Assessment of a Microphysical Ensemble Used to Investigate the OWLeS IOP4 Lake-Effect Storm

LAURIANA C. GAUDET,[a] KARA J. SULIA,[a] TZU-CHIN TSAI,[b] JEN-PING CHEN,[b] AND JESSICA P. BLAIR[c,d,e]

[a] *Atmospheric Sciences Research Center, University at Albany, State University of New York, Albany, New York*
[b] *National Taiwan University, Taipei, Taiwan*
[c] *Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma*
[d] *School of Meteorology, University of Oklahoma, Norman, Oklahoma*
[e] *NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma*

(Manuscript received 13 February 2020, in final form 12 February 2021)

ABSTRACT: Microphysical processes within mixed-phase convective clouds can have cascading impacts on cloud properties and resultant precipitation. This paper investigates the role of microphysics in the lake-effect storm (LES) observed during intensive observing period 4 of the Ontario Winter Lake-effect Systems field campaign. A microphysical ensemble is composed of 24 simulations that differ in the microphysics scheme used (e.g., Weather Research and Forecasting Model microphysics options or a choice of two bulk adaptive habit models) along with changes in the representation of aerosol and potential ice nuclei concentrations, ice nucleation parameterizations, rain and ice fall speeds, spectral indices, ice habit assumptions, and the number of moments used for modeling ice-phase hydrometeors in each adaptive habit model. Each of these changes to microphysics resulted in varied precipitation types at the surface; 15 members forecast a mixture of snow, ice, and graupel, 7 members forecast only snow and ice, and the remaining 2 members forecast a combination of snow, ice, graupel, and rain. Observations from an optical disdrometer positioned to the south of the LES core indicate that 92% of the observed particles were snow and ice, 5% were graupel, and 3% were rain and drizzle. Analysis of observations spanning more than a point location, such as polarimetric radar observations and aircraft measurements of liquid water content, provides insight into cloud composition and processes leading to the differences at the surface. Ensemble spread is controlled by hydrometeor type differences spurred by processes or parameters (e.g., ice fall speed) that affect graupel mass.

KEYWORDS: Cloud microphysics; In situ atmospheric observations; Remote sensing; Ensembles; Cloud resolving models; Lake effects

1. Introduction

Parameterized physical processes within numerical weather prediction (NWP) models have varied forecast impacts, which have been investigated via sensitivity studies and/or ensembles. Many studies have investigated the usefulness and skill of ensembles derived from initial condition uncertainty (IC; Golding et al. 2016), physics uncertainty (PHYS), or a combination of both (IC-PHYS; e.g., Stensrud et al. 2000; Fujita et al. 2007; Golding et al. 2016). PHYS ensembles can include uncertainties in one or many processes in the land surface, the boundary layer, radiative transfer, cloud microphysics, and cumulus parameterizations. Although results can differ depending on the season (e.g., warm vs cold; Stensrud et al. 2000; Meng and Zhang 2007), some PHYS ensembles provide a larger envelope of solutions (i.e., forecast spread) encompassing atmospheric variability, compared to IC ensembles. An ensemble forecast that is reliable and encompasses atmospheric variability should randomly sample from the same probability density function (PDF) as the observation (Hamill 2001). Interestingly, IC-PHYS ensembles surpass the capabilities of individual ensembles due to representation of different portions of the atmospheric PDF; IC-PHYS returns the largest forecast spread in thermodynamic variables while also following closest to observations (Fujita et al. 2007). However, situations may arise where model error leads to forecast biases, potentially skewing the range of forecast possibilities. Meng and Zhang (2007) demonstrated that a combination of cumulus schemes both with and without PHYS model error strengthens ensemble performance due to reduced model biases. Saslo and Greybush (2017) used an ensemble-based framework focused on IC and boundary condition (BC) perturbations, different combinations of microphysics and boundary layer schemes, and data assimilation within lake-effect storm (LES) forecasts downwind of Lake Ontario. Changes in ICs and BCs as well as various environmental variables, such as low-level wind, largely influenced precipitation forecasts. The PHYS ensembles suggested that different physics scheme choices can lead to varied precipitation intensities.

Considerable uncertainties surrounding physics representation in NWP currently exist, specifically when considering precipitation on the meso and synoptic scales and hydrometeor phases (i.e., liquid and ice). Liu et al. (2011) found that planetary boundary layer, land surface, and radiative transfer schemes only weakly influenced cold-season precipitation forecasts, while microphysics schemes provided considerable uncertainty. This partially stems from an underdeveloped observational understanding of microphysics due to the innate difficulty with such investigations, which affects the representative parameterizations. Liu et al. (2011) noted that inconsistencies are

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JAS-D-20-0045.1.

Corresponding author: Lauriana C. Gaudet, lgaudet@albany.edu

DOI: 10.1175/JAS-D-20-0045.1

even evident in the implementation of the same exact parameterizations, such as the Bigg (1953) method to represent the freezing of rain, among different schemes. For example, wide-ranging differences among precipitation efficiencies (e.g., creation and growth of graupel) in mesoscale winter storms were discovered when comparing multiple microphysics schemes (Reeves and Dawson 2013; McMillen and Steenburgh 2015). These and many other studies have been conducted solely investigating precipitation sensitivity to physics choices in NWP. Identification of sensitivities to user-defined choices within microphysical models by means of both single and multimodel ensembles is integral to build upon understanding of such processes and their effects on precipitating systems, as well as to recognize which processes produce forecast uncertainty.

This paper explores the ensemble spread produced when using different microphysics, either by varying schemes (e.g., Morrison et al. 2009; Thompson et al. 2008) or by altering a parameter or process physics (e.g., vapor deposition rate) within a microphysics scheme. It also builds upon the exploration of LES quantitative precipitation forecast (QPF) sensitivity to the choice of ice nucleation parameterization and subsequent spherical or nonspherical mode of growth in Gaudet et al. (2019). Only slight forecast differences existed among the QPF magnitude, but there were intriguing changes in the hydrometeor distribution and spatial QPF that open the door for further investigation in this paper.

The remainder of this paper is organized as follows. The microphysical models used herein are described in section 2 and details regarding the simulations within the ensemble in section 3. The following sections address the observations and ensemble forecast of the Ontario Winter Lake-effect System (OWLeS; Kristovich et al. 2017) intensive observing period 4 (IOP4) in a synergistic manner. Section 4 includes an overview of the LES mesoscale characteristics and associated precipitation. A breakdown of the precipitation type and its influence on the LES morphology is in section 5. Section 6 brings focus to the cloud microphysical features, aiming to provide context on to which ensemble members, if any, fully capture the evolution of IOP4. Finally, section 7 closes with a discussion of the strengths and weaknesses of the ensemble for this case.

2. Model descriptions

a. Adaptive habit microphysics model

The adaptive growth of nonspherical ice crystals in varying environmental conditions is tracked in the three-moment, bulk adaptive habit microphysics model (AHM; Harrington et al. 2013a) by means of vapor diffusional mass growth and the mass redistribution hypothesis (Chen and Lamb 1994a). The AHM considers ice crystal habits to be spheroidal, thereby resulting in either prolate spheroids (i.e., columnar ice) or oblate spheroids (i.e., platelike ice) when nonspherical growth freely evolves. Spherical growth of crystals is forced when the temperature-dependent inherent growth ratio, which describes the distribution of mass among the major and minor crystal dimensions, is set to one. Crystal nucleation can be represented by either condensation and deposition freezing (Meyers et al. 1992) or condensation and immersion freezing (DeMott et al. 2015), as these heterogeneous freezing modes are currently understood to contribute the greatest pristine number concentrations within clouds (Kanji et al. 2017). The AHM has been used to investigate mixed-phase cloud processes in both ideal and real Weather Research and Forecasting (WRF) Model cases (Sulia et al. 2014; Sulia and Kumjian 2017a,b; Jensen et al. 2018; Gaudet et al. 2019; Sulia et al. 2020).

b. National Taiwan University microphysics model

Like the AHM, the three-moment National Taiwan University (NTU) model allows for the evolution of ice crystal habit following the parameterization of Chen and Lamb (1994a). It also allows for snow shape to vary, with the assumption that snow is an oblate spheroid. The NTU model describes habit with a single variable, the volume-weighted aspect ratio (Tsai and Chen 2020). Bulk volume and mass are tracked separately to provide a variable bulk density for ice, snow, and graupel. Additionally, the calculation of fall speeds and collision efficiencies depend on the shape and density of ice and snow (Tsai and Chen 2020). The default setup of the NTU model (NTU-DEF) is outlined as follows. The initial cloud condensation nuclei (CCN) distribution is homogeneous in the horizontal and decreases exponentially in the vertical with a scale height of 3.57 km except for the lowest three sigma levels, or below 850 hPa (Cheng et al. 2007, cf.). CCN composition is assumed to be ammonium sulfate, while its size distribution is trimodal lognormal of clean continental type (Whitby 1978). The primary production of ice crystals follows DeMott et al. (2010) for deposition and condensation-freezing nucleation with a given potential ice nuclei (IN) number concentration of $400\,\mathrm{L}^{-1}$ (Georgii and Kleinjung 1967; Chen and Lamb 1994b). Also, raindrop fall speed calculation matches that of Chen and Liu (2004), and the representation of crystal properties (aspherical shape and variable apparent density) is based on the bulk parameterization of adaptive growth habit together with the triple-moment bulk closure method (Chen and Tsai 2016). The fall speeds for solid-phase hydrometeors (pristine ice, snow aggregates, graupel, and hail) follow the theoretical parameterization of Mitchell and Heymsfield (2005).

c. Publicly available WRF microphysics

The WRF Model version 3.7.1 includes several microphysics schemes. The following schemes are used herein: Purdue Lin (LIN; Chen and Sun 2002), WRF single-moment 6-class (WSM6; Hong and Lim 2006), Goddard (GCE; Tao et al. 1989, 2016), WRF double-moment 6-class (WDM6; Lim and Hong 2010), Morrison two moment (M2M; Morrison et al. 2009), Milbrandt–Yau double moment (MY2; Milbrandt and Yau 2005a,b), Thompson (THOM; Thompson et al. 2008), and CAM V5.1 two-moment 5-class (CAM; Eaton 2011). The numerous combinations of physical process representation within these schemes makes it difficult to attribute a specific physical process to the largest discrepancy in a forecast, if the response is indeed linear. The aim of this work is not this type of attribution, but instead to understand the underlying variability derived from microphysics during a mesoscale winter precipitating event.

3. Data and methodology

Gaudet et al. (2019) investigated the LES sensitivity to ice nucleation and growth mode using the AHM. This assemblage is used

TABLE 1. Reference table of all members comprising the UA-NTU ensemble with their associated name and change, if any, to their respective microphysics code. If applicable, the WRF microphysics option is provided for all public members in parenthesis in the setup column.

No.	Name	Reference	Setup (option)
1	AHM-MEY92H	Harrington et al. (2013a,b)	Nonspherical ice with Meyers
2	AHM-MEY92S	Harrington et al. (2013a,b)	Spherical ice with Meyers
3	AHM-DEM15H	Harrington et al. (2013a,b)	Nonspherical ice with DeMott
4	AHM-DEM15S	Harrington et al. (2013a,b)	Spherical ice with DeMott
5	M2M	Morrison et al. (2009)	Semi-2M(10)
6	CAM	Eaton (2011)	Semi-2M(11)
7	GCE	Tao et al. (1989, 2016)	1M(7)
8	LIN	Chen and Sun (2002)	1M(2)
9	MY2	Milbrandt and Yau (2005a,b)	2M(9)
10	THOM	Thompson et al. (2008)	Semi-2M(28)
11	WDM6	Lim and Hong (2010)	Semi-2M(16)
12	WSM6	Hong and Lim (2006)	1M(6)
13	NTU-2SPH	Tsai and Chen (2020)	Spherical ice crystal with 2M
14	NTU-SPH	Tsai and Chen (2020)	Spherical ice crystal
15	NTU-FSI	Tsai and Chen (2020)	Fixed spectral index
16	NTU-FRFS	Tsai and Chen (2020)	Fixed relations of fall speed
17	NTU-DEF	Tsai and Chen (2020)	Default
18	NTU-RFS	Tsai and Chen (2020)	Raindrop fall speed
19	NTU-INH	Tsai and Chen (2020)	High potential IN concentration 40 000 L^{-1}
20	NTU-INL	Tsai and Chen (2020)	Low potential IN concentration 4 L^{-1}
21	NTU-MAERO	Tsai and Chen (2020)	Marine aerosol
22	NTU-PAERO	Tsai and Chen (2020)	Polluted aerosol
23	NTU-DSI	Tsai and Chen (2020)	Diagnosed spectral index
24	NTU-2TRAD	Tsai and Chen (2020)	Traditional parameterization

as an initial framework to better identify which processes and parameters contribute to forecast variations. Analyzing an event through multiple simulations using various physics schemes is considered to be a multiphysics ensemble approach. As such, the 24-member ensemble used herein follows other similar multiphysics ensemble investigations in which the total number of members ranges from 8 to 48 (Jankov et al. 2017; Imran et al. 2018; Yang et al. 2019).

a. *Microphysical ensemble composition*

In addition to using publicly available WRF microphysics modules, potential sources of uncertainty including ice nucleation, mode of ice crystal growth, aerosol concentration, potential IN concentration, fall speeds, and spectral indices were isolated and changed within the AHM and NTU to build a microphysical ensemble. These were chosen based on existing nucleation and growth sensitivities identified in Gaudet et al. (2019) and a prespecified ensemble from NTU. The 24 ensemble simulations and their associated modeling options diverging from the standard namelist are summarized in Table 1. Members 1–4 are referred to as AHM, 5–12 as public, and 13–24 as NTU in the remainder of this paper. Public members were used without any modification.

AHM ensemble member perturbations use a combination of Meyers et al. (1992) (MEY92) or DeMott et al. (2015) (DEM15) ice nucleation parameterization with spherical (AHM-MEY92S, AHM-DEM15S) or nonspherical (AHM-MEY92H, AHM-DEM15H) ice growth. Nonspherical ice growth is calculated and tracked in the AHM through volume-weighted major and minor crystal axis lengths and ice crystal bulk density. A thorough description of the AHM and nucleation parameterizations is provided in Gaudet et al. (2019). The NTU ensemble is built from individual changes to the various representations of specific microphysical processes and parameterizations within NTU-DEF. Initial CN distributions are separately changed to marine and polluted types in NTU-MAERO and NTU-PAERO, respectively (Whitby 1978, cf.). The given potential IN concentration is 4 L^{-1} in NTU-INL and 40 000 L^{-1} in NTU-INH. Raindrop fall speed calculation uses the empirical relation $v = 841.997D^{0.8}$ in NTU-RFS (Liu and Orville 1969). Variations of spectral index (α, shape parameter of the size spectrum) are diagnosed from size in NTU-DSI or fixed as constants ($\alpha = 3$ for pristine ice, $\alpha = 0$ for other hydrometeor categories) in NTU-FSI, which employs a bulk two-moment method. Ice fall speed calculations are replaced with empirical size relations such as $v = aD^b$ in NTU-FRFS, where a and b are constants, v is the fall speed of cloud ice, and D is the spherical equivalent diameter. Pristine ice and aggregate shape is assumed spherical in NTU-SPH; this assumption is combined with the bulk two-moment method in NTU-2SPH. Finally, traditional parameterizations of crystal properties (spherical with fixed density) and fall speed calculations (empirical relations) together with the bulk two-moment method define NTU-2TRAD.

Each ensemble member was run with WRFv3.7.1 for 12–17 December 2013, allowing for sufficient spinup time before the LES observed during OWLeS IOP4 initiated around 1800 UTC 15 December and dissipated around 0900 UTC 16 December. All LES simulations are run with three, two-way nested domains centered at 43.605°N, 76.721°W, seen in Fig. 1a,

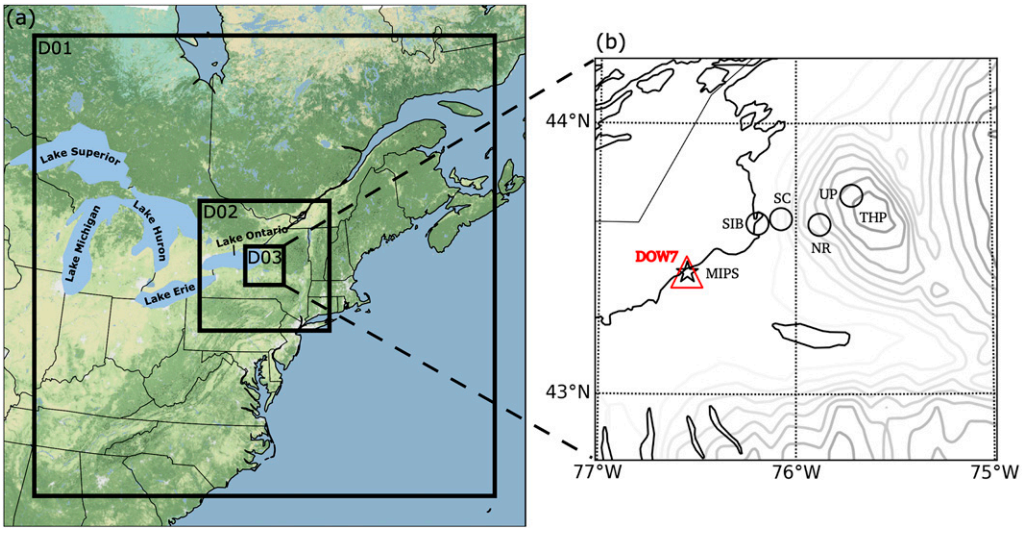

FIG. 1. (a) Nested domains used in WRF ensemble and (b) domain 3 with location of OWLeS IOP4 deployment sites for Micro Rain Radars (circles) at Sandy Island Beach (SIB), Sandy Creek (SC), North Redfield (NR), and the upper plateau (UP); MIPS instrument suite (star); and Doppler on Wheels 7 (DOW7; triangle). Topography is contoured every 50 m beginning at 100 m, with the Tug Hill Plateau (THP) annotated at the highest elevation shown.

varying in horizontal grid spacing from 25 km (domain 1, D01), 5 km (domain 2, D02), and 1 km (domain 3, D03), and 30 nonlinear vertical levels, extending to about 15 km. Other than microphysics, all remaining physics namelist options are outlined in Table 2.

b. Machine comparison

The University at Albany, State University of New York (SUNY, UA), and NTU simulated WRF on two separate machines to produce the ensemble members listed in Table 1. UA simulated members 1–5 while NTU simulated both member 5 and members 13–24. Member 5 was intentionally run on both UA and NTU machines to perform a quantitative comparison and ensure that any differences in the following analysis are solely a result of microphysical perturbations, rather than machine hardware/software differences. A Spearman correlation coefficient of 0.9985 with a p value of 0 indicated no statistically significant difference between member 5 QPFs in D03. This specific test does not assume a normal data distribution.

Additionally, the Spearman correlation coefficients computed for each mass mixing ratio in D03 is as follows: 0.98 for ice, 0.97 for snow, 0.48 for graupel, 0.97 for cloud, and 0.97 for rain mass mixing ratios. The data in Fig. 2 show the differences between the UA and NTU member 5 mixing ratios ($M2M_{NTU} - M2M_{UA}$) at each grid point in D03 every 3 h from 1200 UTC 15 December to 1200 UTC 16 December 2013. Figure 2 demonstrates that snow and cloud mass differences dominate, while ice and rain differences have lesser contributions. On average, these discrepancies are on the order of 1% of the total forecast for ice and rain, 0.1% for cloud, and 0.01% for snow. The only exception is graupel, with differences ranging between 14.8% and 17.4%. A possible explanation of this relatively large difference is the low occurrence (zeroth moment) of graupel in M2M (as judged by Fig. 8). While forecasts of the same member are not compared again in this paper, readers should be aware that there is variation in the graupel forecasts. All hydrometeor difference distributions center around zero, implying that a balance exists between the number of grid points in D03 forecasting more or less liquid or frozen mass relative to the NTU-simulated member 5. Since the differences in the remaining hydrometeor mixing ratios vary at each 3-h output and do not increase or decrease throughout time, it is sufficient to run the simulations on separate machines.

c. Observational and validation datasets

Several observational datasets are available through the National Center for Atmospheric Research (NCAR) Earth Observing Laboratory (EOL) for OWLeS. Table 3 includes the instruments, their deployment locations (Fig. 1b), and observation types used to analyze IOP4. Hourly measurements of snow-water liquid-equivalent (SWLE) precipitation and 6-h snow depth were recorded at two stations, North Redfield (Steenburgh et al. 2014a) and Sandy Creek (Steenburgh et al. 2014b), east of Lake Ontario. To provide a more spatially extensive accumulation dataset, a radar-derived and rain gauge-corrected 24-h quantitative precipitation estimate (QPE) valid

TABLE 2. Reference table of the physical scheme setup for WRF simulations discussed in this prospectus.

Physical process	Scheme	Reference
Longwave radiation	RRTMG	Iacono et al. (2008)
Shortwave radiation	Dudhia	Dudhia (1989)
Boundary layer	YSU	Hong and Noh (2006)
Cumulus (only D01, D02)	Kain–Fritsch	Kain (2004)

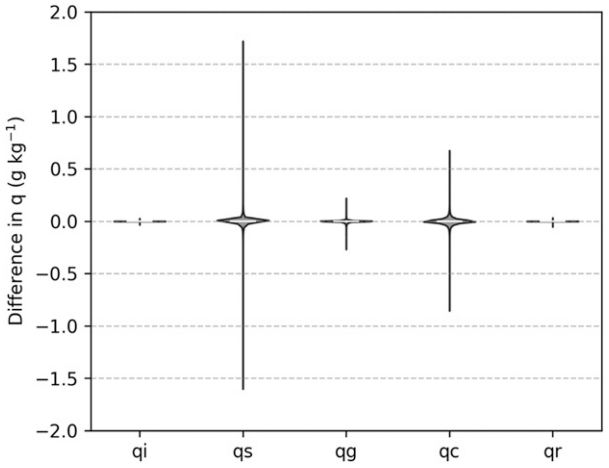

FIG. 2. Violin plot of mass mixing ratio (q) differences (g kg^{-1}) between NTU- and UA-simulated member 5 for ice (qi), snow (qs), graupel (qg), cloud (qc), and rain (qr) mixing ratios at each grid point in D03 and at 3-h intervals during the time period of 1200 UTC 15 Dec to 1200 UTC 16 Dec 2013. The median (horizontal white line) and distribution of differences (shading) is provided by a kernel density plot for each mass mixing ratio. The vertical lines of each plot represent the range of differences.

at 1200 UTC 16 December 2013 was obtained from the Advanced Hydrologic Prediction Service (AHPS; National Weather Service 2020). As these data are not directly measured at the surface, the magnitude of the QPE is relatively uncertain. Therefore, the AHPS data are only used for forecast location validation.

Center for Severe Weather Research (CSWR) Doppler on Wheels (DOW; Wurman 2001) polarimetric X-band radar data provide observations to infer microphysical information about the LES. Out of the three DOWs deployed during IOP4, DOW7 provided the best coverage of the LES, especially over Lake Ontario. As such, only the DOW7 data are used in this paper. These are supplemented by operational KTYX (NOAA 2014) 0.5° plan position indicator (PPI) scans. Four Micro Rain Radars (MRRs; Steenburgh et al. 2014c) were deployed east of Lake Ontario allowing for analysis of LES structure and characteristics with increasing inland distance. These raw data were processed with the Maahn and Kollias (2012) algorithm and averaged to a time resolution of 60 s, as in Minder et al. (2015). The Mobile Integrated Profiling System (MIPS) was also operational during OWLeS, providing data used herein from the Particle Size Velocity (PARSIVEL) optical disdrometer (Phillips and Knupp 2014a) and the X-band profiling radar (XPR; Phillips and Knupp 2014b). The disdrometer was designed to measure rain characteristics and therefore is built on assumptions such as sphericity for particles < 1 mm. While it can and has been used to characterize snow, these assumptions can cause data issues particularly for particles with a diameter < 1 mm (Yuter et al. 2006). MIPS was deployed at SUNY Oswego on the southern shore of Lake Ontario (Fig. 1b, star) due to the likelihood it would detect the LES convective core. However, the storm developed north of the operational forecast location, leaving the MIPS instruments too far south to observe the LES convective core for a majority of its lifetime. As such, ensemble data were averaged over a 20 km $\times$ 20 km area surrounding the MIPS location to mitigate any simulated cloud location issues within comparisons of the ensemble and MIPS observations. Finally, the University of Wyoming King Air (UWKA) flew through the LES in both across and along band directions. Aircraft measurements of liquid water content (LWC) were provided by the Gerber Particle Volume Monitor (PVM; University of Wyoming Flight Center 1977).

4. Precipitation observations and forecasts

A departing cold front associated with a low pressure system off the New England coast introduced an Arctic air mass over the eastern Great Lakes. Strong lake-effect convection ensued over Lake Ontario beginning at 1800 UTC 15 December 2013 and continued until about 0900 UTC 16 December 2013, dumping roughly 18 mm of SWLE near the Tug Hill Plateau (Fig. 1b), according to the AHPS estimates. A thorough description of the associated synoptic and mesoscale features is available in Gaudet et al. (2019). An in-depth analysis of the OWLeS IOP4 observations and the ability of the ensemble to accurately simulate the LES properties leading to the precipitation forecast follows.

a. Assessment of ensemble forecast spread and error

Spearman correlation coefficients were calculated between each ensemble member and the remaining 23 members for the 24-h QPF throughout D03 valid at 1200 UTC 16 December 2013 (Fig. 3). The smaller the correlation, the greater potential difference in QPF with an increased likelihood that a member substantially increases the ensemble spread. In Fig. 3, all NTU members except for NTU-FRFS and NTU-2TRAD are clustered toward higher correlation coefficients, indicating that they are relatively similar to other member forecasts. Since these NTU members comprise half of the ensemble, they are effectively being compared to each other. AHM members using DEM15 are also clustered but those using MEY92 provide

TABLE 3. Instruments deployed during OWLeS IOP4 that provided data used for analysis in this paper.

Instrument	Location	Details	Observation-type used
DOW7	Fig. 1b, triangle	PPI scans	Polarimetric
MRR	Fig. 1b, circles	Vertically pointing	Effective reflectivity, vertical velocity
XPR	MIPS, Fig. 1b, star	Surface	Effective reflectivity, radial velocity
PARSIVEL disdrometer	MIPS, Fig. 1b, star	Surface	Particle size and velocity
UWKA PVM	Fig. 13b	Aircraft	Liquid water content

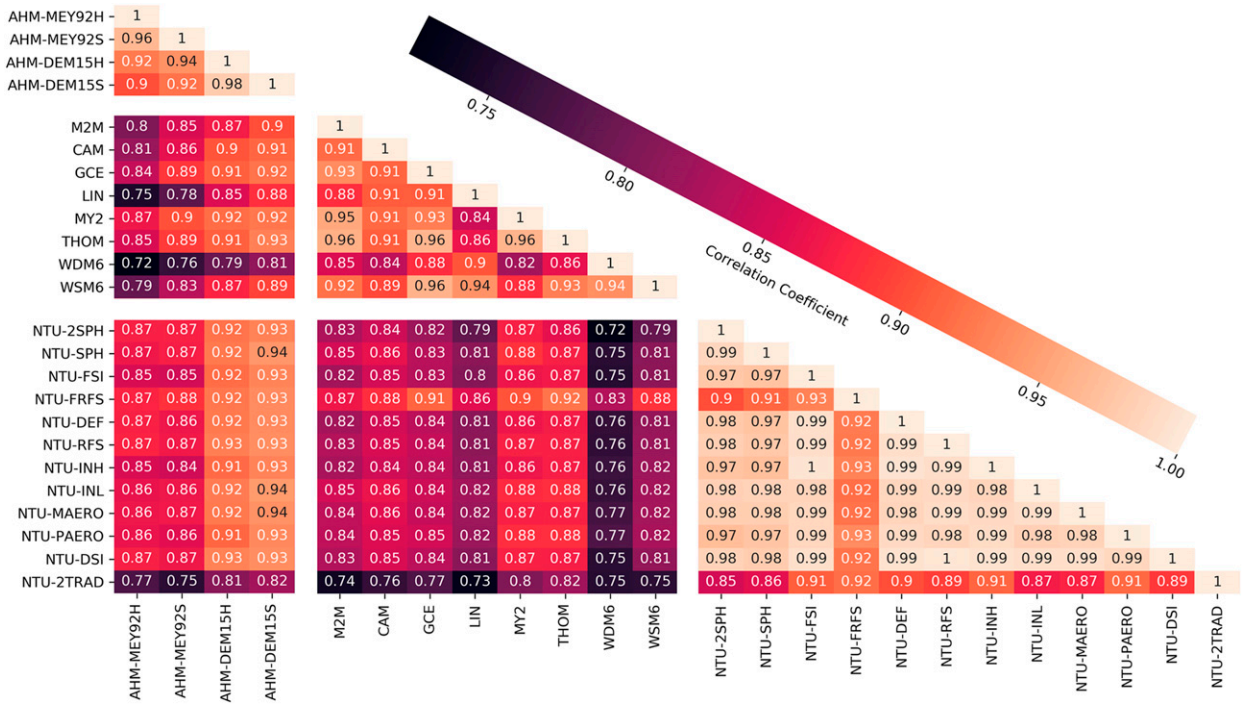

FIG. 3. Spatial correlation coefficients between 24-h QPFs valid at 1200 UTC 16 Dec 2013 of each ensemble member in D03. The correlation is annotated and shaded by color for each ensemble member on both axes. AHM, public, and NTU members are grouped together on both axes by the use of horizontal and vertical white space.

greater QPF differences. Additionally, decreases in correlation coefficient associated with the ice nucleation parameterization changes in the AHM point toward IN concentration potentially controlling some forecast variation. Since correlation coefficient is a proxy for QPF differences, any disparity in correlation coefficient between the AHM-MEY92 and AHM-DEM15 members and the remaining ensemble members stems from the choice of nucleation parameterization. This choice ultimately alters the forecast enough to cause differences between the median correlation coefficient in both AHM-MEY92 members (0.86) and both AHM-DEM15 members (0.92). Finally, the publicly available microphysics schemes in WRF, such as LIN and WDM6, have some of the lowest correlation coefficients, indicating that a change in microphysics scheme may elicit greater ensemble spread than a physical parameter change within the AHM or NTU. This is not a surprising result as physical process representation varies extensively from one scheme to another, making it difficult to pinpoint the root cause for the spread. Smaller correlation coefficients increase forecast spread, which helps identify potential forecast uncertainties and widen the envelope of event forecasts, and also indicate sensitivity to microphysics; identifying that sensitivity is the purpose of this study. This ensemble provides enough variation to warrant an investigation into the reasons driving those forecast differences.

To assess forecast accuracy during 1200 UTC 15 December–1200 UTC 16 December 2013, the hourly SWLE observations at Sandy Creek and North Redfield, New York, were resampled to 3-h accumulations and compared to each member's 3-h (liquid-equivalent) QPF both qualitatively and quantitatively, by use of a root-mean-square error (RMSE; Fig. 4) calculated from eight observed and simulated values. A time series of observations and forecasts (not shown) indicates that most ensemble members overestimate the precipitation at North Redfield before the peak precipitation at 0600 UTC by as much as 5.0 mm while the opposite problem exists at Sandy Creek, underestimating by as much as 6.3 mm. At both locations, public members exhibit a substantial forecast range (i.e., 1.1–9.6 mm at Sandy Creek, 1.7–14.5 mm at North Redfield), AHM members tend toward a relatively moderate forecast range (i.e., 5.0–10.7 mm at Sandy Creek, 8.6–16.3 mm at North Redfield), and the NTU members provide the greatest SWLE forecast (i.e., 6.5–14.8 mm at Sandy Creek, 13.1–17.6 mm at North Redfield), which matches best with the observations at 0600 UTC 16 December (i.e., 15.0 mm at Sandy Creek, 14.7 mm at North Redfield). The RMSE analysis mirrors the preceding qualitative analysis. The closer to zero the RMSE, the less the observations and forecasts deviate, while RMSE > 0 means that the QPF was either under or over the observed amount. The RMSE values in Fig. 4 indicate a range of imperfection within the ensemble at both locations, ranging from 1.1 to 5.4 mm at Sandy Creek and 0.4–4.6 mm at North Redfield. Note that the mean absolute error (MAE) ranges from 0.71 mm (NTU-2SPH, 13) to 2.88 mm (WSM6, 12) at Sandy Creek and 0.30 mm (MY2, 9) to 1.95 mm (WDM6, 11) at North Redfield. At North Redfield, the maximum RMSE and MAE are associated with MY2 (9) and the minimum RMSE and MAE are from WDM6 (11). Although the same is not true at Sandy Creek, the members with maximum (WDM6, 11) and minimum RMSE (NTU-2SPH,

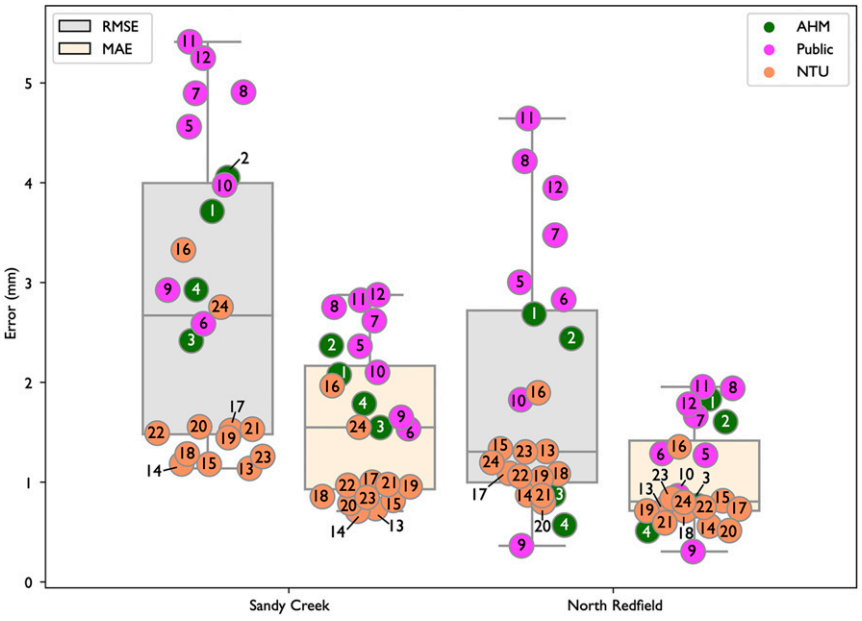

FIG. 4. Root-mean-square error (RMSE; mm, gray) and mean absolute error (MAE; mm, light orange) of SWLE precipitation between observations at Sandy Creek and North Redfield and individual ensemble members from 1200 UTC 15 Dec to 1200 UTC 16 Dec 2013. The interquartile range is represented by each box, the median is the horizontal line within the box, and the remainder of the distribution is denoted by the lines extending from each box. Individual member RMSE and MAE are overlaid as green, pink, and orange circles denoting AHM, public, and NTU members, respectively. Annotated numbers correspond to the simulation number in Table 1.

13) only marginally differ from the members with the maximum and minimum MAE. A majority of NTU members and AHM-DEM15H (3) have RMSE values lesser than the median and lower quartile at both Sandy Creek and North Redfield. The large spread among public members indicates a high uncertainty in bulk microphysical treatments. Interestingly, NTU-FRFS (16) has RMSE values on par with most AHM and public members at both Sandy Creek and North Redfield. As introduced in section 3a, NTU-FRFS replaces ice fall speed calculations with empirical size relations, which signifies that simplifying this calculation adversely impacts QPF error due to a mass flux change toward the surface. At Sandy Creek, most NTU members best forecast 3-h QPF (RMSE closer to zero) while a majority of public members are furthest from observations, which is a factor of spatial forecast differences[1] in the 3-h QPF among the ensemble members. The distribution of RMSE at North Redfield is slightly different; while NTU members provide relatively low RMSE compared to almost all public members and both AHM-MEY members (1 and 2), the lowest RMSE is from a public member, MY2 (9). This analysis allows for the understanding of which clusters of the ensemble lead to the greatest forecast error while also exemplifying the spatial variability of ensemble verification between Sandy Creek and North Redfield.

[1] Defined as the longitudinal and/or latitudinal differences relative to where the forecast field is located for each respective ensemble member.

b. *Spatial and temporal characteristics*

The MIPS disdrometer (see Fig. 1b for location) indicated upticks in precipitation intensity (Fig. 5a) ranging from 1 to 5 mm h^{-1} around 2215 and 2315 UTC 15 December, as well as 0100, 0200, 0330, and slightly before 0530 UTC 16 December, which coincide with increases in reflectivity, ranging from 10 to 25 dBZ (Fig. 5b, stars). Precipitation intensifies slightly before and immediately after 0600 UTC 16 December, peaking at 23.5 mm h^{-1}. This is considerably larger than the preceding intensity observations because the number of observed particles more than doubled and increased in diameter (not shown). The precipitation intensity peak also coincides with a reflectivity maximum of 38.6 dBZ. Based on analysis of KTYX 0.5° PPI scans (Figs. 5c,d), the cloud band propagated southward before 0600 UTC and brought with it large reflectivity values and the sudden increase in heavy precipitation. Lighter precipitation was observed earlier because the disdrometer was sensing precipitation on the periphery of the quasi-zonal convective core (Figs. 5c,d).

The 24-h QPE provided by the AHPS and the 24-h QPFs for all ensemble members valid at 1200 UTC 16 December 2013 are presented in Fig. 6. Each member forecast captures LES precipitation downwind (i.e., east) of Lake Ontario, with some variance in latitudinal placement. Although the QPFs are about double the AHPS observations (Fig. 6, upper left panel), of which the validity is questionable (see Gaudet et al. 2019), some agreement exists among all members as to the expected magnitude of precipitation during IOP4. All AHM and NTU

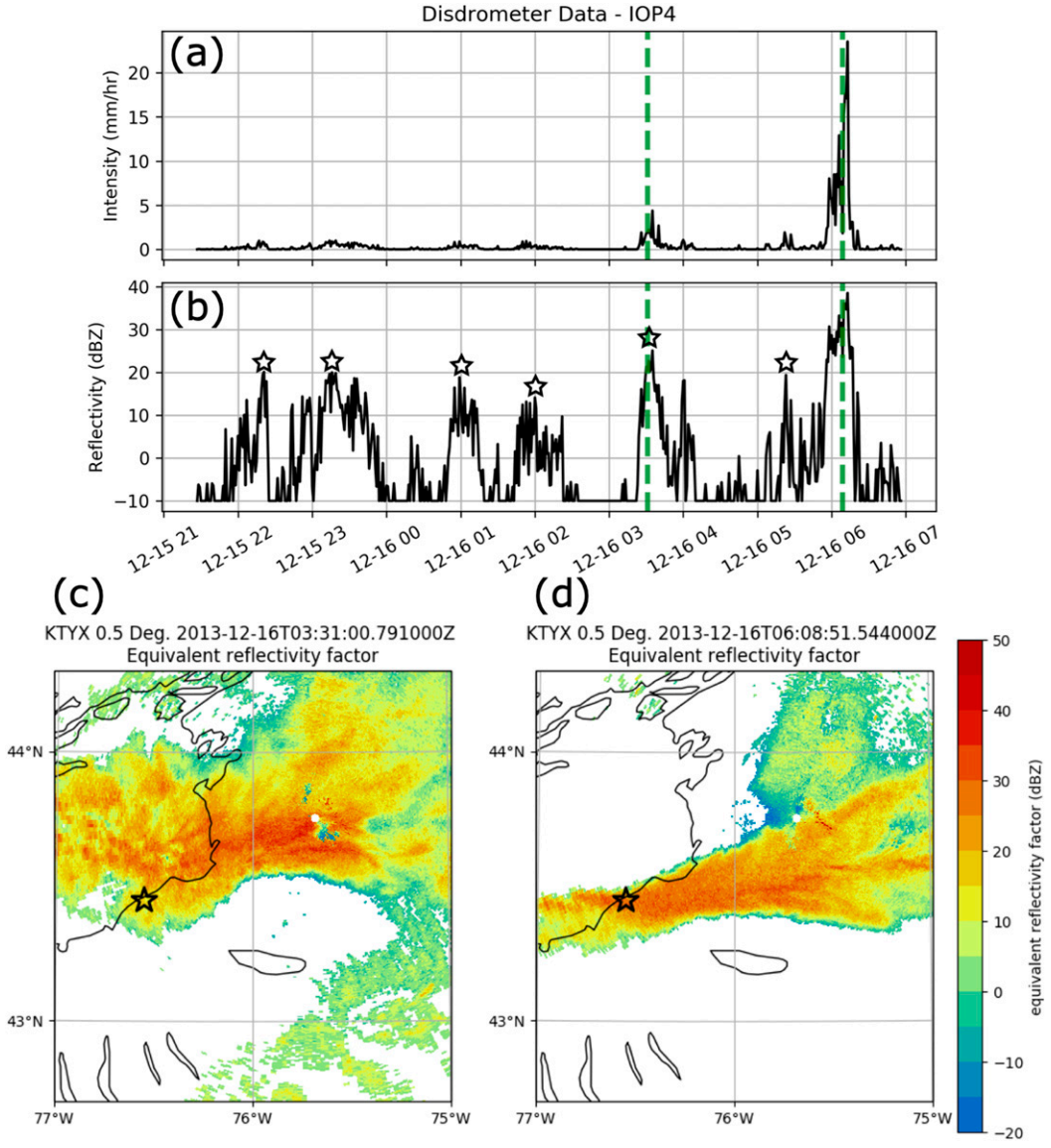

FIG. 5. (a) Intensity (mm h^{-1}) and (b) reflectivity (dBZ) derived by the MIPS disdrometer during the period of 2127–0700 UTC 16 Dec 2013. Vertical dashed green lines correspond to the 0.5° KTYX horizontal reflectivity (dBZ) PPIs valid at (c) 0331:52 and (d) 0609:43 UTC 16 Dec 2013. The stars in (b) indicate the reflectivity enhancements associated with the increases in precipitation intensity at the following times: 2215 and 2315 UTC 15 Dec 2013 and 0100, 0200, 0330, and around 0530 UTC 16 Dec 2013. The black star in (c) and (d) indicates the location of the disdrometer and other MIPS instruments.

members produce a thin band of precipitation southwest of the main area of interest, which was not present in the observations. The mean QPF (D03 average of all horizontal grid points) ranges from 3.5 mm (WDM6) to 6.1 mm (NTU-FRFS); NTU and AHM members generally produce greater mean precipitation than the public members (Fig. 6). While the ensemble accurately forecasts the location of maximum precipitation, the abundance of precipitation elsewhere reduces confidence that this ensemble is able to completely represent IOP4. The following analyses will provide insight into the sources of ensemble forecast variation and how those contribute to the inability to accurately forecast QPF.

During 1800 UTC 15 December–1200 UTC 16 December, measurements of snow totaled 300 mm (24 mm SWLE) at Sandy Creek and 420 mm (22.5 mm SWLE) at North Redfield. The snow-to-liquid ratios (mm mm^{-1}) were 12.5 and 18.67 (within the climatological range in upstate New York; Baxter et al. 2005), respectively, confirming that the accumulated snow had a greater water content at Sandy Creek. So, the snow accumulations increased and SWLE decreased slightly with inland distance. From analysis of almost 30 additional OWLeS-observed LES events, Minder et al. (2015) found that it was common for these storms to become less convective and show

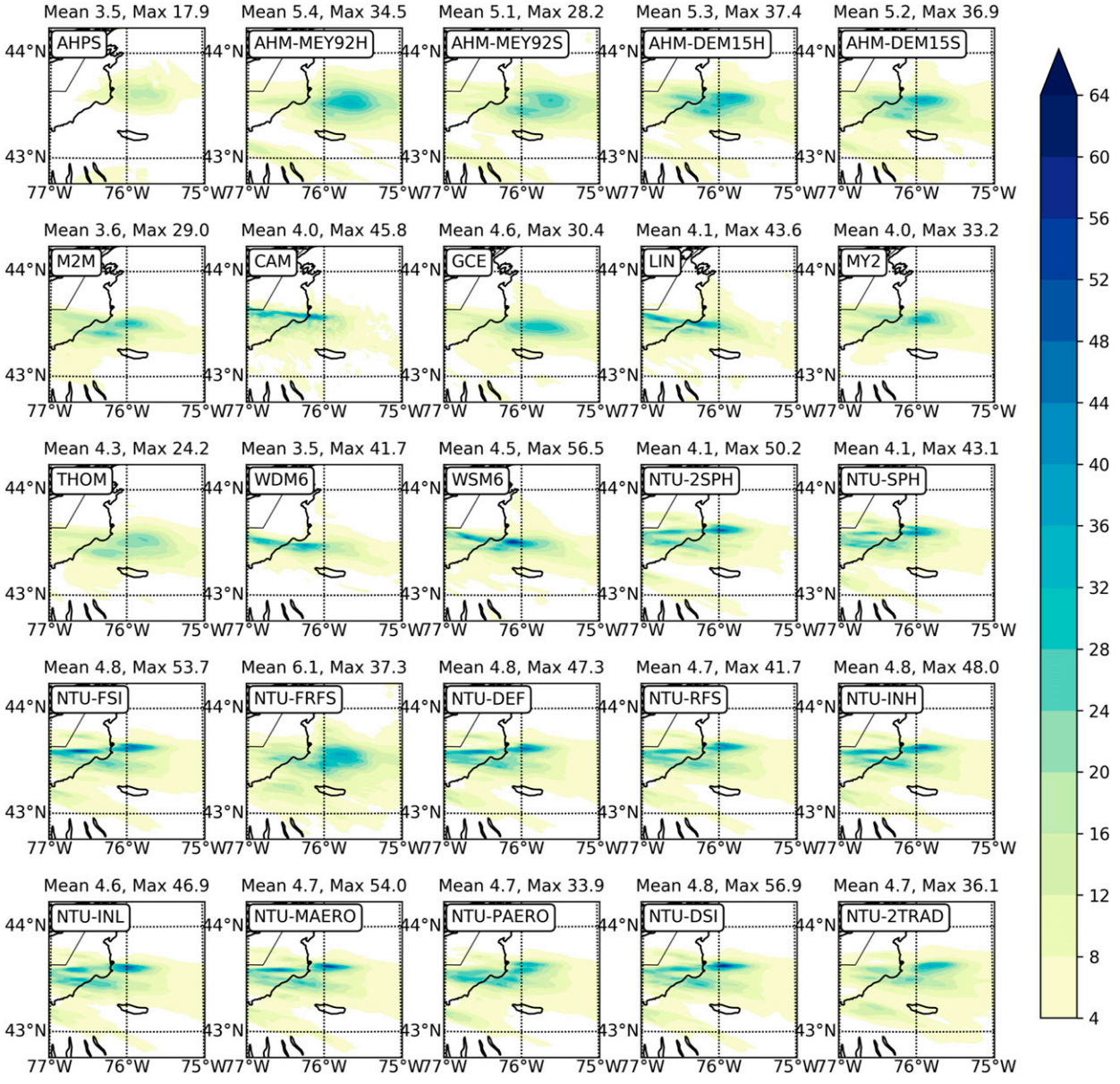

FIG. 6. The 24-h model QPF (mm) in D03, valid at 1200 UTC 16 Dec for each ensemble member outlined in Table 1. Respective mean and maximum QPF (mm) are provided above each panel. (top left) AHPS QPE are presented for reference.

characteristics of a convective-to-stratiform transition with increasing inland extent. Unsurprisingly, the MRR data also suggest that such a transition may have also occurred in IOP4. Contoured frequency with altitude diagrams (CFADs; Yuter and Houze 1995) of effective reflectivity (Z_e) normalized by the number of observations taken by vertically pointing MRRs deployed at Sandy Island Beach, Sandy Creek, North Redfield, and the upper plateau during IOP4 are provided in Fig. 7. The CFAD construction follows the detailed methodology of Minder et al. (2015). Note that data with low observation counts (defined as below the 5th percentile of the total counts at each site) are not shown. Along with a general decrease in maximum reflectivity at all altitudes, the LES echo top lowers with increasing distance inland (Figs. 7a–d), indicative of weakening convection as upward vertical motion becomes less intense (Figs. 7e–h). The negative tilt of the modal Z_e with altitude in all CFADs suggests either melting processes or a size increase of particles detected by the MRRs within the LES (Minder et al. 2015). As the subfreezing temperatures present throughout the atmospheric column during IOP4 would not allow for considerable melting, the tilt of the modal Z_e is likely the product of highly efficient hydrometeor growth processes.

At Sandy Island Beach and Sandy Creek, the modal Z_e in the lowest 1 km of the cloud is >20 dBZ_e and decreases to about 15 dBZ_e at North Redfield and upper plateau. The larger modal Z_e at lake-proximate locations likely corresponds with an

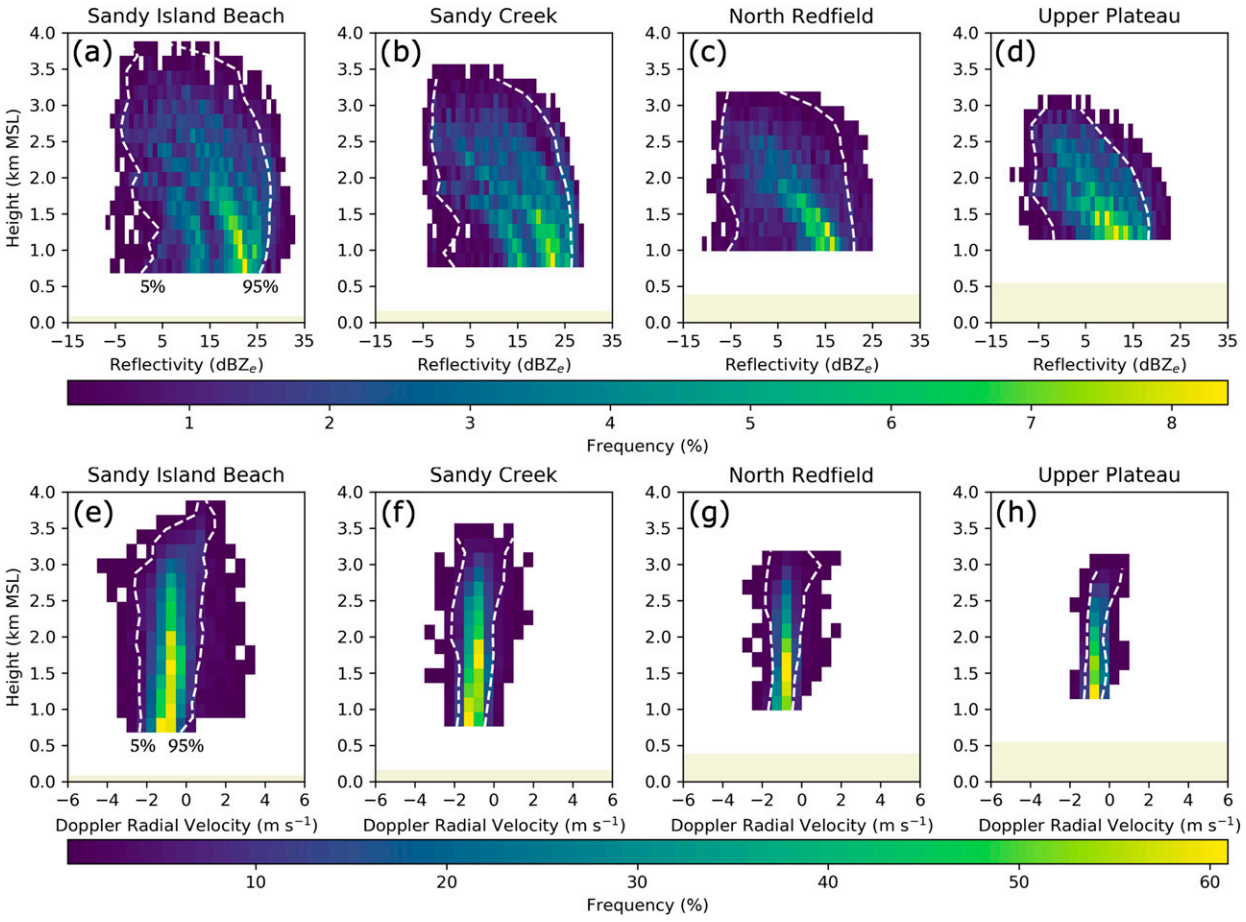

FIG. 7. CFADs of (a)–(d) effective reflectivity (dBZ_e) and (e)–(h) Doppler radial velocity (m s^{-1}) measured by MRRs from 1800 UTC 15 Dec to 0630 UTC 16 Dec 2013 at (a),(e) Sandy Island Beach, (b),(f) Sandy Creek, (c),(g) North Redfield, and (d),(h) the upper plateau in New York. The beige shading at the bottom of each panel indicates the terrain height at each site. The bin sizes for the reflectivity and Doppler radial velocity CFADs at each site are 200 m × 1 dBZ_e and 200 m × 0.5 m s^{-1}, respectively, where the vertical spacing is 200 m. The dashed lines on each panel are the 5th and 95th percentile of the (a)–(d) reflectivity and (e)–(h) radial velocity at each height.

increase in particle density leading to a larger dielectric constant, hydrometeor number concentration, hydrometeor size, or a combination of these (Figs. 7a–d). There is likely a lower concentration of hydrometeors with less LWC at the easternmost locations due to their shifted frequency distributions to smaller Z_e. Additionally, increases in particle density may be supported by increased LWC; an east–west UWKA flight leg at an average altitude of 1.68 km confirms that while there is greater variability in LWC measurements near Lake Ontario, LWC generally decreases with increasing distance from the lake (not shown). This corroborates the manual snow observations: a greater amount of SWLE was observed at Sandy Creek than at North Redfield, meaning that the densities of frozen hydrometeors were greater and/or more cloud liquid water was present at Sandy Creek, as evidenced by the large frequency of high Z_e (Fig. 7b). Meanwhile, precipitation processes were efficient at North Redfield due to a greater observed snow accumulation with a relatively shallow vertical cloud extent and decreased modal Z_e (Fig. 7c).

Further, Figs. 7e–h highlight the 5th- and 95th-percentile Doppler radial velocities[2] (i.e., hydrometeor vertical motion) observed by each MRR throughout the vertical extent of the LES. The intensity of the hydrometeor vertical motion decreases with increasing inland extent. The prevalent upward hydrometeor motion coupled with the direct moisture source at locations closest to Lake Ontario may sustain a greater supersaturation and particle residence time in the cloud system, resulting in greater particle growth and density changes via deposition, aggregation, and/or riming. Since riming would increase particle density and thereby downward hydrometeor motion, it is postulated that riming led to the notable difference

[2] Here, radial velocity as detected by the MRR is defined as the Doppler velocity contributed to by both hydrometeors and air motions, meaning that situations could arise where the hydrometeors and air motions are moving in opposing directions.

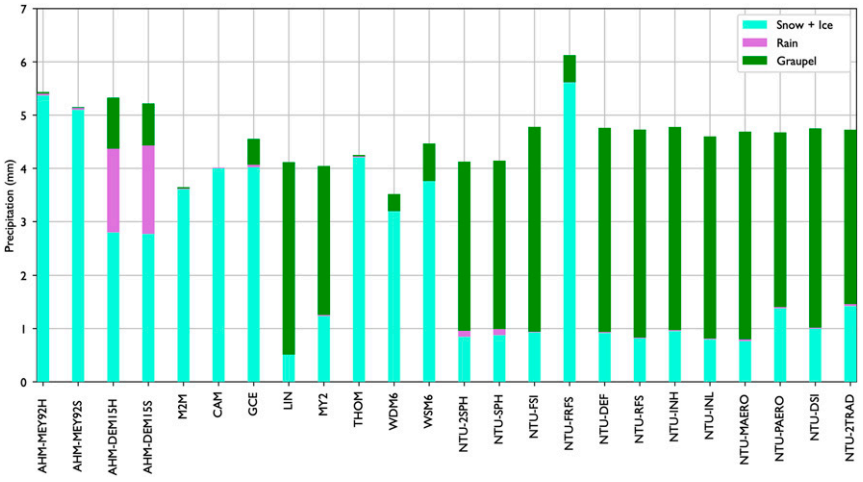

FIG. 8. Bar plot of 24-h D03-average QPF (mm) for each ensemble member, valid at 1200 UTC 16 Dec 2013. The contribution of snow and ice, rain, and graupel in each QPF are represented by turquoise, purple, and green bars, respectively. Note that these bars are *stacked* and not accumulating, and so where the bar begins relative to where it ends must be considered when determining the magnitude.

in this motion between the near-lake Sandy Island Beach MRR observations (Fig. 7e) and those at the upper plateau (Fig. 7h). Analysis of NTU members (not shown) support this hypothesis, as graupel accretion rates among the members increase from the upper plateau to Sandy Creek by 58.2%, but then decrease slightly by 5.3% at Sandy Island Beach; despite that reduction, the accretion rate at Sandy Island Beach ($3.37 \times 10^{-9}\,\mathrm{kg\,kg^{-1}\,s^{-1}}$) surpasses that at upper plateau ($2.25 \times 10^{-9}\,\mathrm{kg\,kg^{-1}\,s^{-1}}$). While graupel density differences of <1% do not sufficiently support differences in precipitation accumulation at the surface, the accretion increases do provide support to these differences. Further, NTU member hail (large graupel that exceeds the Schuman–Ludlum limit during growth; Tsai and Chen 2020) mass-weighted fall speeds are 299% greater at lake-proximate locations. Due to these effects, spatial characteristics of QPF also depend on proximity to Lake Ontario, as this can control the vertical extent of the LES and its convective intensity, which in turn influences hydrometeor growth and sedimentation.

5. Analysis of precipitation type

Ensemble members can be grouped into three subsets based on precipitation types accumulating at the surface (Fig. 8) during the 24-h period ending 1200 UTC 16 December: snow and ice (SI); snow, ice, and graupel (SIG); and snow, ice, graupel, and rain (SIGR). An ensemble member is placed into a group if at least 10% of the D03-averaged accumulation is attributed to any of these precipitation types. Seven members predict SI (AHM-MEY92H, AHM-MEY92S, M2M, CAM, THOM, WDM6, and NTU-FRFS), 15 forecast SIG (GCE, LIN, MY2, WSM6, all NTU members except for NTU-FRFS), and two predict SIGR (AHM-DEM15H, AHM-DEM15S). Autoconversion of cloud droplets to rain and melting of ice crystals at any near-surface and in-cloud grid point where the temperature is above freezing are sources of rain in SIGR simulations.

a. *Ensemble representation of graupel processes*

As with other ice-phase hydrometeors, there are several differences in how each of these microphysics models represent the same graupel production or growth processes, if they are represented at all. CAM is the only member that does not model graupel. Parameterizations used to represent certain physical processes, such as the initiation and growth of graupel, and the manner in which they were implemented into the schemes can largely influence graupel production (Liu et al. 2011; Reeves and Dawson 2013; McMillen and Steenburgh 2015). First, there are differences in the graupel particle size distribution (PSD) intercept parameter: LIN, WSM6, WDM6, and GCE assign a constant value ($4.0 \times 10^4\,\mathrm{m^{-4}}$ in LIN and $4.0 \times 10^6\,\mathrm{m^{-4}}$ in WSM6, WDM6, and GCE), whereas M2M, MY2, THOM, and all AHM and NTU members calculate a value that is based on the graupel mixing ratio and number concentration, in addition to cross-sectional area in the NTU members. Additionally, there are many potential graupel sources, including riming, the freezing of rain, and various combinations of collection between frozen hydrometeors as well as between frozen and liquid hydrometeors. There are criticisms of how these processes are coded within the microphysics models, the efficiency of the collection processes, and if some processes (e.g., collection of snow by graupel) should even be included within the graupel production terms (Liu et al. 2011). MY2 imposes a threshold for conversion of snow to graupel that is different than the other ensemble members; the conversion only occurs when the snow riming rate is at least 3 times greater than the snow deposition rate (Milbrandt and Morrison 2013). THOM, forecasting 0.55% of its QPF as graupel, allows the freezing of rain to contribute to either cloud ice or graupel, depending on the size of the rain

TABLE 4. Percentage of total accumulations broken down by precipitation type observed by the disdrometer and forecast by each ensemble member. The disdrometer data are valid during disdrometer observation period of 2127 UTC 15 Dec to 0657 UTC 16 Dec 2013 and the ensemble time period is valid during 2100 UTC 15 Dec to 0900 UTC 16 Dec 2013.

Label	Snow and ice	Graupel	Rain and drizzle
Disdrometer	92.14	5.31	2.54
AHM-MEY92H	99.45	0.42	0.13
AHM-MEY92S	99.82	0.18	0.00
AHM-DEM15H	54.09	31.23	14.69
AHM-DEM15S	51.32	36.11	12.57
M2M	99.79	0.21	0.00
CAM	99.95	0.00	0.05
GCE	58.21	41.75	0.03
LIN	0.38	99.62	0.00
MY2	12.93	86.95	0.12
THOM	99.72	0.28	0.00
WDM6	71.39	28.61	0.00
WSM6	58.96	41.04	0.00
NTU-2SPH	11.99	87.57	0.44
NTU-SPH	11.33	87.89	0.77
NTU-FSI	10.94	89.04	0.03
NTU-FRFS	89.56	10.44	0.00
NTU-DEF	8.00	91.99	0.01
NTU-RFS	5.12	94.88	0.01
NTU-INH	8.94	91.05	0.00
NTU-INL	5.54	94.46	0.01
NTU-MAERO	7.64	92.30	0.06
NTU-PAERO	18.30	81.63	0.08
NTU-DSI	11.48	88.51	0.02
NTU-2TRAD	8.98	90.93	0.09

(Thompson et al. 2008). AHM members add all frozen rain mass to cloud ice, whereas LIN, WSM6, WDM6, M2M, MY2, GCE, and NTU members add it to graupel. The only member that includes collection of snow by graupel is LIN. All members include collection of rain by snow, but THOM does not instantly convert rain to graupel.

The combination of a low graupel PSD intercept parameter, all frozen rain converted to graupel, the inclusion of graupel and snow collisions, and highly efficient rain and snow collisions resulted in the considerable graupel production accounting for 87.6% of QPF in LIN. NTU-FRFS has slower ice fall speeds, which consequently leads to considerably less graupel production through lessened riming processes. Therefore, the riming efficiency from either the collision efficiency or the fall speed may be too high, as they lead to very active graupel production in all other NTU members. It is less clear why MY2, forecasting 68.9% of its QPF as graupel, followed closely to LIN, seeing as there was no consideration of graupel and snow collisions and the member had a PSD intercept parameter that matched the value of $4.0 \times 10^6\,\mathrm{m}^{-4}$ as in WSM6, WDM6, and GCE. Since there are differences in these microphysical models that extend beyond the representation of graupel, there were likely processes occurring in other hydrometeors that led to cascading effects on graupel. For example, graupel is present in the AHM-DEM15 simulations because the choice of ice nucleation parameterization from DeMott et al. (2015) initiated less pristine cloud ice, allowing for the simultaneous growth of ice and cloud droplets, eventually leading to riming processes (Gaudet et al. 2019). Implementation differences that can lead to considerable changes in mass transfer and ultimately govern precipitation type are important to keep in mind when comparing model output that make use of different microphysics schemes, such as those in this paper.

b. *Precipitation type verification and effects*

Precipitation type was classified at the surface by means of Meteorological Terminal Aviation Routine Weather Report (METAR) codes recorded by the MIPS disdrometer. During IOP4, the following METAR codes were recorded: light, moderate, and heavy snow (−SN, SN, and +SN, respectively), light and heavy soft hail (−GS and +GS, respectively), moderate hail (GR), light and heavy drizzle with rain (−RADZ and +RADZ, respectively), and light, moderate, and heavy drizzle (−DZ, DZ, and +DZ, respectively). Subcategories of hydrometeors and their respective intensities were merged into larger groups; snow combines −SN, SN, and +SN, graupel includes −GS, +GS, and GR, and rain accounts for −RADZ, +RADZ, −DZ, DZ, and +DZ. Between 2127 UTC 15 December and 0657 UTC 16 December, the report count of each code was summed and added to their respective major hydrometeor group. The counts were weighted by the precipitation intensity of each major category, effectively calculating the amount of precipitation each hour while removing the assumption that all hydrometeor categories are contributing to the accumulation equally. The percent contribution of each hydrometeor category was computed for the disdrometer observations during the aforementioned period and for each of the ensemble members during the time period of 2100 UTC 15 December–0600 UTC 16 December (Table 4). According to the recorded METAR codes, 92% of precipitation observed at the MIPS site (Fig. 1b) was snow and ice, with a 5% contribution from graupel and the remaining 3% from rain and drizzle. With the exception of NTU-FRFS, NTU members stray the furthest from observations, with forecasts of snow and ice ranging from 5% to 18%, an exorbitant graupel forecast from 82% to 95%, and rain and drizzle ranging from 0% to 1%. The NTU scheme defines snow as pure aggregates and graupel as rimed ice, in that rimed ice may contain rimed snow while larger pristine ice is what other schemes traditionally define as snow. This may serve to inflate the graupel contributions in most NTU members. GCE, LIN, MY2, WDM6, WSM6, and AHM-DEM15 members also produce an aggressive amount of graupel at the MIPS site (Table 4), ranging from 29% (WDM6) to 100% (LIN) of the total accumulation, for reasons discussed earlier in this section. Ensemble members providing a solution closer to observations included AHM-MEY92 members, M2M, CAM, and THOM with snow contributions around 99%. While these amounts leave little room for graupel, drizzle, and rain, they are far from the extremely unlikely forecasts calling for an equal or dominant contribution from graupel. However, NTU-FRFS best captured the relative hydrometeor contributions, with a 90% forecast of snow,

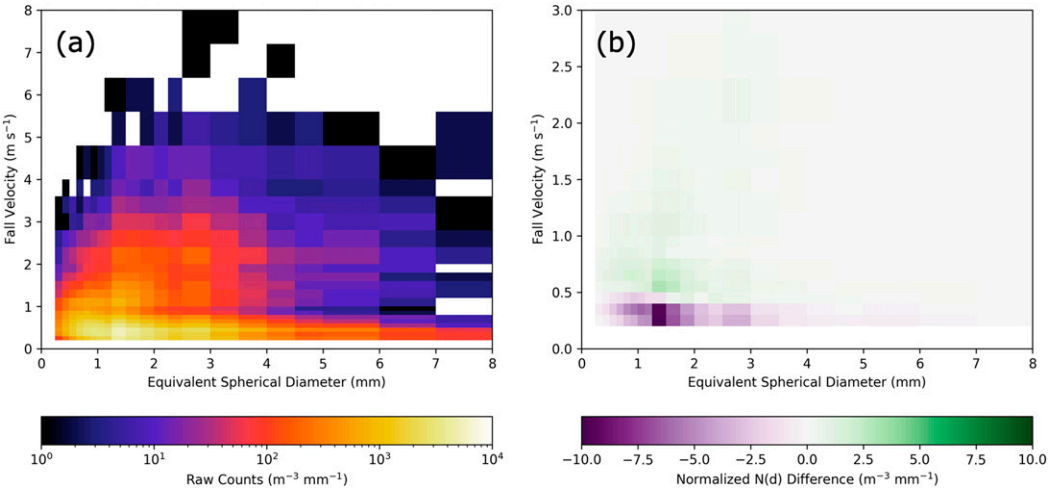

FIG. 9. Joint particle equivalent spherical diameter (mm) and velocity (m s^{-1}) distribution observed by the disdrometer (a) aggregated during the time period of 0000–0657 UTC 16 Dec 2013 and (b) as a difference between the time periods of 0600–0655 and 0300–0400 UTC. These data are normalized by the average particle count observed during each respective time period before calculating the difference shown in (b).

10% forecast of graupel, and no prediction of rain and drizzle. However, this member was also the one of the only NTU members to have an RMSE above the median at both Sandy Creek and North Redfield (Fig. 4). Though this analysis is a point comparison due to the deployment of a single disdrometer, it is a step toward understanding surface precipitation types. As the majority of these data were observed while the LES convective core was to the north of the MIPS site, the hydrometeor composition may not be fully representative of that portion of the LES producing the greatest amount of precipitation. This can lead to discrepancies among the following analyses, in which MIPS observations do not prompt the same conclusions (e.g., LES only composed of snow and ice) as domain-wide comparisons.

Additional understanding of the dominant hydrometeor species accumulating at the surface stems from the joint particle size–velocity distribution (Fig. 9a). However, uncertainties within these disdrometer data exist due to observational noise, sampling effects, and assumptions about the type of particles being sensed. While research shows these uncertainties can be quantified for both snow (Battaglia et al. 2010) and rain (Jaffrain and Berne 2011) events, there is not a standard for quantifying uncertainties for mixed-phase events such as IOP4. Due to this unknown effect, uncertainties exist within these disdrometer data but the actual quantity of which is not accounted for in this work. The greatest number of particles converge around diameters roughly at 1.5 mm and fall speeds of 0.5 m s^{-1}. The extension of a high incidence of particles with diameters > 4 mm may be skewed by the issues of oversizing snow particles with an equivalent spherical diameter < 2 mm (Battaglia et al. 2010). Observations of graupel tend to be in the tail of the fall velocity distribution (>3 m s^{-1}) but do not have much bearing on the total joint distribution since they were infrequently observed. To further understand the particles falling during the LES core passage, which may be representative of some LES convective cores influencing areas east of Lake Ontario earlier in its lifetime, a difference joint distribution normalized by the average particle count during each respective time period was calculated between 0600 and 0655 UTC (high intensity precipitation) and 0300–0400 UTC (low–moderate intensity precipitation) 16 December (Fig. 9b). The latter period shifts the joint distribution to faster velocities, which is hypothesized to correspond well with potentially rimed particles falling at faster speeds due to the precipitation intensification during this time (Fig. 5a).

Both observed and simulated precipitation type mixtures at the ground stem from in-cloud liquid–ice partitioning differences. The presence, or lack thereof, of certain species (e.g., liquid water) can alter remotely retrieved cloud location and expanse due to properties such as particle density and composition. As an example of remotely sensed differences, point probabilities of simulated horizontal reflectivity (Z_H) $>$ 15 dBZ interpolated to a 0.5° PPI scan in D03 at 0600 UTC 16 December are provided for both the entire ensemble and each hydrometeor group and are compared to KTYX-observed 15 dBZ contour in the 0.5° PPI scan in Fig. 10. AHM and NTU members derive simulated Z_H by means of the polarimetric radar operator developed by Ryzhkov et al. (2011) while public members use Z_H calculations provided by the Python package WRF-Python (Ladwig 2017). To calculate these probabilities at each grid point in D03, the number of ensemble members with simulated $Z_H > 15$ dBZ ($N_{Z_H>15\text{dB}Z}$) was divided by the total number of members in the ensemble, ($N_{Z_H>15\text{dB}Z}$)/24, or divided by the total in each respective hydrometeor group, ($N_{Z_H>15\text{dB}Z}$)/(N_{hgroup}). The resulting value at each grid point is considered to be the probability of that Z_H threshold being reached or surpassed and is contoured in Fig. 10.

Each hydrometeor group is well within the observed 15 dBZ contour, but some differences exist among the relative LES

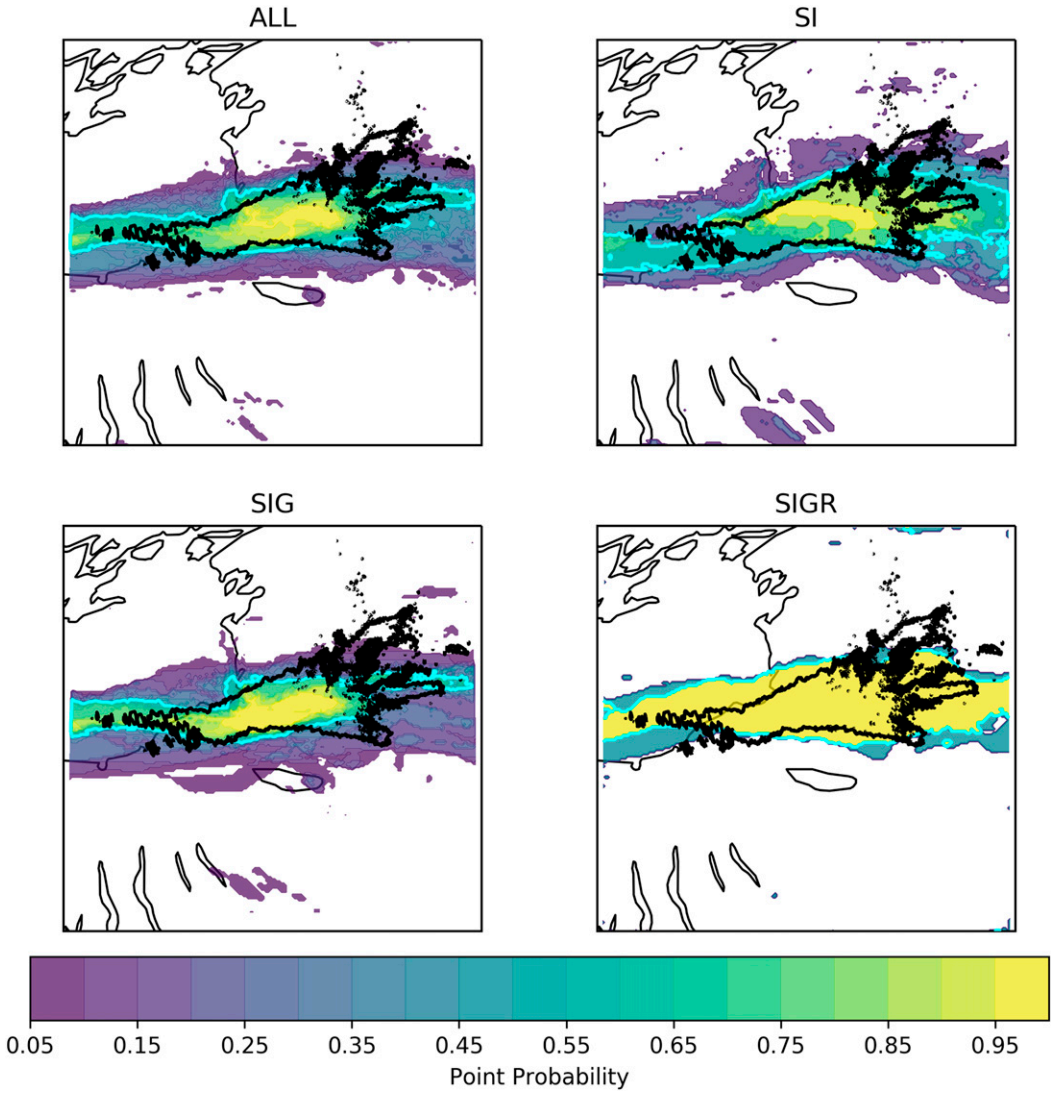

FIG. 10. Point probabilities of $Z_h > 15$ dBZ interpolated to the 0.5° PPI scan strategy at KTYX at 0600 UTC 16 Dec 2013 for (a) all simulations ($N = 24$), as well as simulations that produce (b) snow and ice ($N = 7$), (c) snow, ice, and graupel ($N = 15$), and (d) snow, ice, graupel, and rain ($N = 2$). Superimposed on each panel is the KTYX-observed 15 dBZ contour (black) at the 0.5° PPI scan and 50% point probability contour (cyan).

positioning. The western LES position is slightly to the south of observations in SI and marginally to the north in SIGR. Another interesting feature is the spatial expanse of the 50% probabilities among the groups. SI filled 84.2% of the 15 dBZ KTYX reflectivity contour, but that only comprised 51.3% of the total area forecast to reach the 15 dBZ threshold. Similarly, SIGR filled 83.4% of the observed contour, but that only represented 46.1% of its vast forecast area; 59.3% of the observed contour was filled by SIG, which accounted for 62.0% of its total forecast area. While SIG does not surpass the 71.3% of the observed contour that the entire ensemble forecast, it was higher than the 60.8% of the observed portion of its total forecast area. In both quantitative and qualitative lenses, ensemble members in SIG best match the observations at this time due to the less expansive radar echoes to the north and south of the main LES. This confirms that the precipitation type(s) (e.g., SI, SIG, SIGR; Fig. 8) within the cloud system does have an influence on the spatial extent and location of the LES, specifically when graupel and/or rain are prevalent.

While low-level KTYX PPI analysis provides an overview of the system, the MIPS XPR can resolve details about the vertical profile of reflectivity at the MIPS location, near-surface (which is missed by KTYX) through cloud top. The vertical geometry also provides an opportunity to parse the dynamic vertical structure. The IQR of hourly Z_e data ≥ -5 dBZ_e observed by the XPR is provided in Fig. 11. Note that since the LES is mixed phase, X-band attenuation of any liquid present in the cloud impact these Z_e data. However, any attenuation is likely minimal given that the sensible precipitation is primarily frozen. During the majority of the LES lifetime, there is an

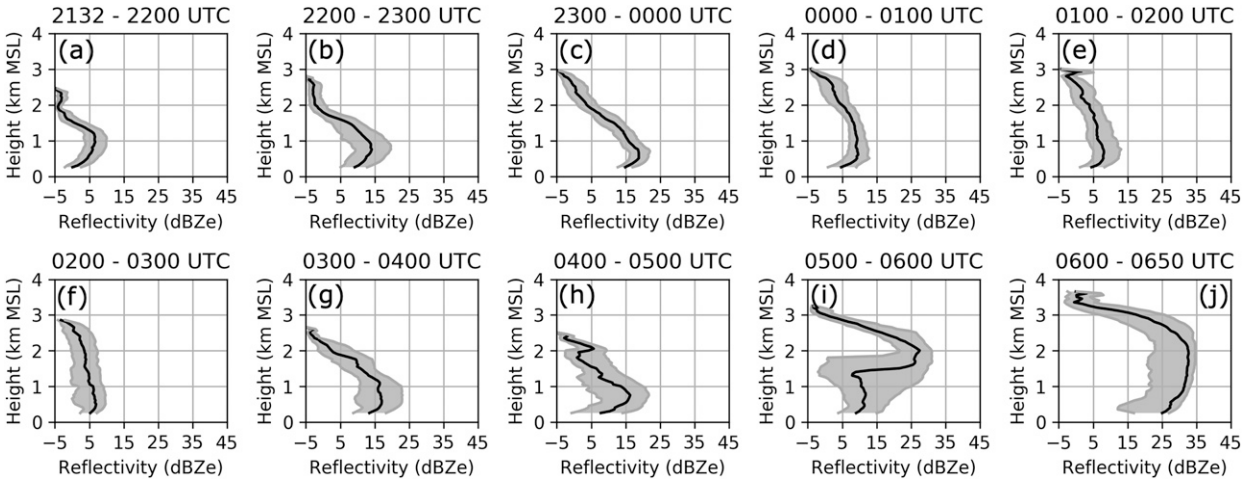

FIG. 11. The interquartile range (shaded gray area) and median effective reflectivity (dBZ_e, solid black line) measured by the XPR are plotted at each height and hourly interval between 2132 UTC 15 Dec and 0650 UTC 16 Dec 2013 for $Z_e \geq -5$ dBZ_e. Each panel represents XPR data from the time periods of (a) 2132–2200, (b) 2200–2300, (c) 2300–0000, (d) 0000–0100, (e) 0100–0200, (f) 0200–0300, (g) 0300–0400, (h) 0400–0500, (i) 0500–0600, and (j) 0600–0650 UTC.

increase in Z_e with increasing proximity at varying levels near the surface (0.5–1 km), suggesting hydrometeor growth, followed by a decrease in Z_e toward the surface (Figs. 11a–h) suggesting sublimation/evaporation. Between 0500 and 0650 UTC (Figs. 11i,j) there is a pronounced Z_e signature maximizing at approximately 2 km, followed by a sharp decrease near 1.5 km between 0500 and 0600 UTC (Fig. 11i). Discussion of how these Z_e data compare to process rates leading to particle growth and decay within the ensemble is provided in the online supplemental materials. The increased altitude at which the maximum Z_e is found during 0500–0650 UTC emanates from the collocation of the MIPS deployment site (Fig. 5d) and the relatively intense LES convective core at the end of its lifespan. However, the Z_e decrease near 1.5 km in Fig. 11i is an artifact of the hourly time intervals; it exists solely because of the arrival of the LES core slightly before 0600 UTC. The level of maximum Z_e during 0500–0650 UTC corresponds with that of the largest radial velocity observed during IOP4, 5 m s^{-1} (not shown). Therefore, hydrometeor lofting is suggested by the collocation of strong radial velocities, considerably increased Z_e especially at relatively higher altitudes, and a decrease of Z_e below due to hydrometeor displacement.

6. Microphysical features

Now that an understanding of surface precipitation type has been established, it will be helpful to explore liquid–ice partitioning within the cloud through remote and in situ observations. Analysis of both observed DOW7 and PVM data as well as concurrent comparison to the ensemble will allow for key inferences to be made.

a. DOW7 polarimetric data analysis and comparison

Polarimetric data collected by DOW7 (situated on the southeastern shore of Lake Ontario, Fig. 1b) during a majority of the LES lifetime allowed for cross-band analysis via 0° azimuthal cross sections. To efficiently evaluate more than one analysis time, CFADs were computed from these cross sections at 30-min intervals during 0000–0600 UTC 16 December. Forward simulation of polarimetric variables (Ryzhkov et al. 2011) was only completed for the AHM-MEY92H and AHM-DEM15H simulations due to the availability of hydrometeor shape information required for computation of polarimetric variables, such as differential phase shift (K_{DP}), differential reflectivity (Z_{DR}), and correlation coefficient (ρ_{hv}).

Much like the XPR (Fig. 11), the DOW7 CFAD in Fig. 12a indicates an increase of Z_H with decreasing altitude to approximately 1 km, which is a classic signature of hydrometeor growth. This is followed by a decrease in Z_H in the lowest 1 km, which may be a result of hydrometeor lofting or sublimation. At each altitude below 3 km, the upper range of reflectivity values in AHM-MEY92H (Fig. 12b) matches best to the DOW7 CFAD. Meanwhile, AHM-DEM15H (Fig. 12c) simulates Z_H values close to double those observed in the 0–3 km layer. Through calculation of the individual contributions of ice, snow, graupel, and rain to the simulation of horizontal reflectivity for both AHM-MEY92H and AHM-DEM15H, it was found that the higher-end Z_H field simulated by AHM-MEY92H is most impacted by snow and ice, whereas the signature of $Z_H > 20$ dBZ simulated by AHM-DEM15H is dominated by snow and graupel.

Observed K_{DP} frequencies center around 0° km^{-1}, but extend to larger positive and negative values throughout the profile (Fig. 12d). The K_{DP} values of 0° km^{-1} imply that a differential phase shift did not occur in the presence of spherical particles while positive (negative) K_{DP} values indicate that the horizontal (vertical) phase shift is larger than the vertical (horizontal), indicative of nonspherical particles. AHM-MEY92H has frequencies of K_{DP} extending past 2° km^{-1} in the 1–3 km layer due to the presence of oblate and prolate ice crystals (Fig. 12e). In the 0–1 km layer, these frequencies start to decrease in magnitude likely due to aggregation scavenging

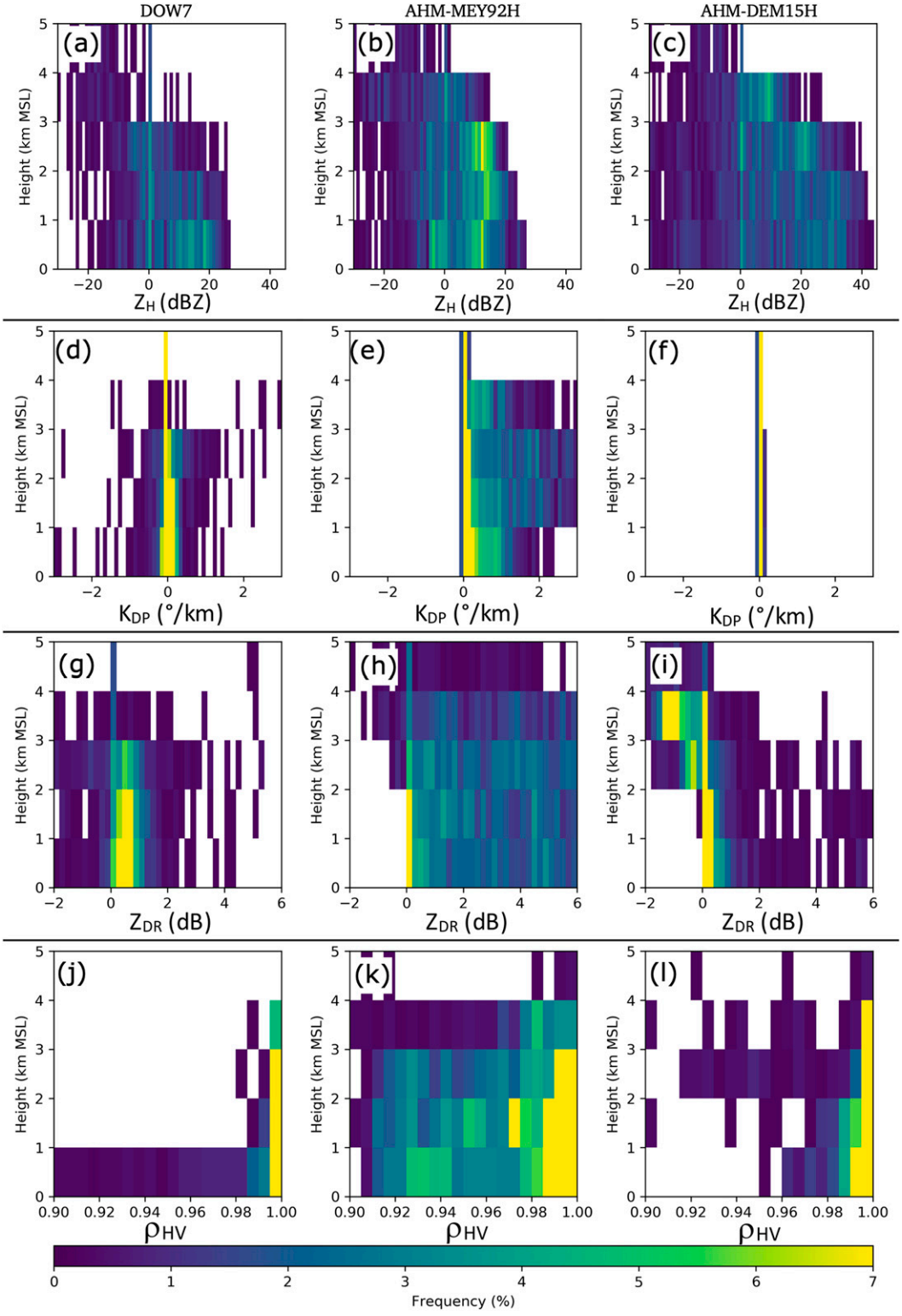

FIG. 12. CFADs of polarimetric radar variables, including (a)–(c) horizontal reflectivity, (d)–(f) differential phase shift, (e)–(i) differential reflectivity, and (j)–(l) correlation coefficient for (left) DOW7, (center) AHM-MEY92H, and (right) AHM-DEM15H simulations. All CFADs were produced at 30-min intervals during 0000–0600 UTC 16 Dec.

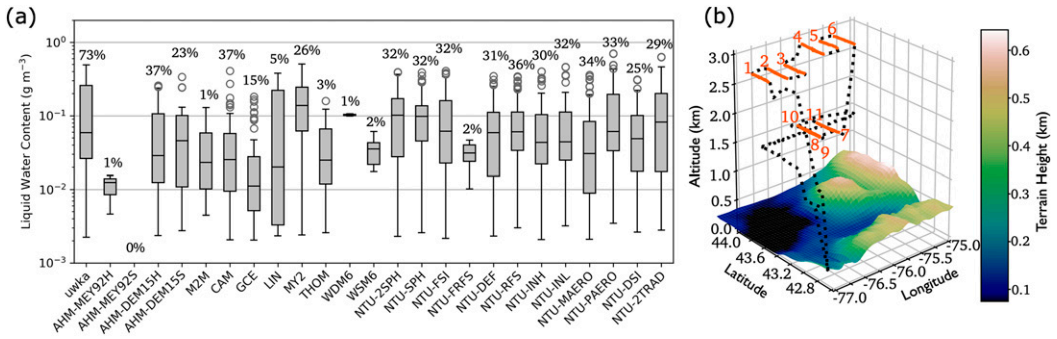

FIG. 13. (a) Box-and-whisker plots of LWC (g m^{-3}) measured by the UWKA PVM during 11 flight legs taking place during 2312 UTC 15 Dec–0103 UTC 16 Dec 2013 and simulated by ensemble members at 0000 UTC 16 Dec 2013. Each box extends from the lower to upper quartiles, has a horizontal line at the median value, includes whiskers to show the maximum and minimum values, and circles to represent outliers beyond $\pm 1.5 \times$ IQR, if applicable. The percentages over each box-and-whisker plot represent the percentage of LWC observations or forecasts in all of the 11 flight legs that are included in the analysis after filtering. (b) IOP4 UWKA flight and innermost domain topography in a three-dimensional space. The highlighted orange segments and accompanying orange annotations indicate the flight leg locations and numbers (1–11) at which the displayed data were observed. Note that some flight legs overlap multiple times.

these crystals, moving them to the snow category, which does not consider shape. Within AHM-DEM15H, K_{DP} values are centered at 0° km^{-1} throughout the 0–5 km layer (Fig. 12f) since quasi-spherical ice crystals were prominent (not shown), resulting in a negligible phase shift.

DOW7 Z_{DR} primarily ranges from 0 to 1 dB up to 2 km with some less frequently observed positive and negative values throughout the lowest 4 km layer (Fig. 12g). Note that due to the relatively low modal reflectivity (Fig. 12a), the Z_{DR} data (Fig. 12g) are less reliable and potentially dominated by noise, therefore resulting in too high of a frequency near 0 dB. Observations above 3 km suggest low concentrations of aspherical particles while increased frequencies of 0 dB below this level are indicative of these aspherical particles aggregating efficiently. Generally, aggregation serves to decrease the exaggerated difference between the horizontal and vertical dimension of the aggregate snow particle (Przybylo et al. 2019), moving the differential reflectivity closer to 0 dB. However, graupel particles can also produce Z_{DR} ranging from −0.5 to 1–2 dB (Straka et al. 2020). Within AHM-MEY92H, a primary concentration around 0 dB exists in the lowest 2 km, but larger frequencies of values up to 6 dB are present, primarily in the lowest 4 km (Fig. 12h). This suggests the considerable presence of aspherical ice particles that have a greater horizontal backscatter. AHM-DEM15H has too large of a frequency of Z_{DR} near 0 dB above 3 km, but it does not present the issue of relatively large frequencies of high Z_{DR} near the surface (Fig. 12i). Also, both simulations produce $Z_{DR} < 0$ dB above 2 km, which better represents observations. Finally, note that AHM-DEM15H concurrently shows high Z_{DR} and K_{DP} around 0° km^{-1}. This is possible due to the different dependencies of these polarimetric variables; Z_{DR} could be responding to a small concentration of aspherical particles which would not be large enough for an appreciable K_{DP} signal.

Lower frequencies of observed ρ_{hv} extend down to 0.90, with higher frequencies clustered roughly above 0.98 in the lowest 1 km layer (Fig. 12j). Extending to the 1–4 km layer yields $\rho_{hv} >$ 0.98. Therefore, hydrometeors sensed by the DOW are relatively homogeneous above 3 km and become increasingly varied toward the surface. The overwhelming presence of snow and ice observed by the disdrometer (Table 4) fits well with the narrative of the DOW7 frequency of ρ_{hv} in the lowest layer. AHM-MEY92H has a wider ρ_{hv} distribution due to the largely ice-dominated cloud system it simulates (Fig. 12k). The greater distribution of ice aspect ratios (not shown) results in a population of ice crystals of various habits that overtakes the particle shape dependence of ρ_{hv}. Through box model simulations, Sulia and Kumjian (2017a) found that larger aspherical ice concentrations can cause greater extremes in ρ_{hv}, perhaps explaining the breadth of the AHM-MEY92H distribution. The highest frequencies of the AHM-DEM15H distribution (Fig. 12l) are more narrowly concentrated to $\rho_{hv} > 0.95$, resembling the rightmost distribution of DOW7 (Fig. 12j). The ρ_{hv} values are closer to unity in AHM-DEM15H because, in addition to the greater abundance of spherical rain, the modal ice aspect ratio only varies between 0.1 and 1 (not shown). However, the different hydrometeor composition in AHM-DEM15H serves to decrease ρ_{hv}.

Neither AHM-MEY92H nor AHM-DEM15H are fully comparable to the DOW7 polarimetric observations as the simulations tended to be on either extreme of the observations. Considerable concentrations of aspherical ice in AHM-MEY92H are likely causing strong K_{DP} and Z_{DR} returns. If more ice to snow conversion through processes such as aggregation occurred within AHM-MEY92H, the resulting K_{DP} and Z_{DR} fields may be more comparable to observations.

b. PVM LWC analysis and comparison

For each ensemble member, LWC values > 2 g m^{-3} at each of the grid points with the same latitude, longitude, and height of 11 UWKA flight legs perpendicular to the southern shoreline of Lake Ontario were aggregated into distributions to understand the range of liquid present in the cloud system

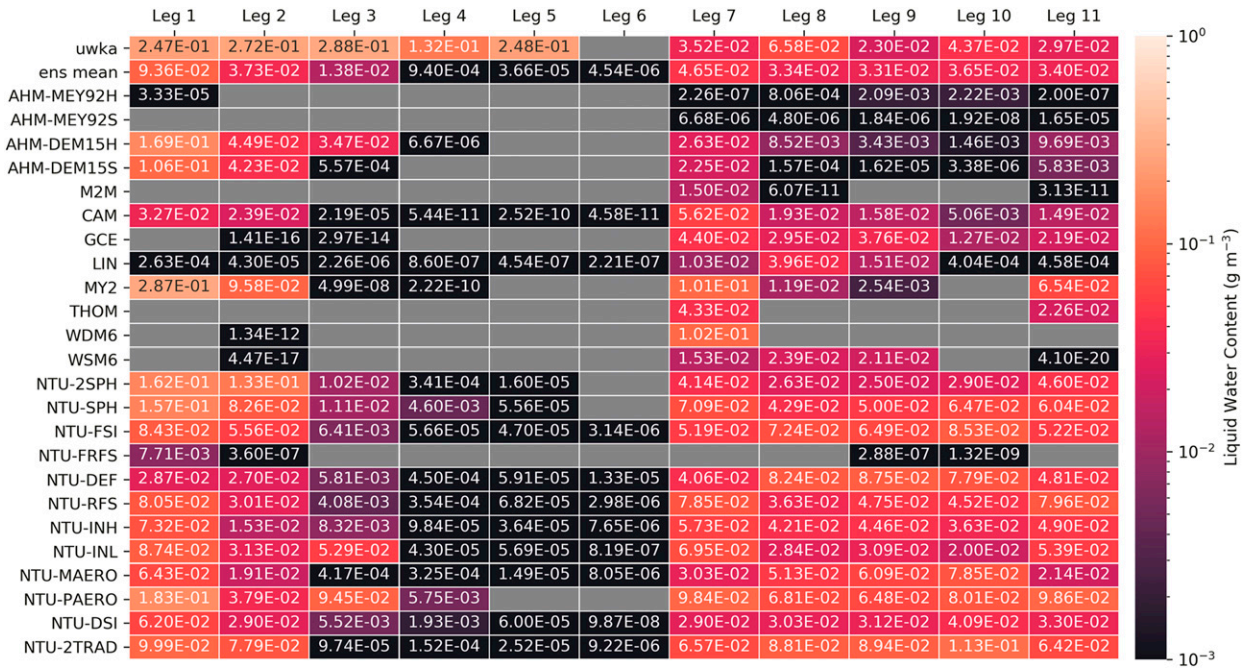

FIG. 14. Heat map of average observed and simulated LWC for the 11 flight legs analyzed in Fig. 13. Flight legs are represented on the x axis, with (top row) UWKA PVM observations, (second row) the ensemble mean LWC, and (remaining rows) the average LWC values for each ensemble member along the same flight path at 0000 UTC 16 Dec 2013. Values are annotated on each box of the heat map. Gray-filled boxes indicate a lack of observed or simulated LWC values $>0\,\mathrm{g\,m^{-3}}$.

(Fig. 13a). The 0000 UTC 16 December output is used for this analysis as it is closest to the UWKA flight time. The PVM LWC data were resampled for each flight leg to compare a similar amount of data to the ensemble forecast LWC data. Resampling time was calculated by dividing the ensemble member grid spacing along the simulated flight path by the average UWKA airspeed along that flight path. Throughout the 11 flight paths spanning approximately three altitudes and legs (Fig. 13b), the PVM measures LWC ranging from 2.24×10^{-3} to $\sim 4.94 \times 10^{-1}\,\mathrm{g\,m^{-3}}$. LWC within this range was measured 73% of the time during its operation on the flight legs. Simulated LWC values below the lower PVM measurement threshold ($2 \times 10^{-3}\,\mathrm{g\,m^{-3}}$; Gerber 2003) are ignored (Fig. 13a); 79% of ensemble members simulated a median within the observed IQR. Five of these members (MY2, NTU-2SPH, NTU-FSI, NTU-PAERO, NTU-2TRAD), all of which are in the SIG group, have distributions that extend outside of the observed IQR. As these members best simulated the persistence of LWC, it is likely that this presence of LWC allowed for graupel production to occur. Note that not all SIG members are comparable to the observed LWC distribution, meaning that graupel production was likely not as zealous as those forecasts suggest. This comparison to observations encapsulates the difficulty of fully capturing the liquid–ice partitioning in the cloud system.

As in Garvert et al. (2005) and Morrison et al. (2015), comparisons are made between the observed and average LWC for each of the 11 flight legs to delve into the variation of LWC (Fig. 14). These model data are not imposed with the lower limit of $2 \times 10^{-3}\,\mathrm{g\,m^{-3}}$ so that it is easier to see which members produced LWC and allowed its persistence during each flight leg. The average flight altitudes were 2.83 km for flight legs 1–3, 3.12 km for legs 4–6, 1.69 km for legs 7–9, and 1.68 km for legs 10–11. The observations on the top row of Fig. 14 indicate that the average LWC is of similar magnitude during the first five flight legs, around $2.9 \times 10^{-1}\,\mathrm{g\,m^{-3}}$. Flight leg 6 did not have any observations $>2 \times 10^{-3}\,\mathrm{g\,m^{-3}}$, which was likely due to the combined effect of the aircraft flying slightly above the cloud top and the flight leg located the furthest inland of those included in this analysis. LWC values decreased by an order of magnitude at the flight levels of 1.69 and 1.68 km, ranging from 2.41×10^{-2} to $6.74 \times 10^{-2}\,\mathrm{g\,m^{-3}}$ in flight legs 9 and 8, respectively. As with Fig. 13, each of the ensemble members represents the observed LWC to a different extent and interesting new details are elucidated by Fig. 14. On average, legs 1–3 were simulated approximately one order of magnitude less than observations. The ensemble members with LWC present did not adequately capture the highest LWC values in flight legs 4 and 5 due to the altitude increase, which was not well resolved at this time in the ensemble. LWC was absent from nine members during leg 4 and 12 members during leg 5. Only leg 7 was well simulated by almost the entire ensemble, except for NTU-FRFS and both AHM-MEY92 members. AHM-MEY92S does not allow for LWC persistence due to its abundance of frozen hydrometeors (Fig. 8). However, LWC does persist in AHM-MEY92H specifically in the lower altitude flight legs 8–10. The AHM-DEM15 members allow for greater persistence of LWC, but still forecast up

to one order of magnitude less LWC in legs 1–3, are on par with leg 7, and again produce up to four orders of magnitude less LWC in legs 8–11. With the exception of CAM and MY2 during flight legs 1 and 2, all public members either produce far too little or no LWC. CAM, GCE, LIN, and MY2 compare much better to observations during the lower-altitude flight legs 7–11. With the exception of NTU-FRFS, NTU members are able to resolve LWC for flight legs 1–3 and 7–11, but do not systematically over or under produce LWC, except for at the highest flight levels (flight legs 7–9). A majority of NTU members allow enough in-cloud persistence of LWC to be comparable to the observations, which is notable considering their underperformance in the precipitation-type comparative analysis.

7. Conclusions

A 24-member microphysical ensemble was built from multiple microphysics schemes available within WRF v3.7.1 and two bulk adaptive habit models, each with varied parameters and representations of physical processes within, including aerosol and potential ice nuclei concentrations, ice nucleation parameterizations, rain and ice fall speeds, spectral indices, ice habit assumptions, and two- or three-moment methods of modeling ice-phase hydrometeors. The ensemble was used to investigate microphysical characteristics and resultant precipitation observed during OWLeS IOP4 with operational and research radar observations as well as in situ surface and aircraft observations. The eight simulations using publicly available microphysics schemes and the 12 members using a bulk adaptive habit model were run at National Taiwan University (NTU). The remaining four simulations using a different bulk adaptive habit model were run at the University at Albany, SUNY (UA). For this reason, a machine difference was conducted for one of these simulations at both UA and NTU. The analysis determined existing differences between these simulations to be negligible, giving credence to continue using the other ''remotely'' simulated ensemble members.

The 24-h QPF valid at 1200 UTC 16 December 2013 of each ensemble member was statistically compared to those of the remaining ensemble members to assess the present forecast spread and member intercomparison. The RMSE of each member was also investigated to assess the ensemble accuracy at Sandy Creek and North Redfield, New York. Almost all NTU members except for NTU-FRFS resulted in lower forecast errors than most ensemble members at Sandy Creek and errors below the median at North Redfield.

Several remote and in situ observations were analyzed to stitch together the in-cloud microphysical processes and surface precipitation. Analysis of MRR data suggest particle growth and/or aggregation at all four sites as well as increased downward hydrometeor vertical motion with increasing proximity to Lake Ontario due to increased riming. Collectively, the ensemble was able to detect areas highly impacted by precipitation as well as correctly simulate LES morphology. The ensemble tended to produce a QPF with two maxima, but only one directly east of Lake Ontario was observed.

The most intriguing result was the various mixtures of hydrometeor types that were forecast to accumulate at the surface: 7 members produced snow and ice, 15 produced snow, ice, and graupel, and 2 forecast snow, ice, graupel, and rain. A disdrometer stationed south of the convective core for the majority of the LES lifetime predominantly observed snow, with $<8\%$ attributed to graupel and/or rain. As the convective core passed over the MIPS deployment site, the disdrometer detected increased fall velocities of larger hydrometeors, as indicated by its PSDs and joint-distributions of equivalent spherical diameter and fall velocity. Within the ensemble, graupel was either grossly overforecast (AHM-DEM15H, AHM-DEM15S, GCE, LIN, MY2, WDM6, WSM6, and all NTU members except NTU-FRFS) or underforecast (AHM-MEY92H, AHM-MEY92S, M2M, CAM, THOM, NTU-FRFS) during the LES event. The same ensemble members that did not predict enough graupel as compared to the disdrometer observations also forecast the most snow. Ensemble members forecasting SI and SIGR increased the areal coverage of the simulated radar reflectivity in D03, thereby slightly changing the locales affected by precipitation.

The signature of hydrometeor growth and/or aggregation observed in the MRR data was also present in the XPR effective reflectivity data. Near the peak intensity of the LES the XPR data showed considerable variability in the vertical, suggesting lofting via strong upward vertical motion, decreasing effective reflectivity in the low levels and increasing effective reflectivity aloft due to the upward flux of hydrometeor mass. Earlier XPR CFAD signatures are confirmed by DOW7 reflectivity vertical profiles, which are similarly compared to the forward operator-simulated polarimetric quantities of AHM-MEY92H (SI) and AHM-DEM15H (SIGR). This analysis elucidated the ability of both simulations to capture some of the observed polarimetric signatures, but neither was fully representative of the observations.

Last, in situ LWC data from 11 UWKA flight legs were compared to the LWC in the ensemble, demonstrating the ability of some SIG members to model LWC distributions comparable to observations. The ensemble members struggle to capture the variability of these LWC data during the first five flight legs, but resolve the lack of LWC in the sixth. Some public members and the majority of NTU members are able to model LWC on the order of the observed magnitude during the final five flight legs.

Processes that directly (through parameterizations) or indirectly affect the creation or growth (e.g., riming) of graupel lead to hydrometeor type differences within the ensemble. The NTU microphysics scheme was most sensitive to the change of ice fall speed parameterization (NTU-FRFS), as use of the empirical size relation reduced riming rates, allowing for in-cloud persistence of snow and ice and depletion of LWC. This resulted in a poor comparison to in-cloud observations and 3-h QPF, but the best ensemble member comparison to precipitation type as observed by the MIPS disdrometer. The remaining NTU members were relatively insensitive to the changes made within the scheme, which resulted in forecasts of SIG and did not negatively affect the RMSE, as this group provided some of the lowest RMSE at both Sandy Creek and North Redfield. Additionally, the faster ice fall speeds indirectly resulted in LWC magnitudes that were on par with observations. Public schemes provided numerous forecast differences due to

different numbers of moments, parameters and parameterization choices, and hydrometeor definitions. This led to a wide range of RMSE, resulting in both the lowest and highest RMSE at North Redfield. These members forecast either SI or SIG; those forecasting SI have poor LWC persistence compared to observations, while members forecasting SIG perform slightly better. Finally, AHM-MEY92 members produced greater RMSE at both locations and poor representation of LWC due to their forecasts of SI. AHM-DEM15 members have lower RMSE at both locations due to their forecasts of SIG, allowing LWC magnitudes to be more comparable to observations.

This work has identified considerable sensitivities of in-cloud hydrometeor partitioning extending to surface precipitation type during a cold-season event, as a direct result of microphysical processes in bulk models. Ensemble difficulty in correctly diagnosing hydrometeor type for IOP4 had implications on remote sensing of cloud LWC and precipitation magnitude, among other characteristics, contributing to the forecast spread. However, comparisons to remote and in situ observations indicated that at times the ensemble was able to capture some characteristics (e.g., spatial, microphysical) of the storm. Even slight changes or seemingly minute choices made within the bulk microphysical models that comprise this ensemble had considerable impacts on this LES forecast.

Acknowledgments. Authors Gaudet and Sulia were supported by the National Science Foundation Partnership for International Research and Education (PIRE) Program between the United States and Taiwan, OISE-1545917, awarded to the University at Albany, SUNY. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation (NSF). Author Blair conducted this research as an undergraduate student and was also supported by the PIRE Program. Authors Gaudet and Sulia would like to extend appreciation to Justin Minder and Massey Bartolini for fruitful discussions about this research. The authors are grateful for the comments and suggestions of three anonymous reviewers, which significantly improved this manuscript. OWLeS data were provided by NCAR/EOL under the sponsorship of the NSF. Figures 5c and 5d were created with the Py-ART package in Python (Helmus and Collis 2016).

REFERENCES

Battaglia, A., E. Rustemeier, A. Tokay, U. Blahak, and C. Simmer, 2010: PARSIVEL snow observations: A critical assessment. *J. Atmos. Oceanic Technol.*, **27**, 333–344, https://doi.org/10.1175/2009JTECHA1332.1.

Baxter, M. A., C. E. Graves, and J. T. Moore, 2005: A climatology of snow-to-liquid ratio for the contiguous United States. *Wea. Forecasting*, **20**, 729–744, https://doi.org/10.1175/WAF856.1.

Bigg, E. K., 1953: The supercooling of water. *Proc. Phys. Soc. London*, **66B**, 688–694, https://doi.org/10.1088/0370-1301/66/8/309.

Chen, J. P., and D. Lamb, 1994a: The theoretical basis for the parameterization of ice crystal habits: Growth by vapor deposition. *J. Atmos. Sci.*, **51**, 1206–1222, https://doi.org/10.1175/1520-0469(1994)051<1206:TTBFTP>2.0.CO;2.

——, and ——, 1994b: Simulation of cloud microphysical and chemical processes using a multicomponent framework. Part I: Description of the microphysical model. *J. Atmos. Sci.*, **51**, 2613–2630, https://doi.org/10.1175/1520-0469(1994)051<2613:SOCMAC>2.0.CO;2.

——, and S. T. Liu, 2004: Physically based two-moment bulk water parametrization for warm-cloud microphysics. *Quart. J. Roy. Meteor. Soc.*, **130**, 51–78, https://doi.org/10.1256/qj.03.41.

——, and T.-C. Tsai, 2016: Triple-moment modal parameterization for the adaptive growth habit of pristine ice crystals. *J. Atmos. Sci.*, **73**, 2105–2122, https://doi.org/10.1175/JAS-D-15-0220.1.

Chen, S.-H., and W.-Y. Sun, 2002: A one-dimensional time dependent cloud model. *J. Meteor. Soc. Japan*, **80**, 99–118, https://doi.org/10.2151/jmsj.80.99.

Cheng, C. T., W. C. Wang, and J. P. Chen, 2007: A modeling study of aerosol impacts on cloud microphysics and radiative properties. *Quart. J. Roy. Meteor. Soc.*, **133**, 283–297, https://doi.org/10.1002/qj.25.

DeMott, P. J., and Coauthors, 2010: Predicting global atmospheric ice nuclei distributions and their impacts on climate. *Proc. Natl. Acad. Sci. USA*, **107**, 11 217–11 222, https://doi.org/10.1073/pnas.0910818107.

——, and Coauthors, 2015: Integrating laboratory and field data to quantify the immersion freezing ice nucleation activity of mineral dust particles. *Atmos. Chem. Phys.*, **15**, 393–409, https://doi.org/10.5194/acp-15-393-2015.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Eaton, B., 2011: User's guide to the Community Atmosphere Model CAM-5.1. NCAR Doc., 40 pp., http://www.cesm.ucar.edu/models/cesm1.0/cam.

Fujita, T., D. J. Stensrud, and D. C. Dowell, 2007: Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physics uncertainties. *Mon. Wea. Rev.*, **135**, 1846–1868, https://doi.org/10.1175/MWR3391.1.

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. *J. Atmos. Sci.*, **62**, 3520–3534, https://doi.org/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. *J. Atmos. Sci.*, **76**, 3411–3434, https://doi.org/10.1175/JAS-D-19-0004.1.

Georgii, H. W., and E. Kleinjung, 1967: Relations between the chemical composition of atmospheric aerosol particles and the concentration of natural ice nuclei. *J. Rech. Atmos.*, **3**, 145–156.

Gerber, 2003: Particulate volume monitor model PVM-100A. Gerber Scientific, accessed 5 May 2020, http://www.gerberscience.com/pvmaspecs.html.

Golding, B., N. Roberts, G. Leoncini, K. Mylne, and R. Swinbank, 2016: MOGREPS-UK convection-permitting ensemble products for surface water flood forecasting: Rationale and first results. *J. Hydrometeor.*, **17**, 1383–1406, https://doi.org/10.1175/JHM-D-15-0083.1.

Hamill, T. M., 2001: Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Wea. Rev.*, **129**, 550–560, https://doi.org/10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2.

Harrington, J. Y., K. Sulia, and H. Morrison, 2013a: A method for adaptive habit prediction in bulk microphysical models. Part I: Theoretical development. *J. Atmos. Sci.*, **70**, 349–364, https://doi.org/10.1175/JAS-D-12-040.1.

——, ——, and ——, 2013b: A method for adaptive habit prediction in bulk microphysical models. Part II: Parcel model corroboration. *J. Atmos. Sci.*, **70**, 365–376, https://doi.org/10.1175/JAS-D-12-0152.1.

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. *J. Open Res. Software*, **4**, e25, https://doi.org/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**, 129–151.

——, and Y. Noh, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, https://doi.org/10.1175/MWR3199.1.

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. *J. Geophys. Res.*, **113**, D13103, https://doi.org/10.1029/2008JD009944.

Imran, H. M., J. Kara, A. W. M. Ng, and S. Muthukumaran, 2018: An evaluation of the performance of a WRF multi-physics ensemble for heatwave events over the city of Melbourne in southeast Australia. *Climate Dyn.*, **50**, 2553–2586, https://doi.org/10.1007/s00382-017-3758-y.

Jaffrain, J., and A. Berne, 2011: Experimental quantification of the sampling uncertainty associated with measurements from PARSIVEL disdrometers. *J. Hydrometeor.*, **12**, 352–370, https://doi.org/10.1175/2010JHM1244.1.

Jankov, I., and Coauthors, 2017: A performance comparison between multiphysics and stochastic approaches within a North American RAP ensemble. *Mon. Wea. Rev.*, **145**, 1161–1179, https://doi.org/10.1175/MWR-D-16-0160.1.

Jensen, A. A., J. Y. Harrington, and H. Morrison, 2018: Microphysical characteristics of squall-line stratiform precipitation and transition zones simulated using an ice particle property-evolving model. *Mon. Wea. Rev.*, **146**, 723–743, https://doi.org/10.1175/MWR-D-17-0215.1.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181, https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Kanji, Z. A., L. A. Ladino, H. Wex, Y. Boose, M. Burkert-Kohn, D. J. Cziczo, and M. Kramer, 2017: Overview of ice nucleating particles. *Ice Formation and Evolution in Clouds and Precipitation: Measurement and Modeling Challenges*, *Meteor. Monogr.*, No. 58, Amer. Meteor. Soc., https://doi.org/10.1175/AMSMONOGRAPHS-D-16-0006.1.

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. *Bull. Amer. Meteor. Soc.*, **98**, 315–332, https://doi.org/10.1175/BAMS-D-15-00034.1.

Ladwig, W., 2017: WRF-Python Version 1.3.2. UCAR/NCAR, https://doi.org/10.5065/D6W094P1.

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. *Mon. Wea. Rev.*, **138**, 1587–1612, https://doi.org/10.1175/2009MWR2968.1.

Liu, C., K. Ikeda, G. Thompson, R. Rasmussen, and J. Dudhia, 2011: High-resolution simulations of wintertime precipitation in the Colorado headwaters region: Sensitivity to physics parameterizations. *Mon. Wea. Rev.*, **139**, 3533–3553, https://doi.org/10.1175/MWR-D-11-00009.1.

Liu, J. Y., and H. D. Orville, 1969: Numerical modeling of precipitation and cloud shadow effects on mountain-induced cumuli. *J. Atmos. Sci.*, **26**, 1283–1298, https://doi.org/10.1175/1520-0469(1969)026<1283:NMOPAC>2.0.CO;2.

Maahn, M., and P. Kollias, 2012: Improved Micro Rain Radar snow measurements using Doppler spectra post-processing. *Atmos. Meas. Tech.*, **5**, 2661–2673, https://doi.org/10.5194/amt-5-2661-2012.

McMillen, J. D., and W. J. Steenburgh, 2015: Impact of microphysics parameterizations on simulations of the 27 October 2010 Great Salt Lake–effect snowstorm. *Wea. Forecasting*, **30**, 136–152, https://doi.org/10.1175/WAF-D-14-00060.1.

Meng, Z., and F. Zhang, 2007: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Wea. Rev.*, **135**, 1403–1423, https://doi.org/10.1175/MWR3352.1.

Meyers, M. P., P. J. DeMott, and W. R. Cotton, 1992: New primary ice-nucleation parameterizations in an explicit cloud model. *J. Appl. Meteor.*, **31**, 708–721, https://doi.org/10.1175/1520-0450(1992)031<0708:NPINPI>2.0.CO;2.

Milbrandt, J. A., and M. K. Yau, 2005a: A multimoment bulk microphysics parameterization. Part I: Analysis of the role of the spectral shape parameter. *J. Atmos. Sci.*, **62**, 3051–3064, https://doi.org/10.1175/JAS3534.1.

——, and ——, 2005b: A multimoment bulk microphysics parameterization. Part II: A proposed three-moment closure and scheme description. *J. Atmos. Sci.*, **62**, 3065–3081, https://doi.org/10.1175/JAS3535.1.

——, and H. Morrison, 2013: Prediction of graupel density in a bulk microphysics scheme. *J. Atmos. Sci.*, **70**, 410–429, https://doi.org/10.1175/JAS-D-12-0204.1.

Minder, J. R., T. W. Letcher, L. S. Campbell, P. G. Veals, and W. J. Steenburgh, 2015: The evolution of lake-effect convection during landfall and orographic uplift as observed by profiling radars. *Mon. Wea. Rev.*, **143**, 4422–4442, https://doi.org/10.1175/MWR-D-15-0117.1.

Mitchell, D. L., and A. J. Heymsfield, 2005: Refinements in the treatment of ice particle terminal velocities, highlighting aggregates. *J. Atmos. Sci.*, **62**, 1637–1644, https://doi.org/10.1175/JAS3413.1.

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. *Mon. Wea. Rev.*, **137**, 991–1007, https://doi.org/10.1175/2008MWR2556.1.

——, J. A. Milbrandt, G. H. Bryan, K. Ikeda, S. A. Tessendorf, and G. Thompson, 2015: Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. Part II: Case study comparisons with observations and other schemes. *J. Atmos. Sci.*, **72**, 312–339, https://doi.org/10.1175/JAS-D-14-0066.1.

National Weather Service, 2020: Stage III daily accumulations. Advanced Hydrologic Prediction Service, accessed 8 December 2017, https://water.weather.gov/precip/download.php.

NOAA, 2014: NOAA Next Generation Radar (NEXRAD) level 2 base data. Subset used: 15 December 2013–16 December 2013. NOAA National Centers for Environmental Information, accessed 27 July 2017, https://doi.org/10.7289/V5W9574V.

Phillips, D., and K. Knupp, 2014a: UAH MIPS PARSIVEL disdrometer data, version 1.0. UCAR/NCAR Earth Observing Laboratory, accessed 13 September 2019, https://doi.org/10.26023/33H3-H9EE-7M0F.

——, and ——, 2014b: UAH MIPS X-band profiling radar (XPR) data, version 1.0. UCAR/NCAR Earth Observing

Laboratory, accessed 28 September 2019, https://doi.org/10.26023/7RWC-FC33-M50C.

Przybylo, V. M., K. J. Sulia, C. G. Schmitt, Z. J. Lebo, and W. C. May, 2019: The Ice Particle and Aggregate Simulator (IPAS). Part I: Extracting dimensional properties of ice-ice aggregates for microphysical parameterization. *J. Atmos. Sci.*, **76**, 1661–1676, https://doi.org/10.1175/JAS-D-18-0187.1.

Reeves, H. D., and D. T. Dawson, 2013: The dependence of QPF on the choice of microphysical parameterization for lake-effect snowstorms. *J. Appl. Meteor. Climatol.*, **52**, 363–377, https://doi.org/10.1175/JAMC-D-12-019.1.

Ryzhkov, A., M. Pinsky, A. Pokrovsky, and A. Khain, 2011: Polarimetric radar observation operator for a cloud model with spectral microphysics. *J. Appl. Meteor. Climatol.*, **50**, 873–894, https://doi.org/10.1175/2010JAMC2363.1.

Saslo, S., and S. J. Greybush, 2017: Prediction of lake-effect snow using convection-allowing ensemble forecasts and regional data assimilation. *Wea. Forecasting*, **32**, 1727–1744, https://doi.org/10.1175/WAF-D-16-0206.1.

Steenburgh, J., L. Campbell, and P. Veals, 2014a: North Redfield snow study station data, version 1.0. UCAR/NCAR Earth Observing Laboratory, accessed 17 October 2019, https://doi.org/10.5065/D68051CN.

——, ——, and ——, 2014b: Sandy Creek snow study station data, version 1.0. UCAR/NCAR Earth Observing Laboratory, accessed 17 October 2019, https://doi.org/10.5065/D68K77WG.

——, ——, ——, J. Minder, and T. Letcher, 2014c: Micro Rain Radar post processed data, version 1.0. UCAR/NCAR Earth Observing Laboratory, accessed 12 September 2019, https://doi.org/10.26023/E60X-0SEH-TT0Q.

Stensrud, D. J., J.-W. Bao, and T. T. Warner, 2000: Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 2077–2107, https://doi.org/10.1175/1520-0493(2000)128<2077:UICAMP>2.0.CO;2.

Straka, J. M., D. S. Zrnić, and A. V. Ryzhkov, 2020: Bulk hydrometeor classification and quantification using polarimetric radar data: Synthesis of relations. *J. Appl. Meteor.*, **39**, 1341–1372, https://doi.org/10.1175/1520-0450(2000)039<1341:BHCAQU>2.0.CO;2.

Sulia, K. J., and M. R. Kumjian, 2017a: Simulated polarimetric fields of ice vapor growth using the adaptive habit model. Part I: Large-eddy simulations. *Mon. Wea. Rev.*, **145**, 2281–2302, https://doi.org/10.1175/MWR-D-16-0061.1.

——, and ——, 2017b: Simulated polarimetric fields of ice vapor growth using the adaptive habit model. Part II: A case study from the FROST experiment. *Mon. Wea. Rev.*, **145**, 2303–2323, https://doi.org/10.1175/MWR-D-16-0062.1.

——, H. Morrison, and J. Y. Harrington, 2014: Dynamical and microphysical evolution during mixed-phase cloud glaciation simulated using the bulk adaptive habit prediction model. *J. Atmos. Sci.*, **71**, 4158–4180, https://doi.org/10.1175/JAS-D-14-0070.1.

——, Z. J. Lebo, V. M. Przybylo, and C. G. Schmitt, 2020: A new method for ice–ice aggregation in the adaptive habit model. *J. Atmos. Sci.*, **78**, 133–154, https://doi.org/10.1175/JAS-D-20-0020.1.

Tao, W.-K., J. Simpson, and M. McCumber, 1989: An ice–water saturation adjustment. *Mon. Wea. Rev.*, **117**, 231–235, https://doi.org/10.1175/1520-0493(1989)117<0231:AIWSA>2.0.CO;2.

——, 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. *J. Geophys. Res. Atmos.*, **121**, 1278–1305, https://doi.org/10.1002/2015JD023986.

Thompson, G., 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. *Mon. Wea. Rev.*, **136**, 5095–5115, https://doi.org/10.1175/2008MWR2387.1.

Tsai, T., and J. Chen, 2020: Multimoment ice bulk microphysics scheme with consideration for particle shape and apparent density. Part I: Methodology and idealized simulation. *J. Atmos. Sci.*, **77**, 1821–1850, https://doi.org/10.1175/JAS-D-19-0125.1.

University of Wyoming Flight Center, 1977: King Air 1 Hz files. Subset used: OWLeS IOP4. University of Wyoming Dept. of Atmospheric Science, accessed 29 October 2019, http://flights.uwyo.edu/projects/owles13/order1hz.shtml.

Whitby, K. T., 1978: The physical characteristics of sulfur aerosols. *Atmos. Environ.*, **12**, 135–159, https://doi.org/10.1016/0004-6981(78)90196-8.

Wurman, J., 2001: The DOW mobile multiple Doppler network. Preprints, *30th Int. Conf. on Radar Meteorology*, Munich, Germany, Amer. Meteor. Soc., 95–97.

Yang, L., S. Wang, J. Tang, X. Niu, and C. Fu, 2019: Evaluation of the effects of a multiphysics ensemble on the simulation of an extremely hot summer in 2003 over the CORDEX-EA-II region. *Int. J. Climatol.*, **39**, 3413–3430, https://doi.org/10.1002/joc.6028.

Yuter, S. E., and R. A. Houze Jr., 1995: Three-dimensional kinematic and microphysical evolution of Florida cumulonimbus. Part II: Frequency distributions of vertical velocity, reflectivity, and differential reflectivity. *Mon. Wea. Rev.*, **123**, 1941–1963, https://doi.org/10.1175/1520-0493(1995)123<1941:TDKAME>2.0.CO;2.

——, D. E. Kingsmill, L. B. Nance, and M. Loffler-Mang, 2006: Observations of precipitation size and fall speed characteristics within coexisting rain and wet snow. *J. Appl. Meteor. Climatol.*, **45**, 1450–1464, https://doi.org/10.1175/JAM2406.1.