Transportation Research Interdisciplinary Perspectives 8 (2020) 100253

Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

journal homepage: www.elsevier.com/locate/trip

Impacts of assimilating observations from connected vehicles into a numerical weather prediction model

Amanda Siems-Anderson *, Jared A. Lee, Barbara Brown, Gerry Wiener, Seth Linden

Research Applications Laboratory, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307, USA

ARTICLE INFO

Keywords:
Road weather
Numerical weather prediction
Connected vehicles

ABSTRACT

Numerical weather prediction (NWP), a method of weather forecasting using equations that describe atmospheric flows and behavior, is used extensively to help predict weather conditions along roadways and has important implications for maintaining safety and efficiency on the transportation network. An insufficiently dense surface observation network presents a major limitation in obtaining the current state of the atmosphere for NWP at high spatial and temporal resolutions, such as required by road weather applications. Connected vehicle technologies, where public, private, and commercial vehicles serve as weather-observing platforms, can be used to fill in these gaps in the surface weather observation network.

A pilot study was formed to quantify the impact of a dense network of observations along roadways, such as would exist with fully implemented connected vehicle technologies, on NWP. First, a simulated vehicle probe dataset was created. This dataset was assimilated into the Weather Research and Forecasting (WRF) model for select case studies, and the resulting output compared against observations and a baseline WRF run without the vehicle data assimilation. It was found that these observations had an overall positive impact on precipitation and other surface variable outputs, though in several cases the impact was slight. Major improvements occurred when using wiper status as a proxy for precipitation to apply forward error correction to a point model forecast of probability of precipitation and quantitative precipitation forecasts.

1. Introduction

During the last several decades, the atmospheric science community has committed significant resources to improving numerical weather prediction (NWP), including the process of data assimilation (DA; e.g., [71,9,21,3,80]). DA is the process by which observations are used to correct the initial state of the model to a more accurate representation of the true weather conditions. Modern NWP depends on the availability of a large number of observations to properly represent current conditions and initialize the model predictions. Such models are useful for a variety of applications, including standard weather forecasting [30], renewable energy [43], wildfire prediction and response [14], and surface transportation safety and mobility [67]. Forecasts of road-specific weather conditions are used throughout the transportation sector [57], and such forecasts are critical for operations including staff management and timing of plowing operations [79].

Proper initial conditions for the state of the atmosphere at the start of any NWP model run are critical to ensure accurate, useful model output, and atmospheric observations can be used to improve this initial condition via DA techniques. As NWP models move toward finer and finer resolution in an effort to improve forecast usability and skill [44,65], the current state of the atmosphere must also be sampled at higher resolutions to properly resolve features at the same scale as the model [38]. While significant strides have been made to increase coverage, both in surface observing density (e.g., [64,47]) and via remote sensing techniques (e.g., [16,83]), the current observing network is far coarser than high-resolution NWP models with 1–4-km horizontal grid spacing, particularly in more rural areas.

Application of high-resolution surface observations for hazardous weather assessment and prediction along roadways has also been an active area of research [42,41,19]. Hazardous weather conditions contribute to 1.2 million crashes annually in the U.S., with 5,376 fatalities and 418,005 injuries [7]. These totals account for 21% of all crashes. This is far higher than the percent of crashes that would be expected to occur in hazardous weather conditions by chance alone (less than 10%; [35]). Economic impacts are severe as well, with $8.659 billion in annual economic losses to trucking companies alone as a result of weather-related delays [35].

To provide road-specific observations such as pavement temperature, Road Weather Information System (RWIS) Environmental Sensor Stations (ESS) have been implemented in most U.S. states [59]. These

* Corresponding author.
E-mail address: aander@ucar.edu (A. Siems-Anderson).

https://doi.org/10.1016/j.trip.2020.100253
Received 10 April 2020; Revised 7 August 2020; Accepted 26 October 2020

2590-1982/© 2020 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

stations provide critical data about weather conditions at the road location, along with road-specific observations like pavement condition (e.g., wet or slick roads) necessary for maintaining roadway safety and mobility [74]. However, conditions may change rapidly between stations, which are often tens of kilometers apart, and many localized conditions such as patchy black ice or dense fog can exist between stations. The weather data that could be collected by connected vehicle (CV) technologies [78] from the millions of vehicles on the roadways may be used to fill the spatial and temporal gaps between fixed weather stations. This can impact not only road weather information and forecasting, but also larger-scale applications such as NWP [42]. Once implemented outside transportation agencies and in the general consumer vehicle fleet, there is a potential for millions of new observations along the nation's roadways. Anderson et al. [2] showed that air temperatures from vehicles' built-in sensors can be of useful quality for meteorological applications, and other important atmospheric variables such as dewpoint temperature and air pressure may be obtained from vehicles as well with externally-mounted sensors. Additionally, non-traditional observations such as windshield wiper speed are used in road weather applications as a proxy for weather observations [66], and these proxies may be useful for NWP applications if properly converted into the types of observations accepted by DA systems.

The usability of crowd-sourced observations is being investigated across a variety of applications, including meteorological. Examples include Mass and Madaus [45] and Madaus and Mass [40], who explored the use of smartphone surface pressure observations in NWP, and Muller et al. [53], who used a number of crowd-sourced observing platforms to study urban heat island effects. Like other non-traditional data source studies, this paper focuses on the use of CV data outside the typical transportation applications and into a wider weather context. Specifically, this pilot study generated CV observations and assimilated them into an NWP model for five case studies over Minnesota and Michigan for a variety of weather conditions. Representing a first step in applying CV data in NWP, the results of the case studies are used to investigate the potential utility of these data and identify future efforts needed to fully explore and refine the use of this upcoming source of high-resolution surface observations.

2. Data and methods

2.1. Cases

A case study approach was used for this initial investigation, with the idea that more extensive studies can be designed based on the results of this pilot study. Cases were chosen to cover a variety of weather and observation scenarios to meet the goal of examining what, if any, benefits to forecast quality are obtained by assimilating vehicle data into NWP models. The study approach included five case days, each chosen to represent a different weather regime. These regimes were: heavy rain/flooding, heavy snow, light snow, no precipitation with some fog, and mixed-phase precipitation (rain changing to snow).

For each case, a representative 24-hour period was chosen with an additional 24 h prior to the start of each period used for model spin-up. The two model domains covered (i) the state of Minnesota and (ii) the state of Michigan. These regions were selected to leverage prior work done as part of the Integrated Mobile Observations (IMO) project [12,10]. The details of each case can be found in Table 1.

2.2. Vehicle data translator simulations

While several projects are underway to integrate CV into current transportation operations (e.g., [24]), CV technology is limited in deployment, and thus the amount of vehicle data available for case analysis remains very small [28]. Assimilating currently available vehicle data, which would be received from approximately 20 to 200 unique vehicles per day [12], would neither have a likely impact on the model nor be representative of the usefulness of a fully deployed CV system. Because the necessary volume of observations to assess any potential impacts, even in a pilot study, are not currently available, it was necessary to simulate the CV data based on neighboring atmospheric observations. To produce such a dataset that is large enough to assess the impacts of CV data from a fully deployed system, a vehicle data simulator was used. This approach is similar to an Observation System Simulation Experiment (OSSE; [4]) in that the vehicle observation system was simulated. However, this study differed from an OSSE because the observations were simulated from true observations of the atmospheric state on case days, rather than from output of another NWP model run (i.e., a "nature run"). OSSEs are commonly used in atmospheric science research [1,84] to evaluate the usefulness of a proposed observing network or platform (e.g., a new weather-observing satellite) before it is deployed and observations are available (e.g., [39,13,36,29]).

Vehicle observations were generated using a simulator that was built and is available as part of the Pikalert® system's Vehicle Data Translator (VDT; [66,67]). The user specifies which observation sets to use (radar, background model, surface station observations, etc.), the update cycle, a road segment configuration file, and the number of vehicle observations per road segment. The number of vehicle observations per segment can be varied by road type, geographic location, and time of day. Each road segment is assigned an observation based on the input data available within a time window and radius around the segment using a nearest neighbor approach. Once mean values of observations are assigned to the segment, the specified numbers of vehicle observations are created for each observation type using a random selection of values from a Gaussian distribution about the mean. This method was used to simulate a realistic distribution of different values that may be measured by different vehicles due to inconsistency in instrumentation types and lack of instrument maintenance.

For this experiment, gridded analysis data from the Real-Time Mesoscale Analysis (RTMA; [18]) and surface station data (i.e., Automated Surface Observation System [ASOS; [55]] stations and RWIS ESS) were used to simulate vehicle observations during the 24-hour spin up period and the 24-hour simulation period for each case study. A gridded radar product from the Multi-Radar Multi-Sensor (MRMS) product suite [85] was used to simulate vehicle wiper status by using radar reflectivity ranges consistent with different wiper speed categories (off, intermittent, low, high) used in the Pikalert system's Road Weather Hazard (RWH) module [66].

A road segment Network Common Data Format (NetCDF; [61]) file was generated for both Minnesota and Michigan using Geographic Information System (GIS) shape files obtained from the Minnesota Department of Transportation (MnDOT; [51]) and the Michigan Department of Technology, Management, & Budget [48], respectively. Two files for each state were created: a less dense case that included interstates and U.S. highways, and a more dense case that included interstates, U.S. highways, and state highways.

To represent the variability among geographic areas, the number of vehicle observations varied by urban area and time of day. A list of urban areas with populations over 50,000 was generated for each state based on the 2010 U.S. Census [77]. Any city in qualitatively close proximity to another on the list was combined into a metro area. For example, cities such as Warren, Sterling Heights, Dearborn, and Livonia, MI, are all close suburbs of the city of Detroit. These and other suburbs with populations over 50,000 were combined into one Detroit Metro Area for the purpose of the simulations. Once the cities were identified and metro areas grouped, the area of each was defined subjectively from a map based on a center latitude/longitude of the city or metro along with a radius about this center that encompassed most of

Table 1
Case descriptions.

Case	Date (UTC)	Case Type	State	Case Rationale	Description
1	0500 31 May – 0500 1 Jun 2014	Rain/ Flooding	MN	Benefits of mobile data on a high impact rain event	Waves of thunderstorms moved through the domain, leading to high rain totals and flooding
2	0000 6 Apr – 0000 7 Apr 2014	Dry/Fog	MN	Benefits of mobile data on a mostly dry case/low visibility	Ridging with seasonable temperature and dry conditions; a few areas of fog briefly in the early morning hours
3	1200 31 Mar – 1200 1 Apr 2014	Mixed Phase	MN	Benefits of mobile data on a rain/snow transition event	A low pressure system with accompanying cold front forced rain switching to snow; wind speeds around 25 to 35 knots
4	0500 5 Jan – 0500 6 Jan 2014	Heavy Snow	MI	Benefits of mobile data on a high impact snow event	A cold front forced widespread moderate to heavy snow across the domain with temperatures dropping to −10 °C; wind speed around 15 to 20 knots
5	0000 25 Feb – 0000 26 Feb 2014	Light Snow	MI	Benefits of mobile data on a low precipitation case, but which may have a high impact on the roadway	Spotty light snow forced by a shortwave trough

the city/metro area. Any area of a state not within these city/metro areas was defined as "rural".

Annual average traffic density (AATD) measurements were used to determine the maximum number of vehicles that may be expected along a five-mile segment of interstate in each city/metro during peak travel times. These measurements were obtained from MnDOT [52] and the Michigan DOT (MDOT; [49]). However, CVs cannot be assumed to make up even the majority of vehicles on the road any time in the near future. Goodall [25] found that if a federal mandate were issued for inclusion of CV technology in all newly manufactured vehicles, then the technology penetration would be expected to be around 5% in the first year and around 30% after 6 to 7 years. These values (5% and 30%) were used for the less dense and more dense simulations, respectively. For example, if a given road segment was expected to have 20 vehicles travel over it in a given 5-minute period based on AATD, then 1 vehicle observation would be produced for the less dense simulation and 6 vehicle observations would be produced for the more dense simulation. Table 2 lists the cities used in the simulations and their associated centers/radii, along with the number of vehicle observations per five-mile road segment for both the less dense and more dense cases.

Along with location and technology penetration, the density of vehicle observations would vary considerably by type of road (e.g., interstate vs. state highway) and time of day. To account for this, hourly data from each road type were obtained from MnDOT for each urban area and representative rural areas. Distributions of hourly traffic rates were used to determine the hour of day with the highest density of vehicles. After this, approximate percentage reductions were calculated for each off hour. To account for the time zone difference between Michigan and Minnesota, the percentages used for Michigan were offset one hour from Minnesota (in UTC). The hourly percentages are shown in Fig. 1. The AATD was also used to compare the number of vehicles on different road types (interstate, U.S. highway, and state highway). Overall, interstates had the highest traffic counts, followed by U.S. highways, and state highways had the lowest traffic counts. U.S. and state highway traffic counts were compared against interstate traffic counts in order to set the number of vehicles relative to the interstates. They were set to 70% and 35% of the traffic counts on interstates, respectively.

The number of vehicle observations located in each city/metro area in Table 2 was assigned to a five-minute period for the interstate roadway at peak traffic time (2200 UTC for Minnesota and 2100 UTC for Michigan) at that location; both 5% (less dense) and 30% (more dense) vehicle penetration values were assigned. From there, the road type and time of day determined the final number of vehicle observations simulated per road segment. For example, the Minneapolis-St. Paul Metro was assigned 14 vehicles on an interstate at the peak travel hour for the less dense case. The number of vehicle observations at 0300 UTC on a U.S. highway in this metro area would be vehicles*hour*road, or 14*0.5*0.7 = 4.9, rounded up to 5 vehicle observations simulated.

After the vehicle simulations were complete, a median value was taken for each road segment. This process results in values that would be available from Pikalert VDT output and gives a more manageable number of meaningful values rather than providing thousands of sub-grid resolution observations. Simulating a varying number of vehicles on each segment for different areas and times provided realistic variability and error to the road segment median values.

It is important to note that the simulations of the number of vehicles per road segment represents the ideal distribution of vehicles on the roadway – a homogeneous one. In a fully deployed CV network, it is likely CV-equipped cars will not be homogeneously distributed.

Table 2
List of urban areas included in the vehicle data simulation with city center and area radius definitions along with the number of vehicle observations generated on an interstate road segment at peak traffic density hour for the less dense and more dense case types.

City	Center Lat/Lon	Radius	Less Dense	More Dense
Detroit Metro	42.45°N,83.19°W	30 km	24	144
Grand Rapids Metro	42.94°N,85.68°W	7 km	12	74
Lansing Metro	42.71°N,84.55°W	7 km	5	30
Ann Arbor	42.27°N,83.73°W	5 km	12	73
Ypsilanti	42.24°N,83.62°W	3 km	16	96
Flint	43.01°N,83.66°W	6 km	7	44
Kalamazoo	42.27°N,85.59°W	5 km	6	35
Battle Creek	42.34°N,85.18°W	4.5 km	8	46
Saginaw	43.43°N,83.94°W	4.5 km	9	53
Minneapolis-St. Paul Metro	44.97°N,93.22°W	30 km	14	83
Rochester	44.02°N,92.47°W	6 km	5	31
Duluth	46.79°N,92.10°W	8 km	7	40
St. Cloud	45.55°N,94.19°W	5 km	9	53
Rural	n/a	n/a	3	18

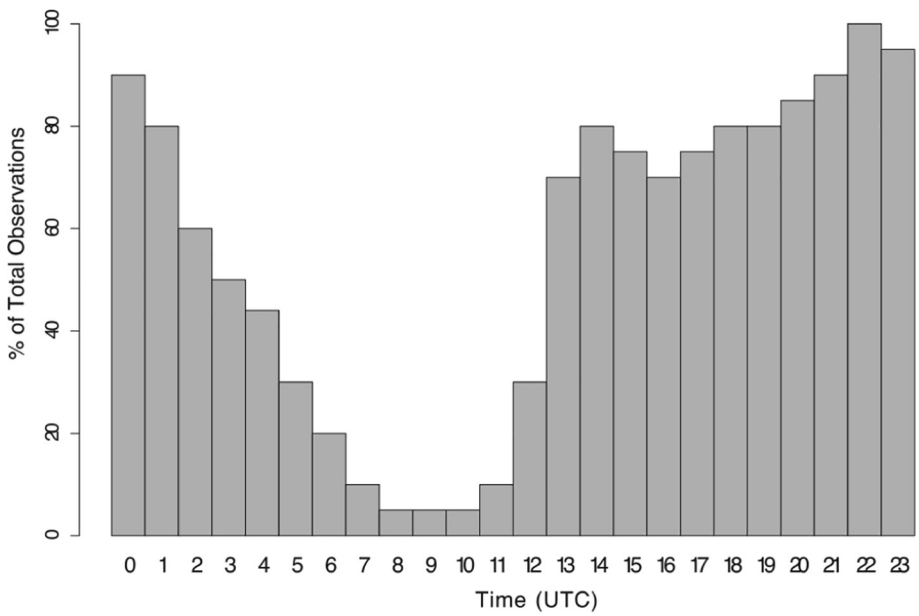

Fig. 1. Percentage of total observations by hour for Minnesota, normalized by the maximum hourly observations. Michigan percentages are shifted one hour earlier (to the left) due to local time zone difference.

Therefore, for the 5% technology penetration case, it is possible for two neighboring road segments to have a 10% and 1% technology penetration respectively, even if the overall technology penetration across the U.S. is 5%. While it was not possible to accurately estimate what this heterogeneity would look like in a future 5% or 30% penetration of CV-equipped vehicles for the cases presented in this pilot study, it is important to keep in mind that the homogeneous approach herein represents the highest impact scenario. Any heterogeneity is likely to reduce the number of road segments with a CV observation, which would in turn decrease the impact of this new dataset on NWP forecasts.

2.3. Numerical weather prediction models

2.3.1. The weather research and forecasting model

The Weather Research and Forecasting Advanced Research WRF (WRF-ARW) model version 3.6.1 [68,60] was chosen as the NWP model for this experiment. The two model domains (over Minnesota and Michigan) were each configured with a 250 × 250 grid point outer domain (Δx = 4 km) and a time step of 15 s, and a 481 × 481 grid point nested domain (Δx = 1.333 km) with a time step of 5 s. The inner domain was approximately centered in the parent domain and covered the entirety of Minnesota (with the exception of Flag Island in the northernmost part of the state) and the entirety of Michigan (Fig. 2). Both domains were configured with 51 vertical levels. The National Oceanic and Atmospheric Administration's (NOAA) Rapid Refresh (RAP) assimilation and model system [5] provided model initial conditions.

For the most part, the physics parameterization schemes used in the WRF model for these case studies were the same as those used in the RAP model at the time the WRF simulations were conducted. Specifically, both used the Thompson microphysics scheme [76], the Rapid Radiation Transfer Model Global (RRTMG) scheme for both longwave and shortwave radiation [32], and the Mellor-Yamada-Nakanishi-Niino (MYNN) planetary boundary layer and surface layer schemes [54]. For the cumulus (convection) parameterization scheme, the RAP uses the Grell-Dévényi scheme [26] while this study's WRF model used the updated Grell three-dimensional (3D) cumulus scheme [27]. The Grell 3D scheme allows for subsidence spreading in neighboring grid columns, which is a useful property in high-resolution simulations

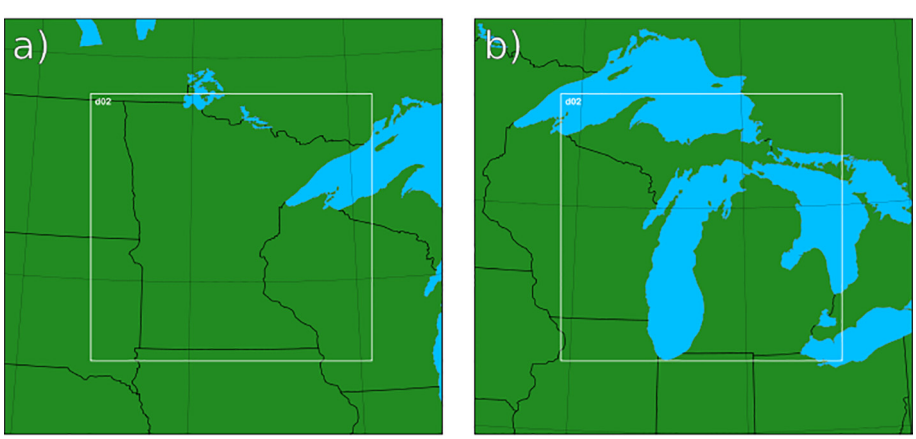

Fig. 2. Model domain for Cases 4 and 5 (left) and Cases 1, 2, and 3 (right). The outer domain is marked by the border of the image, and the inner domain (d02) is marked by a white border.

as subsidence typically occurs over a wider area than convective updrafts. The cumulus scheme was only used on the outer 4-km domain, while explicit convection was permitted on the inner 1.333-km domain where cumulus parameterization would not work well at this fine grid spacing [31]. For the land surface scheme, the RAP uses the Rapid Update Cycle (RUC) scheme [70] while the WRF model in this study used the Noah land surface model [75].

The Four-Dimensional Data Assimilation (FDDA) observation nudging scheme [71–73] was used to assimilate both surface station and simulated vehicle observations. FDDA first compares an observation with the model-predicted state (temperature, moisture/humidity, or wind). An artificial nudging term is then added onto the equations of motion to gradually nudge the model toward the observation state. For the observations assimilated, a time window half-width of 30 min and a radius of influence of 20 km for both domains were used, as recommended by W. Cheng of NCAR (personal communication, 7 Jul 2015). Observations were given a quality-check flag of 0–10 by the FDDA system, with 10 being the highest quality and values of 5–10 retained for assimilation.

The FDDA assimilated non-mobile observations for all case studies from the National Centers for Environmental Prediction (NCEP) Meteorological Assimilation Data Ingest System (MADIS; [50,46], https://madis.ncep.noaa.gov). This dataset includes standard surface, radiosonde, and aircraft observations, as well as satellite wind, profiler, marine buoy, and land mesonet observations. This configuration was defined as the baseline experiment for all the case studies, with the impacts of adding vehicle observations examined in subsequent runs.

After a 24-hour spin up period, FDDA analyses were created for the 24-hour period of interest. At hours 0, 6, 12, and 18 of the period of interest, 6-hour free forecasts using the WRF model were initialized from the WRF-FDDA analyses, meaning that at these four time periods, the WRF model was initialized using data assimilation but allowed to run out six hours without additional FDDA analyses.

2.3.2. The road weather forecast system

Some observations, such as precipitation (derived from wiper status), cannot easily be assimilated into the WRF model using the FDDA technique. More advanced techniques, such as three-dimensional variational analysis (3DVAR; [23]), have been used to assimilate radar reflectivity observations. However, assimilation of wiper status point observations converted into a radar reflectivity suitable for assimilation was outside the scope of this pilot project and is left for future research.

Although not able to be assimilated directly into the WRF model using existing DA methods, observations of wiper status may still be valuable for improving forecasts during the first several hours of a model run if used via a post-processing technique. A post-processing method currently used as part of the Pikalert system involves forward error correction (FEC) of output from the Road Weather Forecast System (RWFS) point-based consensus forecasting system [58,6].

The RWFS consists of two main engines: the Dynamic Integrated Forecast system (DICast®; [43]) and the Model of the Environment and Temperature of Roads (METRo; [15]). The atmospheric portion of the model works by ingesting output from multiple forecast models and using a series of weights based on model performance to produce a consensus forecast. This forecast can then be adjusted using observed data, such that the forecast is fit to the observation at hour 0, then the influence of the observation decreases through time until it decays completely. For example, if the consensus forecast for air temperature at lead time 0 is 15 °C, but observations at that location indicate the temperature is 10 °C, then the first few hours of the forecast are all adjusted colder to correct for this error.

This FEC method was used to analyze the potential usefulness of mobile observations that cannot readily be assimilated into the WRF model, specifically wiper status. RWFS forecasts from over Minnesota and Michigan during the IMO project were leveraged, and the simulated wiper status was used to FEC the RWFS point forecasts on each interstate segment (non-interstate segments were not available during IMO). Because the FEC was applied to real-time forecasts that had already used FEC on other variables such as air temperature, the contribution of the wiper status was isolated. The Quantitative Precipitation Forecast (QPF) FEC methodology was developed using the radar bounds from the vehicle data simulations and the default NWS Weather Surveillance Radar 1988 Doppler (WSR-88D) Z-R relationship of $Z = 300R^{1.4}$, where Z is radar reflectivity and R is rainfall rate [22]. This provided a mechanism for turning the wiper status of a vehicle into an approximate precipitation accumulation value. The probability of precipitation (PoP) methodology simply used a binary on/off of wipers and a degraded approach such that PoP for hours 0 and 1 was 100% if wipers were on, degrading by 25% through hour 3 and with no correction applied to hours 4 to 6.

2.4. Experiments and verification

The WRF model was run with assimilation for four different experiments. For all the experiments, air temperature, dewpoint temperature, and wind speed and direction observations were assimilated. Barometric pressure observations were used indirectly to convert the dewpoint temperature observations into a specific humidity value, as barometric pressure is not a state variable and thus cannot be assimilated with FDDA. The four experiments were as follows:

1. “Baseline” – Only MADIS observations were assimilated, with no vehicle observations
2. “VehObs Less” – Both MADIS observations and vehicle observations were assimilated, with the vehicle observations numbering in the “less dense” scenario (5% penetration, no state highways)
3. “VehObs More” – Both MADIS observations and vehicle observations were assimilated, with the vehicle observations numbering in the “more dense” scenario (30% penetration, including state highways)
4. “VehObs More Wind” – As with “VehObs More,” but with the addition of a theoretical wind speed and direction obtained from vehicle observations

The first experiment, the “Baseline,” included only those atmospheric observations that are currently available to NWP users to serve as a control run against which the experiments that included simulated vehicle observations could be compared.

For all four experiments, the four free forecasts defined in Section 2.3.1 were analyzed. A fifth experiment was also run, where two runs of the RWFS were compared: the real-time output of the RWFS for each case (“No Wipers” hereafter) and the RWFS output with FEC applied from vehicle wiper status (“Wipers” hereafter).

Two datasets served as the verification data for the model runs. For precipitation, the NCEP Stage IV product [37] was used. This is a CONUS-wide hourly analysis of Quantitative Precipitation Estimates (QPEs) produced by deriving a QPE from radar observations, then correcting the estimated precipitation totals with available rain gauge measurements. Stage IV is a standard QPE product for verification in the modeling community and is widely used (e.g., [17,82,11]). Although a standard for verification, the limitations of Stage IV associated with radar-based precipitation estimation uncertainties, errors near and below the freezing level, and quality control and processing discontinuities between regions must be kept in mind while interpreting Stage IV analyses [69,56] and thus the verification results based on them. Observations from the North American Model (NAM) Data Assimilation System (NDAS) Prepbufr data [63] were used for all other variables. The Prepbufr data include a variety of observations (e.g., from ASOS stations, upper-air soundings) and are quality checked before being disseminated.

Standard verification statistics [34,81] were produced using the Model Evaluation Tools (MET) software package, which is freely available from the Development Testbed Center (DTC; [8]). The WRF output was regridded to match the Stage IV grid, and bias-corrected root mean square error (BCRMSE), mean error (ME), and contingency table statistics (precipitation only) were calculated. The BCRMSE and ME were used because BCRMSE is independent of ME (and RMSE is not); thus, BCRMSE is the standard deviation of the errors and representative of the spread around the mean error, while ME represents the overall bias. Contingency table statistics (probability of detection, POD; false alarm ratio, FAR; critical success index, CSI; and bias) allow an analysis of the spatial prediction of precipitation regions compared to the observed precipitation areas. For the RWFS precipitation forecasts, the forecast was matched with the nearest Stage IV precipitation observation and Brier score, receiver operating characteristic (ROC) plots, and contingency table statistics were calculated. The Brier score and ROC plots were chosen for their ability to evaluate the performance of probability forecasts. To facilitate interpretation of the results, BCRMSE, ME, and Brier score were compared using the percent improvement over the "Baseline" relative to the perfect score (Eq. 1) of the different experiments' verification metrics (i.e., a skill measurement). The results here focus on the surface observations only, where surface-based vehicle data are expected to have the largest impact.

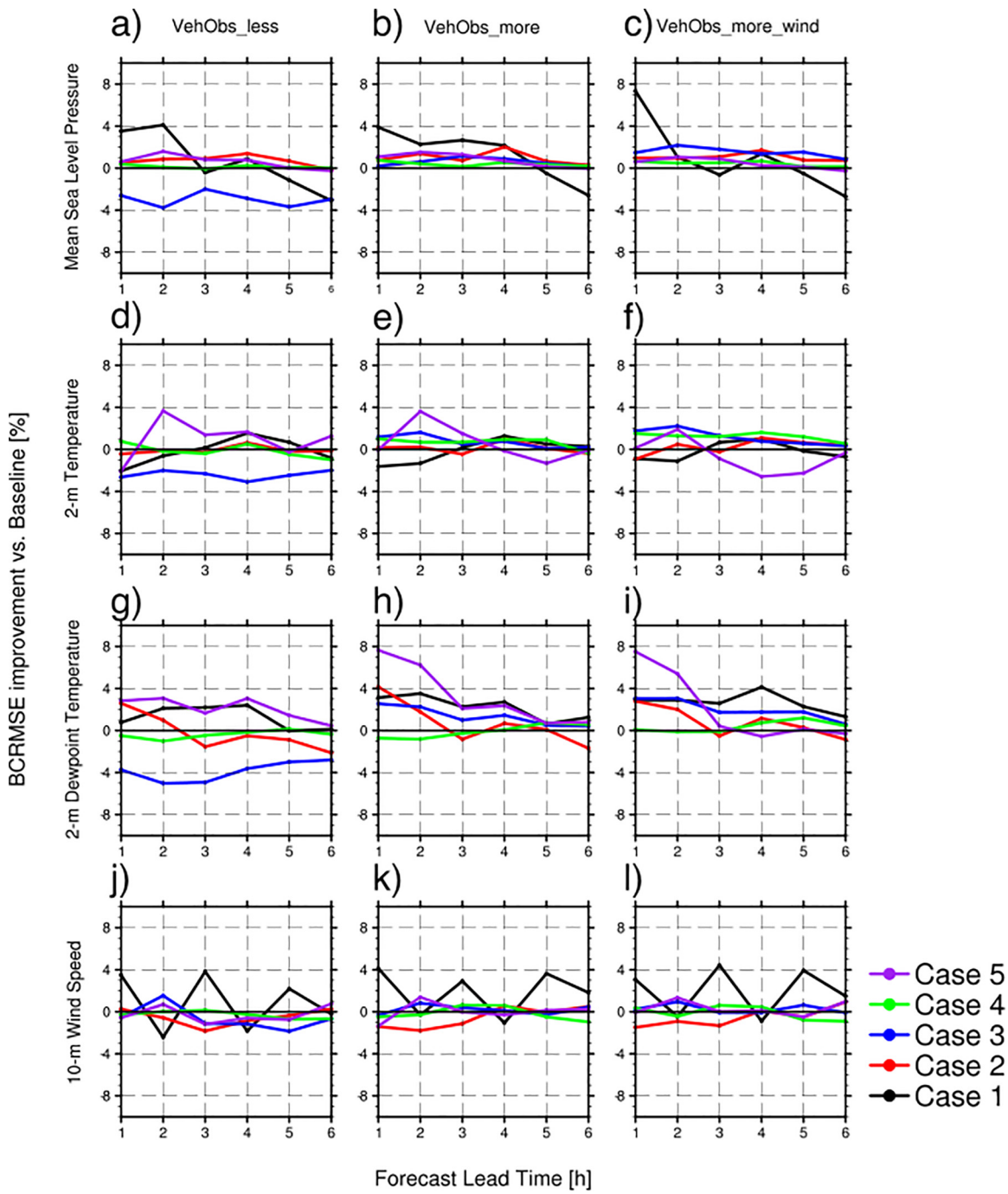

Fig. 3. Time series of percentage change in BCRMSE between the Baseline case and cases including vehicle data assimilation. A positive change indicates the BCRMSE is closer to 0, or improved performance, for the runs with vehicle data assimilation.

$$100 * \left(1 - \left[\frac{abs(ME_{model} - ME_{perfect})}{abs(ME_{baseline} - ME_{perfect})}\right]\right) \quad (1)$$

3. Results

Due to the small sample size (five cases), it is recognized that the below results are not statistically robust for all scenarios and there is a large amount of uncertainty in the results. However, it is hoped that the initial results from this pilot study will lead to larger confirmatory studies with representatively large samples. While no statistical conclusions can be reached with these results, they demonstrate the potential impacts of vehicle-based observation DA.

3.1. WRF atmospheric variables

For all cases, WRF forecasts of the 2-m air temperature, 2-m dewpoint temperature, mean sea level pressure (MSLP), and 10-m wind speed forecasts were verified against the NDAS Prepbufr observations.

In general, the BCRMSE (Fig. 3) was improved by 1%–3% for most of the cases, lead times, and data assimilation tests for each variable. The "VehObs More" (Fig. 3b, e, h, k) and "VehObs More Wind" tests (Fig. 3c, f, i, l), which had a 30% technology penetration versus 5% penetration for the "VehObs Less" test, performed better in comparison with the "Baseline" case for all four variables. This is most marked with Case 3, where the "VehObs Less" test had a performance decrease

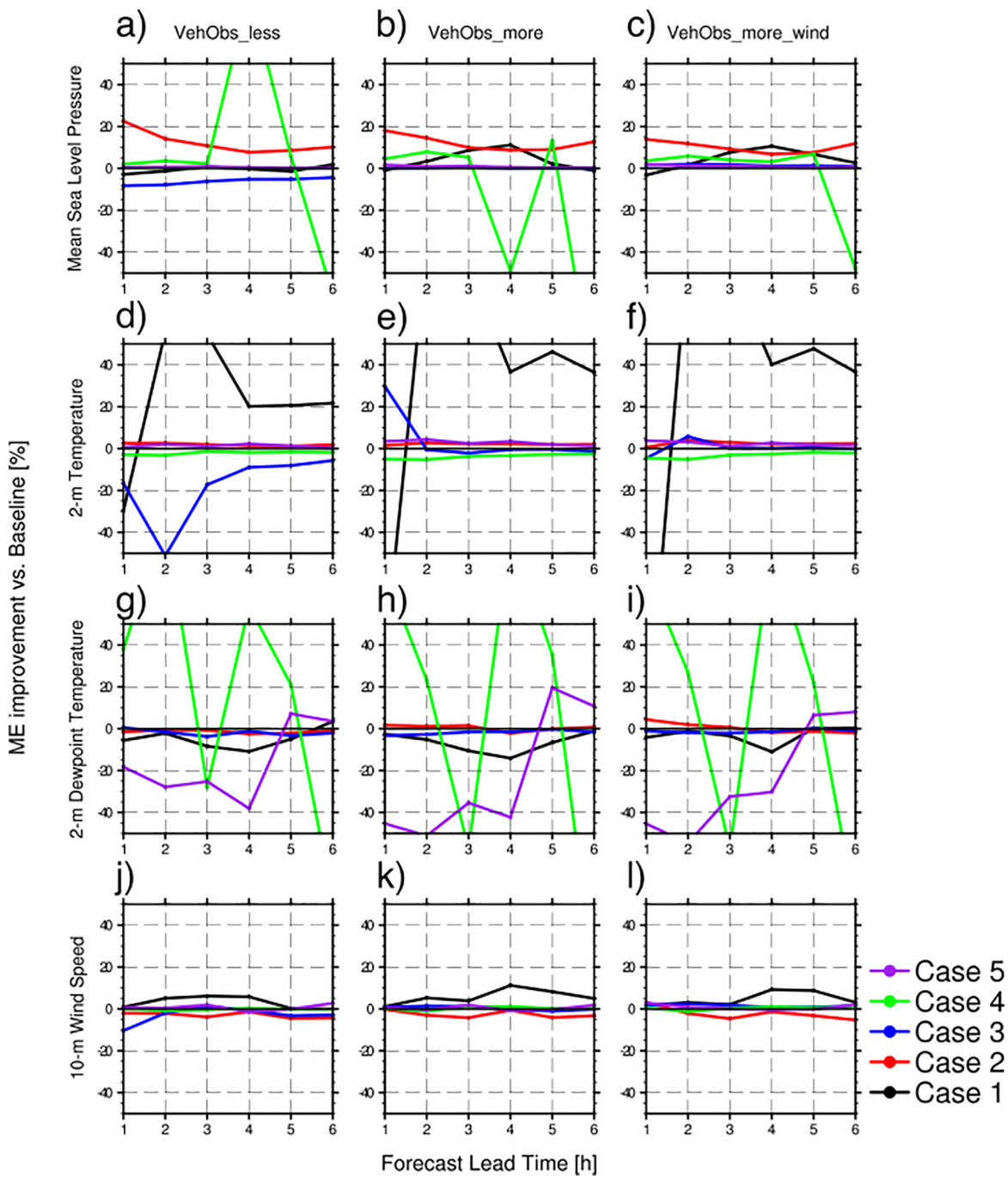

Fig. 4. Time series of percentage change in ME between the Baseline case and cases including vehicle data assimilation. A positive change indicates the ME is closer to 0, or improved performance, for the runs with vehicle data assimilation.

of 1%–4% in comparison to the "Baseline" case (Fig. 3a, d, g, j), but the two denser tests showed 1%–3% performance improvements (Fig. 3b, c, e, f, h, i, k, l). The performance improvements tended to be flat through time, meaning there was not a strong dependence on lead hour. However, a few tests did have performance decreasing with increasing lead hour. An example in Fig. 3 is Case 5 (purple time series), where the "VehObs More" and "VehObs More Wind" tests had a nearly 8% improvement over the "Baseline" at the first lead time for 2-m temperature (Fig. 3e, f), which decayed to 0%–1% by lead hours 5 and 3, respectively.

The results for ME are shown in Fig. 4. Overall, there was little difference in ME between the three assimilation types, and in most instances ME was improved at least a small amount over the "Baseline" in the vehicle data assimilation cases. Particularly large improvements can be seen for MSLP forecasts for Case 2 (Fig. 4a, b, c; red series), 2-m temperature for Case 2 (Fig. 4d, e, f; red series), and 10-m wind speed for Case 1 (Fig. 4j, k, l; black series). Decreased ME performance is evident for the "VehObs Less" for Case 3 for all observation types (Fig. 4a, d, g, j; blue series), and a large performance reduction through forecast hour 4 for 2-m dewpoint temperature for Case 5 (Fig. 4g, h, i; purple series). Percent improvement in ME is vulnerable to large, erratic values when the absolute value of the denominator (i.e., the ME for the baseline) is small, which explains the seemingly volatile time series for Case 1 for 2-m temperature (Fig. 4d, e, f; black series) and Case 4 for both MSLP (Fig. 4a, b, c; green series) and 2-m dewpoint (Fig. 4g, h, i; green series).

3.2. WRF QPF

Contingency table statistics and skill scores were used to assess the QPF produced with and without vehicle data assimilation. Although all cases were evaluated, for brevity only Case 1 (the heavy rain case) plots and results are presented here.

Performance diagrams [62] were used to visualize the contingency table verification measures. These plots take advantage of the mathematical relationships between probability of detection (POD), false alarm ratio (FAR), critical success index (CSI), and frequency bias (FB) to plot all four of these statistical measures in a single diagram. In general, points located in the upper right part of the diagram have the most skill while those in the lower left have the least skill. Points in the upper left have a high POD but also high FAR and FB, while points in the lower right have low POD, FAR, and FB. CSI values increase from the lower left to the upper right corners. At first glance (Fig. 5), there is little difference between the QPF verification statistics for the Baseline and other models with vehicle data assimilation, as most points are clustered together. While performance differed little between experiments, the "VehObs More" runs slightly outperformed the "Baseline" and "VehObs Less" runs, particularly in thresholds

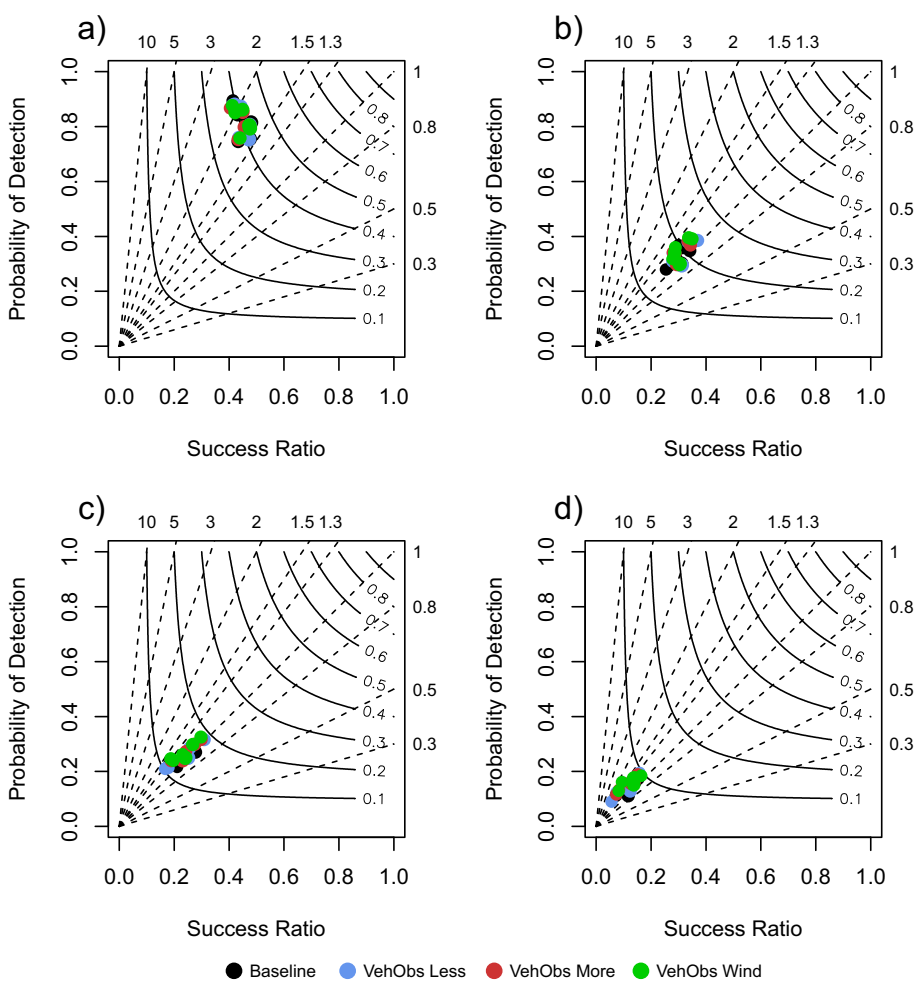

Fig. 5. Performance diagrams for Case 1 by precipitation thresholds: a) >0 mm, b) >1 mm, c) >2 mm, and d) >5 mm. Each point represents a different forecast lead through six hours, and points are scaled based on amount of active weather on a grid point (larger dot = fewer null case grid points at that lead hour).

>1 mm (Fig. 5, b–d; red and green points), at small values of about ~0.02 increase in CSI.

3.3. RWFS PoP and QPF

Percent improvements (Eq. 1) in Brier scores associated with the application of the FEC for a PoP threshold of >0 mm are shown in Fig. 6. For most cases and lead times, a large increase in Brier score performance for PoP forecasts was achieved over the "Baseline" RWFS forecasts. This result is most evident for Case 1, the heavy rain case, with as much as a 30% performance improvement. Results for three of the cases indicated a decrease in PoP performance for the analysis (zero) hour with the addition of connected vehicle data, but all cases were associated with improved performance by either the first or second lead hour. Due to the design of the algorithm, no differences in performance at lead hour four and beyond could be expected because the wiper status no longer impacted the model.

A receiver operating characteristic (ROC) plot (Fig. 7) also analyzes the performance of the probabilistic forecasts at various probability thresholds (for every 10% increment of probability value from 0 to 100%). For the majority of cases and thresholds, the FEC of PoP with wiper status improved the performance relative to the original RWFS forecast. The exception is Case 4, the heavy snow case, where the ROC lines are very closely aligned through a PoP of 70%, and PoP of 80% to 100% did not perform as well after FEC. The ROC results can be quantified by calculating the area under the ROC curve, shown in Table 3. The areas for Cases 1, 2, and 3 improved around 20% (consistent with the Brier scores), while the scores for Case 4 decreased by 69% (though this was only a 0.02 actual change in magnitude), and Case 5 improved only marginally, by 1.4%.

Precipitation accumulation forecasts (QPFs) were also evaluated, and a performance diagram comparing contingency table statistics between the original RWFS forecast and the FEC using wipers is shown in Fig. 8. The arrows connect the original RWFS run value (base of arrow) to the FEC using wipers (head of arrow). Arrows pointing to the right and/or to the top of the diagram indicate improvements in false alarms and detections of precipitation respectively. As was found in the PoP evaluation, the results for all cases except Case 4 indicate a substantial improvement in the QPF when FEC using wipers was applied to the RWFS output. This performance improvement held for all analyzed precipitation thresholds and is demonstrated by the movement of the arrows toward the upper right of the performance diagrams in Fig. 8.

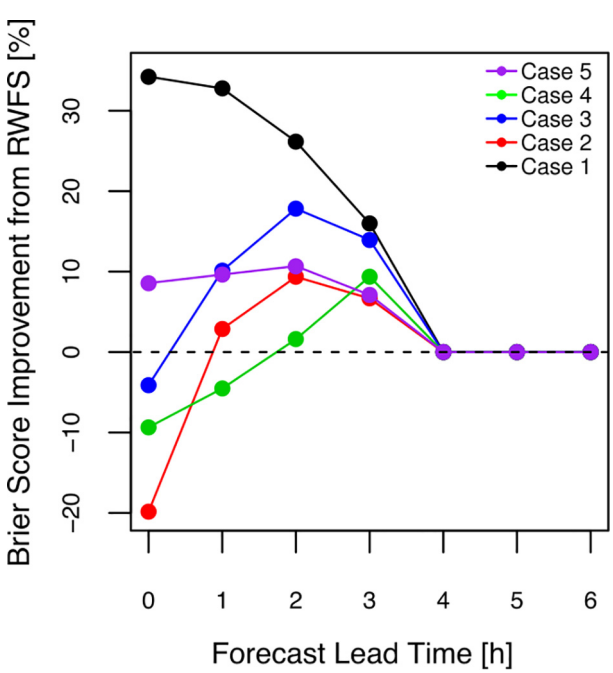

Fig. 6. Brier Score by lead hour for each case. A positive change indicates a Brier Score closer to 0, or improved performance, for the run with wiper status forward error correction.

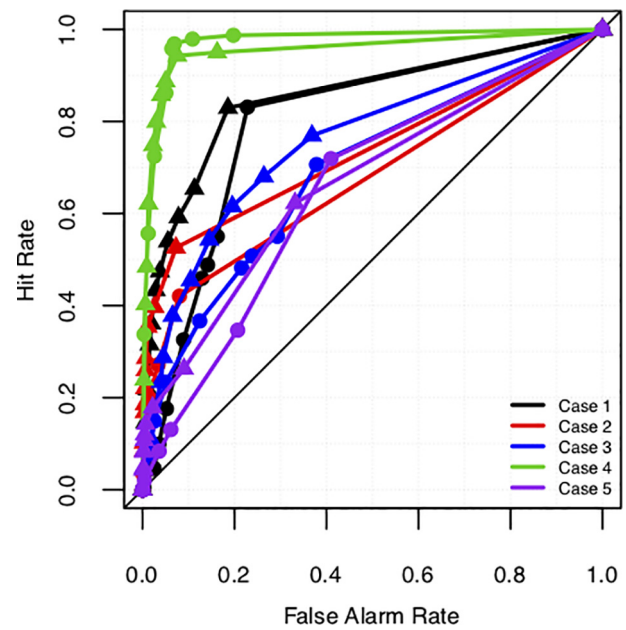

Fig. 7. Receiver Operating Characteristic (ROC) curves for each case for the original RWFS Probability of Precipitation outputs (circles) and the PoP after wiper status forward error correction (triangles). Curves closest to the upper left show best performance. The color scheme is the same as that for Fig. 6.

Table 3
Area under the curve for ROC analysis of RWFS probability of precipitation with and without wiper status forward error correction.

Case	Without Wipers	With Wipers	Percent Improvement
1	0.798	0.855	28.2
2	0.676	0.736	18.5
3	0.692	0.758	21.4
4	0.974	0.956	−69.2
5	0.655	0.660	1.4

4. Discussion

This paper presents the earliest steps in assimilating observations from CVs into an NWP model and quantifying their impact on model performance. While verification results were mixed for this pilot project, overall this new dataset made a modest but appreciable impact on WRF model output. BCRMSE and ME of WRF surface variables (2-m temperature, 2-m dewpoint, 10-m wind speed, and MSLP) were generally improved by 1%–5%, and WRF QPF skill scores were slightly higher for the more dense vehicle data assimilation cases. Because CV observations are all located at the surface, it is not expected that they would make significant impacts on larger-scale (*meso*-alpha to synoptic) flow patterns. This is especially true for limited-area domains as in the present study. Using larger domains with vehicle data available throughout the domain, instead of restricted to a single state or area as in this study, would likely be beneficial for forecast performance. That change would reduce the impacts of lateral boundary condition errors sweeping across the domain, and also allow for more of the domain to

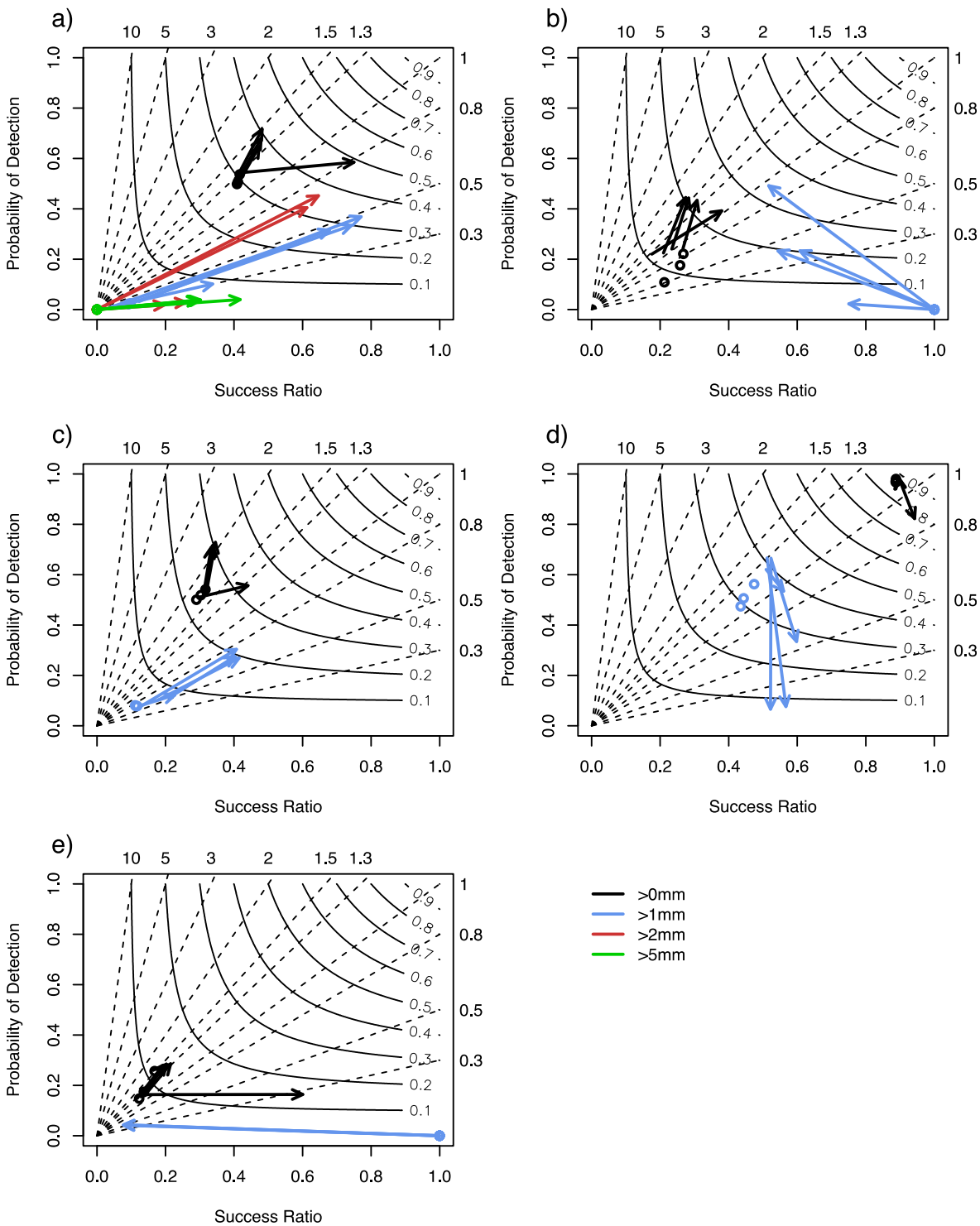

Fig. 8. Performance diagrams for RWFS QPF for Cases 1 – 5 (a – e). Arrows point from the original QPF score to the QPF score after wiper status forward error correction. Circles show little change between the two scores.

benefit from upstream and neighborhood impacts and corrections resulting from DA of vehicle observations.

The FEC using wiper status showed the most dramatic benefit of vehicle observations, with PoP performance increasing around 10% to 20% for Cases 1–3 and QPF categorical scores also improving drastically, with CSI increasing by between 0.1 and 0.4 in most instances. It is encouraging that the FEC using wiper status, rather than precipitation accumulation, led to the most improvement in the case with the heaviest and most widespread precipitation. Leveraging wiper status data could hold particular promise for forecast improvements in areas where radar coverage is poor or non-existent, such as in areas of complex terrain and rural states like Alaska.

These results are reflective of ideal conditions in terms of distribution of CV observations and timeliness of the observations. It is assumed that CV observations are homogeneously distributed across each urban area defined, and that there is less than a five-minute

latency associated with obtaining observations. In a future deployment, it is more likely that CV-equipped vehicles will be heterogeneously distributed and some observations may have an intolerably long latency. As long as an accurate location and date/time stamp are associated with every observation collected, the only effect of these issues will be to decrease the number of unique CV-based observations that are available for DA, rather than introducing inaccuracies to the final model output.

The results of this pilot study support the use of CV observations for applications outside strictly vehicle-to-vehicle safety communications or even road weather itself into a wider weather enterprise context. This represents a significant reach of CV benefits for not only the safety of individuals traveling the roadways, but also into weather prediction on larger scales. Given the safety implications of CVs on the roadways, safety related to road weather, and the economic impacts of extreme weather events [33], CV technology and observations have the potential to save lives and ease economic impacts of weather as they become more and more deployed. This is especially important in light of the Federal Communications Commission's (FCC's) exploration of opening up the 5.9 GHz band previously reserved for safety applications of CV technology [20], which could prove detrimental at best to the ongoing deployment efforts [24].

Future work should include limiting the assessment of vehicle observation impacts to data-sparse areas, where the road network may provide a much-needed boost in surface observations that is drowned out in more data-rich areas such as major urban centers. Different FDDA parameters or different assimilation methods, such as variational or ensemble Kalman filter techniques, should also be tested to determine the optimal way to handle this unique dataset that has irregular horizontal spacing, massive volume, quality control challenges, and non-traditional observing platforms. As more and more CVs come online, testing with real rather than simulated data would also be highly beneficial.

Although much work remains before it will be possible to optimally incorporate CV observations into NWP, this pilot project shows that these high-resolution observations have great potential to improve model initial conditions and short-range forecasts, especially in areas with few current surface observations.

CRediT authorship contribution statement

Amanda Siems-Anderson: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing - original draft, Visualization. **Jared A. Lee:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - review & editing, Visualization. **Barbara Brown:** Conceptualization, Methodology, Writing - review & editing, Supervision. **Gerry Wiener:** Software, Formal analysis, Resources, Data curation, Writing - review & editing. **Seth Linden:** Software, Formal analysis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments.

This work was supported by the U.S. Department of Transportation Road Weather Management Program and the Intelligent Transportation Systems Joint Program Office under contract DTH61-08-D-00012, where specification of the study design was made. Additional funding for preparing this publication for submission was provided by NCAR's Research Applications Laboratory (RAL). This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977. NCAR is sponsored by the National Science Foundation. Francois Vandenberghe and Will Cheng of NCAR/RAL provided generous support in the setup and troubleshooting of the FDDA system. Leigh Sturges and Ralph Patterson of Narwhal Group and Paul Pisano of the Federal Highway Administration provided valuable direction in the experiment design.

References

Atlas, R.,]. Atmospheric observations and experiments to assess their usefulness in data assimilation.. J. Meteorol. Soc. Japan 75, 111–130.

Anderson, A.R.S., Chapman, M., Drobot, S.D., Tadesse, A., Lambi, B., Wiener, G., Pisano, P.,]. Quality of mobile air temperature and atmospheric pressure observations from the 2010 Development Test Environment experiment. J. Atmos. Ocean. Technol. 51, 691–701. https://doi.org/10.1175/JAMC-D-11-0126.1.

Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., Avellano, A.,]. The data assimilation research testbed: a community facility. Bull. Am. Meteorol. Soc. 90, 1283–1296. https://doi.org/10.11175/2009BAMS2618.1.

Arnold Jr., C.P., Dey, C.H.,]. Observing-systems simulation experiments: past, present, and future. Bull. Am. Meteorol. Soc. 67, 687–695. https://doi.org/10.1175/1520-0477(1986)067 < 0687:OSSEPP > 2.0.CO;2.

Benjamin, S.G., Weygandt, S.S., Brown, J.M., Hu, M., Alexander, C.R., Smirnova, T.G., Olson, J.B., James, E.P., Dowell, D.C., Grell, G.A., Lin, H., Peckham, S.E., Smith, T.L., Moniger, W.R., Kenyon, J.S., Manikin, G.S.,]. A North American hourly assimilation and model forecast cycle: the rapid refresh. Mon. Weather Rev. 144, 1669–1694. https://doi.org/10.1175/MWR-D-15-0242.1.

Bernstein, B.C., Wolff, J.K., Linden, S., Myers, W., 2004. Verification of the Road Weather Forecast System run during the 2003 demonstration at the Iowa Department of Transportation. 6th Int. Symp. on Snow Remov. and Ice Control Technol., Spokane, WA, pp. 109–126, https://trid.trb.org/view.aspx?id = 702012.

Booz Allen Hamilton, 2018. Weather-related crash statistics. Accessed 18 March 2020, http://www.ops.fhwa.dot.gov/weather/q1_roadimpact.htm.

Brown, B.G., Gotway, J.H., Bullock, R., Gilleland, E., Ahijevych, D., 2009. The Model Evaluation Tools (MET): Community tools for forecast evaluation. 25th Conf. on Interact. Syst. (IIPS) for Meteorol., Ocean., and Hydrol., Phoenix, AZ, 9A.6, https://ams.confex.com/ams/pdfpapers/151349.pdf.

Bouttier, F., Kelly, G.,]. Observing-system experiments in the ECMWF 4D-Var data assimilation system. Q. J. Royal Met. Soc. 127 (574), 1469–1488. https://doi.org/10.1002/qj.49712757419.

Boyce, B.C., Pisano, P., Guevara, G.N., 2015. FHWA Road Weather Management Program Integrated Mobile Observations (IMO) project. 31st Conf. on Environ. Inf. Process. Technol., Phoenix, AZ, 7B.1, https://ams.confex.com/ams/95Annual/webprogram/Paper270297.html.

Bytheway, J.L., Kummerow, C.D., 1]. Toward an object-based assessment of high-resolution forecasts of long-lived convective precipitation in the central U.S.: HRRR assessment. J. Adv. Model. Earth Syst. 7 (3), 1248–1264. https://doi.org/10.1002/2015MS000497.

Chapman, M., Drobot, S., Anderson, A., Burghardt, C., 2013. Results from the Integrated Mobile Observations Study. FHWA Tech. Rep. FHWA-JPO-13-066, 157 pp, http://ntl.bts.gov/lib/48000/48300/48314/Final_Report_Task_5b_5-31-131.pdf.

Cintineo, R.M., Otkin, J.A., Jones, T.A., Koch, S., Stensrud, D.J., 2016. Assimilation of Synthetic GOES-R ABI Infrared Brightness Temperatures and WSR-88D Radar Observations in a High-Resolution OSSE. Mon. Weather Rev. 144, pp. 3159–3180, doi: 10.1175/MWR-D-15-0366.1.

Coen, J.L., Cameron, M., Michalakes, J., Patton, E.G., Riggan, P.J., Yedinak, K.M., 2013. WRF-Fire: Coupled weather-wildland fire modeling with the Weather Research and Forecasting model. J. Appl. Meteorol. Climatol. 52, pp. 16–38, doi: 10.1175/JAMC-D-12-023.1.

Crevier, L., Delage, Y., 5]. METRo: a new model for road-condition forecasting in Canada. J. Appl. Meteorol. 40, 2026–2037. https://doi.org/10.1175/1520-0450(2001)040 < 2026:MANMFR > 2.0.CO;2.

Crook, N.A., Sun, J., 6]. Assimilating radar, surface, and profiler data for the Sydney 2000 Forecast Demonstration Project. J. Atmos. Ocean. Technol. 19, 888–898. https://doi.org/10.1175/1520-0426(2002)019 < 0888:ARSAPD > 2.0.CO;2.

Davis, C., Brown, B., Bullock, R., 7]. Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. Mon. Weather Rev. 134, 1772–1784. https://doi.org/10.1175/MWR3145.1.

De Pondeca, M.S.F.V., Manikin, G.S., DiMego, G., Benjamin, S.G., Parrish, D.F., Purser, R.J., Wu, W.-S., Horel, J.D., Myrick, D.T., Lin, Y., Aune, R.M., Keyser, D., Colman, B., Mann, G., Vavra, J., 8]. The real-time mesoscale analysis at NOAA's National Centers for environmental prediction: current status and development. Weather Forecast. 26, 593–612. https://doi.org/10.1175/WAF-D-10-05037.1.

Dey, K.C., Mishra, A., Chowdhury, M., 9]. Potential of intelligent transportation systems in mitigating adverse weather impacts on road mobility: a review. IEEE Trans. Intell. Transport. Syst. 16 (3), 1107–1119. https://doi.org/10.1109/TITS.2014.2371455.

Federal Communications Commission, 2019. FCC seeks to promote innovation in the 5.9 Ghz band. Accessed 23 Mar 2020, https://www.fcc.gov/document/fcc-seeks-promote-innovation-59-ghz-band.

Fujita, T., Stensrud, D.J., Dowell, D.C., 1]. Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physics uncertainties. Mon. Weather Rev. 135, 184–1868. https://doi.org/10.1175/MWR3391.1.

Fulton, R.A., Breidenbach, J.P., Seo, D.-J., Miller, D.A., O'Bannon, T., 2]. The WSR-88D rainfall algorithm. Weather and Forecast. 13, 377–395. https://doi.org/10.1175/1520-0434(1998)013%3C0377:TWRA%3E2.0.CO;2.

Gao, J., Xue, M., Shapiro, A., Droegemeier, K., 3]. A variational method for the analysis of three-dimensional wind fields from two Doppler radars. Mon. Weather Rev. 127, 2128–2142. https://doi.org/10.1175/1520-0493(1999)127%3C2128:AVMFTA%3E2.0.CO;2.

Goldin, P., 4]. Meet the USDOT Connected Vehicle Pilot Deployment Sites. ITSdigest. https://www.itsdigest.com/meet-usdot-connected-vehicle-pilot-deployment-sites.

Goodall, N.J., 2013. Real-time prediction of vehicle locations in a connected vehicle environment. USDOT Tech. Rep. FHWA/VCTIR 14-R4, 53 pp., http://www.virginiadot.org/vtrc/main/online_reports/pdf/14-r4.pdf.

Grell, G.A., Dévényi, D., 6]. A generalized approach to parameterizing convection combining ensemble and data assimilation techniques: parameterizing convection combining ensemble and data assimilation techniques. Geophys. Res. Lett. 29 (14), 38-1–38-4. https://doi.org/10.1029/2002GL015311.

Grell, G.A., Freitas, S.R., 7]. A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. Atmos. Chem. Phys. 14, 5233–5250. doi: 105194/acp-14-5233-2014.

Harding, J., Powell, G., Yoon, R., Fikentscher, J., Doyle, C., Sade, D., Lukuc, M., Simons, J., Wang, J., 2014. Vehicle-to-vehicle communications: readiness of V2V technology for application. USDOT Tech. Rep. DOT HS 812 014, 327 pp, http://www.nhtsa.gov/staticfiles/rulemaking/pdf/V2V/Readiness-of-V2V-Technology-for-Application-812014.pdf.

Halliwell, G.R., Goni, G.J., Mehari, M.F., Kourafalou, V.H., Baringer, M., Atlas, R., 9]. OSSE assessment of underwater glider arrays to improve ocean model initialization for tropical cyclone prediction. J. Atmos. Ocean. Technol. 37, 467–487. https://doi.org/10.1175/JTECH-D-18-0195.1.

Harper, K., Uccellini, L.W., Kalnay, E., Carey, K., Morone, L., 0]. 50th anniversary of operational numerical weather prediction. Bull. Am. Meteorol. Soc. 88, 639–650. https://doi.org/10.1175/BAMS-88-5-639.

Hong, S.-Y., Dudhia, J., 1]. Next-generation numerical weather prediction: bridging parameterization, explicit clouds, and large eddies. Bull. Am. Meteor. Soc. 93 (1), ES6–ES9. https://doi.org/10.1175/2011BAMS3224.1.

Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shephard, M.W., Clough, S.A., Collins, W.D., 2]. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. J. Geophys. Res. 113 (D13). https://doi.org/10.1029/2008JD009944.

Jahn, M., 3]. Economics of extreme weather events: terminology and regional impact models. Weather and Climate Extremes 10, 29–39. https://doi.org/10.1016/j.wace:2015.08.005.

Jolliffe, I.T., Stephenson, D.B., 2012. Forecast Verification: A Practitioner's Guide in Atmospheric Science, second ed. Wiley, 292 pp.

Krechmer, D., N. Perry, L. Destro, S. Gallaher, 2012. Weather delay costs to trucking. FHWA Tech. Rep. FHWA-JPO-13-023, 72 pp, http://transops.s3.amazonaws.com/uploaded_files/FHWA-JPO-13-023-Weather-Delay-Costs-to-Trucking.pdf.

Leidner, S.M., Annane, B., McNoldy, B., Hoffman, R., Atlas, R., 6]. Variational analysis of simulated ocean surface winds from the cyclone global navigation satellite system (CYGNSS) and evaluation using a regional OSSE. J. Atmos. Ocean. Technol. 35, 1571–1584. https://doi.org/10.1175/JTECH-D-17-0136.1.

Lin, Y., Mitchell, K.E., 2005. The NCEP stage II/IV hourly precipitation analyses: development and applications. Preprints, 19th Conf. on Hydrol., San Diego, CA., 1.2, http://www.emc.ncep.noaa.gov/mmb/ylin/pcpanl/refs/stage2-4.19hydro.pdf.

Liu, Z., Rabier, F., 8]. The interaction between model resolution, observation resolution and observation density in data assimilation: a one-dimensional study. Q. J. Royal Meteorol. Soc. 128, 1367–1386. https://doi.org/10.1256/003590002320373337.

Ma, Z., Riishøjgaard, L.P., Masutani, M., Woollen, J.S., Emmitt, G.D., 9]. Impact of different satellite wind Lidar telescope configurations on NCEP GFS forecast skill in observing system simulation experiments. J. Atmos. Ocean. Technol. 32, 478–495. https://doi.org/10.1175/JTECH-D-14-00057.1.

Madaus, L.E., Mass, C.F., 0]. Evaluating smartphone pressure observations for mesoscale analyses and forecasts. Weather Forecast. 32, 511–531. https://doi.org/10.1175/WAF-D-16-0135.1.

Mahoney, W.P., O'Sullivan, J.M., 1]. Realizing the potential of vehicle-based observations. Bull. Am. Meteorol. Soc. 94, 1007–1018. https://doi.org/10.1175/BAMS-D-12-00044.1.

Mahoney, W.P., Drobot, S., Pisano, P., McKeever, B., O'Sullivan, J., 2]. Vehicles as mobile weather observation systems. Bull. Am. Meteorol. Soc. 91, 1179–1182. https://doi.org/10.1175/2010BAMS2954.1.

Mahoney, W.P., Parks, K., Wiener, G., Liu, Y., Myers, W.L., Sun, J., Delle Monache, L., Hopson, T., Johnson, D., Haupt, S.E., 3]. A wind power forecasting system to optimize grid integration. IEEE Trans. Sustain. Energy 3 (4), 670–682. https://doi.org/10.1109/TSTE.2012.2201758.

Mass, C.F., Ovens, D., Westrick, K., Colle, B.A., 4]. Does increasing horizontal resolution produce more skillful forecasts?. Bull. Am. Meteorol. Soc. 83, 407–430. https://doi.org/10.1175/1520-0477(2002)083 < 0407:DIHRPM > 2.3.CO;2.

Mass, C., Madaus, L., 5]. Surface pressure observations from smartphones: a potential revolution for high-resolution weather prediction?. Bull. Am. Meteorol. Soc. 95, 1343–1349. https://doi.org/10.1175/BAMS-D-13-00188.1.

McNitt, J.R., Facundo, J., O'Sullivan, J., Campbell, M., 2007. Transitioning the Meteorological Assimilation Data Ingest System (MADIS) into NWS operations. 23rd Int. Conf. on Interact. Inf. and Process. Syst., San Antonio, TX, 3B.7, https://ams.confex.com/ams/87ANNUAL/techprogram/paper_116942.htm.

McPherson, R.A., Fiebrich, C.A., Crawford, K.C., Kilby, J.R., Grimsley, D.L., Martinez, J.E., Basara, J.B., Illston, B.G., Morris, D.A., Kloesel, K.A., Melvin, A.D., Shrivastava, H., Wolfinbarger, J.M., Bostic, J.P., Demko, D.B., 7]. Statewide monitoring of the mesoscale environment: a technical update on the Oklahoma Mesonet. J. Atmos. Ocean. Technol. 24, 301–321. https://doi.org/10.1175/JTECH1976.1.

Michigan Department of Technology, Management, & Budget, 2014. MI Geographic Data Library. Accessed 20 March 2020, http://gis-michigan.opendata.arcgis.com/search?tags = transportation.

Michigan Department of Transportation (MDOT), 2014. Traffic Monitoring Information System (TMIS). Accessed 20 March 2020, http://mdotnetpublic.state.mi.us/tmispublic.

Miller, P.A., Barth, M.F., Benjamin, L.A., 2005. An update on MADIS observation ingest, integration, quality control and distribution capabilities. Preprints, 21st Int. Conf. on Interact. Inf. and Process. Syst., San Diego, CA, J7.12, https://ams.confex.com/ams/Annual2005/techprogram/paper_86703.htm.

Minnesota Department of Transportation (MnDOT), 2014a. Traffic Forecasting & Analysis. Accessed 20 March 2020, http://www.dot.state.mn.us/traffic/data.

Minnesota Department of Transportation (MnDOT), 2014b. GIS Data Products. Accessed 20 March 2020, http://www.dot.state.mn.us/maps/gdma/cart-products.html.

Muller, C.L., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A., Leigh, R.R., 3]. Crowdsourcing for climate and atmospheric sciences: current status and future potential: crowdsourcing for climate and atmospheric sciences. Int. J. Climatol. 35 (11), 3185–3203. https://doi.org/10.1002/joc.4210.

Nakanishi, M., Niino, H., 4]. An improved Mellor–Yamada Level-3 model: its numerical stability and application to a regional prediction of advection fog. Boundary-Layer Meteorol. 119 (2), 397–407. https://doi.org/10.1007/s10546-005-9030-8.

National Weather Service (NWS), 1998. Automated Surface Observation System (ASOS) user's guide. National Oceanic and Atmospheric Administration (NOAA), Washington, DC, 61 pp, http://www.nws.noaa.gov/asos.

Nelson, B.R., Prat, O.P., Seo, D.-J., Habib, E., 6]. Assessment and implications of NCEP Stage IV quantitative precipitation estimated for product intercomparisons. Weather Forecast. 31, 371–394. https://doi.org/10.1175/WAF-D-14-00112.1.

Pisano, P., Goodwin, L.C., 2002. Surface Transportation Weather Applications. ITE 2002 Annual Meeting and Exhibit, Philadelphia, PA, https://trid.trb.org/view/732438.

Pisano, P.A., Stern, A.D., Mahoney III, W.P., Myers, W.L., Burkheimer, D., 2004. Winter road maintenance decision support system project: Overview and status. 6th Int. Symp. on Snow Remov. and Ice Control Technol., Spokane, WA, pp. 3–17, https://trid.trb.org/view.aspx?id = 702004.

Pisano, P.A., Pol, J.S., Stern, A.D., Boyce, B.C., Garrett, J.K., 2007. Evolution of the U.S. Department of Transportation Clarus Initiative: Project Status and Future Plans. Preprints, 23rd Conf. on Interact. Syst. (IIPS) for Meteorol., Ocean., and Hydrol., San Antonio, TX, 4A.5, https://ams.confex.com/ams/87ANNUAL/techprogram/paper_119018.htm.

Powers, J.G., Klemp, J.B., Skamarock, W.C., Davis, C.A., Dudhia, J., Gill, D.O., Coen, J.L., Gochis, D.J., 2017. The Weather Research and Forecasting model: overview, system efforts, and future directions. Bull. Am. Meteorol. Soc. 98, pp. 1717–1737. doi: 10.1175/BAMS-D-15-00308.1.

Rew, R., Davis, G., 1]. NetCDF: an interface for scientific data access. IEEE Comput. Grap. Appl. 10 (4), 76–82. https://doi.org/10.1109/38.56302.

Roebber, P.J., 2009. Visualizing multiple measures of forecast quality. Weather Forecast. 24, pp. 601–608, doi: 10.1175/2008WAF2222159.1

Rogers, E., DiMego, G., Black, T., Ek, M., Ferrier, B., Gayno, G., Janjic, Z., Lin, Y., Pyle, M., Wong, V., Wu, W.S., Carley, J., 2009. The NCEP North American mesoscale modeling system: Recent changes and future plans. Preprints, 23rd Conf. on Weather Analysis and Forecasting, Omaha, NE, 2A.4, https://ams.confex.com/ams/23WAF19NWP/techprogram/paper_154114.htm.

Schroeder, J.L., Burgett, W.S., Haynie, K.B., Sonmez, I., Skwira, G.D., Doggett, A.L., Lipe, J.W., 2005. The West Texas Mesonet: a technical overview. J. Atmos. Ocean. Technol. 22, pp. 211–222, doi: 10.1175/JTECH-1690.1.

Schwartz, C.S., Romine, G.S., Fossell, K.R., Sobash, R.A., Weisman, M.L., 5]. Toward 1-km ensemble forecasts over large domains. Mon. Wea. Rev. 145 (8), 2943–2969. https://doi.org/10.1175/MWR-D-16-0410.1.

Siems-Anderson, A.R., Wiener, G., Brummet, T., Linden, S., Petzke, W., McCarthy, P., Garcia, V., Ragan, A., English, T., Gopalakrishna, D., Hsu, E., 2017. Use of the Pikalert(R) System in the Wyoming Department of Transportation Connected Vehicle Pilot Deployment. Routes/Roads No 375, 4th quarter, pp. 27-30, https://www.piarc.org/ressources/publications/10/27658,RR375-Routes-Roads-Magazine-Dossiers-EN-pp27-30.pdf.

Siems-Anderson, A.R., Walker, C.L., Wiener, G., Mahoney III, W.P., Haupt, S.E., 7]. An adaptive big data weather system for surface transportation. Transp. Res. Interdisciplinary Perspectives 3, 100071. https://doi.org/10.1016/j.trip.2019.100071.

Skamarock, W.C., Klemp, J.B. Dudhia, J., Gill, D.O., Barker, D., Duda, Huang, X.-Y., Wang, W., Powers, J.G., 2008. A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475 + STR, 113 pp, https://opensky.ucar.edu/islandora/object/technotes:500.

Smalley, M., L'Ecuyer, T., Lebsock, M., Haynes, J., 2014. A comparison of precipitation occurrence from the NCEP Stage IV QPE product and the CloudSat cloud profiling radar. J. Hydrometeorol. 15, pp. 444–458, doi: 10.1175/JHM-D-13-048.1.

Smirnova, T.G., Brown, J.M., Benjamin, S.G., 0]. Performance of different soil model configurations in simulating ground surface temperature and surface fluxes. Mon. Weather Rev. 125, 1870–1884. https://doi.org/10.1175/1520-0493(1997)125 < 1870LPODSMC > 2.0CO;2.

Stauffer, D.R., Seaman, N.L., 1]. Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. Mon. Weather

Rev. 118, 1250–1277. https://doi.org/10.1175/1520-0493(1990)118<1250:UOFDDA>2.0.CO;2.

Stauffer, D.R., Seaman, N.L., 2]. Multiscale four-dimensional data assimilation. J. Appl. Meteorol. 33, 416–434. https://doi.org/10.1175/1502-0450(1994)033<0416:MFDDA>2.0.CO;2.

Stauffer, D.R., Seaman, N.L., Binkowski, F.S., 3]. Use of four-dimensional data assimilation in a limited-area mesoscale model. Part II: Effects of data assimilation within the planetary boundary layer. Mon. Weather Rev. 119, 734–754. https://doi.org/10.1175/1520-0493(1991)119<0734:UOFDDA>2.0.CO;2.

Strong, C.K., Shi, X., 4]. Integrating weather into transportation operations: A Utah Department of Transportation case study. Transp. Res. Circ. E-C126, 318–333.

Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M.A., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., Cuenca, R.H., 2004. Implementation and verification of the unified Noah land surface model in the WRF model. 20th Conf. on Weather Analy. and Forecast./16th Conf. on Numerical Weather Pred., Paper 14.2A, Seattle, WA, 14 Jan 2004, https://ams.confex.com/ams/pdfpapers/69061.pdf.

Thompson, G., Field, P.R., Rasmussen, R.M., Hall, W.D., 2008. Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. Mon. Weather Rev. 136, pp. 5095–5115, doi: 10.1175/2008MWR2387.1

United States Census Bureau, 2012: Population and Housing Unit Counts: 2010. Accessed 23 Mar 2020, https://www2.census.gov/library/publications/decennial/2010/cph-2/cph-2-1.pdf.

United States Department of Transportation, 2017. Connected Vehicles. Accessed 23 Mar 2020, https://www.its.dot.gov/cv_basics/.

Walker, C.L., Steinkruger, D., Gholizadeh, P., Hasanzadeh, S., Anderson, M.R., Esmaeili, B., 2019. Developing a department of transportation winter severity index. J. Appl. Meteorol. Climatol. 58, 1779–1798. doi: 10.1175/JAMC-D-18-0240.1.

Warner, T.T., 0]. Quality assurance in atmospheric modeling. Bull. Amer. Meteor. Soc. 92 (12), 1601–1610. https://doi.org/10.1175/BAMS-D-11-00054.1.

Wilks, D.S., 1]. Statistical Methods in the Atmospheric Sciences. Academic Press, p. 676.

Yuan, H., Gao, X., S.L. Mullen, Sorooshian, S., Du, J., Juang,, H.-M.H., 2007. Calibration of probabilistic quantitative precipitation forecasts with an artificial neural network. Weather Forecast. 22, pp. 1287–1303, doi: 10.1175/2007WAF2006114.1.

Zapotocny, T.H., Jung, J.A., Le Marshall, J.F., Treadon, R.E., 2007. A two-season impact study of satellite and in situ data in the NCEP Global Data Assimilation System. Weather Forecast. 22, pp. 887–909, doi: 10.1175/WAF1025.1

Zeng, X., Atlas, R., Birk, R.J., Carr, F.H., Carrier, M.J., Cucurull, L., Hooke, W.H., Kalnay, E., Murtugudde, R., Posselt, D.J., Russell, J.L., Tyndall, D.P., Weller, R.A., Zhang, F., 2020. Use of Observing System Simulation Experiments in the U.S.. Bull. Amer. Meteor. Soc., doi: 10.1175/BAMS-D-19-0155.1.

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cocks, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., 5]. Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimations: Initial operating capabilities. Bull. Am. Meteorol. Soc. 97, 621–638. https://doi.org/10.1175/BAMS-D-14-00174.1.