CONVECTION-PERMITTING ENSEMBLE FORECASTS OF THE 10–12 DECEMBER 2013 LAKE-EFFECT SNOW EVENT:

SENSITIVITY TO MICROPHYSICAL, PLANETARY BOUNDARY LAYER, AND SURFACE LAYER PARAMETERIZATIONS

by

William Massey Bartolini

A Thesis

Submitted to the University at Albany, State University of New York

in Partial Fulfillment of

the Requirements for the Degree of

Master of Science

College of Arts and Sciences

Department of Atmospheric and Environmental Sciences

2019

ABSTRACT

Lake-effect snow (LeS) presents a substantial forecast challenge for convection-permitting models, due in part to uncertainties in the parameterization of microphysical (MP) and planetary boundary layer / surface layer (PBL/SL) processes. Here we focus on understanding these uncertainties for a LeS event that occurred during 10–12 December 2013 during the Ontario Winter Lake-effect Systems (OWLeS) field campaign. Throughout this event, long-lake-axis-parallel snowbands persisted downwind of the eastern shore of Lake Ontario, leading to snowfall accumulations as high as 105 cm (liquid precipitation equivalent of 64.5 mm) on the Tug Hill Plateau. We run nested simulations of the 10–12 December 2013 LeS event at 12-, 4-, and 1.33-km horizontal grid spacing using the Weather Research and Forecasting (WRF) model configured similarly to the operational High-Resolution Rapid Refresh (HRRR) model. Two suites of sensitivity experiments are conducted for the event, first, with nine different MP schemes, and second, with eleven different PBL/SL schemes. Large differences between the MP experiments are found in the LeS band intensity and precipitation type, with smaller differences in the band timing, position, and morphology. Maximum storm-total liquid precipitation equivalent amounts among MP ensemble members range from 35 to 62 mm. Results from the WRF simulations are compared to detailed observations from OWLeS, including scanning and profiling radar data and surface snowfall and crystal habit observations. Additionally, measurements from the University of Wyoming King Air aircraft, including vertically pointing cloud radars and in-situ flight-level thermodynamic and microphysical observations, are used to compare modeled and observed cloud structures. By analyzing cloud microphysics, such as supercooled liquid water content, hy-

drometeor size distributions, and precipitation fallspeed, we compare modeled and observed lake-effect cloud properties and assess which MP schemes are most accurately simulating cloud processes during this LeS event. Motivated by substantial differences in the model-observational comparisons, we also perform several sensitivity experiments with the aerosol-aware Thompson MP scheme used in the HRRR, modifying the snow size distribution and fallspeed within WRF to quantify single-scheme precipitation uncertainty. These sensitivity experiments vary maximum liquid precipitation equivalent amounts over the Tug Hill Plateau by 5 to 15 mm relative to the Thompson control simulation. In contrast to the MP ensemble experiments, the PBL/SL ensemble primarily shows differences in the LeS band intensity and morphology. Maximum storm-total liquid precipitation equivalent amounts among PBL/SL ensemble members range from 54 to 99 mm, which is a larger range than in the MP ensemble. PBL/SL precipitation forecasts are compared to observations of manual and radar-estimated precipitation to identify which PBL/SL schemes produce excessive QPF. Several sensitivity experiments are also performed with the Quasi-Normal-Scale-Elimination (QNSE) SL scheme, which produces the highest fluxes and snowfall, by varying the Prandtl number and critical threshold for friction velocity to investigate surface flux sensitivity.

ACKNOWLEDGMENTS

The work presented in this thesis would not have been possible without the support of a team of individuals helping me throughout graduate school. First and foremost, thanks to Justin Minder, my primary advisor, whose advice, suggestions, and patience throughout the iterative research, presentation, and writing process made this work possible. I also thank Ryan Torn and Daniel Keyser, my other two co-advisors, for their helpful suggestions and feedback on my research. My National Weather Service (NWS) focal points, David Zaff of NWS Buffalo, NY, and Joseph Villani of NWS Albany, NY, generously hosted me during visits to their two respective forecast offices, sharing many insights into the practical lake-effect forecast challenges faced by today's operational forecasters. Joseph Olson and the rest of the NOAA Earth System Research Laboratory (ESRL) – High-Resolution Rapid Refresh (HRRR) development team provided me with the latest versions of the HRRR code to use in this research and also hosted me during two separate visits to ESRL, giving me valuable feedback from the operational model developer perspective. Thanks to the National Center for Atmospheric Research for accepting me into their 2018 Advanced Study Program Summer Colloquium, titled "Synthesis of Observations and Models in Studies of Shallow and Deep Clouds", where several mentors helped two other graduate students (Vanessa Pryzbylo and Ana Sena) and me to develop the particle size distribution comparison code that I later adapted for use in this study. This reseach was supported by the NOAA/NWS Collaborative Science, Technology, and Applied Research (CSTAR) Program at the University at Albany, under NOAA Award NA16NWS4680005. Computing support was provided by the University at Albany RIT Snow Cluster and the Department of Atmospheric and Environmental

Sciences IT staff. Last but not least, thanks to my mother, Morgan Bartolini, who I am indebted to for her unwavering encouragement and sacrifice during several difficult years of dealing with my health challenges, and to my entire family for their support throughout my academic pursuit of atmospheric science.

CONTENTS

1. Introduction

Lake-effect snowstorms are a major part of the cool-season hydroclimate of the Laurentian Great Lakes region, often producing intense, heavy snowfall over localized areas downwind of the lakeshore (Peace and Sykes 1966; Niziol et al. 1995). While these snowstorms positively contribute to winter recreation and water resource storage, they can negatively impact transportation and commerce across the region, among other impacts to society (Schmidlin 1993). Lake-effect snow (LeS) forms as Arctic airmasses cross the Great Lakes, often in the wake of a mid-latitude cyclone. Initially cold, dry Arctic air is warmed and moistened from below via turbulent lake-atmosphere fluxes over ice-free or partially ice-covered portions of the lakes, driving boundary-layer convection, clouds, and lake-effect precipitation (Lenschow 1973; Gerbush et al. 2008). Factors such as the amount of over-lake wind fetch, lake-induced instability, vertical wind shear, and lake shoreline geometries cause a variety of precipitation band morphologies (e.g., Braham and Kelly 1982; Forbes and Merritt 1984; Niziol et al. 1995; Steiger et al. 2013). Pattern recognition of favorable synoptic-scale features for LeS by forecasters can help anticipate a storm as much as a week in advance or more, but mesoscale details such as the timing and amount of substantial LeS impacting specific locations are often unclear at lead times of one day or less (D. Zaff, NOAA/NWS Buffalo, 2017, personal communication).

1.1 Background and motivation

1.1.1 Lake-effect snowstorm dynamics

Overall, LeS storms are driven by upward heat and moisture fluxes over the lakes. Boundary-layer convective rolls develop, forming cloud streets, and often grow upscale into coherent LeS bands over a lake's downwind shoreline (e.g., Kristovich et al. 1999, 2003). Depending on the prevailing wind direction and amount of over-lake fetch affecting upscale growth, LeS bands can remain separate as horizontal convective rolls or aggregate to form

a larger long-lake-axis-parallel (LLAP, Steiger et al. 2013) band, among other morphologies. Furthermore, the lake-modified convective boundary layer and subsequent LeS band formation can be influenced by air parcel trajectories over multiple Great Lakes, often referred to as a multi-lake connection (Niziol et al. 1995; Kristovich et al. 2018). This process is further complicated by the presence of lake ice cover in late winter. While the Great Lakes' shoreline geometries affect precipitation patterns throughout the year, the effective shoreline changes during the winter as the lakes partially freeze over and ice coverage is distributed by the wind. Although latent and sensible heat fluxes nearly cease from a completely frozen lake, Gerbush et al. (2008) analyzed upward heat fluxes from aircraft flights over ice-covered Lake Erie, where large fluxes were observed even in regions of 70% partial ice cover. Therefore, the presence of partial lake ice cover can still allow intense LeS to occur.

Lake-effect snowbands are further intensified by shoreline friction and especially orographic lift, across the Allegheny Plateau downwind of Lake Erie and over the Tug Hill Plateau (hereafter Tug Hill) and western Adirondack Mountains downwind of Lake Ontario. Mean annual snowfall in the immediate vicity of Tug Hill, which rises 500 m above the surface of Lake Ontario, ranges from 300 to 700 cm according to a 13-year climatology by Veals and Steenburgh (2015). From their climatology of days with at least 2 h of lake-effect, a substantial fraction (61–76%) of the total seasonal snowfall is associated with lake-effect, with a smaller contribution to the total liquid precipitation equivalent (24–37%).

Motivated to better understand the dynamical and microphysical processes associated with cool-season lake-effect convection and boundary-layer evolution, including LeS leading to extreme snowfall amounts over Tug Hill, the Ontario Winter Lake-effect Systems field campaign (OWLeS, Kristovich et al. 2017) took place during the winter of 2013–2014. Building on a previous field campaign (LOWS, Reinking et al. 1993), OWLeS focused on understanding many aspects of lake-effect precipitation processes, using networks of ground-based snow observation measurements, multiple scanning and profiling radar platforms, and in situ aircraft measurements to study boundary-layer evolution, cloud microphysics, and orographic

enhancement. Recent literature has already investigated several Intensive Observing Periods (IOPs) from OWLeS to better understand lake-effect snowband dynamics.

The majority of research has focused on OWLeS IOP2b, a lake-effect snowstorm occurring during 10–12 December 2013, which was a particularly well-observed event during OWLeS. The event occurred under synoptic westerly flow, maximing the over-water wind fetch across Lake Ontario and causing a single, long-lake-axis-parallel (LLAP) band extending downwind of Lake Ontario. The quasi-stationary position of the LLAP band, in combination with inland and orographic enhancement, resulted in localized but intense snowfall amounts in excess of 102.5 cm on Tug Hill (Minder et al. 2015; Campbell et al. 2016).

Minder et al. (2015) examined the processes controlling the inland evolution and intensification of LeS over Tug Hill, using OWLeS IOP2b as a case study. Using a west-east-oriented transect of four K-band profiling radars, they determined that inland intensification was not associated with an increase in the depth or intensity of lake-effect convection, and instead followed an over-lake convective to inland-stratiform transition during the IOP2b event with an inland increase in echo frequency. They also analyzed composites from twenty-nine OWLeS cases with profiling radar data, which showed a similar inland evolution as IOP2b. Welsh et al. (2016) synthesized observations from all available scanning and profiling radars during IOP2b, including data from a profiling, aircraft-mounted, W-band radar, to analyze inland intensification during this event. From the aircraft observations, they found convective updrafts over the lake (sometimes creating a bounded weak echo region) locally coincident at flight level with heavily rimed snow particles aloft. Over land, Welsh et al. (2016) noted a transition to broad stratiform ascent, with ice crystals observed aloft and dry aggregates at low levels, inferred from radar data. Campbell et al. (2016) studied surface snowfall observations and scanning NEXRAD radar data during IOP2b and found near-equal amounts of banded (often a single, long-lake-axis-parallel, or LLAP, band) and non-banded (broad coverage) precipitation periods occurred during IOP2b. While LLAP banded periods had heavier precipitation rates overall, broad coverage periods had weaker precipitation rates but

stronger enhancement of precipitation over Tug Hill.

To understand the forcing mechanisms associated with inland intensification of the LeS bands, Campbell and Steenburgh (2017) ran two WRF simulations of IOP2b with and without the Tug Hill terrain. Their experiments showed that both orographic and non-orographic features contributed to the Tug Hill precipitation maximum: low-level convergence along differential surface heating boundaries (termed land-breeze fronts), enhanced depositional growth and accretion of cloud water on snow associated with orographic uplift, and reduced low-level sublimation over the higher terrain. Steenburgh and Campbell (2017) performed additional diagnostics using RAP analyses and Campbell and Steenburgh (2017)'s WRF simulation with terrain to study the land-breeze fronts. Two land-breeze fronts and an additional convergence zone developed during IOP2b due to westerly winds blowing parallel to the shoreline of Lake Ontario, with all three boundaries intersecting over Tug Hill during the event. Using the WRF simulation, Steenburgh and Campbell (2017) determined the land-breeze fronts strengthened via kinematic frontogenesis due to convergence and deformation generated by differential roughness along the shoreline, with minimal contributions from diabatic frontogenesis.

Besides the land-breeze front and orographic forcing mechanisms during IOP2b, internal LeS dynamics also helped to focus heavy convective precipitation during periods of the most well-organized LLAP banding. Bergmaier et al. (2017) performed a dual-dopper analysis of the W-band aircraft radar data and found a secondary circulation during a period of singular LLAP banding over Lake Ontario that also persisted downwind, with low-level inflow and upper-level outflow flanking a central updraft. Using a WRF simulation, they found that the band-scale secondary circulation was driven by solenoidal (buoyancy) forcing from latent heat release. All of the above studies of OWLeS IOP2b have increased our understanding of the complex lake-land-atmosphere interactions associated with extreme lake-effect storms, and the role of orographic processes and band-scale circulations in organizing convection.

1.1.2 Numerical weather prediction requirements for operational forecasting of lake-effect

Given the mesoscale nature of processes contributing to intense lake-effect snowstorms, limited-area numerical weather prediction (NWP) models with sufficiently fine horizontal grid spacing (convection-permitting resolution, where convective clouds are treated explicitly instead of parameterized, with horizontal grid spacing less than about 5 km) to resolve lake-effect precipitation have only become routinely operational within the past decade. Legacy ingredients-based approaches developed in the 1980s and 1990s continue to be used by National Weather Service (NWS) forecast offices to characterize global NWP model forecasts of environmental quantities favorable for lake-effect, such as lake-induced instability and boundary-layer wind direction and shear (Niziol 1987; Niziol et al. 1995). However, successful convection-permitting NWP forecasts of lake-effect remain challenging, since accurate prediction depends on many factors: correct initialization and prediction of the synoptic and mesoscale flow over the Great Lakes, well-resolved analyses of lake-surface temperature and ice cover, adequate horizontal grid spacing to resolve long, narrow lake-effect snow bands, and realistic parameterization of turbulent lake-atmosphere fluxes and cloud microphysics. Precise, convection-permitting NWP forecasts of LeS are also important since NWS Buffalo, NY, began experimenting with time-evolving, polygon-based LeS warnings in 2016, to reduce false alarm ratios in the current county-based LeS warning paradigm (D. Zaff, NOAA/NWS Buffalo, 2017, personal communication).

Several studies have already investigated mesoscale NWP model performance for LeS downwind of the Great Salt Lake in Utah and the Great Lakes. McMillen and Steenburgh (2015) ran deterministic WRF simulations of nineteen banded and non-banded Great Salt Lake-effect storms at approximately 12–24h lead times, and found that the simulations were able to produce realistic LeS structures but were biased in band morphology, precipitation amounts, and location. They concluded that deterministic simulations of LeS were of limited use for their case studies, and that model postprocessing or convection-permitting ensembles

were needed.

To explore the optimal convection-permitting ensemble model configuration for LeS forecasting in the short- to medium-range (24–48+ h lead times) for the OWLeS IOP2b case, Saslo and Greybush (2017) ran several limited-area initial and boundary condition ensembles as well as a mixed physics ensemble, consisting of different combinations of microphysics (MP) and planetary boundary layer / surface layer (PBL/SL) parameterization schemes. They found that their mixed physics ensemble impacted forecasting of mostly boundary-layer processes, with large spread in precipitation intensity but not LeS band position. Overall, the IOP2b forecast LeS band position uncertainty was driven by the synoptic scales. In their study, the ensemble with the most optimal ensemble dispersion was an initial and boundary condition ensemble using Global Ensemble Forecast System (GEFS) forecasts for the boundary conditions.

1.1.3 Sensitivity of lake-effect storms to microphysical parameterizations

NWP model parameterizations of cloud and precipitation MP have been extensively studied in cool-season precipitation regimes including lake-effect snowstorms. Reeves and Dawson (2013) analyzed a suite of WRF MP experiments for a Lake Erie case study, finding that precipitation amounts were strongly sensitive to the choice of MP scheme (factor of two differences between the wettest and driest members) with a subset of graupel-dominant members having precipitation maxima closer to the lake shore and a subset of snow-dominant members having more precipitation displaced inland. McMillen and Steenburgh (2015) analyzed a suite of MP experiments for a Great Salt Lake-effect snowstorm in northern Utah, finding that the Thompson scheme (Thompson et al. 2008) was most accurate relative to radar-based liquid-equivalent precipitation estimates, while other schemes like WRF Double-Moment 6-class (WDM6, Lim and Hong 2010) produced excessive precipitation largely due to an increase in graupel.

1.1.4 Sensitivity of lake-effect storms to planetary boundary layer / surface layer parameterizations

Field studies have also focused on understanding the dynamics of the boundary layer governing lake-effect snow processes, including horizontal convective roll structures (e.g., Kristovich et al. 1999, 2003, 2018) and the importance of correctly modeling upward sensible heat and moisture fluxes (Conrick et al. 2015; Fujisaki-Manome et al. 2017). Additional NWP complexity arises during the latter half of the winter when the Great Lakes partially freeze over, potentially requiring treatment of fractional lake ice cover in model initializations or coupling with a lake model for more accurate predictions across both weather and climate timescales (e.g., Powers and Stoelinga 2000; Xue et al. 2017).

To investigate PBL/SL scheme sensitivity, Conrick et al. (2015) studied the same Lake Erie case as Reeves and Dawson (2013) using a suite of WRF experiments varying the PBL/SL scheme. They found that the primary drivers of upward sensible and latent heat flux differences (as large as 400 W m^{-2} between the highest-flux and lowest-flux schemes, leading to liquid precipitation differences as large as 20 mm) were the SL schemes' choice of Prandtl number and similarity stability function. Fujisaki-Manome et al. (2017) studied a different Lake Erie case during November 2014 using a combination of atmospheric and lake models, and compared the simulated lake temperature and sensible/latent heat flux forecasts to observed over-lake turbulent flux measurements from Lake Erie. Their results also showed a range in maximum forecast fluxes (as large as 300-400 W m^{-2} for sensible heat flux), but that their Finite-Volume Community Ocean Model (FVCOM, Chen et al. 2006) simulations with the Coupled Ocean-Atmosphere Response Experiment (COARE, Fairall et al. 1996a,b) flux algorithm most closely matched the observations. However, all the models appeared to substantially over-predict the fluxes at a lighthouse flux observation site, with smaller model over-predictions of fluxes at a second site close to the upwind shore of Lake Erie. PBL evolution and NWP performance has also been investigated for other types of cool-season convective environments (e.g., Coniglio et al. 2013; Cohen et al. 2015).

1.1.5 Additional operational needs for lake-effect forecasting: Inland extent

While all Great Lakes' NWS offices have NWP lake-effect forecast concerns, several inland offices (such as NWS Binghamton, NY, and NWS Albany, NY) have additional challenges due to uncertainty in the inland extent of LeS. Villani et al. (2017) examined factors controlling inland extent using a four-year climatology of LeS bands using NEXRAD radar data and NAM analysis soundings. They found that a multi-lake connection (upstream moisture source air trajectory), lake-air temperature difference, 0-1 km bulk wind shear, and mean mixed-layer wind speed were the factors with the highest correlation to inland extent. Using their analyzed variables and a stepwise and backwards selection algorithm, they created a predictive equation for forecasting inland extent used in their office.

Eipper et al. (2018) also studied inland extent (termed inland penetration, or InPen, in their study) of 34 samples of LeS bands during all OWLeS cases, using NEXRAD radar composites and North American Regional Reanalysis (NARR) proximity soundings. Using regression models comprised of parameters from NEXRAD and NARR data, they studied three hypotheses for InPen: pure advection of a LeS band downwind of the lake (advection-only), inland consolidation of a LeS band along a narrowing moisture plume (inland plume-focusing), and inland invigoration of LeS bands (inland invigoration). While their advection-only and inland plume-focusing statistical models were unable to fully explain InPen during the OWLeS cases, their inland invigoration model performed best, suggesting that boundary layer destabilization by vertically differential cold air advection is a key driver of InPen. Eipper et al. (2018) hypothesized that a combination of previously-studied factors, including both vertically differential cold air advection and a multi-lake connection may contribute to larger InPen (Niziol et al. 1995; Villani et al. 2017).

A third study by Veals et al. (2018) investigated other climatological factors affecting inland extent and orographic enhancement, using a nine-year sample of NEXRAD radar data and NARR soundings. Similar to Villani et al. (2017), they focused on parameters like mean wind speed and lake-induced convective available potential energy (LCAPE), but

also stratified their dataset by lake-effect mode (non-banded, weakly banded, and banded). In general, they found that higher mean wind speeds allowed for more overall precipitation and a larger degree of inland and orographic enhancement, while higher LCAPE resulted in more lake-effect precipitation closer to the lakeshore. They also found that periods of banded precipitation had higher precipitation amounts and a precipitation maximum closer to the lakeshore, supporting the results from Campbell et al. (2016).

1.2 Research objectives

In the past few years surrounding the OWLeS field campaign, research efforts have revealed new insights into the structure and dynamics of LeS storms and their inland intensification over Tug Hill. Modeling studies of lake-effect and orographic precipitation from the Great Lakes, Utah, and the Cascades, among other locations, have long demonstrated parameterization sensitivities and deficiencies in models' MP and PBL/SL schemes that strongly affect the location and amount of forecast precipitation. However, many of the model physics sensitivity studies of LeS (Reeves and Dawson 2013; McMillen and Steenburgh 2015; Conrick et al. 2015) were only able to do minimal validation, limited by operational scanning radar coverage, a lack of precise snowfall/precipitation measurements, and very few microphysical or lake-air flux measurements. From an operational perspective, forecasters still rely on a combination of pattern recognition techniques, post-processing tools, and convection-permitting model output, since high-resolution operational models are relatively new and that research has demonstrated the large precipitation uncertainties in the model forecasts. The datasets (especially microphysical and precipitation observations) from OWLeS provide an opportunity to revisit the mesoscale model forecast challenges posed by intense LeS storms, such as OWLeS IOP2b. This study builds off the previous model sensitivity work on LeS by running two WRF MP and PBL/SL physics ensembles along with several additional sensitivity experiments, but also includes extensive comparisons with observational data from OWLeS to constrain MP/PBL/SL parameterization uncertainty. Our

specific research objectives are:

- Quantify precipitation uncertainty during IOP2b due to the choice of MP and PBL/SL parameterization.
- Identify schemes that clearly over- or under-forecast precipitation totals, using OWLeS snowfall and radar-estimated precipitation amounts for validation.
- Characterize schemes that predict cloud and precipitation properties inconsistent with OWLeS in-situ and remotely-sensed cloud physics measurements, and ground-based precipitation-type observations.
- Isolate specific sources of uncertainty within MP and PBL/SL schemes leading to precipitation forecast sensitivity.

In this manner, we provide guidance to operational forecasters and model developers on which schemes provide the most realistic precipitation forecasts during OWLeS IOP2b. Chapter 2 discusses the data and methods used in this study. Chapter 3 presents an overview of the OWLeS IOP2b case. Chapters 4 and 5 describe results from the MP and PBL/SL experiments. Chapter 6 presents a summary and discusses overall conclusions from this research.

2. Data and methods

2.1 Model configuration

For this study, we focus on an intense period of lake-effect precipitation occuring between 2100 UTC 10 December and 0600 UTC 12 December 2013 during OWLeS IOP2b. We run limited-area, triply-nested Weather Research and Forecasting (WRF v3.9, Skamarock et al. 2008) model simulations of this event at 12-, 4-, and 1.33-km horizontal grid spacing (see Fig. 2.1 for domain locations). All analyses herein are performed on the innermost domain. Initial conditions (ICs) and boundary conditions (BCs) are prescribed using Rapid Refresh (RAP) model 0-h analyses, to minimize the sensitivity of our results to BC errors. In their ensemble configuration study of IOP2b, Saslo and Greybush (2017) demonstrate that synoptic-scale flow uncertainties at the BCs have the greatest effect on ensemble spread at 1–3 day forecast lead times. In our study, analysis nudging every 3-h of wind, temperature, and water vapor is used to further limit forecast BC errors and thus isolate forecast sensitivity to MP and PBL/SL parameterizations. Uniform nudging of all wavelengths is performed above the boundary layer and only over the outermost domain.

All WRF experiments are initialized at 1200 UTC 10 December, to allow for sufficient spin-up of lake-effect precipitation. Together with the choice of RAP ICs/BCs, model configuration and physics options were chosen to emulate the latest operational version (v3) of the High-Resolution Rapid Refresh (HRRR, Benjamin et al. 2016). While convection is parameterized over the outer domain using the Grell-Freitas deep convection scheme (Grell and Freitas 2014), both inner domains are convection-permitting over the entire Great Lakes region. The HRRR model configuration includes Thompson aerosol-aware microphysics (Thompson and Eidhammer 2014) with climatological aerosol concentrations, Mellor-Yamada-Nakanishi-Niino Level 2.5 planetary boundary layer and surface layer schemes (MYNN, Nakanishi and Niino 2006), RRTMG longwave and shortwave radiation schemes (Iacono et al. 2008), and

RUC land surface model (Benjamin et al. 2004). Bottom BCs over the Great Lakes in our WRF simulations are modified from the original RAP data, and updated using fixed 0-h analysis fields for lake-surface temperature and ice cover from the NOAA Great Lakes Operational Forecast System lake model (GLOFS, Chu et al. 2011) forecast initialized at 1200 UTC 10 December, representing the most accurate lake conditions.

A control WRF simulation (CTRL) is performed using the above HRRR physics options, as well as two separate model physics ensembles varying either the MP or the PBL/SL parameterization, summarized in Table 2.1 and Table 2.2, respectively. The ensemble member using Morrison MP repeatedly crashed early in the simulation when run with the RUC land surface model, so that member instead uses the Noah land surface model (Ek et al. 2003). Microphysics schemes for these experiments were chosen from a number of commonly used schemes in research and operational forecasting, and include several single- and hybrid/double-moment options. PBL/SL schemes were chosen in the same manner, and include both local (K-profile parameterization, or KPP, where "K" refers to eddy diffusivity) and non-local (turbulent kinetic energy, or TKE) methods of turbulence parameterization, noted as "PBL type" in Table 2.2.

Operational model configurations used by the NWS High Resolution Forecasting System version 2.1 (HREF) ensemble are summarized in Table 2.3. HREF uses a diverse set of MP, PBL/SL, and initial and boundary conditions from different models, providing many configuration differences between HREF members that can contribute to spread in precipitation forecasts. Here we focus on the differences due to the MP and PBL/SL schemes individually, including parameterizations used in HREF as part of our physics ensembles based on a HRRR-like configuration. Therefore, it is important to note that these separate experiments do not describe the full range of uncertainty in HREF forecasts. For example, this study does not investigate sensitivity to initial and boundary conditions which can lead to position uncertainty of LeS bands (e.g., Saslo and Greybush 2017). Note also that HREF members also use horizontal grid spacing of either 3.0 km (HRRR, NAM Nest) or 3.2 km

(HRW ARW, HRW NMMB, HRW NSSL), whereas the inner domain in this study, used for all analyses, is higher-resolution (1.33 km) than all HREF members.

Table 2.1: Summary of WRF model configurations for the multi-microphysics ensemble experiments. Mellor-Yamada Nakanishi Niino Level 2.5 PBL/SL (MYNN, Nakanishi and Niino 2004, 2009) is used in all members.

Experiment	MP scheme
CTRL	Thompson Aerosol-Aware[1]
WSM6	WRF Single Moment 6-Class[2]
P3	Predicted Particle Properties[3]
MYAU	Milbrandt-Yau[4]
MORR	Morrison[5,10]
SBUL	Stony Brook Univ. - Lin[6]
FERR	Ferrier[7]
GDRD	Goddard[8]
NSSL	NSSL 2-Moment[9]

[1] Thompson et al. (2008); Thompson and Eidhammer (2014)
[2] Hong and Lim (2006)
[3] Morrison and Milbrandt (2015)
[4] Milbrandt and Yau (2005a,b)
[5] Morrison et al. (2009)
[6] Lin and Colle (2011)
[7] No paper available.
[8] Tao et al. (2016)
[9] Mansell et al. (2010)
[10] Uses Noah land surface model (Ek et al. 2003) instead of RUC land surface model (Benjamin et al. 2004).

2.2 Observations

2.2.1 Surface snowfall

During OWLeS IOP2b, two special observing sites on the western side of Tug Hill were outfitted with both automated and manual precipitation measurements, located at Sandy Creek (SC) and North Redfield (NR), NY, shown in Figure 2.2 (Steenburgh 2014a,b). While the manual observations of accumulated depth and liquid precipitation equivalent (LPE) at 6-h intervals were reliable, automated measurements from shielded Noah ETI weighing

Table 2.2: Summary of WRF model configurations for the multi-PBL/SL ensemble experiments. Thompson Aerosol-Aware MP (Thompson et al. 2008; Thompson and Eidhammer 2014) is used in all members. The "PBL type" column indicates whether the PBL scheme uses a local, K-profile parameterization (KPP) or non-local, turbulent kinetic energy (TKE) formulation.

Experiment	PBL scheme	PBL type	SL scheme
CTRL	Mellor-Yamada-Nakanishi-Niino Level 2.5[1]	TKE	MYNN
QNSE	Quasi-Normal-Scale-Elimination[2]	TKE	QNSE
MYJ	Mellor-Yamada-Janjic[3]	TKE	MYJ
YSU_MM5	Yonsei University[4]	KPP	revised MM5[11]
TEMF	Total Energy – Mass Flux[5]	TKE	TEMF
BOLA_MM5	Bougeault-Lacarrére[6]	TKE	revised MM5[11]
SH_MM5	Shin-Hong[7]	KPP	revised MM5[11]
ACM_MM5	Asymmetric Convective Model 2[8]	KPP	revised MM5[11]
UW_MM5	University of Washington[9]	TKE	revised MM5[11]
GBM_MM5	Greiner-Bretherton-McCaa[10]	TKE	revised MM5[11]
MYNN_MM5	Mellor-Yamada-Nakanishi-Niino Level 2.5[2]	TKE	revised MM5[11]

[1] Nakanishi and Niino (2004, 2009)
[2] Mellor and Yamada (1982); Janjic (2002); Sukoriansky et al. (2005)
[3] Mellor and Yamada (1982); Janjić (1994); Janjic (2002)
[4] Hong and Lim (2006)
[5] Angevine et al. (2010)
[6] Bougeault and Lacarrere (1989)
[7] Shin and Hong (2015)
[8] Pleim (2007)
[9] Bretherton and Park (2009)
[10] Grenier and Bretherton (2001)
[11] Jiménez et al. (2012)

precipitation gauges were affected by undercatch (Campbell et al. 2016).

In addition to the snowfall measurements, snowflake crystal habits were recorded by OWLeS scientists at SC and NR using handheld and automated cameras. Automated camera images were recorded at NR by a Hydrometeor Videosonde (HYVIS, Murakami and Matsuo 1990) mounted above the snow surface. Also, a Particle Size Velocity (Parsivel) optical disdrometer (Löffler-Mang and Joss 2000) was located at SC, as part of the University of Alabama - Huntsville Mobile Integrated Profiling System (MIPS) instrument platform.

The Parsivel detects precipitation using a horizontal laser beam sheet that is intercepted by falling particles, briefly reducing the voltage measured by the receiver, with an amplitude decrease proportional to a particle's size while the duration of the decrease is

Table 2.3: Summary of parameterization configurations used in NWS operational convection-permitting models.

Model	MP scheme	PBL/SL	ICs	BCs
HRRR	Thompson Aerosol-Aware	MYNN	RAP-1h	RAP-1h
NAM Nest	Ferrier-Aligo[1]	MYJ	NAM	NAM
HRW ARW	WSM6	YSU	RAP	GFS-6h
HRW NMMB	Ferrier-Aligo[1]	MYJ	RAP	GFS-6h
HRW NSSL	WSM6	MYJ	NAM	NAM-6h

[1] Aligo et al. (2018)

proportional to a particle's fallspeed. Since the Parsivel was designed to measure liquid precipitation, the use of the instrument can be subject to biases when measuring non-spherical, frozen hydrometeors due to measurement limitations such as particles only partially filling the Parsivel laser beam. These issues include small negative biases in the sizes recorded, as compared to measurements from a two-dimensional video disdrometer (Battaglia et al. 2010). Another study of mixed-precipitation events ignored Parsivel observations of all particles with diameters smaller than 1 mm due to partial beam-filling artifacts (Yuter et al. 2006). This study uses the same methodology as Yuter et al. (2006), discarding particles with $D < 1$ mm. The Parsivel manufacturer also recommends the instrument to not be used in windy conditions, although it is unclear what threshold of wind speed is considered "windy". During IOP2b, MIPS surface observations recorded average wind speeds between 3 and 6 m s^{-1} with only occasional gusts above 10 m s^{-1}. Therefore, all Parsivel data during IOP2b is used in this study, although this choice may retain small instrument biases due to measurement uncertainty in non-calm conditions.

2.2.2 Radars

To obtain a gridded observed precipitation estimate, we use data from the NEXRAD WSR-88D radar located at Montague, NY (KTYX) along with the OWLeS manual measurements at SC and NR (Fig. 2.2). KTYX Level II base reflectivity data (0.5-degree elevation angle) were downloaded from the National Centers for Environmental Information NEXRAD archive hosted by Amazon Web Services. KTYX data were available at approximately 5-min

intervals and were interpolated to a Cartesian grid with gridboxes 1 km $\times$ 1 km in size using Py-ART (Helmus and Collis 2016). A reflectivity - precipitation rate (Z-R) relation typically used for Western U.S. cool-season precipitation ($Z = 75R^2$) was chosen because the Z-R relation best matched the manual OWLeS observations (Campbell et al. 2016). Typically, $Z = 180R^2$ is used for lake-effect snow, or $Z = 130R^2$ for Eastern U.S. cool-season precipitation (Campbell et al. 2016). Precipitation rate estimates were computed for each KTYX radar scan assuming a constant precipitation rate between scans, multiplying the precipitation rate by the time interval of a single scan using centered differencing, and summed to obtain a gridded estimate of the event-total LPE over the region surrounding the Tug Hill Plateau.

While the gridded reflectivity and LPE is relatively accurate near and over Tug Hill in the vicinity of the OWLeS surface snowfall observations, radar beam overshooting of the shallow LeS bands becomes an issue at greater distances from KTYX. For example, the 0.5° elevation scan from KTYX reaches 1 km AGL at the eastern shoreline of Lake Ontario (Brown et al. 2007). In addition, the Maple Ridge wind farm east-northeast of KTYX (see Fig. 2.2 for location) causes interference with the radar beam, artifically increasing reflectivity and radar-estimated LPE on the northeastern part of Tug Hill.

In addition to the operational NEXRAD network of scanning radars, five ground-based profiling radars were deployed during OWLeS, including an X-Band Profiling Radar (XPR) at SC (Fig. 2.2). A part of the MIPS platform, XPR operates at a wavelength of about 3 cm (9.41 GHz) and a beamwidth of 1.2° (Phillips 2009). Data from the lowest three XPR range gates are discarded due to near-field effects, following Minder et al. (2015). Minder et al. (2015) also note that mean particle sizes measured by the MIPS Parsivel disdrometer during the event were about $D = 1$ mm, with particle sizes occasionally exceeding 5 mm. These particles were large enough to cross from a Rayleigh scattering to Mie scattering regime occasionally during IOP2b.

2.2.3 Profiling radiometer

In addition to the XPR and Parsivel, the MIPS instrument suite includes a vertically-pointing passive microwave profiling radiometer (MPR). The MPR is a 12-channel radiometer that estimates profiles of temperature and water vapor, and also derives a profile of cloud liquid water content by measuring passive microwave radiances at several frequencies (Ware et al. 2003). Note that radiometer observations in some cases may be suspect or biased, but Welsh et al. (2016, their Fig. 10) show that MPR values at several altitudes are in good agreement with UWKA in-situ measurements during the aircraft flight, so the MIPS radiometer data appear reliable in this case.

2.2.4 Upper air

Upper-air soundings from several university teams were launched near Tug Hill, along the southern shore of Lake Ontario, and upwind of the lake at Darlington, Ontario, during the event. In this study, we focus on GRAW DFM-09 GPS-based soundings taken at NR within the core of the LeS band (Steenburgh et al. 2014). Soundings from other upwind locations outside the LeS bands are not discussed here, since they are outside the focus of this study.

2.2.5 Airborne cloud physics

To complement the ground-based observations, the University of Wyoming King Air (UWKA) provided airborne radar and in situ measurements during OWLeS IOP2b (flight track shown in Fig. 2.2). The Wyoming Cloud Radar (WCR, Wang et al. 2012) provides measurements using zenith, nadir, and forward-slanted antennas (3-mm wavelength, W-band). Among other instruments onboard the aircraft, two types of optical array probes providing hydrometeor particle size distribution (PSD) information are mounted on the wings: a Cloud Imaging Probe (CIP) and 2D-Precipitation probe (2D-P). The CIP has a bin width of 25 μm and bin sizes ranging from 0.01 to 2.51 mm, while the 2D-P has a bin width of 200 μm and bin sizes ranging from 0.1 to 20.1 mm.

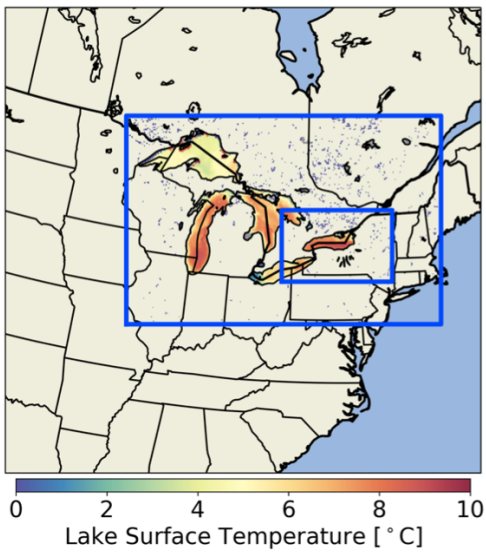

Figure 2.1: Fixed lake-surface temperature (colored shading) and ice cover (gray shading) used in all WRF simulations, from the GLOFS 0-h lake model analysis at 1200 UTC 10 December 2013. Figure boundaries denote the extent of the outermost WRF domain, with two nested domain locations shown with blue rectangles. Lake-surface temperature and ice cover are only plotted from domain 2.

However, it is important to note that observational measurement uncertainties potentially limit the smallest resolvable PSD bin data of both the CIP and 2D-P probes. These uncertainties include ice shattering on the probe tips, and depth-of-field definition and out-of-focus error limitations of the instruments (Korolev et al. 2013). Following the recommendations of Korolev et al. (2013), data from bins with median diameters smaller than four laser beam pixels for each respective instrument ($D < 60$ μm for the CIP and $D < 800$ μm for the 2D-P) are masked in this study, consistent with the PSD analyses from (Welsh et al. 2016).

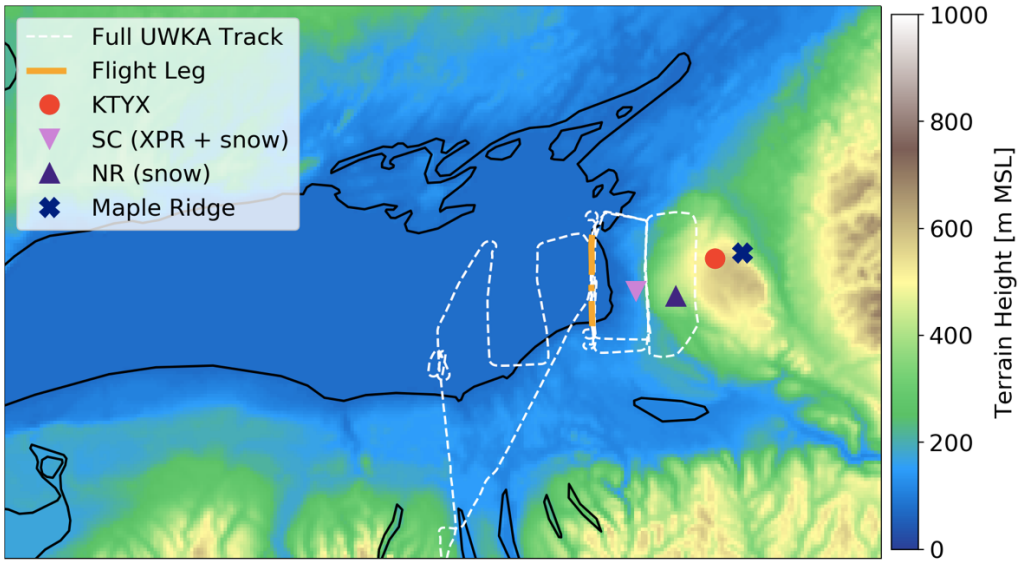

Figure 2.2: Geography of eastern Lake Ontario and the Tug Hill Plateau region, with terrain shaded (m MSL) and operational and research instrument locations during OWLeS IOP2b. The full University of Wyoming King Air flight path during the IOP is shown by the white dashed line, although this study focuses on three vertically-stacked flight tracks denoted by the orange line. Manual snow observation locations are noted by purple triangles, with a downward-pointing triangle denoting a location with both XPR and manual snow observations. KTYX NEXRAD radar and Maple Ridge wind farm locations are shown by a red circle and navy cross, respectively.

3. Event overview

The 10–12 December 2013 LeS event occurred during a period of broad upper-level troughing over eastern North America, transporting cold Arctic air across the Great Lakes under westerly flow (Fig. 3.1a,b). Two shortwave troughs were embedded in the zonal flow during this period, aiding large-scale ascent. During the first of these trough passages on 10 December, an elongated area of low pressure skirted the east coast of the U.S.. As surface high pressure built in behind the system with cold air advection aloft, LeS developed under westerly flow across all of the Great Lakes which were largely ice-free with only partial ice cover in several bays, including Green Bay, Saginaw Bay, and northwestern Georgian Bay (Fig. 2.1). Lake surface temperatures were generally between +5 to +9 °C during the event, with the warmest temperatures in southern Lake Michigan and central Lake Ontario. Combined with 850-hPa temperatures around -15 °C over Lake Ontario throughout the event (Fig. 3.1b), lake-induced instability (estimated by the lake surface to 850-hPa temperature difference) was around +24 °C, well above the +13 °C threshold used by NWS forecasters to indicate potential for significant LeS (Niziol et al. 1995). Downwind of Lakes Erie and Ontario, a modest period of lake-effect snow developed early in the day on 10 December, subsiding by 2300 UTC (OWLeS IOP2a) before the onset of the primary LeS event (OWLeS IOP2b).

LeS redeveloped around 0000 UTC 11 December downwind of Lake Ontario, with the heaviest precipitation remaining quasi-stationary over Tug Hill until 0000 UTC 12 December (Fig. 3.2). During 11 December, westerly 850-hPa winds around 15-20 m s^{-1} over the eastern Great Lakes (Fig. 3.1b) maximized the length of open-water fetch across Lake Ontario, allowing intense LeS to develop and extend inland as far east as the southwestern Adirondack Mountains. LeS band structures from KTYX radar observations during the event exhibited a variety of morphologies, including periods of long-lake-axis-parallel and broad coverage banding as defined by Veals and Steenburgh (2015). The majority of the event was

characterized by weakly-banded and non-banded broad coverage (Fig. 3.2b-d); however, two extended periods of more-organized LLAP banding occurred between 0050 and 0320 UTC 11 December and between 1745 and 2101 UTC 11 December (Fig. 3.2e), as classified by Campbell et al. (2016). After 0030 UTC 12 December, a gradual shift to west-northwesterly boundary-layer winds behind the second shortwave trough passage caused the primary band downwind of Lake Ontario to collapse southward (Fig. 3.2f), dissipating along the southern shoreline of Lake Ontario by 0700 UTC.

A sequence of observed soundings launched from NR (often within the LeS band) on 11 December depict the thermodynamic environment in and near the banded convection (Fig. 3.3, adapted from Campbell et al. 2016, their Fig. 6). Note that dewpoint data in several soundings may be suspect due to sensor icing, as the dewpoint lines closely parallel several of the temperature curves (particularly Fig. 3.3c-e). Throughout the event, a convective boundary layer was present, varying in depth (occasionally as deep as 650 hPa) with a capping inversion located above this feature (Fig. 3.3). The height of the capping inversion varied with time, due in part to variations in the strength of synoptic-scale subsidence behind each of the shortwave troughs and variations in band morphology affecting the depth of lake-effect convection.

Time-height plots of reflectivity and vertical (radial) velocity from the MIPS XPR at SC show the changes in LeS band intensity and depth throughout the event 3.4. During the two extended, well-organized LLAP band periods (0050–0320 UTC and 1745–2101 UTC), the XPR measured an increase in band depth with echoes extending above 3 km MSL. These periods also had stronger reflectivity cores exceeding 30–35 dBZ_e, along with an increase in frequency and intensity of mesoscale updrafts, with upward vertical velocities occasionally exceeding +2 m s^{-1}. Intermediate periods between 0330 and 1730 UTC 11 December had shallower, less-intense convection measured by the XPR, either due to broad coverage banding becoming less-organized and shallower in depth or due to LeS bands occasionally shifting north or south of the XPR location.

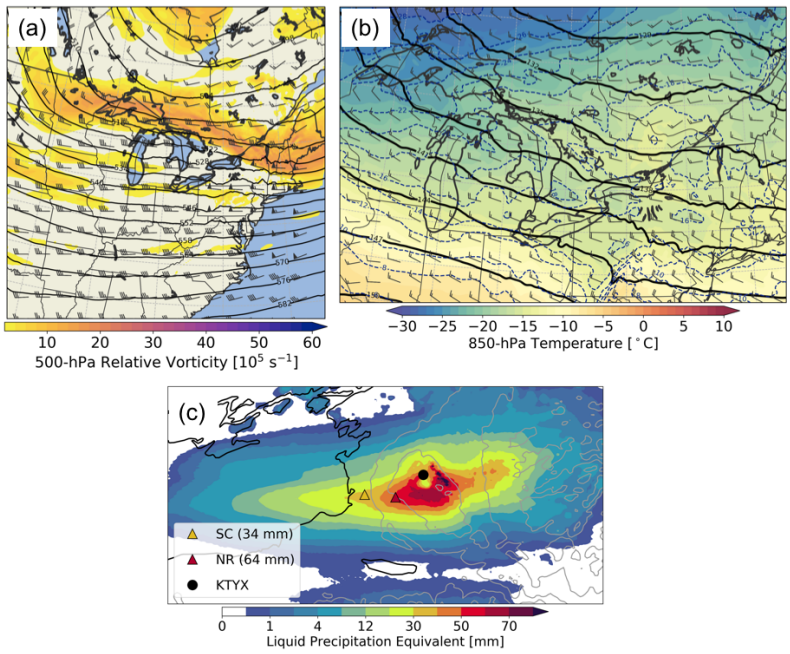

Figure 3.1: WRF control simulation (a) 500-hPa geopotential height (contours, dam), wind (barbs, m s^{-1}, full barb every 10 m s^{-1}), and relative vorticity (shading, 10^5 s^{-1}), (b) 850-hPa geopotential height (solid contours, dam), wind (barbs, m s^{-1}), and temperature (shading, °C), valid at 0600 UTC 11 December 2013. (c) KTYX-estimated, 24-h liquid precipitation equivalent (LPE, shading) from 11 December with manual surface snow LPE measurements at SC and NR shown as triangles, following the same color scale. KTYX location in (c) is denoted by a black circle.

Due to the persistent LeS banding over Tug Hill, the heaviest radar-estimated LPE amounts were quite localized along a west-east line intersecting the central Tug Hill and southwestern Adirondacks, with a maximum around 64 mm on Tug Hill (Fig. 3.1c). KTYX radar beam interference with the Maple Ridge Wind Farm (see Fig. 2.2 for location) caused spurious LPE values higher than 70 mm on the northeastern portion of Tug Hill. As compared to the windward side of Tug Hill, radar-estimated precipitation amounts in the lee of Tug Hill and the southwestern Adirondacks were less certain due to several factors, including radar beam overshooting of the shallow LeS bands and a lack of manual precipitation observations in the sparsely populated region to corroborate the radar-estimated LPE.

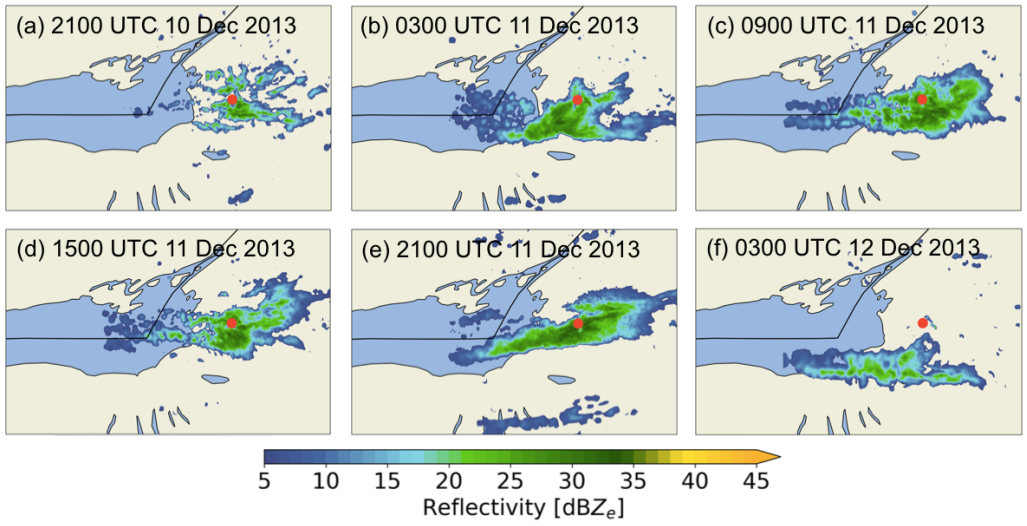

Figure 3.2: Equivalent radar reflectivity (0.5° elevation, shaded) from KTYX NEXRAD radar (red circle) at (a) 2100 UTC 10 December, (b) 0300 UTC 11 December, (c) 0900 UTC 11 December, (d) 1500 UTC 11 December, (e) 2100 UTC 11 December, and (f) 0300 UTC 12 December 2013.

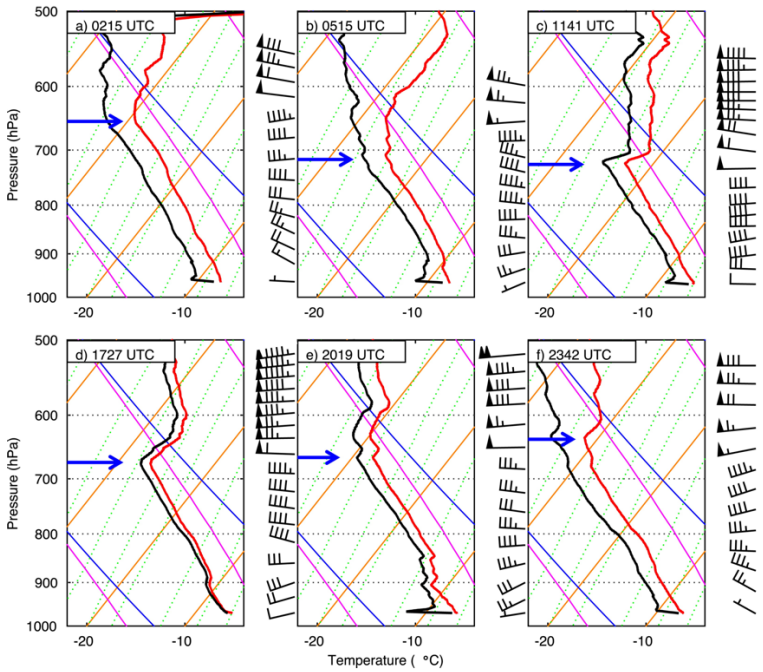

Figure 3.3: Skew T–logp diagrams [temperature (red), dewpoint (black), and wind barbs (full and half barbs denote 5 and 2.5 m s^{-1}, respectively)] at NR for (a) 0215, (b) 0515, (c) 1141, (d) 1727, (e) 2019, and (f) 2342 UTC 11 Dec 2013. Blue arrows denote subjectively identified convective boundary layer top height. Figure and caption adapted from Campbell et al. (2016, their Fig. 6). ©American Meteorological Society. Used with permission.

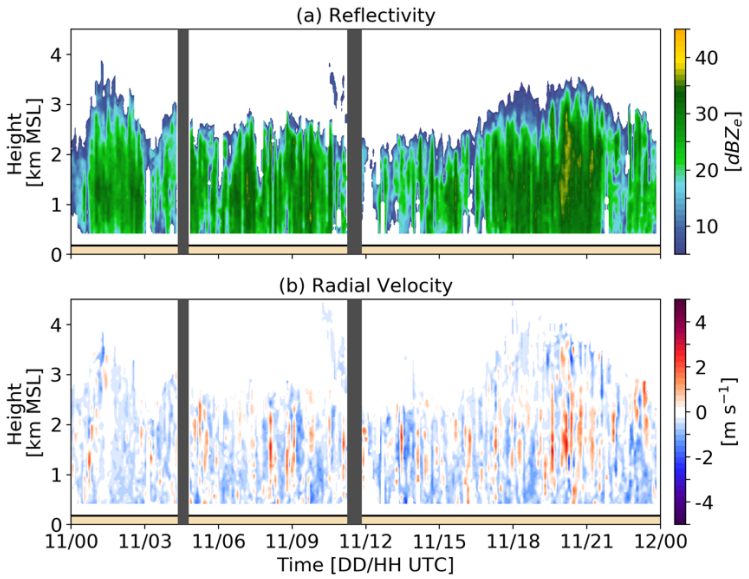

Figure 3.4: Time-height profiles of (a) equivalent radar reflectivity factor and (b) radial velocity from the MIPS XPR profiling radar at Sandy Creek for the 24-h period ending at 0000 UTC 12 December 2013. Gray boxes denote periods with missing data.

4. Sensitivity to microphysical parameterizations

4.1 Evaluation against precipitation and radar observations

Compared to the observed LeS bands, the MP ensemble accurately simulates the large-scale band structures during the event. Figure 4.1 shows a comparison of the KTYX observed base reflectivity and model-simulated reflectivity interpolated to 1 km MSL. At 0900 UTC 11 December, broad coverage banding is present in both the observations and model forecasts. However, the LeS band morphology differs between schemes, with some schemes exhibiting more linear bands (e.g., CTRL, MORR, WSM6) and others depicting more cellular convection (e.g., MYAU, NSSL, P3). LeS band intensity also varies between MP schemes, with several members producing reflectivity values in excess of 35-40 dBZ_e within the band cores (especially MORR). KTYX observations depicted a mixture of linear and cellular convection during the event (Fig. 3.2). Other forecast hours (not shown) show morphology and intensity differences generally consistent with Figure 4.1.

Qualititatively, no single MP scheme performed better than the others, with all schemes exhibiting minor morphology errors at different times during the event (e.g., biased toward linear or toward cellular morphology). Overall, position differences are minimal among all MP ensemble members. Recall that, by design, the simulations are configured to have minimal synoptic-scale forecast errors affecting LeS band position since initial and boundary conditions are taken from RAP analyses and nudging is performed on the outer domain.

Since LeS band position errors in the ensemble are relatively small, the spatial distribution of the event-total quantitative precipitation forecast (QPF) and the location of the QPF maximum also closely match the radar-estimated 24-h precipitation ending at 0000 UTC 12 December (Fig. 4.2). All members have the axis of maximum QPF extending from west to east across Tug Hill, with a maximum over the higher terrain of Tug Hill, and only a slight southward bias of the heaviest QPF region. However, QPF amounts are substantially

different between the MP ensemble members and KTYX observations. All MP schemes underforecast the maximum over Tug Hill, ranging from 35 to 62 mm (observed maximum LPE of 64.5 mm). Schemes such as CTRL, FERR, NSSL, and WSM6 produce relatively more LPE over Tug Hill, while schemes such as MYAU, MORR, P3, and SBUL produce less LPE. The north-south extent of measureable QPF greater than 0.5 mm in the MP ensemble members is also slightly narrower than in the radar-estimated precipitation.

Overall, CTRL had a less negative QPF bias than many of the other MP members. Figure 4.3 shows the QPF amount differences relative to CTRL, further highlighting the substantial underforecast bias by several schemes (e.g., MYAU). A few schemes (e.g., WSM6) exhibit a dipole signature in the QPF difference field shown in Figure 4.3, with an increase in precipitation over the windward slopes of Tug Hill (reduced bias at NR and SC) and a decrease in leeward precipitation.

4.1.1 Verification of LeS band properties

To quantify differences in the modeled and observed LeS band structures leading to differences in event-total QPF, LeS band statistics are performed over an area on the western slope of Tug Hill upwind of KTYX. Analyses are limited to a relatively small rectangular area within 100 km of the radar to minimize the effects of radar beam overshooting at longer distances (see inset in Fig. 4.4a for location). A small area encompassing the Maple Ridge wind farm is also masked in both the observations and WRF output, since the radar observations are affected by wind turbine artifacts. The following analyses use the hourly precipitation output from the ensemble members and the radar-estimated precipitation.

Figure 4.4a shows a time series of LeS intensity, calculated as the 99th percentile of the hourly precipitation within the box, with MP ensemble members showing a wide range of forecast intensities (colored lines) above and below the observed intensity (black dashed lines). Note that the observed (radar-estimated) line is uncertain and based on the choice of Z-R relation for western U.S. cool-season precipitation ($Z = 75R^2$), so two other Z-R

relations commonly used in the eastern U.S. ($Z = 180R^2$ and $Z = 130R^2$) are also plotted using gray shading to emphasize the uncertainty in the radar estimate.

Observed intensities increase as the event begins around 0000 UTC on 11 December, and reach an approximate steady-state around 3.5 mm hr^{-1} until 0000 UTC 12 December, when the LeS bands weaken. The majority of MP ensemble members have similar maximum LeS intensities, with minor fluctuations in precipitation rate during the event. However, several members substantially overforecast LeS intensity, outside the range of observational uncertainty (especially WSM6). Therefore, the choice of MP scheme has a large effect on the precipitation intensity in this case study, with differences up to a factor of two as seen in the event-total QPF.

Similar to intensity, Figure 4.4b shows a time series of total band area, calculated for the same box by considering grid points with an hourly QPF greater than 0.5 mm. During most of IOP2b, all MP ensemble members forecast bands about 1000-1500 km^2 smaller than observed, particularly during the middle of the event between 0600 and 2200 UTC on 11 December. A sensitivity experiment rerunning CTRL without the innermost 1.33-km domain and calculating the total band area for the 4-km domain has a similar small area bias (not shown), suggesting that the small bias result may not be dependent on the choice of 1.33-km horizontal grid spacing for the innermost domain. The underforecast area bias occurs while the observed LeS band was quasi-stationary over Tug Hill, during a period when Campbell et al. (2016) classify most of the radar observations as weakly banded or non-banded. Higher observed orographic ratios, or fraction of orographic enhancement, accompanied these non-banded periods.

Using the same methodology as above, box-average precipitation is computed in Figure 4.4c. The box-average precipitation is too low across all members, due in part to the small band area bias between 0600 and 2100 UTC 11 December. Note that several schemes (e.g., WSM6, NSSL) have compensating biases with over-forecast intensity and under-forecast area, reducing their overall box-average precipitation bias during this period. Before 0600

UTC, the areal extent and intensity of the LeS band is well-forecast by the MP ensemble. After 0000 UTC 12 December, the observed and modeled LeS bands are weakening and shifting south due to a change in wind direction from westerly to northwesterly, reducing the over-lake wind fetch.

Finally, mean band position is quantified for both the KTYX observations and MP ensemble members by calculating the hourly precipitation mass centroid latitude (y_p),

$$y_p = \frac{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} p_{i,j} \; y_{i,j}}{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} p_{i,j}} \tag{4.1}$$

where $p_{i,j}$ and $y_{i,j}$ are the hourly precipitation and latitude at a single grid point, respectively. N_x and N_y are the number of gridpoints in the x- and y-directions for the red bounding box in Figure 4.4d. To visualize any north-south position differences between the MP ensemble and observations, time series of the mass centroid latitudes are plotted in Figure 4.4d. For this analysis, centroid calculations use a much larger box than in Figures 4.4a-c to capture the full extent of all LeS bands downwind of Lake Ontario, including later times during 12 December when the LeS bands began shifting south.

Compared to the observed centroid latitude, which remains steady around 43.7 °N during much of the event, the ensemble member centroid latitudes are generally similar except for a southward oscillation around 1200 UTC 11 December. Furthermore, the MP ensemble members are clustered in their north-south centroid positions and are more similar to each other than to the observations. Thus, for this case study, the choice of MP scheme mostly affects LeS intensity and does not affect band position or size.

4.2 Evaluation against observed sounding

Compared to the observed sounding from NR at 1727 UTC 11 December, modeled soundings from 1700 UTC at the nearest gridpoint to NR show a similar thermodynamic

profile in the lower troposphere (Fig. 4.5a). However, while convective, well-mixed planetary boundary layers are present in both the observations and MP ensemble member forecasts, the height of the capping inversion is about 720 hPa in all the model forecasts versus 670 hPa observed. The observed inversion is also much sharper than the model forecast inversions. Temperatures at the base of the inversion are around -21 °C in the model forecasts versus -26 °C observed. While not directly comparable to cloud top temperature, the warmer temperatures at the inversion base (by about 5 °C in the model forecasts) could have microphysical implications for ice nucleation and growth, since lake-effect clouds grow tall enough to reach the capping inversion during IOP2b (Minder et al. 2015, their Fig. 6). Model parameterizations for ice nucleation are often a function of temperature, supersaturation, aerosol concentration, or a combination of the three, and LeS forecast QPF amounts and precipitation types can be sensitive to the choice of parameterization (Gaudet et al. 2019). Also, glaciation times for mixed-phase clouds have been shown to be strongly sensitive to temperature and ice nuclei concentrations (e.g., Korolev and Isaac 2003; Sulia and Harrington 2011), and the thermodynamic profile in Figure 4.5a passes through several regions where non-spherical, efficient crystal growth is favored [e.g., dendritic growth zone (DGZ) between approximately -12 °C and -18 °C].

To understand whether MP ensemble members have similar forecast kinematic conditions important for condensation and deposition rates, vertical velocity profiles are plotted from each of the simulations. Figure 4.5b shows 24-h average vertical velocity profiles from each of the MP simulations at the nearest gridpoint to SC. The level of maximum upward velocity occurs around 1.2 km MSL in each of the ensemble members, within the DGZ located between 1 and 3 km MSL during IOP2b. However, the average vertical velocity magnitude varies between members, with several schemes (WSM6, CTRL, GDRD) exceeding 0.4 m s^{-1}. These differences in vertical velocity may correspond to cloud physics differences between MP ensemble members; for example, a stronger-updraft member like WSM6 might produce more cloud water aloft if deposition rates are unable to keep up and convert cloud water to

snow and graupel.

4.3 Evaluation against surface precipitation-type observations

After characterizing the MP ensemble precipitation forecasts and their position and intensity biases, we seek to better evaluate the cloud and precipitation properties predicted or assumed by each MP scheme against observations to evaluate their physical realism, beginning with surface precipitation-type observations. Within the observed LeS bands at the surface, mainly dendrites and clusters of dendritic aggregates were recorded by snow measurement teams at SC and NR (e.g., Fig. 4.6f) and other research groups launching soundings during the event. A Hydrometeor Video Sonde (HYVIS) snowflake camera located at NR also captured dendrites and aggregates as the primary crystal habits (Fig. 4.6a-c). However, occasional episodes of partially-rimed particles and graupel were observed infrequently (Fig. 4.6d,e,g) during periods of more intense, deeper convective precipitation (e.g., 1800-2100 UTC 11 December) observed by XPR (Fig. 3.4) and KTYX (Fig. 3.2e), as discussed in Welsh et al. (2016). Overall, high snow-to-liquid ratios (SLRs) observed at SC and NR (14.1 and 16.3, respectively, for the 24-h event) corroborate the manual and automated camera observations of primarily unrimed snow during this event, consistent with habit and SLR observations from other northeast U.S. winter storms (e.g., Colle et al. 2014).

Complementing the camera observations of crystal habits, the MIPS Parsivel disdrometer at SC also measured hydrometeor properties throughout IOP2b. Figure 4.7 shows a 2-D histogram of Parsivel-measured particle counts normalized by bin size in velocity-diameter (V-D) space during IOP2b. Solid, dashed, and alternating dash-dotted green lines indicate CTRL (Thompson MP) V-D relationships for snow, graupel, and rain. Gray solid (dashed) lines indicate V-D relations for snow (graupel) from several other MP schemes used in this study (MORR, MYAU, NSSL, SBUL, and WSM6 snow; and WSM6 graupel). Note that all of the plotted snow V-D lines are quite similar, resulting in nearly indistinguishable snow V-D relationships. Model V-D relationships for different precipitation types typically follow

a power-law formulation [$V(D) = aD^b$, where a and b are fixed coefficients and D is the particle diameter] in most schemes in this study, although some schemes like P3 (not shown in Fig. 4.7) use more complex parameterizations to account for variable riming, dynamically changing ice particle fallspeeds. The thick black line denotes the 24-h median observed V-D values from all of the Parsivel data during IOP2b, while the thin black lines represent the interquartile range of the V-D observations over the 24-h period.

Compared to the V-D relations in CTRL, the observed V-D falls in between snow and graupel, but is closer to snow velocities as represented in Thompson (and in the other MP members), suggestive of snow with light to occasionally moderate riming. Furthermore, the observed V-D is not constant with time as shown by the median and interquartile lines. These Parsivel observations are consistent with the crystal habit observations of mainly dendritic aggregates with intermittent riming. Most schemes' V-D relations for snow (graupel) have fallspeeds that are too small (large) relative to the observed fallspeeds characteristic of partially rimed snow, which most schemes are not explicitly representing.

The differences between observed and modeled V-D highlight one of the limitations of many operational bulk MP schemes, which use separate categories to predict snow and graupel, although newer schemes such as SBUL and P3 remove this artifical distinction by having one evolving ice category with variable properties, such as riming (Lin and Colle 2011; Morrison and Milbrandt 2015). Thompson MP applies a simpler scaling method to the V-D relation for snow to account for partial riming, using a constant multiplication factor varying between 1.1 and 1.5 if a certain threshold for riming fraction is reached. In Thompson, the riming fraction is defined as the ratio of the process rate of snow collecting cloud water to the rate of snow deposition, with the upper limit for partial riming set equal to that needed for the scheme to convert all snow to graupel. However, this scaling of the snow V-D to account for partial riming is rarely activated in CTRL. Thus, the choice of a constant V-D relation for both pure snow and graupel as in Thompson potentially introduces error into the model forecast for IOP2b, since the observed velocities vary with time due to factors such as

variations in riming and crystal habit associated with different microphysical environments within the LeS band.

Compared to the observations, the distribution of modeled precipitation types varies widely among MP ensemble members, mainly comprised of snow and graupel. Rainfall is negligible in all ensemble members during this IOP, with surface temperatures around -4 °C measured by the MIPS at SC and -5 to -8 °C on Tug Hill during the event, with slightly warmer conditions over eastern Lake Ontario around -1 to -3 °C. Figure 4.8 shows the spatial distribution of the fraction of the total precipitation falling as graupel from each MP ensemble member. Since P3 and SBUL do not explicitly calculate graupel, the instantaneous riming fraction in each member (for SBUL, rime mass mixing ratio divided by the total mixing ratio) is multiplied by the hourly precipitation output and summed to approximate the graupel fraction in these schemes. Therefore, they are directly comparable to each other, but not to the other members.

Overall, the majority of members produced very little graupel ($< 15\%$), while several members (MYAU, SBUL, WSM6) had swaths of high graupel fraction in excess of 40-70%. Substantial differences in riming and graupel fraction between MP schemes appear to be important for explaining differences between members in the spatial distribution of the heaviest QPF near Tug Hill. The primary graupel-producing members (MYAU, SBUL, WSM6) have the largest precipitation differences relative to CTRL (Fig. 4.3e,i,j) with less precipitation over the lee slopes and immediately downwind of Tug Hill. Only WSM6 (Fig. 4.3j) also produces substantially more precipitation over the windward slopes of Tug Hill, forming a dipole pattern.

Since WSM6 and MYAU produce lower ratios of leeward to windward QPF relative to CTRL, these particular schemes exhibit consistent performance with the results of (Reeves and Dawson 2013), where WSM6 and MYAU had greater shoreline versus inland QPF as compared to Thompson. However, SBUL also appears to have a lower leeward/windward QPF ratio in this case than in (Reeves and Dawson 2013). In their Lake Erie case study,

(Reeves and Dawson 2013) showed that MP members with higher QPF near the lakeshore had shorter in-cloud precipitation residence times due to more rapid sedimentation of graupel than snow. However, unlike their results, graupel-producing schemes in this case did not consistently produce higher maximum QPF amounts near the lakeshore, since our results show maximum precipitation amounts among graupel-producing MP schemes vary by a factor of two with some members (such as MYAU) severely underforecasting QPF over Tug Hill. While we do not have data to accurately characterize graupel fractions across the Tug Hill region, Parsivel and SLR observations support that at least MYAU and SBUL and possibly WSM6 produce too much graupel in this case. The sensitivity to precipitation type among the MP ensemble members in this case (for example, causing a reduction in precipitation leeward of Tug Hill for WSM6 but not for Thompson), could have implications for forecasting the inland extent of the heaviest LeS amounts if these precipitation-type differences remain in other case studies.

4.4 Evaluation against cloud physics observations

To investigate why some MP ensemble members produce excessive graupel, we analyze the amount of supercooled liquid water present in the model output since several sources of cloud water observations were available during OWLeS. Figure 4.9 shows the spatial distribution of the cloud liquid water mixing ratio (q_{cl}) vertically integrated over all model levels, or cloud liquid water path (CLWP), from each of the MP ensemble members, averaged over the 24-h period ending at 0000 UTC 12 December 2013,

$$CLWP = -\frac{1}{\rho_w g} \int q_{cl} \; dp. \tag{4.2}$$

All members except WSM6 have swaths of low, but non-zero, CLWP within the simulated LeS bands. CLWP in WSM6 is non-zero but does not reach the lowest plotted contour level of 0.05 mm, and only occasionally exceeds 0.05 mm at times in localized areas over the lake. The majority of members have more CLWP over Lake Ontario where updrafts are strongest,

decaying with eastward extent as cloud water is converted to snow and graupel. However, several members (MORR and SBUL, and especially MYAU and P3) have the opposite trend, with increasing CLWP downwind over Tug Hill.

In order to evaluate the model forecasts of CLWP, we compare to point measurements of observed CLWP from the MPR located at SC. Figure 4.10 shows time series of CLWP (solid lines) from the MPR and from the nearest gridpoint in each of the MP ensemble members. Since the vertically-pointing radiometer takes measurements every five minutes and is a noisy, derived measurement, hourly-average values are calculated to compare to the hourly output from the MP ensemble. Shading on the observed panel shows the full temporal range of values from the MPR, while shading on the model panels depicts the spatial range of values over a 20-gridpoint north-south range (26.6 km) centered on SC. Since CLWP varies over short distances (as seen in Fig. 4.9), the range is shown in the MP ensemble member panels as a rough estimate of potential impacts of biases in the N-S position of forecast LeS bands on simulated CLWP. Choosing a different range between 5 and 40 gridpoints does not substantially affect the model results. Despite the sampling differences between model and observations, several MP ensemble members clearly over- or under-forecast CLWP relative to MPR observations.

Despite the range of CLWP and graupel amounts in each of the MP ensemble members, there is not a simple correlation between high CLWP and high graupel fraction in the model forecasts. For example, while both MYAU and MORR often produce CLWP in excess of 0.2 mm at SC, graupel fraction often exceeds 70% in MYAU while being negligible in MORR (mostly snow). Low-CLWP members such as CTRL (Thompson aerosol-aware) and WSM6 also have different precipitation-type partitioning, since CTRL simulates negligible graupel fraction (mostly snow) while WSM6 simulates fractions that often exceed 20%.

Further analysis of microphysical processes leading to precipitation-type differences would require output of mixing ratio tendencies and conversion terms from each of the ensemble members. For their Lake Erie case study, Reeves and Dawson (2013) analyzed a

graupel tendency budget from high-graupel and low-graupel WRF simulations (WSM6 and GDRD MP schemes). They found the most substantial differences in the budget were due to the accretion of snow by rain term (P_{racs}). In their case study, P_{racs} was an order of magnitude larger in WSM6 than in GDRD, leading to much larger graupel amounts in the WSM6 member.

Similar to the Reeves and Dawson (2013) case, WSM6 produces a substantial amount of graupel while GDRD produces mostly snow during IOP2b. Since the thermodynamic environments between IOP2b and the Reeves and Dawson (2013) case are fairly similar as well, differences in P_{racs} between the two schemes likely explain the differences in precipitation type between the two schemes. Other MP schemes such as Thompson have a slightly different formula for P_{racs}, involving varied formulations for the terminal velocity difference between rain and snow, so graupel budgets for those schemes may yield different results (Reeves and Dawson 2013).

While a complete understanding of the model-simulated microphysical evolution of this OWLeS case is limited without outputting additional MP scheme information, several conclusions can be drawn from the graupel fraction and CLWP analyses. Compared to the surface snowfall observations of high SLRs and dendritic aggregates with minimal riming, MP ensemble members (such as MYAU, SBUL, and perhaps WSM6) that produce excessive graupel are clearly suspect in this case. Several of these schemes also produce excessive CLWP relative to MIPS radiometer observations. In contrast, the HRRR-like CTRL member with Thompson aerosol-aware MP is one of the better members with respect to graupel fraction, producing mostly snow, but is biased towards too little CLWP. No schemes have a perfect representation of the observations of low CLWP and mostly unrimed snow (likely low graupel fraction), but the moderate-CLWP, snow-dominant schemes appear closer to reality in this case.

4.5 Evaluation against profiling radar observations

Since CLWP, snow, and graupel all contribute to the vertical profile of precipitation in the MP ensemble, we compare the model-derived reflectivity to observed equivalent reflectivity (Z_e) profiles from the MIPS XPR at SC (Fig. 3.4) to further evaluate MP scheme realism (Fig. 4.11). FERR and SBUL are omitted from Figure 4.11 because reflectivity is not calculated within these MP schemes. One way of summarizing the time evolution of these profiles in a single plot is a contoured frequency by altitude diagram (CFAD, Yuter and Houze 1995), which show the frequency distribution of reflectivity [per bin size (250 m $\times$ 1.5 dBZ_e)] as a function of height. All CFAD values are normalized by the total number of observation points in an analysis sample, giving units of percentage frequency per bin size. Since the normalization technique is constant with height and consistent across models and observations, CFAD values at different heights and locations can be directly compared. The observed CFAD uses the full temporal resolution (1-second frequency) of the XPR. Model CFADs use the hourly output at all gridpoints within a 30-km $\times$ 30-km box centered on SC, to account for small band position biases in MP ensemble members and to increase the sample size for the CFADs. Choosing a smaller radius for the model CFADs leads to fewer data points and gaps in the CFAD diagram, while a larger radius leads to a shift in the distribution towards lower reflectivity values since more points outside the modeled LeS band cores are sampled.

As seen in Figure 4.11a, the observed CFAD from the XPR has the highest frequency of high Z_e values around 1.0–1.5 km MSL, decreasing aloft and near the surface. At higher altitudes, the increase in Z_e with decreasing height is consistent with the growth of falling hydrometeors, from cloud top at 3.5–4 km MSL downward to 1.5–2 km. Below 1 km MSL, a decrease in Z_e with decreasing height near the surface may be due to below-cloud sublimation or lofting/advection of hydrometeors (Campbell and Steenburgh 2017; Welsh et al. 2016; Minder et al. 2015).

Compared to the observed XPR CFAD, the MP ensemble CFADs (Fig. 4.11b-h) have

profiles that vary in both the depth and frequency of stronger convection. The model CFADs generally lack the higher frequencies of larger Z_e values around 1–1.5 km shown in the XPR CFAD (Fig. 4.11a), in part because the model spatial sampling technique permits profiles outside the band core to be included in the CFAD. In the region of hydrometeor growth above 1.5 km, the majority of MP ensemble members have increasing Z_e with decreasing height, although the slope of the median Z_e with height varies between members. All schemes have LeS bands that are biased shallow relative to the depth of the observed LeS bands, related to the low bias in the modeled height of the capping inversion, with the observed CFAD showing relatively high frequencies of values around 5 dBZ_e above 3 km MSL as compared to the modeled CFADs. Also, the median and width of the modeled CFAD distributions varies substantially above 2 km MSL between MP schemes (GDRD, MYAU, and P3 appear to have distributions that are too wide), perhaps due to microphysical process differences between schemes, such as differences in ice nucleation and diffusional growth.

Near the surface, some ensemble members display a clear decrease in Z_e with decreasing height (e.g., GDRD) below 1 km, which is more consistent with XPR observations. While this decrease in Z_e might be caused in part by hydrometeor lofting or low-level sublimation (perhaps greater for snow), differences in precipitation type between MP members do not explain the near-surface decrease in Z_e found in some schemes. Overall, several MP scheme CFADs (e.g., NSSL, WSM6) are more consistent with observations while others are more biased (e.g., MYAU, P3), both near-surface and aloft.

4.6 Thompson sensitivity experiments

While the multi-MP ensemble has a range of QPF amounts and uncertainty, an in-depth interpretation of the causes of QPF and precipitation-type differences is difficult due to the numerous and complex MP differences between schemes. To simplify the analysis, we perform additional diagnostics on the CTRL member and run several experiments modifying only the Thompson MP scheme. Thompson MP is chosen for further study because it is

commonly used in operational forecasting (e.g., HRRR) and because it performed relatively well compared to other MP schemes in terms of its QPF distribution and snow-dominant cloud physics properties (negligible graupel) during IOP2b, although it underforecast graupel and CLWP. Sensitivity experiments, described in Table 4.1, are performed to understand some of the single-scheme microphysical controls on the QPF distribution during this event. In addition, the following experiments address several parameter values related to snow parameterization that differ among the various MP schemes discussed previously.

4.6.1 Comparison of modeled and observed PSDs

During OWLeS IOP2b, the UWKA was flying in a "lawnmower" pattern with north-south flight legs at several altitudes (approximately 3 km, 1.7 km, and 1 km MSL) through the primary LeS band (Fig. 2.2). Using the output from the model and in-situ observations from the UWKA aircraft, modeled and observed PSDs can be compared. Figure 4.12 shows PSDs along three vertically-stacked legs over the eastern shoreline of Lake Ontario (orange line in Fig. 2.2), between 1905 UTC and 2104 UTC 11 December. Welsh et al. (2016) shows additional analysis of PSDs from the remaining legs flown during IOP2b. Observed PSDs are calculated within the LeS "band core" by averaging over a 15-km interval centered on the instantaneous maximum updraft measured at flight level, following a similar methodology as Welsh et al. (2016).

Model band core PSDs are defined similarly, but averaging over the 15-km interval centered on the instantaneous maximum updraft in the layer between flight level and the surface. This distinction allows for more accurate identification of the primary updraft for the 3-km flight legs where the aircraft flew in-cloud near the top of the observed LeS band, but very close to cloud top in the model LeS bands, which were shallower than observed. While the choice of a 15-km-wide band core is arbitrary, the following PSD results are not particularly sensitive to the horizontal distance used to define the band core between the range of 5 and 45 km, so the 15-km definition is used for consistency with Welsh et al. (2016).

This band core methodology also allows for differences in the position of the observed and modeled LeS bands that may affect the PSD comparison. Also, model output from 2000 UTC is used in all PSD plots, causing a slight temporal mismatch for the 3-km and 1-km UWKA flight legs, but the choice of model output times between 1900 and 2100 UTC does not affect the PSD results since the LeS bands were in a quasi-steady state during this time period.

In the Thompson MP scheme, the snow PSD, $N_s(D)$, is calculated using a formula summing two different distribution shapes, incorporating both an exponential distribution for small ice particles and a gamma distribution for larger particles, following the work of Field et al. (2005),

$$N_s(D) = \frac{{M_2}^4}{{M_3}^3} \left[\kappa_0 e^{-\frac{M_2}{M_3}\Lambda_0 D} + \kappa_1 \left(\frac{M_2}{M_3} D \right)^{\mu_s} e^{-\frac{M_2}{M_3}\Lambda_1 D} \right]. \tag{4.3}$$

Fixed coefficients of the two distribution terms are the intercept and slope parameters for each distribution, namely κ_0 and Λ_0 for the exponential term and κ_1 and Λ_1 for the gamma term, and the shape parameter, μ_s, for the gamma term. These fixed coefficients were derived from 9000 samples of ice PSDs from aircraft flights over England within mid-latitude stratiform clouds (Field et al. 2005). The M terms are different moments of the size distribution (moment order denoted by subscript), with the second moment calculated directly as a function of the snow mixing ratio, q_s. Appendix C in Thompson et al. (2008) describes the moment calculations in more detail. Only the snow PSD is plotted since rain is negligible in the simulations, and cloud water, cloud ice, and graupel mixing ratios are 1–2 orders of magnitude smaller than snow mixing ratios along the flight legs, so all have minimal contributions to the modeled total PSD.

Using the output mixing ratios interpolated along the flight track, modeled PSDs are calculated for the nearest forecast time to the UWKA flight tracks (2000 UTC 11 December), overlaid with the observations from the CIP and 2D-P (Fig. 4.12). At 3 km MSL, observed

particle concentrations between 100 and 1000 μm are similar within an order of magnitude, before rapidly decreasing for particle bins greater than 1000 μm (Fig. 4.12a). Over the range of particle sizes between 1000 and 2000 μm measured by both the CIP and 2D-P, particle concentration measurements are comparable between the two instruments. At the intermediate altitude (1.7 km), small particle concentrations for sizes less than 1000 μm are similar to those at 3 km, but particle concentrations for sizes greater than 1000 μm are about 10 times larger at 1.7 km than at 3 km (Fig. 4.12b). The increase in concentration of larger particles between 3 km and 1.7 km MSL may be due to processes such as diffusional growth and aggregation, consistent with an increase in Z_e with decreasing height as observed by the XPR (Fig. 4.11a). At the lowest flight leg altitude (1 km MSL), particle concentrations for sizes less than 1000 μm are about 10 times smaller than at 1.7 and 3 km, but concentrations of particles greater than 1000 μm are similar to concentrations at 1.7 km (Fig. 4.12c). Overall, observed PSDs follow a broadly similar distribution in shape between the three flight legs, with small particle concentrations decreasing and large particle concentrations increasing with decreasing flight leg altitude.

When considering the observed PSDs, recall that observational measurement uncertainties are large for the smallest resolvable bin data of both the CIP and 2D-P probes, including artifacts such as ice shattering (Korolev et al. 2013). These uncertainties include ice shattering on the probe tips, and depth-of-field definition and out-of-focus error limitations of the instruments (Korolev et al. 2013). While data from the smallest CIP and 2D-P bins are masked following the recommendations of Korolev et al. (2013), PSD concentrations remain large for the smallest two unmasked CIP size bins in Figure 4.12 and may be biased despite being above the 60 μm threshold used for masking CIP data.

In comparison to the observed PSDs, the model correctly forecasts concentrations of medium particles between 500 μm and 3000 μm but underforecasts concentrations of large particles greater than 3000 μm at 1.7 and 1 km (Fig. 4.12b-c). Concentrations of the smallest particles around 90 μm are also forecast correctly at 1.7 and 1 km due to the exponential

portion of the model PSD formula, although particles between 100 μm and 500 μm are underforecast at 1.0 km (Fig. 4.12c). In contrast, the modeled PSD at 3 km has the correct shape but has about 10 times too few particles at all diameters (Fig. 4.12a). A portion of the 3-km PSD forecast error may be due to the shallow bias of the modeled LeS bands, with modeled PSDs being sampled closer to the cloud top than in the UWKA observations. Also, given the consistent underforecasting of larger particles greater than 2-3 mm at all altitudes, the parameterization used in the Thompson MP scheme may be misrepresenting processes important to the formation of larger snowflakes and aggregates, visible as errors in the evolution of the PSD shape. Since the snow PSD parameterization developed by Field et al. (2005) is a function of snow mixing ratio and temperature, and was empirically derived based on observations from stratiform clouds in mid-latitude cyclones, it may be unrepresentative of the shallow convective cloud environment associated with LeS.

4.6.2 Sensitivity to snow PSD

Given the biases uncovered by the PSD comparison, an additional experiment is run to test the QPF sensitivity to the snow PSD parameterization used in the Thompson scheme, summarized in Table 4.1. This experiment, THOM_PSD, has a modified slope parameter, Λ_1, in the gamma term of the PSD, simulating a snow PSD that better represents the large end of the PSD spectra. A value of $\Lambda_1 = 1.6$ was chosen to attempt to reduce the PSD biases evident in CTRL. Figures 4.12d-f show the comparison between the observed and THOM_PSD results, with the same observed band core PSDs as in Figures 4.12a-c. While the modified model PSDs reduce the underforecast biases for all diameters at 3 km (Fig. 4.12d) and for particles greater than 3000 μm at 1.7 km (Fig. 4.12e), THOM_PSD has an overforecast bias for all diameters at 1 km (Fig. 4.12e). Therefore, this simple modification to Λ_1 is an imperfect solution to the CTRL PSD biases.

Nonetheless, how do these snow PSD modifications impact the QPF distribution across Tug Hill? Figure 4.13c shows the total QPF difference of THOM_PSD relative to CTRL

(Fig. 4.13b). Note that both the CTRL (Fig. 4.13b) forecast and the KTYX-estimated LPE (Fig. 4.13a) are underestimates of precipitation amounts compared to the OWLeS surface snow observations at SC and NR windward of Tug Hill. These amounts and biases are quantified in Table 4.2. Relative to CTRL, THOM_PSD produces more precipitation windward of Tug Hill and less precipitation leeward.

Due to the snow PSD slope modification towards larger particles, mean particle diameters are also shifted towards larger sizes. Larger particles are prescribed a faster fallspeed in the model [$V(D) = aD^b$], thus increasing the sedimentation rate of snow over Tug Hill, resulting in less precipitation over the lee slopes and immediately downwind of Tug Hill. While CTRL has a substantial 50% underforecast QPF bias at both SC and NR (LPE of 17.0 and 31.9 mm, respectively), the THOM_PSD experiment reduces the underforecast bias by about 12% (LPE of 20.9 and 40.7 mm), increasing QPF amounts by 5-15 mm across Tug Hill relative to CTRL (Fig. 4.13c). This analysis indicates that improving snow PSD assumptions may help to improve the spatial distribution of QPF, including the maximum QPF amounts and the inland extent of the heaviest precipitation.

4.6.3 Sensitivity to crystal habit

In addition to the THOM_PSD experiment, more simulations are run to test the sensitivity of LeS to the parameterized snow crystal habit in Thompson MP. In the model, coefficients in two equations are chosen to determine the bulk properties of snow, which implicitly correspond to assumptions about crystal habit and the degree of riming: first, the mass-diameter

$$M_s(D) = a_{m_s} D^{b_{m_s}} \tag{4.4}$$

and second, the fall velocity-diameter

$$V_s(D) = a_{v_s} D^{b_{v_s}} e^{-f_{v_s} D} \tag{4.5}$$

relations for snow. M_s is the mass for a given particle diameter (D) and a_{m_s} and b_{m_s} are empirically-derived constants for snow. Similarly, V_s is the terminal fall velocity for a given particle diameter and a_{v_s}, b_{v_s}, and f_{v_s} are the corresponding empirical constants. Different MP parameterizations use different values for these relation constants, as shown in Figure 4.7. The following experiments are configured the same as CTRL, but with a_{m_s}, b_{m_s}, a_{v_s}, b_{v_s}, and f_{v_s} altered (see Table 4.1).

Different values of the coefficients in M_s and V_s can be used to approximate various types of snow crystal habits, including dendrites and columns, compiled from numerous observational studies by Woods et al. (2007, their Table 1). Figure 4.14 shows several of the M_s and V_s relations, with the greatest differences between crystal habit M_s relations (Fig. 4.14a) and V_s relations (Fig. 4.14b) for large snow particle sizes greater than 2 mm. The Thompson scheme V_s equation (Fig. 4.14b, green solid line) uses an exponential term to reduce the fallspeeds of large snow particles (Thompson et al. 2008; Ferrier 1994), shifting its V_s from one similar to column-like habits (red line) to one similar to bullet-like/cold-type/sector-like habits used in several other MP schemes (gray lines).

As discussed previously, observed crystal habits during IOP2b were mostly characterized as dendritic aggregates, with only occasional periods of heavily rimed particles (Fig. 4.6). MIPS Parsivel observations at SC (Fig. 4.14b) have the majority of particles clustered between the Thompson velocity curves for graupel and snow (dashed and solid green lines, respectively), but closer to Thompson snow, consistent with aggregates being the primary habit with occasional riming. The V_s relations used in Thompson and other MP schemes (gray lines) appear to be a compromise between slowly-falling, dendrite-like crystals (blue line) and faster-falling habits, or a representation of partially-rimed snowflakes that are not rimed enough to be classified as graupel.

To test the sensitivity of QPF to the choice of parameterized crystal habit, two simulations are run, one at each extreme of the habit spectrum, with modifications to both M_s and V_s coefficients. These experiments are based on the observed snow crys-

tal habits. One simulation, THOM_MFV_COLM, uses coefficients representative of relatively fast-falling, column-like crystal habits (red lines in Fig. 4.14). Another simulation, THOM_MFV_DDRT, uses coefficients representative of slow-falling, dendrite-like habits (blue lines in Fig. 4.14). According to the Parsivel data, the THOM_MFV_COLM experiment should be more consistent with observations than THOM_MFV_DDRT, which assumes a mixture of dendrites and dendritic aggregates with no riming (sometimes consistent with HYVIS imagery). A third simulation, THOM_FV_NOEXP, tests the sensitivity to V_s alone by removing the exponential term in the CTRL V_s equation, Eqn. 4.5 (orange lines in Fig. 4.14), making V_s similar to the relation for columns.

To evaluate the realism of V_s throughout the depth of the LeS bands in the above experiments, simulated hydrometeor vertical velocity profiles are compared to radial velocity profiles from the MIPS XPR at SC. Simulated hydrometeor vertical velocities are the sum of the vertical air velocity plus the reflectivity-weighted precipitation terminal fall velocity output from the Thompson MP scheme, to faciliate comparison with the radar observations. Using the same CFAD sampling methodologies for model and observations as in Figure 4.11, with velocity bins of 250 m $\times$ 0.1 m s^{-1} in size, observed and modeled CFADs for velocity are shown in Figure 4.15. Thick solid black lines denote the 50th percentile of the velocity distribution, while thin black lines denote the 25th and 75th percentiles. Observed 25th, 50th, and 75th percentiles from the MIPS XPR radial velocity data follow the same convention, and are overlaid on each panel in magenta for comparison with the model velocities.

Throughout the depth of the approximately 3.5-km deep LeS bands, XPR median velocities remain nearly constant with height around 0.5 m s^{-1} downward, consistent with mostly unrimed snow, with faster downward velocities below 1 km as precipitation particles increase in size (Fig. 4.15a). Below 2 km MSL, a larger spread in outlier values above 1 m s^{-1} upward and below 1 m s^{-1} downward indicates the presence of occasional mesoscale updrafts and downdrafts within the LeS bands. Compared to XPR, CTRL underestimates median velocities above 3.3 km MSL, likely related to the band-depth bias. However, CTRL

has similar median velocities as XPR around 2.5–3 km and below 1 km MSL, with a small overestimate of 0.3 m s^{-1} around 1.8 km (Fig. 4.15b). While CTRL velocities appear relatively unbiased near the surface compared to XPR, Parsivel observations suggested that snow velocities in CTRL were too slow (Fig. 4.14). These differences in velocity bias relative to Parsivel and XPR observations may be due to differences in the vertical location of precipitation sampling by the Parsivel and XPR (around 200 m AGL for the lowest XPR range gate) or biases in the XPR- and Parsivel-observed velocities themselves.

Compared to CTRL, THOM_FV_NOEXP has a marginal increase in downward velocities due to the removal of the exponential V_s term, resulting in an increase in bias relative to XPR (Fig. 4.15c), but is otherwise similar to CTRL. THOM_MFV_COLM (Fig. 4.15d) has a substantial increase in median velocities below 2 km MSL relative to CTRL, along with a widening of the distribution near the surface. In THOM_MFV_COLM, downward velocites below 2 km are about 1 m s^{-1} too large relative to XPR due to the assumed columnar habits, which are inconsistent with Parsivel observations and limited manual and HYVIS crystal images. In contrast, THOM_MFV_DDRT (Fig. 4.15e) has a small decrease in median velocities below 2 km MSL relative to CTRL. While THOM_MFV_DDRT has less-biased velocities around 2 km relative to the XPR observations, it has velocities that are slightly too slow near the surface. Although CTRL and THOM_MFV_DDRT velocities appear to have the smallest biases between 2 and 3 km MSL, the larger near-surface biases below 1 km in THOM_MFV_DDRT indicate that dendrite-like assumptions are poorest there, with faster-falling observed particles perhaps experiencing more riming or less lofting at low levels. Note that changes in model hydrometeor vertical velocity are almost entirely due to changes in precipitation terminal fall velocity, since air vertical velocity CFADs are almost identical between experiments (not shown).

To assess the sensitivity of the model QPF distribution over Tug Hill due to changes in crystal habit assumptions, QPF difference plots are calculated relative to CTRL. Figure 4.13 shows the QPF differences relative to CTRL for THOM_FV_NOEXP (Fig. 4.13d),

THOM_MFV_DDRT (Fig. 4.13e), and THOM_MFV_COLM (Fig. 4.13f). Increasing V_s for large snow particles in THOM_FV_NOEXP slightly increases the amount of windward precipitation on Tug Hill (2 mm at SC and 5 mm at NR relative to CTRL), as snowflakes sediment more quickly. The dipole pattern in the QPF difference is similar but less pronounced than the result from THOM_PSD, which instead increases the amount of large snow particles, increasing bulk V_s without changing the formulation for V_s itself.

Similar to the THOM_FV_NOEXP results, THOM_MFV_COLM has a dipole pattern of the same sign but much larger magnitude (Fig. 4.13f). This indicates that changing both M_s and V_s to more column-like habits (THOM_MFV_COLM) is a more drastic change in this case than just changing V_s alone (THOM_FV_NOEXP), with the changes in M_s likely affecting process rates such as diffusional growth. In contrast to THOM_MFV_COLM, THOM_MFV_DDRT using the reduced M_s and V_s relations for dendrites has a dipole pattern of similar magnitude but opposite sign, dramatically reducing the windward precipitation over Tug Hill and increasing precipitation downwind near the Adirondacks (Fig. 4.13e). Since larger ice particles are falling more slowly in THOM_MFV_DDRT, they are able to be advected further downwind of Lake Ontario.

While the unrimed, dendrite-like snow in THOM_MFV_DDRT is more representative of HYVIS crystal habit observations of mostly dendritic aggregates, the simulation produces a more biased windward QPF pattern than THOM_MFV_COLM. This is likely due to the slowly-falling snow represented in THOM_MFV_DDRT, inconsistent with the faster-falling, partially-rimed snow observed by the Parsivel. Note that it is unclear which simulation performs best downwind of Tug Hill, since radar-estimated LPE amounts downwind are more uncertain due to the lack of manual snowfall observations to constrain the estimated amounts. Note that QPF differences in the four experiments (including THOM_PSD) are solely due to changes in bulk snow properties, since none of the above experiments increase the negligible amount of graupel produced by CTRL. While demonstrating the QPF forecast sensitivity to bulk snow properties, these simple Thompson MP crystal habit experiments (motivated

by observational uncertainty in V_s) are somewhat crude since they assume temporally fixed coefficients whereas the HYVIS and Parsivel observations show mostly aggregate habits with occasional periods of riming, causing temporal variability in V_s.

Although these modifications to Thompson MP likely exaggerate the microphysical and QPF uncertainty during OWLeS IOP2b, they demonstrate the sensitivity of the forecast QPF to changes in snow parameterization. Maximum event-total QPF amounts range between 39 and 64 mm for the single-scheme MP experiments and from 35 to 62 mm for the multiple-scheme MP ensemble. Therefore, the physical uncertainty of the LeS event microphysics is demonstrated with a range of QPF amounts from the single-scheme MP experiments that is almost as large as the full MP ensemble employing numerous MP schemes. However, the above Thompson scheme experiments do not have the same diversity of precipitation types as the full MP ensemble (snow and graupel), although that could be investigated with additional experiments modifying other portions of the MP scheme sensitive to graupel, such as increasing the production of cloud water.

4.7 Discussion

While the specific MP results presented above are limited to a single case study of OWLeS IOP2b, some of the results are consistent with long-standing MP scheme deficiencies noted in diverse microphysical environments. These NWP prediction challenges have been the focus of major field campaigns collecting cloud and precipitation observations for improving MP schemes in various precipitation regimes: orographic precipitation in the western U.S. and Europe (e.g., MAP, Bougeault et al. 2001; IPEX, Schultz et al. 2002; IMPROVE-2, Stoelinga et al. 2003; OLYMPEX, Houze et al. 2017), midlatitude cyclones (e.g., PLOWS, Rauber et al. 2014; GCPEX, Skofronick-Jackson et al. 2015), and lake-effect snow (LOWS, Reinking et al. 1993; OWLeS, Kristovich et al. 2017). Studies comparing high-resolution model output with observations from these field campaigns have demonstrated key areas for improvement, driving continuous refinements to existing MP schemes and development of

new schemes.

For example, orographic precipitation research from IMPROVE-2 case studies over the Cascade Mountains revealed numerous MP sensitivities and biases. Garvert et al. (2005) found that some schemes (e.g., Reisner-2, Reisner et al. 1998) produced too much leeward precipitation due to an overprediction of snow aloft being advected leeward, with an underprediction of cloud liquid water (CLW) near the mountain crest and a windward overprediction of CLW at low levels. Colle et al. (2005) performed a microphysical budget showing that water vapor loss was primarily associated with condensation and snow deposition. In another Cascades case study, Lin and Colle (2009) demonstrated that several MP schemes [e.g., Thompson (Thompson et al. 2008) and WSM6 (Hong and Lim 2006)] similarly produced too much snow aloft at the expense of CLW, while the older version of Thompson (Thompson et al. 2004) was more representative of aircraft observations of CLW. Lin and Colle (2009) did several additional sensitivity experiments to modify graupel collection equations in the snow-biased schemes, but found that reducing the snow bias aloft led to excessive windward graupel production in their study. For a sample of midlatitude cyclones affecting the Northeastern U.S., Molthan et al. (2016) demonstrated MP scheme sensitivity in regions of the storm with supercooled liquid water aloft, with MP schemes performing better in periods of moderate riming rather than light riming of snow observed at the surface. In the present study, we did not investigate whether snow was overpredicted aloft at the expense of CLW (although several schemes had deficits in CLWP, including Thompson and WSM6), which could be an area for future investigation of MP scheme sensitivity in this LeS case with occasional riming.

As discussed above, the MP sensitivity case study of LeS downwind of Lake Erie by Reeves and Dawson (2013) had generally similar results with respect to schemes that produce more snow (e.g. Thompson) and more graupel (e.g., WSM6), with similar microphysical environments between the Lake Erie study and OWLeS IOP2b likely causing similar MP sensitivity to graupel production. However, their study had more differences in the location

and amount of maximum QPF relative to the lakeshore between subsets of snow-dominant and graupel-dominant members, perhaps related to differences in LeS band orientation and interaction with terrain between their case and IOP2b. The results from this case study of IOP2b are also similar to those from McMillen and Steenburgh (2015) who did multi-MP ensemble simulations for an October 2010 Great Salt Lake case study in Utah, and found Thompson MP had the most accurate QPF results as compared to manual and radar-estimated precipitation observations.

Overall, environments with light to moderate riming seem to be a NWP challenge for many MP schemes, for both multiple-ice category (e.g., Thompson) and single-ice category with fractional riming (e.g., P3) methods, in both IOP2b and other types of winter storms. This represents a microphysical continuum in need of further research for future parameterization improvements.

For example, one possible improvement pathway is through the development of stochastic physics, to better describe physical parameterization uncertainty within individual parameterization schemes. Stochastic physics operates by adding a pattern of random noise, which varies in amplitude, space, and time, to uncertain parameters or tendencies within existing parameterizations, and a different noise pattern can be applied to each member in a single-physics ensemble. While previous convection-permitting model studies of stochastic physics have mostly focused on warm-season convection (e.g., Berner et al. 2017; Jankov et al. 2017, 2019), very little research has been done on stochastic physics forecasts of cool-season, mixed-precipitation events.

Table 4.1: Summary of WRF model configurations for Thompson MP single-physics sensitivity experiments, with coefficients corresponding to those in snow PSD, mass-diameter, and velocity-diameter formulas given by Equations 4.3, 4.4, and 4.5, respectively.

Experiment	Snow PSD coeffs.	Mass-diameter coeffs.	Fall velocity-diameter coeffs.
CTRL	$\Lambda_1 = 3.29$	$a_{m_s} = 0.069,\ b_{m_s} = 2.0$	$a_{v_s} = 40.0,\ b_{v_s} = 0.55,\ f_{v_s} = 100.0$
THOM_PSD	$\Lambda_1 = 1.6$	$a_{m_s} = 0.069,\ b_{m_s} = 2.0$	$a_{v_s} = 40.0,\ b_{v_s} = 0.55,\ f_{v_s} = 100.0$
THOM_FV_NOEXP	$\Lambda_1 = 3.29$	$a_{m_s} = 0.069,\ b_{m_s} = 2.0$	$a_{v_s} = 40.0,\ b_{v_s} = 0.55,\ f_{v_s} = 0.0$
THOM_MFV_DDRT	$\Lambda_1 = 3.29$	$a_{m_s} = 0.052,\ b_{m_s} = 2.19$	$a_{v_s} = 1.683,\ b_{v_s} = 0.22,\ f_{v_s} = 0.0$
THOM_MFV_COLM	$\Lambda_1 = 3.29$	$a_{m_s} = 45.0,\ b_{m_s} = 3.0$	$a_{v_s} = 73.528,\ b_{v_s} = 0.65,\ f_{v_s} = 0.0$

Table 4.2: Summary of SC and NR precipitation amounts and biases (relative to OWLeS manual observations) for Thompson MP single-physics sensitivity experiments. All precipitation amounts are for the 24-h period ending at 0000 UTC 12 December 2013.

Experiment	SC precip.	SC bias	Pct. of obs.	NR precip.	NR bias	Pct. of obs.
Observed	34.0 mm			64.5 mm		
KTYX-estimated	33.1 mm	-0.9 mm	97.4%	51.3 mm	-13.2 mm	79.5%
CTRL	17.0 mm	-17.0 mm	50.0%	31.9 mm	-32.6 mm	49.5%
THOM_PSD	20.9 mm	-13.1 mm	61.5%	40.7 mm	-23.8 mm	63.1%
THOM_FV_NOEXP	19.1 mm	-14.9 mm	56.2%	36.6 mm	-27.9 mm	56.7%
THOM_MFV_DDRT	10.4 mm	-23.6 mm	30.5%	22.7 mm	-41.8 mm	35.2%
THOM_MFV_COLM	28.3 mm	-5.7 mm	83.2%	54.5 mm	-10.0 mm	84.5%

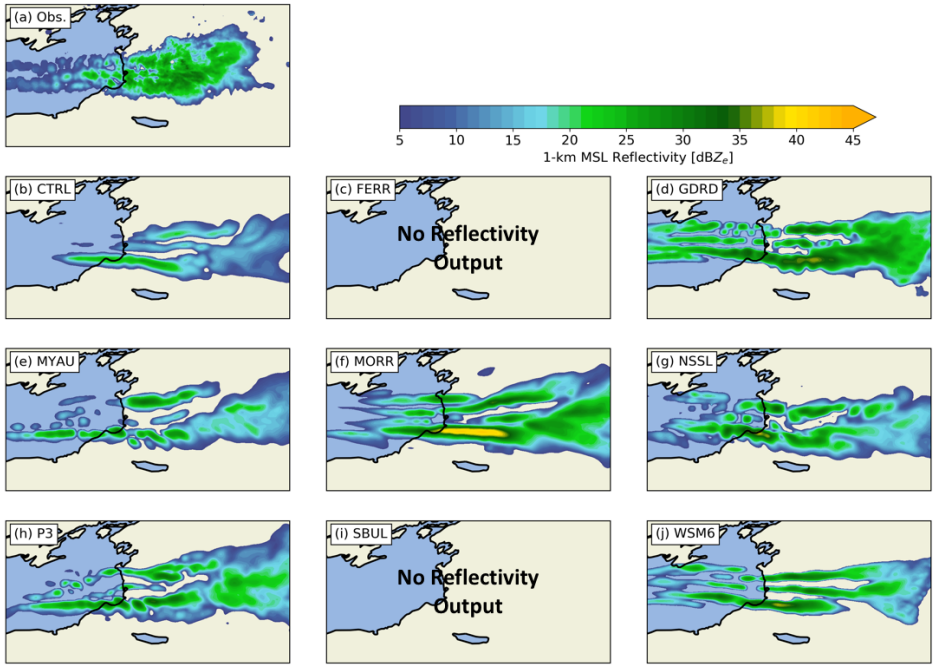

Figure 4.1: Observed and WRF-simulated equivalent radar reflectivity factor, dBZ_e, at 0900 UTC 11 December 2013. (a) KTYX NEXRAD radar-observed dBZ_e, as in Fig. 3.2c, (b)-(j) WRF-simulated dBZ_e at 1-km MSL from each of the microphysics (MP) ensemble members.

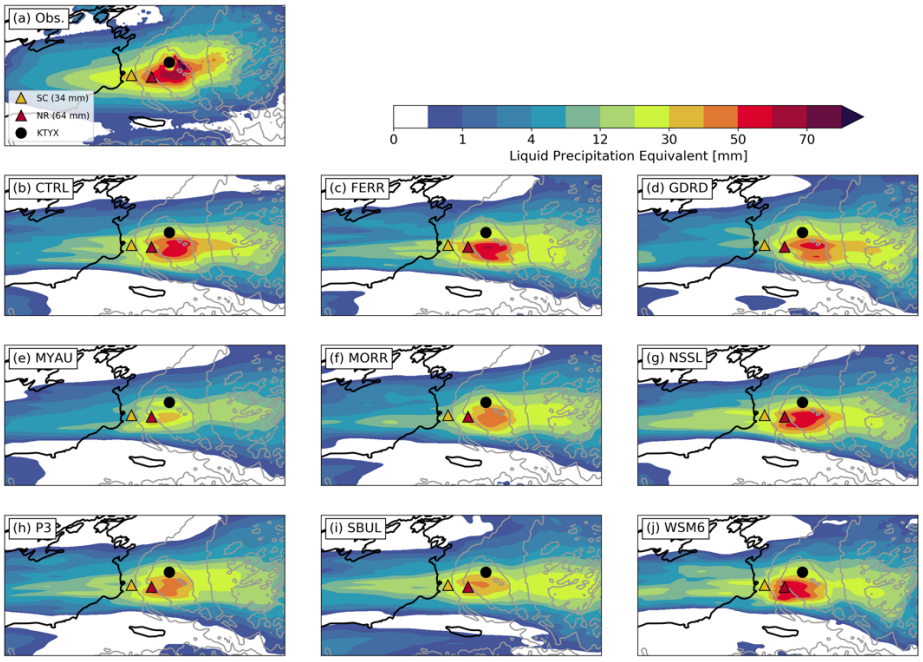

Figure 4.2: Observed and WRF-simulated 24-h accumulated precipitation ending at 0000 UTC 12 December 2013. (a) Precipitation measured by OWLeS manual observations (filled triangles) and estimated by KTYX radar (shading). (b)-(j) WRF quantitative precipitation forecast (QPF) over the same 24-h period for each MP ensemble member, with OWLeS manual observation amounts denoted by triangles.

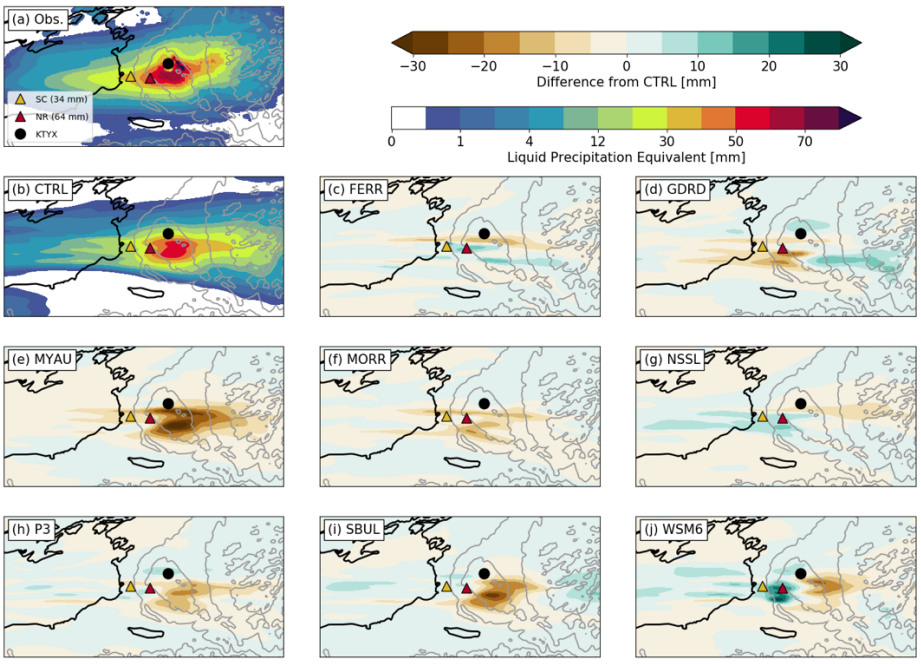

Figure 4.3: (a)-(b) Same as Fig. 4.2. (c)-(j) As in Fig. 4.2, but for accumulated precipitation difference relative to CTRL member (b).

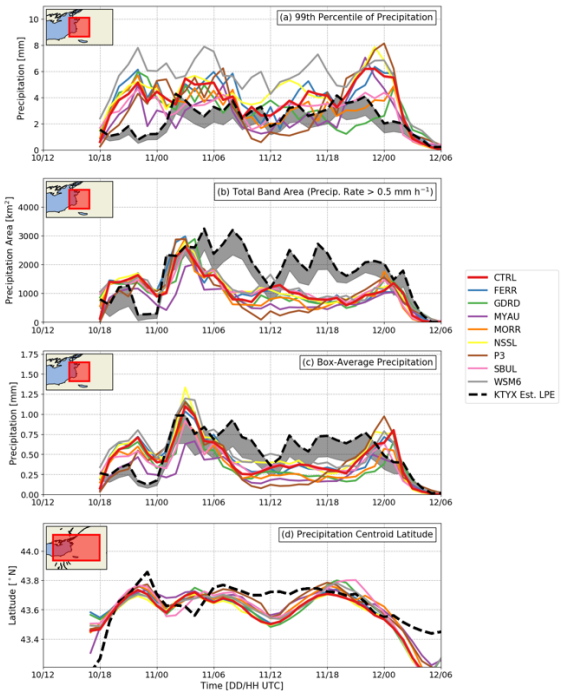

Figure 4.4: Time series comparison of model- (solid lines) and KTYX radar-derived (dashed line) LeS band properties: (a) 99th percentile of hourly precipitation, (b) total area of hourly precipitation greater than 0.5 mm h^{-1}, (c) box-average precipitation, and (d) latitude of hourly precipitation mass centroid. All statistics computed over the inset boxes shown in each panel. Gray shading in (a)-(c) shows a range of possible radar-derived hourly values using three different reflectivity-precipitation (Z-R) relations.

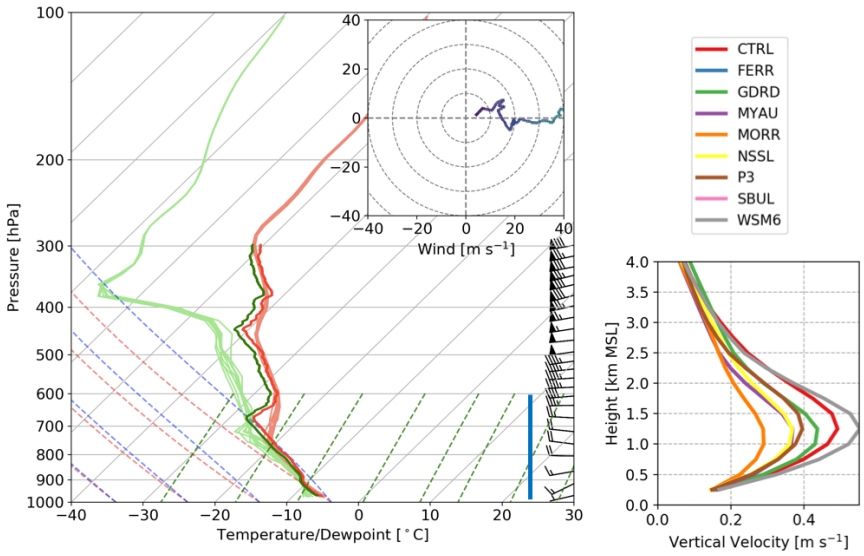

Figure 4.5: (a) Skew T-logp diagram at North Redfield (NR) within the LeS band, with observed [temperature (red), dewpoint (green), wind barbs (full and half barbs denote 10 and 5 m s^{-1}, respectively), and hodograph] and MP ensemble member [temperature (light red) and dewpoint (light green)] data plotted. The observed sounding was launched at 1727 UTC 11 December and modeled soundings are from 1700 UTC 11 December 2013. WRF soundings are computed from the nearest model gridpoint to NR. (b) MP ensemble member 24-h average vertical velocity profiles from Sandy Creek during 11 December, calculated from the nearest model gridpoint. A vertical blue line in (a) denotes the y-axis range in (b).

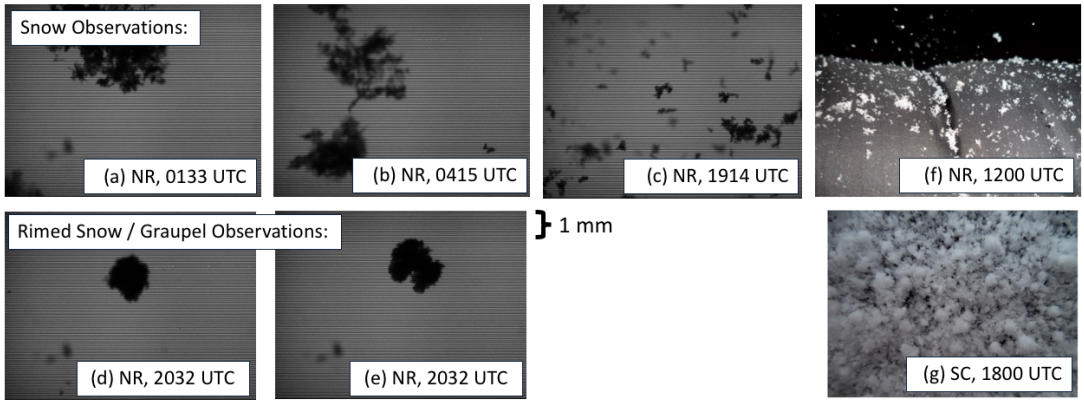

Figure 4.6: Examples of observed crystal habit imagery during 11 December 2013, with location abbreviations and times noted for each image. (a)-(e) Images captured by the HYVIS automated camera located at North Redfield (NR). (f)-(g) Pictures taken by research scientists using handheld cameras. HYVIS images all have the same scale as noted.

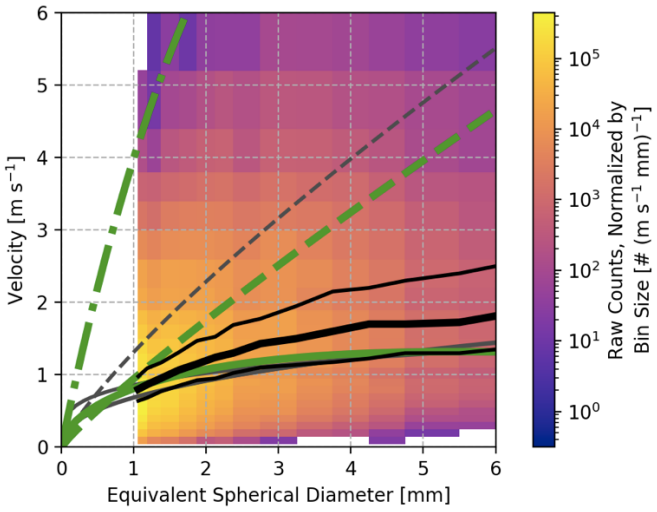

Figure 4.7: Observed 2-D histogram of all Parsivel hydrometeor counts, normalized by bin size, in velocity-diameter space during the 24-h period ending at 0000 UTC 12 December 2013 at Sandy Creek. The thick black line represents the 50th percentile of the observed velocities, with thin black lines denoting the 25th and 75th percentiles. Solid gray lines indicate velocity-diameter relations for snow from all of the MP schemes used in this study that categorically predict snow, with Thompson MP highlighted in green. Dashed and alternating dash-dot green lines represent the Thompson MP velocity-diameter relations for graupel and rain, respectively, while the gray dashed line represents the WSM6 graupel relation.

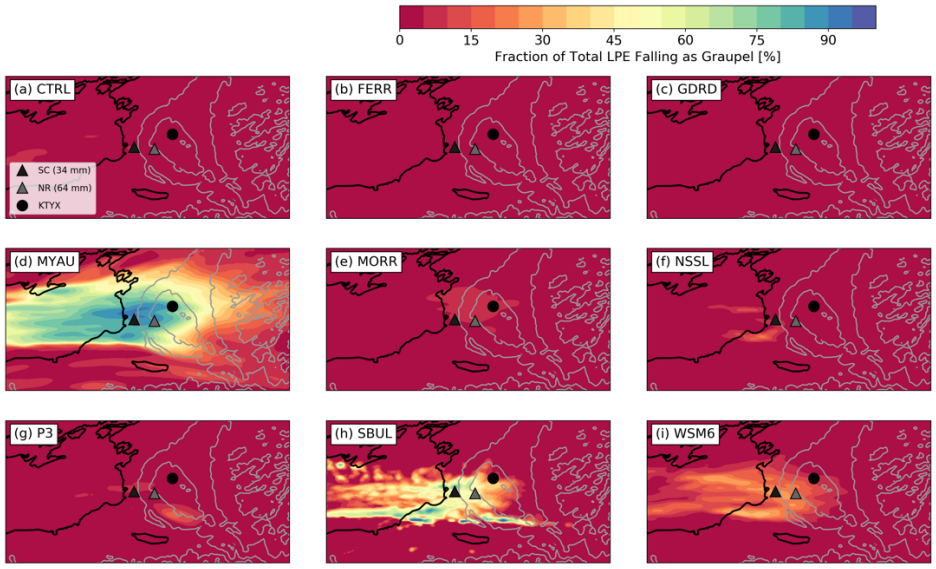

Figure 4.8: Fraction of 24-h accumulated precipitation falling as graupel from each of the MP ensemble members (a)-(i), ending at 0000 UTC 12 December 2013.

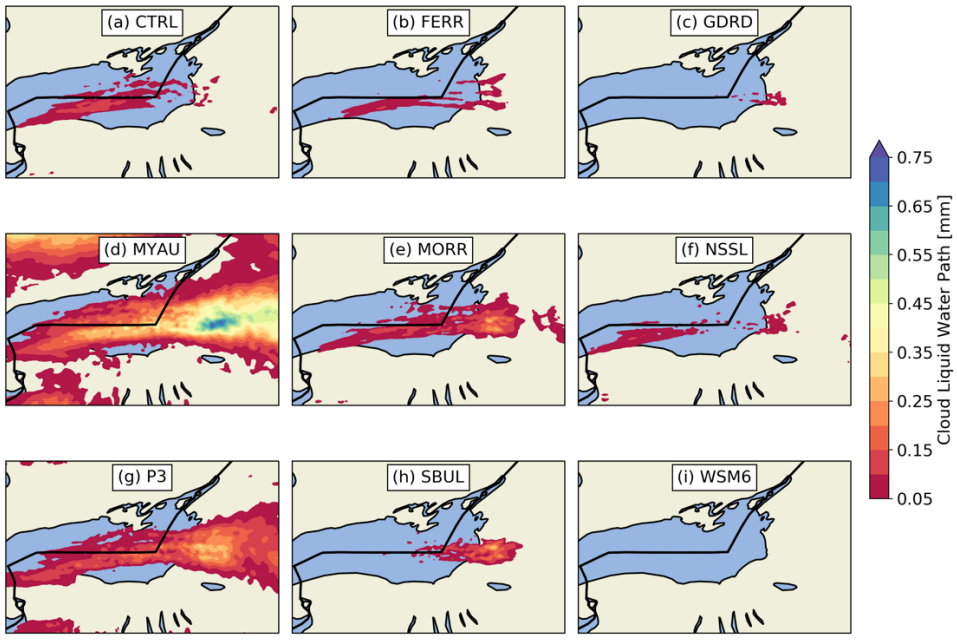

Figure 4.9: Model-simulated cloud liquid water path averaged over the 24-h period ending at 0000 UTC 12 December 2013 from each of the MP ensemble members.

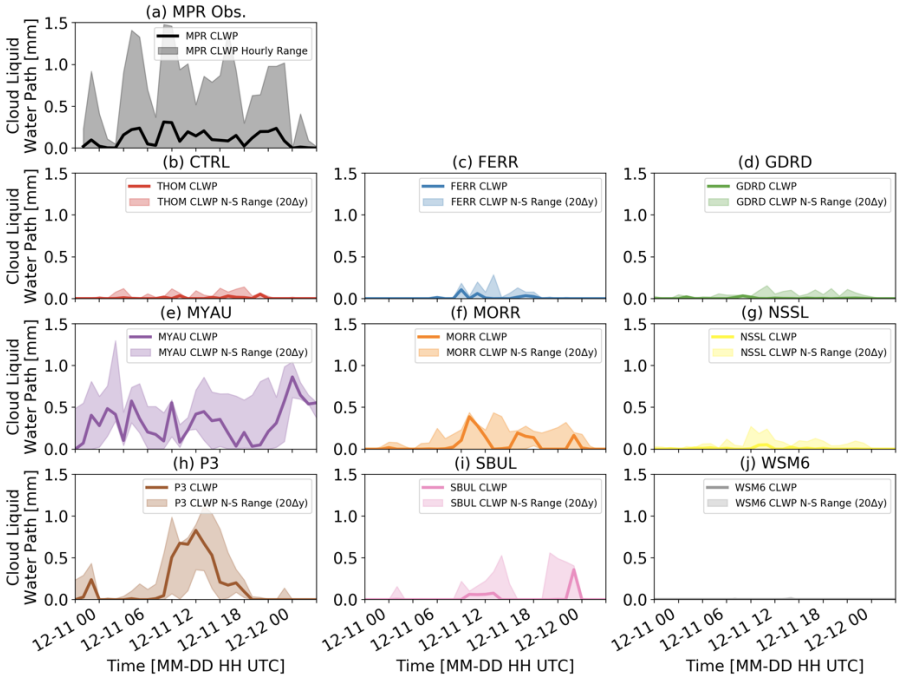

Figure 4.10: Time series comparison of observed and WRF-simulated cloud liquid water path (CLWP) at Sandy Creek (SC). (a) Observed hourly average CLWP from the MPR (solid line), shading shows full hourly range of 5-min measurements, (b)-(j) Model-simulated instantaneous hourly CLWP from each microphysics ensemble member, calculated at the nearest gridpoint to SC (solid lines), shading shows 20-gridpoint (26.6 km) north-south range of CLWP values centered on SC.

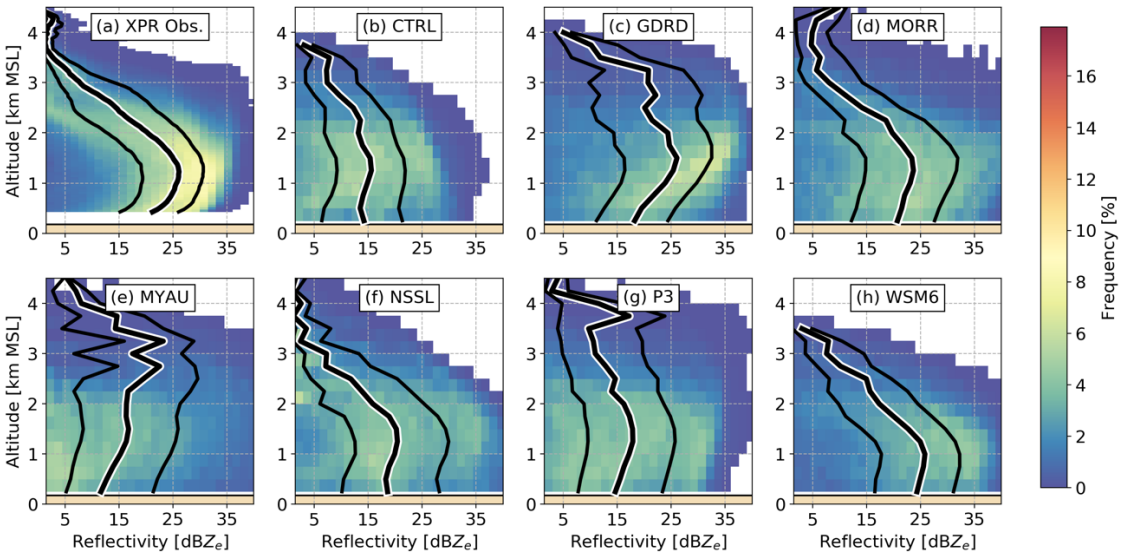

Figure 4.11: Observed and WRF-simulated equivalent radar reflectivity contoured frequency by altitude diagrams (CFADs) at Sandy Creek, for the 24-h period ending at 0000 UTC 12 December 2013: (a) Observed CFAD from the MIPS XPR, (b)-(h) Modeled CFADs from each of the MP ensemble members. Black lines are overlaid to indicate the reflectivity quantiles from the XPR and model simulations, with thick lines denoting the 50th percentile and thin lines denoting the 25th and 75th percentiles. Model CFADs sample all gridpoints within a 30-km $\times$ 30-km box centered on SC, and model reflectivity profile values are interpolated to a constant height grid with 250-m vertical grid spacing for the CFAD analysis.

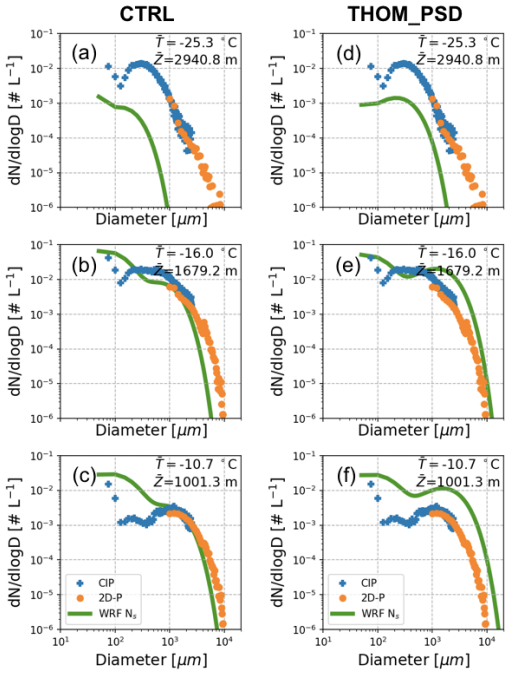

Figure 4.12: (a)-(c) Observed and WRF CTRL forecast snow particle size distributions (PSDs) within the LeS band core along three vertically-stacked University of Wyoming King Air (UWKA) flight legs between 1917 and 2104 UTC 11 December 2013. UWKA observations are sampled from CIP and 2D-P probes (blue circles and orange pluses), with the mean altitude and temperature of the flight leg shown in each subplot. Model-derived PSDs from 2000 UTC (solid lines) are interpolated to the horizontal location and mean altitude of each flight leg. (d)-(f) As in (a)-(c), but model PSDs are from the THOM_PSD experiment.

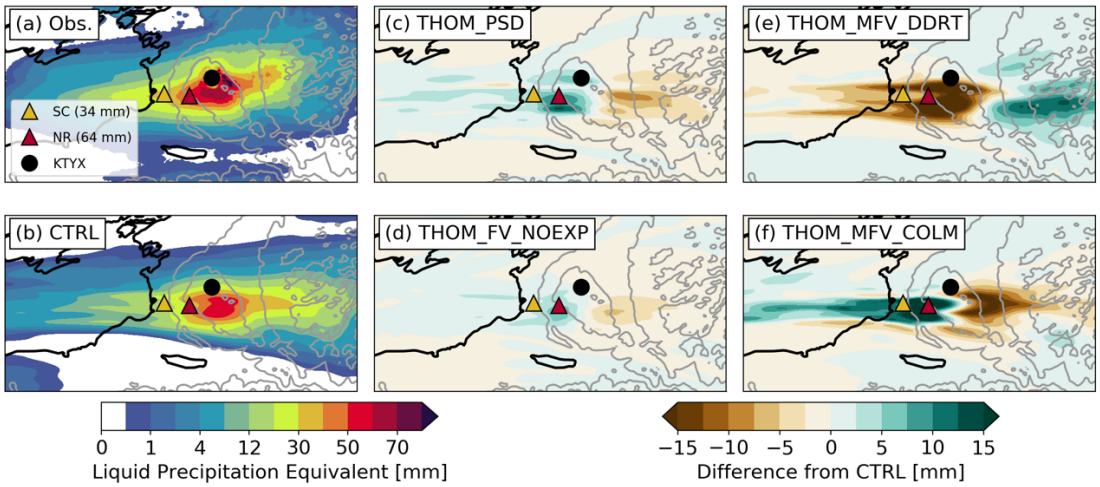

Figure 4.13: Observed and WRF-simulated 24-h accumulated precipitation ending at 0000 UTC 12 December 2013. (a)-(b) Same as Fig. 4.2. (c)-(f) Accumulated precipitation difference relative to CTRL member, for single-scheme sensitivity experiments with Thompson MP.

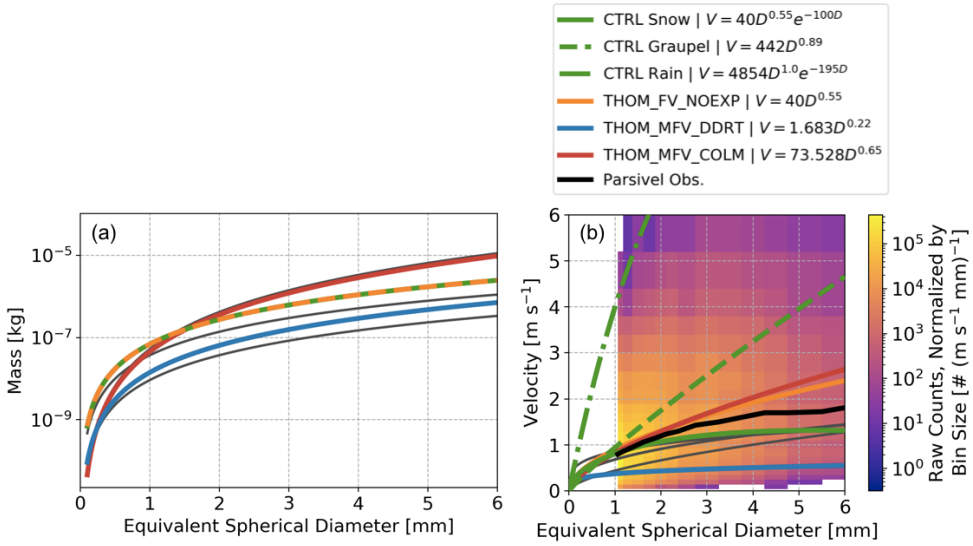

Figure 4.14: Comparison of differences in crystal habit experiments described in Table 4.1, for (a) mass-diameter and (b) fall velocity-diameter relations. Dashed lines in (a) consisting of two alternating colors indicate where two experiments share the same relation. Parsivel data shading in (b) is the same as in Fig. 4.7, but with more velocity-diameter relations plotted. Dashed and alternating dash-dot green lines in (b) represent Thompson MP velocity-diameter relations for graupel and rain, respectively, but the corresponding mass-diameter relations are not plotted in (a) for clarity. Additional gray lines in both (a) and (b) indicate other habit relations from observational data (Woods et al. 2007).

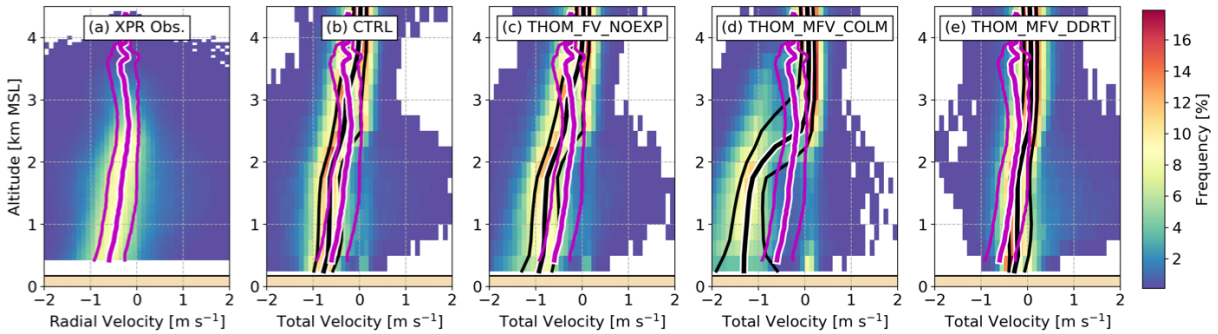

Figure 4.15: CFADs of observed radial velocity and WRF-simulated model total velocity (vertical velocity plus precipitation terminal fall velocity) at Sandy Creek for (a) XPR, (b) CTRL, (c) THOM_FV_NOEXP, (d) THOM_MFV_COLM, and (e) THOM_MFV_DDRT. Solid lines are overlaid to indicate the vertical velocity quantiles from the XPR and model simulations (magenta and black lines, respectively), with thick lines denoting the 50th percentile and thin lines denoting the 25th and 75th percentiles.

5. Sensitivity to planetary boundary layer and surface layer parameterizations

Results presented in this chapter are included in a manuscript under review[1].

5.1 Evaluation against precipitation and radar observations

Similar to the MP ensemble, the PBL/SL ensemble also generally performs well at forecasting realistic LeS band structures and position over Tug Hill during IOP2b. Figure 5.1 shows a comparison of observed KTYX reflectivity with reflectivity output from each of the PBL/SL members at 0900 UTC 11 December (the same time shown in Fig. 4.1). At 0900 UTC, the observations show LeS band structure with some indication of multiple parallel bands or convective cells aggregated together, which is reflected in most of the PBL/SL ensemble members. Several members (e.g., CTRL, ACM_MM5, UW_MM5) depict LeS bands that are more linear while other members (e.g., MYJ, YSU_MM5, TEMF) depict more cellular structures. However, LeS morphologies vary with time during the event, with both observations and models transitioning between linear and cellular structures.

As evident in Figure 5.1, the primary differences between members are the morphology and especially the intensity of the LeS bands. For example, QNSE, MYJ, and TEMF have 1-km MSL reflectivity values exceeding 35 dBZ_e in the most intense LeS cells (Fig. 5.1), with these differences in LeS band intensity between members persisting throughout IOP2b (not shown). Therefore, in this case, the choice of PBL/SL scheme appears to affect both the morphology and intensity of LeS, while the choice of MP scheme primarily affects LeS intensity and precipitation type. Note that the north-south positions of the LeS bands in all PBL/SL members are consistent due to the WRF simulations using RAP analysis BCs and nudging on the outermost domain, as in the MP ensemble.

Due to the minimal position differences but large range of forecast LeS band intensi-

[1] Minder, J. R., and W. M. Bartolini, C. Spence, N. R. Hedstrom, P. D. Blanken, and J. D. Lenters, 2019 (in revision): Characterizing and constraining uncertainty in simulations of lake-effect snowfall associated with surface and boundary layer turbulence. *Weather and Forecasting*.

ties from each PBL/SL ensemble member, event-total QPF amounts also vary between the members. Figure 5.2 shows the KTYX-estimated LPE and PBL/SL ensemble member QPF, highlighting a range of maximum forecast precipitation amounts over Tug Hill between 54 and 99 mm. This range is larger than that for the multi-MP ensemble (35–62 mm) and the PBL/SL distribution is shifted towards more precipitation over and especially upwind of Tug Hill. Overall, CTRL and most of the other PBL/SL members underforecast maximum QPF amounts by 10–20 mm relative to the manual observations at SC and NR and the radar-estimated LPE (Fig. 5.2). The only exceptions are MYJ and especially QNSE, which overforecast QPF, particularly on the windward side of Tug Hill. Consistent with the more-intense LeS bands forecast by QNSE and MYJ (Fig. 5.1c-d), these members also have the highest QPF amounts (Fig. 5.2c-d) compared to the other members.

Difference plots of each PBL/SL ensemble member's QPF relative to the CTRL forecast also highlight the spread in total QPF amounts between members (Fig. 5.3). The higher QPF amounts in QNSE and MYJ are accentuated (Fig. 5.3c-d) especially on the windward side of Tug Hill, while most other members have similar locations and magnitudes of the maximum QPF amounts relative to CTRL. However, TEMF has a slight event-average southward shift in the axis of maximum QPF (Fig. 5.3f) while BOLA_MM5 has a slight northward QPF shift (Fig. 5.3g) relative to CTRL.

5.1.1 Verification of LeS band properties

As presented for the MP ensemble, time series of LeS band statistics for the PBL/SL ensemble include the intensity (99th percentile of precipitation), total band area, box-average precipitation, and mass centroid latitude, shown in Figure 5.4. Similar to the results from the MP ensemble statistics, intensities vary in time and between PBL/SL ensemble members (Fig. 5.4a). Overall, most members have generally accurate intensities throughout the event (but with a modest positive bias), peaking in intensity at roughly the correct times between 0000 and 0600 UTC on 11 December and between 1800 UTC 11 December and 0000 UTC

12 December. However, QNSE and MYJ are biased high throughout IOP2b, consistent with their excessive QPF totals (Fig. 5.2).

In comparison, total band area has less spread between members (Fig. 5.4b). All the PBL/SL ensemble members except for QNSE have similar total LeS band area to KTYX observations early (around 0300 UTC 11 December) and late (around 0000 UTC 12 December) in the event, but are all biased toward LeS bands that are 1000–1500 km^2 too small during the middle of the event between 0600 and 2100 UTC 11 December. However, QNSE has total band area exceeding 1500 km^2 throughout the event, more consistent with observations. Other than QNSE, these results share the same small-band bias found in the MP ensemble (Fig. 4.4), suggesting that most PBL/SL and MP schemes are not able to correct the small area bias during the middle period of IOP2b.

As in the intensity and total band area statistics, box-average precipitation also reveals differences between ensemble members and biases relative to KTYX observations (Fig. 5.4c). Similar to intensity, box-averaged precipitation in QNSE (and MYJ at times) is biased high due in part to its too-intense LeS bands. QNSE is particularly high-biased relative to the observed box-average precipitation, in part because it has the largest total band area (relatively accurate) of the PBL/SL members during much the event. MYJ has minimal bias in box-average precipitation during most of the event, except between 0300 and 0600 UTC 11 December and between 2200 UTC 11 December and 0200 UTC 12 December, but this appears to be due to high-intensity and small-area biases cancelling each other. Similar to MYJ, most other PBL/SL schemes (except for QNSE) have compensating biases in intensity and area that act to reduce biases in box-average precipitation. Therefore, schemes like MYNN (CTRL) appear to perform better in this case since they have less overall bias across all three metrics relative to observations.

Finally, north-south LeS position variability is quantified by the mass centroid latitude computed for each PBL/SL ensemble member (Fig. 5.4d). Similar to the results from the MP ensemble (Fig. 4.4d), most PBL/SL members are tightly clustered in position throughout

IOP2b. All members except BOLA_MM5 have minimal position biases during most of the event, except the period between 0900 UTC and 1500 UTC 11 December, when the model forecasts are shifted too far south. During this time, the observations have somewhat of a parallel-band morphology over the lake, collapsing into a single diffuse band over Tug Hill (Fig. 3.2c-d). Most of the model forecasts maintain a double-band structure, with all members except BOLA_MM5 keeping a dominant southern band while BOLA_MM5 has a dominant northern band (not shown), with the latter having a better north-south position relative to observations (Fig. 5.4d).

5.2 Comparison of sensible and latent heat fluxes

Overall, differences between PBL/SL members in LeS band morphology, QPF, and hourly statistics are primarily driven by uncertainty in the amount of parameterized heat and moisture fluxes off of Lake Ontario. 24-h average surface sensible and latent heat fluxes from each of the PBL/SL members are plotted in Figures 5.5 and 5.6, respectively. In general, higher-flux members usually have higher forecast total precipitation amounts. For example, QNSE, the PBL/SL member with the most intense LeS band forecast reflectivity values (Fig. 5.1c) and highest QPF (Fig. 5.2c), has average sensible (Fig. 5.5b) and latent (Fig. 5.6b) heat fluxes that are 400 W m^{-2} higher over southern Lake Ontario than the other members. MYJ is the other primary outlier with higher fluxes and QPF, although less than QNSE, relative to the rest of the PBL/SL ensemble. Physically, an increase in surface fluxes leads to more over-lake thermodynamic instability and moisture, causing more-intense LeS. Unfortunately, no flux observations over the lake were gathered during OWLeS IOP2b, so no direct comparisons of modeled and observed fluxes can be made for evaluation in this case, only secondary comparisons of modeled and observed precipitation. Nonetheless, QNSE and MYJ are biased with too much QPF relative to both the radar-estimated and manual observations of LPE, whereas many of the other PBL/SL members have too little QPF.

5.3 QNSE sensitivity experiments

Motivated by the results from Conrick et al. (2015), several sensitivity experiments are performed to modify the SL scheme in QNSE to investigate the differences in fluxes and QPF in the outlier PBL/SL members (QNSE and MYJ). These SL experiments, summarized in Table 5.1, involve two parameters, the first being the Prandtl number (P_R) at neutral stability where P_R is defined as the ratio of the eddy diffusivities of momentum and heat. The second parameter is the critical threshold of friction velocity (u_*) used to define transitions between three different roughness regimes in the QNSE SL scheme (Janjić 1994).

Since many other SL schemes use a value for P_R near 1, we also perform an experiment setting $P_R = 1$, QNSE_PR1, as was done in Conrick et al. (2015). A second experiment, QNSE_PR1_USTC1, sets $P_R = 1$ while also increasing the critical u_* threshold value to 1.0 m s^{-1} needed for the scheme to transition to a "rough with spray" turbulence regime. During times when this "rough with spray" regime is reached in QNSE throughout IOP2b, fluxes abruptly increase by several hundred W m^{-2}. In particular, the spatial and temporal abruptness of the flux increases that accompany the transitions to the "rough with spray" regime appear to be unrealistic (not shown). Therefore, QNSE_PR1_USTC1 tests the forecast sensitivity by preventing the scheme from entering this regime.

The QNSE_PR1 experiment shows a reduction in lake-average, 24-h average surface sensible and latent heat fluxes, with a sensible heat flux reduction of 150 W m^{-2} (Fig. 5.7f) relative to QNSE (Fig. 5.7e). QNSE_PR1_USTC1 also shows a similar reduction in average sensible heat fluxes of 200 W m^{-2} relative to QNSE (Fig. 5.7g). In this case, changing P_R appears to have relatively more impact in reducing surface fluxes than changing the critical u_* for the "rough with spray" regime, since QNSE_PR1 shows approximately 75% of the reduction in surface fluxes as in QNSE_PR1_USTC1. The combined QNSE_PR1_USTC1 experiment reduces average sensible heat fluxes to around 280 W m^{-2}, comparable to CTRL and most of the other PBL/SL ensemble members. Along with the decrease in surface fluxes, the QNSE experiments have a corresponding reduction in QPF, shown in Figure

5.7b-d. For example, the maximum QPF is reduced from 99 mm in QNSE to 61 mm in QNSE_PR1_USTC1, with the latter being more consistent with the manual and radar-estimated LPE observations.

While the QNSE scheme is not currently used in NWS operational models, MYJ is used in operations (NAM Nest and HRW NMMB, see Table 2.3). MYJ uses a similar roughness formulation as QNSE, with abrupt transitions between flux regimes, and also crosses into a "rough with spray" regime at times during IOP2b, causing very large fluxes. While MYJ by default uses a higher P_R than QNSE ($P_R = 1$), an MYJ sensitivity experiment increasing the critical u_* to 1.0 m s^{-1} also reduces fluxes and QPF (MYJ_USTC1, not shown), causing MYJ_USTC1 to perform more similarly to the rest of the PBL/SL members.

Table 5.1: Summary of WRF model configurations for the QNSE SL sensitivity experiments. Aerosol-aware Thompson MP (Thompson et al. 2008; Thompson and Eidhammer 2014) is used in all members.

Experiment	PBL scheme	SL scheme	P_R	Roughness
QNSE	Quasi-Normal-Scale-Elimination[1]	QNSE	0.72	critical $u_* = 0.7$
QNSE_PR1	Quasi-Normal-Scale-Elimination[1]	QNSE	1	critical $u_* = 0.7$
QNSE_PR1_USTC1	Quasi-Normal-Scale-Elimination[1]	QNSE	1	critical $u_* = 1$

[1] Mellor and Yamada (1982); Janjic (2002); Sukoriansky et al. (2005)

5.4 Discussion

Overall, the PBL/SL results presented in this study are generally consistent with past research on other intense Great Lakes LeS cases (Conrick et al. 2015; Fujisaki-Manome et al. 2017). In comparison with the December 2009 Lake Erie case studied by Conrick et al. (2015), relative differences in surface fluxes and total QPF are very similar between CTRL (MYNN), MYJ, and QNSE, with the latter having much larger fluxes and precipitation. Conrick et al. (2015) also found that other schemes (e.g., ACM, BOLA, YSU) had intermediate fluxes and QPF, similar to MYJ but more than MYNN. However, in this study of OWLeS IOP2b, QNSE and MYJ are the primary outliers with higher fluxes and precipitation in our PBL/SL ensemble, while the remaining schemes are similar to CTRL (MYNN).

The sensitivity experiments to QNSE and MYJ performed in this study are also similar to those studied by Conrick et al. (2015). In particular, increasing P_R in QNSE reduces surface fluxes and QPF in this case, as in their study. However, Conrick et al. (2015) attributed the increased fluxes in MYJ to differences in the formulations of similarity stability functions between MYJ and MYNN, but we do not find that to be the case in this study. While the high-flux "rough with spray" regime transition was not examined in Conrick et al. (2015), the difference in performance of MYJ during this event might be due in part to MYJ entering the "rough with spray" regime more often than in the Conrick et al. (2015) case study.

Our results demonstrating the sensitivity of LeS to the choice of PBL/SL scheme in IOP2b are also generally consistent with the results from Fujisaki-Manome et al. (2017), who studied a different Lake Erie case from November 2014. They found a large range of forecast fluxes and QPF across several combinations of atmospheric and coupled lake models of different resolutions. Similar to the variety of model forecasts examined by Fujisaki-Manome et al. (2017), the spatial and temporal variability of surface sensible and latent heat fluxes is generally consistent across the PBL/SL ensemble for IOP2b (except for QNSE and MYJ during the "rough with spray" regime), with flux magnitudes being the primary difference between members, affecting QPF. While differences in lake-atmosphere thermodynamics, flow orientation, and lakeshore geometries vary between the three cases (each have complex roles contributing to forecast uncertainty), all of the above results demonstrate the importance of accurate PBL/SL parameterizations, along with a need for accurate NWP lake-surface temperature and ice cover analyses which are also vital to forecasting the correct amount of surface fluxes.

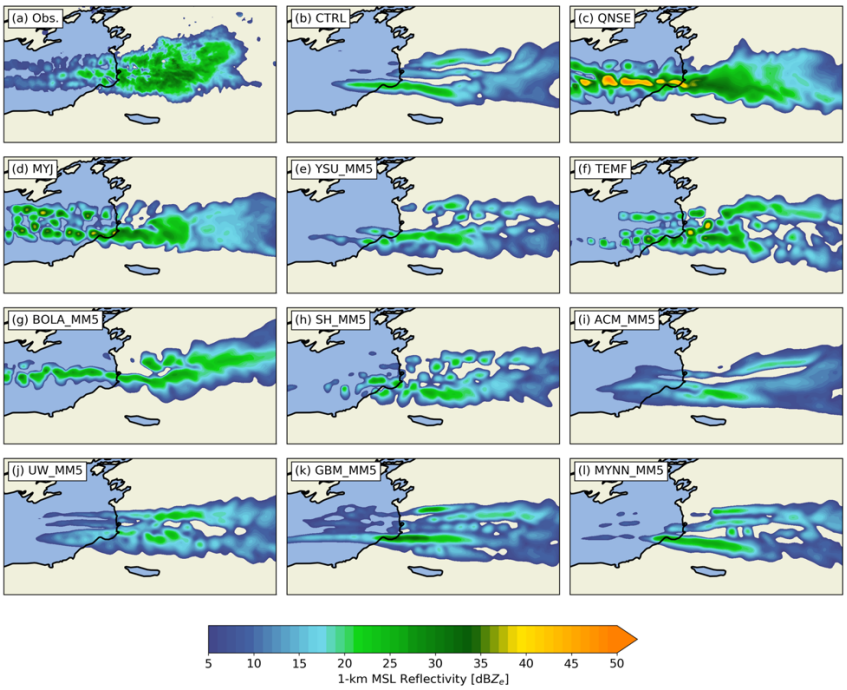

Figure 5.1: Comparison of observed and WRF-simulated equivalent radar reflectivity factor, dBZ_e, at 0900 UTC 11 December 2013. (a) KTYX NEXRAD radar-observed dBZ_e, as in Fig. 3.2c, (b)-(l) WRF-simulated dBZ_e at 1-km MSL for each planetary boundary layer / surface layer (PBL/SL) ensemble member.

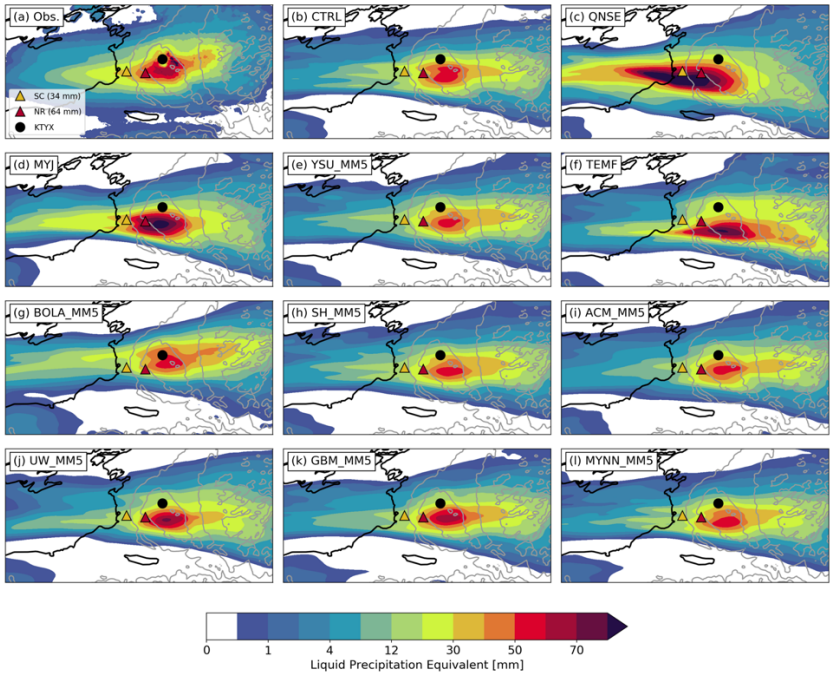

Figure 5.2: Comparison of observed and WRF-simulated 24-h accumulated precipitation ending at 0000 UTC 12 December 2013. (a) Precipitation measured by OWLeS manual observations (filled triangles) and estimated by KTYX radar (shading). (b)-(l) WRF quantitative precipitation forecast (QPF) over the same 24-h period for each PBL/SL ensemble member, with OWLeS manual observation amounts denoted by triangles.

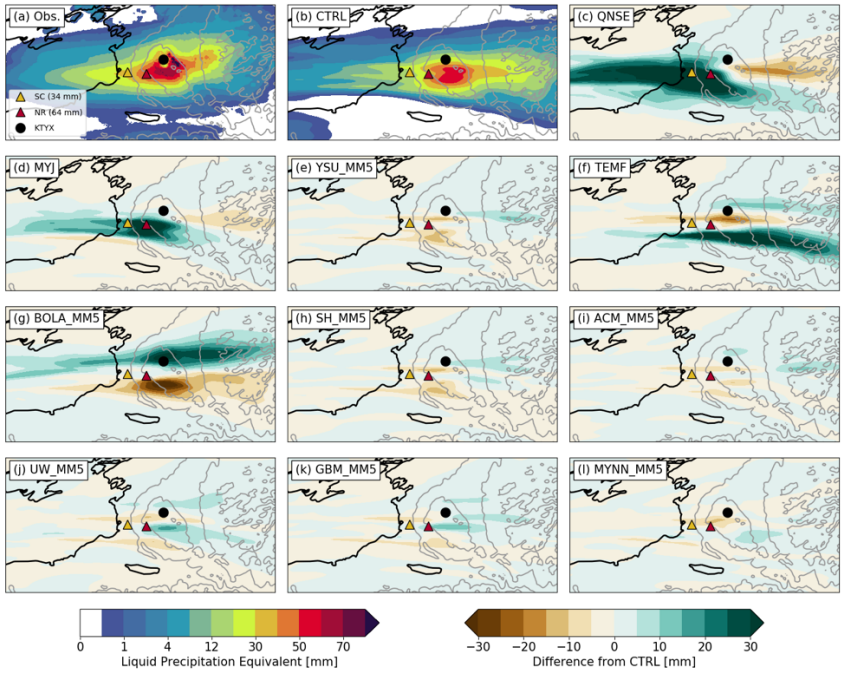

Figure 5.3: (a)-(b) Same as Fig. 5.2. (c)-(j) As in Fig. 5.2, but for accumulated precipitation difference relative to CTRL member (b).

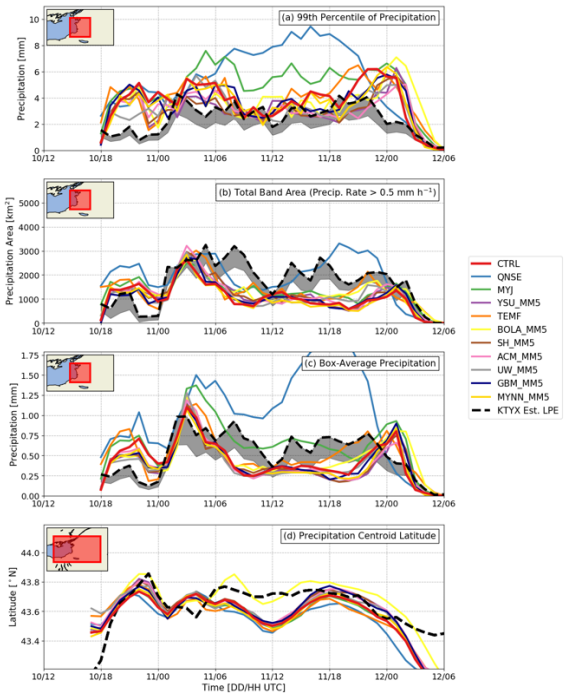

Figure 5.4: Time series comparison of model- (solid lines) and KTYX radar-derived (dashed line) LeS band properties: (a) 99th percentile of hourly precipitation, (b) total area of hourly precipitation greater than 0.5 mm h^{-1}, (c) box-average precipitation, and (d) latitude of hourly precipitation mass centroid. All statistics computed over the inset boxes shown in each panel.

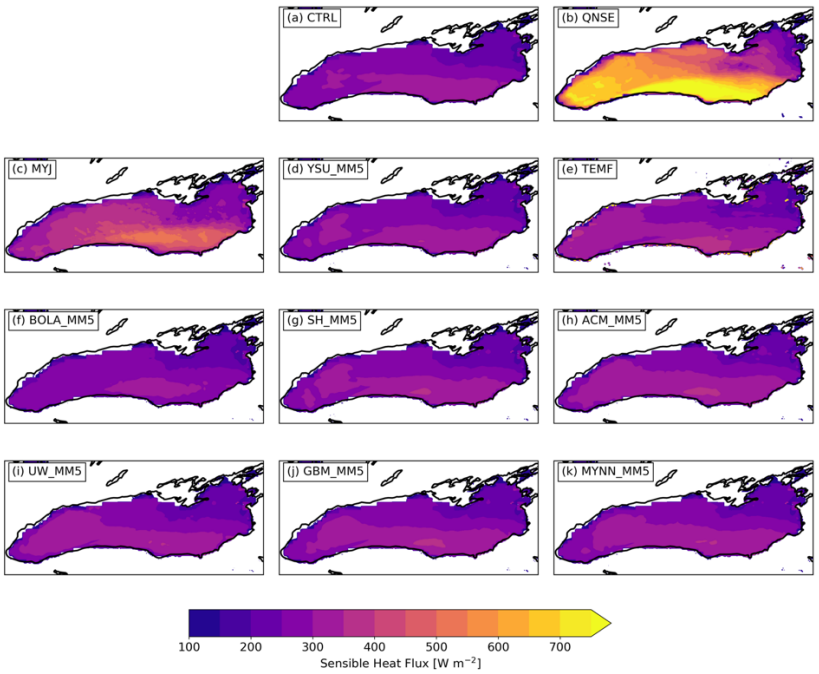

Figure 5.5: Model-simulated sensible heat flux averaged over the 24-h period ending at 0000 UTC 12 December 2013 from each of the PBL/SL ensemble members.

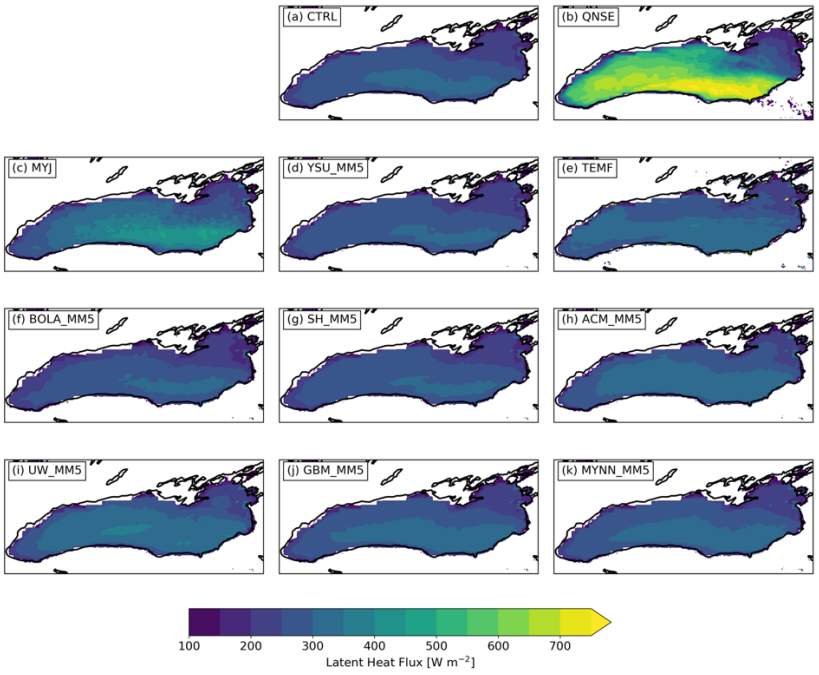

Figure 5.6: Model-simulated latent heat flux averaged over the 24-h period ending at 0000 UTC 12 December 2013 from each of the PBL/SL ensemble members.

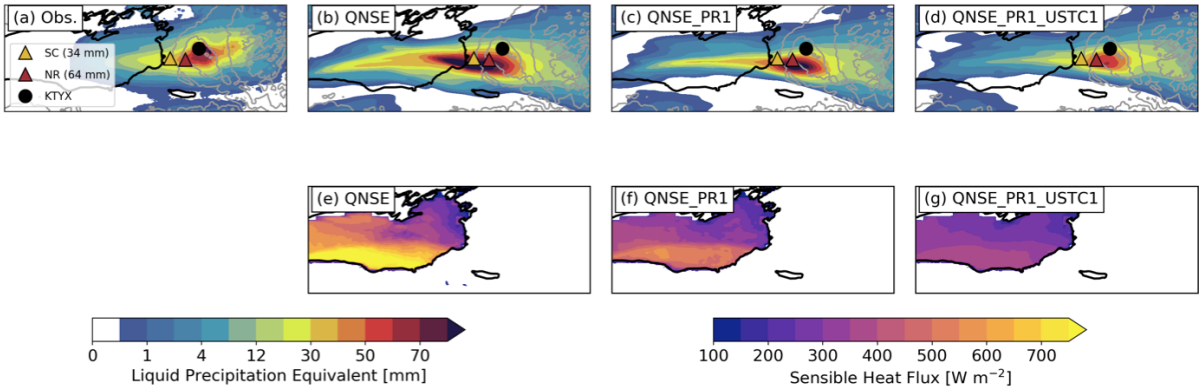

Figure 5.7: Observed and WRF-simulated 24-h accumulated precipitation and WRF-simulated 24-h average surface sensible heat flux, for the period ending at 0000 UTC 12 December 2013. (a) Precipitation measured by OWLeS manual observations (filled triangles) and estimated by KTYX radar (shading). (b)-(d) WRF quantitative precipitation forecast (QPF) over the same 24-h period for each QNSE sensitivity experiment, with OWLeS manual observation amounts denoted by triangles. (e)-(g) As in (b)-(d), but for 24-h average surface sensible heat flux from each of the QNSE experiments.

6. Summary and conclusions

Two multi-physics Weather Research and Forecasting (WRF) model ensembles were performed to assess the forecast sensitivity of lake-effect snow (LeS) during the Ontario Winter Lake-effect Systems (OWLeS) IOP2b event to the choice of microphysical (MP) and planetary boundary layer / surface layer (PBL/SL) parameterizations. These ensembles included a nine-member set of MP experiments and eleven-member set of PBL/SL experiments. The 10–12 December 2013 case was one of the most well-observed events during the two-month OWLeS field campaign, with an observed precipitation maximum of 64.5 mm measured at North Redfield (NR). This event was a classic westerly flow event impacting the region downwind of Lake Ontario including Tug Hill, with a combination of banded and non-banded lake-effect precipitation.

Simulation results from the MP ensemble revealed the primary differences between MP schemes were the LeS intensity and precipitation type, with relatively small impacts of the MP scheme choice on LeS band morphology and position, leading to a range in 24-h maximum quantitative precipitation forecast (QPF) amounts between 35 and 62 mm. However, all MP schemes suffered from some common biases, including LeS bands that were consistently 1000–1500 km^2 too small in areal extent during much of the event compared to radar observations from OWLeS.

Analysis of precipitation-type differences between MP members revealed that some members tended to produce mostly snow (e.g., Thompson), while other members produced almost as much graupel as snow (e.g., WSM6) over the windward slope of Tug Hill during the event. Profiling radiometer observations of modest amounts of cloud water aloft were generally in agreement with most MP scheme forecasts of cloud water; however, there was not a clear relationship between forecasts of high cloud water and high graupel in each of the MP members due to numerous parameterization differences between schemes. Comparisons

with ground-based observations of snow-to-liquid ratios (SLR), crystal habits, and Parsivel disdrometer fallspeed-diameter (V-D) relationships revealed that only a small fraction of the observed precipitation was likely graupel in this case. Therefore, the more graupel-dominant MP schemes were inconsistent with observations.

While it is unknown how the results presented from this case study would change in other LeS events downwind of Lake Ontario or in other regions, including different thermodynamic and dynamic environments, the range of precipitation types in each MP scheme are broadly similar to results from other case studies of LeS (e.g., Reeves and Dawson 2013; McMillen and Steenburgh 2015). Future studies could investigate MP scheme performance in other OWLeS events with different thermodynamic and kinematic environments affecting LeS band intensity, morphology, and precipitation type.

In this study, we also focused on understanding microphysical sensitivities within the CTRL member (Thompson MP), relying on additional observations to constrain the choice of parameter values used in sensitivity experiments. Additional diagnostics from Thompson MP were output to assess its performance during IOP2b in more detail, including offline computation of particle size distributions (PSDs) and online output of precipitation terminal fall velocities. Comparisons with University of Wyoming King Air (UWKA) in-situ data over eastern Lake Ontaro revealed biases in the CTRL PSDs with too few large particles (diameter $>$ 2–3 mm). Thus, a sensitivity experiment was performed changing the assumed PSD shape to be more representative of the large end of the PSD spectra. Due to the modification of the snow PSD, bulk snow fall velocities were increased, resulting in more precipitation relative to CTRL on the windward side of Tug Hill and a reduction in the underforecast precipitation bias there. This suggests that uncertainties in snow PSD shape are important in predicting the inland extent of the heaviest precipitation amounts, and that improved treatment of PSDs may lead to forecast improvements.

In addition to the PSD comparisons, vertical velocity and precipitation fall velocity data from X-band profiling radar (XPR) and Parsivel disdrometer instruments were

compared to the CTRL model output velocities. These comparisons suggested the CTRL precipitation fall velocities were relatively unbiased overall for the snow-dominant CTRL simulation, although a range of fall velocities occurred in reality due to temporal differences in riming and crystal habits. Several experiments were designed to test the sensitivity of Thompson MP to the assumed crystal habit by changing the mass-diameter and V-D relations prescribed for snow. While a dendrite-/aggregate-like experiment was more representative of dendritic aggregate habits observed at Sandy Creek (SC) and NR, it produced even less precipitation windward of Tug Hill than CTRL resulting in an increased negative precipitation bias. Meanwhile, a column-like experiment had more realistic precipitation windward of Tug Hill, but compared poorly to XPR velocity profiles and was less representative of crystal habit observations. These habit experiments showed that the Thompson MP scheme is sensitive to the choice of parameters used to determine snow PSD and fall velocity in the model for this case study.

Besides studying the properties of snow (e.g., crystal habit, riming), future work should investigate the sensitivity (or lack thereof) of LeS forecasts to differences in the dominant precipitation type between snow and graupel forecast by various MP schemes. In particular, future studies could focus on the performance of newer MP schemes that incorporate variable riming, such as P3, which did not perform particularly well in this case study. Overall, cold-cloud microphysics has many complex processes being parameterized, from collection of snow and graupel all the way down to the smallest-scale processes such as cloud water and cloud ice nucleation, which are themselves sensitive to temperature and aerosol concentrations, among other factors. Many of these processes are uncertain in LeS and other storms, and could be studied in additional sensitivity experiments not performed for this case.

In comparison to the results from the MP ensemble, the PBL/SL ensemble revealed that the primary differences between PBL/SL schemes were both the LeS intensity and morphology, with relatively small impacts on LeS band position and precipitation type. These differences lead to a larger spread and higher median in maximum QPF amounts in the

PBL/SL ensemble (54–99 mm), as compared to the MP ensemble (35–62 mm). However, PBL/SL schemes suffered from some similar biases as the MP schemes, including LeS bands that were 1000–1500 km^2 too small for much of the event. The primary processes accounting for differences between PBL/SL members were the amount of sensible and latent heat fluxes generated from Lake Ontario, with high-flux members such as MYJ and especially QNSE producing too much precipitation windward of Tug Hill. The spread in surface fluxes and maximum QPF among PBL/SL schemes is consistent with past case studies (Conrick et al. 2015; Fujisaki-Manome et al. 2017). Future work should investigate PBL/SL scheme performance in additional LeS cases, especially those with available over-lake surface flux measurements to help constrain the flux uncertainty in SL schemes.

While the results presented in this study are specific to OWLeS IOP2b and must be evaluated in other cases to verify their generalizability to other LeS forecasts, several key conclusions are summarized below for both operational forecasters and model developers.

6.1 Relevance for operational forecasters

- PBL/SL ensemble has more spread in max QPF amounts than MP ensemble, although both ensembles show uncertainty in LeS intensity, areal extent, and inland extent of the heaviest precipitation amounts, important factors for a forecaster making decisions about LeS warning polygons.

- MP ensemble differences are due in part to uncertainty in the relative amounts of forecast precipitation types (snow versus graupel), while PBL/SL ensemble differences are primarily driven by uncertainty in the amount of surface sensible and latent heat fluxes, so a forecaster might want to view model precipitation types and surface fluxes to help understand QPF differences between HREF members.

- Based on these convection-permitting WRF simulations of OWLeS IOP2b with a HRRR-like physics configuration, several schemes exhibit clear biases:

  - Thompson MP (HRRR) is too snow-dominant while WSM6 (HRW ARW, HRW NSSL) is too graupel-dominant, reality is somewhere in between but closer to Thompson.
  - WSM6 has more total QPF windward of Tug Hill relative to Thompson, with too-intense LeS rates relative to radar-estimated precipitation rates.
  - MYJ PBL/SL (HRW NSSL, HRW NMMB, NAM Nest) has larger fluxes than MYNN PBL/SL (HRRR), increasing total QPF and leading to too-intense LeS rates in MYJ relative to radar-estimated precipitation rates. However, total QPF in MYJ may be slightly better than in MYNN, which is too low.
  - Forecasters might want to consider the high biases in LeS intensity when viewing HREF members using WSM6 MP (HRW ARW, HRW NSSL) and MYJ PBL/SL (HRW NSSL, HRW NMMB, NAM Nest).

- Consistent small area biases in both MP and PBL/SL ensembles appear to be independent of the physics schemes tested, and forecasters might want to consider this bias when determining the areal extent of LeS warning polygons.

6.2 Relevance for NWP model developers

- While a multi-physics ensemble can provide large QPF spread and precipitation-type diversity, many assumptions and parameterizations varying across numerous MP, PBL, and SL schemes greatly complicate interpretation of physical processes and QPF differences between schemes. Some schemes could have systematic biases (e.g., too much graupel, too-high sensible heat fluxes) that may limit ensemble performance.

- As in a multi-physics ensemble, a future stochastic physics-based ensemble suitable for forecasting frozen precipitation needs to predict a range of QPF amounts and precipitation types to better quantify forecast uncertainty due in part to MP, PBL, and SL schemes.

- When compared with observations from OWLeS IOP2b, some fixed parameter value assumptions in Thompson MP (e.g., PSD, V-D relations for snow) do not accurately describe uncertainty in and temporal variability of frozen precipitation characteristics, especially with complications such as non-spherical ice and environments with partial riming like OWLeS IOP2b.

- Observationally derived, fixed parameters (e.g., snow PSD coefficients, Prandtl number) are good candidates for further study of stochastic physics (specifically, stochastic parameter perturbations, or SPP) in convection-permitting ensemble model development, to better describe physical parameterization uncertainty within individual parameterization schemes. Perturbations to these parameters can generate large spread that has a physical basis, and these parameter values can be somewhat constrained by observations.

BIBLIOGRAPHY

Aligo, E. A., B. Ferrier, and J. R. Carley, 2018: Modified NAM Microphysics for Forecasts of Deep Convective Storms. *Monthly Weather Review*, **146 (12)**, 4115–4153, doi:10.1175/MWR-D-17-0277.1.

Angevine, W. M., H. Jiang, and T. Mauritsen, 2010: Performance of an Eddy Diffusivity–Mass Flux Scheme for Shallow Cumulus Boundary Layers. *Monthly Weather Review*, **138 (7)**, 2895–2912, doi:10.1175/2010MWR3142.1.

Battaglia, A., E. Rustemeier, A. Tokay, U. Blahak, and C. Simmer, 2010: PARSIVEL Snow Observations: A Critical Assessment. *Journal of Atmospheric and Oceanic Technology*, **27 (2)**, 333–344, doi:10.1175/2009JTECHA1332.1.

Benjamin, S. G., G. A. Grell, J. M. Brown, T. G. Smirnova, and R. Bleck, 2004: Mesoscale Weather Prediction with the RUC Hybrid Isentropic–Terrain-Following Coordinate Model. *Monthly Weather Review*, **132 (2)**, 473–494, doi:10.1175/1520-0493(2004)132<0473:MWPWTR>2.0.CO;2.

Benjamin, S. G., and Coauthors, 2016: A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Monthly Weather Review*, **144 (4)**, 1669–1694, doi:10.1175/MWR-D-15-0242.1.

Bergmaier, P. T., B. Geerts, L. S. Campbell, and W. J. Steenburgh, 2017: The OWLeS IOP2b Lake-Effect Snowstorm: Dynamics of the Secondary Circulation. *Monthly Weather Review*, **145 (7)**, 2437–2459, doi:10.1175/MWR-D-16-0462.1.

Berner, J., and Coauthors, 2017: Stochastic Parameterization: Toward a New View of Weather and Climate Models. *Bulletin of the American Meteorological Society*, **98 (3)**, 565–588, doi:10.1175/BAMS-D-15-00268.1.

Bougeault, P., and P. Lacarrere, 1989: Parameterization of Orography-Induced Turbulence in a Mesobeta–Scale Model. *Monthly Weather Review*, **117 (8)**, 1872–1890, doi:10.1175/1520-0493(1989)117<1872:POOITI>2.0.CO;2.

Bougeault, P., and Coauthors, 2001: The MAP Special Observing Period. *Bulletin of the American Meteorological Society*, **82 (3)**, 433–462, doi:10.1175/1520-0477(2001)082<0433:TMSOP>2.3.CO;2.

Braham, R. R., and R. D. Kelly, 1982: *Cloud Dynamics*. D. Reidel, 87–101 pp.

Bretherton, C. S., and S. Park, 2009: A New Moist Turbulence Parameterization in the Community Atmosphere Model. *Journal of Climate*, **22 (12)**, 3422–3448, doi:10.1175/2008JCLI2556.1.

Brown, R. A., T. A. Niziol, N. R. Donaldson, P. I. Joe, and V. T. Wood, 2007: Improved Detection Using Negative Elevation Angles for Mountaintop WSR-88Ds. Part III: Simulations of Shallow Convective Activity over and around Lake Ontario. *Weather and Forecasting*, **22 (4)**, 839–852, doi:10.1175/WAF1019.1.

Campbell, L. S., and W. J. Steenburgh, 2017: The OWLeS IOP2b Lake-Effect Snowstorm: Mechanisms Contributing to the Tug Hill Precipitation Maximum. *Monthly Weather Review*, **145 (7)**, 2461–2478, doi:10.1175/MWR-D-16-0461.1.

Campbell, L. S., W. J. Steenburgh, P. G. Veals, T. W. Letcher, and J. R. Minder, 2016: Lake-

Effect Mode and Precipitation Enhancement over the Tug Hill Plateau during OWLeS IOP2b. *Monthly Weather Review*, **144 (5)**, 1729–1748, doi:10.1175/MWR-D-15-0412.1.

Chen, C., R. Beardsley, and G. Cowles, 2006: An Unstructured Grid, Finite-Volume Coastal Ocean Model (FVCOM) System. *Oceanography*, **19 (1)**, 78–89, doi:10.5670/oceanog.2006.92.

Chu, P. Y., J. G. W. Kelley, G. V. Mott, A. Zhang, and G. A. Lang, 2011: Development, implementation, and skill assessment of the NOAA/NOS Great Lakes Operational Forecast System. *Ocean Dynamics*, **61 (9)**, 1305–1316, doi:10.1007/s10236-011-0424-5.

Cohen, A. E., S. M. Cavallo, M. C. Coniglio, and H. E. Brooks, 2015: A Review of Planetary Boundary Layer Parameterization Schemes and Their Sensitivity in Simulating Southeastern U.S. Cold Season Severe Weather Environments. *Weather and Forecasting*, **30 (3)**, 591–612, doi:10.1175/WAF-D-14-00105.1.

Colle, B. A., M. F. Garvert, J. B. Wolfe, C. F. Mass, and C. P. Woods, 2005: The 13–14 December 2001 IMPROVE-2 Event. Part III: Simulated Microphysical Budgets and Sensitivity Studies. *Journal of the Atmospheric Sciences*, **62 (10)**, 3535–3558, doi:10.1175/JAS3552.1.

Colle, B. A., D. Stark, and S. E. Yuter, 2014: Surface Microphysical Observations within East Coast Winter Storms on Long Island, New York. *Monthly Weather Review*, **142 (9)**, 3126–3146, doi:10.1175/MWR-D-14-00035.1.

Coniglio, M. C., J. Correia, P. T. Marsh, and F. Kong, 2013: Verification of Convection-Allowing WRF Model Forecasts of the Planetary Boundary Layer Using Sounding Observations. *Weather and Forecasting*, **28 (3)**, 842–862, doi:10.1175/WAF-D-12-00103.1.

Conrick, R., H. D. Reeves, and S. Zhong, 2015: The Dependence of QPF on the Choice of Boundary- and Surface-Layer Parameterization for a Lake-Effect Snowstorm. *Journal of Applied Meteorology and Climatology*, **54 (6)**, 1177–1190, doi:10.1175/JAMC-D-14-0291.1.

Eipper, D. T., G. S. Young, S. J. Greybush, S. Saslo, T. D. Sikora, and R. D. Clark, 2018: Predicting the Inland Penetration of Long-Lake-Axis-Parallel Snowbands. *Weather and Forecasting*, **33 (5)**, 1435–1451, doi:10.1175/WAF-D-18-0033.1.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *Journal of Geophysical Research: Atmospheres*, **108 (D22)**, 2002JD003 296, doi:10.1029/2002JD003296.

Fairall, C. W., E. F. Bradley, J. S. Godfrey, G. A. Wick, J. B. Edson, and G. S. Young, 1996a: Cool-skin and warm-layer effects on sea surface temperature. *Journal of Geophysical Research: Oceans*, **101 (C1)**, 1295–1308, doi:10.1029/95JC03190.

Fairall, C. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S. Young, 1996b: Bulk parameterization of air-sea fluxes for Tropical Ocean-Global Atmosphere Coupled-Ocean Atmosphere Response Experiment. *Journal of Geophysical Research: Oceans*, **101 (C2)**, 3747–3764, doi:10.1029/95JC03205.

Ferrier, B. S., 1994: A Double-Moment Multiple-Phase Four-Class Bulk Ice Scheme. Part I: Description. *Journal of the Atmospheric Sciences*, **51 (2)**, 249–280, doi:10.1175/1520-0469(1994)051<0249:ADMMPF>2.0.CO;2.

Field, P. R., R. J. Hogan, P. R. A. Brown, A. J. Illingworth, T. W. Choularton, and R. J.

Cotton, 2005: Parametrization of ice-particle size distributions for mid-latitude stratiform cloud. *Quarterly Journal of the Royal Meteorological Society*, **131 (609)**, 1997–2017, doi: 10.1256/qj.04.134.

Forbes, G. S., and J. H. Merritt, 1984: Mesoscale Vortices over the Great Lakes in Wintertime. *Monthly Weather Review*, **112 (2)**, 377–381, doi:10.1175/1520-0493(1984)112<0377: MVOTGL>2.0.CO;2.

Fujisaki-Manome, A., and Coauthors, 2017: Turbulent Heat Fluxes during an Extreme Lake-Effect Snow Event. *Journal of Hydrometeorology*, **18 (12)**, 3145–3163, doi:10.1175/ JHM-D-17-0062.1.

Garvert, M. F., C. P. Woods, B. A. Colle, C. F. Mass, P. V. Hobbs, M. T. Stoelinga, and J. B. Wolfe, 2005: The 13–14 December 2001 IMPROVE-2 Event. Part II: Comparisons of MM5 Model Simulations of Clouds and Precipitation with Observations. *Journal of the Atmospheric Sciences*, **62 (10)**, 3520–3534, doi:10.1175/JAS3551.1.

Gaudet, L. C., K. J. Sulia, F. Yu, and G. Luo, 2019: Sensitivity of Lake-Effect Cloud Microphysical Processes to Ice Crystal Habit and Nucleation during OWLeS IOP4. *Journal of the Atmospheric Sciences*, doi:10.1175/jas-d-19-0004.1.

Gerbush, M. R., D. A. R. Kristovich, and N. F. Laird, 2008: Mesoscale Boundary Layer and Heat Flux Variations over Pack Ice–Covered Lake Erie. *Journal of Applied Meteorology and Climatology*, **47 (2)**, 668–682, doi:10.1175/2007JAMC1479.1.

Grell, G. A., and S. R. Freitas, 2014: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmos. Chem. Phys*, **14**, 5233–5250, doi:10.5194/acp-14-5233-2014.

Grenier, H., and C. S. Bretherton, 2001: A Moist PBL Parameterization for Large-Scale Models and Its Application to Subtropical Cloud-Topped Marine Boundary Layers. *Monthly Weather Review*, **129 (3)**, 357–377, doi:10.1175/1520-0493(2001)129<0357:AMPPFL>2.0.CO;2.

Helmus, J. J., and S. M. Collis, 2016: The Python ARM Radar Toolkit (Py-ART), a Library for Working with Weather Radar Data in the Python Programming Language. *Journal of Open Research Software*, **4 (1)**, doi:10.5334/jors.119.

Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc*, **42 (2)**, 129–151.

Houze, R. A., and Coauthors, 2017: The Olympic Mountains Experiment (OLYMPEX). *Bulletin of the American Meteorological Society*, **98 (10)**, 2167–2188, doi:10.1175/BAMS-D-16-0182.1.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, **113 (D13)**, D13 103, doi:10.1029/2008JD009944.

Janjić, Z. I., 1994: The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes. *Monthly Weather Review*, **122 (5)**, 927–945, doi:10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.

Janjic, Z. I., 2002: Nonsingular implementation of the Mellor-Yamada Level 2.5 Scheme in the NCEP Meso model. Tech. rep., NCEP Office Note No. 437, 61 pp.

Jankov, I., J. Beck, J. Wolff, M. Harrold, J. B. Olson, T. Smirnova, C. Alexander, and J. Berner, 2019: Stochastically Perturbed Parameterizations in an HRRR-Based Ensemble. *Monthly Weather Review*, **147 (1)**, 153–173, doi:10.1175/MWR-D-18-0092.1.

Jankov, I., and Coauthors, 2017: A Performance Comparison between Multiphysics and Stochastic Approaches within a North American RAP Ensemble. *Monthly Weather Review*, **145 (4)**, 1161–1179, doi:10.1175/MWR-D-16-0160.1.

Jiménez, P. A., J. Dudhia, J. F. González-Rouco, J. Navarro, J. P. Montávez, and E. García-Bustamante, 2012: A Revised Scheme for the WRF Surface Layer Formulation. *Monthly Weather Review*, **140 (3)**, 898–918, doi:10.1175/MWR-D-11-00056.1.

Korolev, A., and G. Isaac, 2003: Phase transformation of mixed-phase clouds. *Quarterly Journal of the Royal Meteorological Society*, **129 (587)**, 19–38, doi:10.1256/qj.01.203.

Korolev, A. V., E. F. Emery, J. W. Strapp, S. G. Cober, and G. A. Isaac, 2013: Quantification of the Effects of Shattering on Airborne Ice Particle Measurements. *Journal of Atmospheric and Oceanic Technology*, **30 (11)**, 2527–2553, doi:10.1175/JTECH-D-13-00115.1.

Kristovich, D. A. R., L. Bard, L. Stoecker, and B. Geerts, 2018: Influence of Lake Erie on a Lake Ontario Lake-Effect Snowstorm. *Journal of Applied Meteorology and Climatology*, **57 (9)**, 2019–2033, doi:10.1175/JAMC-D-17-0349.1.

Kristovich, D. A. R., N. F. Laird, and M. R. Hjelmfelt, 2003: Convective Evolution across Lake Michigan during a Widespread Lake-Effect Snow Event. *Monthly Weather Review*, **131 (4)**, 643–655, doi:10.1175/1520-0493(2003)131<0643:CEALMD>2.0.CO;2.

Kristovich, D. A. R., N. F. Laird, M. R. Hjelmfelt, R. G. Derickson, and K. A. Cooper, 1999: Transitions in Boundary Layer Meso- γ Convective Structures:An Observational

Case Study. *Monthly Weather Review*, **127 (12)**, 2895–2909, doi:10.1175/1520-0493(1999)127<2895:TIBLMC>2.0.CO;2.

Kristovich, D. A. R., and Coauthors, 2017: The Ontario Winter Lake-Effect Systems Field Campaign: Scientific and Educational Adventures to Further Our Knowledge and Prediction of Lake-Effect Storms. *Bulletin of the American Meteorological Society*, **98 (2)**, 315–332, doi:10.1175/BAMS-D-15-00034.1.

Lenschow, D. H., 1973: Two Examples of Planetary Boundary Layer Modification Over the Great Lakes. *Journal of the Atmospheric Sciences*, **30 (4)**, 568–581, doi:10.1175/1520-0469(1973)030<0568:TEOPBL>2.0.CO;2.

Lim, K.-S. S., and S.-Y. Hong, 2010: Development of an Effective Double-Moment Cloud Microphysics Scheme with Prognostic Cloud Condensation Nuclei (CCN) for Weather and Climate Models. *Monthly Weather Review*, **138 (5)**, 1587–1612, doi:10.1175/2009MWR2968.1.

Lin, Y., and B. A. Colle, 2009: The 4–5 December 2001 IMPROVE-2 Event: Observed Microphysics and Comparisons with the Weather Research and Forecasting Model. *Monthly Weather Review*, **137 (4)**, 1372–1392, doi:10.1175/2008MWR2653.1.

Lin, Y., and B. A. Colle, 2011: A New Bulk Microphysical Scheme That Includes Riming Intensity and Temperature-Dependent Ice Characteristics. *Monthly Weather Review*, **139 (3)**, 1013–1035, doi:10.1175/2010MWR3293.1.

Löffler-Mang, M., and J. Joss, 2000: An Optical Disdrometer for Measuring Size and Velocity of Hydrometeors. *Journal of Atmospheric and Oceanic Technology*, **17 (2)**, 130–139, doi:10.1175/1520-0426(2000)017<0130:AODFMS>2.0.CO;2.

Mansell, E. R., C. L. Ziegler, and E. C. Bruning, 2010: Simulated Electrification of a Small Thunderstorm with Two-Moment Bulk Microphysics. *Journal of the Atmospheric Sciences*, **67 (1)**, 171–194, doi:10.1175/2009JAS2965.1.

McMillen, J. D., and W. J. Steenburgh, 2015: Impact of Microphysics Parameterizations on Simulations of the 27 October 2010 Great Salt Lake–Effect Snowstorm. *Weather and Forecasting*, **30 (1)**, 136–152, doi:10.1175/WAF-D-14-00060.1.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics*, **20 (4)**, 851, doi:10.1029/RG020i004p00851.

Milbrandt, J. A., and M. K. Yau, 2005a: A Multimoment Bulk Microphysics Parameterization. Part I: Analysis of the Role of the Spectral Shape Parameter. *Journal of the Atmospheric Sciences*, **62 (9)**, 3051–3064, doi:10.1175/JAS3534.1.

Milbrandt, J. A., and M. K. Yau, 2005b: A Multimoment Bulk Microphysics Parameterization. Part II: A Proposed Three-Moment Closure and Scheme Description. *Journal of the Atmospheric Sciences*, **62 (9)**, 3065–3081, doi:10.1175/JAS3535.1.

Minder, J. R., and Coauthors, 2015: The Evolution of Lake-Effect Convection during Landfall and Orographic Uplift as Observed by Profiling Radars. *Monthly Weather Review*, **143 (11)**, 4422–4442, doi:10.1175/MWR-D-15-0117.1.

Molthan, A. L., B. A. Colle, S. E. Yuter, and D. Stark, 2016: Comparisons of Modeled and Observed Reflectivities and Fall Speeds for Snowfall of Varied Riming Degrees during Winter Storms on Long Island, New York. *Monthly Weather Review*, **144 (11)**, 4327–4347, doi:10.1175/MWR-D-15-0397.1.

Morrison, H., and J. A. Milbrandt, 2015: Parameterization of Cloud Microphysics Based on the Prediction of Bulk Ice Particle Properties. Part I: Scheme Description and Idealized Tests. *Journal of the Atmospheric Sciences*, **72 (1)**, 287–311, doi:10.1175/JAS-D-14-0065.1.

Morrison, H., G. Thompson, and V. Tatarskii, 2009: Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One- and Two-Moment Schemes. *Monthly Weather Review*, **137 (3)**, 991–1007, doi:10.1175/2008MWR2556.1.

Murakami, M., and T. Matsuo, 1990: Development of the Hydrometeor Videosonde. *Journal of Atmospheric and Oceanic Technology*, **7 (5)**, 613–620, doi:10.1175/1520-0426(1990)007<0613:DOTHV>2.0.CO;2.

Nakanishi, M., and H. Niino, 2004: An Improved Mellor–Yamada Level-3 Model with Condensation Physics: Its Design and Verification. *Boundary-Layer Meteorology*, **112 (1)**, 1–31, doi:10.1023/B:BOUN.0000020164.04146.98.

Nakanishi, M., and H. Niino, 2006: An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog. *Boundary-Layer Meteorology*, **119 (2)**, 397–407, doi:10.1007/s10546-005-9030-8.

Nakanishi, M., and H. Niino, 2009: Development of an Improved Turbulence Closure Model for the Atmospheric Boundary Layer. *Journal of the Meteorological Society of Japan*, **87 (5)**, 895–912, doi:10.2151/jmsj.87.895.

Niziol, T. A., 1987: Operational Forecasting of Lake Effect Snowfall in Western and Cen-

tral New York. *Weather and Forecasting*, **2 (4)**, 310–321, doi:10.1175/1520-0434(1987)002<0310:OFOLES>2.0.CO;2.

Niziol, T. A., W. R. Snyder, and J. S. Waldstreicher, 1995: Winter Weather Forecasting throughout the Eastern United States. Part IV: Lake Effect Snow. *Weather and Forecasting*, **10 (1)**, 61–77, doi:10.1175/1520-0434(1995)010<0061:WWFTTE>2.0.CO;2.

Peace, R. L., and R. B. Sykes, 1966: Mesoscale Study of a Lake Effect Snow Storm. *Monthly Weather Review*, **94 (8)**, 495–507, doi:10.1175/1520-0493(1966)094<0495:MSOALE>2.3.CO;2.

Phillips, D. W., 2009: The UAHuntsville X-Band Profiling Radar (XPR). *34th Conf. on Radar Meteorology*, American Meteorological Society, Williamsburg, VA, P11.7.

Pleim, J. E., 2007: A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part I: Model Description and Testing. *Journal of Applied Meteorology and Climatology*, **46 (9)**, 1383–1395, doi:10.1175/JAM2539.1.

Powers, J. G., and M. T. Stoelinga, 2000: A Coupled Air–Sea Mesoscale Model: Experiments in Atmospheric Sensitivity to Marine Roughness. *Monthly Weather Review*, **128 (1)**, 208–228, doi:10.1175/1520-0493(2000)128<0208:ACASMM>2.0.CO;2.

Rauber, R. M., and Coauthors, 2014: Stability and Charging Characteristics of the Comma Head Region of Continental Winter Cyclones. *Journal of the Atmospheric Sciences*, **71 (5)**, 1559–1582, doi:10.1175/JAS-D-13-0253.1.

Reeves, H. D., and D. T. Dawson, 2013: The Dependence of QPF on the Choice of Microphysical Parameterization for Lake-Effect Snowstorms. *Journal of Applied Meteorology and Climatology*, **52 (2)**, 363–377, doi:10.1175/JAMC-D-12-019.1.

Reinking, R. F., and Coauthors, 1993: The Lake Ontario Winter Storms (LOWS) Project. *Bulletin of the American Meteorological Society*, **74 (10)**, 1828–1850, doi: 10.1175/1520-0477-74-10-1828.

Reisner, J., R. M. Rasmussen, and R. T. Bruintjes, 1998: Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model. *Quarterly Journal of the Royal Meteorological Society*, **124 (548)**, 1071–1107, doi:10.1002/qj.49712454804.

Saslo, S., and S. J. Greybush, 2017: Prediction of Lake-Effect Snow Using Convection-Allowing Ensemble Forecasts and Regional Data Assimilation. *Weather and Forecasting*, **32 (5)**, 1727–1744, doi:10.1175/WAF-D-16-0206.1.

Schmidlin, T. W., 1993: Impacts of Severe Winter Weather during December 1989 in the Lake Erie Snowbelt. *Journal of Climate*, **6 (4)**, 759–767, doi:10.1175/1520-0442(1993)006<0759:IOSWWD>2.0.CO;2.

Schultz, D. M., and Coauthors, 2002: Understanding Utah Winter Storms: The Intermountain Precipitation Experiment. *Bulletin of the American Meteorological Society*, **83 (2)**, 189–210, doi:10.1175/1520-0477(2002)083<0189:UUWSTI>2.3.CO;2.

Shin, H. H., and S.-Y. Hong, 2015: Representation of the Subgrid-Scale Turbulent Transport in Convective Boundary Layers at Gray-Zone Resolutions. *Monthly Weather Review*, **143 (1)**, 250–271, doi:10.1175/MWR-D-14-00116.1.

Skamarock, W. C., and Coauthors, 2008: A Description of the Advanced Research WRF Version 3. Tech. rep., NCAR Tech. Note NCAR/TN-475+STR, 113 pp. doi:10.5065/D68S4MVH.

Skofronick-Jackson, G., and Coauthors, 2015: Global Precipitation Measurement Cold Season Precipitation Experiment (GCPEX): For Measurement's Sake, Let It Snow. *Bulletin of the American Meteorological Society*, **96 (10)**, 1719–1741, doi:10.1175/BAMS-D-13-00262.1.

Steenburgh, J., 2014a: North Redfield Snow Study Station Data. Version 1.0. UCAR/NCAR - Earth Observing Laboratory, doi:10.5065/d68051cn.

Steenburgh, J., 2014b: Sandy Creek Snow Study Station Data. Version 1.0. UCAR/NCAR - Earth Observing Laboratory, doi:10.5065/D68K77WG.

Steenburgh, W. J., and L. S. Campbell, 2017: The OWLeS IOP2b Lake-Effect Snowstorm: Shoreline Geometry and the Mesoscale Forcing of Precipitation. *Monthly Weather Review*, **145 (7)**, 2421–2436, doi:10.1175/MWR-D-16-0460.1.

Steenburgh, W. J., L. S. Campbell, and P. G. Veals, 2014: University of Utah North Redfield Radiosonde Data. Version 1.0. UCAR/NCAR - Earth Observing Laboratory, doi:10.26023/RH2Y-8E1R-G905.

Steiger, S. M., and Coauthors, 2013: Circulations, Bounded Weak Echo Regions, and Horizontal Vortices Observed within Long-Lake-Axis-Parallel–Lake-Effect Storms by the Doppler on Wheels*. *Monthly Weather Review*, **141 (8)**, 2821–2840, doi:10.1175/MWR-D-12-00226.1.

Stoelinga, M. T., and Coauthors, 2003: Improvement of Microphysical Parameterization through Observational Verification Experiment. *Bulletin of the American Meteorological Society*, **84 (12)**, 1807–1826, doi:10.1175/BAMS-84-12-1807.

Sukoriansky, S., B. Galperin, and V. Perov, 2005: 'Application of a New Spectral Theory of Stably Stratified Turbulence to the Atmospheric Boundary Layer over Sea Ice'. *Boundary-Layer Meteorology*, **117 (2)**, 231–257, doi:10.1007/s10546-004-6848-4.

Sulia, K. J., and J. Y. Harrington, 2011: Ice aspect ratio influences on mixed-phase clouds: Impacts on phase partitioning in parcel models. *Journal of Geophysical Research: Atmospheres*, **116 (D21)**, doi:10.1029/2011JD016298.

Tao, W.-K., D. Wu, S. Lang, J.-D. Chern, C. Peters-Lidard, A. Fridlind, and T. Matsui, 2016: High-resolution NU-WRF simulations of a deep convective-precipitation system during MC3E: Further improvements and comparisons between Goddard microphysics schemes and observations. *Journal of Geophysical Research: Atmospheres*, **121 (3)**, 1278–1305, doi:10.1002/2015JD023986.

Thompson, G., and T. Eidhammer, 2014: A Study of Aerosol Impacts on Clouds and Precipitation Development in a Large Winter Cyclone. *Journal of the Atmospheric Sciences*, **71 (10)**, 3636–3658, doi:10.1175/JAS-D-13-0305.1.

Thompson, G., P. R. Field, R. M. Rasmussen, W. D. Hall, G. Thompson, P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Monthly Weather Review*, **136 (12)**, 5095–5115, doi:10.1175/2008MWR2387.1.

Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part I: Description and Sensitivity Analysis. *Monthly Weather Review*, **132 (2)**, 519–542, doi:10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2.

Veals, P. G., and W. J. Steenburgh, 2015: Climatological Characteristics and Orographic Enhancement of Lake-Effect Precipitation East of Lake Ontario and over the Tug Hill Plateau. *Monthly Weather Review*, **143 (9)**, 3591–3609, doi:10.1175/MWR-D-15-0009.1.

Veals, P. G., W. J. Steenburgh, and L. S. Campbell, 2018: Factors Affecting the Inland and Orographic Enhancement of Lake-Effect Precipitation over the Tug Hill Plateau. *Monthly Weather Review*, **146 (6)**, 1745–1762, doi:10.1175/MWR-D-17-0385.1.

Villani, J. P., M. L. Jurewicz, and K. Reinhold, 2017: Forecasting the Inland Extent of Lake Effect Snow Bands Downwind of Lake Ontario. *Journal of Operational Meteorology*, **5 (5)**, 53–70, doi:10.15191/nwajom.2017.0505.

Wang, Z., and Coauthors, 2012: Single Aircraft Integration of Remote Sensing and In Situ Sampling for the Study of Cloud Microphysics and Dynamics. *Bulletin of the American Meteorological Society*, **93 (5)**, 653–668, doi:10.1175/BAMS-D-11-00044.1.

Ware, R., R. Carpenter, J. Güldner, J. Liljegren, T. Nehrkorn, F. Solheim, and F. Vandenberghe, 2003: A multichannel radiometric profiler of temperature, humidity, and cloud liquid. *Radio Science*, **38 (4)**, n/a–n/a, doi:10.1029/2002RS002856.

Welsh, D., B. Geerts, X. Jing, P. T. Bergmaier, and J. R. Minder, 2016: Understanding Heavy Lake-Effect Snowfall: The Vertical Structure of Radar Reflectivity in a Deep Snowband over and downwind of Lake Ontario. *Monthly Weather Review*, **144 (11)**, 4221–4244, doi:10.1175/MWR-D-16-0057.1.

Woods, C. P., M. T. Stoelinga, and J. D. Locatelli, 2007: The IMPROVE-1 Storm of 1–2 February 2001. Part III: Sensitivity of a Mesoscale Model Simulation to the Representation of Snow Particle Types and Testing of a Bulk Microphysical Scheme with

Snow Habit Prediction. *Journal of the Atmospheric Sciences*, **64 (11)**, 3927–3948, doi: 10.1175/2007JAS2239.1.

Xue, P., J. S. Pal, X. Ye, J. D. Lenters, C. Huang, and P. Y. Chu, 2017: Improving the Simulation of Large Lakes in Regional Climate Modeling: Two-Way Lake–Atmosphere Coupling with a 3D Hydrodynamic Model of the Great Lakes. *Journal of Climate*, **30 (5)**, 1605–1627, doi:10.1175/JCLI-D-16-0225.1.

Yuter, S. E., and R. A. Houze, 1995: Three-Dimensional Kinematic and Microphysical Evolution of Florida Cumulonimbus. Part II: Frequency Distributions of Vertical Velocity, Reflectivity, and Differential Reflectivity. *Monthly Weather Review*, **123 (7)**, 1941–1963, doi:10.1175/1520-0493(1995)123<1941:TDKAME>2.0.CO;2.

Yuter, S. E., D. E. Kingsmill, L. B. Nance, and M. Löffler-Mang, 2006: Observations of Precipitation Size and Fall Speed Characteristics within Coexisting Rain and Wet Snow. *Journal of Applied Meteorology and Climatology*, **45 (10)**, 1450–1464, doi:10.1175/JAM2406.1.