Impacts of Northeastern Pacific Buoy Surface Pressure Observations

CAROLYN A. REYNOLDS,[a] REBECCA E. STONE,[b] JAMES D. DOYLE,[a] NANCY L. BAKER,[a] ANNA M. WILSON,[c] F. MARTIN RALPH,[c] DAVID A. LAVERS,[d] ANEESH C. SUBRAMANIAN,[e] AND LUCA CENTURIONI[f]

[a] *Marine Meteorology Division, U.S. Naval Research Laboratory, Monterey, California*
[b] *Science Applications International Corporation, Monterey, California*
[c] *Center for Western Weather and Water Extremes, Scripps Institution of Oceanography, University of California, San Diego, San Diego, California*
[d] *European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom*
[e] *University of Colorado Boulder, Boulder, Colorado*
[f] *Scripps Institution of Oceanography, University of California, San Diego, San Diego, California*

(Manuscript received 27 April 2022, in final form 4 October 2022)

ABSTRACT: Under the Atmospheric River Reconnaissance (AR Recon) Program, ocean drifting buoys (drifters) that provide surface pressure observations were deployed in the northeastern Pacific Ocean to improve forecasts of U.S. West Coast high-impact weather. We examine the impacts of both AR Recon and non-AR Recon drifter observations in the U.S. Navy's global atmospheric data assimilation (DA) and forecast system using data-denial experiments and forecast sensitivity observation impact (FSOI) analysis, which estimates the impact of each observation on the 24-h global forecast error total energy. Considering all drifters in the eastern North Pacific for the 2020 AR Recon season, FSOI indicates that most of the beneficial impacts come from observations in the lowest quartile of observed surface pressure values, particularly those taken late in the DA window. Observations in the upper quartile have near-neutral impacts on average and are slightly nonbeneficial when taken late in the DA window. This may occur because the DA configuration used here does not account for model biases, and innovation statistics show that the forecast model has a low pressure bias at high pressures. Case studies and other analyses indicate large beneficial impacts coming from observations in regions with large surface pressure gradients and integrated vapor transport, such as fronts and ARs. Data-denial experiments indicate that the assimilation of AR Recon drifter observations results in a better-constrained analysis at nearby non-AR Recon drifter locations and counteracts the NAVGEM pressure bias. Assimilating the AR Recon drifter observations improves 72- and 96-h Northern Hemisphere forecasts of winds in the lower and middle troposphere, and geopotential height in the lower, middle, and upper troposphere.

SIGNIFICANCE STATEMENT: The purpose of this study is to understand how observations of atmospheric pressure at the ocean surface provided by drifting buoys impact weather forecasts. Some of these drifting buoys were deployed under a program to study atmospheric rivers (ARs) to improve forecasts of high-impact weather on the West Coast. We find that these observations are most effective at reducing forecast errors when taken in regions near fronts and cyclones. The additional drifting buoys deployed under the AR Reconnaissance project reduce forecast errors at 72 and 96 h over North America and the Northern Hemisphere. These results are important because they illustrate the potential for improving forecasts by increasing the number of drifting buoy surface pressure observations over the world oceans.

KEYWORDS: Atmospheric river; Buoy observations; Forecast verification/skill; Numerical weather prediction/forecasting; Data assimilation

1. Introduction

Atmospheric rivers (ARs), elongated corridors of horizontal moisture transport, account for more than 90% of the extratropical vapor transport (Zhu and Newell 1998) and are associated with both beneficial and hazardous impacts at landfall (Ralph and Dettinger 2012). ARs account for a large fraction of the annual precipitation and floods over the western United States (Dettinger et al. 2011; Dettinger 2013) and are primary drivers of flood damage (Corringham et al. 2019). ARs are also often associated with other hazardous impacts such as landslides (Cordeira et al. 2019), high winds (Waliser and Guan 2017), and avalanches (Hatchett et al. 2017). Despite steady progress in numerical weather prediction over the years, there are still significant forecast errors in short-range prediction of ARs in terms of integrated vapor transport (IVT) and landfall location (Nardi et al. 2018). Given the potentially extreme impacts, it is important to have AR forecasts as accurate as possible.

The AR Reconnaissance (AR Recon) project is a research and operations partnership developed to address the need to improve AR forecasts (Ralph et al. 2020). As part of AR Recon, dropsonde observations are taken in regions where it is determined that more accurate analyses are likely to improve forecasts of the ARs and their impacts (Cobb et al. 2023). These AR Recon dropsondes have been shown to improve forecasts of ARs (Zheng et al. 2021a), and to have comparable impact on

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/MWR-D-22-0124.s1.

Corresponding author: Carolyn Reynolds, carolyn.reynolds@nrlmry.navy.mil

DOI: 10.1175/MWR-D-22-0124.1

24-h global forecast errors relative to the North American radiosonde network (Stone et al. 2020). The AR Recon dropsondes have also been used to diagnose and assess biases and errors in reanalyses (Cobb et al. 2021) and analyses and short-term forecasts (Lavers et al. 2018, 2020a; Stone et al. 2020). These observations tend to be very impactful as they are taken in regions where there are few other in situ observations, and often are taken in cloudy and precipitating regions where satellite observations provide less information than in clear regions (Zheng et al. 2021b).

In addition to the deployment of dropsondes, AR Recon has also deployed drifting buoys, or drifters (Centurioni 2018), in partnership with the Global Drifter Program, (Lavers et al. 2020b; Ralph et al. 2020) that provide surface pressure observations (Centurioni et al. 2017b) in the northeastern Pacific. The hourly surface pressure observations from the AR Recon drifters serve as a complement to the episodic dropsonde profile observations and substantially enhance the network of non-AR Recon drifters of the same type that are maintained by the Global Drifter program. Previous studies find substantial impacts of the global drifter surface pressure observations on forecast skill. Centurioni et al. (2017b) find substantial degradations in mean SLP forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecast System (IFS) when the global drifter surface pressure observations are not assimilated. The impacts are seen primarily over ocean regions at short lead times (e.g., out to 72 h), but extend over land by 96 and 120 h. Denial of the drifter observations also degrades forecasts of wind vectors in the troposphere out to three days. Centurioni et al. (2017b) also employ the forecast sensitivity observation impact technique (FSOI; Langland and Baker 2004), which quantifies the value of each observation in reducing the forecast error, and find that the per-observation impact of the drifters is larger than that of any other component of the observing system. In a more detailed examination of the data-denial experiments reported in Centurioni et al. (2017b); Horányi et al. (2017) find the largest fractional impact of the drifter surface pressure observations in the Arctic, the poorly observed Southern Ocean, and dynamically active regions such as the North Atlantic Ocean near Greenland and the North Pacific Ocean near Alaska. The drifter surface observations contribute approximately 3% to the total forecast error reduction on average; however, their impact can be much larger regionally (over 50%) during complex or rapidly evolving cyclogenesis cases. In a cost-benefit observing system study using FSOI, only aircraft reports were found to have a larger benefit per cost than the drifter surface pressure observations (Eyre and Reid 2014).

As noted in Ingleby and Isaksen (2018), the large impact from the drifter surface pressure observations is not surprising given that surface pressure provides information on synoptic-scale variability, the observations are often within in situ data-sparse regions, and satellite data provide little information about surface pressure. In their FSOI examination, Ingleby and Isaksen (2018), find the largest drifter surface pressure impacts in the ECMWF IFS occur in the North Atlantic and North Pacific storm tracks and over the Southern Ocean, regions that are spatially well correlated with baroclinicity. In a series of data-denial experiments, denial of the Northern Hemisphere drifter observations leads to significant degradations in midlatitude tropospheric geopotential height anomaly correlations at short ranges. The largest changes in the analyses occur in the western parts of the Pacific and Atlantic. Given these results, they advocate for upgrading drifters to include surface pressure observations when possible, and would prioritize additional drifters in regions of baroclinic development, where surface pressure observations are particularly impactful.

Motivated by the large impact of the drifter surface pressure observations shown in previous studies, we examine the impact of drifter surface pressure observations in the northeastern Pacific in the U.S. Navy's global prediction system. This study will allow us to identify weaknesses in the prediction system, understand the conditions under which the drifter observations are most impactful, and quantify the impact of the AR Recon drifters on analyses and forecasts. This paper is a complement to the investigation of the impact of the AR Recon Dropsondes in the Navy global system (Stone et al. 2020). We employ both FSOI and data-denial techniques in this study, as these techniques provide independent complementary information on observation impacts. Data-denial experiments quantify the impact of removing a set of observations on forecast error, while FSOI measures the contribution of individual observations (or sets of observations) to short-term forecast error reduction when those observations are assimilated. See Eyre (2021) for a theoretical study exploring the differences between the two metrics. In the first part of this study, we use FSOI and observation minus background and observation minus analysis statistics applied to all (AR Recon and non-AR Recon) drifter surface pressure observations in the northeastern Pacific to highlight model error and explore how drifter impact varies by location, observation value, and time in the data assimilation (DA) window. We separate the FSOI and innovation statistics into observation value quartiles to shed light on how atmospheric dynamics and model biases affect observation impact. We consider case studies to help pinpoint specific physical features where additional observations are most impactful. In the second part of the study, we perform data-denial experiments to examine how the assimilation of the AR-Recon drifters affects the analyses and forecasts.

While we build on previous work, we present several findings that, to our knowledge, are new to the literature. Previous work has related drifter pressure observations to environmental conditions in a time-averaged and/or area-averaged sense (Ingleby and Isaksen 2018; Horanyi et al. 2017). Here, we relate the impact of drifter observations to the specific environmental conditions at the drifter site at the time the observation is taken, both through quartile analysis and through case studies. Previous work has shown that satellite observations taken later in the DA window have a larger beneficial impact on the forecast error than satellite observations taken earlier in the DA window (McNally 2019). We will show this is true for the drifter observations on average (broadening the generality of this finding), but our analysis goes further to show that this is dependent on the observation value. Previous work has examined the impact of withholding full sets of sea surface pressure observations. In this study, we withhold only the AR Recon drifter observations while continuing to

assimilate observations from other drifters in the region. This allows us to assess the impact of these special observations when added to a complete standard network of observations (including those from non-AR drifters).

In section 2, we describe the drifter surface pressure observations, data assimilation and forecast system, and experimental design and diagnostics. In section 3, we present the results from the FSOI analyses and data-denial experiments. We provide a brief summary and discussion in section 4.

2. Observations, models, and experimental design

a. Buoy surface pressure observations

As part of the AR Recon project, 32 drifters with surface pressure observations were deployed ahead of the 2019 Northern Hemisphere winter season, with an additional 64 drifters deployed ahead of the 2020 season (Ralph et al. 2020). The drifters deployed for the 2019 season, and 48 of the 64 drifters deployed for the 2020 season are Surface Velocity Program Barometer drifters (SVPB; https://gdp.ucsd.edu/ldl/svpb/) that provide sea surface temperature and barometric pressure. The barometer is located inside the buoy at its equator and therefore, on average, is located at sea level. Sixteen of the 64 drifters deployed for the 2020 season are Directional Wave Spectra Barometer drifters (DWSB; https://gdp.ucsd.edu/ldl/dwsbd/), which also provide a directional wave spectrum (Centurioni et al. 2017a; Lavers et al. 2020b). 2020 AR Recon drifters were deployed by a ship of opportunity between 10 and 22 January 2020 and by 53rd Weather Reconnaissance Squadron, U.S. Air Force Reserve Command, WC130-J flights on 11, 17, and 24 February. Given the staggered deployment and the limited drifter life span (drifters have a median life span of approximately 1.5 years), the number of drifters within our northeastern Pacific domain (115°–165°W, 15°–60°N) varied through the AR Recon 2020 season. There were 51 AR Recon drifters within the domain at the beginning of the AR Recon 2020 season (0000 UTC 22 January 2020) and 60 AR Recon drifters at the end of this AR Recon season (0000 UTC 12 March 2020). Additionally, there were 22 and 23 Global Drifter Program non-AR Recon drifters supplying surface pressure observations within this domain at the beginning and end of the AR Recon 2020 season, respectively. Most drifters report within a few minutes of the top of the hour, so each drifter typically contributes six observations within each 6-h DA cycle. The observations are assimilated during all four of the update cycles (centered on 0000, 0600, 1200, and 1800 UTC) each day. The instrument error contribution to the surface pressure observation errors from the drifters is characterized as 0.3–0.4-hPa RMSE and as such is considered more accurate than the 0.5-hPa errors associated with ship-automated systems (Centurioni 2018; Centurioni et al. 2019). This accuracy specification comes from the company making the sensor (Honeywell), and each sensor is calibrated and provided with a calibration certificate.

b. Data assimilation and forecast system

We are using the same U.S. Navy global operational atmospheric model and DA system used in Stone et al. (2020), and the system description here follows the description in that article closely. The system consists of the Navy Global Environmental Model (NAVGEM; Hogan et al. 2014) and the four-dimensional hybrid variational-ensemble DA system, the NRL Atmospheric Variational DA System-Accelerated Representer (NAVDAS-AR; Xu et al. 2005; Rosmond and Xu 2006; Kuhl et al. 2013). The system is very similar to the operational system run by Fleet Numerical Meteorology and Oceanography Center at that time. The forecast model is run at a T425L60 resolution, or approximately 31-km horizontal resolution, with a model top at 0.04 hPa. The adjoint and tangent linear models are run at a reduced resolution (T119), producing analysis increments of approximately 100-km resolution. For the experiments performed here, NAVDAS-AR is run as a strong constraint variational assimilation (although the option to run it in the weak constraint formulation exists). Over 100 million observations are processed in every 6-h DA cycle, with approximately 3.4 million observations assimilated after quality control (QC) and data thinning to create the final analysis. Only the observations retained after quality control and data thinning steps are used in our study. The observation types routinely assimilated are those listed in Table 2 of Stone et al. (2020) with two exceptions. In this study we are not assimilating the Ozone Mapping and Profiler Suite (OMPS) ozone retrievals because format version 6, used in the prior study, was discontinued and format version 8 was not immediately adopted. The other difference is the assimilation of the Geostationary Operational Environmental Satellite Advanced Baseline Imager (GOES ABI) clear sky radiances, which became available after the time period considered in the earlier study.

Pressure observations from drifters are subjected to QC prescreening checks prior to forming the innovation vector, as well as later being subjected to a three-sigma rejection criterion within the NAVDAS DA system. The QC prescreening measures include duplicate checking, track checking, and checking for errors related to time and elevation. A broad discussion of these issues can be found in Pauley and Ingleby (2022). The criterion for the three-sigma check, which operates on the innovation, is based on specified observation error and estimated background error, which includes a static component that varies with latitude and an ensemble component that varies over time. In our study, the drifter surface pressure observation errors range between 1.03 and 1.21 hPa. This includes the 0.3–0.4-hPa RMSE instrument error as well as representativeness error. The static portion of the background error ranges from 0.29 to 0.65 hPa. As detailed in Kuhl et al. (2013), the ensemble covariances used in the hybrid DA scheme are localized such that the 50% covariance localization is approximately 20° in both latitude and longitude, and approximately 225 hPa in the vertical. There is no bias correction applied to the observed drifter surface pressure.

Compo et al. (2011) and Slivinski et al. (2021) note in their description and evaluation of the Twentieth Century Reanalysis Project that it is possible to produce reliable tropospheric reanalyses by assimilating only surface pressure observations using an ensemble Kalman filter DA system and prescribing sea surface temperature, sea ice concentration, and radiative forcing. As Compo et al. (2011) note, outside the tropics, geostrophic balance allows one to describe a reasonable approximation to

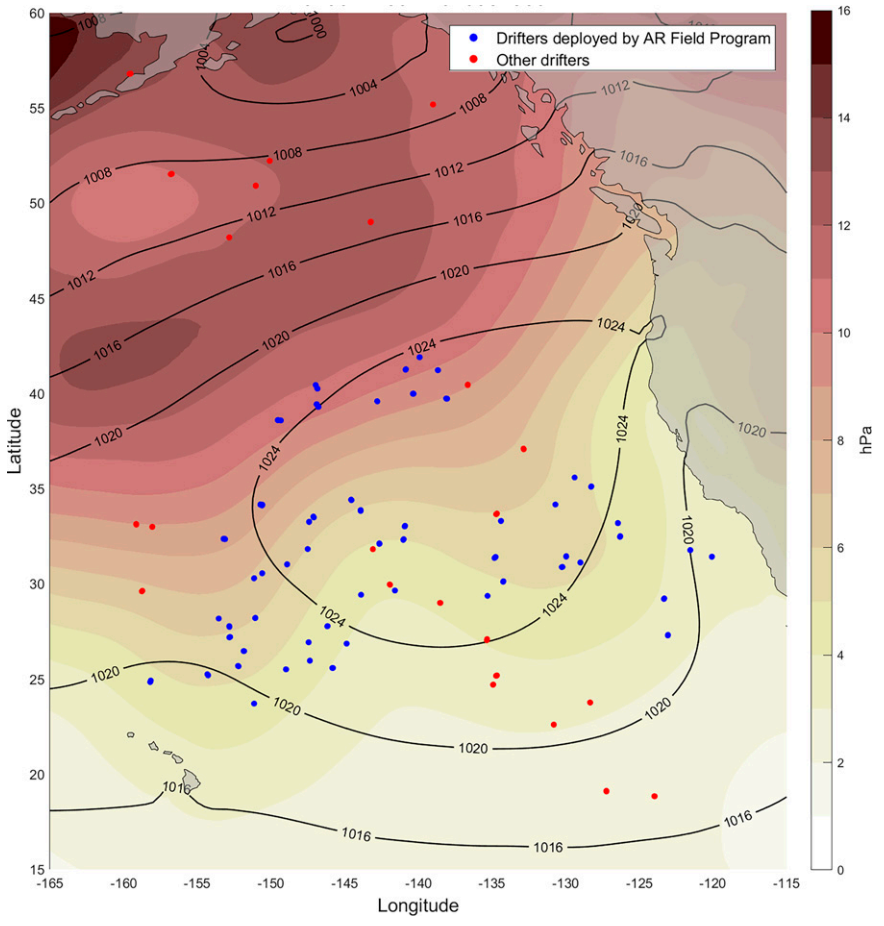

FIG. 1. Location of the 60 AR Recon (blue) and 23 non-AR Recon (red) drifters with pressure sensors at the end of the AR Recon 2020 field project season. Also shown is NAVGEM analyzed sea level pressure (hPa) average (contours) and standard deviation (shading) over the AR Recon 2020 field project season (0000 UTC 22 Jan–0000 UTC 13 Mar 2022).

the barotropic part of the flow, and surface pressure tendencies provide information about tropospheric circulations. Slivinski et al. (2021) find that version 3 of this reanalysis has 500-hPa geopotential height skill comparable to modern 3–4-day operational forecasts. The success of the Twentieth Century Reanalysis illustrates the potential impact of surface pressure observations in constraining the flow deep into the troposphere. In our system, assimilation of the AR Recon drifters after just one update cycle modifies the temperature analysis by over 0.05°C from the surface to 700 hPa, with a maximum of over 0.15°C (Fig. S1 in the online supplemental material). This is a small but nonnegligible impact when compared with the full analysis increment (background minus analysis), which reaches values of 1.2°C in this region.

The global forecasting and DA system also has the capability of computing forecast sensitivity observation impact (FSOI). As described in Stone et al. (2020), FSOI is used to quantify the contribution of individual observations or sets of observations to a reduction in the 24-h forecast error (Langland and Baker 2004). Forecast error here is measured in terms of a global moist total energy norm, including both kinetic and moist static energy (Ehrendorfer 2000). We apply FSOI to investigate the impact of the surface pressure observations from the drifters in the northeastern Pacific on forecast error. We use FSOI to understand how the impact of the observations varies as a function of location, both in a time mean sense and for case studies, and as a function of when the observations are taken within the DA window.

c. *Experimental design and diagnostics*

For the control experiments, we run the DA update cycle assimilating all data, including the AR Recon drifter surface pressure observations and dropsonde observations, for the AR Recon 2020 season, starting on 0000 UTC 22 January–0000 UTC 12 March. The 120-h forecasts are run from every 0000 and 1200 UTC analyses during that period. We run a parallel update cycle and forecast system in which the AR Recon drifter surface pressure observations are not assimilated, and refer to this as the denial experiment. To learn about model biases and the impact of the AR Recon drifter observations on the analyses, we examine DA statistics including the innovations (observations minus the 6-h background forecast, or OmB), and

observations minus the analysis (OmA). We examine the impact of both the AR Recon and non-AR Recon drifter surface pressure observations using the FSOI technique described above, computed to examine the impact on 24-h forecasts produced from the 0000, 0600, 1200, and 1800 UTC analysis times. We also compare the control and denial forecast errors (measured in terms of RMSE and anomaly correlation) over North America and the Northern Hemisphere extratropics as verified against ECMWF operational analyses for several metrics.

3. Results

a. *DA statistics and FSOI for all northeastern Pacific drifter surface pressure observations*

The location of the AR Recon and non-AR Recon Global Drifter Program drifters that provide surface pressure observations at the end of the AR Recon 2020 season (on 0000 UTC 13 March 2020) are shown in Fig. 1. The AR Recon drifters were deployed in regions meant to intersect atmospheric rivers that affect the U.S. West Coast, and are thus clustered between Hawaii and North America. The amount of drift over the AR Recon time period varies from nearly stationary to over 600 km, with the drifters in the Gulf of Alaska drifting east northeastward, and those on the eastern side of the subtropical high drifting south southwestward (see Fig. S2 in the online supplemental material). Figure 1 also shows the average and standard deviation of sea level pressure (SLP) for the 2020 AR Recon season. The AR Recon drifters are situated mostly under the time-average high pressure area and its periphery, while the non-AR Recon drifters are more widespread, extending from the subtropics to near the Aleutian low in the Gulf of Alaska. The SLP standard deviation indicates that most of the storm track activity occurs to the north and northwest of many of the AR Recon drifters for this particular season, although case studies indicate that the drifters successfully sampled ARs during the season.

We use statistics from the DA system to identify potential biases in our model. To investigate these biases and determine if the biases are conditional on the observed pressure values, we divide the observations into equally populated quartiles for subsequent analyses. The first quartile contains the lowest observed pressures (below 1019.5 hPa). The second quartile contains the observed pressures between 1019.5 and 1024.2 hPa. The third quartile contains observed pressures between 1024.2 and 1028.4 hPa, and the fourth quartile contains observations above 1028.4 hPa. We note that what we refer to as model bias in this paper is the "signal" seen by the data assimilation system as the systematic difference between the background forecasts and the assimilated observations. It is not an independent estimate of model bias by unassimilated observations.

The distribution of surface pressure innovations (OmB) differs depending on the observation quartile considered (Fig. 2). For all the quartiles, the distributions (solid lines) differ slightly from a Gaussian distribution (dashed line) with the same variance. All quartile distributions have positive kurtosis, indicating heavy tails. These heavy tail innovation distributions are common, as noted in Tavolato and Isaksen (2015). The observations in the lowest quartile have a heavy tail on the lower side of the distribution

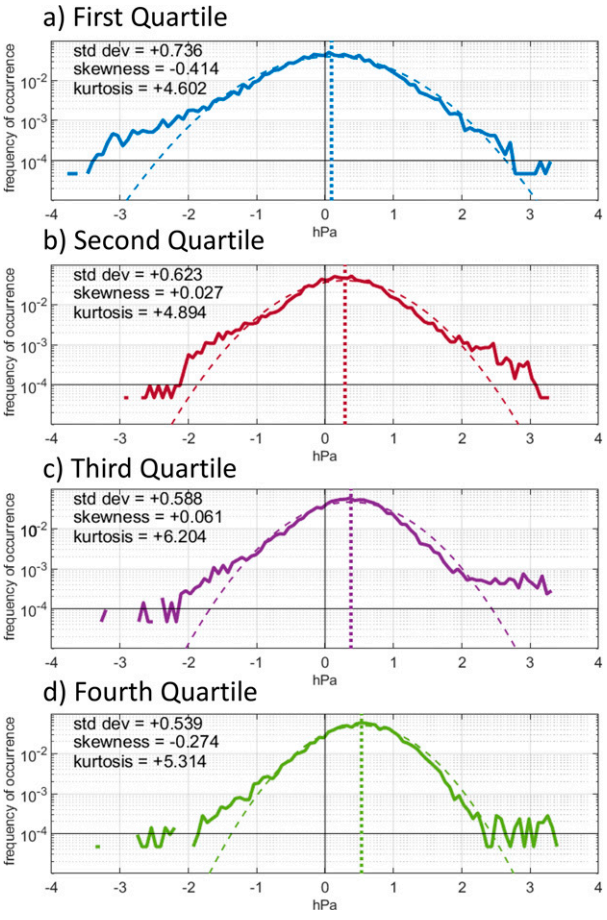

FIG. 2. Frequency of drifter surface pressure innovation (observation − background; hPa) for the four different quartiles ranging from (a) the first quartile with the lowest observed values (blue) to (d) the fourth quartile with the highest observed values (green). Frequency of the actual innovations is denoted by the solid curves, and a Gaussian distribution that matches the variance of the innovations is given by the dashed curves. The vertical dotted lines show the innovation mean value. Standard deviation, skewness, and kurtosis are provided in the panels.

(where pressure observations are lower than the background) and have the largest magnitude skewness. This may be related to the inability of the coarse resolution NAVGEM model and DA system to capture extreme low pressures with strong pressure gradients (and may also be due to the model failing to represent the lowest observed pressures that occur between grid points). Time series for observed and background values for a drifter reporting very low pressures (not shown) indicate that the model responds too slowly to observations of rapid pressure falls and the three-sigma check blocks the use of some valid information. The center of the distribution of the innovations shifts to the right for the higher quartiles, indicative of a conditional bias in NAVGEM (the low pressure bias becomes more pronounced at higher pressures). Results for quartiles determined by background pressure rather than observed pressure show some quantitative differences but are qualitatively very similar (Fig. S3

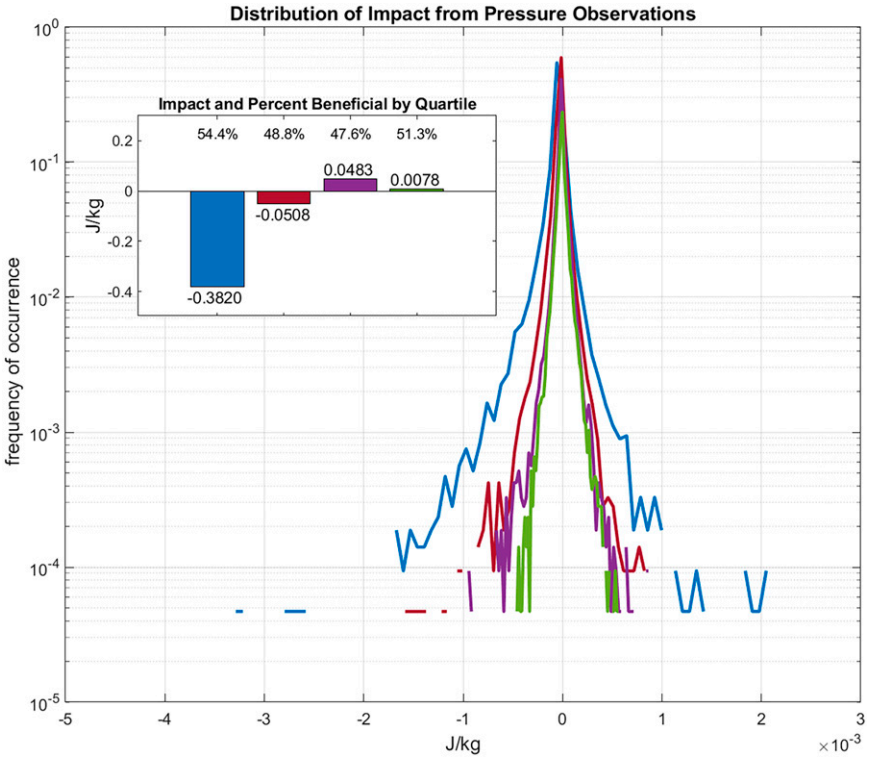

FIG. 3. Frequency distribution of drifter surface pressure observation impact (10^{-3} J kg^{-1}) for the different observation quartiles. The quartile with the lowest pressure observations is in blue, the quartile with the second-lowest pressure observations is in maroon, the quartile with the second-highest pressure observations is in purple, and the quartile with the highest pressure observations is in green. Negative values indicate beneficial impacts, and positive values indicate nonbeneficial impacts. The inset shows aggregate impact (J kg^{-1}) and percentage of observations that are beneficial for each quartile.

in the online supplemental material). While the kurtosis values remain similar for the observation-based and background-based quartiles, the skewness does change substantially, particularly for the fourth quartile, where it changes from −0.274 to −0.656 hPa.

The heavy tail to the higher side of the distribution for the middle quartiles (and to a lesser extent for the highest quartile) in Fig. 2 suggests that NAVGEM is also not capturing the extreme high pressures. Because high pressure centers are typically broad in geographic extent, this NAVGEM conditional bias may not be due to model resolution. Time series for observed and background values for a drifter reporting very high pressures (not shown) indicate that the model is able to respond to the more slowly varying changes associated with the high pressure, such that the observations are not rejected (in contrast with some rapid pressure falls associated with cyclones). Rather, the innovations in high pressure regions are small, but positive on average. Further study would be required to determine the cause of this model low bias at high pressures.

Horanyi et al. (2017) find that aggregate drifter surface pressure observation impact in the North Atlantic can spike during periods of cyclogenesis, when the observed pressure values are relatively low. We examine the FSOI frequency distributions as a function of observation quartile to see if lower pressure observations are more impactful than higher pressure observations in our northeastern Pacific domain. The frequency distribution of FSOI impact is tightly clustered near zero for all quartiles (Fig. 3). The wider distribution for the observations in the lowest quartile (blue curve) than for the other quartiles confirms that the lower pressure observations have more frequent large impacts on the 24-h forecast error than the observations in the other quartiles. These impacts are both beneficial (negative values) and nonbeneficial (positive values). On average, the observations provide a beneficial impact on the forecast, as indicated by the inset showing the aggregate impact by quartile. Most of the beneficial impact comes from observations in the lowest quartile. In contrast, observations in the other quartiles have near neutral impacts in aggregate. Similar results are found when considering quartiles based on background pressures rather than observed pressures (Fig. S4 in the online supplemental material). To examine the statistical significance of this result, we evaluate the differences between the impact of the observations in the lowest pressure quartile and the impact of the observations in the other quartiles at each analysis time. The impact from the observations in the lowest pressure quartile is significantly more beneficial than the impact from the observations in the other three quartiles at the 95% level (Figs. S5a and S6a in the online supplemental material).

More than half of all drifter observations are beneficial, although when considered by quartile, the largest fraction of beneficial observations (54.4%) occurs for the observations in the lowest quartile. The fraction of beneficial observations in the other quartiles ranges from 47.6% to 51.3%. The finding of just over half of the observations having a positive impact is typical of operational DA systems (e.g., Gelaro et al. 2010). Lorenc and Marriott (2014) use a simple model to explore why the fraction of observations improving a forecast are just over 50% and find this is due to multiple factors, including observational errors, errors in the verifying forecasts, imperfect background error statistics, and growing modes.

Previous work has indicated that the observation impact, in addition to varying by observation value, may also vary by the observation time within the DA window. In the ECMWF 4D-Var system, McNally (2019) uses data-denial experiments to show that excluding the satellite observations taken later within the 12-h DA window has a larger detrimental impact on the analyses and forecasts than does excluding the observations taken earlier in the 12-h window. In fact, excluding observations taken in the last three hours of the 12-h window has a bigger detrimental impact than excluding observations taken in the first 6 hours of the window. In McNally (2019) this is attributed to observations taken later in the window providing the most up-to-date information on the atmosphere, as well as providing additional dynamical information via feature advection or wind tracing. That is, mismatches between observations and the background state at the start of the window can only be addressed through local changes, while mismatches between observations and the background state at the end of the window may also be addressed by remote changes that evolve to influence the local state.

We are interested in determining if the impact of the buoy surface pressure observations is likewise a function of the observation time within the DA window. Aggregating the observation impact for all drifters as both a function of observation quartile, and window within the DA system (Fig. 4) shows that for all observation values (orange bars) the biggest impact comes from observations take in the last hour (+3) of the DA window. The smallest impacts, however, come from observations taken near the center of the DA window (−1 and +1) and not at the beginning of the window. Separating the impacts by observation value quartile explains this pattern. For the observations in the lowest quartile, observations taken at any time during the DA window are beneficial on average, but larger benefits are obtained by observations taken in the second half of the DA window as opposed to the first half of the DA window, and the largest benefits come from observations taken in the last hour. The observations in the second quartile are near-neutral or beneficial throughout the window. The observations taken in the third and fourth quartiles are near neutral or slightly beneficial in hours −3 and −2 but become nonbeneficial for the hours from −1 through +3. When the impacts of all quartiles are summed together, this results in the smallest impact occurring during the center of the window.

To evaluate the statistical significance of the varying impacts across the time window, we take the difference between the +3-h observation impact and the −3-h observation impact for each analysis time, then evaluate if the distribution of these differences is significantly different from zero. We repeat the process such that we can evaluate the difference between the +3- h observation impact and the impact from every other time window. When considering impacts for all the observations (orange bars in Fig. 4), the +3-h observations are not significantly more beneficial than the observations at other hours in the DA window at the 95% level (Fig. S7 in the online supplemental material). However, the +3-h observations are more beneficial than the −1-h observations and the +2-h observations at the 90% level. In contrast, when considering only the observations in the lowest pressure quartile (blue bars in Fig. 4), the +3-h observations are significantly more beneficial than the observations taken at every other hour in the DA window at the 95% level (Fig. S8 in the online supplemental material).

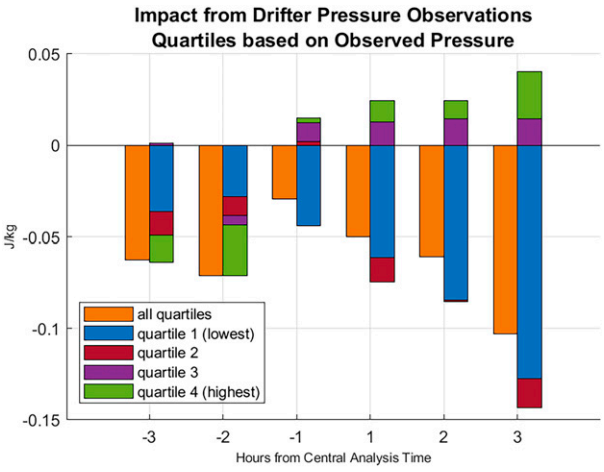

FIG. 4. Total impact from all northeastern Pacific drifter surface pressure observations (orange) as a function of hour within the 6-h DA cycle. Impact is also broken out by observed pressure quartile, with the lowest quartile in blue and the highest quartile in green.

The observation impact sensitivity to DA window time that emerges from this analysis is more nuanced than the simple explanation of later observations having a larger beneficial impact than earlier observations. The relatively small total impact that comes from the observations in the middle of the DA window, at hours −1 and +1, is most likely specific to the Navy system. At these hours in the middle of the DA cycle, much of the beneficial impact of the observations from the lower quartile is canceled out by the nonbeneficial impact of the observations in the third and fourth quartiles. McNally (2019) hypothesizes that model error might make it more difficult for a 4D-Var system to make good use of observations taken later in the update cycle. As denoted by the vertical lines in Fig. 2, the bias is near zero for the first quartile, but increases to almost 0.5 hPa for the highest pressure quartile. This bias may be responsible for the increasingly nonbeneficial impacts for the observations taken in the highest quartile as the time in the DA window increases. Mean innovations for the highest quartile (not shown) are the smallest during the first 2 h of the DA window and largest during the last two hours of the DA window. While we are currently using a

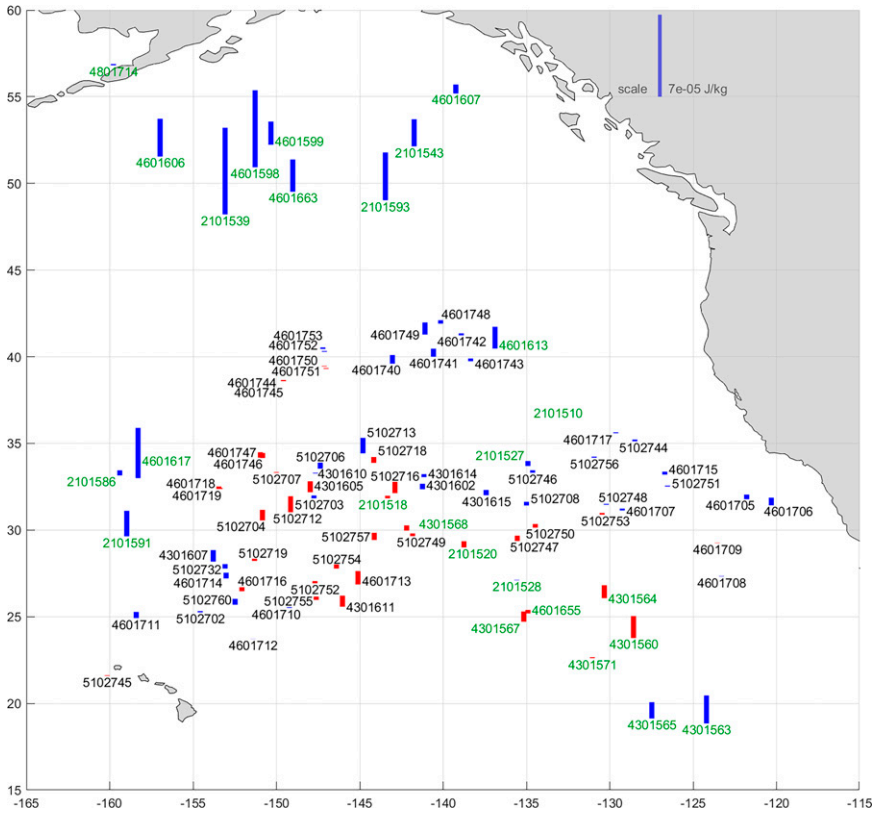

FIG. 5. Average observation impact for the northeastern Pacific drifter surface pressures (J kg^{-1}; the scale is provided at the top right). AR drifter identification numbers are in black text, and non-AR drifter numbers for drifters reporting surface pressure are in green text. Beneficial impacts are denoted by blue bars, and nonbeneficial impacts are denoted by red bars.

strong-constraint version of NAVDAS-AR, it is possible that using the weak constraint version that accounts for model error may mitigate this issue, and is an avenue for future research.

b. *Case studies*

As may be expected from earlier work (Centurioni et al. 2017b; Horanyi et al. 2017; Ingleby and Isaksen 2018), the impact of the drifter observations varies substantially by location (Fig. 5). Drifters that are in the northern part of the domain, near the Aleutian low in the Gulf of Alaska, have the largest beneficial impact when averaged over the entire AR Recon 2020 time period, followed by drifters in the far western and southern parts of the domain. Observations in the northern part of the domain occur in a region with relatively low pressure and large pressure variability as compared with the rest of the domain. The relatively large impact from observations in this region is consistent with the high impact in the low pressure quartile shown in Fig. 3, as well as prior results showing the largest beneficial FSOI impacts of the drifter surface pressure observations in the ECMWF IFS system in storm track regions (Ingleby and Isaksen 2018). These results are also consistent with an ensemble-sensitivity study of West Coast cyclones showing lower sensitivity for storms that make landfall south of 40°N than those that make landfall farther north (McMurdie and Ancell 2014). The drifters in the central part of the domain, under the time-averaged high pressure, have near neutral (small beneficial or small nonbeneficial) average impacts.

A visual inspection of Fig. 5 suggests higher impacts from relatively isolated drifters, or drifters situated on the edges of drifter clusters, than from drifters in the center of clusters, consistent with the idealized work of Baker (2000) and the results in Baker and Langland (2009). To quantify this isolation effect, we separate observation impact into quartiles based on the average distance of the drifter to its four nearest drifter neighbors (Fig. S9 in the online supplemental material). This is not a perfect measure of observation isolation, and does not take into account other types of observations. However, on average, the drifters that are relatively far from their nearest neighbors have larger impacts than drifters that are relatively close to their neighbors. In data assimilation, the spread of information from the observations to the surrounding model grid points is controlled by both the observation density and the structure of the observation and background error covariances. Specific factors related to the error covariances can include the forward/adjoint observation operators, the vertical and horizontal background error correlation length scales and cross-variable covariances, and correlated observation errors.

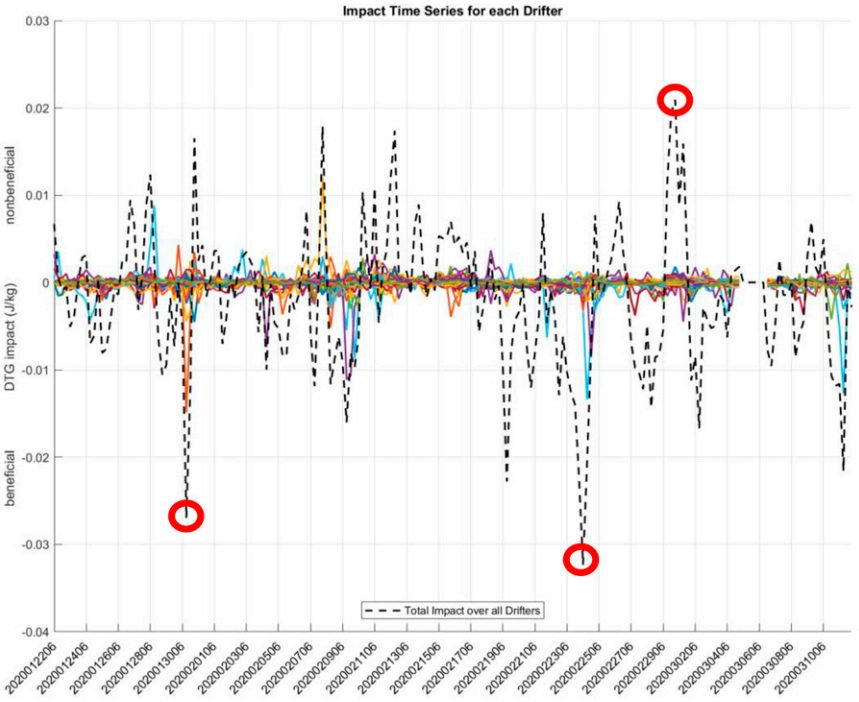

FIG. 6. Impact for all northeastern Pacific drifter surface pressure observations (dashed line) as a function of the analysis time. The impact from the individual drifters is denoted by individual solid curves. The dates with the two largest beneficial impacts (1200 UTC 30 Jan 2020 and 0600 UTC 24 Feb 2020) and the date with the largest nonbeneficial impact (0000 UTC 1 Mar 2020) are circled in red and are considered in detail in Fig. 7, below.

Behind the time-average impact there is large case-to-case variability. Both the impact of individual drifters, as well as aggregate drifter impact, vary considerably from day to day, as is apparent from plotting the drifter impact as a function of analysis date and time (Fig. 6). This is consistent with the large case-to-case variability in drifter impact shown for the North Atlantic in the ECMWF IFS system (Horanyi et al. 2017, their Fig. 12). For each analysis date–time group in Fig. 6, color lines representing individual drifters occur on both the positive and negative side of zero, indicating both nonbeneficial and beneficial impacts from different drifters at any one analysis time. This is true even for analysis times with relatively large cumulative impacts (denoted by the dashed line), which are usually not dominated by an individual drifter. The analysis times with the two largest beneficial impacts (1200 UTC 30 January 2020 and 0600 UTC 24 February 2020) and the analysis time with the largest nonbeneficial impact (0000 UTC 1 March 2020) all occurred close to AR Recon intensive observing periods (IOPs). These include IOP3 at 0000 UTC 30 January 2020, IOP 11 at 0000 UTC 24 February 2020, and IOP12 at 0000 UTC 2 March 2020.

Horanyi et al. (2017) find that the aggregate impact of the drifters over the North Atlantic is much larger than the time average impact (sometimes by a factor of 4) during periods of evolving complex or fast deepening cyclogenesis. We examine the impact of each individual drifter for a few cases (highlighted by red in Fig. 6) in order to pinpoint the specific features and regions that are associated with large drifter impact. The impacts of the drifter surface pressure observations for these individual cases (Fig. 7) exhibit impact patterns that are distinct from the time-average patterns. For the two large beneficial impacts (1200 UTC 30 January 2020 and 0600 UTC 24 February 2020, Figs. 7a–d), the biggest impacts come from drifters in the regions associated with large pressure gradients (e.g., frontal regions) and strong horizontal vapor transport in the atmospheric rivers. Large impacts also come from observations within or on the periphery of the parent low pressure center. Many of the drifters northeast of Hawaii that have large impacts for these two analyses times have near neutral impacts when averaged over the entire season (cf. Figs. 7 with 5). The impact of the observations located under the high pressure centers to the east of the fronts and ARs have much smaller impacts, some of which are beneficial and some of which are nonbeneficial. Inspection of other analysis times associated with large beneficial impacts (not shown) are consistent with the scenario of beneficial observations occurring in regions of strong pressure gradients. For example, the large beneficial impact at 1200 UTC 19 February is associated with observations just south of a strong low pressure center in the Gulf of Alaska, and the large beneficial impact at 1200 UTC 11 March is primarily associated with observations from two drifters on the northeastern edge of an intensifying low pressure center. In contrast, the large aggregate nonbeneficial impact for the 0000 UTC 1 March 2020 case (Figs. 7e,f) does not

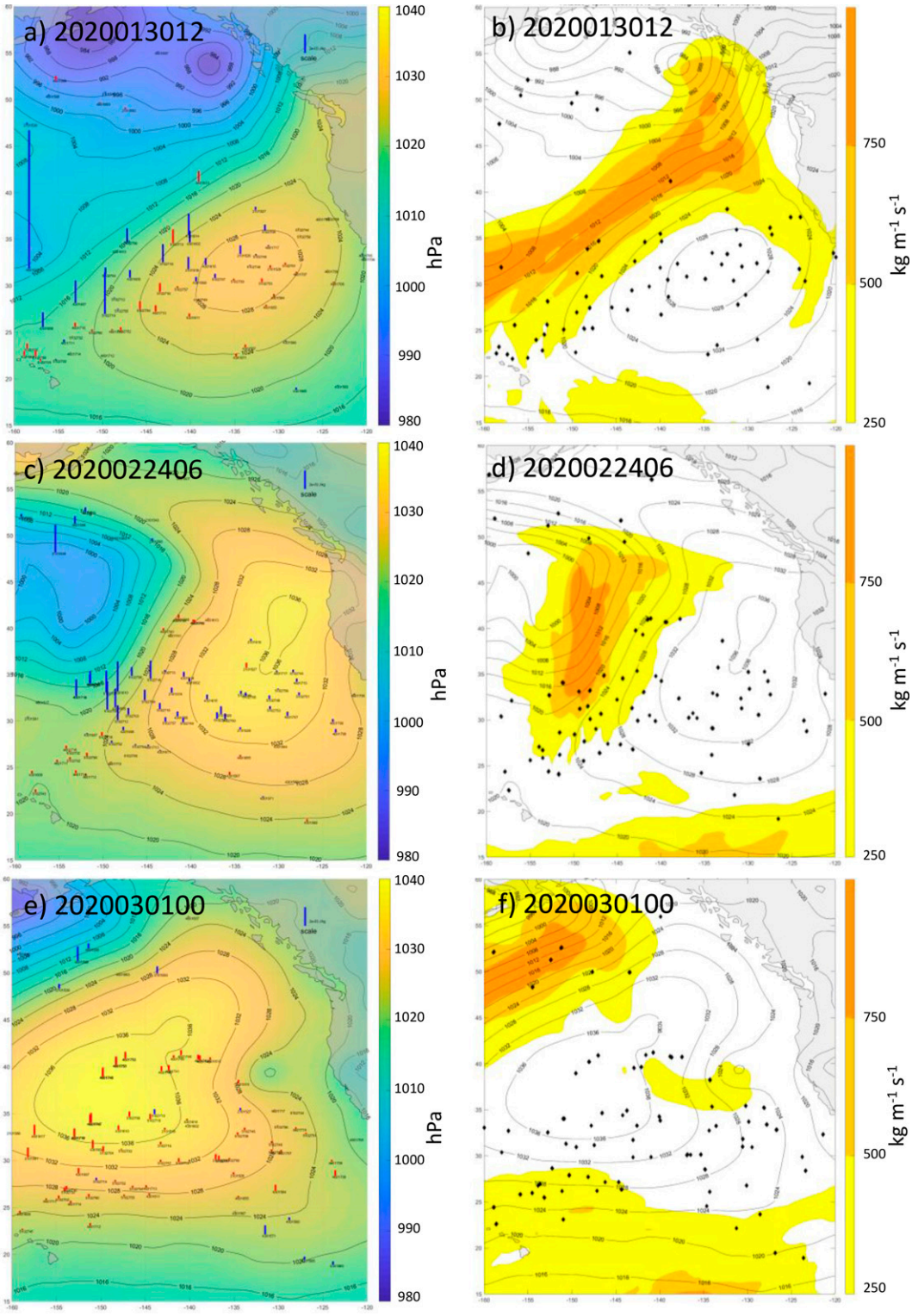

FIG. 7. Analyzed surface pressure (contours and shading; hPa) and observation impact of the drifter surface pressures [J kg^{-1}; the scale is provided at the top right of (a), (c), and (e)] valid at (a) 1200 UTC 30 Jan 2020, (c) 0600 UTC 24 Feb 2020, and (e) 0000 UTC 1 Mar 2020 summed over the 6-h DA window. Beneficial observations are denoted by blue bars, and nonbeneficial observations are denoted by red bars. Also shown is integrated vapor transport (IVT; kg m^{-1} s^{-1}; color shading) and analyzed surface pressure (contours; hPa) for (b) 1200 UTC 30 Jan 2020, (d) 0600 UTC 24 Feb 2020, and (f) 0000 UTC 1 Mar 2020. Drifter locations are noted by black dots in (b), (d), and (f).

come from a few drifters with large impacts, but rather is the result of the accumulation of many drifters with small nonbeneficial impacts located under the high pressure center that dominates the northeastern Pacific domain at this time. The few drifter observations that do have moderate beneficial impacts are in the far northwest corner of the domain, in the vicinity of an AR and in a region of relatively low pressures (consistent with the quartile analysis shown in Fig. 3).

The case study results are consistent with the results shown in section 3a where the largest beneficial impacts are coming primarily from observations taken in the lowest surface pressure quartile (below 1020 hPa), with near neutral impact coming from higher pressure observations. These results support the importance of observations in dynamically active baroclinic regions (Ingleby and Isaksen 2018; Horanyi et al. 2017). They are also consistent with prior studies showing large forecast sensitivity to changes in the initial state in baroclinic zones (Kleist and Morgan 2005; Doyle et al. 2019), shortwave troughs (Zheng et al. 2013), frontal regions (Doyle et al. 2014), atmospheric rivers (Reynolds et al. 2019) and potential vorticity features (Reynolds et al. 2001; Torn and Romine 2015). To quantify these relationships in a systematic way for the entire period, we have separated observation impact into quartiles based on the analyzed pressure gradient (Fig. S10 in the online supplemental material) and analyzed IVT (Fig. S11 in the online supplemental material) at the drifter observation location. These results confirm that larger beneficial impacts come more from the observations that occur in regions that have the strongest pressure gradients and largest IVT than from observations in the other quartiles. The impact for the quartile corresponding to the largest pressure gradients is -0.3424 J kg^{-1} with 53.9% beneficial observations. The impact for the quartile corresponding to the largest IVT values is -0.3329 J kg^{-1}, with 53.3% beneficial observations. These values are comparable to, but slightly smaller than, the average impact from observations in the lowest pressure observation quartile (-0.3820 J kg^{-1}, with 54.4% beneficial). Average impacts for the other quartiles are near neutral. Some of the similarity of the results for the different quartile definitions is probably due to overlap in the observation sets (e.g., regions of high IVT are also often regions of strong pressure gradients).

The nonbeneficial impacts of the observations under the high pressure center in the 0000 UTC 1 March 2020 case may be related to the strong-constraint 4DVar suboptimal use of observations when there is significant model bias. However, this is not a universal finding. Observation impacts under the high pressure centers for 1200 UTC 30 January 2020 and 0600 UTC 24 February 2020 are small, but some are beneficial.

c. *Data-denial results*

The analysis in the previous subsections was based on the control experiment in which both non-AR Recon and AR Recon drifter observations are assimilated, and did not distinguish between observations from non-AR Recon and AR Recon drifters. In this section, we compare results from the control experiment and the experiment in which the AR Recon drifter observations are denied so as to examine how the addition of the AR Recon drifter observations affects both the analyses and the forecasts. The non-AR Recon drifter observations are assimilated in the denial experiment (it is only the AR Recon drifter observations that are denied).

Figure 8a shows the location of the non-AR Recon drifters with drifter identification numbers in blue, and the location of the AR Recon drifters in green, at 0000 UTC 12 March 2020. Figure 8b shows the innovation (OmB) mean and standard deviation at each of the non-AR Recon drifter locations for the DA update cycles in which the AR Recon drifters are not assimilated (in orange) and are assimilated (in blue). The red circles in Fig. 8b indicate the non-AR Recon drifters for which there is a discernible lowering of the innovation magnitudes by the assimilation of the AR Recon drifters. The identification numbers of these drifters are also underlined in red in Fig. 8a. Not surprisingly, the non-AR Recon drifter locations (underlined in red) that are impacted by the assimilation of the AR Recon drifters are the ones that are in close proximity to the AR Recon drifters (denoted by green dots).

We examine non-AR Recon drifter number 210518 (located in the middle of the domain, at 143°W, 32°N) in more detail, as this drifter shows the biggest differences from the assimilation of the AR Recon drifters. The distributions for the control and denial update cycles shown in the left and right boxplots, respectively, in Figs. 9a,b indicate that the assimilation of the AR Recon drifter observations in the control case reduces the median of both the OmA and OmB values and tightens the distribution as compared with the denial case. It is expected that the match between the observations and the analysis will be better (differences will be smaller) than the match between the observations and the background, and this is seen in both the control and denial cases, but there is a bigger difference between OmB and OmA in the control case than in the denial case. The scatterplots of OmB and OmA versus observed pressure (Figs. 9c,d) indicate that the assimilation of the AR Recon drifters helps correct the low pressure bias of NAVGEM in the analyses and background, particularly at high pressure values. This is indicated by the blue (control) dots being closer to zero than the red (denial) dots, and the reduction in the slope of the linear fit line when going from red (denial) to blue (control).

Given that the assimilation of the AR Recon drifters better constrains analyses at the non-AR Recon drifter locations, we may expect that the assimilation of these drifters will have a positive impact on forecast skill. NAVGEM performance as measured by RMSE and Anomaly Correlation over North America and the Northern Hemisphere extratropics (Fig. 10) indicates that this is indeed the case. Statistically significant (at the 95% level) improvements are found for various metrics, including 500-hPa anomaly correlation over the Northern Hemisphere at 72 and 96 h, and over North America from 72 to 120 h, using ECMWF operational analyses as verification. These metrics are part of the scorecard used by Fleet Numerical Meteorology and Oceanography Center to decide on implementing model upgrades. There are also reductions in vector RMSE for 850-, 500-, and 200-hPa winds over the Northern Hemisphere at 72 and 96 h. The percent reductions are very small on average (less than 1%), with the largest percent reduction being 2.76% for the 200-hPa

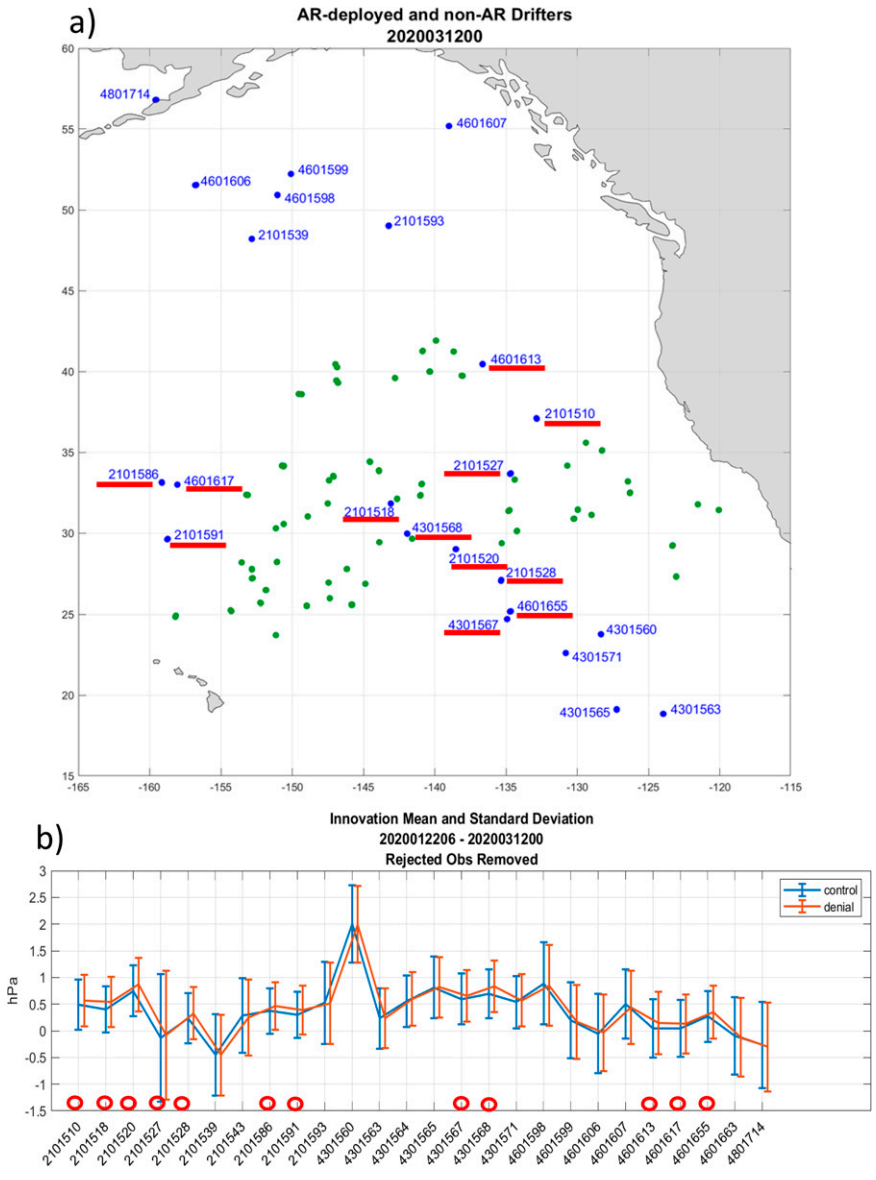

FIG. 8. (a) Location and number of AR Recon (green dots) and non-AR Recon (blue dots and numbers) drifters providing surface pressure observations. (b) Mean and standard deviation of the innovations at each non-AR Recon drifter (denoted by drifter number on the x axis) when the AR Recon drifters are assimilated (blue) and not assimilated (orange). Blue and orange bars are slightly offset for the sake of clarity. The non-AR Recon drifters that show the largest differences in mean innovation from assimilation of the AR Recon drifters are underlined in (a) and circled in (b) in red.

geopotential height RMSE over North America at 96 h. For most fields, the differences at analysis time do not meet the significance threshold. However, these initial differences sufficiently dampen error growth such that they result in a significant error reduction two to four days later. (The reason for the slight increase in Northern Hemisphere 200-hPa geopotential height RMSE at analyses time is unknown.) These results are consistent with the nonlocal beneficial impacts on forecast skill shown in the drifter data-denial experiments in the ECMWF IFS system (Horanyi et al. 2017; Ingleby and Isaksen 2018). However, unlike these previous studies, which withheld global or Northern Hemisphere drifter observations, in this study, the large-scale impacts are coming only from withholding a subset of the drifter observations in the northeastern Pacific. The impact through the depth of the troposphere comes from the ability of advanced (cross covariance) DA systems to extract information about the tropospheric circulation through the evolution of surface pressure observations (Compo et al. 2011; Slivinski et al. 2021).

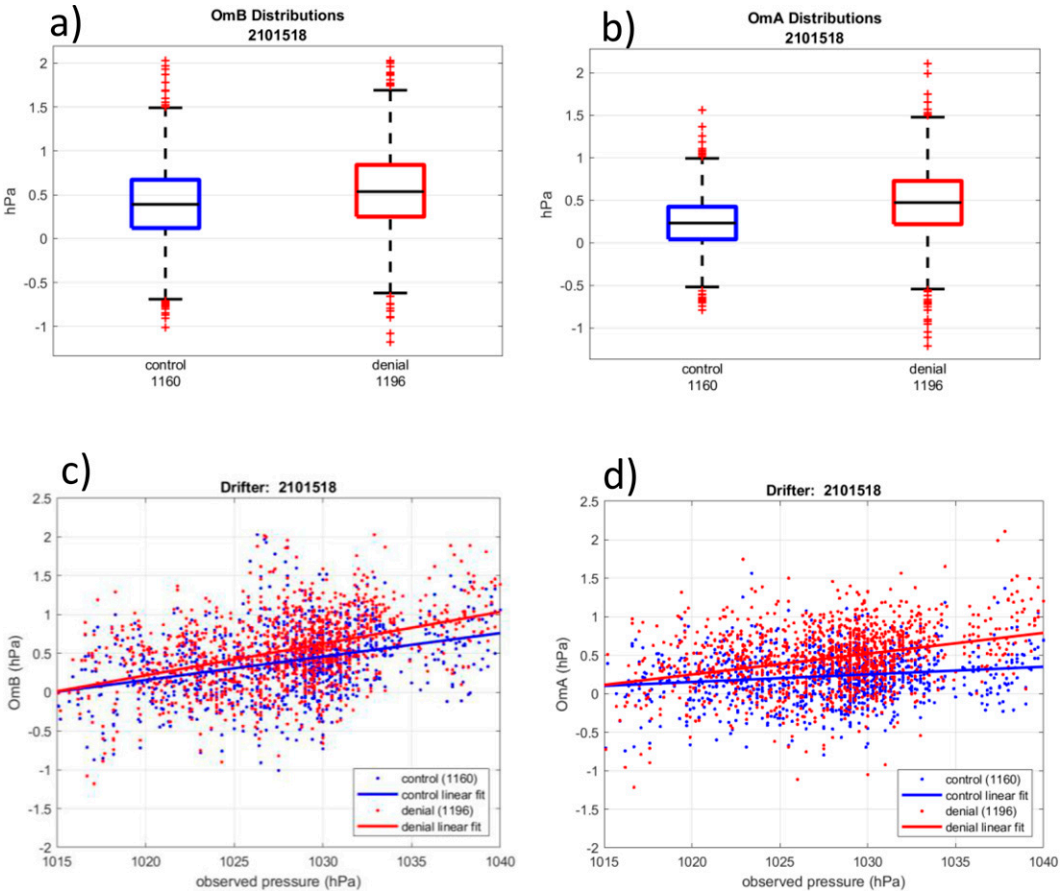

FIG. 9. Distributions of (a) observation minus background (OmB; hPa) and (b) observation minus analysis (OmA; hPa) for non-AR Recon drifter 210518 (32°N, 143°W) when AR Recon drifters are assimilated (control; left boxplots with blue boxes) and are not assimilated (denial; right boxplots with red boxes). Also shown are (c) OmB and (d) OmA as functions of observed pressure for non-AR Recon drifter 210518 when AR Recon drifters are assimilated (blue) and are not assimilated (red), along with linear fits (blue and red lines). In the boxplots, the center black line indicates the median and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to approximately 99.3% coverage if the data are normally distributed.

4. Summary and discussion

We evaluate the impact of drifting buoy surface pressure observations in the northeastern Pacific on analyses and forecasts in the Navy global atmospheric forecasting system during the AR Recon 2020 season using two methods: FSOI and data denial. The FSOI analysis considering both AR Recon and non-AR Recon drifters provides information on how the impact of the observations varies with timing, placement, and observation value. The drifters located in the north of the domain, near the time-averaged low pressure region in the Gulf of Alaska, have larger average impacts than those located near or in the time-averaged high pressure center between Hawaii and U.S. West Coast. The drifter observations in the lowest quartile of observed surface pressures have much larger (beneficial and nonbeneficial) forecast impact than observations at higher pressures, and account for almost all the net beneficial impacts of all the drifters. Drifter observations in locations of large pressure gradients and IVT also have more beneficial impacts than drifter observations in less dynamically active regions. This is consistent with the findings of Ingleby and Isaksen (2018), who find that FSOI drifter impact is spatially correlated with baroclinic instability, and Horanyi et al. (2017), who find that FSOI drifter impact is larger in dynamically active areas such as close to Alaska and near Greenland. Our study also illustrates how drifters that are more isolated, including those on the western and southern peripheries of the cluster of drifters in the central northeastern Pacific, tend to have larger impacts than the drifters in the center of the cluster, consistent with the idealized work of Baker (2000) and the findings in Baker and Langland (2009).

We find large case-to-case variability in the FSOI impact for the northeastern Pacific drifters. Horanyi et al. (2017) find similar variability for North Atlantic drifters and that the spatial aggregate of the drifter impact is largest during periods of complex cyclone evolution or rapid cyclogenesis. Our case studies looking at the impact of the individual drifters allow us to pinpoint the

Reference	Level	Metric	Variable	Level type	Region	0	24	48	72	96	120
ECMWF Analysis	200.0	RMS Error	Geopotential Height	pressure	North America	🙂	🙂	🙂	🙂	🙂	🙂
ECMWF Analysis	200.0	RMS Error	Geopotential Height	pressure	Northern Hemisphere	☹	🙂	🙂	🙂	🙂	🙂
ECMWF Analysis	200.0	Vector RMS Error	Wind	pressure	Northern Hemisphere	🙂	🙂	🙂	🙂	🙂	🙂
ECMWF Analysis	500.0	Anomaly Correlation	Geopotential Height	pressure	North America	🙂	🙂	🙂	🙂	🙂	🙂
ECMWF Analysis	500.0	Anomaly Correlation	Geopotential Height	pressure	Northern Hemisphere	🙂	🙂	🙂	🙂	🙂	🙂
ECMWF Analysis	500.0	Vector RMS Error	Wind	pressure	Northern Hemisphere	🙂	🙂	🙂	🙂	🙂	🙂
ECMWF Analysis	850.0	Vector RMS Error	Wind	pressure	Northern Hemisphere	🙂	🙂	🙂	🙂	🙂	🙂
ECMWF Analysis	1000.0	Anomaly Correlation	Geopotential Height	pressure	North America	🙂	🙂	🙂	🙂	🙂	🙂
ECMWF Analysis	1000.0	Anomaly Correlation	Geopotential Height	pressure	Northern Hemisphere	🙂	🙂	🙂	🙂	🙂	🙂

FIG. 10. Standard scorecard metrics for North America and Northern Hemisphere NAVGEM forecasts as a function of forecast hour as verified against ECMWF operational analyses for forecast start times of 0000 UTC 22 Jan 2020–0000 UTC 13 Mar 2020. Green colors indicate improvements in the metric with the assimilation of the AR-Recon drifter surface pressure observations that are statistically significant at the 95% level. Pink colors indicate degradations at the 95% level.

largest impacts to regions associated with tight pressure gradients and strong IVT (i.e., fronts and atmospheric rivers). These results highlight how drifters that do not have a large benefit in a time-averaged sense may become very beneficial in potentially high-impact situations. These findings are also consistent with adjoint and ensemble-based analyses showing that the forecasted evolution of cyclones is very sensitive to small changes in the initial state in troughs, frontal regions and regions of strong moisture transport (e.g., Kleist and Morgan 2005; Zheng et al. 2013; Doyle et al. 2014; Reynolds et al. 2019; Torn and Romine 2015). The results are also consistent with the adaptive observing strategy of the AR Recon field project (Ralph et al. 2020; Cobb et al. 2023). For a review of previous targeted observing efforts and their impacts, see Majumdar (2017).

When analyzing drifter observation impact as a function of time within the 6-h DA window, we find that the aggregate impact for all drifters is largest during the last hour within the window (+3 h). This is consistent with the satellite data-denial experiments of McNally (2019). This effect is attributed to observations later in the DA cycle providing the most up-to-date information, and because the evolution of covariances within the 4DVAR system as the DA window progresses allows for more nonlocal influences on local background–observation mismatches. However, the aggregate impact for all the drifters is smallest during the middle of the DA window (from −1 to +1 h). This is likely system dependent and not necessarily a general result. The reason for this becomes apparent when we break the impacts down by observation quartile. The observations in the lowest quartile have small beneficial impacts during the first half of the DA window, and then have increasingly larger beneficial impacts during the second half of the DA window. In contrast, the small beneficial impacts of the upper quartile observations early in the DA window become nonbeneficial as the window progresses. This may well be due to the strong-constraint version of NAVDAS-AR being unable to account for the conditional NAVGEM model low pressure bias at high pressures, as seen in the OmB and OmA statistics. In future work, it would be interesting to see if weak-constraint DA systems that account for model error can alleviate this issue of nonbeneficial observations in the upper quartiles later in the DA window.

The data-denial experiments of the AR Recon drifters shows that these drifters add value to both the analyses and forecasts. The assimilation of the AR Recon drifters results in a better-constrained analysis at nearby non AR-Recon locations, and helps correct for the NAVGEM low bias at high pressure. It also results in a small but statistically significant improvement to lower and middle-tropospheric winds and lower-, middle-, and upper-tropospheric geopotential height forecasts over the Northern Hemisphere at 72 and 96 h. Similar results are found in the ECMWF IFS system by Ingleby and Isaksen (2018) and Horanyi et al. (2017), but unlike those experiments, in which drifter observations are denied over the globe or Northern Hemisphere, here we are only denying a subset of the drifters in the northeastern Pacific. A potential avenue of future research is to decrease the error assigned to the drifter surface pressures. In the NAVGEM system this observation error (which includes both instrument error and representativeness error) is currently over 1.0 hPa, while, as noted in Centurioni (2018) and Centurioni et al. (2019), the drifter observation instrument errors are characterized as 0.3–0.4 hPa. In the future it would be worthwhile to assess the impact of allowing the DA system to draw more closely to the drifter observations, although this will not overcome difficulties in assimilating observations in regions where the model is biased. It would also be of interest to evaluate the drifter impact using AR-relevant metrics such as those used in Cordeira and Ralph (2021).

In the subsequent AR Recon seasons of 2021 and 2022, additional drifters have been deployed, expanding both the density and spatial coverage of the drifter network in the northeastern Pacific. The 2022 AR Recon deployment included drifters in the Gulf of Alaska (Lavers et al. 2022), motivated in part by the results of this analysis. Additional buoys providing surface pressure observations have also been deployed off southern Greenland under the Targeted Experiment to Reconcile Increased Freshwater with Increased Convection (Lavers et al. 2021). Our work indicates that more isolated drifters tend to have a bigger impact on average than drifters that are closer to other drifters, which is consistent with the need for additional drifters over larger expanses of the ocean. This work also indicates that it would be beneficial to have drifters spaced such that they can capture pressure gradient patterns associated with fronts and cyclones. Future work on determining optimal placement and spacing of the buoy network would serve as very useful guidance for the design of future observing

networks, both for general NWP and specific goals such as improving AR forecasts.

Acknowledgments. Authors Reynolds, Doyle, and Baker were supported by the Chief of Naval Research through the NRL Base Program, PE 0601153N (NRL 6.1 Atmospheric Rivers and NRL 6.1 Multi-Scale, Non Gaussian, and Non Linear: Data Assimilation for Clouds and High resolution Models Projects). Authors Wilson, Subramanian, and Ralph were supported by the California Department of Water Resources AR program (Award 4600013361) and the U.S. Army Engineer Research and Development Center (Awards 609 W912HZ-15-2-0019 and W912HZ-19-2-0023). Author Centurioni, the drifter equipment, and the air deployment equipment were supported by NOAA Grant NA20OAR 4320278/304593 "The Global Drifter Program." We thank the 53rd Weather Reconnaissance Squadron, U.S. Air Force Reserve Command, WC130-J flights for the drifter deployment. We thank three anonymous reviewers for their thorough and insightful reviews that helped us to improve the paper. We thank Drs. Patricia Pauley, Elizabeth Satterfield, and David Kuhl for information on the data assimilation system.

Data availability statement. The drifter surface pressure observations are available from the Global Telecommunication System. Archived drifter observations are also available from the International Comprehensive Ocean–Atmosphere Data Set (ICOADS; https://icoads.noaa.gov). The NAVGEM simulation data have not been released for public access. There are limited NAVGEM operational forecast fields available on the U.S. GODAE server (https://usgodae.org/cgi-bin/datalist.pl?summary=Go&dset=fnmoc_forcing_navgem).

REFERENCES

Baker, N. L., 2000: Observation adjoint sensitivity and the adaptive observation-targeting problem. Ph.D. dissertation, Naval Postgraduate School, 265 pp.

——, and R. H. Langland, 2009: Diagnostics for evaluating the impact of satellite observations. *Data Assimilation for Atmospheric, Oceanic and Hydrological Applications*, S. K. Park and L. Xu, Eds., Springer, 177–196.

Centurioni, L., 2018: Drifter technology and impacts for sea surface temperature, sea-level pressure, and ocean circulation studies. *Observing the Oceans in Real Time*, R. Venkatesan et al., Eds., Springer, 37–57.

——, L. Braasch, E. D. Lauro, P. Contestabile, F. D. Leo, R. Casotti, L. Franco, and D. Vicinanza, 2017a: A new strategic wave measurement station off Naples port main breakwater. *Coastal Eng. Proc.*, **1**, 36, https://doi.org/10.9753/icce.v35.waves.36.

——, A. Horányi, C. Cardinali, E. Charpentier, and R. Lumpkin, 2017b: A global ocean observing system for measuring sea level atmospheric pressure: Effects and impacts on numerical weather prediction. *Bull. Amer. Meteor. Soc.*, **98**, 231–238, https://doi.org/10.1175/BAMS-D-15-00080.1.

——, and Coauthors, 2019: Global in situ observations of essential climate and ocean variables at the air-sea interface. *Front. Mar. Sci.*, **6**, 419, https://doi.org/10.3389/fmars.2019.00419.

Cobb, A., L. Delle Monache, F. Cannon, and F. M. Ralph, 2021: Representation of dropsonde-observed atmospheric river conditions in reanalyses. *Geophys. Res. Lett.*, **48**, e2021GL0993357, https://doi.org/10.1029/2021GL093357.

——, and Coauthors, 2023: Atmospheric river reconnaissance 2021: A review. *Wea. Forecasting*, https://doi.org/10.1175/WAF-D-21-0164.1, in press.

Compo, G. P., and Coauthors, 2011: The twentieth century reanalysis project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28, https://doi.org/10.1002/qj.776.

Cordeira, J. M., and F. M. Ralph, 2021: A summary of GFS ensemble integrated water vapor transport forecasts and skill along the U.S. West Coast during water years 2017–20. *Wea. Forecasting*, **36**, 361–377, https://doi.org/10.1175/WAF-D-20-0121.1.

——, J. Stock, M. D. Dettinger, A. M. Young, J. F. Kalansky, and F. M Ralph, 2019: A 142-year climatology of Northern California landslides and atmospheric rivers. *Bull. Amer. Meteor. Soc.*, **100**, 1499–1509, https://doi.org/10.1175/BAMS-D-18-0158.1.

Corringham, T. W., F. M. Ralph, A. Gershunov, D. R. Cayan, and C. A. Talbot, 2019: Atmospheric rivers drive flood damages in the western United States. *Sci. Adv.*, **5**, eeax4631, https://doi.org/10.1126/sciadv.aax4631.

Dettinger, M. D., 2013: Atmospheric rivers as drought busters on the U.S. West Coast. *J. Hydrometeor.*, **14**, 1721–1732, https://doi.org/10.1175/JHM-D-13-02.1.

——, F. M. Ralph, T. Das, P. J. Neiman, and D. R. Cayan, 2011: Atmospheric rivers, floods, and the water resources of California. *Water*, **3**, 445–478, https://doi.org/10.3390/w3020445.

Doyle, J. D., C. Amerault, C. A. Reynolds, and P. A. Reinecke, 2014: Initial condition sensitivity and predictability of a severe extratropical cyclone using a moist adjoint. *Mon. Wea. Rev.*, **142**, 320–342, https://doi.org/10.1175/MWR-D-13-00201.1.

——, C. A. Reynolds, and C. Amerault, 2019: Adjoint sensitivity analysis of high-impact extratropical cyclones. *Mon. Wea. Rev.*, **147**, 4511–4532, https://doi.org/10.1175/MWR-D-19-0055.1.

Ehrendorfer, M., 2000: The total energy norm in a quasigeostrophic model. *J. Atmos. Sci.*, **57**, 3443–3451, https://doi.org/10.1175/1520-0469(2000)057<3443:NACTEN>2.0.CO;2.

Eyre, J. R., 2021: Observation impact metrics in NWP: A theoretical study. Part I: Optimal systems. *Quart. J. Roy. Meteor. Soc.*, **147**, 3180–3200, https://doi.org/10.1002/qj.4123.

—— and R. Reid, 2014: Cost-benefit studies for observing systems. Met Office Forecasting Research Tech. Rep. 593, 11 pp., https://www.semanticscholar.org/paper/Cost-benefit-studies-for-observing-systems-Eyre-Reid/156286e87d0dffd5a3bb833337822de66678a014.

Gelaro, R., R. H. Langland, S. Pellerin, and R. Todling, 2010: The THORPEX Observation Impact Intercomparison Experiment. *Mon. Wea. Rev.*, **138**, 4009–4025, https://doi.org/10.1175/2010MWR3393.1.

Hatchett, B. J., S. Burak, J. J. Rutz, N. S. Oakley, E. H. Bair, and M.L. Kaplan , 2017: Avalanche fatalities during atmospheric river events in the western United States. *J. Hydrometeor.*, **18**, 1359–1374, https://doi.org/10.1175/JHM-D-16-0219.1.

Hogan, T. F., and Coauthors, 2014: The Navy Global Environmental Model. *Oceanography*, **27**, 116–125, https://doi.org/10.5670/oceanog.2014.73.

Horányi, A., C. Cardinali, and L. Centurioni, 2017: The global numerical weather prediction impact of mean sea level pressure observations from drifting buoys. *Quart. J. Roy. Meteor. Soc.*, **143**, 974–985, https://doi.org/10.1002/qj.2981.

Ingleby, B., and L. Isaksen, 2018: Drifting buoy pressures: Impact on NWP. *Atmos. Sci. Lett.*, **19**, e822, https://doi.org/10.1002/asl.822.

Kleist, D. T., and M. C. Morgan, 2005: Application of adjoint-derived forecast sensitivities to the 24–25 January 200 U.S. East Coast snowstorm. *Mon. Wea. Rev.*, **133**, 3148–3175, https://doi.org/10.1175/MWR3023.1.

Kuhl, D. D., T. E. Rosmond, C. H. Bishop, J. McLay and N. L. Baker, 2013: Comparison of hybrid ensemble/4DVar and 4DVAR within the NAVDAS-AR data assimilation framework. *Mon. Wea. Rev.*, **141**, 2740–2758, https://doi.org/10.1175/MWR-D-12-00182.1.

Langland, R. H., and N. L. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation system. *Tellus*, **56A**, 189–201, https://doi.org/10.3402/tellusa.v56i3.14413.

Lavers, D., M. J. Rodwell, D. S. Richardson, F. M.Ralph, J. D. Doyle, C. A. Reynolds, V. Tallapragada, and F. Pappenberger, 2018: The gauging and modeling of rivers in the sky. *Geophys. Res. Lett.*, **45**, 7828–7834, https://doi.org/10.1029/2018GL079019.

——, and Coauthors, 2020a: Forecast errors and uncertainties in atmospheric rivers. *Wea. Forecasting*, **35**, 1447–1458, https://doi.org/10.1175/WAF-D-20-0049.1.

——, B. Ingleby, L. Centurioni, A. Wilson, R. Ralph, and A. Subramanian, 2020b: Drifting buoys deployed in the northeast Pacific. *ECMWF Newsletter*, No. 163, ECMWF, Reading, United Kingdom, 10, https://www.ecmwf.int/sites/default/files/elibrary/2020/19508-newsletter-no-163-spring-2020.pdf.

——, E. Frajka-Williams, L. Centurioni, A. Wilson, M. Ralph, and J. Sugier, 2021: Drifting buoys deployed off Greenland. *ECMWF Newsletter*, No. 169, ECMWF, Reading, United Kingdom, 4, https://www.ecmwf.int/sites/default/files/elibrary/2021/20225-newsletter-no-169-autumn-2021.pdf.

——, A. Wilson, M. Ralph, R. Torn, and F. Pappenberger, 2022: ECMWF participates in Atmospheric River Reconnaissance. *ECMWF Newsletter*, No. 171, ECMWF, Reading, United Kingdom, 8, https://www.ecmwf.int/sites/default/files/elibrary/2022/20361-newsletter-no-171-spring-2022.pdf.

Lorenc, A. C., and R. T. Marriott, 2014: Forecast sensitivity to observations in the Met Office global numerical weather prediction system. *Quart. J. Roy. Meteor. Soc.*, **140**, 209–224, https://doi.org/10.1002/qj.2122.

Majumdar, S. J., 2017: A review of targeted observations. *Bull. Amer. Meteor. Soc.*, **97**, 2287–2303, https://doi.org/10.1175/BAMS-D-14-00259.1.

McMurdie, L. A., and B. Ancell, 2014: Predictability characteristics of landfalling cyclones along the North American west coast. *Mon. Wea. Rev.*, **142**, 301–319, https://doi.org/10.1175/MWR-D-13-00141.1.

McNally, A. P., 2019: On the sensitivity of a 4D-Var analysis system to satellite observations located at different times within the assimilation window. *Quart. J. Roy. Meteor. Soc.*, **145**, 2806–2816, https://doi.org/10.1002/qj.3596.

Nardi, K. M., E. A. Barnes, and F. M. Ralph, 2018: Assessment of numerical weather prediction model reforecasts of the occurrence, intensity, and location of atmospheric rivers along the west coast of North America. *Mon. Wea. Rev.*, **146**, 3343–3362, https://doi.org/10.1175/MWR-D-18-0060.1.

Pauley, P. M., and B. Ingleby, 2022: Assimilation of in-situ observations. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*, S. K. Park and L. Xu, Eds., Vol. IV, Springer, 293–371, https://doi.org/10.1007/978-3-030-77722-7_12.

Ralph, F. M., and M. D. Dettinger, 2012: Historical and national perspectives on extreme West Coast precipitation associated with atmospheric rivers during December 2010. *Bull. Amer. Meteor. Soc.*, **93**, 783–790, https://doi.org/10.1175/BAMS-D-11-00188.1.

——, and Coauthors, 2020: West Coast forecast challenges and development of atmospheric river reconnaissance. *Bull. Amer. Meteor. Soc.*, **101**, E1357–E1377, https://doi.org/10.1175/BAMS-D-19-0183.1.

Reynolds, C. A., J. D. Doyle, F. M. Ralph, and R. Demirdjian, 2019: Adjoint sensitivity of North Pacific atmospheric river forecasts. *Mon. Wea. Rev.*, **147**, 1871–1897, https://doi.org/10.1175/MWR-D-18-0347.1.

——, R. Gelaro, and J. D. Doyle, 2001: Relationship between singular vector and transient features in the background flow. *Quart. J. Roy. Meteor. Soc.*, **127**, 1731–1760, https://doi.org/10.1002/qj.49712757514.

Rosmond, T., and L. Xu, 2006: Development of NAVDAS-AR: Non-linear formulation and outer loop tests. *Tellus*, **58A**, 45–58, https://doi.org/10.1111/j.1600-0870.2006.00148.x.

Slivinski, L. C., and Coauthors, 2021: An evaluation of the performance of the Twentieth Century Reanalysis version 3. *J. Climate*, **34**, 1417–1438, https://doi.org/10.1175/JCLI-D-20-0505.1.

Stone, R. E., C. A. Reynolds, J. D. Doyle, R. H. Langland, N. L. Baker, D. A. Lavers, and F. M. Ralph, 2020: Atmospheric river reconnaissance observation impact in the Navy Global Forecast System. *Mon. Wea. Rev.*, **148**, 763–782, https://doi.org/10.1175/MWR-D-19-0101.1.

Tavolato, C., and L. Isaksen, 2015: On the use of a Huber norm for observation quality control in the ECMWF 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **141**, 1514–1527, https://doi.org/10.1002/qj.2440.

Torn, R. D., and G. S. Romine, 2015: Sensitivity of central Oklahoma convection forecasts to upstream potential vorticity anomalies during two strongly forced cases during MPEX. *Mon. Wea. Rev.*, **143**, 4064–4087, https://doi.org/10.1175/MWR-D-15-0085.1.

Waliser, D., and B. Guan, 2017: Extreme winds and precipitation during landfall of atmospheric rivers. *Nat. Geosci.*, **10**, 179–183, https://doi.org/10.1038/ngeo2894.

Xu, L., T. Rosmond, and R. Daley, 2005: Development of NAVDAS-AR: Formulation and initial tests of the linear problem. *Tellus*, **57A**, 546–559, https://doi.org/10.3402/tellusa.v57i4.14710.

Zheng, M., E. K. M. Chang, and B. Colle, 2013: Ensemble sensitivity tools for assessing extratropical cyclone intensity and track predictability. *Wea. Forecasting*, **28**, 1133–1156, https://doi.org/10.1175/WAF-D-12-00132.1.

——, and Coauthors, 2021a: Improved forecast skill through the assimilation of dropsonde observations from the atmospheric river reconnaissance program. *J. Geophys. Res. Atmos.*, **126**, e2021JD034967, https://doi.org/10.1029/2021JD034967.

——, and Coauthors, 2021b: Data gaps within atmospheric rivers over the northeastern Pacific. *Bull. Amer. Meteor. Soc.*, **102**, E492–E524, https://doi.org/10.1175/BAMS-D-19-0287.1.

Zhu, Y., and R. E. Newell, 1998: A proposed algorithm for moisture fluxes from atmospheric rivers. *Mon. Wea. Rev.*, **126**, 725–735, https://doi.org/10.1175/1520-0493(1998)126<0725:APAFMF>2.0.CO;2.