A REVIEW OF TARGETED OBSERVATIONS

by Sharanya J. Majumdar

Targeted observations to improve numerical forecasts of high-impact weather events over the past two decades, particularly during the THORPEX era (2005–14), are evaluated.

The concept of targeted observations lies in the question of “Where and when should one deploy and assimilate observations, in order to improve a numerical forecast of a weather event that is important to society?” The premise is that forecast errors of a particular event can grow, sometimes rapidly, when initial conditions are not properly constrained by observational data. To curtail rapidly growing forecast errors, regions in which analysis errors amplify into these large forecast errors must be identified. Observations can then be targeted and assimilated to reduce analysis errors in these important regions. In addition, observations in regions of initially large analysis errors that may evolve (perhaps slowly) into large forecast errors can also be targeted.

Targeted observations make up a subset of the broader field of “adaptive sampling,” which is widely used in areas including electrical and computer engineering applications, and in detecting behavior in localized animal clusters. Within the context of weather, adaptive observations refer to any supplemental observations that can be deployed at will, based on a certain phenomenon to be studied or predicted. The first aircraft reconnaissance missions into hurricanes and typhoons in the 1940s can be considered to be adaptive observations. One of the earliest examples of adaptive observations with an emphasis on weather forecasting was the launching of extra radiosondes in the United States and neighboring regions ahead of a tropical cyclone (TC) in the 1950s (B. Sheets 2015, personal communication). While these radiosondes were deployed adaptively, there was no technique to target the observations in particular areas. Later, during the National Oceanic and Atmospheric Administration’s (NOAA) Synoptic Flow Experiment between 1982 and 1996, aircraft-based dropwindsonde observations were targeted around the TC in regions based on synoptic reasoning (Burpee et al. 1996).

The first large-scale attempt to investigate the utility of targeted observations in the extratropics was the multinational Fronts and Atlantic Storm Track Experiment (FASTEX) field campaign in 1997 (Snyder 1996; Joly et al. 1999), where one goal was to test the hypothesis that observations in selected “target
shortcomings of targeted observations, and to assess whether the original grand vision has been realized. This article aims to offer a broad perspective that is more accessible to a wider readership than Majumdar et al. (2011). In the next section (“Techniques for targeted observing guidance”), the techniques used to select target regions are reviewed, followed by a summary of the prominent field campaigns (“Field campaigns” section). A synopsis of evaluations of the influence of assimilating targeted observations is then provided in the “Results from evaluations” section. Finally, conclusions and recommendations for the future are provided. Although the results summarized in this paper are synthesized from many studies, references are restricted to a subset of selected papers to preserve the readability. A full bibliography of around 200 papers is provided in the online supplement to this paper (http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-14-00259.2).

An explanation of how the targeted observations are selected is presented in the first sidebar.

TECHNIQUES FOR TARGETED OBSERVING GUIDANCE. To predict the optimal locations and times for targeting prior to deployment, several mathematical techniques have been developed and widely published since the mid-1990s. Given the process of an operational assimilation–forecast cycle, a strategy for targeted observations would ideally account for i) the forecast uncertainty of a high-impact event at time $t_a$, ii) all observations to be assimilated up to and including the targeted analysis time $t_v$, iii) the characteristics of the data assimilation scheme, iv) the type and accuracy of the targeted observations, and v) the projected influence of assimilating the targeted observations at time $t_v$ on a future forecast of a specified metric at time $t_d$. Because of the prohibitively high computational effort that is required to account for all of these facets, it has been necessary to make crude assumptions such as the linear evolution of errors.

The main techniques employed prior to and during the THORPEX era are summarized in Table 1. Additionally, a large body of the literature on the techniques is listed in the accompanying bibliography (papers labeled with an M). The earliest techniques developed for use in FASTEX were based on dry dynamics and analysis sensitivity, in other words, how a modification to the analysis at time $t_a$ is estimated to affect forecast errors at time $t_d$. Over the subsequent few years, the observations and data assimilation scheme began to be incorporated into existing or new methodologies. Recent advances include the addition

1 THORPEX, which was in operation between 2005 and 2014 under the auspices of the WMO, was “A World Weather Research Programme accelerating improvements in the accuracy of one day to two week high-impact weather forecasts for the benefit of society, the economy and the environment” (Shapiro and Thorpe 2004).

2 The WMO/THORPEX report on targeted observations (Majumdar et al. 2011) is available at www.wmo.int/pages/prog/arep/wrrp/new/documents/THORPEX_No_15.pdf.
HOW ARE TARGETED OBSERVATIONS SELECTED?

The procedure to select targeted observations is complex and imperfect. A basic example of a common forecast case scenario is illustrated in Fig. SB1. The case is selected by a forecaster or lead investigator, who at time $t_i$ identifies a potential high-impact weather event of importance to society at a future verification time $t_v$ (top-left map). A verification region is selected (red box). The objective is to target observations at a future analysis time $t_a$ to improve a forecast between $t_d$ and $t_v$ within this verification region. Using targeting guidance products based on models initialized at time $t_i$ (shading in bottom map), a decision on whether and where to deploy is then issued at the decision time $t_d$. Because of the lead time involved with mission planning, this decision normally needs to be made more than a day prior to $t_d$. Observations such as those from dropwindsondes are then targeted in the sensitive area at time $t_d$.

After the conclusion of the experiment, the influence of assimilating the targeted observations at time $t_d$ is evaluated within the verification region at time $t_v$. In observing system experiments (OSEs), this is accomplished by comparing a forecast with routine and targeted data (green contours; top-right map) versus a forecast that is the same except that the targeted data are withheld (black contours; top-right map). Finally, it should be noted that many elements can be varied beyond the example illustrated in Fig. SB1, such as the weather system that is targeted, the types of guidance products used, the forecast window ($t_v - t_i$), the types of observations to be targeted, and the evaluation procedure.

Diagram of targeted observing procedure:

- $t_i$: Case selection at $t_i$
- $t_d$: Targeting guidance using models initialized at $t_i$
- $t_a$: Decision at $t_a$
- $t_v$: Targeted data assimilated at $t_v$
- $t_f$: Forecasts with and without targeted data, initialized at $t_i$ and valid at $t_v$

Fig. SB1. Schematic of the targeted observing procedure.

Of moisture, with applications to mesoscale weather. All of the techniques to date employ either adjoint methods or ensemble forecasts. Some are based on the direct sensitivity of a response function or forecast metric (e.g., total energy within the verification region) to observations or changes to the analysis, whereas others make quantitative predictions of the effect of assimilating targeted observations. Hybrids of certain techniques have been proposed, such as the inclusion of ensemble-based error covariance information within a singular vector (SV) framework. While these techniques possess theoretical similarities and differences, the consensus view is that they all possess limitations, especially when nonlinearity is significant.

The meteorological characteristics of the guidance products used during the field campaigns have been investigated and compared. In the extratropics, the widely used adjoint, SV, and ensemble transform Kalman filter (ETKF) products commonly identified baroclinic zones and jet features as targets, though smaller-scale aspects differed. As the forecast window ($t_v - t_i$) increased, the targets could be traced farther upstream in the storm track, with multiple targets often appearing. For TCs, the guidance products often disagreed. For example, SV guidance identified the sensitivity in an annulus around the TC and upstream locations in the midlatitude trough for recurving TCs, whereas the ETKF identified features such as adjacent ridges and troughs, and regions downstream in the midlatitude storm track (Wu et al. 2009). The results were also dependent on the model used. The target regions were often of synoptic scale, which is usually beyond the range of coverage of a single aircraft mission. The consensus among the community is that the target regions generally tended to be in meteorologically sensible regions, though spurious targets could sometimes exist. Additionally, a limited number of studies have demonstrated that the assimilation of observations targeted within target regions has had a larger positive impact on forecasts than equivalent observations outside the target regions or selected at random.

After two decades of development and use, it is prudent to ask what lessons have been learned from the targeting techniques. First, the techniques have all proven to be practicable for real-time implementation during field campaigns. In many cases, this had been accomplished by using an adjoint model of lower resolution, or an ensemble of model output with limited fields. To expedite the mission planning, several field campaigns used automated software to draw or select flight tracks that traversed the maxima of the
Table 1. Summary of targeting methods used during field campaigns. The color code refers to subjective methods (pink), analysis sensitivity methods (green), and observation sensitivity methods (orange). All of these techniques provide guidance based on forecasts initialized at time $t_a$, prior to the targeted analysis time $t_x$.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Summary</th>
<th>Key references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synoptic reasoning</td>
<td>Target areas are selected subjectively, based on human understanding and experience. Avoids theoretical deficiencies in objective techniques. Targets may be missed (e.g., regions of initially small errors that grow rapidly).</td>
<td>Burpee et al. (1996)</td>
</tr>
<tr>
<td>Ensemble variance</td>
<td>Predicts regions of high analysis error variance at future time $t_x$; in other words, observations are targeted in areas of potentially large analysis error. Very easy to compute given any ensemble forecast. Typical variables used include horizontal winds and temperature. Does not consider forecasts at time $t_a$.</td>
<td>Lorenz and Emanuel (1998); Aberson (2003)</td>
</tr>
<tr>
<td>Adjoint sensitivity</td>
<td>Uses an adjoint model to predict the response of a forecast aspect at $t_a$ to changes in any variable at $t_x$. Traditionally uses “dry” variables, though moisture has recently been included for mesoscale targeting.</td>
<td>Langland et al. (1999a); Bergot et al. (1999); Doyle et al. (2014)</td>
</tr>
<tr>
<td>Quasi-inverse linear method</td>
<td>Uses a quasi-inverse linear operator to identify a region of origin at time $t_a$ of a forecast difference at time $t_x$.</td>
<td>Pu et al. (1997); Pu and Kalnay (1999)</td>
</tr>
<tr>
<td>Singular vectors (SVs)</td>
<td>Identifies perturbation structures that grow optimally from time $t_a$ into a selected verification region at time $t_x$. Several analysis and verification metrics are possible, the most common being total energy at both times. Advances include analysis error covariance SVs, which incorporate analysis error estimates.</td>
<td>Palmer et al. (1998); Buizza and Montani (1999); Gelaro et al. (2002)</td>
</tr>
<tr>
<td>Ensemble transform (ET) technique</td>
<td>Uses ensemble forecasts to estimate the reduction in forecast error variance at time $t_x$, based on changes to the analysis error variance at time $t_a$. Theoretically equivalent to a special case of the ETKF.</td>
<td>Bishop and Toth (1999)</td>
</tr>
<tr>
<td>Ensemble transform Kalman filter (ETKF)</td>
<td>Uses an ensemble-based data assimilation methodology to quantitatively predict the influence of assimilating a given set of observations at $t_a$ on forecast error variance at $t_x$. Ideally, error covariance information is consistent with that of the actual data assimilation scheme.</td>
<td>Bishop et al. (2001); Majumdar et al. (2002)</td>
</tr>
<tr>
<td>Forecast sensitivity to observations (FSOs)</td>
<td>Uses the adjoint of a data assimilation system to estimate the sensitivity of a forecast aspect to observations. A similar methodology is also used after the observations are collected to evaluate the impact of assimilating observations on forecast errors.</td>
<td>Baker and Daley (2000); Langland and Baker (2004)</td>
</tr>
<tr>
<td>Adjoint-derived sensitivity steering vector (ADSSV)</td>
<td>Uses an adjoint model to predict the response of the tropical cyclone steering flow at $t_x$ to variables at $t_a$.</td>
<td>Wu et al. (2007)</td>
</tr>
<tr>
<td>Ensemble sensitivity</td>
<td>Uses an ensemble-based data assimilation scheme to quantitatively estimate the sensitivity of a forecast response function at $t_a$ to analysis changes or observations at $t_x$.</td>
<td>Ancell and Hakim (2007); Torn and Hakim (2008)</td>
</tr>
</tbody>
</table>
guidance products and accounted for the constraints of the aircraft. Second, while all the techniques possess limitations, they have mostly demonstrated the ability to provide reliable guidance under operational time constraints, and they show potential in predicting the impact of assimilating observations on selected forecasts. However, some crucial practical issues exist, such as whether the observations sample the full extent of the target areas, and whether the assimilation scheme and model are able to cleanly translate improved analyses within the target areas into better forecasts.

FIELD CAMPAIGNS. Around 20 field campaigns have included a component in which observations were targeted to improve predictions of a weather event. Additionally, a handful of relevant studies that did not involve a field campaign have been performed. A chronology of the field campaigns is provided in Table 2, together with a summary of their goals, instrumentation, the targeting guidance products used, evaluation results, and key publications. Many further publications that focus on the field campaigns are listed in the supplemental online bibliography (labeled with an F). The goals and new developments associated with the major experiments together with lessons learned are summarized in this section.

Pre-THORPEX. The establishment of NOAA’s Hurricane Synoptic Flow experiments between 1982 and 1996 was driven by the need to improve forecasts of TC track. Dropwindsonde observations were targeted in the synoptic environment of the TC to augment the satellite data that were assimilated into numerical models. The success of these early experiments led to NOAA establishing an annual operational “synoptic surveillance” program in 1997 with a dedicated new aircraft, the Gulfstream IV (G-IV) jet. The goal of this program has been to release dropwindsondes in critical areas, in order for 12–60-h forecasts to inherit the benefits from the data prior to and during the issuance of watches and warnings. This program also motivated the initiation of Taiwan’s annual Dropsonde Observations for Typhoon Surveillance near the Taiwan Region (DOTSTAR) program in 2003. The consensus view is that TC surveillance is practicable operationally, and that the benefits of better preparations and responses due to improved forecasts justify the costs. A side benefit is that many published studies have resulted from the surveillance data.

In the extratropics, the multiple successes of FASTEX followed by the more focused North Pacific Experiment (NORPEX) experiment, which coincided with an El Niño event in early 1998, led to NOAA establishing the annual Winter Storm Reconnaissance (WSR) program in 1999. The objective of WSR was to deploy targeted dropwindsondes over the northern Pacific Ocean to improve 1–5-day forecasts of events over the United States. Forecasters from the National Weather Service selected events of interest each day such as a potential snowstorm, assigned a priority to each event, and made the call on where to deploy the aircraft (if at all). The smooth implementation of WSR demonstrated the practicability of an operational end-to-end system that efficiently combined human decision-making with automated components such as the targeting guidance and subsequent flight track selection. The collection of data from over 5000 dropwindsondes through the duration of WSR (1999–2013) provided a large sample of high quality observations relevant to developing storms over the Pacific Ocean. Although these data could be utilized for a variety of studies, their actual use beyond a small number of targeting papers appears to be limited.

The THORPEX era. The Atlantic Observing-System Research and Predictability Experiment (THORPEX) Regional Campaign (A-TReC) in 2003, sponsored by the European Meteorological Network (EUMETNET) Composite Observing System (EUCOS), was the first extratropical field campaign over the northern Atlantic Ocean to employ targeted observations since FASTEX. Its goal was to test the feasibility of quasi-operational targeting with a large variety of in situ and remotely sensed observational data. More refined versions of the targeting techniques were used, including SVs with norms based on analysis errors, adjoint models that included moisture, observation sensitivity, and an upgraded ETKF. In 2006, a THORPEX component of the African Monsoon Multidisciplinary Analyses (AMMA) was aimed at improving forecasts of African rainfall and easterly waves that may lead to tropical cyclogenesis. An innovation was the “driftsonde” (Drobinski et al. 2013), comprising a large stratospheric balloon and gondola carrying up to 40 dropwindsondes, which could be released on command in a target region. The driftsonde was also employed during Concordiasi in 2010, as part of the THORPEX International Polar Year efforts. These campaigns all demonstrated the successful large-scale coordination of a variety of observing platforms between several nations.

In the mid-2000s, a EUCOS study suggested that the degradation of the ground-based observational network would have a significant negative impact on the forecast skill. This paved the way for the
### Table 2. Summary of field campaigns and studies that included a targeting component. The descriptions, results, and references give a brief summary with a specific focus on the aspect of targeted observations as opposed to the full scope of the field experiment. Campaigns listed in italics were or are held annually. The color code refers to tropical cyclones (orange), midlatitude systems (blue), precipitation and convective weather (green), and high-latitude and polar weather (purple).

<table>
<thead>
<tr>
<th>Name</th>
<th>Periods and cases</th>
<th>Forecasts</th>
<th>Techniques</th>
<th>Targeted observations</th>
<th>Main results of evaluations</th>
<th>Key references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synoptic flow</td>
<td>1982–96; 18 experiments</td>
<td>12–60-h TC track in Atlantic</td>
<td>Synoptic reasoning</td>
<td>NOAA P-3 aircraft</td>
<td>Mean errors in 12–60-h track forecasts in NOAA models reduced by 16%–30%</td>
<td>Burpee et al. (1996)</td>
</tr>
<tr>
<td>Hurricane synoptic surveillance</td>
<td>1997–present; 10–50 missions per year, depending on TC activity</td>
<td>12–60-h TC track, mostly in Atlantic</td>
<td>Synoptic reasoning; ensemble variance</td>
<td>NOAA G-IV and U.S. Air Force (USAF) C-130 aircraft; 20–30 dropwindsondes per flight</td>
<td>10%–15% average improvement in NCEP GFS track forecasts through 60 h; negligible improvements beyond 72 h; minimal impact on Geophysical Fluid Dynamics Laboratory (GFDL) forecasts</td>
<td>Aberson (2010) and references therein</td>
</tr>
<tr>
<td>FASTEX</td>
<td>Jan–Feb 1997; 19 intensive observing periods</td>
<td>1–3-day Atlantic midlatitude cyclones</td>
<td>SV; ET; adjoint; quasi-inverse</td>
<td>Aircraft based in Ireland and North America; ships, soundings, surface data, and satellites; ~400 dropwindsondes</td>
<td>Positive impact over Atlantic and western Europe in short range (2 days or less); around 10%–15% for most modeling and assimilation systems</td>
<td>Joly et al. (1999); Langland (2005) and references therein</td>
</tr>
<tr>
<td>NORPEX-98</td>
<td>Jan–Feb 1998; 27 days; 38 missions</td>
<td>1–3-day winter storms over North America and Mexico</td>
<td>SV; ET</td>
<td>~700 dropwindsondes; winds from geostationary satellites</td>
<td>Improved 2-day Navy Operational Global Atmospheric Prediction System (NOGAPS) forecasts by 10% on average; relatively small improvement in ECMWF</td>
<td>Langland et al. (1999b)</td>
</tr>
<tr>
<td>WSR</td>
<td>Annual, Jan–Mar, 1999–2013; 20–30 cases per year</td>
<td>1–5-day winter weather conditions over North America</td>
<td>ETKF</td>
<td>NOAA G-IV aircraft and USAF C-130s in AK, HI, and sometimes Japan; ~700 dropwindsondes per year</td>
<td>RMS surface pressure errors during 1999 and 2000 reduced by 10%–25% in low-resolution NCEP GFS; approximately 70% of cases have been improved on average through the past decade of WSR programs</td>
<td>Szunyogh et al. (2002)</td>
</tr>
<tr>
<td>DOTSTAR</td>
<td>Annual, 2003–present; 75 cases through the end of 2015</td>
<td>I–4-day TC tracks in the western North Pacific basin</td>
<td>Ensemble sensitivity; ETKF; ADSSV</td>
<td>Astra aircraft stationed in Taiwan; 13–20 dropwindsondes per mission</td>
<td>&gt;14% average improvement in NCEP GFS, NOGAPS, and JMA 1–3-day forecast track errors (10 cases in 2004); 10%–20% average improvement in NCEP GFS 1–5-day track forecasts, with 60% of all cases improved (42 cases in 2003–09); minor improvements and degradations in ECMWF</td>
<td>Wu et al. (2005); Chou et al. (2011)</td>
</tr>
<tr>
<td>DEEPWAVE</td>
<td>Jun–Jul 2014; 16 cases, not all for targeting</td>
<td>Gravity wave predictability near New Zealand</td>
<td>Adjoint sensitivity</td>
<td>Evaluations to be completed</td>
<td></td>
<td>Fritts et al. (2016)</td>
</tr>
<tr>
<td>Name</td>
<td>Periods and cases</td>
<td>Forecasts</td>
<td>Techniques</td>
<td>Targeted observations</td>
<td>Main results of evaluations</td>
<td>Key references</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------</td>
<td>----------------------------------------</td>
<td>-------------------------------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>SHOUT</td>
<td>Aug–Sep 2015; two TCs</td>
<td>TC track, structure, and intensity; all forecast ranges</td>
<td>Ensemble sensitivity</td>
<td>NOAA-dedicated missions aboard National Aeronautics and Space Administration (NASA) Global Hawk unmanned aircraft; dropwindsondes and remote sensing</td>
<td>Evaluations to be completed</td>
<td>Langland (2005); Rabier et al. (2008)</td>
</tr>
<tr>
<td>A-TReC</td>
<td>Oct–Dec 2003; 32 events</td>
<td>1–3-day high-impact weather events, mostly over Europe; a few storms affecting eastern North America</td>
<td>SV; ETKF; adjoint (dry and moist); observation sensitivity</td>
<td>Dropwindsondes from four aircraft; special rawinsondes; drifting buoys; AMDAR; airborne Doppler Wind Lidar (DWL); rapid-scan Atmospheric Motion Vectors AMVs</td>
<td>Small positive impact over large domains; overall improvement in 32% of 38 forecasts using the Met Office (UKMO) system; for ECMWF, forecasts of mean sea level pressure were improved (by at least 10%) in 24% of all cases; NOGAPS FSO showed the highest impact per observation by targeted drop-windsondes</td>
<td>Buizza et al. (2007); Cardinali et al. (2007); Bauer et al. (2011)</td>
</tr>
<tr>
<td>ECMWF studies (no field component)</td>
<td>Winter 2003–04 (92 cases); summer 2004 (91 cases); Jul–Sep 2008; Dec 2008–Feb 2009</td>
<td>2 days over North America and Europe, 3–4 days in the Southern Hemisphere</td>
<td>SV</td>
<td>Operationally assimilated observations removed from target or random areas; Density of radiance data increased in target areas</td>
<td>Removing targeted observations over Pacific (Atlantic) reduced 2-day forecast errors of 500-hPa Z by 4.0% (2.0%); greater reduction than removing observations in random locations; increasing the radiance data density in SV target areas improved forecasts at all levels up to 3 days in the Southern Hemisphere</td>
<td>Faccani et al. (2009); Agusti-Panareda et al. (2010)</td>
</tr>
<tr>
<td>AMMA (THORPEX component)</td>
<td>Aug 2006</td>
<td>1–3-day African weather events, including easterly waves</td>
<td>Adjoint; ETKF</td>
<td>Rawinsondes over Africa; dropwindsondes launched from driftsonde gondolas</td>
<td>Large impact on analysis fields over Africa and improvement of 1-day precipitation over the central Sahel; positive downstream impact over Europe in the 2–3-day range</td>
<td>Faccani et al. (2009); Agusti-Panareda et al. (2010)</td>
</tr>
<tr>
<td>Greenland Flow Distortion Experiment (GFDeX)</td>
<td>Feb–Mar 2007; four cases</td>
<td>1–2-day forecasts across northwest Europe</td>
<td>SV; ETKF</td>
<td>Supplemental rawinsondes; dropwindsondes from aircraft around southern Greenland and Iceland</td>
<td>Neutral to small forecast improvements, max 5%</td>
<td>Renfrew et al. (2008); Irvine et al. (2009)</td>
</tr>
<tr>
<td>Convective and Orographically Induced Precipitation Study (COPS)/European THORPEX Regional Campaign (E-TReC)</td>
<td>Jun–Aug 2007; during COPS, 25 flights for water vapor lidar; during E-TReC, simultaneous aircraft missions upstream</td>
<td>24–36-h precipitation over France and Germany</td>
<td>Adjoint; SV; ETKF for E-TReC</td>
<td>DWL, airborne water vapor lidar (COPS), dropwindsondes from aircraft (E-TReC); EUCOS rawinsondes and enhanced AMDAR over central and southern Europe</td>
<td>COPS airborne water vapor lidar produced a positive impact on forecasts of 6-hourly precipitation out to 24 h; precipitation sum forecasts over the entire period improved by 10%</td>
<td>Wulfmeyer et al. (2008); Bielli et al. (2012)</td>
</tr>
<tr>
<td>Name</td>
<td>Periods and cases</td>
<td>Forecasts</td>
<td>Techniques</td>
<td>Targeted observations</td>
<td>Main results of evaluations</td>
<td>Key references</td>
</tr>
<tr>
<td>------</td>
<td>------------------</td>
<td>-----------</td>
<td>------------</td>
<td>-----------------------</td>
<td>-----------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>International Polar Year/THORPEX</td>
<td>Mar 2008; three flights</td>
<td>1–2-day polar lows</td>
<td>SV; ETKF</td>
<td>Dropwindsondes released from DLR aircraft</td>
<td>Variable improvements depending on forecast time and verification metric</td>
<td>Irvine et al. (2011)</td>
</tr>
<tr>
<td>T-PARC Summer Phase/Tropical Cyclone Structure 2008 (TCS-08)</td>
<td>Aug–Sep 2008</td>
<td>1–4-day TCs and extratropical transition in the western North Pacific; a few non-TC cases</td>
<td>SV; ETKF; adjoint; ADSSV; ensemble variance</td>
<td>&gt;1500 dropwindsondes from four aircraft; DWL and water vapor lidar; rawinsondes and observations on research vessels, driftsondes, rapid-scan geostationary AMVs</td>
<td>20%–40% improvement to NCEP GFS and KMA WRF track forecasts; modest improvements to forecasts up to 3 days in ECMWF and JMA</td>
<td>Weissmann et al. (2011)</td>
</tr>
<tr>
<td>T-PARC Winter Phase</td>
<td>Jan–Mar 2009</td>
<td>1–5-day winter storms over North America</td>
<td>ETKF</td>
<td>Dropwindsondes from NOAA and USAF aircraft; rawinsondes over Russia</td>
<td>75% of the 52 forecast cases of 1–5 days were improved; magnitude of improvement not published</td>
<td>Y. Song (2011, personal communication)</td>
</tr>
<tr>
<td>EURORISK PREVIEW/MEDEX</td>
<td>For PREVIEW, Feb–Dec 2008; 54 events; for MEDEX, Oct–Dec 2009; 132 cases</td>
<td>1–2-day high-impact weather events over Europe and especially the Mediterranean</td>
<td>SV; ETKF; Kalman filter sensitivity</td>
<td>For PREVIEW, 1402 land stations; 226 European Meteorological Network (EUMETNET) EU-Automated Shipboard Aerological Programme (E-ASAP) ship-based measurements. For MEDEX, 484 additional rawinsondes in Europe and Algeria</td>
<td>Modest average improvements (2%) with supplemental rawinsondes; more substantial improvements with additional targeted satellite data (9%) during MEDEX; impacts were more modest during PREVIEW (for a subset of autumn cases)</td>
<td>Jansa et al. (2014); Campins et al. (2013)</td>
</tr>
<tr>
<td>Concordiasi</td>
<td>Sep–Nov 2010</td>
<td>1–4-day events in Antarctica</td>
<td>SV</td>
<td>25% of 644 dropwindsondes launched from 13 driftsondes were targeted</td>
<td>Improvement of short-range forecasts in Naval Research Laboratory (NRL), NASA, ECMWF, and Météo-France, using FSO; dropwindsonde impact was small compared with satellite data, though large impacts were seen in upper-tropospheric winds and lower-tropospheric temperatures; qualitative impact on forecast fields in ECMWF and Météo-France systems using OSEs</td>
<td>Rabier et al. (2013); Boullot et al. (2016)</td>
</tr>
<tr>
<td>HyMeX</td>
<td>2012–present</td>
<td>Droughts and heat waves; mesoscale heavy precipitation</td>
<td>SV</td>
<td>American and European research aircrafts, rawinsondes, AMDAR, and drifting boundary layer (BL) balloons</td>
<td>Evaluations to be completed</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2. Continued.**

DECEMBER 2016 | DECEMBER 2016
establishment of an adaptively controlled operational observing system in Europe, with increased flexibility in the land-based rawinsonde network plus in situ components such as Aircraft Meteorological Data Relay (AMDR) and rawinsondes from merchant ships. To coordinate requests for multiple types of targeted observations during field operations, the EURORISK PREVIEW Data Targeting System (DTS) was established. The web-based facility allows registered users to identify potential high-impact weather events, request sensitive area calculations for chosen cases, identify and issue requests for targeted observations, and monitor the requested observations. Successful examples of the use of the PREVIEW DTS were during the Mediterranean Experiment (MEDEX) and the THORPEX Pacific Asian Regional Campaign (T-PARC).

T-PARC possessed an ambitious scope. The summer phase in 2008 was aimed at investigating the science and predictability of the life cycle of TCs in the western North Pacific basin, from formation through to recurvature and extratropical transition, including tropical–extratropical interactions and the flow far downstream in the midlatitude storm track. As had been the case for FASTEX, this campaign stood out because of the remarkable number of international collaborations and publications generated, ranging from predictability and processes through to targeting methods and evaluations (Parsons et al. 2017). The winter phase in early 2009 was aimed at targeting observations to improve forecasts out to a week over North America. During both phases, the PREVIEW DTS facilitated the comparison between different targeting guidance products and improved the efficiency of the decision-making process for multiple observation types.

Given the increased emphasis on mesoscale prediction in the community in the 2010s, several field campaigns have followed suit. Examples include the Hydrological Cycle in the Mediterranean Experiment (HyMeX), the Mesoscale Predictability Experiment (MPEX), and the Deep Propagating Gravity Wave Experiment (DEEPWAVE) in Europe, the United States, and New Zealand, respectively.

Targeted observations have also been tested in the ocean, using mobile platforms such as underwater gliders (Curtin and Bellingham 2009). These programs served to initiate collaborations across multiple disciplines (e.g., physical and biological oceanography) that may not otherwise have occurred, as well as coordinating multiple types of underwater vehicles, and establishing new data sharing and formatting protocols.

Overall, much has been developed and learned through the implementation of these field campaigns, which took place across six continents (Europe, Asia, Africa, Australasia, Antarctica, and North America). There were several programmatic successes: i) an adaptively controlled data targeting system could be run smoothly; ii) new targeting techniques were developed and employed in real time; iii) existing instrumentation was adapted and new instrumentation was tested; iv) long-term multinational collaborations were established; v) students and scientists in dynamics, data assimilation, and modeling were brought together with forecasters and exposed to operational models; and vi) databases of observations, targeting guidance, and model outputs were created and shared. Challenges included suboptimal sampling due to the limited range of the aircraft, and the need to rely on guidance initialized over 2 days prior to $t_0$ because of the lead time required for deployment. What has yet to be determined in this paper is whether the targeted observations improved the forecasts, and that is the topic of the next section. A discussion of how the targeted observations are evaluated is also provided in the second sidebar.

**RESULTS FROM EVALUATIONS.** At the conclusion of a field campaign, operational and research teams usually perform evaluations of the influence of targeted observations on numerical forecasts. The general results for each field campaign, together with a key reference, are summarized in Table 2, with many more references on evaluations available in the online bibliography (labeled with an E). The results are difficult to interpret or generalize in many situations. A net improvement using one variable for evaluation (e.g., 500-hPa geopotential height) does not imply that another variable (e.g., precipitation) will be improved. Furthermore, the results may be sensitive to the sample of cases, the room for improvement in the sample, the forecast window ($t_f - t_0$), the model and data assimilation scheme, and the choice of verification region. The following subsections are organized by common themes.

**Field campaigns in the extratropics.** To evaluate the impact of assimilating the FASTEX targeted observations, several groups associated with major operational centers performed Observing System Experiments (OSEs) in the late 1990s. While the results varied by center, the average improvements of the forecast events in global models were found to be around 10%–15%. Importantly, it was suggested by Bergot (1999) that the results depend substantially on the
data assimilation scheme, a notion that is now widely accepted. Similar results were found for NORPEX. The early evaluations of WSR found improvements to low-resolution National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) forecasts within the verification regions in around 70% of all cases, with an average magnitude that was considered equivalent to a 12-h gain in forecast lead time. Forecasts of sea level pressure and total energy were used, though precipitation forecasts were not verified. No evaluations were published for the 2002–10 WSR programs, though NCEP suggested that the same proportions of forecasts were improved with an average improvement of 10%–20% for cases selected as a high priority (Y. Song 2011, personal communication). Recent studies in the United States and Europe found that the influence of assimilating WSR data, using the latest assimilation systems, was negligibly small (Hamill et al. 2013). These conclusions led to NOAA suspending the WSR program from 2014, though occasional missions are still requested.

The targeted WSR data were usually collected in features such as upper-tropospheric waves and mature cyclones. Such areas are often cloudy, limiting the influence of satellite data and suggesting the need for detailed in situ sampling of the vertical structure. Generally, the assimilation of observations sampled within the selected target regions produced a larger improvement in the forecasts than those outside the target regions, even if the overall magnitude of the improvement was often small. The propagation of the effect of the targeted observations was found to be associated with processes such as downstream baroclinic development (Szunyogh et al. 2002).

At least six centers used OSEs to evaluate the influence of assimilating the targeted A-TReC data. Unlike FASTEX, the targeted data made little difference to the forecasts, owing to the forecasts over Europe being largely accurate without the targeted observations. These neutral results were corroborated by studies using the forecast sensitivity to observations (FSO) method in which small improvements were found, even though the A-TReC dropwindsondes had been deployed in an area of relatively high sensitivity to observations. Langland (2005) suggested that improvements to regular analyses and forecasts due to “substantial increases in the amounts of regular observations from aircraft and satellites” during the 6-yr interval between FASTEX and A-TReC led to a reduction in the potential impact from the targeted observations. Evaluations from the next large extratropical THORPEX campaign, the winter phase of T-PARC in 2009, remain unpublished.

For Concordiasi, OSEs and FSO studies show qualitative and quantitative positive impacts per dropwindsonde, especially for lower-tropospheric temperature (where radiance data are less reliable) and upper-tropospheric wind. It was suggested that these improvements were in areas of high instability (seen in the targeted SVs) and low data coverage. In smaller extratropical campaigns, the few evaluations provided results that were interesting on a case-by-case basis, though without decisive conclusions.

The key lessons learned from targeting in the extratropics is that the magnitudes of the forecast improvements, while largely positive, are usually small and difficult to interpret scientifically and in terms of socioeconomic value. This could be for several reasons. In many cases, the forecast errors were small even without the targeted observations, leaving minimal room for improvement. This suggests the need for a technique that quantitatively predicts the error evolution prior to deployment, to identify cases in which targeted observations would likely be redundant. Several targeting methods, together with estimates of forecast uncertainty based on the ensemble variance, provide a framework for making such estimates. Additionally, given the lack of in-depth evaluations, a clear scientific understanding of how the data assimilation scheme is using the targeted data to improve the analysis, and the subsequent propagation of this improvement for future forecasts, remains elusive. The inability to sample the synoptic-scale extent of the target regions with aircraft also suggests that other types of targeted data with a broader range and frequency of coverage would be useful.

Field campaigns in the tropics. Evaluations of targeted observations aimed at improving TC track forecasts have almost entirely been conducted using OSEs. In NOAA’s early Hurricane Synoptic Flow experiments, the average track forecast improvements from the dropwindsonde data were very large (30%). During the first decade of synoptic surveillance (1997–2006) the NCEP GFS track forecasts were improved, though by a smaller amount (10%–15%). Similar improvements were found in the GFS for TCs worldwide during an active 6-week period in 2008 in which NOAA missions in the Atlantic and T-PARC missions in the western North Pacific took place [10%; Aberson (2011)]. Results from DOTSTAR have shown similar improvements to TC track forecasts. Individual case studies in the 2010s have shown a modest positive impact on the track forecasts. A key multi-agency paper from T-PARC illustrated the dependence on the modeling and assimilation
HOW ARE TARGETED OBSERVATIONS EVALUATED?

The most widely used method for observation evaluation is OSE. First, a “control” assimilation–forecast cycle is run through the period of interest with all operationally assimilated observations. Next, a parallel cycle, which is identical to the control except that the dataset in question is either added to or withheld from the assimilation, is run. The difference between the two forecasts integrated from analyses valid at the same time in the two cycles is the “data impact.” The improvement in the forecast is defined as the difference between the errors of the two forecasts integrated from analyses valid at the same time in the two cycles.

OSEs are standard in preoperational testing of new data types and have been used in all targeted observing campaigns. They are limited in that they are computationally expensive, particularly if one wishes to test different components of the datasets. The more recently developed forecast sensitivity to observations (FSO) method uses the adjoint of a data assimilation system to estimate the contribution of any subset of observations to the reduction in forecast error (Langland and Baker 2004). This technique is more efficient than the OSE and offers the capability to compute the impact on the forecast for any selected data type, location, or channel. While FSO does not quantitatively reproduce the results of an OSE, it has been demonstrated to yield qualitatively similar results. Several operational centers now use the FSO method to monitor the global observing system and the impacts on their models (Gelaro et al. 2010). However, this approach is limited by the tangent linear assumption and so is restricted to short-range forecasts. Recently, ensemble-based FSO methods have been developed and used.

The OSE and FSO methods complement each other, with each method providing unique and helpful information about the impact of observations.

system (Weissmann et al. 2011). NCEP and the Korea Meteorological Administration (KMA), who used a three-dimensional variational data assimilation (3DVAR) scheme, showed a large average improvement (20%–40%) in their 1–5-day track forecasts due to the targeted dropwindsonde data. In contrast, the European Centre for Medium-Range Weather Forecasts (ECMWF) and the Japan Meteorological Agency (JMA), who employed a four-dimensional variational data assimilation scheme, showed a lower improvement for the same sample, due in part to the forecasts without the targeted data having lower errors and leaving less room for improvement. The same paper also investigated tropical influences on midlatitude forecast skill and predictability using targeted T-PARC data. The 4–5-day cycled ECMWF forecasts of 500-hPa geopotential height in the mid-latitudes were found to improve modestly because of the dropwindsonde observations around two TCs. This was attributed to the aforementioned improvements in the corresponding TC track forecasts, which subsequently improved the long-wave pattern downstream.

Extending beyond the limited spatial and temporal coverage of aircraft data, a few studies have been conducted on the influence of assimilating extra satellite data. The benefits to forecasts of TC track and structure of assimilating specially processed rapid-scan atmospheric motion vectors were demonstrated during T-PARC (Berger et al. 2011). Further OSEs demonstrated that extra satellite data in SV target regions provided a consistent positive improvement to ECMWF TC track forecasts (Reynolds et al. 2013).

Given recent improvements in models and data assimilation schemes, as well as substantial efforts to advance regional, convection-permitting models, attention is being directed at predictions of TC structure and intensity. While not “targeted” with objective techniques, recent studies have demonstrated the positive impact of assimilating observations such as airborne Doppler radar in the TC inner core (e.g., Zhang and Weng 2015).

For other tropical weather, evaluations are limited beyond the AMMA studies cited in Table 2.

The main conclusion in the evaluations of TC track forecasts is that the targeted observations have on average improved forecasts by around 10% during the THORPEX era, though the results depend on the model and assimilation system. Although evaluations in the 2010s are limited, the continuation of the aircraft surveillance programs suggests that improvements to track forecasts of around 10% are sufficient to justify the need for targeted observations. Several of the lessons learned within the context of targeting for extratropical weather are also appropriate here, such as the need to discriminate between scenarios of high versus low room for forecast improvement, a need to understand scientifically and predict how the impact of the targeted observations is propagated forward, and the importance of sampling the extent of the target regions.
Mesoscale weather. The results from the MEDEX phases in 2008 and 2009 were encouraging. The targeted rawinsondes, which were often deployed upstream of the weather events, improved regional model forecasts of most parameters in the Mediterranean, including precipitation. Crucially, the improvements were found to depend upon the predictability of the forecast case. Targeted satellite data often provided additional improvements.

Results from 15 aircraft missions during MPEX, in which dropwindsondes were spaced about 150 km apart, found an overall small and positive impact from the targeted data (Romine et al. 2016). These results were consistent with the earlier synoptic-scale evaluations and also possess substantial case-to-case variability.

It is worth emphasizing that a wide variety of assimilation studies using finescale models and observations of high spatial and temporal resolution have been conducted in the past decade. Observations include radars and wind profilers, aimed at improving local, short-range forecasts of severe weather. Although targeting with objective guidance has not been used in these studies, similar concepts can be extended for use.

Targeting of routinely available data. Evaluations of targeted observations have also been done independently of field campaigns, including a series of OSEs conducted by the ECMWF. Buizza et al. (2007) demonstrated that observations in SV target areas reduced the errors in verification regions by a greater amount than observations in randomly selected areas. Furthermore, the impact of the targeted observations was found to be dependent on the region, season, baseline observing system, and flow regime. Bauer et al. (2011) asked the important question of whether additional routine data that are thinned in operations add value to forecasts. They found that if the densities of radiance data in SV target areas were increased, ECMWF forecasts up to 3 days were improved in the Southern Hemisphere.

The FSO method facilitates the comparison of specific observation types. The impacts of major observation types on 24-h forecast errors were found to be similar in different global modeling systems, though regional details varied. The largest error reductions were due to the assimilation of satellite radiances, geostationary satellite winds, rawinsondes, and commercial aircraft. A conclusion across the different modeling systems was that only a small majority (50%–54%) of the total observations assimilated improved the forecast (Gelaro et al. 2010). Most of the improvement resulted from a large number of observations that had relatively small impacts per observation. This finding amplifies the need to optimize the use of satellite data, and that regional targeting on a continuous basis may be more effective than occasional, limited-area sampling.

Targeting of new observational data. In the majority of field experiments, rawinsondes and dropwindsondes were the primary types of targeted observations. However, one of the goals of THORPEX had been to explore the utility of new observation types. In the THORPEX Implementation Plan, a new in situ sensors considered for targeting included stratospheric balloons, piloted and unmanned aircraft, rocketsondes, and bidirectional radiosondes. The last two were not developed, perhaps because of technological and financial challenges. The vision of targeting with unmanned aircraft is only beginning to be realized 10 years later [e.g., NOAA’s Sensing Hazards with Operational Unmanned Technology (SHOUT) program, which began in 2015]. Two other new types of observing platform have been developed, tested, and evaluated. The first is the driftsonde balloon, whose ability to release drop-windsondes in selected regions offers a new observational capability over the oceans without the need for aircraft. Results from evaluations remain very limited.

The second is an airborne Doppler wind lidar. During A-TReC, the mean 2–4-day ECMWF forecast errors in 500-hPa geopotential height over Europe were reduced by 3% as a result of the assimilation of around 1600 wind profiles (Weissmann and Cardinali 2007). During T-PARC, the mean 2–4-day ECMWF TC track forecast errors were reduced by 50 km through the cumulative assimilation of 2500 wind profiles (Weissmann et al. 2012). Additional new observation types may be beneficial in the future, though structured studies to investigate their potential prior to deployment may be more cost effective and yield greater forecast benefits.

CONCLUSIONS AND DISCUSSION. The encouraging results in the 1990s led to optimism that targeted observations would benefit forecasts to the extent that adaptive networks would regularly supplement the global observational structure. Accordingly, a series of field campaigns and projects were initiated prior to and during the THORPEX era (2005–14), with observations targeted to improve 1–5-day global model forecasts of high-impact weather events. The PREVIEW DTS and aircraft surveillance programs
impressively demonstrated how segments of the available observing network can be coordinated across multiple nations and made adaptable on a daily basis. Evaluations in the extratropics found that the influence of assimilating targeted observations was positive though small. In the tropics, the targeted observations usually improved TC track forecasts, though evaluations in the 2010s are lacking. In the multiagency collaborative OSE studies, the conclusions were found to depend on the model, the data assimilation scheme, and the treatment of the observations. In FSO studies, the impact per targeted observation was found to be high, but their cumulative benefit was small given the vast quantity of routine observations.

The fact that many results are inconclusive is not surprising, for several reasons. First, synoptic-scale NWP is maturing. Advances in the Global Observing System, the improved use of observations by advanced data assimilation schemes, and better model physics and resolution are raising the bar for forecast improvement. As the forecast skill increases, the average marginal impact of any individual observing system decreases. Second, the limited range of in situ observations restricts the data coverage to a fraction of the size of typical target regions, suggesting the need to consider the targeting of more broadly available observations such as satellite radiances. Third, the results depend on the flow regime and may vary beyond the limited samples tested during the field experiments or evaluation studies. The large number of cases in which forecast errors are low without the targeted data suggests that low-predictability cases with substantial room for improvement via targeted observations are not common. Additionally, it is possible in some situations that errors in the model tend to obfuscate the positive improvements to the initial conditions. Fourth, the results can be inconclusive as a result of inconsistencies in the evaluation methods and ambiguity in the perceived value of the verification metric. As an example, the most conclusive results to date have been related to TC track, whose forecast improvement is relatively straightforward to appreciate in terms of socioeconomic gain. On the other hand, an “average 4% improvement” using an area-averaged norm of an atmospheric variable does not inform the decision-maker about whether the targeted data are cost effective. Finally, the coordination of evaluations across multiple centers with common metrics is challenging, given the different priorities of each agency and the labor and computational expenses of OSEs. It should be noted that one of the most significant successes was a multiagency coordination of FSO studies.

The types of instrumentation used for targeted observations have mostly been supplemental rawinsondes and aircraft-borne dropwindsondes, whose coverage is limited. Some of the new observation platforms proposed at the start of the THORPEX era were developed and put into action (e.g., the driftsonde), whereas others did not materialize (e.g., the bidirectional radiosonde). It remains an open question as to which new platforms would be most useful and cost effective for targeting using the next generation of numerical weather prediction systems. Within the context of short-range, high-resolution forecasts of weather hazards, platforms such as small unmanned aircraft may be useful. However, many technical, financial, and logistical challenges are faced in the development of more comprehensive platforms that can be deployed to important areas (e.g., to sample lower-tropospheric profiles of wind and moisture to capture convective initiation). Within the context of satellite observations, there have been very few studies on how to optimize the use of routinely available data, despite the promising results of Bauer et al. (2011). OSE and FSO studies can be conducted to address this issue. Additionally, new developments such as cost-effective constellations of microsatellites (e.g., Ruf et al. 2016) require investigation. Observing system simulation experiments (OSSEs; Hoffman and Atlas 2016) can be utilized to address questions concerning future observing systems prior to their development or launch. Combining OSSEs with predictability studies, a strategic plan can be designed to prioritize platforms for routine and targeted observations that can provide the largest improvements to societally relevant forecasts.

The objective techniques used to identify preferred locations for targeted observations have been shown to be superior to selecting targets subjectively or randomly. The techniques are likely not the first-order problem, though there is room for improvement, such as in the treatment of nonlinearities, incorporating error characteristics from the data assimilation scheme, and quantitatively predicting the potential forecast impact in order to decide whether substantial forecast gains can be achievable.

The effectiveness of targeted observations depends on the data assimilation scheme, as has been reported in several publications since FASTEX. During the THORPEX era, important advances have been made in variational and ensemble-based data assimilation, as reported in Parsons et al. (2017). As data assimilation schemes advance, so do the impacts of assimilating observations. In this changing landscape, in which forecast models and the observational network are also advancing, the results from the THORPEX era (e.g.,
those that used 3DVAR) may not be directly relevant to planning over the next decade. It is also an interesting question whether improvements in data assimilation will render targeted observations more or less useful. One can argue that there is less room for improvement if the superior assimilation of routine observations yields smaller analysis and forecast errors. On the other hand, one can also argue that an improved data assimilation scheme is able to better exploit the targeted data, spreading their influence within regions in which analysis errors may be large or may grow quickly.

The relationship between forecast improvements found by the evaluation studies and the corresponding socioeconomic values is largely unknown (though it is more tangible for TC tracks). Research designed to quantify the benefits would help sharpen the development and usage of observing systems. Furthermore, forecast improvement has commonly been quantified for deterministic forecasts. Given the increased emphasis of ensemble forecasts as part of an integrated forecasting system, the impact of targeted observations on narrowing the forecast uncertainty and improving the probability density function would be another avenue for investigation.

In summary, the initial vision that the Global Observing System would ultimately be supplemented by networks on an adaptive basis, or even optimized into a fully adaptive network, has not been realized. In 2016, the only targeted observations deployed operationally are for TC surveillance. The proliferation of studies has slowed down, as is evident in the bibliography included in the online supplement to this paper. Nevertheless, the author believes that the concept of targeting observations to improve forecasts of high-impact weather events remains useful, even though a compelling justification is elusive. Several ideas from over a decade ago, such as the assimilation of targeted satellite data (Langland 2005) remain relevant. However, one constraint is that targeted observational networks are expensive and time-consuming to design. A strong justification is also needed to propose an alteration to an operational NWP system to automatically assimilate observations in target areas. In light of these conclusions, the author suggests a creative rethinking of the approach taken to accomplish the goal of using targeted observations in a cost-effective manner, using state-of-the-art and future prediction systems. Some recommendations aimed at this goal are provided in the final section.

**Recommendations.** Based on annual meetings and discussions with the DAOS Working Group, as well as with numerous personnel in the field, several recommendations are made. They are based partially on the conclusions and lessons learned during the field campaigns and studies, and also via the identification of deficiencies that had not been addressed during those field campaigns, along with parallel advances in observations, modeling, and data assimilation.

**Optimize use of the global observational network.**

i) Exploit observational resources that are *routinely available* and potentially adaptable, such as special processing, targeting, and/or thinning of satellite radiances and winds. Facilities such as the PREVIEW DTS can be used for this purpose. Systematic, regime-based targeting can be continuously conducted on a broader scale over days to weeks, especially during low-predictability regimes.

ii) Explore the utility of *new observing platforms*, such as space-based lidar; high-altitude, long-endurance unmanned aircraft; and surface data such as pressure readings from smart phones. Rapid response and assimilation would be essential.

iii) With an increasing emphasis on *small-scale prediction*, explore the potential for targeting the use of convective-scale models. Potential observational resources include mobile mesonets for forecasts of severe weather, and airborne Doppler radar and unmanned aircraft for forecasts of tropical cyclone structure and intensity. New techniques for targeting guidance may need to be developed.

**Invest in quantitative evaluations and understanding.**

i) Field programs, whether in research or operations, require *routine evaluations*.

ii) Explore concepts of future targeted observations via OSSEs. The OSSE uses the OSE concept to evaluate the potential impact of assimilating *synthetic observations*, sampled from a high-resolution “nature run,” which also serves as the “truth” against which forecasts are evaluated. OSSEs are presently being used to quantify the impact of future satellite observing systems on NWP. They possess strengths and limitations.

iii) Assess the role of the *data assimilation scheme*, including the treatment of targeted observations versus routinely available observations, and the use of flow-dependent spatial structure functions that control the spread of the influence of the observations.

iv) Quantify the *reduction in forecast uncertainty* realized by the assimilation of targeted observations, using ensemble-based probabilistic forecasts.
Assess the value of improved forecasts.

i) Identify the low-predictability regimes in which forecasts are most likely to “bust” because of initial condition errors and, therefore, where targeted observations are expected to yield the highest benefit. Identify priorities for phenomena whose prediction requires improvement. In parallel, improve the basis for quantitatively predicting the influence of targeted observations on forecast errors.

ii) Improve our understanding of the socioeconomic impact of assimilating targeted observations, via user-focused measures of the value of forecast improvements.

In closing, considerable progress has been made toward the goal of supplementing the Global Observing System with targeted observations, with enhanced international cooperation fostered through THORPEX. However, many open and fundamental questions remain, most notably the overall cost effectiveness and benefits to society of creating a network on a more adaptive basis. Investments in the collaborative design and evaluation of strategies to target observations are necessary to ensure that modifications to the Global Observing System are of value to society.

ACKNOWLEDGMENTS. The advancement of targeted observations would not have been possible without the financial support from several nations, and the substantial efforts of all involved in the planning, execution, and evaluation of field campaigns and related initiatives. The THORPEX Program Office enabled many of these activities. The author thanks the WMO/WWRP DAOS Working Group for suggesting and reviewing this article. Finally, the comments and criticisms of three anonymous reviewers greatly improved the article.

REFERENCES


HURRICANE PIONEER
Memoirs of Bob Simpson
Robert H. Simpson with Neal M. Dorst

In 1951, Bob Simpson rode a plane directly into the wall of a hurricane—just one of his many pioneering explorations. This autobiography of the first director of the National Hurricane Research Project and co-creator of the Saffir-Simpson Hurricane Scale starts with childhood remembrance and ends in first-hand account of a revolutionary

© 2014, PAPERBACK
ISBN: 978-1-935704-75-1
LIST $30  MEMBER $20

CLIMATE CONUNDRUMS
What the Climate Debate Reveals About Us
William B. Gail

This is a journey through how we think, individually and collectively, derived from the climate change debate. With wit and wisdom, Gail explores several questions: Can we make nature better? Could science and religion reconcile? Insights from such issues can help us better understand who we are and help

© 2014, PAPERBACK
ISBN: 978-1-935704-74-4
LIST $30  MEMBER $20

AN OBSERVER’S GUIDE TO CLOUDS AND WEATHER
A Northeast Primer on Prediction
Toby Carlson, Paul Knight, and Celia Wyckoff

With help from Penn State experts, start at the beginning and go deep. This primer for enthusiasts and new students alike will leave you with both refined observation skills and an understanding of the complex science behind the weather: the ingredients for making reliable predictions of your own.

© 2014, PAPERBACK
LIST $35  MEMBER $20

AMS Books are available to groups and booksellers, and desk copies may be obtained, through our distributor, The University of Chicago Press: 1-800-621-2736 or custserv@press.uchicago.edu.