1. INTRODUCTION

Weather forecasting is a complex intellectual task due to the complexity of atmospheric phenomena and the vast amount of data and tools used to measure, model, and present these phenomena. New interactive computing tools are continually being made available to meteorologists, and existing tools are frequently upgraded. While operational meteorologists are often involved in the design of these new tools, no comprehensive task analysis of the forecasting process could be found in the literature. Understanding the overall forecasting process should be fundamental to a new system’s user interface design as well as incorporation of new tools into the comprehensive operational meteorology environment.

1.1 Methodology

This study involved observations of forecasters at two National Weather Service (NWS) offices. A series of interviews, observations, and follow-up discussions with meteorologists at two NWS offices (Greer, SC, and Blacksburg, VA) were vital to this study. We are indebted to several meteorologists at both offices who volunteered their time and expertise to the project.

Hierarchical Task Analysis (HTA) methods were used to gather information through a process of research, interviews, and data assessment. From this process, an overall task analysis was performed for the production of public area forecasts.

The task analysis revealed fundamental knowledge structures used by forecasters: pattern abstraction, trend recognition, change detection, four-dimensional data, time budgeting, forecast heuristics, parameter computation/data conversions, and forecast verification and feedback.

Additionally, fundamental processing scripts were identified including: initial data sampling, communication/coordination, data relevancy filtering, and errant data elimination. This paper presents an overview of the task analysis, knowledge structures, mental processing scripts, and implications for forecast production tools based on these observations.

1.2 Importance

Presenting information in a timely and task-oriented manner can improve the speed of data collection, quality of data analysis, and ultimately the quality of the forecast product. Many improvements to user interfaces have already been made for tools used by operational meteorologists. AWIPS (Advanced Weather Interactive Processing System) has made countless improvements in how information is displayed, entered, integrated, and processed by operational meteorologists. Additionally, a vast array of new systems are being developed and deployed with the goal of adding features and data products for forecasters to use. The questions that must be addressed are: “Can a forecaster effectively use all these tools?” or “Will these new tools actually result in better forecast products or will information overload prevent the potential improvements that new systems promise?”

To address these questions, it is important for designers of forecast tools to have a thorough understanding of the overall forecast process. Thus, we have completed an initial Task Analysis (TA) that is as “tool-independent” as possible. The question we posed was: “What fundamental tasks must be performed by an operational meteorologist regardless of the computing tools being used?” The results of this project provide: 1) an overview of forecast tasks that helps the system designer/developer be aware of the larger context into which his/her system will be used and 2) a set of “architectural components” that can provide a checklist of issues that should be
considered in almost any system design effort targeted for use by operational meteorologists.

2. BACKGROUND

To begin the Task Analysis (TA), it was important to examine current forecast methodologies and forecast workstation (e.g., AWIPS) functionality. For our purposes, forecast workstation functionality is an artifact providing evidence of tasks that must be performed by an operational meteorologist.

2.1 Forecast Funnel

The forecast funnel method initially focuses the forecasters’ attention on the global scale, moves to a synoptic scale, then the meso-scale. The fundamental concept is that the forecaster should first get the “big context” before focusing on the more immediate matters in the short-term (meso-scale or micro-scale issues). The forecast funnel is linked to a time pyramid which indicates how much time on average is spent during each phase; the greatest amount of time being inversely proportional to the size of the forecast domain, as seen in Figure 1.

"With so many new tools and data sources becoming available, it is likely that the forecaster will be unable to process all of the information at her/his disposal. As a result, we expect forecasters will need to be trained on what tools are most relevant for particular weather scenarios. Although a forecast funnel approach will still be valid, we believe that the above scientific and technological advances point to a critical need to update the aging forecast process module to make it relevant to contemporary forecast operations." Byrd (2005) Thus, while the forecast funnel is useful for understanding and teaching forecasting, it is being used in operational meteorology less and less for several reasons including: 1) Some tasks that were time-consuming and tedious ten years ago are now automated. 2) Some tasks are more accurately performed by automation. 3) Improved modeling often allows the forecaster to use computer-generated analysis when previously hand analysis would have been required.

Also, with improved communication and networking tools, it is now more reasonable for the forecast funnel process to be distributed across different forecasters located in different offices. For example, not every forecaster should be required to invest significant time doing global analysis; an area forecaster can take the analysis from a national office and focus on implications for their area. The danger of course is that important large details can be lost in communication; also, “dissenting” opinions can be lost in a distributed scenario.

2.2 AWIPS

AWIPS (Advanced Weather Interactive Processing System) has been deployed in NWS offices during the past few years. “AWIPS is used to:

- Provide computational and display functions at operational NWS sites;
- Provide open access, via NOAAPORT, to extensive NOAA data sets that are centrally collected and/or produced;
- Acquire and process data from an array of meteorological sensors (e.g., Weather Surveillance Radar-88 Doppler, Geostationary Operational Environmental Satellite, and Automated Surface Observing System) and local sources;
- Provide an interactive communications system to interconnect NWS operations sites and to broadcast data to these sites; and,
- Disseminate warnings and forecasts in a rapid, highly reliable manner.” AWIPS OneStop (2004)

2.3 Task Analysis (TA)

TA is a process used to elicit descriptions of the tasks people perform and the knowledge they possess to perform their assigned tasks. Sheperd (2001), Sutcliffe (1997). Furthermore, a task analysis provides a structure for a description of the activities performed which then makes it easier to describe how these activities form a cohesive task. This study used Hierarchical Task Analysis (HTA), which is well suited for analyzing manual tasks and Task Knowledge Structure (TKS) that was designed for intellectual tasks. Only the HTA is presented in this paper.

HTA divides the target task into subtasks, using a top-down approach. This top-down approach illustrates the hierarchical structure of the tasks and how they interrelate to each other. The first set of subtasks, called primary subtasks, is further decomposed into smaller subtasks. HTA focuses on how the subject perceives the task, how the task is preformed, and the logic behind the task. HTA is extremely useful when concerned with issues such as interface design and work organization Sheperd (2001). In a HTA, each subtask is expressed as a verb or action phrase, which has a number associated with it. This helps determine the tree structure of the activities and how they interact.

3. METHODOLOGY

With the integration of AWIPS as the NWS main forecasting tool, methods used for developing forecasts are changing. The forecast funnel is well documented but is becoming more of a teaching tool than a viable forecast methodology (as described above). In many offices, forecast products are separated into discrete, specific time and product groups. These product groups are assigned to forecasters on duty; however, in the two NWS offices that were visited, more flexible divisions of the work were observed. We have called these methodologies: the "three-way split method" and "team method".

3.1 Three-way-split method

The three-way-split method consists of splitting the forecast into three sections: short-range (usually days 1-2), mid-range (usually days 3-4) and long-range (usually days 5-7). The short-range section normally contains forecast products not related to the public forecast, such as aviation weather, fire weather, and severe weather.

The forecaster workload is divided between the three sections. The forecasters work together to create the mid-range forecast. Next, forecasters come to agreement on the midrange forecast and then work on the short-range or long-range sections independently (typically one forecaster on each section).

Forecasters use the mid-range forecast as a common point from which the two remaining forecast sections are created. The short-range forecaster will use the mid-range as an endpoint; the long-range forecaster uses the mid-range forecast as a starting point, resulting in a seamless product.

3.2 Team Method

The team method consists of splitting the forecast products (public forecast, aviation forecast, fire weather, severe weather, etc.) as evenly as possible among the forecasters working the shift (two or three meteorologists). This distribution of workload is either made by the Shift Leader or ahead of time by office management (usually the Science and Operations Officer and Meteorologist-in-Charge). Once the work is divided and distributed to the forecasters, each forecaster works somewhat independently on their assigned products. Of course, the forecasters must coordinate with each other so that their forecasts products are seamless.

4. TASK ANALYSIS RESULTS

4.1 Goals and Operations

The results of the TA are summarized in Appendix A. This chart presents the fundamental tasks that are performed in the production of a public forecast. We have attempted to describe the tasks independent of any particular system of forecast production. The chart shows a hierarchical subdivision of the major tasks into subtasks.

The chart should NOT be understood as a flowchart, even though the general flow in the process is from the left of the chart to the right. In fact, the most striking issue related to the task analysis itself is the vast differences in the order in which the forecast subtasks are done. The ordering of the tasks can be dramatically different as a result of: 1) personal preferences, 2) the nature of the weather event being forecast, and 3) time constraints. Furthermore, tasks already performed may be subsequently revisited for
reanalysis for verification or because additional data has arrived.

4.1 Knowledge Structures

Knowledge structures are mental representations of information that forecasters typically process. The sheer volume of information necessitates that some organization must be imposed on the data in order for the forecaster to remember and process the information.

4.1.1 Pattern Abstraction

An experienced forecaster needs relatively few data points to understand the weather pattern. This ability is derived from the forecaster’s knowledge of physics and the behavior of weather systems. Pattern Abstraction allows the forecaster to process and predict large complex weather patterns at global, synoptic, and meso-scales without having to actually remember all the specific data associated with the pattern.

This information structure is essentially what Donald Norman calls memory through explanation and memory for meaningful relationships. Norman (1988). “Memory through explanation” is where data is not stored explicitly “in the head”; rather, it is derived from other data through some explanatory mechanism. “Memory for meaningful relationships” is again where data is not “in the head” but is derived from meaningful relationships with data that is in a person’s memory. The number of data points required in memory depends upon the experience level of the forecaster and the complexity of the weather event.

For example, an inexperienced forecaster may have to examine and remember parameters from many sounding or model layers in order to analyze a potential snow event. However, an experienced forecaster might need only a few data points and infer the rest (even though experienced forecasters might examine other points to verify their understanding of the event).

4.1.2 Trend Recognition

Recognizing trends in the data allows a forecaster to reason backward in time to identify causes or work forward in time to identify subsequent effects. Trend recognition requires that the forecaster’s attention focus become the rate of change rather than the individual data elements. Some trend recognition is more easily accomplished automatically (such as storm cell projected movement); however, some trend recognition involves discovering trends amid considerable “data noise.” Automation in the midst of significant data noise requires human observation or at least artificial intelligence automation techniques. Sometimes the trend to be recognized is so rarely of interest that it would not be reasonable to invest the effort required to automate the process.

Examples of automated trend recognition include radar storm cell tracking and identification of source air regions. However, trends can be observed in a wide variety data sets such as: 1) how well a particular model handles a particular weather event, 2) the gradual change in the computer model forecasts in a series of model runs handling a specific event, and 3) the rate of change in temperature, pressure, or some other parameter over time.

4.1.3 Change Detection

Change detection is the recognition of a marked change in the value of some data item or set of data items. Change detection can in some situations be very easy if the change is large, if it is anticipated, or if it produces a significant sensory result. However, change detection can be very difficult in situations where the change is small, the level of “noise data” is high, false alarms are often mixed with actual changes, or where one does not expect the change. Often evidence of change is easily dismissed by humans when they have a vested interest in the continuation of the current trend; this is often the case with a forecast—evidence contrary to one’s forecast would be evidence of an error in analysis and subsequently require much work to change the forecast.

Here is just one example where change detection is important. Suppose the going forecast is for a particular sequence of events; however, upstream data or even data from early stages of the event are not consistent with the current forecast. The issues involved in the perception and correction reaction to this change include: 1) which data elements would be true indicators of a change, 2) whether the perceived change evidence of errant data, a difference in only the timing of the event, or a real change that needs to be responded to, and 3) too hasty a reaction can result in flip-flopping the forecast.
4.1.4 Four Dimensional Data

Weather exists in four dimensions, the three physical planes (longitude, latitude, and elevation) and the temporal plane. One can also think of weather occurring in $n>4$ dimensions if we considered different parameters (temperature, humidity, etc.) each as a different dimension. Current display devices are limited to two dimensions. The ongoing challenge for designers is to map data that naturally exists in 4-D onto a 2-D display. Typical mechanisms for this mapping are animations and color-coded overlays of more than one level. Also, many common meteorological parameters are actually means to collapse an entire dimension into a single data element; for example, height and thickness data can be thought of as a way to collapse elevation temperature data into a single number that can be displayed over latitudinal and longitudinal axes effectively showing 3-D data on a 2-D display.

4.1.5 Time Budgeting

Forecasters budget time so that forecast deliverables are delivered on schedule. The actual amount of time required to produce a public forecast can vary greatly depending on the level of uncertainty, the granularity of meso- or micro-scale effects resulting in a wide variety of conditions across the forecast area, and other factors. Also, the amount of time available to complete a product can vary greatly from shift to shift depending on staffing, answering phone calls (media, public, etc.), administrative work, and the amount of time required to produce other forecast products. With experience, forecasters develop a good sense of how much time is required to complete subtasks and when some subtask will require more time than usual.

4.1.6 Forecast Heuristics

Forecasters have large collections of heuristics (rules of thumb) that are used to simplify the programming process. Here are a few examples:

- The use of model guidance is filtered through a system of heuristics based on previous performance.
- Adjustments to the forecast because of the terrain or season are heuristics.
- Judgments regarding the effect of various features based on shape, position, strength, etc. come in the form of heuristics.

Heuristics play an important role in the analysis and resulting forecast. As a forecaster becomes more experienced, new heuristics are developed and existing heuristics are refined and rules regarding when to apply a particular heuristic become more sophisticated.

4.1.7 Parameter Computation; Data Conversion

Hundreds if not thousands of formulas are in the heads of experienced forecasters. Of course the less frequently a formula is used, the more likely that it will not be remembered. Many of these computations are automated in current systems; still, conversion between Fahrenheit and Celsius, millibars and inches of mercury, meters and feet and many others are made “on the fly” in the mind of a forecaster.

4.1.8 Forecast Verification and Feedback

Forecast verification and feedback requires an extremely large data set; it must include forecasts for 16-20 parameters over a large forecast area for seven days compared to actual conditions. Multiply that by the fact that forecasts are being issued four times a day (or more). The verification data archive is enormous. There is no way this data could be maintained in the mind of the forecaster. A forecaster will remember a major blown forecast all too well for years, but more important subtle trends such as averaging two degrees above actual for a temperature forecast for some particular local can continue almost imperceptibly. In addition to the sheer size of this dataset, other factors make feedback data problematic including 1) the internal resistance that most humans (not just forecasters) have to recognizing their own errors and 2) the fear that this data would be used in a way that punishes the forecaster. In spite of the challenges with access and use of feedback information, it is vital to continued improvement of forecast products. Administrators and system designers must consider carefully how to present this data so that
it indeed improves forecast quality instead of it being a demoralizing factor.

4.2 Scripts

Scripts are processes that are automatically performed by a person. Some of these scripts are formally learned, but many become a part of a person’s habits gradually and informally. Some are personal preferences; others are imposed by operational requirements or cultural constraints. For example, when people go into a typical restaurant, most know the script to follow (e.g., wait to be seated, be directed to a table by the host, get a menu, give the drink and appetizer order, etc.) without being told or even consciously attending to the process.

Scripting is a common aspect of expert behavior in virtually every workplace. We are sure that many scripts exist in forecast offices that are not described here; however, we have identified the most common ones. Also, the details of these scripts often vary from forecaster to forecaster and from office to office; the descriptions here give only a general characterization.

4.2.1 Data Sampling

Forecasters generally have data sampling scripts that give them an overview of the current state of issues related to the forecast. Data sampling scripts can be in the context of current conditions (micro-, meso-, or synoptic scales), a survey of model guidance, etc. Forecasters have preferred sources and preferred sequences from which they gain the overall perspective of the issue at hand. These scripts vary greatly between forecasters and the type of event under consideration.

4.2.2 Communication/Coordination

Communications between forecasters is vital to producing accurate forecast that have a high level of continuity between regions and issuances. We identified the following communication channels:

- Intra-office communication as the forecast is being created
- Inter-office (usually surrounding offices) as the forecast is being created
- With national offices
- The shift-change briefing.

Established protocols determine the timing and medium for this communication; however, also it is also important to understand informal rules and the effects of using various means of communication (instant messaging, email, phone, etc.) on the communication process. For example, one forecaster explained that, when he has a strong opinion about the evolution of weather patterns in the forecast, he tries to get his forecast posted first resulting in forecasters in surrounding offices following his lead.

AWIPS highlights parameters in adjacent (or nearly adjacent cells) that are out of tolerance. This is another form of communication between forecasters from different offices; it is also a signal that further communication is needed.

Communication processes may be the most important area for future study in order to understand forecast processes and improve forecast products.

4.2.3 Data Relevancy Filtering

Associated with most weather events, forecasters have processes by which they determine what data is relevant. This process is vital to the forecast process because of the volume of data (models, satellite, radar, soundings, etc.) available at any given point in time. These filters include decisions regarding what parameters to study, what levels of the atmosphere will be important, source regions, etc.

Evidence of these filters comes in the form of event checklists, such as the “A Comprehensive Winter Forecast Checklist” Gordon (no date). However, with use, these checklists become a part of the forecaster’s mental tools.

4.2.4 Errant Data Elimination

Frequently, a data set is corrupted because of poor initialization, equipment failure, or communications failure. An important process in a variety of settings is identifying and correcting for errant data. For example, to identify a bad model run the forecaster must compare the initialization to observations. As a result, the forecaster might eliminate the model from consideration altogether or make mental adjustments as to what the guidance would look like if the initialization had been correct. Another example is an errant satellite or radar image; in this case the forecaster might simply take the errant image out of the loop being studied.
5. IMPLICATIONS FOR DESIGN

In general, there are two important implications of this task analysis for forecast tool user interface designers. First, the designer of a new tool should consider where the tool fits into the overall process of operational meteorology. Consideration of the overall process can guide decisions as to how and when access to the tool is made to the forecaster, the level of detail at which data should be presented, and how the data is presented. The forecaster is already overwhelmed at times by the volume of data and tools; simply adding more tools and data will not automatically make the forecast process more efficient or improve the quality of products.

Second, while the above lists of knowledge structures and scripts are by no means exhaustive, the important issue is that tool user interface designers should catalogue and consider these items. As the lists become more comprehensive, they will provide a collection of “architectural components” for tools targeted to operational meteorologists. Because they are generic components, it is likely that several will be relevant in any design context. The user interface designer should consider each component in the list for relevance to the tool being development.

The following is just an initial list of implications for the design of computing tools targeting operational meteorology. It is obvious that many of these issues have been addressed in the development (and continued enhancements) of AWIPS.

5.1 Customizability

The resounding theme from operational meteorologists is “one size does NOT fit all.” While we have identified the top-level components of the forecast process, the actual order in which tasks are accomplished varies widely based on: geographical location, the type and severity of the weather event being forecasts, office culture, forecast methodology, personal preferences, time pressures, etc. Any user interface design decision that limits the ordering or choices available to forecasters must be made with great care.

Often the best a user interface designer can do is to make the interface customizable. Customization can take the form of interface preferences associated with a user login. Parameters that could be customized are: 1) menu names, options, and ordering, 2) the ability to record and save actions into macros that can be catalogued and replayed by the forecaster, 3) startup options, 4) bookmarks, and 5) how to present interface options (color, formats, units of measure, etc.). Having the ability to construct macros is of particular importance since it provides direct support for many of the scripting issues above.

5.2 Hyperlink Anticipation

At virtually any point in the forecast process, a robust system may be able to predict or even suggest what the forecaster will want to view next. For example, if the forecaster was looking at the 48-hour gfs 500 mb height chart, the system could make shortcut hyperlinks available the following charts: times adjacent to 48 hours, higher or lower resolution charts, the same chart from other available models, height data below 500 mb, etc. The decision about what shortcuts should be available could be made by a database of common “what’s next” lists or could done by live analysis of forecaster behavior (analogous to Amazon’s “many customers who have purchased this book have also purchased…”).

5.3 4-D Animation

A common means of addressing the issue of displaying of 4-D data on 2-D displays is by animation. A rich set of animation tools and views should be available. As much as possible, the forecaster should be able to designate the x- and y-axes, the data to be displayed on the graph, and then allow the viewer to move along another axis by means of animation.

5.4 Drawing Overlays

One mechanism that could be used to assist with Trend Recognition would be to have a drawing window overlaying other windows. For example, a drawing window overlaying an animation window would allow the forecaster to mark points in the animation in order to more easily identify trends in the data.

5.5 Simplify Controls

Designers sometimes add functionality that is neither needed nor used. The needs of the user should dictate the complexity of the controls. Elimination of unused or unneeded controls would make the interface easier to use and minimize the errors in use.
For example, one commonly used animation tool, has controls for lengthening or shortening the amount of time the animation “dwells” on the first and last frame in the set of graphics. We have never seen those controls used. Let’s assume that a control is used once in 100,000 uses of the tool. The designer should consider if that level of use justifies the screen space used and the added complexity of the visible interface—or better yet, make the tool’s controls customizable.

5.6 Large, Flexible Windowing

It is clear that operational meteorologists need lots of screen “real estate” because of the vast amount of data to be accessed and processed. Forecasters need multiple monitors and even multiple screens mapped to a given monitor or window. Furthermore, forecasters need maximum flexibility in organization their screen workspace.

5.7 Data Sampling Templates

Similar to macros, data sampling templates would provide the forecaster with the ability to specify a series of data items to be presented for analysis without having to individually select each item for viewing.

5.8 Automated Checklists

Forecast support systems could provide automated checklists created by the individual forecaster, by the local office, or built into the system itself. Items in the checklist could be hyperlinked to the data of interest making it easier for the forecaster to systematically work through a procedure for forecasting various types of events. Not only would automated checklists make data analysis more efficient; it would also improve chances that all important data items have been analyzed and could also result in more continuity within a particular office for forecasting a particular event.

Another form of checklist would be an “operational checklist” that reinforces the sequence of steps that must be performed to complete some task. The operational checklist should include hyperlinks to relevant sections of the production system at each step.

Automated checklists are particularly helpful in situations where the pace of operation is unusual (abnormally rushed or slow). They also are extremely useful in situations where the person is frequently interrupted from their primary task.

5.9 Progress Meters

To assist the forecaster with scheduling and meeting the product deadline, a forecast production system could display progress meters. Progress meters could take the form of a progress bar that displays the percentage of the forecast progress that has been completed based on this forecaster’s previous time to complete forecast tasks or on the basis of overall measures of the tasks remaining and the amount of time typically used by all forecasters. Color-coding could be used to highlight normal progress, being slightly behind schedule, or significantly behind schedule. The progress meter should not be obtrusive but keep the forecast informed as to where they are in the process compared to normal.

6. CONCLUSIONS

This is an initial study; however, it demonstrates the importance of understanding the fundamental tasks in creating a public forecast. There is still much work yet to be done to understand in detail the processes that go into creating a forecast—refinement of the task analysis, understanding the social context, understanding communication processes (especially now that this context is distributed across many offices) and identifying additional knowledge structures and scripts.

This study also emphasizes the need, when introducing new technology, to consider how that technology should be presented in the context of the entire system (including human and machine processing). The potential improvements from the new technology could go unrealized because of how the new system interacts with other aspects of the forecaster’s task.

Much attention has been paid in recent years on the creation of new computing tools to model and describe weather phenomena and “how” the forecast is entered into the forecast system (now graphically entering nearly twenty parameters as opposed to creating a text product directly). These new systems have without a doubt resulted in higher quality forecasts. The next significant improvement in forecast quality may come from better understanding how people (forecasters) most efficiently and productively fit into the system, molding systems to fit the forecaster’s fundamental tasks and innate abilities rather than training the forecaster to fit the system’s capabilities and limitations.
7. REFERENCES

Richard A. Anthes, et. al., 1997: Assessment of AWIPS after operational testing of the first system build. www.nwas.org/awips.html


Patrice C. Kucera, Scott P. Longmore, and William F. Roberts, 2000: Product Usage Patterns at the AWIPS Build 4.2 OT&E Sites. 16th IIPS Conference, American Meteorological Society.


