System. *Monthly Weather Review*, doi:10.1175/mwr-d-21-0214.1.

As previously mentioned in this dissertation, future scientific advancement of the GFS will warrant the assimilation of radial winds. One such example includes a collaborative effort between EMC and ESRL/PSL, ESRL/GSL and JCSDA which is currently underway to develop and test a Global Rapid Refresh Forecast System (GRR). The proposed GRR would feature an hourly-updating cadence as opposed to the 6-hourly update that is currently used in operations. It is hypothesized that an hourly-updating global system would be able to better constrain rapidly evolving systems such as hurricanes and convective storms and make better use of high temporal resolution observations, such as radial winds. This chapter will discuss the progress of

the GRR system based on Slivinski et al. (2022) and will also include additional results of assimilating radial wind during a landfalling tropical cyclone (TC) event.

4.1 Introduction and Background

The GFS currently uses a 6-hour data assimilation window, updating the state with observations to produce analyses every 6 hours. As of the version 16 upgrade in 2021, the ensemble part of the hybrid 4DEnVar algorithm has been updated to use the gain form of the local ensemble transform Kalman Filter (LETKF; Hunt et al. 2007) with a modulated ensemble (Lei et al. 2018) and employing 4D-Incremental Analysis Updates (4DIAU; Lei and Whitaker 2016). The purpose of the 4DIAU is to suppress unwanted imbalance in the analysis by applying the analysis increments as forcing throughout the duration of the assimilation window as opposed to applying the full increment to the analysis all at once. This system is currently responsible for providing boundary conditions for regional systems such as the hourly cycling Rapid Refresh (RAP; Benjamin et al. 2016) which subsequently provides initial and boundary conditions to other systems such as the High Resolution Rapid Refresh (HRRR; Dowell et al. 2022; James et al. 2022). The RAP/HRRR systems both use a 1-hourly data assimilation window updating the model state with observations every 1 hour which is necessary for resolving the convective scales.

Positional errors in the background of rapidly evolving systems, such as track errors in TCs especially in data sparse regions, can severely degrade the analysis of those features if the track errors are approximately the representative scale of the feature itself (Lawson and Hansen 2005). The 6-hourly assimilation window, in some cases, can be found to be too long to constrain the position of TCs, so there have been

many ad hoc techniques that have been developed to help combat this. For instance, the operational Hurricane Weather Research and Forecasting system (HWRF; Biswas et al. 2018) uses a TC initialization technique which includes three parts: a vortex relocation procedure which assimilates the official advisory minimum sea level pressure observation (Kleist 2011); a storm size correction which uses the radius of maximum surface wind speed, radius of 34-kt winds, and radius of outermost closed isobar; and a storm intensity correction which uses maximum surface wind speed and minimum sea level pressure (Biswas et al. 2018). While these techniques are important to the quality of TC forecasts in the current paradigm, it is preferred to not need any external, ad hoc component to the assimilation of real observations to maintain accurate forecasts. Updating more frequently will provide shorter background forecasts, and thus less impacted by aspects of nonlinearity, to feed into each subsequent cycle, potentially using observations more effectively by assimilating them as soon as they become available. Hourly updates would also provide six times more frequent updates so that forecasters and stakeholders can always have the best and most up-to-date global weather information.

Many WoF projects (i.e., $\leq$ 1 hourly update cadences) have also found a similar issue of biases in storm motion in the background fields (e.g., Flora et al. 2019 and references therein). These biases, or alignment errors of "features" (i.e., TCs, super-cell thunderstorms, etc), are the result of model error which are nonlinear and non-Gaussian thereby reducing the ability for the variational and ensemble based DA methods (which are based on Gaussian and linear assumptions) to perform optimally. EnKF and variational methods are all based on the basic assumptions of Gaussian

background and observation errors and linear observation operators; therefore, these methods only work optimally when the background and observation errors are Gaussian and the model state variables and observed quantities are linearly related. For example, Chen and Snyder (2007) showed that updating the TC vortex location can be done effectively using an EnKF when the original displacement error of the vortex is small (i.e., approximately Gaussian prior errors). They found that when the initial displacement error is comparable to or larger than the vortex size (i.e., non-Gaussian prior errors), the EnKF performs poorly and can even result in creating an additional vortex. One solution to this problem is a feature alignment technique (FAT; Stratman and Potvin 2022) which is based on similar techniques developed by Nehrkorn et al. (2014) and Nehrkorn et al. (2015). Essentially, FAT finds the optimal field of 2D displacement vectors that align the background to the observations resulting in smaller background errors and making the problem less nonlinear and non-Gaussian. Such techniques are not currently employed in operational systems at NCEP, but this could be a subject for future work.

There are three primary issues that need to be addressed in this work. The first is that of data latency. Data latency is defined as:

$$\text{data latency} = \text{receipt time} - \text{report time} \tag{67}$$

where report time is the time at which an observation is valid. The receipt time is the time it was received by the operational center and available for assimilation. Data latency is thus the amount of time between the time an observation is taken and the time it is ready to be used in operations. There is always some amount of data latency for observations because data collection and transmission are not instantaneous.

The second issue is controlling high-frequency noise and imbalances in the analyses. In previous versions of the GFS, the tangent linear normal mode constraint (TLNMC; Kleist et al. 2009a) has been used. More recently in the upgrade to GFSv16, the 4DIAU (Lei and Whitaker 2016) technique is used (in combination with the TLNMC) which is a method to apply the analysis increments to the backgrounds to not excite imbalances causing numerical instability during the model integration. It does this by applying the increments as a model forcing that is applied through the assimilation window. The use of 4DIAU in an hourly overlapping window configuration will be tested.

The third issue is an improved use of observations and incorporating high resolution and high frequency observations in a global, hourly updating system. The new [to the GFS] observation data set considered in this chapter includes radial wind observations. For the purposes of this dissertation, we will investigate some initial results from assimilating Doppler radial wind observations in the context of the GRR.

This chapter has two main objectives. The first objective is to develop a system to implement 1-hourly, global data assimilation and to compare that against the 6-hourly configuration at low resolution (C192 ensemble). The issues of data latency and mitigating noise imbalances in the initialization will be discussed. The second objective of this study bridges the work of Chapters 3 and the first objective of this chapter by testing the assimilation of *real* radial wind observations from the NEXRAD radar network within the GRR during a land-falling TC event in Aug 2021 (Fred; https://www.nhc.noaa.gov/data/tcr/AL062021_Fred.pdf). One of the advantages of an hourly updating system over a system that updates every 6 hours is that high-frequency

observations, such as radial winds, might be used more effectively and make a significant impact on the forecast.

In Section 4.2, the model and data assimilation configurations are described. Results are presented in Section 4.3 including the results of the simplified no satellite, pure LETKF experiment; the more complex full-input, hybrid gain LETKF experiment; additional experiments to help better understand the causes for improvements in the hourly system; and sensitivity tests assimilating radial wind observations. The concluding thoughts and future work are then presented in Section 4.4. First, before discussing the GRR any further, additional background information about the differences between the GFS and GDAS systems is given to help better contextualize the issue of data latency.

4.1.1 The GFS/GDAS Systems

To better understand the issue of data latency, it can be helpful to understand the differences between the GFS and GDAS systems. Both systems use the same underlying model and data assimilation software but differ in their specific implementations. Both are necessary for maintaining *timely* and *accurate* global NWP forecasts (as of the GFSv16 upgrade in 2021, both systems use the hybrid 4DEnVar with LETKF employing 4DIAU). The GFS is the Global *Forecast* System and the GDAS is the Global *Data Assimilation* System. To be clear, both systems are complete data assimilation systems meaning that they both have a forecast step and a data assimilation step. The difference is that the GFS is meant to create a *timely forecast* and the GDAS is meant to create an *accurate first guess background forecast* for the subsequent GFS cycle. Both systems currently use a 6-hour data assimilation window

defined as [-3:00, +3:00] h centered around the nominal analysis time, updating the state with observations to produce analyses every 6 hours at the usual synoptic times of 0000, 0600, 1200 and 1800 UTC. These systems, however, assimilate a different total number of observations because of the issue of data latency and the differences in their specific implementations such as how long they wait for late arriving observations. Fig. 35 shows an example of the differences in observation counts between the GDAS and GFS; the largest difference is near the end of the window due to the GFS's relatively shorter waiting time for observations.

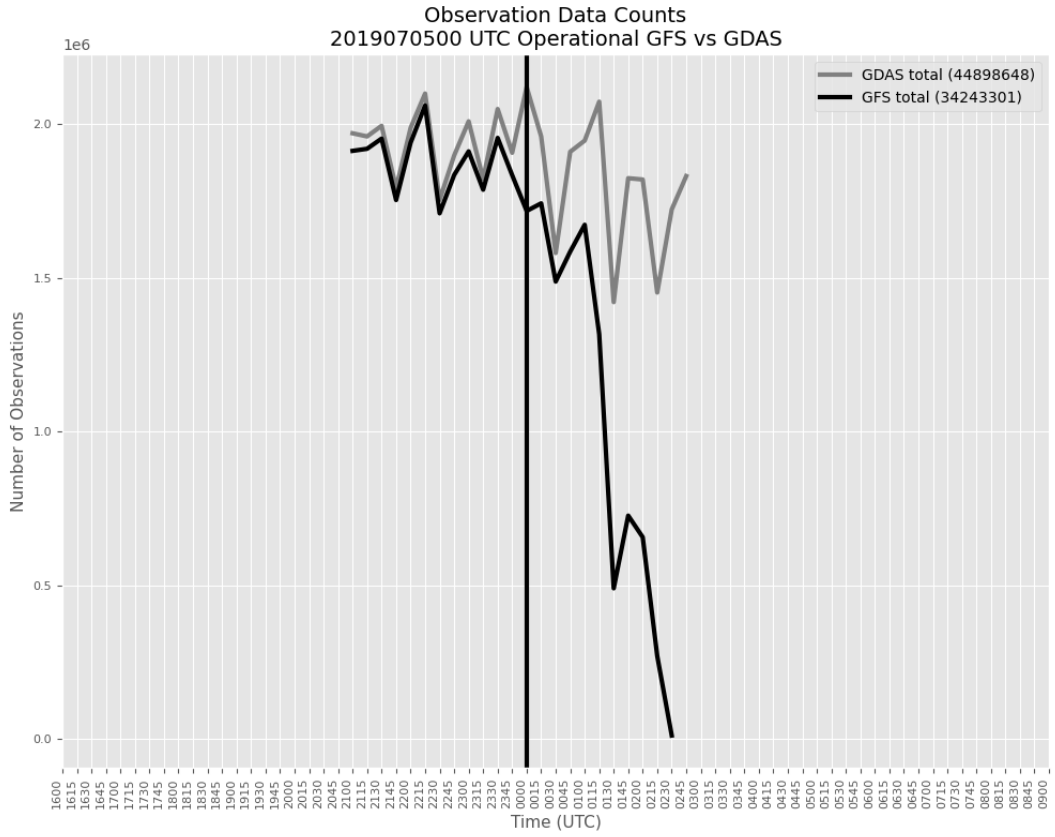

Fig. 35. Histogram plot of observation counts for the GDAS (gray) and GFS (black) depicting the difference in observation counts due to the data receipt cutoff time which are 5:50 h and 2:45 h relative to the nominal analysis time for GDAS and GFS respectively. The x-axis is the observation report time.

Table 9. Reproduction of https://www.nco.ncep.noaa.gov/pmb/nwprod/prodstat/index.html#TARGET for the 0000 UTC GFS/GDAS systems at the time of this writing.

00 UTC GFS		
Event	**Average Start Time**	**Average End Time**
Observation Processing	02:47:07	02:58:30
Analysis	02:58:42	03:28:14
...	...	...
000-384 Products	03:34:40	05:11:27

00 UTC GDAS		
Event	**Average Start Time**	**Average End Time**
Observation Processing	05:50:06	06:02:19
Analysis	06:02:31	06:41:29
Forecast	06:49:10	07:10:00

Table 7 shows the average start/end times for the various steps of the GFS and GDAS systems. The average start time for the GFS for a given cycle is ~2:45 h (after the cycle time) while for the GDAS, the start time is ~5:50 h (after the cycle time). These start times are what we referred to as *data receipt cutoff times* for the GFS and GDAS respectively in this paper. This is because the first step in each system is the observation processing step which prepares all *currently received* observations that are valid within the defined observation window [-3:00, +3:00] h. This means that the GFS will not include any observations with a valid time of 2:45 h or later (e.g., Fig. 35) since the observation processing step has already begun and the data would need to have already been received by this time. The same applies to the GDAS; if an observation would have a data latency higher than 5:50 h but has a valid time within the observation window it would not meet the receipt cutoff time and therefore would not have been available for the operational center to prepare for use. We can write the GFS and GDAS observation windows with their corresponding data receipt cutoff time as

GFS: [-3:00, +3:00; +2:45] h

GDAS: [-3:00, +3:00; +5:50] h (68)

where the first and second numbers (separated by ",") correspond to the beginning and end of the assimilation window (relative to the nominal analysis time), respectively. The third number (following the ";") corresponds to the data receipt cutoff time (relative to the nominal analysis time).

About 95% of all observations (from GDAS files) have a data latency of 3:30 h or less (Fig. 36), so waiting 5:50 h allows enough time for the high latency observations valid near the end of the assimilation window to arrive and be used in the GDAS assimilation but also allowing enough time to finish before the next GFS cycle begins.

To summarize, the GFS is launched earlier to balance waiting long enough for observations to become available and producing a timely forecast. The GDAS is used to handle the issue of data latency by waiting longer for more higher latency observations and then producing a more robust background for the next GFS/GDAS cycles.

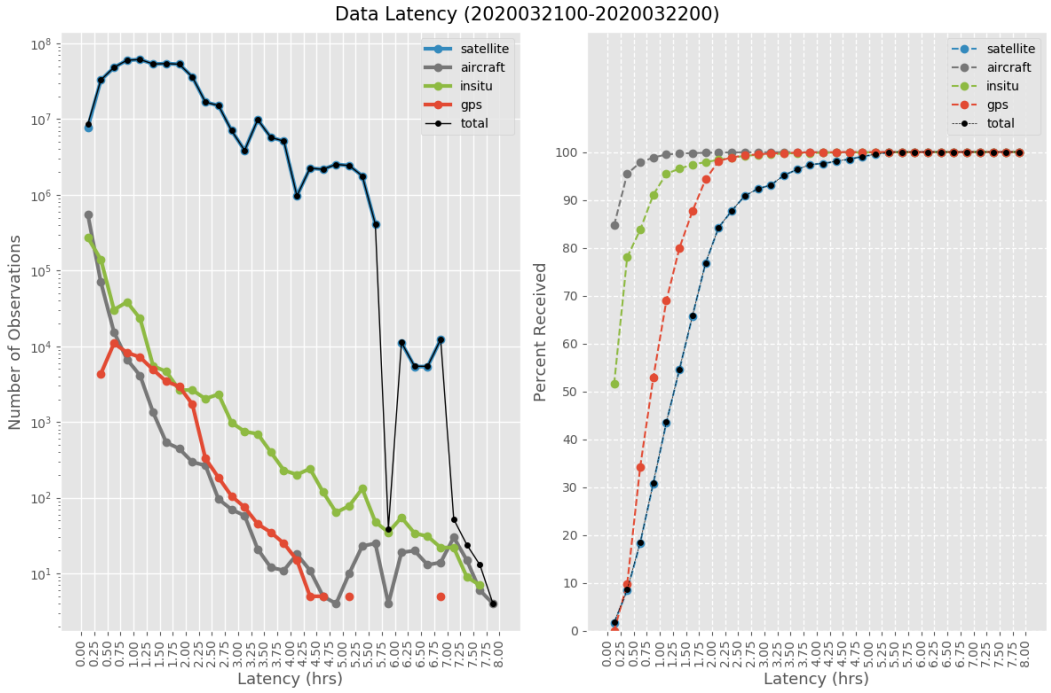

Fig. 36. Observation data latency from the GDAS observation files from 0000 UTC 22 March 2020 – 0000 UTC 23 March 2020. The left panel shows the line histograms of each observation type in 15-minute bins according to each observation's computed data latency. The right panel shows a cumulative distribution function denoting the percent received given a particular data latency period for each observation type. Observations are separated into four types: satellite, aircraft, in situ, and GPS. The total of all observations is also shown.

4.2 Methods

Fig. 36 demonstrates the issue of data latency in the current observing systems and the challenge of implementing an hourly updating system. This figure shows the data latency of all observations in the OW window for 0000 UTC 22 March 2020 – 0000 UTC 23 March 2020 and is computed into 15-min bins. The histograms of latency are shown by the solid lines. The cumulative percent received curves are also shown by different observation categories (e.g., satellite, aircraft, in situ, and GPS). Traditionally, hourly systems have been configured with non-overlapping windows (i.e., [-0:30, +0:30; +0:26] h) which means the times for receiving data do not overlap

between cycles. Such a system would only be able to assimilate up to 30% of all observations since less than 30% of all data have a latency of less than 1 hour (Fig. 36).

Catchup cycles is an approach that is currently used to reset the state of systems where data latency is an issue and usually occurs every 6 hours in the operational systems. This approach is most commonly thought of to be employed in rapidly updating systems such as the 1-hourly updating RAP/HRRR systems, but it is also used in global applications via the GDAS system to reset the GFS system. Rapidly updating systems and the GFS are necessary to provide useful real-time forecast guidance. However, due to the formulation of the assimilation window and the need to initialize the forecast in a timely manner (i.e., relatively short data receipt cutoff times), there would be many observations that go unused (e.g., Fig. 35). This would likely negatively impact the forecasts after some time if these systems were left updating without their respective catch-up cycle.

Overlapping windows is an approach for configuring a data assimilation system to cycle more frequently than would otherwise be possible in a system where observational data latency would be a barrier. For example, many hourly systems are currently designed with a symmetric, non-overlapping window which might have an assimilation window defined as [-0:30, 0:30] h relative to each analysis time. This would not allow for observations that arrive more than 1 h late (for instance, due to delayed communications of observations to the operational center) to be assimilated, either in the current cycle or a later cycle. In overlapping windows, the backward-looking end of the assimilation window is extended to an earlier time which *overlaps* with previous cycles (e.g., -3 h would overlap with three previous analysis windows)

while keeping the forward-looking end of the assimilation window to a time suitable for the current system (e.g., +30 min for a 1-hourly system) to make timely forecasts. The corresponding assimilation window that has been used in the hourly GFS was defined as [-3:00, +0:30] h with a data receipt cutoff time of +26 min. To mitigate duplicate data, a non-overlapping receipt time window of say [-0:34, +0:26] h can be used to filter observations such that only those that were received within that period are used (i.e., observations that were received within the past hour). There are two primary strategies that have been proposed to address the issue of data latency: (1) catch-up cycles and (2) overlapping, asymmetric assimilation windows. The hourly catch-up cycle method is inspired by GDAS/GFS systems and will be used as a benchmark for comparisons with the overlapping windows method. It is hypothesized that the catch-up cycles method will not have any significant advantage over the overlapping windows method. The advantage of the overlapping windows method is two-fold: 1) the state does not have to be reset at regular intervals thereby potentially not losing any advantage from more frequent cycling and 2) it does not require a separately running 6-hourly system to do the state reset.

4.2.1 Overlapping Windows (OW)

The overlapping windows (OW) method was demonstrated in Payne (2017) to overcome the challenge of delayed observations in a rapidly-cycling assimilation system. The method of overlapping windows deals with the issue of data latency by having asymmetric assimilation windows that overlap with the previous cycle's assimilation window and does not require a separately running global system to reset the hourly state at regular intervals. We chose to configure the OW method to feature

a 3-hour backward looking observation window and a 30-min forward looking waiting period because this specification seems to be the most balanced for capturing late arriving observations and producing timely forecasts. Therefore, we can write the observation window with the corresponding data receipt cutoff time for the OW as

$$\text{OW: } [-3{:}00, +0{:}30; +0{:}26] \text{ h} \tag{69}$$

In this way, each cycle can capture late arriving data during the cycle in which it would be first available without needing to wait several hours before launching each hourly cycle forecast as a forecast instead of a hindcast. In fact, the 3-hour backward looking window (i.e., a 3:30 h total window) is enough to capture at least 95% of the available observations according to Fig. 36. Fig. 37 shows an example of the observation counts between the GDAS/GFS (from Fig. 35) and the six consecutive OW cycles (2200 UTC – 0300 UTC). The red line shows the sum of the six OW cycles and shows that cumulatively, the OW will follow the GDAS observation counts as more cycles are included.

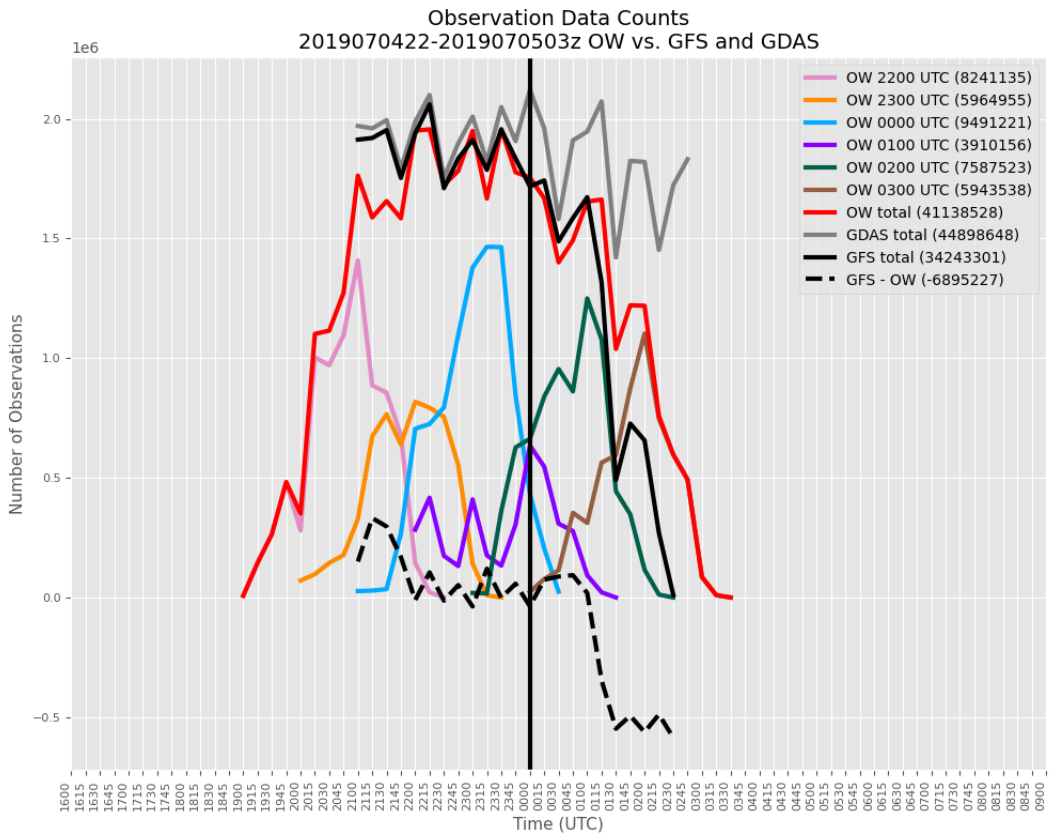

Fig. 37. As in Fig. 35 comparing GDAS and GFS against six consecutive overlapping window (OW) hourly observation files (duplicate filtered) over a period from 2200 UTC – 0300 UTC. The red curve shows the cumulative observation count over the six OW cycles. The difference between the cumulative OW (red) and GFS (black) is also shown (black dashed). The vertical line represents 0000 UTC and is the corresponding time for the GFS/GDAS observation files.

The goal in a real-time system is to have the hourly system begin running approximately 0:30 h after each hour. The OW method will inherently have duplicate data from cycle to cycle; therefore, careful attention is also made to remove the duplicate data (by checking the receipt time) from each cycle to ensure that observations are only assimilated once. *Essentially, the OW method will assimilate all new data within one hour of their receipt* potentially using the data more effectively than what is done currently, especially for low-latency, high-frequency data such as radar observations. For example, the current GDAS/GFS systems assimilates

observations with a valid time window of [-3:00, +3:00] h relative to each analysis time. This lends to necessarily using some low-latency, high-frequency observations (e.g., observations valid between 0900 UTC – 1100 UTC for a 1200 UTC cycle) near the beginning of the assimilation window. We know that observations near the end of the assimilation window have a higher impact to the analysis (McNally 2019). Therefore, the more frequent assimilation of this type of data would naturally place those observations near the end of the assimilation window.

The overlapping window *without* IAU (referred to as OW) is visualized in Fig. 38 (left column). The first step is to run a cold start cycle (0000 UTC; purple box; Fig. 38) which generates the first guess (fg) background states (fg1-5) for the next cycle. The next cycle begins (0400 UTC; green box; Fig. 38) and the nominal analysis time (anl4) is asymmetrically centered by the assimilation window (gray box). In this method, the analyses are calculated for each state within the assimilation window denoted as anl1-5. The analyses anl2-4 become the first guess files fg1-3 respectively which act as part of the background for the subsequent cycle. The model then starts a new 2-hour, unforced forecast from the nominal analysis time (0400 UTC; anl4) and the resulting 1- and 2-hour forecast files act as the remaining set of background files (fg4 and 5) for the subsequent cycle.

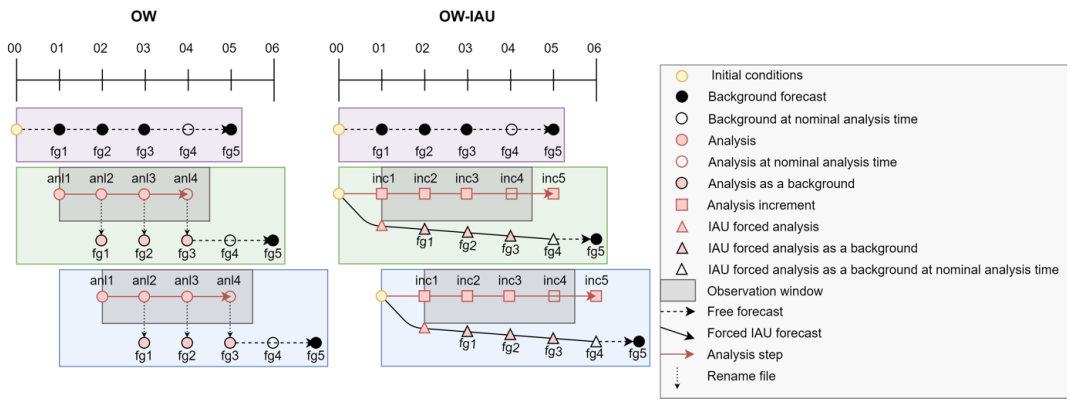

Fig. 38. Diagram of overlapping windows without 4DIAU (OW; left) and overlapping windows with 4DIAU (OW-IAU; right). The purple box demonstrates cycling from a cold start. The green box demonstrates the first full assimilation cycle. The blue box shows the second full assimilation cycle. See text for more details.

The overlapping window *with* 4DIAU (referred to as OW-IAU or OW-4DIAU) is also visualized in Fig. 38 (right column). The first step is to run a cold start cycle (0000 UTC; purple box; Fig. 38) which operates the same as in the OW and generates the first guess background states (fg1-5) for the next cycle. The next cycle begins (0400 UTC; green box; Fig. 38) and the nominal analysis time (anl4) is asymmetrically centered by the assimilation window (gray box). An analysis increment is calculated for each of the states within the assimilation window indicated by (inc1-5). The model then restarts 1-hour prior (at 0000 UTC) to the current time (0100 UTC) of the restart file, runs for 1-hour then begins the IAU and forcing with the analysis increments. After the model has been forced with the fourth increment, the model runs for one additional hour unforced by the IAU. The backgrounds for the subsequent cycle begin with the state fg1 forced by the associated increment inc2 through the unforced state fg5. These backgrounds are denoted as fg1-5. The cycling continues in the same manner on an hourly updating basis (blue box).

To summarize a few key points, another way to think of the OW method is that the observation windows do overlap, but the receipt time windows do not. Therefore, each cycle will only assimilate observations within the receipt window defined as [-0:34, +0:26] h. The non-overlapping receipt time windows is how we ensure that each observation is only assimilated once, and the overlapping valid time windows is how data latency is managed.

4.2.2 Catch-Up Cycles

The overlapping windows method will also be compared against the catch-up cycle algorithm inspired by the RAP and is similar to that described in Djalalova et al. (2016) and to that used in Chapter 2 of this dissertation. The catch-up cycles is also thoroughly described in Slivinski et al. (2022). In such a system, analyses are generated hourly using 1-hour symmetric assimilation windows and data receipt cutoff time defined as

$$\text{Catch-up: } [-0{:}30, +0{:}30; +0{:}26] \text{ h} \tag{70}$$

If such a system were continuously cycled, it would only be able to assimilate low-latency data. At best, it could assimilate observations with 1-hour or less latency which only accounts for up to 30% of all available observations (Fig. 36). To address the issue of data latency, the catch-up cycle system is reinitialized every 6 hours from a separately running global, 6-hourly system which uses a much later receipt cutoff time to capture high latent observations. In practice, this separate 6-hourly system would be the NOAA GDAS system, which has a receipt cutoff time of about 3 hours past the end of the assimilation window and lags real-time by approximately 6 hours (i.e., the 0600 UTC analysis runs at about 1200 UTC) and is currently used to reinitialize the GFS.

Therefore, the catch-up cycle algorithm is designed such that it would be initialized from the end of a GDAS assimilation window (say 0900 UTC) with a run time 3 hours later (1200 UTC), and then cycled hourly for 9 hours. The first four cycles (blue filled circles in Fig. 39) are deemed the actual "catch-up" cycles before running the real-time hourly cycling (unfilled red circles in Fig. 39) beginning at 1300 until 1800 UTC. Note that the control experiment used for reinitialization includes 4DIAU, but the hourly cycling does not because it cannot be implemented in an hourly system without sub-hourly increments.

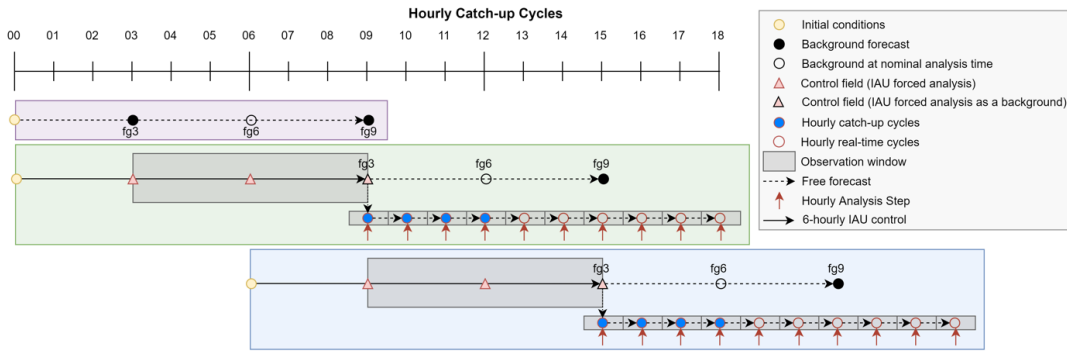

Fig. 39. Catch-up cycles diagram. The purple box demonstrates cycling from a cold start. The separately running 6-hourly system is depicted in the green and blue boxes as the 6-hourly IAU control where the first guess hour 3 (fg3) is used to initialize the hourly catch-up cycles denoted by the blue circles. Four, 1-hourly data assimilation is performed until the nominal analysis time for the control at which point the hourly-real-time analyses begin. The state is reset using the next fg3 in the subsequent hourly catch-up cycle.

4.2.3 The 6-Hourly GFS (control)

Each of the previous configurations will be compared against the 6-hourly GFS system with 4DIAU which serves as the control in these experiments and is like what is used in operations. Note that the control uses the GFS observation files (as opposed to the GDAS files) for reasons described later. Fig. 40 illustrates how the control experiment is implemented. The first cycle (0000 UTC; purple box) generates the first guess (fg) background states (fg3-9) for the next cycle. The next cycle begins (0600

UTC; green box) and the nominal analysis time (anl6) is symmetrically centered by the assimilation window (gray box). Since 4DIAU is used, an analysis increment is calculated for each of the states within the assimilation window indicated by (inc3, 6, and 9; red squares). The model then restarts 3 hours prior (at 0000 UTC) relative to the current time (0300 UTC) in the restart file, it runs for 3 hours, and then it begins the IAU by using the increments as forcing to create the IAU-forced analyses (red triangles). After the final increment has been forced, the model runs unforced for six additional hours generating the backgrounds (fg3-9) for the next assimilation cycle (blue box).

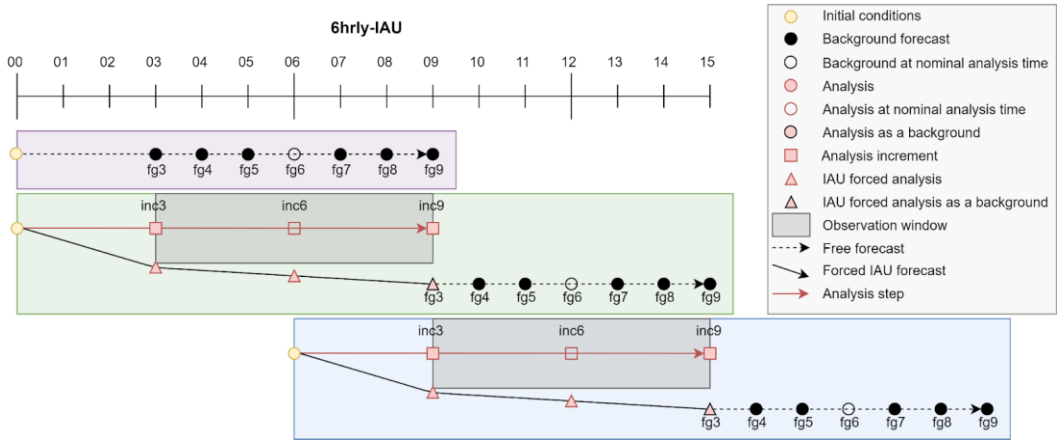

Fig. 40. Diagram of 6-hourly control configuration with 4DIAU and hourly background fields. The purple box demonstrates cycling from a cold start, the green box demonstrates the first full assimilation cycle, and the blue box shows the second full assimilation cycle. See text for more details.

4.2.4 Fairest Comparisons of Experiments

In section 4.1.1, the differences in the implementations between the GDAS and GFS were given. Since then, we have described an alternative method for a 1-hourly system that can account for high latency data: overlapping windows. This poses another challenge in how to perform a fair comparison between a 6-hourly experiment with a

much longer receipt cutoff time than the new 1-hourly system since the 6-hourly system for a given cycle will have seen more observations. For example, Fig. 41 shows two options (orange and red) for comparing the hourly experiment against the control experiment (blue)—focusing on the data receipt cutoff times denoted by the vertical dashed lines in Fig. 41. Notice that there is not a perfect one-to-one comparison for the hourly and the control experiments. When comparing the control against the 0200 UTC hourly cycle, the hourly will have seen fewer observations (a 0:19 h disadvantage). When comparing the control against the 0300 UTC hourly cycle, the hourly will have seen more observations (a 0:41 h advantage).

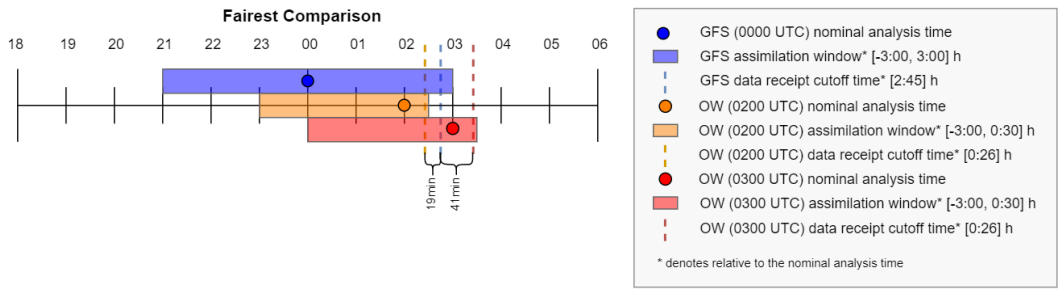

Fig. 41. Fair comparison diagram. Compares the assimilation windows for 6-hourly GFS control (blue; 0000 UTC), hourly conservative option (orange; 0200 UTC), and less conservative option (red; 0300 UTC). There is no perfect option with respect to comparing the control with the hourly experiments in terms of the data cut off. The control has a 0:19 h longer waiting period relative to the 0200 UTC and a 0:41 h shorter waiting period relative to the 0300 UTC. The conservative, 0200 UTC option was used as the fairest comparison in this study; it also gives the advantage to the control, in terms of waiting longer for observations.

The hourly cycle that occurs 2 hours after the 6-hourly cycle (i.e., 0200, 0800, 1400, and 2100 UTC) is used for all comparisons against the 6-hourly cycles (i.e., 0000, 0600, 1200, and 1800 UTC) in this study, respectively. Using the 2-hour lagged cycle comparison gives the fairest comparison giving a 0:19 h advantage to the control in terms of the waiting period for the observations, but also favors the 6-hourly control (with 4DIAU) which will have a 1-hour shorter free-forecast (due to 4DIAU) compared to the 2-hour lagged cycle. Specifically, the 0000 UTC GFS 6-hour forecast valid at

0600 UTC is a 3-hour free-forecast beginning from the end of the analysis window at 0300 UTC (i.e., the IAU is run until the end of the analysis window). The forecast valid at the same time (at 0600 UTC) from the 0200 UTC OW cycle is a 4-hour free forecast from the 0200 UTC analysis.

In summary, the "fairest comparison" method that we have chosen is to compare the OW cycle lagged by 2-hours relative to the control such that both systems would have seen approximately the same observations. The advantage is given to the control in terms of having a 0:19 h longer data receipt cutoff time (more observations) and having a 1-hour shorter free-forecast. We chose to not use GDAS as the control in this study because it would be even more difficult to have a fair comparison since the GDAS would have a 5:24 h advantage in terms of data receipt cutoff time. Also, we used the GFS as the control because it would provide the fairest comparison between fields available to a forecaster at a given time. The effects from differences in observation counts will be examined in additional experiments presented later in this chapter. We will show that the difference in performance from the extra observations is not as impactful in the 6-hourly system; therefore, it is sufficient to use the GFS as the control.

4.2.5 Experimental Setup

Two sets of hourly observation files were created: "overlapping windows" (OW) and "hourly symmetric windows" (used by the catch-up cycle experiment). The OW observation files were created using observations valid within the window defined by Eq. (69) and the hourly symmetric window observation files were created using observations valid within the observation window defined by Eq. (70). The data receipt

cutoff time is also indicated by Eq. (69) and Eq. (70). Duplicate data is removed from the overlapping windows by checking the receipt times (which do not overlap) or other metadata when receipt times are not available.

Table 10. Summary of experiments shown in Section 4.3.

Expt. name	Inflation (RTPS/RTPP)	3DVar weight	Ob types assimilated	Ob window (rel. to nominal analysis time)	Ob receipt cutoff (rel. to nominal analysis time)
NoSat, LETKF control	0.75/0.25	0	All except satellite radiances	[-3:00, +3:00] h	+2:45 h (GFS)
NoSat, LETKF OW	0.75/0.25	0	All except satellite radiances	[-3:00, +0:30] h	+0:26 h
No Sat, LETKF OW-IAU	0.75/0.25	0	All except satellite radiances	[-3:00, +0:30] h	+0:26 h
Full-input control	0.75/0.25	25	All	[-3:00, +3:00] h	+2:45 h (GFS)
Full-input OW	0.9/0.4	15	All	[-3:00, +0:30] h	+0:26 h
Full-input catch-up cycles (hourly)	0.9/0.4	15	All	[-0:30, +0:30] h	+0:26 h
Full-input catch-up cycles (control for reinit.)	0.75/0.25	25	All	[-3:00, +3:00] h	+2:45 h (GFS)
Radial wind	0.9/0.5	20	All + Radial Wind	[-3:00, +0:30] h	+0:26 h

These observation files are then used in a series of experiments with increasing complexity over a four-week period in Mar-Apr 2020 using the GFSv15 at C192 (low resolution; ~50km) for the 80-member ensemble (ensemble mean is used as the control member) using the hybrid-gain, LETKF (Penny 2014). The hybrid-gain LETKF, as opposed to the hybrid 4DEnVar (Kleist and Ide 2015a; NWS 2021) currently used in

GFSv16 (and v15), is used for testing purposes since it is shown to have comparable performance and is computationally less expensive than the hybrid 4DEnVar. Future work will investigate the hybrid 4DEnVar in this system.

For the control, we use the "fairest comparison" method as described in Section 4.2.4 giving advantage to the control via a 0:19 h longer data receipt cutoff time and 1-hour shorter free forecast used as the background. Once again, we use the GFS observation files instead of the GDAS files to have the fairest comparison.

We first demonstrate that the overlapping windows configurations works in a simplified experiment. The simplified experiment consists of running without satellite observations (NoSat) and assigning zero weight to the 3DVar update in the hybrid-gain, LETKF (i.e., pure LETKF). We also exclude aircraft bias correction in the simplified experiment. We then show the impacts of the overlapping windows method in a full-input configuration (now assimilating satellite observations) and using the hybrid-gain LETKF. We also compared the full-input OW experiments against the catch-up cycle method.

Four additional experiments were conducted to investigate further the causes for improvements in the hourly system, and whether they are due to the cumulative difference in observation counts between the GFS and OW or by the hourly updating itself. These experiments ignore data latency to allow us to better understand the causes for improvements.

In a final experiment, radial wind observations were then retrospectively generated using NCEP's global observation generation scripts and were tested in a simple set of retrospective experiments spanning several 1-hourly assimilation cycles

in August of 2021 of a landfalling TC event. The experiments were cold started from the GDAS at the 1200 UTC cycle on 16 August 2021. The hourly cycling begins with the 1800 UTC cycle on 16 August 2021 (1 h before TC Fred was observed to make landfall) and cycled for 20 hours. The system was cold started close to the TC making landfall so that the storm was close enough to be sampled by the land-based radar network.

In this experiment, the control is simply the newly formulated GRR but updated to use GFSv16 and run at the full operational resolution for the ensemble (C384 ensemble; ~25km); there is still no high-resolution control member for this experiment. The experimental run only differs by the radial wind assimilation and parameters specifically related to the use of that data.

The assimilation of observations is performed using a two-pass procedure (e.g., Tong et al. 2020) whereby all non-radial wind observations are assimilated in the first pass of the GSI, then radial wind observations are assimilated in the second pass. This is done to use a variable specific horizontal localization in the variational part of the assimilation and because scale dependent localization is not fully ready in the GSI (Huang et al. 2021). A value of 100 is used for the horizontal localization which translates to approximately 270km for the distance at which the correlations go to zero. For all other observations, this parameter is 343 which corresponds to approximately 1250km. Radar specific localization is also applied in the LETKF part of the assimilation, which was increased from its default value of 10km to 100km for this work owing to the coarse model resolution. Non-radar observations use a localization

scale of 2500km. The super-observation parameters are summarized in Table 4. The system also uses the hybrid-gain LETKF.

Table 11. List of the super-observation parameters and their default and experimental values.

Exp	Azimuth range (degrees) $\Delta\theta$	Elevation angle range (degrees) $\Delta\propto$	Gate spacing (meters) Δr	One-half time range (hours) Δt	Max elevation angle (degrees) $\propto_{max}$	Minimum number of samples N	Radar Data Localization Var/Ens (km)
Fred	6	0.25	25,000	+/- 0.125	5	50	270/ 100

4.3 Results – Comparing Hourly and 6-hourly Global Systems

4.3.1 NoSat, Pure LETKF

Using the "fairest comparison" described in Section 4.2.4, we consider the root-mean-square (RMS) error forecast fit to in-situ observations of each experiment for short forecasts (3–4-hour free forecast). Statistical significance is also computed using a Student's t-test with inflation to account for temporal autocorrelations (Geer 2016). Fig. 42 shows RMS error fit to all in-situ observations for the overlapping windows with and without 4DIAU relative to the 6-hourly control in the NoSat setup. The OW (without 4DIAU) shows improvements in upper-tropospheric winds, but slight degradations in all layers for temperature relative to the control, whereas the OW-IAU (with 4DIAU) shows degradations relative to the control. These results from the simplified experiment are promising except for the OW-IAU. In a later section we will compare the OW and the catch-up cycle method relative to the control in the full-input experiment. For now, we will continue to investigate the differences in the overlapping windows with and without 4DIAU.

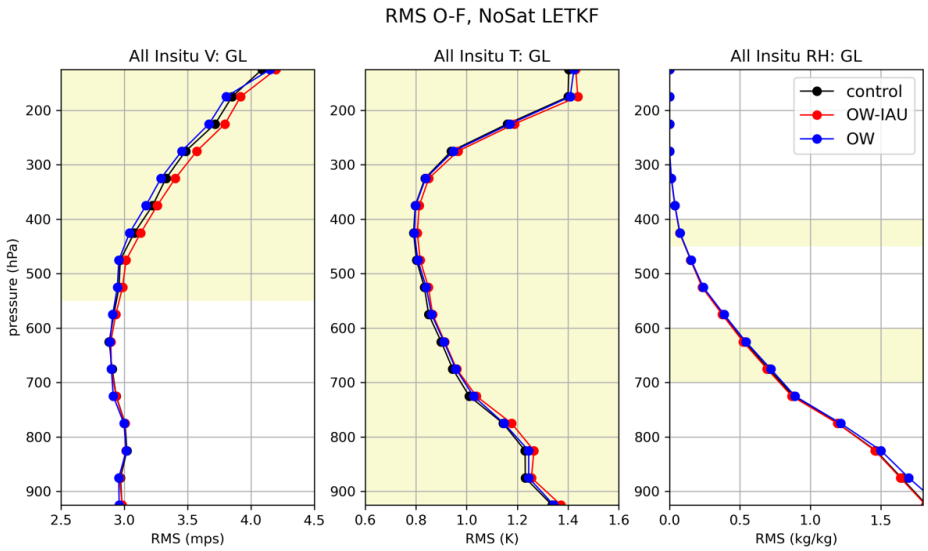

Fig. 42. Vertical profiles of short forecast root-mean-square (RMS) fits to in-situ observations of wind (left) temperature (center), and relative humidity (right) for the control (black), OW (blue), and OW-IAU (red) experiments in the simple, NoSat LETKF setup, averaged over the period 1800 UTC 13 Mar 2020 – 1800 UTC 13 Apr 2020. Yellow shading denotes where the control and OW experiments differ significantly at the 95% level.

4.3.1.1 Issues with overlapping windows and 4DIAU

During the preliminary stages of developing an hourly, overlapping system *with* 4DIAU there was a noticeable degradation in the forecast fit to observations when compared against the control (Fig. 43). There are fundamental challenges for implementing OW with 4DIAU for the full assimilation window (i.e., the version that we also detailed in Slivinski et al. (2022); see appendix A). There are at least two issues with this implementation: 1) the background forecast would come from IAU forced analyses which would not have seen all the observation information and 2) the restart files (used for starting the IAU step) should not be valid before the current cycle's analysis time (this is the case for the formulation of the OW-4DIAU that we used in Slivinski et al. (2022)). It would only be possible to correctly implement OW-4DIAU in a 1-hourly system if there were sub-hourly IAU increments that could be used over a 1-hour period beginning from the nominal analysis time. What we have found is that

the IAU can only be implemented in a forecast system in which the length of the IAU period is at most equal to the frequency of the cycling (e.g., a 1-hour IAU period for 1-hourly updates). To properly implement OW-4DIAU, it would require sub-hourly IAU increments over a 1-hour period.

The purpose of this section is to present the details of this ill-formed configuration of OW-4DIAU over the full assimilation window which exhibits what we refer to as a "loss of information". This loss of information is unique to the hourly, overlapping window and is not present in the 6-hourly control using 4DIAU (e.g., GFSv16). Notice that in the control (6-hourly with 4DIAU; Fig. 40), the 4DIAU runs all the way through before backgrounds fg3,6,9 are generated for the subsequent cycles, whereas in the hourly OW-4DIAU (Fig. 38; right column), the 4DIAU has only started forcing with the second analysis increment and is simultaneously creating backgrounds fg1-4 for the next cycle. We demonstrate that generating backgrounds before the 4DIAU is run to completion will cause a loss of observation information because only a fraction of the increment is applied at each time step, and it is not until the final IAU time step when the remaining portion of the full increment has been applied. Additionally, information is lost by generating the restart file before the IAU runs to completion. This potential loss of information has prompted the development of an hourly, overlapping windows *without* 4DIAU configuration (Fig. 38; left column) and is simply just called OW.

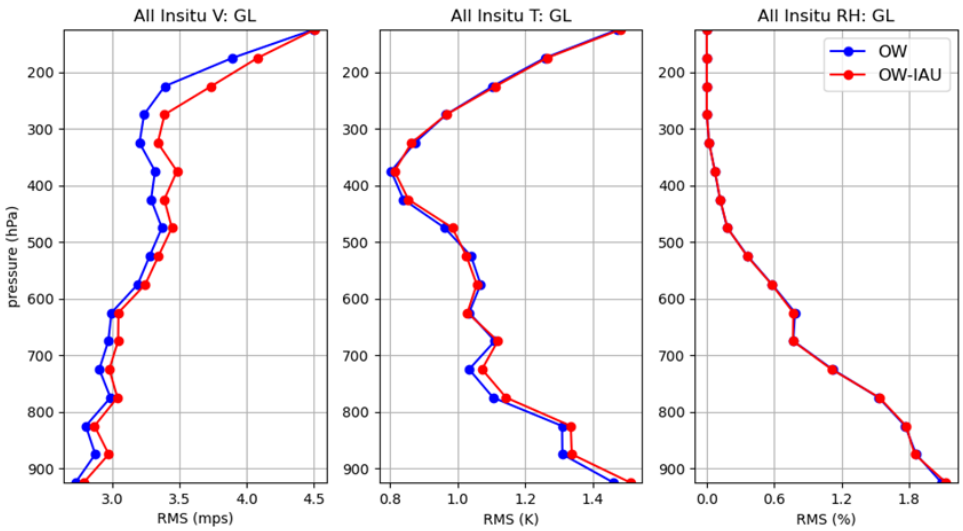

Fig. 43. Vertical profiles of RMS 1-h forecast differences with respect to in-situ observations of wind (left), temperature (center), and relative humidity (right) for OW (blue) and OW-IAU (red).

We tested the hypothesis that the loss of information in OW-4DIAU over the full assimilation window is due to 1) using IAU forced analyses as backgrounds that have not "seen" all of the observation information and 2) the restart file (used for starting the IAU step) being written out prior to the current analysis. First, both versions of the overlapping windows workflow (with and without 4DIAU) were shown to provide reproducible results when no observations were assimilated after the first full cycle and after removing randomness from the forecasts such as that brought about by stochastic physics. Then surface pressure observations were added into the assimilation step. Both configurations were run for one full cycle and the innovation statistics were compared. The configuration with 4DIAU had higher innovation RMS scores (Fig. 43). This simple test along with all testing thus far have consistently shown the hourly with 4DIAU configuration to underperform when compared to both the hourly configuration without 4DIAU and the 6-hourly control. Therefore, we have determined that the OW-4DIAU where the 4D window consists of the full assimilation window is an ill-

proposed implementation. It would be possible to implement the OW-4DIAU with sub-hourly IAU increments forced over a 1-hour period. It would also be possible to implement OW-3DIAU using a single increment over a 1-hour period. Both of these 1-hour periods should start from the nominal analysis time. The filtering properties of these implementations, however, would be very different from the 4DIAU as used currently in the GDAS. Such efforts are not discussed in this dissertation and are left for future work.

The purpose of the 4DIAU is to reduce gravity wave noise and prevent it from dominating the short-term forecast. Since, at this point, we have not been able to successfully implement a system with IAU in an overlapping windows configuration, it is important to assess the presence of imbalance in the OW without IAU configuration. To assess the presence of imbalance we examined the 1-hour global absolute surface pressure tendencies which act as a proxy for gravity wave noise (e.g., He et al. 2020; Lei and Whitaker 2016). In summary, we found that the OW had a 20% higher surface pressure tendency than the OW-4DIAU and 35% higher surface pressure tendency than the control. Although there was an increase in surface pressure tendency suggesting increased gravity wave noise, these imbalances were not found to be enough to substantially degrade the longer forecast.

Furthermore, we found that the 4DIAU led to increased ensemble spread relative to the control and removing the 4DIAU led to decreased ensemble spread relative to the control (Fig. 44). The increased ensemble spread in OW-4DIAU is likely the result of the compounding "loss of information" causing the analyses to be less constrained by the observations. The decreased ensemble spread in the OW is likely

the result of using analyses as backgrounds. In the full-input experiments we attempt to counter the decrease in ensemble spread through inflation.

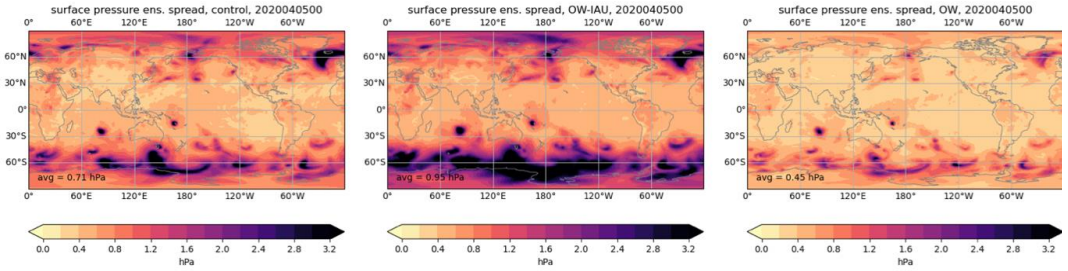

Fig. 44. Ensemble spread of 6-hour background surface pressure (hPa) at 0000 UTC 05 April 2020 for the control (left), OW-IAU (center), and OW (right) experiments.

Despite the increased gravity wave noise (not shown) and decreased ensemble spread (Fig. 44), OW consistently outperforms the OW-4DIAU at all levels and outperforms the control in wind and temperature at upper tropospheric levels (Fig. 42).

4.3.2 Full Input, Hybrid Gain LETKF

The OW system was further evaluated using a more realistic configuration including the full observation dataset with the hybrid gain LETKF and employing radiance and aircraft bias correction (Zhu et al., 2014, 2015). The OW was compared to the control as well as the catch-up cycles method. In addition to the full suite of observations, the experiments also use the updated values for inflation of 0.9 and 0.4 for relaxation to prior spread (RTPS; Whitaker and Hamill 2012) and relaxation to prior perturbation (RTPP; Zhang et al. 2004) (originally 0.75 and 0.25 respectively). RTPS and RTPP inflation is used to counter the decreased ensemble spread (Fig. 44). An updated value of 15% (originally 25%) was used for the alpha parameter for weighting the 3DVar derived increment in the hybrid gain formulation. These values were applied to the hourly systems (OW and catch-up) while these parameters for the control system

were left unchanged. Finally, based on the results of section 4.3.1, OW-IAU was dropped from further evaluation.

4.3.2.1 Overlapping windows and catch-up cycles

Fig. 45 shows the RMS error fit to all in-situ observations for the full-input, hybrid-gain experiments including OW, catch-up cycles, and the control. The OW experiment shows a slightly larger improvement compared to the control than what was found in the simplified NoSat, pure LETKF experiment (Fig. 42). Here, the OW shows improvements in both upper- and lower-tropospheric winds (Fig. 45, left). In temperature, the OW now outperforms the control at most levels (Fig. 45, center). The control still has slightly lower errors for relative humidity (Fig. 45, right).

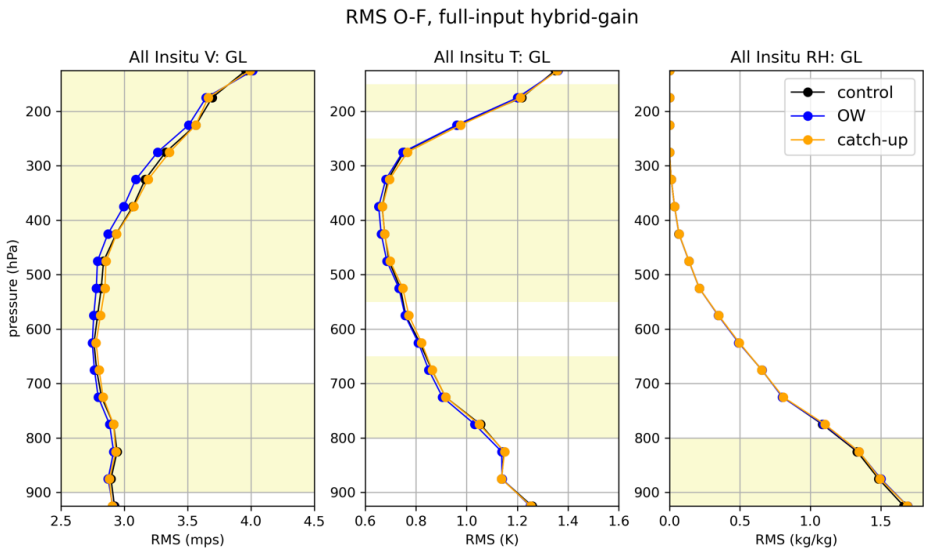

Fig. 45. Vertical profiles of short forecast root-mean-square (RMS) fits to in-situ observations of wind (left) temperature (center), and relative humidity (right) for the control (black), OW (blue), and catch-up (orange) experiments in the full-input hybrid gain setup, averaged over the period 1800 UTC 13 Mar 2020 – 1800 UTC 13 Apr 2020. Yellow shading denotes where the control and OW experiments differ significantly at the 95% level.

The performance of the catch-up cycles is also shown in Fig. 45 relative to the OW and control experiments. The catch-up cycles method very closely resembles the control. We discuss more about the behavior of the catch-up cycles in Slivinski et al.

(2022). In short, the purpose of the catch-up cycles in a global system would be to include more latent data than the 1-hour symmetric window [Eq. (70)] is able to capture. In our configuration, we use the GFS for the control to reinitialize the catch-up cycles with more latent data. In practice, the GDAS system would be used to reinitialize the catch-up cycles; however, it will be shown in a following section that, for a 6-hourly system, the difference between using the GFS and GDAS observation windows is small. Therefore, we hypothesize that the differences between the catch-up cycle and the OW to be due to the OW's ability to assimilate more high-latency data, by assimilating all data within 1-hour of their receipt time, and by updating more frequently without needing to be reset to the control state.

4.3.2.2 Five-day Forecasts

In Slivinski et al. (2022), we also investigated the impacts on the longer range forecast out to 5 days. The 5-day deterministic forecast was run at C384 (~25km) initialized from the downscaled C192 ensemble mean. The 5-day, free forecasts were initiated once per day (to accommodate computational constraints) from the 0000 UTC and 0200 UTC cycle times for the control and OW experiment, respectively. Note that the 0000 UTC and 0200 UTC cycle times correspond to the time at which both systems would have seen approximately the same number of observations and is consistent with our "fairest comparison" method used previously.

Benefits of hourly assimilation over the 6-hourly control are seen in the first 6-12 hours of the forecast as indicated by the blue shading in wind and temperature (Fig. 46). Those benefits mostly disappear after the 6–12-hour period with mostly mixed results thereafter. Note that maximum and minimum differences are on the order of 3%

over the 30 5-day forecasts. More tuning could improve results beyond what has been found here.

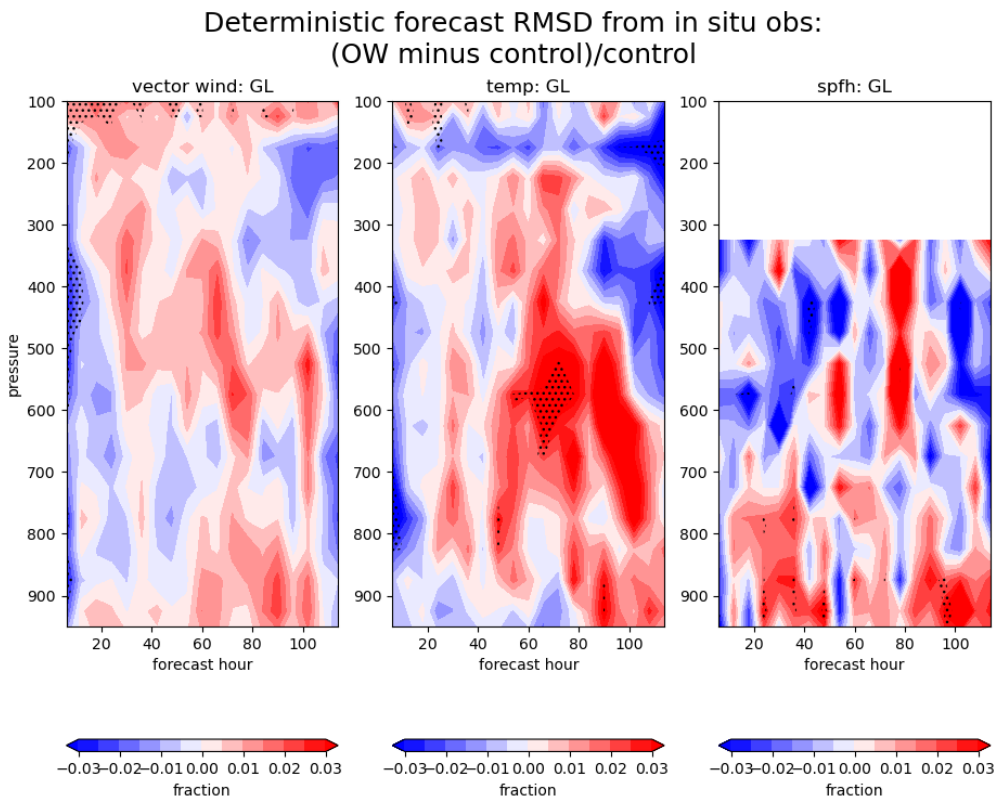

Fig. 46. Contour plot of global relative differences between the long forecast RMSDs of OW and of the control, relative to in-situ observations, as a function of forecast hour and averaged over the period 0600 UTC 14 Mar 2020 – 0600 UTC 12 Apr 2020, for vector wind (left), temperature (center), and specific humidity (right). Blue contours denote that OW has smaller RMSDs than the control, and red colors denote that the control has smaller RMSDs than OW. Stippling denotes where the control and OW experiments differ significantly at the 95% level.

We have previously noted that OW had increased surface pressure tendency than the OW-IAU and control. Fig. 47a shows the 1-month averaged surface pressure tendency as a function of forecast hour. This plot shows that the initial surface pressure tendency is higher in the OW than in the control; however, the initial increased surface pressure tendency eventually dissipates at the longer forecast leads (Fig. 47a).

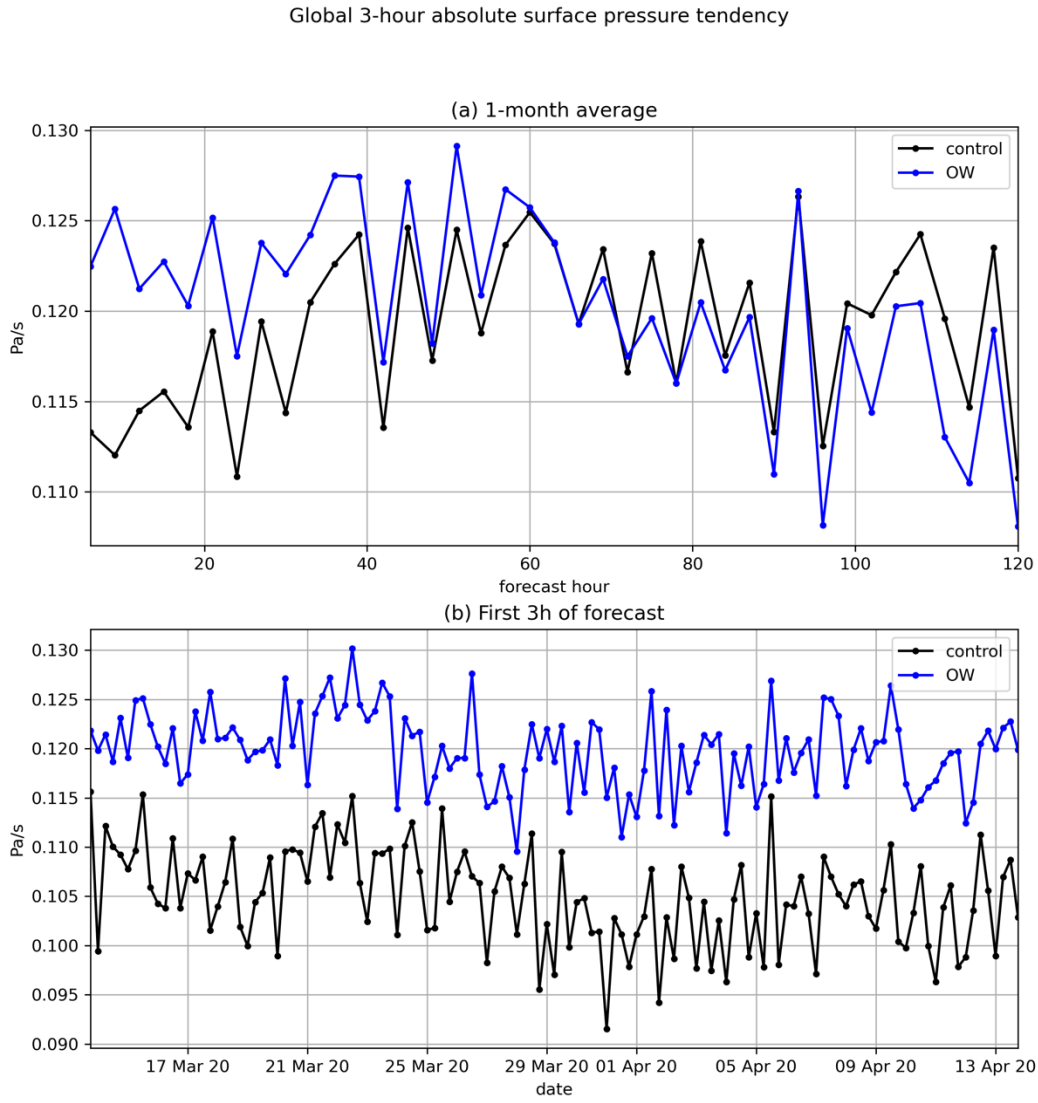

Fig. 47. Globally-averaged 3-hour absolute surface pressure tendency for the control experiment (black) and OW (blue) (a) as a function of forecast lead time, averaged over the 1-month period 14 Mar – 13 Apr 2020 and (b) for the first three hours of the forecast as a function of cycle date for the experiment.

We also compared the surface pressure tendency from the first 3 hours of the forecast as a function of cycle date (Fig. 47b). This figure shows that there is consistently a larger surface pressure tendency in the initial forecast of each cycle of the hourly experiment relative to the control experiment. We found that while the OW has higher surface pressure tendency, it does not consistently increase from cycle to

cycle over the 1-month period that we tested. This further supports that IAU may not be necessary in our hourly cycling system with overlapping windows.

4.3.3 Causes for Improvement in an Hourly System

We hypothesized that there were two main factors as to why the OW shows significant improvements over the control: 1) the OW cumulatively assimilates more observations than the control or catch-up cycles and 2) the hourly frequency of updates. The OW cumulatively assimilates more observations than the GFS control because the 3-hour backward looking time window overlaps with previous cycles. Therefore, it is always waiting for high latency data and obtains roughly 95% of all GDAS observations. The GFS does not have overlapping windows and misses a lot of observations near the end of the assimilation window due to its data receipt cutoff time being 0:15 h before the end of the assimilation window (Fig. 35).

The second factor may also be simply due to more frequent updates. More frequent updates necessarily involve shorter background forecasts which reduces the time for the background error to grow nonlinearly. It may also more effectively use observations by assimilating them closer to their receipt time and potentially closer to the end of the assimilation window which we know will allow them to have a greater impact on the analysis (McNally 2019).

These two factors were investigated in a set of additional experiments using the full-input, hybrid-gain LETKF as in Section 4.3.2 but with RTPS and RTPP inflation parameters of 0.9 and 0.5 respectively and a weighting factor of 20% is assigned to the 3DVar gain. For this exercise we have four configurations which are outlined in Table 12. Two control experiments and two hourly experiments each of which uses either the

GDAS or GFS observations, respectively. In each of these hourly experiments, the GDAS and GFS observations were separated into hourly files. This allows us to ignore data latency and focus strictly on the impact of more frequent updates.

Table 12. Summary of experiments shown in Section 4.4.

Expt. name	Update Freq. (h)	Inflation (RTPS/ RTPP)	3DVar weight	Ob types assim.	Ob window (rel. to nominal analysis time)	Ob receipt cutoff (relative to nominal analysis
control_gfs	6	0.9/0.5	20	All	[-3:00, +3:00] h	+2:45 h (GFS)
control_gdas	6	0.9/0.5	20	All	[-3:00, +3:00] h	+5:50 h (GDAS)
hourly_gfs	1	0.9/0.5	20	All	[-1:00, +0:30] h @ 0400, 1000, 1600, and 2200 UTC; [-0:30, 0:00] h @ 0300, 0900, 1500, and 2100 UTC; [-0:30, 0:30] h @ all other times	+2:45 h (GFS) @ 0000, 0600, 1200, and 1800 UTC
hourly_gdas	1	0.9/0.5	20	All	[-1:00, +0:30] h @ 0400, 1000, 1600, and 2200 UTC; [-0:30, +0:00] h @ 0300, 0900, 1500, and 2100 UTC; [-0:30, +0:30] h @ all other times	+5:50 h (GDAS) @ 0000, 0600, 1200, and 1800 UTC

Fig. 48 shows the RMS fit to observation differences of each of these four experiments for wind (left) and temperature (right). The black curve (control_gdas – control_gfs) shows that for the 6-hourly system the use of the GDAS observation data set over the GFS has slight, but consistent benefits in improving the RMS fit to observations for both wind and temperature. The pink curve (hourly_gdas – hourly_gfs) shows that the use of GDAS observations data set over the GFS has a much larger impact for a 1-hourly updating system especially in upper-tropospheric winds.

The green curve (hourly_gdas – control_gdas) shows the difference between the 1-hourly and 6-hourly system assimilating the same GDAS observations and shows the 1-hourly system has lower RMS fit to observations in both wind and temperature. These results suggest that the use of GDAS observations plays a larger role in the 1-hourly system than in a 6-hourly system and that the improvements are also found due to more frequent updates.

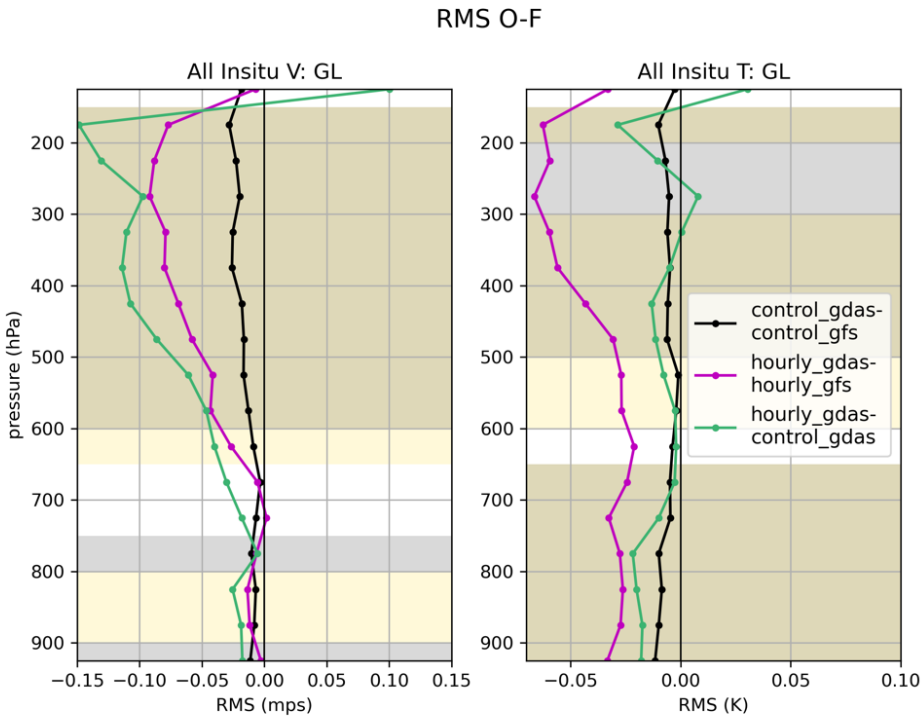

Fig. 48. Vertical profiles of differences of short forecast RMSDs with respect to global in-situ observations of vector wind (m s-1, left) and temperature (K, right) for control_gdas minus control_gfs (black), hourly_gdas minus hourly_gfs (pink), and hourly_gdas minus control_gdas (green) in the full-input hybrid-gain LETKF setup, averaged over the period 1200 UTC 13 Mar 2020 – 1800 UTC 22 Mar 2020. Gray shading denotes where control_gfs and control_gdas differ significantly at 95% level, and yellow shading denotes where hourly_gfs and hourly_gdas differ significantly at 95% level.

These additional experiments removed the issue of data latency to test the theoretical reasons for improvement. Now, in a real system where data latency cannot be ignored, it is useful to determine what improvements are possible by comparing the

experiments without data latency against the experiments with real data latency. Therefore, we compare the OW experiment (with data latency) and the hourly_gdas (essentially the OW experiment without data latency). Fig. 49 shows the RMS fit to observations between the OW and hourly_gdas and suggests that there is room for improvement in such a system solely by reducing the real-world data latency problem, but overall, the OW works well for mitigating the data latency issue.

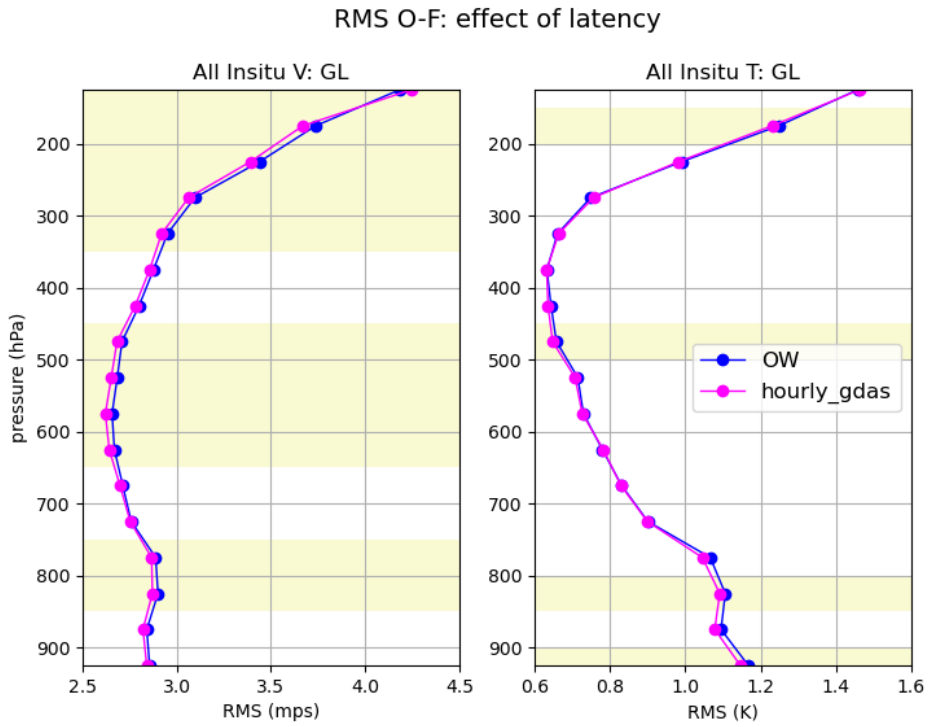

Fig. 49. Vertical profiles of short forecast comparisons with respect to global in-situ observations of vector wind (left) and temperature (right) for OW (blue) and hourly_gdas (pink) experiments in the full-input hybrid-gain LETKF setup, averaged over the period 1200 UTC 13 Mar 2020 – 1800 UTC 22 Mar 2020, to demonstrate the effects of data latency in an hourly system. Yellow shading denotes where the two experiments differ significantly at the 95% level.

4.3.4 Sensitivity Tests Using Radial Wind Observations

To investigate the impacts of radial wind observations in the GRR, we look at the impacts on forecast precipitation, track error, and intensity error of the TC event.

Fig. 50 shows a snapshot of the two simulations of TC Fred (Aug 2021) as it is making landfall (a) without radial wind assimilation (the control), (b) with radial wind assimilation, and (c) the associated increments from the second pass of the GSI (effects of only radial wind assimilation). Both experiments were cold started from the 6-hourly GDAS ICs and began fully cycling at 1-hourly intervals 1 h before TC Fred made landfall.

The location of the central low pressure in the control has a larger displacement error than that in the radial wind experiment (marked by the black "x" in Fig. 50a, b respectively compared against the blue dot NHC time/space interpolated best track). This difference in the storm location in each experiment is the result of hourly updating with and without radial wind observations. Comparing the surface pressure increments between the first and second pass of the GSI (Fig. 50c), we can confirm that the improved track forecast is the result of the assimilation of the radial wind observations. The surface pressure increments reveal that there is a deepening of the storm toward the location of the NHC best track and rising surface pressure near the location of the storm in the background. As a result of this, we notice a degradation of the storm intensity forecast in the radial wind experiment. It is noted that there is a relatively broader (weaker) storm in Fig. 50b noted by the elongated surface pressure contours which is a byproduct of radial wind observations relocating the position of the storm.

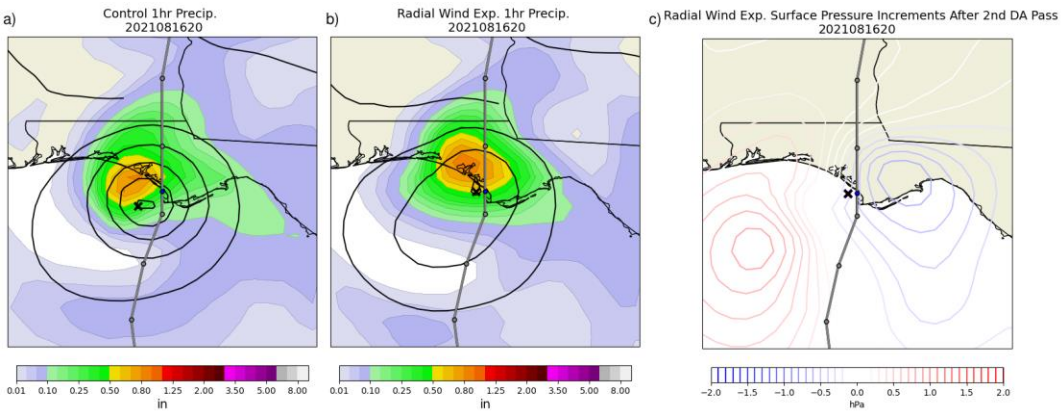

Fig. 50. Snapshot of TC Fred making landfall with surface pressure contours and one hour precipitation (a and b) with increments due to radial wind (second pass of GSI) (c). Blue increment contours indicate the radial wind assimilation is deepening the surface pressure at that location. The gray dotted line indicates the 6-hourly NHC best track of the TC where the blue dot indicates the time/space interpolated 1-hourly NHC best track of the TC. The black "x" indicates the location of the central low in the model forecast.

To further investigate the effects of the radial wind assimilation, a time series of the intensity and track error of the control and radial wind experiment is shown by Fig. 51. Both experiments have a negative intensity bias for the TC through almost all of the simulated period. As the TC begins making landfall (denoted by the black vertical line at 1900 UTC on 16 August 2021), there is a clear degradation in the intensity error (and much weaker storm by about 7 hPa in the experiment (solid red) than in the control (solid blue). This is hypothesized to be due to the relocation of the storm by the radial winds since there is a sharp decrease in track error in the experiment (dashed red) after making landfall relative to the control (dashed blue). This aligns with our interpretation that the radial winds are moving the storm location closer to the observed track. The track error is reduced by nearly 30km at the beginning of the experiment. After the 7th DA cycle (at 0000 UTC 2020 Aug 17), the radial wind experiment begins to have a consistently improved intensity forecast relative to the control, but still often too low compared to the observed storm intensity.

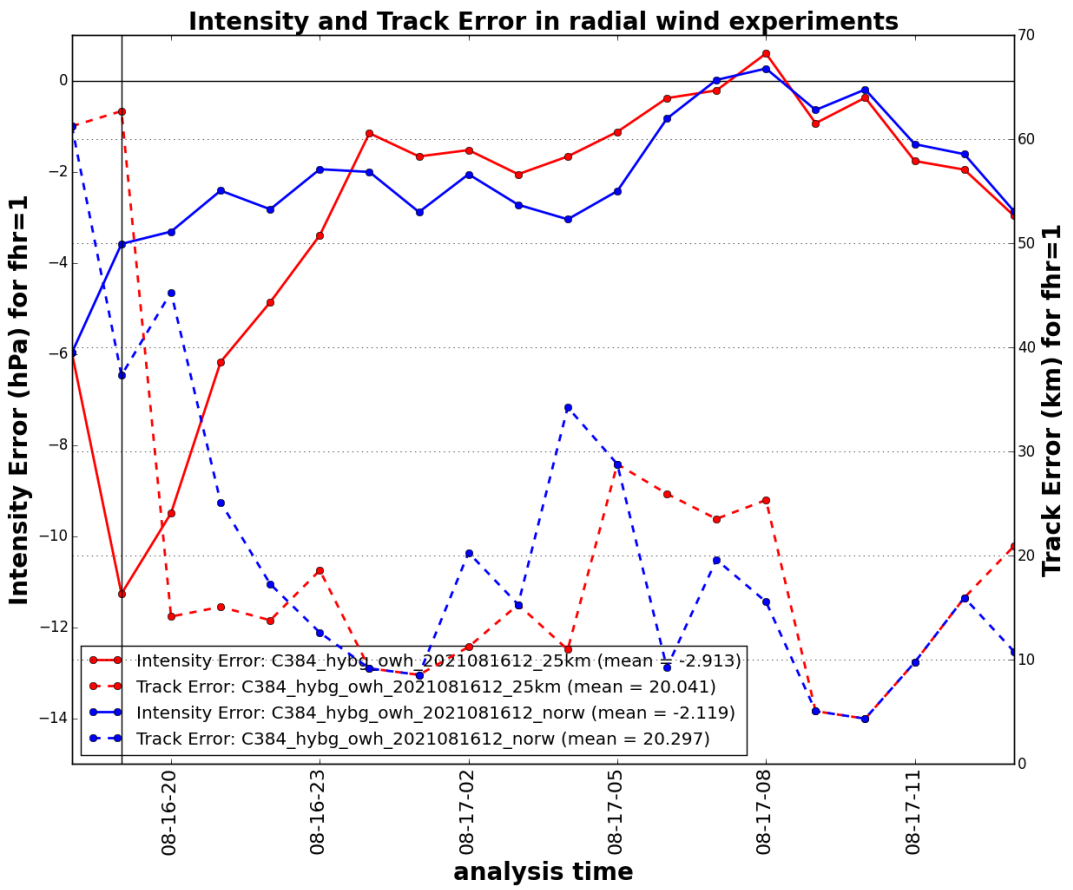

Fig. 51. Time series of intensity (solid) and track (dashed) error of control (blue) and radial wind experiment (red) compared to the NHC best track estimate for 1-h forecast. Intensity error is computed as observation minus background thus negative values indicates the forecasted storm has higher surface pressure (weaker storm) than observed. The vertical line indicates the moment of landfall.

To investigate the sensitivity of assimilating the radial wind observation on other aspects of the forecast, the control and radial wind experiment ensemble mean precipitation totals are also compared to the total stage IV precipitation (Lin and Mitchell 2005) observations (Fig. 52). A good precipitation forecast requires all aspects of the forecast model state (including TC location and intensity) coming together correctly and is therefore a good indicator for overall forecast improvement. The sensitivity to precipitation largely appears in the regions of heaviest precipitation, especially over the Florida panhandle, Georgia, and South Carolina – all locations that correspond to the TC path. There is a noticeable increase in precipitation for the radial

wind experiment relative to the control near the point of landfall in the Florida panhandle where the observations show a small area that reached 5.5+ in. of precipitation. The sensitivity to the precipitation is subjectively better in many areas. Based on these results, more cases should be run for more statistically robust results and more thorough analyses.

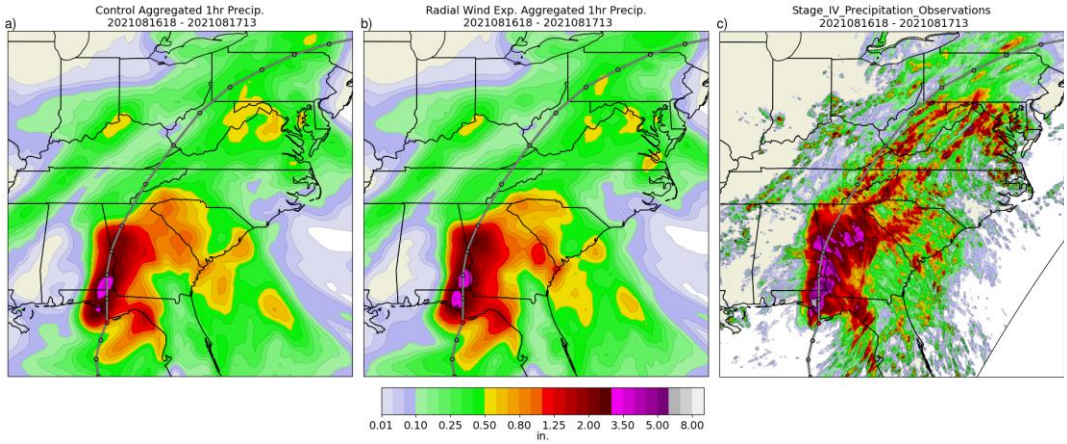

Fig. 52. Total 1-hour precipitation forecasts (ensemble mean) aggregated over the period 18 Aug 1800 UTC – 17 Aug 1300 UTC 2021 (TC Fred Landfall): a) control, b) GRR with radial wind assimilation, c) Stage IV observations. NHC best track is denoted by the gray dotted line.

4.4 Summary, Conclusions, and Future Work

A global hourly data assimilation system using an asymmetric, overlapping assimilation window method is tested in the GFSv15 at reduced resolution and using the hybrid-gain 3DVar/LETKF data assimilation algorithm. In Slivinski et al. (2022) we found that even with minimal system tuning, the overlapping windows (OW) method improved upon the 6-hourly assimilation system at least within the context of the 1-4 h background fit to observations of in-situ observations. Furthermore, we found that the OW system did not significantly degrade the 5-day forecasts. These results, however, were only evaluated over a one-month period and the long-term forecasts

were issued only once per day, so longer experiments would be needed to have more robust results.

Our initial attempts at implementing overlapping windows with 4DIAU (OW-IAU) were unsuccessful which have been fully documented here. Since our current implementation of the OW system does not use 4DIAU we have investigated the resulting impacts on the gravity wave noise. We have found that our OW system using the hybrid-gain assimilation method has a slight increase in gravity wave noise without the 4DIAU but does not degrade the performance of the short-term fit to observations. This is due to the hourly cycling having smaller analysis increments relative to a 6-hourly system, in which smoothing via 4DIAU or otherwise would be required.

The OW system was also compared against the catch-up cycles method which reinitializes an hourly system at regular intervals (6-hourly intervals were used in this study). The catch-up cycle method had comparable results to the 6-hourly control likely because resetting the state every 6 hours removes any benefits that hourly cycling might have had. Furthermore, the hourly cycling used 1-hour (i.e., [-0:30, +0:30; +0:26] h) symmetric assimilation windows which misses a lot of high latency data until the state can be reinitialized by the 6-hourly system to get those high latency data into the system.

Some additional experiments were also performed which split GFS and GDAS observation files into 1-hourly binned files ignoring data latency. These experiments suggest that more frequent assimilation cycles in concert with the additional data that the OW can assimilate are the reasons that the OW method outperforms the 6-hourly control and the hourly catch-up cycles method. The OW can more effectively use

observations relative to the 6-hourly system because it uses observations within 1-hour of their receipt time. This is particularly motivating for future work to incorporate high-frequency observations within an hourly overlapping windows system.

Overall, the OW method is a promising technique that should be considered in future work in both global and regional applications. More parameter tuning for inflation and localization, static background error, and bias correction is needed in future work and will likely improve results beyond what has been found in this study. It would also be worthwhile to test the OW method in the Rapid Refresh Forecast System (RRFS), which is NOAA's planned next-generation rapidly-updating convection-allowing ensemble forecast system, and compare against the partial cycling (i.e., catchup cycling) algorithm as it is currently implemented.

We need to further investigate the issue of the gravity wave noise imbalances and look for more solutions for implementing IAU. What we have found is that *the IAU can only be implemented in a forecast system in which the length of the IAU period is at most equal to the frequency of the cycling* (e.g., a 1-hour IAU period for 1-hourly updates). To properly implement OW-4DIAU, it would require sub-hourly IAU increments over a 1-hour period. For a 1-hourly system with 1-hourly background files, only OW-3DIAU (requiring a single IAU increment) can be implemented and is the next logical choice to try.

The overlapping windows technique developed here is a generic framework which could also be extended for use with sub-hourly updates. In a sub-hourly system, say with 30-min updates, it would likely still be required to maintain the -3 h backwards

looking overlapping window to mitigate data latency which would also require twice the number of background fields as the 1-hourly system.

Evaluating the assimilation of new and currently unused high-frequency observational data sets would also be worthwhile. For example, additional radial wind experiments (and other radar observations), as well as hourly tropical cyclone vitals (TC vitals), which uses the advanced Dvorak technique (ADT; Olander and Velden 2019), and high resolution atmospheric motion vectors (AMVs) could be tested and is left for future work.

An experiment was conducted where the assimilation of a high-frequency observations type that is new [to] the GFS: Doppler radial winds. Real data experiments were performed during a landfalling tropical cyclone case and the track and intensity errors were evaluated during the forecast periods. The control consisted of an upgraded version of the GRR as used in Slivinski et al. (2022). Both configurations used the 1-hourly data assimilation with overlapping windows using the hybrid gain data assimilation method with an 80-member ensemble run at C384 with no high resolution control member.

From this radial wind experiment, we have found that the track and intensity forecasts of landfalling tropical cyclones are sensitive to the assimilation of radial winds. The assimilation of radial wind observations was shown to reduce the initial track errors of a landfalling TC while also degrading the TC intensity (creating a weaker storm). After a few cycles, the storm was better located, and the intensity forecast was improved relative to the control. Furthermore, there is also some sensitivity to the precipitation forecasts mostly in areas with the highest amounts. In

precipitation, the radial wind experiment showed slight increases and had subjectively better placement and quantities relative to the control as measured against the Stage IV precipitation observations. These results, which are based on a single case study, motivate testing with additional cases for a more statistically robust assessment as a part of future development.

Overall, the assimilation of radial wind assimilation in the 1-hourly global overlapping windows system is promising and could help improve the track and intensity forecasts of landfalling tropical cyclones. An important task for the future of global radial wind assimilation will be to obtain radar observations from outside the U.S. For example, the Caribbean radar network would be of high priority for the U.S. since many TCs that make landfall to the U.S., at some point, pass over many of the Caribbean islands which will help in better initializing the storm locations and intensities within the model before making landfall to the U.S. thereby providing better forecast guidance to users.

Chapter 5: Conclusion

This chapter will conclude this dissertation by summarizing the key research findings as well as discussing their value and contribution to the science. It will also propose opportunities for future study.

5.1 Summary

Global forecast systems will continue to progress to higher spatial and temporal resolutions required to resolve convective-scale phenomena. Global models currently generate ensembles with relatively coarse resolution, while regional models produce high-resolution ensembles but require lateral boundary conditions from global systems. Skillful multi-scale forecasts require forecasts with high accuracy at synoptic and global scales and frequent assimilation of convective-scale observations such as radial winds (among other types) to provide accurate initial conditions for predicting phenomena such as fronts, convective storms, tropical cyclones, etc. Therefore, it is critical to improve the use of radial winds, which is but one type of convective-scale observation, in both regional and global systems.

Doppler radial winds have been an underutilized observation in U.S. operational forecast systems. This has typically been due to limitations in formulation of the observation operator, the amount of data thinning via super-obbing, or exclusion from assimilation in global modeling systems. In this work we explored some of the more feasible aspects of radial wind assimilation that could more readily be applied to the operational systems with the main goal to improve the use of radial winds in the operational forecast systems used by NOAA. It was shown that improvements could

be made to systems not only operating at the convective scale, but also that global systems could benefit from the assimilation of radial winds. Experiments featuring results from the regional NAM version 4 forecast system along with the GFS versions 15 and 16 were shown. The GFS experiments explore the potential radial wind impact via hypothetical observing networks tested with observing system simulation experiments. We further extended this study to a real-data case with a land falling tropical cyclone event in a novel, rapidly-updated version of the GFS.

This work is a first step toward improving the use of radial wind observations and tests their use for the very first time in the GFS. This effort demonstrates potential for radial wind assimilation in the GFS, which may become increasingly important as update cadences and spatial resolution are increased.

5.2 Relationship with Previous DA Systems with Long Assimilation Windows

Atmospheric DA has long been an optimization problem which has relied on the balance of waiting as long as possible (i.e., the so called "cut-off" time) for late arriving data and making the most use out of efficient, and less accurate, algorithms to be able to meet the timely dissemination requirements within the short timeframe imposed by the late cut-off time. For example, operational systems have relied on a two-stage suite consisting of an early-delivery (e.g., GFS) and a delayed cut-off (e.g., GDAS) paradigm. In this study, we investigated the methodology of overlapping windows for a rapidly updating (1-hourly) cycles in global DA which allows for the assimilation of late arriving data while also being able to update more frequently. This method removes the requirement for a two-stage system and the need to have a long cut-off time which was previously required to assimilate as many observations as

possible. The overlapping windows technique is inherently always waiting for late observations using observations as soon as they become available. There are clear benefits to having more up-to-date and accurate forecasts in the hours between the standard synoptic times (e.g., Payne 2017; Slivinski et al. 2022). However, the increased frequency of assimilation still suffers from having to complete the analysis and forecast steps within a short timeframe but not as a result of late cut-off time. Therefore, efficient algorithms are required in this framework. One of the main challenges will be in the scalability of the problem and to make the algorithms more efficient.

Previous studies, such as the continuous DA methodology for global NWP at the ECMWF by Lean et al. (2021), have shown another method for overcoming data latency and issuing timely forecasts that also eliminates the need for the two-stage system. This system combines the addition of new observations in later outer loops (i.e., a continuous data stream) with an extended assimilation window (e.g., 12 h long at ECMWF). One main advantage of the continuous DA is that it allows for the use of more expensive and accurate atmospheric DA configurations and so they were able to increase the number of outer loops in their 4DVar configuration while still able to deliver timely forecasts. Under this framework, instead of having more frequent updates, the consideration is on the improvement of the analyses and forecasts at the usual synoptic times by feeding in recent observations during an iteration of outer-loop analysis calculations.

For global models, the benefits of using 4DVar outweigh any negative impacts of using the linear models (Buehner et al. 2010). For high resolution, limited-area

models, and presumably global models as they progress to 3-km grid spacing, Gustafsson and Bojarova (2014) suggests that 4DEnVar (over En4DVar) to provide the best results owing to the high nonlinearity of the convective-scale problem. It would be worthwhile to compare the relative performances of OW with the continuous DA/long assimilation window method.

Overcoming the issue of data latency is an important and practical aspect for operational systems; however, there are also more general differences that could be considered for long vs. short assimilation windows. For example, Fisher et al. (2005) showed that the quality of the 4DVar analyses could be improved by extending the assimilation window (on the order of up to about ten days) to produce analyses equal in quality to those of the EKF (at least in a simplified model; Lorenz 1996). Fisher et al. (2005) also showed that the current ECMWF operational 4DVar system retains memory of earlier observations and states over a period of between four to ten days; however, an assimilation window of such length would not be computationally viable for operational purposes, at least not in the near future. For reference, the ECMWF currently uses a 12 h assimilation window. An important aspect to consider for defining the assimilation window length is the assumption of linearity as this will impose some theoretical limits to how long the window can be even for large scale cases. While such a system would not yet be useful for operational purposes, having the ability to run such a system (run at the optimal window length; whatever that may be) would be valuable for assessing the relative performance of the suboptimal and more efficient algorithms used for meeting the needs of operational centers.

5.3 Future Work

The science of the assimilation of radial wind in operational systems is not complete; there is still much work to be done in both regional and global systems. The results of this dissertation show that furthering this part of the science will be challenging but likely worthwhile. There are many aspects to the NWP systems that could be updated to improve the use of not only radial wind observations, but also other convective-scale observations. Some ideas for future work are discussed below.

One of the most important aspects of data assimilation is the specification of the BEC. In this dissertation, the BEC was not re-computed to include vertical velocity or re-calibrated for the specific applications used (e.g., the GRR) and would be an important aspect to consider in the future. For example, in Chapter 2, the analysis of vertical velocity was only treated as a sink term due to the use of the NMMB dynamic core, and a constant value was prescribed for the vertical velocity in the BEC. With the transition to the FV3, as the unified dynamic core across both regional and global systems, the analysis of vertical velocity could play a larger role than found in the experiments within this dissertation. This would also require computing the BEC statistics for vertical velocity. The review paper Bannister (2008) provides information on methods for error covariance modeling including the NMC method which has been the method previously used by NCEP.

The NMC method (Parrish and Derber 1992) has been a popular choice for estimating the background error by using differences between forecasts at varied lengths that verify at the same time (usually 24 and 48 h forecast pairs). Usually, these lagged forecast pairs vary in length by 24 hours to reduce the chance of the model's

diurnal cycle being incorrectly interpreted as background errors. Furthermore, the short forecast is typically set to be at least 12 h to reduce the risk incorporating spin up effects into the estimate. Fisher (2003) discussed limitation of the NMC method may underestimate the variance of the background error in poorly observed regions since few observations may lead to minimal change in the later forecast and thus resulting in similar forecasts. Furthermore, the 24 and 48 h forecasts used for estimating the background error are significantly longer than the 6 h (or 1 h) background used in the analysis.

Wang and Wang (2021) used the analysis-ensemble method of Fisher (2003) to develop a convective-scale static BEC. The analysis-ensemble method essentially propagates multiple ensemble members through an analysis/forecast system over many assimilation cycles (e.g., 1 month) using perturbed observations according to their error statistics. The result are pairs of perturbed background fields valid at the same time of forecast length equal to that of the backgrounds used in the analysis system from which to compute forecast differences. This method is less susceptible to problems related to poorly observed regions by not using lagged forecast pairs and has the advantage over the NMC method due to using forecasts of the same lead time as used for the background (e.g., 1 or 6 h) in the analysis system (Bannister 2008).

The convective scale BEC of Wang and Wang (2021) was created to improve the direct assimilation of radar reflectivity and thus includes an updated set of control variables such as reflectivity and all hydrometeor variables as well as vertical velocity and meridional and zonal wind for the horizontal momentum control variables. Cross-variable correlations have been included with this BEC. Future studies should consider

the impacts of this updated BEC will make for convective-scale DA especially in the context of radial wind assimilation.

In addition to the BEC, we must also consider the observation error statistics. Observation errors are typically assumed uncorrelated, and observations are thinned and super-obbed to help satisfy that assumption. In the future, if we want to make the best use of these high-resolution, high-frequency observations, one aspect will be to revisit the observation error statistics since they need to be accurately estimated in the assimilation system. One aspect might be to include correlated observation errors, which are errors that arise from an imperfect observation operator, errors of representativity (i.e., errors due to the observations being able to represent scales that the model cannot), and preprocessing errors (e.g., super-obbing). Waller et al. (2016) showed that radial wind observation errors have large spatial correlations and are much larger than the operational thinning distance. Therefore, in order to assimilate this high-resolution observations correctly, the inclusion of correlated observation errors in the assimilation system would be important. The inclusion of correlated observation errors allows for less thinning of the data and thus assimilating the observations more effectively. Further revisions of the thinning/super-obbing techniques and observation quality control would need to follow.

In Chapter 2, one of the main focuses was on updating the observation operator. The observation operator was updated to include vertical velocity; however, it still contains many simplifications (Fabry 2010). One part of the radial wind observation routine used in the GSI is the estimation of the radar beam at a given distance from the radar. The current method assumes standard atmospheric refraction (i.e., the 4/3rds

rule). The alternative to using this approximation would require the vertical profile of refractive index. There have been studies which seek to estimate the low level refractivity profile using radars and point targets (e.g., Feng et al. 2016). In the future, the 4/3rds rule could be modified by a correction factor for improving the height estimation of radar observations by relying on known heights of point targets.

In Chapter 3, an identical-twin OSSE assimilating only radial winds was used. We found that discrete boundaries associated with Doppler radar networks present a potential for aliasing. Aliasing is the misinterpretation of a signal frequency, perhaps through partial sampling, which then results in an incorrect increment in the analysis. Another aspect of this is likely the misrepresentation of the background error. This phenomenon might not have been an issue if the full suite of observations were included in the OSSE, which may obviate issues with signal aliasing associated with discrete observation networks having definitive boundaries. Regardless, this exercise exposed potential future problems for radial wind assimilation that should be an area of focus. Early results with a prototype regional data assimilation system suggests similar issues may indeed be present in real-data, full observation applications. We therefore recommend further tests with a full suite of observations to confirm if such an issue persists as an area of future work. This exercise has also brought into view other deficiencies in the data assimilation systems that could be improved. For example, a larger ensemble could allow for better utilization of observations of partially sampled features.

There are over 800 radars in the World Radar Database (WRD) [Available online: https://wrd.mgm.gov.tr/Home/Wrd], yet there is no global sharing of this data.

This data, which exists today, could be shared among national centers for the purposes of assimilation. Based on results from this study, such data has the potential to improve global forecasts systems, perhaps especially when such systems began to operate at convective scales. Sharing this data, however, presents a huge problem in terms of data collection, transmission, processing, storage, etc., yet in an era of cloud computing this should be addressable. In the short term, for the TC problem, obtaining sharing agreements from the Caribbean Islands would likely prove to be most useful to the U.S. for the assimilation of that data for improving the forecast of TCs prior to making landfall. Further testing could also be done to explore the benefits of assimilating radial winds within global models at higher resolution (e.g., 3 km global resolution) with consideration for impacts on precipitating systems.

This work highlighted a major advancement in global modeling (the GRR; Chapter 4) which featured hourly updating (via overlapping assimilation windows; OW) which would play a major role in making better use of all observations. The OW approach could be extended for use in a few different ways. First, it could be applied to the regional systems such as the RRFS to use observations more effectively and overcome the issue of data latency. Second, it could be extended to sub-hourly DA, which has shown great benefit for the assimilation of radial wind in non-operational systems (Gao and Stensrud 2014; Jones et al. 2015; Jones et al. 2016; Lawson et al. 2018; Wheatley et al. 2015). Practical, real-time implementation of such systems would likely require improvements in the NWP software components and access to hardware to support such advancements. In this work, we were unable to implement an OW configuration using IAU which is meant to help constrain imbalances in the analysis.

As we found, in our current configuration of the OW, the initial imbalances tended to dissipate over some time; however, it might still be important to include methods to help improve initialization in a future system, such as IAU. As previously noted, *the IAU can only be implemented in forecast systems in which the length of the IAU period is at most equal to the frequency of the cycling* (i.e., a 1-hour IAU period for a 1-hourly update). Early work has shown that implementing 3DIAU (having one increment to force the IAU) may be effective. A 3DIAU technique would involve using an averaged increment between the nominal analysis time and the analysis at 1-hour after the nominal analysis time to represent the increment at the center of that 1-hour time window. Then, using a 1-hour IAU period starting from the nominal analysis time, apply the increment forcing at 30 min following the analysis time. On the other hand, 4DIAU could be implemented in the 1-hourly system so long as there are sub-hourly increments. The filtering properties of such implementations would be different from that which is used in the current 6-hourly system; however, such configurations might still be worthwhile.

In addition to using radar networks that currently exist, such as the full GLOBAL network in Chapter 3, emerging radar technologies could further enhance the current networks such as phased array radar (PAR; Zrnic et al. 2007). PARs do not use a rotating antenna, and thus can provide volume scans much more frequently than with the WSR-88Ds. Furthermore, gap filling technologies such as bi-static radar (Rinehart 2004; p. 237; Wurman 1994) could be used to enhance the current coverage of radars. Over the U.S., the average spacing is about 230km where radial winds are only used out to 100km, so there are many regions that go unobserved, especially near

the ground and in complex terrain. The bistatic radar simply uses additional receivers around a central radar to collect the energy scattered in the direction of the receivers. Since bistatic radars is comprised of only additional receivers, this makes them more cost effective and easier to maintain than complete radar systems such as the WSR-88D or PAR systems. Employing such radar technologies would greatly enhance the current radar observing networks and provide better and more robust data coverage. While having the best data is important, it is also just as important to be able to use that data effectively; therefore, all the techniques and ideas proposed will become more relevant as our observing network is improved.

This dissertation was motivated by the successes of the assimilation of radial winds in previous idealized work (3-km grid spacing or less with sub-hourly updates), and yet often unclear impacts in the current operational systems (3-km grid spacing with 1-hourly updates). The main goal of this dissertation was to understand how the assimilation of radial wind observations could lead to improvements in the operational regional and global systems. There is still much challenging work to do on the front of radial wind assimilation in the operational systems beyond what was covered in this dissertation; however, this work will help point us in the right direction for global and regional applications as these highly dense, convective-scale observations become increasingly prevalent.

Bibliography

Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A Multicase Comparative Assessment of the Ensemble Kalman Filter for Assimilation of Radar Observations. Part I: Storm-Scale Analyses. *Monthly Weather Review*, **137,** 1805-1824.

Alexander, C., and Coauthors, 2020: Rapid Refresh (RAP) and High Resolution Rapid Refresh (HRRR) Model Development. *100th American Meteorological Society Annual Meeting.*

Alpert, J. C., and V. K. Kumar, 2007: Radial Wind Super-Obs from the WSR-88D Radars in the NCEP Operational Assimilation System. *Monthly Weather Review,* **135,** 1090-1109.

Armijo, L., 1969: A Theory for the Determination of Wind and Precipitation Velocities with Doppler Radars. *Journal of Atmospheric Sciences*, **26,** 570-573.

Bannister, R. N., 2008: A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances. *Quarterly Journal of the Royal Meteorological Society*, **134,** 1951-1970.

——, 2017: A review of operational methods of variational and ensemble-variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **143,** 607-633.

Battan, L., J., 1973: Radar observation of the atmosphere. L. J. Battan (The University of Chicago Press) 1973. PP X, 324; 125 figures, 21 tables. £7·15. *Quarterly Journal of the Royal Meteorological Society*, **99,** 793-793.

Benjamin, S. G., and Coauthors, 2004: An Hourly Assimilation–Forecast Cycle: The RUC. *Monthly Weather Review*, **132,** 495-518.

Benjamin, S. G., and Coauthors, 2016: A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Monthly Weather Review,* **144,** 1669-1694.

Bieringer, P., and P. S. Ray, 1996: A Comparison of Tornado Warning Lead Times with and without NEXRAD Doppler Radar. *Weather and Forecasting*, **11,** 47-52.

Bishop, C. H., and E. A. Satterfield, 2013: Hidden Error Variance Theory. Part I: Exposition and Analytic Model. *Monthly Weather Review*, **141,** 1454-1468.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Monthly Weather Review*, **129,** 420-436.

Biswas, M. K., and Coauthors, 2018: Hurricane Weather Research and Forecasting (HWRF) Model: 2017 Scientific Documentation(No. NCAR/TN-544+STR).

Bluestein, H. B., K. J. Thiem, J. C. Snyder, and J. B. Houser, 2019: Tornadogenesis and early tornado evolution in the El Reno, Oklahoma, supercell on 31 May 2013. *Monthly Weather Review*, **147,** 2045-2066.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part I: Description and Single-Observation Experiments. *Monthly Weather Review*, **138,** 1550-1566.

Burgers, G., P. Jan van Leeuwen, and G. Evensen, 1998: Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review*, **126,** 1719-1724.

Carley, J. R., 2012: Hybrid ensemble-3DVar radar data assimilation for the short-term prediction of convective storms, Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, 205 pp.
Caya, A., J. Sun, and C. Snyder, 2005: A Comparison between the 4DVAR and the Ensemble Kalman Filter Techniques for Radar Data Assimilation. *Monthly Weather Review*, **133,** 3081-3094.
Chen, J.-H., and S.-J. Lin, 2013: Seasonal Predictions of Tropical Cyclones Using a 25-km-Resolution General Circulation Model. *Journal of Climate*, **26,** 380-398.
Chen, Y., and C. Snyder, 2007: Assimilating Vortex Position with an Ensemble Kalman Filter. *Monthly Weather Review*, **135,** 1828-1845.
Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, **120,** 1367-1387.
Crum, T. D., and R. L. Alberty, 1993: The WSR-88D and the WSR-88D Operational Support Facility. *Bulletin of the American Meteorological Society*, **74,** 1669-1688.
Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press.
Dawson, D. T., L. J. Wicker, E. R. Mansell, and R. L. Tanamachi, 2012: Impact of the environmental low-level wind profile on ensemble forecasts of the 4 May 2007 Greensburg, Kansas, tornadic storm and associated mesocyclones. *Mon. Wea. Rev.*, **140,** 696-716.
De Pondeca, M. S. F. V., and Coauthors, 2011: The Real-Time Mesoscale Analysis at NOAA's National Centers for Environmental Prediction: Current Status and Development. *Weather and Forecasting*, **26,** 593-612.
Derber, J., and A. Rosati, 1989: A global oceanic data assimilation system. 1333-1347.
Djalalova, I. V., and Coauthors, 2016: The POWER Experiment: Impact of Assimilation of a Network of Coastal Wind Profiling Radars on Simulating Offshore Winds in and above the Wind Turbine Layer. *Weather and Forecasting*, **31,** 1071-1091.
Dong, J., M. Xue, and K. Droegemeier, 2011: The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteor. Atmos. Phys.*, **112,** 41-61.
Dowell, D. C., and L. J. Wicker, 2009: Additive Noise for Storm-Scale Ensemble Data Assimilation. *Journal of Atmospheric and Oceanic Technology*, **26,** 911-927.
Dowell, D. C., L. J. Wicker, and C. Snyder, 2011: Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.*, **139,** 272-294.
Dowell, D. C., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and Temperature Retrievals in the 17 May 1981 Arcadia, Oklahoma, Supercell: Ensemble Kalman Filter Experiments. *Monthly Weather Review*, **132,** 1982-2005.
Dowell, D. C., and Coauthors, 2022: The High-Resolution Rapid Refresh (HRRR): An Hourly Updating Convection-Allowing Forecast Model. Part 1: Motivation and System Description. *Weather and Forecasting*.
DTC, 2018: The Developmental Testbed Center (DTC) Annual Operational Plan (AOP) 2017 Project for the National Oceanic and Atmospheric Administration (NOAA) Office of Oceanic and Atmospheric Research (OAR), 87 pp.
Duda, J. D., X. Wang, Y. Wang, and J. R. Carley, 2019: Comparing the Assimilation of Radar Reflectivity Using the Direct GSI-Based Ensemble–Variational (EnVar) and

Indirect Cloud Analysis Methods in Convection-Allowing Forecasts over the Continental United States. *Monthly Weather Review*, **147,** 1655-1678.
Einicke, G. A., and L. B. White, 1999: Robust extended Kalman filtering. *IEEE Transactions on Signal Processing*, **47,** 2596-2599.
Eliassen, A., 1954: Provisional report on calculation of spatial covariance and autocorrelation of the pressure field. Rep. 5, 12 pp.
Emerick, A. A., and A. C. Reynolds, 2012: History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations. *Computational Geosciences*, **16,** 639-659.
Errico, R. M., and N. C. Privé, 2018: *Some general and fundamental requirements for designing observing system simulation experiments (OSSEs).* Vol. 8, WMO Rep. WWRP 2018-, 33 pp.
Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, **99,** 10143-10162.
Evensen, G., and P. J. van Leeuwen, 2000: An Ensemble Kalman Smoother for Nonlinear Dynamics. *Monthly Weather Review*, **128,** 1852-1867.
Fabry, F., 2010: Radial velocity measurement simulations: Common errors, approximations, or omissions and their impact on estimation accuracy. *Proc. Sixth European Conf. on Radar in Meteorology and Hydrology*, ERAD, 17.2.
Fabry, F., and V. Meunier, 2020: Why Are Radar Data so Difficult to Assimilate Skillfully? *Monthly Weather Review*, **148,** 2819-2836.
Feng, Y.-C., F. Fabry, and T. M. Weckwerth, 2016: Improving Radar Refractivity Retrieval by Considering the Change in the Refractivity Profile and the Varying Altitudes of Ground Targets. *Journal of Atmospheric and Oceanic Technology*, **33,** 989-1004.
Fisher, M., 2003: Background error covariance modelling. *Seminar on Recent developments in data assimilation for atmosphere and ocean, 8-12 September 2003*, Shinfield Park, Reading, ECMWF.
Fisher, M., M. Leutbecher, and G. A. Kelly, 2005: On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **131,** 3235-3246.
Flora, M. L., P. S. Skinner, C. K. Potvin, A. E. Reinhart, T. A. Jones, N. Yussouf, and K. H. Knopfmeier, 2019: Object-Based Verification of Short-Term, Storm-Scale Probabilistic Mesocyclone Guidance from an Experimental Warn-on-Forecast System. *Weather and Forecasting*, **34,** 1721-1739.
Gal-Chen, T., 1978: A Method for the Initialization of the Anelastic Equations: Implications for Matching Models with Observations. *Monthly Weather Review*, **106,** 587-606.
Gandin, L. S., 1963: *Objective Analysis of Meteorological Fields.* Israel Program for Scientific Translations, 1965, 285 pp.
Gao, J., and D. J. Stensrud, 2014: Some Observing System Simulation Experiments with a Hybrid 3DEnVAR System for Storm-Scale Radar Data Assimilation. *Monthly Weather Review*, **142,** 3326-3346.

Gao, J., M. Xue, K. Brewster, and K. K. Droegemeier, 2004: A Three-Dimensional Variational Data Analysis Method with Recursive Filter for Doppler Radars. *Journal of Atmospheric and Oceanic Technology*, **21,** 457-469.
Ge, G., J. Gao, and M. Xue, 2012: Diagnostic Pressure Equation as a Weak Constraint in a Storm-Scale Three-Dimensional Variational Radar Data Assimilation System. *Journal of Atmospheric and Oceanic Technology*, **29,** 1075-1092.
Ge, G., J. Gao, K. Brewster, and M. Xue, 2010: Impacts of Beam Broadening and Earth Curvature on Storm-Scale 3D Variational Data Assimilation of Radial Velocity with Two Doppler Radars. *Journal of Atmospheric and Oceanic Technology*, **27,** 617-636.
Geer, A. J., 2016: Significance of changes in medium-range forecast scores. *Tellus A: Dynamic Meteorology and Oceanography*, **68,** 30229.
Gustafsson, N., 2007: Discussion on '4D-Var or EnKF?'. *Tellus A: Dynamic Meteorology and Oceanography*, **59,** 774-777.
Gustafsson, N., and J. Bojarova, 2014: Four-dimensional ensemble variational (4D-En-Var) data assimilation for the HIgh Resolution Limited Area Model (HIRLAM). *Nonlin. Processes Geophys.*, **21,** 745-762.
Gustafsson, N., and Coauthors, 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quarterly Journal of the Royal Meteorological Society*, **144,** 1218-1256.
Hamill, T. M., and C. Snyder, 2000: A Hybrid Ensemble Kalman Filter–3D Variational Analysis Scheme. *Monthly Weather Review*, **128,** 2905-2919.
Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter. *Monthly Weather Review*, **129,** 2776-2790.
Hane, C. E., and B. C. Scott, 1978: Temperature and Pressure Perturbations within Convective Clouds Derived from Detailed Air Motion Information: Preliminary Testing. *Monthly Weather Review*, **106,** 654-661.
Harlim, J., and B. R. Hunt, 2005: Local ensemble transform kalman filter: An efficient scheme for assimilating atmospheric data. *preprint*.
Harris, L., X. Chen, L. Zhou, and J.-H. Chen, 2020: The Nonhydrostatic Solver of the GFDL Finite-Volume Cubed-Sphere Dynamical Core. *NOAA Technical Memorandum OAR GFDL*, **2020-003**.
Harris, L. M., and S.-J. Lin, 2013: A Two-Way Nested Global-Regional Dynamical Core on the Cubed-Sphere Grid. *Monthly Weather Review*, **141,** 283-306.
He, H., L. Lei, J. S. Whitaker, and Z.-M. Tan, 2020: Impacts of Assimilation Frequency on Ensemble Kalman Filter Data Assimilation and Imbalances. *Journal of Advances in Modeling Earth Systems*, **12,** e2020MS002187.
Heiss, W. H., D. L. McGrew, and D. Sirmans, 1990: Nexrad: next generation weather radar (WSR-88D). *Microwave Journal*, 79+.
Houtekamer, P. L., and H. L. Mitchell, 1998: Data Assimilation Using an Ensemble Kalman Filter Technique. *Monthly Weather Review*, **126,** 796-811.
——, 2001: A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review*, **129,** 123-137.
Hu, M., M. Xue, and K. Brewster, 2006: 3DVAR and Cloud Analysis with WSR-88D Level-II Data for the Prediction of the Fort Worth, Texas, Tornadic Thunderstorms. Part I: Cloud Analysis and Its Impact. *Monthly Weather Review*, **134,** 675-698.

Hu, M., S. G. Benjamin, T. T. Ladwig, D. C. Dowell, S. S. Weygandt, C. R. Alexander, and J. S. Whitaker, 2017: GSI Three-Dimensional Ensemble–Variational Hybrid Data Assimilation Using a Global Ensemble for the Regional Rapid Refresh Model. *Monthly Weather Review*, **145,** 4205-4225.
Huang, B., X. Wang, D. T. Kleist, and T. Lei, 2021: A Simultaneous Multiscale Data Assimilation Using Scale-Dependent Localization in GSI-Based Hybrid 4DEnVar for NCEP FV3-Based GFS. *Monthly Weather Review*, **149,** 479-501.
Huber, M., and J. Trapp, 2009: A Review of NEXRAD Level II: Data, Distribution, and Applications. *Journal of Terrestrial Observation*, **1,** 5-15.
Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, **230,** 112-126.
Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, 1997: Unified Notation for Data Assimilation : Operational, Sequential and Variational (gtSpecial IssueltData Assimilation in Meteology and Oceanography: Theory and Practice). *Journal of the Meteorological Society of Japan. Ser. II*, **75,** 181-189.
Ishikawa, Y., and K. Koizumi, 2006: Doppler radar wind data assimilation with JMA meso 4DVar. *WMO WGNE Research Activities in Atmosphere and Oceanic Modeling 2006*, Report No. 36, Section 01, 11-12.
James, E. P., and Coauthors, 2022: The High-Resolution Rapid Refresh (HRRR): An Hourly Updating Convection-Allowing Forecast Model. Part 2: Forecast Performance. *Weather and Forecasting*.
Janjić, T., and Coauthors, 2018: On the representation error in data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **144,** 1257-1278.
Janjić, Z. I., and R. Gall, 2012: Scientific documentation of the NCEP nonhydrostatic multiscale model on the B grid (NMMB). Part 1: Dynamics. . *NCAR Tech. Note***,** 74.
JDOP, 1979: Final report on the Joint Doppler Operational Project (JDOP) 1976-1978.
Johnson, A., X. Wang, J. R. Carley, L. J. Wicker, and C. Karstens, 2015: A Comparison of Multiscale GSI-Based EnKF and 3DVar Data Assimilation Using Radar and Conventional Observations for Midlatitude Convective-Scale Precipitation Forecasts. *Monthly Weather Review*, **143,** 3087-3108.
Jones, T. A., D. Stensrud, L. Wicker, P. Minnis, and R. Palikonda, 2015: Simultaneous Radar and Satellite Data Storm-Scale Assimilation Using an Ensemble Kalman Filter Approach for 24 May 2011. *Monthly Weather Review*, **143,** 165-194.
Jones, T. A., K. Knopfmeier, D. Wheatley, G. Creager, P. Minnis, and R. Palikonda, 2016: Storm-Scale Data Assimilation and Ensemble Forecasting with the NSSL Experimental Warn-on-Forecast System. Part II: Combined Radar and Satellite Data Experiments. *Weather and Forecasting*, **31,** 297-327.
Kalman, R. E., 1960: A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, **82,** 35-45.
Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press.
Kalnay, E., H. Li, T. Miyoshi, S.-C. Yang, and J. Ballabrera, 2007a: Response to the discussion on ``4-D-Var or EnKF?" by Nils Gustafsson. *Tellus Series A*, **59,** 778-780.

Kalnay, E., H. Li, T. Miyoshi, S.-C. Yang, and J. Ballabrera-Poy, 2007b: 4-D-Var or ensemble Kalman filter? *Tellus A: Dynamic Meteorology and Oceanography*, **59,** 758-773.
Kim, S.-M., and H. M. Kim, 2019: Forecast Sensitivity Observation Impact in the 4DVAR and Hybrid-4DVAR Data Assimilation Systems. *Journal of Atmospheric and Oceanic Technology*, **36,** 1563-1575.
Kleist, D. T., 2011: Assimilation of Tropical Cyclone Advisory Minimum Sea Level Pressure in the NCEP Global Data Assimilation System. *Weather and Forecasting*, **26,** 1085-1091.
Kleist, D. T., and K. Ide, 2015a: An OSSE-Based Evaluation of Hybrid Variational–Ensemble Data Assimilation for the NCEP GFS. Part II: 4DEnVar and Hybrid Variants. *Monthly Weather Review*, **143,** 452-470.
——, 2015b: An OSSE-Based Evaluation of Hybrid Variational–Ensemble Data Assimilation for the NCEP GFS. Part I: System Description and 3D-Hybrid Results. *Monthly Weather Review*, **143,** 433-451.
Kleist, D. T., R. Mahajan, and C. Thomas, 2018: Data assimilation in the Next-Generation Global Prediction System (NGGPS) Era: Initial implementation of FV3-based Global Forecast System (GFS). *JCSDA Quarterly*, **61**.
Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009a: Improving Incremental Balance in the GSI 3DVAR Analysis System. *Monthly Weather Review*, **137,** 1046-1060.
Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009b: Introduction of the GSI into the NCEP Global Data Assimilation System. *Weather and Forecasting*, **24,** 1691-1705.
Kurosawa, K., and J. Poterjoy, 2021: Data Assimilation Challenges Posed by Nonlinear Operators: A Comparative Study of Ensemble and Variational Filters and Smoothers. *Monthly Weather Review*, **149,** 2369-2389.
Lawson, J. R., J. S. Kain, N. Yussouf, D. C. Dowell, D. M. Wheatley, K. H. Knopfmeier, and T. A. Jones, 2018: Advancing from Convection-Allowing NWP to Warn-on-Forecast: Evidence of Progress. *Weather and Forecasting*, **33**.
Lawson, W. G., and J. A. Hansen, 2005: Alignment Error Models and Ensemble-Based Data Assimilation. *Monthly Weather Review*, **133,** 1687-1709.
Lean, P., E. V. Hólm, M. Bonavita, N. Bormann, A. P. McNally, and H. Järvinen, 2021: Continuous data assimilation for global numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, **147,** 273-288.
Lei, L., and J. S. Whitaker, 2016: A Four-Dimensional Incremental Analysis Update for the Ensemble Kalman Filter. *Monthly Weather Review*, **144,** 2605-2621.
Lei, L., J. S. Whitaker, and C. Bishop, 2018: Improving Assimilation of Radiance Observations by Implementing Model Space Localization in an Ensemble Kalman Filter. *Journal of Advances in Modeling Earth Systems*, **10,** 3221-3232.
Li, Y., X. Wang, and M. Xue, 2012: Assimilation of Radar Radial Velocity Data with the WRF Hybrid Ensemble–3DVAR System for the Prediction of Hurricane Ike (2008). *Monthly Weather Review*, **140,** 3507-3524.
Lin, K.-J., S.-C. Yang, and S. S. Chen, 2018: Reducing TC Position Uncertainty in an Ensemble Data Assimilation and Prediction System: A Case Study of Typhoon Fanapi (2010). *Weather and Forecasting*, **33,** 561-582.

Lin, S.-J., 2004: A "Vertically Lagrangian" Finite-Volume Dynamical Core for Global Models. *Monthly Weather Review*, **132,** 2293-2307.
Lin, Y., and K. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses: Development and applications.
Lin, Y., P. S. Ray, and K. W. Johnson, 1993: Initialization of a Modeled Convective Storm Using Doppler Radar–derived Fields. *Monthly Weather Review*, **121,** 2757-2775.
Lindskog, M., K. Salonen, H. Järvinen, and D. B. Michelson, 2004: Doppler Radar Wind Data Assimilation with HIRLAM 3DVAR. *Monthly Weather Review*, **132,** 1081-1092.
Lippi, D. E., D. T. Kleist, and J. R. Carley, 2016: Assimilation of radar radial velocity with GSI-based hybrid 3DEnVar system for improved precipitation forecasts, Dept. of Atmospheric and Oceanic Science, University of Maryland, 45 pp.
Lippi, D. E., J. R. Carley, and D. T. Kleist, 2019: Improvements to the Assimilation of Doppler Radial Winds for Convection-Permitting Forecasts of a Heavy Rain Event. *Monthly Weather Review*, **147,** 3609-3632.
Liu, C., M. Xue, and R. Kong, 2020a: Direct Variational Assimilation of Radar Reflectivity and Radial Velocity Data: Issues with Nonlinear Reflectivity Operator and Solutions. *Monthly Weather Review*, **148,** 1483-1502.
Liu, Q., and Coauthors, 2020b: Vortex Initialization in the NCEP Operational Hurricane Models. *Atmosphere*, **11,** 968.
Liu, S., M. Xue, J. Gao, and D. Parrish, 2005: Analysis and impact of super-obbed Doppler radial velocity in the NCEP Grid-point Statistical Interpolation (GSI) analysis system. *21st Conf. on Weather Analysis and Forecasting/17th Conf. on Numerical Weather Prediction*, 13A.4.
Liu, S., and Coauthors, 2016: WSR-88D Radar Data Processing at NCEP. *Weather and Forecasting*, **31,** 2047-2055.
Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, **112,** 1177-1194.
Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP—a comparison with 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, **129,** 3183-3203.
Lorenc, A. C., 2013: Recommended nomenclature for EnVar data assimilation methods. *Research Activities in Atmospheric and Oceanic Modeling*, **5**.
Lorenz, E. N., 1963: Deterministic Nonperiodic Flow. *Journal of Atmospheric Sciences*, **20,** 130-141.
——, 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21,** 289-307.
Lorenz, E. N., 1996: Predictability: A problem partly solved. *Proc. Seminar on predictability*.
Marshall, J. S., and W. M. K. Palmer, 1948: THE DISTRIBUTION OF RAINDROPS WITH SIZE. *Journal of Atmospheric Sciences*, **5,** 165-166.
McNally, A. P., 2019: On the sensitivity of a 4D-Var analysis system to satellite observations located at different times within the assimilation window. *Quarterly Journal of the Royal Meteorological Society*, **145,** 2806-2816.

Miller, L. J., and J. Sun, 2003: Initialization and forecasting of thunderstorms: Specification of radar measurement errors.
Navon, I. M., and D. M. Legler, 1987: Conjugate-gradient methods for large-scale minimization in meteorology. 1479-1502.
Nehrkorn, T., B. Woods, T. Auligné, and R. N. Hoffman, 2014: Application of Feature Calibration and Alignment to High-Resolution Analysis: Examples Using Observations Sensitive to Cloud and Water Vapor. *Monthly Weather Review*, **142,** 686-702.
Nehrkorn, T., B. K. Woods, R. N. Hoffman, and T. Auligné, 2015: Correcting for Position Errors in Variational Data Assimilation. *Monthly Weather Review*, **143,** 1368-1381.
Newman, K., and Coauthors, 2022: The MET Version 10.1.2 User's Guide. Developmental Testbed Center.
NWS, 2014: Technical Implementation Notice 14-46 Corrected.
——, 2019: Service Change Notice 19-40.
——, 2021: Service Change Notice 21-20.
OFCM, 2017: Federal Meteorologist Handbook No. 11 WSR-88D Meteorological Observations Part C. OFCM, Ed.
Oke, P. R., P. Sakov, and S. P. Corney, 2007: Impacts of localisation in the EnKF and EnOI: experiments with a small model. *Ocean Dynamics*, **57,** 32-45.
Olander, T. L., and C. S. Velden, 2019: The Advanced Dvorak Technique (ADT) for Estimating Tropical Cyclone Intensity: Update and New Capabilities. *Weather and Forecasting*, **34,** 905-922.
Ott, E., B. Hunt, I. Szunyogh, M. Corazza, E. Kalnay, D. Patil, and J. Yorke, 2002: Exploiting local low dimensionality of the atmospheric dynamics for efficient Kalman filtering.
Ott, E., and Coauthors, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus A: Dynamic Meteorology and Oceanography*, **56,** 415-428.
Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Monthly Weather Review*, **120,** 1747-1763.
Payne, T. J., 2017: Rapid update cycling with delayed observations. *Tellus A: Dynamic Meteorology and Oceanography*, **69,** 1409061.
Peckham, S. E., T. G. Smirnova, S. G. Benjamin, J. M. Brown, and J. S. Kenyon, 2015: Implementation of a Digital Filter Initialization in the WRF Model and Its Application in the Rapid Refresh. *Monthly Weather Review*, **144,** 99-106.
Penny, S. G., 2014: The hybrid local ensemble transform Kalman filter. *Monthly Weather Review*, **142,** 2139-2149.
Poterjoy, J., and F. Zhang, 2015: Systematic Comparison of Four-Dimensional Data Assimilation Methods With and Without the Tangent Linear Model Using Hybrid Background Error Covariance: E4DVar versus 4DEnVar. *Monthly Weather Review*, **143,** 1601-1621.
Poterjoy, J., L. Wicker, and M. Buehner, 2019: Progress toward the Application of a Localized Particle Filter for Numerical Weather Prediction. *Monthly Weather Review*, **147,** 1107-1126.
Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical Aspects of the Application of Recursive Filters to Variational Statistical Analysis. Part I:

Spatially Homogeneous and Isotropic Gaussian Covariances. *Monthly Weather Review*, **131,** 1524-1535.
——, 2003b: Numerical Aspects of the Application of Recursive Filters to Variational Statistical Analysis. Part II: Spatially Inhomogeneous and Anisotropic General Covariances. *Monthly Weather Review*, **131,** 1536-1548.
Putman, W. M., and S.-J. Lin, 2007: Finite-volume transport on various cubed-sphere grids. *Journal of Computational Physics*, **227,** 55-78.
Rihan, F. A., C. G. Collier, S. P. Ballard, and S. J. Swarbrick, 2008: Assimilation of Doppler radial winds into a 3D-Var system: Errors and impact of radial velocities on the variational analysis and model forecasts. *Quarterly Journal of the Royal Meteorological Society*, **134,** 1701-1716.
Rinehart, R. E., 2004: *Radar for Meteorologists*. 4 ed. Rinehart Publications.
Roberts, N. M., 2005: An investigation of the ability of a storm scale configuration of the Met Office NWP model to predict flood-producing rainfall.Met Office Tech. Rep. 455, 80 pp.
Roberts, N. M., and H. W. Lean, 2008: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Monthly Weather Review*, **136,** 78-97.
Rogers, E., and Coauthors, 2017: Upgrades to the NCEP North American Mesoscale (NAM) System. *Research activities in atmospheric and oceanic modelling, CAS/JSC Working Group on Numerical Experimentation. Report No. 47. WCRP Report No.12/2017. WMO, Geneva*.
Saltikoff, E., and Coauthors, 2019: OPERA the Radar Project. *Atmosphere*, **10,** 320.
Shapiro, A., S. Ellis, and J. Shaw, 1995: Single-Doppler Velocity Retrievals with Phoenix II Data: Clear Air and Microburst Wind Retrievals in the Planetary Boundary Layer. *Journal of Atmospheric Sciences*, **52,** 1265-1287.
Simonin, D., S. P. Ballard, and Z. Li, 2014: Doppler radar radial wind assimilation using an hourly cycling 3D-Var with a 1.5 km resolution version of the Met Office Unified Model for nowcasting. *Quarterly Journal of the Royal Meteorological Society*, **140,** 2298-2314.
Slivinski, L. C., D. E. Lippi, J. S. Whitaker, G. Ge, J. R. Carley, C. R. Alexander, and G. P. Compo, 2022: Overlapping Windows in a Global Hourly Data Assimilation System. *Monthly Weather Review*, **150,** 1317-1334.
Snook, N., M. Xue, and Y. Jung, 2012: Ensemble Probabilistic Forecasts of a Tornadic Mesoscale Convective System from Ensemble Kalman Filter Analyses Using WSR-88D and CASA Radar Data. *Monthly Weather Review*, **140,** 2126-2146.
Snyder, C., and F. Zhang, 2003: Assimilation of Simulated Doppler Radar Observations with an Ensemble Kalman Filter. *Monthly Weather Review*, **131,** 1663-1677.
Spilhaus, A. F., 1948: RAINDROP SIZE, SHAPE AND FALLING SPEED. *Journal of Atmospheric Sciences*, **5,** 108-110.
Stensrud, D. J., and Coauthors, 2009: Convective-Scale Warn-on-Forecast System: A Vision for 2020. *Bulletin of the American Meteorological Society*, **90,** 1487-1500.
Stratman, D. R., and C. K. Potvin, 2022: Testing the Feature Alignment Technique (FAT) in an Ensemble-Based Data Assimilation and Forecast System with Multiple-Storm Scenarios. *Monthly Weather Review*, **150,** 2033-2054.

Sun, J., 2005a: Initialization and Numerical Forecasting of a Supercell Storm Observed during STEPS. *Monthly Weather Review*, **133,** 793-813.
——, 2005b: Convective-scale assimilation of radar data: progress and challenges. *Quarterly Journal of the Royal Meteorological Society*, **131,** 3439-3463.
Sun, J., and N. A. Crook, 1997: Dynamical and Microphysical Retrieval from Doppler Radar Observations Using a Cloud Model and Its Adjoint. Part I: Model Development and Simulated Data Experiments. *Journal of the Atmospheric Sciences*, **54,** 1642-1661.
——, 1998: Dynamical and Microphysical Retrieval from Doppler Radar Observations Using a Cloud Model and Its Adjoint. Part II: Retrieval Experiments of an Observed Florida Convective Storm. *Journal of the Atmospheric Sciences*, **55,** 835-852.
——, 2001: Real-Time Low-Level Wind and Temperature Analysis Using Single WSR-88D Data. *Weather and Forecasting*, **16,** 117-132.
Sun, J., and H. Wang, 2013: Radar Data Assimilation with WRF 4D-Var. Part II: Comparison with 3D-Var for a Squall Line over the U.S. Great Plains. *Monthly Weather Review*, **141,** 2245-2264.
Thépaut, J.-N., R. N. Hoffman, and P. Courtier, 1993a: Interactions of Dynamics and Observations in a Four-Dimensional Variational Assimilation. *Monthly Weather Review*, **121,** 3393-3414.
Thépaut, J.-N., D. Vasiljevic, P. Courtier, and J. Pailleux, 1993b: Variational assimilation of conventional meteorological observations with a multilevel primitive-equation model. *Quarterly Journal of the Royal Meteorological Society*, **119,** 153-186.
Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble Square Root Filters. *Monthly Weather Review*, **131,** 1485-1490.
Tong, C.-C., Y. Jung, M. Xue, and C. Liu, 2020: Direct Assimilation of Radar Data With Ensemble Kalman Filter and Hybrid Ensemble-Variational Method in the National Weather Service Operational Data Assimilation System GSI for the Stand-Alone Regional FV3 Model at a Convection-Allowing Resolution. *Geophysical Research Letters*, **47,** e2020GL090179.
Tong, M., and M. Xue, 2005: Ensemble Kalman Filter Assimilation of Doppler Radar Data with a Compressible Nonhydrostatic Model: OSS Experiments. *Monthly Weather Review*, **133,** 1789-1807.
——, 2008: Simultaneous Estimation of Microphysical Parameters and Atmospheric State with Simulated Radar Data and Ensemble Square Root Kalman Filter. Part I: Sensitivity Analysis and Parameter Identifiability. *Monthly Weather Review*, **136,** 1630-1648.
van Leeuwen, P. J., and G. Evensen, 1996: Data Assimilation and Inverse Methods in Terms of a Probabilistic Formulation. *Monthly Weather Review*, **124,** 2898-2913.
Waller, J. A., D. Simonin, S. L. Dance, N. K. Nichols, and S. P. Ballard, 2016: Diagnosing Observation Error Correlations for Doppler Radar Radial Winds in the Met Office UKV Model Using Observation-Minus-Background and Observation-Minus-Analysis Statistics. *Monthly Weather Review*, **144,** 3533-3551.
Wang, X., 2010: Incorporating Ensemble Covariance in the Gridpoint Statistical Interpolation Variational Minimization: A Mathematical Framework. *Monthly Weather Review*, **138,** 2990-2995.

Wang, X., and C. H. Bishop, 2003: A Comparison of Breeding and Ensemble Transform Kalman Filter Ensemble Forecast Schemes. *Journal of the Atmospheric Sciences*, **60,** 1140-1158.
Wang, X., and T. Lei, 2014: GSI-Based Four-Dimensional Ensemble–Variational (4DEnsVar) Data Assimilation: Formulation and Single-Resolution Experiments with Real Data for NCEP Global Forecast System. *Monthly Weather Review*, **142,** 3303-3325.
Wang, X., C. H. Bishop, and S. J. Julier, 2004: Which Is Better, an Ensemble of Positive–Negative Pairs or a Centered Spherical Simplex Ensemble? *Monthly Weather Review*, **132,** 1590-1605.
Wang, X., D. Parrish, D. Kleist, and J. Whitaker, 2013: GSI 3DVar-Based Ensemble–Variational Hybrid Data Assimilation for NCEP Global Forecast System: Single-Resolution Experiments. *Monthly Weather Review*, **141,** 4098-4117.
Wang, Y., and X. Wang, 2021: Development of Convective-Scale Static Background Error Covariance within GSI-Based Hybrid EnVar System for Direct Radar Reflectivity Data Assimilation. *Monthly Weather Review*, **149,** 2713-2736.
Weygandt, S. S., A. Shapiro, and K. K. Droegemeier, 2002: Retrieval of Model Initial Fields from Single-Doppler Observations of a Supercell Thunderstorm. Part I: Single-Doppler Velocity Retrieval. *Monthly Weather Review*, **130,** 433-453.
Wheatley, D. M., K. H. Knopfmeier, T. A. Jones, and G. J. Creager, 2015: Storm-Scale Data Assimilation and Ensemble Forecasting with the NSSL Experimental Warn-on-Forecast System. Part I: Radar Data Experiments. *Weather and Forecasting*, **30,** 1795-1817.
Whitaker, J. S., 2020: Hybrid-Gain Versus Hybrid-Covariance Data Assimilation. *24th Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*.
Whitaker, J. S., and T. M. Hamill, 2002: Ensemble Data Assimilation without Perturbed Observations. *Monthly Weather Review*, **130,** 1913-1924.
——, 2012: Evaluating Methods to Account for System Errors in Ensemble Data Assimilation. *Monthly Weather Review*, **140,** 3078-3089.
Whiton, R. C., P. L. Smith, S. G. Bigler, K. E. Wilk, and A. C. Harbuck, 1998a: History of Operational Use of Weather Radar by U.S. Weather Services. Part II: Development of Operational Doppler Weather Radars. *Weather and Forecasting*, **13,** 244-252.
——, 1998b: History of Operational Use of Weather Radar by U.S. Weather Services. Part I: The Pre-NEXRAD Era. *Weather and Forecasting*, **13,** 219-243.
Wilks, D. S., 2011: *Statistical methods in the atmospheric sciences.* Vol. 100, Academic press.
Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. *Monthly Weather Review*, **130,** 2905-2916.
Wu, W.-S., D. F. Parrish, E. Rogers, and Y. Lin, 2017: Regional Ensemble–Variational Data Assimilation Using Global Ensemble Forecasts. *Weather and Forecasting*, **32,** 83-96.
Wurman, J., 1994: Vector Winds from a Single-Transmitter Bistatic Dual-Doppler Radar Network. *Bulletin of the American Meteorological Society*, **75,** 983-994.

Xiao, Q., Y.-H. Kuo, J. Sun, W.-C. Lee, E. Lim, Y.-R. Guo, and D. M. Barker, 2005: Assimilation of Doppler Radar Observations with a Regional 3DVAR System: Impact of Doppler Velocities on Forecasts of a Heavy Rainfall Case. *Journal of Applied Meteorology*, **44,** 768-788.
Xie, Y., C. Lu, and G. L. Browning, 2002: Impact of Formulation of Cost Function and Constraints on Three-Dimensional Variational Data Assimilation. *Monthly Weather Review*, **130,** 2433-2447.
Xu, Q., and J. Gong, 2003: Background error covariance functions for Doppler radial-wind analysis. *Quarterly Journal of the Royal Meteorological Society*, **129,** 1703-1720.
Xue, M., M. Tong, and K. K. Droegemeier, 2006: An OSSE Framework Based on the Ensemble Square Root Kalman Filter for Evaluating the Impact of Data from Radar Networks on Thunderstorm Analysis and Forecasting. *Journal of Atmospheric and Oceanic Technology*, **23,** 46-66.
Xue, M., and Coauthors, 2001: The Advanced Regional Prediction System (ARPS) – A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteorology and Atmospheric Physics*, **76,** 143-165.
Yakubu, M. L., Z. Yusop, and M. A. Fulazzaky, 2016: The influence of rain intensity on raindrop diameter and the kinetics of tropical rainfall: case study of Skudai, Malaysia. *Hydrological Sciences Journal*, **61,** 944-951.
Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of Initial Estimate and Observation Availability on Convective-Scale Data Assimilation with an Ensemble Kalman Filter. *Monthly Weather Review*, **132,** 1238-1253.
Zhao, Q., and F. H. Carr, 1997: A Prognostic Cloud Scheme for Operational NWP Models. *Monthly Weather Review*, **125,** 1931-1953.
Zhou, L., S.-J. Lin, J.-H. Chen, L. M. Harris, X. Chen, and S. L. Rees, 2019: Toward Convective-Scale Prediction within the Next Generation Global Prediction System. *Bulletin of the American Meteorological Society*, **100,** 1225-1243.
Zrnic, D. S., and Coauthors, 2007: Agile-Beam Phased Array Radar for Weather Observations. *Bulletin of the American Meteorological Society*, **88,** 1753-1766.