Optimal Network Design Applied to Monitoring and Forecasting Surface Temperature in Antarctica

GREGORY J. HAKIM

Department of Atmospheric Sciences, University of Washington, Seattle, Washington

KARIN A. BUMBACO

Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, Washington

ROBERT TARDIF

Department of Atmospheric Sciences, University of Washington, Seattle, Washington

JORDAN G. POWERS

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 6 April 2019, in final form 11 November 2019)

ABSTRACT

As harsh weather conditions in Antarctica make it difficult to support a dense weather observing network there, it is critical to place new weather stations in locations that are optimal for a given monitoring goal. Here we demonstrate a network design algorithm that uses ensemble sensitivity to identify optimal locations for new automatic weather stations in Antarctica. We define the optimal location as one that maximizes the reduction in total variance of a given spatial field. Using WRF Model forecast output from the Antarctic Mesoscale Prediction System (AMPS), we identify the best locations for observations across the continent by considering two spatial fields: (i) the daily 0000 UTC 2-m temperature analysis field and (ii) the daily 0000 UTC 2-m air temperature 24-h forecast field. We explore the impact of spatial localization on the results, finding that a covariance length scale of 3000 km is appropriate for these metrics. We find optimal locations assuming that no stations exist on the continent (blank slate) and conditional on existing stations (CD90). In the “blank slate” scenario, the Megadunes region emerges as the most important location to both monitor temperature and reduce temperature forecast errors, with the Ronne Coast and the Siple Coast following. Results for the monitoring and forecasting metrics are similar for the CD90 subset as well, indicating that additional stations could benefit multiple performance goals. Considering the CD90 subset, Wilkes Land–Adelie Coast, Ellsworth Land, and Queen Maud Land–Interior are identified as regions to consider installing new stations for optimizing network performance.

1. Introduction

The remote environment and extreme conditions on the Antarctic continent, home to a range of vital Earth science research and field work, make travel and logistics for the science and its support challenging. An important part of maintaining safety and supporting operations is having a strong observing network to identify hazardous weather and provide information for weather forecasts. The latter is necessary for both data assimilation to improve numerical weather prediction (NWP) and input for operational meteorologists’ forecasts. Short- to medium-range forecasts, for example, are critical in flight planning and guidance for air and ground operations, all of which impact project schedules and outcomes (e.g., Powers et al. 2003, 2012). A majority of the current weather stations over Antarctica are maintained by the University of Wisconsin’s Antarctic Meteorological Research Center, but even when combined with observing sites from other countries, the monitoring network is sparse given the logistical needs and challenges (Lazzara et al. 2012).

Previous studies have confirmed that increasing the number of Antarctic observations would benefit
For both fixed and targeted optimal network design, a variety of ensemble Kalman filter (EnKF) approaches have been used, and the actual impact of observations may differ from the optimal prediction due to a variety of factors, including differences in the assimilation scheme (e.g., Langland 2005). The general approach in ensemble-sampling methods is to use ensemble forecasts to calculate model covariances and estimate the impact of new observations (either fixed or targeted) on a forecast metric without using an adjoint model. Ensemble methods are useful when an adjoint version of the model is not available and for effectively assessing many performance goals. One ensemble method, the Retrospective Design Algorithm (RDA) described in Khare and Anderson (2006), tests the impact of sets of fixed networks by maximizing an objective function of covariance matrices that have been updated to reflect the potential new networks. Bishop et al. (2001) and Majumdar et al. (2002b) developed an ensemble transform Kalman filter (ETKF) approach to find optimal observations. The ETKF approach has been used to design both adaptive and fixed observational networks in both simple model and fieldwork implementations. The ETKF approach also uses archived forecasts to find optimal observations for any desired forecast lead time, and it does so by transforming the forecast error covariance matrix within an EnKF framework into an orthonormal vector space to quickly solve for optimal locations. Many subsequent studies have successfully used the ETKF for optimizing and developing a wide variety of networks (e.g., Majumdar et al. 2002a; Aberson 2003; Xue et al. 2006). Similar to the ETKF approach is the ensemble sensitivity approach developed in Ancell and Hakim (2007), which we use for our experiments in the current study. The ensemble sensitivity approach does not transform covariance matrices, and uses the sensitivity of state variables to a generic metric function (which can involve nonstate variables) to find optimal measurements. The ensemble sensitivity approach has been tested in theoretical modeling for time-averaged observations (Huntley and Hakim 2010), used for an example climate observing network for the Pacific Northwest (Mauger et al. 2013), shown as an example for continent-averaged surface temperature monitoring in Antarctica (Bumbaco et al. 2014), and applied to find salinity and sea surface paleoclimate estimates from corals (Comboul et al. 2015).

The optimal network design approach here uses an ensemble of existing deterministic model analyses and forecasts (not output from forecast ensembles) to find the best locations for stations by identifying the location
that reduces the climatological variance the most in the chosen metric. Once that location is chosen, the statistics for the entire metric field are updated using the Kalman update equation, as described below, to reflect the new hypothetical measurement, and the process then repeats. This approach allows the optimal network to be found sequentially without needing to run a model, and the flexible methodology can also address other concerns, such as excluding locations from the optimal design calculation (e.g., mountainous terrain) that may not feasibly support a weather station. We apply spatial localization to address sample error, which otherwise removes too much ensemble spread, affecting the choice of subsequent station locations.

The outline of the remainder of this paper is as follows. Section 2 describes the data used, the theoretical underpinnings of the optimal network design approach, and reviews an assessment of the current Antarctic surface observing network. In addition to the design algorithm, we discuss the implementation and statistical approach to account for uncertainty in the calculations. The results of three optimal network calculations and the impact of spatial localization are presented in section 3. The first is a “blank slate” calculation, in which we assume no weather stations exist on the continent. The second and third calculations are a set of augmented networks conditional on two different configurations of existing stations. These three optimal network design calculations are performed with both the monitoring and forecasting metrics, while the impact of using various localization length scales is explored with the monitoring metric. Finally, section 4 provides a concluding discussion of the results and suggestions for further work.

2. Data and methods

a. Antarctic mesoscale prediction system

The primary data used in this study are output from the Advanced Research version of the Weather Research and Forecasting (WRF) Model (Skamarock et al. 2008) from the Antarctic Mesoscale Prediction System (AMPS) (Powers et al. 2012). AMPS is a real-time numerical weather prediction system run by the National Center for Atmospheric Research in support of the weather forecasting needs of the U.S. Antarctic Program (USAP) and USAP’s scientific and logistical operations. For the period of the model output used here, AMPS ran a polar-modified version of WRF, and the forecasts were initialized using the WRF data assimilation (WRFDA) system (Barker et al. 2004; Huang et al. 2009) employing a three-dimensional variational (3DVAR) approach. AMPS ran a WRF nested-domain configuration over Antarctica and the Southern Ocean (Powers et al. 2012), and the model output used in this study reflects a continental domain with 15-km horizontal grid spacing, nested within an outer 45-km domain (Fig. 1).

Forecasts from AMPS have been verified and evaluated in a number of published studies (Powers et al. 2003; Bromwich et al. 2005; Monaghan et al. 2005; Lazzara et al. 2012; Bromwich et al. 2013), and thus the model output used here is regarded as sufficiently reliable. Furthermore, most relevant to the applicability of network design, the pointwise statistics of the AMPS output and observations have been shown to be comparable. For example, as a precursor to this study, Bumbaco et al. (2014) found good agreement for point correlations between station observations and AMPS output (both the analyses and forecasts) for 2-m temperature and surface pressure, lending confidence that AMPS output can be used for optimal network design. Here, we use 2-m air temperature field output from the 15-km AMPS grid over the period 31 September 2008–1 October 2012. Data beyond October 2012 were not used due to a change in the WRF grid spacings used in the AMPS domains. All data reflect the 0000 UTC analyses, and forecasts initialized at 0000 UTC with a 24-h lead time are considered. We note that by focusing on the 0000 UTC data, we are ignoring diurnal temperature variations over the continent and only optimizing for the 0000 UTC snapshot, both for the sake of simplicity.

b. Antarctic surface observing network

The current Antarctic surface network consists of automatic weather stations (AWSs), a majority of which are maintained by the University of Wisconsin’s Antarctic Meteorological Research Center, but includes additional stations maintained by other efforts and countries, as well as staffed bases. The network provides crucial observational data over the entire continent that are assimilated into AMPS WRF forecasts. In addition to real-time observations, the stations have provided historical data for many Antarctic weather and climate studies (e.g., Lazzara et al. 2012), and the network has primarily grown to serve the purposes of the USAP and the research efforts of the National Science Foundation. While the network is important for furthering these goals, both existing and potential new stations may also be valuable for the analysis of specific phenomena, forecasting for a specific area of operations, or various other monitoring purposes. We consider the locations of weather stations that are assimilated into WRF as defined in Bumbaco et al. (2014). Namely, we use the locations of the weather stations that reported temperature
(i) on at least 90% of the days (defined as “CD90” stations for 90% “complete data”) and (ii) on at least 75% of the days (defined as “CD75” stations for 75% “complete data”) over the 2008–12 period. The 18 (45) stations in the CD90 (CD75) subset are assumed to represent the existing temperature network over the continent (Bumbaco et al. 2014).

c. Network design theory

We take an ensemble sensitivity approach to find optimal station locations. To define locations, a performance measure is needed (e.g., Rabier et al. 1996) to define the optimization. Typically a scalar measure is chosen, and evaluated for new hypothetical measurements; the location that has the greatest impact on the measure is the one chosen (e.g., Ancell and Hakim 2007). To determine a second location, the impact of the first new measurement must be taken into account. Continuing in this way, one can proceed to determine an ordered list of new hypothetical measurements that optimally improve performance as measured by the metric. Some applications require optimization for more than one monitoring goal, such as, for example, temperature and wind speed. While such applications can be accommodated with the theory for a scalar, it is helpful to have an approach that explicitly introduces more than one performance measure. Here we generalize the scalar approach to a vector having an arbitrary number of measures to be optimized simultaneously. We then solve the problem using an ensemble technique for the placement of Antarctic AWS stations.

The multivariate ensemble-based approach used here is similar to the ETKF method described in Bishop et al. (2001) and Majumdar et al. (2002b). Our approach differs from the ETKF method in that it can be used to optimize for an arbitrary number of general functions of the state and/or nonstate variables (e.g., model parameters) and in the solution method, which involves no matrix computations, and the option for smooth spatial covariance localization functions. We begin by defining the vector containing measures for the individual monitoring criteria:

![Diagram of the WRF Model grids in AMPS, including the 15-km grid used here. Terrain height is shaded.]
where \( \mathbf{x} \) is the state vector of dimension \( m \), \( \delta \mathbf{x} \) is the perturbation state vector about the ensemble mean, and superscript \( T \) represents the transpose operation. We note that the metric sensitivity, \( \left[ \frac{\partial \mathbf{J}}{\partial \mathbf{x}} \right] \), is an \( n \times m \) matrix, which compares with the \( n \times 1 \) vector sensitivity in Ancell and Hakim (2007).

The covariance of \( \mathbf{J} \) about the mean value, is then approximated by

\[
\Sigma^2 = \{ \delta \mathbf{J} \delta \mathbf{J}^T \} \approx \left[ \frac{\partial \mathbf{J}}{\partial \mathbf{x}} \right]^T \mathbf{A} \left[ \frac{\partial \mathbf{J}}{\partial \mathbf{x}} \right] , \quad \text{where} \quad \mathbf{A} = \{ \delta \mathbf{x} \delta \mathbf{x}^T \}
\]

is the state covariance matrix and the curly braces denote an expectation operation. We note that \( \Sigma^2 \) is an \( n \times n \) matrix representing the generalization of the scalar version \( \sigma_j^2 \) in Ancell and Hakim (2007).

When a new observation is incorporated, the state covariance \( \mathbf{A} \) changes to \( \mathbf{A}' \), and the metric covariance matrix also changes. Therefore, from (3), we have that the change in the metric covariance is

\[
\delta \Sigma^2 = \left[ \frac{\partial \mathbf{J}}{\partial \mathbf{x}} \right]^T (\mathbf{A}' - \mathbf{A}) \left[ \frac{\partial \mathbf{J}}{\partial \mathbf{x}} \right].
\]

The optimal location is the one that maximizes a measure of \( \delta \Sigma^2 \), but a practical solution of (4) is enabled by two additional steps. First, consistent with the linear and Gaussian approximations, the change in the state covariance is given by the update step in the Kalman filter as

\[
\mathbf{A}' - \mathbf{A} = -\mathbf{K} \mathbf{H} \mathbf{A}, \quad \text{where} \quad \mathbf{K} = \mathbf{A} \mathbf{H}^T (\mathbf{H} \mathbf{A}^T + \mathbf{R})^{-1}
\]

is the Kalman gain, and \( \mathbf{H} \) maps from the state space to the measurement space. Using (5) and (6) in (4) gives

\[
\delta \Sigma^2 = - \left[ \frac{\partial \mathbf{J}}{\partial \mathbf{x}} \right]^T \mathbf{A} \mathbf{H}^T [\mathbf{H} \mathbf{A}^T + \mathbf{R}]^{-1} \mathbf{H} \left[ \frac{\partial \mathbf{J}}{\partial \mathbf{x}} \right].
\]
with the traditional scalar approach based on a state vector, but in any case, it is straightforward with the method described here. In the second case, where the optimization aims to constrain both the analysis and the 24-h forecast fields, the elements of $J$ will contain gridpoint values over the continent for both fields. Again, as we describe above, to rank order a list for a decision, a scalar measure is needed by any approach, and the trace of $\delta \Sigma^2$ serves that purpose here. In the event that the elements of $J$ do not share units, they may be standardized in a way that preserves spatial variability (e.g., by dividing each field by a constant value). Alternatively, one may transform the fields to a unifying variable, if appropriate, that links all aspects of interest into a single quantity, such as energy.

Here we solve the optimization problem using an ensemble sampling approach, where covariance expectations are estimated from finite samples. The state and metric perturbations are updated using the square root form of the ensemble Kalman filter (Whitaker and Hamill 2002),

$$\delta \mathbf{x}^a = \delta \mathbf{x}^p - \mathbf{K} \delta \mathbf{x}^e$$ (11)

where the superscript $a$ denotes the value after assimilating a new observation, and $p$ the value prior to assimilation. The modified Kalman gain $\mathbf{K}$ is given by

$$\mathbf{K} = a \mathbf{K} = \alpha \mathbf{A}^T \mathbf{H} \mathbf{E}^{-1}, \quad \text{where}$$

$$\alpha = \left(1 + \frac{r}{\mathbf{E}}\right)^{-1},$$

and $r$ is the error variance associated with the new observation.

Once the state and metric are updated to reflect the new observation, the process repeats until the described number of stations is reached. This procedure ensures that each station, with the exception of the first, is conditional on previous optimal stations.

To summarize, the general procedure for finding optimal locations uses the following algorithm:

1) Choose a vector metric that quantifies the desired aspects of the system of interest.
2) Calculate the total variance for the metric (trace of the metric covariance matrix).
3) Calculate the change in the total variance for all possible stations.
4) Choose as the optimal measurement location the one that maximizes the change in total variance from (10).
5) Update the metric and state to reflect the chosen measurement using (11).
6) Repeat the procedure for the next measurement, using the updated state and metric to find the next location conditional on the previous measurement (i.e., repeat steps 3–5).

d. Ensemble sampling, performance measures, and observing strategies

The algorithm for finding optimal observing locations depends upon an ensemble sample estimate for evaluating the selection algorithm. We randomly draw a 250-member ensemble sample from the 0000 UTC AMPS 2-m temperature analysis field during 31 September 2008–1 October 2012 to define the state and metric perturbations (i.e., $\delta \mathbf{x}$ and $\delta \mathbf{J}$, respectively).

We tested a variety of ensembles sizes, from 30 members to 1000 members, and 250-member ensembles gave results that were quantitatively consistent with the largest ensembles. Smaller ensembles did not produce results that were clustered geographically, resembling instead a random selection of observations. An ensemble size of 250 was also close to the computational rank of the temperature covariance matrix (when every grid point across the entire data period was used to calculate the covariance matrix). Hence, a smaller ensemble size might underestimate covariances, while using an ensemble larger than 250 does not provide enough additional information to justify the computational cost of using a very large ensemble.

To account for uncertainty in the network calculation, we use a Monte Carlo bootstrap approach where we repeat the calculation for a different, randomly drawn, 250-member ensemble. For each 250-member ensemble, the 20 locations for the idealized network are found as described above. This process is repeated 10 000 times, providing 10 000 sets of 20 locations, allowing for statistics on station location.

Our vector metrics $J$ in this work are defined for two separate network-design objectives: (i) the model analysis 0000 UTC 2-m temperature and (ii) the 24-h 2-m forecast temperature errors. Forecast errors for the second objective are defined as the difference between AMPS forecasts and analyses (i.e., 0-h forecasts) valid at the same date and hour.

Network design experiments are conducted for “blank slate” and “conditional” observing strategies. For the blank slate observing strategy, we assume no weather stations exist on the continent. For the conditional observing strategy, we consider existing observation networks by performing additional calculations between steps 1 and 2 described in the previous section. These calculations consist of applying (11) at every observation location in the existing network and updating the state and metric vectors, reflecting the
cumulative impact of all of these observations. The resulting state and metric vectors then serve as input to remaining steps in the network design calculations (i.e., steps 2–6).

To reduce the size of the calculations, we subsample the AMPS grid from the original resolution. Specifically, the state vector $\mathbf{x}$ consists of every fifth model grid point on the 15 km grid, which defines the set of possible observation locations (blue and gray dots on Fig. 2); there are 2446 such locations. We will refer to these potential observation locations as “stations” in the remainder of this paper. A further subsampling is used in the evaluation of the metric $J$ where every twentieth grid point on the 15-km grid is chosen because it is the same distance (300 km) as the shortest correlation length scales over the continent (Bumbaco et al. 2014), and it makes the calculations computationally tractable (155 blue dots on Fig. 2).

All of the optimal network calculations are conducted for a total of 20 stations to meet the monitoring goal. This number was found to provide insight into the performance of the optimal design calculation beyond just a few stations, but not so many as to be computationally prohibitive. In real-world situations, the number of new stations will likely be affected by practical constraints, such as budgetary or logistical limitations.

3. Results for monitoring and forecasting networks

In this section we explore the results of designing optimal networks with and without spatial localization to identify the best locations for observations across the continent by considering two spatial fields: (i) daily 0000 UTC 2-m temperature (the “monitoring network”) and, (ii) the daily 0000 UTC 2-m air temperature 24-h forecast field (“forecast error reduction network”). For both cases, we develop networks assuming various configurations of existing stations on the continent. The first calculations are done assuming that no other stations currently exist (i.e., blank slate). For the second and third calculations—the augmented networks—the influence on the metric from the subsets of existing stations identified in Bumbaco et al. (2014) that report for 90% of the record (CD90) and report for 75% of the record (CD75) are taken into account before optimally placing 20 additional stations. The influence of the existing stations is removed as described in section 2c.

The “augmented network” experiments show how an optimal network that is conditional on the current network can be produced by identifying the best locations to add new stations. Figure 2 shows locations referenced subsequently.

Interpretation of results is aided by two quantities shown in Fig. 3 for 2 m air temperature: spatial correlation length scale, and variance. These quantities are computed from AMPS data during the time period under investigation using the method described in Bumbaco et al. (2014). Specifically, correlation length scale is defined as the $e$-folding distance of the squared autocorrelation at each point. Generally, length scales are longest over East Antarctica and shortest over the Antarctic Peninsula. Variance is largest near Wilkes Land–Knox Coast and the Ronne and Ross Ice Shelves. Favored locations for new stations are those having large variance and long length scales.

a. Monitoring network without localization

The first network example using the optimal design algorithm assumes no stations exist on the continent, that is, a blank slate calculation, and optimizes for the monitoring metric. Figure 4 shows the first three stations selected to reduce the temperature analysis uncertainty the most across the continent. The first station location chosen is in the Megadunes region in East Antarctica, and the pixels identifying this location are fairly concentrated in only that one area, lending confidence that this reduces the most variance (54%) in the temperature metric throughout the continent. The closest existing station geographically in the CD90 subset is Vostok, but the point chosen most often is within a large gap in the existing network in East Antarctica with long length scales (Fig. 3a), especially for austral summer...
Locations with long correlation length scales allow information from measurements to affect a large area, which yields a relatively large change in the total variance compared to locations with shorter correlation length scales. This same region was chosen as the first station in an example presented in Bumbaco et al. (2014) using the mean daily East Antarctica temperature as a metric, indicating that sampling this region could be important for various monitoring goals.

The second station, conditional on placement of a station in the Megadunes region, is chosen about 16% of the time over 10,000 simulations at a single grid point, and is on the Siple Coast in Marie Byrd Land in West Antarctica. For placement of a third station, three different areas are frequently chosen: Queen Maud Land closer to the coast in East Antarctica, the Antarctic Plateau near South Pole, and the Siple Coast location again. The plateau location (subsequently called “South Pole,” though not directly over the pole) was chosen most frequently, representing geographical diversity in station locations that are most impactful for monitoring 2-m temperature across the continent.

### b. Localization

For the monitoring network example shown in Fig. 4, the first station removes much more of the variance of the analysis uncertainty compared to the subsequent stations. Spurious correlations may occur between distant locations that in reality are uncorrelated due to sampling error or other effects. For instance, a point on the Antarctic Peninsula may have a spurious correlation with a point in the interior of East Antarctica on daily time scales. However, these two points are distant, separated by significant topography, and have very different weather regimes on short time scales (i.e., coastal vs plateau). Furthermore, the climatological correlation length scales on the peninsula are much shorter than the distance to East Antarctica (Fig. 3), likely resulting in some part of the correlation being a statistical anomaly. Covariance localization seeks to remove or minimize these spurious long-distance correlations. This is done by weighting covariances using a function of distance; covariances between points that are far apart are weighted less, and covariances between points closer together are weighted more.

The impact of using various localization length scales is illustrated in Fig. 5 by applying covariance localization as described in Gaspari and Cohn (1999) to the variance calculation by multiplying this function with the Kalman gain in (11). These length scales are Gaspari–Cohn zero/cutoff distances, and not the e-folding squared correlation decay distances used in Bumbaco et al. (2014), and so they are not equivalent. The e-folding distance is about 60% of the zero distance, and the correspondence is noted when relevant. In Fig. 5a, the lettered regions on the map represent seven 450 km-radius areas of preferred locations that were chosen for each of the length scales ranging from 1000 to 8000 km. These regions are also identified as areas of interest for potential station locations in Fig. 4. For the first station chosen over various length scales (Fig. 5b), the Siple Coast region (B) was chosen most often for length scales less than 2000 km, while the Megadunes region (A) was chosen most often for

---

**Fig. 3.** (a) Spatial correlation length e-folding length scales (km) and (b) 2-m air temperature variance (K²).
length scales longer than this (i.e., 2200–8000 km). The frequency with which the Megadunes region is chosen for medium and longer length scales lends confidence to the result in Fig. 4 that this is an important region to observe to account for continent-wide temperature variability. Given that this region has the longest correlation length scales on the continent on a climatological basis (Fig. 3a), we suspect that the results pointing to Siple Coast (B) for short localization length scales are spurious.

For the second station (Fig. 5c), Siple Coast (B) is frequently chosen for the shorter length scales (<2600 km) and again for longer length scales (>5000 km), indicating that there is quite a bit of temperature variability at that location. For the middle range in length scales (3000–4000 km), the Ronne Coast (G) region is chosen most frequently. As is introduced below, both of these regions have significant temperature variance (Fig. 3b) so it is not too surprising that these locations are chosen repeatedly. For the third station (Fig. 5d), the South Pole (C) region, Siple Coast (B) region, and Queen Maud Land–Coastal (D) regions are chosen most frequently. Siple Coast (B) appears for all three of the station selections at various length scales, indicating that the ordinal position in which it is chosen is sensitive to the length scale. On the other hand, the Megadunes region (A) is only preferred at medium and long length scales for the first station and not at all for the second or third stations; this shows that there is strong consensus that Megadunes is the location that reduces the most variance, and thus is chosen as the first station. While the Wilkes Land–Adelie Coast (E) and Wilkes Land–Knox Coast (F) regions are never the most chosen among the top three stations in Fig. 5, these areas will reappear in other calculations presented below.

Figure 6 illustrates the impact of spatial localization on the change in the analysis uncertainty with each station selected. Without localization (Fig. 6; bold, black line), the first station reduces the analysis uncertainty variance by about a factor of 10, leaving less variance to be explained by the subsequent stations. Localization with 4000–8000 km zero distance still removes much more variance with the first station, while a localization of 1000 km is too short. A length scale of 2000–3000 km allows a more gradual decrease in the amount of variance explained by each station selection. Thus, it is likely that many more stations could be found before the results are no longer statistically different from that of randomly chosen stations. How localization should be applied (i.e., the

Fig. 4. Analysis of preferred regions for the first three optimally sited stations for monitoring 0000 UTC 2-m temperature over the entire continent, assuming no observation network currently exists. Cells that are colored indicate that the location was chosen at least once, and the color indicates the frequency with which that location is chosen over 10,000 iterations. The grid points shaded on the plots were made larger for easy viewing; finer resolution was used for the calculation, as described in the text. Note that the scale of the colorbars varies on each panel.
appropriate radius, whether it is adaptive, how to apply localization in time for forecast errors, etc.) is dependent on the particular covariance structures of a system of interest. To make decisions on where to place new stations in future work, localization should be understood in the specific context of a desired network. For the subsequent examples, based on Figs. 5 and 6, we use a localization length scale of 3000 km, which is also consistent with the climatological $e$-folding length scales of 1200 km averaged across the continent found in Bumbaco et al. (2014) (see also Fig. 3a).

c. Monitoring networks with localization

1) BLANK SLATE NETWORK

With a covariance localization of 3000 km, the results assuming that no stations currently exist on the continent are shown in Fig. 7. Although the main affect on station ordering is evident from Fig. 5, the spatial distribution of selected locations provides additional information on the source for these differences. Similar to the results without localization (Fig. 4), the Megadunes region is still chosen as the location that explains the

**Fig. 5.** (a) Location of seven 450-km radius regions corresponding to preferred locations of optimal stations chosen over 10,000 iterations using localization length scales ranging from 1000 to 8000 km. The maximum percentage chosen over each of the seven regions for each localization length is shown for the (b) first, (c) second, and (d) third stations chosen by network design calculations.
most temperature variance. There is a reduction in the amount of variance explained by this region when compared to the results without localization, as expected, from about 54% to 13%. The results for the second station with localization employed differs from Fig. 4, with the Ronne Coast region chosen most frequently. For the third station, the Siple Coast area is frequently chosen, although Queen Maud Land and South Pole are also indicated, but less frequently.

2) AUGMENTED NETWORKS

The previous calculations pertained to starting with a blank slate, that is, there are no existing stations. Figure 8 shows the results for the monitoring metric after removing the variance explained by including the CD90 (Figs. 8a–c) and the CD75 (Figs. 8d–f) subsets. For the CD90 results, the percentage of time that the first three stations are chosen is high in comparison to the blank slate calculations, indicating that there is more certainty that these locations capture the remaining variance in the temperature signal. For the first station, an area not sampled by the CD90 network is chosen in Wilkes Land–Adelie Coast, which is a region that emerged as frequently chosen when testing the different localization length scales (cf. Fig. 5). The second station identified is on the Peninsula near Ellsworth Land, an area that did not previously emerge in results, as well as a location in Queen Maud Land (less frequently) that we call Queen Maud Land–Interior. For the third station, the Queen Maud Land–Interior location is chosen marginally over the second station near Ellsworth Land location. This indicates that these two sites are interchangeably chosen as either the second or third station in the 10 000 Monte Carlo simulations, explaining much of the remaining temperature variance after the CD90 network.

Figure 8 also shows the placement of the first three stations after considering the larger set of CD75 stations (45). Since more existing stations are taken into account in this calculation, more than double the number in the
CD90 set, there is less remaining variance to be explained and each region is selected less frequently. As for the first station, there are two locations that have been comparably chosen among the 10 000 simulations (Fig. 8d). Either the Queen Maud Land–Interior location (similar to the second or third station chosen with the CD90 set) or a region in Marie Byrd Land near the Bakutis Coast is chosen for the first station, and then the one not chosen as the first is chosen as the second (Fig. 8e). A relatively small amount of variance (only 0.6%) is explained by the third station and several different regions are chosen in the Monte Carlo simulations (Fig. 8f). There are two different areas within Queen Maud Land that are possibilities for an additional station along with sites in Marie Byrd Land near Pine Island Glacier and the Siple Coast. The Marie Byrd Land location explains slightly more variance than the others averaged over the 10 000 simulations, although the Siple Coast location is chosen a bit more frequently over the runs.

Though percentages of variance explained by additional stations on top of the CD90 or CD75 subsets of stations are relatively small, these additional stations do incrementally reduce uncertainty in the temperature metric. We illustrate this from a spatial perspective in Fig. 9 which shows the median temperature variance for several different network configurations. The optimal network designs of 20 stations, whether conditional on the existing CD90 stations (Fig. 9d) or placed assuming no stations exist (Fig. 9c), account for more temperature variance throughout the continent compared to the variance accounted for by the existing CD90 subset alone (Fig. 9b). Additionally, designing a 20-station blank slate network (Fig. 9c) accounts for nearly the same amount of temperature variance as a 38-station network (i.e., CD90 subset plus 20 optimally placed stations) (Fig. 9d). This shows that objectively installing 20 stations to optimize the given metric would be more valuable than a more ad hoc network of 38 stations, showcasing the benefits of optimal network design.

d. Forecast error reduction network with localization

Here we present results for the second metric described in section 2d: 24-h forecast errors in 2-m temperature. Again, we use a covariance localization length scale of 3000 km, and consider the same three cases as in the previous section: no existing network, new locations conditional on the CD90 subset, and new locations conditional on the CD75 subset. The regions highlighted to reduce forecast errors (Fig. 10) are similar to those identified for optimizing the surface temperature climate monitoring network, indicating that stations in similar regions could be useful for two different monitoring goals.
In the case of no existing network, again the Megadunes region is chosen for placement of the first station reducing about 16% of the variance (Fig. 10a), demonstrating that this is the single most valuable location for reducing 24-h, 2-m temperature forecast error. For the second station (Fig. 10b), three regions were chosen frequently over the 10 000 simulations: the South Pole area, Wilkes Land–Adelie Coast, and the Ronne Coast, with the South Pole area reducing variance the most. For the third station (Fig. 10c), the Siple Coast is chosen most frequently, though all of the regions that were highlighted for the second station selection were emphasized again, showing that they would likely be identified in subsequent station selection.

When taking the existing CD90 subset into account (Figs. 10d–f), the results to minimize 2-m temperature forecast error are remarkably similar to those for the monitoring network (Figs. 8a–c). Wilkes Land–Adelie Coast was chosen most frequently as the first station location, while sites on the peninsula (Ellsworth Land)
and Queen Maud Land–Interior alternated between the second or third station selected depending on the simulation.

The results after taking the CD75 subset of stations into account (Figs. 10g–i) are similar for both metrics, though there are some slight differences. For the first and second stations, the Queen Maud Land–Interior and Marie Byrd Land Coast are preferred and are interchangeable as the first or second station. In contrast to the climate monitoring results, however, a larger area of the Marie Byrd Land coastline is chosen, although the area near the Bakutis Coast is still chosen most frequently. For the third station, all of the regions highlighted—two locations on Queen Maud Land, Marie Byrd Land near Pine Island Glacier, the Siple Coast, and Wilkes Land near Sabrina Coast—explain nearly the same amount of variance, indicating that any of those locations would be good choices for the third station. Contrary to the monitoring network analysis where the Siple Coast location was chosen most frequently, the Queen Maud Land–Coastal location was chosen most frequently as the third station for the forecast network.

4. Summary and discussion

Accurate atmospheric observations are crucial for supporting science and operations over the Antarctic continent. Beyond assisting in basic logistical support, a robust weather observing network in Antarctica supports scientific campaigns across the region. An augmented observing network on the Antarctic continent
could improve forecasts throughout the Southern Hemisphere (Boullot et al. 2016), verify current forecasts (Jung et al. 2016), help ground-truth reanalysis products (Bracegirdle and Marshall 2012), improve historical simulations of Antarctic climate (Shu et al. 2015), and advance other scientific aims. In climate research, thinning glaciers in West and East Antarctica have been linked to warm ocean water (e.g., Jenkins et al. 2010; Greenbaum et al. 2015), and more weather observations would play an indispensable role in assessing climate change over the continent. These challenges, and the relative youth of the current observing network, motivate the testing of an objective technique for optimal surface observing network design. When expanding a network, especially in an expensive and dangerous environment such as Antarctica, it is critical to weigh the potential costs and benefits of where to place new stations. Optimal network design can maximize beneficial coverage and the information gained from the observing network, via an objective methodology for the siting decision process.

Two target surface observing networks have been considered here: one for monitoring daily 2-m temperature over the continent and one for reducing 24-h, 2-m temperature forecast errors. The data used to perform these calculations come from WRF Model forecasts from AMPS over a period of 4 years. A multivariate ensemble sensitivity approach with Monte Carlo bootstrapping is used to find optimal locations, which are those that maximally reduce the total variance in 2-m air temperature. Three types of calculations were performed for each monitoring goal. First, a blank slate calculation was done to find optimal locations assuming no stations exist in Antarctica. The second and third calculations were augmented networks, first taking into account the impact of the CD90 and CD75 subsets of stations, respectively, before determining the next stations that would be optimal.

We have explored the impact of spatial covariance localization on the results, and find that a distance of roughly 3000 km works well, which is consistent with climatological correlation length scales identified previously by Bumbaco et al. (2014). With localization employed, the results for the blank slate monitoring network (Fig. 7) are similar to those without localization (Fig. 4). The Megadunes region is identified as the optimal region for the first observing site in the results with and without localization, as well as the site for optimizing daily temperature monitoring averaged over eastern Antarctica, as found in Bumbaco et al. (2014). This emphasizes the importance of this location for monitoring 2-m temperature across the continent, under the assumption that no other stations exist. The two other regions identified as the second and third stations were the Ronne Coast and the Siple Coast.

When adjusting for the influence of the existing CD90 and CD75 stations on the monitoring metric and using localization (Fig. 8), different regions are highlighted as potential high-impact new station locations, but there are similarities between the regions regardless of the sets of existing stations considered. The top three stations for augmentation of the CD90 network are Wilkes Land–Adelie Coast, Ellsworth Land, and Queen Maud Land–Interior. For augmentation of the CD75 network, the top three stations are Queen Maud Land–Interior, Marie Byrd Land near the Bakutis Coast, and Marie Byrd Land near Pine Island Glacier. Thus, Queen Maud Land is clearly a region that could be targeted for observation to better monitor temperatures across the continent.

The analyses for designing Antarctic networks optimized to reduce 24-h forecast errors in 2-m temperature (Fig. 10) highlighted some of the same areas identified from the climate monitoring networks. Megadunes, once again, appeared as the optimal location to reduce these temperature forecast errors assuming that no other stations exist on the continent. South Pole and Siple Coast were the second and third station for this blank slate condition. The CD90 locations (Wilkes Land–Adelie Coast, Ellsworth Land, and Queen Maud Land–Interior) and CD75 locations (Queen Maud Land–Interior, Marie Byrd Land Coast, and Queen Maud Land–Coastal) highlight familiar regions. Many of the locations found to optimize the forecast network are the same as those for the monitoring network, showing that potentially installing new stations in some of these regions would be beneficial for two different goals.

These results could theoretically be used to prioritize the direction of resources to support current weather stations. As an example, the region on the Antarctic Peninsula near Ellsworth Land is chosen for the second station to optimize for both the monitoring (Fig. 8b) and forecasting (Fig. 10e) metrics when the CD90 stations are taken into account. There are two stations relatively close to that location in the CD75 set (Limbert and Sky-Blu) that already explain enough temperature variance such that Ellsworth Land was not chosen among the top three stations for the CD75 calculations. Dedicating resources to not only ensure that Limbert and Sky-Blu report consistently, but also that they are always assimilated into WRF for AMPS forecasts, would improve both monitoring and forecasting Antarctic surface temperature, and would likely be more cost-effective than installing and maintaining a new station in the area.

Relating to cost considerations, there are three different areas within Queen Maud Land that are
frequently identified throughout these calculations. However, this region contains areas that present complex topography and challenges to access, and thus, as with most of the continent, there are practical problems posed in installing new stations. These sorts of practical issues can be taken into account with the optimal network design technique presented here, by masking out specific regions that present enhanced practical challenges, so that the objective method chooses an alternative location that explains the highest amount of variance in the given metric. Regardless of these practical considerations, a blank slate 20 station, optimally placed network accounts for more temperature variance than the current CD90 network and is comparable to a 38-station network of 18 CD90 stations plus 20 more optimal stations conditional on the CD90 set (Fig. 9). This demonstrates, along with previous work (Huntley and Hakim 2010; Mauger et al. 2013), that optimizing station placement for a specific monitoring goal is more cost-effective than an ad hoc approach since the goal can be achieved with fewer stations. We emphasize that this information provides only one contributing element to the decision process, which often will have overriding concerns, such as for air transport, that may supersede the objective results from a network design calculation. Moreover, these optimal calculations for both monitoring and forecasting networks provide a demonstration of the approach, and real decisions must be based on additional factors and calculations for the issues at hand. Further work in testing the impact of introducing observations in the identified locations would advance Antarctic application specifically and the methodology for observing network design in general.

Acknowledgments. We are grateful for the constructive feedback of three anonymous reviewers, which improved the clarity of the paper. Three anonymous reviewers provided valuable feedback on a previous version of this paper, and Natalia Hryniv contributed to the calculations for the previous version. We thank Kevin Manning (NCAR) for providing access to the AMPS gridded data and observations over the period studied here, and Guillaume Mauger (UW) for sharing software related to network design. This research was supported by the National Science Foundation through Grant PLR-1542766 made to the University of Washington.

REFERENCES


FEBRUARY 2020
HAKIM ET AL.


