National Weather Service Data Needs for Short-Term Forecasts and the Role of Unmanned Aircraft in Filling the Gap

Results from a Nationwide Survey

Adam L. Houston, Lisa M. PytlikZillig, and Janell C. Walther

ABSTRACT: Inclusion of unmanned aircraft systems (UAS) into the weather surveillance network has the potential to improve short-term (<1 day) weather forecasts through direct integration of UAS-collected data into the forecast process and assimilation into numerical weather prediction models. However, one of the primary means by which the value of any new sensing platform can be assessed is through consultation with principal stakeholders. National Weather Service (NWS) forecasters are principal stakeholders responsible for the issuance of short-term forecasts. The purpose of the work presented here is to use results from a survey of 630 NWS forecasters to assess critical data gaps that impact short-term forecast accuracy and explore the potential role of UAS in filling these gaps. NWS forecasters view winter precipitation, icing, flood, lake-effect/lake-enhanced snow, turbulence, and waves as the phenomena principally impacted by data gaps. Of the 10 high-priority weather-related characteristics that need to be observed to fill critical data gaps, 7 are either measures of precipitation or related to precipitation-producing phenomena. The three most important UAS capabilities/characteristics required for useful data for weather forecasting are real-time or near-real-time data, the ability to integrate UAS data with additional data gathered by other systems, and UASs equipped with cameras to verify forecasts and monitor weather. Of the three operation modes offered for forecasters to consider, targeted surveillance is considered to be the most important compared to fixed site profiling or transects between fixed sites.

<https://doi.org/10.1175/BAMS-D-20-0183.1>

Corresponding author: Adam L. Houston, ahouston2@unl.edu

In final form 5 July 2021

©2021 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

Assessments of the value of any new sensing capability under consideration for integration into the U.S. weather surveillance network should be guided by the data requirements of the principal stakeholders [National Research Council (NRC); NRC 2009; Heinselman et al. 2012; Houston et al. 2020]. The identification of data gaps for short-term (<1 day) forecasting has been the focus of prior reviews (e.g., Dabberdt et al. 2005; NRC 2009); but, as noted by Houston et al. (2020), these were rarely (if ever) directly based on surveys of primary users. Unmanned aircraft systems (UAS) have been proposed as an observing system capable of filling data gaps in the lower atmosphere that impede both manual and numerical weather prediction (NRC 2009; Koch et al. 2018; National Academies of Sciences, Engineering, and Medicine 2018; Vömel et al. 2018; McFarquhar et al. 2020). The potential impact of UAS on numerical weather prediction has been explicitly considered via observing system simulation experiments (Chilson et al. 2019) and observing system experiments (Flagg et al. 2018; Cione et al. 2020; Jensen et al. 2021; Leuenberger et al. 2020). The value of UAS for operational meteorology is likely to extend beyond their impact on NWP (Houston et al. 2020). Ultimately, with a strong theoretical case based on expert reviews, numerical studies, and surveys of end users, there is at least theoretical justification to consider the technical and economic challenges involved in weather surveillance network modernization with UAS.

In this article, results are presented from the analysis of responses to a survey distributed to NWS forecasters that aimed to determine critical data gaps for short-term forecasting and the role that UAS might play in filling these gaps. Specifically, this survey aimed to quantitatively gauge 1) which atmospheric phenomena are difficult to forecast owing to data gaps, 2) which weather-related characteristics need to be observed to fill in these data gaps, and 3) what roles might UAS play in filling these gaps. The focus of the analysis is on a national assessment of data gaps. Regionalities are present and important to consider, but in the interest of brevity, regional analyses will be described in a companion article.

The survey was developed based on qualitative results from focus groups and interviews of a smaller number of NWS forecasters conducted by the authors (Houston et al. 2020). These focus groups and interviews enabled the development of the survey questions by providing a preliminary list of phenomena that forecasters have difficulty forecasting with “acceptable” precision and/or accuracy due to data gaps, a comprehensive list of atmospheric features that may need to be observed to assist in gap filling, and initial considerations from the forecaster perspective for the integration of UAS into the weather surveillance network (Houston et al. 2020).

Methodology

The survey¹ was composed of five components. The first component involved questions that allowed an assessment of the representativeness of survey results across geographic areas and NWS personnel roles or positions. The second component gauged forecaster experience forecasting various atmospheric phenomena. The third component asked forecasters to rate the extent to which data gaps impeded their ability to forecast the phenomena assessed in the second component. In the fourth component forecasters were asked to prioritize weather-related characteristics (e.g., storm damage, hydrometeor

¹ The present study and all of its methods and measures were reviewed and approved by the University of Nebraska Institutional Review Board for the Ethical Treatment of Human Subjects (RII Track-2 FEC: Unmanned Aircraft System for Atmospheric Physics, IRB Approval 20151115696 EX).

type, low-level jet) that would need to be observed to fill data gaps. Finally, forecasters were asked to rate the importance of specific UAS capabilities, UAS data characteristics, and UAS operation modes for collecting “useful data” for weather forecasts. The survey was administered online via Qualtrics. Participants were recruited via an email sent to all NWS employees by John Murphy, NWS Chief Operations Officer.

Results

Representativeness. Forecasters from all six NWS regions along with those from several national centers participated in the survey (Table 1). The overall participation rate was estimated at 31% based on the number of respondents and the total number of NWS forecasters (not the total number of NWS employees) at the time of the survey. The roles of the participants are listed in Table 2. The mean years of experience as a professional forecaster was 17.6 at the time of the survey with a fairly flat distribution of experience across the range from 0 to ~20 years (Fig. 1).

Forecasting type. The aim of the “forecasting type” question is to broadly characterize the types of weather forecasted by the NWS. Regional differences in the frequency with which these phenomena are forecasted serves as a validity check on the results. More importantly, these responses allow consideration of the relative importance of data gaps associated with these phenomena based on how frequently forecasters actually forecast a particular phenomenon.

Seven phenomena were included in the survey based on focus group results: fire, flood, hurricane, thunderstorms, severe thunderstorms, lake-effect/lake-enhanced snow, and winter precipitation. Forecasters were asked if they “often,” “sometimes,” or “never” forecasted each phenomenon. The survey design deliberately omitted guidance on the numeric frequencies that should constitute “often” or “sometimes.” This omission was based on our contention that numeric frequencies would be difficult for forecasters to estimate accurately. Overlaps are expected when using qualitative expressions (Bocklisch et al. 2012) but our expectation is that, regardless of the actual numeric frequencies that a forecaster implicitly associates with each category, “sometimes” will always be less than “often.” Of the seven default phenomena all but “thunderstorms” are included in the analysis below. The lack of independence in the responses to thunderstorms and severe thunderstorms was deemed to be too small to justify separately reporting these results. The broadness of these “default” phenomena was a deliberate decision aimed at keeping these categories as region agnostic as possible. For

Table 1. Participation rate from the six NWS forecast regions and national centers.

Region	Respondents	Participation rate
Alaska	26	38%
Central	199	36%
Eastern	117	35%
National	37	23%
Pacific	9	21%
Southern	125	25%
Western	117	31%

Table 2. Positions of respondents.

Position	Respondents
Data acquisition program manager and observation program lead	11
General forecaster	232
Hydrometeorological technician	6
Lead/senior forecaster	211
Meteorologist intern	16
Meteorologist in charge	52
Science and operations officer	49
Warning coordination meteorologist	48

Fig. 1. Distribution of respondents' years of professional forecasting experience.

Fig. 2. Regional distributions of the relative occurrence of the three forecasting frequencies. In each inset, the relative occurrence (expressed as a percentage) of forecasters from a particular region who said they either often (dark blue), sometimes (light blue), or never (gray) forecasted a particular phenomenon.

example, no distinction is made between flash flooding, river flooding, and coastal flooding. Regionalities are present for the expected phenomena (e.g., lake-effect snow, hurricane; Fig. 2). Nevertheless, since the principal motivation for this work is to offer an assessment of the value of UAS for NWS forecasting to guide *national* decisions concerning the integration of UAS into the weather surveillance network, the generality of several of these phenomena seems justified. Across all regions (Fig. 3) the relative occurrence (expressed as a percentage) of each of the three frequencies (“often,” “sometimes,” or “never”) for winter precipitation, severe thunderstorms, floods, and fire are very similar.

Acknowledging the potential incompleteness of the default list, forecasters were allowed the opportunity to add and rate up to four additional phenomena. Additional phenomena (a total of 574 suggestions were offered by 284 respondents) have been consolidated into 14 categories (Fig. 4). Of the 574 original suggestions, 63 are best classified as falling in the category of one of the default phenomena. In each of these cases, the larger forecasting frequency between the one attributed to the “additional” phenomenon and the one attributed to the default

Fig. 3. The relative occurrence (%) of all forecasters (all regions) participating in the survey who said they either often (dark blue), sometimes (light blue), or never (gray) forecasted a particular phenomenon.

phenomenon is assigned to the default phenomenon. Suggestions that are too vague to be categorized as a phenomenon (e.g., “aviation” or “marine”) are also ignored (114 suggestions). Nonthunderstorm winds and visibility² received the most responses among the 14 additional phenomena (Fig. 4).

Data gaps. For each of the six default phenomena (fire, flood, hurricane, severe thunderstorms, lake-effect/lake-enhanced snow, and winter precipitation) and any additional phenomena from the previous question³ that a forecaster indicated forecasting at least “sometimes,”⁴ the forecaster was asked to rate the frequency with which data gaps (a lack of appropriate usable data) impeded his/her ability to forecast that phenomenon. Forecasters were given three options for their rating: “often,” “sometimes,” or “never.” Our analysis considers only those additional phenomena included by at least 10 forecasters. Relative occurrences of each frequency are reported as percentages based on the number of forecasters who responded to a particular phenomenon, not the total number of respondents. As such, since the goal of the survey was explicitly stated in the survey [“identify data gaps for short-term (<1 day) forecasting”], it is possible that some of the additional phenomena might have been included by forecasters because of their perceived data gaps. Therefore, it is possible that the relative occurrences associated with additional phenomena are biased to reflect more frequently occurring data gaps than those of the default phenomena.

The highest percentage of respondents to report that data gaps “often” impeded their ability to forecast one of the default phenomena is associated with winter precipitation (Fig. 5). Among the additional phenomena, waves and turbulence were viewed as often associated with data gaps by the largest percentage of responding forecasters. In an effort to rank the phenomena, each frequency is assigned a score of 0 (never), 1 (sometimes), or 2 (often) and these frequency scores are averaged across all responding forecasters for a given phenomenon, e.g.,

$$f_p = \frac{\sum_i^N s_{p,i}}{N}, \quad (1)$$

where f_p is the mean frequency score for the p th phenomenon, $s_{p,i}$ is the frequency score for the p th phenomenon from the i th forecaster out of a total of N forecasters. The mean frequency scores, f_p , are then averaged across all phenomena to get a “threshold mean frequency score” ϕ , e.g.,

$$\phi = \frac{1}{n_p} \sum_p^{n_p} f_p, \quad (2)$$

Fig. 4. The number of responses across all regions for additional phenomena broken down by frequency of forecasting: often (dark orange) and sometimes (light orange).

² Blowing dust is treated separately from visibility since blowing dust is presumed to present hazards that go beyond the reduction in visibility.

³ Similar to the filtering performed for forecasting frequency, if a respondent’s additional contribution is deemed to actually be a duplication of a default phenomenon, the maximum data gap frequency is assigned to the corresponding default phenomenon.

⁴ A forecaster was not given the option to answer the data gaps question for a phenomenon that they never forecasted.

Fig. 5. (left) The relative occurrence (%) of forecasters who felt that data gaps impeded their ability to forecast the listed phenomena often (dark green), sometimes (light green), or never (gray). (right) The mean frequency scores for each phenomenon along with threshold values (vertical discontinuous lines) for overall frequent gaps (mean frequency score of 1.27) and overall infrequent data gaps (mean frequency score of 0.94), and ± 1 standard deviation (whiskers). The check marks and $\times$ symbols to the left of the mean frequency scores indicate whether a particular gap was “overall frequent” or “overall infrequent,” respectively.

where n_p is the number of phenomena included in the average.⁵ If the mean frequency score for a particular phenomenon is larger than this threshold mean frequency score then the phenomenon is interpreted to have “frequent data gaps overall.” If the mean frequency score for a particular phenomenon is one standard deviation or more below this threshold mean frequency score, then the given phenomenon is interpreted to have “infrequent data gaps overall.”

Among the default phenomena, flood and winter precipitation represent phenomena with frequent data gaps overall and only hurricanes are associated with infrequent data gaps overall (Fig. 5). Of the additional phenomena, icing, turbulence, visibility, and waves are categorized as associated with frequent data gaps overall, while extreme temperatures falls in the infrequent category (Fig. 5).

For a different perspective on the phenomena associated with data gaps, mean frequency scores are only considered from forecasters who report that they often forecast a particular phenomenon [this alters N in Eq. (1)].⁶ Ostensibly, these forecasters would be the most knowledgeable of the presence of data gaps. Based on this filtering, results change in important ways. Of the default phenomena, on average, often-forecasters rate lake-effect/lake-enhanced snow, not winter precipitation, as most frequently possessing data gaps (based on the mean frequency scores), though winter precipitation is second (Fig. 6). Of the additional phenomena, air quality, not waves nor turbulence, has the highest mean frequency score with nonthunderstorm winds and icing second and third, respectively (Fig. 6).

Following the methodology described above for categorizing phenomena as “infrequently overall” or “frequently overall” associated with data gaps, among the often-forecasters, flood and winter precipitation once again fall in the frequent category for the default phenomena

⁵ Only phenomena with more than two responses are included in the average.

⁶ Phenomena with at least five responses by “often-forecasters” are included in this analysis.

Fig. 6. (left) The relative occurrence (%) of forecasters who forecast a particular phenomenon "often" who felt that data gaps impeded their ability to forecast the listed phenomenon often (dark green), sometimes (light green), or never (gray). (right) The mean frequency scores for each phenomenon along with threshold values (vertical discontinuous lines) for overall frequent gaps (mean frequency score of 1.27) and overall infrequent data gaps (mean frequency score of 0.89), and ± 1 standard deviation (whiskers). The check marks and $\times$ symbols to the left of the mean frequency scores indicate whether a particular gap was "overall frequent" or "overall infrequent," respectively.

as does icing for the additional phenomena (Fig. 6). New to the frequent category, compared to when all forecasters are considered, are lake-effect/lake-enhanced snow, air quality, and nonthunderstorm winds. Moreover, neither turbulence, visibility, nor waves are considered frequent suffers from data gaps by often-forecasters. Considering the infrequent overall phenomena, hurricanes no longer fall in this category among the default phenomena. Among the additional phenomena, extreme temperatures are once again categorized as having infrequent data gaps, as is visibility.

Data need prioritization. In the fourth component of the survey, forecasters were asked to prioritize specific weather-related characteristics that would need to be observed to fill data gaps. Twenty-four default characteristics derived from the focus groups and interviews referenced previously were included (Fig. 7). Forecasters were also provided the opportunity to add up to seven additional characteristics. As reflected in the default characteristics, the additions were expected to be very specific. As such, of the 334 suggestions that do not overlap default characteristics, 137 unique categories have been formed. Of these unique categories, 96 have only one response and 26 have three or more responses. These 26 categories (Fig. 8) include the majority (62%) of the 334 original suggestions and are considered in this analysis.

The default characteristic receiving the most high-priority responses is temperature profile in mixed precipitation followed by hydrometeor type in winter, radar gaps, ground conditions for flooding, and near-storm vertical wind profiles (Fig. 7). Among the additional characteristics, the most responses are associated with waves, fog/stratus, wind over open water, surface observations, precipitation distribution/accumulation/rate, and water temperature (Fig. 9).

Fig. 7. (left) The relative occurrence (%) of forecasters who rate the priority of data needs for the default characteristics as high (dark orange), medium (light orange), or low (gray). **(right)** The mean priority scores for each characteristic along with threshold values (vertical discontinuous lines) for overall high priority (mean priority score of 1.35) and overall low priority (mean priority score of 0.95), and ± 1 standard deviation (whiskers). The check marks and $\times$ symbols to the left of the mean priority scores indicate whether a particular observation is a “overall high priority” or “overall low priority,” respectively.

Among the default characteristics, those deemed to be an “overall high priority,” defined as having a mean priority score greater than the mean plus a half standard deviation (this is a somewhat more stringent condition than used for analysis of gaps), are (listed in descending mean priority)

- 1) temperature profile in mixed precipitation,
- 2) hydrometeor type in winter,
- 3) radar gaps,
- 4) near-storm vertical wind profile,
- 5) ground conditions for flooding,
- 6) snow accumulation,
- 7) wind shear of the preconvective environment, and
- 8) wildfires.

Default characteristics in the overall low priority category are thunderstorm outflow, cold air drainage, ice jams, and levee breach.

Considering those *additional* characteristics with the most responses (listed above), waves, fog/stratus, precipitation distribution/accumulation/rate, and surface observations have mean priority scores that qualify as overall high priorities (defined as having priority scores above

the mean). (This analysis deliberately avoids assigning any of the additional characteristics to the low-priority category because of the small numbers of respondents for most characteristics.)

Although no explicit attempt was made in the survey to solicit a ranking of the characteristics to enable a relative prioritization, this ranking can be inferred from the responses. For each respondent, the N characteristics receiving the highest priority are given a value of $1/N$ (this is effectively splitting a first place vote across N characteristics). The votes of all respondents are then added and divided by the total number of respondents. The result is a percentage ranking of the highest-priority characteristics (Table 3). While the top 10 characteristics are the same whether the ranking is done using mean priority score or highest priority votes, radar gaps not temperature profile in mixed precipitation is the highest ranked when using highest-priority votes (Table 3).

UAS capabilities/characteristics and operation mode. In the final component of the survey, forecasters were asked to rate the importance of a set of UAS capabilities/characteristics that emerged from the focus groups and interviews (Fig. 10) along with the importance of typical UAS operation modes (Fig. 11). They were prompted with the scenario that UAS have been integrated into the weather surveillance network to provide data for weather forecasts. Among the nine capabilities/characteristics, the ones rated as overall important (Fig. 10), following

Fig. 8. The number of forecasters who included the listed additional characteristics. Boldface characteristics correspond to those with the most responses.

Table 3. Top 10 rankings of the default weather-related characteristics based on mean priority score (center column) and highest-priority votes (right column).

Ranking	Ranked characteristics	
	Based on mean priority score	Based on highest-priority votes
1	Temperature profile in mixed precipitation	Radar gaps
2	Hydrometeor type in winter	Temperature profile in mixed precipitation
3	Radar gaps	Hydrometeor type in winter
4	Near-storm vertical wind profile	Ground conditions for flooding
5	Ground conditions for flooding	Snow accumulation
6	Snow accumulation	Near-storm vertical wind profile
7	Wind shear of the preconvective environment	Wildfires
8	Wildfires	Nocturnal thunderstorms
9	Nocturnal thunderstorms	Storm damage
10	Storm damage	Wind shear of the preconvective environment

Fig. 9. (left) The relative occurrence (%) of forecasters who rate the priority of data needs for additional characteristics as high (dark orange), medium (light orange), or low (gray). Boldface characteristics correspond to those with the most responses (Fig. 8). (right) The mean priority scores for each characteristic along with the threshold value (vertical discontinuous line) for overall high priority (mean priority score of 1.76), and ± 1 standard deviation (whiskers). The check marks to the left of the mean priority scores indicate whether a particular observation is a “overall high priority.”

the methodology adopted previously, are the following (listed in descending order of importance):

- 1) Real-time or near-real-time data
- 2) Ability to integrate UAS data with additional data gathered by other systems
- 3) UASs equipped with cameras to verify forecasts and monitor weather
- 4) Regulations for use of UAS technology to provide special allowances for weather forecasting
- 5) Ability to get UASs out and assisting with forecasts on-demand, in a short amount of time
- 6) Provision of geocoded data

Those rated as overall unimportant are “positive public opinion of UAS use for weather forecasting” and “NWS pilots control the UAS and where and when it gathers data.”

Considering the mode of UAS operations, forecasters were provided three options: targeted surveillance, fixed-site profiling, and transects between fixed sites. Definitions of the modes were not provided because the nomenclature was assumed to be relatively self-explanatory. Targeted surveillance is uniformly considered to be an overall important operation mode

Fig. 10. (left) The relative occurrence (%) of forecasters who rate the importance of specific UAS capabilities/characteristics as very important (dark orange), somewhat important (light orange), or not important (gray). (right) The mean importance scores for each capability/characteristic along with the threshold value (vertical discontinuous line) for overall important (mean importance score of 1.45) and overall not important (mean importance score of 1.24), and ± 1 standard deviation (whiskers). The check marks and $\times$ symbols to the left of the mean importance scores indicate whether a particular capability/characteristic is “overall important” or “overall unimportant,” respectively.

(Fig. 11). Both fixed-site profiling and transects between fixed sites are rated as markedly less important even though the mean importance score of the operating modes is not less than the threshold for unimportance (Fig. 11).

Although respondents were not asked to rank the three operation modes, as described in the “Data need prioritization” section, this ranking can be inferred from the responses. Applying this approach to operation modes, seven possible combinations of relative importance are possible: all modes have equal importance, two modes have higher importance than the third, or one mode has higher importance than the other two. The number of respondents who fall into one of these seven combinations can be plotted on a ternary diagram (Fig. 12) and the mean preference can be plotted (black dot on Fig. 12). The proximity of the mean preference to one of the modes illustrates its cumulative preference compared to the other two modes. Once again, a preference for targeted surveillance is apparent.

Another way these data can be examined follows the specific analysis described in the “Data need prioritization” section wherein the fraction of respondents who rank a specific option the highest are calculated. For the UAS operation modes, 23.2% of respondents rank fixed-site profiling the most important and 22.4% rank transects between fixed sites as the most important but a majority of respondents (54.4%) rank targeted surveillance the most important. On balance, it is clear that targeted surveillance is generally favored over either fixed-site profiling or transects between fixed sites.

Summary and discussion

Analysis of responses to a survey of 630 NWS forecasters reveals the existence of critical data gaps that impede their ability to forecast a range of phenomena. Analysis focused on the national (all-regions) results. While the generality of the summaries offered below are limited because of the somewhat arbitrary nature of the thresholds chosen, some of the more obvious findings are as follows. NWS forecasters view winter precipitation, icing, flood, lake-effect/

Fig. 11. (left) The relative occurrence (%) of forecasters across all regions who rate the importance of specific UAS operation modes as very important (dark orange), somewhat important (light orange), or not important (gray). **(right)** The mean importance scores for each operation mode along with the threshold value (vertical discontinuous line) for overall important (mean importance score of 1.45) and overall not important (mean importance score of 1.24), and ± 1 standard deviation (whiskers). The check marks and $\times$ symbols to the left of the mean importance scores indicate whether a particular operation mode is “overall important” or “overall unimportant,” respectively.

lake-enhanced snow, turbulence, and waves as the phenomena principally impacted by data gaps (Fig. 13). They also generally consider the most important characteristics that need to be observed to fill these gaps to be temperature profiles in mixed precipitation, radar gaps, hydrometeor type in winter, near-storm vertical wind profiles, ground conditions for flooding, snow accumulation, waves, fog/stratus, precipitation distribution/accumulation/rate, and surface observations. It is clear that forecasters view precipitation as a significant data gap: 3 of the 6 most important phenomena directly relate to precipitation and 7 of the 10 most important characteristics that need to be observed are measures of precipitation or indirectly related to precipitation-producing phenomena. Moreover, there is an obvious preference for observations related to cool-season phenomena: 3 of the 6 most important phenomena are cool-season phenomena (winter precipitation, icing, and lake-effect/lake-enhanced snow), 2 have no obvious seasonal preference (turbulence and waves), and the last could be argued to be a primarily (though not exclusively) a warm-season phenomenon (flood).

Analysis of survey responses also illustrates that some phenomena generally do not suffer from data gaps. Extreme temperature was generally rated as infrequently suffering from data gaps. This was true when all forecasters were considered and when considering only those forecasters who often forecast extreme temperature. Similarly, hurricanes were rated as infrequently suffering from data gaps. When the analysis was limited to forecasters who forecasted hurricanes often, the mean frequency scores fell in the range of values associated with neither frequent nor infrequent data gaps. Of the default characteristics offered to respondents of the survey, thunderstorm outflow, cold-air drainage, ice jams, and levee breach were

Fig. 12. Ternary diagram illustrating the overall preference of UAS operation modes (black dot). Blue dots represent the possible positions of individual forecaster preferences; dots halfway between two modes represent the situation when a forecaster prefers those two modes over the third, and the dot near the center of the diagram represents the situation when a forecaster has no preference. Numbers near each of the blue dots represent the responses for a particular preference.

nearly uniformly considered to be low-priority focuses for data collection.

A number of observational gaps identified by NWS forecasters do not require measurements above the surface, e.g., hydrometeor type in winter, ground conditions during flooding, snow accumulation, storm damage, waves, precipitation distribution/accumulation/rate, and surface observations. Each of these could be measured through remote sensing from an above-ground platform and such airborne-based observations are likely to provide superior spatial coverage compared to ground-based platforms. Nevertheless, the above-ground in situ observations which are the focus of most UAS atmospheric science applications are not required for many of the observation gaps identified by NWS forecasters.

When considering the capabilities/characteristics of UAS, the top three most important were real-time or near-real-time data, the ability to integrate UAS data with additional data gathered by other systems, and UASs equipped with cameras to verify forecasts and monitor weather. The least important characteristics related to NWS pilots controlling the UAS and where and when it gathers data, and positive public opinion of UAS use for weather forecasting. The desire for cameras on UAS and the perceived relatively small importance of positive public opinion presents a challenging intersection that would need to be addressed if UAS are to be integrated into the weather surveillance network. In light of prior research, privacy concerns of the general public are likely to be heightened for federally operated UAS equipped with cameras (Walther et al. 2019), even though UAS operations focused on atmospheric science are generally viewed more favorably than less “noble” applications (Walther et al. 2019). Moreover, historically, public acceptance of new technologies has drastically impacted technology use and implementation (Gaskell et al. 1999; Gupta et al. 2012). This suggests that receiving buy-in from the general public for weather surveillance UAS will be critically important.

Of the three operation modes offered for forecasters to consider, targeted surveillance was considered to be the most important compared to fixed site profiling or transects between fixed sites. While probably ill-advised to use this conclusion to deemphasize the potential value of fixed-site profiling, this result does seem to underscore the value forecasters place on data retrieval at those times when, and those locations where, critical data gaps exist. Moreover, the preference likely reflects forecasters embrace of the unique capabilities of mobile platforms compared to conventional platforms such as WSR-88Ds or ASOS/AWOS.

As noted previously, allowing forecasters to include additional phenomena and characteristics was necessary to ensure a more complete assessment of data gaps and observation priorities than if only default phenomena and characteristics were assessed. However, because these additional phenomena and characteristics contributed by individual forecasters were not available to all forecasters, the associated statistics may not be representative of the population of forecasters. One goal of future work is to repeat the survey with inclusion of these additional phenomena and characteristics.

	All Forecasters	Often Forecasters
Fire		
Flood	✓	✓
Hurricane	✗	
Severe thunderstorm		
Lake effect snow		✓
Winter precipitation	✓	✓
Air quality		✓
Blowing dust		
Icing	✓	✓
Temperatures	✗	✗
Turbulence	✓	
Visibility	✓	✗
Waves	✓	
Winds		✓

Fig. 13. Summary of data gaps: check marks represent frequent gaps, × symbols represent infrequent data gaps, and open cells represent neither frequent nor infrequent gaps.

A deeper examination of the preferred UAS operation mode is also warranted in future work. This could be achieved by providing forecasters with specific forecasting scenarios and example datasets (either synthetic or real) for the different operation modes (e.g., Heinselman et al. 2012). This would provide the concrete examples that are required for informed assessment of the value of the different operation modes.

There is clearly an interest within the operational and research communities to explore the potential of UAS for filling observational gaps. With contributions toward an identification of the critical observational gaps from studies like this one, focused technological development can proceed and the feasibility evaluated. UAS capabilities are rapidly evolving through federal and commercial investments. There is a natural appeal in matching current UAS capabilities with needed observation gaps to assess the current feasibility of UAS to fill these gaps. There is certainly precedential work that could be used for such an assessment. However, in general, significant advancements in UAS technology are still required before such inherently complex feasibility assessments make sense. Toward this end, we offer the following recommendations informed by the analysis presented here:

- Radar gaps are clearly a critical issue. UAS-borne gap filling radars may offer a flexible, cost-efficient solution but present a number of potentially significant engineering challenges whose solutions need to be demonstrated.
- Environmental conditions in winter precipitation are highly valued by NWS forecasters. Icing will be a problem for any airborne platform and needs to be addressed.
- Targeted surveillance is a preferred mode of operation, and its feasibility needs to be explored. Solutions to the unique technological and regulatory challenges need to be demonstrated.

Acknowledgments. This work was supported by the National Science Foundation Grant OIA-1539070. The authors are grateful to the 630 forecasters that responded to the survey. Thanks also go to John Murphy, who approved and distributed this survey. We also greatly appreciate the reviews of editor Jeff Waldstreicher and two anonymous reviewers whose evaluation of earlier versions of this manuscript improved the final product.

References

- Bocklisch, F., S. F. Bocklisch, and J. F. Krems, 2012: Sometimes, often, and always: Exploring the vague meanings of frequency expressions. *Behav. Res. Methods*, **44**, 144–157, <https://doi.org/10.3758/s13428-011-0130-8>.
- Chilson, P. B., and Coauthors, 2019: Moving towards a network of autonomous UAS atmospheric profiling stations for observations in the Earth's lower atmosphere: The 3D Mesonet concept. *Sensors*, **19**, 2720, <https://doi.org/10.3390/s19122720>.
- Cione, J. J., and Coauthors, 2020: Eye of the storm: Observing hurricanes with a small unmanned aircraft system. *Bull. Amer. Meteor. Soc.*, **101**, E186–E205, <https://doi.org/10.1175/BAMS-D-19-0169.1>.
- Dabberdt, W. F., and Coauthors, 2005: Multifunctional mesoscale observing networks. *Bull. Amer. Meteor. Soc.*, **86**, 961–982, <https://doi.org/10.1175/BAMS-86-7-961>.
- Flagg, D. D., and Coauthors, 2018: On the impact of unmanned aerial system observations on numerical weather prediction in the coastal zone. *Mon. Wea. Rev.*, **146**, 599–622, <https://doi.org/10.1175/MWR-D-17-0028.1>.
- Gaskell, G., M. W. Bauer, J. Durant, and N. C. Allum, 1999: Worlds apart? The reception of genetically modified foods in Europe and the US. *Science*, **285**, 384–387, <https://doi.org/10.1126/science.285.5426.384>.
- Gupta, N., A. R. Fischer, and L. J. Frewer, 2012: Socio-psychological determinants of public acceptance of technologies: A review. *Public Understanding Sci.*, **21**, 782–795, <https://doi.org/10.1177/0963662510392485>.
- Heinselman, P. L., D. S. LaDue, and H. Lazrus, 2012: Exploring impacts of rapid-scan radar data on NWS warning decisions. *Wea. Forecasting*, **27**, 1031–1044, <https://doi.org/10.1175/WAF-D-11-00145.1>.
- Houston, A. L., J. C. Walther, L. M. PytlíkZillig, and J. Kawamoto, 2020: Initial assessment of unmanned aircraft system characteristics required to fill data gaps for short-term forecasts: Results from focus groups and interviews. *J. Oper. Meteor.*, **8**, 111–120, <https://doi.org/10.15191/nwajom.2020.0809>.
- Jensen, A. A., and Coauthors, 2021: Assimilation of a coordinated fleet of uncrewed aircraft system observations in complex terrain: EnKF system design and preliminary assessment. *Mon. Wea. Rev.*, **149**, 1459–1480, <https://doi.org/10.1175/MWR-D-20-0359.1>.
- Koch, S. E., M. Fengler, P. B. Chilson, K. L. Elmore, B. Argrow, D. L. Andra Jr., and T. Lindley, 2018: On the use of unmanned aircraft for sampling mesoscale phenomena in the preconvective boundary layer. *J. Atmos. Oceanic Technol.*, **35**, 2265–2288, <https://doi.org/10.1175/JTECH-D-18-0101.1>.
- Leuenberger, D., A. Haebele, N. Omanovic, M. Fengler, G. Martucci, B. Calpini, and O. Fuhrer, 2020: Improving high-impact numerical weather prediction with lidar and drone observations. *Bull. Amer. Meteor. Soc.*, **101**, E1036–E1051, <https://doi.org/10.1175/BAMS-D-19-0119.1>.
- McFarquhar, G. M., and Coauthors, 2020: Current and future uses of UAS for improved forecasts/warnings and scientific studies. *Bull. Amer. Meteor. Soc.*, **101**, E1322–E1328, <https://doi.org/10.1175/BAMS-D-20-0015.1>.
- National Academies of Sciences, Engineering, and Medicine, 2018: *The Future of Atmospheric Boundary Layer Observing, Understanding, and Modeling: Proceedings of a Workshop*. National Academies Press, 70 pp., <https://doi.org/10.17226/25138>.
- NRC, 2009: *Observing Weather and Climate from the Ground Up: A Nationwide Network of Networks*. National Academies Press, 250 pp., <https://doi.org/10.17226/12540>.
- Vömel, H., and Coauthors, 2018: *NCAR/EOL Community Workshop on Unmanned Aircraft Systems for Atmospheric Research*. Boulder, CO, NCAR, <https://doi.org/10.5065/D6X9292S>.
- Walther, J., L. PytlíkZillig, C. Detweiler, and A. Houston, 2019: How people make sense of drones used for atmospheric science (and other purposes): Hopes, concerns, and recommendations. *J. Unmanned Veh. Syst.*, **7**, 219–234, <https://doi.org/10.1139/juvs-2019-0003>.