atmosphere

MDPI

Article

Sustained Wind Forecasts from the High-Resolution Rapid Refresh Model: Skill Assessment and Bias Mitigation

Robert G. Fovell [1,*] and Scott B. Capps [2]

[1] Department of Atmospheric and Environmental Sciences, University at Albany, State University of New York, Albany, NY 12222, USA
[2] Technosylva, Inc., La Jolla, CA 92037, USA; scapps@technosylva.com
* Correspondence: rfovell@albany.edu; Tel.: +1-518-442-4479

Abstract: We examine the skill associated with sustained wind forecasts in the High-Resolution Rapid Refresh (HRRR) model, extending and enhancing previous work. Some utilities use numerical weather prediction models like the HRRR to anticipate electrical transmission line shutdowns for public safety reasons, increasing the importance of forecast accuracy and motivating the need to understand sources of bias and differences among observation networks. We demonstrate that the HRRR forecasts for airport stations are very good albeit with a tendency to underpredict the highest wind speeds and at the windiest locations. Forecasts for non-airport networks are much less accurate owing to a variety of factors, including differences in the way winds are measured and the environments they are measured in, and this results in predictions with excessive temporal variation relative to observations. We demonstrate a practical approach to modifying sustained wind forecasts so that they are more useful proxies for conditions being observed in the field.

Keywords: wind forecasting; model verification; High-Resolution Rapid Refresh; Weather Research and Forecasting model

check for updates

Academic Editor: Jimy Dudhia

Received: 19 November 2024
Revised: 20 December 2024
Accepted: 24 December 2024
Published: 27 December 2024

Citation: Fovell, R.G.; Capps, S.B. Sustained Wind Forecasts from the High-Resolution Rapid Refresh Model: Skill Assessment and Bias Mitigation. *Atmosphere* **2025**, *16*, 16. https://doi.org/10.3390/atmos16010016

1. Introduction

There are many needs and uses for accurate wind forecasts, both of short-period gusts and sustained winds, which are wind samples averaged over a period of time. High-resolution numerical weather prediction (NWP) models are employed for wind energy forecasting e.g., [1–5] and to drive outage forecasting and wildfire spread models e.g., [6], among many other uses. Public utilities in California such as Pacific Gas and Electric (PGE), Southern California Edison (SCE), and San Diego Gas and Electric (SDGE) use NWP forecasts to determine where and when to implement a Public Safety Power Shutoff [PSPS, [7]], an attempt to prevent the initiation of fires associated with power generation equipment. Several notable past wildfires in the U.S., such as the 2007 Witch fire [8,9], the 2017 Thomas and Tubbs fires [10,11], the 2018 Camp fire [12,13], and the 2023 Lahaina fire [14] may have been sparked or exacerbated by power line failures.

The use of NWP wind forecasts in PSPS decision-making significantly elevates the necessity of obtaining accurate wind predictions and an understanding of any forecast biases that might exist. The need for situational awareness has encouraged an enormous expansion of surface weather networks, including those operated by PGE and SCE and commencing with the deployment of the SDGE mesonet in 2010 [15]. In the utility mesonets, many stations represent circuits, and observations and forecasts for those stations determine when a PSPS event is declared or terminated. The wind energy industry also requires skillful forecasts to determine resource availability, a prerequisite to keeping energy supply and demand synchronized [1]. All of this has motivated the development of comprehensive field programs such as the two Wind Forecast Improvement Projects (WFIP; [16–18]), which focused on the wind generation region of the U.S. Great Plains (WFIP1) and on predictions in complex terrain (WFIP2), and efforts such as the present study.

As the National Oceanic and Atmospheric Administration's premiere fine-scale forecasting system, the High-Resolution Rapid Refresh (HRRR) model [19,20] has been the subject of numerous validation studies, including [17,21–25], to name but a few. At this writing, the HRRR employs 3 km grid spacing and is based on the Weather Research and Forecasting (WRF) model's Advanced Research WRF core [26]. The HRRR launches new cycles hourly. As of Version 4 (HRRRV4), which became operational in December 2020, the 00, 06, 12, and 18 UTC cycles extend to 48 h, while the forecast length is 18 h for the other cycles.

This study extends our prior HRRR wind forecast verification work. In [27], we compared HRRR version 3 forecasts to observations from radiosondes and ASOS (Automated Surface Observing System) stations across the conterminous United States (CONUS), focusing on a spring month (April 2019). We demonstrated that sustained wind forecasts were biased in a manner that led to overprediction of the wind speed at slower wind stations and underspecification of the wind at windier locations, consistent with our own WRF model experiments conducted for Southern California [15,28]. The study of [29] examined ASOS wind forecasts for two April months (2019 and 2021), the latter predictions coming from HRRRV4. That paper determined that the wind forecasts were quite skillful overall but had some deficiencies "in both temporal and spatial variability, with significant errors occurring during local nighttime hours in all regions and in forested environments for all hours of the day".

In this work, we expand our focus to include sustained wind observations from non-ASOS networks, such as RAWS (Remote Automated Weather Stations) and the western U.S. networks maintained by PGE and SCE, and again examine New York State Mesonet (NYSM; [30]) observations. The month of October 2021 is highlighted as it included challenging PSPS-related weather conditions in PGE's service territory. However, there are some difficulties involved with making forecast/observation comparisons that are not routinely appreciated. Models like the HRRR produce sustained wind forecasts valid at 10 m above ground level (AGL), involving the stability-adjusted log wind profile and the wind speed at the lowest half-sigma (scalar and horizontal wind) model level. Since the HRRR places that level close to 10 m, this adjustment is small. However, the substantial majority of near-surface observations routinely reported to MADIS come from anemometers mounted even closer to the surface. It is well-understood that wind speeds are typically slower closer to the ground owing to surface drag, so it can be assumed that alterations to the raw (unmodified) HRRR forecasts are necessary. We will show a straightforward adjustment for anemometer mounting height does not suffice, and does not fully explain the differences that emerge when forecasting for stations in various networks, and thus a different strategy is needed.

The structure of this paper is as follows. Section 2 discusses our data sources and processing. Section 3 presents forecast/observation comparisons for the unmodified HRRR forecasts and considers some potential adjustments to enhance predictive skill. Some important influences on observed wind measurements that vary among networks are considered in Section 4. Finally, Section 5 discusses our proposed adjustment strategy and a summary is provided in Section 6.

2. Data and Methodology

2.1. Observation Data and Processing

Information regarding the networks employed in this study is presented in Table 1. ASOS observations were obtained from the National Centers for Environmental Information (NCEI). Each 1-min record consisted of the station's 2-min average sustained wind and the fastest 3-s gust sample since the last report. 788 stations across the CONUS were available for October 2021 (Figure 1a), unsurprisingly oversampling the urban areas of the CONUS [cf. [29]]. A small subset of ASOS stations are not located at airports and can be very obstructed (e.g., [29]). As in our prior work, these stations were retained in our analysis. ASOS installations presently employ sonic anemometers but not all instruments are mounted at 10 m above ground level (AGL). For comparison with the forecasts, observations closest to the top of each hour were identified and used. Given the 1-min observation interval, selected observations were almost always on the hour.

RAWS and public utility observations were downloaded from MADIS (Meteorological Assimilation Data Ingest System). There were 2153 RAWS stations available for the target month, many located in forested areas (Figure 1b). The RAWS reports were hourly, consisting of the hour's fastest gust and the sustained wind averaged over the prior 10 min. Reporting times varied among sites with few coinciding with the top of the hour, so each observation was assigned to the closest hour. Nominal anemometer mounting height was 20 ft (6.1 m). Anemometer hardware varied among installations and included cup, propeller, and sonic instruments.

Table 1. Summary of observation network information.

Network [Source]	ASOS [NCEI]	NYSM [NYSM]	RAWS [MADIS]	Utility (PGE and SCE) [MADIS]	Affects Mean?	Affects Variance?
Reporting interval (min)	1	5 [3-s raw data]	60	10	no	no
Sustained wind averaging interval (min)	2	5	10	10	no	yes
Nominal instrument(s)	Sonic	Sonic and propeller pairs	Cup, propeller, sonic	Propeller	yes	yes
Nominal mounting height (m)	10	10	6.1 (20 ft.)	Varies (6.1 m target)	yes	yes
Nominal siting and situation	Airports, usually flat and unobstructed	Varies	Sloped terrain, often obstructed	Varies, often obstructed	yes	yes
Starting threshold in source (m s^{-1})	0.51	0.1 [5-min data] 1.0 (propeller) or 0.3 (sonic) [raw data]	0.45	0.0005	yes	yes
Percent calm observations (%)	0.9	1.8 (propeller) 0.1 (sonic)	12.2	0.1	–	–
Number of stations	788	121	2138	1278 PGE 1459 SCE	– –	– –
Station mean (median) HRRR surface roughness (m)	0.29 (0.28)	0.48 (0.45)	0.39 (0.31)	0.34 (0.25) PGE 0.25 (0.21) SCE	–	–
HRRR network-average slope (m s^{-1})	1.00	0.64 (sonic) 0.60 (propeller)	0.62	0.57 PGE 0.52 SCE	–	–

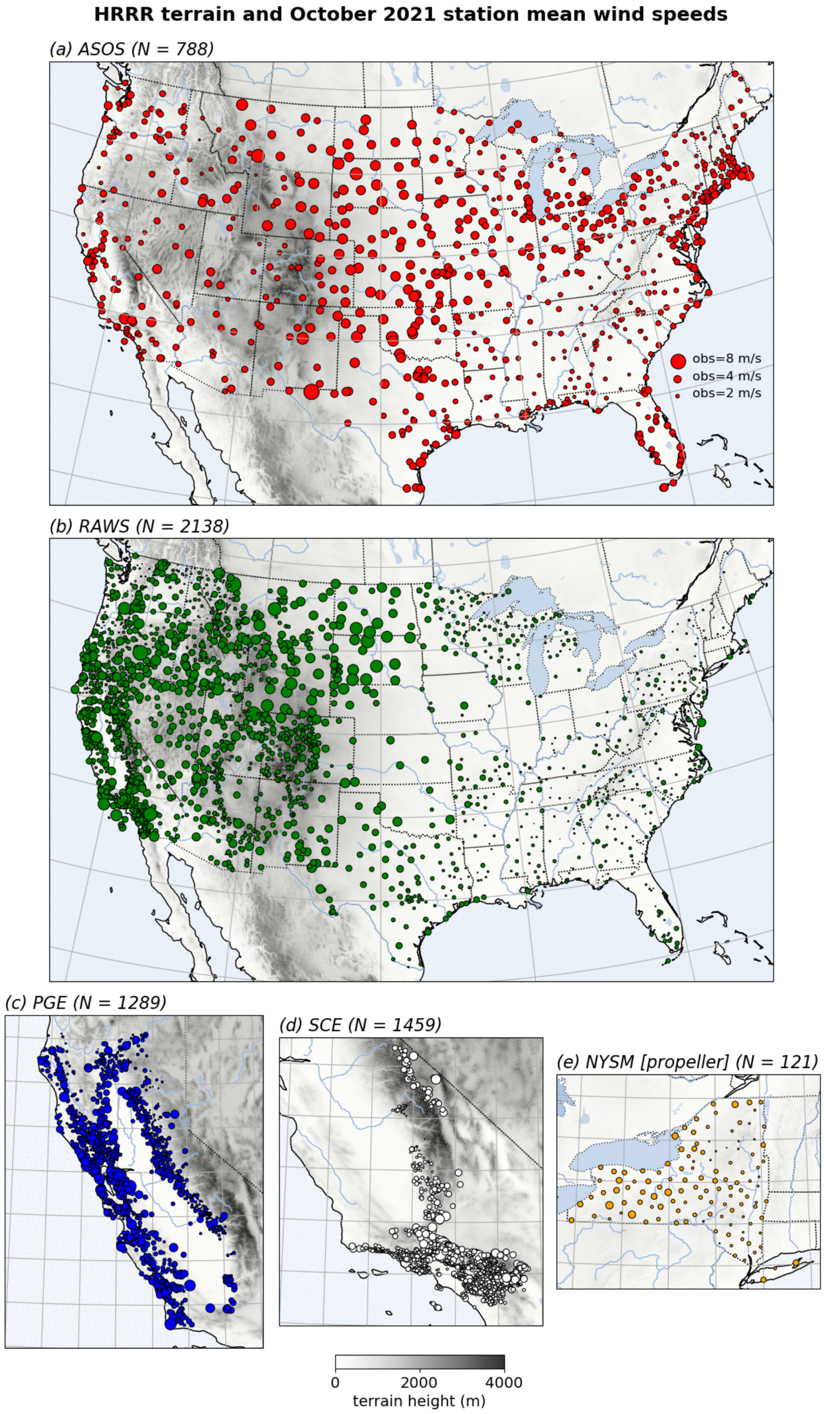

Figure 1. HRRRV4 model terrain (shaded) superimposed with (**a**) ASOS, (**b**) RAWS, (**c**) PGE, (**d**) SCE, and (**e**) NYSM stations, with marker sizes representing average sustained wind speed for October 2021. For (**e**) the values for the propeller instrument are shown. Same marker sizing used in all panels.

The utility observations analyzed herein came from the PGE and SCE mesonets. These two propeller-based networks are very dense, consisting at the time of 1288 and 1459 sites, respectively, concentrating stations on the mountain slopes of California (Figure 1c,d). These stations reported 10-min average sustained winds every 10 min. As with ASOS, observations closest to the top of the hour were used, nearly always representing on the hour. In these two networks, the anemometers were usually installed on power poles and mounting height varied among sites. Analyses performed on the SDGE mesonet are not discussed as the results were not materially distinct.

The NYSM [30] samples the landscapes of the state (Figure 1e). Observations were obtained directly from the mesonet, with permission. NYSM stations possess sonic and propeller anemometer pairs, mounted at 10 m above the surface. These datasets will be referenced as NYSMS and NYSMP, respectively. For five of the sites, that surface is a rooftop; these New York City stations were excluded, leaving 121 locations. We obtained both 5-min and 3-s observations for our target month; unless otherwise indicated, observations closest to the top of the hour from the former were used. For both datasets, the propeller and sonic anemometer observations were merged and records missing reports from either instrument were discarded.

Although we only discuss sustained winds herein, we also gathered observations for other standard meteorological variables (temperature, relative humidity or dewpoint, and gust) for each network. Records with missing or invalid data in at least one variable were excluded. In our most recent relevant study (e.g., [29]), we used the ASOS 1-min data to compute hourly mean winds centered on the top of the hour. However, at least with respect to station mean sustained winds, we also reported there was no sensitivity to the averaging interval adopted. As there is insufficient information among the RAWS observations to compute hourly means, we reverted to using closest-to-hour observations in all analyses reported herein, as in [27]. In our other work (e.g., [10,23]) we also considered Citizen Weather Observer Program (CWOP) observations. Unfortunately, those data were incomplete in the MADIS database for the target month.

2.2. HRRR Model Outputs and Forecast-Observation Comparisons

HRRRV4 model outputs for the month of October 2021 were obtained from Amazon Web Services. Analyses presented herein represented the 49 forecast hours (including the 0-h analysis) from the 00 and 12 UTC cycles of the HRRR for days between 1 and 29 October, inclusive, a total of 2842 maximum possible forecast hours. Observations and forecast comparisons were performed using the Model Evaluation Tools (MET; [31]). MET provides several ways of spatially interpolating forecasts to station locations, including the nearest neighbor and least squares fit strategies. We employed both for each network considered, finding no important differences. Note that because forecast periods overlap, most observations were used multiple times in the comparisons. Stations with fewer than 100 forecast-observation pairs for the month have been removed, reducing (in particular) the number of RAWS sites to 2138.

3. Results

3.1. Network Analysis

Figure 2 presents network-averaged results for the two CONUS-spanning networks. These are “unmodified” forecasts of the 10 m AGL wind speed compared to the closest-to-hour sustained wind observations. Each marker is a forecast hour, representing an average of all observations and forecasts across the network valid for that time.

For ASOS (top row), we see a nearly perfect relationship between network-mean forecast and observed wind speeds among all hours, thereby clearly independent of forecast hour or time of day (Figure 2a). The coefficient of determination (R^2) is 0.96 among the T = 2837 available forecast times. The remaining panels of the top row (Figure 2b,c) present mean forecast and observed wind speed vs. forecast bias, defined here as forecast minus observation. There is no correlation seen in either panel. The model was equally skillful whether the CONUS-mean wind was predicted or observed to be relatively faster or slower. These results are consistent with the findings of [29], who showed that, in the mean, ASOS network-averaged sustained wind forecasts had negligible bias at all forecast hours and local times of day.

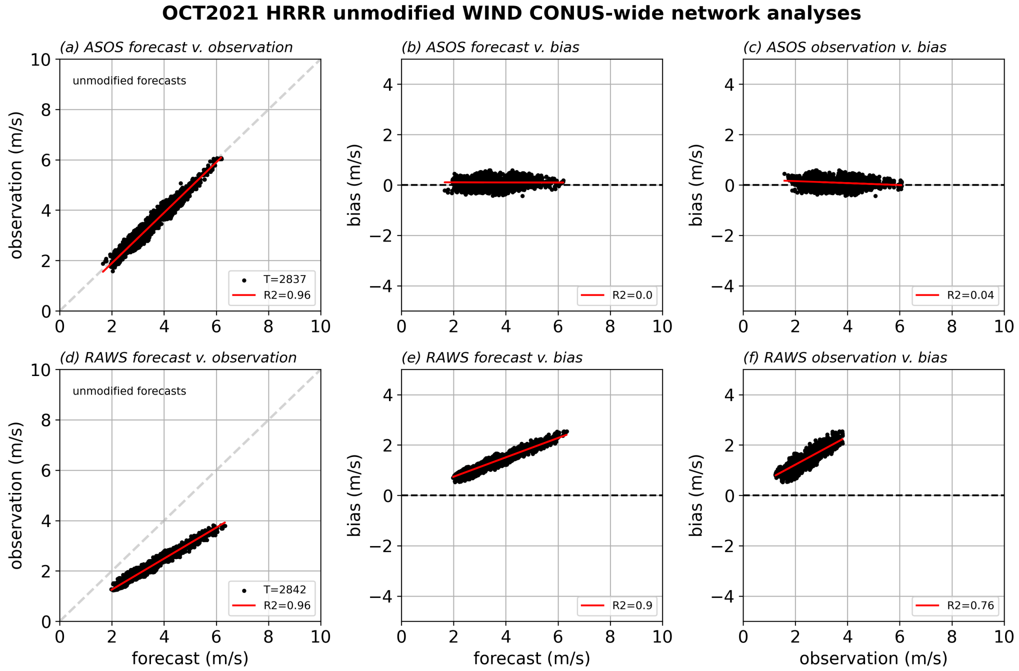

Figure 2. Network analyses for ASOS (**top row**) and RAWS (**bottom row**). Each marker is a forecast hour, totaling T = 2837 and 2842 for ASOS and RAWS, respectively, representing an average of all values available for that hour. Red lines show least squares fits. The forecasts were extracted from HRRR outputs and not subsequently modified.

The situation with the RAWS network (Figure 2, bottom row) is markedly different. The relationship between network-averaged forecast and observed winds is still very high (R^2 again 0.96) but now there is a pronounced overprediction (Figure 2d), with the slope of the least squares fit being 0.62 m s^{-1}. As the intercept of the least squares fit is nearly zero (0.03 m s^{-1}), the HRRR was overpredicting CONUS-averaged RAWS sustained wind speeds by almost 40%, independent of time of day. As a consequence of this, both network-averaged forecasts and observations were significantly and positively correlated with bias (Figure 2e,f).

As summarized in Table 1, there are important differences between the ASOS and RAWS networks. The HRRR is providing forecasts for 10 m AGL but RAWS measurements are made closer to the surface (at 6.1 m). In neutral conditions, this can be expected to result in a 10–20% reduction in wind speed, depending on the roughness of the surface, according to the logarithmic (log) wind profile (e.g., [15,32]). In the HRRR for October 2021, mean and median roughness lengths (z_0) assigned to grid cells containing RAWS stations (0.39 and 0.31 m) were notably longer than those assigned to ASOS (0.29 and 0.28 m), translating to a roughly 15% reduction between 10 and 6.1 m AGL.

However, the discrepancies seen here are much larger than can be explained by mounting height alone. Certainly, the differences with respect to situation and exposure are important; unlike the generally flat and well-exposed environments characteristic of most airport ASOS stations, RAWS installations are often placed in complex terrain and in the vicinity of obstructions, including trees and buildings. This is one of the factors influencing roughness length. In their study of Hurricane Andrew's landfall, [32] emphasized the importance of converting wind observations taken in areas with different roughnesses to a reference (open) exposure, demonstrating that a z_0 of 0.5 m could slow the wind by about 40% compared to open exposure (defined as z_0 = 0.03 m). Additionally, significant obstacles might be treated via a zero-plane displacement [33], something that is not done in the HRRR. The networks also use different reporting and averaging intervals, but these are not found to affect the mean sustained wind [cf. [10,29,34]].

Before exploring these important differences further, we examine the three regional networks included in this study (Figure 3). All three also present significant overpredictions, similar to or exceeding that seen in the RAWS results. R^2 values are a little lower, perhaps a consequence of these networks' smaller spatial extents, but still very high. Although mounting heights vary among stations in the PGE and SCE mesonets, most instruments are closer to the ground than 10 m AGL. That being said, the NYSM stations included in this study all represent 10 m installations and their overpredictions (Figure 3, bottom row) are comparable to those seen with RAWS. Furthermore, note that despite time-matched records, the propeller and sonic instruments in the NYSM yielded different results, with the former (black markers in Figure 3g–i) being about 5% slower than the sonic (grey markers), thereby resulting in even larger positive wind speed biases. This is the first indication in this analysis that instrument type, which varies among the networks examined (Table 1), is also a relevant factor.

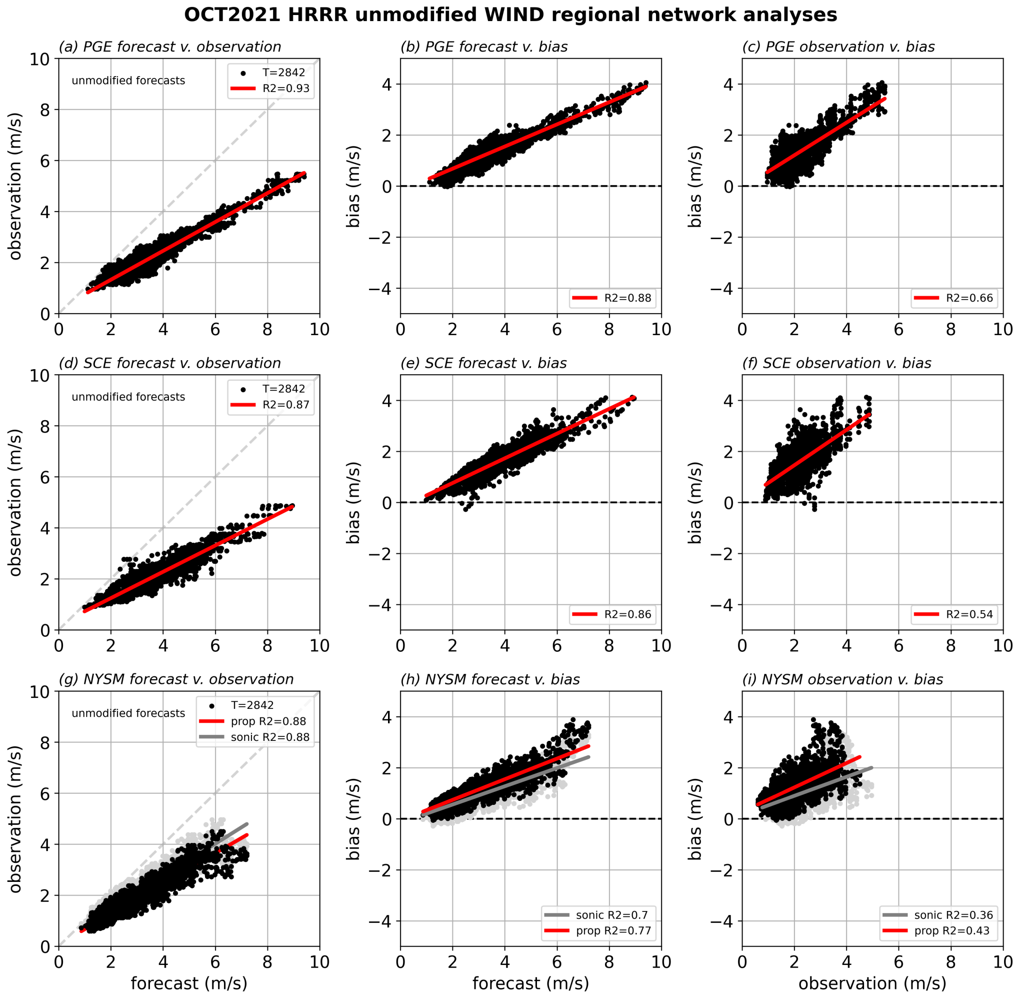

Figure 3. Similar to Figure 2 but for the PGE (**top row**), SCE (**middle row**), and NYSM (**bottom row**) network analyses. For NYSM, results using propeller (black markers, red regression line) and sonic (grey markers and regression line) observations are shown.

3.2. Station Analysis

The station analyses for ASOS and RAWS are presented in Figure 4. Each marker now represents station mean values, averaged over all available (up to 2842) forecasts. For ASOS (top row), the relationship between mean forecast and observed wind speeds for the $N = 788$ stations (Figure 4a) is very good, albeit with more scatter around the 1:1 correspondence (grey dashed) line than in the network analysis (Figure 2a), and not unlike the results found for different months and model versions by [29]. There are clearly some outliers, and they are the same stations flagged by [29]: KGDP (Guadalupe Pass, Texas, being the windiest site and possessing the largest negative bias), KMEH (Meacham, Oregon, the least windy location), and KMHS (Mt. Shasta, California, having the largest positive bias). These installations are not at airports and satellite imagery suggests the latter two are very obstructed.

As in our prior study, there was no association found between mean forecast wind speed and bias among stations (Figure 4b) but a sizable negative one ($R^2 = 0.31$, correlation $r = -0.56$) between mean observed speed and bias (Figure 4c), indicating that slower locales were being systemically overpredicted while windier sites were underforecasted. As described in [15], "the bias was itself biased" (e.g., a "bias-bias"). We note that the negative correlation was actually smaller in magnitude in this analysis than in the one presented in [29] ($R^2 = 0.51$; see their Figure 8b) but is still a matter of concern. Furthermore, although the mean station bias was close to zero (+0.11 m s^{-1}), 57% of the stations had positive biases.

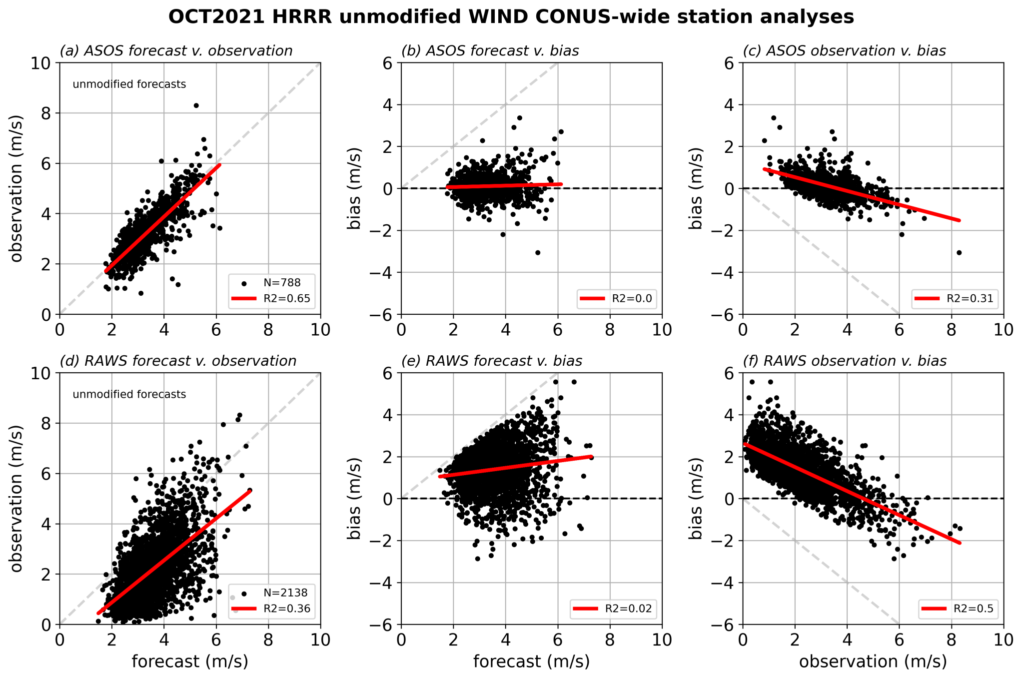

Figure 4. Station analyses for ASOS (**top row**) and RAWS (**bottom row**). Each marker is a station, totaling $N = 788$ and 2138 for ASOS and RAWS, respectively, representing an average of all available values for that site. Red lines show least squares fits. Grey dashed diagonals in middle and right columns represent forecast bias limits imposed by the non-negativity constraint of wind speed. The forecasts were extracted from HRRR outputs and not subsequently modified.

Regarding RAWS (Figure 4, bottom row), the association between station mean forecast and observed wind was significantly less skillful [panel (a)] than for ASOS. The variance of the forecasts was also noticeably narrower (by about 40%) than that of the observations. Despite that, nearly all (90% of $N = 2138$ stations) of the markers fell below the 1:1 line, indicating overprediction. The grey diagonal lines in panels (e) and (f) indicate the limits to forecast bias imposed by the constraint that wind speeds are non-negative. Forecast bias reaches the upper limit when the observed wind approaches zero, and the lower bound is attained when the mean forecast wind is calm. There was a sizable number of sites with mean wind speeds < 2 m s^{-1} for the target month (12% of all RAWS sustained winds during this month were exactly zero; cf. Table 1) but very few forecasts were that slow. Still, for the subset of windier sites, the mean forecast wind speeds were too low. As a consequence, there was again a strong, negative relationship (R^2 = 0.5, r = −0.71) between bias and observed wind speed (Figure 4c).

The situation is similar for the regional networks (Figure 5). Despite being spatially concentrated, the PGE (top row) results resemble RAWS in that this network had a narrower range of forecast than observed mean wind speed [panel (a)], and there was no correlation between forecast and bias [panel (b)] but a substantial negative one between observed wind and bias. The mean station bias was +1.26 m s^{-1}, with 91% of sites being overpredicted in the mean. The mean bias was even larger for SCE (+1.42 m s^{-1}), with 97% biased positively. For this network, the correlations between bias and forecast and observed winds were somewhat different, but again calmer stations were more likely to be overpredicted. For clarity, only the propeller data for NYSM is being shown (bottom row). The mean bias was +1.22 m s^{-1} and all but 4 sites were overforecasted.

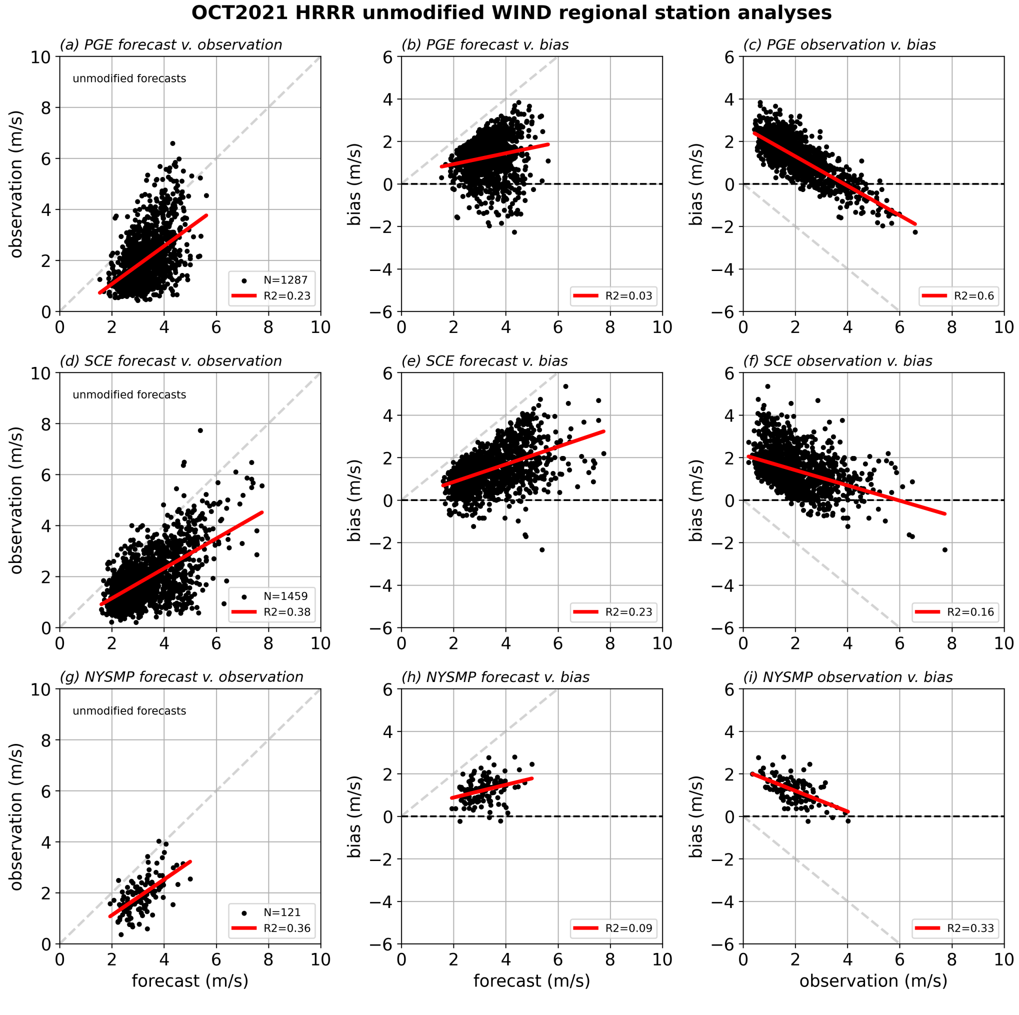

Figure 5. Similar to Figure 4 but for the PGE (**top row**), SCE (**middle row**), and NYSM (**bottom row**) station analyses. For NYSM, only the propeller results (NYSMP) are shown for clarity.

Histograms of all forecasts and observations used in the comparisons for the five networks are shown in Figure 6. The coarser precision and higher starting thresholds of ASOS and RAWS observations are reflected in these histograms, as is the pronounced tendency for RAWS sites to report calm winds (Table 1). At first glance, the distributions of forecasts (blue) and observations (red) appear very similar for ASOS (Figure 6a), underscoring the model's skillfulness. Closer inspection reveals the standard deviation of the forecasts was about 10% smaller than that for the observations. As the means were essentially identical, this indicates both smaller and larger values appeared slightly more frequently in the observations than among the forecasts.

For the other networks, the forecast and observation distributions were very different. It is notable that these non-ASOS observations shared fairly similar shapes, while their forecasts resembled those produced for ASOS stations. What is different about the non-ASOS networks is how their observations were distributed. To a very large degree, the HRRR was providing ASOS-like forecasts for every location, for every network. This is clearly inadequate.

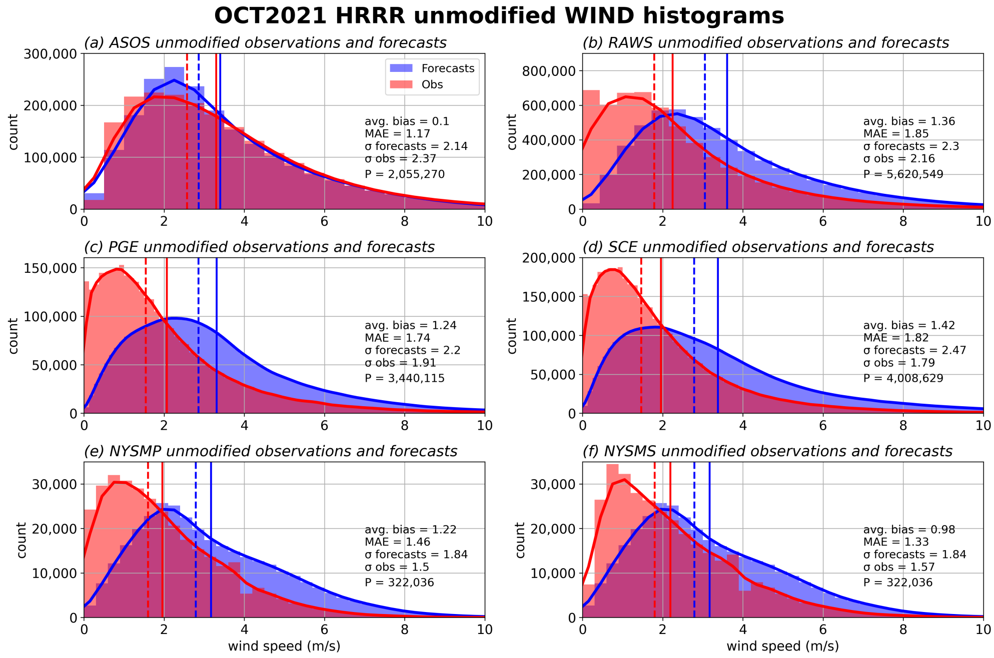

Figure 6. Histograms of all observations (red) and unmodified HRRR forecasts (blue) employed in the network and station analyses, representing the (**a**) ASOS, (**b**) RAWS, (**c**) PGE, (**d**) SCE, (**e**) NYSMP (propeller), and (**f**) NYSMS (sonic) networks. The vertical solid and dashed lines indicate mean and median values, respectively. Bin interval determined by observation precision. Each panel provides information on the number of forecast/observation pairs (P), bias, mean absolute error (MAE), and standard deviation (σ) of the forecasts and observations.

3.3. Initial Adjustment Attempts

We have seen that the *same model that is producing very skillful network-averaged wind speed predictions for ASOS stations is overspecifying the wind speed in other networks by sizable amounts* (Figures 2 and 3). But, perhaps the forecasts just need a simple adjustment. Given the very high R^2 evidenced in the scatterplots of network-averaged observations vs. forecasts (Figures 2 and 3), one might be tempted to fashion revised HRRR forecasts based on the least squares slopes between the forecasts and observations. Multiplying the unmodified HRRR forecasts by network adjustment factors suggested by those slopes (0.62 for RAWS and 0.57 for PGE, as reported in Table 1) produced the "naive adjustment" network and station analyses presented in Figure 7.

For the network-averaged analysis (top two rows), multiplying the forecasts by the slopes unsurprisingly brought the adjusted values and observations in line, removing the previously-existing correlations with respect to bias. However, the defects previously noted in the station analysis have been made much worse (Figure 7, bottom two rows). The naive network adjustment has further compressed the already too-narrow range of the forecasts, actually sharpening the negative correlation between observed wind and adjusted bias (Figure 7i,l). The compression can also be seen in the histograms (Figure 8, top row). The forecast distributions look better now relative to the observations, but are still markedly different.

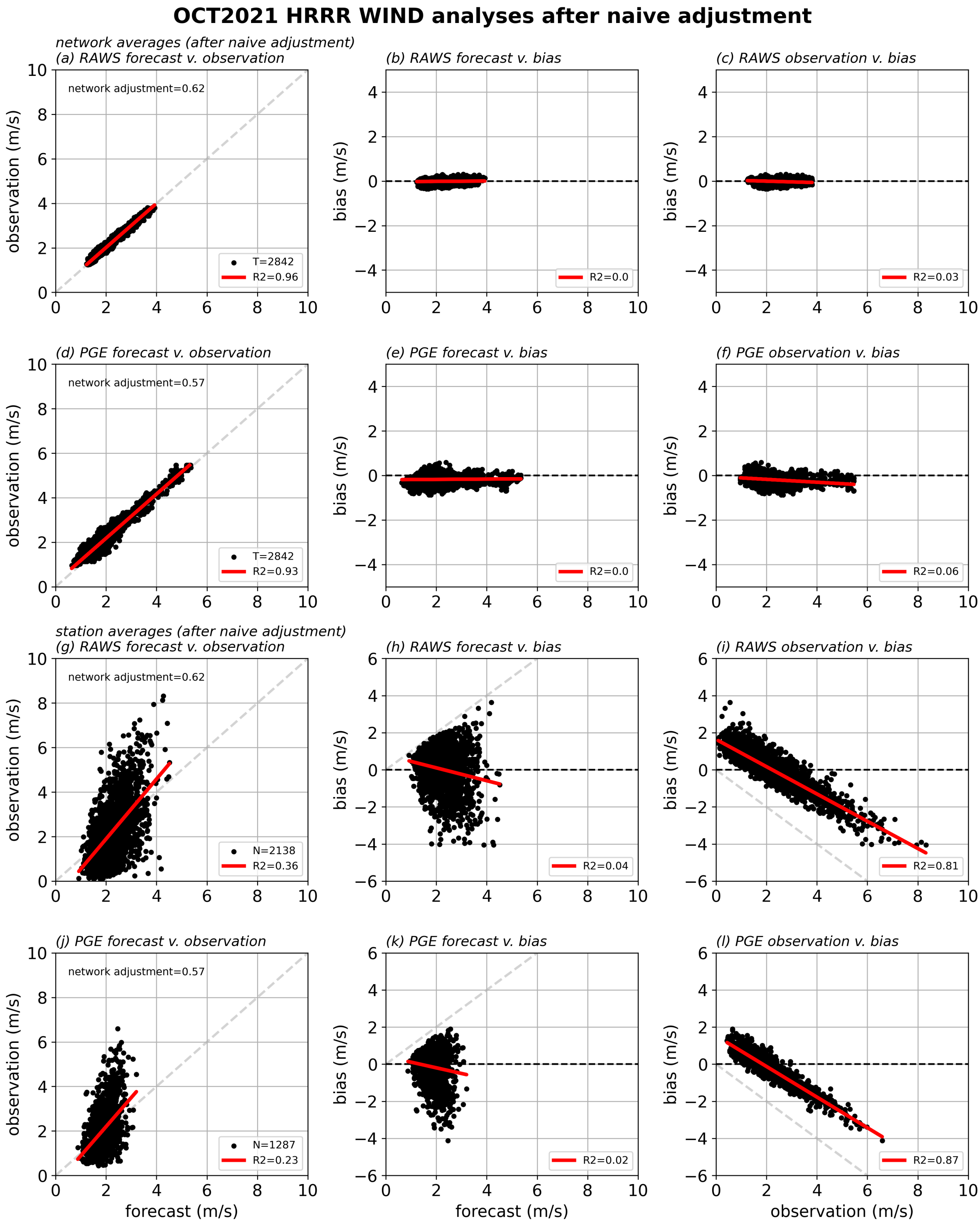

Figure 7. Similar to Figures 2 and 4 but showing the network and station analyses following naive forecast adjustment.

Perhaps adjustments should be made to the observations instead. This strategy, inspired by [32], transformed observations into "exposure-adjusted" equivalents and consisted of several steps. First, for RAWS and utility stations, the observed wind *U* at anemometer height Z was transformed into a 10 m wind U_{10} using the neutral log wind profile

$$U_{10} = U\left[\frac{\ln(10/z_0)}{\ln(Z/z_0)}\right], \tag{1}$$

where Z is 6.1 m for RAWS and varies (but is generally below 10 m) for PGE and z_0 was provided by the HRRR grid cell in which the station resides. No displacement height was used. The friction velocity u_* representing this 10 m height-adjusted wind was then computed, using

$$u_* = \frac{\kappa U_{10}}{\ln(10/z_0)}, \tag{2}$$

where κ is von Karman's constant. The next step transformed this u_* to a value more appropriate for open terrain (u_{*S}) via the expression

$$u_{*S} = u_*\left[\frac{z_{0S}}{z_0}\right]^{0.0706}, \tag{3}$$

in which [following [32]] $z_{0S} \approx 0.03$ m was adopted as the open terrain roughness. Finally, the adjusted 10 m wind observation was obtained from u_{*S} via the neutral log wind profile. As discussed in [32], this can produce an increase in the wind speed over that observed in rough terrain by about 40%, approximately the degree of overprediction seen in our non-ASOS analyses.

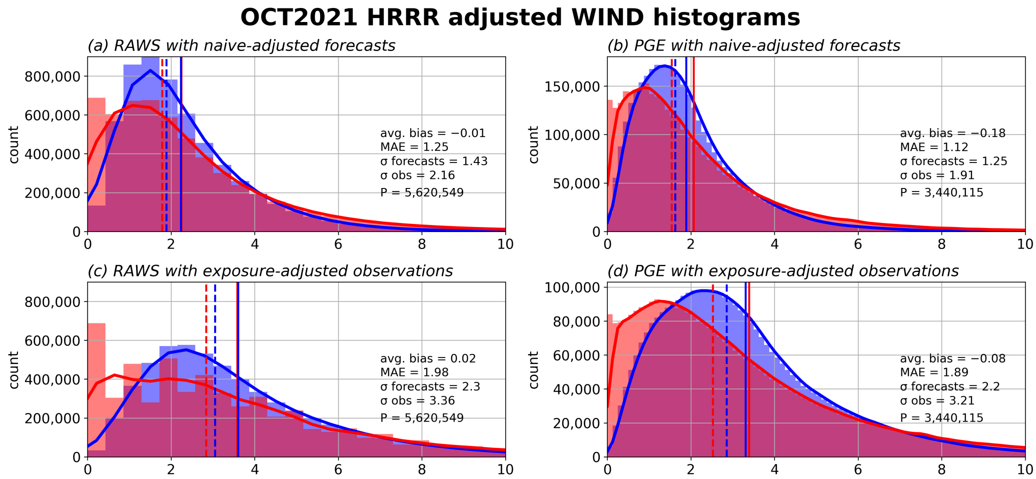

Figure 8. Similar to Figure 6 but after initial adjustment attempts for RAWS (**left**) and PGE (**right**) forecasts and observations. **Top row**: naive adjustment of forecasts; **bottom row**: exposure-modified observations.

Figure 9 presents the comparison of exposure-adjusted, 10 m versions of the RAWS and PGE observations with their corresponding HRRR forecasts. The top two rows demonstrate this technique can effectively remove the substantial positive bias with respect to the original readings. The mean bias is now zero in both networks. Yet the issues with the station analyses (bottom two rows) remain. There is not much skill in the association between station mean forecast and adjusted observation [panels (g,j)], particularly when compared to the ASOS results. The ranges of the observed and predicted winds remain clearly different, this time because the former was expanded by this adjustment. The exposure correction cannot compensate for the high frequency of calm observations in the RAWS network, which is partially a consequence of restricted exposure. Furthermore, the negative correlation between adjusted bias and adjusted observation [panels (i,l)] remains quite substantial and the histograms (Figure 8, bottom row) are still distinctly different. Adopting alternative z_0 values for open terrain, including the average value used by the model for ASOS stations (0.29 m; see Table 1), was not found to resolve these issues.

Clearly, a different approach is needed. In prior work [28,29], we made use of the gust factor (the gust to sustained wind ratio) as a means of mitigating bias in the sustained wind forecasts for the purposes of creating more skillful gust predictions. In this study, we attempt to isolate the root cause of the problem, which turns out to lie in the sustained wind observations themselves.

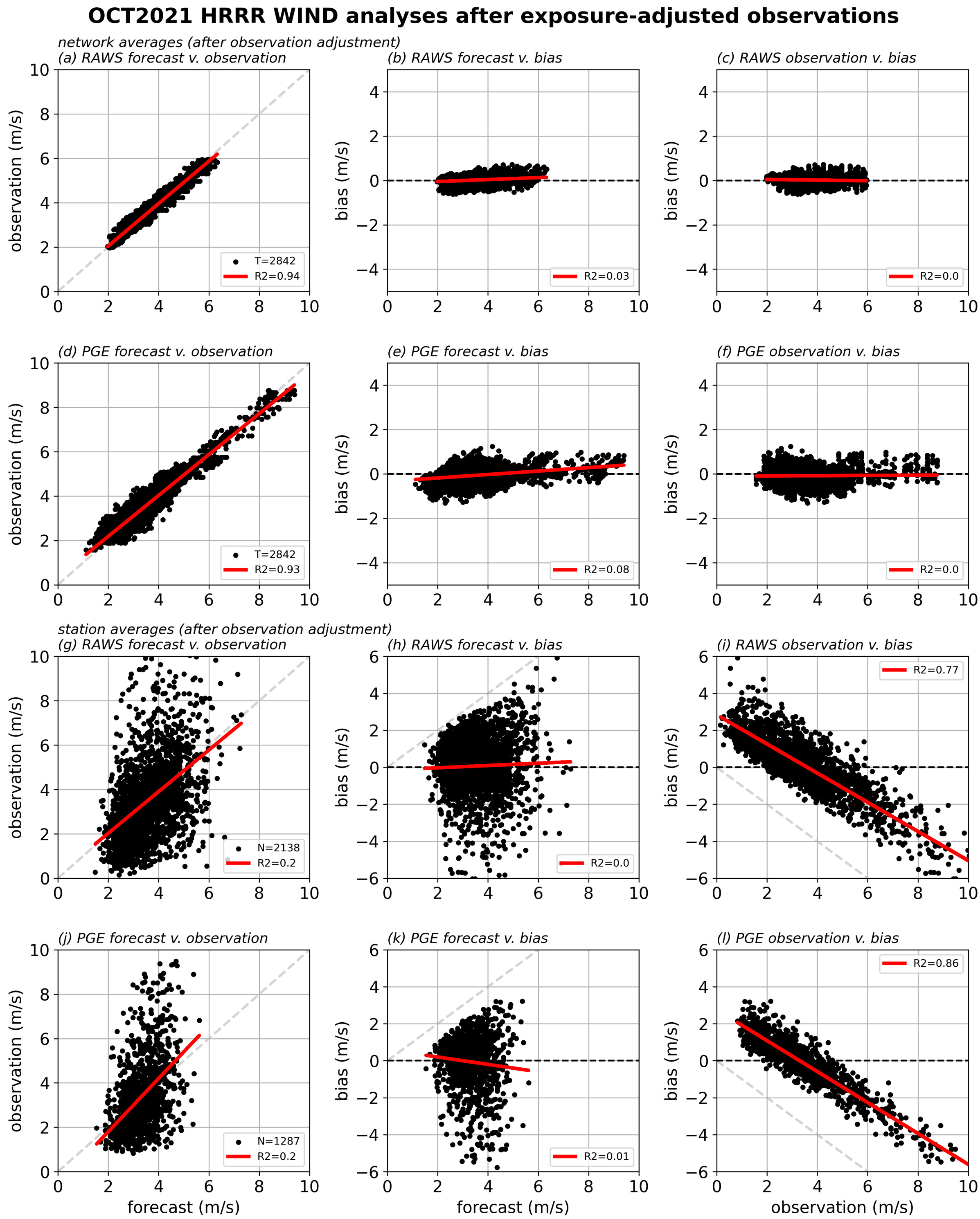

Figure 9. Similar to Figure 7 but after exposure adjustment to the observations. In bottom two rows, this adjustment has caused some station values to fall outside the ranges shown.

4. Influences on Anemometer Readings

It is well known that there are numerous ways in which sustained wind observations can vary among observing networks. Among them are averaging interval, anemometer hardware, mounting height, reporting interval, and situation and exposure. We will consider the first four in turn, emphasizing their influence not only on the mean wind but also on its temporal variation, so we can appreciate the influence of situation and exposure isolated from these other factors.

4.1. Averaging Interval

The averaging interval is the time period over which the sustained wind is computed. ASOS uses a 2-min interval while NYSM adopted a 5-min period and RAWS and our utility networks employed 10-min intervals (Table 1). Ref. [29] transformed their ASOS data into 60-min means. The averaging interval length is known to not to influence the true mean wind but does affect the temporal variance [34]. This is explored herein using the high-frequency (3-s), quality-controlled propeller observations from the NYSM. As explained in [29], these observations form a gapless dataset, consisting of 3-s average winds available every 3 seconds.

From these raw, 3-s samples, averaging intervals of various lengths were constructed, as illustrated in Figure 10. For each interval considered, the observational record was divided into overlapping segments of length n-min. Any segment with missing data was discarded. Each of those segments ended with a 3-s sample, $\bar{V}$(3-s), which was compared to that segment's n-min mean, $\bar{V}$(n-min). For each segment, the difference D_n = $\bar{V}$(n-min)− $\bar{V}$(3-s) was computed, and compared to both sustained wind estimates. (It is important to note that a 3-s sample drawn from an n-min segment is *not* a gust. The gust is the fastest 3-s sample in that segment).

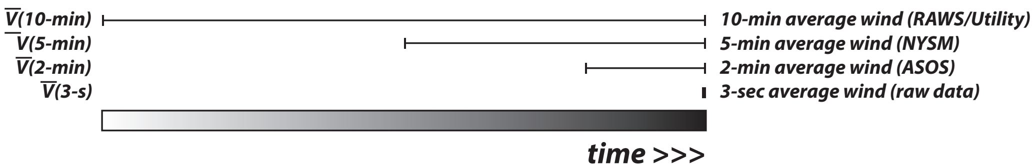

Figure 10. Strategy employed in the averaging interval experiment.

Table 2 and Figure 11 present an example assessment conducted for the station in Penn Yan (PENN), one of the NYSM's windiest sites during October 2021. The table reveals that, as expected, the mean was unaffected by the length of the averaging interval. However, the standard deviation (σ) decreased with interval length. We form the standard deviation ratio (SDR) by dividing the interval's σ with the 3-s data's value. By the 60-min averaging period, the SDR for this station had decreased to 0.90, representing a variance reduction of about 19%. The maximum wind speed remaining in the dataset also decreased, substantially, as the interval lengthened, from about 15 m s^{-1} in the 3-s dataset to 10 m s^{-1} in the 60-min version.

Thus the right, or high wind speed, tail of the distributions was being progressively lost as the averaging interval increased. That is very obvious in the plots of $\bar{V}$(3-s) vs. $\bar{V}$(n-min) displayed in the left column of Figure 11. The 12 m s^{-1} value is highlighted on the ordinates of those scatterplots to emphasize the loss of information occurring at higher wind speeds. Contraction of the high tail is also very obvious when distributions are viewed in histogram format (Figure 12a). There are relatively few observations in the high wind tail but the shrinkage is not negligible.

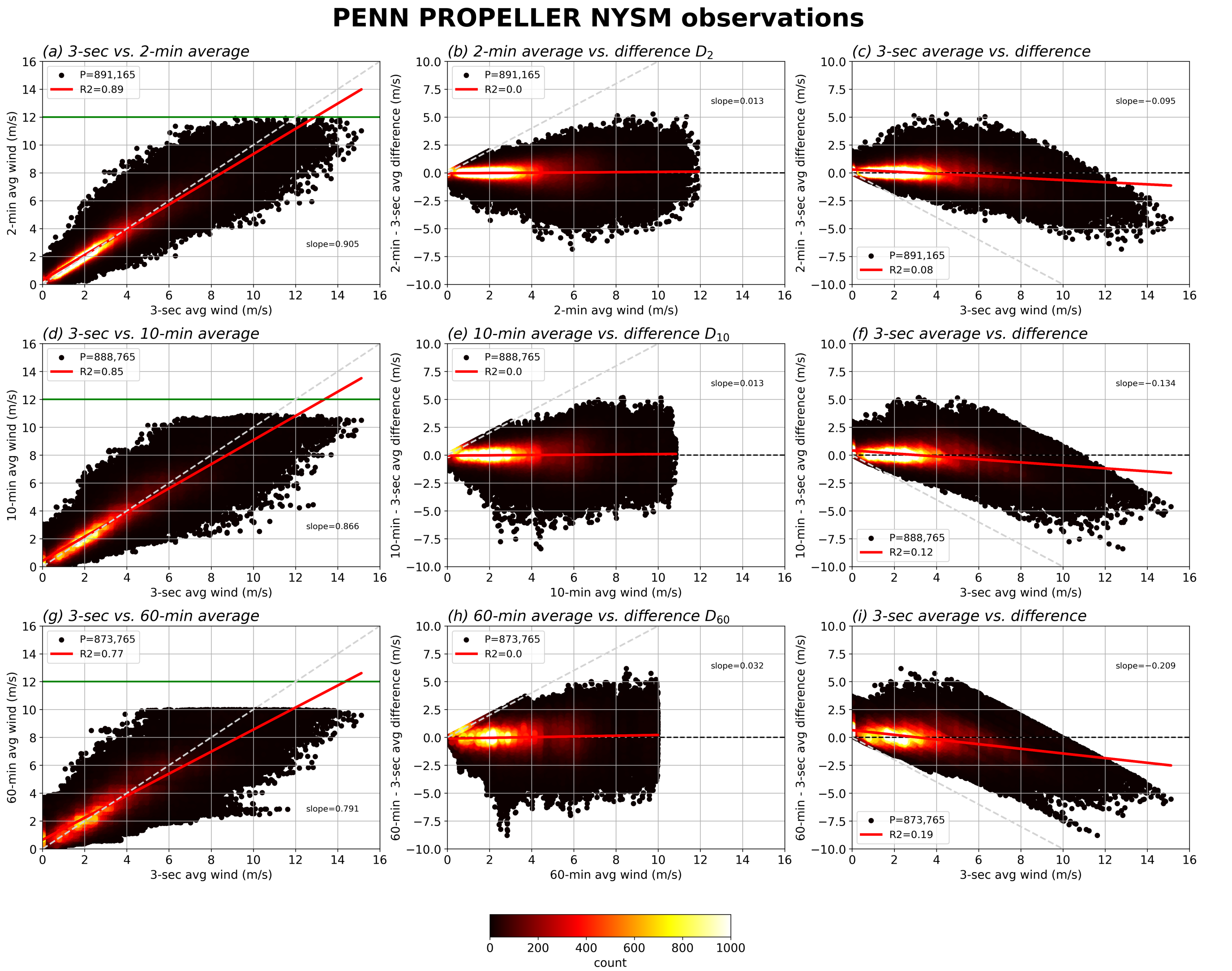

Figure 11. Scatterplots illustrating influence of averaging interval on sustained wind speeds using NYSM Penn Yan (PENN) propeller observations. Left column compares segments representing n = 2-, 10-, and 60-min averaging intervals with the 3-s sustained winds. On the ordinate, the 12 m s^{-1} value is highlighted for reference. Middle column compares the same averaging intervals to D_n, the difference between the n-min and 3-s wind speeds, with the grey dashed diagonals representing maximum possible values. Right column presents these differences against 3-s wind speeds, with the grey dashed diagonals being minimum possible values. Regression lines shown in red.

Although more difficult to discern, even in the histogram, the slow wind (left) tail was also shifting away from zero (calm) as the averaging interval lenghtened, which is important owing to the large number of low wind observations in the dataset. Figure 12b shows the differences between the 60-min and raw distributions as a function of wind speed. Speeds slower than about 2 m s^{-1} were much less common in the $\bar{V}$ (60-min) dataset; the averaging pushed those observations towards the mean (vertical dashed line). There were also fewer instances above 6 m s^{-1}. The consequences of this are clear: longer averaging intervals do not affect the mean, but do decrease the variance, forging narrower frequency distributions.

Because wind speed is non-negative, D_n has an upper bound of $\bar{V}$(n-min) as the 3-s reading approaches zero and a lower bound of $-\bar{V}$(3-s), these bounds being illustrated in Figure 11 with dashed grey lines. As a consequence, the distributions of $\bar{V}$(3-s) and $\bar{V}$(n-min) with respect to D_n (middle and right columns of Figure 11) also differ. These constraints contribute to the tendency for segments with low $\bar{V}$(3-s) values to be associated with faster $\bar{V}$(n-min) winds and those with higher $\bar{V}$(3-s) readings to occur alongside slower $\bar{V}$(n-min) values. Thus, the constraints contribute to forging the negative correlation seen in Figure 11's right column, which strongly resembles the aforementioned "bias-bias". The SDR therefore decreases as the averaging interval lengthens, as documented in Table 2.

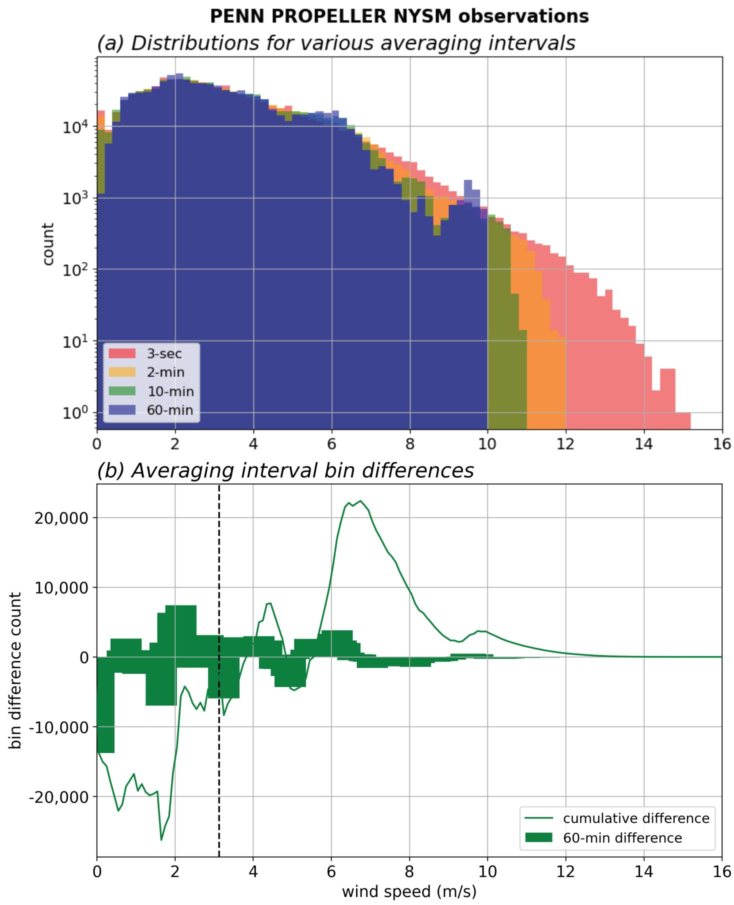

Figure 12. PENN propeller observations: (**a**) histograms of 3-s and 2-, 10-, and 60-min wind speeds; (**b**) differences between the 60-min and 3-s distributions, expressed on a per-bin basis (bars) and as a cumulative total across wind speeds (curve).

As noted above, PENN was one of the windiest stations during October 2021. Regarding SDR, it had one of the network's smallest decreases with lengthening averaging interval (again, 0.90 for an averaging period of 60 min). In contrast, the network's smallest SDR for the hourly averaged wind was about 0.60. That the largest reductions tended to occur at the most obstructed sites is demonstrated in Figure 13a, in which the stations are color-coded by the WMO classification [35] determined by Mesonet staff. The classification system runs from 1 (most exposed, blue) to 5 (most obstructed, red). More obstructed locations tended to be less windy, so SDR was also correlated with station mean wind speed [panel (b)]. Also unsurprisingly, temporal σ was also well-associated with the mean wind; since speeds are non-negative, measurements at less windy locations cannot vary as much. Forecast bias, which was positive for nearly all of the sites, was negatively correlated with station temporal standard deviation [panel (e)] and the association with SDR was even stronger [panel (c)].

This is relevant to the subject of model-observation comparisons if the model forecasts have an inherent temporal variation (which is influenced by spatial and temporal filtering) that is different than that of the corresponding observations (which is impacted by averaging interval) to which they are being compared. The bias can even be zero without the distributions of the forecasts and observations being the same, which can result in the "bias-bias".

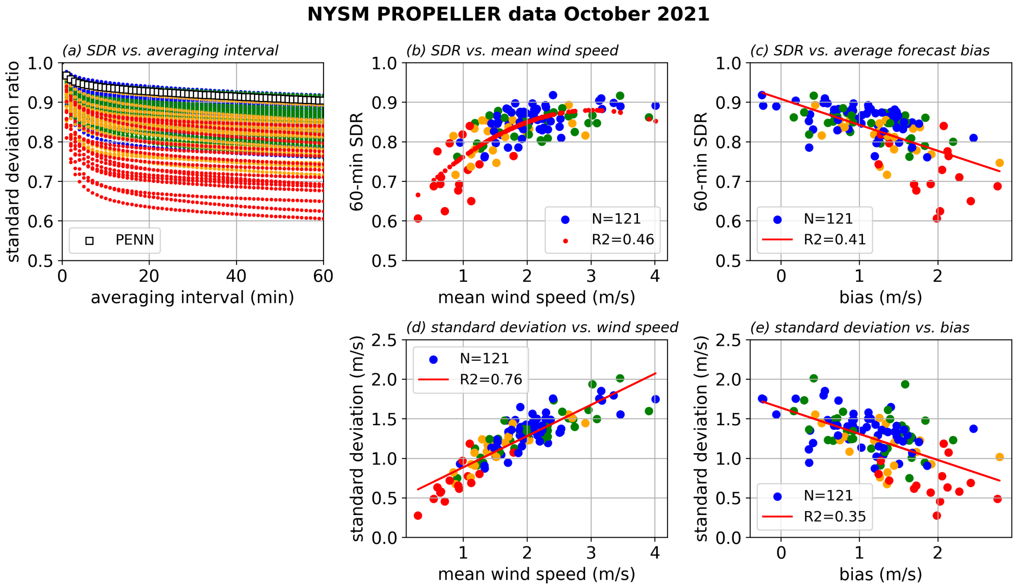

Figure 13. NYSM propeller network averaging interval analysis: (**a**) standard deviation ratio (SDR) vs. averaging interval for each of the N = 121 stations; (**b**) 60-min SDR vs. mean wind speed; (**c**) 60-min SDR vs. bias; (**d**) standard deviation vs. station mean wind speed; and (**e**) standard deviation vs. bias. Markers are color-coded by WMO site classification for wind: 1 = blue, 2 = green, 3 = orange, and 4 and 5 = red. However, in (**a**), PENN is indicated by the white squares. In (**b**–**e**), red lines/curves indicate least squares fits.

Table 2. Summary of averaging interval assessment of NYSM PENN propeller observations.

Averaging Interval	3-s [Raw]	2-min [ASOS-like]	10-min [RAWS-like]	60-min
Number of segments (P)	891,765	891,165	888,765	873,765
Mean (m s^{-1})	3.13	3.13	3.13	3.13
Standard deviation σ (m s^{-1})	1.94	1.86	1.82	1.75
Standard deviation ratio (SDR)	–	0.96	0.94	0.90
Maximum value (m s^{-1})	15.12	11.92	10.85	10.00
First quartile (m s^{-1})	1.73	1.78	1.79	1.84
Third quartile (m s^{-1})	4.18	4.21	4.19	4.15

4.2. Anemometer Hardware

Among our networks, the most common anemometer types are sonic and propeller. These instruments have different characteristics [cf. [36–39]] that can influence the wind speeds being recorded as already hinted by Figure 3g–i. In this subsection, we more closely compare 5-min data from the NYSM's sonic and propeller instruments, which were mounted side by side at 10 m AGL. This analysis involves the N = 121 non-rooftop sites.

Figure 14 presents a comparison of P = 88,772 time-matched, on-the-hour sonic and propeller pairs in the dataset. As expected, there was a very high correlation between the two (R^2 = 0.99 in Figure 14a). Perhaps owing to the absence of moving parts, however, the sonic wind speeds were nearly always faster. The mean difference was about 5%, consistent with that seen in the network-average comparison (Figure 3g) and the findings of [40] (see their Figure 3). Also consistent with [40], the propeller-sonic differences became more clearly negative as the propeller wind speed increased (Figure 14b). This can influence the "bias-bias" via right tail contraction. The standard deviation of the sonic data was also about 5% larger (1.62 vs. 1.54 m s^{-1}).

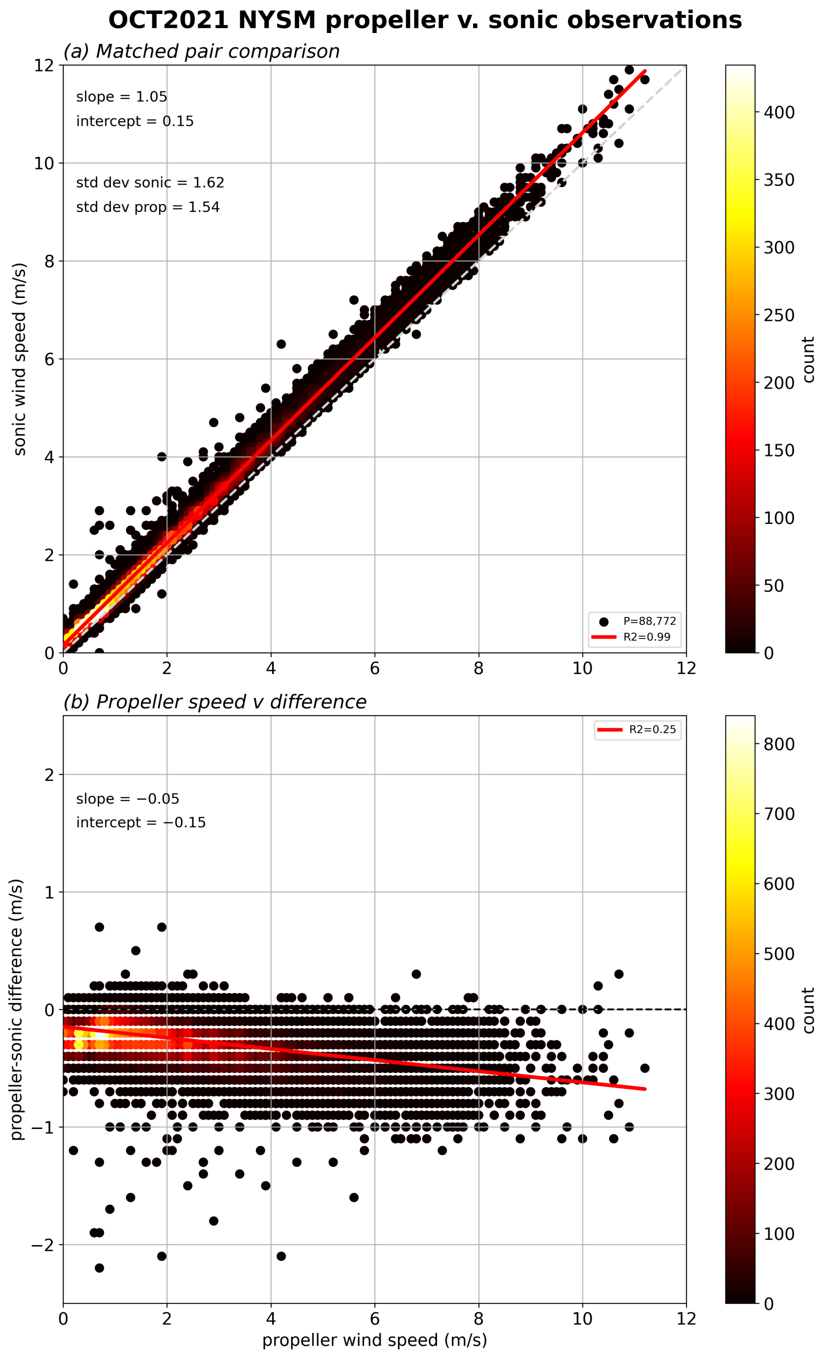

Figure 14. Comparison of 5-min observations from sonic and propeller anemometers representing N = 121 NYSM stations spanning P = 88,772 reports, plotting propeller wind speeds vs. (**a**) sonic wind speeds, and (**b**) the difference between propeller and sonic speeds. Regression lines shown in red.

4.3. Mounting Height and Reporting Interval

It is well appreciated that wind speed decreases towards the surface owing to drag, so anemometers mounted closer to the ground should yield slower readings, other factors being equal. For completeness, we demonstrate the effect of mounting height with fixed tower data. The National Wind Technology Center (NWTC) Flatirons Campus, located south of Boulder, Colorado, operates a tower with anemometers placed at 2, 5, 10, 20, 50, and 80 m AGL. Figure 15a presents October 2021 mean wind speed (black) and temporal standard deviation (red) computed from 1-min average wind observations (S = 44,640). The mean 5 m wind speed was about 9% slower than the 10 m average and the standard deviation was 11% smaller. As anticipated, temporal standard deviation increased with height above the ground [panel (a)]. A least squares fit to the mean wind profile (Figure 15a black dashed curve) suggests a surface roughness of 0.013 m, considerably smaller than the value extracted from HRRR forecasts (0.094 m). Compared to 10 m winds, readings taken closer to the surface resulted in speed distributions with slower means and narrower ranges (and thus thinner high wind tails) as illustrated in Figure 15b.

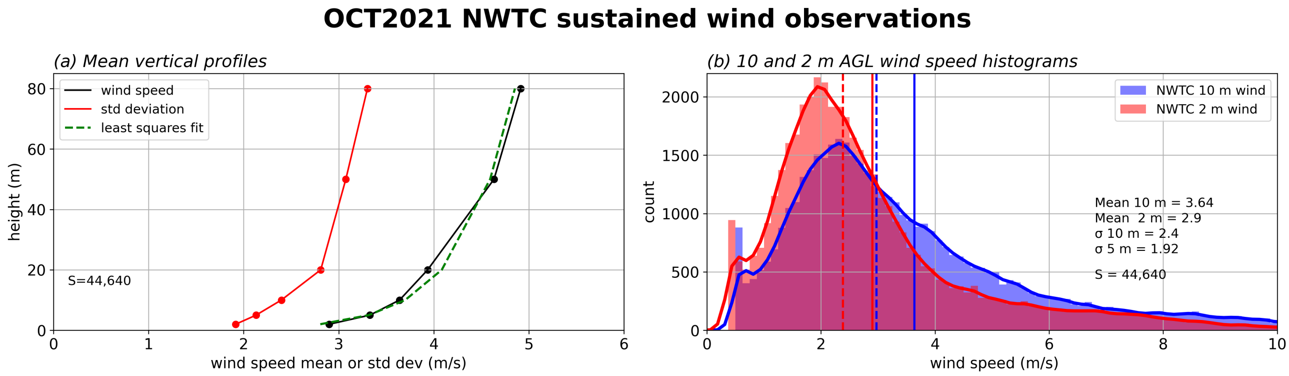

Figure 15. Observations from the National Wind Technology Center (NWTC) Boulder tower for October 2021, showing (**a**) vertical profiles of monthly mean wind (black) and temporal standard deviation (red), and (**b**) histograms of 10 m (blue) and 2 m (red) sustained wind speeds. In (**a**), the black dashed curve represents a logarithmic fit using least squares. In (**b**), the vertical solid and dashed lines indicate mean and median values, respectively.

The reporting interval adopted for a particular network or station does not appear to influence either the mean or the variance of the sustained wind. This was assessed again using the raw, 3-s PENN propeller data. The full record of 891,765 observations was sampled at intervals ranging from 1 to 60 min, creating subset datasets. These revealed no trends in either subset mean or σ as a function of reporting interval.

5. Network-Agnostic Forecast Corrections

We have demonstrated that the HRRR produced forecasts of 10 m AGL sustained wind speed for the ASOS network with high skill with respect to network- and station-average means albeit with a tendency to overpredict at less windy sites and underpredict at locations with faster winds. With other networks, overpredictions were much more common, even if the anemometer measurements were made at 10 m, and the "bias-bias" was even more pronounced. In prior work (e.g., [28,29]), we used station gust information to mitigate the flaws in the sustained wind forecasts to produce skillful and unbiased gust predictions. However, we also wish to improve the relationship between the sustained wind forecasts and the predictions. That is the goal of this section.

As summarized in Table 1, instrument type and mounting height have been determined to influence the mean wind and thus cause forecast errors if left uncorrected. Both also impact the temporal variance as can the averaging interval. We considered making adjustments to the forecasts to compensate for each of these factors separately but have elected instead to pursue a different strategy, one that also results in an effective correction for situation and exposure.

Figure 16 presents distributions of the temporal standard deviation (σ) computed from observed and forecast winds for ASOS (top row) and non-ASOS (bottom row) stations for the target month. Among the N = 788 ASOS stations (Figure 16a), the two distributions look quite similar, with the mean σ for the observations (red) and forecasts (blue) being 2.08 and 1.92 m s^{-1}, respectively. Forecasts are somewhat overrepresented at smaller values of σ and underrepresented at larger values. A local minimum around 2.5 m s^{-1} appears in both distributions for unclear reasons.

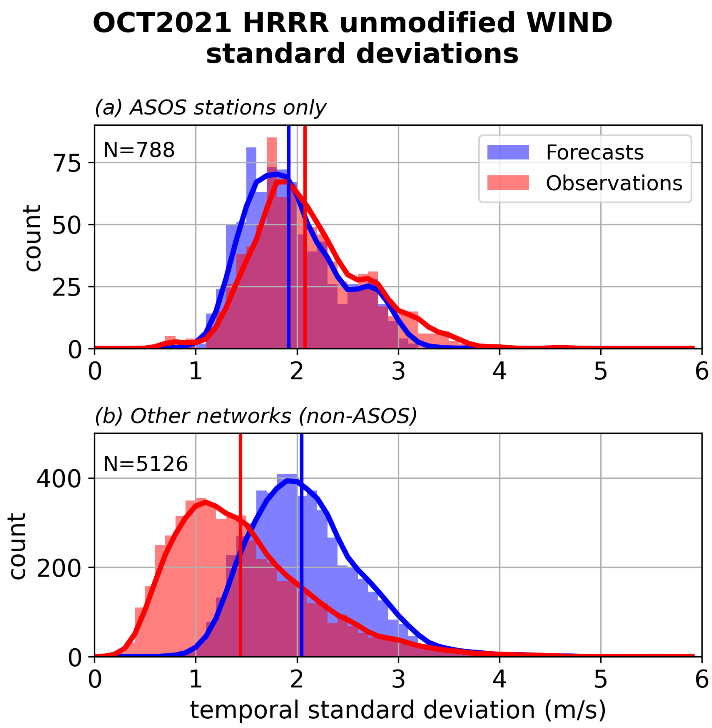

Figure 16. Histograms of station temporal standard deviations from forecasts (blue) and observations (red) from the (**a**) ASOS network (N = 788), and (**b**) non-ASOS (N = 5126) networks. The vertical solid and dashed lines indicate mean and median values, respectively.

The non-ASOS networks (RAWS, PGE, SCE, and both NYSM instruments) form a dataset consisting of N = 5126 stations. The standard deviation distribution for their forecasts closely resembles that seen among the ASOS stations, absent the local minimum about 2.5 m s^{-1} (Figure 16b). The distribution for their observations, however, is radically different. The mean observed σ of 1.44 m s^{-1} is 30% smaller than the forecast value (2.05 m s^{-1}). The HRRR is providing quite skillful forecasts for ASOS stations, albeit with some "bias-bias", but—again—the important point is that every station is essentially treated as an ASOS installation. This results in forecasts for non-ASOS sites that are biased high and have excessively large range and temporal variance relative to the observations.

Whatever distinguishes ASOS from non-ASOS stations is not something the model is presently resolving or capturing. We believe these unresolved factors are manifested in the temporal standard deviation, as was also hinted at in Figure 8d of [29]. The stations of these non-ASOS networks are more likely to be mounted closer to the ground, use non-sonic anemometers, and construct reports having longer averaging intervals, all of which encourage smaller temporal variations. These sites are also much more likely to be situated in less exposed terrain, further reducing σ.

Figure 17 combines all of our observations into a single dataset with N = 5916 stations, with markers color-coded by network membership. Viewing the station mean forecast winds vs. the biases (Figure 17a), we again see little correlation superimposed upon a positive average bias. When plotted against station mean observations (Figure 17b), the "bias-bias" emerges, with a aggregate R^2 of 0.47 ($r = -0.69$). In panel (c), we see an even stronger negative relationship ($R^2 = 0.69$, $r = -0.83$) between bias and the ratio of the observed to forecast standard deviation, $\frac{\sigma_o}{\sigma_f}$. Clearly, sites with ratios < 1 were largely overpredicted and those with ratios exceeding 1 had primarily negative biases. Temporal variation is reduced when readings are taken closer to the ground and/or in less exposed locations, and averaged longer and/or recorded with a propeller instrument, and this contributes to overprediction. That said, some stations are underpredicted and those had a strong tendency to vary more in time than both other stations in the non-ASOS catalogue and their model forecasts.

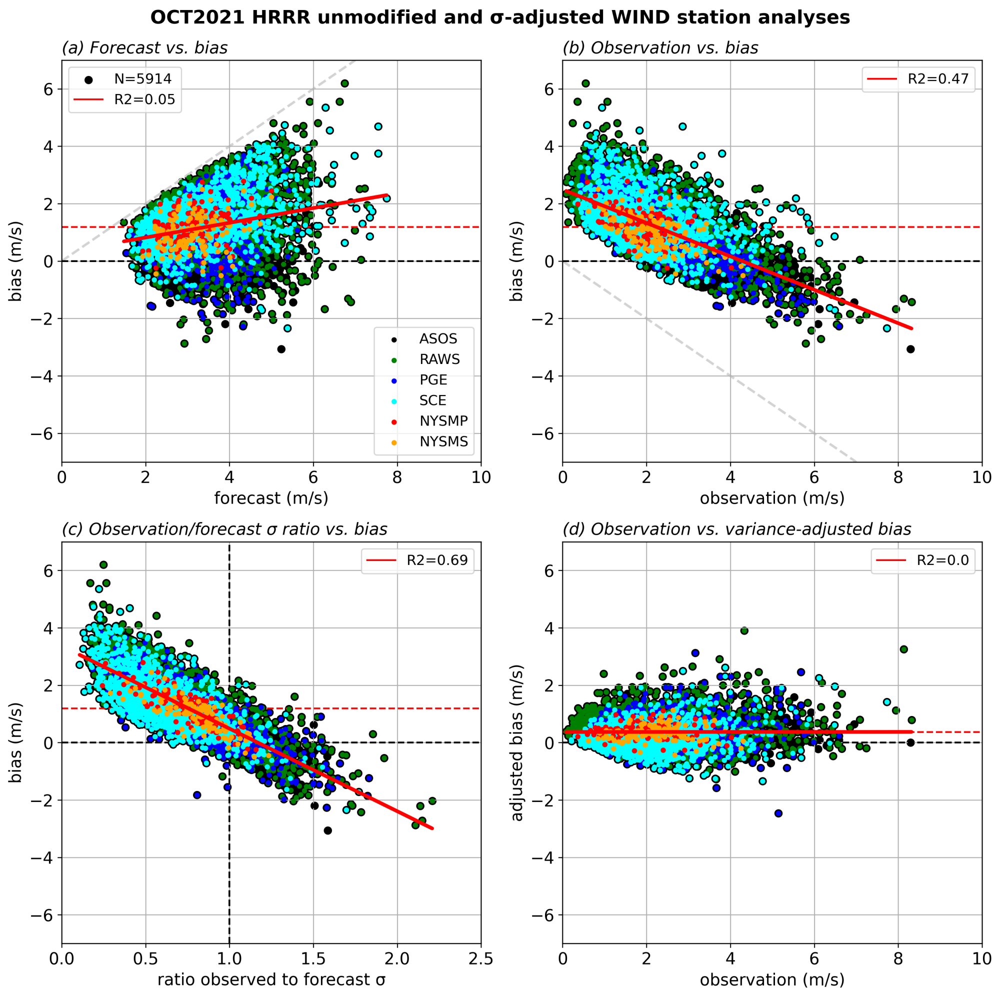

Figure 17. Station analysis of unmodified and σ-adjusted forecasts combining all networks: (**a**) forecast vs. bias; (**b**) observed wind vs. bias; (**c**) observation to forecast σ ratio vs. bias; and (**d**) observed wind vs. bias after adjustment. Markers are color-coded by network membership. Grey dashed diagonals represent limits imposed by the non-negativity constraint of wind speed. Regression lines shown in red.

Thus, as a practical corrective, we propose to transform the standard deviations from the forecasts to match those from the observations, on a station by station basis, to see if that resolves the problem. This is a postprocessing adjustment that is perhaps similar to what a machine learning algorithm operating on these data might discover. We applied to the raw HRRR forecasts at every station and forecast time a correction factor consisting of the $\frac{\sigma_o}{\sigma_f}$ ratio computed for the station, producing a variance-adjusted forecast that is guaranteed to possess the σ observed at the site. Variance-adjusted biases were then computed from these revised forecasts. In this transformed dataset, there is no relationship between the mean station observed wind speed and adjusted bias (Figure 17d). The network and station analysis for the σ-adjusted dataset (Figures 18 and 19) reveal that some unwanted small errors and correlations remain but overall there has been substantial improvement and the model is now providing more useful proxies for conditions being observed in the field, especially at non-ASOS sites.

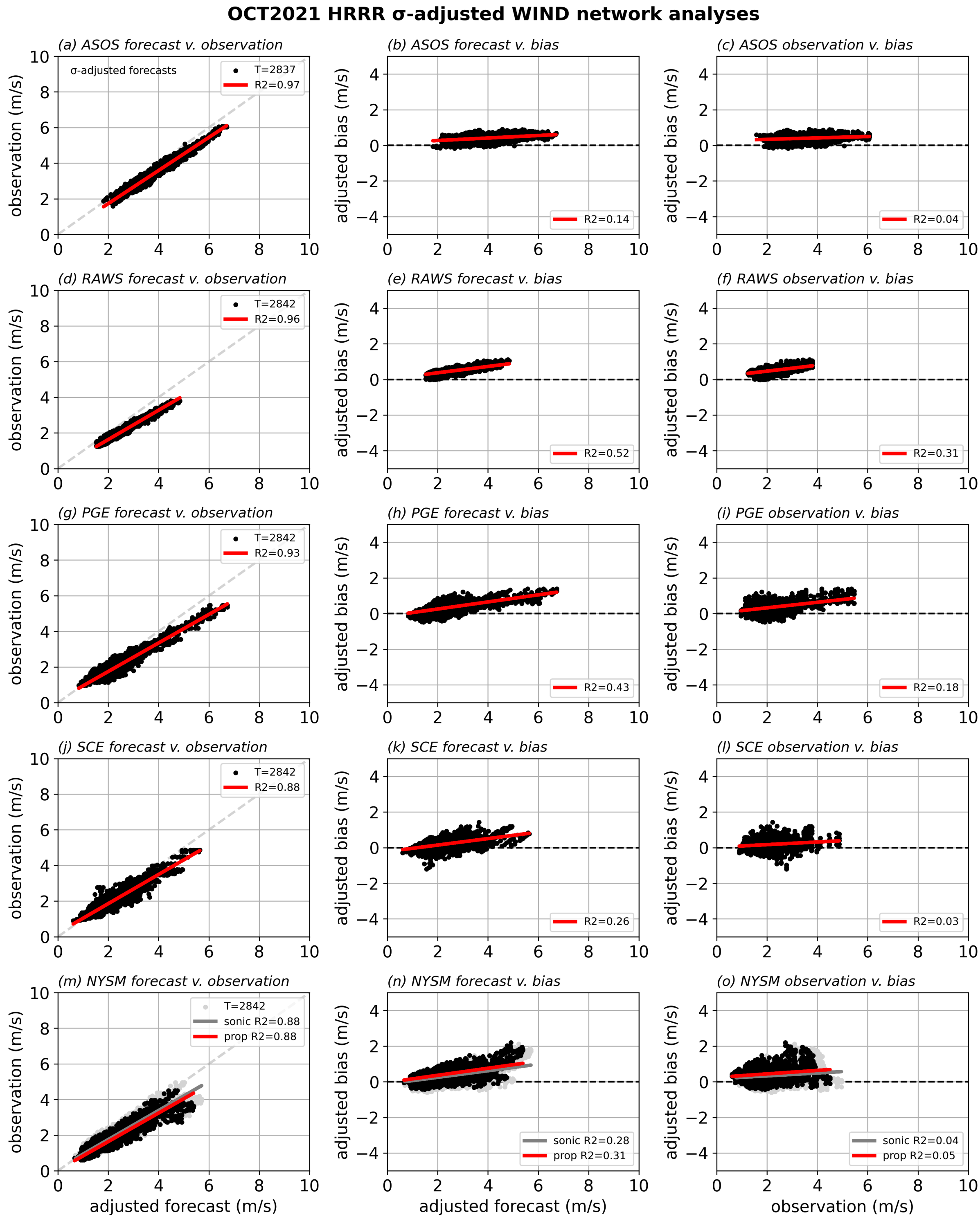

Figure 18. Similar to Figure 2 but for the σ-adjusted network analyses.

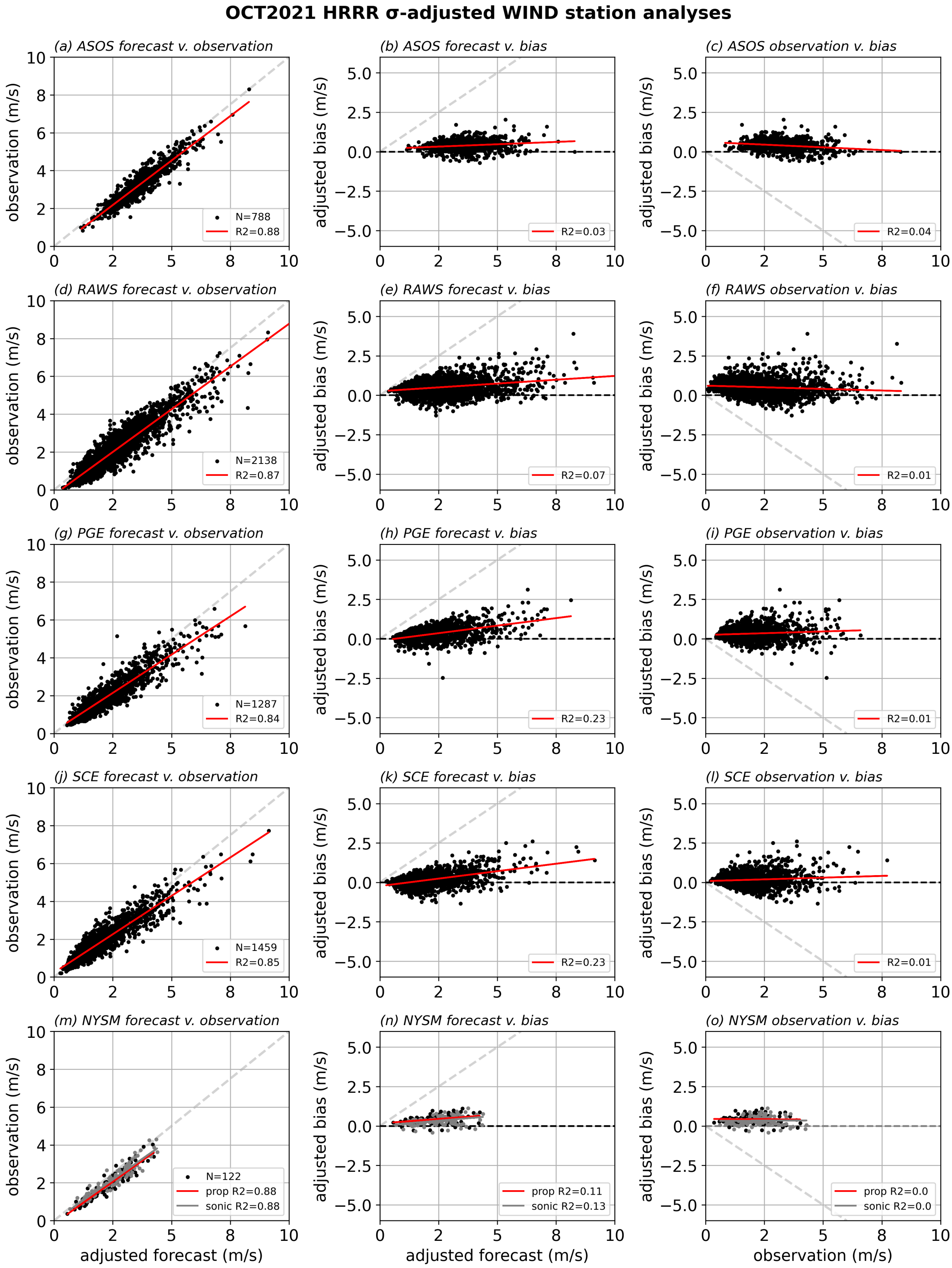

Figure 19. Similar to Figure 4 but for the σ-adjusted station analyses.

We emphasize that we did not perform any explicit and individual corrections for mounting height, instrument type, averaging interval, and/or situation and exposure, and did not make any direct mean adjustment or bias correction. We simply made sure the forecasts for a given station had the same temporal variation as observed at that site. The biases that remain might perhaps be addressed by reconsidering the surface roughnesses in some areas of the landscape. It does not seem reasonable that the mean roughness values that characterize the PGE and SCE networks in the HRRR are actually smaller than those for ASOS (Table 1). As noted by [32], “roughness estimation is very much an art”, and z_0 has both an empirical as well as theoretical role to play in the wind simulations.

6. Summary

We have presented a comprehensive verification of HRRR model forecasts against routinely measured near-surface observations of the sustained wind. Wind speeds are measured by anemometers of a variety of types and mounted at a range of heights above the surface. Sustained winds are then created by averaging samples over some temporal length—intervals from 2 to 10 min are common—and reported at intervals that vary among networks. Stations also differ enormously with respect to their situation and exposure. While a subset of observations are taken in the relatively well-exposed environments that characterize many airports, others are sited in more challenging conditions, such as on hillsides or in canyons, and/or near trees and buildings, that cannot be represented properly on the grid of operational numerical weather prediction models currently in use.

Our analysis focused on sustained winds measured in both CONUS-wide (ASOS, RAWS) and regional networks, the latter including the New York State Mesonet (NYSM) and two utility mesonets established by California public utilities (PGE and SCE). Our focus was on October 2021 and forecasts from the 00 and 12 UTC cycles of the operational HRRR (version 4) model. Forecasts and observations were compared using the Model Evaluation Tools (MET) and aggregated into network analyses (averaged across stations) and station analyses (averaged across forecast times).

Consistent with our prior work that examined different time periods [27,29], we found that the HRRR provided excellent network-averaged forecasts for ASOS stations, independent of time of day, over the $T \approx 2800$ forecast hours included in our study. The CONUS-mean forecasts and observations fell along the 1:1 correspondence line with minimal scatter ($R^2 \approx 1$). There was no correlation between bias (forecast minus observation) with temporal mean forecasted or observed wind speed, which means that the errors were functions of neither predicted nor measured wind speed. The station analysis, which compared temporal mean forecasts and observations for the $N \approx 800$ sites, also resembled past findings in that there was general (albeit lower) correspondence between station means of forecasts and observations with a tendency for windier locations to be underpredicted and for positive biases where observed winds were slower. We referred to this as the “bias-bias” because the forecast wind bias was itself intrinsically biased, and in a manner that reduced the usefulness of the predictions.

There were significant shortcomings in the raw (unadjusted) model forecasts for the other networks examined herein. While there was still a very good relationship ($R^2 \approx 1$) between network-averaged forecast and observation for all of the networks, mean winds were overpredicted by roughly 40% for all non-ASOS networks examined. This implied the existence of large correlations between bias and both forecast and observed winds. Station analyses presented lower skill, with forecasts generally having markedly narrower frequency distributions than the observations. One consequence of that was the aforementioned “bias-bias” was substantially more pronounced.

Simple ways of adjusting either the forecasts or observations were considered and discarded, as these did not resolve (or even exacerbated) the “bias-bias”. We hypothesized that forecasting issues, even with respect to station means, were a consequence of discrepancies between the observed and modeled temporal standard deviations (σ). The variability of the wind measured in the field is influenced by many factors, including mounting height, instrument type, and the time interval used to compute the sustained wind as well as the station's environment (situation and exposure). In making its forecasts, the numerical model is not taking into account the first three listed factors and the last is being only partially addressed via parameters such as roughness length (z_0).

Ultimately, we elected to adjust the forecasts in post-processing by giving each station's forecasts the temporal standard deviation (σ) observed at that location. This approach was found to remove nearly all of the mean error and "bias-bias" among the adjusted forecasts, rendering the predictions and observations more compatible. The result was that the forecasts became much more accurate assessments of what was actually measured in the field. We note that we have examined other time periods (months, seasons) and while in some cases particulars can differ the main conclusions we came to herein are robust.

In this example, October 2021 forecasts were adjusted using σ values computed from October 2021 observations. In an operational setting, more sophisticated approaches would likely be used, perhaps including having σ vary by season and/or computed from long-term records. We also stipulate that we have provided a solution to the problem but not *the* solution. One very important reason we use NWP models is to obtain guidance for locations that do not have an observing site. A useful advance would be to be able to predict from the model itself what a reasonable temporal standard deviation would be for a particular site, instead of relying on external (observation-based) corrections. Perhaps the model-derived correction will be a function of landuse category or subgrid orography or other factors. This is a topic for future research.

Author Contributions: Conceptualization, R.G.F.; Data curation, R.G.F.; Formal analysis, R.G.F. and S.B.C.; Funding acquisition, R.G.F. and S.B.C.; Methodology, R.G.F.; Writing—original draft, R.G.F.; Writing—review and editing, R.G.F. and S.B.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Atmospheric Data Solutions LLC.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: HRRR model outputs were obtained from Amazon Web Services, accessed in May 2023. Observation data sources included MADIS (https://madis-data.ncep.noaa.gov), NCEI (https://ncei.noaa.gov), NWTC (http://dx.doi.org/10.5439/1052222), and the New York State Mesonet (https://nysmesonet.org; with permission), accessed on 17 May 2023.

Conflicts of Interest: The authors declare that this study received funding from Atmospheric Data Solutions LLC. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Abbreviations

The following abbreviations and acronyms are used in this manuscript:

AGL	Above Ground Level
ASOS	Automated Surface Observing System
CONUS	Conterminous United States
CWOP	Citizen Weather Observer Program
HRRR	High-Resolution Rapid Refresh model
MADIS	Meteorological Assimilation Data Ingest System
NCEI	National Centers for Environmental Information
NOAA	National Oceanic and Atmospheric Administration
NWTC	National Wind Technology Center
NYSM	New York State Mesonet

NYSMP	New York State Mesonet propeller
NYSMS	New York State Mesonet sonic
PGE	Pacific Gas and Electric
RAWS	Remote Automated Weather Stations
SCE	Southern California Edison
SDGE	San Diego Gas and Electric
SLP	Sea Level Pressure
WRF	Weather Research and Forecasting model
WMO	World Meteorological Organization

References

1. Siuta, D.; West, G.; Stull, R. WRF hub-height wind forecast sensitivity to PBL scheme, grid length, and initial condition choice in complex terrain. *Weather Forecast.* **2017**, *32*, 493–509. [CrossRef]
2. Salfate, I.; Marin, J.C.; Cuevas, O.; Montecinos, S. Improving wind speed forecasts from the Weather Research and Forecasting model at a wind farm in the semiarid Coquimbo region in central Chile. *Wind Energy* **2020**, *23*, 1939–1954. [CrossRef]
3. Tan, E.; Mentes, S.S.; Unal, E.; Unal, Y.; Efe, B.; Barutcu, B.; Onol, B.; Topcu, H.S.; Incecik, S. Short term wind energy resource prediction using WRF model for a location in western part of Turkey. *J. Renew. Sustain. Energy* **2021**, *13*, 013303. [CrossRef]
4. Xu, W.; Liu, P.; Cheng, L.; Zhou, Y.; Xia, Q.; Gong, Y.; Liu, Y. Multi-step wind speed prediction by combining a WRF simulation and an error correction strategy. *Renew. Energy* **2021**, *113*, 772–782. [CrossRef]
5. Ye, F.; Brodie, J.; Miles, T.; Aziz Ezzat, A. AIRU-WRF: A physics-guided spatio-temporal wind forecasting model and its application to the U.S. Mid Atlantic offshore wind energy areas. *Renew. Energy* **2024**, *223*, 119934. [CrossRef]
6. Carpenter, R.L.; Gowan, T.A.; Lillo, S.P.; Strenfel, S.J.; Eiserloh, A.J.; Duffey, E.J.; Qu, X.; Capps, S.B.; Liu, R.; Zhuang, W. Improving a WRF-based high-impact weather forecast system for a Northern California power utility. *Atmosphere* **2024**, *15*, 1244. [CrossRef]
7. Public Utilities Commission of the State of California. *Resolution WSEB-1*; Public Utilities Commission of the State of California: San Francisco, CA, USA, 2023. Available online: https://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M520/K467/520467882.PDF (accessed on 21 December 2023).
8. Fovell, R.G.; Cao, Y. The Santa Ana winds of Southern California: Winds, gusts, and the 2007 Witch fire. *Wind Struct.* **2017**, *113*, 529–564. [CrossRef]
9. Keeley, J.E.; Safford, H.; Fotheringham, C.J.; Franklin, J.; Moritz, M. The 2007 Southern California wildfires: Lessons in complexity. *J. For.* **2009**, *107*, 287–296. [CrossRef]
10. Fovell, R.G.; Gallagher, A. Winds and gusts during the Thomas fire. *Fire* **2018**, *1*, 47. [CrossRef]
11. Mass, C.F.; Ovens, D. The Northern California wildfires of 8–9 October 2017. *Bull. Am. Meteor. Soc.* **2019**, *100*, 235–256. [CrossRef]
12. Brewer, M.J.; Clements, C.B. The 2018 Camp fire: Meteorological analysis using in situ observations and numerical simulations. *Atmosphere* **2019**, *11*, 47. [CrossRef]
13. Mass, C.F.; Ovens, D. The synoptic and mesoscale evolution accompanying the 2018 Camp fire of Northern California. *Bull. Am. Meteor. Soc.* **2020**, *102*, E168–E192. [CrossRef]
14. Mass, C.F.; Ovens, D. The Meteorology of the August 2023 Maui Wildfire. *Weather Forecast.* **2024**, *39*, 1097–1115. [CrossRef]
15. Cao, Y.; Fovell, R.G. Downslope windstorms of San Diego County. Part I: A case study. *Mon. Weather. Rev.* **2016**, *144*, 529–552. [CrossRef]
16. Wilczak, J.; Finley, C.; Freedman, J.; Cline, J.; Bianco, L.; Olson, J.; Djalalova, I.; Sheridan, L.; Ahlstrom, M.; Manobianco, J.; et al. The Wind Forecast Improvement Project (WFIP): A Public–Private Partnership Addressing Wind Energy Forecast Needs. *Bull. Am. Meteor. Soc.* **2015**, *96*, 1699–1718. [CrossRef]
17. Wilczak, J.; Stoelinga, M.; Berg, L.K.; Sharp, J.; Draxl, C.; McCaffrey, K.; Banta, R.M.; Bianco, L.; Djalalova, I.; Lundquist, J.K.; et al. The Second Wind Forecast Improvement Project (WFIP2): Observational field campaign. *Bull. Am. Meteor. Soc.* **2019**, *100*, 1701–1723. [CrossRef]
18. Shaw, W.J.; Berg, L.K.; Cline, J.; Draxl, C.; Djalalova, I.; Grimit, E.P.; Lundquist, J.K.; Marquis, M.; McCaa, J.; Olson, J.B.; et al. The Second Wind Forecast Improvement Project (WFIP2): General Overview. *Bull. Am. Meteor. Soc.* **2019**, *100*, 1687–1699. [CrossRef]
19. Dowell, D.C.; Alexander, C.R.; Dowell, D.C.; Weygandt, S.S.; Benjamin, S.G.; Manikin, G.S.; Brown, J.M.; Olson, J.B.; Hu, M.; Smirnova, T.G.; et al. The High-Resolution Rapid Refresh (HRRR): An hourly updating convection-allowing forecast model. Part 1: Motivation and system description. *Weather Forecast.* **2022**, *37*, 1371–1395. [CrossRef]
20. James, E.P.; Alexander, C.R.; Dowell, D.C.; Weygandt, S.S.; Benjamin, S.G.; Manikin, G.S.; Brown, J.M.; Olson, J.B.; Hu, M.; Smirnova, T.G.; et al. The High-Resolution Rapid Refresh (HRRR): An hourly updating convection-allowing forecast model. Part II: Forecast performance. *Weather Forecast.* **2022**, *37*, 1397–1417. [CrossRef]

21. Pichugina, Y.L.; Banta, R.M.; Bonin, T.; Brewer, W.A.; Choukulkar, A.; McCarty, B.J.; Baidar, S.; Draxl, C.; Fernando, H.J.S.; Kenyon, J.; et al. Spatial variability of winds and HRRR–NCEP model error statistics at three Doppler-lidar sites in the wind-energy generation region of the Columbia River Basin. *J. Appl. Meteor. Clim.* **2019**, *58*, 1633–1656. [CrossRef]
22. Fovell, R.G.; Brewer, M.J.; Garmong, R.J. The December 2021 Marshall fire: Predictability and gust forecasts from operational models. *Atmosphere* **2022**, *13*, 765. [CrossRef]
23. Fovell, R.G.; Brewer, M.J. Sundowner Winds at Montecito during the Sundowner Winds Experiment. *Atmosphere* **2024**, *15*, 810. [CrossRef]
24. Collins, E.; Lebo, Z.J.; Cox, R.; Hammer, C.; Brothers, M.; Geerts, B.; Capella, R.; McCorkle, S. Forecasting high wind events in the HRRR model over Wyoming and Colorado. Part I: Evaluation of wind speeds and gusts. *Weather Forecast.* **2024**, *39*, 705–723. [CrossRef]
25. Collins, E.; Lebo, Z.J.; Cox, R.; Hammer, C.; Brothers, M.; Geerts, B.; Capella, R.; McCorkle, S. Forecasting high wind events in the HRRR model over Wyoming and Colorado. Part II: Sensitivity of surface wind speeds to model resolution and physics. *Weather Forecast.* **2024**, *39*, 725–743. [CrossRef]
26. Skamarock, W.C.; Klemp, J.B.; Dudhia, J.; Gill, D.O.; Liu, Z.; Berner, J.; Wang, W.; Powers, J.G.; Duda, M.G.; Barker, D.M.; et al. *A Description of the Advanced Research WRF Version 4*; NCAR Tech. Note NCAR/TN-556+STR; National Center for Atmospheric Research: Boulder, CO, USA, 2019; p. 145. [CrossRef]
27. Fovell, R.G.; Gallagher, A. Boundary layer and surface verification of the High-Resolution Rapid Refresh, version 3. *Weather Forecast.* **2020**, *35*, 2255–2278. [CrossRef]
28. Cao, Y.; Fovell, R.G. Downslope windstorms of San Diego County. Part II: Physics ensemble analyses and gust forecasting. *Weather Forecast.* **2018**, *33*, 539–559. [CrossRef]
29. Fovell, R.G.; Gallagher, A. An evaluation of surface wind and gust forecasts from the High-Resolution Rapid Refresh model. *Weather Forecast.* **2022**, *37*, 1049–1068. [CrossRef]
30. Brotzge, J.A.; Wang, J.; Thorncroft, C.D.; Joseph, E.; Bain, N.; Bassill, N.; Farruggio, N.; Freedman, J.M.; Hemker, K.; Johnston, D.; et al. A technical overview of the New York State Mesonet standard network. *J. Atmos. Ocean. Tech.* **2020**, *37*, 1827–1845. [CrossRef]
31. Bullock, R.; Fowler, T.; Gotway, J.H.; Newman, K.; Jensen, T.; Meier-fleischer, K.; Carslaw, D.C.; Ropkins, K. *Model Evaluation Tools Version 6.1 (METv6.1): Users Guide*; Developmental Testbed Center (DTC): Boulder, CO, USA, 2017. [CrossRef]
32. Powell, M.D.; Houston, S.H.; Reinhold, T.A. Hurricane Andrew's landfall in south Florida. Part I: Standardizing measurements for documentation of surface wind fields. *Weather Forecast.* **1996**, *11*, 304–328. [CrossRef]
33. Wieringa, J. Roughness-dependent geographical interpolation of surface wind speed averages. *Quart. J. Roy. Meteor. Soc.* **1986**, *112*, 867–889. [CrossRef]
34. Harper, B.A.; Kepert, J.D.; Ginger, J.D. *Guidelines for Converting Between Various Wind Averaging Periods in Tropical Cyclone Conditions*; WMO Tech. Doc. TD-1555; WMO: Geneva, Switzerland, 2010; p. 64. Available online: http://www.systemsengineeringaustralia.com.au/download/WMO_TC_Wind_Averaging_27_Aug_2010.pdf (accessed on 15 August 2018).
35. World Meteorological Organization. *Guide to Meteorological Instruments and Methods of Observation*; WMO Tech. Doc. 8; WMO: Geneva, Switzerland, 2018; p. 573. Available online: https://library.wmo.int/idurl/4/41650 (accessed on 1 September 2021).
36. Wieringa, J. Cup, propeller, vane, and sonic anemometers in turbulence research. *Ann. Rev. Fluid Mech.* **1981**, *13*, 399–423. [CrossRef]
37. Finkelstein, P.L.; Kaimal, J.C.; Gaynor, J.E.; Graves, M.E.; Lockhart, T.J. Comparison of wind monitoring systems. Part I: In situ sensors. *J. Atmos. Ocean. Tech.* **1986**, *3*, 583–593. [CrossRef]
38. Bowen, B.M. Improved wind and turbulence measurements using a low-cost 3-D sonic anemometer at a low-wind site. *Open Atmos. Sci. J.* **2008**, *2*, 131–138. [CrossRef]
39. Suomi, I.; Vihma, T. Wind gust measurement techniques—From traditional anemometry to new possibilities. *Sensors* **2008**, *18*, 1300. [CrossRef] [PubMed]
40. Wang, J.; Brotzge, J.; Shultis, J.; Bain, N. Enhancing icing detection and characterization using the New York State Mesonet. *J. Atmos. Ocean. Tech.* **2021**, *38*, 1499–1514. [CrossRef]