On the Limits of Estimating the Maximum Wind Speeds in Hurricanes

DAVID S. NOLAN
Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida

JUN A. ZHANG
Cooperative Institute for Marine and Atmospheric Science, University of Miami, and Hurricane Research Division, NOAA/AOML, Miami, Florida

ERIC W. UHLHORN
Hurricane Research Division, NOAA/AOML, Miami, Florida

(Manuscript received 25 October 2013, in final form 17 April 2014)

ABSTRACT
This study uses an observing system simulation experiment (OSSE) approach to test the limitations of even nearly ideal observing systems to capture the peak wind speed occurring within a tropical storm or hurricane. The dataset is provided by a 1-km resolution simulation of an Atlantic hurricane with surface wind speeds saved every 10 s. An optimal observing system consisting of a dense field of anemometers provides perfect measurements of the peak 1-min wind speed as well as the average peak wind speed. Suboptimal observing systems consisting of a small number of anemometers are sampled and compared to the truth provided by the optimal observing system. Results show that a single, perfect anemometer experiencing a direct hit by the right side of the eyewall will underestimate the actual peak intensity by 10%–20%. Even an unusually large number of anemometers (e.g., 3–5) experiencing direct hits by the storm together will underestimate the peak wind speeds by 5%–10%. However, the peak winds of just one or two anemometers will provide on average a good estimate of the average peak intensity over several hours. Enhancing the variability of the simulated winds to better match observed winds does not change the results. Adding observational errors generally increases the reported peak winds, thus reducing the underestimates. If the average underestimate (negative bias) were known perfectly for each case, it could be used to correct the wind speeds, leaving only mean absolute errors of 3%–5%.

1. Introduction
Various meteorological centers around the world have the responsibility to report to the public the present and future locations and intensities of tropical cyclones (known in the North Atlantic as tropical depressions, tropical storms, and hurricanes). The “intensity” of a tropical cyclone is usually defined as the maximum 1-min or 10-min mean wind speed at the 10-m observing level that is associated with the weather system at that time (Harper et al. 2010; Office of the Federal Coordinator for Meteorological Services and Supporting Research 2012). Even in the modern era, direct measurements of the winds in hurricanes or other tropical cyclones are quite rare. Occasionally, storms may pass over islands, ships, fixed towers, or buoys with sufficient technologies to report such winds. In the North Atlantic, various aircraft are sent to penetrate the storms (Aberson et al. 2006), but they do not make in situ measurements near the surface of the ocean. Winds reported by dropsondes, parachuted instruments released by the aircraft, are an exception, but they only report surface winds for an instant and at an uncontrolled location (Hock and Franklin 1999; Franklin et al. 2003). Alternatively, dropsondes released in the center of the storm can accurately report surface pressure, which in turn can be related to peak wind speeds (Atkinson and Holliday 1977; Landsea et al. 2004; Knaff and Zehr 2007).

Unfortunately, the benefit of aircraft reconnaissance applies to only about 30% of 6-hourly intensity estimates...
in the North Atlantic (Rappaport et al. 2009) and to virtually none from around the rest of the world. Thus, most estimates of hurricane intensity are a combination of inferences from other, indirectly measured quantities, such as wind speeds measured by satellite-borne scatterometers (Brennan et al. 2009), the structure and symmetry of the infrared satellite images (Dvorak 1984; Velden et al. 2006), or the temperature anomaly of the core of the storm as detected by microwave satellites (Demuth et al. 2004, 2006). In postseason analyses, often referred to as “best tracks,” all available types of information are used in consideration with their known limitations and biases, to produce an estimate of the actual intensity. This best track value is meant to be representative of the intensity of the cyclone for the 6 h around the valid synoptic time (e.g., 1200 UTC), rather than the actual intensity at the exact time given (Landsea and Franklin 2013).

The best-track intensity records have many uses, such as verification of real-time forecasts, reconstruction of past weather events, and providing a climatological history of tropical cyclone activity in the various ocean basins. Given the importance of these records, it is important to consider their accuracy. Two recent papers have addressed this topic with very different methods. Landsea and Franklin (2013) estimate best-track uncertainty through two surveys, about 11 years apart, of the confidence of meteorologists at the National Hurricane Center [those who perform both the real-time and postseason (best track) analyses] in the accuracies of their analyses. From these surveys they estimate the average uncertainty in best-track intensity is about 10 kt (5 m s$^{-1}$). In a separate study, Torn and Snyder (2012) compare best-track intensities for times when aircraft reconnaissance did occur in a storm (providing at least flight-level winds) to times in which it did not. From these comparisons, they estimate a best-track uncertainty of about 8 kt (about 4 m s$^{-1}$), in fairly close agreement with the results of Landsea and Franklin (2013).

One of the most useful instruments developed in recent years for estimating hurricane intensity is the Stepped Frequency Microwave Radiometer (SFMR) (Uhlhorn and Black 2003; Uhlhorn et al. 2007). The SFMR measures the apparent brightness temperature $T_b$ of the ocean surface, which generally increases with surface wind speed. Measurements are made at six closely spaced frequency channels along the flight track, which allows the often significant $T_b$ contribution to changes in $T_b$ caused by rain to be removed. As a result, the surface wind and path-averaged rain rate are simultaneously retrieved at a 1-Hz rate, and retrievals are averaged using a 10-s running mean filter. This instrument, now mounted on both the National Oceanic and Atmospheric Administration (NOAA) P3 research aircraft and the Air Force C-130J reconnaissance aircraft, provides a continuous estimate of the surface wind speed directly below the aircraft as it flies through the storm, usually in a “figure 4” pattern that crosses through the center twice at right angles. Although such a flight pattern provides multiple radial profiles of the surface wind speed across the storm center, the chance of observing the actual peak wind speed during any time period remains quite small, leading to a potential underestimate of the maximum wind. In past years, meteorologists have compensated for this underestimate by reporting a slightly higher intensity than the fastest representative SFMR wind.

Recently, Uhlhorn and Nolan (2012, hereafter UN12) attempted to quantify this underestimate through an approach that borrows from the framework of an observing system simulation experiment (OSSE). They used a previously validated, high-resolution numerical model simulation of Hurricane Isabel (2003) as a reference dataset (sometimes referred to as a “nature run”), and computed synthetic data from a hypothetical SFMR instrument on an aircraft flying through the storm. They found that, on average, the highest SFMR wind from a single figure-4 pattern flown through a major hurricane will underestimate the true peak 1-min wind speed by 8.5%.

Of course, the best estimates of wind speeds in hurricanes come from fixed instruments near the ocean surface. All of the aforementioned techniques are validated to some extent against such measurements, either “locally” (e.g., a scatterometer wind is compared to a buoy report) or “globally” in the sense that a best-track intensity estimate may have the benefit of a well-placed surface wind measurement. Given the fairly large average underestimate determined by UN12 for the SFMR data, it seems certain that a single, or even several, anemometers are very unlikely to experience the maximum winds.

This paper investigates the limits of estimating hurricane intensity from surface instruments. In particular, the primary goal of this study is to estimate the mean absolute errors and biases between one or more surface instruments that directly observe the most damaging part of a storm (such as a primary rainband or eyewall) and the true peak wind speeds and surface pressures that may be occurring in nearby locations but are not directly observed. Our approach is very similar to UN12 in that we use a high-resolution simulation of a hurricane to provide a realistic, time-evolving hurricane wind field. The results of this study will provide guidance for those estimating hurricane intensity, either in real time or retrospectively, for the occasions when surface instruments do survive direct contact with a tropical storm or hurricane.
Section 2 describes the hurricane nature run used for the analysis. Section 3 describes the winds reported by single anemometers embedded in the storm and by “optimal” arrays of anemometers that are capable of reporting the true, peak wind speed. Section 4 demonstrates the mean biases and errors of “suboptimal” arrays with one or just a few anemometers. Section 5 describes the effects on the results of enhancing the simulated winds with more realistic fluctuations and also by adding observational errors. Conclusions are provided in section 6.

2. Simulated hurricane wind fields

The hurricane nature run (HNR) is a simulation of a hurricane previously created and analyzed by Nolan et al. (2013). The simulation was produced using the Weather Research and Forecasting Model (WRF) version 3.2.1. The model domain covers much of the tropical Atlantic with a 27-km grid and uses multiply nested grids that follow the center of the cyclone with 9-, 3-, and 1-km grid spacings. Three-dimensional wind, temperature, and moisture fields from the 27-, 9-, and 3-km grids were saved every 30 min over the 13-day period from 0000 UTC 29 July to 0000 UTC 10 August 2005. The simulation depicts the entire life cycle of an Atlantic hurricane, from genesis from an African easterly wave to recurvature into the North Atlantic. Data from the 1-km grid were saved every 6 min over the entire 13-day period.

The realism of the hurricane simulated in the HNR, hereafter referred to as NRH1, was validated by extensive comparison of the storm properties to various observational datasets. These validated properties include the pressure–wind relationship, the boundary layer structure, the size and slope of the eyewall, and the frequency distributions of vertical velocity and simulated reflectivity in the eyewall region. NRH1 compares remarkably well to real hurricanes in most of these aspects, with the exceptions being that its eyewall slopes outward more than storms of similar sizes [outward slope and size are correlated; see Stern and Nolan (2009)], and the vertical variations of its vertical velocity distribution have some features that are not seen in observations.

Figure 1 shows the track (center locations), minimum surface pressure, and peak instantaneous 10-m wind speed every 3 h for NRH1 as computed from the 1-km grid. As in Nolan et al. (2013), the storm centers were computed as the location of the minimum value of surface pressure after the pressure field has been smoothed 100 times with a 1–1–1 filter in both directions.

After the HNR was completed, additional datasets were generated by restarting the simulation at four different times, each corresponding to different stages of the hurricane life cycle: a sheared tropical storm at 0000 UTC 2 August; immediately after a period of rapid intensification at 0000 UTC 4 August; a mature, steady-state hurricane at 0000 UTC 6 August; and during

---

1 Note that these dates and times are arbitrary, because the HNR is downscaled from a 13-month simulation of global weather generated by the European Centre for Medium-Range Weather Forecasts (ECMWF). However, for convenience and clarity we will describe the simulated storm using such dates.

2 At the present time, all model output from the HNR is freely available; interested parties should contact the first author.
curvature as the storm accelerates toward the North Atlantic at 0000 UTC 10 August. Hereafter these periods are referred to as STS, ARI, MTR, and REC, respectively. Information about each period is provided in Table 1.

The model output stream was modified so that the surface pressure and the zonal and meridional winds at 10 m on the 1-km grid were saved every 10 s. The center positions, minimum pressures, and peak total wind speeds every 10 min from these high-frequency output datasets are also overlaid onto the plots in Fig. 1. The plots show that sampling a hurricane simulation or forecast at higher frequency can report peak winds significantly higher or lower than what might be obtained, by chance, from infrequent sampling (such as every 3 or 6 h). This issue was discussed in more detail in section 4.3 of Nolan et al. (2013).

3. Peak winds from optimal observing systems

a. Configuration of the optimal observing systems

We begin with NRH1 in stage MTR at 0300 UTC 6 August, exactly halfway through the 6-h period of high-frequency output. A hypothetical optimal observing system is constructed as a network of regularly spaced anemometers fixed over the ocean and in the direct path of NRH1. The anemometers record perfect measurements of 10-m wind speed and surface pressure. Figure 2 shows a map of surface wind speeds at 0300 UTC, along with the positions of anemometers distributed between 23.1° and 24.7°N, 62.2° and 60.6°W at intervals of 0.1° latitude or longitude between them. This network consists of a grid of 17 × 17 anemometers approximately spaced by 11 km, centered on the storm at this time.

The latitude and longitude points that define the network do not necessarily coincide with the locations of model grid points. An additional complication is that the grid is moving in time. Therefore, for each output time, the wind speed of each anemometer is taken from the model grid point closest to the given latitude and longitude location. Since the model grid spacing is 1 km, the differences between the defined anemometer coordinates and the gridpoint coordinates will be less than 0.005°. When the model grid relocates, new grid points are relocated onto the same latitude and longitude held previously by other grid points. Therefore, the offset between the defined anemometer location and its nearest grid point is always the same, and discontinuities or “jumps” in the wind speed do not occur when the nested grids move.

Now suppose we are tasked with determining the intensity (the 1-min wind speed that is associated with the storm) for NRH1 at the time 0300 UTC, in the sense of the best-track methodology. (In practice, best tracks are computed for times such as 0000 and 0600 UTC, but since our datasets range from 0000 to 0600 UTC, we will shift our perspective by 3 h.) In other words, what wind speed is most representative of the intensity at the time of 0300 UTC, plus or minus 3 h?

We first consider the maximum 1-min wind that is occurring in the storm at a given time, \( V_{1\text{peak}} \), and its
global maximum over the 6-h interval, $V_{1\text{maxpeak}}$. One might expect that the aforementioned observing network of 289 perfect anemometers would be likely to provide a very close estimate of these values. Figure 3 shows time series from 9 of the 289 anemometers, each of which is the center anemometer of rows 1, 3, 5, and so on to row 17 at the north edge of the array. Also shown are the maximum at each time over all 289 anemometers (magenta curve), a “perfect” $V_{1peak}$ to be explained shortly (thick black curve), and the $V_{1peak}$ estimated by applying the gust factor formula of Nolan et al. (2013) to the instantaneous model output (red curve). This plot illustrates the measurement limitations of a small number of anemometers. Even for a north–south column of nine anemometers centered exactly on the storm, very few of them come close the true intensity, which is best represented by the thick black curve. Even though the center of the storm passes very close to the anemometer from row 9, it does not record the fastest wind speed. The anemometer from row 11 does appear to coincide with $V_{1peak}$ at $t = 98$ min, but $V_{1peak}$ often rises $5 \text{ m s}^{-1}$ or more above this value at numerous other times in the interval. The highest wind speed is observed by the anemometer from row 13 (on the right side of the eyewall) at $t = 155$ min, but this is still a few meters per second below the maximum at this time. The magenta curve shows that even a network consisting of all 289 anemometers in the $17 \times 17$ array can underestimate $V_{1peak}$ by as much as $5 \text{ m s}^{-1}$ at any given time.

How many anemometers does it take to converge to the truth? $V_{1\text{maxpeak}}$ for the $17 \times 17$ array is $63.73 \text{ m s}^{-1}$. For identically configured arrays with $33 \times 33$, $65 \times 65$, and $129 \times 129$ anemometers, the values for $V_{1\text{maxpeak}}$ are $66.30$, $66.30$, and $66.61 \text{ m s}^{-1}$, respectively. Although there is a tiny increase for the $129 \times 129$ array, we will hereafter use arrays of size $65 \times 65$, with anemometers separated by $0.025^\circ$ of latitude and longitude, as defining the perfect observing system. Although the model grid spacing is 1 km, both numerical and explicit diffusion in the model prohibit significant variations on scales less than 5–7 km, as discussed in Skamarock (2004). A real hurricane wind field could have significant variability on smaller scales, although it is not immediately evident if this variability would carry over to the 1-min mean wind. We will return to this point in later sections.

The thick black curve in Fig. 3 shows $V_{1peak}$ from the perfect observing system (the $65 \times 65$ array). We can also see from Fig. 3 that the wind speed estimate using the gust factor formula of Nolan et al. (2013) quite consistently overestimates the 1-min wind, occasionally by as much as $4 \text{ m s}^{-1}$. The scatter in the data used to develop such formulas is very large [see Fig. 12 of Nolan et al. (2013) and Fig. 4 of UN12] and, therefore, significant differences are not surprising.

While the NHC in the United States uses 1-min wind speeds to describe hurricane intensity, other international centers such as the Japan Meteorological Agency and the Bureau of Meteorology in Australia use 10-min wind speeds (Harper et al. 2010). Figure 4a is identical to Fig. 3, but uses wind speeds computed from 10-min running averages (i.e., $V_{10\text{peak}}$ and $V_{10\text{maxpeak}}$). Not surprisingly, the observed values and their variability are considerably less than for 1-min winds. In this case, the anemometer from row 13 does come very close to reporting the actual $V_{10\text{peak}}$ at $t = 160$ min, and this maximum is not far off from $V_{10\text{maxpeak}}$. Still, many of the other anemometers that are on the right side of the storm, but not as ideally located, report $V_{10\text{peak}}$ values that are occurring in the storm.

Figure 4b shows the same analysis for instantaneous surface pressure. Only anemometers that experience the eye of the storm (as suggested here by winds less than $20 \text{ m s}^{-1}$; see Fig. 2 and Fig. 3) come close to the global pressure minimum. Note here and throughout the
we presume that all our perfect anemometers are conjoined to perfect pressures sensors that also provide data every 10 s.

Figure 5 shows surface wind speeds and anemometer arrays for the three other times of interest. For stage ARI, the array was centered close to the center of the storm, as for MTR. For REC and STS, the array was chosen subjectively so that the peak winds for several hours before and after 0300 UTC would be observed by as many of the anemometers as possible.

b. Comparison of simulated winds to tower observations

As noted in the introduction, the results above and those that follow are only meaningful if the simulated hurricane wind fields are sufficiently realistic. Unfortunately, there are no direct observations of
instantaneous, spatially varying wind fields in hurricanes. Surface wind fields from operational satellite-based scatterometers do provide spatially varying winds, but not at resolutions less than 12.5 km (Brennan et al. 2009). The SFMR instruments that have been mounted on the NOAA P-3 research aircraft continuously since 1998 only measure surface wind speeds in a path about 1 km wide directly below the aircraft and, thus, cannot be used to reconstruct the spatial variability of the surface winds. Spatially varying horizontal wind fields above the surface have been derived from dual-Doppler analyses of winds observed by the Doppler radar mounted on the P3 aircraft, but these also are limited to resolutions of 2–3 km and do not represent the instantaneous wind fields (Marks and Houze 1984; Reasor et al. 2000, 2005).

However, we can compare the simulated anemometer time series to actual surface observations in hurricanes. The observations that we use are from four portable towers deployed at locations near the paths of Hurricanes Frances (2004), Ivan (2004), and Jeanne (2004) at landfall (Zhu et al. 2010; Zhang et al. 2011). Here we use data from when the towers encountered the highest wind speed during the observation period. The wind data were measured by a 3D Gill propeller anemometer installed at 10 m with 10-Hz sampling rate. Detailed descriptions of the instrumentation are provided in Masters (2004). Since all of the towers were on land, it is likely that the intensity of the fluctuations due to turbulence and small-scale features such as streamwise rolls would be larger for the towers than for similar instruments over the ocean (see Zhang et al. 2011). However, all towers were placed in open terrain conditions (e.g., airports) 5 km or less from the coastline and the wind directions were primarily (though not exactly) onshore during the peak wind periods. Thus, we believe that the intensity of the fluctuations measured by the towers is only a modest overestimate of intensity of fluctuations over the ocean.

Data from one of the towers, reported every 0.1 s, and the wind reported by a simulated anemometer, reported every 10 s, are compared in Fig. 6a. Substantial high-frequency variability (gustiness) is evident in the tower data and clearly not present in the simulated winds. Figure 6b shows the same data averaged to 1-min running means computed every 10 s. The amplitude of the variability on the time scales of a few minutes is quite similar, but there is still greater high-frequency variability in the tower data. These differences are illustrated further with power spectrum analyses. Figure 6c shows the spectral power density averaged over all anemometers reporting from the optimal (65 × 65) observing systems for each of the four storm days, as well as for the four towers. The power in the variability of simulated winds at frequencies around 0.017 Hz (1 min) is clearly less, while at 0.0017 Hz (10 min) the observed and simulated spectra nearly match. Figure 6d compares the power spectra of the towers to a few anemometers with winds exceeding 55 m s\(^{-1}\) at some time during the interval. The power around 0.017 Hz is closer, but still lacking.

The analysis shows that the surface winds generated in the hurricane nature run have less variability than those that occur in real storms, even for 1-min means. This suggests that the results of this study will underestimate the differences in wind speeds between anemometers and the truth since real winds undoubtedly have even greater spatial variability at small scales. This matter will be revisited in section 5.

4. Peak winds from suboptimal observing systems

4a. Suboptimal sampling at the mature stage

We consider the following thought experiment. Suppose a meteorologist is attempting to assess the intensity of NRH1 during stage MTR at 0300 UTC (recall we have shifted our synoptic analysis times by 3 h). Let us assume that in situ measurements from either the Air Force or NOAA reconnaissance aircraft are not available, and for the purposes of our thought experiment, let us put aside any objective or subjective satellite intensity estimates (that are almost always available). But, suppose NRH1 has happened to travel directly over a fixed tower in the middle of the Atlantic that reports wind speeds and surface pressure, and the tower data are known to be very accurate. The tower survives direct contact with NRH1 and reports a full record of data over the period from 0000 to 0600 UTC. Finally, suppose this tower appears to be ideally located to report the actual intensity: it is on the right side of the track of the storm and clearly falls inside the RMW. Without aircraft reconnaissance the meteorologist would not necessarily have an accurate estimate of the RMW, but let us give him or her the benefit of having an accurate estimate on hand.

The towers from rows 11, 13, and perhaps 15 as shown in Figs. 2 and 3 would meet these criteria, and yet Fig. 3 already shows that the peak wind reported by any of these towers would underestimate the global maximum intensity \(V_{\text{max peak}}\) by 5 m s\(^{-1}\) or more. What would the average low-intensity bias be for a single, perfect anemometer that appears to be ideally suited to measure the intensity of the hurricane? To answer this question, we record the peak 1-min wind speed of randomly chosen anemometers that meet these criteria: they are
on the right side of the track and the center of the storm passes within 50 km. A total of 1364 of the 4225 anemometers meet these criteria.

This sampling strategy is depicted in Fig. 7, which shows the maximum wind speed recorded by each anemometer in the 65 × 65 perfect observing system: such a figure is typically referred to as a “wind swath.” The figure shows that there is considerable structure to the wind swath on scales of 5–10 km. As discussed in Nolan et al. (2009), local 1-min wind maxima are caused by small-scale vortices, also known as eyewall vorticity maxima (EVMs) (Marks et al. 2008), that are swept around the eyewall. The location of $V_{1\text{max peak}}$, indicated by the large plus sign, is on the right side of the eyewall and appears to be associated with the passage of one of these features.

Also shown in Fig. 7 is the center of the storm at 10-min intervals and 50 randomly chosen anemometers that meet the above criteria. The track shows the familiar trochoidal wobble discussed by Nolan et al. (2001). Points are considered “right of track” if they lie north of a line that runs through the first and last points of the track over the 6-h period. We randomly choose 2000 such points that also fall within 50 km of the track. In this selection strategy we do not exclude choosing some points repeatedly and, since there are only 1364 points available, many points will be chosen more than once. We take this approach for consistency with tests using multiple anemometers, as will be described shortly. The number 2000 is chosen so that numerical results are consistent to within less than 1% for repeated tests and that the results (to be shown)
vary consistently as the number of anemometers increases.

The average of the maximum wind speeds reported by a single anemometer after 2000 random selections is $57.5 \text{ m s}^{-1}$. Comparing to $V_{\text{1maxpeak}}$ ($66.3 \text{ m s}^{-1}$) indicates that, on average, a single anemometer will report a low-wind speed bias of $8.8 \text{ m s}^{-1}$. The standard deviation of the reported maximum is $3.4 \text{ m s}^{-1}$. Therefore, we find that an intensity estimate for a mature hurricane that is based on a single anemometer, even one that appears to have been ideally located, will be on average biased 17 kt (about $9 \text{ m s}^{-1}$), or 12%, below the actual intensity. A low bias of 23 kt (about $12 \text{ m s}^{-1}$) (the mean plus one standard deviation) would not be unlikely.

We describe the previous thought experiment as corresponding to $N = 1$, where $N$ is the number of anemometers in the network. To continue our experiment, suppose the meteorologist is even more fortunate, and there were two highly reliable anemometers in the direct path of NRH1 (i.e., $N = 2$). We now consider a suboptimal observing network of $N = 2$ distinct anemometers, both of which still meet the same criteria as above. The estimate of $V_{\text{1maxpeak}}$ is the maximum of either anemometer. After constructing 2000 networks of two distinct anemometers, we find that a network of two anemometers would on average report a maximum wind speed of $59.5 \text{ m s}^{-1}$, with a standard deviation of $2.6 \text{ m s}^{-1}$.

The result of continuing this exercise up $N = 10$ is shown in Fig. 8. There is a large decrease in the negative bias as the number of anemometers increases from 1 to 4, but then the bias begins to asymptote to about $-4 \text{ m s}^{-1}$. Not surprisingly, the standard deviation decreases steadily as $N$ increases. In all cases, the true maximum intensity ($V_{\text{1maxpeak}}$) stays outside the range of two standard deviations.

These large underestimates are relative to $V_{\text{1maxpeak}}$. If the interpretation of the representative intensity at 0300 UTC is the average of the maximum wind speed in the storm over the period, hereafter, $V_{\text{1meanpeak}}$, then we should compare to this same measure as reported by the perfect observing system (i.e., the mean of the solid black line in Fig. 3). Note that this is an apples-to-oranges comparison: we are comparing the maximum wind speed from a suboptimal network at any time in the interval to the time mean of the optimal network over that interval. We could compare to the time mean over the suboptimal network, but many of the anemometers will experience the eye of the storm, and thus a network with just a few anemometers will very frequently report anomalously low mean winds over the interval. Conceivably, the eye period could be excluded, but this would require an additional selection process.

The $V_{\text{1meanpeak}}$ reported by the optimal network, $59.3 \text{ m s}^{-1}$, is shown as the dashed line in Fig. 8. The mean bias of a single anemometer is only $-1.8 \text{ m s}^{-1}$ (3.0%) and it is nearly zero for two anemometers. For more than two anemometers reporting data from direct hits during the same 6-h period (which in fact would be extremely rare), the bias becomes significantly positive.

Figure 9 shows results for 10-min wind and minimum surface pressure. Compared to $V_{\text{10maxpeak}}$, the bias in the
10-min wind begins at \(-5.5 \text{ m s}^{-1}\) for \(N = 1\) and decreases to \(-1.9 \text{ m s}^{-1}\) for \(N = 10\). The comparisons to the mean maximum wind \(V_{10\text{meanpeak}}\) are much better, with a bias of \(-1.8 \text{ m s}^{-1}\) for \(N = 1\), nearly zero bias for \(N = 2\), and positive bias for \(N > 2\).

For surface pressure we expect that a meteorologist would consider minimum reported values to be representative only if the anemometer experienced the eye of the storm. For pressure, the sampling strategy is changed to include only anemometers that experience a minimum wind of less than \(5 \text{ m s}^{-1}\) (about 10 kt) during the period (365 anemometers meet this criterion). Figure 9b shows that the biases in minimum pressure are positive and very small, beginning at \(1.1 \text{ hPa}\) for \(N = 1\) and decreasing to \(0.5 \text{ hPa}\) for \(N = 10\).

Returning to the 1-min winds, we consider the effect of limiting the observations to a narrower window around 0300 UTC. As an example, Figure 10 shows the reported winds when the interval is limited from 1.5 h before to 1.5 h after 0300 UTC. The average wind speeds reported by the networks are slightly decreased, leading to larger negative bias as compared to \(V_{1\text{maxpeak}}\). Note that the \(V_{1\text{meanpeak}}\) is actually slightly greater in this period (see Fig. 3) so the biases for \(N = 1\) or \(N = 2\) compared to \(V_{1\text{meanpeak}}\) are slightly increased. Generally, we find that shortening the interval increases the negative biases because it gives the anemometers less opportunity to measure higher wind speed.

b. Suboptimal sampling at other stages

Figure 11 shows results of the same analysis applied to stage ARI. The wind swath shows that at this time the highest wind speeds of NRH1 are confined to a much smaller region, with a faster decrease in wind speeds outside of the RMW, which at this time is about 35 km (see our Fig. 11a, or Fig. 17b of Nolan et al. 2013). The location of \(V_{1\text{maxpeak}}\) is again associated with a wind feature that is rotating around the right-front quadrant of the eyewall, which in this case stands out as much more anomalous compared to the surrounding winds. This suggests it would be even less likely for a limited observing system to report winds that are close to the maximum.

The remaining plots in Fig. 11 shows the average reported 1-min wind, 10-min wind, and minimum pressure for networks with 1 to 10 anemometers. The stronger,
more localized global maximum causes an even larger gap between $V_{1\text{maxpeak}}$ and $V_{1\text{meanpeak}}$, which are 63.2 and 53.9 m s$^{-1}$, respectively, and causes the bias for $N = 1$ to be $-11.2$ m s$^{-1}$, or 17%. However, comparisons to $V_{1\text{meanpeak}}$ are similar to before, with biases of only $-1.9$ and $-0.2$ m s$^{-1}$ for $N = 1$ and 2, respectively.

The 10-min averaging has the effect of substantially reducing the very localized global maximum, and thus greatly reduces the low biases, making them similar in percentage to the results for MTR. Pressure biases are again just a few hectopascals.

Figure 12 shows results for the stage REC. For this case, we sample from anemometers that are within 60 km of the storm center, and they are considered right of track if they are east of the line that connects the center positions at 0000 and 0600 UTC. Initial calculations using the entire 6-h period produced a very large bias, $-11.4$ m s$^{-1}$, for $N = 1$. This is because of the fairly singular $V_{1\text{maxpeak}}$ that appears in the southwest corner of the wind swath (Fig. 12a) that occurs during the first hour of the period. A large drop in peak wind speed after this time can also be inferred from Fig. 1c. To mitigate the effect of this intensity change (and to give our hypothetical meteorologist every chance for success in analyzing the correct intensity), the results shown in Fig. 12 were computed by limiting the period of interest from 0100 to 0500 UTC, producing a 1-min wind speed bias for 1 anemometer of $-6.8$ m s$^{-1}$, or 15%. The biases against $V_{1\text{meanpeak}}$ are slightly larger than above, $-2.7$ and $-0.8$ m s$^{-1}$ ($-6.3\%$ and $-2.2\%$) for 1 and 2 anemometers, respectively. The results for 10-min wind and minimum surface pressure are consistent with the previous results.
For stage STS, choosing how to sample anemometers becomes more complicated. From the wind field at 0300 UTC (Fig. 5c), it is clear that the peak winds are not on the right side of the track. At this time the highest winds are associated with a large area of active convection that is located on the south side of the cyclone and is considerably displaced from the surface pressure center (see also Fig. 5b of Nolan et al. 2013). Again, we will give the meteorologist the benefit of knowing that the highest winds are in this region. Other than being in the region directly affected by the mesoscale convective band, there do not seem to be any other criteria that would identify an anemometer as having had a good chance to experience the maximum winds of the storm. Here we will presume that the meteorologist would know from the history and current structure of the system that it is at least of tropical storm strength. Therefore, our criterion for sampling representative anemometers is that they report a peak 1-min wind of at least 18 m s\(^{-1}\) (about 35 kt), making 3579 anemometers available.

The results are shown in Fig. 13. The \(N = 1\) bias against \(V_{1\text{maxpeak}}\) is \(-5.3\) m s\(^{-1}\) (\(-18.6\%)\), while the bias against \(V_{1\text{meanpeak}}\) is \(-2.6\) m s\(^{-1}\) (\(-10.0\%)\).

Furthermore, we cannot use the optimal observing network depicted previously in Fig. 5c to compute pressure biases, since the pressure minimum moves outside of this network during the period (note the storm is moving to the south-southwest at this time). For the pressure bias we used a similar network of anemometers (mounted with pressure sensors) repositioned to be over the center of NRH1 at 0300 UTC (not shown). Using the full period, 0000–0600 UTC, results in the pressure biases shown in Fig. 13d. In this case, the pressure bias has become very large, ranging from 8 to...
For all four cases, the biases relative to related wind fields are representative of real hurricane from about 3% to 1.5%. To the extent that our simu-
for MTR and ARI), the MAE values are less, ranging around 2.5% for
MAE values start around 5% for the known bias. For all four stages, the 1-min wind
would remain, due to the inherent variability of the correct for the bias. Nonetheless, some mean error
correct, and could be extended to tropical cyclones of
(SErse et al. 2004; Landsea and Franklin 2013).
Currently, such adjustments are made subjectively
son would then adjust the best-track intensity upward.
provided by a small number of anemometers. This per-
the significant negative biases associated with records
the negative biases are less than 5% for
spatially. Thus, the 1- and 10-min winds are reported
every 10 s. In reality, most instruments that report mean
winds only do so at the time interval of the mean (e.g.,
1-min winds every 1 min), or even less often.
Many of the results of section 4 were repeated only
using the 1-min or 10-min winds every 1 or 10 min,
respectively. This was found to decrease the mean reported
winds and thus cause small increases in the negative biases. For example, for MTR the 1-min wind bias for
a single-anemometer network was changed from −13.5% to
−14.5%. For the 10-min wind, the bias changed from
−9.5% to −10.6%. This change is due to the fact that reporting only every 1 or 10 min produces less chance for the anemometer to report a slightly faster wind that might occur in between the reporting intervals. The standard deviations of the winds and biases were also decreased by only a few percent (not shown). Although these effects are small, our further calculations in this section will continue with data reported only every 1 or 10 min, to be more consistent with reporting by real instruments.
Next we make the reported winds more realistic by adding the effects of measurement error. While the “instrument” error in the 1-min wind of a properly exposed anemometer on a fixed tower might be very small, we should also consider errors associated with less stable instruments, such as an anemometer on a buoy that pitches and rolls owing to large ocean waves (Howden et al. 2008). Standards for meteorological measurements required by the World Meteorological Organization (WMO) are defined in terms of percentages of the wind speed, with recent requirements for errors of less than 10% for observations to be acceptable for use (WMO 2008). Most instruments have significantly less error in reporting wind speed (e.g., Gilhousen 1987).
FIG. 13. Results for STS: (a) wind swath, center track (barely visible in the top-left corner), and randomly selected eligible anemometers; (b) average reported peak 1-min wind speeds; (c) average reported peak 10-min wind speeds; (d) average reported minimum surface pressures from 0000 to 0600 UTC; (e) surface pressure field and wind vectors at 0550 UTC; and (f) average reported minimum surface pressures from 0000 to 0300 UTC. In (a), the large plus sign shows the location where the fastest 1-min wind ($V_{1\text{maxpeak}}$) occurs.
FIG. 14. Mean bias and bias-corrected errors as a function of $N$ (the number of anemometers in the network) for (a) 1-min wind for MTR, (b) 10-min wind for MTR, (c) 1-min wind for ARI, (d) 10-min wind for ARI, (e) 1-min wind for REC, and (f) 1-min wind for STS.
To illustrate the effects of random measurement error, we first add random errors that are normally distributed with zero mean and standard deviation that is 5% of the current wind speed. In this implementation, the errors are completely decorrelated on the time scale of the 1- or 10-min means. Adding such errors to each of the 10-s observations would result in a much lower variance for the 1- and 10-min winds owing to time averaging over the errors. Instead we choose this approach so that we can explicitly define the errors of the 1- or 10-min mean winds. While 5% is much larger than expected errors for observing 1-min winds in optimal conditions, we start with this arbitrarily large value so that its effect can be clearly seen in the results.

As noted above, mean winds and errors from the truth are computed by sampling 2000 randomly chosen networks of one or more anemometers in the path of the storm. Since we are adding random measurement errors, the most comprehensive approach would be to accumulate the statistics of the biases over a large number of measurements with different random errors for each of the 2000 random networks, and then to average these statistics over all 2000 networks. Unfortunately, generating several thousand sets of time series with random errors for each anemometer in the 2000 random networks is computationally impractical. Instead, we take the shortcut of adding one set of different random errors to the anemometer time series each time a random measurement error is chosen.

For stage MTR, Fig. 15 shows the results of adding 5% measurement error to the 1-min and 10-min mean wind speeds. For both metrics, adding measurement error increases the mean reported wind speed for networks of all sizes. For $N = 1$ the increase is $1.6 \text{ m s}^{-1}$. The standard deviation also increases by just $0.08 \text{ m s}^{-1}$. The changes for reported 10-min winds are similar, but diminished in amplitude. These increases occur because adding random errors to the observations causes some of the faster wind speeds to be even faster than the fastest wind observed by perfect anemometers. Random observational errors do not “average out” because we are selecting the peak wind speed reported by any one of one or more anemometers, not an average over time.

As above, these results can be recast into average biases and errors, also shown in Fig. 15. Generally, adding measurement error to the simulated observations increases the measured winds and thus lessens the negative bias from $V_{1 \text{maxpeak}}$. For comparisons to $V_{1 \text{meanpeak}}$, the biases are shifted toward positive values, leading to significant positive biases for $N = 2$ or greater. The very small increases in variance caused by the random errors also lead to very small increases in the bias-corrected errors (e.g., for $N = 1$, MAE increases from 4.2% to 4.3%).

For a network of a given size, the magnitude of this positive bias caused by measurement error can be shown as a function of the error itself. Changes in the bias and bias-corrected error for $N = 1$ (by far the most likely case) are shown in Fig. 16 for errors ranging from 0% to 10%. The 0% case is included to confirm that, as mentioned above, reducing the reporting frequency from every 10 s to 1 or 10 min causes the average reported wind speed to decrease. As expected, Fig. 16a shows a slightly greater negative bias for one anemometer with 0% error than Fig. 14a. The negative biases become less negative as the error increases, rising from $-14\%$ to $-5\%$ for 10% error for comparison to $V_{1 \text{maxpeak}}$. For comparison to $V_{1 \text{meanpeak}}$, adding a measurement error of 6% or more causes the bias to become positive. The results for 10-min winds (Fig. 16b) are similar, though again the variations are diminished. Also shown in Fig. 16 are results for increasing measurement error for networks with $N = 2$. Much like the results for 10-min winds, the biases with $N = 2$ and 0% error are less negative and become less negative (and ultimately, positive, for comparison to $V_{1 \text{meanpeak}}$) as the measurement error increases.

Results for varying measurement error during REC are nearly identical to those for MTR and are not shown. As noted above, the 1-min negative biases were greater for ARI and STS, due to their more isolated wind maxima. Results for $N = 1$ for these two periods are shown in Fig. 17. The shift of the bias from negative toward positive by increasing error is similar for both cases to the results for MTR (cf. Fig. 16a), although it is somewhat larger for STS.

To summarize, we find the somewhat counterintuitive result that adding measurement errors to the instruments lessens the negative bias for the peak wind speeds, while causing only small increases in the variance and the bias-corrected MAE. For the range of likely errors in 1- and 10-min winds, perhaps between 0.5% and 5%, this effect ranges from being negligible to causing a small decrease in the negative bias for comparison to $V_{1 \text{maxpeak}}$ (e.g., in Fig. 16a, from $-9.5\%$ to $-7.3\%$). For comparison to $V_{1 \text{meanpeak}}$, measurement error can bring the bias to near zero or to positive values. Results for networks with $N = 2$ are similar, but like results for 10-min winds, the variations as a function of measurement error are less.

b. Enhancing the simulated winds

As discussed in section 3b, the simulated wind fields of NRH1 do not show sufficient fluctuations on time scales of less than a few minutes. However, it is possible to add such fluctuations to our simulated time series. Our approach is to generate time series of random fluctuations that have the same power spectrum as the
part of the power spectrum that is missing from the simulated winds as can be seen in Fig. 6.

We start with data from the four towers during their strongest wind periods, with mean wind speeds ranging from 31 to 34 m s\(^{-1}\). Their power spectra are computed and then averaged together. A polynomial curve is then fit to the averaged spectrum of the towers. The power spectra of the towers and the fitted curve are shown in Fig. 18a. The formula for this curve is

\[
S_{10}(f) = \frac{630}{(1 + 33f)^{5/3}},
\]

where \(S_{10}\) is the spectral power density of the 10-m wind speed as a function of the frequency \(f\). The form of Eq. (5.1) is modeled after equations for power spectra provided by Kaimal and Finnigan (1994, chapter 2). However, the extension of this curve to low frequencies would add power to the low-frequency variations of the wind field that are already present in the simulated winds. Therefore, we multiply the expression in Eq. (5.1) by an arbitrary function that rapidly forces the power to small values between frequencies of 0.01 and 0.001 Hz:

\[
S_{\text{mod}}(f) = \exp\left(-\frac{\log_{10}(f)^6}{2.5}\right)S_{10}.
\]

This formula is shown as the magenta curve in Fig. 18a. To generate each random time series we take the square root of the power \(S_{\text{mod}}\) at each frequency, multiply each of these amplitudes by a complex number with amplitude 1 and different, random phase angles between \(0\) and \(2\pi\), and then invert the Fourier transform of this set of complex coefficients. These calculations generate random fluctuations at the same frequency as the tower data (i.e., data every 0.1 s). These fluctuations are 

---

**Fig. 15.** Results for adding 5% measurement errors to the anemometer winds during MTR: (a) reported and true peak 1-min winds, (b) reported and true peak 10-min winds, (c) mean biases and bias-corrected errors for 1-min winds, and (d) mean biases and bias-corrected errors for 10-min winds.
averaged to 10-s means to match the output frequency of the model data.

A single realization of the simulated winds with enhanced fluctuations can be generated by adding a random time series of fluctuations to each of the 4225 anemometer time series. However, the power spectrum used to define the fluctuations is derived from tower data with an overall mean wind speed of only 32.5 m s\(^{-1}\). Clearly, the simulated winds vary widely from this value. Therefore, an additional intermediate step is incorporated where the amplitudes of the fluctuations are modulated by the running mean wind speed, conveniently provided by the 10-min wind speed:

\[
V_{\text{enh}} = V_{10s} + \frac{V_{\text{10min}}}{32.5 \text{ m s}^{-1}} F_{10s},
\]

where \(F_{10s}\) are the fluctuations, \(V_{10s}\) and \(V_{\text{10min}}\) are the 10-s and 10-min wind speeds, and \(V_{\text{enh}}\) are the final, enhanced winds. Certainly it is possible that amplitudes of the fluctuations vary nonlinearly with the mean wind speed. However, without a comprehensive dataset of tower-observed wind speeds in hurricanes ranging from low winds up to 70 m s\(^{-1}\), we can only speculate on the relationship between the intensity of the unresolved fluctuations and the 10-min wind speed. If the relationship were nonlinear, for example with the intensity of the turbulence increasing as the square of the wind speed, then the turbulent fluctuations would be even larger for wind speeds above 32.5 m s\(^{-1}\). Linear proportionality is the simplest choice and we use it here.

An example of how the fluctuations change the simulated time series are shown in Fig. 18c. The anemometer
and the time period shown are selected to illustrate how the amplitude of the fluctuations varies with wind speed. Although the fluctuations are damped significantly when the 10-s data are averaged to 1-min wind speed, they still can add or subtract a few meters per second to the peak wind of the anemometer. Averaging to 10-min winds almost completely eliminates the fluctuations (not shown), so further results will not consider the 10-min winds. Note that our method of adding random fluctuations to the simulated anemometers does not appear to enhance the realism of the structure of the wind field. Wind swath plots with the fluctuations added (not shown) only appear as noisier (and louder) versions of the wind swaths shown above.

Since each set of fluctuations is a realization of a random process, our results should be the average over many such realizations for each random network of anemometers. As with the random observational errors, the computation time required to average over many sets of fluctuations for each of the 2000 random networks is quite large. Therefore, we use a similar procedure as for errors, where different random fluctuations are added to the anemometer winds only once for each of the random networks. However, the time required to generate a full set of 4225 fluctuation time series requires an additional shortcut: the full set of 4225 fluctuations is generated only once. For each random network, a new set of simulated winds is generated by “scrambling” the order with which the fluctuations are added to the anemometers. In other words, there is only one set of fluctuation time series, but each one is added to a (very likely) different anemometer every time, during which Eq. (5.3) is applied.

During stage MTR, Fig. 19a shows average reported winds for networks of varying sizes with fluctuations added to the winds. This plot should be compared to the results without fluctuations in Fig. 8 (note that the range on the y axis is greater for the new plot). For \( N = 1 \), the average reported wind speed has increased by about 3.5 m s\(^{-1}\). However, both \( V_{1\text{max peak}} \) and \( V_{1\text{mean peak}} \) have also increased, by about 6.5 and 3 m s\(^{-1}\), respectively. Note that, when fluctuations are added to the simulated winds, these metrics become different for each random realization. The lines shown in Fig. 19 are averages over 10000 such realizations. Figure 19b shows the biases and bias-corrected MAEs. As a result of the large increase in \( V_{1\text{max peak}} \), enhancing the winds leads to an increase in the negative bias (from \(-13.5\%\) to \(-16.0\%) for \( N = 1 \) (cf. Fig. 14a). In contrast, the bias against \( V_{1\text{mean peak}} \) and the bias-corrected errors are hardly changed.

Results for stage STS are also shown in Fig. 19. In contrast to MTR, the increase in \( V_{1\text{max peak}} \) is not greater than the increase in the reported winds; all wind speeds are shifted upward by about 3 m s\(^{-1}\). Thus, the negative bias against \( V_{1\text{max peak}} \) which was largest for STS among the four time periods of the storm (see Fig. 14e), is slightly mitigated from \(-17.8\%\) to \(-17.1\%). The negative bias against \( V_{1\text{mean peak}} \) is also reduced, changing from \(-9.3\%\) to \(-7.5\%). Surprisingly, the results for ARI and REC were more similar to STS: the average \( V_{1\text{max peak}} \) and \( V_{1\text{mean peak}} \) increased by only a little more than the average reported winds, so in terms of percentages the negative biases decreased slightly (not shown).

### c. Results with enhanced winds and observational errors

Finally, we put together the two enhancements described in the previous sections. Considering the previous results, we would expect adding observational
error to shift the reported wind speeds to higher values (despite that reporting every 1 min instead of every 10 s reduces the peak winds), while $V_{\text{1maxpeak}}$ and $V_{\text{1meanpeak}}$ (here averaged over many fluctuations) remain increased by the same amount, thus lessening the bias. This is confirmed in Fig. 20, which shows the reported winds as a function of $N$ with measurement error 5%, and the mean biases and bias-corrected errors as a function of measurement error for $N = 1$, for both MTR and STS. With increasing measurement error, the wind speeds are increased, their standard deviations are slightly larger, the biases shift toward more positive values, and the bias-corrected errors increase.

6. Conclusions
This study attempts to provide quantitative assessments of the accuracy to which the intensity of a hurricane, as...
defined by the peak 1-min or 10-min wind speed associated with the storm, can be estimated by a limited number of instruments at the ocean surface. Our method is to compute synthetic measurements from randomly arranged networks of anemometers in the path of a simulated hurricane. The primary result of this paper is that for a small number \(N \leq 1\) or 2) of anemometers that survive a direct encounter with the strongest part of the storm, the peak 1-min wind speed during a 6-h time interval (defined as \(V_{1\text{maxpeak}}\)) will be underestimated by 10%–20%. The peak 10-min wind speed (\(V_{10\text{maxpeak}}\)) will be underestimated by 5%–10%. As the number of anemometers \(N\) increases from 1 to 10 the bias is reduced, typically changing from about –15% to about –8%. To our knowledge, there are never more than a handful of reliable, contemporaneous reports of surface winds in hurricanes over the ocean.

The biases are significantly different if we choose as our metric of storm intensity the average of the peak 1-min wind over the period (defined as \(V_{1\text{meanpeak}}\)). This metric is in fact more consistent with the meaning of the intensity values reported in the best-track reports that are provided by operational centers. For \(N = 1\), the negative bias between the peak wind reported at any time by an anemometer and \(V_{1\text{meanpeak}}\) can be less than 5%. This negative bias is somewhat greater, 5%–10%, for the more “difficult” cases of a small hurricane or an asymmetric tropical storm. For \(N = 2\) or more, the bias can be close to zero or even positive.

It is worth noting that negative bias values of 5% are less than the negative biases estimated by UN12 for an SFMR instrument flown through a hurricane in a figure-4 pattern (around 8.5%). This indicates that a surface instrument experiencing a “direct hit” by the storm will, on average, provide a better estimate of intensity than a reconnaissance flight with SFMR, provided that we are making the apple-to-oranges conversion of the anemometer peak reported wind to the average peak wind (\(V_{1\text{meanpeak}}\)).

---

**Fig. 19.** The effects of adding fluctuations: (a) reported 1-min winds for MTR, (b) mean biases and bias-corrected errors for MTR, (c) reported 1-min winds for STS, and (d) biases and bias-corrected errors for STS.
Provided that one or more anemometers comes close to the center of the storm (defined here as experiencing winds of less than 5 m s$^{-1}$), average overestimates of minimum surface pressure are very small, ranging from 1 to 3 hPa. An exception can occur for a developing tropical storm, where lower surface pressures can briefly occur in convection displaced from the center. However, such pressure values are not representative of the storm intensity.

If the average biases computed here were sufficiently accurate for real storms, or if they could be estimated on a case-by-case basis, the reported winds could be adjusted to remove the bias. This would leave only random errors in intensity that would represent the true limit of how well hurricane intensity can be estimated from a limited number of surface instruments. For 1-min winds, these bias-corrected mean absolute errors range from around 5% for $N = 1$ to 2.5% for $N = 10$. These errors are less for 10-min winds, ranging from around 3% to 1.5%.

Adding random measurement errors has the effect of bringing the reported peak winds upward and generally closer to the truth, with the interesting consequence of reducing negative biases (or in some cases making them more positive). This effect can be significant: for example, for $N = 1$ during stage MTR (shown in Fig. 16a), the bias against $V_{\text{maxpeak}}$ is $-14.3\%$ for 0% measurement error, $-11.0\%$ for 5% error, and $-3.7\%$ for 10% error. For the bias against $V_{\text{meanpeak}}$, the bias changes from $-3.7\%$ to $+0.4\%$ and then to $+6.9\%$, respectively.

As shown by power spectrum analyses, the time series of winds reported by anemometers in our simulated hurricane do not have as much variability as real time series, even when both are averaged to 1-min means. In an attempt to account for this lack of variability, wind...
speed fluctuations with the same power spectrum as observations were added to the simulated time series. This had the effect of, on average, increasing all the 1-min wind reports ($V_{\text{1minpeak}}$, $V_{\text{1meanpeak}}$ and reported winds) by generally the same amount ($3–5\ \text{m}\ \text{s}^{-1}$) so that the biases and bias-corrected errors were only slightly changed. Using the enhanced winds combined with the effects of observational errors did not produce any significant differences from the results with observational errors alone.

It is important to note that this study and its results do not directly apply to many of the methods currently used by operational centers to estimate hurricane intensity, both in real time, or in the development of post-season best-track reports. Here we have addressed the narrower questions of the limits of estimating intensity from reliable surface instruments, even in unrealistically favorable cases (such as five perfect anemometers experiencing and surviving direct contact with the eyewall of a major hurricane). The results show that the accuracies of these estimates are highly sensitive to the definition of hurricane intensity. Measurements of 10-min winds show considerably less negative bias and variance than measurements of 1-min wind. More importantly, measurements of peak winds show very little bias compared to averages of the actual peak winds over a period of several hours. This further justifies the use of the average value of the peak wind speed over a period of several hours (i.e., $V_{\text{1meanpeak}}$ or $V_{\text{10meanpeak}}$) as the definition of intensity in operational analyses and best-track reports.

Acknowledgments. D. Nolan was supported in part by the NOAA/Office of Weather and Air Quality (OWAQ) through its funding of the OSSE Testbed at the Atlantic Oceanographic and Meteorological Laboratory, by the NOAA/Unmanned Aerial Systems (UAS) Program, and by the Hurricane Forecast Improvement Program (HFIP). J. Zhang was also supported by HFIP. The authors thank the Florida Coastal Monitoring Program (FCMP) for collecting the tower data.

REFERENCES
——, J. A. Zhang, and D. P. Stern, 2009: Evaluation of planetary boundary layer parameterizations in tropical cyclones by


