BAMS Article

Evaluating the Economic Impacts of Improvements to the High-Resolution Rapid Refresh (HRRR) Numerical Weather Prediction Model

David D. Turner, Harvey Cutler, Martin Shields, Rebecca Hill, Brad Hartman, Yuchen Hu, Tao Lu, and Hwayoung Jeon

ABSTRACT: Forecasts from numerical weather prediction (NWP) models play a critical role in many sectors of the U.S. economy. Improvements to operational NWP model forecasts are generally assumed to provide significant economic savings through better decision-making. But is this true? Since 2014, several new versions of the High-Resolution Rapid Refresh (HRRR) model were released into operation within the National Weather Service. Practically, forecasts have an economic impact only if they lead to a different action than what would be taken under an alternative information set. And in many sectors, these decisions only need to be considered during certain weather conditions. We estimate the economic impacts of improvements made to the HRRR, using 12-h wind, precipitation, and temperature forecasts in several cases where they can have "economically meaningful" behavioral consequences. We examine three different components of the U.S. economy where such information matters: 1) better integration of wind energy resources into the electric grid, 2) increased worker output due to better precipitation forecasts that allow workers to arrive to their jobs on time, and 3) better decisions by agricultural producers in preparing for freezing conditions. These applications demonstrate some of the challenges in ascertaining the economic impacts of improved weather forecasts, including highlighting key assumptions that must be made to make the problem tractable. For these sectors, we demonstrate that there was a marked economic gain for the United States between HRRR versions 1 and 2 and a smaller, but still appreciable economic gain between versions 2 and 3.

https://doi.org/10.1175/BAMS-D-20-0099.1
Corresponding author: Dave Turner, dave.turner@noaa.gov
In final form 31 August 2021

Improved weather forecasts, resulting from continued development of the HRRR, can change behaviors and hence have an economic impact. Here, we quantify that impact in three areas.

AFFILIATIONS: Turner—NOAA/Global Systems Laboratory, Boulder, Colorado; **Cutler, Shields, Hill, Hartman, Hu, Lu,** and **Jeon**—Colorado State University, Fort Collins, Colorado

Weather forecasts play an essential role in society (Lazo et al. 2009). Each day, decision-makers use weather forecasts to make a wide range of choices, from the trivial (e.g., "Should I plan on a picnic tomorrow?") to very impactful (e.g., "Should we evacuate our community due to possible heavy rain?"). Indeed, the mission of the U.S. National Weather Service (NWS) is to "provide weather, water, and climate data, forecasts and warnings for the protection of life and property and enhancement of the national economy." While such decisions can be greatly consequential, it is difficult to gauge the economic value of improved weather forecasts.

Perhaps the most intuitive way is to show the impact of two different forecasts, recognizing that a forecast will only have an effect if it changes individual behavior (Demuth et al. 2011). For example, would you change your decision to drive to your grandmother's house if the forecasted weather conditions were rainy and 70°F versus rainy and 80°F? What if the forecasted weather conditions were rainy and 25°F versus rainy and 35°F? Both cases have the same forecast difference (namely, 10°F), but the latter case is likely to have a much higher probability of changing your decision to make that drive. Would you have changed your decision in either case if the forecast was sunny instead of rainy?

Researchers have estimated the economic impact of both forecasts and various weather events for decades. Thompson and Brier (1955) provided early guidance on how to characterize economic impact, using both probabilistic and categorial forecasts. Murphy (1969) used a cost-loss ratio as a way to evaluate whether or not to take a precautionary, costly action based upon a weather forecast that prevents or reduces the effects of an adverse outcome, dependent on expected payoffs. Lazo et al. (2011) looked at the aggregated economic impact of weather variability in the United States and how the sensitivity of the economic impact depends on the region of the country, while Strobl (2011) estimated the impact of hurricanes on the economic growth in coastal U.S. communities. There are many other examples demonstrating that the economic impacts are often localized and frequently episodic. However, learning how to better disseminate weather forecast information (e.g., as part of the "Weather Ready Nation" effort within the NWS) can reduce the economic impact of severe weather (e.g., Lazo et al. 2020).

Estimating the economic implications of a decision—including those informed by weather forecasts—is seldom straightforward. While many facets of the economy warrant evaluation, we focus on three: 1) wind energy, 2) morning commutes, and 3) specialty crop agriculture, each of which use wind, precipitation, and temperature forecasts, respectively. We chose these areas because there is a relatively intuitive, though somewhat complex, methodology to ascertain the economic benefits of improved weather forecasts. For the commuting and agricultural analyses, we applied economic models, described later, to quantify the differences in the economic impact of different weather forecasts for each of the eight Bureau of Economic Analysis (BEA)[1] regions in the United States (Table 1). By looking at these subnational regions, we are able to allow for regional heterogeneity in both economic structure and weather.

The High-Resolution Rapid Refresh (HRRR; Benjamin et al. 2016) is one of the operational forecast models used by the NWS for short-term forecasts. HRRR is under nearly continuous

[1] www.bea.gov

Table 1. States, number of MSAs, total turbine capacity, and number of turbines within each BEA region, as well as the turbine capacity and number of turbines that are within 20 km of a METAR station within each BEA (data source: Hoen et al. 2018).

BEA	States	Number of MSAs	Total turbine capacity (kW)	Total number of turbines	Turbine capacity within 20-km buffer zone (kW)	Number of turbines within 20-km buffer zone
New England	CT, ME, MA, NH, RI, VT	10	1,434,065	650	210,765	123
Mideast	DE, DC, MD, NJ, NY, PA	14	3,408,460	1,873	171,275	79
Great Lakes	IL, IN, MI, OH, WI	33	9,800,118	5,629	1,204,288	709
Plains	IA, KS, MN, MO, NE, ND, SD	22	22,555,838	12,948	2,533,352	1,461
Southeast	AL, AR, FL, GA, KY, LA, MS, NC, SC, TN, VA, WV	59	922,880	499	83,280	53
Southwest	AZ, NM, OK, TX	29	32,786,460	17,789	3,755,350	2,184
Rocky Mountain	CO, ID, MT, UT, WY	14	6,699,610	4,229	700,370	404
Far West	AK, CA, HI, NV, OR, WA	25	12,316,156	10,811	3,048,176	3,422

development, with model updates transitioned to NWS operations approximately every two years. We estimate selected economic impacts resulting from improvements to the HRRR's wind, temperature, and precipitation forecasts. As part of the transition between versions, there are time periods when the new model (version $X + 1$) is running simultaneously with the operational model (version X); we focus on the periods when both versions provide forecasts for the same actual weather conditions. The ground truth data—used to evaluate the two different forecasts—are provided by observations.

The evolution of the HRRR

The HRRR is a storm-scale model, with a 3-km horizontal grid spacing that is initialized hourly, that provides forecasts over the conterminous United States. It is based on the Advanced Research version of the Weather Research and Forecast (WRF-ARW) Model (Skamarock et al. 2008), using a fixed set of physical parameterizations within that modeling suite (Benjamin et al. 2016). Observations from surface meteorological stations, radiosondes, aircraft, and scanning precipitation radars (the WSR-88D network; Crum and Alberty 1993) are assimilated using the Gridpoint Statistical Interpolation data assimilation system (Kleist et al. 2009) to initialize the atmospheric state of the model. Scientists at the National Oceanic and Atmospheric Administration (NOAA) Global Systems Laboratory (GSL) continually improve both the representation of the physical processes in the model (e.g., Olson et al. 2019; Angevine et al. 2018; Smirnova et al. 2016) and the initialization methods (e.g., Peckham et al. 2016). The HRRR is regularly evaluated against a range of different observation types during its development (Turner et al. 2020).

HRRR model development started in 2008, with the first version (HRRR1) becoming operational at the NWS National Centers for Environmental Prediction (NCEP) in September 2014. Since then, additional versions of the model were released on 23 August 2016 (HRRR2), 12 July 2018 (HRRR3), and 2 December 2020 (HRRR4). As part of the transition to a new version, there is an overlap period of many months, when the new version runs simultaneously with the current operational version. We focus on the 12-h forecasts (from all initialization times; i.e., hourly) from these overlap periods between HRRR1 and HRRR2 and between HRRR2 and HRRR3, as both models provide forecasts for the same weather events. The overlap periods analyzed and the primary differences between the model versions are provided in Table 2. Note that this approach does not allow us to analyze HRRR1 against HRRR3, as these two models were not run over the same weather events. Further, the length and seasonal period of overlap differs between HRRR1 against HRRR2 and HRRR2 against HRRR3.

Table 2. Primary improvements made to the HRRR2 and HRRR3, relative to the previous version, and the overlap period used for the evaluation of the different versions.

Version	Primary updates in new version	Overlap period
HRRR1 → HRRR2	First inclusion of subgrid-scale clouds, aerosol particles included in cloud and precipitation processes, full cycling of the land surface model	1 Jun 2015–1 Aug 2016
HRRR2 → HRRR3	Updated turbulence scheme to use nonlocal mixing, more realistic treatment of subgrid-scale clouds, improved vertical coordinate for simulation above complex terrain, improved data assimilation approach to help retain stratiform clouds	1 Jul 2017–1 Jun 2018

To give a sense of improvements in the HRRR's accuracy, consider Fig. 1, where we compare performance across the operational versions for 2-m relative humidity (RH) forecasts across the eastern part of the United States. The bias and root-mean-square (RMS) difference between the forecasts and observed RH values were clearly much larger and more variable for HRRR1 relative to the other versions, and show a strong seasonal dependence. However, HRRR2 and HRRR3 show almost no seasonal dependence in the RMS, and generally have a very similar character to the seasonal evolution of the bias. Based on this single metric, one might infer that there was a dramatic improvement from HRRR1 to HRRR2, but a smaller improvement from HRRR2 to HRRR3. In practice, the HRRR's performance is measured by more than simple statistics in a single geophysical variable over a broad domain, but the general idea that there was a larger improvement going from HRRR1 to HRRR2 compared to the gains from HRRR2 to HRRR3 presents itself in the economic analysis.

Economic impact analyses

Although we look at three unique sectors of economic impact, each analysis follows the same basic intuition, depicted in the decision tree shown in Fig. 2. A decision-maker's objective is to make choices that maximize their expected payoff, which could be gauged in either profit or utility. Ex ante, they use all available information—including weather forecasts—when deciding whether or not to take a specific mitigating action (e.g., to preemptively protect a crop from a frost event). The decision-maker's ex post payoff depends on the action taken—which may be costly—and the consequences of the actual weather, which are not known at the time the decision is made, but can be anticipated. Payoffs can depend on whether or not an action was taken, and are typically greatest when there is no adverse weather forecasted and the

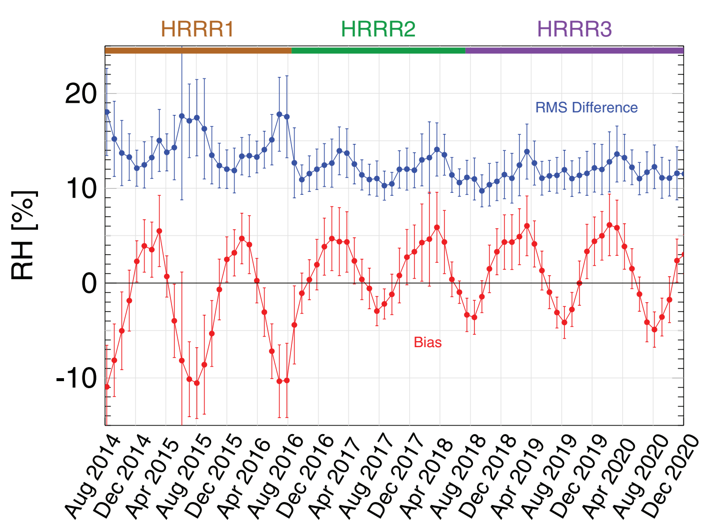

Fig. 1. The 30-day average of the bias (red) and RMS difference (blue) of 2-m relative humidity between the operational HRRR 12-h forecast and METAR observations over the eastern United States (east of 100°W). The periods when the HRRR1, HRRR2, and HRRR3 were operational at NCEP are indicated along the top.

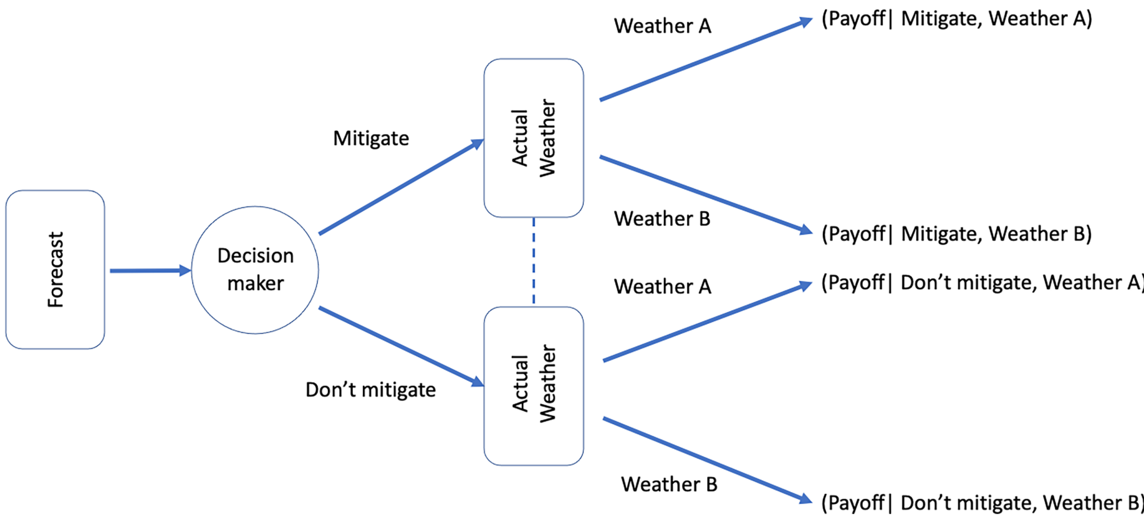

Fig. 2. The decision tree on whether or not to mitigate. The dashed line represents the uncertainty about future weather events when the decision-maker must take action. Note that the four payoffs on the right often have markedly different values.

decision-maker does not take a costly action (e.g., not protecting crops from frost on a warm day). Realized payoffs are generally lowest when the decision-maker does not take a mitigating action (perhaps due to a "no freeze" forecast) but the adverse weather event actually occurs, causing economic harm. Improved weather forecasts reduce the likelihood of such "mistakes."

Impact on wind-generated renewable energy. In the United States, electric utilities produce power from a variety of different sources in their portfolio, including both nonrenewable (e.g., fossil fuels, nuclear) and renewable sources (e.g., wind, solar). Production costs vary greatly by source (e.g., Ray 2019), and some of these sources are more "reliable" or readily available than others. Each day, a utility forecasts how much electricity its customers will demand the next day, and the decision-maker must schedule how much electricity it will produce from each source. Because wind energy is relatively low-cost and has zero emissions, it is often a preferred choice; however, its variability of the energy produced depends on the weather, which makes its full integration challenging. This task becomes less daunting when utility managers have accurate wind forecasts for the lowest part of the atmospheric boundary layer. The HRRR is one of the several weather forecasting models the energy community uses, characterizing the winds at the turbine hub height, both over terrain and offshore (James et al. 2017).

In practice, utilities model the anticipated power that their turbines will produce given the forecasted winds.[2] There are two potential costs associated with "errors" in day-ahead wind forecasts. One occurs if the actual wind speed is less than what was forecasted (i.e., "overprediction" of the amount of wind energy that would be produced). In such a case, the utility may need to bring a reserve fossil fuel generator online quickly or purchase electricity from the spot market; both of these options are relatively expensive. The second cost occurs if the actual wind speed is greater than what was forecast (i.e., "underprediction"), allowing the utility to produce more electricity than anticipated. If the wind speed had been accurately predicted, then the utility could have reduced the amount of electricity it generated from fossil fuels; this is wasteful, but not nearly as impactful as overprediction. Thus, improved wind forecasts allow utility managers to better plan their generation needs, which can lead to substantial cost savings.

[2] Wilczak et al. (2019) provide an example of this wind-to-power conversion formula, which we use to estimate both predicted and actual wind energy.

We compare the 12-h wind forecasts from the two versions in each overlap period relative to actual surface-based METAR observations.[3] Although we know the location of all U.S. wind turbines (Table 1), we restrict our analysis to those turbines that are within 20 km of a METAR station, representing about 15.5% of the total turbines in the conterminous United States.[4] The 20-km radius is somewhat arbitrary, but we do this because wind speeds vary greatly over even small distances, and the METAR wind speed observations are made at a single geographical point.[5]

As noted above, when a utility fails to produce enough electricity to meet demand, they must often purchase the shortfall on the spot market. To estimate the cost of wind (and thus wind-generated power) overprediction, we multiply the realized shortage (i.e., the difference between the forecasted energy and the realized energy from wind) by the average quarter-hourly, real-time market prices in the relevant regional wholesale electricity market. For our overlap periods, average spot market prices ranged from \$20.46 to \$33.77 MWh^{-1}.[6]

We provide aggregated estimates of wind overprediction costs in Table 3. In each overlap period, both the "old" and "new" models have periods where they provide a better wind speed forecasts than the other. In the first overlap period, HRRR2 is markedly more accurate than HRRR1, resulting in a \$49.9 million cost savings. However, in the second overlap period, HRRR3 provides only slightly better wind forecasts than HRRR2, resulting in a cost savings of \$17.7 million. This is explained by the many cases of overproduction due to wind speeds being biased high for HRRR1 (Table 3), which was improved by the physics adjustments moving to HRRR2 (Table 2). The improvement in the near-surface wind speeds between HRRR2 and HRRR3 are much smaller, similar to the 2-m relative humidity improvements seen in Fig. 1.

Unlike the "overprediction" case, wind "underprediction" creates problems as utilities expect to produce less wind energy than was actually generated, and thus committed to use more costly, nonrenewable energy sources to meet the anticipated demand. To evaluate the wind underprediction costs, we multiplied the difference between the potential realized energy from wind and the forecasted production by the marginal cost of fossil fuel of \$28 MWh^{-1}, which is "the midpoint of the marginal cost of operating fully deprecated gas combined cycle and coal facilities" (Ray 2019). By considering the marginal cost of wind generation at \$2 MWh^{-1}, we have a marginal cost of wind underprediction of \$26 MWh^{-1}. We provide aggregated estimates of wind underprediction costs for both overlap periods in Table 4, which shows that HRRR2 provides a \$46.6 million cost savings over HRRR1 while

[3] While utility managers may rely on ensemble forecasts in practice, for this study we assume the deterministic 12-h HRRR forecast offers preeminent guidance in predicting wind power generation.

[4] Note that in some cases a wind turbine fell within 20 km of two (or more) METAR stations. To avoid double counting, we associate a turbine with the closest METAR station.

[5] In this way, a key assumption is that the reported wind speeds are consistent for all wind turbines within the 20-km zones surrounding a given METAR station. Note that had we expanded the radius, we would have captured more turbines, but we would be less confident in the representativeness of observed wind at the METAR station and the actual wind at the turbine. We also assume that the 10-m comparison of the HRRR forecast and METAR observations is representative of the bias at hub height (e.g., 80 m).

[6] All dollar amounts listed in this paper have been adjusted for inflation.

Table 3. Wind energy financial losses due to overprediction for the two overlap periods indicated in Table 2.

	Electricity generated (thousands of MW)	Extra costs (millions of dollars)	Potential savings (millions of dollars)
"Actual" in overlap period 1	10,713.5	—	—
HRRR1 forecast	2,485.5	63.4	49.9
HRRR2 forecast	528.9	13.4	
"Actual" in overlap period 2	9,176.0	—	—
HRRR2 forecast	1,118.2	37.0	17.7
HRRR3 forecast	583.0	19.3	

Table 4. As in Table 3, but for the financial losses due to wind energy underprediction.

	Electricity generated (thousands of MW)	Extra costs to supply the overgeneration (millions of dollars)	Potential savings (millions of dollars)
"Actual" in overlap period 1	10,713.5	—	—
HRRR1 forecast	3,738.6	97.2	46.6
HRRR2 forecast	1,947.5	50.6	
"Actual" in overlap period 2	9,176.0	—	—
HRRR2 forecast	1,970.7	51.2	14.3
HRRR3 forecast	1,419.5	36.9	

HRRR3 provides a $14.3 million cost savings over HRRR2. Thus, the combined savings from both reducing wind over- and underprediction errors is $96.5 million by upgrading to HRRR2 and $32.0 million by upgrading to HRRR3. There are two things to note about these cost estimates. The first is that these estimates are only for the overlap periods between the two models, which is 14 months for HRRR1 to HRRR2 and 11 months for HRRR2 to HRRR3 (Table 2). Second, these economic impact estimates apply only to 15.5% of the total array of U.S. wind turbines, which were located close enough to METAR observations as described above. Although we cannot say with certainty, it is plausible that the predictions for the remaining wind turbines have similarly improved as HRRR has evolved, generating additional (perhaps 5 times more) cost savings. If true, the estimated cost savings we present here may be quite conservative.

Impact on morning commute times during precipitation events.[7] Adverse weather and forecasts of the same can change travel behavior and outcomes (Khattak and De Palma 1997; Kilpeläinen and Summala 2007). In this section, we examine how accurate precipitation forecasts can mitigate economic losses due to missed work time. Because we are only interested in cases where the forecasts would result in different consequential behaviors, we limit our analysis to the cases where one version of the HRRR model accurately predicts "economically meaningful" precipitation while the other version does not.

We assume that commuters use the 12-h HRRR forecast the evening before work to plan their departure time the following morning. If the forecast calls for nontrivial rain, a worker (i.e., decision-maker) will leave earlier than normal the next morning (i.e., mitigate), allowing them to arrive to their job on time. Economic losses arise when informed workers do not leave early, yet it unexpectedly rains, making the worker late for their job.

Because most economic activity in the United States occurs in cities,[8] we focus on observed precipitation over large U.S. metropolitan statistical areas (MSAs; the number of MSAs in each region of the country is given in Table 1). In practice, there is no universal standard for connecting rainfall intensity in terms of accumulation per hour to observed conditions, as is done with the Beaufort scale and wind speeds.[9] We used accumulation thresholds of 0.25 and 6.25 mm h^{-1} for "moderate" and "heavy" precipitation in our study, which is consistent with the methodology from Tsapakis et al. (2013), who used the same accumulation cutoff values.[10] Note that we do not separately consider the impact of precipitation on commuting when the temperature is near freezing; these conditions provide a difficult situation for forecasters in providing useful messaging to the public (Walker et al. 2019).

[7] Full details of this analysis are in Hartman et al. (2021).

[8] According to the U.S. BEA, metropolitan areas accounted for nearly 90% of U.S. GDP in 2017 (www.bea.gov/system/files/2018-09/gdp_metro0918_0.pdf).

[9] Royal Meteorological Society. The Beaufort scale. Retrieved from www.rmets.org/resource/beaufort-scale; retrieved on 19 July 2018.

[10] Stern et al. (2003) provides an upper-bound estimate of travel delays due to inclement weather that is approximately 3–3.5 times the amount Tsapakis et al. (2013). Therefore, the results of forecast errors can be multiplied by these scale factors.

We assume that the morning commute occurs between 0600 and 1000 local time, with the total number of drivers in each MSA uniformly distributed over this 4-h period.[11] We calculate "labor hours lost" by aggregating the increased travel time across all commuters due to underforecasted precipitation over the entirety of each hourly HRRR overlap period. Based on Tsapakis et al. (2013), we impose a 2% increase in average commuting times for an MSA when there is moderate precipitation, and a 5% increase when precipitation is heavy. The total MSA labor hours lost for any hourly underprediction of meaningful precipitation is the increase in average commuting time multiplied by the number of workers commuting in that MSA during that hour. This is done for each MSA, and is summed up to the BEA region level, resulting in an estimate of total regional labor hours lost due to the underprediction of precipitation in each HRRR variant.

[11] MSA-level commuting data are available from the American Community Survey.

Economists use a variety of tools to quantify a decision's economic impact, including computable general equilibrium (CGE) models (see sidebar). To calculate the economic impacts of underforecasted precipitation, we change the labor supply in our eight regional models by the amount of labor hours lost due to underestimated precipitation in each forecast model. Reducing labor supply negatively impacts an economy's ability to produce goods and services, thus reducing the gross domestic product (GDP). Table 5 shows the economic impacts of differences in anticipated commuting behavior under the alternative precipitation forecasts. While we have eight regional CGE models, we report only the aggregated impacts between the different HRRR versions for the entire conterminous United States.

Our results suggest that had HRRR2 been operational rather than HRRR1 during the overlap period, economic losses due to underforecasted precipitation would have been reduced by \$106.5 million for the conterminous 48 states, the equivalent output from 1,419 workers. However, the economic impact of using HRRR3 relative to HRRR2 is minus \$10.5 million; in other words, the precipitation forecasts for the morning rush hour were slightly worse economically in HRRR3 than in HRRR2.

Why is this the case? One possible answer is that, even though HRRR2 and HRRR3 have similar critical success index (CSI) scores—as seen on the performance diagram (Roebber 2009) in Fig. 3—HRRR2 has a much larger frequency bias than HRRR3 during the second overlap period (Fig. 3), thus was overforecasting the number of precipitation events generally.

Note that we did not assign any economic impact to a false precipitation forecast. If a commuter expects precipitation that never materializes, then they will arrive early at work, resulting in lost leisure time as they wait for their shift to begin. Accordingly, overforecasts actually work to give a false impression (from an economic point of view) that HRRR2 is better than HRRR3, due to the asymmetric cost of the two outcomes. This illustrates one of many difficulties in performing these types of economic impact studies, as what is the cost to society if a person loses a few minutes of their personal time to a false positive forecast?

Impact on agriculture during freeze events. Weather is one of the largest risk variables in agricultural production and there is a breadth of literature exploring the connections between weather and agricultural productivity (e.g., Wang et al. 2018; Eccel et al. 2009; Deschênes and Greenstone 2007). Specifically, according to the Food and Agricultural Organization, more economic losses have been caused by crop freezes in the United States than by any other

Table 5. Economic savings from the two overlap periods for precipitation during the morning commute.

	FTE lost	GDP savings (millions of dollars)	Tax savings (millions of dollars)
HRRR1 forecast	1,189.5	106.5	5.9
HRRR2 forecast	328.2		
HRRR2 forecast	356.0	−10.5	−1.3
HRRR3 forecast	490.9		

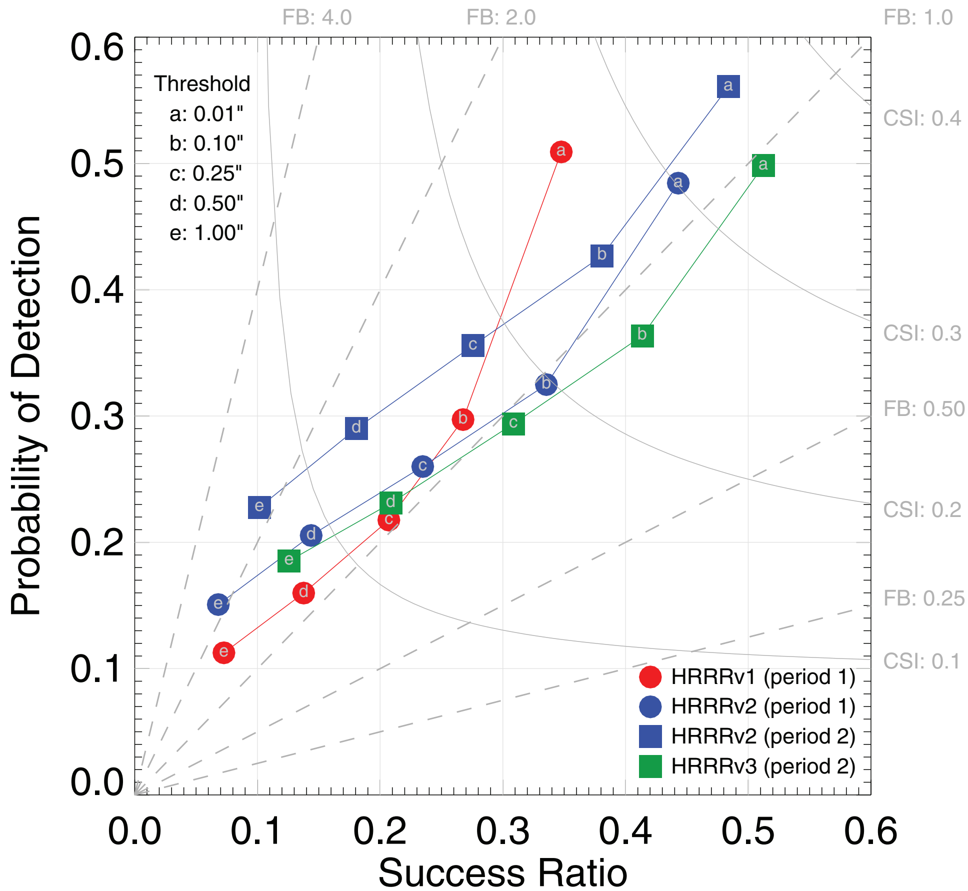

Fig. 3. Performance diagram for 1-h accumulated precipitation over 6 h for the thresholds 0.01, 0.10, 0.25, 0.50, and 1.0 in. (dots labeled "a" through "e," respectively; 1 in. = 25.4 mm) for HRRR forecasts from overlap period 1 (circles) and overlap period 2 (squares). HRRR1, HRRR2, and HRRR3 results are show in red, blue, and green, respectively. Stage IV (Nelson et al. 2016) served as the truth. The observations and model output were degraded to 20-km neighborhoods. The success ratio is defined as 1 minus the false alarm ratio. Dashed diagonal lines represent the frequency bias (FB) from 0.25, 0.50, 1.0, 2.0, and 4.0, and the curved gray lines denote the critical success index (CSI) for values 0.1, 0.2, 0.3, and 0.4.

weather hazard (Snyder and de Melo-Abreu 2005). Skilled weather forecasts have been shown to lead to better outcomes for agricultural producers, if they are aware of the consequences of their decisions and they are rational decision-makers (e.g., Klockow et al. 2010; Meza et al. 2008; Kusunose and Mahmood 2016).

In this section, we evaluate the economic savings from improved freeze forecasts, focusing on specialty crops, such as tree nuts, fruits, and vegetables. Because these crops are "high value," producers often employ preemptive protective measures (e.g., smudge pot heaters and wind machines to induce vertical mixing) when freeze and frost are forecasted. When such measures are taken, the farmer protects their crop, and receives the market value of their product less the cost of mitigation. When the farmer does not protect, the output from their crop is lost or reduced. Because such mitigation measures are costly to implement, decision-makers must carefully consider when to undertake protective actions.

We evaluate the HRRR 12-h forecasts from each hourly initialization, compared to actual temperature observations, to quantify the economic impact of informed crop protection. If the farmer trusts a forecast of an upcoming freeze event and that forecast is correct, then we assume that the farmer mitigates, and losses are (partially) averted. Conversely, if the forecast is "too warm" and the freeze event happens without protective action, then we assume economic loss.

Because we are only interested in "economically meaningful" events, we limit our analysis to documented cases where U.S. producers were adversely affected by a freeze or frost event and received an indemnity payment, comparing the two overlapping HRRR models in such instances.[12] We identify these cases using insurance claims information published by the U.S. Department of Agriculture's Risk Management Agency (USDA RMA; USDA RMA 2016). This dataset summarizes crop indemnity payments, which serve as a proxy for the monetary cost of crop damage and are broken down by both the time of the loss and cause of loss.

We focus only on regions and times where a specialty crop suffered losses. Because not all farmers have crop insurance, we assume that only 74% of the total U.S. specialty crop was covered by insurance (USDA RMA 2016). We consider only the initial freeze event, ignoring continuous, multiday events. Like the first two analyses, we assume that the 12-h HRRR forecast is the only information available to the farmer. Finally, we assume a risk-averse farmer will take protective measures if the HRRR's 2-m temperature forecast is 35°F or lower, and that no action would be taken if the forecasted temperature was above this threshold. We narrow our analysis to cases where one version of the HRRR correctly forecast the freeze event while the other did not.

We consider the reduced agricultural output due to freeze damage as a revenue loss for the agricultural sector. We introduce this to the CGE model as a reduction in export demand for agricultural products.

Like the commuting study, there are marked differences in the economic impact for the different BEA regions. These differences arise from regional differences in specialty crop mixes and the geographic distribution of freeze events (not shown). Because yearly weather events and the timeframe of the HRRR version overlap period plays a significant role in the total economic impact between the various model versions, it is not appropriate to compare the results from the HRRR1 versus HRRR2 analysis and the HRRR2 versus HRRR3 analysis. Overall, we find that had agricultural producers made decisions based on temperature forecasts from HRRR2 instead of HRRR1 during the overlap period, the U.S. economy would have saved about $8.4 million in real GDP. Using HRRR3 instead of HRRR2 in the overlap period would have resulted in a $3.9 million savings.

Summary

Each year, the U.S. government makes substantial investments to improve weather forecast modeling, with the goal of saving lives, protecting property, and increasing economic activity. However, only a few efforts have specifically quantified the economic value of such investments. We have examined how improvements in the HRRR model can translate into additional economic activity in three select aspects of the U.S. economy, enlisting a series of strong, yet plausible assumptions. Overall, we find that there was a marked economic gain for the U.S. economy between HRRR1 and HRRR2, and a smaller, but still appreciable economic gain between HRRR2 and HRRR3.

The economic impacts we estimate are limited in scope, and we do not address some (potentially) important dimensions. For example, in the commuting example, we look at the losses in economic output when a worker arrives late, but not losses to their utility when they leave earlier than necessary due to forecasted rain that never materialized. Additionally, in our agricultural application we do not adjust for the potential costs to producers of false alarms (i.e., instances when the weather forecast incorrectly predicts a freeze event). Accordingly, such findings can be considered a "one-way" analysis.

[12] Note that due to the wide range in freeze protection techniques and their associated costs, we do not consider the cases where a version of the HRRR model predicted a freeze that did not happen. In such cases, a decision-maker may unnecessarily choose to mitigate, absorbing an unnecessary cost. With the large geographic area in this study, we are unable to determine the frost protection method that would best be employed in each instance. Typically frost protection is associated with high upfront fixed costs and lower variable costs, meaning in many instances the cost associated with a false alarm is relatively low. Improved forecasts that reduce the instances of incorrect predictions of freeze events could also lead to economic benefits to agricultural producers.

Further, we make several strong assumptions in each case, and changing these assumptions may have implications on the economic impact, perhaps markedly so. One critical assumption is that all decision-makers use only the HRRR model forecast, and act based on these forecasts. This is seldom the case, as decision-makers often have other sources and information, as well as their own personal biases, that can influence their actions. Further, there are ways to correct for systematic errors in weather forecasts (e.g., Cui et al. 2012; Cho et al. 2020), and thus applying such a correction to a given version may impact the decision made from that forecast.[13] As such, our results provide a theoretical upper bound on the value of improvements in the HRRR model in these three economic sectors.

That said, there are many other examples where improvements to the HRRR model may have important economic impacts that we do not explore here. For example, because they are extremely infrequent and hard to characterize with any statistical confidence, we do not consider large, highly impactful events, such as hurricanes. Similarly, we did not look at conditions where lives may have been lost or saved due to a decision from a forecast, such as a flash flood. And our list of sectors is far from exhaustive. For example, the aviation community reaps substantial benefits from accurate forecasts of both convection (which impacts routing of the aircraft) and ceiling heights (which impacts the operational availability of airports); we do not consider these effects. Future work will explore some of these effects, as well as estimate any gains in moving from HRRR3 to HRRR4, the latter of which became operational at NCEP in December 2020.

[13] Due to the relatively frequent biyearly updates of the HRRR at NCEP, long multiyear datasets needed to derive regional bias corrections for different weather conditions are not available. Thus, there is no operational postprocessing of the HRRR forecasts to remove biases.

Acknowledgments. We thank Kevin Kelleher and Jenn Mahoney, the former and current directors of the NOAA Global Systems Laboratory, for their encouragement and support of this work. We would also like to thank Christian Kummerow, the director of the Cooperative Institute for Research in the Atmosphere at Colorado State University. We thank Matt Wandishin for comments on an earlier version of this paper. This project was partially supported by the NOAA Atmospheric Science for Renewable Energy (ASRE) Program.

Data availability statement. The HRRR model output is very voluminous. Only the two-dimensional HRRR output (i.e., surface fields) were used in this analysis, which can be found on the HRRR archive at the University of Utah (http://hrrr.chpc.utah.edu). These data can also be found on the Google Cloud Platform (https://console.cloud.google.com/marketplace/product/noaa-public/hrrr) or the Amazon Web Services (https://registry.opendata.aws/noaa-hrrr-pds) HRRR archive pages.

CGE models

Economists commonly use computable general equilibrium (CGE) models to describe an economy and analyze the effects of some change to the system. CGE models allow analysts to quantify predicted changes in a variety of economic indicators, including employment, income, and federal and state tax revenue. Partridge and Rickman (2010) describe general CGE modeling when applied to regional analysis. Specific studies include Ballard et al. (1985) and Cutler et al. (2018), who used CGE models to examine optimal tax policy. Rose and Liao (2005), Kajitani and Tatano (2018), and Attary et al. (2020) used CGE models to analyze the economic impacts of natural disasters.

The CGE framework is a numerical model founded in microeconomic theory that characterizes the economic interactions between regional households, the private sector, and the government, including how each respond to some change. In practice, CGE models are semiempirical models built using data from federal, state, and local sources that describe the economic behavior of households, firms, and the government. These data underlie a social accounting matrix (SAM) that quantifies the regional flow of resources. In Fig. SB1, "households" represent all those who live in the region and supply their labor to firms; these

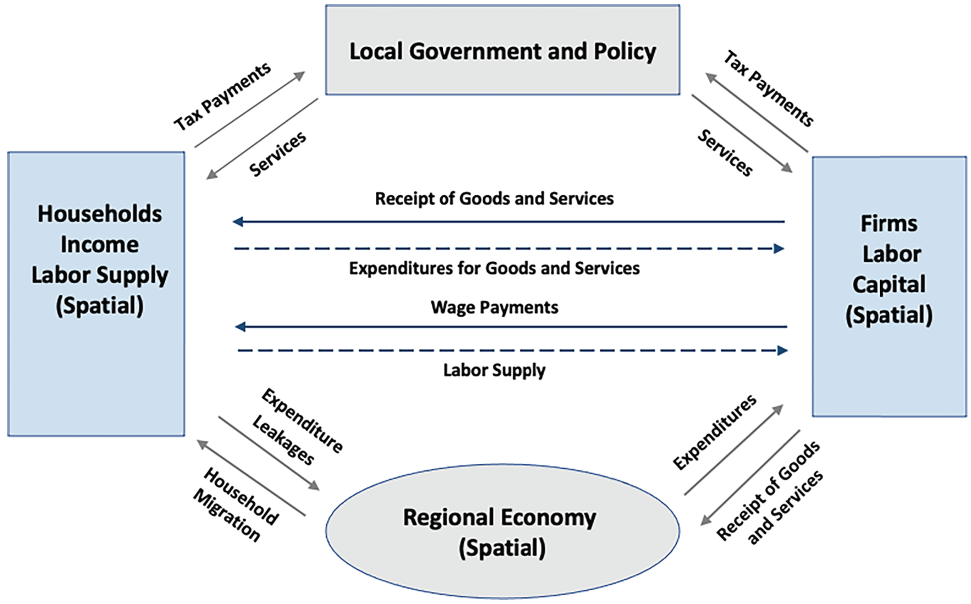

Fig. SB1. Graphical representation of the regional CGE model.

households are differentiated by income. "Firms" use labor and capital to produce goods and services, making capital and labor payments to households in exchange. The households in turn purchase the goods and services produced by firms. Both actors pay taxes to the government in exchange for public services, such as infrastructure, education, police and fire protection, and national defense.

With respect to the commuting example, a forecast error where precipitation is not predicted results in a worker being late to work, causing the supply of labor to fall. This results in lower output worker earnings. For the agricultural scenario, unexpected freezing temperatures results in crop losses, and thus a fall in farm revenue.

References

Angevine, W. M., J. B. Olson, J. Kenyon, W. I. Gustafson Jr., S. Endo, K. Suselj, and D. D. Turner, 2018: Shallow cumulus in WRF parameterizations evaluated against LASSO large-eddy simulation. *Mon. Wea. Rev.*, **146**, 4303–4322, https://doi.org/10.1175/MWR-D-18-0115.1.

Attary, N., H. Cutler, M. Shields, and J. van de Lindt, 2020: The economic effects of financial relief delays following a natural disaster. *Econ. Res. Syst.*, **32**, 351–377, https://doi.org/10.1080/09535314.2020.1713729.

Ballard, C. L., J. B. Shoven, and J. Whalley, 1985: General equilibrium computations of the marginal welfare costs of taxes in the United States. *Amer. Econ. Rev.*, **75**, 128–138.

Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Cho, D., C. Yoo, J. Im, and D.-H. Cha, 2020: Comparative assessment of various machine learning based bias correction methods for numerical weather prediction model forecasts of extreme air temperatures in urban areas. *Earth Space Sci.*, **7**, e2019EA000740, https://doi.org/10.1029/2019EA000740.

Crum, T. D., and R. L. Alberty, 1993: The WSR-88D and the WSR-88D Operational Support Facility. *Bull. Amer. Meteor. Soc.*, **74**, 1669–1687, https://doi.org/10.1175/1520-0477(1993)074<1669:TWATWO>2.0.CO;2.

Cui, B., Z. Toth, Y. Zhu, and D. Hou, 2012: Bias correction for global ensemble forecast. *Wea. Forecasting*, **27**, 396–410, https://doi.org/10.1175/WAF-D-11-00011.1.

Cutler, H., M. Shields, and S. Davies, 2018: Can state tax policy increase economic activity and reduce inequality? *Growth Change*, **49**, 142–164, https://doi.org/10.1111/grow.12216.

Demuth, J., J. Lazo, and R. Morss, 2011: Exploring variations in people's sources, uses, and perceptions of weather forecasts. *Wea. Climate Soc.*, **3**, 177–192, https://doi.org/10.1175/2011WCAS1061.1.

Deschênes, O., and M. Greenstone, 2007: The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather. *Amer. Econ. Rev.*, **97**, 354–385, https://doi.org/10.1257/aer.97.1.354.

Eccel, E., R. Rea, A. Caffarra, and A. Crisci, 2009: Risk of spring frost to apple production under future climate scenarios: The role of phenological acclimation. *Int. J. Biometeor.*, **53**, 273–286, https://doi.org/10.1007/s00484-009-0213-8.

Hartman, B., H. Cutler, M. Shields, and D. D. Turner, 2021: The economic effects of improved precipitation forecasts in the United States due to better commuting decisions. *Growth Change*, **52**, 2149–2171, https://doi.org/10.1111/grow.12542.

Hoen, B. D., J. E. Diffendorfer, J. T. Rand, L. A. Kramer, C. P. Garrity, and H. E. Hunt, 2018: United States wind turbine database, version 1.0. U.S. Geological Survey, accessed 12 August 2020, https://doi.org/10.5066/F7TX3DN0.

James, E. P., S. G. Benjamin, and M. Marquis, 2017: A unified high-resolution wind and solar dataset from a rapidly updating numerical weather prediction model. *Renewable Energy*, **102**, 390–405, https://doi.org/10.1016/j.renene.2016.10.059.

Kajitani, Y., and H. Tatano, 2018: Applicability of a spatial computable general equilibrium model to assess the short-term economic impact of natural disasters. *Econ. Syst. Res.*, **30**, 289–312, https://doi.org/10.1080/09535314.2017.1369010.

Khattak, A. J., and A. De Palma, 1997: The impact of adverse weather conditions on the propensity to change travel decisions: A survey of Brussels commuters. *Transp. Res.*, **31A**, 181–203, https://doi.org/10.1016/S0965-8564(96)00025-0.

Kilpeläinen, M., and H. Summala, 2007: Effects of weather and weather forecasts on driver behaviour. *Transp. Res.*, **10F**, 288–299, https://doi.org/10.1016/j.trf.2006.11.002.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea. Forecasting*, **24**, 1691–1705, https://doi.org/10.1175/2009WAF2222201.1.

Klockow, K., R. McPherson, and D. Sutter, 2010: On the economic nature of crop production decisions using the Oklahoma Mesonet. *Wea. Climate Soc.*, **2**, 224–236, https://doi.org/10.1175/2010WCAS1034.1.

Kusunose, Y., and R. Mahmood, 2016: Imperfect forecasts and decision making in agriculture. *Agric. Syst.*, **146**, 103–110, https://doi.org/10.1016/j.agsy.2016.04.006.

Lazo, J. K., R. E. Morss, and J. L. Demuth, 2009: 300 billion served: Sources, perceptions, uses, and values of weather forecasts. *Bull. Amer. Meteor. Soc.*, **90**, 785–798, https://doi.org/10.1175/2008BAMS2604.1.

——, M. Lawson, P. H. Larsen, and D. M. Waldman, 2011: U.S. economic sensitivity of weather variability. *Bull. Amer. Meteor. Soc.*, **92**, 709–720, https://doi.org/10.1175/2011BAMS2928.1.

——, H. R. Hosterman, J. M. Sprague-Hilderbrand, and J. E. Adkins, 2020: Impact-based decision support services and the socioeconomic impact of winter storms. *Bull. Amer. Meteor. Soc.*, **101**, E626–E639, https://doi.org/10.1175/BAMS-D-18-0153.1.

Meza, F., J. Hansen, and D. Osgood, 2008: Economic value of seasonal climate forecasts for agriculture: Review of ex-ante assessments and recommendations for future research. *J. Appl. Meteor. Climatol.*, **47**, 1269–1286, https://doi.org/10.1175/2007JAMC1540.1.

Murphy, A. H., 1969: On expected-utility measures in cost-loss ratio decision situations. *J. Appl. Meteor.*, **8**, 989–991, https://doi.org/10.1175/1520-0450(1969)008<0989:OEUMIC>2.0.CO;2.

Nelson, B. R., O. P. Prat, D.-J. Seo, and E. Habib, 2016: Assessment and implications of NCEP stage IV quantitative precipitation estimates for product intercomparisons. *Wea. Forecasting*, **31**, 371–394, https://doi.org/10.1175/WAF-D-14-00112.1.

Olson, J. B., J. S. Kenyon, W. A. Angevine, J. M. Brown, M. Pagowski, and K. Siuselj, 2019: A description of the MYNN-EDMF scheme and coupling to other components in WRF-ARW. NOAA Tech. Memo. OAR GSD-61, 42 pp., https://doi.org/10.25923/n9wm-be49.

Partridge, M., and D. Rickman, 2010: Computable general equilibrium (CGE) modeling for regional economic development analysis. *Reg. Stud.*, **44**, 1311–1328, https://doi.org/10.1080/00343400701654236.

Peckham, S. E., T. G. Smirnova, S. G. Benjamin, J. M. Brown, and J. S. Kenyon, 2016: Implementation of a digital filter initialization in the WRF Model and its application in the Rapid Refresh. *Mon. Wea. Rev.*, **144**, 99–106, https://doi.org/10.1175/MWR-D-15-0219.1.

Ray, D., 2019: Lazard's Levelized Cost of Energy Analysis, version 13.0. Lazard, www.lazard.com/perspective/lcoe2019.

Roebber, P., 2009: Visualizing multiple measures of forecast quality. *Wea. Forecasting*, **24**, 601–608, https://doi.org/10.1175/2008WAF2222159.1.

Rose, A., and S. Y. Liao, 2005: Modeling regional economic resilience to disasters: A computable general equilibrium analysis of water service disruptions. *J. Reg. Sci.*, **45**, 75–112, https://doi.org/10.1111/j.0022-4146.2005.00365.x.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp.

Smirnova, T. G., J. M. Brown, S. G. Benjamin, and J. S. Kenyon, 2016: Modifications to the Rapid Update Cycle Land Surface Model (RUC LSM) available in the Weather Research and Forecasting (WRF) Model. *Mon. Wea. Rev.*, **144**, 1851–1865, https://doi.org/10.1175/MWR-D-15-0198.1.

Snyder, R. L., and J. P. de Melo-Abreu, 2005: Frost protection: Fundamentals, practice, and economics. FAO Environment and Natural Resources Service Series Rep., Vol. 1, 223 pp., www.fao.org/3/y7223e/y7223e00.htm.

Stern, A. D., V. Shah, L. Goodwin, and P. Pisano, 2003: Analysis of weather impacts on traffic flow in metropolitan Washington, DC. Federal Highway Administration Tech. Rep., 20 pp., https://rosap.ntl.bts.gov/view/dot/51762.

Strobl, E., 2011: The economic growth impact of hurricanes: Evidence from US coastal counties. *Rev. Econ. Stat.*, **93**, 575–589, https://doi.org/10.1162/REST_a_00082.

Thompson, J. C., and G. W. Brier, 1955: The economic utility of weather forecasts. *Mon. Wea. Rev.*, **83**, 249–253, https://doi.org/10.1175/1520-0493(1955)083<0249:TEUOWF>2.0.CO;2.

Tsapakis, I., T. Cheng, and A. Bolbol, 2013: Impact of weather conditions on macroscopic urban travel times. *J. Transp. Geogr.*, **28**, 204–211, https://doi.org/10.1016/j.jtrangeo.2012.11.003.

Turner, D. D., and Coauthors, 2020: A verification approach used in developing the Rapid Refresh and other numerical weather prediction models. *J. Oper. Meteor.*, **8**, 39–53, https://doi.org/10.15191/nwajom.2020.0803.

USDA RMA, 2016: RMA quick facts. United States Department of Agriculture Doc., 4 pp., www.rma.usda.gov/en/Fact-Sheets/National-Fact-Sheets/RMA-Quick-Facts.

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

Wang, S. L., E. Ball, R. Nehring, R. Williams, and T. Chau, 2018: Impacts of climate change and extreme weather on US agricultural productivity: Evidence and projection. *Agricultural Productivity and Producer Behavior*, University of Chicago Press, 41–75, www.nber.org/books-and-chapters/agricultural-productivity-and-producer-behavior/impacts-climate-change-and-extreme-weather-us-agricultural-productivity-evidence-and-projection.

Wilczak J. M., and Coauthors, 2019: Data assimilation impact of in-situ and remote sensing meteorological observations on wind power forecasts during the first Wind Forecast Improvement Project (WFIP). *Wind Energy*, **22**, 932–944, https://doi.org/10.1002/we.2332.