ARCTIC
VOL. 67, NO. 1 (MARCH 2014) P. 104–112

Quality Assessment of Meteorological Data for the Beaufort and Chukchi Sea Coastal Region using Automated Routines

MARTHA D. SHULSKI,[1,2] JINSHENG YOU,[1] JEREMY R. KRIEGER,[3] WILLIAM BAULE,[1] JING ZHANG,[4]
XIANGDONG ZHANG[5] and WARREN HOROWITZ[6]

(*Received 10 December 2012; accepted in revised form 9 August 2013*)

ABSTRACT. Meteorological observations from more than 250 stations in the Beaufort and Chukchi Sea coastal, interior, and offshore regions were gathered and quality-controlled for the period 1979 through 2009. These stations represent many different observing networks that operate in the region for the purposes of aviation, fire weather, coastal weather, climate, surface radiation, and hydrology and report data hourly or sub-hourly. A unified data quality control (QC) has been applied to these multi-resource data, incorporating three main QC procedures: the threshold test (identifying instances of an observation falling outside of a normal range); the step change test (identifying consecutive values that are excessively different); and the persistence test (flagging instances of excessively high or low variability in the observations). Methods previously developed for daily data QC do not work well for hourly data because they flag too many data entries. Improvements were developed to obtain the proper limits for hourly data QC. These QC procedures are able to identify the suspect data while producing far fewer Type I errors (the erroneous flagging of valid data). The fraction of flagged data for the entire database illustrates that the persistence test was failed the most often (1.34%), followed by the threshold (0.99%) and step change tests (0.02%). Comparisons based on neighboring stations were not performed for the database; however, correlations between nearby stations show promise, indicating that this type of check may be a viable option in such cases. This integrated high temporal resolution dataset will be invaluable for weather and climate analysis, as well as regional modeling applications, in an area that is undergoing significant climatic change.

Key words: western Arctic, meteorological observations, data quality, automated quality control, Beaufort Sea, Chukchi Sea, Alaska

RÉSUMÉ. Des observations météorologiques provenant de plus de 250 stations des régions côtières, intérieures et extracôtières de la mer de Beaufort et de la mer des Tchouktches ont été recueillies pendant la période allant de 1979 à 2009, puis elles ont fait l’objet d’un contrôle de la qualité. Ces stations relèvent de plusieurs réseaux d’observation différents qui existent dans la région à des fins d’aviation, de météorologie forestière, de météorologie côtière, de climat, de rayonnement de surface et d’hydrologie, et elles fournissent des données horaires ou subhoraires. Un contrôle de la qualité (CQ) unifié des données a été appliqué à ces données provenant de sources multiples en faisant appel à trois méthodes principales de CQ, soit le test d’acceptabilité (qui a permis de déterminer dans quels cas une observation ne faisait pas partie de la gamme normale); le test de la variation discrète (qui a permis de détecter les valeurs consécutives qui sont excessivement différentes); et le test de la persistance (qui a permis de repérer les cas de variabilité excessivement élevée ou basse). Les anciennes méthodes de CQ des données quotidiennes ne donnent pas de bons résultats dans le cas des données horaires parce qu’elles se trouvent à signaler un trop grand nombre d’entrées de données. Des améliorations ont été apportées afin d’obtenir les bonnes limites en vue du CQ des données horaires. Ces méthodes de CQ permettent de repérer les données douteuses et produisent beaucoup moins d’erreurs de type I (le signalement erroné de données valables). La fraction de données signalées pour l’ensemble de la base de données illustre que le test de persistance a échoué le plus souvent (1,34 %), suivi du test d’acceptabilité (0,99 %) et des tests de la variation discrète (0,02 %). Des comparaisons effectuées avec les données de stations avoisinantes n’ont pas été effectuées pour la base de données. Cependant, des corrélations entre les stations annexes s’avéraient prometteuses, ce qui a laissé entendre que ce type de vérification pourrait présenter une option viable dans de tels cas. Cet ensemble de données

---

[1] High Plains Regional Climate Center, School of Natural Resources, University of Nebraska – Lincoln, Lincoln, Nebraska 68583, USA
[2] Corresponding author: mshulski3@unl.edu
[3] Arctic Region Supercomputing Center, University of Alaska Fairbanks, Fairbanks, Alaska 99775, USA
[4] Department of Physics, and Department of Energy & Environmental Systems, North Carolina A&T State University, Greensboro, North Carolina 27411, USA
[5] International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, Alaska 99775, USA
[6] Bureau of Ocean Energy Management (BOEM) Alaska Region, 3801 Centerpoint Drive, Anchorage, Alaska 99503, USA
http://dx.dor.org/10.14430/arctic4367

intégrées à haute résolution temporelle aura une très grande valeur pour l'analyse météorologique et climatique ainsi que pour les applications de modélisation régionale dans une région où le changement climatique est important.

Mots clés : Arctique de l'Ouest, observations météorologiques, qualité des données, contrôle de la qualité automatisé, mer de Beaufort, mer des Tchouktches, Alaska

Traduit pour la revue *Arctic* par Nicole Giguère.

INTRODUCTION

Meteorological data from observing stations in the Arctic have lower spatial and temporal resolution than data from stations at lower latitudes, which is not surprising given the remote and harsh environmental conditions that complicate the collection of reliable observations. This region is currently experiencing the most rapid climatic changes to be found anywhere on the globe, resulting in significant environmental and social impacts (Hinzman et al., 2005; Jorgenson et al., 2006; Sakakibara, 2011). The western Arctic region encompassing the Beaufort and Chukchi Seas, an area of particular interest to this study, is being affected by seasonal declines in the extent of perennial ice cover (Comiso, 2012; Stroeve et al., 2012). The region is also influenced by both Arctic and extratropical storm systems. In the north of the study area, the Beaufort High pressure system dominates for much of the year, though it weakens in summer (Overland, 2009). At the same time, the Aleutian Low located over the south of the study area also displays clear seasonal variability, reaching its peak intensity in fall and diminishing in spring. Winds in the region are thus typically strong because of a steep pressure gradient between the northern high-pressure and southern low-pressure systems (Stegall and Zhang, 2012). In the snow-free season, mountain upslope and sea breeze circulations have been shown to occur along the Alaskan coast and North Slope in the absence of a dominant synoptic pattern (Moritz, 1977; Kozo, 1982a, b).

Offshore resource extraction activities have been increasing in the region and are expected to continue to do so, given the trend in seasonal ice cover. Of particular importance for drilling, shipping, and oil spill response is an understanding of the wind field, which is the main driving force for surface currents. In addition, coastal and offshore winds are of paramount importance for subsistence hunting and fishing activities in the region. High winds can compromise the safety and success of the local residents in hunting and fishing, with significant impact on the nutrition and cultural needs of the community. For offshore locations, in situ wind observational records are short-lived and associated with exploration platforms, meteorological buoys, and ships. Therefore, to better understand the spatial character of winds in this environment, it is necessary to use numerical modeling, in which accurate in situ observations are of prime importance for both initialization and validation of models. We conducted this study to develop a meteorological database for the Beaufort and Chukchi Sea coastal region that was subject to uniform quality control procedures. This database has been used in the production of the Chukchi-Beaufort Seas High-Resolution Atmospheric Reanalysis (CBHAR), a model-based data assimilation effort to better understand the climatology, variability, and changes in the region's surface wind field (Zhang et al., 2013).

DESCRIPTION OF METEOROLOGICAL DATA FOR THE WESTERN ARCTIC

In this study, surface meteorological data from 254 stations in a variety of observing networks and from various stand-alone projects were obtained for a Beaufort/Chukchi Seas Mesoscale Meteorological Modeling Study, which provides an improved representation of meteorological state in the region (Zhang, 2013). All available meteorological data for the period 1 January 1979 to 31 December 2009 were obtained (Fig. 1, Table 1). Only a subset of stations reported for this entire time period, and many stations have a relatively short period of record, of less than five years (Fig. 2). The following meteorological variables were included in the database: surface air temperature (˚C), dew point temperature (˚C), relative humidity (%), wind speed (m/s), wind direction (degrees), station pressure (hPa), sea level pressure (hPa), altimeter (in Hg), incoming shortwave radiation (W/m^2), incoming longwave radiation (W/m^2), total one-hour precipitation (mm), total six-hour precipitation (mm), and total 24-hour precipitation (mm). Height above ground for the observation sensors, such as the anemometer, varied somewhat by network as well as for some stations within specific networks. Sensor height is dependent on specifications set forth by the various sponsoring agencies and related to the purpose of the particular network, and specific details about sensor placement are generally obtainable from the sponsoring agency. All observational time stamps have been transformed to Coordinated Universal Time (UTC) and missing observations are identified as −999.0.

The majority of stations in this area are represented in the Integrated Surface Hourly (ISH) dataset obtained from the National Climatic Data Center (NCDC). At the 102 ISH stations in the region, data are measured at sub-hourly, hourly, or three-hourly intervals, depending on the station. Elevation for these sites ranges from 0 to 991 m above sea level. Data from this source have gone through quality control procedures at NCDC before being posted for distribution; however, they were still subjected to our quality control (QC) routines to maintain uniformity. ISH stations

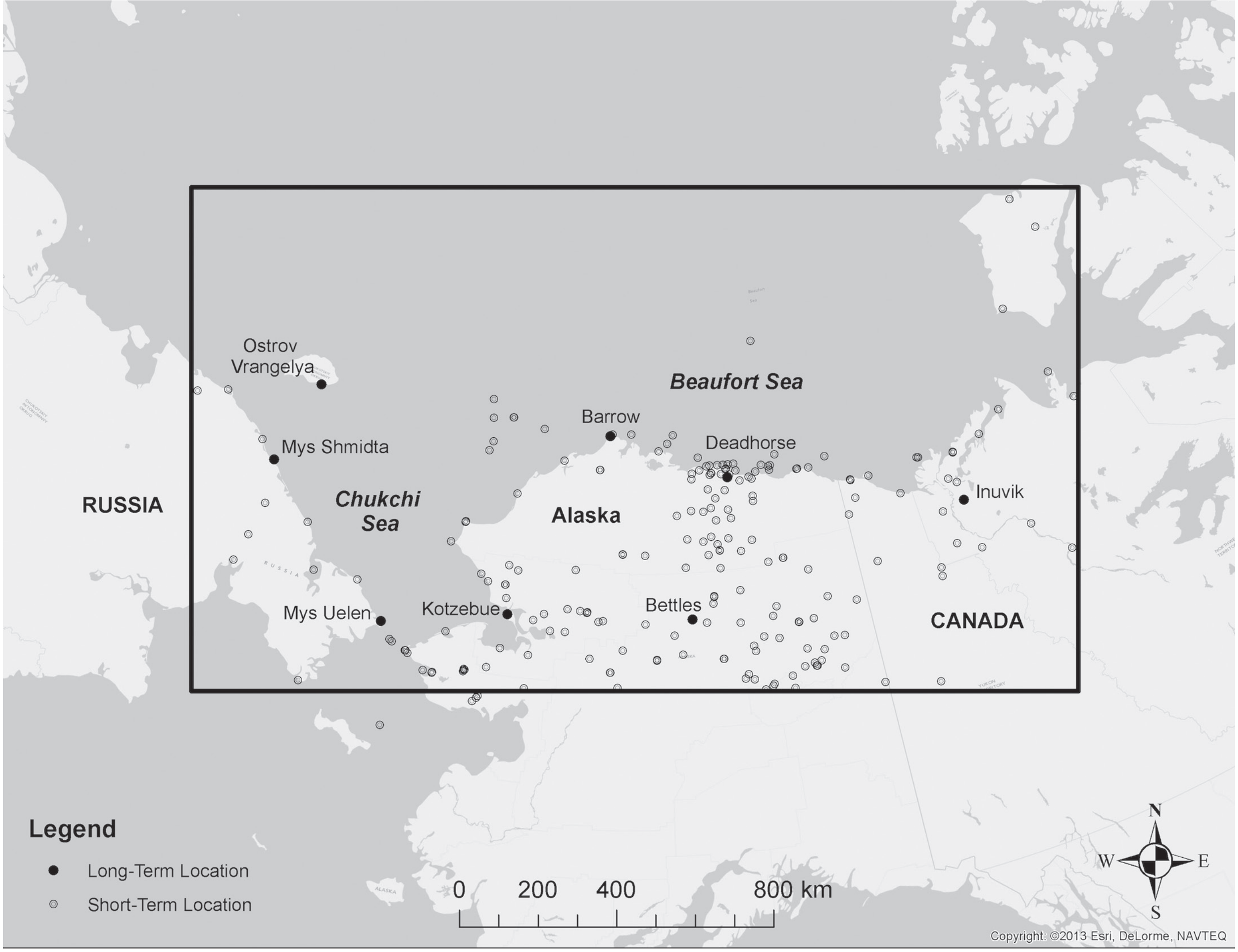

FIG. 1. Map of station locations in the database. Black dots represent long-term locations (stations with 30 years of data and suitable for a climatology), and open dots show stations with shorter reporting periods. The rectangle encloses the meteorological modeling domain for the accompanying CBHAR study.

TABLE 1. Network-specific information for stations used in the study.

Network	Abbreviation	Purpose	Number of stations	URL
Atmospheric Radiation Network	ARM	Radiation monitoring	2	http://www.arm.gov
Bureau of Ocean Energy Management	BOEM	Coastal weather monitoring	7	http://www.boem.gov
Buoy	Buoy	Offshore monitoring	6	N/A
Coastal-Marine Automated Network	C-MAN	Coastal weather monitoring	2	http://www.ndbc.noaa.gov
Long Term Ecological Research Network	LTER	Ecological monitoring	1	http://www.lternet.edu/sites/arc
NCDC Integrated Surface Hourly	NCDC	Weather monitoring, aviation	102	http://www.ncdc.noaa.gov
Remote Automated Weather Stations	RAWS	Fire weather monitoring	32	http://www.raws.dri.edu
Ship	Ship	Weather monitoring	5	N/A
Wellsite	Wellsite	Coastal/Offshore monitoring	17	N/A
Water and Environmental Research Center	WERC	Hydrologic monitoring	30	http://ine.uaf.edu/werc

are located primarily at airports across the region, and instrumentation height is standardized. Alaska has a good number of stations in this network overall, but a dearth of stations across the Arctic Coastal Plain. Hourly data from a total of 32 Remote Automated Weather Stations (RAWS) in Alaska were obtained through the Western Regional Climate Center. As this network is designed for the purpose of fire weather monitoring, most stations are located south of the Brooks Range, where fire danger is highest. Elevation of these sites ranges from 45 to 853 m above sea level, and instrumentation height is standardized. For some of these stations, the record is biased toward the warm season, or fire season, and winter observations are missing.

Data for seven sites in the vicinity of Prudhoe Bay were obtained from the Bureau of Ocean Energy Management (BOEM). These stations are located on or near the Beaufort

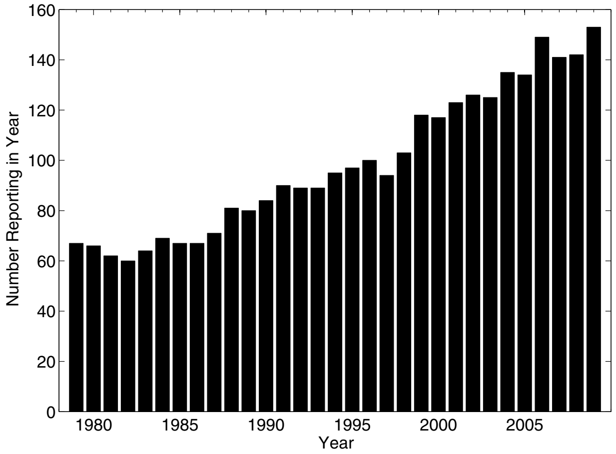

FIG. 2. The number of stations reporting each year from 1979 through 2009.

Sea coast at elevations ranging from 5 to 21 m above sea level. The Arctic coast is also home to two stations (Barrow and Atqasuk) within the Atmospheric Radiation Measurement (ARM) network. As the name implies, this network is primarily focused on measuring incoming and outgoing radiation for various spectral wavebands; however, supplementary meteorological data are also recorded. Data from two coastal stations (Prudhoe Bay and Red Dog Dock) were obtained from the National Data Buoy Center Coastal-Marine Automated Network (C-MAN) program. These stations are situated at sea level, and data recorded every six minutes from 2005 to 2008 are available.

The Water and Environmental Research Center (WERC) at the University of Alaska Fairbanks (UAF) maintains an extensive network of 30 surface meteorological stations, which are also located primarily in the Prudhoe Bay region and sites to the south along the Dalton Highway. The few stations not in this area are located on the Seward Peninsula. Several stations are located at higher altitudes, with elevations of more than 1000 m near the foothills of the Brooks Range. A station in the northern foothills of the Brooks Range, located at UAF's Toolik Field Station, is part of the Arctic Long Term Ecological Research (LTER) site, with meteorological data available since 1999.

Short-term data are available from a variety of well sites (primarily exploration wells and a few development wells) that operated in the Beaufort and Chukchi Seas from 1986 to 2009. In the Chukchi Sea, these exploration wells were in operation for one to three months during the open water season between 1989 and 1991. In the Beaufort Sea, meteorological data were also collected from exploration platforms for one to three months during the open water season, as well as during the winter months when the platforms occupied landfast ice. Meteorological data were also gathered from three buoys and one ship operated by the ConocoPhillips and Shell Companies in the domain of interest. Again these represent short-term sites, with data collected for just a few months.

Data from the scientific research vessel *Mirai*, which surveyed the region intermittently, were obtained from the Japan Agency for Marine-Earth Science and Technology (JAMSTEC, 2011). These data cover the months of September and October for various years from 1999 through 2009. Offshore data are also available from the Sea Ice Experiment – Dynamic Nature of the Arctic (SEDNA), a field campaign conducted at an ice camp in the Beaufort Sea (IARC, 2011). For this experiment, a meteorological tower was installed on the pack ice and observations were taken in April 2007. Lastly, the BOEM-supported project that conducted the study discussed in this paper launched a buoy with the assistance of the icebreaker USCGC *Healy*. This buoy operated in the western Beaufort Sea, collecting meteorological data every five minutes from early August to mid-September of 2009 (http://knik.iarc.uaf.edu/buoy09/).

AUTOMATED QC METHODS

A critical step in the development of a high-quality database is the performance of quality control procedures that serve to identify errors in the observations. This step is especially important for a remote, Arctic coastal environment, which can pose many challenges for the reliable operation of meteorological instrumentation. While manual procedures are quite valuable, particularly for shorter-term datasets in which the complete metadata are known, the use of manual procedures alone is not possible for longer-term datasets that would require heavy manual workloads. One objective of using automatic QC procedures for climatological data is to reduce the manual workload of human validators.

Quality control procedures have long been applied by NCDC (Guttman and Quayle, 1990) in a mix of manual and automatic checks to assess the validity of weather data from long-term stations. The statistical literature is replete with general guidance about identifying outliers in data (e.g., Barnett and Lewis, 1994), but literature concerning the application of techniques specific to the quality assessment of climatological data is scant, particularly for high temporal resolution meteorological data. General testing approaches, such as using threshold and step change criteria, have been designed for the review of single-station data to detect potential outliers (Wade, 1987; Reek et al., 1992; Meek and Hatfield, 1994; Eischeid et al., 1995; Shafer et al., 2000; Hubbard et al., 2005). Recently, QC has been expanding from procedures based on in-station checking to include procedures for inter-station checking (Wade, 1987; Gandin, 1988; Eischeid et al., 1995; Hubbard et al., 2005). The inter-station checks are based on reference estimates made using spatial techniques such as inverse weighting or statistical regressions between stations. Fiebrich et al. (2010) provide a recent review of advances and techniques used by various weather and climate networks.

Various examples exist of quality control performed on high temporal resolution data. The Oklahoma Mesonet (http://www.mesonet.org/) measures and archives weather conditions at five-minute intervals. The Climate Reference

Network (CRN) has installed multiple sensors for each observed variable to guarantee the continuous operation of the weather station, and thus the quality control can draw on multiple measurements of a single variable. This method efficiently detects instrument failures or other disturbances in the observational record; however, the cost of such a network is prohibitive for most applications.

Many different types of networks in the coastal areas of the Beaufort and Chukchi Seas record meteorological data at high temporal resolutions. The period of record at these stations varies greatly (ranging from less than one year to the full 30 years of the study period), though some of the short-term sites (< five years) represent key regions such as offshore areas. Techniques for quality control of these datasets are not available within an automatic system. Our work introduces an enhanced QC system, based on methods presented in Hubbard et al. (2005), to ensure a high-quality dataset. Although some of the data (e.g., the ISH database) had gone through basic QC procedures prior to being made public, all observations underwent the same QC techniques in this study in order to uniformly examine the database.

Hubbard et al. (2005) summarized the threshold, step change, and persistence tests, which check, respectively, whether an observational value crosses a given threshold, the change of the value between adjacent time steps, and whether the short-term variability of the value falls within certain limits. The *upper and lower threshold* test determines whether a given variable falls within a specified range for the month in question. These threshold limits are determined on the basis of distribution statistics, using a technique called the sigma test (Guttman et al., 1988) as in equation (1) of Hubbard et al. (2005). However, for stations without a long historical record, the climatic extremes for a given area are often used instead (Shafer et al., 2000). If a data entry is flagged by an $f = 3$, indicating a value more than three standard deviations away from the mean of the time period in question, this value represents 99.73% confidence that the data entry is an outlier; $f = 4$ represents 99.99% confidence. This procedure allows an informed choice regarding how many data points are flagged in the natural data stream. If the data stream itself contained no errors, the values being flagged would thus be Type I errors (the erroneous flagging of valid data) as defined in Hubbard et al. (2005). In operational use, data flagged as potential Type I errors should be considered suspect and subjected to further manual checking.

The *step change* (SC) test examines consecutively observed values in the data stream to determine whether their difference falls within an expected limit, based on the stations' climatology. In this case, the difference between the values for consecutive time intervals is tested.

The *persistence* test checks the variability of the measurements as in Equation (2) of Hubbard et al. (2005). When a sensor fails, it often reports a constant value, causing the standard deviation to become excessively small. Should the sensor fail for a long enough time period, the standard deviation will eventually become zero. For the study region, the persistence test is of particular importance because in an Arctic environment, the icing of anemometers making automated observations is common. In these cases, wind speed observations fall to zero for an extended and unrealistic period of time, which is particularly noteworthy during the cold season when winds are typically quite high. This unnatural period of reported calm can be followed in the record by a sudden increase in wind speed as the ice is either mechanically or thermally removed.

The basic QC procedures described above were originally developed and tested for daily data and thus need improvement for use with the higher temporal resolution data collected in this study. The modification entails first calculating the daily maximum and minimum observations from the high-resolution (hourly or sub-hourly) data. The means and standard deviations of the maximum and the minimum can then be calculated from the time series as (u_{max}, s_{max}) and (u_{min}, s_{min}), respectively. Limits can thus be formed by modifying equation (1) of Hubbard et al. (2005) as:

$$u_{min} - f \cdot s_{min} < x < u_{max} + f \cdot s_{max} \qquad (1)$$

This equation forms limits defined by the upper limits of the daily maximum and the lower limits of the daily minimum. Values falling outside of these limits will be flagged as outliers and subjected to further manual checking, where f can be taken as 3 (99.73% confidence) or 4 (99.99% confidence), as defined previously. The diurnal change of a variable (e.g., temperature) was calculated from the high-resolution (hourly or sub-hourly) data. The means and standard deviations calculated from the diurnal changes are then used to form the limits for each variable. Since both positive and negative changes are assumed to have the same probability, the equation can be written as:

$$u_{sc} - f s_{sc} < sc_{hr} < u_{sc} + f s_{sc} \qquad (2)$$

where sc_{hr} is the value change in a single time step, u_{sc} is the mean daily range, and s_{sc} the standard deviation daily range. Any value change between time steps that falls beyond the limits fails the test and is flagged as an outlier for further manual checking.

Statistics were obtained from the entire dataset (1979–2009) to determine the limits for each parameter in the automated QC procedures. The mean and standard deviation of the daily time series of the variables' maximum, mean, and daily range were calculated to form the limits expressed in equations (1) and (2). Several limits were calculated from the daily time series. In this application, only air temperature, dew point, and pressure have a normal distribution, which can be quality controlled for both the low and high ends of the distribution. The other variables, including wind speed and precipitation, do not, and thus only the highest value in the distribution can be tested. For the step change test, a minimum number of 15 consecutive valid step changes within a 72 data-entry window is required, otherwise a flag is assigned to indicate the need for additional manual checks. As suggested by Hubbard et al. (2005), QC procedures were applied using both $f = 3$ and $f = 4$ for the variables other than precipitation, for which an $f = 6$ was used instead. The $f = 4$ is most suitable for an extreme event analysis, while the $f = 3$ choice is most suitable for a variability analysis.

In addition to the above limits, hard limits of 0 and 360 degrees were used for wind direction. After scanning the entire dataset, a 254 mm (10 inch) value was used as a hard upper limit for precipitation, in addition to the parameters obtained from the observational record. This limit was added to override any possible excessively high precipitation amounts in the dataset. The station pressure was also specially treated. The minimum and maximum pressure were obtained from all stations, and after manual examination, only those having reasonable values were retained in order to calculate the standard deviation of the minimum and maximum pressures.

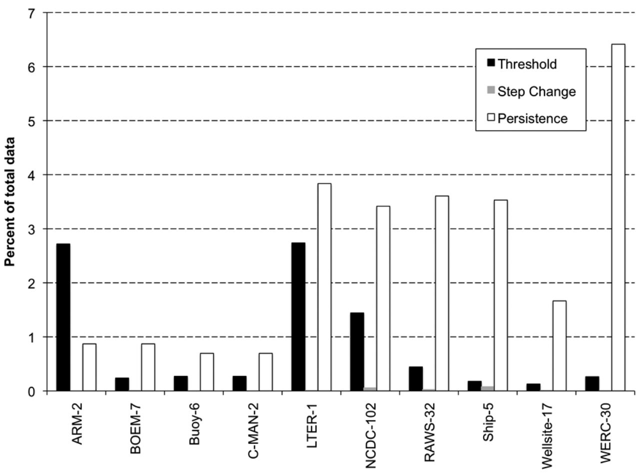

FIG. 3. The percentage of data (all variables combined) flagged using the three QC procedures, by network. The number of stations is shown after each network name.

AUTOMATED QC RESULTS

The QC methods used in this study detect three types of errors in the observational data (threshold, step change, and persistence). Values are flagged in automated procedures as the identified outliers are determined to be statistically significant relative to both the surrounding temporal data and the climatologically defined limits. In aggregate, the percentages of observations flagged in the database for each of the QC tests are as follows: 0.99% for the threshold test, 0.02% for the step change test, and 1.34% for the persistence test (Fig. 3). The relatively small fraction of flagged data illustrates the overall high quality of both the observations and the networks. It also demonstrates that the QC methods adopted in this work meet our expectations, since a manual validity reexamination is not needed for a large number of flagged data entries.

Analysis of the identified errors shows that most of the flagged data are the result of either instrumentation failures or the miscalibration of sensors. A few observation types have a much larger than average percentage of flagged data; these include daily precipitation, shortwave radiation, and one- and six-hour precipitation (NCDC). It should be noted that the short-term observational records (less than five years) of some of the stations may not be efficiently quality controlled by using only the three tests described in this work. The efficiency is lower because such stations lack the long history with which to set limits and determine standard deviations and other parameters used in the three different QC checks. One example of this weakness is the QC of dew point measurements at the station Iultin-in-Chukot, Russia, in the NCDC network. More than 90% of its dew point measurements were flagged. The reason for this high percentage of flagged data is that the parameters for quality controlling the variable use regionally derived values, which do not accurately reflect this station. A spatially based method, such as comparison with the nearest neighbor, would likely improve the performance in this instance.

The automated checks work equally well regardless of network or region, provided that the stations have a sufficiently long period of record. A station-by-station look at the results shows that there are problem stations, rather than problem networks or regions. There appear to be no major systematic differences in the QC performance with respect to network or region. Some particularly poor-performing stations are the sites with very short and highly discontinuous records, both because they lack a long station history with which to determine in-station limits and because the step change and persistence tests require runs of continuous data in order to perform properly. For some of the short-term sites, this is not the case. As an example, the NCDC period of record for the Tuktoyaktuk, Canada, site is January 1979 to April 1981; however, only 14% of the hourly observations are reported, and the remaining 86% of the hours have no data. For sites such as these, apart from use of the threshold test, it is nearly impossible to automate QC routines. The only alternative would be to check hour by hour with a neighboring station, if available.

In total, less than 5% of the data were flagged by any of the three QC tests. This relatively low percentage of flagged data is indicative of the overall good quality of the observations. For the most part, these data had gone through some form of quality check by their respective maintaining agencies prior to acquisition for this study. In terms of overall performance, the ARM, BOEM, Buoy, C-MAN, and Well site stations performed better than the LTER, NCDC, RAWS, Ship, and WERC stations (Fig. 3). The persistence test failed at a higher percentage than either the threshold or the step change test for nearly all stations (except for ARM).

Qualitatively, the shorter-term stations with records of less than five years do not appear to have more data problems, overall, than do the longer-term stations. For this database, these short-term stations comprise the majority and are therefore highly important for the examination of spatial patterns in the meteorological data. From the quality checks performed here and supplementary data reports provided, there is no indication that the overall quality of the short-term stations is any lower than the others. In fact, for certain stations (e.g., the SEDNA ice camp), the quality may be higher than for some of the long-term sites because they have fewer missing observations and periodic manual checks of the stations and sensors.

When delineated by variable, wind speed was flagged nearly 2.5% of the time (Fig. 4). In randomized manual checks of the flagged data, it appears that cases of zero wind speed or wind direction (or both) lasting for a period of several days were caught by these tests. It is assumed that sensor failure due to icing events was the cause of these errors. Manual checks also identified cases of persistent high wind speeds that were flagged by the persistence test. However, given the climatology of the coastal region, such cases do, in fact, occur periodically. It is therefore recommended that users of the database manually check these instances when using periods of consistently high wind speed in which there was persistence test failure.

DISCUSSION AND CONCLUSIONS

The traditional automated quality control methods developed for use with daily observations have been improved for examining weather stations covering a wide geographic area that produce high temporal resolution data, in order to avoid intensive manual reviewing, which is neither timely nor cost-effective. The identified problems in the observational dataset examined demonstrate that the improved methods successfully identified a variety of errors in the raw data. The newly developed QC tools significantly enhance our confidence in the dataset after performing random manual checks of the flagged data. Two different cutoff limits (99.73% confidence and 99.99% confidence) used in the QC algorithms will assist users in determining whether the flagged data are valid or invalid for their own purposes. As mentioned previously, the automated routines give users a guide by which to check the flagged data manually. Manual checks can be done by a combination of several different methods, including comparison to neighboring stations (where available), investigation of the synoptic meteorological conditions occurring around the time of the flagged observation(s), and comparison to available model output. The advantage of automated routines is that they can check all the data, which cannot be done manually given the sheer number of observations in this newly compiled dataset. Randomized checking of the automated QC results indicates that flagged data are indeed erroneous; however, further systematic manual checking, though time consuming, is nevertheless optimal.

This work illustrates that applied in-station limit tests can successfully identify outliers in the dataset. However, spatial tests based on information from neighboring stations can, in some cases, be more robust at identifying errors or outliers in the dataset when a strong correlation exists. Figure 5 shows a promising example of a high correlation between temperatures at West Dock, a coastal station, and Betty Pingo, located 15 km inland and 18 km from West Dock, two stations operating in the WERC network. The strong relationship between the measurements at the two stations demonstrates a potential opportunity to apply the spatial regression test (SRT, Hubbard et al., 2005) successfully to nearby stations that measure the same variables, such as the air temperature and wind speed. Further testing of this method is required in order to perform this type of test on the entire dataset and to find candidates for neighboring stations.

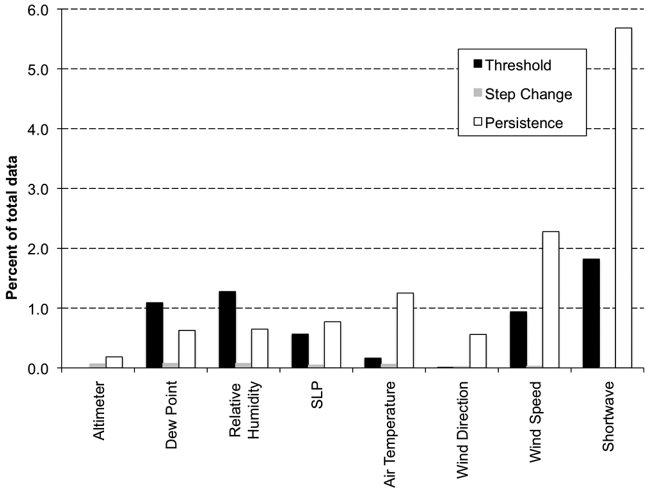

FIG. 4. The percentage of data (all stations combined) flagged using the three QC procedures, by variable.

The short-term observational records of some stations may not be efficiently quality controlled by using only the three basic automated methods described in this work. In such cases, the parameters used for quality control of different meteorological variables are region-wide parameters, which may not accurately reflect the climatological conditions of a particular station. The spatially based method would likely improve the performance of the automated routines in these cases.

The automated routines do not take into account any sort of manual flagging. For example, if a station, sensor, or time period is known to have problems with one or more particular variables, this additional information could be used to enhance the database quality. In most instances, this information would be known to individuals and organizations that maintain and service the stations and network, rather than included with normal metadata information distributed to the public. The inclusion of such information in the QC procedure could represent an additional type of quality assurance to enhance the database.

Overall, the newly created and quality-controlled database represents a unique blend of stations and networks for the Beaufort and Chukchi Sea coastal area. These data have never before been combined into a single dataset and

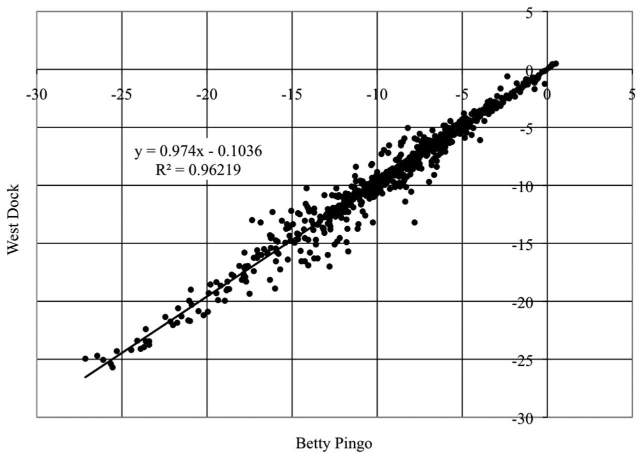

FIG. 5. Example of strong correlation between two neighboring stations, showing hourly air temperatures (˚C) for May 2000 at Betty Pingo and West Dock, near Prudhoe Bay, Alaska (WERC network).

subjected to uniform QC procedures. Even though the periods of record for these stations are highly variable, this type of database has significant value for improving model initialization (and by extension, simulation), as well as validation of model output, particularly as it includes offshore stations. Ultimately, this quality-controlled dataset will lead to an improved representation of the area's regional atmospheric conditions and overall climatology, which will include being incorporated into the construction of the high-resolution regional reanalysis CBHAR, as well as a greater understanding of multiscale regional climate processes in the Arctic.

ACKNOWLEDGEMENTS

This work was supported by the Bureau of Ocean Energy Management, Department of Interior, under contract M06PC00018. Data storage resources were provided by the Arctic Region Supercomputing Center at the University of Alaska Fairbanks.

REFERENCES

Barnett, V., and Lewis, T. 1994. Outliers in statistical data, 3rd ed. Hoboken, New Jersey: Wiley and Sons.

Comiso, J.C. 2012. Large decadal decline of the Arctic multiyear ice cover. Journal of Climate 25(4):1176–1193. http://dx.doi.org/10.1175/JCLI-D-11-00113.1

Eischeid, J.K., Baker, C.B., Karl, T.R., and Diaz, H.F. 1995. The quality control of long-term climatological data using objective data analysis. Journal of Applied Meteorology 34(12):2787–2795. http://dx.doi.org/10.1175/1520-0450(1995)034<2787:TQCOLT>2.0.CO;2

Fiebrich, C.A., Morgan, C.R., McCombs, A.G., Hall, P.K., Jr., and McPherson, R.A. 2010. Quality assurance procedures for mesoscale meteorological data. Journal of Atmospheric and Oceanic Technology 27(10):1565–1582. http://dx.doi.org/10.1175/2010JTECHA1433.1

Gandin, L.S. 1988. Complex quality control of meteorological observations. Monthly Weather Review 116(5):1137–1156. http://dx.doi.org/10.1175/1520-0493(1988)116<1137:CQCOMO>2.0.CO;2

Guttman N.B., and Quayle, R.G. 1990. A review of cooperative temperature data validation. Journal of Atmospheric and Oceanic Technology 7(2):334–339. http://dx.doi.org/10.1175/1520-0426(1990)007<0334:AROCTD>2.0.CO;2

Guttman, N., Karl, C., Reek, T., and Shuler, V. 1988. Measuring the performance of data validators. Bulletin of the American Meteorological Society 69(12):1448–1452. http://dx.doi.org/10.1175/1520-0477(1988)069<1448:MTPODV>2.0.CO;2

Hinzman, L.D., Bettez, N.D., Bolton, W.R., Chapin, F.S., Dyurgerov, M.B., Fastie, C.L., Griffith, B., et al. 2005. Evidence and implications of recent climate change in northern Alaska and other Arctic regions. Climatic Change 72(3):251–298. http://dx.doi.org/10.1007/s10584-005-5352-2

Hubbard, K.G., Goddard, S., Sorensen, W.D., Wells, N., and Osugi, T.T. 2005. Performance of quality assurance procedures for an Applied Climate Information System. Journal of Atmospheric and Oceanic Technology 22(1):105–112. http://dx.doi.org/10.1175/JTECH-1657.1

Jorgenson, M.T., Shur, Y.L., and Pullman, E.R. 2006. Abrupt increase in permafrost degradation in Arctic Alaska. Geophysical Research Letters 33, L02503. http://dx.doi.org/10.1029/2005GL024960

Kozo, T.L. 1982a. An observational study of sea breezes along the Alaskan Beaufort Sea coast: Part I. Journal of Applied Meteorology 21(7):891–905. http://dx.doi.org/10.1175/1520-0450(1982)021<0891:AOSOSB>2.0.CO;2

———. 1982b. A mathematical model of sea breezes along the Alaskan Beaufort Sea coast: Part II. Journal of Applied Meteorology 21(7):906–924. http://dx.doi.org/10.1175/1520-0450(1982)021<0906:AMMOSB>2.0.CO;2

Meek, D.W., and Hatfield, J.L. 1994. Data quality checking for single station meteorological databases. Agricultural and Forest Meteorology 69(1-2):85–109. http://dx.doi.org/10.1016/0168-1923(94)90083-3

Moritz, R.E. 1977. On a possible sea-breeze circulation near Barrow, Alaska. Arctic and Alpine Research 9(4):427–431. http://dx.doi.org/10.2307/1550535

Overland, J.E. 2009. Meteorology of the Beaufort Sea. Journal of Geophysical Research 114, C00A07. http://dx.doi.org/10.1029/2008JC004861

Reek, T., Doty, S.R., and Owen, T.W. 1992. A deterministic approach to the validation of historical daily temperature and precipitation data from the Cooperative Network. Bulletin of the American Meteorological Society 73(6):753–762.
http://dx.doi.org/10.1175/1520-0477(1992)073<0753:ADATTV>2.0.CO;2

Sakakibara, C. 2011. Climate change and cultural survival in the Arctic: People of the whales and muktuk politics. Weather, Climate, and Society 3(2):76–89.
http://dx.doi.org/10.1175/WCAS-D-10-05007.1

Shafer, M.A., Fiebrich, C.A., Arndt, D.S., Fredrickson, S.E., and Hughes, T.W. 2000. Quality assurance procedures in the Oklahoma Mesonetwork. Journal of Atmospheric and Oceanic Technology 17(4):474–494.
http://dx.doi.org/10.1175/1520-0426(2000)017<0474:QAPITO>2.0.CO;2

Stegall, S.T., and Zhang, J. 2012. Wind field climatology, changes, and extremes in the Chukchi–Beaufort Seas and Alaska North Slope during 1979–2009. Journal of Climate 25(23):8075–8089.
http://dx.doi.org/10.1175/JCLI-D-11-00532.1

Stroeve, J.C., Serreze, M.C., Holland, M.M., Kay, J.E., Maslanik, J., and Barrett, A.P. 2012. The Arctic's rapidly shrinking sea ice cover: A research synthesis. Climatic Change 110(3-4):1005–1027.
http://dx.doi.org/10.1007/s10584-011-0101-1

Wade, C.G. 1987. A quality control program for surface mesometeorological data. Journal of Atmospheric and Oceanic Technology 4(3):435–453.
http://dx.doi.org/10.1175/1520-0426(1987)004<0435:AQCPFS>2.0.CO;2

Zhang, X., Zhang, J., Krieger, J., Shulski, M., Liu, F., Stegall, S., Tao, W., You, J., Baule, W., and Potter, B. 2013. Beaufort and Chukchi Seas mesoscale meteorology modeling study, Final Report. OCS Study BOEM 2013-0119. Anchorage, Alaska: U.S. Department of the Interior, Bureau of Ocean Energy Management, Alaska OCS Region. 203 p.