Atmospheric River Reconnaissance Observation Impact in the Navy Global Forecast System

REBECCA E. STONE
SAIC, U.S. Naval Research Laboratory, Monterey, California

CAROLYN A. REYNOLDS, JAMES D. DOYLE, ROLF H. LANGLAND, AND NANCY L. BAKER
Marine Meteorology Division, Naval Research Laboratory, Monterey, California

DAVID A. LAVERS
European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom

F. MARTIN RALPH
Center for Western Weather and Water Extremes, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California

(Manuscript received 5 April 2019, in final form 12 November 2019)

ABSTRACT

Atmospheric rivers, often associated with impactful weather along the west coast of North America, can be a challenge to forecast even on short time scales. This is attributed, at least in part, to the scarcity of eastern Pacific in situ observations. We examine the impact of assimilating dropsonde observations collected during the Atmospheric River (AR) Reconnaissance 2018 field program on the Navy Global Environmental Model (NAVGEM) analyses and forecasts. We compare NAVGEM’s representation of the ARs to the observations, and examine whether the observation–background difference statistics are similar to the observation error variance specified in the data assimilation system. Forecast sensitivity observation impact is determined for each dropsonde variable, and compared to the impacts of the North American radiosonde network. We find that the reconnaissance soundings have significant beneficial impact, with per observation impact more than double that of the North American radiosonde network. Temperature and wind observations have larger total and per observation impact than moisture observations. In our experiment, the 24-h global forecast error reduction from the reconnaissance soundings can be comparable to the reduction from the North American radiosonde network for the field program dates that include at least two flights.

1. Introduction

Atmospheric rivers (ARs) are relatively narrow and long corridors of horizontal moisture transport that account for more than 90% of the water vapor transport outside of the tropics (Zhu and Newell 1998; Ralph et al. 2005, 2016). They are particularly impactful on the U.S. West Coast, where they account for a large fraction of annual precipitation, and bring beneficial impacts such as drought mitigation (e.g., Dettinger 2013). However, ARs are also often associated with extreme flooding and high wind events (e.g., Ralph et al. 2006, 2016; Dettinger et al. 2011; Lavers et al. 2011; Neiman et al. 2011; Waliser and Guan 2017). As such, they are of great interest from scientific, decision-making, and societal impact perspectives. While weather forecasts continue to gain in accuracy due to improved data assimilation and forecast models and methods, there can still be substantial forecast errors in short-term forecasts. For example, DeFlorio et al. (2019) find that the percentage of ensemble members that correctly capture AR landfall within a 250-km distance drops below 75% at 48-h lead times in the European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble system. Given these large impacts and remaining forecast challenges,
several field campaigns have been devoted to observing ARs and improving AR forecasts, including the California Land-falling Jets Experiment (CALJET), the Pacific Land-falling Jets Experiment (PACJET), and the California Water Service (Calwater) field campaigns (Ralph et al. 2005, 2016). The most recent of these field programs is an AR Reconnaissance (AR RECON) program that took place in early 2018.

The 2018 AR Reconnaissance program (CW3E 2018) was a multiagency effort led by the Center for Western Weather and Water Extremes of the Scripps Institution of Oceanography at the University of California San Diego, with the goal of deploying dropsondes from aircraft to observe ARs and related sensitive regions in order to improve forecasts of landfalling AR and their attendant impacts on the west coast of North America. Six intensive operating periods (IOPs) took place between 27 January and 28 February 2018. Three of these IOPs saw the deployment of three aircraft (two Air Force C130s and the NOAA GIV); one IOP had the two Air Force C130s, and two IOPs had one Air Force C130 deployed. See Table 1 for a summary of aircraft used on each mission date. Flight tracks were planned using several sources of information, including forecasts from operational and research centers, and adjoint sensitivity from the Navy Coupled Ocean–Atmosphere Mesoscale Prediction System (COAMPS) forecast and adjoint system (Hodur 1997; Amerault et al. 2008). Previous work using the COAMPS adjoint system has shown strong sensitivity of short-term forecasts of winds and precipitation to the detailed structures of the moisture in ARs (Doyle et al. 2014; Reynolds et al. 2019; Doyle et al. 2019).

The goal of the field project was to increase the accuracy of analyses in these sensitive regions through the assimilation of the dropsonde observations, thereby decreasing the error of the ARs and their impacts (flooding, high winds) in short-term (24–48 h) forecasts. The Center for Western Weather and Water Extremes website (CW3E 2018) describes the field program in greater detail.

These AR field program dropsonde data (NOAA/AOC and USAF 53 WRS 2018) may be used to both improve short-term forecasts of west coast landfalling ARs, and for diagnostics and verification of the data assimilation and forecasting system. Lavers et al. (2018) have used the 2018 AR RECON program dropsondes to evaluate how well the ECMWF Integrated Forecast System (IFS) represents AR structure and water vapor fluxes. They found that short-range forecast errors of water vapor fluxes are about 22% as large as the mean observed flux values, and primarily related to uncertainties in the winds and humidity near the top of the boundary layer. Ongoing work at NRL that will be reported on in a separate publication is examining the impact of assimilating these AR dropsondes on various forecast metrics. In this study, we use the dropsondes to help us better understand the behavior of our DA and modeling system when assimilating these observations, and quantify the impact of these dropsondes relative to the standard radiosonde network (across North America). We find that the targeting of sensitive regions did provide high impact observations to the forecast system, and that the impact was beneficial. We also find that the impact of the humidity observations is smaller than the impact of the temperature and wind observations, and we identify approaches that might help to extract more utility from these observations.

### 2. Methodology

#### a. Modeling

The work described in this paper was done using the Navy global forecast suite, which consists of the Navy Global Environmental Model (NAVGEM; Hogan et al. 2014) and the four-dimensional hybrid variational-ensemble data assimilation system called the NRL Atmospheric Variational Data Assimilation System-Accelerated Representor (NAVDAS-AR; Xu et al. 2005; Rosmond and Xu 2006). NAVGEM is the Navy’s global weather prediction system, and it provides Navy and Marine Corps users with high-resolution 180-h forecasts every 6 h and twice-daily 16-day guidance using a 20-member global ensemble. In this study we run the forecast model at the same resolution as the deterministic operational forecast, T425L60, with a forecast resolution of approximately 34 km. The resolution of the analysis increments is approximately 100 km, due to the lower resolutions of the adjoint and tangent-linear models (T119). Thus the analysis increments are capable of capturing the major features of the atmospheric rivers, but may miss the stronger gradients or small-scale variations in the ARs. The NAVDAS-AR system has the capability to perform either weak or strong constraint variational assimilation; computations are formulated in

<table>
<thead>
<tr>
<th>Mission date (format is yyyyymmddhh)</th>
<th>Intensive operating period (IOP)</th>
<th>C-130 out of Travis AFB</th>
<th>C-130 out of Hickam AFB</th>
<th>G-IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018012700</td>
<td>1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2018012900</td>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2018020100</td>
<td>3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2018020300</td>
<td>4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2018022600</td>
<td>5</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2018022800</td>
<td>6</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
TABLE 2. Listing of NAVGEM routinely assimilated observation types during the 2018–19 AR RECON experimental runs. Additional information on these sensors and platforms can be found at the World Meteorological Organization Observing Systems Capability Analysis and Review site (WMO OSCAR 2019).

<table>
<thead>
<tr>
<th>Type of data</th>
<th>Sensor/nomenclature</th>
<th>Notes</th>
<th>Platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional observations</td>
<td>Raob</td>
<td>Ascents and descents if transmitted</td>
<td>Balloon launches</td>
</tr>
<tr>
<td></td>
<td>PIBAL</td>
<td>Winds only</td>
<td>Balloon launches</td>
</tr>
<tr>
<td></td>
<td>RECO</td>
<td></td>
<td>Dropsondes</td>
</tr>
<tr>
<td></td>
<td>Surface observations</td>
<td></td>
<td>Ships, buoys, and land stations</td>
</tr>
<tr>
<td>Aircraft observations</td>
<td>AMDAR</td>
<td>Aircraft Communications Addressing and Reporting System observations during level flight, ascents, descents</td>
<td>Commercial aircraft</td>
</tr>
<tr>
<td></td>
<td>AIREP</td>
<td>Air reports flight level observations</td>
<td>Transoceanic aircraft routes</td>
</tr>
<tr>
<td></td>
<td>ADS</td>
<td>Automatic dependent surveillance flight level observations</td>
<td>Transoceanic aircraft routes</td>
</tr>
<tr>
<td>Feature tracked winds, geostationary</td>
<td>SEVIRI</td>
<td>Spinning Enhanced Visible InfraRed Imager multiple channels</td>
<td>Meteosat-8, Meteosat-11</td>
</tr>
<tr>
<td></td>
<td>AHI</td>
<td>Advanced Himawari Imager multiple channels</td>
<td>Himawari-8</td>
</tr>
<tr>
<td></td>
<td>ABI</td>
<td>Advanced Baseline Imager multiple channels</td>
<td>Geostationary Operational Environmental Satellites (GOES)-East, GOES-West</td>
</tr>
<tr>
<td>Feature tracked winds, polar orbiters</td>
<td>AVHRR</td>
<td>Advanced Very High Resolution Radiometer</td>
<td>NOAA-18, NOAA-19, METOP-A, METOP-B SNPP, NOAA-20</td>
</tr>
<tr>
<td></td>
<td>VIIRS</td>
<td>Visible/Infrared Imager Radiometer Suite</td>
<td>METOP series pairs and triplets</td>
</tr>
<tr>
<td></td>
<td>Global AVHRR</td>
<td>Combined views of METOP series</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MODIS</td>
<td>Moderate Resolution Imaging Spectroradiometer</td>
<td>Aqua, Terra</td>
</tr>
<tr>
<td></td>
<td>LEOGEO</td>
<td>Uses composite views of low Earth orbit (LEO) polar orbiters and geostationary (GEO) satellites</td>
<td></td>
</tr>
<tr>
<td>Occultation soundings</td>
<td>GRAS</td>
<td>Global Navigation Satellite System (GNSS) Receiver for Atmospheric Sounding</td>
<td>METOP-A, METOP-B, METOP-C</td>
</tr>
<tr>
<td></td>
<td>GPS-RO</td>
<td>Integrated GPS Occultation Receiver</td>
<td>TerraSAR-X, TanDEM-X</td>
</tr>
<tr>
<td>Ozone profiles</td>
<td>SBUV</td>
<td>Solar Backscatter Ultraviolet spectrometer</td>
<td>NOAA-18, NOAA-19</td>
</tr>
<tr>
<td></td>
<td>OMPS</td>
<td>Ozone Mapping and Profiler Suite</td>
<td>NOAA-20, SNPP</td>
</tr>
<tr>
<td>Surface wind speeds</td>
<td>SSM/I</td>
<td>Special Sensor Microwave Imager</td>
<td>F15</td>
</tr>
<tr>
<td></td>
<td>SSMIS</td>
<td>Special Sensor Microwave Imager/Sounder</td>
<td>F16, F17, F18</td>
</tr>
<tr>
<td>Surface winds</td>
<td>ASCAT</td>
<td>Advanced Scatterometer</td>
<td>METOP-A, METOP-B, SCATSAT-1</td>
</tr>
<tr>
<td></td>
<td>OSCAT</td>
<td>OceanSat Scatterometer</td>
<td></td>
</tr>
<tr>
<td>Radiances</td>
<td>IASI</td>
<td>Infrared Atmospheric Sounding Interferometer</td>
<td>METOP-A, METOP-B</td>
</tr>
<tr>
<td></td>
<td>AIRS</td>
<td>Atmospheric Infrared Sounder</td>
<td>Aqua</td>
</tr>
<tr>
<td></td>
<td>MHS</td>
<td>Microwave Humidity Sounder</td>
<td>NOAA-19, METOP-A, METOP-B</td>
</tr>
<tr>
<td></td>
<td>MHS RARS</td>
<td>Regional ATOVS Retransmission Services of Microwave Humidity Sounder</td>
<td>NOAA-19, METOP-A, METOP-B</td>
</tr>
<tr>
<td></td>
<td>ATMS</td>
<td>Advanced Technology Microwave Sounder</td>
<td>Suomi National Polar-Orbiting Partnership (SNPP)</td>
</tr>
<tr>
<td></td>
<td>CrIS</td>
<td>Cross-track Infrared Sounder</td>
<td>Suomi National Polar-Orbiting Partnership (SNPP)</td>
</tr>
<tr>
<td>Synthetic observations</td>
<td>TC synthetics</td>
<td>Observations generated around a tropical cyclone, based on official TC warning and the Rankine vortex</td>
<td></td>
</tr>
</tbody>
</table>
FIG. 1. The dropsonde locations and the background (i.e., 6-h forecast field) column total precipitable water \((10^{-3} \text{ m})\), and 850-hPa heights (m) for each IOP: (a) 2018012700, (b) 2018012900, (c) 2018020100, (d) 2018020300, (e) 2018022600, and (f) 2018022800.
observation space, which generally has a much smaller number of dimensions than the corresponding model state space. Hybrid 4DVar data assimilation (Kuhl et al. 2013) is implemented within the NAVDAS-AR data assimilation system, as well as in the operational ensemble forecasting system (McLay et al. 2008, 2010). NAVDAS-AR processes over 100 million observations in every 6-h data assimilation window, but after quality control and data thinning approximately 3.4 million observations are assimilated to create the final analysis. The observations types routinely assimilated are listed in Table 2.

In this study, we make use of an important capability of NAVDAS-AR, the forecast sensitivity observation impact (FSOI) computation. FSOI is a mathematical method to quantify the contribution of individual observations or sets of observations to a reduction in the 24-h forecast error (Langland and Baker 2004). The measure of error used in our FSOI is a moist energy norm (Ehrendorfer 2000), which is sensitive to both kinetic and moist static energy. Factors such as incorrect moisture placement, errors in position and intensity of midlatitude cyclones, high pressure centers, and jet streams all affect the norm; it provides a measure of the ability of the data assimilation system to use observations to generate an improved forecast. FSOI is usually quantified per observation subset (e.g., separate channels, or separate satellites). In this study we will separate the impact of the AR RECON observations for examination.

b. Flight tracks and observations

The dropsonde locations for each of the six IOPs are shown in Fig. 1. The background field shown with color shading is the column total water vapor in units of millimeters of water equivalent; background field 850-hPa heights are contoured. In each case the sampled region included the central core of approaching moisture, and off-axis regions when sufficient aircraft were available. As stated above, flight track planning incorporated forecasts and COAMPS adjoint sensitivity estimates, as well as logistical considerations. Table 1 summarizes the aircraft used for each IOP: IOPs 1, 3, and 4 included flights from two Air Force C130s and the NOAA GIV; IOP2 included flights from the two Air Force C130s; IOP 5 and 6 had one Air Force C130 deployed. Missions flown with the NOAA GIV had dropsonde data starting at approximately 150 hPa, while missions flown with only the C130s had dropsonde data starting at approximately 300 hPa. Dropsondes were instrumented with the Vaisala RSS903 sensor module, which includes the RS92 Heated HUMICAP Humidity Sensor.

3. Innovations and effects on analysis

a. Innovations

The observations received from the dropsondes were observations of temperature, zonal wind component
Humidity observations are transformed into relative humidity values, using the background temperature rather than the observed temperature, to compute the saturation mixing ratio; the resultant quantity is called pseudo–relative humidity (PRH). PRH has advantages over specific humidity, log specific humidity, and relative humidity for atmospheric moisture analysis systems (Dee and da Silva 2003), and was implemented in NAVDAS-AR as the analysis variable for moisture. Observations are processed through QC and formed into innovations, or differences from the background (observation–background). Figure 2 shows the sign, magnitude, and spatial distribution of the temperature innovations for IOP 2018022800. The sign of the innovation is indicated by the color of the spheres: red means the observation was warmer than the background; blue means the observation was cooler than the background. The size of the sphere indicates the magnitude of the difference. The map below the spheres shows the 700-hPa heights and column-integrated water vapor, for reference. The pattern of innovation signs and magnitudes shown in Fig. 2 reveals that the background is cooler than the observations at low levels in the AR inflow region, is mainly warmer than the observations at midlevels, and is very close to observations aloft. Three-dimensional views such as in Fig. 2 can help us visualize the changes made in the representation of an AR, but we can also use the innovations to reveal biases in the analyses and short-term forecasts over the relative in situ
data void region of the east Pacific. Figure 3 shows profiles of innovations for the 2018012700 case and aggregated over the six IOPs, along with the mean profile computed in 50-hPa layers. Temperature, $u$ wind, and $v$ wind all have mean profiles close to zero, but PRH is predominantly left of the zero line; observed PRH–background PRH is negative on average. The model state shows a moist bias relative to the

Fig. 4. Wind speed difference (observation − background, m s$^{-1}$) characteristics aggregated over the six IOPs. (a) Wind speed differences and mean difference profile; (b) Wind speed differences and mean differences plotted vs observed wind speed.

Fig. 5. Locations of all dropsonde observations during the data window for 2018012700 are shown as a function of pressure and time within the data window. Rejected data are marked with squares, and the number rejected is tabulated in the legend.
observations above 850 hPa, and especially above 300 hPa (Fig. 3c).

Figure 4 shows wind speed bias characteristics of the model state for the AR RECON region for these events. The wind speed difference (observed − background) has a mean profile close to zero, except above 300 hPa (Fig. 4a), where there is slow bias of a few meters per second. When the differences are plotted as a function of observed wind speed, we see that the slow bias exists in the higher wind speeds (40 m s$^{-1}$ and higher), while there is a fast speed bias where observed winds are 10 m s$^{-1}$ and lower. The mismatch in scale between the point observations taken from the instrument and the model values which represent relatively large grid boxes can at least partially account for these trends, and is usually referred to as representativeness error. Although not unexpected, given the relatively coarse resolution of the global model compared to observations that include finescale features and fluctuations present in the actual winds, these bias characteristics do impact ARs in the model, because they limit the accuracy with which vapor transport is represented.

b. Treatment of data in the data assimilation system

Two key aspects of how data may affect a forecast system are its handling in quality control screening procedures and its weighting in the assimilation step. Both aspects are governed by the estimates of observation error and background error that are specified in the data assimilation system. The innovation error estimate is a function of the specified observation error variance and the specified background error variance. Specifically it is the square root of the sum of the specified observation error variance and the specified background error variance. The specified innovation error variance determines how much weight a particular observation is given in the analysis, and sets the threshold for rejection of an innovation. Rejections occur within the assimilation system when the innovation magnitude is greater than 3$\sigma$ where $\sigma$ is the specified innovation error standard deviation.

Each set of soundings was examined to see what data were rejected. Figure 5 shows the RECON dropsonde observations over the data window for the 2018012700 analysis, with rejected data marked by squares, and tallied in the legend. PRH observations were more often rejected than temperature or wind observations. The localization of the wind rejections suggests the rejections could be feature-related (i.e., the observations could be correctly reporting a localized extreme value that is rejected by the assimilation system because the innovation is too large). Figure 6 shows the geographic locations of the rejected wind observations in relation to the analyzed wind speed field. Although the rejections do occur in regions with relatively higher wind speeds, the locations assure us that the peak intensity of the flow is not being systematically diminished by the rejection of these observations. The rejection rates for the dropsonde observations were 1.2% for temperature, 3.8% for PRH, and 1.4% for winds. If 3 times the specified innovation error variance encompassed about 95% of the observations (as for a Gaussian error distribution), then 5% of the observations would be characterized as outliers and rejected. Our system identified fewer than 5% of the observations as outliers; this could mean that the specified observation error variance is larger than the actual
observation error variance, but is also likely due to non-Gaussian innovation distributions (since the background, with biases and other non-Gaussian errors is included in the innovation).

Figure 7 shows the dropsonde observations’ RMS error variance values used in NAVDAS-AR. Figure 8 shows IOP 2018012700 innovations and rejections for each variable as a function of pressure. Colored markers denote each innovation; black markers denote 3 times the assigned innovation error for each innovation. The innovations marked with squares are innovations that were rejected by the departure check (i.e., innovations whose magnitude is more than 3 times the assigned innovation error). The cluster of rejections in Fig. 8a on the left of the plot consists of innovations that would tend to move the model to a drier state. This tendency in PRH innovations is seen in five of the six IOPs; the model has a moist bias. Figure 8c shows wind innovation vector differences rather than \( u \) or \( v \) wind innovations. The departure check for \( u \) and \( v \) innovations uses the wind speed differences rather than the individual \( u \) and \( v \) innovations, and is limited by 3 times the wind speed \( \sigma \) rather than 3 times the individual wind components\(^2\) \( \sigma \). Comparison of the assigned innovation error (black markers) with the colored lines indicating 3 times the standard deviation of the innovations gives an indication of the appropriateness of the specified innovation error variance. For the dropsonde observations during this IOP, the specified temperature innovation error variance corresponds closely to the temperature innovation variance; the specified PRH innovation error variance is smaller than the PRH innovation variance; the specified wind innovation error variance is larger than the wind innovation variance. These relationships hold for the aggregate of the six IOPs as well.

c. Effects on analyses

For each of the IOPs, the observations were assimilated and produced changes in the representation of the AR and the surrounding area. Two cases are shown here to illustrate these changes. Figures 9 and 10 show contours of vapor transport for the background (black) and the analysis (green), and the differences in vapor transport between the background and the analysis (red and blue shading) for two IOPs. The differences between the background and analysis are due to the total set of assimilated observations, so are not limited to the

\[ \sigma \] for \( u \) and \( v \). Square markers indicate rejected innovations. (a) Pseudo–relative humidity innovations (%); (b) temperature innovations (K); (c) wind vector difference magnitudes (m s\(^{-1}\)).

FIG. 8. Profiles of innovations for IOP 2018012700. Colored points show each innovation. Black lines show the mean of the assimilated innovations, computed for 100-hPa vertical bins. The specified error variance associated with each innovation is plotted with a black marker. Colored lines are plotted where the 3\( \sigma \) departure check falls, computed for 100-hPa vertical bins. For PRH and temperature it is simply 3 times the standard deviation of the innovations; for wind vector difference the line is plotted at 3\( \sqrt{2} \) times the standard deviation of the assimilated innovations, because the 3\( \sigma \) check uses wind speed \( \sigma \) rather than the

observation error variance, but is also likely due to non-Gaussian innovation distributions (since the background, with biases and other non-Gaussian errors is included in the innovation).
area immediately surrounding the AR dropsondes. The pressure levels and meridional slice locations presented in Figs. 9 and 10 were chosen to depict the locations with the greatest changes in AR vapor transport. Vapor transport, as plotted here, is calculated as wind speed × specific humidity at each point in the $1/2^\circ$ output fields and on pressure levels every 50 hPa. For IOP 2018012700 (Fig. 9), there were two separate regions of the AR where the background and analysis differed, one in the inflow region near 35$^\circ$N, 155$^\circ$W at 950 hPa (Figs. 9a,b), and another along the southern limit of the AR near 37$^\circ$N, 140$^\circ$W at 700 hPa (Figs. 9c,d). The sense of the changes was to reduce the net transport near the inflow region, and to reduce the river's width as it flowed eastward. Notice that the main axis of the river is unchanged in the analysis; only the southern edges show a reduction in transport. For IOP 2018012900 (Fig. 10), the strength of the main core of the river was reduced in the analysis, and there were also changes in the details of the vertical structure. The northern edge of the river near 42$^\circ$N is analyzed to be more vertically homogeneous than the background (Fig. 10a). The main core and the broad extent of the river at 750 hPa are also reduced in the analysis (Fig. 10b). Some of the reduction in transport at 750 hPa is offset by increases at 950 hPa (Figs. 10c,d).

4. Observation impacts and effects on forecasts

a. Impact differences among variables

The three assimilated variable types from the dropsondes—PRH, temperature, and winds—have...
differing amounts of impact on the forecast error norm. Although \( u \) wind and \( v \) wind are assimilated separately, we will present their impacts together, since they are always and only assimilated as matching pairs. Figure 11a shows the impact for each observation from the dropsondes for IOP 2018012900, plotted as colored dots. The legend and Fig. 11b indicate the total for each variable over the analysis window; negative values indicate error reduction (i.e., beneficial impact). For each variable, there is scatter about the \( x \) axis (zero impact line). This is normal not only for dropsondes and radiosondes, but for all observing systems, since the background and the observations are both good quality estimates of the atmospheric state. However, totaled over the 6-h data window, each variable makes a net error reduction for this IOP. Winds have the largest beneficial impact; temperature has the next largest; and PRH has the smallest beneficial impact of the three variable types. Figure 12a shows the net impact from the AR dropsondes for each variable over the data window, for each of the six IOPs. Winds often give the greatest error reduction among the three variables, but there is quite a bit of variability in both the day-to-day relative contributions among the variables, and the day-to-day overall impact. PRH observations not only give the smallest error reduction overall, but in some cases their contribution is not beneficial. The last two IOPs had only one aircraft available, so the overall AR RECON impact on those dates was limited by the smaller number of observations. Although there is significant day-to-day variability in the error reduction percentages by variable (Fig. 12a), the percentages by variable for the aggregate of the six dates are very
similar to the percentages from the global and North America sets of radiosondes (Fig. 13).

**b. Impact differences by geographic region**

Another way to visualize the observations and their impacts is shown in Fig. 14. The AR RECON region (outlined with red dashed lines) contains many more observations with large impact than does North America, although these large impacts may be either positive or negative. Temperature impacts have similar behavior. In Fig. 14, individual observation impacts are relatively small for North America, but mainly because

![Fig. 11. Impacts for IOP 2018012900: (a) observation impact for each AR dropsonde observation within the assimilation window ($10^{-3}$ J kg$^{-1}$); (b) net observation impact (J kg$^{-1}$) over the assimilation window for each variable type.](http://journals.ametsoc.org/doi/pdf/10.1175/MWR-D-19-0101.1)

![Fig. 12. AR RECON dropsonde impacts from each variable type: (a) for each of the six IOPs (J kg$^{-1}$); (b) per observations for the six IOPs together ($10^{-3}$ J kg$^{-1}$).](http://journals.ametsoc.org/doi/pdf/10.1175/MWR-D-19-0101.1)
there are many in situ aircraft ascents and descents being assimilated. The individual AR RECON observation impacts are large partly because they are in the relative data void over the east Pacific and partly because of increased sensitivity associated with the AR. To examine more quantitatively the impact distributions for $u$ wind, $v$ wind, temperature, and PRH, semilog plots of impact versus frequency are shown in Fig. 15. Note that the AR PRH distribution differs significantly from the other observed parameters’ impact distributions. For PRH impacts (Fig. 15a), the blue line (AR dropsondes) sits inside the yellow line (global RAOBs); for temperature, wind, and total impacts (Figs. 15b,c,d), the blue line sits outside the yellow line. AR PRH observations have fewer large magnitude impacts than the global dataset, while AR wind and temperature observations have more large magnitude impacts than the global dataset. Possible explanations for this difference are discussed in section 4.

The total impact of AR dropsonde observations for each of the six IOPs is shown in Fig. 16. There is quite a bit of day-to-day variability in total impact, as well as the fraction of impact due to the AR dropsondes. Part of the variability is due to the differing amounts of AR dropsonde data available (e.g., the last two IOPs involved only one aircraft). For just the first four IOPs, which had at least two flights each, the average total observation impact is about equal to the total observation impact of the NA radiosonde network (exclusive of the dropsondes). The impact of the RECON flights during the last two IOPs with only one flight each were considerably smaller than the impact of the NA radiosonde network. However, part of the variability in total impact is due to the day-to-day differences in sensitivity to the larger radiosonde observing system; the NA radiosonde observation (raob) impact for IOP 2018022800 is more than 6 times that for IOP 2018012700.

The aircraft tracks were specifically chosen to sample regions of the atmosphere with large forecast sensitivities to initial conditions, so we might expect the AR dropsondes to account for a disproportionately large amount of the forecast error reduction. In Fig. 17 we see that the AR dropsonde impact (represented by the blue wedges), is smaller than the North America (NA) radiosonde network (represented by the red wedges), for all three of the observed variables, when averaged over all six IOPs. AR dropsonde PRH impact is $3/4$ that of the NA radiosonde network (Fig. 17a); impacts from AR dropsonde temperature and winds are smaller fractions of the NA radiosonde network totals, but neither is less than 45% of the NA radiosonde network (Figs. 17b,c). However, when considering only the first 4 IOPs, when at least 2 AR recon flights were flown, the impacts of the AR dropsondes are proportionally larger. The PRH dropsonde impact is more than 6 times that of the NA network (Fig. 17d); temperature dropsonde impact is twice that of the NA network (Fig. 17e); and dropsonde wind impact is comparable although less than that from NA radiosondes (Fig. 17f). Considering both the AR observation set and the North America radiosonde observation set, we also see that the share of the global total error reduction due to PRH (7%) is less than half the share due to temperature (19%) and winds (16%). However, we do note that the AR PRH reduction is only 25% smaller than the North America PRH reduction (3% versus 4%), while the AR temperature and wind

FIG. 13. Percentage of forecast error reduction due to observations of each variable type for the aggregate of the six IOPs: (a) AR RECON dropsondes; (b) North America radiosondes; (c) global radiosonde network.
reductions are about half as large as the corresponding North America temperature and wind reductions.

Considering PRH, temperature, and wind observations together, AR dropsonde observations account for 7% of the combined AR and North America radiosonde observations (by count), but they are responsible for 33% of the combined AR and North America forecast error reduction. The per observation forecast error reduction is shown in Fig. 18. There is quite a bit of variability among the IOPs of this study, but for the six IOPs together, the AR error reduction per observation is more than twice that for the global RAOBs.

Figure 19a shows the AR dropsondes versus North America radiosondes comparison of per observation forecast error reduction. The AR dropsondes (which are in a region with relatively sparse in situ data) consistently
have greater impact per observation than the North America radiosonde observations (where there are many in situ observations from ascending and descending aircraft). The fluctuations in per observation impact do not appear to be directly related to observation counts (cf. to Fig. 19b).

c. Forecast differences

An in-depth investigation of data-denial experiments is currently ongoing and will be reported on in a future paper. However, it is instructive to provide a first look at the impact of assimilating the dropsondes on the forecasts of precipitation. Figure 20 shows four examples of differences in forecasts of accumulated precipitation along the west coast of North America; these cases were selected to show a variety of forecast changes that were driven by the assimilation of the additional dropsonde data. For the AR RECON 2018 mission dates, the strongest differences were found in the Pacific Northwest rather than the California coast, due to the orientation of the river and flow pattern. In Fig. 20a, for IOP 2018012900, the experiment’s 3-h forecast of accumulated total precipitation is shifted southwestward relative to the control; the experiment has slowed the movement of a region of precipitation coincident with the atmospheric river. This could be related to the fast wind speed bias at low levels seen in Fig. 4b. The difference

![Fig. 15. Impact distributions (J kg⁻¹) for the aggregate of the six IOPs: (a) PRH; (b) temperature; (c) winds (u and v combined); (d) PRH, temperature, and winds combined.](image)

![Fig. 16. Comparison of AR RECON and North America raob error reduction total impacts (J kg⁻¹) for each of the six IOPs, and for the average of IOPs 1–4.](image)
patterns shown in Figs. 20a–d show that the dropsonde observations caused changes in locations of maximum precipitation, timing of forecast precipitation, and angle of approach for the rainbands.

5. Discussion

The observations that were gathered in the AR RECON flights were effective in sampling the AR and adjusting the model state to more accurately represent the specific details of the flow. The differences between the observations and the background model state (the innovations) were not atypical from other radiosondes in terms of mean bias and variance. Wind and temperature observations from the AR RECON dropsondes show consistent beneficial impact (Fig. 19). They have a greater proportion of large magnitude impacts than the global dataset, and this tendency is more prominent in the beneficial impacts than nonbeneficial (Fig. 15d). This confirms that observing ARs and related sensitive regions provides high impact beneficial observations to the forecast system.

Adjoint sensitivity studies (Doyle et al. 2014; Reynolds et al. 2019; Doyle et al. 2019) have found large sensitivity of midlatitude cyclones and ARs to the finescale structure in the lower-mid tropospheric moisture in and on the edges of the ARs. Thus it was somewhat unexpected that the relative humidity observations had smaller (sometimes even nonbeneficial) forecast error reduction impact than wind and temperature observations. The relatively small impact of moisture observations compared to wind and temperature observations is even more pronounced for the North America radiosonde network than for the AR dropsonde observations. One reason for low PRH impacts may be that our metric of FSOI is based on a global error norm; changes in this norm due to observations in one localized area may be overshadowed by observations in another area. The studies mentioned above, however, examined sensitivity to changes in the initial state, not the impact of observations, and in particular, identified sensitivity of forecasts of low level winds and precipitation in landfalling storms to changes in the initial state. Another difference between our work and the earlier studies is that in our current work, we include radiances from moisture-sensitive channels in the control observations, while the earlier studies did not assimilate radiances; it is
possible that this difference accounts for some of the difference in the importance of the dropsonde humidity observations. Model bias may also result in relatively small impacts from PRH observations. The bias shown by the negative mean PRH innovation in Fig. 3c indicates that each set of observations acts to dry the model state, but that model tendencies return the model to a too moist condition very quickly. The difference between the 24- and 30-h error norm may show no benefit from the humidity observations because by forecast hour 24 the model tendency may have overwhelmed that benefit. A third possible reason for limited impact from PRH observations is that, since the specified error variance for PRH is smaller than the actual innovation variance (Fig. 8b), useful PRH observations may have been rejected. On the other hand, underestimating the PRH innovation error may result in giving the PRH observations more weight in the analysis than is optimal. Finally, the relatively coarse resolution of NAVGEM will result in a mismatch between the finescale structures in the actual AR moisture field that are sampled by the dropsondes, and the relatively smooth AR structure in the model representation. This representativeness error would preclude the NAVGEM system from making optimal use of the observations. Further work testing both bias removal and specified observation error is needed to try to extract more forecast benefit from the humidity observations. In addition, tests will be performed with COAMPS to see if the finer resolution of the mesoscale system allows for a more effective use of the observations.

6. Summary and conclusions

In this study we examined observations and impacts for six dates where we had in situ dropsonde observations collected during the Atmospheric River (AR) Reconnaissance 2018 field program. We used the Navy Global Environmental Model (NAVGEM) analyses and forecasts, including the NAVDAS-AR capability of computing forecast sensitivity observation impact (FSOI).

We were able to use the dropsondes to characterize aspects of our DA and modeling system when assimilating these observations, and we compared the impact
of these dropsondes to impacts from the standard radiosonde network.

Assimilation of the additional in situ observations resulted in modifications to the detailed structure of each AR. While the details varied in each case, we also found, by examining the aggregate of the observations, that there were typical model errors that the observations were attempting to correct. In this way we used the innovations to reveal tendencies of the model state over the relative data-void region of the east Pacific. Wind innovations, both $u$ and $v$ wind, had mean profiles close to zero; little to no bias was found in these parameters. However, observed PRH–background PRH was negative on average throughout most of the atmospheric column; the model state had a moist bias even at forecast hour 6. The wind speed difference (observed − background) did exhibit some bias characteristics. There was a slow bias aloft, which was present in the higher wind speeds (40 m s$^{-1}$ and higher); a fast speed bias was present where observed winds were 10 m s$^{-1}$ and lower. These wind speed biases are not unexpected given the relatively coarse resolution of the NAVGEM system; however, they will have an impact on the integrated vapor transport and subsequent hydrological impacts at landfall.

We also compared the specified innovation error variances (specified within NAVDAS-AR) to variances of the actual innovations. The specified temperature innovation error variance corresponds closely to the temperature innovation variance; the specified PRH innovation error variance was smaller than the PRH innovation variance; the specified wind innovation error variance was larger than the wind innovation variance.

FIG. 20. Differences in forecast accumulated precipitation between runs with and without AR RECON dropsondes. Shading is the 3-h accumulated precipitation in the control forecast, and the red and blue contour lines depict locations where the forecast that assimilated the dropsondes had more (red) or less (blue) accumulation. (a) 2018012900 3-h forecast; (b) 2018020100 27-h forecast; (c) 2018020300 24-h forecast; (d) 2018022800 30-h forecast.
This suggests that some tuning of the specified error variances may result in improvement to the system. The forecast sensitivity observation impact for each dropsonde variable was compared to the impacts of the North American radiosonde network. We found that the AR RECON soundings had significant beneficial impact on the global moist total energy error with per observation impact more than double that of the North American radiosonde network or the global radiosonde network. The total reduction in the global forecast error due to the AR dropsondes was of comparable size to the reduction due to the North America radiosondes for the first 4 IOPs in which there were two or three aircraft. The targeting of sensitive regions did provide high impact observations to the forecast system, and the impact was beneficial on average although for two IOPs the impact from the AR dropsonde PRH observations was slightly negative. The AR dropsonde observations had a greater proportion of large magnitude impacts than the global dataset, and this tendency was more prominent in the beneficial impacts than non-beneficial. We found that the impact of the humidity observations was smaller than the impact of the temperature and wind observations, and approaches were identified that might help to extract more utility from humidity observations, such as adjusting the specified observation error.

ARs making landfall on the North American west coast present a challenging forecast problem partly due to the relative scarcity of in situ observational data over the Eastern Pacific. This work has shown that the additional in situ observations of ARs and related sensitive regions collected during the AR RECON 2018 field program proved beneficial to the NAVGEM forecast system in terms of global total energy 24-h forecast error. Future work will examine the impact of the observations on metrics that are more specific to landfalling ARs, such as precipitation along the North American west coast.

Acknowledgments. We gratefully acknowledge the support of the Chief of Naval Research through the NRL Base Program, PE 0601153N. Computational resources were provided by the Navy DSRC in Stennis, Mississippi. Rebecca E. Stone is grateful for the discussions with and helpful suggestions from Drs. Dan Hodys and Patricia Pauley.

REFERENCES


Ralph, F. M., P. J. Neiman, and R. Rotunno, 2005: Dropsonde observations in low-level jets over the northeastern Pacific


