Generated using the official AMS LaTeX template v6.1

Impacts of marine surface pressure observations from a spaceborne differential absorption radar investigated with an observing system simulation experiment

N.C. Privé[a,b] Matthew McLinden,[b] Bing Lin,[c] Isaac Moradi,[b,d] Meta Sienkiewicz,[b,e] G. M. Heymsfield, [b] Will McCarty[b]

[a] *Goddard Earth Sciences Technology and Research II, Morgan State University, Baltimore MD*

[b] *Goddard Space Flight Center, National Aeronautics and Space Administration, Greenbelt MD*

[c] *Langley Research Center, National Aeronautics and Space Administration, Hampton, VA*

[d] *Earth System Science Interdisciplinary Center, University of Maryland College Park, College Park MD*

[e] *Science Systems and Applications, Inc., Lanham MD*

Corresponding author: N.C. Privé, nikki.prive@nasa.gov

ABSTRACT: A new instrument has been proposed for measuring surface air pressure over the marine surface with a combined active/passive scanning multi-channel differential absorption radar (DAR) to provide an estimate of the total atmospheric column oxygen content. A demonstrator instrument, the Microwave Barometric Radar and Sounder (MBARS), has been funded by the National Aeronautics and Space Administration (NASA) for airborne test missions. Here, a proof-of-concept study to evaluate the potential impact of spaceborne surface pressure data on numerical weather prediction is performed using the Goddard Modeling and Assimilation Office global observing system simulation experiment (OSSE) framework. This OSSE framework employs the Goddard Earth Observing System model and the hybrid 4D ensemble variational Gridpoint Statistical Interpolation data assimilation system.

Multiple flight and scanning configurations of potential spaceborne orbits are examined. Swath width and observation spacing for the surface pressure data are varied to explore a range of sampling strategies. For wider swaths, the addition of surface pressures reduces the root mean square surface pressure analysis error by as much as 20% over some ocean regions. The forecast sensitivity observation impact tool estimates impacts on the Pacific Ocean basin boundary layer 24-hour forecast temperatures for spaceborne surface pressures on par with rawinsondes and aircraft, and greater impacts than the current network of ships and buoys. The largest forecast impacts are found in the southern hemisphere extratropics.

1. Introduction

The global observing network consists of millions of observations that are ingested into numerical weather prediction (NWP) models each day. Some fields, such as temperature, are relatively well sampled by remote instruments in conjunction with in-situ observations. Other fields, such as atmospheric winds, are much less well sampled and are largely indirectly observed through balance with temperature observations.

Surface pressure is a function of the column total mass of the atmosphere, and can have sharp gradients and subtle features. While tropical cyclones are a classic example of strong surface pressure gradients and associated extreme winds, other atmospheric phenomena such as baroclinic lows are highly impactful to daily life and weather forecasts in the midlatitudes. Although some continental regions feature well-developed networks of surface pressure observations, vast regions of the global oceans have only sparse in-situ surface pressure observations.

The primary in situ marine surface pressure observations are taken by drifting or moored buoys, ships, or ocean platforms such as oil rigs. These in-situ observations act to anchor the surface pressure field in areas that are mainly observed by remote sensing platforms. Moored buoys and platforms are generally located along the coastlines, with ship observations confined to shipping lanes. The global drifting buoy network fills in some gaps in the open ocean, however approximately half of the buoys do not report surface pressures, and funding for the drifting buoys is limited (Centurion et al. 2017).

Figure 1 shows an example of the distribution of in-situ marine surface pressure observations used in the National Aeronautics and Space Administration Global Modeling Office (NASA/GMAO) numerical weather prediction system on 1 September 2020. In this example, there are approximately 200 drifting buoy measurements, 2250 ship observations, and 23750 observations from moored platforms during the 6-hour window. With the exception of Antarctica, Greenland, and parts of Africa, the land areas are very well observed. While the coastal oceans near North America and Europe are densely observed, much of the open ocean is very poorly observed, in particular the central and eastern Pacific Ocean, where there may be thousands of kilometers between in situ observations. More than 70% of the global ocean is greater than 150 km from the nearest in-situ surface pressure observation in this example, with 25% of the ocean surface being more than 500 km from the nearest observation.

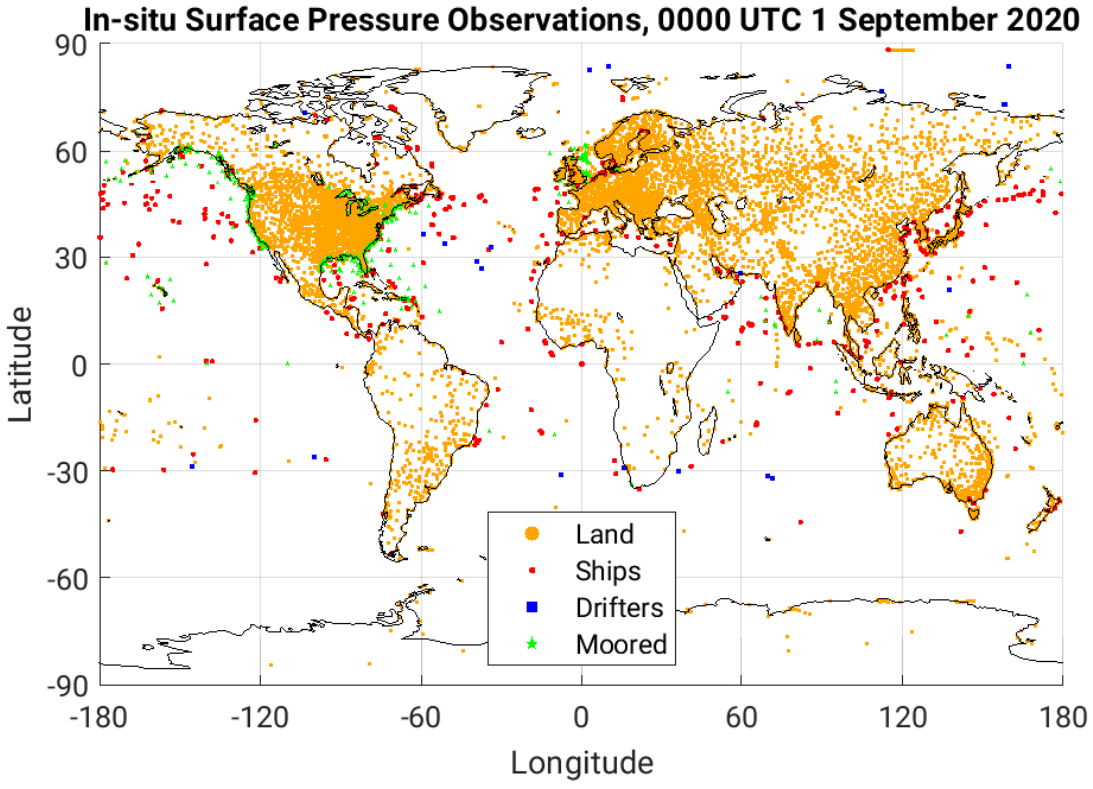


FIG. 1. Map of locations of surface pressure observations used in the semi-operational GMAO NWP system for 0000 UTC 1 September 2020. Orange dots indicate land surface measurements, red dots indicate ship measurements, blue squares indicate the location of drifting buoys, and green stars indicate moored platforms.

Several previous studies have examined the impact of current marine surface observations on numerical weather prediction. Radnóti et al. (2012) performed data denial experiments that showed significant forecast degradation over the Southern Hemisphere when all marine surface observations were excluded from the ECMWF operational system. They also studied a case of cyclogenesis in the North Atlantic, and found that drifting buoy data caused a greater than 50% reduction in short term forecast error in the vicinity of the cyclone. Horányi et al. (2017) looked at the impact of data denial of drifting buoys from the ECMWF system, and found that the inclusion of approximately 660 drifting buoys improved the total global forecast error for up to two to three forecast days, with impacts up to 250 hPa in height. Removal of the drifting buoys resulted in a increase in root mean square mean surface level pressure error of up to 30% in some regions. Ingleby and Isaksen (2018)

performed data denial experiments of surface pressures from drifting buoys, and found significant impact from the drifter surface pressures, particularly in the extratropics.

Remotely sensed surface pressure observations are not currently available, but several potential methods of producing surface pressure observations have been investigated. Global Navigation Satellite System radio occultation (GNSS-RO) has been proposed to retrieve surface pressure data, but this method has very coarse horizontal resolution and can also result in biased observations (Healy 2013). Scatterometry has also been proposed as a method of determining surface pressures through hydrostatic balance with surface winds. However, this method is limited to a modest range of wind speeds, and provides pressure gradient measurements rather than actual surface pressures.

a. Sea Surface Air Pressure Measurement and Instrument

Sea surface air pressure (or sea level pressure, SLP) is caused by the gravity force of combined atmospheric column masses of dry air and moisture. The column moisture mass is primarily determined by column water vapor (CWV) along with small contributions from cloud liquid and ice water paths, and precipitating hydrometeor amounts in no- or light-rain conditions. Over oceans, the atmospheric moisture content has been well observed from space for decades using both optical and microwave remote sensing techniques (e.g., Lin and Rossow (1994), Lin and Rossow (1997), Minnis et al. (2007), and references therein). Since random errors of satellite CWV retrievals are small (within 1.0 kg/m^2; Mears et al. (2018); or equivalently < 0.1 hPa when pressure is considered) and errors in other moisture content would also produce very small mass uncertainties for SLP retrievals, the key for accurate SLP remote sensing would be the measurement of column dry air mass.

The considered novel atmospheric pressure sensing system combines active and passive microwave sensors operating at the O_2 absorption V-band (64-70 GHz) to measure atmospheric column O_2 amounts in no- or light-rain conditions using a differential absorption technique (Lin and Hu (2005); Lawrence et al. (2011); Lin et al. (2023)). Note that heavy rain conditions are excluded from the SLP microwave remote sensing because precipitating hydrometeors such as large ice, snow, hailstones, graupel, and raindrops would produce significant scattering and absorption (or attenuation) to microwave signals (see Lin and Rossow (1997)) and make radar SLP retrieval unrealistic. Since oxygen is well mixed in the atmosphere, the measured O_2 amounts would be

proportional to the column dry air mass. Thus, surface air pressure could be retrieved from the dry air mass along with atmospheric moisture measurements. The sounder element of the system will be used to retrieve atmospheric temperature vertical profiles, and then to derive atmospheric pressure profiles based on the hypsometric equation constrained by the retrieved SLP.

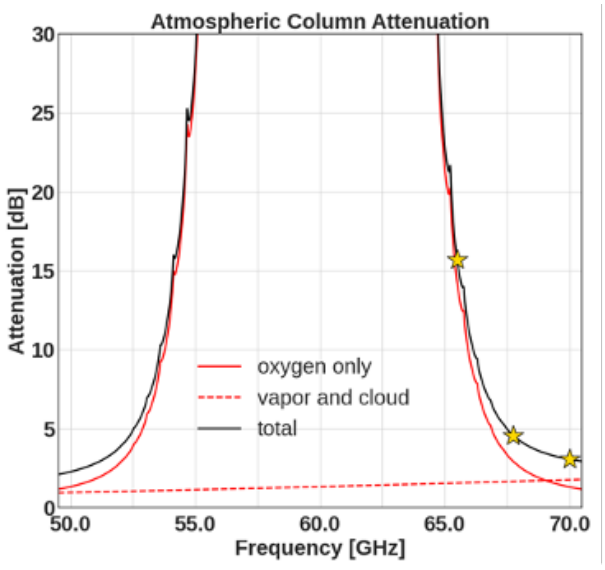


FIG. 2. Atmospheric column attenuation characteristics at the V-band O_2 absorption complex. The instrument will operate at the upper sideband (64 - 70 GHz) of the O_2 absorption complex with three-frequency transmission (indicated by stars) to estimate atmospheric column O_2 amounts while mitigating the impact of water vapor and hydrometeors. An estimate of the atmospheric O_2 provides the dry-air mass, and thus the dry-air surface pressure.

Figure 2 shows the atmospheric column attenuation characteristics at the O_2 V-band. The total attenuation of the atmosphere (black curve) varies from a minimum of approximately 1 dB in the far wing frequency region to much more than 30 dB near the V-band center. This total attenuation is primarily decided by the O_2 absorption (red curve) with non-negligible contributions from atmospheric water vapor and clouds (dashed line). Generally, the instrument design needs to avoid the central spectral region of this O_2-absorption complex to circumvent excessive attenuation. Instead of the frequency selection from the lower sideband of this absorption complex as in previous work (e.g., Lin and Hu (2005)), three upper sideband radar frequencies (yellow stars) are considered to eliminate potential interference of passive microwave sensors (such as Advanced Microwave

Sounding Unit (AMSU)) operating at the same lower sideband for atmospheric temperature vertical profile measurements.

The radar frequencies are chosen to be closely-spaced such that the water vapor (V) and cloud hydrometer (H) attenuation and surface reflection are similar, but the difference in O_2 absorption is substantial. Thus, when differential radar returns from paired frequencies are measured, the effects of V, H, and surface reflectivity on the paired radar signals are significantly reduced and the differential loss due to atmospheric O_2 is dominant and can be measured. As seen in Figure 2, the atmospheric attenuation from water vapor and cloud hydrometers increases linearly (in logarithmic units) with frequency. When the three frequencies chosen (70.0, 67.75, and 65.5 GHz in the order of increased O_2 absorption) are evenly spaced and their returns are organized into 70.0/67.75 and 67.75/65.5 two pairs, non-oxygen common factors affecting radar return signals such as the residual small differential loss from water vapor and hydrometers can be further reduced to a negligible level or eliminated using the difference of the differential losses of the two pairs. Likewise, this two-pair (or three-frequency) approach would remove effects of sea surface reflectance differences on SLP retrievals too due to quasi-linear features of the reflectance at those closely-spaced wavelengths. As the total differential absorption signal associated with SLP changes is small, many independent samples at each radar frequency are needed to reduce the fading noise of the surface backscatter. This is achieved through the use of frequency diversity, with multiple transmitted and received waveforms per radar pulse repetition interval (PRI). Thus, high precision SLP retrievals are expected.

In addition to the DAR, the spaceborne instrument concept uses microwave radiometry at the same V-band frequency range to estimate the atmospheric temperature. This will further improve the surface pressure estimate, as well as provide a means to achieve vertical pressure information through the Hypsometric equation (Salby 1995).

To understand the impact of spaceborne marine surface pressure (SMSP) retrievals on global weather forecasts of NWP models, various architecture and operation potentials of spaceborne systems are considered. For example, the transmitted power is to be 40 to 50 dBm, and the antenna diameter will be 1 to 2.5 m. These radar system parameters are chosen to be consistent with existing radar technologies such as those from TRMM, GPM, CloudSat, and EarthCare space missions. Spacecraft would be in LEO orbits with altitudes ranging from low about 400 km of TRMM-like

orbits to high around 705 km of A-Train-like orbits. Sea surface spatial resolution for reported SLP averages could vary from 5 km to 50 km, depending on mission requirements and system designs. Viewing modes would be both narrow nadir viewing only for limited spatial sampling and wide cross-track scanning with a viewing angle as large as ±15°. This type of instrument on a polar or inclined orbit could provide additional SLP inputs to NWP models to fill in gaps in surface pressure measurements over the oceans as illustrated in Figure 1.

The precision of SMSP retrievals is targeted at 1–2 hPa, which is obtained based on current in-situ observational capability and NWP model needs. While in-situ sensors can be very accurate, Ingleby (2010) estimated drifting buoy SLP observation errors to be 0.9 hPa. Kent and Berry (2005) found that global average random errors from ship-based pressure measurements were greater than 1.5 hPa even for 12 month running means. These results indicate the errors in in-situ SLP inputs for NWP models could be even bigger, especially in local and regional scales. Thus, the precision of our spaceborne observation goal for remote-sensed SLP is assumed to be 1–2 hPa, a level current instrumentation could reach.

In 2022 the NASA Earth Science Technology Office (ESTO) funded the development of an airborne demonstrator to advance the technical maturity of this instrument concept. The Microwave Barometric Radar and Sounder (MBARS) instrument is being developed by NASA/Goddard Space Flight Center (GSFC), NASA/Langley Research Center (LaRC), and Tomorrow.io under the Instrument Incubator Program (IIP). The MBARS instrument is planned to be completed for engineering test flights on the NASA ER-2 high-altitude aircraft by summer 2024. MBARS will be configured to fly either Nadir-pointing or in a cross-track scan configuration, and will contain the V-band DAR and radiometric temperature sounder. Co-incident radiometric water vapor sounding will be performed by a separate microwave radiometer.

b. Observing System Simulation Experiments

Observing system simulation experiments (OSSEs) are a well-known method of simulating proposed new observing systems and testing the potential impact in a realistic and sophisticated framework (Arnold and Dey (1986), Hoffman and Atlas (2016), Errico and Privé (2018)). In an OSSE, the real world is replaced by a long, free model run of a high-resolution NWP model, referred to as the Nature Run (NR). This NR is the basis for simulating all data types used operationally from

the global observing network, and is also used for the verification fields when calculating analysis and forecast errors. These simulated observations, along with simulations of the proposed new instrument data, are then ingested into a different NWP model using a modern data assimilation system (DAS). The OSSE should be calibrated and validated to ensure that the characteristic behaviors of the DAS are well replicated compared to the real world. By comparing forecasts with and without ingestion of the simulated new observing type, the impact of the proposed instrument on NWP can be estimated. As the OSSE framework is entirely simulated, different instrument configurations and sampling strategies can be compared to inform design decisions.

In this work, an OSSE is performed to demonstrate proof-of-concept for the use of SMSP observations in NWP. A traditional observing system experiment (OSE) is not possible because there is no available data since the instrument has not yet been built or tested. This initial study of the impact of air pressure observations on numerical weather forecasts will focus on SLP retrievals, i.e., the active sensing portion of the instrument system. The NASA/GMAO global OSSE framework is used to conduct preliminary evaluation of different possible sampling designs for a spaceborne instrument for measuring marine surface pressures. The GMAO OSSE has been used for a number of previous OSSE studies (Privé et al. (2014), McCarty et al. (2021)), and the behavior has been well characterized (Errico et al. (2013), Privé et al. (2013b), Privé and Errico (2019), Privé et al. (2021)). One of the key tools available for the GMAO OSSE is adjoint-based forecast sensitivity to observation impact (FSOI) that can be used to compare the relative contributions of different observation types to improving forecast skill. The goal of this initial study is to demonstrate the potential impacts of idealized SMSP data on operational NWP as well as to identify areas of the DAS that may require additional development in order to optimize use of marine surface pressure observations. In this initial work, the SMSP observations are treated in a simplified fashion as if they were in-situ point measurements of surface pressure without added simulated observation errors in order to examine the impacts of the information content of the observations on NWP. Future studies are expected to explore the instrument requirements such as observation errors, effective field of view, influence of rain contamination, and selection of orbits.

An OSSE was chosen for these experiments due to the availability of a mature, validated OSSE framework at the GMAO. Another available method of testing the impact of proposed new observing systems is an ensemble of data assimilations (EDA) which also uses simulated observations to

estimate observation impacts by measuring the ability of the simulated observations to affect the ensemble spread and short-term forecasts. EDAs were first developed by Tan et al. (2007) for evaluation of the Aeolus wind-profiling LIDAR, and have been also used to evaluate radio occultations (Harnisch et al. 2013). One advantage of the OSSE over EDA methods is the ability to examine observation impacts on medium range forecast skill.

A description of the GMAO OSSE system, the simulation of observations, and validation of the OSSE is given in Section 2. The results of the OSSE experiments are given in Section 3, and a discussion of the findings as well as future work is found in Section 4.

2. Method

The NASA/GMAO NWP OSSE framework has been well-documented and validated for use in evaluating proposed new data types. Full details of the OSSE are described in Errico et al. (2017); a brief overview along with the specific design of the surface pressure observation experiments is given here.The method of validation of the overall OSSE framework is described by Errico et al. (2013) and Privé et al. (2013b). Examples of validation of the version of the OSSE framework used in these experiments can be found in Privé et al. (2020), Privé et al. (2021), and Privé et al. (2022).

a. OSSE Framework

The GMAO OSSE uses the 'G5NR' Nature Run developed at NASA/GMAO for simulating observations and verification of experimental forecasts. This Nature Run is a 24 month free run of the Goddard Earth Observing System (GEOS) forecast model at approximately 7 km (C1440) horizontal resolution with 72 vertical levels and 30 minute output frequency. The sea surface temperature (SST), carbon levels, and sea ice used in the G5NR were from the period May 2005 to May 2007, however there is no expectation that the synoptic state of the G5NR should correspond to the synoptic state of the real world during this timeframe. The G5NR has been extensively validated (Gelaro et al. 2014) to ensure that the model behavior is representative of the real atmosphere. As described in Errico and Privé (2018), validation methods depend on what is being validated, and can range from bulk statistical comparisons with reanalyses or operational models to careful analysis of phenomenological case studies.

Simulated or 'synthetic' observations were generated for most data types used in operational NWP during 2015. These simulated data were derived from the NR fields in order to generate an observational suite representing how the NR would have been observed by the global network of conventional and remote sensing instruments. For most conventional data types, such as aircraft and surface observations, the NR fields were spatiotemporally interpolated to the locations of real observations on corresponding dates in 2015. Rawinsondes were 'launched' at the time of real rawinsondes in 2015, but were advected by the NR wind fields. Atmospheric motion vector locations were simulated based on the cloud and water vapor fields of the NR, with an effort to match the statistical distribution of real observations in space and time (Errico et al. 2020). Radiance types such as AMSU-A, ATMS, HIRS4, SSMIS, MHS, AIRS, CrIS, and IASI were simulated using the Community Radiance Transfer Model (CRTM, Han et al. (2006)) including cloud contamination from the NR cloud fields. GNSS-RO data were simulated using the two-dimensional radio occultation processing package (Culverwell et al. 2015).

As simulated observations have reduced intrinsic errors compared to real observations that include not only instrument error, but also additional errors such as representativeness error, simulated errors were added to the synthetic observations (Errico et al. (2017), Privé et al. (2021)). The simulated errors were calibrated for each data type so that the statistics of the variance of observation innovation (O-F) in the OSSE match the statistics from real data. Random uncorrelated errors were added to all data types, and random correlated errors were added to selected data types. Vertically correlated errors were added to rawinsondes, AMVs, and GNSS-RO, channel correlated errors were added to the hyperspectral instruments IASI, AIRS, and CrIS, and horizontally correlated errors were added to AMVs, AMSU-A, HIRS4, SSMIS, and MHS. Aircraft, conventional surface types, and scatterometer data had only uncorrelated errors added.

The simulated observations were ingested into a semi-operational NWP model in order to perform the OSSE experiments. These simulated observations are handled by the data assimilation in exactly the same way as real observations. The DAS used in the GMAO OSSE was the 4-dimensional hybrid ensemble variational (4D-EnVar) Gridpoint Statistical Interpolation (GSI). The NWP model used was the GEOS model version employed semi-operationally in 2018 with approximately 25 km horizontal resolution (C360) and 72 vertical levels. This model version differs from the model used to generate the G5NR in that it is has lower horizontal resolution, a

different (two-moment, Barahona et al. (2014)) microphysics scheme, and several different physics parametrizations that affect the lower troposphere. These alternate parametrizations were chosen to maximize differences in model processes between the Nature Run and the NWP model. At the time of running these experiments, the 2018 model version was not significantly different from the then-operational model version, and was deemed to be a reasonable model choice.

b. Validation

This model setup would be considered a 'fraternal twin' OSSE, where a similar but not identical model is used for the NR and the NWP experiments. The primary consequence of using a fraternal twin setup is reduced model error compared to the real world - while OSSEs in general tend to have weaker model error, the issue is exacerbated for fraternal twin OSSEs. This results in slower model error growth, unrealistically good forecast skill, smaller analysis increments, and weaker forecast impacts. It is important therefore to perform a thorough validation of the OSSE performance to understand how the OSSE results might differ from the real world. This hybrid 4D-EnVar version of the GMAO OSSE has been extensively documented and validated (El Akkraoui et al. 2023). Additional validation specific to the surface pressure fields pertinent for the SMSP instrument has been conducted for these experiments as described below.

A Control run is first performed for the OSSE - the simulated data for observing types that are used currently in semi-operational practice are ingested for the month of July in the second year of the Nature Run, without any added SMSP observations. In order to place the experiment results into context, the quality of the Control surface pressure analysis is of interest. The analysis error of the Control is calculated by comparing the Control fields with the Nature run fields regridded to the same resolution, because the Nature Run acts as the "truth" in the OSSE. Figure 3 shows the bias and standard deviation of the surface pressure errors in the OSSE Control for the month of July. Over the marine surfaces, there is generally a positive bias in surface pressure in both the analysis and background, although there are some regions of negative surface pressure bias off the coast of South America and Africa, and over the Maritime Continent. There are regions of particularly large positive surface pressure bias in the Control, such as over the subtropical Pacific and along the Antarctic sea ice edge, while the relatively well-observed Atlantic has smaller bias.

These regional biases in surface pressure can be the result of different model climatology between the NR and the forecast model.

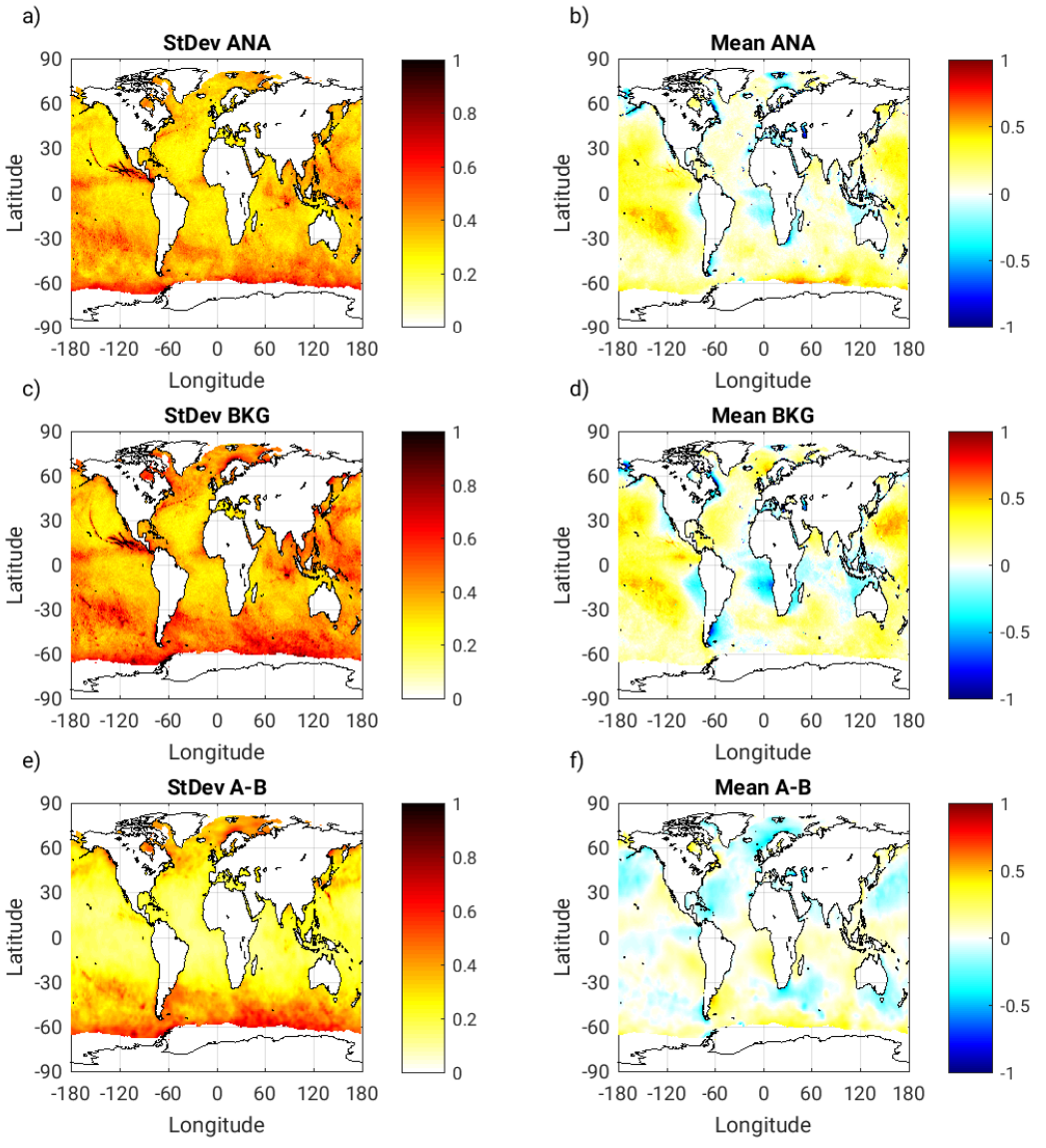
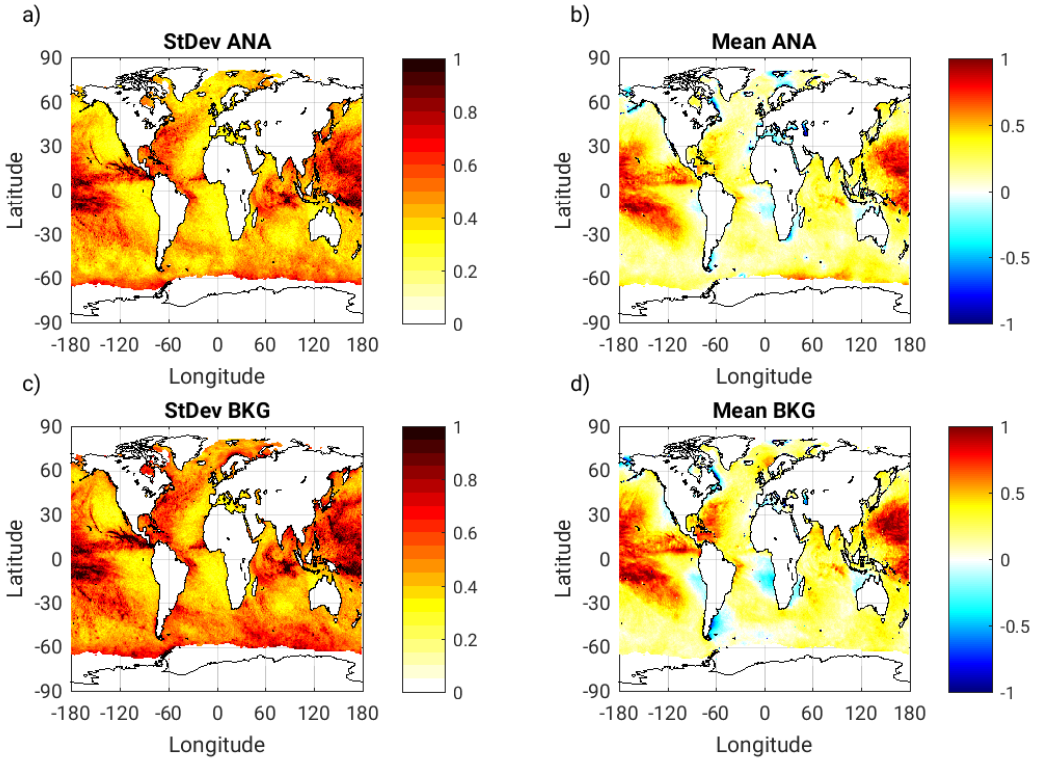


FIG. 3. OSSE Control surface pressure errors (hPa): a) standard deviation analysis error; b) time mean analysis error; c) standard deviation background error; d) time mean background error; e) standard deviation analysis increment; f) time mean analysis increment.

The total surface pressure includes both the pressure due to the dry atmospheric mass as well as a contribution from moisture in the form of vapor and condensates. In order to determine the cause of the surface pressure bias, the dry mass surface pressure is compared to the NR for the Control analysis and background in Figure 4. The dry mass is calculated as in Trenberth and Smith (2005), removing the wet components of total column pressure from the total surface pressure. It is seen that there is a larger bias versus the NR in the dry surface pressure than there is for the total surface pressure. This indicates that the Control has less total precipitable water than the NR, but more dry atmospheric mass. While the differences in total precipitable water can be related to changes in the model physics in the GEOS experiment model compared to the NR model, the dry mass is conserved by the forward model integration and is instead constrained by the dry mass of the initial condition of each run. The dry mass constraint is implemented as a modification to the Incremental Analysis Update (IAU) procedure used with the GMAO GSI as detailed in Takacs et al. (2015) and Gelaro et al. (2017). This constraint was added in order to alleviate undesirable precipitation/evaporation feedbacks that occurred in response to the ingestion of observational data. In the experiment model, the GSI is set to constrain to a global mean surface pressure of 983.24 hPa, which was chosen based on a comparison of several reanalyses (Takacs et al. 2015). Changes in dry surface pressure between DAS versions and differences between reanalysis datasets have been on the order of 0.2 hPa, thus the bias seen between the analysis and the NR is considered to be a bias of reasonable magnitude that might occur in a real world NWP system. The dry mass constraint was not changed to match exactly with the G5NR dry mass, as this would contribute to the lack of sufficient model error in the OSSE.

Examination of the background bias in surface pressure in Figure 3, as well as the time-mean analysis increment for the Control, shows that in relatively well-observed areas of the ocean, such as the North Pacific and the Atlantic, the baseline observing system acts to reduce the magnitude of the surface pressure bias. However in some poorly observed areas, such as along the Antarctic sea ice edge, the analysis increment acts to strengthen the surface pressure bias. Due to the IAU dry mass constraint, if the dry surface pressure is reduced in one region, there must be a corresponding increase in dry pressure elsewhere. Since there is a significant difference in the total dry mass of the NR compared to the experiment, an effort to draw toward the lower NR surface pressure in some areas will result in an increase in surface pressure error elsewhere. This can be a drawback of



322 FIG. 4. OSSE Control dry mass surface pressure errors (hPa): a) standard deviation analysis error; b) time
323 mean analysis error; c) standard deviation background error; d) time mean background error.

333 the implementation of a dry mass constraint when there is non-uniform global coverage of surface
334 observations, especially when the true global dry mass of the atmosphere is uncertain.

335 This global bias in surface pressure has the advantage of providing a surface pressure error for
336 observations to correct. A common difficulty with OSSEs is that insufficient model error results in
337 background fields that are more accurate than in the real world. For a preliminary proof-of-concept
338 OSSE, having a sufficiently large background error to test the proposed instrument is advantageous.
339 This bias also allows different instrument configurations to be compared with a relatively short
340 experiment period of one month.

341 Figure 3 also shows the standard deviation of the surface pressure analysis and background errors,
342 along with the analysis increment. The regions of large surface pressure error standard deviation
343 are over the Southern Ocean, and the eastern and western tropical Pacific. These are regions where
344 there may be active low pressure systems such as baroclinic lows or tropical cyclones, and where

the model is relatively poorly constrained. The standard deviation of the analysis increment shows little signal in the tropical Pacific, but a larger signal over the Southern Ocean, likely due to a combination of ships, buoys, and radiance observations. As seen in Figure 1, the tropical and subtropical Pacific is sparsely observed by in-situ instruments.

The existence of a surface pressure bias in the Control also affects the statistics of how observations are handled by the DAS. Figure 5 compares the observation innovations (O-F) of marine surface pressure data for the month of July for the Real and Control cases. The OSSE Control has a negative skew with a thick negative tail, and also smaller observation count than Real. The lower count is due to greater rejection of observations with large O-F in the Control case. The negative skew is a result of the lower surface pressures in the G5NR that are reflected in the observations compared to the background state of the Control. Due to the availability of the NR for verification of the observation impacts and analysis quality, this negative skew in the simulated surface pressure observations can be considered in context with the known biases when assessing the SMSP observation impacts. However, the standard deviation of the simulated marine surface pressure observation innovations in the Control is similar to that of the Real innovations, which is the goal of the OSSE calibration.

c. Spaceborne Surface Pressure Simulation

In these preliminary experiments, a simple method of simulating SMSP observations was used to show proof-of-concept. First, the footprint of an orbital instrument was based on the real footprint of the ATMS instrument. This allowed the observation locations to have realistic qualities without the need to develop an orbital simulator. As will be described in further detail in Section 2d, the ATMS scan locations were strategically chosen to test different possible scan configurations of a spaceborne instrument. An example of the selected orbit can be seen in Figures 7 and 8. In order to achieve the desired spacing of observations, every other scan angle from the ATMS instrument was omitted, and along-track scans were selected to produce approximately 50 km spacing between scans.

To generate the SMSP observations, the G5NR surface pressure was simply interpolated to each scan location over marine surfaces as a point measurement. Simulated errors were not added to the synthetic SMSP observations, and as such these experiments can be considered to examine

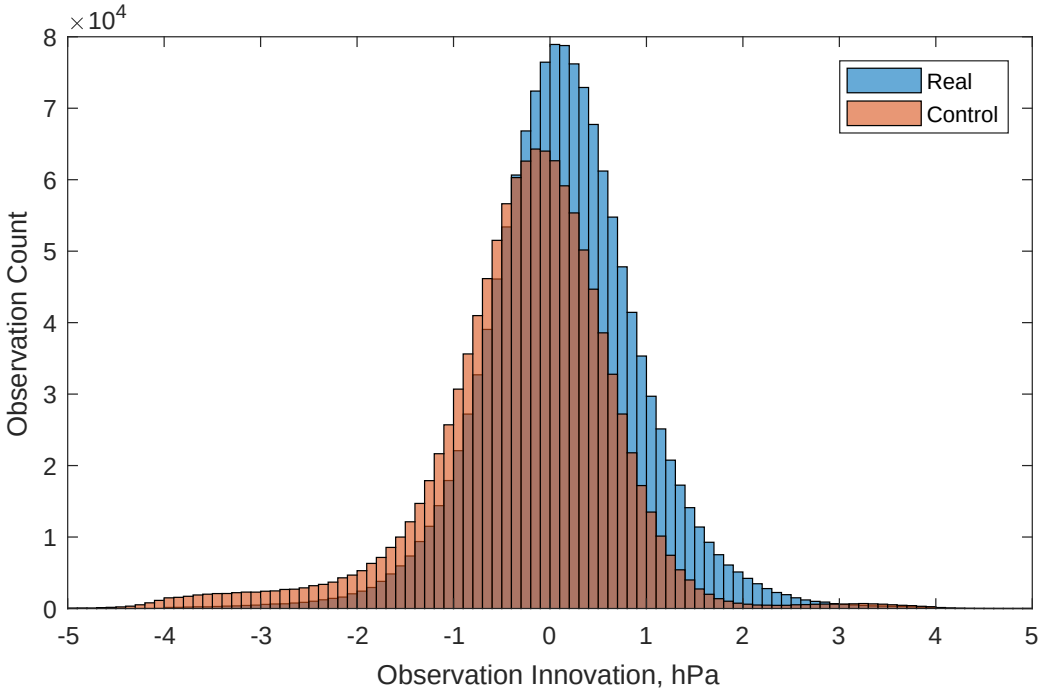


FIG. 5. Histogram of observation innovations (O-F, hPa) for Real (blue) and OSSE Control (orange) for marine surface pressure types including ships, buoys, and platforms, total for the month of July.

the information content of the SMSP observations. The error characteristics of an actual SMSP instrument are not yet known, so the results of these experiments illustrate the greatest possible observation impact from ingesting marine surface pressures. As the ingestion of actual SMSP data into a DAS would require the development of a suitably designed observation operator, for these simple tests the SMSP observations were treated by the GSI as conventional marine surface observations with assumed error standard deviation of 2.0 hPa for the purpose of weighting by the DAS. This GSI weighting is twice that of what is assigned to the other marine surface pressure observation types to allow for greater errors from a spaceborne instrument compared to in situ types. While simplified compared to real SMSP observations, this methodology is expected to capture those aspects of the SMSP data that are most critical for data assimilation, such as the spatiotemporal distribution and count of observations.

d. Experiment Setup

The DAS can be run in two different modes: 1) 'standalone' mode or 2) full cycling. The typical operation of NWP DAS is full cycling, in which observations are ingested by the DAS and combined with a prior forecast or 'background' state to produce a new 'analysis' state, which is the best estimate of the current state of the atmosphere. This analysis is then used as the initial condition for a short forecast cycle that produces a new background for the next cycle time, usually 6 to 12 hours later.

In standalone mode, a baseline run is performed with full cycling, and this baseline run provides all the background states for the standalone run. The baseline run used here for the standalone tests is the OSSE Control run, which includes the set of simulated data based on the 2015 global observational network, but does not include SMSP observations. At each cycle time in standalone mode, the Control background is used as the first guess and combined by the DAS with an observational dataset that includes both the Control observations and the SMSP observations to produce a new analysis. However, the new analysis is not used to initialize a new forecast. A diagram of the two DAS modes is illustrated in Figure 6. The standalone mode has the advantage of being inexpensive to perform, but the background state is not informed by the SMSP observations being ingested. However, the impact of the SMSP observations on the analysis is very similar in the standalone runs to the impacts seen in the full cycling runs. Also, the standalone runs allow the impact of the SMSP observations to be exactly identified without interference of random error growth.

Two sets of experiments were performed for preliminary evaluation of the SMSP instrument. In the first set of experiments, the standalone mode was used to test multiple different configurations of possible scanning methods. The results of these standalone tests were used to inform the selection of a smaller set of tests that were then performed with full cycling.

Table 1 lists the different experiments that were performed using simulated SMSP data. Four different scan widths were tested: 250 km, 500 km, 1000 km, and 2700 km, with the latter being the full width of the ATMS scan on which the simulated data was based. The 250 km scan is the most realistic width for a single instrument, with the wider scan width experiments performed to explore more fully the information content of surface pressure observations in these proof-of-concept studies. The widest swaths are more representative of the kind of observational

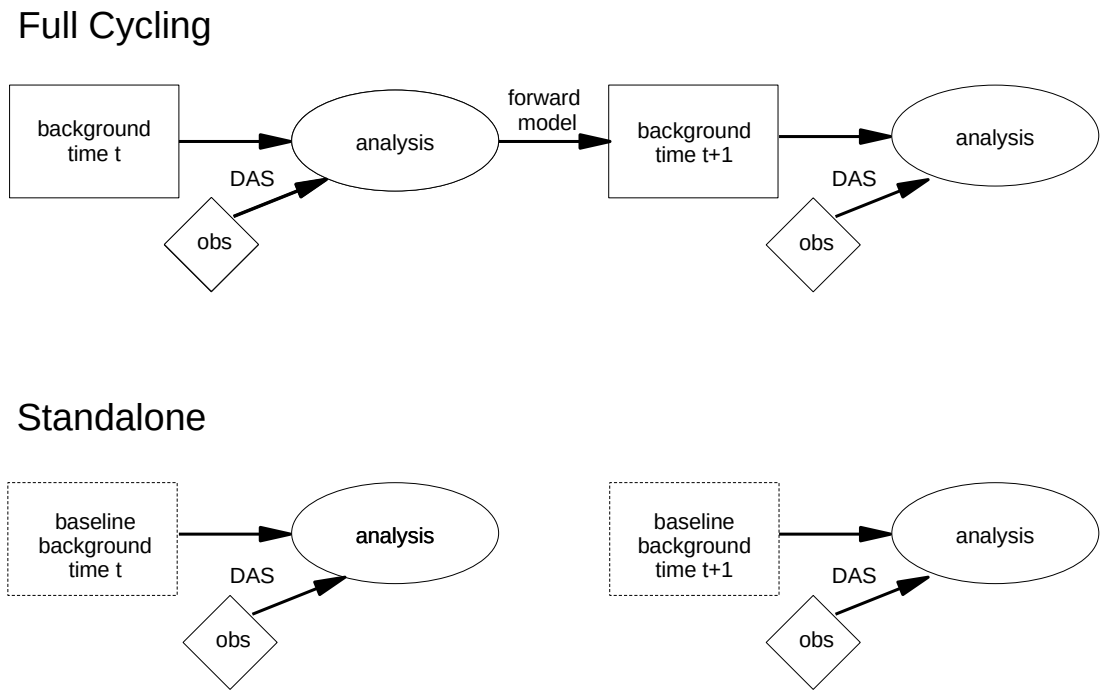


FIG. 6. Diagram of DAS cycling procedure for full cycling (top) and standalone (bottom) configurations. In the standalone configuration, the background states are taken from a prior baseline run. The DAS ingests the observations ('obs') along with the background to produce a new analysis state.

coverage that would be possible with a constellation of SMSP instruments. Examples of the SMSP observations for 0000 UTC 10 July are illustrated for the different scan widths in Figure 7.

3. Results

All SMSP experiments were performed for the month of July in the second year of the Nature Run, with the initial background state taken from the Control case on 0000 UTC 1 July. Comparison of this period to longer test runs of up to three months have shown similar results, indicating that the one-month period is sufficient, especially for calculations of analysis errors. While longer experiment periods would be ideal, this choice of experiment length allowed for multiple scenarios to be tested at moderate computational expense.

TABLE 1. List of experiments.

Name	Scan Width	Thinning	Exp Type
S250	250 km	No	Standalone
S500	500 km	No	Standalone
S500T	500 km	Yes	Standalone
S1000	1000 km	No	Standalone
S1000T	1000 km	Yes	Standalone
S2700	2700 km	No	Standalone
S2700T	2700 km	Yes	Standalone
C250	250 km	No	Full Cycling
C500	500 km	No	Full Cycling
C1000	1000 km	No	Full Cycling
C2700	2700 km	No	Full Cycling

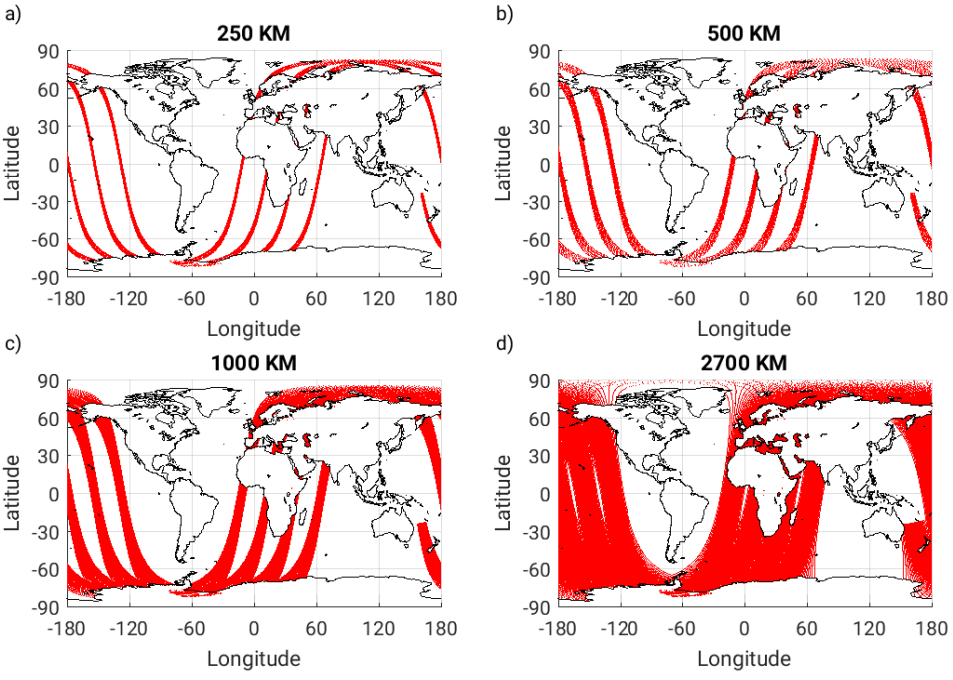


FIG. 7. Maps of locations of SMSP observations (red dots) at 0000 UTC 10 July for the four different scan width configurations: a) 250 km, b) 500 km, c) 1000 km, and d) 2700 km.

For statistics of the analysis and background fields, the time mean, standard deviation, and root-temporal-mean square (RTMS) errors were calculated by comparing the model state with the NR fields regridded to the same horizontal resolution, as the NR is considered the complete truth for

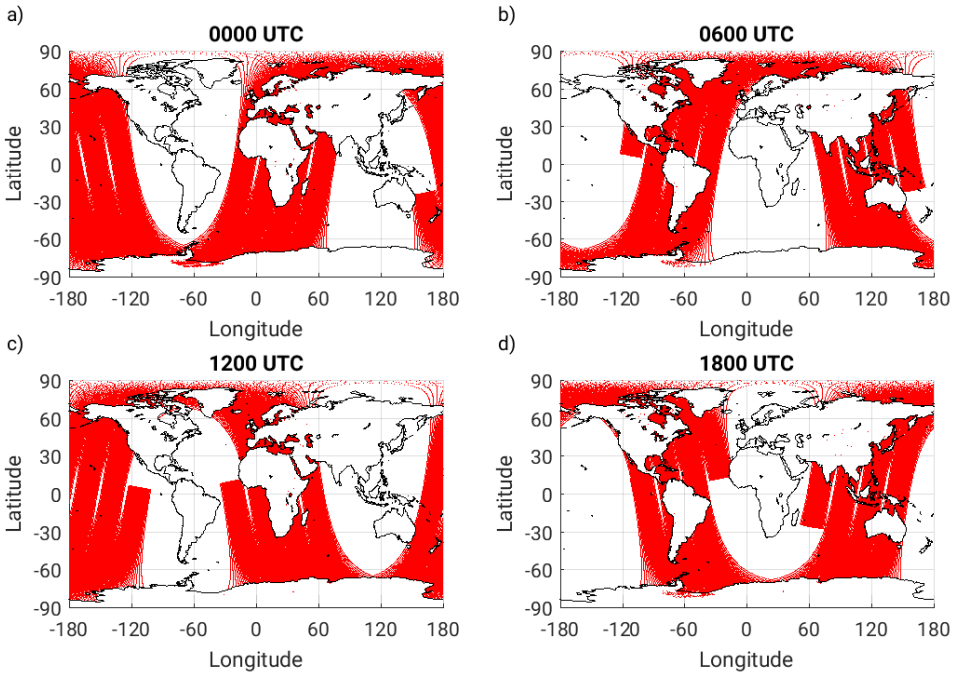
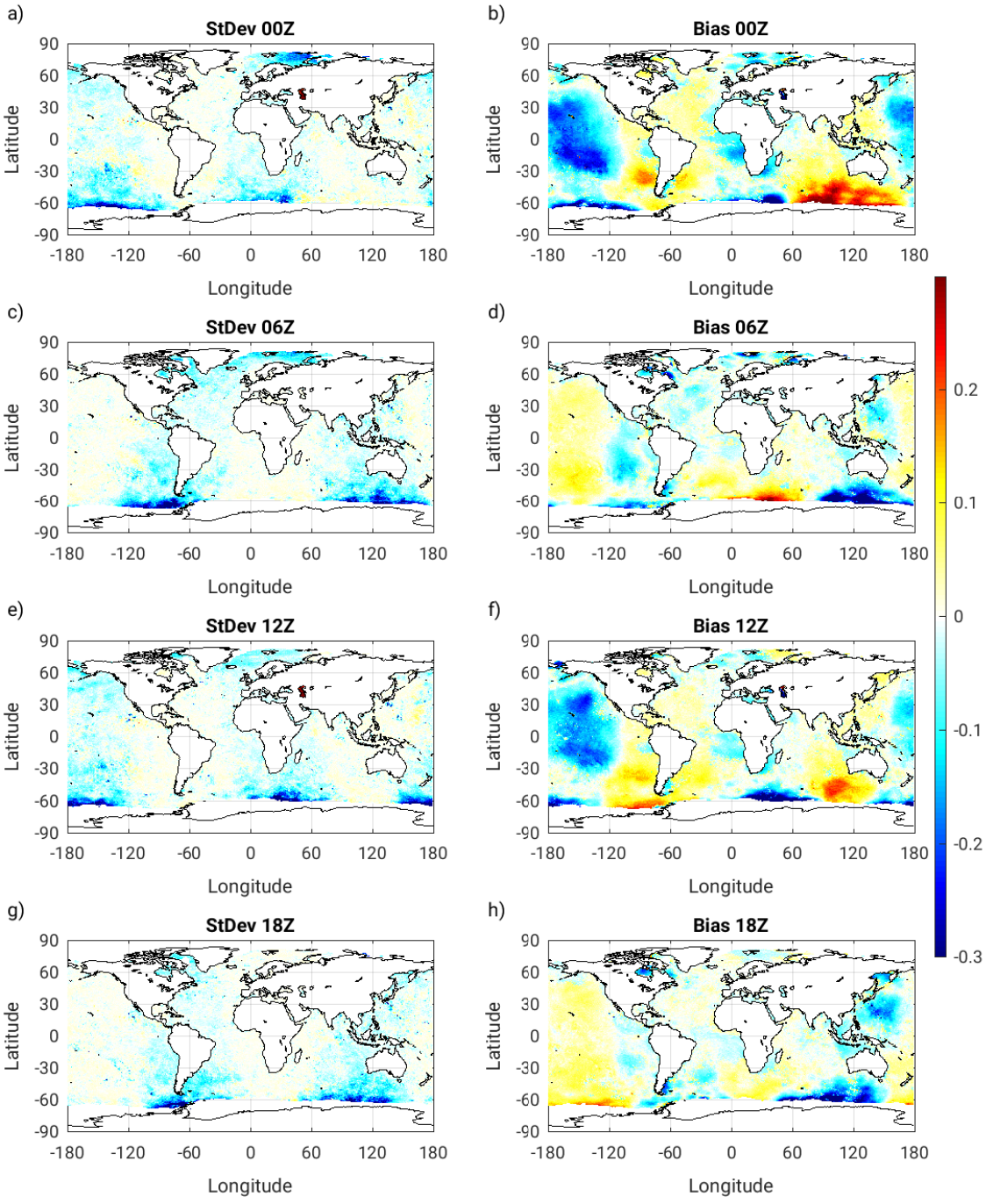


FIG. 8. Maps of locations of SMSP observations (red dots) at different cycle times on 10 July for the S2700 case: a) 0000 UTC, b) 0600 UTC, c) 1200 UTC, and d) 1800 UTC.

the OSSE experiments. Results are shown with areas of land surface and sea ice masked out, as these areas are not observed by the spaceborne instrument and can be excessively noisy due to orographic effects.

a. Standalone

The SMSP swath has variable coverage of ocean surfaces at different cycle times (Figure 8). At 0000 UTC and 1200 UTC, there is SMSP coverage over the central Pacific and southern Atlantic oceans, while at 0600 UTC and 1800 UTC, there is coverage over the western Pacific, Indian, and north Atlantic oceans. Figure 9 compares the impact of SMSP on analysis surface pressure error in the S1000 case at each cycle time for both error standard deviation and absolute time mean error versus Control, where negative values indicate a reduction in error for the S1000 case. The standard deviation of surface pressure error is decreased in the S1000 case compared to Control under the SMSP swath, with largely neutral impacts outside of the swath.



448 FIG. 9. Difference in surface pressure analysis errors for the S1000 Standalone case compared to Control at
449 0000 UTC (a,b); 0600 UTC (c,d), 1200 UTC (e,f), and 1800 UTC (g,h) for the month of July. a,c,e,g) Analysis
450 error standard deviation, hPa; b,d,f,h) analysis absolute bias, hPa.

However, for the absolute time mean error impacts (absolute bias), there is a significant reduction in bias under the SMSP swath but an increase in surface pressure bias outside of the swath coverage. This is due to the combination of the dry mass constraint with the global difference in surface pressure between the background and the NR. Since the marine surface pressure in the background is higher than the NR, the SMSP observations tend to decrease the surface pressures under the swath, resulting in a corresponding increase in surface pressure elsewhere. This increase in surface pressure primarily occurs in poorly observed regions such as near the Antarctic sea ice edge and south of Australia. When these poorly observed regions are covered by an SMSP swath (0600 and 1800 UTC), the increased surface pressure is spread out over a large region of the Pacific Ocean. In a test in which the dry mass constraint was removed, these regions of degradation were absent.

When statistics are calculated using four times daily cycles, the alternating swath coverage with reduced and then increased surface pressures at every other cycle time affects the standard deviation of surface pressure error. Instead, the RTMS surface pressure error versus Control for four times daily cycle times is shown in Figure 10 for the S250, S500, S1000, and S2700 cases.

In general, Figure 10 shows that the SMSP impacts increase with wider scan angle (and correspondingly greater number of SMSP observations). The analysis RTMS surface pressure error is decreased over the Pacific, much of the Atlantic, and along the Antarctic sea ice edge. There are some regions of slightly increased RTMS error in the Southern Ocean which are a result of the dry mass constraint.

1) Thinning

One question of interest is whether the greater SMSP impacts seen with wider scan configurations are due to the larger areal coverage or due to the larger number of SMSP observations. In order to test this, three additional configurations are tested. These standalone experiments have scan widths of 500 km, 1000 km, and 2700 km, but the observations are 'thinned' to have the same number of observations as the S250 case.

Figure 11 compares the RTMS surface pressure error for the thinned experiments compared to Control (left column) and to the corresponding un-thinned standalone cases (right column). For the S500T case, the improvement in RTMS surface pressure error is only slightly weaker than in the un-thinned case, with more improvement than the S250 case. There is a greater difference

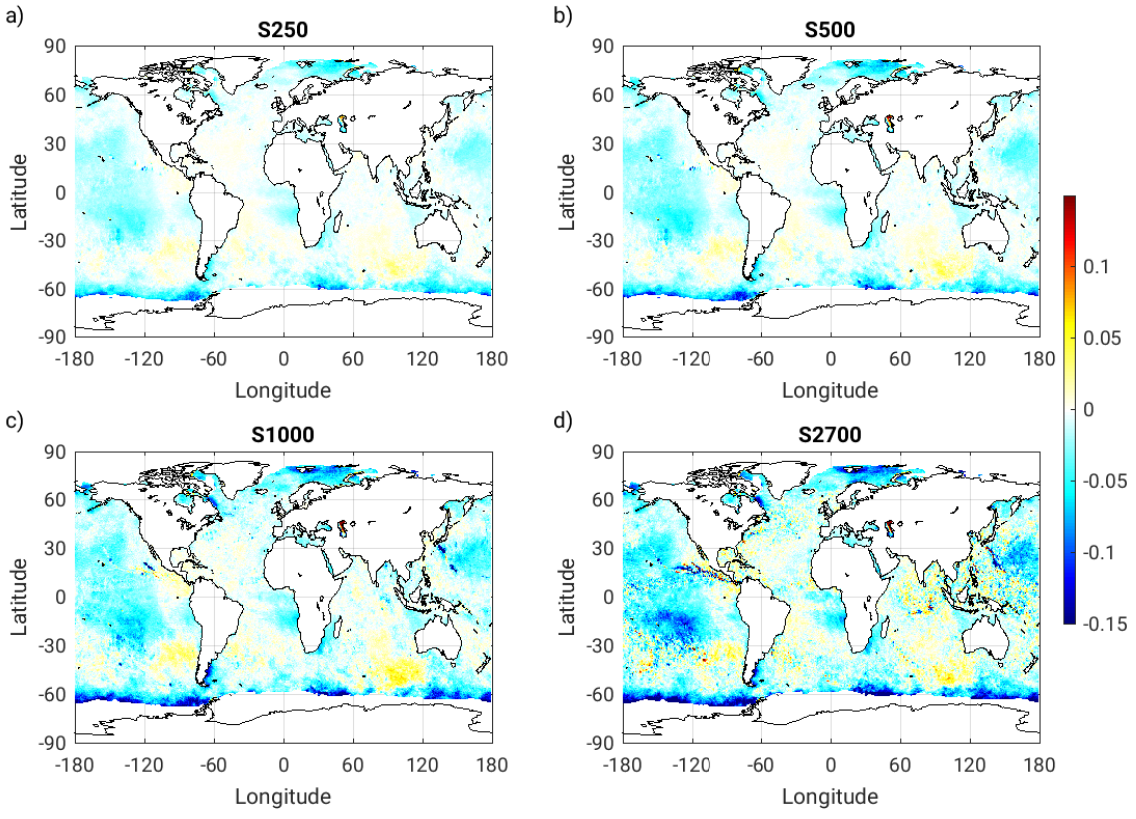
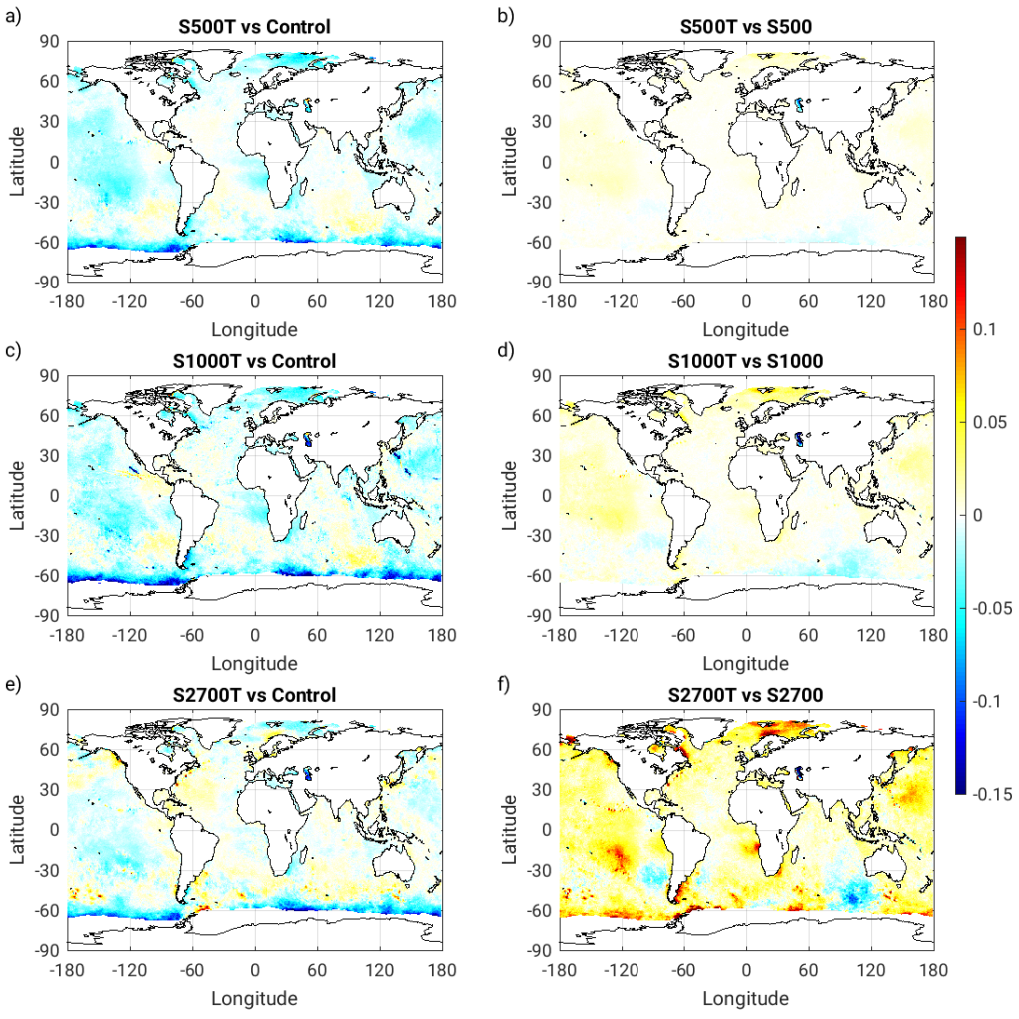


FIG. 10. Difference RTMS Surface pressure analysis errors (hPa) for the unthinned Standalone cases versus Control, four times daily for the month of July. a) S250, b) S500, c) S1000, d) S2700.

between the thinned and unthinned 1000 km scan case than was seen for the 500 km case. The S500T and S1000T cases have similar impact of SMSP in the tropical and subtropical oceans, but the wider scan configurations have greater beneficial impacts along Antarctic sea ice edge.

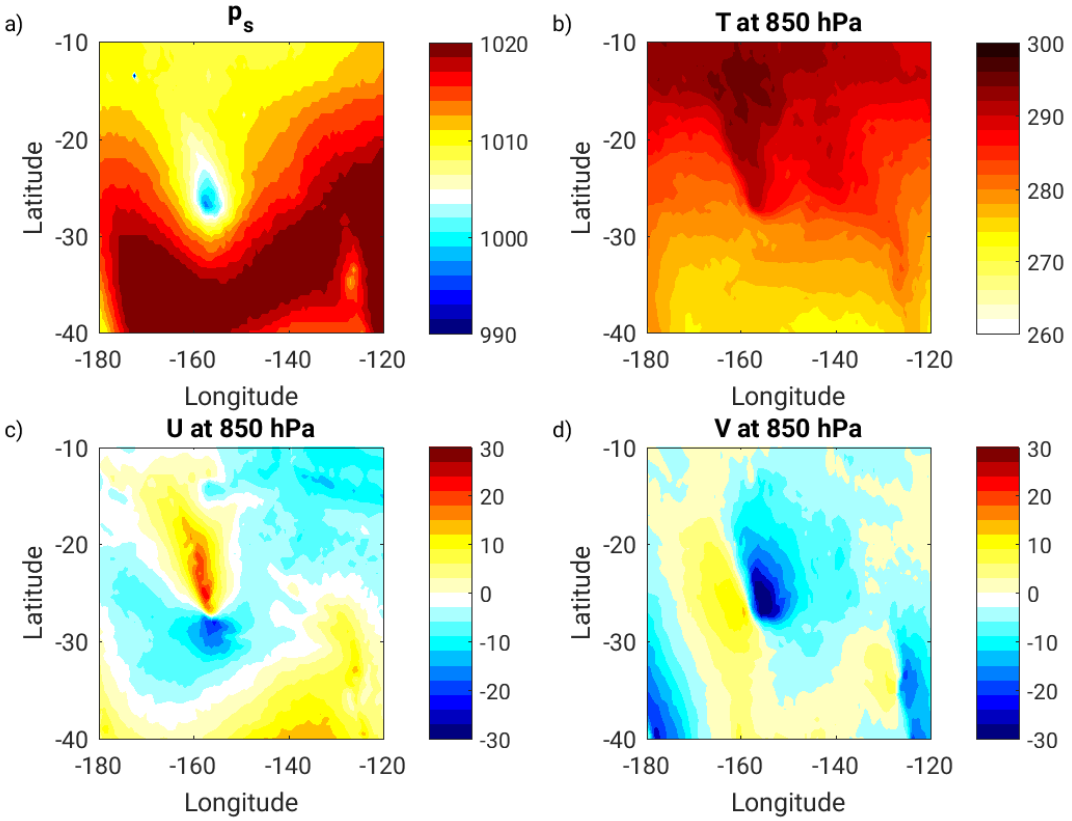
For the S2700T case, the RTMS surface pressure impacts from SMSP are weaker than the S500T and S1000T cases, and much weaker than the unthinned S2700 case. For this thinned configuration, the observations are quite widely spaced. These results imply that there is an optimal combination of observation spacing and scan width if there is limited capability for scan angles, with the 500 or 1000 km scan widths closer to the ideal configuration.



485 FIG. 11. RTMS Surface pressure analysis errors for the thinned Standalone cases four times daily for the month
486 of July a,c,e) compared to Control; b,d,f) compared to unthinned SMSP. a,b) S500T, c,d) S1000T, e,f) S2700T.

492 ### 2) CASE STUDY

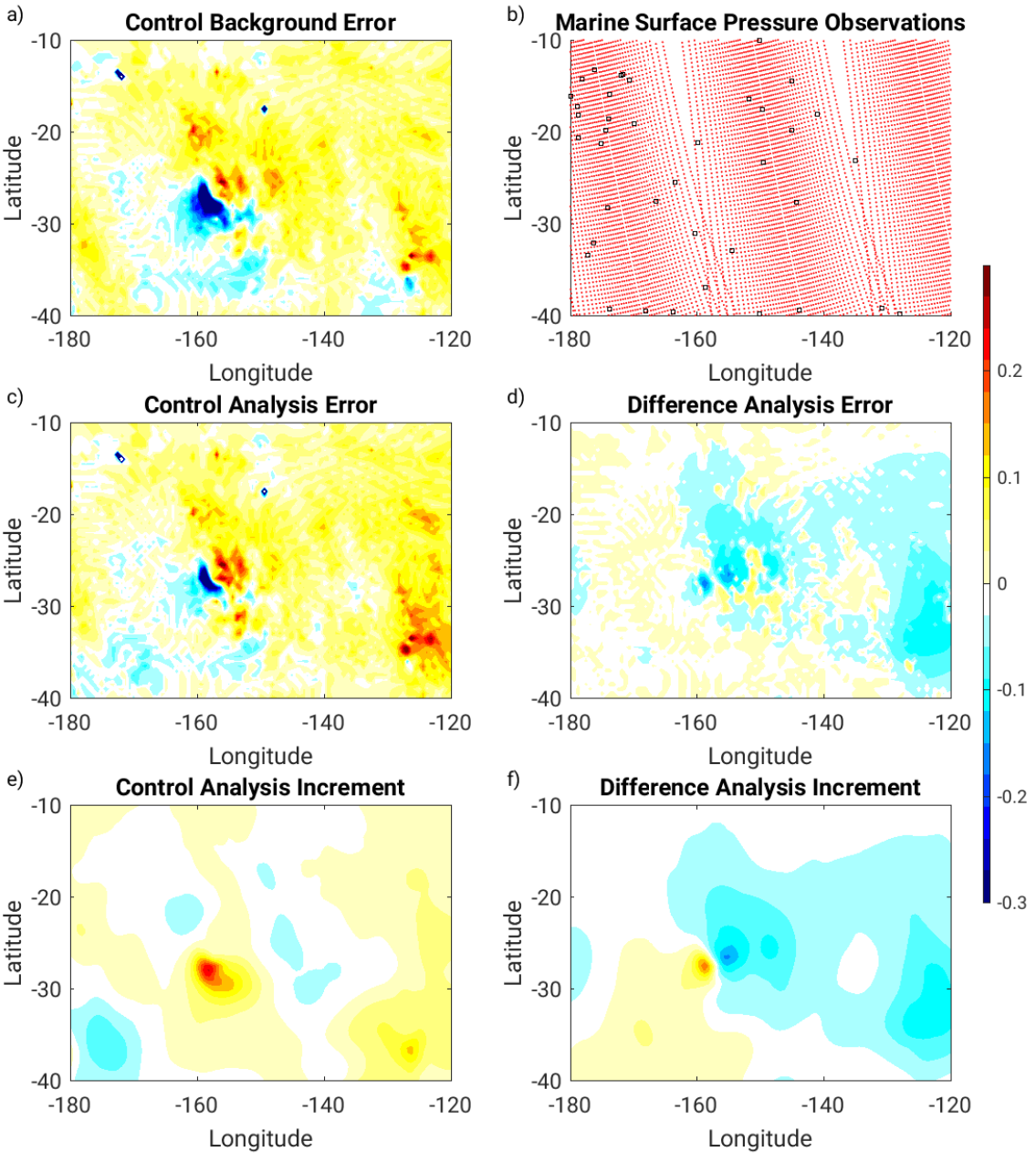
493 A baroclinic low in the central Pacific is selected as a case study to examine how surface pressure
494 measurements can affect the accuracy of particular synoptic systems. For maximum observational
495 coverage, the S2700 experiment is selected. A baroclinic low located near 155W and 27S is found
496 to have significant background errors on 0000 UTC 10 July. The lower tropospheric structure of
497 this system is illustrated in Figure 12, showing the surface pressure field, temperature, and winds



501 FIG. 12. Synoptic situation for the case study of a baroclinic low in the central Pacific, 0000 UTC 10 July,
502 Nature Run fields. a) Surface pressure (hPa); b) temperature at 850 hPa (K); c) zonal wind at 850 hPa ($m\,s^{-1}$);
503 d) meridional wind at 850 hPa ($m\,s^{-1}$)

498 at 850 hPa in the NR. The low is associated with a strong temperature anomaly and cold front
499 extending from 15S to 27S near 160W. The wind field has an elongated vortex structure at low
500 levels, and is associated with an upper level Rossby wave in the winter jet stream (not shown).

504 The error structure of the pressure field for this baroclinic low for both the Control and the S2700
505 experiment is shown in Figure 13. The background state - which is not altered in the standalone
506 experiment - features a displacement of the surface low to the southwest of the NR location, with
507 the low having a minimum surface pressure slightly lower than the NR. The coverage of both SMSP
508 and other marine in-situ surface observations for this cycle time is illustrated in Figure 13b. The
509 ships and buoys are widely spaced and only minimally sample the vicinity of the surface low, while
510 there is nearly complete coverage of the region by SMSP.



511 FIG. 13. Comparison of Control and S2700 experiments for the case study baroclinic low for 0000 UTC 10
512 July. a) surface pressure background error, hPa; b) marine surface pressure observations for 0000 UTC cycle
513 time, conventional marine surface types for both Control and S2700 (black boxes) and SMSP (red dots, S2700
514 only) c) Control surface pressure analysis error, hPa; d) difference in analysis error between S2700 and Control;
515 e) analysis increment (A-B) for surface pressure, Control (hPa); f) difference in analysis increment for surface
516 pressure, S2700-Control (hPa).

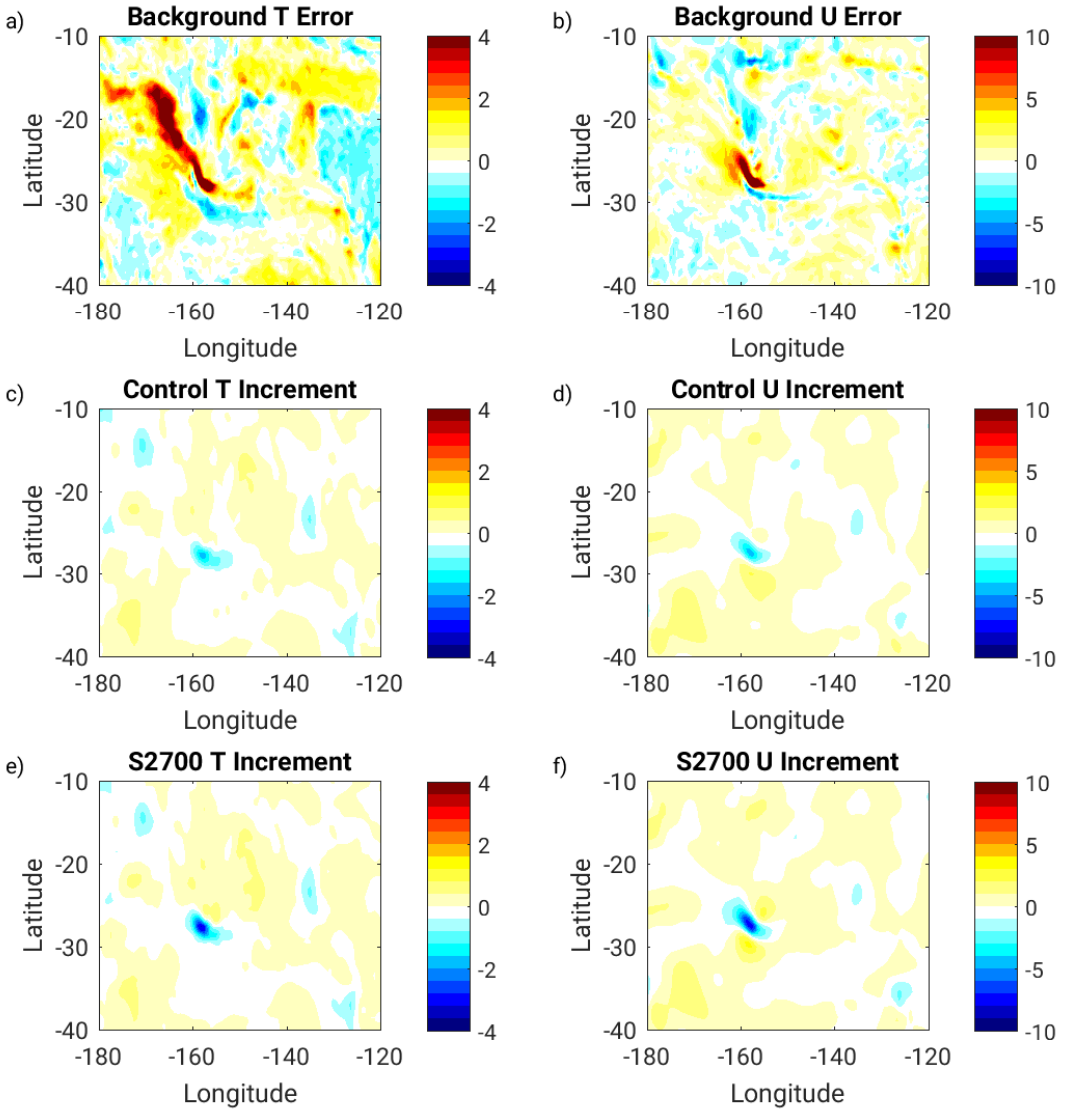
In the Control case, the available observations are able to partially correct the displacement of the surface low, largely by increasing the surface pressure near the center of the low, as shown by the analysis increment in Figure 13e. The resulting analysis error for the Control tends to weaken the low overall, but does not act to shift the low to the correct position to the northeast, as evidenced by the dipole error structure in Figure 13c.

The inclusion of SMSP observations acts to correct the large scale pressure field, and also tends to shift the baroclinic low closer to the correct location. The difference in analysis error for the S2700 case compared to Control is shown in Figure 13d, with reduction in error across a large area. The difference in the analysis increment for the S2700 case compared to Control shows a dipole structure, indicating that the SMSP observations induce a movement of the low rather than only the weakening seen in the Control.

Figure 14 shows the background error and difference in analysis increments between the S2700 case and Control for temperature and zonal wind at 850 hPa. For temperature, the background error field shows a displacement of the cold front to the west of the true location. The temperature increments due to surface pressure observations are only able to affect the temperature field near the location of the baroclinic low, and not along the length of the front. The inclusion of SMSP data results in a slightly stronger temperature increment at 850 hPa but very little difference at 500 hPa or higher (not shown).

Similarly, the wind field background error shows a displacement of the low level wind field associated with displacement of the baroclinic low and frontal system, although this feature has greatest errors close to the surface low. At 850 hPa, the analysis increment acts to correct the wind field near the surface low, with slightly stronger increments for the S2700 case compared to Control. Again, very little difference is observed at altitudes of 500 hPa or above (not shown).

This experiment does not consider the effects of rain contamination on the SMSP observations. In the vicinity of this system, there is a narrow band of heavy rain associated with the front. The SMSP observations in the area of that rainband would likely be contaminated. If rain contamination were considered here, it is expected that the analysis increment due to the SMSP would be similar in location and character, but weaker in magnitude.



535 FIG. 14. Case study comparison of 850 hPa temperature and winds at 0000 UTC 10 July. a) Background
536 temperature error (K); b) Background zonal wind error ($m\,s^{-1}$); c) Control temperature analysis increment (K);
537 d) Control zonal wind analysis increment ($m\,s^{-1}$); e) S2700 temperature analysis increment (K); f) S2700 zonal
538 wind analysis increment ($m\,s^{-1}$)

549 *b. Full OSSE*

550 The unthinned scenarios with scan width of 250, 500, 1000, and 2700 km were further tested
551 running the GSI/GEOS system in the regular cycling mode, designated as C250, C500, C1000, and

C2700 respectively. In addition to the cycling analyses and FSOI calculations that were performed for all cycling runs, the C1000 case was selected for the generation of 4-day forecasts in order to investigate the impact of SMSP on forecast skill.

1) Analysis Impact

As with the standalone experiments, the analyses for the cycling experiments may be directly compared with the NR to calculate analysis errors. Unlike the standalone experiments, the background fields used by the GSI to generate the analysis are informed by SMSP data from prior cycling times, allowing some accumulation of information in the background state (Figure 6).

The impact of SMSP on the RTMS surface pressure error is shown in Figure 15, where negative values indicate a reduction in error with the addition of SMSP observations. The greatest reductions in error are found along the Antarctic sea ice edge, in the central and western tropical and subtropical Pacific, and in the vicinity of the Barents Sea. Some beneficial impacts are also seen in the tropical eastern Pacific that are related to tropical cyclone tracks. As with the standalone experiments, there are some regions of analysis degradation, most notably to the southwest of Australia. These are regions that poorly observed by in situ observations, and where mass is redistributed by the dry mass constraint when the SMSP orbit is not overhead, especially during the 0000 and 1200 UTC cycles.

Comparing the cycling results in Figure 15 with the equivalent standalone case results in Figure 10, the impacts in the cycling cases have slightly higher magnitude than the corresponding standalone cases, but with overall similar character and distribution. This indicates that some surface pressure information is retained between cycles for a cumulative impact when the DAS is cycled. The overall similarity between the standalone and cycling cases supports the utility of the inexpensive standalone methodology when testing many different design options in the OSSE.

To put the RTMS error impacts into context with the total surface pressure error, Figure 16 shows the fraction of the Control analysis RTMS surface pressure that is changed in the SMSP experiment cases. In some regions such as the central Pacific, Barents sea, and southern ocean, the surface pressure error is reduced by more than 20% for the C1000 case, and greater than 25% in the C2700 case.

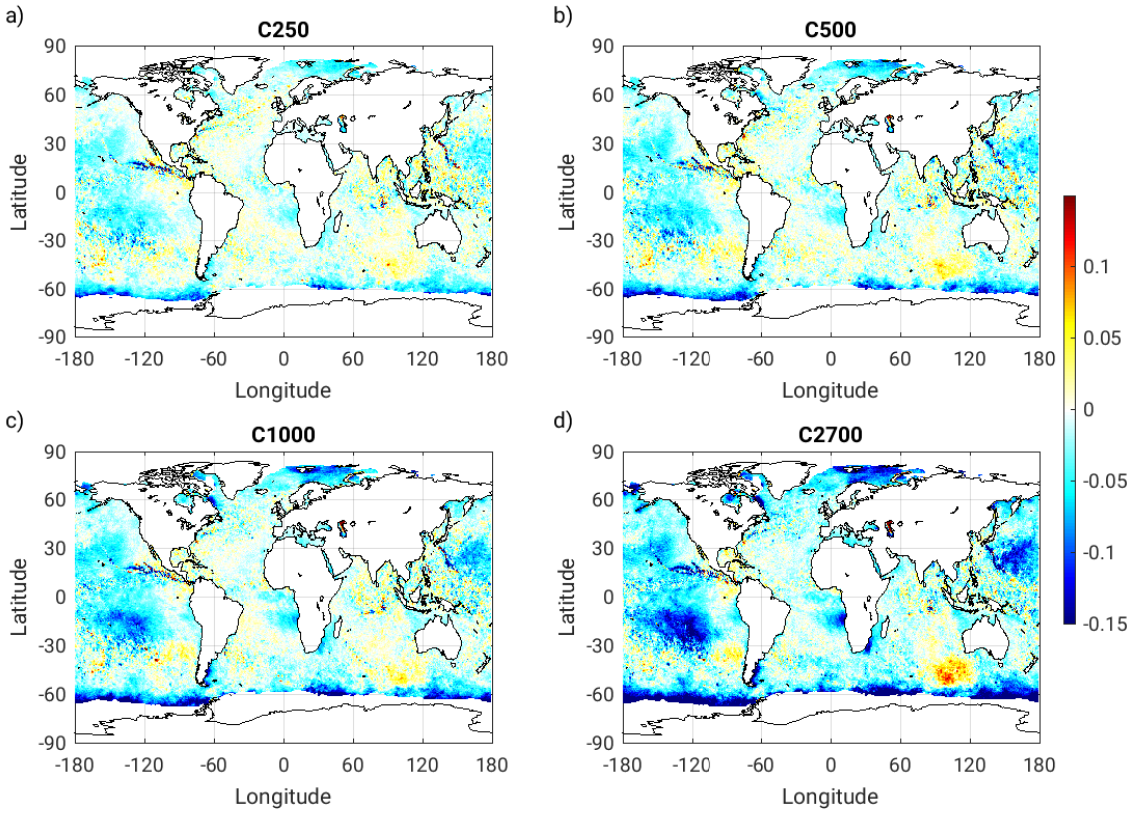
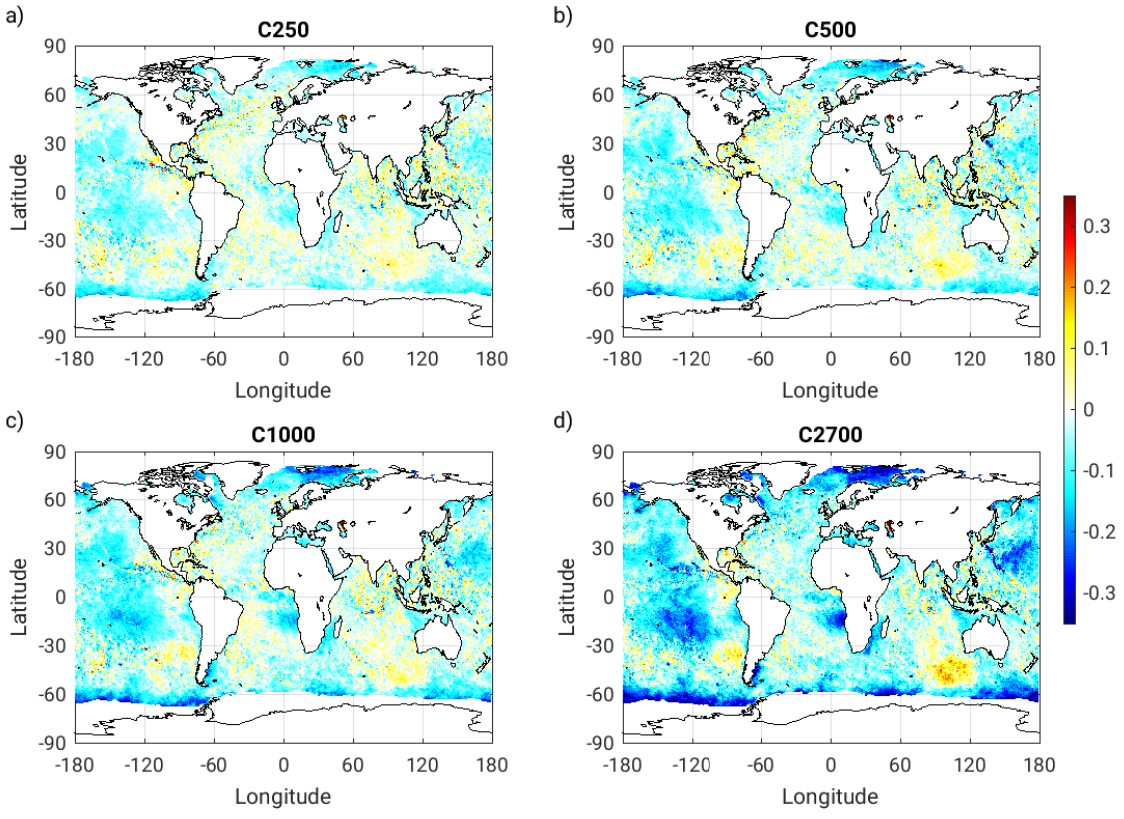


FIG. 15. RTMS surface pressure analysis error (hPa) compared to Control for cycling cases, four times daily for the month of July. a) C250, b) C500, c) C1000, d) C2700.

2) FORECAST IMPACT

Forecasts were performed once daily starting at 0000 UTC for the C1000 case, as well as for the Control. Figure 17 shows a comparison of the root-areal mean square (RAMS) forecast error for C1000 vs Control, normalized by the Control RAMS error, where negative values are a reduction of error in C1000 and stippling indicates a 90th percentile significance using a Pearson paired test. Three regions were examined: 20N-70N (NHEX), 20S-70S (SHEX) and 20S-20N (Tropics).

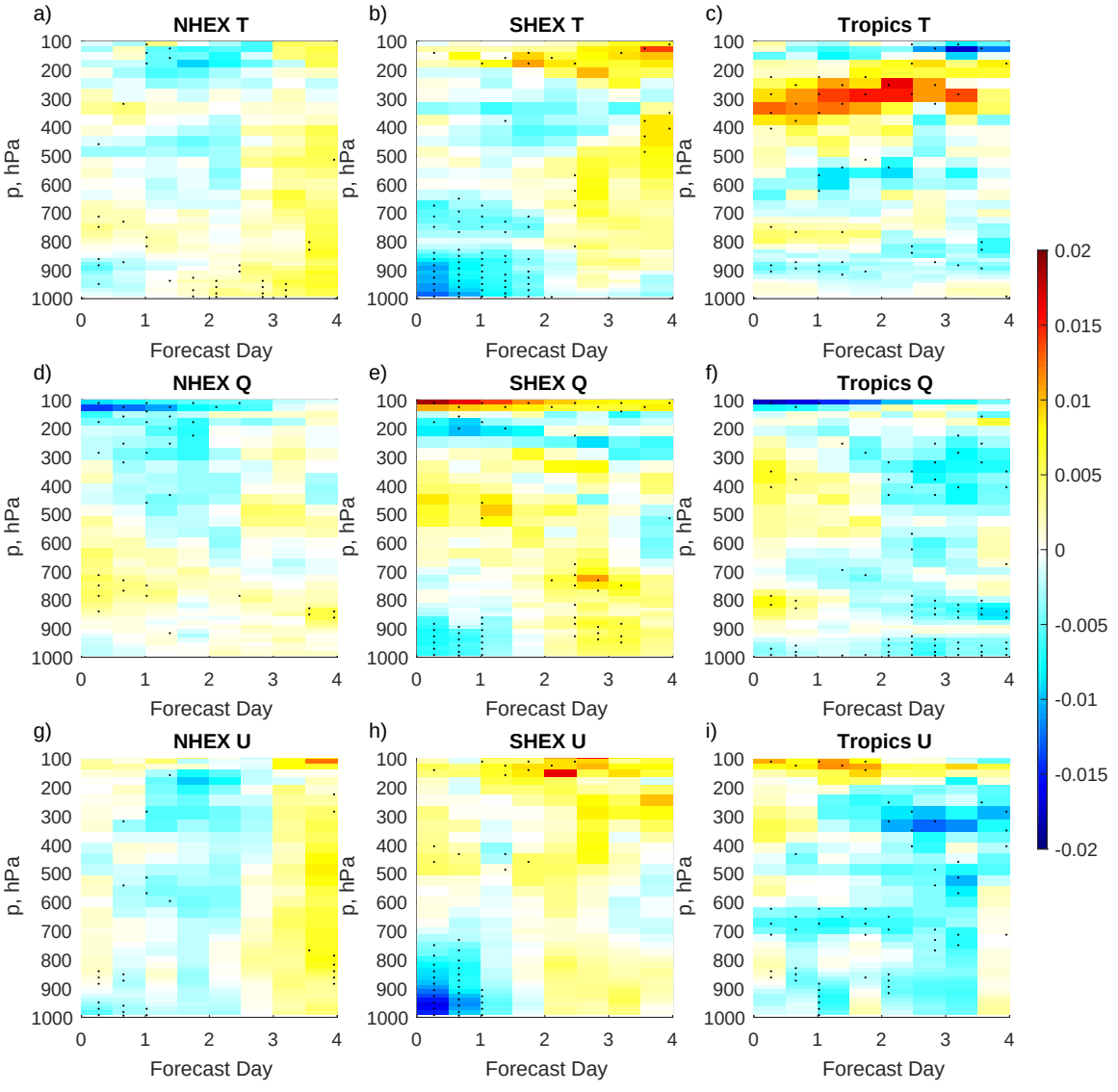
In the Tropics and NHEX regions, there is a slight improvement of temperature at 900 hPa during the initial forecast period, but otherwise changes to the forecast skill are minimal. The upper tropospheric degradation in tropical temperature is likely related to deep convective processes. As has been observed for other data types (Privé et al. 2022), adjustments to the analysis field in the tropics can lead to a short-lived spike in precipitation in the initial forecast period and resultant



582 FIG. 16. Fractional RTMS surface pressure analysis error change compared to Control error for cycling cases,
583 four times daily for the month of July. a) C250, b) C500, c) C1000, d) C2700.

598 heating of the upper troposphere that degrades the forecast skill in that region. However, for the
599 SHEX region, there is a more significant improvement to temperatures of up to 1-2% below 700
600 hPa and lasting for approximately 48 hours. Forecasts in the SHEX region may show greater
601 improvement compared to other regions as it is the winter season, with associated baroclinic lows,
602 and also because of the large fraction of surface area observed by SMSP.

603 The forecast skill in the SHEX region for specific humidity and zonal winds show similar impacts
604 as for temperature, with significant improvements in the lower troposphere out to 24 hours. Similar
605 to temperature, the wind and humidity forecast errors are not significantly affected by SMSP
606 observations in the NHEX and Tropics regions. There is some significant improvement of specific
607 humidity below 900 hPa and zonal wind below 700 hPa during the first 24 hour forecast period in
608 the SHEX region. These improvements in the SHEX region are not retained as long for humidity



590 FIG. 17. Fractional change in areal averaged RAMS forecast error for the C1000 case compared to Control as
591 a function of model level height equivalent for a,b,c) temperature; d,e,f) specific humidity, and g,h,i) zonal wind.
592 Stippling indicates 90th percentile statistical significance.

609 and wind as for temperature. Due to the use of normalized impacts, the humidity impacts near the
610 tropopause and lower stratosphere can be very noisy where humidities are low, as seen by the large
611 magnitude impacts near 100 hPa.

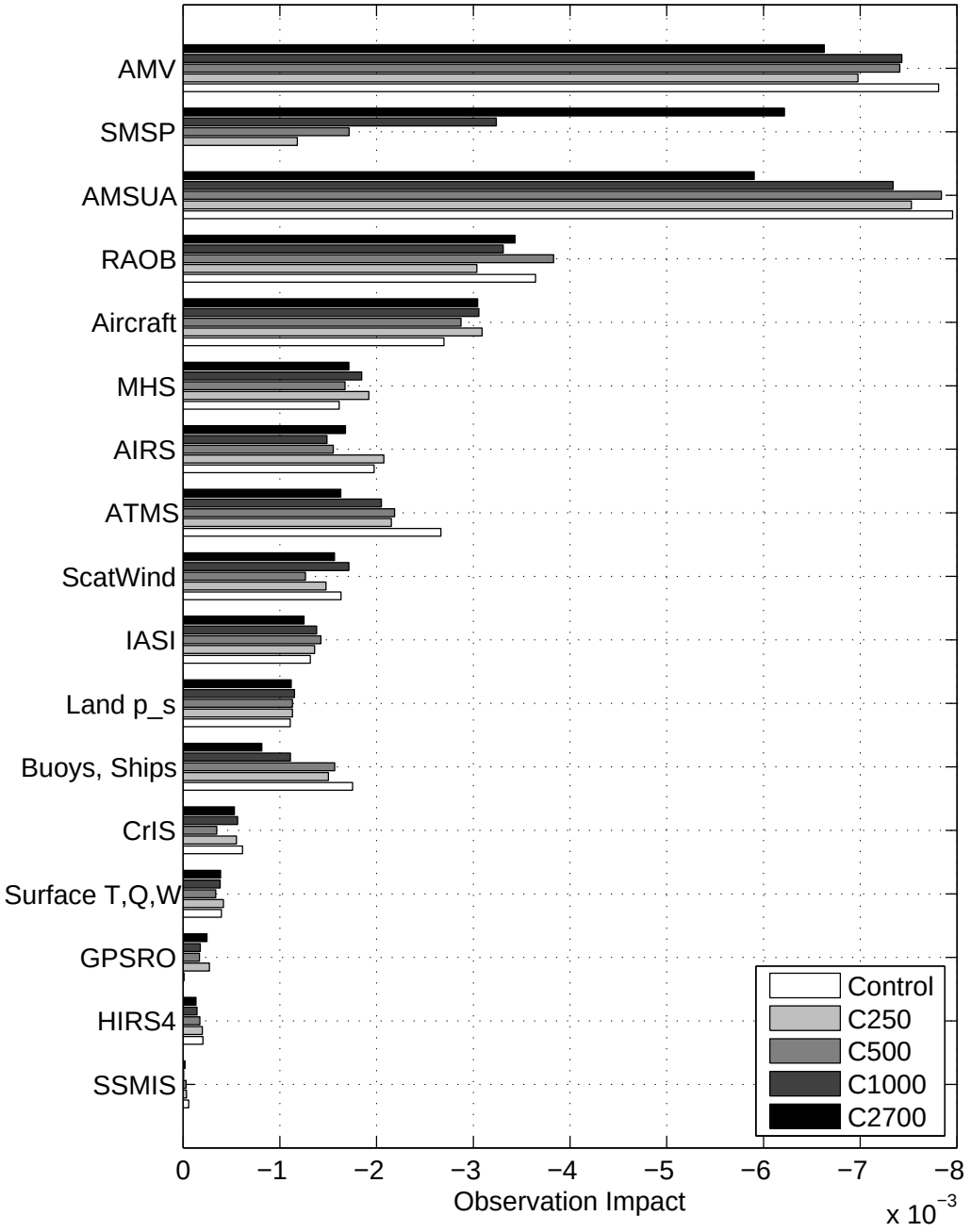
Due to the relatively small number of forecasts and the use of the 90th percentile for determining significance, many of the regions that appear to have significance in Figure 17 after the initial forecast period or in the upper troposphere are considered likely to be artifacts.

3) Forecast Sensitivity to Observation Impact

The GEOS/GSI system at the GMAO includes a forecast sensitivity to observation impact (FSOI) tool for estimating the contribution of observations to the reduction of a 24-hour forecast error norm. The FSOI uses a linearized version of the NWP model (including moist processes (Holdaway et al. 2014)) to separately calculate the impact of each ingested observation on the 24-hour forecast error, determined by comparing the forecast fields with the NR fields. The impacts of each data type are calculated by summing the individual impacts of each observation of that type. Because a linearized model is used for the FSOI calculations, only approximately 75% of the total impact is captured in these calculations (Privé et al. 2020). While the FSOI yields similar results to those found from data-denial experiments (Gelaro and Zhu 2009), it cannot be used to compare the relative forecast skill in different experiments and does not definitively inform as to redundancies between observing types.

For these experiments, the selected error norm is the RAMS temperature error between the surface and the 850 hPa model level over a box in the Pacific Ocean from 140E to 115E and 50S to 50N. This norm was chosen to represent a region that is overpassed by the SMSP swath for the 0000 UTC forecast cycle. The boundary layer temperature was selected as a norm as the lower troposphere shows a large forecast impact in Figure 17 that persists to and beyond 24 hours, unlike other variables such as wind or humidity for which significant impacts do not persist beyond 24 hours. Pairs of forecasts starting at 1800 UTC and the following 0000 UTC were used for the FSOI calculations. The OSSE FSOI performance has been validated against real observation FSOI (Privé et al. (2021), Privé et al. (2022)). While the net impacts in the OSSE are of smaller magnitude than for real observations due to insufficient model error in the OSSE, the relative impact of different observing types is well-preserved so that individual observing types can be placed into context with the rest of the global observing network.

Figure 18 compares the FSOI estimates of observation impacts for the C250, C500, C1000, and C2700 cases as well as Control. As expected, larger impacts are seen for increased scan width.



639 FIG. 18. FSOI estimates of observation impact on 24 hour boundary layer temperature error over the Pacific
640 Ocean for cycling cases: Control, C250, C500, C1000, and C2700.

With a scan width of 500 km or less, the SMSP impacts are similar in magnitude to the global network of ships and buoys, MHS, AIRS, or IASI. For the C1000 case, the SMSP data have an impact that is on par with rawinsondes and aircraft observations. The widest scan of C2700 has an impact similar in magnitude to that of AMSU-A. Notably, AMSU-A consists of six satellite platforms, and MHS consists of three platforms. There is some reduction in impact for buoys and ships as the SMSP scan width is increased - this may be due to a combination of redundancy with the SMSP data and less error growth during the initial forecast period. Compared to the Control FSOI, inclusion of the SMSP data resulted in decreased impacts for in situ marine surface observations and ATMS. The particular decrease in ATMS impact is likely due to the use of the same orbit for the SMSP and ATMS, allowing for greater chance of redundancy of the observations compared to other data types.

4. Discussion

These experiments are a preliminary examination of the potential impact of SMSP observations. There are several important limitations of these experiments. First, the SMSP observations have no simulated observation errors and as such are expected to be 'too good' relative to real SMSP data. The expected error characteristics of SMSP is currently unknown, but the development and demonstrations of the MBARS instrument should provide a better understanding of potential error magnitudes and correlations. Previous work (Privé et al. (2013a), Privé et al. (2021)) has shown that including simulated observation errors can result in small but statistically significant changes to observation impacts on the forecast skill and FSOI metrics, particularly for correlated errors. Given the magnitude of the impacts seen for simulated SMSP observations, it is likely that unless SMSP errors are substantially larger than expected, the SMSP data would still have a significant beneficial impact on the marine surface pressure field if observation errors were included.

A second limitation is that the SMSP observations are simulated and assimilated based on in-situ marine surface pressure types rather than having the characteristics of the remote-sensing instrumentation. For example, ships and buoys are point measurements, while spaceborne surface pressure observations would have a considerably larger footprint, likely on the order of 50 km in diameter.

Third, these initial experiments do not consider the effects of rain contamination on the SMSP observations. The observations could reasonably be expected to be contaminated by rain rates of greater than 1-2 mm/hr based on the instrument design. Initial testing to simulate rain contamination using the Nature Run rain fields shows that fewer than 5% of SMSP observations would be contaminated by rain. The location of rain-contaminated observations is geographically widespread, with contamination generally occurring only in localized areas. Owing to the distribution of rain contamination it is expected that the geographic distribution of overall observation impacts would be similar, with slightly weaker impact magnitudes. However, rain contamination would likely occur near certain types of phenomena of interest such as baroclinic lows and tropical cyclones. Depending on the spatial extent of the areas of heavy rainfall, this could reduce the beneficial impact of SMSP observations near low pressure systems such as the case study described in this work.

These experiments were performed only for the month of July. It would be reasonable to expect some regional difference in observation impacts if a different season, such as boreal winter, were tested. While the predictability of the Tropics does not have a strong annual cycle, the extratropics do vary in predictability between the baroclinically-dominated winter and the more convectively-dominated summer periods. In these experiments, the primary impacts of the SMSP data are observed in the boundary layer during the initial forecast period, with some large impacts along the Antarctic sea ice edge. As the sea ice retreats in austral summer, the observation impacts near the ice edge might be expected to change.

One question is whether the dry mass surface pressure bias between the experiment Control and the NR is realistic, and if not, whether this affects the performance of the SMSP data. The total mass of the atmosphere is constrained by the DAS to a set value. This value is changed in operations when a new orographic dataset is used, but there is a degree of uncertainty as to the most appropriate value. This uncertainty is true both for the real world and for the OSSE. Estimates of the total mass of the real atmosphere have been made using reanalyses (Trenberth and Smith 2005), but reanalyses also have uncertainties in orography and mass constraints. The magnitude of recent changes to the GEOS DAS dry mass constraint value have been on the order of 0.2 hPa, which is in line with the magnitude of the global difference in mean surface pressure between the Control and the NR. Therefore, it is expected that the dry mass bias seen in the Control is not necessarily

unrealistic, and that some of the mass redistribution effects that occur in the SMSP OSSE might also be seen with real SMSP data if there were a bias of similar magnitude between the operational DAS/forward model and the real world mass.

The use of the dry mass constraint clearly affects the impact of the SMSP observations. While regions overpassed by the SMSP orbit see beneficial impacts from the radar observations, regions that are not observed by the SMSP can have degraded surface pressures in the analysis. Poorly observed marine regions such as the southern ocean are strongly affected by degradation from the dry mass constraint. Although land and ice-covered regions are not shown for these experiments due to noisy errors from regridding over orography, low-lying regions of Antarctica and the ice-covered southern ocean are also subject to substantial degradations of the surface pressure field. Degradation over poorly observed areas of Africa and Greenland is minimal, possibly due to the presence of at least some in-situ observations such as aircraft. These experiment results bring into question the value of the dry mass constraint for the DAS, and suggest that the use of the dry mass constraint be reconsidered in future DAS development.

The results of these experiments show that spaceborne marine surface pressure data can have a beneficial impact on not just the surface pressure analysis, but also boundary layer mass and wind fields. It is also clear from these results that further development of the data assimilation system to make better use of surface pressure data would be advantageous in order to extend the beneficial impacts throughout the depth of the troposphere and thus allow retention of improvements further into the medium range forecast period. While the global observing network would be expected to change between the versions used for this study and the versions when the instrument might be launched in a future decade, there is minimal overlap between the SMSP and other spaceborne platform data, due to the unique nature of the observations. Thus the observation impacts should be relatively unaffected by changes to the global observing network as there is little chance that new instruments would be redundant with the SMSP.

The thinned scan experiments are intended to compare more realistic observation distributions to the more idealised unthinned scanning configurations. The SMSP instrument could realistically be expected to have a minimum 50 km across-scan spacing. It has been found that very widely spaced surface pressure observations with wide scanning are less beneficial than more densely spaced observations with a narrower scan width. For the configurations tested here, the 500 km or

1000 km width scans with 30-90 km spacing between observations had the best improvement of the thinned scan experiments, with less benefit for scans with greater than 100 km spacing between observations, although this result may be at least somewhat dependent on the GSI treatment of surface pressure observations.

a. Future Work

Additional OSSE experiments are planned for the SMSP instrument as part of the MBARS project, exploring different potential orbits including constellation configurations, expected field of view, and scan widths. Error-added SMSP observations will also be tested. The experiment period will be extended to evaluate the impact of SMSP observations on tropical cyclone prediction.

Several improvements to the OSSE framework are expected to produce more robust testing of the SMSP instrument. First, the GMAO OSSE has recently been upgraded to a newer version of both the GSI and the GEOS model, including boundary layer physics that should provide more realistic differences between the G5NR and the OSSE forecast model. This new system has a smaller global dry mass surface pressure bias compared to the G5NR, but has regional surface pressure differences of similar magnitude to the current system. It is hoped that the new system will have boundary layer errors that more faithfully represent the characteristics of operational NWP errors, whereas the boundary layer physics in the current OSSE framework is very similar to the NR, so that boundary layer errors may be underestimated.

One powerful capability of the OSSE is the ability to test the behavior of the DAS in a situation where the analysis and background errors can be directly calculated. Development and testing of an observation operator for the SMSP instrument in the OSSE framework can both aid in further OSSE impact studies and also expedite the eventual use of real SMSP data in operational NWP.

Acknowledgments. Support for this project was provided by NASA Earth Science Technology Office (ESTO) and Goddard Space Flight Center (GSFC). Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at GSFC. The software for simulating GPSRO observations was provided by the Radio Occultation Processing Package (ROPP) of the Radio Occultation Meteorology (ROM) Satellite Applications Facility (SAF) of EUMETSAT, with the assistance of Sean Healy at ECMWF. Comments from two anonymous reviewers led to substantial improvements to this manuscript.

Data availability statement. The dataset on which this paper is based is too large to be retained or publicly archived with available resources. Documentation and methods used to support this study are available from Nikki Privé at NASA/GMAO.

References

Arnold, C. J., and C. Dey, 1986: Observing-system simulation experiments: past, present, and future. *Bull. of the Amer. Met. Soc.*, **67**, 687–695.

Barahona, D., A. Molod, J. Bacmeister, A. Nenes, A. Gettelman, H. Morrison, V. Phillips, and A. Eichmann, 2014: Development of two-moment cloud microphysics for liquid and ice within the NASA Goddard Earth Observing System Model (GEOS-5). *Geoscience Model Development*, **7**, 1733–1766, doi: 10.5194/gmd-7-1733-2014.

Centurion, L., A. Horányi, C. Cardinali, E. Charpentier, and R. Lumpkin, 2017: A global ocean observing system for measuring sea level atmospheric pressure: effects and impacts on numerical weather prediction. *Bull. of the Amer. Met. Soc.*, **98**, 231–238, doi:10.1175/BAMS-D-15-00080.1.

Culverwell, I., H. Lewis, D. Offiler, C. Marquardt, and C. Burrows, 2015: The radio occultation processing package, ROPP. *Atmos. Meas. Tech.*, **8**, 1887–1899, https://doi.org/10.5194/amt-8-1887-2015.

El Akkraoui, A., N. Privé, R. Errico, and R. Todling, 2023: The GMAO hybrid 4D-EnVar observing system simulation experiment framework. *Mon. Wea. Rev.*, in press. doi: 10.1175/MWR-D-22-0254.1.

Errico, R., and Coauthors, 2017: Description of the GMAO OSSE for Weather Analysis software package: Version 3. Tech. Rep. 48, National Aeronautics and Space Administration, 156 pp. NASA/TM-2017-104606.

Errico, R. M., D. Carvalho, N. C. Privé, and M. Sienkiewicz, 2020: Simulation of atmospheric motion vectors for an observing system simulation experiment. *J. Atmos. Ocean Tech.*, **37**, 489–505, doi:10.1175/JTECH-D-19-0079.1.

786 Errico, R. M., and N. C. Privé, 2018: Some general and fundamental requirements for designing
787 observing system simulation experiments (osses). *WMO Rep. WWRP 2018-8*, 33 pp., URL
788 https://www.wmo.int/pages/prog/arep/wwrp/new/documents/Final_WWRP_2018_8.pdf.

789 Errico, R. M., R. Yang, N. Privé, K.-S. Tai, R. Todling, M. Sienkiewicz, and J. Guo, 2013: Vali-
790 dation of version one of the Observing System Simulation Experiments at the Global Modeling
791 and Assimilation Office. *Quart. J. Roy. Meteor. Soc.*, **139**, 1162–1178, doi: 10.1002/qj.2027.

792 Gelaro, R., and Y. Zhu, 2009: Examination of observation impacts derived from observing system
793 experiments (OSEs) and adjoint models. *Tellus*, **61A**, 179–193.

794 Gelaro, R., and Coauthors, 2014: Evaluation of the 7-km GEOS-5 nature run. NASA/TM–2014-
795 104606, 36, NASA.

796 Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Ap-
797 plications, Version-2 (MERRA-2). *J. Climate*, **30**, 5419–5454, doi:10.1175/JCLI-D-16-0758.1.

798 Han, Y., P. van Delst, Q. Liu, F. Weng, B. Yan, R. Treadon, and J. Derber, 2006: JCSDA Community
799 Radiative Transfer Model (CRTM) - version 1. OAA Tech. Report 122.

800 Harnisch, F., S. Healy, P. Bauer, and S. English, 2013: Scaling of GNSS radio occultation impact
801 with observation number using an ensemble of data assimilations. *Mon. Wea. Rev.*, 4395–4413,
802 doi;10.1175/MWR-D-13-00098.1.

803 Healy, S., 2013: Surface pressure information retrieved from GPS radio occultation measurements.
804 *Quart. J. Roy. Meteor. Soc.*, **139**, 2108–2118.

805 Hoffman, R., and R. Atlas, 2016: Future observing system simulation experiments. *Bull. of the*
806 *Amer. Met. Soc.*, **97**, 1601–1616, doi:10.1175/BAMS-D-15-00200.1.

807 Holdaway, D., R. Errico, R. Gelaro, and J. Kim, 2014: Inclusion of linearized moist physics
808 in NASA's Goddard Earth Observing System data assimilation tools. *Mon. Wea. Rev.*, **142**,
809 414–433.

810 Horányi, A., C. Cardinali, and L. Centurion, 2017: The global numerical weather prediction impact
811 of mean sea level pressure observations from drifting buoys. *Quart. J. Roy. Meteor. Soc.*, **143**,
812 974–985, doi:10.1002/qj.2981.

Ingleby, B., 2010: Factors affecting ship and buoy data quality: a data assimilation perspective. *J. Atmos. Ocean Tech.*, **27**, 1476–1489.

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

Kent, E., and D. Berry, 2005: Quantifying random errors in voluntary observing ships meteorological observations. *Intl. J. Climatol.*, **25**, 843–856.

Lawrence, R., B. Lin, S. Harrah, Y. Hu, P. Hunt, and C. Lipp, 2011: Initial flight test results of differential absorption barometric radar for remote sensing of sea surface air pressure. *J. Quant. Spectrosc. Ra.*, **12**, 247–253.

Lin, B., and Y. Hu, 2005: Numerical simulations of radar surface air pressure measurements at O2 bands. *IEEE T. Geosci. Remote*, **2**, 324–328.

Lin, B., and W. Rossow, 1994: Observations of cloud liquid water path over oceans: Optical and microwave remote sensing methods. *J. Geophys. Res.*, **99**, 20 907–20 927.

Lin, B., and W. Rossow, 1997: Precipitation water path and rainfall rate estimates over oceans using SSM/I and ISCCP data. *J. Geophys. Res.*, **102**, 9359–9374.

Lin, B., and Coauthors, 2023: Simulations of sea surface reflection for V-band O2 differential absorption radar barometry. *Frontiers in Remote Sensing*, **14**, doi:10.3389/frsen.2023.1105627.

McCarty, W., D. Carvalho, I. Moradi, and N. C. Prive, 2021: Observing system simulation experiments investigating atmospheric motion vectors and radiances from a constellation of 4-5 micro-m infrared sounders. *J. Atmos. Ocean Tech.*, **38**, 331–347, 10.1175/JTECH-D-20-0109.1.

Mears, C. A., D. Smith, L. Ricciardulli, J. Wang, H. Huelsing, and F. Wentz, 2018: Construction and uncertainty estimation of a satellite-derived total precipitable water data record over the world's oceans. *Earth and Space Science*, **5**, 197–210, doi.org/10.1002/2018EA000363.

Minnis, P., J. Huang, B. Lin, Y. Yi, R. Arduini, T.-F. Fan, J. Ayers, and G. Mace, 2007: Ice cloud properties in ice-over-water cloud systems using TRMM VIRS and TMI data. *J. Geophys. Res.*, **112**, doi:10.1029/2006JD007626.

Privé, N., and R. Errico, 2019: Uncertainty of observation impact estimation in an adjoint model investigated with an observing system simulation experiment. *Mon. Wea. Rev.*, **147**, 3191–3204, doi: 10.1175/MWR-D-19-0097.1.

Privé, N., R. Errico, and K.-S. Tai, 2013a: The influence of observation errors on analysis error and forecast skill investigated with an observing system simulation experiment. *J. Geophys. Res.*, **118**, 5332–5346, doi: 10.1002/jgrd.50452.

Privé, N., R. Errico, and K.-S. Tai, 2013b: Validation of forecast skill of the Global Modeling and Assimilation Office observing system simulation experiment. *Quart. J. Roy. Meteor. Soc.*, **139**, 1354–1363, doi: 10.1002/qj.2029.

Privé, N., R. Errico, and K.-S. Tai, 2014: The impact of increased frequency of rawinsonde observations on forecast skill investigated with an observing system simulation experiment. *Mon. Wea. Rev.*, **142**, 1823–1834, doi: 10.1175/MWR-D-13-00237.1.

Privé, N. C., R. M. Errico, and A. E. Akkraoui, 2022: Investigation of the potential saturation of information from global navigation satellite system radio occultation observations with an observing system simulation experiment. *Mon. Wea. Rev.*, **150**, 1293–1316, doi:10.1175/mwr-d-20-0256.1.

Privé, N. C., R. M. Errico, and W. McCarty, 2021: The importance of simulated errors in observing system simulation experiments. *Tellus A*, **73**, 1–117, doi:10.1080/16000870.2021.1886795.

Privé, N. C., R. M. Errico, R. Todling, and A. E. Akkraoui, 2020: Evaluation of adjoint-based observation impacts as a function of forecast length using an observing system simulation experiment. *Quart. J. Roy. Meteor. Soc.*, **147**, 121–138, doi:10.1002/qj.3909.

Radnóti, G., P. Bauer, A. McNally, and A. Horányi, 2012: ECMWF study to quantify the interaction between terrestrial and space-based observing systems on numerical weather prediction skill. Tech. Rep. ECMWF Technical Memorandum 679.

Salby, M., 1995: *Fundamentals of atmospheric physics*. Acedemic Press, doi: 10.1016/s0074-6142(96)x8037-9.

Takacs, L., M. Suarez, and R. Todling, 2015: Maintaining atmospheric mass and water balance within reanalysis. Tech. Rep. 37, NASA. NASA/TM-2014-104606.

867 Tan, D., E. Andersson, M. Fisher, and L. Isaksen, 2007: Observing-system impact assessment
868 using a data assimilation ensemble technique: application to the adm–aeolus wind profiling
869 mission. *Quart. J. Roy. Meteor. Soc.*, **133**, 381–390, doi:10.1002/qj.43.

870 Trenberth, K., and L. Smith, 2005: The mass of the atmosphere: a constraint on global analyses.
871 *J. Climate*, **18**, 864–875.