High-quality in situ measurements from radiosondes and dropsondes are the gold standard for vertical profiles of fundamental atmospheric measurements such as wind, temperature ($T$), and relative humidity (RH). Satellite-borne remote sensors provide much-needed global, long-term coverage; however, they do not match the ability of sondes to capture sharp transitions and fine vertical structure, and have significant performance limitations (e.g., the inability of infrared sounders to penetrate clouds, poor accuracy in the boundary layer). Sondes are also a trusted means to calibrate and validate remote sensors. However, it is challenging to launch radiosondes from remote locations such as the ocean surface or the interior of Antarctica. Aircraft release dropsondes above such locations but are limited by the range and endurance of the aircraft. The driftsonde system fills an important gap in our ability to use sondes to measure atmospheric profiles in remote locations. Although its creation was motivated by The Observing System Research and Predictability Experiment (THORPEX; e.g., Shapiro and Thorpe 2004) to optimize the global observing system, it has contributed to varied investigations ranging from understanding the development of tropical cyclones to validating satellite retrievals in Antarctica.

The driftsonde is a unique balloonborne instrument that releases dropsondes to provide high-resolution in situ profiles of atmospheric temperature, humidity ($H$), pressure ($P$), and winds from the lower stratosphere down to the surface. It is ideal for applications over oceans and remote polar and continental regions, filling critical gaps in data coverage where the release of surface-based radiosondes is not possible. Figure 1 shows the driftsonde system concept in which a stratospheric balloon-borne dropsonde platform fills an important gap in in-situ research measurement capabilities by delivering high-resolution, MIST dropsondes to remote locations from heights unobtainable by research aircraft.

Fig. 1. The driftsonde system concept.
The driftsonde system consists of the flight train shown in Fig. 2, ground control software that is web based, and ground control servers and associated hardware located at NCAR in Boulder, Colorado. Web-based software adds flexibility, so that an experiment operations center can be located worldwide or it may rotate through several locations to support the continuous (24 hours a day, seven days a week) nature of balloon flight operations. The latest version of the gondola structure (Fig. 3), made of insulating foam, contains up to 54 MIST dropsondes, a custom electronics motherboard that acts as the brain of the system, lithium batteries to power the gondola for the expected flight duration, radio equipment to communicate with both a released dropsonde and Iridium satellites, and electric heaters to maintain the gondola electronics and batteries at an operational temperature. The heaters are powered by solar panels mounted outside the gondola. Heating the sonde electronics and batteries before they are released from the gondola ensures the sensors will operate normally. Early driftsonde tests were done with several ballooning partners. Subsequent deployments have been a close collaboration between NCAR and the French Centre National d’Etudes Spatiales (CNES), with NCAR developing the driftsonde measurement capability (gondola, MIST sondes, communications, data quality, etc.) and CNES having responsibility for all ballooning development and flight operations.

Characteristics of the driftsonde’s MIST dropsondes are shown in Table 1. They are physically smaller and lighter than current aircraft dropsondes but both use the Vaisala RS921 sensor module, with the same pressure, temperature, and humidity sensors as the well-documented RS92-SGP radiosonde (Vaisala 2013). Each MIST dropsonde undergoes a calibration verification step at NCAR. While the MIST dropsondes make similar measurements to aircraft dropsondes, the two platforms—driftsonde gondola and aircraft—have notable differences. Capabilities of the stratospheric balloons used to lift driftsonde are central to its strengths and limitations. Aircraft are maneuverable but can remain aloft for only a few hours. They also can precisely target specific locations. The driftsonde is not maneuverable but can remain aloft for several months. Sondes
are released from much higher altitudes than most aircraft (Fig. 4), and multiple driftsondes can be in flight simultaneously. In the just completed Concordiasi experiment, a constellation of 13 driftsondes were aloft simultaneously for about two months, with drops controlled from the ground in McMurdo Station, Antarctica; Toulouse, France; and Boulder, Colorado. In general, because stratospheric balloons drift with the wind and have long duration, driftsonde data can provide synoptic-scale or finer observations with wide-ranging geographical coverage that would be difficult to obtain with research aircraft. On the other hand, precise targeting of drops is limited by the accuracy of balloon trajectory forecasts.

FIELD EXPERIMENTS AND SCIENCE APPLICATIONS. As the driftsonde system was developed (Fig. 5), it was deployed in three field experiments associated with THORPEX activities (Table 2). Each revealed and led to needed improvements, and from these experiments we also learned how to take better advantage of the system’s strengths for varied science applications. Details of the driftsonde system performance and scientific applications in each experiment are presented in the following sections.

2 AMMA, based on a French initiative, was organized by an international scientific group and is currently funded by a large number of agencies, especially from France, the United Kingdom, the United States, and Africa. It has been the beneficiary of a major financial contribution from the European Community’s Sixth Framework Programme. Detailed information on scientific coordination and funding is available on the AMMA International website.

Driftsonde observations during the African Monsoon Multidisciplinary Analyses. The first field project experience with driftsonde was in the African Monsoon Multidisciplinary Analyses (AMMA; www.amma-international.org) project,2 both as a rigorous field test and for its scientific value. AMMA was organized to advance understanding of the West African monsoon
system and to improve predictions of its variability and the associated wide range of societal impacts. It is a major international program led by France but involves agencies and scientists located in the United Kingdom, Germany, the United States, and other countries across Africa and Europe. The driftsonde deployment supported AMMA’s research focus on high-impact weather and was undertaken through a collaboration between AMMA and THORPEX. The program is summarized in Redelsperger et al. (2006), with the driftsonde-observing strategy described in Rabier et al. (2008).

The AMMA measurement strategy included long-term observations from 2002 through 2010 to investigate the interannual variability of the West African monsoon. Within this period was an extended observing period (EOP) from 2005 to 2007 to document the annual cycle, and four special observing periods (SOPs) during 2006 to provide specific observations of physical processes and weather systems. The driftsonde operations took place during the fourth SOP, covering the late monsoon period in August and September 2006.

The driftsonde deployment was considered a THORPEX observing system test (e.g., Shapiro and Thorpe 2004). Thus, a large component of driftsonde operations concentrated on engineering tests, including the first major test of NCAR’s new smaller and lighter dropsonde called MIST, which at 175 g was less than half the weight of the previous dropsondes and was developed specifically for use in the driftsonde. Miniaturization of the dropsonde was necessary for ballooning, where weight is more critical than for aircraft deployments. For AMMA, the driftsonde gondola held 49 MIST dropsondes.

The AMMA deployment was the first scientific use of the new CNES 12-m superpressure balloons, the new NCAR gondolas, and the MIST dropsondes.

Eight driftsondes were launched from Zinder, Niger, and floated at about 20 km as they drifted eastward, reaching the Atlantic Ocean. The location was chosen to allow investigators to study both African easterly waves over central and western Africa and the potential intensification of

<table>
<thead>
<tr>
<th>Table 1. MIST dropsonde characteristics.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIST sonde</td>
</tr>
<tr>
<td>Weight: 175 g</td>
</tr>
<tr>
<td>Length: 30.5 cm</td>
</tr>
<tr>
<td>Diameter: 4.6 cm</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pressure, temperature, and dual-humidity sensors (used for T-PARC and Concordiasi, not AMMA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vaisala module RSS921 (same as in RS92 radiosonde)</td>
</tr>
<tr>
<td>Sample rate: 0.5 s</td>
</tr>
<tr>
<td>Resolution: P: 0.01 hPa, T: 0.01°, RH: 0.1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wind speed and direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPS</td>
</tr>
<tr>
<td>Sample rate: 0.25 s</td>
</tr>
<tr>
<td>Resolution: 0.01 m s^{-1}</td>
</tr>
<tr>
<td>and 0.1°, respectively</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sonde fall speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approx 90 m s^{-1} at 30 km, 45 m s^{-1} at 20 km, and 10 m s^{-1} at sea level</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sonde fall time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approx 19 min from 30 km, 16 min from 20 km</td>
</tr>
</tbody>
</table>

Fig. 4. Sample driftsonde temperature profiles, with the drop altitudes of driftsonde and the T-PARC aircraft.
these waves into tropical disturbances or even hurricanes over the subtropical Atlantic. Flight trajectories and the dropsonde locations are shown in Fig. 6. The dropsondes were able to sample both the Saharan air layer that is often advected over the tropical Atlantic and precursor environments, and the near-storm environments associated with 2006 Tropical Storm Florence and Hurricanes Gordon and Helen. The first balloon was launched on 28 August 2006, and the termination of the final balloon over the central subtropical Atlantic occurred on 22 September. Two balloons had mission durations in excess of eight days. Sondes were typically deployed near 0000 and 1200 UTC, as well as on demand for promising weather conditions. For further details on the driftsonde operations during AMMA, a description of the challenges associated with balloon and dropsonde design, and preliminary scientific results, refer to the overview of the deployment presented in Drobinski et al. (2006, 2013a).

**Fig. 5.** Four images of driftsonde deployments. (a) Launch from McMurdo Station, Antarctica, during Concordiasi (2010). (b) Shortly after launch during Concordiasi. Dropsondes are visible at the perimeter of the driftsonde gondola below the CNES superpressure balloon. (c) Launch from Hawaii during T-PARC (2008). Driftsonde gondola is on its deployment sled. (d) Driftsonde before launch during AMMA (2006). This earlier version of the driftsonde gondola was constructed from cardboard rather than hard foam.

<table>
<thead>
<tr>
<th>Table 2. Driftsonde evolution through three field projects.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AMMA, Aug–Sep 2006</strong></td>
</tr>
<tr>
<td>• Launch site: Zinder, Niger</td>
</tr>
<tr>
<td>• Superpressure balloon: 8 flights, 3–18-day duration, 20-km float level</td>
</tr>
<tr>
<td>• 178 MIST sondes with T, RH, GPS winds; no pressure sensor</td>
</tr>
<tr>
<td>• Ground control through a terminal modem program with simple text commands and manual operation</td>
</tr>
<tr>
<td><strong>T-PARC, Aug–Oct 2008</strong></td>
</tr>
<tr>
<td>• Launch site: Hawaii (the Big Island)</td>
</tr>
<tr>
<td>• Zero-pressure balloon: 15 flights, 3–6-day duration, 30-km float level</td>
</tr>
<tr>
<td>• 254 MIST sondes with P, T, RH, GPS winds (equivalent to dropsonde sensor suite)</td>
</tr>
<tr>
<td>• Web-based ground control for drops and display position and sounding data</td>
</tr>
<tr>
<td><strong>Concordiasi field experiment, Sep–Dec 2010</strong></td>
</tr>
<tr>
<td>• Launch site: McMurdo Station, Antarctica</td>
</tr>
<tr>
<td>• Superpressure balloon: 13 flights, 50+-day duration, 18-km float level</td>
</tr>
<tr>
<td>• 644 MIST sondes with P, T, RH, GPS winds (equivalent to dropsonde sensor suite)</td>
</tr>
<tr>
<td>• Enhanced web-based ground control to schedule automatic drops and display position and sounding data</td>
</tr>
</tbody>
</table>
Much was learned about the system’s operation and performance as 124 sondes were successfully deployed from the eight driftsonde gondolas. While there was one premature failure of a stratospheric balloon and lessons learned about driftsonde launch procedures, most engineering challenges highlighted the need to improve aspects of the driftsonde gondola design and usability of system software. There were many cases where sondes failed to launch, as well as periods of lost communication between the ground control station and the gondola. Another significant conclusion was the need to redesign the MIST sonde to include a pressure sensor. A pressure sensor was not included in the AMMA version of the MIST sonde because of the incorrect assumption that the pressure for the dropsonde profile could be obtained from knowledge of the pressure and GPS altitude at launch and the hydrostatic equation. However, the flight-level pressure sensor on the gondolas did not have sufficient accuracy, so small errors in the initial pressure were magnified by the downward integration of the hydrostatic equation.

A success of AMMA was the demonstrated ability to target measurements with driftsondes launched 4–5 days before the sampling window. The successful sampling of storms was due to the accuracy of the upper-level winds predicted from operational NWP models, the quasi-nondivergent nature of the flow at 20 km, and the successful tropical storm genesis forecasts by the AMMA team through combining experimental and operation products for storm genesis. Despite technical challenges, AMMA demonstrated the scientific value of driftsonde, in particular to evaluate operational model performance and to identify specific areas for model improvement. Drobinski et al. (2013b) use driftsonde and other data to evaluate the performance of the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecast System (IFS) and the two versions of the Météo-France Action de Recherche Petite Echelle Grande Echelle (ARPEGE) operational forecast system through comparison of the dropsonde observations and the model analysis and prediction. The concept was to improve performance in key regions and different flow regimes using the special dropsonde observations to supplement the assessment procedure employed by the operational centers. This technique was extended to evaluate the impacts of recent upgrades in model physics and assimilation techniques.

The findings from Drobinski et al. (2013b) include that these models well represent the complex vertical structure of humidity associated with the Saharan air layer (Fig. 7a). However, the comparison identified temperature errors of several degrees Celsius in both modeling systems near the base of the Saharan air layer (Fig. 7b). The relatively large errors are likely because of the lack of dust and the associated radiative impacts within NWP models. This result suggests shortcomings in the assimilation system, perhaps due to the vertical resolution of the satellite data, and argues for the inclusion of aerosols and their radiative effects in NWP models. In this case, static stability errors in the vicinity of the Saharan dust could, in turn, impact the likelihood, intensity, and structure of convection. Considerable debate currently exists in determining and explaining the potential impacts of the Saharan air layer on tropical cyclones [e.g., see Braun (2010) and references therein]. It was also found that within the analyses, the zonal and meridional winds in the Saharan air layer cases have significant errors (Figs. 7c,d). The driftsonde observations were also valuable as “ground truth”...
in data impact and data assimilation experiments (Drobinski et al. 2013b).

**Driftsonde observations during the THORPEX Pacific Asian Regional Campaign.** The driftsonde experience in AMMA was a success both in identifying technical issues that needed attention after the campaign and collecting scientifically valuable observations. Prior to the next large field use in the THORPEX Pacific Asian Regional Campaign (T-PARC) in 2008, many parts of the system were upgraded. In particular, a pressure measurement was added to the MIST sonde, robustness of the satellite communication link was improved, and reliability of the sonde separation from the gondola when a drop is commanded was also improved.

As a multinational field campaign and research initiative, T-PARC addressed the shorter-range dynamics and forecast skill of one region (eastern Asian and the western North Pacific) and its downstream impact on the medium-range dynamics and forecast skill of another region (eastern North Pacific and North America). High-impact weather events over the regions examined in T-PARC have strong dynamical links downstream. For example, persistent deep tropical convection or the extratropical transition of tropical cyclones can trigger downstream responses over the eastern North Pacific, North America, and beyond via upper-tropospheric wave packets on the primary midlatitude waveguides (Anwender et al. 2008, Harr et al. 2008). Then, wave packets can be invigorated by subsequent downstream cyclogenesis events that are often associated with reduced predictability. High-impact weather events over North America driven by these processes can include intense extratropical cyclones, floods, severe weather, and hot, dry winds that increase the risk of wildfires and the severity of droughts. While T-PARC objectives encompassed mesoscale and synoptic-scale processes associated with tropical cyclones over the western North Pacific and eastern Asia, they also addressed medium-range forecast skill associated with downstream impacts across the North Pacific and beyond.

Dropsondes were an important contribution to T-PARC. In addition to use of driftsonde, four aircraft from three countries [United States: U.S. Air Force (USAF) WC-130J and Naval Research Laboratory (NRL) P-3, Taiwan: Dropwindsonde

---

3 T-PARC was supported financially by Germany, Canada, Japan, Australia, France, Korea, Taiwan, the United Kingdom, ECMWF, and the United States.
Observations for Typhoon Surveillance near the Taiwan Region (DOTSTAR) Astra (Wu et al. 2005, 2007b), and Germany: Deutsches Zentrum für Luft- und Raumfahrt (DLR) Falcon] were used to take special observations. The collaborative program required an experimental design that covered a very wide geographic range to address three primary components: 1) a tropical measurement strategy examined circulations of the tropical western North Pacific monsoon environment as they related to enhanced and reduced periods of widespread deep convection, tropical cyclone formation, tropical cyclone intensification, and tropical cyclone structure change; 2) a measurement strategy for the extratropical transition and its downstream impacts followed the poleward movement of a decaying tropical cyclone and the intense cyclogenesis that results from its interaction with the midlatitude circulation; and 3) a targeted observation strategy focused on regions in which extra observations may reduce numerical forecast error growth (Wu et al. 2007a, 2009; Harnisch and Weissmann 2010; Reynolds et al. 2010; Weissmann et al. 2011; Chou et al. 2011). In T-PARC, the targeted observations were aimed at reducing errors associated with tropical cyclone track forecasts, which included whether a tropical cyclone would recurve, the longitude of recurvature, and the orientation and speed along the track following recurvature.

To accomplish the primary objectives of T-PARC, a complete tropical-to-extratropical measurement strategy was necessary. For example, predictability associated with extratropical transition depends on the intensity and structure of the tropical cyclone, where and when the tropical cyclone arrives in the midlatitude westerlies, the characteristics of the midlatitude waveguide that impact the extratropical-transition-related cyclogenesis, and the downstream propagation and evolution of the wave packets.

The motivation for deployment of driftsondes in T-PARC was to provide measurements over data-sparse regions of the tropical central Pacific. The data from the driftsonde complemented satellite observations and provided calibration and validation data for new satellite-based observations (Hawkins and Velden 2011) and global reanalysis products (Wang et al. 2010). During T-PARC, 16 driftsondes were launched from the southern end of the Big Island of Hawaii between 15 August and 30 September 2008. Thirteen of the driftsondes traveled at an altitude of about 30 km for up to five days to reach the western North Pacific and the primary T-PARC observation region (Fig. 8). Throughout T-PARC, 254 dropsondes were deployed from the driftsondes. The location and timing of the dropsonde deployments were coordinated from the T-PARC operations center at the Naval Postgraduate School in Monterey, California, taking advantage of the now web-based driftsonde control and display software. During each balloon flight, data were relayed to the T-PARC operations center, quality controlled, and transmitted to the GTS for use at operational weather centers.

As an example of driftsonde use during T-PARC, the fourth Driftsonde was launched on 24 August 2008 and on 29 August it reached a tropical disturbance that was being investigated by the T-PARC aircraft (Fig. 9). While the driftsonde was overflying the tropical disturbance in the lower stratosphere, two aircraft were deploying dropsondes from their respective flight-level altitudes. Seven dropsondes were deployed from the driftsonde and provided measurements of two upper-tropospheric cyclonic systems that were preventing the development of the tropical disturbance.

Driftsondes in T-PARC were flown with zero-pressure balloons. These were designed to float much higher than the superpressure balloons used for AMMA (and later in Concordiasi), but they

Fig. 8. Locations of dropsondes deployed from the 16 driftsonde balloons launched from the Big Island of Hawaii during T-PARC. Each square indicates the location of a dropsonde release.
also had a shorter flight lifetime. Because of light winds and a flaw in the ballooning technique, several flights failed to advect far enough westward to enter the most interesting measurement region. However, overall, data obtained from dropsondes released from the driftsondes provided valuable measurements of persistent deep convection, with special emphasis on the detailed vertical structure, impacts of vertical wind shear, and upper-level divergent outflow. Because of the driftsonde launch location and trajectories, data were instrumental in monitoring tropical cloud clusters that migrated over the data-sparse region of the tropical central Pacific until the clusters reached the region of T-PARC aircraft operations.

**Driftsonde observations during the Concordiasi field experiment.** The third major driftsonde deployment was in 2010 for the Concordiasi field experiment (Rabier et al. 2010, 2013), a multidisciplinary effort jointly conducted by several groups in France and the United States to study the lower stratosphere and troposphere above Antarctica. Concordiasi was one of the cluster of THORPEX projects associated with the International Polar Year (e.g., Renfrew et al. 2008; Hanesiak et al. 2010; Kristjánsson et al. 2011). The primary focus of Concordiasi was to validate the use of satellite observations and to document which observing systems are most relevant for numerical weather prediction over the polar areas. Concordiasi field experiments took place in austral springs 2008–10, including surface measurements and radiosoundings at the Concordia Antarctica station at Dome C, and radiosoundings at the Dumont d’Urville and Rothera sites on Antarctica. In 2010 driftsonde was part of an innovative constellation of balloons that provided a unique set of measurements spanning a large spatial extent (both horizontal and vertical) and time. The balloons drifted for several months in the lower stratosphere around 18 km, circling over Antarctica in the polar vortex. The balloon flotilla formed a regional observatory of the atmosphere. As in the earlier driftsonde experiments, hundreds of soundings were performed on command. The launch campaign took place from the U.S. McMurdo Station, located at 78°S latitude. Nineteen balloons were launched between 8 September and 26 October 2010. The mean flight duration was 69 days, while the longest flight lasted 95 days. Thirteen balloons carried a driftsonde, and six carried other instruments for Concordiasi. The long flight duration of the superpressure balloons used for Concordiasi, months rather than about a week for the previous driftsonde use, was critical to enable the project’s science.

To prepare for Concordiasi, the driftsonde system was modified for the much longer duration flights and challenging range of thermal conditions it would encounter. Early in this high-latitude project, the gondolas were in total darkness, and later in the project they transitioned to full sunlight. The software was

---

4 Concordiasi was organized by an international scientific group and supported by the following agencies: Météo-France, CNES, the Institut Polaire Français (IPEV), the Progetto Nazionale Ricerche in Antartide (PNRA), CNRS’s Institut national des sciences de l’Univers (INSU), the National Science Foundation (NSF), NCAR, the Concordia consortium, the University of Wyoming, Purdue University, and the University of Colorado. ECMWF also contributed to the project through computer resources and support, and scientific expertise. The two operational polar agencies—PNRA and IPEV—are thanked for their support at Concordia station.
also enhanced to allow for drops at prescheduled times. Overall, the 13 driftsonde gondolas returned 644 high-quality profiles, with only 14 failed drops. This is a much higher success rate than in either AMMA or T-PARC, and resulted in an excellent spatial distribution of observations both over the Antarctic continent and the surrounding ocean. Figure 10 shows the comprehensive coverage and distribution of drop locations over the full experiment, as well as an example of the coverage of the constellation on a single day.

Many dropsondes were released to coincide with driftsonde overpasses of Concordia station, allowing for comparison of dropsonde and radiosonde profiles, and also to coincide with overpasses of the Meteorological Operation (MetOp) satellite, allowing comparison with data from the Infrared Atmospheric Sounding Interferometer (IASI). IASI is an advanced infrared sounder that has a large impact on NWP systems in general. However, there are some difficulties in its use over polar areas because the extremely cold polar environment makes it more difficult to extract temperature information from infrared spectra and makes it difficult to detect cloud properties. As a consequence, IASI is currently underutilized over Antarctica.

A number of important results have already come from the 2010 Concordiasi dataset, as described in the Concordiasi workshop report (Rabier et al. 2013). Wang et al. (2013) compare sonde profiles with satellite retrievals, using the National Oceanic and Atmospheric Administration (NOAA) Products Validation System (NPROVS) to match Concordiasi dropsonde and radiosonde profiles with profiles from several satellite products. A comparison of temperature profiles shows a cold bias present in all satellite data. The cold bias has a larger magnitude relative to the dropsonde data than the radiosonde for all satellite products except the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC; Fig. 11). The difference between the radiosonde and dropsonde bias can be traced to a larger cold bias over the Antarctic continent than over the coast and ocean, since all radiosonde stations but two are located along the coast (Fig. 10a). The source of this bias remains a topic of investigation. Aside from the bias, and an inability to resolve detailed thermodynamic structures near the surface and tropopause, the satellite retrievals reproduce the temperature profiles reasonably well.
In addition to temperature and humidity profiles, cloud properties such as cloud-top pressure and cloud effective emissivity (emissivity of an equivalent single-layer cloud) can be retrieved from IASI measurements. These retrievals are highly dependent on the quality of temperature and humidity profiles. As reported in Rabier et al. (2013), detection of cloud properties can be improved by using an accurate atmospheric profile provided by the Concordiasi dropsondes rather than the atmospheric model.

Another result from this dataset comes from the use of Concordiasi driftsonde observations in real time at NWP centers. As noted in Rabier et al. (2013), large systematic differences exist between various NWP analyses and forecasts for temperature over Antarctica, and for winds on the surrounding oceans. Comparison between short-range forecasts and the Concordiasi dropsonde data show that models poorly represent near-surface temperature over the Antarctic high terrain. The strong thermal inversions are challenging because numerical models need very good representations of both turbulent exchange processes and snow processes to simulate this extreme atmospheric behavior. The difference between the dropsonde and the model temperatures at the lowest model level is presented in Fig. 12 for the French global model.
model is too warm over the plateau in Antarctica, and it is too cold over the surrounding sea ice.

The impact of dropsondes on NWP models has also been studied with data denial experiments and advanced data impact diagnostics. Dropsondes have a positive impact on the forecast performance in different models, with an impact of the same order of magnitude as that of radiosondes. Rabier et al. (2013) report that the average error reduction per observation is much larger for dropsondes than for satellite data, and that dropsonde observations have a greater impact when they are closer to the pole. In Fig. 13, the impact of dropsonde temperature and wind profiles is illustrated at high levels (pressure less than 400 hPa) and low levels (pressure greater than 400 hPa), together with the number of observations. Overall, temperature information contributes most at low levels, and wind information contributes more at high levels. However, on a per-observation basis, both wind and temperature have larger impacts at low levels, where there are very few other observations.

These results from driftsonde data in Concordiasi provide insight into improvements to the global observing system that must be achieved to improve NWP over the polar areas. This is important not only to improve forecast performance but also for producing more accurate reanalyses of the atmosphere to document climate change.

CONCLUSIONS. Thanks to the driftsonde system, in situ measurements were obtained in parts of the world that are not accessible by any other means. This has provided invaluable information about model strengths and weaknesses, and about which observations will be needed in the future to monitor the climate, especially in polar areas. The excellent technical success achieved during Concordiasi demonstrates that driftsonde is now a mature, reliable, and productive observing system. Its strengths include the ability to reach difficult parts of the globe; to collect highly accurate, in situ dropsonde profiles; to reliably release up to 54 dropsondes per system from the lower stratosphere; and to be deployed as a constellation with many driftsondes flying simultaneously.

Like aircraft dropsonde systems, field experiments using driftsondes involve significant cost and require much advance planning. It can take months to understand likely flight paths and obtain permissions to overfly many countries. If a balloon drifts near a region where overflight permissions have not been granted, then it must be cut down. The ability of the driftsonde to observe specific phenomena depends critically upon finding a suitable launch site relative to stratospheric wind patterns. For T-PARC, finding a subtropical Pacific island for which stratospheric wind patterns intersected the climatological tracks of tropical cyclones was surprisingly difficult. In contrast, finding a suitable launch site for AMMA within Africa was more straightforward. Once a launch site is selected, the success of the driftsonde to target a specific event depends on accurate forecasts of both the wind field in the stratosphere and the evolution and movement of the event to be targeted. Despite the ability of driftsondes to intercept tropical cyclones during AMMA and to a lesser extent during T-PARC, reliance upon forecasts with lead times of several days is a disadvantage relative to aircraft dropsonde
deployment. At far longer lead times, the driftsonde behavior from Concordiasi and experience from ballooning campaigns during GARP suggest that the balloons tend to be advected into the confluent, more dynamically active regions of the atmosphere. For many aspects of weather research, this behavior is desirable.

Despite these complexities, scientifically, the driftsondes are well suited to numerous applications, especially because of their ability to operate in otherwise data-sparse regions. In particular, they have unique value for verifying and evaluating NWP models, global reanalysis models, and data assimilation approaches. The measurements are also a valuable resource to validate remote sensors, especially on satellites but also airborne or ground-based remote sensors. Driftsondes also can support process studies in otherwise difficult locations. In AMMA and T-PARC, examples include the effects of the SAL and factors that control the development of a tropical disturbance. There is also a potential role for driftsondes operationally, although costs and forecast impacts of this have not been considered in detail. For example, a concept discussed in the early years of driftsonde was that a series of driftsondes could be released at regular intervals from sites in Asia to provide synoptic data over the Pacific Ocean, or from the East Coast of the United States for regular observations over the Atlantic Ocean. In addition to driftsonde flights, the Concordiasi program included measurements of ozone-related processes, microphysics of stratospheric clouds, and remote sensing using GPS occultation. It showed the potential of combining such diverse flight-level measurements with the vertical profiles from dropsondes to carry out a multidisciplinary experiment that would not be possible with aircraft.

The driftsonde system has been discussed as a possible contributor to future field studies. One is a long-duration stratospheric balloon campaign at the equator intended to study, among other goals, the dynamics of the equatorial middle atmosphere with a focus on the quasi-biennial oscillation, and transport, dehydration, and clouds in the tropical tropopause layer. A second field study suggested driftsondes as one of a synergistic set of tools to study the propagation and effects of orographically generated atmospheric gravity waves from near the surface to the upper atmosphere. Those with interest in using driftsondes in their research are encouraged to contact the lead authors.

ACKNOWLEDGMENTS. Development and field use of the driftsonde system is a joint effort of many people, institutions, and nations. We thank all who made the development and deployments to AMMA, T-PARC, and Concordiasi successful. We are grateful to Rolf Langland, Mark Bradford, Joe VanAndel, Dean Lauritsen, Chip Owens, Clayton Arendt, and Mary Hanson for contributions to the development.

Melvyn Shapiro is especially acknowledged for his resurrection of the early ideas of the late Vin Lally, which were developed decades before their time. In many regards, this paper is a testament to Vin’s creativity.

We are grateful for funding support from the National Science Foundation’s Division of Atmospheric and Geospace Sciences and Office of Polar Programs (Grants ATM-0301213, ATM-9732665, ANT-0733007, ANT-1002057, and AGS-0736003), the National Oceanic and Atmospheric Administration (Grant NA17GPI376), the National Science Council of Taiwan (Grants NSC 96-2745-M-002-004, NSC 97-2111-M-002-005, and NSC 97-2111-M-002-016-MY3), the Taiwan Central Weather Bureau (Grant MOTC-CWB-97-6M-01), the Office of Naval Research (Grants N00173-08-1-G007 and N00014-09-WR20008), and the support of the national and international THORPEX project offices. NSF further supported these field projects through their support of the U.S. THORPEX project Office, and the Lower Atmospheric Observing Facilities.

REFERENCES


——, and Coauthors, 2013b: Driftsonde observations to evaluate numerical weather prediction of the late


Although urban areas comprise a very small fraction of Earth’s land cover (Schneider et al. 2009), over half of global population live in urban agglomerations. Therefore, it is important to monitor, understand, and predict the modifications occurring in local weather and climate due to urbanization, particularly for the perspective of accurate high-resolution weather and air-quality forecasting and climate-sensitive urban design and planning. Cities are characterized by a high fraction of impervious surfaces, which modify both surface energy and water balances, further affecting atmospheric boundary layer (ABL) turbulence and weather processes. Cities are also the main “area sources” of air pollutants with detrimental effects on human health and comfort. Cities can generate, modify, and/or amplify many processes behind global changes such as

A dedicated intensive research-grade observational network in Helsinki enables studies of the physical processes in the urban atmosphere at high latitude.
increases in greenhouse gas concentrations, increased water and energy demand, environmental pollution, or change of biodiversity. Moreover, the recent increase of spatial resolution in numerical weather prediction (NWP) models and the improved treatments of model physics and chemistry in NWP and chemical-transport models demand a more realistic representation of urban features, processes, and feedbacks in these models (Kukkonen et al. 2012). Therefore, we need dedicated and long-term monitoring of urban meteorological processes.

Long-term urban ABL observations at different spatial and temporal scales using various instruments have been made in few cities (Rotach et al. 2002; Grimmond 2006), and most studies have not used state-of-the-art equipment for long periods. Since the pioneering Metropolitan Meteorological Experiment (METROMEX) urban study in St. Louis (Changnon et al. 1971), the more recent comprehensive studies include the cities of Basel, Switzerland (Rotach et al. 2005); Marseille, France (Cros et al. 2004); Oklahoma City, United States (Allwine and Flaherty 2006); New York, United States (Hanna et al. 2006); Toulouse, France (Masson et al. 2008); Montreal and Vancouver, Canada (www.epic.ca); and London, United Kingdom (Wood et al. 2009; Harrison et al. 2012). Most studies have focused on specific campaigns, often with less than one year of measurements and have concentrated on midlatitude cities. To date, most high-latitude urban ABL research has used solitary or few point measurements (Eresmaa et al. 2006; Mårtensson et al. 2006; Lemonsu et al. 2008; Vesala et al. 2008; Järvi et al. 2009a; Bergeron and Strachan 2012; Nordbo et al. 2012a). Thus, there is a clear lack of intensive research-grade long-term ABL observations, especially from high-latitude cities, with their associated pronounced annual variations in meteorological conditions and continuous snow cover that can last several months (Lemonsu et al. 2010). These conditions might create extreme meteorological conditions, such as strong ground-based or elevated inversions (Kukkonen et al. 2005) with detrimental air-quality implications.

An extensive mesoscale effort (Helsinki Testbed) centered around the Helsinki metropolitan area has been running several years (core data years 2005–09) covering a 150 × 150 km² area of southern Finland and northern Estonia (Koskinen et al. 2011). In addition to this mesoscale observational network, the Station for Measuring Ecosystem-Atmosphere Relations III (SMEAR-III) urban measurement station has been running in Helsinki since 2004 (Järvi et al. 2009b). The station concentrates on observing surface–atmosphere exchange processes at a micrometeorological scale, in particular using the eddy-covariance (EC) method to directly measure vertical turbulent fluxes of momentum, sensible and latent heat, and carbon dioxide (Vesala et al. 2008; Järvi et al. 2012; Nordbo et al. 2012a,b).

Combining those observations with additional state-of-the-art observations has enabled a new observational research-intensive network, the Helsinki Urban Boundary-Layer Atmosphere Network (UrBAN; http://urban.fmi.fi). Our aim is that the network will improve our understanding of urban ABLs by including measurements of a wide variety of processes at a range of scales. The observations in the network are complemented with model developments, such as NWP and meteorological preprocessing of air-quality (see online supplemental material). The long-term purposes of Helsinki UrBAN are to

1) understand the processes in Helsinki’s ABL, as affected by a range of surfaces within a few kilometers (urban, suburban, and sea) and the strong climatic seasonality;
2) provide better experimental data for developing and evaluating numerical models such as NWP, meteorological preprocessing, and air-quality models (see online supplemental material); and
3) provide results that will support improvements in relevant applications such as city planning, building design, and energy use.

AFFILIATIONS: Wood, JOFFRE, DREBS, VIHMA, HRISIKKO, SUOMI, FORTELIUS, KANGAS, KARPINEN, AND KUKKONEN—Finnish Meteorological Institute, Helsinki, Finland; JÄRVI, NORDBO, HAAFANALA, MOILANEN, AND VESALA—Department of Physics, University of Helsinki, Helsinki, Finland; KOUZNETSOV—Finnish Meteorological Institute, Helsinki, Finland, and Obukhov Institute of Atmospheric Physics, Moscow, Russia; O’CONNOR—Finnish Meteorological Institute, Helsinki, Finland, and Department of Meteorology, University of Reading, Reading, United Kingdom; MOISEV—Finnish Meteorological Institute, Helsinki, Finland, and Department of Physics, University of Helsinki, Helsinki, Finland
*CURRENT AFFILIATION: Forschungszentrum Jülich GmbH, Institut für Energie-und Klimaforschung; Troposphäre (IEK-8), Jülich, Germany
CORRESPONDING AUTHOR: C. R. Wood, Finnish Meteorological Institute, Erik Palmén aukio 1, FI-00101, Helsinki, Finland E-mail: curtis.wood@fmi.fi
The abstract for this article can be found in this issue, following the table of contents.
DOI:10.1175/BAMS-D-12-00146.1
A supplement to this article is available online (10.1175/BAMS-D-12-00146.2)
In final form 12 March 2013 ©2013 American Meteorological Society
The aim of this article is to describe the network, show selected results, and remark on potential future use of the results of the network.

**THE NETWORK.** An overview of observing site locations and instrumentation for studying Helsinki's ABL is displayed in Fig. 1 and Table 1. We define downtown Helsinki as the central area of buildings that are densely packed with a regular layout (Nordbo et al. 2012a): that is, approximately the land area south of 60.18°N (encompassing the observation sites of Torni, Fire Station, Sitra, and Kaisaniemi) (see sidebar on the climate and topography of Helsinki).

*In situ equipment and main sites.* The EC measurements of turbulent fluxes are made at three sites in Helsinki: Kumpula, Fire Station, and Torni.

On the Kumpula campus, next to the Finnish Meteorological Institute (FMI) and University of Helsinki buildings, is the SMEAR-III mast (Järvi et al. 2009b). This semi-urban site, located around 4 km northeast from downtown Helsinki, has developed as a core site of focused activity for intercomparison of instruments and a deeper understanding of processes at one site. EC measurements at Kumpula have been ongoing since 1 December 2005 (Vesala et al. 2008; Järvi et al. 2009b,c). The flux measurements are carried out atop a 31-m-high lattice mast (the base being 26 m MSL), with possible mast flow distortion from 0° to 50°. It is characterized as a semi-urban measurement site, given that three different surface cover areas (road, vegetation, built) can be distinguished within about a kilometer around the mast, allowing separate analysis of fluxes for those upwind surfaces. The mean height of the nearby buildings ($z_H$) in the built sector is 20 m (and only 8.4 m averaged for the whole sector within a 1-km radius), so the flux measurements are carried out at $1.6z_H$ with respect to the nearby buildings. Detailed description of the site and measurements can be found elsewhere (Vesala et al. 2008; Järvi et al. 2009c; Nordbo et al. 2012a).

The other two EC sites, Fire Station and Torni, are located downtown within a distance of only 500 m from each other (Nordbo et al. 2012a). Those locations are among the highest possible locations in downtown Helsinki, and their source areas can be estimated at about a kilometer radius depending on atmospheric stability and wind direction (Nordbo et al. 2012a). At the Fire Station, flux measurements began on 28 June 2010 (the measurements had to stop on 27 January 2011 because of building refurbishments and to date have not resumed yet). Measurements were made on top of the 38-m-high tower (the base is 23 m MSL) with mast flow distortion from 90° to 180°. The tower itself is 8.8 m × 8.8 m, and a pole was installed on top of it in northwest corner so that the measurements were carried out 42 m AGL (above ground level). At Torni, the measurements have been ongoing since 28 September 2010. The measurements are made 60 m AGL (ground at 15 m MSL) with mast flow distortion from 50° to 185°, where a 2.3-m-high pole

---

**Fig. 1. Maps of Helsinki with equipment locations marked.** (a) Shown are profile masts at Kivenlahti and Isosaari island, Vaisala site, and airport. The extent of the area shown in (a) is approximately 38 km latitude by 38 km longitude. (b) Land-use map: urban/paved (white), vegetative (green), and water (blue) (HSY 2008). The red box in (a) shows the area in (b). The sodar moved from Kumpula westward to Pasila. The grid points of the HARMONIE model are 2.5 km spacing on Lambert conformal plane (see online supplemental materials).
is installed. The details of these measurements—including building configuration, instrumental layout, and footprint analysis—can be found in Nordbo et al. (2012a).

The surrounding area of both sites is highly built up with the fraction of impervious surfaces, including buildings and paved surfaces, being above 80% inside a 1-km-radius circle. In downtown Helsinki, buildings are reasonably well packed and uniform and variation as a function of wind direction within the footprint of Torni EC site are as follows: mean building height $z_H = 19–29$ m, aerodynamic roughness length $z_0 = 0.9–1.9$ m, and zero-plane displacement height $z_d = 12–18$ m (following morphological methods from building databases; Nordbo et al. 2012a). The building height and roughness length vary slightly as a function of wind direction. Both downtown EC stations are at about 2–3 times the mean building height and 9–50 $(z - z_d)/z_0$ (Nordbo et al. 2012a). For comparison, although there are EC measurements at a range of heights within and above the urban canopy across the globe, most urban EC sites range between 0.4 and 500 $(z - z_d)/z_0$ and up to 22 $(z - z_d)/z_0$ (Roth 2000; Nordbo et al. 2012a).

At all three EC sites, the fluxes (of momentum, sensible heat, latent heat, CO$_2$, and aerosol particle number) are measured with the well-established EC method (Aubinet et al. 2012). The configurations comprise a three-dimensional sonic anemometer, infrared gas analyzers for measuring the CO$_2$ and water-vapor fluctuations, and water-condensation particle counters for measuring the CO$_2$ and water-vapor fluctuations. Raw eddy-covariance data are logged at 10 Hz for postprocessing, and fluxes are listed in Table 1. The list of the core equipment, their locations (latitude and longitude), and their uses is found in Nordbo et al. (2012a).
Helsinki, the capital city of Finland, had approximately 596,000 inhabitants in 2012 (density of 2789 km$^{-2}$), although the Helsinki metropolitan area (defined as the municipalities of Helsinki, Espoo, Vantaa, and Kauniainen) has about 1.1 million inhabitants (1377 km$^{-2}$) (Population Register Center of Finland 2012). Helsinki is located on the north shore of the Gulf of Finland in the Baltic Sea. Because of the influence of the vicinity of the sea, as well as oceanic-scale features such as the North Atlantic Drift, winter temperatures are higher than the far northern location might suggest, resulting in a humid continental climate (Köppen–Geiger climate classification Dfb). The annual mean air temperature is +5.9°C (period 1981–2010), with the lowest monthly average temperature in February of −4.7°C and the highest monthly average temperature in July of +17.8°C (Pirinen et al. 2012). Monthly average sunshine hours vary from 25 h in December to 280 h in June. Precipitation amounts are moderate, where the average annual precipitation sum is 655 mm, with the highest monthly average in August (80 mm) and the lowest in April (32 mm). The average number of days with precipitation exceeding 1 mm varies all year round from 7 to 12 days per month. Because of the vicinity of the Gulf of Finland, long spells of dry and hot weather are rare, unless synoptic conditions dictate: for example, a long period of easterly flow occurs in the summer.

In winter, the sea is typically ice covered but can be open or partly open. In winter conditions of an open sea during cold-air outbreaks, the sea surface may be 20°–30°C warmer than the air above it. This may result in turbulent fluxes of sensible and latent heat totaling several hundreds of watts per square meter, resulting in a development of a convective boundary layer (Vihma and Brümmer 2002; Tammelin et al. 2012). In summer, the sea surface temperature may vary in time and space because of upwelling in the sea. During coastal upwelling, the sea surface temperature may rapidly drop by up to 10°C, causing stabilization of the ABL and fog formation (Suomi 2004).

Topographical changes in Helsinki are very slight and most of the metropolitan area is less than 50 m above mean sea level (MSL). Helsinki has heterogeneity at the kilometer scale with urban, suburban, forest, grassland, and sea (open or frozen) (Fig. 1). Most of the city is low rise, with few buildings above 60 m.
snow depth, and visibility. Around 10 of these are within the area on Fig. 1a. Two reference sites worthy of mention include Kaisaniemi—where air temperature has been recorded since 1838—and Helsinki–Vantaa airport, useful for inland conditions (Fig. 1).

**Spatially resolving and spatially averaging equipment.** A single-antenna vertical sodar (Kouznetsov 2007, 2009) was located at Kumpula (on the FMI roof) from 14 August 2009 to 5 September 2011 but was moved westward 2.3 km to another semi-urban site (Pasila) from 29 March 2012 onward. The repetition frequency is 5 s, and the sounding range is 20–400 m above the antenna, with a vertical resolution of 10 m. The Doppler shift of the echo signal is used to evaluate the vertical-velocity component. During the echo-signal processing for each range gate, the scattering intensity and the velocity are evaluated. To distinguish the echo signal from the noise, the noise level is evaluated in adjacent frequency bands. The two key outputs are (i) profiles of vertical-velocity variance and (ii) the diagnosis of ABL depth when it is within the sodar’s range using the minimum in backscatter gradient profile (Wood et al. 2012).

A pulsed scanning Doppler lidar (Pearson et al. 2009) has been operating at Kumpula (on FMI roof) since 1 September 2011 operating at 1.5-mm wavelength and equipped with a depolarization channel receiving backscattering signal from aerosol particles and hydrometeors (i.e., cloud/fog droplets and ice crystals). Primary data processing includes noise removal from clear-sky data based on signal-to-noise ratio and a standard self-calibration procedure (O’Connor et al. 2004). The observation range is 90–9600 m, with 30-m resolution. Low aerosol concentration—typical for poleward sites—affects lidar data quality, which requires careful optimization of integration time and beam focus. Data availability is variable but, since typically a 5-s interval or longer time is needed to improve the signal-to-noise ratio, the current integration time is 20 s. In addition to vertical profiles of aerosol and vertical velocity, Doppler scanning (beam swinging) techniques enable the collection of vertical profiles of horizontal wind speed vectors. Novel scanning strategies may enable the study of estimates of a horizontal transect of horizontal wind vectors (Wood et al. 2013b), turbulent mixing and dynamics of the ABL (Barlow et al. 2011), and sensible heat and momentum flux (Collier et al. 2005). Performance of the lidar was investigated during a 2-week intercomparison campaign in Helsinki (Hirsikko et al. 2013) against FMI’s other Doppler lidars showing good agreement for the wind vector profiles. From 2013 onward, these ABL wind and aerosol profiles will be complemented with a radio acoustic sounding system (SODAR-RASS) instrument for temperature and wind profiles.

FMI has five operational ceilometers within 50 km of central Helsinki—Kumpula, Helsinki–Vantaa airport (~14 km north), Isosaari island (~8 km south–east and ~7 km off the Finnish coast), Nurmijärvi (~50 km north of downtown Helsinki), and Porvoo (~50 km east)—but are configured to provide just cloud base. Two ceilometers have continuously logged profile measurements at Kumpula from 22 June 2009 onward (Vaisala CL31) and at the FMI radiosounding observatory of Jokioinen (about 100 km northwest of Helsinki; Vaisala CL51). In addition, some investigations have been performed using ceilometer profile data from specific experiments in suburban Helsinki (Vaisala, Vantaa; Fig. 1) to estimate ABL depth (Eresmaa et al. 2006, 2012).

Two boundary layer scintillometers (BLS900) were installed in central Helsinki. A 4.1-km “city scale” path has been operating since 5 July 2011 on a near south–north path from downtown (Torni) to semi-urban Kumpula (198° bearing) and the beam is about 40–60 m AGL (~10–40 m above building height). The second downtown BLS was installed on 2 March 2012 in a roughly east–west 1.8-km path across downtown Helsinki at a height of 50–70 m AGL (271° bearing) from the Fire Station mast to the Sitra building. The Sitra building is located on the western edge of the densely packed buildings downtown. The source area of the second scintillometer is downtown and thus qualitatively corresponds to much of the source area of the downtown (Torni) EC measurements, where building heights are 19–29 m. These scintillometers give a spatially integrated value of the refractive-index structure parameter \( C_n^2 \) and cross-beam horizontal wind component. The parameter \( C_n^2 \) can readily be converted to \( C_{\text{eq}}^2 \), but conversion to sensible heat flux requires many assumptions and is most readily applied only during free convection (Zeweldi et al. 2009). The path weighting of the instrument is such that most of the signal comes from the center of the beam (Hill and Ochs 1978; Scintec 2011). The scintillometer measurements require good visibility; thus, no data come during fog or precipitation (Uijlenhoet et al. 2011).

An infrared camera is installed on the Fire Station mast with a westward view across downtown in order to cover a part of the footprint of the Torni EC and the downtown scintillometer measurements. The camera has 48 × 47 pixels and 60° × 60° view area. This results in effective horizontal resolution of about
10 m close to the camera. The images are acquired and stored every 5 s. No corrections are yet applied to the camera data. Despite the absolute thermal accuracy being poor (±3 K), thermal inhomogeneities within a fraction of a degree can be measured. The camera is affected by the thermal anisotropy of the urban surface and thus should be used with care. Therefore, it is more useful for relative comparison of thermal sources in any investigated direction in order to help for the analyses of scintillometer data or case studies.

These spatially averaging instruments are a key addition to point measurements in order to assess model grid values. Helsinki UrBAN will support FMI’s goal of developing within the High-Resolution Limited-Area Model (HIRLAM)—Aire Limitée Adaptation Dynamique Développement International (ALADIN) framework its new operational nonhydrostatic HIRLAM–ALADIN Research on Mesoscale Operational NWP in Euro–Mediterranean Partnership (HARMONIE) limited area model with 2.5-km horizontal resolution that includes the urban Town Energy Balance module (Masson 2000). Supporting information on model developments along with more details on instrument manufacturers and a summary of some auxiliary instruments to Helsinki UrBAN are provided in the online supplemental material.

**OBSERVATIONS OF HELSINKI’S ATMOSPHERE.** We now present examples as a testimony to the potential of the network.

**Seasonality and diurnal cycle.** One of the main drivers determining the climate of Helsinki is the strong seasonality of net all-wave radiation $Q^*$. A greater daytime $Q^*$ is observed downtown than at semi-urban Kumpula for most times in the year and day (Fig. 2a). In late winter (January–March), the lower $Q^*$ values are mainly caused by the snow cover, which increases the surface albedo and causes a large amount of $K^\uparrow$ (Fig. 2b), especially at the semi-urban Kumpula site, whereas less snow is observed downtown. We hypothesize that the large difference is due to (i) the extensive snow clearing from roads, footpaths, and roofs downtown and (ii) energy balance differences such as increased long wave radiation from the vertical snow-free walls, building heat storage, and anthropogenic heat emissions. During snow-free periods (May–October), the surface albedo is slightly higher at Kumpula (0.13 ± 0.01) compared to downtown (0.11 ± 0.01). This higher albedo can be explained by the higher fraction of vegetation cover in Kumpula as vegetation has typically higher reflectivity than built surfaces (Oke 1987). Interestingly, only slightly higher upward longwave radiation is measured downtown than Kumpula (not shown). Lower $Q^*$ in Kumpula could also be explained by atmospheric pollutants or humidity, which might affect the downward radiation components especially in spring when typically intensive resuspension of road gravel takes place (Kupiainen et al. 2011). However, since Helsinki’s air pollution is typically relatively low (except in spring; Järvi et al. 2009b), this would require more detailed investigations with respect to, for example, wind direction, thus yielding possible explanations from the prevailing downwind position of Kumpula with regard to downtown during these months and/or from local sources of pollution, moisture, or secondary organic particles (from vegetation). The large wintertime difference in $K^\uparrow$ between downtown and Kumpula emphasizes the importance of snow in the radiation balance and further in the energy partitioning and in energy balance models. Previously, this has been only marginally studied (Lemonsu et al. 2008, 2010; Bergeron and Strachan 2012), and thus our network could provide important

![Fig. 2. Median monthly (a) net all-wave radiation and (b) upwelling shortwave radiation for 2011 at Kumpula and downtown (Torni) calculated for times that are daytime throughout the year (0800–1200 UTC). The error bars show quartile deviations during the plotted hours.](image-url)
data for developing and assessing the parameterization of snow processes in atmospheric models.

The pattern of net all-wave radiation translates into the sensible heat flux $Q_H$, which also experiences strong annual and diurnal variation (Fig. 3). First, it is perhaps obvious that the greatest sensible heat fluxes, of above 150 W m$^{-2}$ at both sites, are observed in May–August during the daytime hours of 0600–1400 UTC. At Kumpula, there is little annual variation of sensible heat flux by night: the average is negative, although the average nocturnal sensible heat flux in winter is near zero. Downtown, the winter nights exhibit a mean positive sensible heat flux, although some cases of negative sensible heat flux can occur (see later case study). The strongest diurnal cycles in sensible heat flux occurred in spring and summer (especially downtown). In winter, there is very little diurnal variation in the mean sensible heat flux at both sites especially under the cloudy snow-free conditions of November–December 2011. This is consistent with the near-zero net all-wave radiation (Fig. 2). When comparing the two sites, we see how the difference in sensible heat fluxes has a patchy pattern: the largest values occurring in late April–early May, between 0800 and 1200 UTC, with fluxes at Kumpula higher by over 50 W m$^{-2}$ than downtown (also in 2012; not shown). At the same time, very high latent heat fluxes were not observed downtown, so we hypothesize these low $Q_H$ values downtown could be caused by heat storage flux to the cold building walls in spring. The storage heat flux typically peaks before midday and thus reduces available energy consumed in $Q_H$ (Grimmond et al. 1991). This can be verified in the future when the seasonal behavior of the storage heat flux will be examined by modeling approaches. A diurnal pattern is also seen when comparing the two sites: greater daytime sensible heat fluxes are observed downtown than for Kumpula; this is likely a result of several factors including higher $Q_*$, anthropogenic heat emissions and heat storage, advection from the sea, and energy partitioning caused by a high fraction of vegetation at Kumpula.

**A clear diurnal-cycle case study: 4 September 2011.** To observe the evolution of the Helsinki ABL characteristics during a diurnal cycle, a day is chosen in the autumn with high atmospheric pressure (maximum 1015 hPa) and fair-weather cumulus (Fig. 4). A clear diurnal cycle is observed in many variables, such as relative humidity and upwelling longwave radiation, notwithstanding some additional synoptic changes, increase in wind speed and veer of wind with time. Patterns in the thermal infrared (IR) camera data clearly show the main features of a warm surface by day and cool by night (Fig. 5). However, there is variability across the urban surface, primarily since surfaces facing different directions receive different solar irradiance. By day, a sharp temperature difference can be seen between the urban surface and the atmosphere, and a gradient in the atmosphere can also be seen, which is consistent with nearby vertical profiles of temperature (Fig. ES1 in the online supplement). The nocturnal IR data show a much more even distribution of temperature across the urban surfaces, compared with the values from the daytime. One possible implication is that modeling of daytime ABL conditions will require either high horizontal resolution or suitable aggregation of inhomogeneous forcing from the surface, while at

![Fig. 3. Mean sensible heat flux for each 30-min period for each month of 2011 at (a) Kumpula and (b) downtown (Torni) and (c) their difference (same color scales in all subplots). Sunrise and sunset times are shown as thick dashed black lines. Zero sensible heat flux is shown as a white line. For all 30-min periods in 2011, negative sensible heat fluxes occurred 48% of the time at Kumpula and 13% of the time downtown (Torni).](image-url)
The evolution of ABL depth follows the expected pattern for clear-sky conditions (Fig. 6). A shallower ABL is observed with the sodar for stable/nighttime conditions between 100 and 150 m (unobservable by the lidar because of blind region of 0–90 m). By day, the ABL grows to a depth of greater than 400 m and is thus not observable with the sodar; the vertical profile of velocity variance in the lidar data indicates that the ABL depth was probably over 1 km (crude estimation based on vertical-velocity variance limit of 0.1 m² s⁻²; Barlow et al. 2011). There is temporal agreement in the turbulence regimes in the lidar and sodar: turbulence maximum by day and quiescent by night. However, because of signal-to-noise constraints in the sodar velocity estimates, a quantitative comparison of vertical velocity cannot be made on this occasion. It is a major strength of an integrated observation network that we are able to observe the majority of shallow and deep ABLs by using both sodar and lidar at the same site but also to quantify on a long-term basis their differentiated observation methods of atmospheric vertical

**Fig. 4.** Conditions in Helsinki on 4 Sep 2011, at Kumpula (black) and downtown (Torni; red) for (a) net all-wave radiation, (b) wind speed, (c) shortwave radiation, (d) wind direction, (e) longwave radiation, and (f) Kumpula pressure (blue) and relative humidity (green). Upwelling and downwelling radiation is marked with dashed and solid lines, respectively.
structure inhomogeneities as a function of synoptic, stability, and surface conditions. Nevertheless, on some occasions the ABL depth cannot be determined since it is below even the height range of the sodar. Thus, a new lidar scanning technique within Helsinki Urban to cover these events is under development, since shallow ABLs are especially important for urban air quality.

The diurnal evolution is clearly seen in the time series of structure parameter of temperature (Fig. 7b). The greatest values (above $10^{-2}$ K$^2$ m$^{-2/3}$) are seen in the daytime, and the greatest values occur downtown as we would expect on the basis of greater positive sensible heat fluxes (Fig. 7a). On the other hand, the scintillometer curve does not follow the average $C_T^2$ values retrieved from downtown (Torni) and Kumpula sonics (at both ends of the scintillometer path), except late morning. We hypothesize that there is always some temperature gradient along the longer scintillometer beam due to the difference in stability cycle between downtown and semi-urban sites. This motivates detailed interpretation of the scintillometer path weighting together with footprint analysis and for additional point measurements between Kumpula and downtown.

The daytime unstable stratification lasts several hours longer downtown, particularly due to a later evening transition. The Kumpula time series of $Q_H$ goes from unstable through neutral (with $C_T^2 < 10^{-4}$ K$^2$ m$^{-2/3}$) to stable (with $C_T^2 = 10^{-3}$ K$^2$ m$^{-2/3}$ after 1900 UTC), while downtown (Torni) the heat flux and $C_T^2$ show a longer period of neutral conditions. Persistence of neutral conditions may have been caused by greater anthropogenic heat release downtown and/or the evening increase in wind velocity observed downtown, but not at Kumpula (Fig. 4b). The
The diurnal maximum in anthropogenic heat is usually observed after working hours, which coincides with the persistent positive $Q_H$ downtown between 1500 and 1800 UTC (Sailor 2011). It is noteworthy that sensible heat flux values are slightly positive or zero from 0000 to 0600, while at 1600–2400 they go negative (stratification increasing) at both downtown and Kumpula sites. On the other hand, $C_T^2$ values are closer together in the morning but strongly differ in the evening with the downtown values plummeting before settling. The semi-urban (Kumpula) $C_T^2$ curve resets to its level as per morning (consistent with the same ABL depth estimates in morning and evening from sodar), while downtown the low $C_T^2$ value might indicate a different ABL height. These data show the interest of combining the different measuring approaches, which can reveal subtle differences even within a few kilometers range.

There are also possible complexities introduced by mesoscale changes such as topographical differences, varying closeness of sea bays along the path. Most scintillometer work has been conducted over homogeneous terrain (Moene et al. 2009). These preliminary results of structure parameters determined from different techniques over heterogeneous terrain indicate an overall qualitative agreement (Wood et al. 2013a); on the other hand, subtle differences open opportunities for more detailed investigations and understanding of intra-urban processes.

Given the evolution of $C_T^2$ and $Q_H$, it is interesting that stable boundary layers might be coupled to a highly urbanized surface (<10% occurrence downtown of negative $Q_H$). Even downtown (Torni) has occasionally stable flow (Nordbo et al. 2012a), perhaps partly explained by a small anthropogenic heat flux, more sustainable use of energy, and climatology.

A stable-atmosphere case study: 3 January 2012. Given the interest in stably stratified flow above a city, a second case study was chosen specifically where the Kumpula station had negative sensible heat fluxes even during daylight hours while the downtown station (Torni) had negative fluxes at night and for a short period in the afternoon (Fig. 8a). Despite the negative $Q_H$, the buoyancy term is very small in the turbulent kinetic energy budget (Fig. 8c) and it does not destroy the mechanically generated turbulence that is balanced out mainly by dissipation. The large mechanical production is urban specific, since the rough surface induces a high momentum flux. As a result, atmospheric stability is suppressed toward neutral in comparison with a less rough surface with an equivalent sensible heat flux. Some of the changes in stability are consistent with changes in cloud cover (Fig. 9), such as the stability going more stable between 0400 and 0700 UTC, when the sky was clear (Illingworth et al. 2007). There are few reports of stable stratification over cities (Fisher et al. 2005; Wood et al. 2010; Bergeron and Strachan 2012); our observations of stable stratification (13% in 2011 at Torni; Fig. 3) motivate further our network due to a link to air quality: the largest particle fluxes and concentrations are observed at times of stability transition from stable to neutral (Figs. 8d,e). Nevertheless, it is hard to distinguish the effect of atmospheric stability versus that of traffic release since the stability transition and rush hour take place around the same time. Analyzing together atmospheric stability with traffic data will lead to an improved ability to predict...
Helsinki’s air quality. One of the main purposes of the network is to improve the description of urban surface and the ABL in air-quality models, since stable urban ABLs are rarely observed in cities.

The lidar’s custom scanning strategy allows a comparison between lidar and mast winds (Kumpula and downtown). The lidar operated a 10-s sample every 5 min pointing due south at zero elevation. The data from the first usable range gate (gate 4) was compared with the northward component of horizontal wind from Kumpula (Fig. 10). Reassuringly a small rmse of 0.5 m s\(^{-1}\) (with respect to 0–8 m s\(^{-1}\) v-component winds on this day) was observed for the 30-min mean values on this day; indeed we would even expect some difference given the different spatial and temporal sampling of the two instruments. This highlights the potential of ground-based scanning lidar to reveal flow features over a city.

Fig. 8. Time series on 3 Jan 2012 of (a) sensible heat flux and (b) Monin–Obukhov stability parameter at Kumpula (black) and downtown (Torni, red), (c) turbulent kinetic energy budget at Kumpula, (d) particle flux at Kumpula, and (e) particle concentration at Kumpula. Data with thick lines and x markers in (a),(b),(d),(e) have stringent quality assurance, such as flux nonstationarity and mast interference (Nordbo et al. 2012a); lighter dashed lines have less stringent quality assurance. The transport term in (c) is calculated as the residual of the other terms: time derivative of turbulent kinetic energy (black square), shear production in the streamwise (blue circle) and crosswind (blue dot) directions, buoyancy (red squares), and dissipation (black crosses).
SUMMARY. The importance of long-term research-grade observations of the urban ABL has led to the development of Helsinki UrBAN (http://urban.fmi.fi). This observation network has begun to improve our understanding of Helsinki’s ABL. Components of the network have been running since 2004, with substantial expansion in 2010–2013. New equipment has been installed during the last few years: EC stations, scintillometers, sodar, lidar, and thermal IR camera.

Naturally, the network also has inherent limitations. (i) Although there are numerous non-urban background reference sites, these either have not been equipped with the same instrumentation as the corresponding urban sites (e.g., the radio tower measurements at the site of Kivenlahti, about 10 km west of downtown Helsinki; see online supplemental material) or are located at a substantial distance (e.g., SMEAR-II station at Hyytiälä, southern Finland, about 200 km north; Suni et al. 2003). (ii) The coastal location of Helsinki complicates the distinction of urban effects (Lowry 1977). However, 38% of global population live within 50 km of the coastline (Kay and Alder 2005), increasingly many of them in urban environments, so, although the meteorology of coastal regions is more challenging to study than that of homogeneous inland sites, there is an evident need to study it (Mestayer et al. 2005). (iii) There are currently no in-canopy-layer measurements that would allow linking ABL climates to the canopy layer. Helsinki was selected as the location of this network mainly since it forms the largest urban agglomeration in Finland (the Helsinki metropolitan area) and also because of logistical reasons (the location of the headquarters of the FMI and the University of Helsinki).

Data were shown to represent the potential of intensive state-of-the-art observations of the urban ABL. The results showed the range of stabilities for diurnal and annual cycles both downtown and at a semi-urban site: in particular, the seasonal cycle is pronounced. These results also show clear differences between sites at even modest distance, thus calling for more studies for better understanding of Helsinki’s ABL processes affected by a range of surfaces. The examples in this paper have highlighted, for example, negative sensible heat fluxes over a city center with very low vegetation fraction, the applicability of urban Fig. 9. Vertical profiles from lidar at Kumpula on 3 Jan 2012. (a) Lidar backscatter values of $10^{-7}$ to $5 \times 10^{-4}$ are typically aerosol (blue and green), while values above $10^{-4}$ are cloud droplets (red). (b) Variance of vertical velocity, with 30-min mean threshold of $0.1 \text{ m}^2 \text{s}^{-2}$ (red) as a crude ABL depth estimator.

Fig. 10. Comparison of 30-min mean northward $v$-component horizontal winds from sonic anemometer at Kumpula and downtown (Torni) with lidar beam pointing from Kumpula due south (range gate 4: 90–120 m) on 3 Jan 2012. Statistics between lidar and sonic anemometer at Kumpula from the 30-min data on this day are correlation coefficient of 0.96, root-mean-square difference of 0.52 m s$^{-1}$, and bias of $-0.32 \text{ m} \text{s}^{-1}$ (sonic anemometer greater than lidar).
scintillometry, applying scanning lidar over a city, and the combination of sodar and lidar to give a fuller range of ABL depth estimates. There is also strong indication of the important role of the snow cover in the heat balance over the city. Previous studies in Helsinki have already reported many findings, such as analyses of fluxes (Vesala et al. 2008), exploitation of morphological datasets (Nordbo et al. 2012a), analysis of several years of CO₂ fluxes (Järvi et al. 2012), and air-quality observations (www.airquality.fi; Pirjola et al. 2012).

These results will hopefully trigger modeling activities to distinguish and prioritize the various processes at hand. Beyond the forthcoming science outputs, we expect to expand this network. We anticipate that others will bring their equipment and/or expertise alongside ours for the development of technology, science, and applications.

ACKNOWLEDGMENTS. EC FP7 ERC Grant 227915 “Atmospheric planetary boundary layers—Physics, modeling and role in Earth system” and the Academy of Finland (Projects 138328, 1118615, and ICOS-Finland, 263149) provided financial support. Kari Riikonen, Erkki Siivola, Pasi Aalto, and Petri Kerönen provided technical support; Kari Niemelä extracted AMDAR data (online supplemental materials). We had useful discussions with Iolanda Iolongo, Mari Kauhaniemi, Juuso Suomi, Jukka Käyhkö, Anu Kousa, Tarja Koskentalo, Heikki Turtiainen, Mikko Laakso, Reijo Roininen, Antti Hellsten, Adriaan Perrels, Mika Komppula, Elena Saltikoff, Sue Grimmond, and Janet Barlow. We are very grateful for the insightful and constructive comments from the reviewers.

REFERENCES


HSY, 2008: SeutuCD. Helsinki Region Environmental Services Authority dataset, CD-ROM.


The atmospheric boundary layer impacts strongly the model performance for temperature and wind, yet stable situations, such as in clear, calm conditions at night or over ice, remain problematic.

The atmospheric boundary layer (ABL) is the lower part of the atmosphere that is in continuous interaction with Earth’s surface owing to friction and heating or cooling. The ABL is generally turbulent and has a pronounced diurnal cycle of temperature, wind, and related variables—in particular over land and ice. Turbulence in the ABL is three-dimensional and chaotic with time scales typically between fractions of a second and an hour. The corresponding length scales are from a millimeter up to the depth of the boundary layer (or more in the case of convective clouds). The depth of the dry ABL varies both in time and space between tens of meters up to kilometers. Over land it has a strong diurnal variation, while over the sea the depth of the ABL is typically a few hundred meters and rather constant on the time scale of a day.

While turbulence is generally strongest in the ABL, most of the troposphere contains turbulent motions. Strong turbulence is also found at high altitudes in, for example, towering cumulus clouds, which may grow into thunderstorms. In this case, the convection produced by the heat released owing to the condensation of water vapor in the cloud reinforces the turbulence. Strong turbulence may also occur in clear air above the ABL; most of this is produced in layers of strong vertical wind shear known as clear-air turbulence.

Turbulent flows in the atmosphere efficiently transport momentum, heat, and matter. The ABL and its turbulence are also important for the momentum and sensible and latent heat transfers between the surface and the atmosphere. These directly affect the diurnal cycle of the near-surface variables and also strongly impact on the life time of synoptic-scale systems. Appropriate representation of the overall effects by turbulence, either inside or outside the atmospheric boundary layer, is thus an essential part of atmospheric models dealing with the prediction and study of weather, climate, air quality, wind energy, and other environmental factors. Because of the small-scale features of atmospheric turbulence, there will always be important effects on the mean flow from scales smaller than the numerical grid cells of the models used.

To bridge the gap in scales, the equations of motions are averaged over the scales of turbulence, known as ensemble or Reynolds averaging (Reynolds 1895). In this process higher-order terms arise out
of the averaged advection terms and this makes the system of equations unclosed. The challenge is to relate the new unknown terms to the forecast variables of the model. This so-called turbulence closure is basically unsolvable on a fundamental level and several approaches have been taken in the past (e.g., Wyngaard 2010).

In its basic form a turbulence closure is based on flux-gradient theory utilizing a proper formulation of an eddy diffusivity. Such a “local” closure is found useful for stably stratified turbulence. For unstable and convective boundary layers, nonlocal mixing effects are also typically required to properly represent the mixing processes (e.g., Holtslag and Moeng 1991). Turbulence closures can be of different order depending on the order of the terms that are parameterized, but in practice few are higher than second order in operational models. A popular version is to combine a prognostic equation for the turbulent kinetic energy (TKE) with a diagnostic equation for the turbulent length scale to derive an eddy diffusivity (e.g., Mellor and Yamada 1974).

While the basic principles underlying the parameterizations are the same, the actual model implementations vary substantially among models developed and used by research groups and by operational centers. This results in large variations for the diurnal cycles of near-surface temperature and wind. The reasons behind this diversity in model formulations and strategies are not that easy to unravel. Most likely, this is for historical reasons owing to the outcome of various tuning exercises and how models have been evaluated with observations in the past (see also discussion by Jakob 2010). In addition, modelers have different standpoints on the complexity needed to represent atmospheric turbulence and vertical diffusion processes (e.g., Delage 1997; Teixeira et al. 2008; Zilitinkevich et al. 2008; Baklanov et al. 2011). An overview of most turbulent mixing and diffusion parameterizations in current use is given by Cuxart et al. (2006).

Figure 1 gives an example of the performance of the leading European Centre for Medium-Range Weather Forecasts (ECMWF) numerical weather prediction model for the 2-m temperature. Here the temperature forecast errors at day and night over Europe are shown as monthly averages during more than 20 years. The results display a rich history with many model changes, including those for the land surface, soil freezing, and vertical mixing. Improvements have been made during the years as indicated by the long-term reduction in biases and standard deviation of the errors. However, at night these measures have increased in recent years. This is due to a change in the vertical diffusion scheme for stratified conditions that was found necessary in the ECMWF model to reduce the dissipation of stratocumulus through erosion of the strong capping inversions (Köhler et al. 2011). This illustrates the difficulty encountered when atmospheric models are improved in one aspect but with unintended implications elsewhere.

Climate models also struggle to represent the correct near-surface parameters and diurnal cycles (e.g., Zhang et al. 2009; Kyselý and Plavcová 2012; Mearns et al. 2012). As an example, Fig. 2 shows errors in 2-m temperature for two versions of the National Center for Atmospheric Research (NCAR) community climate model over land north of 50°N during wintertime. The two model versions share the same dynamical core and resolution as well as land models but have a quite different suite of atmospheric parameterizations. As can be seen in Fig. 2, the 2-m temperature errors are typically positive for the Community Atmosphere Model version 4 (CAM4) and negative for CAM5. They also show large spatial variability over large areas in both models. Furthermore, wind speeds at the lowest model level are rather different in the two versions (not shown). Obviously, not all differences can be attributed to the representation of the boundary layer since cloudiness and land surface factors also play an important role (e.g., Van den Hurk et al. 2011). However, our current understanding
and capability to model stably stratified conditions is limited and most certainly influences the results (Mahrt 1987; McCabe and Brown 2007). Note also that the descriptions of the ABL are rather different in CAM4 (Holtslag and Boville 1993) and in CAM5 (Bretherton and Park 2009).

The diurnal cycles of near-surface variables in CAM4 and CAM5 have been evaluated using many years of data from flux-tower observations by Lindvall et al. (2013). Figure 3 shows an example from a similar and extended analysis with a large number of the climate models participating in the Coupled Model Intercomparison Project, phase 5 (CMIP5; Taylor et al. 2012). The figure represents models with different complexity, boundary layer parameterizations, and numerical grids and includes the two NCAR models. The diurnal cycles for near-surface wind and temperature in Fig. 3 are shown here in comparison with tower data from the Atmospheric Radiation Measurement Program (ARM) Southern Great Plains main site (Fischer et al. 2007). Notice that the diurnal cycle is given with respect to the mean value. It is seen that wind and temperature vary considerably in the individual models even in this relatively flat and homogenous region. The ensemble model median temperatures in Fig. 3 compare well with the observed diurnal cycles, but with large intermodel variations. Many factors influence these variables, such as soil moisture and clouds in addition to vertical mixing in the ABL. The diurnal cycle in the wind is generally underestimated (see Figs. 3b and 3d) and some models are out of phase, especially during summer.

To understand the basis for the various boundary layer parameterizations and to make a critical evaluation of the various schemes in use, model intercomparison studies have been organized within the Global Energy and Water Exchanges (GEWEX) Atmospheric Boundary Layer Study (GABLS; Holtslag 2003, 2006). Specific cases are chosen for which single-column versions of models (SCMs) are run with the same specifications, and the results are compared with observations and/or finescale (large eddy) simulations (LES). Such a strategy has also been found to be very useful for cloudy boundary layers (e.g., Randall et al. 2003; Neggers et al. 2012). Following the former authors, we note that an “SCM is essentially the column physics of an atmospheric model, considered in isolation from the remainder of the atmospheric model” (p. 456). As such, an SCM can be used to make a direct comparison with observations or LES given prescribed values for advection, specific surface conditions, and other
factors. The cases studied so far within GABLS are based on observations taken in the Arctic, in Kansas, and at Cabauw (the Netherlands) during clear skies.

Below, we further introduce the subject following the overview by Holtslag et al. (2012), and summarize the GABLS findings in the sections “Stable atmospheric boundary layers” and “Diurnal cycles.” Final points are provided in “Summary and prospects.”

**Stable Atmospheric Boundary Layers.** Stably stratified conditions occur frequently in the ABL over polar regions and over continental land during night and wintertime. Correct representation of the stable boundary layer (SBL) is difficult owing to the weak and sometimes intermittent behavior of turbulence (e.g., Mauritsen and Svensson 2007) and interaction with other small-scale processes (see below). Overall progress in understanding and model formulations has been slow (e.g., Louis 1979; Beljaars and Holtslag 1991; Fernando and Weil 2010; Baklanov et al. 2011). It has also been known for quite some time that numerical weather prediction and climate models show great sensitivity to the model mixing formulations in these conditions.

Viterbo et al. (1999) performed a sensitivity study with the ECMWF model using (slightly) different formulations for describing the impact of increased stability on the damped turbulence for stable conditions (these formulations are known as the stability functions). Even with the same forcing conditions, they noticed large differences in the mean January 2-m temperatures over the Northern Hemispheric continental areas. The sensitivity study was recently repeated with the 2011 version of the ECMWF model. Results for both model experiments are shown in Fig. 4. The sensitivity experiments were for the 1995/96 winter season starting from 1 October 1995 and applying relaxation to the 6-hourly operational analyses above 500 m from the surface. This is an efficient way of doing “deterministic” seasonal integrations without constraining the boundary

**Fig. 3.** Observed and modeled diurnal cycles of 2-m temperature (°C) and wind speed (m s⁻¹) with respect to their daily means for the ARM SGP main site (36.6°N, 97.5°W) in (a),(b) winter and (c),(d) summer. The model results are from AMIP simulations (atmosphere-only simulations with observed sea surface temperature and sea ice concentration for the period 1999–2008) by CMIP5 models (colors, 16 models for temperature and 12 for wind speed) including the model median (dashed thick line). Median (solid thick line) and 25th and 75th percentiles (gray area) of observed diurnal cycle minus daily mean for the period 2002–09 are also shown.
layer. It appears that the same change in the stability functions has a much larger impact in the 2011 model version than in the 1994 version. In the meantime, many model changes have been made but most likely the different sensitivity is related to the updated soil hydrology scheme (Balsamo et al. 2009) and the new snow scheme (Dutra et al. 2010). In the latter, snow is a much better insulator and therefore the winter temperatures are lower. This illustrates the tight coupling between boundary layer processes and land surface and snow feedbacks, which obviously needs further attention and research (see also Sterk et al. 2013).

Besides a tight interaction of atmospheric turbulence with the land surface, stable boundary layers are influenced by other small-scale processes and phenomena such as radiation (divergence), fog and dew formation, drainage flow, gravity waves, and low-level jets. In addition, the morphology of stable boundary layers is quite diverse—for example, shallow and deep boundary layers with continuous turbulence through most of their depth and boundary layers with very weak and/or intermittent turbulence in very stable cases at night (e.g., Mahrt 1985; Van de Wiel et al. 2003, 2007). On the other hand, surface heterogeneities and topography are factors typically enhancing momentum transport over land (e.g., Cuxart and Jiménez 2007; Martínez et al. 2010; McCabe and Brown 2007).

The first GABLS intercomparison case was designed to document and better understand the differences between the various boundary layer schemes in numerical weather and climate models using an idealized case focusing on the representation of turbulence. The case is based on the results from the Arctic originally presented by Kosović and Curry (2000). The stable boundary layer in the SCMs was driven by an imposed, uniform geostrophic wind of 8 m s⁻¹, with a specified surface-cooling rate of 0.25 K h⁻¹ and an overlying capping inversion. The same case was run by a range of SCMs and LES models and the main results are presented in Cuxart et al. (2006) and Beare et al. (2006), respectively. Overall, it turns out that with the same initial conditions and model forcings, the results of the LES models are surprisingly consistent when a vertical resolution of 6.25 m was utilized (Beare et al. 2006). Thus, the LES results can serve as a suitable reference for the turbulence representation in the single-column models (since other processes like radiation and land surface schemes are not active in this case).

The results by the participating single-column models (colored lines) indicate a large range in vertical structure for the mean temperature and wind magnitude profiles (Figs. 5a,b) in comparison with the LES results (black lines). In addition, the hodographs are shown in Fig. 5c. In the latter figure, a selection of 10 out of the 19 participating SCMs in GABLS1 was made that showed a consistent behavior between the surface and boundary layer, following the analysis by Svensson and Holtslag (2009). The models not selected for Fig. 5c are shown as dashed lines in Figs. 5a,b and these typically show a larger deviation from the LES reference.

In Figs. 6a,b the turbulent heat and momentum fluxes are given and these show a rather large range owing to the various parameterizations. Overall, the models in use at operational weather forecast and climate centers provide deeper boundary layers and allow for “enhanced mixing” (see also below), resulting in larger fluxes, while the research models show less mixing in more agreement with the LES. Note that the complexity of the turbulent scheme does not seem to matter here; even a relatively basic local diffusivity scheme can do well for this simple case (e.g., Steeneveld et al. 2006a). Because of the
enhanced mixing in weather and climate models, these models tend to show too strong surface drag and too deep boundary layers. This typically results in erosion of low-level jets and the underestimation of the turning of wind with height in the lower atmosphere (Svensson and Holtslag 2009). It is clear that this directly leads to errors in any application such as for air quality and wind energy. It also explains why ECMWF has relaxed the mixing in their model to avoid eroding stratocumulus capping inversions (see the above discussion with respect to Fig. 1).

Returning to Fig. 5c for the hodographs, we note a clear lineup such that operational models have the smallest turning of the wind in the boundary layer, followed by the LES results placed in the middle of the research model results. The grid points in the lowest 10% of the SBL (in the surface layer) are indicated with dots to show that even in the surface layer there is a turning of the wind in contrast to what is often assumed. The shape of the spirals depends on how the turbulent stress is parameterized, which varies significantly among the participating models (Cuxart et al. 2006). It can be shown that the angle between the surface wind and the geostrophic wind is directly related to the depth of the turbulent boundary layer such that deeper (shallower) boundary layers have smaller (larger) surface angles (Svensson and Holtslag 2009; Grisogono 2011). The operational models with enhanced mixing and a deeper boundary layer also have a larger integrated cross-isobaric flux, and this directly impacts the larger-scale flow through “Ekman pumping.”

The boundary layer scheme in any atmospheric model is responsible for surface drag that feeds back to the large-scale flow through the momentum budget and through the ageostrophic flow. In general, boundary layer formulations with enhanced mixing tend to give better performance for the larger-scale flow and as such this has been a motivation to use these. Also, the momentum budget aspect is an important contributor to the sensitivity of large-scale scores of weather forecast models to the formulation of the boundary layer scheme. The mechanism behind this sensitivity is, however, not well understood. It is known that large drag damps weather systems and reduces the “activity” of a model, which tends to be good for operational scores possibly by compensating for other deficiencies. By decreasing the mixing and surface drag, a direct impact on the atmospheric dynamics has been noted (e.g., Beljaars and Viterbo 1998). Consequently, cyclones may become too active (e.g., Beare 2007), resulting in too high extremes for wind speed and precipitation (see also Sinclair et al. 2010).

In GABLS1, the ensemble of results by LES models was used as the reference for the single-column model.
results (as indicated in Figs. 5 and 6). The GABLS1 case was, however, highly idealized and not directly comparable to observations. Basu et al. (2012) set up an intercomparison of LES models with observations from the Cabauw tower in the Netherlands as a part of the diurnal cycle study within GABLS3 (see next section). The Cabauw site, with its 200-m meteorological tower, is situated in a relatively flat environment dominated by grassland. On many nights a low-level jet develops because of decoupling and inertial oscillation. The GABLS3–LES case involves a total simulation period of 9 hours (0000–0900 UTC 2 July 2006). This period essentially encompasses the development of a moderately/strongly stratified, baroclinic, midlatitude nighttime boundary layer as well as its transition into a daytime convective boundary layer. Here we show some results for the night time hours and we refer to Basu et al. (2012, 2013, unpublished manuscript) and Moene et al. (2011) for additional information.

The initial conditions for the LES runs at 0000 UTC were created by merging the observed 200-m Cabauw tower data, wind profiler data, and a high-resolution sounding from De Bilt. Time–height-dependent geostrophic wind forcing was derived from a network of surface pressure stations combined with the analysis of a mesoscale weather forecasting model. In a similar fashion, time–height-dependent advection terms were also obtained from the Regional Atmospheric Climate Model (RACMO) and Weather Research and Forecasting Model (WRF) forecasts and observed trends at the 200-m level during night time (Baas et al. 2010). For the LES study, observed (extrapolated) near-surface (0.25 m above ground level) potential

---

**Fig. 7.** LES model results (with a vertical resolution of 6.25 m) and observations for the GABLS3 case study corresponding to 0300–0400 UTC 2 Jul 2006. Results are for (a) potential temperature (K), (b) wind speed magnitude (m s⁻¹), (c) sensible heat flux (W m⁻²), and (d) momentum flux (N m⁻²). The red dots with whiskers represent median and min–max values of the observations from the Cabauw meteorological tower. Data from a wind profiler are depicted by blue squares. The solid black lines, dark gray shaded areas, and the light gray areas correspond to the medians, 25th–75th percentile ranges, and 10th–90th percentile ranges of the LES ensemble-generated output data, respectively. The simulated profiles from a very-high-resolution LES run (vertical resolution of 1 m) are denoted by the green dashed lines. Results adapted from Basu et al. (2012, 2013, unpublished manuscript).
temperature and specific humidity were prescribed as lower boundary conditions. Thus, the differences among the LES models do not depend on the lower boundary conditions but on differences in numerics and subgrid closures.

Eleven LES groups from different international institutes provided results for the GABLS3–LES case. Figures 7a–d show the results by the LES models for 0300–0400 UTC for wind magnitude, temperature, and their corresponding turbulent fluxes. Overall, the mean wind, temperature, and related turbulent flux profiles are captured very well in the simulations, indicating that LES is a useful reference for SCM intercomparison studies in weak to moderately stably stratified boundary layers, as was anticipated in GABLS1.

**DIURNAL CYCLES.** As discussed above, operational weather and climate models have difficulty in representing the diurnal cycle of temperature, wind, and related variables (see also Edwards et al. 2011). This has an impact on the modeled temperature trends by climate models, in particular for the minimum temperature (Zhou et al. 2010; Steeneveld et al. 2011a; McNider et al. 2012). Steeneveld et al. (2008b) compared the performance of three state-of-the-art mesoscale models and noted that all three models underestimate the amplitude of the diurnal temperature cycle and the near-surface wind speed. These findings were achieved by comparing the models with observations taken in Kansas in the early autumn during the 1999 Cooperative Atmosphere–Surface Exchange Study (CASES-99; Poulos et al. 2002). Two consecutive clear days from these data with a strong diurnal cycle over relatively dry land were selected for the intercomparison study, inspired by an earlier single-column study by Steeneveld et al. (2006b).

The CASES-99 dataset was used to set up an intercomparison case for SCMs in GABLS2. The forcing conditions were simplified to facilitate a more straightforward comparison between model closures. As such, a prescribed surface temperature and simplified time-dependent barotropic geostrophic wind forcing was used (Svensson et al. 2011). Nineteen models participated in this SCM intercomparison study, ranging from operational models with first-order closure and a vertical resolution of six grid points within the first 400 m to more advanced models with much higher...
resolution. The analysis of the model results were performed according to their turbulent closure and the height of their first model level, below or above 5 m above the surface. Results from one experimental LES by Kumar et al. (2010) are also available for this case.

It is found that the models produce very different results in all variables and that they all differ substantially from the observations. Although the surface temperature has been prescribed, a large variation is seen in the diurnal cycle of 2-m temperature with most models overestimating the amplitude while the LES has a smaller amplitude than observed (Fig. 8a). The modeled diurnal cycle of the 10-m wind speed does not resemble the observations in many cases, most models overestimate the wind speed during night, and the speed does not increase enough after the morning transition (Fig. 8b). This is also seen in many of the CMIP5 climate models, especially for wind speed (Fig. 3).

The impact of forcing and boundary conditions on the variability of model results for GABLS2 is discussed by Holtslag et al. (2007). Interestingly, it was found that the variation between various model permutations is less when the boundary layer scheme is coupled to a well-performing land surface scheme. Thus, prescribing the surface temperature as in GABLS2 was, in the end, a more critical test for the boundary layer schemes than when allowing for surface interactions. One may therefore speculate to what extent differences in boundary layer schemes are at least partly a result from tuning them together with other process parameterizations in the different weather and climate models.

The experience from the two first GABLS cases led to the setup of the third SCM intercomparison case using data gathered at the Cabauw tower (Baas et al. 2010). In the previous studies it was found that especially the complexity of real-world large-scale forcing and the lack of interaction with the surface hampered a direct comparison of models with observations. Thus, the GABLS3–SCM case involves a more realistic large-scale forcing and allows for interactions with the land surface and atmospheric radiation.

For the GABLS3–SCM case, the early afternoon of 1 July 2006 was chosen as an initial time. The total simulation of 24 h covers the decoupling around sunset with low-level jet formation and the following morning transition. Note that the LES case of GABLS3 as discussed in “Stable atmospheric boundary layers” encompasses part of the SCM simulation period. The observations show an almost clear sky with a reasonably constant geostrophic wind over time of about 7 m s⁻¹, resulting in a turbulent stable boundary layer overnight with a pronounced temperature drop and a well-developed low-level jet around 200 m. To make a valid comparison with observations possible, care was taken to prescribe realistic geostrophic forcing and dynamic tendencies for the SCMs (Baas et al. 2010). The description of the third GABLS–SCM case, details of the selection criteria, and the composition of the large-scale forcing are documented in Bosveld et al. (2013b, manuscript submitted to Bound.-Layer Meteor.).
et al. (2012, 2013a, manuscript submitted to Bound.-Layer Meteor.).

Nineteen models of different complexity from 11 institutes participated in the intercomparison. Twelve of these models also participated in GABLS2. Figure 9a shows time series of the 2-m temperature from the SCMs together with the observations. The general signature of the temperature change is well captured by the models—that is, an initial fast decrease, followed by a more gradual decrease in the subsequent hours, and slightly faster cooling before midnight. Seven out of the 19 models are within 1 K of the observations. The remaining models are up to 5 K colder than observed, which seems mostly related to coupling of the atmosphere to the surface. It appears that variations in thermal land surface coupling among the models explain to a large extent the variations in the minimum 2-m temperature for the GABLS3 case. Variations in turbulent mixing and representation of longwave radiation seem to be of lesser importance for this parameter (Bosveld et al. 2012, 2013a, manuscript submitted to Bound.-Layer Meteor.). This issue obviously needs further research.

Winds at the 200-m level are shown in Fig. 9b. For each model, the first level above 200 m was chosen. This height interval is interesting because in the observations it is decoupled from the surface. After the evening decoupling, the observed wind accelerates much faster than the modeled winds, which is related to the timing of the evening transition and the corresponding wind profiles at that time (see also Van de Wiel et al. 2010). The inertial oscillation is also strongly affected by horizontal momentum advection especially after midnight (Baas et al. 2010). All model wind speeds peak 11 h after the start of the simulation at lower values than observed. Around and after sunrise models start to differ even more, both from each other and from the observations. At the 80-m level, which is well within the turbulent layer, a number of models peak at higher wind speed than observed (not shown).

Finally, Fig. 10 shows time series for temperature and wind at a height of 40 m for the LES models in comparison with the observations for the GABLS3 case. The 40-m height was chosen here as a representative level in the middle part of the nighttime stable boundary layer (typically six grid points away from the surface to ensure that the resolved turbulence is not impacted by the surface). Overall, the agreement between the ensemble of model results and the observations is very good for this realistic case, in particular if one compares the outcome with the results of the SCMs for GABLS2. Note again that in both cases the near-surface temperature was prescribed in accordance with the observations. However, coupling of an LES model to an interactive land surface scheme may result in similar discrepancies as seen in Fig. 9 for the SCMs in GABLS3; this calls for further investigation.

**SUMMARY AND PROSPECTS.** The representation of the atmospheric boundary layer in state-of-the-art weather forecast and climate models has important practical implications for many users within air quality, wind energy, climate, and Earth system studies. In fact, model output of near-surface weather parameters is increasingly being supplied to users either directly or with some statistical postprocessing. Overall, the diurnal cycles of temperature and wind are strongly influenced by processes in the atmospheric boundary layer, in particular by turbulent diffusion and radiation, but also by the thermal coupling to the underlying surface through vegetation and snow (as illustrated here for the GABLS3 case). This contribution elaborates the state-of-the-art in these areas with particular emphasis on stable boundary layers over land and ice at clear skies.

As discussed in the paper, the performance of weather and climate models is sensitive to the details of the boundary layer formulation. Most large-scale
atmospheric models utilize overly diffusive boundary layer schemes in stably stratified conditions with the result that these boundary layers are too thick, have too little wind turning with height, and underestimate the magnitude of the nocturnal jet. Climate projections show large temperature signals at high latitudes where stable boundary layers occur frequently. Findings from investigations with the ECMWF global numerical weather forecast model and analysis of some models participating in the Coupled Model Intercomparison Project phase 5 (CMIP5) are discussed to illustrate the current status and developments.

The performance of NWP and climate models during stably stratified conditions is part of the underpinning for the GEWEX Atmospheric Boundary Layer Study (GABLS) and here findings from the three GABLS model intercomparison studies are presented. Based on these results it is indeed clear that operational models typically have too much mixing in stable conditions. This strongly impacts the diurnal cycles of temperature, wind, and related variables. Enhanced mixing has an impact on the life time of the synoptic systems and thus reduces the “activity” of a model, which improves the operational scores. By decreasing mixing (i.e., reducing the surface drag to more realistic values), a direct impact on the atmospheric dynamics has been noted (e.g., Beljaars and Viterbo 1998; Beljaars et al. 2012a; Sandu et al. 2013).

Another motivation to use enhanced mixing is to prevent models going into an unphysical decoupled mode (i.e., separating the atmosphere from the cool surface). Such a decoupling may lead to a runaway cooling close to the ground (e.g., Louis 1979; Derbyshire 1999; Steeneveld et al. 2006b; Basu et al. 2008). This is one example where the turbulent mixing in stratified flows has an inherent nonlinear character and may trigger unwanted positive feedbacks (e.g., Mahrt and Vickers 2006). Such positive feedbacks, in turn, may cause unexpected transitions between totally different regimes in the stable boundary layer (e.g., McNider et al. 1995; Derbyshire 1999; Delage 1997; Van de Wiel et al. 2007, 2012a,b; Bintanja et al. 2012). The overall representation of the small-scale atmospheric processes and the related “spatial averaging” is highly nontrivial since there are many nonlinear processes involved and because the land surface often displays a heterogeneous character on a variety of scales.

The GABLS cases brought together persons with expertise on LES and observations with academic and operational modeling skills. The cases also inspired new model developments (e.g., Sukoriansky et al. 2005; Mauritsen and Svensson 2007; Buzzi et al. 2011) and are increasingly used for applications like particle dispersion (e.g., Weil 2010). Inspired by the GABLS results, modeling groups at many operational centers—such as ECMWF, the Met Office, Météo-France, National Centers for Environmental Prediction (NCEP), and elsewhere—have been encouraged to study and improve their representation of the atmospheric boundary layer (e.g., Beare 2007; Brown et al. 2008; Bazile et al. 2012). It is clear that this issue is still not fully solved and needs further attention by the modeling centers and within the academic community (see also Jakob 2010). It also appears that changes in the mixing formulation may have strong impacts on the representation of fog and clouds as well as vertical diffusion in the atmosphere above the boundary layer (Bretherton and Park 2009; Köhler et al. 2011; Steeneveld et al. 2011b; Müller et al. 2010). In the sidebar below, an overview of GABLS achievements is given.

Overall, there is still a clear need for a better understanding and a more general description of the atmospheric boundary layer in atmospheric models for weather and climate, in particular under stably stratified conditions (see also Hong and Dudhia 2012). The ultimate goal is to have a full understanding of the complexity of atmospheric boundary layers as well as a unified treatment of turbulent mixing on the different scales and surface types which occur in

**OVERVIEW OF GABLS ACHIEVEMENTS**

- GABLS has inspired academia and operational modeling centers to work together on boundary layer issues.
- The GABLS cases are increasingly used for model testing and benchmarking.
- Large-eddy simulation has become a useful tool to study stable boundary layers.
- Research models are able to represent a realistic stable boundary layer structure.
- Weather forecast and climate models generally have too much vertical mixing in stable conditions, resulting in too deep boundary layers, too less turning of wind with height, too large downward sensible heat fluxes, and too weak low-level jets.
- Operational weather forecast models still need enhanced mixing for good forecast scores but have difficulty in representing the diurnal cycles over land.
- Coupling between the atmosphere and the land surface is key for a good representation of the diurnal cycles of temperature, wind, and other variables.
To model strongly stratified conditions (e.g., Zilitinkevich et al. 2008; Van de Wiel et al. 2012a,b; Sterk et al. 2013). In the future, we see studying boundary layers that have a stronger stratification and lower geostrophic wind speeds (<5 m s⁻¹) as recommended by participants of the ECMWF–GABLS workshop (Beljaars et al. 2012b). Boundary layers over ice and snow in the Arctic and Antarctic as well as boundary layers over heterogeneous landscapes (e.g., in Lindenberg, Germany, and Sodankylä, Finland) provide additional complexities and challenges. The sidebar above gives an overview of future directions and challenges.

Finally, we recommend to study the ABL in interaction with other atmospheric and Earth surface processes (e.g., Ek and Holtslag 2004; Van Heerwaarden et al. 2009; Sterk et al. 2013), and encourage the setup of such studies within the new GEWEX program on Global Atmospheric System Studies (GASS). Neggers et al. (2012) present an interesting way to enhance process understanding by systematic comparing SCM results with observations in an operational suite. Attention should further be paid to integrate the activities with modelers at weather forecast and climate centers—for instance, by facilitating regional model intercomparisons such as in the Arctic Regional Climate Model Intercomparison Project (ARCMIP) (Tjernström et al. 2005) and to acquire and compare short-term forecasts from full weather forecast and climate models for the study points of interest.

ACKNOWLEDGMENTS. We thank all the contributors to the three GABLS model intercomparisons and the many persons involved in gathering and analyzing the field data. We also thank the participants of the GABLS workshops and meetings during the last decade as well as the hosting organizations (ECMWF, KNMI, NCAR, University of Mallorca, Wageningen University, and Stockholm University) and the various AMS Boundary Layer Symposia for providing a GABLS platform. In addition we thank GEWEX and the Working Group on Numerical Experimentation (WGNE) for their support. For CMIP, the U.S. Department of Energy’s Program for Climate Model Diagnosis and Intercomparison provides support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. The paper benefited from discussions at the ECMWF–GABLS workshop in November 2011. Michael Ek, John Edwards, Adrian Lock, and Thorsten Mauritsen are thanked for providing comments on a draft of this paper. SB acknowledges the financial support received from the National Science Foundation by way of Grant AGS-1122315.

REFERENCES


CROSS-CULTURAL INSIGHTS INTO CLIMATE CHANGE SKEPTICISM

by Peter Rudiak-Gould

Because climate change skepticism seems to stem from cultural particulars rather than human universals, climate communicators should target specific ideological obstacles to belief rather than panhuman psychological biases or scientific ignorance.

Climate scientists find themselves vexed by lingering public distrust of the scientific consensus on global climate change (Leiserowitz et al. 2012); a November 2012 poll of likely U.S. voters found that almost as many believe that warming trends are natural (38%) as manmade (41%) (Rasmussen Reports 2013). While many researchers have sought to diagnose reasons for climate change apathy, disbelief, and skepticism (Feinberg and Willer 2011; Feygina et al. 2010; Jamieson 2007; Johnson and Levin 2009; Jamieson 2007; Johnson and Levin 2009; Kahn et al. 2007; Norgaard 2006; Oreskes and Conway 2010; Stoll-Kleemann et al. 2001; Sunstein 2007; Weber 2006), rarely is the question considered in a cross-cultural light. The conjecture that climate change skepticism stems from universal features of human psychology is simply assumed a priori to be true (Jamieson 2007; Johnson and Levin 2009), or variations in climate change acceptance are considered only within a single society, usually the United States (Feinberg and Willer 2011; Feygina et al. 2010; Kahan et al. 2007). A notable and commendable exception is the work of geographer Simon Donner (Donner 2007, 2011), who focuses specifically on the public misconception that human beings cannot influence the weather (as opposed to other climate-skeptics’ arguments, such as the assertion that the planet is not in fact warming or that warming trends are benign). Donner draws from diverse ethnographic and historical sources to assess to what extent such skepticism is ancient and cross-culturally widespread, as opposed to historically novel and culturally specific. Broadly speaking, such cross-cultural comparison could lead to one of two possible conclusions about a particular misconception that science educators hope to dislodge: 1) the misconception is found in all or nearly all societies and historical epochs, and thus is probably rooted in innate or universal human psychology, or 2) the misconception is historically new or confined to a few societies, and thus probably stems from culturally specific worldviews or experiences.

The distinction I am drawing—beliefs rooted in panhuman intuitions, versus beliefs rooted in cultural specifics—obviously simplifies a complex interplay between the universal and the specific, the innate and the learned, and the biological and the cultural. But
a large body of work has demonstrated that, in many realms of research, the distinction remains a viable one. For instance, experimental evidence has shown that “folk physics”—a simple mental model of the mechanics of objects—is cross-culturally universal, develops extremely early in children, and is quickly and effortlessly deployed without formal training—in stark opposition to scientific physics, which is unique to scientifically literate societies, develops late, is deployed slowly and with much effort, and requires formal instruction (McCauley 2011). The question then is whether skepticism of anthropogenic climate change is more like folk physics, automatically commonsensical in any society, or more like scientific physics, requiring considerable cultural scaffolding to support. The implications for the science educator are deep.

In the first case, the educator is warned of the difficulty of her job. Intuitive reasoning, perhaps even innate predispositions, will need to be discarded in favor of a prima facie implausible idea. But the situation is not hopeless. Science, as we know, is radically counter to human intuitions and habits both in its methodology and in its conclusions (McCauley 2011). Yet science can be taught. Many religious belief systems, too, are radically counterintuitive (Barrett 1999), but they can be taught. Where a concept lacks “maturational naturalness” (essentially, an innate intuitiveness to the human mind), “practiced [learned] naturalness” can be built instead (McCauley 2011); indeed this is perhaps the central mission of the educational system (Pinker 2008, p. 439). One strategy is to identify other concepts that are intuitive and cleverly recombine them to communicate counterintuitive truths; Richard Feynman managed to explain, with utmost lucidity, the wave/particle duality of electrons and light (a highly counterintuitive scientific finding if there ever was one) by piggybacking on the more familiar folk physics of waves rippling in water and bullets sprayed from a machine gun (Feynman et al. 2011, chapter 1).

In the second case, in which the misconception is historically novel and cross-culturally rare, panhuman psychology is not the culprit. The work of the educator here appears easier. In some cases she may even have human intuition on her side. Anthropologist Ernest Gellner suggests that many social groups are bound together by the bitter pill of an implausible belief: the painful rite of passage of accepting what at first seems a “great absurdity” is what binds the group together (Gellner 1968, 257–258), and the implausible belief acts as an unmistakable badge of group membership because it is unlikely to be espoused by others. Here the science educator is carving with rather than against the grain of panhuman cognition. But the task is not, in reality, so easy. To say that a belief is specific to a particular society or epoch is not to say that it is brittle or half-heartedly held. When vested interests and cherished group memberships are at stake, belief resilience and confirmation bias (Nickerson 1998) set in. Richard Norgaard shows how true believers in “progress” manage to convince themselves, against mountains of counterevidence, that environmental conditions are stable or improving (Norgaard 2002). But there is still hope. The educator can reinvent himself as cultural agitator, proclaiming rather than disowning his political leanings, gently (or not so gently) challenging the ideologies that lead to scientific error. The educator can also take a less ambitious tack, to reframe the message in such a way that it no longer seems to offend the values and commitments of the target audience.

Thus, the educator should proceed somewhat differently in these two cases. In the case of climate change, the following question therefore becomes crucial: Is skepticism that humans can warm the planet a misconception of the first type, or of the second type? It is at this point that I depart from Simon Donner. He places dismissal of human influence on the climate squarely in the first category. In his earlier article (Donner 2007), Donner is unequivocal: human influence on the climate is “an extremely new concept” and a “major paradigm shift, arguably on the order of the Copernican Revolution” (Donner 2007, p. 233). In his later article (Donner 2011), Donner concedes that the perception is not universal, but nonetheless heavily emphasizes its widespread, long-lasting, and cross-culturally recurrent nature, calling it “the ancient view” and suggesting that it traces back to our hunter–gatherer ancestors (Donner 2011, p. 1298).

My reading of the historical and ethnographic record is different: human influence on the climate is an intuitive, ancient, and widespread idea; it is the notion of a separate meteorological realm, outside of human influence, that is the cultural oddity and historical novelty. Donner correctly notes that many cultures across the globe posit that spirits, gods, God, or other supernatural agents can control the weather. But it is also the case that in most of those cultures, human actions, especially moral and immoral behavior, are seen to influence these beings. The examples are countless and include contemporary and traditional Tibet (Byg and Salick 2009; Huber and Pedersen 1997), India (Gold 1998, p. 174), the Sahel (Togola 2000), Melanesia (Jacka 2009), and
Brazil (Taddei 2008). If bad weather is a punishment and good weather a reward, then the climate is under human influence, even if needing the intermediary of a deity. Moreover, many societies eliminate the middleman, so to speak, believing that human virtue and sin can directly drive weather patterns. This view has been documented among Kalahari San (Hitchcock 2009), the Chinese (Hsu 2000), the Inuit (Leduc 2007), indigenous Siberians (Crate 2008), and indigenous groups of Mexico (Smith 2007), to name only a few. In a wide variety of cultures it is believed that human magicians can direct the weather (see, e.g., Jacka 2009; Lipset 2011) or that powerful foreign human groups can do so (Byg and Salick 2009; Crate 2008; Hitchcock 2009, p. 258). Belief in divine regulation of the climate does not preclude human influence: in a variety of African societies, it is believed that both God and man influence the climate, or that God created the original, “correct” climate that has now been spoiled by human interference (BBC World Service Trust 2012, p. 11). People often find human–climate entanglement so plausible that no description of the mechanism of causation is offered or asked for. Tibetan villagers are convinced that recently experienced local climate change is human caused (either by their own religious infractions or by those of encroaching foreigners) even though, when asked, they give conflicting and uncertain answers about how that causation occurs (Byg and Salick 2009, p. 165). Denizens of a variety of African countries attribute recent climatic changes to human environmental mismanagement or to religious sin, despite often having little or no knowledge of atmospheric science (BBC World Service Trust 2012). Thus, blaming people for meteorological perturbations is so plausible that it is assumed even when it cannot be explained or proven—perhaps the very definition of intuitiveness.1

In the Marshall Islands, a Micronesian society where I conduct ethnographic fieldwork on climate change attitudes, the notion of anthropogenic climate change has been plausible since long before locals heard about greenhouse gases and melting ice caps on the radio. Bad weather can be blamed on human infractions. Magicians from a particular clan are credited with having once been able to both stop and start storms. In a strikingly prophetic statement, Marshall Islanders in the nineteenth century feared that human folly might cause the sea to rise and destroy the low-lying archipelago (von Chamisso 1986, p. 278). While Donner may have correctly identified some historical societies that regard the weather as beyond human tampering, the extreme ease of finding numerous counterexamples in the literature casts doubt on his universalistic claims; other scholars have come to the same conclusion about the cross-cultural pervasiveness of weather as a register for human action (Hulme 2009, 13–14; Rayner 2003, p. 278; Strauss and Orlove 2003, 3–4, 7).

Whether Donner is correct to emphasize the commonness, and hence the intuitiveness, of anthropogenic climate change skepticism, or I am correct to emphasize its uncommonness, we need to think carefully about whether a scientific misperception that we wish to challenge is based primarily on panhuman intuitions or on the cultural inculcation of a particular time and place. The ramifications for public education and citizen–scientist dialogue are far reaching.

If I am correct that the roots of climate change skepticism lie in cultural specifics rather than human universals, what might those cultural specifics be? One suspect is the modern Western notion of “nature” as a realm distinct from and opposed to human culture (see, e.g., Ingold 2008). Such a worldview comes in many flavors, not all of which may militate against acceptance of anthropogenic climate change. For instance, the idea that nature is a place of primal authenticity spoiled by contact with humans (Fairhead and Leach 2008) may be compatible with the idea of an Anthropocene (see, e.g., McKibben 2006), the geological era of massive human interference in the climate system. But the conception of human history distinct from natural history (Chakrabarty 2009), the assumption that there exists a class of issues called “environmental” that are different from economic, social, and cultural issues, may encourage many Westerners to dismiss the social–environmental, natural–cultural nexus that is climate change. Mark Twain’s famous quip that everyone talks about the weather but no one does anything about it encapsulates this Western view; the humor derives from the supposed absurdity of implying that humans might do any more than observe and

---

1 There is another possible interpretation of the cross-cultural evidence: the panhuman intuition is that people can influence local but not global weather. Donner mentions this alternate theory in his latter publication (2011, p. 1298). But the main thrust of his argument elides this local/global distinction. Furthermore, testing this theory against the ethnographic and historical record founders on ambiguity because most of societies in question are only concerned with, or aware of, their own geographical region: weather changing locally is for them weather changing everywhere (globally).
react to the climate. A skeptical Australian tract on manmade warming hinges its argument on the vast
gulf between a Lilliputian humanity, producing only
3% of carbon dioxide emissions, and an enormous,
unyielding, self-correcting “Nature”: “Human pro-
duction of CO₂ is irrelevant and of no consequence
to Nature or Nature’s balance” (Roberts 2013, p. 2).
If the nature–culture dichotomy is at the heart of
anthropogenic climate change skepticism, then cli-
mate educators might proceed by encouraging a more
holistic view of nature and society, stressing their
tangled interconnections. Countless ethnographic
studies from around the world [see, e.g., Århem (1996)
in the Amazon, Strathern (1980) in Melanesia, and
Cruikshank (2005) and Leduc (2007) in native North
America] describe societies that draw no clean line
(arginably, no line at all) between nature and society,
proving that such a worldview is both possible and
reasonable. Educators could also employ some of
the vast ethnographic and archaeological evidence
that apparent “wilderness” often bears the imprint
of human habitation, that “natural” landscapes may
have been created by people (Fairhead and Leach
2008), and that ecosystems cannot be adequately
described without taking into account their human
constituents (Rappaport 2008). Nature and society
lose their separation and the notion that climate has
become a human artifact no longer seems ludicrous.

Another possible culturally specific font of climate
change skepticism is the widespread tendency of
Western social conservatives—a commonly trotted-
out example these days is right-wing Americans—to
justify the social system in which they live and to
dismiss its critics (Feygina et al. 2010). Relatedly, in
some communities people believe in a “just world”
where suffering comes as a deserved punishment
for immorality (Feinberg and Willer 2011). Both
attitudes have recently been shown by experimental
psychologists to work against climate change belief
among Americans (Feinberg and Willer 2011; Feygina
et al. 2010): for system justifiers and believers in a just
world, climate change uncomfortably illustrates the
unsustainability of the status quo and the injustice of
a world where the smallest greenhouse gas emitters
are the most vulnerable to the greenhouse effect.
Many “system justifiers” also subscribe to the notion
of progress: a keystone of Enlightenment thought
and an enormously influential “metanarrative” up
to the present day. While we have no direct evidence
of a link between belief in progress and disbelief
in climate change, such a link seems plausible (see
Rudiak-Gould 2013), and it is has been convinc-
ingly argued for in the more general domain of the
denial of environmental problems (Norgaard 2002). If
this constellation of faiths—in the social system,
in worldly justice, and in progress—is at the root
of anthropogenic climate change skepticism, then
it is not climate change per se that these skeptics
find implausible, but rather the notion of an unjust,
unsustainable, declining world. British journalist
Melanie Phillips makes the link obvious in a 2004
piece titled “The global warming fraud”: climate
change is an “anti-American, anti-west ideology
which goes hand in hand with anti-globalisation
and the belief that everything done by the industri-
alised world is wicked” (Phillips 2013). Conservative
American commentator Jonathan S. Tobin similarly
writes that “many of the warming polemics have
been motivated not so much by ‘science’ as by an
ideological predisposition by some to view capital-
ism and the prosperity-producing economic activity
that it has generated as inherently sinful” (Tobin
2011). Climate change becomes merely a proxy war
in a larger confrontation between traditionalists and
radicals, Polyannas and Chicken Littles.

If this is the true meaning of climate change skep-
ticism, the science communicator could proceed in
two ways. The first would be to attempt to convince
skeptics that the social system is not actually good,
the world not truly just, progress not so assured. This
would be an unforgivingly difficult task, considering
how steadfastly people hold onto core ideological
commitments. Many educators might also find it
ethically dubious, not to mention supremely awkward,
to preach values and proselytize for an ideology
(though many science and technology scholars would
argue that they have been doing so all along, unknow-
ingly). Still, this approach is an option, and the science
communicator who is comfortable seeing himself not
only as a teacher but also as an activist may wish to
attempt it—occasionally, perhaps, with some success.
The second, less ambitious way to proceed would be
to mold climate change to the audience’s worldview
rather than vice versa; to recast climate change belief
as an ideological windfall and victory rather than a
concession to “them” and a betrayal of “us.” If the
ideological obstacle is system justification, extoll the
ability of a capitalist, scientific, industrial society to
pioneer a new generation of energy technology; say
that the point of solving climate change is not to de-
stroy the system but to save it [see Feygina et al. (2010)
on the possibility of “system-sanctioned change”]. If
the obstacle is “just world” belief, and the audience is
a Christian one, say that the universe is indeed just,
but only in the next life, not in this wicked and fallen
physical world. If the issue is progress, say that climate
change will merely slow the steady improvement of life rather than curtail it. Of course the educator may or may not agree with these statements—and cannot simply will herself to believe them—but they are the sort most likely to bear fruit. The fieldwork that I have conducted in the Marshall Islands shows that this approach—framing the threat as ideologically appealing even if physically dangerous—can succeed even in a situation where people are highly motivated to disbelieve (sea level rise promises to make the entire territory of the Republic of the Marshall Islands uninhabitable) and have locally plausible reasons for disbelief (God’s Biblical promise to Noah to never flood the Earth again). By framing climate change as the ultimate confirmation of a preexisting belief in sociocultural decline and the evils of abandoning tradition, the frightening idea is not just widely believed by Marshall Islanders but in a sense embraced (Rudiak-Gould 2012, 2013).

None of this would be easy, or free of ethical dilemmas, for the discussion has now gone far beyond “apolitical” science into the realm of unapologetically ideological debate. But the cultural approach has offered some useful suggestions that would have been overlooked by an approach based on human universals or on “just the facts, ma’am” communication of empirical evidence. The “deficit model” of science education would assume that people reject manmade climate change because they are not sufficiently scientifically literate. A more enlightened perspective, advocated by researchers in Science and Technology Studies (STS), assumes that scientific misconceptions arise not from publics’ ignorance but from their knowledge—the cultural frameworks, intuitive concepts, and moral visions with which people take up and interpret scientific issues (see, e.g., Wynne 1992). STS scholars sometimes go further and suggest that what science educators label as “misconceptions” are actually reasonable statements of values, norms, and symbolism. For instance, the cross-culturally widespread notion that human morality dictates the weather is unfounded or unintelligible from a scientific viewpoint, but is a reasonable interpretation of climate change and indeed conducive to belief and action. There is no need for a crippling relativism here, only an appreciation for the ethical and symbolic dimensions of climate change, and a healthy humility in which scientific experts have no privileged position when it comes to recommending policy or evaluating moral tradeoffs. Citizens may actually surpass scientists in assessing the moral dimensions of the problem because, unlike scientists, they have not been trained to shear scientific issues of their normative content (Jasanoff 2010).

This perspective does not eliminate the ethical quandaries faced by an educator tackling the ideological roots of climate change skepticism, but it offers guidance for navigating those quandaries in a more sensitive manner. In this way, the cross-cultural approach not only helps science communicators better educate the public, but indeed can take them beyond the role of “educator” into the role of participant in democratic dialogue—debating progress and decline, justice and injustice, human specialness—an engagement in which all sides may learn and benefit.

ACKNOWLEDGMENTS. Research that contributed to this article was supported by the Andrew W. Mellon Foundation; Dr. Alun Hughes; Oxford University; All Souls College, Oxford; St. Hugh’s College, Oxford; Jesus College, Oxford; and the Institute of Social and Cultural Anthropology, Oxford.

REFERENCES


NORTHEAST SNOWSTORMS
Volume I: Overview
Volume II: The Cases
Paul J. Kocin and Louis W. Uccellini

Northeast Snowstorms offers the most comprehensive treatment on winter storms ever compiled: more than 50 years of professional experience in the form of a two-volume compendium of insights, examples, photographs, over 200 color figures, and a DVD of added material.

American Meteorological Society; 818 pages, hardbound; AMS code MM54
Price: $100.00 list/$80.00 member/ $60.00 student member

ORDER ONLINE!
www.ametsoc.org/amsbookstore
or see the order form at the back of this issue.
THE TORNADO WARNING PROCESS
A Review of Current Research, Challenges, and Opportunities

by J. Brotzge and W. Donner

A review of the entire warning system, from prediction and detection to public response, reveals such fundamental needs as identifying acceptable risks, improving personal preparation, and personalizing warnings.

One of the scientific community’s greatest achievements in meteorology during the twentieth century has been the development of a largely effective public tornado warning system. Between 1912 and 1936, tornadoes killed an average 260 persons per year, about 1.8 deaths per million people when normalized by population (Brooks and Doswell 2001). Between 1975 and 2000, that number had declined to 54 deaths per year, or 0.12 deaths per million people in 2000 (Brooks and Doswell 2001), a reduction of 93% from 1925. In 1986 the tornado warning lead time was approximately five minutes, with only 25% of tornadoes warned; by 2004, the mean lead time was 13 min, with about 75% of tornadoes warned (Erickson and Brooks 2006).

Far from simple, the tornado warning process is a complex chain of events, encompassing institutional action and individual responses, that utilizes sensing technologies, conceptual models, numerical weather prediction (NWP), forecaster and emergency management (EM) decision making, warning dissemination technologies, and public experience and education (Fig. 1). The sequential steps of this process—forecast, detection, warning decision, dissemination, and public response—are known as the Integrated Warning System (IWS; Leik et al. 1981; Doswell et al. 1999).

This article reviews the end-to-end tornado warning process and related research, considers the challenges to improving the current system, and explores possible next steps. While this article cannot provide a completely comprehensive review of all research in each specific area, the goal is to provide a broad overview of the tornado warning process and a brief summary of the many avenues of research that could contribute to improvements in the current system.

TORNADO PREDICTION. The ability to predict a tornado’s precise path and intensity days in advance could allow for evacuation to take place well ahead of storm development and the predeployment of assets needed to support emergency response and recovery. While restrained to less accurate forecasts by the inherent limitations imposed by atmospheric predictability, the last decade has seen a growing recognition of the connection between large-scale patterns and large-scale tornado outbreaks.

As high-resolution, convection-allowing (≤4-km grid resolution) NWP becomes more accurate at longer time scales, multivariate model output may be used to a greater extent in identifying and predicting
tornado outbreak events. Using observational and modeling analysis, Egentowich et al. (2000a,b,c) identified a series of dynamic precursors during the 6–84 h preceding a major tornado outbreak. Shafer et al. (2009) found that Weather Research and Forecasting (WRF) model output could be used to discriminate between tornadic and nontornadic events up to three days in advance. Using WRF simulation output, Mercer et al. (2009) developed a statistical objective analysis technique to extract relevant predictive variables, yielding statistically significant accuracy scores >0.7 and skill scores >0.5 of these variables one day in advance of storm formation (Shafer et al. 2010).

Ever faster computer processing, and increasing memory and storage capacities combined with advances in parallel computing and code efficiency now enable the routine use of mesoscale forecast ensembles at high-resolution hours or even days in advance. Using WRF simulation output, Mercer et al. (2009) developed a statistical objective analysis technique to extract relevant predictive variables, yielding statistically significant accuracy scores >0.7 and skill scores >0.5 of these variables one day in advance of storm formation (Shafer et al. 2010).

FIG. 1. Summary of the institutional and individual responses that comprise the tornado warning process.

and observations, and may be issued up to several hours in advance of initial convective initiation. The skill level of the tornado watch has continued to improve over the years with increased observations, refined conceptual models, and more accurate and higher-resolution NWP (Pearson and Weiss 1979; Ostby 1999). Two additional, increasingly popular products issued by the SPC are the convective outlooks and mesoscale convective discussions (MCDs; Stough et al. 2012). Convective outlooks are issued up to eight days in advance, highlighting areas of the country with the potential for severe weather. MCDs are used to highlight general areas of concern, often issued just hours ahead of convective initiation or just prior to issuance of a watch. Both convective outlooks and MCDs are composed of a discussion briefing and visual map, and provide additional lead time and probabilistic information.

Currently, all official NWS tornado warnings are issued based upon “detections,” where an immediate tornado threat is observed either directly by spotters and media or inferred from observations (e.g., radar). However, as the accuracy and precision of short-term (0–3 h) storm predictions continue to improve, model output is expected to become an increasingly important basis upon which to issue NWS tornado
warnings. This is the eventual goal of “warn on forecast” (Stensrud et al. 2009), where NWS tornado warnings may be issued based not only on detected tornadoes or observed precursors, but also on model output. Utilizing model output as the basis for some warnings could theoretically extend lead time to tornadogenesis.

Significant advances in computer processing, the utilization of new types and greater numbers of real-time weather observations (NRC 2009), and the development and adoption of new data assimilation (DA) techniques (Kalnay 2003; Park and Xu 2009) are making warn on forecast a reality. Computer processing capabilities continue increasing at an exponential rate, as predicted by Moore’s law (Moore 1965). Faster processing permits higher-resolution NWP, which allows for the direct use of convective-resolving physics, bypassing less accurate parameterization schemes. The use of new observations, such as dual-polarimetric radar (e.g., Jung et al. 2008), wind and temperature profilers (e.g., Otkin et al. 2011), data from aircraft [e.g., Aircraft Communication, Addressing, and Reporting System (ACARS); Benjamin et al. 1991], lightning data (Fierro et al. 2012), and new evolving mobile platforms (e.g., Mahoney et al. 2010) facilitates a more accurate, three-dimensional analysis of the initial conditions. Model initialization also has been made easier with greater access to real-time observations through the use of such systems as the Collaborative Radar Acquisition Field Test (CRAFT; Kelleher et al. 2007), the Meteorological Assimilation Data Ingest System (MADIS; Miller et al. 2007), and Thematic Real-time Environmental Distributed Data Services (THREDDS; Unidata 2012); see “Prediction challenges” for more information.

**TORNADO DETECTION.** Weather radar is the primary tool used by warning forecasters to identify areas of potential tornado development. Radar reflectivity provides forecasters with a clear view of tornadic features, such as the hook echo (Markowski 2002), and Doppler radial velocity shows horizontal wind shear, sometimes an early indicator of tornado formation (Brown et al. 1971). Radar polarimetric data provide storm microphysical information, such as hydrometeor type and shape, that can be used to identify areas of significant low-level wind shear (referred to as \( Z_{DR} \) arcs) and tornado debris (Ryzhkov et al. 2005; Bodine et al. 2013).

To better standardize weather radar coverage across the United States, the national Weather Surveillance Radar-1988 Doppler (WSR-88D) network [known as Next Generation Weather Network (NEXRAD)]

**PREDICTION CHALLENGES**

Several significant challenges remain to be addressed before routine 0–3-h tornado prediction can be realized. These needs include i) faster computer processing to permit even higher-resolution NWP and more robust ensemble systems; ii) the ability to enable real-time DA of even larger volumes of data; iii) greater numbers of observations at high spatial and temporal resolutions; and iv) the ability to predict marginal, less predictable events with greater accuracy and fewer false alarms. Model grid spacing is tightly coupled with the model physics; for example, Bryan et al. (2003) determined that model grid spacing on the order of 100 m is needed to fully resolve subgrid-scale turbulence. Parameterization schemes, such as cloud microphysics, convective, and planetary boundary layer schemes, fail to capture subgrid-scale processes which can lead to large sensitivities in storm-scale NWP results (e.g., Dawson et al. 2010; Bryan and Morrison 2012).

In a similar manner, storm-scale NWP is equally sensitive to model initialization and analysis. Numerical modeling of convective storms has shown sensitivity to model initialization of low-level thermodynamics (Frame and Markowski 2010), low-level wind profiles (Dawson et al. 2012), surface soil moisture (Martin and Xue 2006), and orography (Markowski and Dotzek 2011). Model assimilation sensitivity may be reduced by increasing the number and use of observations in critical areas (Schenkman et al. 2011; Snook et al. 2012) and at critical times (Richter and Bosart 2002). However, the ability to collect, quality control, and properly assimilate all the necessary data in real time at high resolutions is a significant challenge (e.g., Brewster et al. 2008). To address this issue, an optimally designed national observing network is needed to collect the necessary observations at the high resolutions required (e.g., low-level moisture and wind profiles; Dabberdt et al. 2005; National Research Council 2009).

Finally, while our ability to anticipate and predict significant events is relatively good with a POD of nearly 90% for tornado outbreaks (Brotzge and Erickson 2009), the community faces significant challenges in predicting marginal and/or weakly forced tornado events. Brotzge and Erickson (2009) found the first tornado of the day, solitary tornado events, tornadoes from hurricanes, and weak (F0, F1) tornadoes had a much greater chance of not being warned. The FAR has been found to be highest for weakly forced and isolated events (Brotzge et al. 2011). Nonsupercell tornadoes, such as from tropical storms (Schultz and Cecil 2009; Moore and Dixon 2012) and squall lines (Trapp et al. 2005), pose a significant difficulty for prediction because of their often transient nature. Among the greatest remaining challenges for tornado prediction are the ability to predict exactly when a tornado will initiate (Markowski and Richardson 2009), to differentiate between tornadic and nontornadic supercells (Brooks et al. 1994; Stensrud et al. 1997; Mead 1997; Davies 2004; Schultz and Askelson 2012), and to identify threatening nonsupercell tornado storms (Wakimoto and Wilson 1989).
Radar (NEXRAD); Crum and Alberty 1993; Crum et al. 1993, 1998] was deployed (Whiton et al. 1998). The WSR-88D scanning geometry was designed to facilitate complete coverage between 610 m (2000 ft) and ~18 km (60,000 ft) AGL, with minimum height coverage at or below 610 m within a range of 102 km from radar (Leone et al. 1989); the final network provided contiguous coverage across the United States at 3.05 km (10,000 ft) and above (Crum and Alberty 1993). Specific radar site locations were chosen based primarily upon population distribution, severe weather climatology, topography, and proximity to other radars; most radars were sited to provide coverage over and slightly upwind of major metropolitan areas (Leone et al. 1989). As of 2012, 160 WSR-88D (S band) systems comprised the NEXRAD network across the United States and territories.

NEXRAD deployment had an immediate and significant positive impact on tornado warning statistics (Polger et al. 1994; NRC 1995). Bieringer and Ray (1996) found that the probability of detection (POD) increased by 10%–15% and that warning lead times increased by several minutes after installation of the WSR-88D network. Analyzing all tornadoes in the conterminous United States (CONUS) between 1986 and 1999, Simmons and Sutter (2005) estimated that the deployment of NEXRAD increased the percentage of tornadoes warned from 35.0% to 59.7%, increased the lead time from 5.3 to 9.5 min, reduced the false alarm ratio (FAR) from 78.6% to 76.0%, and reduced the number of expected fatalities and injuries by 45% and 40%, respectively. Smith (1999), however, noted that verification procedures changed as the NEXRAD system was deployed, possibly accounting for some of the observed increase in the POD.

For enhanced tornado detection, automated detection algorithms, such as the WSR-88D mesocyclone (Stumpf et al. 1998) and tornado detection algorithms (MDA and TDA, respectively; Mitchell et al. 1998), automatically identify radar-based tornado features and are displayed in real time within the Advanced Weather Interactive Processing System (AWIPS). Radar data can be combined with additional weather information to linearly project storm motion and extrapolate mesocyclone, tornado, and hail core movement (e.g., Smith and Elmore 2004; Lakshmanan et al. 2007; Wang et al. 2008; Ortega et al. 2009; Lakshmanan and Smith 2010; Miller et al. 2013).

Storm reports from individuals in the field can provide timely, critical information to warning officials. Trained “storm spotters” provide a valuable service to the NWS, EMs, and media by providing reliable, real-time information on storm evolution and tornado development (Moller 1978; McCarthy 2002). As well documented by Doswell et al. (1999), storm spotter networks were first organized during World War II largely to protect military installations. By the mid-1960s, spotter groups were organized more formally by the Weather Bureau for more general use under its SKYWARN program. With the advent of cell phone and embedded camera technology, widespread access to the Internet, television station helicopters, volunteer and professional storm chasers, and the rise of social media, warning forecasters now have greater access to real-time information than ever before.

Nearly as important, spotters provide much-needed postevent verification; Brotzge and Erickson (2010) found a systematic increase in the numbers of weak tornadoes verified over densely populated counties when compared with rural counties. However, erroneous reports from the field can impede the warning process; Smith (1999) describes how poor tornado verification overinflates tornado POD and overestimates the FAR. Brotzge et al. (2011) found very high FAR in high-population-density counties and very low FAR in sparsely populated counties, perhaps indicative of lower warning rates across rural areas because of the prevalence (or lack) of field reports available and a subsequent decrease in forecaster confidence for warning in those areas; see “Detection challenges” for more information.

**TORNADO WARNING DECISION.** Once the formation of a tornado is considered likely or is reported already in progress, the NWS issues a tornado warning, the official NWS product used to warn the public of a tornado. The first tornado warning was issued on 25 March 1948 by U.S. Air Force officers E. Fawbush and R. Miller at Tinker Air Force Base in Oklahoma City, Oklahoma, and was remarkably successful (Miller and Crisp 1999b; Maddox and Crisp 1999). In fact, this first warning was so successful that it provided the scientific underpinning for establishment of the Air Weather Service Severe Weather Warning Center (SWWC), the first national severe weather warning program. During its first year of operation in 1951, the SWWC issued 156 (multicounty) tornado warnings, of which 102 (65%) were verified (Miller and Crisp 1999a). Since that time, tornado warnings, now issued by the local NWS Weather Forecast Offices (WFOs), have continued to improve as measured by the total percentage of tornadoes warned.

The final decision by the operational forecaster on whether to issue a warning is based upon a number of
DETECTION CHALLENGES

The most common reasons for operational warning forecasters for not detecting (and thereby not warning) tornadoes prior to touchdown often can be traced to having either too little information available—because of inadequacies in existing technology (e.g., LaDue et al. 2010), limited spotter networks, and incomplete conceptual models—or too much information, that is, data overload.

As the primary tool used for detecting tornadoes, weather radar is critical for seeing low-level to midlevel rotation prior to tornadogenesis. In areas with limited low-level radar coverage, tornado detection (and prediction) is severely hampered. In a root cause analysis study of 146 unwarned tornadoes between 2004 and 2009, “radar sampling,” “no radar signature,” and “radar use” were listed as 3 of the top 10 reasons for failure to warn and were cited in over two-thirds of all missed events (Quoetone et al. 2009). Sampling issues were cited in 19 of 31 false alarm events evaluated. Brotzge and Erickson (2010) found a mean 20% increase in the number of tornadoes not warned with increasing distance from radar, once sorted by population density.

Solutions to improving radar coverage include the use of lower-elevation scans, deployment of gap filling and rapid-scan radars, and an optimization of the radar network configuration. The WSR-88Ds’ lowest scanning angle is 5.5° elevation, as limited by Federal Communications Commission (FCC) regulations. At some mountain sites across the western United States, the WSR-88D radars are located on mountain tops, limiting the views of critical valley areas. One solution now being implemented at a few locations is the use of zero and/or negative elevation angles (R. Brown et al. 2002, 2007; Wood et al. 2003). A second, long-term solution to improve radar coverage is to simply add more radars to the network. However, because of the high cost associated with deploying and operating large antenna, S-band (WSR-88D type) radars, a more cost-effective solution may be to deploy limited numbers of “gap filling” (X or C band) radars to fill in coverage gaps between WSR-88Ds (McLaughlin et al. 2009). Brotzge et al. (2010) and Mahale et al. (2012) demonstrated the value of gap-filling radars for improving detection of tornado radar signatures. A third option for improving radar coverage is to sample more frequently. Replacement of the WSR-88Ds with rapid-scan, phased-array radar (PAR) technology (e.g., Zrnic et al. 2007) could provide 1-min volume scans (or faster single elevation scans), an improvement over the current 4–6-min volume scans provided by the WSR-88Ds. In ongoing evaluations of the impact of PAR data on tornado warnings, Heinselman et al. (2012, 2013) found that the use of faster scans has the potential to extend tornado warning lead times, reduce false alarms, and increase forecaster confidence. Finally, a more rigorous, optimal radar network configuration could improve overall low-level coverage. NEXRAD radars were originally deployed to operate as single autonomous systems; however, merged, multiradar data have proven more effective for extracting severe weather information (Lakshmanan et al. 2006). Geometric, statistical, and genetic algorithm techniques have been developed to optimize the low-level coverage and maximize multi-Doppler overlap (Ray and Sangren 1983; de Elia and Zawadzki 2001; Minciardi et al. 2003; Junyent and Chandrasekar 2009; Kurdzio and Palmer 2012). Nevertheless, the addition and/or replacement of radars will require a significant financial public investment.

Storm spotters provide an equally critical role to the warning forecaster. In the root cause analysis study, a lack of, conflicting or erroneous spotter reports were cited as having contributed to warning failure in nearly two-thirds of all missed events, and a lack of reports contributed to 15 of 31 false alarms (Quoetone et al. 2009). Sustained education and coordination of spotter groups requires dedicated NWS resources. Fortunately, as described previously, access to real-time information and video from the field is becoming easier, with the proliferation of new video and wireless technologies (e.g., Dixon et al. 2012).

A basic understanding of tornado dynamics is still key to good forecasting and detection. In the Quoetone et al. (2009) root cause analysis study, “not anticipated,” “conceptual model failure,” and “environment” were listed among the top six reasons for warning misses. Poor radar, environmental conceptual models, and environment were listed as three of the top four reasons cited for issuing tornado false alarms. “Fits radar conceptual model” was cited in 30 of the 31 false alarm events studied. Continued improvement in the conceptual models requires sustained advances in basic research. Field programs such as the Verification of the Origins of Rotation in Tornados Experiment (VORTEX; Rasmussen et al. 1994) and the second VORTEX project (VORTEX2; Wurman et al. 2012) provide valuable observational data from which to study and improve understanding. Continued meteorological training and education are essential for moving research to operations.

Finally, with the plethora of new sensors and model output now available to the warning forecaster, many are now experiencing data overload, which is hampering warning operations. “Workload” was cited in one-third of all missed warnings, with “distractions” cited in one-quarter of all missed warnings (Quoetone et al. 2009). One solution to this is the use of integrated, “fused” and/or assimilated sensor products (e.g., Wang et al. 2008). A second, complementary solution is the advent of multisensor and three-dimensional visualization (e.g., Gibson Ridge Software, LLC).
tornado warned is essential for public safety; the public is much more likely to take shelter once they have received an official warning (Balluz et al. 2000). However, there is an incentive to keep the warning area size small; the use of smaller warning polygons is estimated to save over $1.9 billion annually in reduced interruption and unnecessary sheltering (Sutter and Erickson 2010). County-based tornado warnings were replaced with storm-based warning polygons in 2007.

As of 2011, the national tornado POD was 0.75, with a mean lead time of 14.6 min, and a FAR of 0.74 (NOAA 2011b). A review of the long-term trends in these statistics reveals that the POD and mean lead time have increased dramatically since the installation of the WSR-88D network and NWS modernization program (Friday 1994), with a POD of 0.48 and a mean lead time of 7.6 min in 1994. However, nearly all of this increase in lead time was a direct result of greater numbers of tornadoes being warned (Erickson and Brooks 2006); all tornadoes not warned were assigned a lead time of zero, and then included in the calculation of the mean lead time. Using data between 1986 and 2004, Erickson and Brooks recalculated tornado lead time without the missed tornadoes included and found a rather steady lead time of around 18.5 minutes. While greater numbers of tornadoes are being warned in advance (possibly because of improved radar technology, conceptual models, and training), lead time on warned tornadoes has not increased, and the FAR has remained steady at around 0.75 as well; see “Warning decision challenges” for more information.

**WARNING DISSEMINATION.** Warning the public remains difficult in large part because the “public” is a largely diverse population with tremendous variation in education, physical abilities, family support, and situational awareness. To overcome these challenges, a variety of communication alert systems are used. Warnings may reach the public directly from the NWS through the National Oceanic and Atmospheric Administration (NOAA) Weather Radio (NWR) and the Internet, or indirectly through media, emergency management, and private sector weather providers. Widely adopted following the April 1974 tornado outbreak (Coleman et al. 2011), NWR allows for an in-home method for waking a person from sleep in case of an emergency through its alert tone. Today, over 1,000 NWR transmitters offer 98% national coverage (Zubrick 2010). The NWS also provides direct information to the general public via the Internet with some WFOs now experimenting with social media to distribute warning information.

The public most commonly receives tornado warnings from local media through television and radio (e.g., Hammer and Schmidlin 2002). Media utilize a host of methods to catch each viewer’s attention and to convey the necessary information, including the use of “cut-ins,” “crawlers,” mobile phone apps, Facebook, and Twitter (Coleman et al. 2011). Storm video and radar imagery provide greater spatial and temporal information regarding storm size, severity, storm motion, and geographic impact. Video media also more easily convey nonverbal cues from the television (TV) meteorologist. Indeed, research demonstrates that local populations often develop profound psychological commitments to specific weather stations or forecasters (Sherman-Morris 2006). Television broadcasts are often simulcast over the radio but without the benefit of the images.
Emergency managers also play a critical role in disseminating weather information to the local community. As part of their responsibilities, EMs operate local warning systems, such as local outdoor warning sirens or reverse 911 systems, and coordinate disaster response and recovery efforts. An instant messaging service called NWSChat was created to facilitate direct communication between the NWS and EMs and to better support EM services. However, there are few consistent criteria applied across jurisdictions for warning dissemination. A number of meteorological (e.g., presence of a wall cloud) and nonweather-related (e.g., public backlash for issuing false alarms) factors influence the judgment of EMs on whether to activate warning systems (Sorensen and Mili 1987; Stewart and Lusk 1994; Donner 2008); see “Warning dissemination challenges” for more information.

PUBLIC RESPONSE. Warning dissemination sets into motion a process of public response, a complex and multidimensional activity. While research on risk and warning response has been conducted since the 1950s, it was not until the 1990s that scholars began to systematize findings into a general model. Mili and Sorensen (1990) and Lindell and Perry (1992) shared the common conclusion that warning response was not a single act, but a set of stages through which the public progressed in responding to disseminated warnings. Before taking action, the public must receive, understand, believe, confirm, and personalize warnings.

Reception. Community members receive warning information through formal and informal channels. Formal communication includes NWS, media, emergency management, and reverse 911, or any official

WARNING DISSEMINATION CHALLENGES

A significant challenge in improving warning dissemination is to integrate new technologies in such a manner that those less able to afford such tools can still be warned. The Commercial Mobile Alert System (CMAS), Wireless Emergency Alerts (WEA), and Interactive NWS (iNWS) were recently created to disseminate warnings to mobile devices. However, many are ill equipped to receive text messaging, and so older warning systems, such as outdoor warning sirens, must still play a critical role within an integrated warning system, even as new, more informative services are made available.

The limitations of dissemination tools must be clearly recognized when building a public warning dissemination system. For example, mobile phone applications fail if and when cell phone towers and communications are disabled, a frequent problem in storm-ravaged areas. Similarly, outdoor warning sirens fail when power is lost to those sirens, such as occurred in some areas during the Alabama tornadoes of 27 April 2011. The use of outdoor sirens also varies significantly among jurisdictions, with some districts using them to warn on all severe thunderstorm (and sometimes nonweather)-related warnings, while other municipalities limit the use of sirens to tornado warnings only. Furthermore, many areas simply do not have sirens available, nor would it be cost effective to install sirens in many regions of the country. However, the consequences of not having sirens can be deadly; two people died in the 2011 Alabama tornadoes when early morning storms knocked out power to their trailer, and because they lived out of range of the nearest sirens, had no warning before they were hit (Ammons 2011). Some jurisdictions have replaced all outdoor sirens with calling systems such as reverse 911. However, these systems have been known to take tens of minutes to call all those in the tornado path, with no guarantee that those called would be alerted prior to tornado impact. A battery-operated NWR provides an immediate and direct warning method, but NWR ownership is low with limited surveys showing ownership of ~10%–33% (Manning 2007; Kupec 2008). NWR often is cited as the least-used method for obtaining warnings; only 3% of 1,650 persons surveyed just after the 3 May 1999 tornado in Moore, Oklahoma, indicated they had received their warning from NWR (S. Brown et al. 2002). While each system has certain limitations, an integrated and redundant dissemination system is more robust. In a survey following the 3 May 1999 Moore, Oklahoma, tornado, 55% of residents interviewed received the warning from more than one source (Hammer and Schmidlin 2002).

Another challenge for the operational forecaster is how to effectively communicate scientific information to the general public. Instantaneous communication and the growth of meteorological support companies have had a significant impact on the warning process (Golden and Adams 2000). As a result, institutions now communicate risk with unprecedented speed. Nevertheless, problems related to the expertise of institutions may affect the process of risk communication. For example, a recent experiment simulating a tornado outbreak tasked EMs with accessing and interpreting radar data (Baumgart et al. 2008). Despite general competence, study participants experienced significant difficulties interpreting wind velocity data and, more importantly, synthesizing multiple forms of radar data to produce overall judgments, which affected the risk communication process.

Effective communication also entails that the public understands and makes effective use of warnings (Lazo 2012). The risk communication process is most effective when those at risk hold a “perceived shared experience” with those already victimized (Aldoorya et al. 2010). When those warned could relate to victims (e.g., similar gender or race), threat acknowledgment and information seeking increased. Thus, risk communication may be taken more seriously if nearby communities are affected. How warnings are communicated also may shape risk perception. Numerical representations of risk often fail to persuade (Lipkus and Hollands 1999). In an experiment on risk perception of flooding, images depicting flood damage reinforced perceived risks (Keller et al. 2006). NOAA is now conducting an impact-based warning experiment (Maximuk and Hudson 2012) to evaluate ways in which to improve NWS communication to motivate improved public response.
warning system. Informal communication includes family, friends, and coworkers. Each form of communication channels warning information to the public, but each does so in dramatically different ways. Formal communication tends to reach members of upper- and middle-class populations, while informal communication often better serves the poor, ethnic minorities, and recent migrants. For instance, warnings issued during the 1987 Saragosa, Texas, tornadoes failed to reach local Hispanic populations (Aguirre 1988; Ahlborn and Franc 2012). Latinos prefer friends and family as sources of warning information (Peguero 2006) and receive tornado warning information from informal networks (Donner 2007). Poorer populations also were less likely to receive formal warnings (Schmidlin and King 1997).

Social networks may play a key role in reception. For instance, Nagarajan et al. (2012) documented the importance of warning dissemination among neighbors in a series of computer simulations. Frequent interaction of family members (Lardry and Rogers 1982), strong community or network involvement (Turner et al. 1979; Sorensen and Gersmehl 1980; Perry and Greene 1983; Rogers and Nehevaqsa 1987; Rogers and Sorensen 1991), regular association with a subculture or voluntary association (Perry et al. 1981), and more frequent community interaction (Scanlon and Frizzell 1979) improved the likelihood of message reception among individuals within the community.

Understanding. How recipients understand and make sense of warning information is deeply connected to human psychology and past experience. With the exception of Quarantelli (1980), research overwhelmingly demonstrates that long-term residents generally tend to hold a better understanding of warning information (Haas et al. 1977; Foster 1980; Perry and Greene 1983; Perry and Lindell 1986; Blanchard-Boehm 1998). Psychologically, the public is more likely to understand warning information if conveyed along with local information and maps (Berry 1999). Multiple warning sources increase chances of comprehension (Mileti and Darlington 1995), while at the same time excessive information within a single message may lead to higher rates of misunderstanding (DiGiovanni et al. 2002). Probability information attached to tornado warnings (e.g., the tornado has a 30% chance of occurring), for instance, may confuse rather than clarify risks for the public (Morss et al. 2010).

One concern is whether individuals understand the difference between warnings and watches. In a study of Austin, Texas, residents, Schultz et al. (2010) found that 90% of the sample could adequately distinguish between watches and warnings. Other studies found similar rates of understanding (Balluz et al. 2000; Biddle and Legates 1999), while others encountered more modest results (Mitchem 2003). Still other research suggests much lower rates of comprehension. In a broad survey of 769 residents across Texas, Oklahoma, and California, only 58% of all participants correctly understood the difference between a watch and a warning, though the percentage improved among residents in Oklahoma and Texas and among older and more educated survey participants (Powell and O’Hair 2008).

Social scientists have identified a number of social and cultural factors that account for variation in warning comprehension between individuals. Education is consistently associated with greater understanding (Turner et al. 1979), and those with a greater familiarity with science and scientific concepts generally hold a stronger understanding of warnings (Blanchard-Boehm 1998). Age, too, shows a direct correlation with understanding (Turner et al. 1979; Blanchard-Boehm 1998).

Belief. After understanding a warning message, the recipient evaluates the credibility of the message. Will there really be a tornado or is the warning a false alarm? In other words, should the message be taken seriously? Rarely, however, at this stage do recipients arrive at a concrete conclusion about whether a tornado will or will not occur. On the contrary, recipients crudely evaluate the probability of severe weather. The psychological qualities, past experiences, and unique demographic characteristics of the individual play a significant role in shaping these judgments of likelihood.

Those closer to a hazard are more likely to believe a warning (Diggory 1956; Sorensen 1982), which may be because of the greater likelihood of experiencing environmental cues (Drabek 1969; Quarantelli 1980; Sorensen 1982; Tierney 1987; Mileti and Fitzpatrick 1993; Hammer and Schmidlin 2002). Additional psychological processes also may play a significant role in the process of believing warnings. There are mixed findings regarding whether certain sources are more or less believable. Some research shows the public places greater faith in “official sources” (e.g., NWS warnings; Li 1991; Drabek 1994), whereas other studies routinely demonstrate “unofficial sources” (e.g., family) to hold greater credibility among the communities (Sorensen 1982; Perry 1983; Li 1991). It may be that the particular source may play a lesser role in credibility when compared to source familiarity. Warning sources to which individuals are personally
or emotionally attached (e.g., a favorite weather forecaster) or with which they are more familiar may appear more credible (Mileti and Fitzpatrick 1993).

Demographic factors have some influence as well. Women appear more likely to believe warnings (Drabek 1969; Farley et al. 1993; Sherman-Morris 2010). Why this is the case may be explained through socialization, as well as the fact that women tend to be caregivers (Perry 1983). Additionally, the higher one’s socioeconomic class, the more likely one is to believe a warning (Sorensen 1982; Perry 1987). Finally, a society’s culture may also play a role in warning response. Finding the Japanese more likely than U.S. residents to respond to volcano warnings, Perry and Hirose (1991, p. 112) explain that Japanese live within a “collectivist culture in which citizens have higher expectations that authorities will provide care in the event of disasters or other disruptions in social life.” Perry and Hirose suggest that the Japanese population has greater trust in government, and thus greater response rates, than Western societies, and that response to warnings among the Japanese might reflect the broader cultural rules of obedience and authority common in Asian societies.

**Confirmation.** A common feature of the warning process (Mileti 1999), confirmation serves to clarify and specify warning information, but at the cost of delaying sheltering. Confirmation has been found to take place among neighbors, rather than through formal channels (Kirschenbaum 1992), with information from media sources more likely the subject of confirmation (Frazier 1979). Confirmation may also be something as simple as visual confirmation of the storm. Whether beneficial or detrimental, confirmation remains a certain feature of the warning process.

**Personalization of risk.** Risk personalization deals with whether community members believe severe weather will affect them personally. In other words, one can believe that a threat exists somewhere, but the threat is not immediate and therefore action is unnecessary. For example, residents may decide that the mountains or rivers surrounding their community protect them from tornadoes, even if they believe local reports that storms may produce tornadoes (Donner et al. 2012).

The psychological elements of risk personalization are well understood. Warning consistency yields greater personalization of risk (McDavid and Marai 1968; Lindell and Perry 1983). Warning specificity (Perry et al. 1981) and sender credibility (Perry 1979; Rogers and Nehnevajas 1987) contribute to personalization. Geographical proximity to a threat appears to be the most important in the literature (Flynn 1979; Perry and Lindell 1986; Rogers and Nehnevajas 1987). With some notable dissent (Mileti and Darlington 1995), most research agrees that past hazards experience leads to a greater likelihood of personalization (Perry 1979; Hansson et al. 1982; Saarinen et al. 1984; Rogers and Nehnevaja 1987).

Demographics also play a role. As with belief, women are more likely to personalize a threat (Flynn 1979; Hodge et al. 1981). Socioeconomic status also may play a role in risk personalization (Flynn 1979; Mileti et al. 1981).

**Action necessary and feasible.** Believing that one is personally at risk sets off a process of determining whether one must and is able to do something to protect oneself. Little research has been conducted in this area of the model. This stage is unique from resource availability, in that resources may be available but the potential victim either does not know about them or does not think them useful for protection.

Protection from severe weather often takes the form of sheltering. Sheltering may be broadly defined as either “in home” or “public.” With in-home sheltering, refuge is typically sought in hallways, closets, underground basements, or, ideally, personal shelters. Those under warning may also choose to seek public shelters, which are typically set up and maintained by local government. Public shelters may be stand-alone shelters, in that their only use is as a shelter, or schools, town halls, or other municipal structures may become “shelters” during storms. Education, possibly through increased income, is most consistently associated with the availability of resources such as shelters (Edwards 1993; Balluz et al. 2000); see “Public response challenges” for more information.

**NEXT STEPS.** All other things being equal, as the U.S. population density increases, tornado fatalities may be expected to increase, calling for a review of the prediction, detection, and communication processes through which tornadoes are warned. Urban populations continue to rise in hazard prone regions, thereby placing greater numbers of people at risk (Brooks and Doswell 2001; Ashley 2007). In addition, the overall population is aging, with increasing numbers living alone (Gusmano and Rodwin 2006). Greater diversity among the population introduces additional challenges, such as warning dissemination to non-English-speaking populations (Donner and Rodriguez 2008). As described herein, a number
Although the determinants of shelter seeking are well documented in the literature, little is known about the sheltering process itself. Personal shelters are ideal, in that sheltering is immediate; traveling to a public shelter may be dangerous, especially in the context of tornadoes that are rapid and violent on onset. For those in mobile homes or similar vulnerable structures without shelters, evacuation may be the only option; mobile homes comprised 7.6% of U.S. housing stock in 2000, but 43.2% of all tornado fatalities between 1985 and 2007 occurred in mobile homes (Sutter and Simmons 2010). In addition to distance, other more “human” factors may shape the use of shelters. Cola (1996) found that people were less likely to use shelters thought uncomfortable. Pet owners also may be less likely to seek shelter (Heath 1999; Pfister 2002). More research is needed to understand shelter use and its relationship to lead time and social factors. Additional work needs to explore the associated needs, optimal locations, and operation of public tornado shelters.

There is also the real inability by some to take shelter because of disability. In the Joplin, Missouri, tornado, three mentally handicapped men died when their home was hit. Also in Joplin, 12 residents and a nursing assistant died at the Greenbriar nursing home, and another 8 patients died when St. John’s Regional Medical Center was hit. Both facilities had been warned and had begun taking storm precautions, but neither had enough lead time to evacuate. In Shooal Creek, Alabama, seven people were killed when an assisted living facility was hit. Additional research is needed to explore the lead time requirements for those who must evacuate (e.g., from trailer homes) or need help sheltering (e.g., those with special needs). Indeed, the public at large requires a continuum of lead times, where for some a warning lead time of 30 min could be essential, whereas for others, a large lead time could lead to apathy and greater danger.

Public Response Challenges

A fundamental question society must ask is, “How much risk are we willing to tolerate?” The answer to this dilemma will set the limit on how much money should be expended toward further research and warning infrastructure. In other words, the public must define its acceptable risks, and its willingness to provide additional resources or reduce existing services or quality to match those risks (Stallings 1990). The public’s level of acceptable risk likely varies across the country as a function of the nature and extent of the risk. This variability calls for an emphasis on local-to-regional decision making, such that any top-down, one-size-fits-all strategy will likely be less than optimal. A dense observing spotter

and warning dissemination network in the plains may vary in function and form from one in the Southeast, whereas neither system may be cost effective in the West or New England.

A second essential subject often overlooked in discussions of the tornado warning process is preparation, both at the organizational and personal levels. Preparation at the organizational level may include the development of public policy regarding the use and availability of public shelters and warning systems, the availability of multilingual warnings, requirements or guidelines for shelters in mobile home parks, building codes, and sheltering procedures. Private preparations may include developing a family disaster plan, copying and storing critical insurance papers and photos in safety deposit boxes, or purchasing a safe room or shelter. Proper preparation at the organizational level can often facilitate the speed and ease of personal decision making during a moment of crisis.

Preparation should focus on maximizing personal safety, minimizing economic loss, and easing recovery efforts. While this article has focused on public safety, total damage estimates from tornadoes between 1950 and 2011 range from $300 billion to $450 billion (U.S. dollars; Simmons et al. 2013). A greater focus on personal mitigation could reduce tornado damage. Sutter et al. (2009) found that low-cost home mitigation could reduce tornado damage by as much as 30%.

Finally, the one common ingredient to a successful end-to-end tornado warning program is the personalization of the warning; to be successful, warnings must evoke a sense of specific and immediate risk. Even days prior to an event, the efforts of the SPC and others are spent narrowing the area of a potential threat; local WFOs narrow the threat further in time and space, issuing warnings over specific regions in time. The most effective warnings are those that communicate clearly to individuals the specific information they need to know with enough time to react. The goals of ensemble NWP, warn on forecast, phased array and gap-fill radars, and storm-based warnings are to provide more detailed data on when and where tornadoes will strike. Many new and innovative warning dissemination tools, many developed and sold by the private sector, convey this detailed information to individuals, through the use of local media, outdoor warning sirens, NOAA Weather Radio, the Internet, smart phones, and pagers. Similarly, preparation for tornadoes needs to be personalized, and specific mitigation information provided at a household

of challenges limit the effectiveness of the current warning system (Table 1). Based upon the preceding literature review and these associated challenges, the warning process can be fundamentally improved with a greater emphasis and understanding of acceptable risk, preparation, and personalization.

A fundamental question society must ask is, “How much risk are we willing to tolerate?” The answer to this dilemma will set the limit on how much money should be expended toward further research and warning infrastructure. In other words, the public must define its acceptable risks, and its willingness to provide additional resources or reduce existing services or quality to match those risks (Stallings 1990). The public’s level of acceptable risk likely varies across the country as a function of the nature and extent of the risk. This variability calls for an emphasis on local-to-regional decision making, such that any top-down, one-size-fits-all strategy will likely be less than optimal. A dense observing spotter
level could see potential dividends in reducing home damage and personal injury.

Social and cultural factors may inhibit personalization of warnings. Long lead times and high false alarm rates tend to depersonalize risk. A continuing program of research and education remains key to systematically improving public response to warnings.

A highly integrated and efficient tornado warning system does not necessarily ensure that no fatalities will ever occur, but it does set a priori standards of warning capability as a function of the community-defined level of acceptable risk, resources, and will. The effectiveness of the best tornado warning system is dependent largely upon the comprehensiveness and manner of preparedness at the organizational and personal levels. This review has demonstrated the value of research and investment at all stages of the warning process for improving the personalization of the warning. In an era of austerity, additional

<table>
<thead>
<tr>
<th>Integrated Warning System</th>
<th>Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prediction</strong></td>
<td>Need higher spatial and temporal observation sampling</td>
</tr>
<tr>
<td></td>
<td>Ability to process and assimilate large volumes of data</td>
</tr>
<tr>
<td></td>
<td>Faster computer processing</td>
</tr>
<tr>
<td></td>
<td>Improve prediction of inherently less-predictable systems</td>
</tr>
<tr>
<td></td>
<td>Improve differentiation between tornadic and nontornadic cells</td>
</tr>
<tr>
<td></td>
<td>Greater accuracy at longer time scales</td>
</tr>
<tr>
<td></td>
<td>Ability to apply ensemble prediction at high resolutions</td>
</tr>
<tr>
<td><strong>Detection</strong></td>
<td>Radar temporal sampling</td>
</tr>
<tr>
<td></td>
<td>Radar spatial gaps, primarily at low levels</td>
</tr>
<tr>
<td></td>
<td>Erroneous, sporadic, or unreliable spotter reports</td>
</tr>
<tr>
<td></td>
<td>Poor or incomplete conceptual models</td>
</tr>
<tr>
<td><strong>Warning decision</strong></td>
<td>Balancing POD with FAR</td>
</tr>
<tr>
<td></td>
<td>Data overload</td>
</tr>
<tr>
<td><strong>Warning dissemination</strong></td>
<td>Cost of dissemination systems</td>
</tr>
<tr>
<td></td>
<td>Maintenance of old systems, adoption of new sensors</td>
</tr>
<tr>
<td></td>
<td>Reception of warning during night and in rural areas</td>
</tr>
<tr>
<td></td>
<td>Consistent use of warning systems and false alarms</td>
</tr>
<tr>
<td></td>
<td>Effective communication of warnings</td>
</tr>
<tr>
<td></td>
<td>Multilingual warnings</td>
</tr>
<tr>
<td></td>
<td>Access of poor to private sector warning methods; e.g., personal digital assistants (PDAs)</td>
</tr>
<tr>
<td><strong>Public response</strong></td>
<td>Inability to shelter because of handicap or age</td>
</tr>
<tr>
<td></td>
<td>Mobile homes</td>
</tr>
<tr>
<td></td>
<td>Cost of sheltering</td>
</tr>
<tr>
<td></td>
<td>Cost of purchasing in-home shelters</td>
</tr>
<tr>
<td></td>
<td>Safety of in-home sheltering vs evacuation</td>
</tr>
<tr>
<td></td>
<td>Impact of warning lead times, false alarms (“cry wolf effect”)</td>
</tr>
<tr>
<td></td>
<td>Response of public facilities (e.g., large venues, schools)</td>
</tr>
<tr>
<td></td>
<td>Demographic and cultural factors</td>
</tr>
<tr>
<td></td>
<td>Mitigation and preparation</td>
</tr>
<tr>
<td></td>
<td>Personalization of risk</td>
</tr>
</tbody>
</table>
investments will need to be strategically focused to further prepare and personalize the tornado threat.

**ACKNOWLEDGMENTS.** This work is supported by the Engineering Research Centers Program of the National Science Foundation under NSF Award 0313747. The authors thank three anonymous reviewers for their insightful comments and suggestions for the improvement of this manuscript. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the National Science Foundation.

**REFERENCES**


—, —, and W. Diaz, 2012: Tornado warnings in three southern states: A qualitative analysis of


Ostby, F., 1999: Improved accuracy in severe storm forecasting by the Severe Local Storms Unit during the last 25 years: Then versus now. *Wea. Forecasting, 14*, 526–543.


Whiton, R. C., P. L. Smith, S. G. Bigler, K. E. Wilk, and A. C. Harbuck, 1998: History of operational use of weather radar by U.S. weather services. Part II:


AMS titles now available as eBooks at springer.com

AMS BOOKS
RESEARCH • APPLICATIONS • HISTORY

www.ametsoc.org/amsbookstore

Scan to see AMS eBook titles at springer.com
The year 1863 was a pivotal year in American history as major battles between Union and Confederate forces marked a gradual and inexorable shift of fortunes in favor of the Union. The American Civil War was in its third year and news of its battles constantly filled the columns of newspapers and magazines. More than 3,000 soldiers died in early May 1863 at the Battle of Chancellorsville. Hundreds more died in the early battles at Vicksburg in late May, adding to the accumulating Civil War death total already in the tens of thousands and unprecedented in American history at that time.

Amid the manmade carnage on land, a natural disaster brought a less extensive but locally severe tragedy of lost lives to the Florida coast. On 28 May 1863 an unprecedented hurricane made landfall in northwest Florida—the only landfalling hurricane now recorded in American history in the month of May. This unusual hurricane arrived in even more unusual times. The absence of any major effect on combat readiness and the relatively localized effects of the worst damage seems to have limited the duration of the memory of this storm even locally. This paper documents its rediscovery and the effects it had on the greater drama prevailing in the country.

**DATA SOURCES.** We relied on the extensive series of official government documents that were printed in the late 1800s and early 1900s for written accounts of the storm. Union ship officers wrote to their superiors and to the board of inquiry investigating the storm concerning the loss of the *USS Amanda*, and there are other accounts from ships in the immediate area of the storm. These accounts provided details that are not available in the extant logbooks. The pertinent accounts were published in the 1903 volume (U.S. Navy Department 1903). We also used the original logbooks of the U.S. Navy located at the National Archives in Washington, D.C. In addition, we made use of land-based weather records held in Records Group 27 at the National Archives in College Park, Maryland, which consists of U.S. Army fort records and cooperative observers of the Smithsonian...
Institution (Fleming 1990). Newspaper accounts are from the U.S. Library of Congress Newspaper Library in Washington, D.C., and the newspaper microfilm collections at the University of Maryland in College Park. Online sources for newspaper records from Louisiana, Georgia, and Florida included the Jefferson Parish Library (www.jefferson.lib.la.us/index.htm), Georgia historical newspapers (http://enquirer.galileo.usg.edu), and the Florida Digital Newspaper Library (http://ufdcweb1.uflib.ufl.edu).

**A MAJOR STORM ON A MINOR FRONT.** In late May 1863, the Gulf coast of northwest Florida was blockaded by the Union Navy’s East Gulf Blockading Squadron, which operated from Pensacola eastward through the Florida Straits in an attempt to slow the flow of goods and war material in and out of Confederate ports (Weddle 2002). At least one Union ship was posted for blockading duties in or near Hurricane Inlet (Panama City), St. Joseph’s Bay, the West and East Passes of St. George Sound, St. Mark’s, Cedar Key, Tampa Bay, and Port Charlotte. Navy facilities in Pensacola and Key West were also in Union possession. Figure 1 depicts these locations and the ships on station as well as those at sea on the morning of 28 May 1863.

St. George Sound (Apalachicola Bay) is the waterway connecting the city of Apalachicola, Florida, to the Gulf of Mexico. The Apalachicola River empties into the sound and steamboats in the nineteenth century traveled as far north as Columbus, Georgia. This city became an important industrial center for the Confederacy during the Civil War and defense against possible Union incursions up the river led to the construction of warships for the Confederate Navy. The CSS Chattahoochee was constructed in late 1861 and 1862 and was finally delivered to the Confederate Navy from the shipyard in November 1862. By late May 1863, she was located at Chattahoochee, Florida, on the border with Georgia (Turner 1988; Watts et al. 1990).

On 24 May, Union troops captured the schooner Fashion, loading cotton on the river north of Apalachicola, and towed it down to the Gulf of Mexico. News of its capture led to the decision to take the Chattahoochee down the river although it was in no position to directly challenge Union forces unless it could pass obstructions placed into the river to prevent Union ships from moving upriver (New York Times, 24 June 1863). On 26 May, the gunboat sailed down to Blountstown Bar and waited overnight for a possible rise in water (Watts et al. 1990), the water being lower than normal at that season of the year following a drier than normal spring. Presumably, if the water level had been high enough the ship might have attempted to navigate around the obstructions to engage the Union fleet, but this can only be speculation at this point.

The following day the Chattahoochee was about to head back from Blountstown to the town of Chattahoochee because of low water in the river when the boilers exploded, killing 14 people outright; several more would die from their wounds soon after. [The gunboat sank and remained in the river below Blountstown until mid-August when the wreck was recovered and returned to Chattahoochee.

---

**Fig. 1.** Map depicting the location and name of U.S. Navy ships blockading the Confederate ports along the Gulf of Mexico from Louisiana to Florida on 28 May 1863 for which logbooks are available and used in this study. The only nonextant logbook shown here is that of the Port Royal. The weather data from the logbooks from these ships, along with newspaper accounts and other sources, were used to reconstruct the track of the hurricane. Cities mentioned in the text and Fort Jefferson, where a weather record was taken and used in this study, are also included.
It was not until about 12 hours after the midday boiler explosion on 27 May that another steamer arrived at the scene to transport survivors back to Chattahoochee. By that night heavy rain and strong winds had overspread the area, which would eventually lead to a rise in all the region’s rivers. The Commander of the Confederate Navy Yard at Columbus wrote in a 15 June 1863 letter “The dead and wounded were taken on shore, where they remained until the next afternoon (May 28), most of the time a terrible storm raging” (Campbell 2008). For the wounded, many lying in the open with little protection from the elements of a gathering hurricane, it was a miserable and painful experience with little medical care that could be provided to those scalded by the water from the boilers.

**WRECK OF THE BARK USS AMANDA.** Four U.S. Navy ships were on station on either side of St. George Sound. The steam gunboat USS Port Royal and gunboat USS Somerset were stationed at the West Pass of the sound while the bark USS Amanda and the steamer USS Hendrick Hudson were stationed at the East Pass of St. George Sound. Ships at both locations experienced hurricane force winds for several hours and at West Pass it dismasted and drove ashore the barge Andrew Manderson onto Sand Island while the sloop Brockenburgh was driven ashore on St. Vincent’s Island. Both the Somerset and Port Royal were able to ride out the hurricane, but not without difficulty (U.S. Navy Department 1903).

On the East Pass, off Dog Island, both the Amanda and Hendrick Hudson experienced even worse hurricane conditions. Figure 2 shows the Hendrick Hudson at anchor, taken sometime between 1862 and 1864. The Amanda initially grounded on the west point of Dog Island (where the sea made a complete breach and carried off all of the coal stored there). However, a shift of wind to the southeast drifted the bark into the sound and carried it on a northwest drift toward the mainland where she grounded in mud flats about 7 a.m. A sense of the confusion and danger that took place as the ship neared shore can be found in the original logbook of the Amanda, where the wind and weather data for the hours of 5 a.m. through 8 a.m. are mistakenly recorded as 8 a.m. through 11 a.m. This could be either an error in the original chalk slate log where the readings would normally be first recorded or in a transcription from the slate log to the logbook. A comparison of the nearly collocated Hendrick Hudson’s logbook confirms the transcription error.

Acting Volunteer Lieutenant George E. Welch quickly, and controversially, ordered the abandonment and firing of the ship, having claimed to see Confederate forces nearby and thereby considering the position of the Amanda indefensible (U.S. Navy Department 1903). The focusing of the crew’s attention on a feared Confederate attack and the firing of the ship can be seen in the subsequent log entries of the weather aboard the Amanda, which are less frequent and complete until the log ends later on the afternoon of 29 May. However, according to a note recorded at the bottom of the entry for 29 May the log entries for the last day were recorded from the log slate onto a sheet of paper, but this was lost in the confusion and recorded from memory after the fact. The poor agreement of the instrumental weather data with that of the collocated Hendrick Hudson logbook casts doubt to the accuracy of the other accounts of the ship’s firing made from memory.

A court of enquiry held in June 1863 in Key West could find no corroborating evidence from any witnesses on either the Amanda or the Hendrick Hudson that any Confederate troops were seen at all. The court found that Lt. Welch’s decision to abandon the ship was probably taken too easily and that Acting Volunteer Lt. Cate, of the Hendrick Hudson, as the senior of the two officers, also seemed to express
The edge of the eyewall appears to have brushed the USS Somerset located off the west end of St. George Island in the West Pass of St. George Sound. A northeast hurricane began to moderate at 3:30 a.m. on 28 May and the wind went around gradually to the west; at daylight during the calmer weather the barge Andrew Manderson was spotted dismasted and ashore on Sand Island (at the tip of the west end of St. George Island) and the sloop Brockenburgh ashore on St. Vincent’s Island. At 6 a.m. the wind again increased to a hurricane, the wind blowing from the west-southwest and the Somerset hove up all anchors and stood to leeward of Sand Island (U.S. Navy Department 1903).

Neither the Amanda nor the Hendrick Hudson experienced any diminished wind force during the hurricane. At 2 a.m. the pressure was 29.20 inches (uncorrected; 29.25” corrected for elevation, gravity, temperature, and latitude; equivalent to ~991 hPa), and the wind from the southeast was at force “11+” with heavy rain. The written comments in the log also indicate that at 3:20 a.m. the wind hauled suddenly to southeast (from east-northeast at 1 a.m.) and the vessel then began to float off the west point of Dog Island to the northwest. The lowest pressure and strongest winds recorded in the Amanda’s logbook were at 5 a.m. with an uncorrected reading of 28.74” (28.79” or 975 hPa corrected) and winds of force “12+” from the SSE. At 7 a.m. the pressure was up slightly to 28.80” (28.85” or 977 hPa corrected) with a south wind and at about this time the Amanda, which had drifted across the sound from the west end of Dog Island, was aground in mud flats on the mainland some three or more miles away from the Hendrick Hudson.

The Hendrick Hudson reported a lowest pressure of 28.74” (29.00” or 982 hPa corrected) at 4 a.m. while at Dog Island, riding out the hurricane. The wind force at 3 a.m. was 12 from the east-southeast and at 8 a.m. was still blowing a force 12 hurricane with the wind from the southwest. At 6 a.m. the pressure had risen to 29.04” (29.30” or 992 hPa corrected). By this

**METEOROLOGICAL ANALYSIS.** The hurricane moved due northward toward the Apalachicola area during the daylight hours of 27 May. Landfall on the barrier islands late that night and the mainland about sunrise 28 May (Fig. 3). The western

![Fig. 3. Track of Hurricane Amanda from 24 to 31 May 1863.](image-url)
temperatures owing in part to less direct or indirect solar radiation and possibly owing to the wetting of thermometers. The remnant low later encountered a cold front in the Mississippi Valley and moved north and then northeast before being absorbed by an extratropical low. A proposed track of Hurricane "Amanda" is given in Fig. 3.

DESTRUCTION AND DEATH ALONG THE FLORIDA COAST. The storm came ashore on Thursday morning, 28 May, and reports of varying detail indicated a severe storm along the coast. In a letter dated 30 May from one of the editors of the Tallahassee Floridian to the editor of the Macon Daily Telegraph, it was stated “We have had a heavy blow here the past week—the heaviest I ever witnessed in Florida at this season of the year. From the coast there are various rumors of loss of life and property. I have just heard that from the Ochlockonee to Peurifoy’s Landing, twenty-one bodies of persons drowned were recovered, and eleven from Goose Creek, making thirty-two” (Macon Daily Telegraph, 5 June 1863). At Tallahassee, it had rained more or less every day for the previous week and so severe a gale at the time of year was not within the memory of the oldest citizen.

Table 1 presents the noon temperature data recorded on 11 U.S. Navy ships from Louisiana to northern Florida on 26–30 May, centered on the date of the landfall of Hurricane Amanda on 28 May. The original value of “63” on 28 May recorded by the USS Colorado is corrected to “73” based on the surrounding data and the likelihood of a misreading of the thermometer or a transcription error in the log. The ship also is much cooler than other ships near Mobile Bay during the entire month. The 28–30 May values for the Amanda are estimated based on its average noon temperature difference from that recorded on the collocated Hendrick Hudson for the 20 days of overlap that month. The Beauregard was at Cedar Key on 26 and 27 May and off the coast of St. Marks by 30 May. All data are taken directly from the original logbooks on file at the U.S. National Archives, Washington, D.C.

<table>
<thead>
<tr>
<th>Ship</th>
<th>Location</th>
<th>26 May</th>
<th>27 May</th>
<th>28 May</th>
<th>29 May</th>
<th>30 May</th>
<th>Average temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amanda</td>
<td>St. George Sound</td>
<td>80</td>
<td>78</td>
<td>75</td>
<td>78</td>
<td>84</td>
<td>79</td>
</tr>
<tr>
<td>Hendrick Hudson</td>
<td>St. George Sound</td>
<td>76</td>
<td>73</td>
<td>71</td>
<td>74</td>
<td>80</td>
<td>74.8</td>
</tr>
<tr>
<td>Fort Henry</td>
<td>Cedar Key</td>
<td>87</td>
<td>80</td>
<td>78</td>
<td>77</td>
<td>79</td>
<td>80.2</td>
</tr>
<tr>
<td>Beauregard</td>
<td>Cedar Key to St. Marks</td>
<td>81</td>
<td>80</td>
<td>79</td>
<td>82</td>
<td>82</td>
<td>80.8</td>
</tr>
<tr>
<td>Colorado</td>
<td>Entrance Mobile Bay</td>
<td>80</td>
<td>74</td>
<td>73</td>
<td>78</td>
<td>74</td>
<td>75.8</td>
</tr>
<tr>
<td>Pocahontas</td>
<td>Entrance Mobile Bay</td>
<td>75</td>
<td>77</td>
<td>74</td>
<td>75</td>
<td>78</td>
<td>75.8</td>
</tr>
<tr>
<td>Potomac</td>
<td>Entrance Mobile Bay</td>
<td>71</td>
<td>79</td>
<td>73</td>
<td>77</td>
<td>73</td>
<td>74.6</td>
</tr>
<tr>
<td>Lackawanna</td>
<td>Entrance Mobile Bay</td>
<td>72</td>
<td>81</td>
<td>78</td>
<td>85</td>
<td>76</td>
<td>79.4</td>
</tr>
<tr>
<td>Vincennes</td>
<td>Ship Island, MS</td>
<td>77</td>
<td>77</td>
<td>74</td>
<td>79</td>
<td>76</td>
<td>76.6</td>
</tr>
<tr>
<td>Portsmouth</td>
<td>New Orleans</td>
<td>78</td>
<td>82</td>
<td>80</td>
<td>90</td>
<td>82</td>
<td>82.4</td>
</tr>
<tr>
<td>Winona</td>
<td>Plaquemines, LA</td>
<td>85</td>
<td>85</td>
<td>85</td>
<td>78</td>
<td>90</td>
<td>84.6</td>
</tr>
<tr>
<td>Average all 11 ships</td>
<td></td>
<td>78.4</td>
<td>78.7</td>
<td>76.4</td>
<td>79.4</td>
<td>79.5</td>
<td>78.5</td>
</tr>
<tr>
<td>Average without warmest and coldest ships</td>
<td></td>
<td>78.2</td>
<td>78.7</td>
<td>76.0</td>
<td>78.8</td>
<td>79.0</td>
<td>78.2</td>
</tr>
</tbody>
</table>

The average temperature of 78.2°F (25.6°C) is consistent with the storm being tropical in nature. We also note that despite steady southerly winds on the day of the hurricane, ships in the area saw lower temperatures owing in part to less direct or indirect solar radiation and possibly owing to the wetting of thermometers. The remnant low later encountered a cold front in the Mississippi Valley and moved north and then northeast before being absorbed by an extratropical low. A proposed track of Hurricane “Amanda” is given in Fig. 3.
Wednesday night (27–28 May) the rain was literally pouring down and accompanied by a severe wind from the northeast, which shifted in the course of the night to the southwest with increased violence. (We note here that all of the available ship’s log data and reports from elsewhere and the space–time continuity do not indicate that the southwest wind is representative of the prevailing winds and that it is either a simple error or associated with local convection.) The paper mentioned that at Dickerson Bay 19 persons were drowned and another 7 to 10 at Goose Creek (Florida Sentinel, 2 June 1863). These latter numbers are probably referring to the same deaths as communicated by the editor of the Floridian. His account of deaths “from the Ochlockonee to Peurifoy’s Landing” would be inclusive of Dickerson Bay, which lies between the two.

Farther east, the salt works near St. Marks and Bayport were entirely destroyed. Large quantities of the salt were lost and 40 persons were drowned. The gale was said to be so strong as to have pushed the waters of the Gulf inland for several miles back into the country, inundating parts of the St. Marks railroad (Macon Daily Telegraph 5 June 1863, citing the Columbus Times of 3 June 1863). Railroad passengers arriving in Savannah reported houses blown down, fences destroyed, and, in some locations, total destruction of crops. One island in the vicinity of St. Marks was said to be submerged and a large number of persons drowned. Some of the names of the drowned were reported back to Savannah (Macon Daily Telegraph 5 June 1863, citing the Savannah News of 4 June 1863). Another letter published in the Macon Daily Telegraph of 5 June 1863, dated 3 June 1863 from Thomasville, Georgia, stated “the gale of Thursday is said to have done much mischief among the salt boilers on the Florida coast. One report says 150 lives were lost and many animals, much stock and salt. Hope it is not so bad; some, though, have certainly perished.”

The Florida Sentinel of 9 June 1863 reported that the destruction of property on the coast was greater than anticipated and stated that 32 lives were lost (probably a reference to the deaths in Goose Creek and Dickerson Bay region). The final death totals from the storm in Florida would appear to be at least 72 lives—40 in the St. Marks area and the other 32 farther west at Goose Creek and Dickerson Bay. The possibility of further storm-associated deaths on land cannot be ruled out but the available information cannot confirm the rumored report of 150 lives being lost. In addition, there were 38 deaths at sea on the Soler on 26 May while located about 45 miles southeast of Mobile Bay on a voyage from Havana (New Orleans Bee, 3 June 1863), which brings a confirmed total of 110 dead as a direct result of the storm. We note that the Soler’s approximate location relative to our estimated center at the time (Fig. 3) suggests a larger area of higher winds might hint at a cold core origin. The ship that picked up the survivors of the Soler at 28°30′N, 86°39′W on the evening of 29 May had encountered a very heavy south-southeast to southwest gale with a heavy sea setting from the west-northwest on 27–29 May while sailing from South Carolina to the area. Since the ship would have come around the Dry Tortugas, but remained in the eastern Gulf of Mexico (probably to stay nearer Union shipping to avoid Confederate privateers) this puts constraints onto the location of the hurricane and contributes to the estimated storm track. So the estimated track allows for the absence of gale reports from ships in the main shipping lane, which is consistent with the subsequent ship reports in the eastern half of the Gulf of Mexico and reports from the press of ships at sea and on land.

**SUMMARY OF WEATHER ON NAVAL PROTAGONISTS.** There is not often glory in combat and the protagonists in both navies were sometimes denied a chance to engage in battle by the prevailing weather conditions. The abandonment of the Amanda in the hurricane of 1863 was a black mark on the career of the commanders of the Amanda and Hendrick Hudson and the subsequent official investigation provides some clues to the mindsets of the Union blockaders during a period of high stress. The sheer boredom of blockading duty, punctuated by shorter intervals of potential engagement with blockade runners, was a normal feature of the war. The hurricane’s unexpected arrival exposed both positive and negative features of the blockading squadron’s readiness.

The Confederate force’s ability to engage in warfare was also tested by the weather. The Chattahoochee’s fate in 1863 appears to have been an unfortunate accident. The men involved performed the rescue and salvaging of the ship as well as could be expected given the circumstances. The gunboat had been plagued by production faults for virtually its entire life and the men were likely victim to an accumulation of human errors that contributed to the faulty pressure gauge reading causing the steamer’s destruction. While rain is likely to have been falling at the time of the disaster the worst of the weather was still some hours away and the contribution of the weather at the time of the boiler explosion is unknown.
The weather further frustrated Confederate efforts to engage the Union forces; on one occasion in 1864 it came close to causing more fatalities and led to the capture of some troops, the loss of all of their boats and provisions, and a forced retreat inland. Once again, the weather further ensured that no significant military battles would be recorded along this stretch of the Gulf coast during the Civil War. However, the absence of notable battles does not detract from the high importance of the naval blockade in Union efforts to contain the Confederate ability to obtain supplies and goods or the Confederacy’s recognition of the need to evade, and frequent success in running, the blockade.

**METEOROLOGICAL SUMMARY AND CONCLUSIONS.** Meteorologically, Hurricane “Amanda” is the only documented U.S. landfalling hurricane in the month of May and is nearly two weeks earlier than the next earliest U.S. landfalling hurricane—that of Hurricane Alma on 9 June 1966 (Fig. 4). Hurricane Amanda is the earliest arriving U.S. landfalling hurricane on record in any hurricane season. This is a reminder to Gulf Coast residents to the possibility, however small, of a hurricane outside of the main hurricane season. Other “near miss” hurricanes along the U.S. coastline in the month of May have occurred on 29 May 1908 (tropical storm force winds felt at Hatteras) and on 21 May 1951 a hurricane was centered off the North Carolina coast (and made an earlier landfall in the northernmost islands of the Bahamas). Tropical Storm Beryl had 60-kt winds at landfall on 28 May 2012 when making landfall at Jacksonville Beach, Florida, and is worthy of mention since it approached hurricane strength and a central pressure of 995 hPa (29.38”). Therefore, the risk of a May hurricane is real and not limited to the Gulf Coast. U.S. Naval forces and the coastal populations of the United States were also fortunate that during the Civil War there were no major hurricanes that made landfall anywhere in the United States.

The hurricane caused the deaths of at least 72 people on land (and 110 on both land and sea when the deaths on the Soler are added). This ranks it at number 27 in the deadliest U.S. hurricanes behind Hurricane Betsy in 1965 (75 deaths) and number 28 if Hurricane Sandy of 2012 is treated as a hurricane at landfall. This total does not include the 17 deaths on the CSS Chattahoochee since the role of weather is uncertain in the events leading up to the boiler explosion. If the high end of 150 deaths can subsequently be confirmed it would tie at number 21 with the Indianola hurricane of 1886. The death toll may have been higher had this storm arrived in peacetime as portions of the coastal population had moved inland because of the Union blockade. The blockade ended normal peacetime commerce and most adult men were serving in the military.

Our results also highlight the well-known gaps in the North Atlantic official tropical cyclone records in the nineteenth century (Landsea et al. 2004). However, our results indicate that incomplete U.S. landfall records are not simply due to the occurrence of tropical cyclones in lightly populated coastal areas. The eye of Hurricane Amanda passed over the town of Apalachicola and through the Union blockading force in the area. Instead, previous hurricane compilers simply missed a storm that was reported in the press, including the *New York Times* of 17 June 1863. This indicates that population-based estimates of missing tropical cyclones (Landsea et al. 2004) are only a starting point for estimating the likelihood of false negatives in storm detection in any tropical cyclone basin. Other factors such as a functioning press and the presence of war or civil strife will also increase the likelihood of undercounting tropical cyclones. Previous researchers have also been restricted by their access to data sources, the tendency to accept previous work as being sufficiently complete and accurate, assumptions about the seasonality of tropical cyclones, and the extensive work required to rediscover historical hurricanes and tropical storms.

There is also caution needed in interpreting tropical cyclone metrics such as season length (Kossin 2008). The latest U.S. landfalling hurricane on 1 December 1925 was changed by the Atlantic
basin Hurricane Database (HURDAT) committee to a tropical storm, which shifts the latest date for a U.S. landfalling hurricane to 21 November 1985 (Hurricane Kate). Our results move forward by 13 days the earliest U.S. landfalling hurricane in the spring. Basinwide statistics are also susceptible to a few outlier storms and otherwise incomplete information.

Our results also reemphasize the always provisional nature of even the most complete datasets such as HURDAT. Paleo-hurricane reconstructions from northwest Florida can benefit from reconsidering this new storm in the interpretation and calibration of existing modern sediment records (e.g., Lane et al. 2011) while other users of HURDAT for Atlantic paleo-hurricane reconstructions regions outside of the United States are working with incomplete records on frequency and intensity that extend into the twentieth century.

ACKNOWLEDGMENTS. We thank Chris Landsea and Joan David for their generous assistance in the production of Figs. 2 and 3. The staff of the Wakkula County Historical Society was very helpful in determining the location of Peurifoy’s Landing, Florida.

REFERENCES