U.S. Department of Transportation
Research and Innovative Technology Administration

Baselining Current Road Weather Information

Results of the 2010 Quality and Importance Survey

Final Report

FHWA Publication No.: FHWA-JPO-11-018

January 2011

Notice

The US Department of Transportation provides high-quality information to serve Government, industry, and the public in a manner that promotes public understanding. Standards and policies are used to ensure and maximize the quality, objectivity, utility, and integrity of its information. The USDOT periodically reviews quality issues and adjusts its programs and processes to ensure continuous quality improvement.

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-JPO-11-018		

4. Title and Subtitle

Baselining Current Road Weather Information: Results of the 2010 Quality and Importance Survey Final Report

5. Report Date
January 31, 2011

6. Performing Organization Code

7. Author(s)
Robert Hart and Leon Osborne (Meridian Environmental Technology, Inc.)

8. Performing Organization Report No.

9. Performing Organization Name and Address
Booz Allen Hamilton
8283 Greensboro Drive
McLean, VA 22102

Meridian Environmental Technology, Inc.
4324 University Avenue
Grand Forks, ND 58203

Iteris, Inc.
107 Carpenter Dr., Suite 230
Sterling, VA 20164

10. Work Unit No. (TRAIS)

11. Contract or Grant No.
DTFH61-06-D-00006; Task Order 2

12. Sponsoring Agency Name and Address
U.S. Department of Transportation
Research and Innovative Technology Administration
1200 New Jersey Avenue
Washington, DC 20590

13. Type of Report and Period Covered
Final Report, 8/25/09-1/31/11

14. Sponsoring Agency Code
HOIT

15. Supplementary Notes
Dr. Roemer Alfelor (COTM)

16. Abstract
This final report contains research findings on the characterization of the quality and value of road weather information resources used by members of the surface transportation community in their decision-making process. This report focuses on the results from the second survey in a series of surveys designed to establish a baseline metric on the quality of road weather information from the perspective of department of transportation (DOT) personnel who actively use the information to support their operational decisions.

The assessment of quality for the 2010 survey utilized the online survey approach established during the 2008 phase of the baselining study. The 2010 survey was modified slightly from the previous version to address some inconsistencies found during the first survey. Survey questions sought input on product components, parameters within specific products routinely used by DOT decision makers. The report details the organization of the survey responses into a quality attribute and importance categories and then analyzes the results of the survey. It addresses variable user responses within and amongst quality attribute classes that illustrate the strengths and weaknesses of specific product components. The product and element categories addressed in the 2008 survey were not part of the question set in 2010; however, various product components were subsets of both the product and element categories. The report provides a logical map to illustrate how the product components responses were composed into responses representative of the product and element results from 2008. The resulting derived product and element measures permitted a mechanism to compare survey results from 2008 to 2010.

The report presents the results of the 2010 survey. It augments the quality markers used for comparison in 2008 and points out road weather parameters that could benefit from improvement. Although slightly different sampling techniques were used between 2008 and 2010 the results provided similar quality metrics. Some variability existed between the two sets of results; however, the composite of all results indicated either no change or a slight increase in quality occurred in the two years. The modification of the survey format and a limited sample size created statistical uncertainties that required consideration. These unknowns were addressed and recommendations were offered to continue to improve the road weather monitoring program and stabilize the metrics.

17. Key Word	18. Distribution Statement
Road weather information; baselining road weather information; quality attribute matrix	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
		113	

Table of Contents

Table of Figures

Index of Tables

Executive Summary

The Federal Highway Administration (FHWA) Baselining Current Road Weather Information project, initiated in 2007, is an ongoing attempt to establish a level-of-quality metric for road weather information used by state departments of transportation (DOT). The first survey and analysis were completed in 2008. The most recent survey and its associated analysis of the results conducted in 2010 are the second in a series of routine assessments of current road weather information quality and the related trend in the quality. Both surveys acquired input from DOT participants across a broad spectrum of management strategies for Advisory, Control, and Treatment responsibilities.

DOT personnel whose operational decisions are impacted by weather or weather-related road conditions typically look to a number of resources to acquire road weather information. These resources usually include *Products,* which may be web sites, bulletins delivered by facsimile, or messages received via phone or pagers. Unique products exist to deliver information for specific time frames (e.g., observations, histories, forecasts) and spatial presentation formats (data for a single site, lists of multiple sites, or maps containing data for sites over a region). Each *Product* contains data for a number of road weather *Elements*, the core weather, pavement, or advisory parameters DOT personnel must monitor or integrate into their operational decisions. Thus, there exist many diverse *Products* containing road weather information to assist DOT personnel in making operational decisions. Examples of *Products* typically used by DOTs include National Weather Service (NWS) current observations, Environmental Sensor Station (ESS) historical information, pavement specific weather forecasts, road condition reports, and severe weather watches and warnings. In preparation for the 2008 survey to assess the quality of road weather information, *Products* and *Elements* were chosen as the instruments for the evaluation of road weather quality.

The analysis of the 2008 results indicated DOT users actually determine the quality of road weather data based upon the components within specific *Products*. Therefore, the 2010 survey questions were modified to request input on *Product Components*, that is, the specific *Elements* within a *Product*. This focused the survey on the resources DOT personnel use to support their operational decisions. Since *Product Components* were also directly related to both *Products* and *Elements*, it was possible to use the answers from the *Product Components* to derive quality score estimates for *Products* and *Elements*. Consequently, the 2010 survey evaluated the quality of *Product Components* and used this data to generate derived results for *Products* and *Elements,* thereby retaining continuity with the results from the 2008 survey.

The project team contacted DOT representatives in 45 states to find individuals willing to participate in the 2010 survey. Surveys were eventually sent to 92 DOT representatives. Forty-five (45) individuals started the survey, and 37 completed the survey, providing the responses used to analyze the 2010 results. Of the 37 completed surveys, 15 were submitted by participants with primary Advisory responsibilities, 5 with primary Control responsibilities, and 17 with primary Treatment responsibilities.

Quality was determined using six attributes that measured different facets of quality:

1. Accuracy/Precision;
2. The Completeness of the information;
3. Relevance to the user's needs;
4. The Currency/Latency of the information (relative to when they are received by the user);
5. Timeliness of the information and Reliable delivery of the required information; and,
6. Ease of Use of the information.

In addition, users were asked to indicate the Importance of each Product Component.

The 2010 survey was a web-based electronic survey, containing a set of six demographic questions followed by a series of multiple-choice questions regarding the quality and importance of the road weather *Product Components*. Survey participants had the opportunity to rate the quality of any of the 92 different *Product Components* they used operationally. Since the DOT participants receive the *Product Components* as part of distinct *Products*, the *Product Components* were organized into question sets organized around the 14 *Products*. *Products* had varying numbers of *Product Components* ranging in number from 1 to 22. At the beginning of each *Product* section, survey participants were provided with a description of the *Product* and a list of the *Product's* components. If it was not a weather support resource normally used in the participant's decision-making process, the participant could completely skip the questions related to that *Product*.

There were seven multiple-choice questions for each *Product Component*, one for each of the six attributes and one for importance. The answer set included quality or importance score options of Very High, High, Medium, Low, Very Low, and Not Applicable. To facilitate answer entry, each *Product Component* for a particular quality attribute or for importance was listed on the left side of the page and the multiple-choice options were placed as columns. Thus, survey participants could select their quality or importance rating by selecting one of the radio buttons for each *Product Component* in the list for that *Product*.

The survey responses were transformed to a five-level scale of Likert scores ranging from 5 for Very High down to 1 for Very Low. These Likert scores and the results from the demographic section of the survey were transferred to a MySQL database. The data were extracted from the relational database and organized into data sets necessary to evaluate quality for different demographic samples. The primary data sets included all responses and three sub-sets for Advisory, Control, and Treatment management strategy users. The Likert scores for each of the data sets were then evaluated using standard variance statistics.

The primary quality metric was the average response for each of the *Product Components*. These averages were computed for each of the six attributes and importance. In addition, the responses for all six attributes were consolidated into an overall quality attribute composite average. Most of the distribution of quality scores comparison was performed on the composite averages. The mean and median composite averages of the quality attribute scores for all *Product Components* combined were 3.94 and 3.95 with a range of averages from 2.97 to 4.36. The average on the survey rating scale was just under High, with the range from Medium to midway between High and Very High. The distribution of scores was fairly compact (standard deviation of 0.90), but tended to skew slightly towards the lower scores. The average scores for the Completeness, Relevance, Timeliness/Reliability, and Ease of Use attributes were all near 4.0. The averages for the Accuracy/Precision and Currency/Latency attributes were lower (3.7 to 3.8). Importance returned a higher average (4.25) with a range of scores from 3.37 to 5. The implication from these results is that road weather information is viewed as having a high to very high level of importance to DOT users and exhibiting a reasonably high level of quality. However, accuracy and expedient delivery of data could be improved.

The *Product Components* were ranked based on their means and color-coded by quartiles. This permitted visual inspection of the distribution of the means, providing an easier assessment of patterns of the means in the overall distribution of scores for each attribute and importance. An analysis of the rankings and their patterns provided these findings:

- Weather Summaries suffered because the key components of interest were considered lower in accuracy and not as current as desired by the DOT personnel.
- DOT personnel found the NWS History information of less value than the information in other Products across all quality attributes.
- ESS Observations have a lower level of importance to the DOT respondents than many of the other resources in this Baselining Study and are viewed as having marginal quality to the users.
    - Data accuracy, timeliness, and reliability need further attention
    - These limitations impact usage of data by DOT users, NWS, weather service providers, media, and the *Clarus* program
- ESS Histories were considered one of the least important tools and they generally received quality rankings that were below the median values.
- Map displays had a middle of the road *Product Components* both in importance and quality attribute scores.
- Pavement Weather Forecasts are the highest priority road weather information tool and DOT participants find the quality of the forecast information well above average except for pavement condition-related *Product Components*.
- Road Weather Alerts are important to DOT users, but they are disappointed with the limited content of the information and its availability on a timely basis.
- Watches and Warnings are an important decision support tool that rates second to Pavement Forecasts in its value to the DOT participants in this study.
- MDSS Treatment Recommendations are viewed as a relevant and reasonably important tool but are thus far not perceived as one of the more accurate, timely, reliable, and easy to use resources.
- Road Condition Reports are considered of average importance by DOT users; these users see a lot of room for improvement in this resource. Flood Warnings are of lesser value to DOT users and the quality of the information provided is lower than with most other road weather information resources.
- Flood Warnings are of lesser value to DOT users and the quality of the information provided is lower than with most other road weather information resources.

- Camera imagery is an important resource for DOT operations and survey participants suggest that this tool could be even more effective with more or better selection of camera views.
- Radar information scores indicate users see the fundamental radar images as easy to use, timely, and reasonably accurate. Future radar and storm track derived services need to improve to gain an equivalent acceptance of the observed radar products.

The *Product Component* averages from the 2010 survey were also converted to *Product* and *Element* scores for each of the three management strategy classifications and the variance statistics were computed. The values from 2010 were then compared to equivalent values for 2008 to assess trends in the quality. The *Element* scores in 2010 showed a slight increase over the scores from 2008, whereas the *Product* scores yielded a slight decrease. The comparison was hampered by a limited number of comparisons caused by a limited number of responses. Thus, the differences cannot be shown statistically as significant changes in quality over the two-year period.

An analysis of the results suggests there are a number of potential sources for uncertainty in the results. The survey is an assessment of subjective quality of road weather information; therefore, it is influenced by a number of human factors associated with the survey process. A second significant source of uncertainty resulted from the modification of the survey from its 2008 format to a new format in 2010. A related concern regarding the survey instrument garnered from survey participants' comments is that the survey was too long.

Evaluation of the potential sources of uncertainty and concerns about survey length led to the following recommendations for ensuing versions of the survey:

1. The sample size must be increased
2. The survey needs to be made clearer so respondents understand exactly what they are evaluating
3. The survey needs to be shortened or divided into a series of shorter surveys that can be performed in stages over a period of months

1.0 Introduction

The Federal Highway Administration (FHWA) project "Baselining Current Road Weather Information" initiated in 2007 continued its evaluation of the quality of road weather information through the implementation of the 2010 survey. This 2010 survey builds on work completed in the initial *Baselining Current Road Weather Information Study*[1] (referred to in this document as the 2008 survey). Booz Allen Hamilton (BAH) and its subcontractors Meridian Environmental Technology, Inc., Iteris, Inc., and the Surface Transportation Weather Research Center at the University of North Dakota performed the research. The principal objective of these studies was to evaluate the quality and importance of existing road weather information used by state departments of transportation (DOT). The determination of appropriate road weather information quality considered the following seven fundamental data and measurement requirements:

1. Data needs to be representative of the current road weather state of practice;
2. Data needs to have minimal human bias;
3. Data needs to be appropriate for a long-term longitudinal study of quality characteristics;
4. Methods need to be open and documented;
5. Methods need to be repeatable;
6. Methods need to provide objective statistical measures; and,
7. Methods need to be accepted by the practitioners and data users.

Two dominant themes in the design of an assessment tool to assess road weather information quality was the stability of the assessment system over a series of surveys (longitudinal study), and the minimization of human biasing factors. These objectives laid the foundation for the 2008 survey design, and the lessons learned from that study were instrumental in modifying the design for the 2010 survey to more effectively address requirements #2 and #7.

As with the 2008 evaluation, the 2010 quality assessment tool utilized a web-based electronic survey. The 2010 survey was performed in June and July of 2010. This report contains the results of the survey. It describes:

- The modified procedure used to assess road weather information quality;
- The structure of the survey instrument;
- The results from the 2010 survey; and,
- A comparison of the results from the 2008 and 2010 surveys.

---

[1] Hart, R.D., L.F. Osborne, and S.M. Conger, 2009, Baselining Current Road Weather Information: Final Report, FHWA-JPO-09-055 retrieved 25 November 2010 at http://ntl.bts.gov/lib/31000/31000/31065/14486_files/14486.pdf

2.0 Baselining Study Design

2.1 DESIGN OF THE QUALITY ASSESSMENT METHODOLOGY

The 2008 survey evaluated the quality of road weather information based on road weather *Products* and *Elements*. *Products* and *Elements* are formal names given to unique classifications of road weather information during the 2008 survey. In this report, all references to these unique classifications will be italicized with the first letter capitalized to differentiate the reference from the more general use of the terms products and elements. *Products* were defined as packages of road weather information that service providers have developed to organize and transfer road weather data to the end user (i.e., transportation agency personnel) to meet specific end user decision-support requirements. Different *Products* have unique characteristics, including:

1. One or more road weather parameters aggregated within each *Product*;
2. A source of origin (e.g., transportation agency, National Weather Service, road weather service provider);
3. A specific time frame of the data (historical, current, forecast); and,
4. A spatial representation (single site, composite list, map display).

Standard *Product* packages have evolved in the communication of data from the various sources to the transportation users for each combination of these four defining criteria. These standard *Products* and their defining characteristics are listed in Table 1. Often, the same road weather information parameters occur in multiple *Products*, but each *Product's* unique character is defined by the specific combination of source agency, their specific time frame, and whether they are for a single location or a group of locations. The components within each of the *Products* are illustrated in Table 2. Transportation users know which packages contain the road weather parameters they need for making specific decisions.

Elements in the 2008 survey were the road weather parameters commonly used by transportation agency decision makers (see Table 3 for a list of the *Elements* defined for the 2008 survey). The *Elements* were general classifications of road weather parameters; they did not differentiate the specific characteristic of a given road weather parameter. For example, the *Element* pavement temperature might exist in each of the following formats:

- An observed value from a single Environmental Sensor Station (ESS);
- A composite regional display of observed pavement temperature values;
- A listing of past observed values of pavement temperature;
- A forecast of probable pavement temperatures for the next several hours.

Transportation representatives who participated in the 2008 survey indicated they had difficulty placing a single set of quality ratings on *Elements*, because their assessment of a given *Element* changes depending upon their decision support requirement. Having a form of the *Element* in the correct time frame was their primary concern. Survey participants indicated that historical, current, and forecast data pertaining to the same road weather parameter merited different quality and importance ratings depending on the specific decision requirement. Therefore, survey participants indicated they were uncertain whether they should assign average quality and importance ratings or rate the *Elements* based upon the form that yielded them the most benefit.

Table 1. *Products* used in the 2008 and 2010 surveys and a designation of the source, spatial format, and time frame that defines the *Product*.

PRODUCT		SPATIAL FORMAT		TIME FRAME		
	SOURCE	SINGLE	MAP	HISTORY	CURRENT	FORECAST
Weather Summary	NWS	◆			◆	
Weather History	NWS	◆		◆		
ESS Current Conditions	DOT	◆			◆	
ESS History	DOT	◆		◆		
Regional Weather Map	STWSP		◆		◆	
Regional Forecast (Zone Forecast)	NWS	◆				◆
Pavement Forecast (511 Forecast)	STWSP	◆				◆
Road Weather Alert	STWSP	◆				◆
Watches and Warnings	NWS		◆			◆
MDSS	STWSP	◆				◆
Road Condition Report	DOT	◆			◆	
Flood Warning	NWS	◆			◆	
Camera Images	DOT	◆			◆	
Radar	NWS & WSP		◆		◆	

After the 2008 survey, the BAH team recognized that user uncertainty regarding the definition of an *Element* was potentially biasing the survey results and making it difficult for users to easily complete the survey. Therefore, the 2010 survey was modified to focus the evaluation of road weather quality on the specific road weather components within each of the *Products*. These components represent the unique formats of the *Elements*, each having a distinct source or origination, a distinct spatial format, and a defined time frame. The components were named *Product Components* to clearly designate they represent the components defined as part of the *Product* in the Quality Assurance Matrix in the 2008 Baselining Report. *Product Components* represent the units of road weather information that transportation agencies use to support their decisions. For example, a maintenance supervisor determining how to treat pavements during a winter event and needing to know the most probable pavement temperature during the event will request the *Product* containing the forecasted pavement temperature for the desired route or local area. This forecast *Product* would be different from the ones containing the past or current pavement temperature information. Thus, to support this particular decision, the forecasted pavement temperature would receive a higher quality and importance rating than the observed pavement temperature information.

Thus, if maintenance users consistently used forecasted pavement temperatures more often than observed pavement temperatures, the survey participants with winter maintenance responsibilities would tend to rate the quality of the Pavement Temperature *Element* based upon their evaluation of the quality of pavement temperature forecasts. Therefore, the use of *Product Components* in the 2010 survey allowed survey participants to rate the quality of the individual components of a given *Element* independently.

Table 2. Components in each *Product*

PRODUCT	COMPONENTS
Weather Summary	Air temperature
	Dew point
	Relative humidity
	Wind direction
	Wind speed
	Weather
	Precipitation type
	Precipitation amount
	Snow amount
Weather History	Air temperature
	Dew point
	Relative humidity
	Wind direction
	Wind speed
	Precipitation amount
	Snow amount
ESS Current Conditions	Air temperature
	Dew point
	Relative humidity
	Wind direction
	Wind speed
	Wind gust
	Precipitation type
	Pavement temperature
	Pavement condition
	Chemical concentration
	Freeze point temperature
ESS Histories	Air temperature
	Dew point
	Relative humidity
	Wind direction
	Wind speed
	Precipitation type
	Precipitation start time
	Precipitation end time
	Pavement temperature
	Pavement condition
	Chemical concentration
	Freeze point temperature
Regional Map	Air temperature
	Dew point
	Relative humidity
	Wind direction
	Wind speed
	Precipitation type
	Pavement temperature
	Pavement condition
	Chemical concentration
	Freeze point temperature
Zone Forecast	Maximum temperature
	Minimum temperature
	Wind direction
	Wind speed
	Weather
	Probability of precip
Pavement Weather Forecast	Air temperature
	Dew point
	Relative humidity
	Wind direction
	Wind speed
	Wind gust
	Weather
	Visibility
	Cloud cover
	Precipitation type
	Precipitation start time
	Precipitation end time
	Probability of precip
	Probability of precip type
	Precipitation rate
	Precipitation amount
	Snow rate
	Snow amount
	Pavement temperature
	Pavement condition
	Chemical concentration
	Freeze point temperature
Road Weather Alerts	Alert parameters
	Road closures
Watches and Warnings	Severe weather watches
	Severe weather warnings
	Special weather statements
MDSS	Treatment recommendation
Road Report	Road condition
Flood Warning	Current flood stage
	Forecasted flood stage
Camera Images	View of road
	View of weather
Radar	Radar loop
	Radar loop with precipitation
	type coloring
	Future radar
	Storm tracks

Table 3. *Elements* evaluated for quality in the 2008 and 2010 surveys

ELEMENT	
Air temperature	Visibility
Dew point temperature	Pavement temperature
Relative humidity	Pavement condition
Wind direction	Chemical concentration
Wind speed	Freeze point temperature
Wind gust	Frost probability
Precipitation type	Treatment recommendation
Precipitation rate	Road closure
Precipitation accumulation	Severe weather advisory
Snow rate	Wind advisory
Snow accumulation	Winter weather advisory
Weather type	Dense fog advisory
Precipitation start time	Flood advisory
Precipitation end time	Flood stage
Precipitation probability	Camera – road conditions
Probability of precipitation type	Camera – weather conditions
Maximum temperature	Camera – traffic
Minimum temperature	Radar images
Cloud cover	

The 2010 survey did not solicit input on *Elements* specifically; however, all of the *Product Components* in the 2010 survey that represented one of the several forms of a given *Element* could be aggregated together to form a derived version of the 2008 *Element*. Therefore, the set of questions on the quality of the attributes of a single *Element* in the 2008 survey should be equivalent to the average of the quality ratings from all of the *Product Components* containing the various forms of that *Element*. To perform an assessment of the quality of a derived set of quality attribute scores for each *Element*, the *Elements* were associated with their source *Product Components*. These relationships are shown in Table 4.

The primary objective of the 2010 survey was the assessment of quality and importance of each *Product Component.* The collection of quality ratings was accomplished by asking survey respondents to rank the quality of the *Product Components* based upon the six quality attributes:

1. Accuracy/precision;
2. The completeness of the information;
3. Relevance to the user's needs
4. The currency/latency of the information (relative to when they are received by the user);
5. Timeliness of the information and reliable delivery of the required information; and,
6. Ease of use of the information to be accessed and applied to the required situation, including the visual effectiveness of the data presentation.

Table 4. *Product Components* from 2010 survey associated with 2008 survey *Elements*

ELEMENT	PRODUCT COMPONENT
Air Temperature	Weather Summary - Air Temperature
	Weather History - Air Temperature
	ESS Observations - Air Temperature
	ESS Histories - Air Temperature
	Map - Air Temperature
	Pavement Forecast - Air Temperature
Dew Point Temperature	Weather Summary - Dew Point
	Weather History - Dew Point
	ESS Observations - Dew Point
	ESS Histories - Dew Point
	Map - Dew Point
	Pavement Forecast - Dew Point
Relative Humidity	Weather Summary - Relative Humidity
	Weather History - Relative Humidity
	ESS Observations - Relative Humidity
	ESS Histories - Relative Humidity
	Map - Relative Humidity
	Pavement Forecast - Relative Humidity
Wind Direction	Weather Summary - Wind Direction
	Weather History - Wind Direction
	ESS Observations - Wind Direction
	ESS Histories - Wind Direction
	Map - Wind Direction
	Zone Forecast - Wind Direction
	Pavement Forecast - Wind Direction
Wind Speed	Weather Summary - Wind Speed
	Weather History - Wind Speed
	ESS Observations - Wind Speed
	ESS Histories - Wind Speed
	Map - Wind Speed
	Zone Forecast - Wind Speed
	Pavement Forecast - Wind Speed
Wind Gust	ESS Observations - Wind Gust
	Pavement Forecast - Wind Gust
Precipitation Type	Weather Summary - Precip Type
	ESS Observations - Precip Type
	ESS Histories - Precip Type
	Map - Precip Type
	Pavement Forecast - Precip Type
Precipitation Rate	Pavement Forecast - Precipitation Rate
Precipitation Accumulation	Weather Summary - Precip Amount
	Weather History - Precip Amount
	Pavement Forecast - Precip Amount
Snow Rate	Pavement Forecast - Snow Rate
Snow Accumulation	Weather Summary - Snow Amount
	Weather History - Snow Amount
	Pavement Forecast - Snow Amount
Weather Type	Weather Summary - Weather
	Zone Forecast - Weather
	Pavement Forecast - Weather
Precipitation Start Time	ESS Histories - Precip Start
	Pavement Forecast - Precip Start
Precipitation End Time	ESS Histories - Precip End
	Pavement Forecast - Precip End
	Zone Forecast - Probability of Precipitation

ELEMENT	PRODUCT COMPONENT
Precipitation Probability	Zone Forecast - Probability of Precipitation
	Pavement Forecast - Probability of Precipitation
Precipitation Type Probability	Pavement Forecast - Probability of Precipitation Type
Maximum Temperature	Zone Forecast - Maximum Temperature
Minimum Temperature	Zone Forecast - Minimum Temperature
Cloud Cover	Pavement Forecast - Cloud Cover
Visibility	Pavement Forecast - Visibility
Pavement Temperature	ESS Observations - Pavement Temperature
	ESS Histories - Pavement Temperature
	Map - Pavement Temperature
	Pavement Forecast - Pavement Temperature
Pavement Condition	ESS Observations - Pavement Condition
	ESS Histories - Pavement Condition
	Map - Pavement Condition
	Pavement Forecast - Pavement Condition
Chemical Concentration	ESS Observations - Chemical Concentration
	ESS Histories - Chemical Concentration
	Map - Chemical Concentration
Freeze Point Temperature	ESS Observations - Freeze Point Temperature
	ESS Histories - Freeze Point Temperature
	Map - Freeze Point Temperature
Frost Probability	Pavement Forecast - Percent Probability of Frost
Treatment Recommendation	MDSS - Treatment Recommendation
Road Closure	Road Weather Alerts - Road Closure
Severe Weather Advisory	Watches and Warnings - Severe Weather Watches
	Watches and Warnings - Severe Weather Warnings
Wind Advisory	Watches and Warnings - Special Weather Statements
Winter Weather Advisory	Road Weather Alerts - Alert Parameters
	Watches and Warnings - Special Weather Statements
Dense Fog Advisory	Road Weather Alerts - Alert Parameters
	Watches and Warnings - Special Weather Statements
Flood Advisory	Flood Warning - Forecasted Flood Stage
Flood Stage	Flood Warning - Current Flood Stage
Camera View of Road	Camera - View of Road
Camera View of Weather	Camera - View of Weather
Camera View of Traffic	Camera - View of Weather
Radar Images	Radar - Radar Loop
	Radar - Radar Loop with Precip Type Coloring
	Radar - Future Radar
	Radar - Storm Tracks

Survey participants indicated their assessment of quality for each of the six attributes by selecting their sense of quality from a series of quality ratings that ranged from very high to very low in five levels. These text ratings were then converted to Likert scores, numerical equivalents of the text ratings (5 = very high, 4 = high, 3 = medium, 2 = low, and 1 = very low).

The Likert scores from all respondents were aggregated for each attribute and for the composite of all six attributes, and then averaged to determine the mean quality response for each of the six attributes and the composite score for each *Product Component*. From inception, the Baselining Study recognized there would likely be different levels of quality assessment for DOT representatives who had management responsibilities in different functional areas (maintenance, traffic, traveler information). The Study utilized the three primary management strategy classifications (Advisory, Control, and Treatment) to differentiate the background of each of the respondents. Therefore, the quality response scores were further separated into these three management strategy classes and statistically analyzed separately.

For each *Product Component*, respondents were also asked to rate the importance of that *Product Component*. The answers regarding importance were analyzed in a similar manner as the quality ratings for all, Advisory, Control, and Treatment strategy responses.

The six quality attributes were established prior to the 2008 survey to assess different facets of the quality of road weather information. This approach has proven extremely beneficial in the assessment of quality characteristics because the differentiation has pointed out specific quality attribute scores that vary significantly from the scores of other attributes. However, the evaluation of road weather quality appears to be a complex process involving numerous human factors. The study team approached the Baselining Study in 2008 with the assumption that the selected survey design was measuring the quality of road weather information, or at least quality was the metric chosen to establish a baseline measure with the ongoing intent to assess measurable changes in this chosen metric. The results from the 2008 survey raised questions regarding whether the survey actually measured quality or value. In itself, whether the quality attributes measured quality or value was not critical. Nevertheless, an understanding of the "quality" metric had important impacts on two aspects of the study. First, it influenced the interpretation of the study results and impacted how various readers would evaluate the results. Second, it influenced the organization and structure of the 2010 survey questions since the questions needed to be formulated in a manner appropriate to extract the desired metric.

Results from the 2008 survey indicated quality was perceived by the survey participants as an attribute, characteristic, or property of road weather information that can be observed, interpreted, approximated (quantified), but cannot be measured directly. This subjective definition of quality also incorporates a sense of value that reflects road weather information's ability to meet the specific requirements of the information consumer and the consumer's willingness to expend time and financial resources to acquire and use the resource.

Table 5 lists the six quality attributes, their definitions, and evaluates whether they are more strongly oriented toward a classic definition of quality or value. This differentiation may seem a moot point; however, close analysis of the survey results indicates the overall value of a *Product Component*, *Product*, or *Element* to a survey participant appears to affect all quality attribute responses for that particular *Product Component*, *Product*, or *Element*. This situation was apparent in the scores for the Currency/Latency and Timeliness/Reliability attributes in the 2008 survey. Components delivered in a given *Product* bulletin should have had essentially the same scores within both the Currency/Latency and Timeliness/Reliability attribute categories since all components were delivered together as part of that bulletin. However, the results showed significant variability between these attribute scores for components from common product types. What was evident was that if a component was ranked high across all attributes then the scores for these two attributes were also high. If the attribute scores in general were low, then the scores for the two attributes were also low. These results indicated the DOT responses were significantly affected by the user's assessment of value of that component in the user's operational requirements.

Table 5. Quality attribute classification into quality and value categories

ATTRIBUTE	DEFINITION	TYPE	EXPLANATION
Accuracy/Precision	"Closeness" between an observed or forecasted condition and the actual condition	Quality	Primarily a measurable aspect of the component; the quality measure of "closeness" is partially influenced by the degree to which that "closeness" allows the user to comfortably use the component
Completeness	Adequacy of information to fulfill users' requirements	Value	Use of the component to satisfy user requirements is the assessment criterion.
Relevance	Fit of the information to the users' needs	Value	Use of the component to meet the user's needs is the assessment criterion.
Currency/Latency	Age of the information	Quality	Primarily a measurable characteristic of the component; the quality measure is partially influenced by the degree to which the data age allows the user to comfortably use the information
Timeliness/ Reliability	Consistent and on-time delivery of information	Quality	Primarily a measurable characteristic of the component but it is influenced by value of information as it ages
Ease of Use	Facility to get, interpret, and use the information	Value	Primarily based upon value

Table 6. *Elements* in each management strategy group in the 2008 survey. *Elements* denoted with asterisks were not used in the 2010 survey or were integrated into another *Element.*

ELEMENTS		
Advisory Strategies	**Control Strategies**	**Treatment Strategies**
• Wind direction • Wind speed • Weather type • Probability of precipitation • Estimated amount of precipitation in ranges* • Maximum temperature • Minimum temperature • Road conditions by highway segment* • Road closure • Severe weather advisory • Wind advisory • Winter weather advisory • Dense fog advisory • Flood advisory • Visibility	• Air temperature • Dew point temperature • Relative humidity • Wind direction • Wind speed • Wind gust • Precipitation rate • Precipitation accumulation • Snow rate • Snow accumulation • Weather type • Type of weather and precipitation* • Probability of precipitation • Probability of precipitation type • Estimated amount of	• Air temperature • Dew point temperature • Relative humidity • Wind direction • Wind speed • Wind gust • Precipitation type • Precipitation rate • Precipitation accumulation • Snow rate • Snow accumulation • Weather type • Precipitation start time • Precipitation end time • Precipitation probability • Probability of precipitation type

ELEMENTS		
Advisory Strategies	**Control Strategies**	**Treatment Strategies**
	precipitation in ranges* • Maximum temperature • Minimum temperature • Cloud cover • Pavement temperature • Pavement condition • Frost probability • Road closure • Severe weather advisory • Wind advisory • Winter weather advisory • Dense fog advisory • Flood advisory • Visibility • Camera – road • Camera – weather • Camera – traffic	• Maximum temperature • Minimum temperature • Cloud cover • Visibility • Pavement temperature • Pavement condition • Chemical concentration • Freeze point temp • Frost probability • Treatment recommendation • Road closure • Severe weather advisory • Wind advisory • Winter weather advisory • Dense fog advisory • Flood advisory • Flood stage • Camera – road • Camera – weather • Camera – traffic • Estimated amt of precip* • Flood potential* • Flow rate* • River stage* • Type of weather & precip*

Human factors such as this not only affected the interpretation of the results and the stability of the baseline quality scores, but they became an important consideration in the redesign of the survey format in 2010. The 2008 survey was actually three different surveys, one for each of the three primary management strategy groups. Each of the separate surveys requested input of quality ratings for Elements and Products specific to the different management strategy groups.

The *Elements* for each management strategy are listed in Table 6 and the *Products* are listed in Table 7. This selected resource approach by strategy posed problems since the survey structure constrained the user's ability to rate all potential road weather resources, and the different result sets made cross-comparisons between the responses from the three management strategy groups difficult. It was determined going into 2010 that a better approach was to permit all respondents the ability to rate all possible road weather information options and let the respondent skip sections of road weather information categories that the user did not use. In this way, the user could answer questions on types of information the design team had not expected from a user in a specific management strategy class.

Table 7. *Products* in each management strategy group in the 2008 survey

PRODUCTS		
Advisory Strategies	**Control Strategies**	**Treatment Strategies**
• Zone Forecast • Route Specific Forecast • Watches and Warnings • Road Condition Report	• Zone Forecast • Pavement Weather Forecast • Watches and Warnings • Road Condition Report • Camera Images	• Weather Summary • Weather History • ESS Current Conditions • ESS Histories • Regional Map • Zone Forecast • Pavement Weather Forecast • Road Weather Alerts • Watches and Warnings • Maintenance Decision Support Systems • Road Condition Report • Flood Warning • Recorded Road Weather Forecast

Based on feedback from the 2008 survey, the 2010 survey was modified to assess the quality and importance of *Product Components* instead of *Products* and *Elements*. *Product Components* are the fundamental pieces of road weather information DOT personnel gather to support their decision processes; therefore, the acquisition of input on the quality of *Product Components* offered a more reliable approach to creating a stable metric for the longitudinal assessment of road weather quality and trends in this quality. At the same time, it was imperative to retain the value of the 2008 survey results and the ability to assess trends in the quality of road weather information from comparison of the 2008 results to the 2010 and subsequent survey results. Table 2 and Table 4 illustrate the relationship between the *Product Component* responses and the *Element* and *Product* response categories from the 2008 survey. These relationships become the mechanism for generating derived *Element* and *Product* results from the *Product Component* results. Since all users have access to all *Product Component* questions and may respond with quality ratings to any question, it is now possible to generate *Element* and *Product* results in management strategy categories that the 2008 survey design had precluded. The design of the 2010 survey provides some modifications that will benefit the analysis of the Baselining Study results:

- The establishment of metrics based upon *Product Components*;
- A method to compare 2008 *Product* results with results from the 2010 surveys and subsequent surveys;
- A method to compare 2008 *Element* results with results from the 2010 survey and subsequent surveys; and,
- An expansion of the *Product* and *Element* results categories in all management strategies to a universal set of categories that is consistent across all strategies.

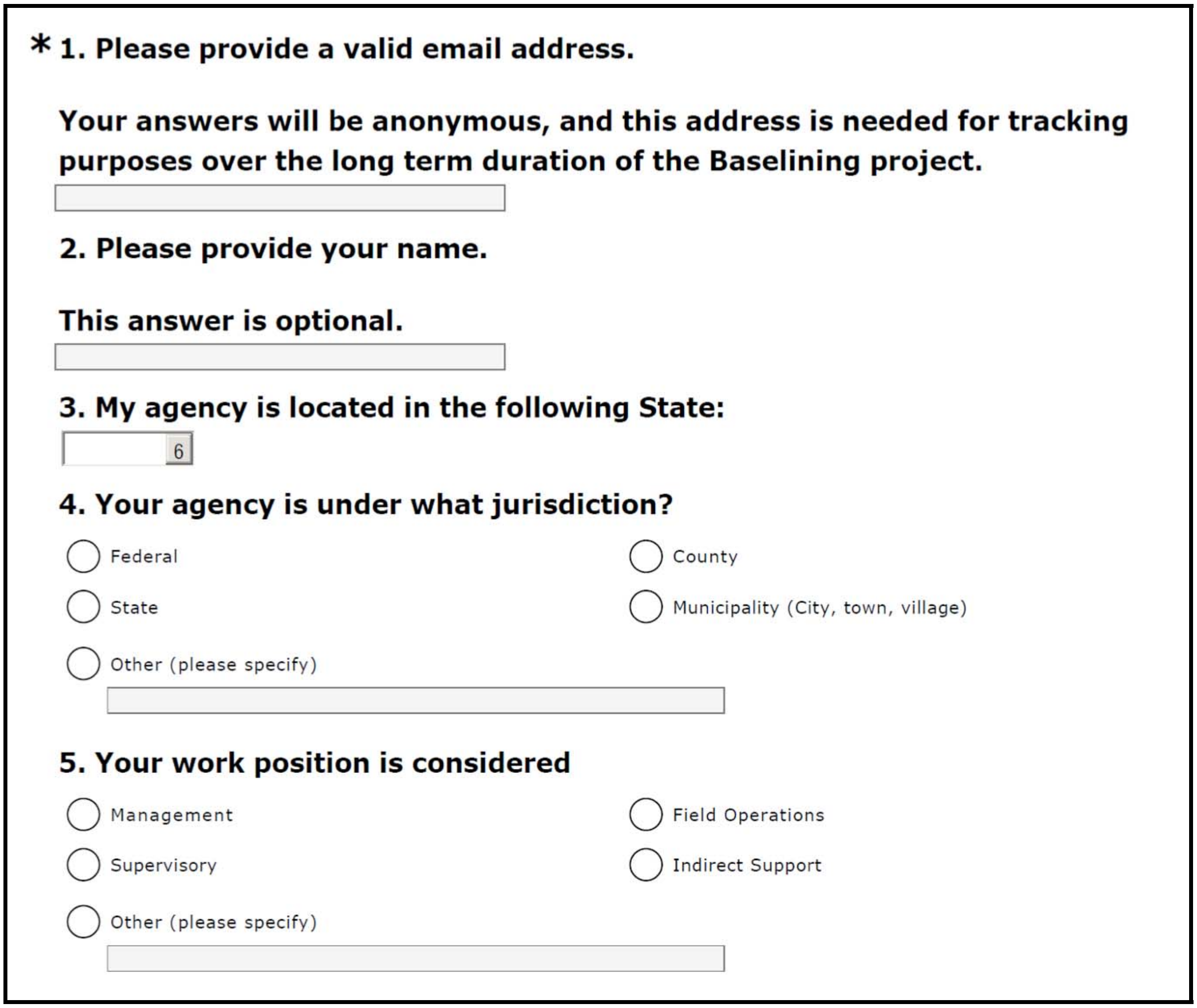
2.2 2010 SURVEY AND DATABASE

The goal of the 2010 Baselining survey was to acquire input on the quality and importance of road weather information used to support various operational activities. The survey was divided into two sections. The first section collected general demographic information from the respondents plus information regarding their experience and familiarity with road weather management. The second section contained questions seeking the respondent's quality valuation of each *Product Component* based on each of the six quality attributes plus the *Product Component's* importance.

The first section contained a combination of multiple-choice and text entry questions. These questions appeared on pages 1 and 2 of the web-based electronic survey (see Figure 1 and Figure 2). The requested information included:

- Email address
- Name (optional)
- State
- Agency jurisdiction (federal, state, county, municipal)
- Job classification level (management, supervisory, field operations, other)
- Management strategy class percentage (Advisory, Control, Treatment)
- Years of experience using road weather information

Although it is not apparent in Figure 1, question 3 had a drop down window containing the names of all the states, and users selected the state name from the list to answer the question. An answer was required for question 1 in order to proceed, but responses to the remaining demographic questions were optional. Figure 2 contains the survey questions up to the day before the survey was released. Question 6 was dropped to reduce the number of questions in the survey, and question 7 was changed to permit participants to indicate the amount of their time they allocated to each strategy. The survey instrument did check to assure that the three numbers entered summed to 100%. This modification seemed reasonable since the expectation was that the primary management strategy could be extracted from the modified question 7. This decision turned out to be a tactical mistake because a number of respondents selected 50% in two management strategies or some other combination that did not result in a single dominant strategy selection. This impacted the ability to classify a number of participants as definite members of the Advisory, Control, or Treatment group. A definitive strategy classification was necessary to perform the statistical analyses and compare the results from the 2008 survey to the 2010 survey, and the lack of a specific strategy selection meant manual intervention was necessary to assure a discrete digital value be placed in the database.



* 1. Please provide a valid email address.

Your answers will be anonymous, and this address is needed for tracking purposes over the long term duration of the Baselining project.

2. Please provide your name.

This answer is optional.

3. My agency is located in the following State:

4. Your agency is under what jurisdiction?

- Federal
- County
- State
- Municipality (City, town, village)
- Other (please specify)

5. Your work position is considered

- Management
- Field Operations
- Supervisory
- Indirect Support
- Other (please specify)

Figure 1. Page 1 of the demographic questions in the 2010 survey.

6. Your work responsibilities fall predominantly within which management strategy?

Advisory: Provide information on prevailing and predicted conditions to both transportation managers and motorists (e.g., Variable Message Signs, HAR, 511)

Control: Alter the state of roadway devices to permit or restrict traffic flow and regulate roadway capacity (e.g., Traffic Signal Timing, Ramp Metering)

Treatment: Supply resources to roads to minimize or eliminate weather impacts (e.g., Snow and Ice Control, Winter Maintenance)

- ◯ Advisory
- ◯ Control
- ◯ Treatment

7. Please indicate the relative amount of time you put into each management strategy. (Total for all columns should be 100%)

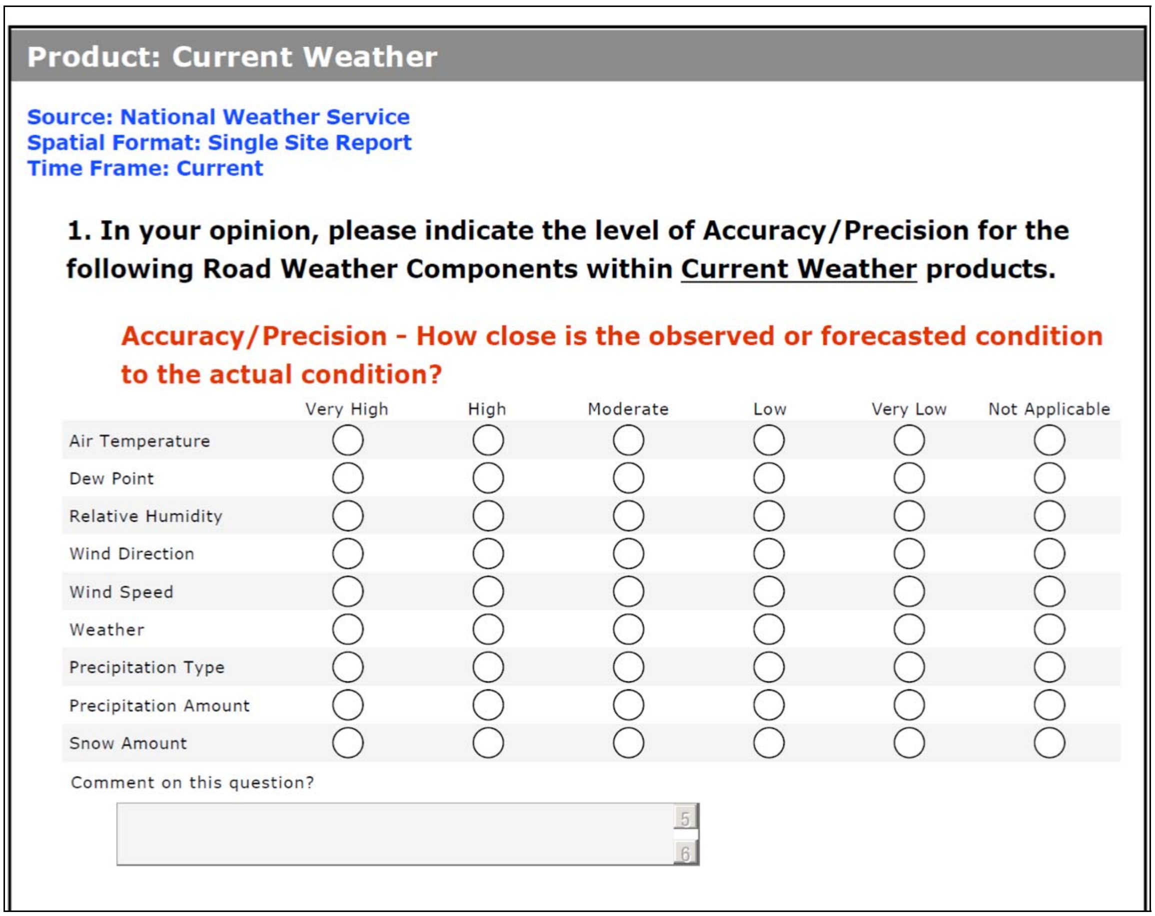
	Advisory	Control	Treatment
0 %	◯	◯	◯
20 %	◯	◯	◯
40 %	◯	◯	◯
60 %	◯	◯	◯
80 %	◯	◯	◯
100 %	◯	◯	◯

8. How long have you been using road weather information, including both your current and previous jobs?

- ◯ Less than 5 years
- ◯ 5 - 10 years
- ◯ 11 - 15 years
- ◯ 16 - 20 years
- ◯ 21 - 25 years
- ◯ 26 - 30 years
- ◯ More than 30 years

Figure 2. Page 2 of the demographic questions in the 2010 survey just prior to releasing the survey to the DOT participants

The second part of the survey was the core and largest part of the survey. It contained questions regarding the quality of each of the six quality attributes plus the importance rating for each of the *Product Components*. The core of the 2010 survey was organized into sections, each section containing the *Product Components* within one of the *Products.* This organizational structure kept all of the *Product Components* that exhibited the same unique characteristics (i.e., source, spatial domain, and time frame) on one page and in one set of multiple-choice questions. This structure was used for two reasons. First, at the beginning of each section survey participants were provided information about the *Product* and the *Product Components* within the *Product*. They then had the option to skip all of the questions related to that *Product* if they did not use the *Product*. Second, by organizing all *Product Components* with similar characteristics (source, spatial format, and time frame), survey participants could answer all the questions in a given *Product* section using a consistent mental frame of reference. The survey format is shown in Figure 3.



Product: Current Weather

Source: National Weather Service
Spatial Format: Single Site Report
Time Frame: Current

1. In your opinion, please indicate the level of Accuracy/Precision for the following Road Weather Components within Current Weather products.

Accuracy/Precision - How close is the observed or forecasted condition to the actual condition?

	Very High	High	Moderate	Low	Very Low	Not Applicable
Air Temperature						
Dew Point						
Relative Humidity						
Wind Direction						
Wind Speed						
Weather						
Precipitation Type						
Precipitation Amount						
Snow Amount						

Comment on this question?

Figure 3. Sample survey question

Within each *Product* section, there were seven (7) sets of questions, one set for each of the six (6) quality attributes and one for the input on importance. Each set contained a list of all components within that *Product*, and each component followed a series of six possible mutually exclusive radio buttons representing the five rating levels or a "not applicable" option. The quality assessment range of "very high" to "very low" permitted the use of a five-point Likert scale to facilitate a numerical evaluation of user responses. “Very high” was set to a Likert score of 5, and “very low” to 1.

The format in Figure 3 represents one page in a commercial web-based survey tool illustrating the Precision/Accuracy attribute questions for the Current Weather *Product*. Therefore, there were seven pages of questions in each *Product* section. As indicated previously, survey participants could make the decision at the beginning of a *Product* section to skip all of the questions related to this *Product,* which then caused the survey tool to totally skip all seven pages associated with this Product. In addition, answers on any given page were not mandatory, so participants could skip past *Product Components* they did not use or chose not to answer. At the end of each page there was a box which allowed participants to provide a free form comment regarding their answers on the page.

The survey was identical for all survey participants, which was a significant change from the 2008 survey. Regardless of their primary management strategy, this allowed all users to rate the quality of all the *Products* they wished to address. This not only permitted a greater ability to compare the assessments of quality and value across the Advisory, Control, and Treatment management strategies, but also provided an indication of the level of variability in usage of different *Products* within specific management strategies.

The project team contacted individuals in each of the road weather management strategy areas (Advisory, Control, and Treatment) in forty-five (45) DOTs to identify their willingness to participate in the quality assessment. An awareness of varying types of road weather conditions across differing geographical areas was an important decision in identifying the potential DOT participants. Of the original list of DOTs contacted (Figure 4), twenty-eight (28) DOTs completed the survey (Figure 5).

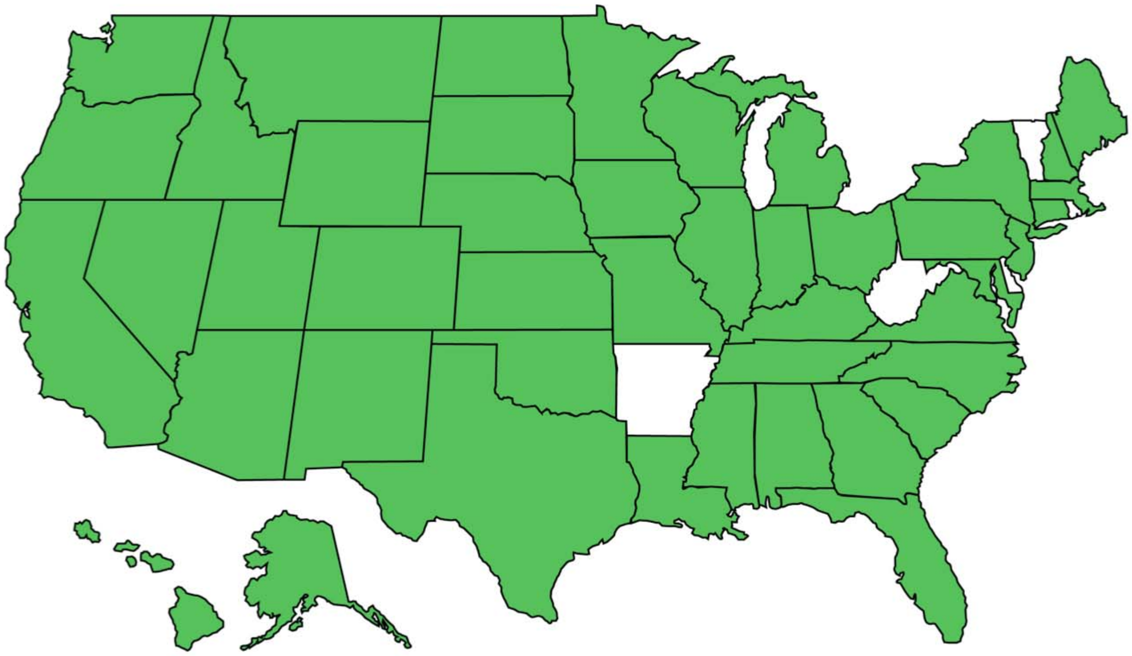


Figure 4. Shaded states indicate those states contacted

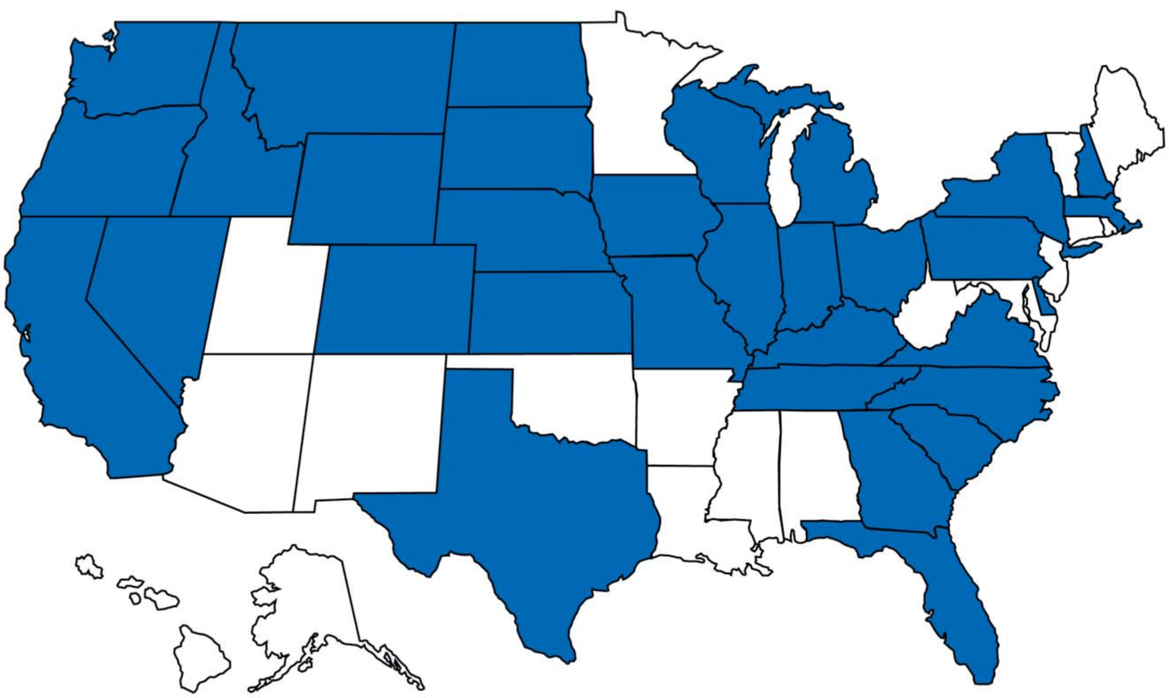


Figure 5. Shaded states indicate those states participating in the survey

Upon completion of the 2010 survey, the results were transferred to a relational database. The relational database management system chosen for the Baseline Study database was MySQL, the world's most widely used open source database. The database structure follows the entity-relationship (ER) modeling methodology utilized by software engineers to produce a conceptual schema or semantic data model of a system, such as this relational database, in a top-down fashion. The entities in the Baseline Study database are the primary survey components (*Product Components, Products, Elements,* quality attributes). A relationship captures how two or more entities are related to one another. Entities and relationships can have attributes specifying the specific characteristics of the entity or relationship. ER diagrams are used to graphically represent the logical database structure. Boxes in the ER diagrams represent entities; diamonds represent relationships; and ovals represent attributes. In the text discussion, the database entities, relationships, and attributes will be denoted as bold-faced text (e.g., **respondent**, an ER entity).

The intent of the database was to store the survey responses associated with the quality attribute assessments and importance of road weather parameters in an electronic database. The responses represent users' input to the set of survey questions. The responses may come from the 2008 survey or the revised 2010 survey format that will form the structure for the 2010 results and ensuing surveys. All future survey responses will be stored using this database structure.

Each question within the surveys either relates to information about the participant or directly requests input on the quality and/or importance of road weather information parameters. A **respondent** refers to the individual who participated in the survey. The survey instrument maintains a link to this participant; therefore, each completed survey is attributable to given participant. This identity is not used in the generation of the survey results or divulged in any form; rather it provides a background mechanism to track continuity of the quality assessment from repeat users over time. The information about the respondent is stored as a **respondent** entity. A **respondent** entity ties together the survey data submitted for a single survey, including quality and importance attributes, demographic information, and survey-specific information. The **respondent** entities will be linked to a specific respondent using the **participant ID** attribute. The attributes associated with the **respondent** entity in the 2010 survey are shown diagrammatically in Figure 6.

The 2008 survey required a more complex database structure because of the separate surveys based upon the three primary management strategy classes and some additional questions asked in the survey that were dropped from the 2010 survey. The **respondent** entity for the 2008 survey (Figure 7) is similar to the one for the 2010 survey; however, it does not include the **Job Class, Experience, Advisory, Control, and Treatment** attributes. The first two reflect additions to the 2010 survey and the last three are needed to store the percentage involvement in each of the three management strategy classes.

The **responds** relationship (Figure 8) contains both **Product** and **Element** entities and the database structure links the **respondent**, **product**, **element**, and **quality attribute** entities with the selected response value and any comments from the respondent (**select** and **comment** attributes). For the purposes of storing data from the 2008 survey the schema includes questions about the source of the data the participants used, whether the data was fee-based or free, and the perceived cost-benefit of the data. The responses were captured in the database in a format described by the ER-diagram of the **uses** relationship (Figure 9).

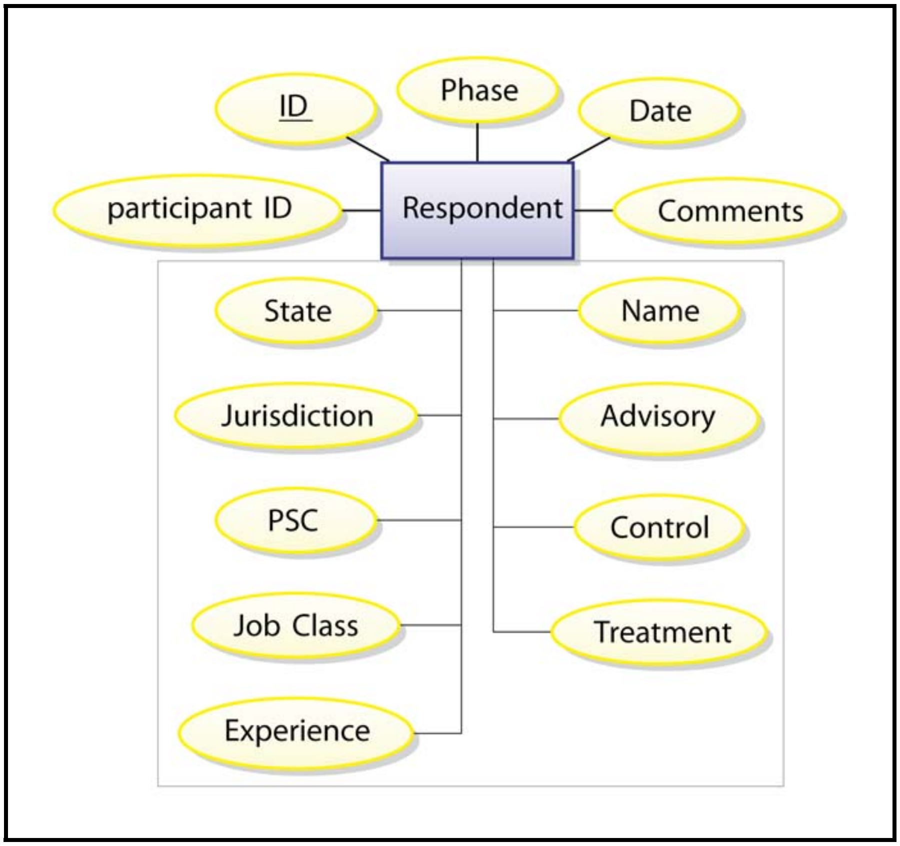


Figure 6. Entity relationship diagram of a respondent for the 2010 survey

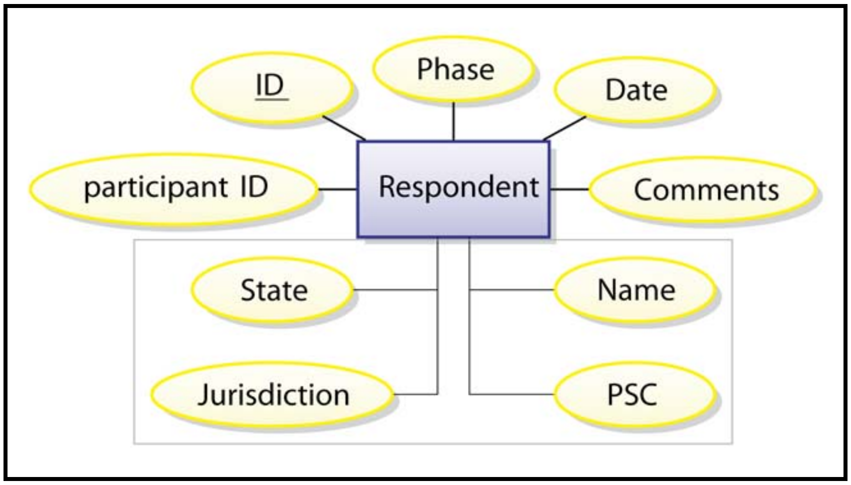


Figure 7. Entity relationship diagram of a respondent in the 2008 survey

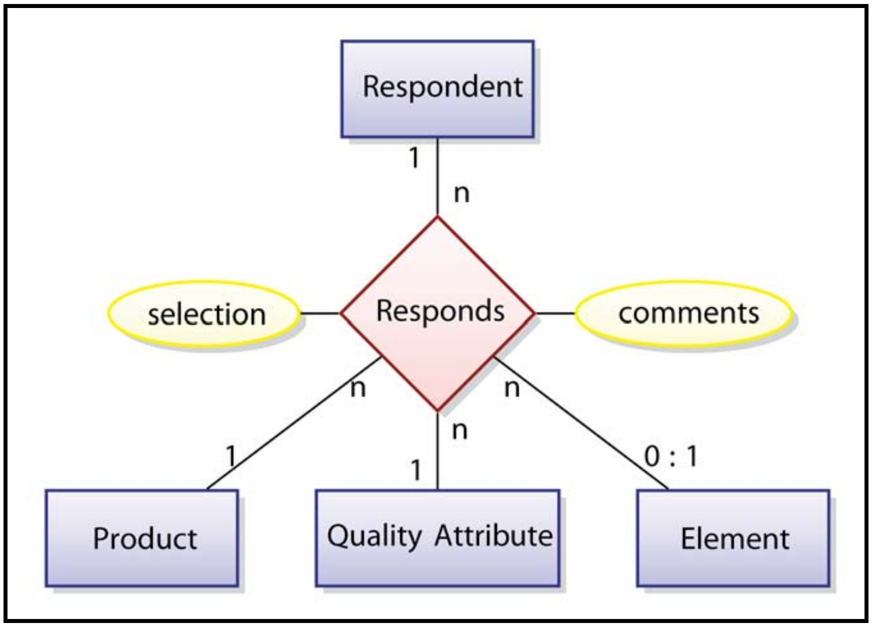


Figure 8. Entity relationship diagram of the responds relationship in the 2008 survey

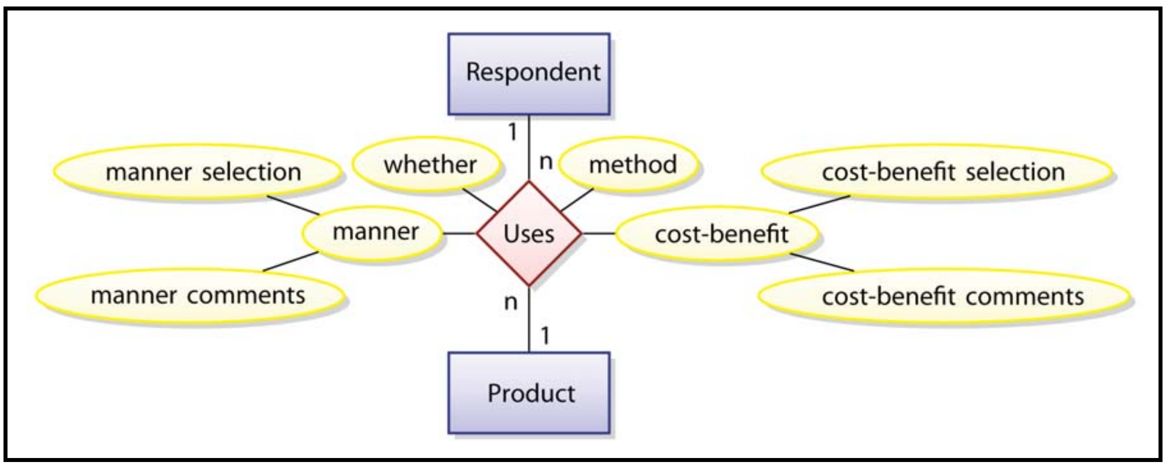


Figure 9. Entity relationship diagram of a uses relationship

The MySQL syntax to create the entity and relationship database entries may be found in Appendix A.

2.3 ANALYTICAL COMPUTATIONS

To gain the full benefit of the results from the 2008 and 2010 surveys, it was necessary to analyze the data stored in the MySQL relational database to establish the baseline level of quality and assess the trends of these quality metrics over time. A standard set of analysis routines was developed for the 2008 results, and these analyses have been expanded to permit the comparison of the results between the 2008 and 2010 surveys. The standard set addresses the quality measures of the three core data objects (*Product Components*, *Elements*, and *Products*) for all respondents and for each of the three primary management strategy classes (Advisory, Control, and Treatment). Commencing with this 2010 survey report, the standard set also includes analyses of the comparison or results between surveys and as trends. The **respondent** entity in the MySQL database contains a number of demographic attributes that can permit further evaluation of the data beyond the standard analyses performed for this study.

The Analytical Computations may be divided into three computational sections, each dealing with one of the three core data objects. These core data objects are:

- *Product Components*: Individual road weather parameters delivered by an unique source and fitting into specified spatial and temporal categories
- *Products*: Packages of road weather parameters composed by a transportation or meteorological agency which contain multiple road weather parameters organized to meet specific user needs
- *Elements*: Types of information utilized by transportation users to support operational needs; they may be weather or transportation-related parameters

The first stage of the standard computational analysis is the generation of results for each survey year required in the analysis. Since the Baselining Study is interested in the trend in road weather quality, the analysis should include the generation of results for all survey years. The modification of the survey between survey years 2008 and 2010 means the process to compute results for 2008 is different from the process needed for 2010 and the ensuing survey years. To simplify the discussion regarding the multi-year survey process in this section, the 2008 survey is denoted as Phase 1 and the 2010 and ensuing surveys are called Phase n where $n = 2$ through the total number of surveys. For Phase 1 of the study, the assessment of quality was done for each of the three primary management strategy classes and different surveys were composed for each strategy class. The Phase 1 survey also contained questions only on *Products* and *Elements*. Thus, there were eight (8) separate sets of results, four for *Product* responses with a set for each of the three strategies plus a composite set and four similar sets for each of the Element responses. With the change of the survey beginning with Phase 2 to address *Product Components*, the number of direct computation sets was reduced to four (4), one for each of the primary management strategy classes plus a composite of all responses. The computation process for Phase 1 is shown in the top portion of Figure 10 and the computation of the statistical analyses of the *Product Component* results are indicated as the top four primary results in the Phase n section.

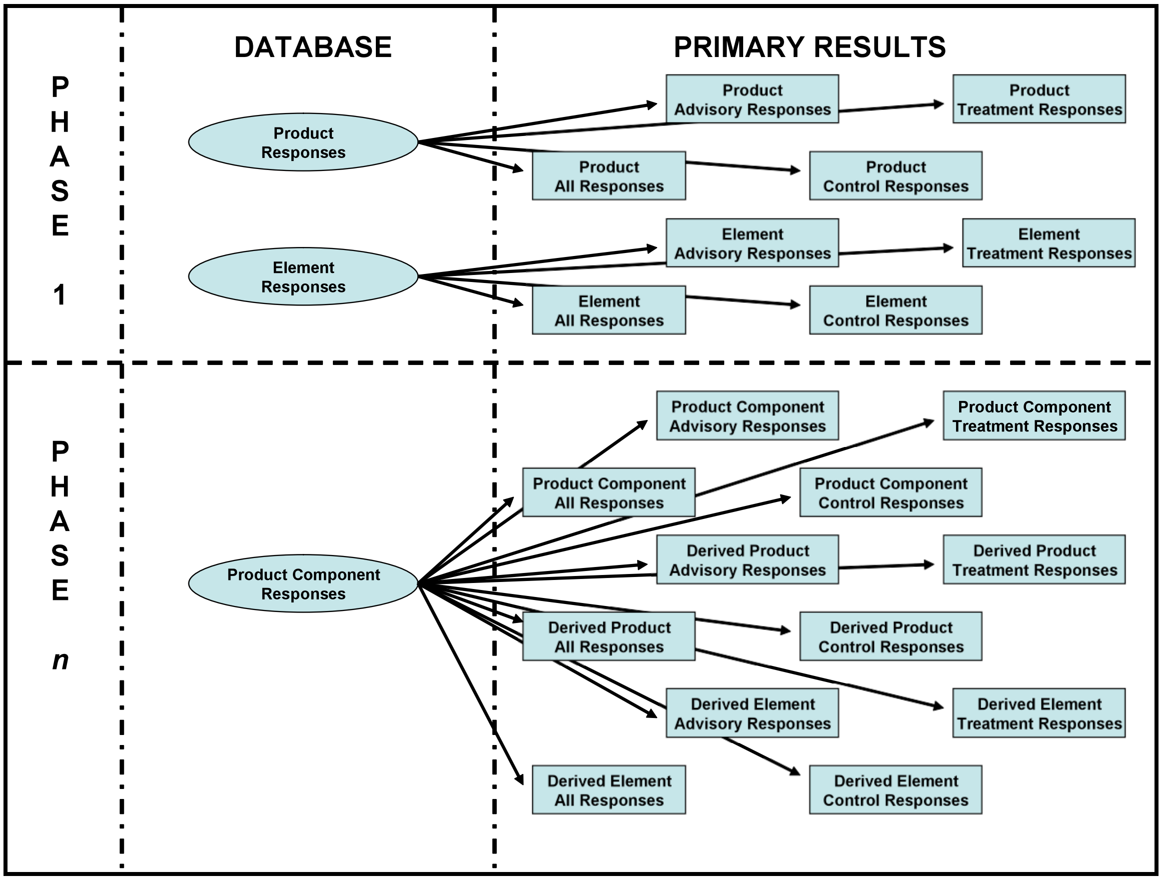


Figure 10. Primary computation sets generated from the data stored in the MySQL database

To retain continuity between the Phase n surveys and the Phase 1 survey, it was also necessary to generate *Product* and *Element* results derived from the *Product Component* responses. The process to accomplish this derivation was discussed in the 'Design of the Quality Assessment Methodology' section. Four sets of *Product* results were generated for the three primary management strategy groups and one for all respondents combined. In like manner, four sets of *Element* results were generated. This set of eight derived *Product* and *Element* results will be part of the standard computations in subsequent surveys. This process is shown diagrammatically in the lower portion of the Phase *n* section of Figure 10.

These computations formed the direct statistical results from the user responses or the derived statistical results in the case of Phase *n Products* and *Elements*. The statistical parameters included the mean, median, standard deviation, maximum value, minimum value, and the skewness of the distribution. The means were then ranked from highest to lowest to determine which *Product Components*, *Products, and Elements* received higher or lower quality and importance scores.

The primary statistical results from a given survey have value in understanding the responses for that phase of the Baseline Study, but the real interest of the road weather information quality baselining process is the longitudinal analysis of the results. This includes trends in the perception of road weather information quality, the relative importance of specific *Product*

Components, the ongoing interrelationships between the quality attribute measures, and different perspectives of users based upon their different management strategy obligations. To accomplish these comparisons, it is necessary to evaluate the results from two or more surveys. The computations necessary to generate these comparisons yield secondary results derived from comparing two or more sets of primary results.

The secondary results associated with the *Product Component* analyses are presented diagrammatically in Figure 11. The secondary trends and interrelationships are derived from the primary results taken from multiple surveys for common management strategy classifications or all responses combined. The stacked boxes in the diagram indicate results from multiple years. It should be noted these comparisons can be done starting with the results from the 2010 survey (Phase 2). The standard comparisons will include comparisons of the means for the each *Product Component* between two surveys or the mathematical trends of the mean values over several surveys. It is anticipated there will likely be fluctuations in the quality and importance means from survey to survey, but these survey-to-survey fluctuations may not be apparent if all of the differences are averaged. The BAH team anticipates the average differences or average trend values will become the primary baseline markers for assessing the change in road weather information quality over time. Still, the differences for individual *Product Components* or groups of *Product Components* will provide insight into specific areas where road weather information quality anomalies are most apparent.

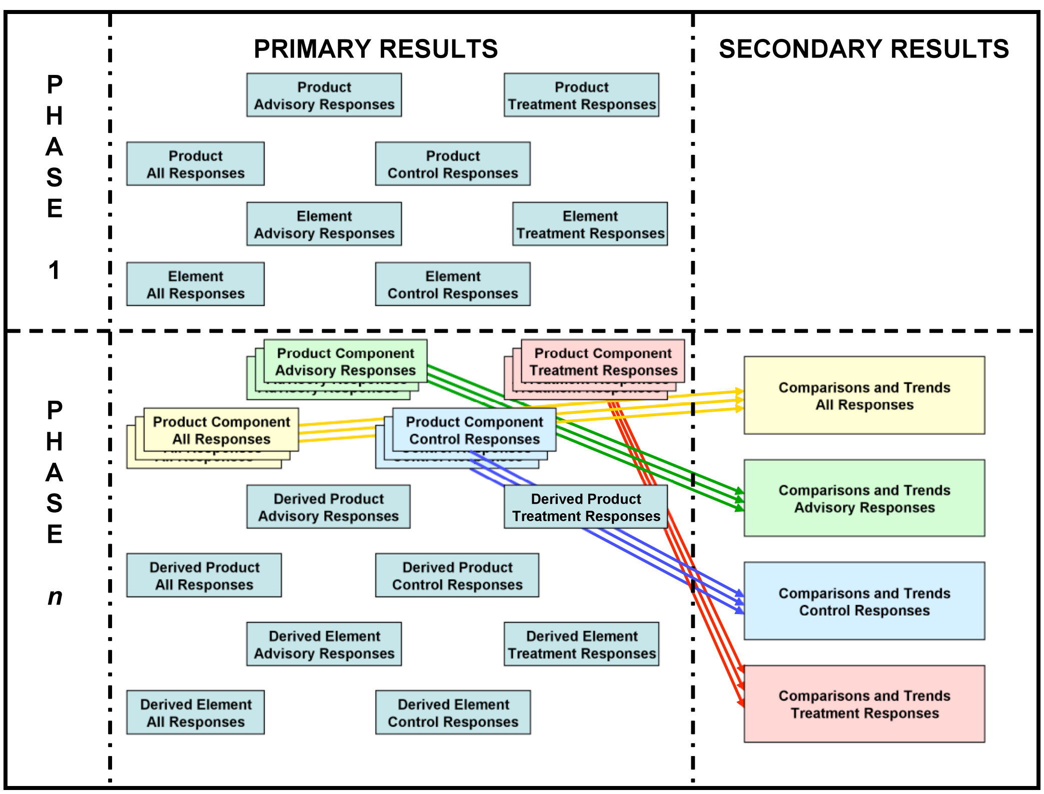


Figure 11. Secondary computation sets derived from the *Product Component* primary results

The second major comparison of results over multiple years addresses the results based upon *Product* and *Element* responses. Figure 10 illustrates the process of generating derived *Product* and *Element* results from *Product Component* results. The derived *Product* results can then be compared against Phase 1 *Product* results or with the results from any other phase of the Baselining study. Likewise, *Element* derived results from a Phase *n* survey can be compared to the results from Phase 1 *Element* results or the results from any other Phase *n* survey. Figure 12 illustrates the generation of secondary *Product* results and Figure 13 shows the process of creating secondary *Element* results.

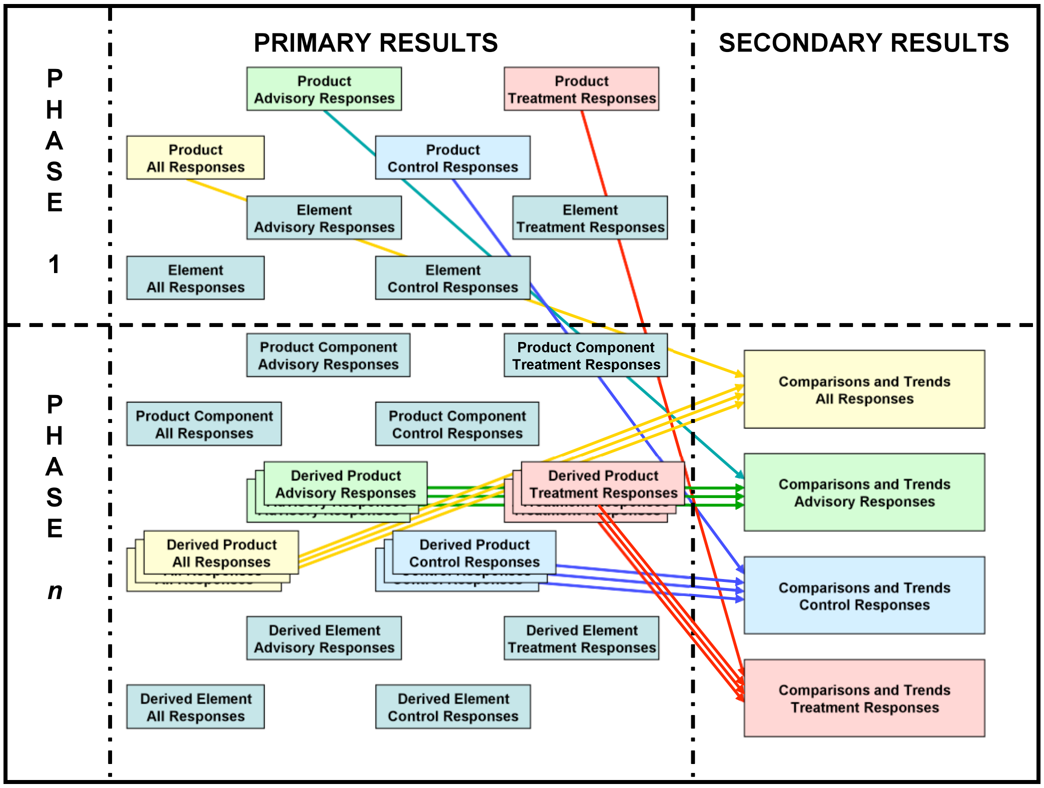


Figure 12. Secondary computation sets derived from the primary results dealing with *Products*

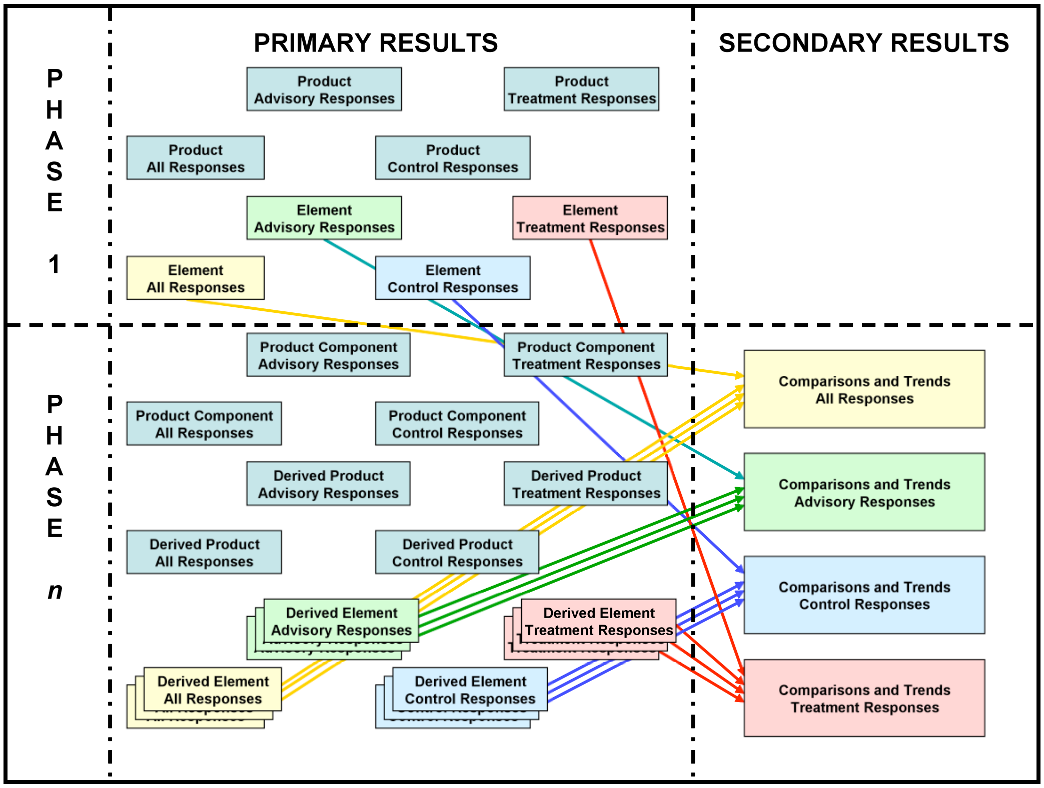


Figure 13. Secondary computation sets derived from the primary results dealing with *Elements*

3.0 Results from the 2010 Survey

3.1 Survey Participation

Forty-five (45) different individuals started the 2010 survey. Thirty-seven (37) provided answers to the questions for one or more of the *Product* groups and the majority of the 37 respondents submitted answers to questions in half or more of the *Product* categories. One individual indicated use of a number of the *Products*, but failed to answer any of the quality assessment questions. The responses from the 37 survey participants who answered the quality assessment questions form the basis for the Results analysis in this report.

3.2 Product Component Results

The quality assessment responses from the 37 respondents were stored by *Product Component*. The responses from the seven (7) attribute responses (six quality attributes and importance) associated with each *Product Component* were converted to their Likert score equivalents. Average response scores were computed for each *Product Component*/Attribute pair using all non-null Likert responses. Null scores included questions with *no response* or a *Not Applicable* response. Table 8 indicates the average scores for the seven attributes for each *Product Component*. The table also includes the composite average for each *Product Component* computed by taking the average of all Likert scores submitted for all six of the quality attributes. The composite averages were computed from the actual Likert scores from all participants across the six attributes and not by averaging the mean values in the six attribute columns in the table. The list of *Product Components* in Table 8 retains the order presented in Table 2. The mean values for each of the *Product Components* and their associated quality attributes and importance were then analyzed to determine a statistical mean, median, standard deviation, maximum, minimum, and skewness for each quality attribute, the composite average, and the importance. The summary statistics from these computations based upon responses from all users are shown in Table 9.

In the demographic section of the 2010 survey, all participants were asked to indicate the percentage of their involvement in each of the three primary management strategies. Nineteen (19) respondents indicated their primary management responsibility was advisory in nature, five (5) said control, and nineteen (19) selected treatment. Two individuals did not select a management category. In addition, six (6) participants did not answer questions beyond the demographic section. This reduced the number of respondents who answered the quality section of the survey in each management category to 15 in Advisory, 5 in Control, and 17 in Treatment. Since survey participants could skip questions dealing with products they did not use, the number of responses to compute several averages was considerably lower than the possible maximum of 37 responses. Table 10 indicates the total number of scores used to compute each average using the responses from all survey participants. Since a significant number of the *Product Component* averages were computed using 15 scores or less (using the responses from all participants), the numbers in Table 10 present a strong argument for not analyzing averages by management strategy. Therefore, all *Product Component* results were performed upon the composite set of results from all respondents.

Table 8. Average scores for quality attributes, composite average, and importance computed from scores submitted by all participants

PC ID	PRODUCT COMPONENT	Accuracy / Precision	Completeness	Relevance	Currency / Latency	Timeliness / Reliability	Ease of Use	Average Composite Attribute Score		Importance
1	Weather Summary - Air Temperature	4.00	4.23	4.23	3.97	3.93	4.14	4.08		4.17
2	Weather Summary - Dew Point	3.88	4.15	3.74	3.92	3.79	3.80	3.88		3.70
3	Weather Summary - Relative Humidity	3.78	4.16	3.52	3.92	3.84	3.76	3.83		3.37
4	Weather Summary - Wind Direction	4.03	4.19	4.06	3.90	3.83	4.03	4.01		3.80
5	Weather Summary - Wind Speed	4.00	4.19	4.23	3.94	3.83	4.10	4.05		4.30
6	Weather Summary - Weather	3.68	3.77	4.00	3.77	3.72	3.97	3.82		4.43
7	Weather Summary - Precip Type	3.77	3.81	4.23	3.77	3.69	3.90	3.86		4.60
8	Weather Summary - Precip Amount	3.33	3.67	3.97	3.67	3.68	3.83	3.69		4.62
9	Weather Summary - Snow Amount	3.11	3.57	4.14	3.79	3.67	3.89	3.70		4.61
10	Weather History - Air Temperature	3.85	4.17	3.83	3.82	3.64	3.91	3.87		4.27
11	Weather History - Dew Point	3.92	4.17	3.67	3.73	3.64	3.82	3.83		4.09
12	Weather History - Relative Humidity	3.83	4.08	3.42	3.73	3.55	3.82	3.74		3.73
13	Weather History - Wind Direction	4.00	4.17	3.67	3.82	3.64	3.91	3.87		4.36
14	Weather History - Wind Speed	3.85	4.08	3.75	3.82	3.64	3.91	3.84		4.36
15	Weather History - Precip Amount	3.15	3.50	3.67	3.45	3.36	3.82	3.49		4.64
16	Weather History - Snow Amount	3.08	3.55	3.64	3.50	3.40	3.90	3.50		4.73
17	ESS Observations - Air Temperature	4.17	4.28	4.16	3.80	4.00	4.08	4.08		4.12
18	ESS Observations - Dew Point	4.05	4.08	4.21	3.79	3.96	3.83	3.99		4.04
19	ESS Observations - Relative Humidity	4.04	4.16	3.92	3.76	4.00	3.84	3.95		3.46
20	ESS Observations - Wind Direction	4.21	4.16	4.12	3.72	4.00	3.92	4.02		3.96
21	ESS Observations - Wind Speed	4.08	4.16	4.28	3.76	4.00	4.00	4.05		4.16
22	ESS Observations - Wind Gust	4.04	4.16	4.12	3.76	3.96	4.00	4.01		4.04
23	ESS Observations - Precip Type	3.33	3.33	4.04	3.50	3.71	3.83	3.63		4.57
24	ESS Observations - Pavement Temperature	3.75	3.92	4.32	3.72	3.84	4.08	3.94		4.56
25	ESS Observations - Pavement Condition	3.13	3.28	3.64	3.68	3.72	3.79	3.54		4.40
26	ESS Observations - Chemical Concentration	2.20	2.68	3.19	3.38	3.38	3.10	2.99		3.52
27	ESS Observations - Freeze Point Temperature	2.50	3.14	3.71	3.65	3.55	3.43	3.33		4.24
28	ESS Histories - Air Temperature	3.88	4.00	3.67	3.88	4.29	4.00	3.94		3.80
29	ESS Histories - Dew Point	3.86	4.00	3.75	3.86	4.33	4.00	3.95		3.89
30	ESS Histories - Relative Humidity	3.88	4.00	3.67	3.88	4.29	4.00	3.94		3.80
31	ESS Histories - Wind Direction	3.88	4.00	3.78	3.88	4.29	4.00	3.96		3.90
32	ESS Histories - Wind Speed	3.88	4.00	3.78	3.88	4.29	4.00	3.96		3.90
33	ESS Histories - Precip Type	3.63	3.78	4.13	4.00	4.29	3.67	3.90		4.20
34	ESS Histories - Precip Start	3.25	3.75	4.13	3.63	4.29	3.75	3.79		4.10
35	ESS Histories - Precip End	3.25	3.75	4.13	3.63	4.29	3.75	3.79		4.10
36	ESS Histories - Pavement Temperature	3.50	3.89	4.11	3.88	4.29	3.89	3.92		4.20
37	ESS Histories - Pavement Condition	2.75	3.88	3.67	3.63	4.29	3.00	3.51		3.80

PC ID	PRODUCT COMPONENT	Accuracy / Precision	Completeness	Relevance	Currency / Latency	Timeliness / Reliability	Ease of Use	Average Composite Attribute Score		Importance
38	ESS Histories - Chemical Concentration	1.83	3.17	3.00	3.33	4.20	2.57	2.97		3.50
39	ESS Histories - Freeze Point Temperature	2.00	3.17	3.33	3.60	4.25	2.83	3.16		3.50
40	Map - Air Temperature	3.91	4.23	3.95	3.75	4.19	4.19	4.04		4.10
41	Map - Dew Point	3.85	4.20	3.95	3.75	4.10	4.15	4.00		4.00
42	Map - Relative Humidity	3.81	4.24	3.86	3.75	4.10	4.20	3.99		3.70
43	Map - Wind Direction	3.95	4.23	4.18	3.75	4.14	4.24	4.09		4.00
44	Map - Wind Speed	3.95	4.23	4.18	3.75	4.14	4.19	4.08		4.19
45	Map - Precip Type	3.67	4.05	4.29	3.85	4.05	3.95	3.98		4.29
46	Map - Pavement Temperature	3.79	4.05	4.25	3.74	4.11	4.11	4.01		4.33
47	Map - Pavement Condition	3.53	3.76	3.94	3.83	4.11	4.06	3.88		4.29
48	Map - Chemical Concentration	3.15	3.50	3.33	3.64	4.00	3.85	3.58		3.56
49	Map - Freeze Point Temperature	3.47	3.87	3.79	3.75	4.13	4.00	3.83		4.06
50	Zone Forecast - Maximum Temperature	3.67	4.25	3.33	3.83	3.73	3.92	3.79		3.75
51	Zone Forecast - Minimum Temperature	3.58	4.25	3.50	3.83	3.73	3.92	3.80		3.92
52	Zone Forecast - Wind Direction	3.42	4.25	3.50	3.75	3.82	3.92	3.77		4.17
53	Zone Forecast - Wind Speed	3.50	4.25	3.58	3.67	3.82	3.92	3.79		4.17
54	Zone Forecast - Weather	3.33	4.08	3.67	3.83	3.73	3.83	3.75		4.42
55	Zone Forecast - Probability of Precipitation	3.17	4.00	3.67	3.83	3.73	3.75	3.69		4.33
56	Pavement Forecast - Air Temperature	4.13	4.50	4.20	3.93	4.36	4.67	4.30		4.57
57	Pavement Forecast - Dew Point	4.00	4.36	4.27	3.93	4.36	4.50	4.24		4.57
58	Pavement Forecast - Relative Humidity	4.00	4.36	4.07	3.93	4.36	4.50	4.20		4.07
59	Pavement Forecast - Wind Direction	4.20	4.50	4.40	4.00	4.36	4.67	4.36		4.57
60	Pavement Forecast - Wind Speed	4.07	4.50	4.47	4.00	4.36	4.67	4.34		4.64
61	Pavement Forecast - Wind gust	3.87	4.50	4.33	3.93	4.36	4.67	4.28		4.21
62	Pavement Forecast - Weather	4.07	4.50	4.47	4.00	4.36	4.60	4.33		4.62
63	Pavement Forecast - Visibility	3.71	4.31	3.93	3.85	4.23	4.43	4.07		4.31
64	Pavement Forecast - Cloud Cover	4.08	4.23	3.93	3.92	4.25	4.43	4.14		3.93
65	Pavement Forecast - Precipitation Type	4.00	4.27	4.53	4.00	4.36	4.57	4.28		5.00
66	Pavement Forecast - Precipitation Start Time	3.73	4.43	4.60	3.93	4.36	4.50	4.26		4.93
67	Pavement Forecast - Precipitation End Time	3.67	4.43	4.60	3.93	4.36	4.50	4.24		4.93
68	Pavement Forecast - Probability of Precipitation	3.88	4.40	4.47	4.00	4.36	4.36	4.24		4.79
69	Pavement Forecast - Probability of Precipitation Type	3.88	4.33	4.40	4.00	4.36	4.36	4.22		4.64
70	Pavement Forecast - Precipitation rate	3.60	4.36	4.40	4.00	4.36	4.50	4.20		4.71
71	Pavement Forecast - Precipitation accumulation	3.93	4.36	4.47	4.00	4.36	4.43	4.26		4.86
72	Pavement Forecast - Snow Rate	3.80	4.36	4.47	4.00	4.38	4.43	4.24		4.92
73	Pavement Forecast - Snow Amount	3.73	4.27	4.40	4.00	4.38	4.50	4.21		4.93
74	Pavement Forecast - Pavement Temperature	3.88	4.40	4.50	3.93	4.21	4.53	4.24		4.87
75	Pavement Forecast - Pavement Condition	3.60	4.33	4.44	3.93	4.13	4.40	4.14		4.60
76	Pavement Forecast - Percent Probability of Frost	3.44	4.21	4.00	3.64	3.93	4.13	3.89		4.33

PC ID	PRODUCT COMPONENT	Accuracy / Precision	Completeness	Relevance	Currency / Latency	Timeliness / Reliability	Ease of Use	Average Composite Attribute Score		Importance
77	Pavement Forecast - Chemical Concentration	2.90	3.14	3.70	3.44	4.13	4.00	3.55		4.09
78	Road Weather Alerts - Alert parameters	3.77	3.85	4.00	4.08	3.77	3.85	3.88		4.54
79	Road Weather Alerts - Road closure	3.89	3.78	4.11	4.00	3.78	4.11	3.94		4.78
80	Watches and Warnings - Severe weather watches	3.77	4.27	4.05	4.14	4.14	4.24	4.10		4.38
81	Watches and Warnings - Severe weather warnings	4.00	4.17	4.22	4.18	4.23	4.09	4.15		4.52
82	Watches and Warnings - Special weather statements	3.90	4.19	3.95	4.10	4.15	4.20	4.08		4.30
83	MDSS - Treatment recommendations	3.50	3.90	4.20	3.70	4.00	3.80	3.85		4.30
84	Road Condition Report - Road condition	3.60	3.53	3.87	3.20	3.29	3.64	3.52		4.14
85	Flood Warning - Current flood stage	3.67	3.89	3.67	3.67	4.00	4.11	3.83		4.22
86	Flood Warning - Forecasted flood stage	3.33	3.78	3.56	3.44	4.00	4.00	3.69		4.00
87	Camera - View of road	4.18	3.86	4.19	4.04	4.04	4.21	4.08		4.63
88	Camera - View of weather	4.07	3.78	4.04	3.96	3.96	4.15	3.99		4.50
89	Radar - Radar loop	3.90	4.05	4.10	4.29	4.29	4.10	4.12		4.40
90	Radar - Radar loop with precip type coloring	3.71	4.00	4.19	4.29	4.29	4.00	4.08		4.33
91	Radar - Future radar	3.16	3.47	3.84	4.05	4.16	3.79	3.75		3.90
92	Radar - Storm tracks	3.59	3.82	4.12	4.12	4.06	4.00	3.95		4.16
	QUALITY ATTRIBUTE MEAN	3.71	4.01	4.03	3.84	4.00	4.04	3.94		4.25
	STANDARD DEVIATION	0.85	0.86	0.92	0.98	0.86	0.89	0.90		0.92

Table 9. Statistical analysis summary for all *Product Components*

Statistic	Quality Attribute								
	Accuracy / Precision	Completeness	Relevance	Currency / Latency	Timeliness / Reliability	Ease of Use	Average Composite Attribute Score		Importance
Mean of QA Averages	3.71	4.01	4.03	3.84	4.00	4.04	3.94		4.25
Standard Deviation	0.85	0.86	0.92	0.98	0.86	0.89	0.90		0.92
Maximum Product Component Average	4.21	4.50	4.60	4.29	4.38	4.67	4.36		5.00
Minimum Product Component Average	1.83	2.68	3.00	3.20	3.29	2.57	2.97		3.37
Median Product Component Average	3.79	4.12	4.04	3.83	4.10	4.00	3.95		4.26
Skewness of QA Averages	-1.89	-1.20	-0.40	-0.49	-0.62	-1.05	-1.10		-0.15

Table 8 and Table 9 indicate respondents rate the quality of road weather information as high (Likert score of 4) or just below high. The quality attributes of Completeness, Relevance, Timeliness/Reliability, and Ease of Use had means right at 4 whereas Accuracy/Precision and Currency/Latency were rated a bit lower with means of 3.71 and 3.84 respectively. The variance statistics associated with the quality scores for all attributes point to a range of responses from moderate quality to about midway between high and very high quality, although the Accuracy/Precision attribute had a couple of Product Component averages near or below the low quality score. The response distributions were not strictly normal but were skewed toward the lower end of the distribution. The range of scores supports this skewed distribution with all but two of the quality attributes having minimum quality average values below moderate (3.0), and all categories having maximum values around 4.5 even though the means and medians were close to 4.0.

The lower assessment of Accuracy/Precision duplicates the results obtained in the 2008 survey, but the lower Currency/Latency averages in 2010 were not evident in 2008. Rather, Timeliness/Reliability was the other quality attribute in 2008 that had lower averages than the other quality attributes.

The importance scores averaged higher than the quality attribute scores. The statistics indicate the entire distribution is shifted to higher scores and becomes less skewed toward the lower scores. The lowest importance average was 3.37, which is between a moderate and high score. This score was associated with the relative humidity component in the Weather Summary *Product*. All scores for relative humidity observations and histories (NWS and ESS) ranged from 3.4 to 3.8, indicating relative humidity is likely the least important weather parameter evaluated. At the other end of the importance score spectrum was forecasted precipitation type. All fourteen (14) of the respondents rated the importance of forecasted precipitation type as very high (5.0). The high importance of forecasted precipitation type was closely followed by the forecasted precipitation start and end times, snow rate, and snow accumulation parameters in the pavement forecast product. All of these products had averages over 4.9.

Table 10. The number of valid scores used to compute averages in Table 8

PC ID	PRODUCT COMPONENT	Accuracy / Precision	Completeness	Relevance	Currency / Latency	Timeliness / Reliability	Ease of Use	Average Composite Attribute Score	Importance
1	Weather Summary - Air Temperature	30	30	30	30	28	29	177	30
2	Weather Summary - Dew Point	24	26	27	26	24	25	152	27
3	Weather Summary - Relative Humidity	23	25	27	26	25	25	151	27
4	Weather Summary - Wind Direction	31	31	31	31	29	30	183	30
5	Weather Summary - Wind Speed	31	31	31	31	29	30	183	30
6	Weather Summary - Weather	31	31	31	31	29	30	183	30
7	Weather Summary - Precip Type	31	31	30	31	29	30	182	30
8	Weather Summary - Precip Amount	30	30	30	30	28	29	177	29
9	Weather Summary - Snow Amount	27	28	28	29	27	27	166	28
10	Weather History - Air Temperature	13	12	12	11	11	11	70	11
11	Weather History - Dew Point	12	12	12	11	11	11	69	11
12	Weather History - Relative Humidity	12	12	12	11	11	11	69	11
13	Weather History - Wind Direction	13	12	12	11	11	11	70	11
14	Weather History - Wind Speed	13	12	12	11	11	11	70	11
15	Weather History - Precip Amount	13	12	12	11	11	11	70	11
16	Weather History - Snow Amount	12	11	11	10	10	10	64	11
17	ESS Observations - Air Temperature	24	25	25	25	25	25	149	25
18	ESS Observations - Dew Point	22	24	24	24	24	24	142	24
19	ESS Observations - Relative Humidity	23	25	25	25	25	25	148	24
20	ESS Observations - Wind Direction	24	25	25	25	25	25	149	25
21	ESS Observations - Wind Speed	24	25	25	25	25	25	149	25
22	ESS Observations - Wind Gust	24	25	25	25	25	25	149	24
23	ESS Observations - Precip Type	21	24	23	24	24	23	139	23
24	ESS Observations - Pavement Temperature	24	25	25	25	25	25	149	25
25	ESS Observations - Pavement Condition	23	25	25	25	25	24	147	25
26	ESS Observations - Chemical Concentration	20	22	21	21	21	21	126	21
27	ESS Observations - Freeze Point Temperature	20	21	21	20	20	21	123	21
28	ESS Histories - Air Temperature	8	9	9	8	7	9	50	10
29	ESS Histories - Dew Point	7	8	8	7	6	8	44	9
30	ESS Histories - Relative Humidity	8	9	9	8	7	9	50	10
31	ESS Histories - Wind Direction	8	9	9	8	7	9	50	10
32	ESS Histories - Wind Speed	8	9	9	8	7	9	50	10
33	ESS Histories - Precip Type	8	9	8	7	7	9	48	10
34	ESS Histories - Precip Start	8	8	8	8	7	8	47	10
35	ESS Histories - Precip End	8	8	8	8	7	8	47	10
36	ESS Histories - Pavement Temperature	8	9	9	8	7	9	50	10
37	ESS Histories - Pavement Condition	8	8	9	8	7	9	49	10
38	ESS Histories - Chemical Concentration	6	6	7	6	5	7	37	8

PC ID	PRODUCT COMPONENT	Accuracy / Precision	Completeness	Relevance	Currency / Latency	Timeliness / Reliability	Ease of Use	Average Composite Attribute Score		Importance
39	ESS Histories - Freeze Point Temperature	5	6	6	5	4	6	32		8
40	Map - Air Temperature	22	22	22	20	21	21	128		21
41	Map - Dew Point	20	20	20	20	21	20	121		19
42	Map - Relative Humidity	21	21	21	20	21	20	124		20
43	Map - Wind Direction	22	22	22	20	21	21	128		21
44	Map - Wind Speed	22	22	22	20	21	21	128		21
45	Map - Precip Type	21	21	21	20	21	20	124		21
46	Map - Pavement Temperature	19	20	20	19	19	19	116		21
47	Map - Pavement Condition	17	17	18	18	18	17	105		21
48	Map - Chemical Concentration	13	14	12	14	13	13	79		16
49	Map - Freeze Point Temperature	15	15	14	16	15	15	90		17
50	Zone Forecast - Maximum Temperature	12	12	12	12	11	12	71		12
51	Zone Forecast - Minimum Temperature	12	12	12	12	11	12	71		12
52	Zone Forecast - Wind Direction	12	12	12	12	11	12	71		12
53	Zone Forecast - Wind Speed	12	12	12	12	11	12	71		12
54	Zone Forecast - Weather	12	12	12	12	11	12	71		12
55	Zone Forecast - Probability of Precipitation	12	12	12	12	11	12	71		12
56	Pavement Forecast - Air Temperature	15	14	15	14	14	15	87		14
57	Pavement Forecast - Dew Point	14	14	15	14	14	14	85		14
58	Pavement Forecast - Relative Humidity	14	14	15	14	14	14	85		14
59	Pavement Forecast - Wind Direction	15	14	15	14	14	15	87		14
60	Pavement Forecast - Wind Speed	15	14	15	14	14	15	87		14
61	Pavement Forecast - Wind gust	15	14	15	14	14	15	87		14
62	Pavement Forecast - Weather	15	14	15	14	14	15	87		13
63	Pavement Forecast - Visibility	14	13	14	13	13	14	81		13
64	Pavement Forecast - Cloud Cover	13	13	14	13	12	14	79		14
65	Pavement Forecast - Precipitation Type	16	15	15	14	14	14	88		14
66	Pavement Forecast - Precipitation Start Time	15	14	15	14	14	14	86		14
67	Pavement Forecast - Precipitation End Time	15	14	15	14	14	14	86		14
68	Pavement Forecast - Probability of Precipitation	16	15	15	14	14	14	88		14
69	Pavement Forecast - Probability of Precipitation Type	16	15	15	14	14	14	88		14
70	Pavement Forecast - Precipitation rate	15	14	15	14	14	14	86		14
71	Pavement Forecast - Precipitation accumulation	15	14	15	14	14	14	86		14
72	Pavement Forecast - Snow Rate	15	14	15	14	13	14	85		13
73	Pavement Forecast - Snow Amount	15	15	15	14	13	14	86		14
74	Pavement Forecast - Pavement Temperature	16	15	16	15	14	15	91		15
75	Pavement Forecast - Pavement Condition	15	15	16	15	15	15	91		15
76	Pavement Forecast - Percent Probability of Frost	16	14	15	14	15	15	89		15
77	Pavement Forecast - Chemical Concentration	10	7	10	9	8	9	53		11

PC ID	PRODUCT COMPONENT	Accuracy / Precision	Completeness	Relevance	Currency / Latency	Timeliness / Reliability	Ease of Use	Average Composite Attribute Score		Importance
78	Road Weather Alerts - Alert parameters	13	13	13	13	13	13	78		13
79	Road Weather Alerts - Road closure	9	9	9	9	9	9	54		9
80	Watches and Warnings - Severe weather watches	22	22	22	21	21	21	129		21
81	Watches and Warnings - Severe weather warnings	23	23	23	22	22	22	135		21
82	Watches and Warnings - Special weather statements	20	21	21	20	20	20	122		20
83	MDSS - Treatment recommendations	10	10	10	10	10	10	60		10
84	Road Condition Report - Road condition	15	15	15	15	14	14	88		14
85	Flood Warning - Current flood stage	9	9	9	9	9	9	54		9
86	Flood Warning - Forecasted flood stage	9	9	9	9	9	9	54		9
87	Camera - View of road	28	28	27	28	28	28	167		27
88	Camera - View of weather	27	27	26	27	27	27	161		26
89	Radar - Radar loop	21	21	21	21	21	21	126		20
90	Radar - Radar loop with precip type coloring	21	21	21	21	21	21	126		21
91	Radar - Future radar	19	19	19	19	19	19	114		20
92	Radar - Storm tracks	17	17	17	17	17	17	102		19

A similar set of questions regarding importance was completed in 2008. The importance question was part of the *Product* section, asking respondents to indicate the level of importance for each component within the *Product* being evaluated. This was a direct corollary to the importance question in the 2010 survey, although it was addressed in a different context in the survey layout. Because the *Products* and *Elements* in the 2008 survey varied between management strategies, the results from 2008 are not a composite from all management classes. However, a set of high and low rated values similar to the *Product Component* importance values in the 2010 survey can be taken from the Treatment survey results in the 2008 survey. The average importance scores of the *Product* components from the 2008 Treatment survey were ranked in descending order from 1 to 44 (with 1 being the highest average importance and 44 the lowest value). The rankings in Table 11 come from Table 20 in the report *Characterization of Road Weather Information*[2]. The low importance for relative humidity in the 2008 survey (39th out of 44 components) correlates with relative humidity being rated as the lowest importance in the 2010 survey. Likewise, most of the parameters listed at the top of the importance list in 2010 survey also received high importance scores in 2008 survey. The one exception was precipitation end time, which was rated next to last in the 2008 survey.

Table 11. Rank of the corollary product components from the 2008 Treatment survey results

Precipitation Type	8	Relative Humidity	39
Precipitation Start Time	2	Snow Rate	10
Precipitation End Time	43	Snow Accumulation	4

[2] Hart, R.D. and Leon F. Osborne, 2008, Characterization of Road Weather Information, unpublished FHWA report under contract DTFH61-06-D-00006

Table 8 through Table 10 provide the raw information necessary to evaluate the quality of road weather information based upon the *Product Component* questions asked during the 2010 survey. In particular, Table 8 contains the averages of all of the *Product Components* for each of the attributes and the composite average of the attributes. Close scrutiny of this table of means shows there are a few clusters of higher and/or lower mean values that do stand out. These clusters are typically associated with a particular *Product*. For example, the composite averages for pavement forecasts, watches and warnings, camera images, and current radar reflect higher quality scores than the other *Products*. On the opposite side, historical information and anything dealing with chemical concentration received lower quality values. Another interesting relationship occurs in conjunction with the comparison of the high and low *Product Component* importance values. Observations and history records of precipitation parameters received importance scores around 4.5; however, the composite quality attribute score of these same parameters fell in the range of 3.6 to 3.8 and the quality scores under the accuracy/precision attribute were in the 3.0 to 3.5 range. This implies user expectations regarding observed and archived precipitation information are apparently not being met. These examples are rather obvious patterns appearing in the data, but it is difficult to visualize all of the means and comprehend distinctive patterns in the averages. Ranking the means within each of the seven attribute categories and for the composite average of all quality attributes helps highlight patterns. However, before addressing the ranking process, it is important to look at the number of answers that went into computing each of the means in Table 8.

The number of valid scores used to compute each of the quality attribute means is shown in Table 10. The number of valid scores should be close to the number of respondents; however, individual survey participants did enter scores for some attributes but not for others. The differences in valid scores are apparent when viewing the numbers in the first six columns after the *Product Component* names in Table 10. The 'Average Composite Attribute Score' column contains the total number of valid scores entered in the computation of the composite average and should be close to six times the number in any of the attribute columns. The number of scores used to compute the statistics in Table 8 and Table 9 are an indicator of the sample size used to perform the analytical computations.

Since the number of respondents to each *Product Component* question is a composite of the responses from individuals in all three management strategy classes, it is reasonable that the number of *Product Component* scores should vary across *Products*. The average scores suggest most respondents use NWS and ESS observations, maps that contain these observed values, watches and warning messages, radar imagery, and camera views. A significantly smaller group of respondents answered the questions on ESS histories, road weather alerts, MDSS, and flood warnings. The usage level of the remaining *Products* fell between these two groups. Seventy (70) percent of the responses in the pavement forecast *Product* section came from individuals in the Treatment management strategy group, and nearly all the responses for the MDSS question were in the Treatment group. Individuals from the Advisory management strategy group were the primary source of the answers for responses in the zone forecast, road weather alerts, and flood warning product sections. The rest of the *Product* responses were derived from answers in all three of the management strategy categories.

The ranking of the mean scores in Table 8 seems to provide a more reasonable method to visually organize these average values. The rankings permit relationships in the level of quality and importance to become more obvious visually. This visual representation approach is applied in Table 12. The averages in each of the columns in Table 8 were ranked in descending order and assigned ranking order from 1 to 92, with 1 being the highest average and 92 being the lowest. The rankings were then placed back into the *Product Component* listing order used in the previous tables. This created a scrambled list of rankings. To make sense out of this presentation, the means with Likert scores in the top quartile of each Quality Attribute-ranked-list (rankings 1 – 23) were highlighted by filling the background for those cell positions in the table with green. Similarly, the second quartile entries (rankings 24 – 46) were highlighted with cyan, the third quartile (47 – 69) with yellow, and the fourth quartile (70 – 92) with orange. This permits the reader to see which *Product Component* means aggregated at the top, bottom, and middle of the list of mean values.

Visual inspection of Table 12 permits color patterns associated with this quartile ranking scheme to appear. By looking at the green highlighted cells in the table, it is possible to pick out a number of ranking groups that have mean values in the top quartile. They include:

- Pavement Forecast – nearly all attributes
- Camera – Accuracy and Currency/Latency
- Road Weather Alerts – Currency/Latency
- Watches and Warnings – Currency/Latency
- ESS Observations (Weather Data only) – Accuracy
- ESS Histories (Weather Data only) – Reliability

Ranking groups with averages in the last quartile include:

- Weather Summary – Completeness (precipitation factors) and Reliability
- Weather History – Relevance and Reliability
- ESS Observations – Accuracy (pavement parameters), Completeness (pavement parameters), Currency/Latency, Reliability, and Ease of Use
- ESS Histories – most attributes, especially related to precipitation and pavement information
- Zone Forecast – Accuracy, Relevance, and Reliability
- MDSS – Accuracy, Currency, Ease of Use
- Road Conditions Report – Completeness, Currency, Reliability, and Ease of Use
- Flood Warning – nearly all attributes
- Camera – Completeness
- Radar (future radar) – Accuracy, Completeness, and Ease of Use

To further simplify the analysis, the rankings for each of the *Product Components* in a specific *Product* in Table 12 were transferred into Table 13 that describes the ranking range within that specific *Product*. For example, if the attribute rankings of all of the *Product Components* in that *Product* were high (i.e., ranking of 1-23), the attribute value for that cell would be defined as high. Likewise, if the attribute values were in the second quartile (rankings of 24-46), the Product cell for that attribute was marked Med/High. Following this logic, *Product Components* having predominantly third quartile ratings (47-69) were designated as Med/Low. Finally, if all the *Product Component* rankings were in the last quartile (rankings of 70-92) the cell was filled with Low. A number of the individual *Product Component* rankings for a given product clustered around the median ranking, so a separate classification was developed for *Product* /Attribute cells that had a number of rankings in quartiles 2 and 3. This special classification was called "Medium." Finally, several of the individual *Product Components* in a given *Product* were scattered over two or more quartiles. To describe these distribution patterns, the following classes were also specified: High-Med/High, Medium-Low, and Med/Low-Low. To visually differentiate the names, the cells containing the *Product*/Attribute ranks were highlighted with unique colors for each category. The category names and their respective colors are:

High:	Green
High – Med/High:	Olive green
Med/High:	Light green
Medium:	Blue
Medium –Low:	Yellow
Med/Low:	Light yellow
Med/Low – Low:	Orange
Low:	Red

Several of the *Product*s had *Product Components* that had a set of *Product Components* rankings that were distinctly different from the rest of the components in their group. For example, in the Weather Summary product the weather components such as air temperature, relative humidity, and wind information received high quality scores while the weather type and precipitation parameters had a number of low ranking scores. To fit this disparity into the table, the distinctly different categories within a *Product* were separated and the ranking classification was done separately.

Table 12. Quality attribute means from Table 8 ranked from the highest value (1) to lowest (92)

RANK: 1-23 | 24-46 | 47-69 | 70-92

PRODUCT COMPONENT	Accuracy / Precision	Completeness	Relevance	Currency / Latency	Timeliness / Reliability	Ease of Use	Average Composite Attribute Score	Importance
Weather Summary - Air Temperature	15	27	21	22	63	30	24	53
Weather Summary - Dew Point	30	46	68	33	70	78	56	85
Weather Summary - Relative Humidity	47	41	84	34	64	82	65	92
Weather Summary - Wind Direction	14	34	44	36	66	41	35	79
Weather Summary - Wind Speed	16	35	23	24	67	35	31	41
Weather Summary - Weather	56	75	48	56	77	54	67	29
Weather Summary - Precip Type	48	70	22	57	80	64	60	19
Weather Summary - Precip Amount	74	79	51	75	81	73	79	16
Weather Summary - Snow Amount	85	80	33	54	82	66	77	18
Weather History - Air Temperature	41	38	62	50	83	61	58	46
Weather History - Dew Point	25	39	71	69	84	75	66	63
Weather History - Relative Humidity	43	47	87	70	88	76	76	84
Weather History - Wind Direction	17	40	72	51	85	62	59	34
Weather History - Wind Speed	42	48	66	52	86	63	62	35
Weather History - Precip Amount	82	83	73	87	91	77	88	14
Weather History - Snow Amount	86	81	81	85	89	65	87	10
ESS Observations - Air Temperature	4	18	32	53	51	38	27	59
ESS Observations - Dew Point	11	49	25	55	61	71	41	67
ESS Observations - Relative Humidity	12	42	58	58	52	70	46	91
ESS Observations - Wind Direction	1	43	37	71	53	56	34	72
ESS Observations - Wind Speed	6	44	18	59	54	42	32	56
ESS Observations - Wind Gust	13	45	38	60	60	43	37	68
ESS Observations - Precip Type	75	86	46	86	79	74	81	24
ESS Observations - Pavement Temperature	51	61	16	72	65	39	51	25
ESS Observations - Pavement Condition	84	87	80	74	78	80	84	31
ESS Observations - Chemical Concentration	90	92	91	90	90	89	91	88
ESS Observations - Freeze Point Temperature	89	90	69	78	87	88	89	47
ESS Histories - Air Temperature	31	54	74	37	18	44	49	80
ESS Histories - Dew Point	39	55	67	42	17	45	45	78
ESS Histories - Relative Humidity	32	56	75	38	19	46	50	81
ESS Histories - Wind Direction	33	57	64	39	20	47	43	75
ESS Histories - Wind Speed	34	58	65	40	21	48	44	76
ESS Histories - Precip Type	61	71	34	10	22	86	53	50

RANK: 1-23, 24-46, 47-69, 70-92

PRODUCT COMPONENT	Accuracy / Precision	Completeness	Relevance	Currency / Latency	Timeliness / Reliability	Ease of Use	Average Composite Attribute Score		Importance
ESS Histories - Precip Start	78	77	35	81	23	83	71		60
ESS Histories - Precip End	79	78	36	82	24	84	72		61
ESS Histories - Pavement Temperature	68	63	40	41	25	67	52		51
ESS Histories - Pavement Condition	88	65	76	83	26	90	86		82
ESS Histories - Chemical Concentration	92	88	92	91	34	92	92		89
ESS Histories - Freeze Point Temperature	91	89	88	84	29	91	90		90
Map - Air Temperature	26	29	52	61	35	26	33		62
Map - Dew Point	40	33	54	62	46	28	38		69
Map - Relative Humidity	44	26	60	63	47	24	40		86
Map - Wind Direction	22	30	30	64	38	21	23		70
Map - Wind Speed	23	31	31	65	39	27	29		52
Map - Precip Type	57	52	17	43	49	55	42		44
Map - Pavement Temperature	46	51	20	68	45	34	36		36
Map - Pavement Condition	67	76	55	45	44	40	57		45
Map - Chemical Concentration	83	84	89	79	55	68	82		87
Map - Freeze Point Temperature	71	66	63	66	41	49	63		66
Zone Forecast - Maximum Temperature	58	22	90	46	73	57	69		83
Zone Forecast - Minimum Temperature	66	23	85	47	74	58	68		74
Zone Forecast - Wind Direction	73	24	86	67	68	59	73		54
Zone Forecast - Wind Speed	69	25	82	76	69	60	70		55
Zone Forecast - Weather	76	50	77	48	75	72	74		30
Zone Forecast - Probability of Precipitation	80	59	78	49	76	85	78		37
Pavement Forecast - Air Temperature	5	1	26	27	3	1	4		21
Pavement Forecast - Dew Point	18	10	19	28	4	8	12		22
Pavement Forecast - Relative Humidity	19	11	43	29	5	9	16		65
Pavement Forecast - Wind Direction	2	2	11	11	6	2	1		23
Pavement Forecast - Wind Speed	9	3	5	12	7	3	2		12
Pavement Forecast - Wind gust	38	4	15	30	8	4	6		49
Pavement Forecast - Weather	10	5	6	13	9	5	3		17
Pavement Forecast - Visibility	54	17	56	44	31	14	30		40
Pavement Forecast - Cloud Cover	7	28	57	35	30	15	20		73
Pavement Forecast - Precipitation Type	20	20	3	14	10	6	5		1
Pavement Forecast - Precipitation Start Time	52	6	1	31	11	10	7		2
Pavement Forecast - Precipitation End Time	59	7	2	32	12	11	9		3

RANK: 1-23 (light green), 24-46 (light blue), 47-69 (yellow), 70-92 (orange)

PRODUCT COMPONENT	Accuracy / Precision	Completeness	Relevance	Currency / Latency	Timeliness / Reliability	Ease of Use	Average Composite Attribute Score	Importance
Pavement Forecast - Probability of Precipitation	35	8	7	15	13	19	11	8
Pavement Forecast - Probability of Precipitation Type	36	15	12	16	14	20	14	13
Pavement Forecast - Precipitation rate	62	12	13	17	15	12	17	11
Pavement Forecast - Precipitation accumulation	24	13	8	18	16	16	8	7
Pavement Forecast - Snow Rate	45	14	9	19	1	17	13	5
Pavement Forecast - Snow Amount	53	21	14	20	2	13	15	4
Pavement Forecast - Pavement Temperature	37	9	4	25	33	7	10	6
Pavement Forecast - Pavement Condition	63	16	10	26	42	18	19	20
Pavement Forecast - Percent Probability of Frost	72	32	49	80	62	31	54	38
Pavement Forecast - Chemical Concentration	87	91	70	88	43	50	83	64
Road Weather Alerts - Alert parameters	50	68	50	7	72	69	55	26
Road Weather Alerts - Road closure	29	72	41	21	71	32	48	9
Watches and Warnings - Severe weather watches	49	19	45	4	40	22	22	33
Watches and Warnings - Severe weather warnings	21	37	24	3	32	37	18	27
Watches and Warnings - Special weather statements	28	36	53	6	37	25	26	42
MDSS - Treatment recommendations	70	62	27	73	56	79	61	43
Road Condition Report - Road condition	64	82	59	92	92	87	85	58
Flood Warning - Current flood stage	60	64	79	77	57	33	64	48
Flood Warning - Forecasted flood stage	77	73	83	89	58	51	80	71
Camera - View of road	3	67	29	9	50	23	25	15
Camera - View of weather	8	74	47	23	59	29	39	28
Radar - Radar loop	27	53	42	1	27	36	21	32
Radar - Radar loop with precip type coloring	55	60	28	2	28	52	28	39
Radar - Future radar	81	85	61	8	36	81	75	77
Radar - Storm tracks	65	69	39	5	48	53	47	57
QUALITY ATTRIBUTE MEAN	3.71	4.01	4.03	3.84	4.00	4.04	3.94	4.25
STANDARD DEVIATION	0.85	0.86	0.92	0.98	0.86	0.89	0.90	0.92

Table 13. *Product* ranking ranges

Product		QUALITY ATTRIBUTE					
		Accuracy	Completeness	Relevance	Currency	Reliability	Ease of Use
Weather Summary	Air Temp, RH, winds	High-Med/High	Med/High	Medium	Med/High	Med/Low	Medium
	Weather and Precipitation	Low	Low	Medium	Low	Med/Low-Low	Medium-Low
Weather History	Air Temp, RH, winds	Med/High	Medium	Med/Low-Low	Med/Low-Low	Low	Med/Low-Low
	Precip and Snow Amounts	Low	Low	Med/Low-Low	Med/Low-Low	Low	Med/Low-Low
ESS Current Conditions	Air Temp, RH, winds	High	Med/High	Medium	Low	Med/Low-Low	Med/Low-Low
	Precip Type and Pavement Info	Low	Low	Low	Low	Med/Low-Low	Med/Low-Low
ESS Histories	Air Temp, RH, winds	Med/High	Med/Low-Low	Medium-Low	Medium-Low	High-Med/High	Medium-Low
	Precip Type and Pavement Info	Low	Med/Low-Low	Medium-Low	Medium-Low	High-Med/High	Medium-Low
Regional Map	Air Temp, RH, winds	Medium	High-Med/High	Medium	Med/Low	Medium	Medium
	Precip Type and Pavement Info	Medium	Med/Low	Medium	Med/Low	Medium	Medium
Zone Forecast	Air Temp, RH, winds	Med/Low-Low	Medium	Low	Medium-Low	Medium-Low	Medium-Low
	Weather and Precipitation	Med/Low-Low	Med/Low	Low	Medium-Low	Medium-Low	Medium-Low
Pavement Weather Forecast	All except frost & chem conc	High-Med/High	High	High	High	High	High
	Frost & Chemical Concentration	Low	Medium-Low	Med/Low-Low	Low	Med/Low	Medium
Road Weather Alerts		Medium	Medium-Low	Medium	High	Low	Medium
Watches & Warnings		Medium	High-Med/High	Medium	High	Medium	High-Med/High
MDSS		Low	Med/Low	Med/High	Low	Med/Low	Low
Road Condition Report		Med/Low	Low	Med/Low	Low	Low	Low
Flood Warning		Medium-Low	Medium-Low	Medium-Low	Medium-Low	Medium-Low	Medium-Low
Camera Images		High	Med/Low-Low	Medium	High	Medium	High-Med/High
Radar		Medium	Medium-Low	Medium	Medium	Medium	Medium

It is imperative to state that the ranking categories reflect the position on the ranking scales of the attribute averages. They do not reference the quality level of the Product Components, per se. The average of all quality attribute average scores was near the Likert score of 4 (related to the survey rating HIGH) and the range of the averages extended from 3 (MEDIUM) to roughly 4.5 (midway between VERY HIGH and HIGH). Therefore, rankings that are considered low in Table 13 or at the bottom of quartile 4 in Table 12 reflect quality average scores that lie above medium in the survey answer options.

3.3 Interpretation of Product Component Results

Table 8 through Table 13 organize the results from the survey questions on *Product Components* in several different formats permitting a more thorough interpretation of the quality and importance assessments made by the DOT participants. Since the structure of the survey was organized around road weather *Products*, this report uses this organizational structure to review the results of the *Product Components*. This approach permits the evaluation to focus on the influence of the four key factors defining the character of the information used by DOT users in support of their ultimate transportation-related operational decisions. The four factors defining each *Product Component* are road weather parameter (called *Element* in the 2008 survey), its source or provider, its timeframe, and its spatial representation. The results in the tables reflect how DOT participants rate each *Product Component*, but more specifically how the participants respond to the influence of these factors on the different forms of road weather information.

The evaluation method is to list the *Product* name; specify the source, time frame, and spatial format; and list the components (road weather parameters) in the *Product*. This will be followed by a discussion of the quality and importance results associated with the *Product Components* with each *Product*.

Weather Summary

Source: National Weather Service
Time Frame: Current
Spatial Format: Single Site
Components:

1. Air temperature
2. Dew point
3. Relative humidity
4. Wind direction
5. Wind speed
6. Weather type
7. Precipitation type
8. Precipitation amount
9. Snow amount

The precipitation components in the components list (6 – 9) were rated as high in importance while the components 2 – 4 were considered on the low end of the importance rating scale. However, the quality scores were reversed with the four weather type/precipitation components being ranked in the lowest two quartiles and air temperature, dew point, and wind data) in the middle to upper portion of the quality rankings.

Under the Currency/Latency attribute, all of the *Product Components* had averages around 3.75 placing these attribute scores in the low to med/low range. In addition, the quality of the Precipitation and Snow Amount *Product Components* were rated low for all attributes except Relevance.

The *Product Component* means in the Currency/Latency quality attribute and their associated rankings indicate a dichotomy between the precipitation parameters and the other weather parameters. This may be an indication that survey participants feel the NWS does not update precipitation-related information often enough. Typically, the NWS updates the air temperature, dew point, RH, and wind data at least once every 20 minutes. However, the precipitation data is typically only updated once an hour. The low ratings may serve as an indication that DOT users would like to see precipitation data updated routinely with the other weather information. The two tiers in the Currency/Latency quality may also be an indication of a biasing factor that was noted in the 2008 survey. Generally, the National Weather Service issues its weather summary bulletins once an hour. The summaries contain all nine of the components in the list above. It would seem intuitive that if all parameters were updated at the same time, the average currency/latency scores of all *Product Components* within the weather summary should be roughly the same. The fact those *Product Components* that received low scores in other attributes also received low scores in the Currency/Latency attribute, and those that received high scores in other attributes also received high scores in Currency/Latency, suggests there is a tendency for survey participants to assess an overall value of a given *Participant Component* and permit this value assessment to influence quality attribute scores across all attributes.

GENERAL SUMMARY: The Weather Summary components were evaluated as being in the moderate to low range of the quality ranks. Weather Summaries suffered because the key components of interest were considered lower in accuracy and not as current as desired by the DOT personnel.

Weather History
Source: National Weather Service
Time Frame: Past
Spatial Format: Single Site
Components:

1. Air temperature
2. Dew point
3. Relative humidity
4. Wind direction
5. Wind speed
6. Precipitation amount
7. Snow amount

As with the Weather Summary *Product*, the two precipitation elements in the list of History components were important to the users while the other elements were not as important, especially relative humidity. Other than the medium/high rankings in the Accuracy/Precision attribute column for the top five weather elements in the components list, all *Product Components* in this *Product* received quality ratings that were in the lower half of all ratings. Of particular note are the averages under the Timeliness/Reliability attribute. They are all near the lower end of the distribution of averages for all Product Components, and all are nearly one standard deviation below the mean value of all Timeliness/Reliability attribute averages. It would seem archived history data stored by the NWS would be a reliable resource and as timely as any other resource. ESS History components did not receive similar low scores for the Timeliness/Reliability attribute; therefore, DOT users may be indicating it is time consuming to dig out the historical data they need from NWS resources. In general, Weather History components had quality ratings lower than most of the other *Product Components* for all attributes except isolated components in the Accuracy/Precision and Completeness attributes.

GENERAL SUMMARY: DOT personnel found the NWS history information of less value than the information in other *Products* across all quality attributes. The low scores under the Currency/Latency and Timeliness/Reliability attributes seem to support the sense that DOT users find the NWS historical information difficult to access or possibly of limited value in meeting their operational requirements.

ESS Observations
Source: DOT
Time Frame: Current
Spatial Format: Single Site
Components:

1. Air temperature
2. Dew point
3. Relative humidity
4. Wind direction
5. Wind speed
6. Wind gust
7. Precipitation type
8. Pavement temperature
9. Pavement condition
10. Chemical concentration
11. Freeze point temperature

Other than precipitation type, pavement temperature, and pavement condition all of the components had importance scores in the lower half of the importance rankings with three of these in the lowest quartile. The Accuracy/Precision of the weather components (1 – 6) were considered high (near the top of the rankings), but the accuracy of the last five components (7 - 11) were rated in the lowest quartile with chemical concentration near the bottom of the entire list. Chemical concentration received either the lowest quality score or very close to it for all six of the quality attributes with half of its average scores being less than 3.0. Pavement temperature that has been the basis of the Road Weather Information System (RWIS) was considered as one of the more relevant parameters but it had an Accuracy/Precision average score of 3.5 that put it 0.2 Likert units below the average Accuracy/Precision score of 3.71. This placed pavement temperature at a ranking of 51 out of the 92 components for accuracy.

The average score rankings in the Currency/Latency and Timeliness/Reliability attributes were all in the third and fourth quartiles. These results suggest DOT users have concerns about the time it takes to collect the data from the field and make it available to the users. The Reliability attribute may also indicate DOT users see issues with the reliability of the data from ESS sites. Finally, the rankings in the lower half of the distribution for the Ease of Use attribute indicate either DOT users find the user interface more difficult to use than what they experience for other *Products,* or the information contained within the ESS observations is hard for them to assimilate and use. The fact that the Pavement Condition, Chemical Concentration, and Freeze Point Temperature components are at the bottom end of the *Product Component* averages may be a good indication that the Ease of Use attribute indicates a difficulty to comprehend ESS data. These three parameters are more abstract representations of road weather conditions than more concrete metrics such as air temperature. Moreover, the existing measurement tools have performance limitations making the data suspect at times and require significant background knowledge to utilize effectively.

GENERAL SUMMARY: ESS Observations have a lower level of importance to the DOT respondents than many of the other resources in this Baselining Study, and are viewed as having marginal quality to the users. This finding from the survey results needs to be highlighted within the community and solutions need to be addressed to resolve the apparent limitations with RWIS. DOT personnel have been the primary users of the ESS data and their concerns need to be recognized. This needs to be a concern of the *Clarus* program and the intended integration of the ESS data into the larger network of weather data sources. It appears more attention is needed regarding the validity of the observations, their timeliness and reliability, and the way the data are presented to the end users.

ESS Histories

Source: DOT
Time Frame: Past
Spatial Format: Single Site
Components:

1. Air temperature
2. Dew point
3. Relative humidity
4. Wind direction
5. Wind speed
6. Precipitation type
7. Precipitation start time
8. Precipitation end time
9. Pavement temperature
10. Pavement condition
11. Chemical concentration
12. Freeze point temperature

All of the components in this list were considered to have medium-low to low importance and, in general, the top five components in the list had quality rankings that were in the top two quartiles while the precipitation and pavement components had rankings in the bottom two quartiles. Interestingly, all of the components received above average quality scores in the Timeliness/Reliability attribute ranking even though the ESS observations had scores for the same attribute that were near the bottom of the ranking for that attribute. The weather parameters in the Accuracy/Precision attribute had quality averages that were above the attribute mean but in quartile 2. This perception of accuracy in the historical weather parameters was above the median but not as high in the ranking scale as the same components in the ESS Observations *Product*. The road-related components received quality scores near or at the bottom of the Accuracy distribution. Elsewhere, all of the *Product Components* received lower than average quality rankings in the Currency/Latency, Timeliness/Reliability, and Ease of Use attributes.

GENERAL SUMMARY: Histories were considered one of the less important tools and generally received quality rankings that were below the median values.

Map Display
Source: Surface Transportation Weather Service Provider
Time Frame: Current
Spatial Format: Regional display of multiple sites
Components:

1. Air temperature
2. Dew point
3. Relative humidity
4. Wind direction
5. Wind speed
6. Precipitation type
7. Pavement temperature
8. Pavement condition
9. Chemical concentration
10. Freeze point temperature

Most of the map components had importance scores that clustered around the mean averages of all *Product Component* attributes. The two exceptions were Relative Humidity and Chemical Concentration, which had averages close to the bottom of the importance ranking. In general, all of the quality ratings clustered around the middle of the rating distribution. The two exceptions were the chemical concentration and the freeze point temperature, which received quality scores in the fourth quartile. The averages in the Ease of Use attribute may provide a clue about presentation modes. The Ease of Use averages for the Map Display components are higher than for the single site Weather Summary and ESS Observation *Products*. This may indicate that DOT personnel find more use and/or value in the spatial display of data rather than receiving data as single site data or bulletins with textual lists of data.

GENERAL SUMMARY: Map displays were conspicuously a middle of the road *Product* in both importance and quality attribute scores.

Zone Forecast
Source: National Weather Service
Time Frame: Future
Spatial Format: Single Site
Components:

1. Maximum air temperature
2. Minimum air temperature
3. Wind direction
4. Wind speed
5. Weather type
6. Probability of precipitation

The Weather Type and Probability of Precipitation components were rated as having above average (quartile 2) importance; however, the remaining parameters were of below average with maximum and minimum temperature forecasts falling in the fourth quartile. Nearly all of the quality rankings were below average with many in the lower quadrant. An obvious pattern that appears in Table 12 and Table 13 is the Relevance attribute scores near the bottom of the rankings for that attribute. The Timeliness/Reliability averages fall along the boundary between the third and fourth quartiles. The author's experience has been that the NWS delivers the zone forecasts on a regular schedule and on time with few exceptions; therefore, the lower ratings in this attribute may suggest that Zone or Regional Forecasts are not updated frequently enough to meet the requirements of DOT users. The Timeliness/Reliability scores in the following Pavement Forecast section seem to indicate that the more frequent updates in that *Product* are favored by the DOT.

The results in the Ease of Use column were not totally expected and raise questions regarding the interpretation of the term 'Ease of Use'. Table 5 defines Ease of Use as "the facility to get, interpret, and use the information." The NWS Zone Forecast format has been around for years and seems well accepted. The forecasts are easy to get and reasonably easy to interpret. The implication in the Zone Forecast scores is DOT users do not find the information as easy to use to support their decision process as they do with other tools.

GENERAL SUMMARY: Zone Forecasts provide information in a format that is not especially relevant to the operational needs of the DOT and the subjective quality (value) of Zone Forecasts suffers.

Pavement Forecast

Source: Surface Transportation Weather Service Provider
Time Frame: Future
Spatial Format: Single Site
Components:

1. Air temperature
2. Dew point
3. Relative humidity
4. Wind direction
5. Wind speed
6. Wind gust
7. Weather type
8. Visibility
9. Cloud cover
10. Precipitation type
11. Precipitation start time
12. Precipitation end time
13. Probability of precipitation
14. Probability of precipitation type
15. Precipitation rate
16. Precipitation accumulation
17. Snow rate
18. Snow amount
19. Pavement temperature

20. Pavement condition
21. Percent probability of frost
22. Chemical concentration

All of the components in this *Product* were rated in the top quartile of the importance attribute averages with the exception of cloud cover, relative humidity, chemical concentration, and wind gusts. Nearly all of the quality scores were above the median, and most scores fell in the top quartile in the rankings for five of the six attributes. The two exceptions were the probability of precipitation and chemical concentration components. The one attribute with a significant number of averages below the median value for that attribute was Accuracy/Precision. Approximately 1/3 of the quality averages for this attribute were less than the median. A closer look indicates DOT respondents perceived the accuracy of the forecasted weather conditions and pavement temperature was in the top quartile of all Accuracy attribute responses. However, the accuracy of the precipitation factors was considered as average (averages in both quartiles 2 and 3) and the forecast of pavement condition, probability of frost, and chemical concentration were rated as being moderate to low in the accuracy ranking scale.

It is important to note all of the precipitation parameters and the pavement temperature in this forecast *Product* were rated in the top ten in the importance rating. These *Product Components* were selected as the most critical pieces of road weather information to serve as resources for DOT decision makers. The attributes dealing with delivery of the information indicated the delivery and display mechanisms are satisfactory; however, accuracy/precision was the only characteristic of the Pavement Forecast data where survey participants saw a need for improvement. This will continue to be an area that needs attention.

GENERAL SUMMARY: The results suggest pavement weather forecasts specific to the transportation-related decision making is the highest priority road weather information tool and that the DOT participants find the quality of the forecast information well above average except in regards to pavement-specific parameters other than pavement temperature.

<u>Road Weather Alerts</u>

Source: Surface Transportation Weather Service Provider and/or DOT
Time Frame: Current
Spatial Format: Single Route or Route Segment
Components:

1. Alert parameters
2. Road closures

Both of these elements were considered important for decision support requirements. However, the quality across all of the attributes fell into an average ranking level. Of particular note is the fact the Completeness and Timeliness/Reliability averages fell into the lowest quartile of the response distribution.

GENERAL SUMMARY: There is a need for road weather alert information, but users are disappointed with the extent of the information and its availability on a timely basis.

Watches and Warnings
Source: National Weather Service
Time Frame: Future
Spatial Format: Map
Components:

1. Severe weather watches
2. Severe weather warnings
3. Special weather statements

The importance scores of all three of these components fell in quartile 2 of the rankings. The quality ratings across all attributes were above the median average and about a third of the averages fell into the top quartile.

GENERAL SUMMARY: Watches and warnings are an important decision support tool that rates second to pavement forecasts in its value to the DOT participants in this study.

Maintenance Decision Support System
Source: Surface Transportation Weather Service Provider
Time Frame: Future
Spatial Format: Single Site
Component:

1. Treatment recommendations

The importance of the treatment recommendation component was in the middle of the response rankings. The quality scores across the attributes were all below the median of the entire list of *Product Component* averages. The only exception was the quality score for Relevance. The 2008 survey asked questions about MDSS as an entire package that included historical, current, and forecasted road weather information. However, in order not to duplicate the questions about the display of the observed and forecasted road weather data available in other *Products*, the MDSS section in the 2010 survey only addressed the treatment recommendations. Therefore, it is important to note that the Maintenance Decision Support System *Product* does not represent the quality of the overall MDSS service, rather only the treatment recommendation portion of MDSS.

GENERAL SUMMARY: MDSS Treatment Recommendations are viewed as a relevant and reasonably important tool but are thus far not perceived as one of the more accurate, timely, reliable, and easy to use resources.

Road Condition Report
Source: DOT
Time Frame: Current
Spatial Format: specific route segment
Component:

1. Road condition

The importance score for road condition reports was just below the median score in the importance rankings and the quality of information scored lower than the quality for most other *Product Components* across all attributes. The quality averages for Currency/Latency and Timeliness/Reliability were both at the bottom of the list in their respective attribute lists and the quality averages for Completeness and Ease of Use were slightly higher in their attribute rankings but still well into the fourth quartile in the rankings.

GENERAL SUMMARY: The scores indicate that Road Condition Reports are considered a tool of average importance by DOT users; however, users see a lot of room for improvement in this resource.

Flood Warning
Source: National Weather Service
Time Frame: Future
Spatial Format: Single Site and Regional
Components:

1. Current flood stage
2. Forecasted flood stage

The scores for all attributes (importance plus the six quality attributes) fell in quartiles 3 and 4.

GENERAL SUMMARY: Flood Warnings are of lesser value to DOT users and the quality of the information provided is lower than with most other road weather information resources.

Camera
Source: DOT
Time Frame: Current
Spatial Format: Single Site
Components:

1. View of road
2. View of weather

The importance scores of the two components fell either side of the separation point between quartiles 1 and 2. Quality scores for Accuracy/Precision and Currency/Latency ranked in the top quartile for each of those attributes with the two Camera components having rankings of 3 and 8 on the Accuracy/Precision quality attribute ranking scale. The rest of the attribute averages clustered around the median averages for each of the attributes. A lower score for View of Weather in the Completeness attribute is probably not a negative statement, but rather an indication that the fixed camera images or the camera stops do not provide views yielding a good assessment of the weather conditions when viewed at a remote location.

GENERAL SUMMARY: The scores indicate that cameras are another tool that serves as a good information resource to support DOT operations and the inference from the Completeness response suggests this tool could be more effective with more or better selection of camera views.

Radar

Source: National Weather Service
Time Frame: Current
Spatial Format: Single Site
Components:

1. Radar loop
2. Radar loop with precipitation type coloring
3. Future radar
4. Storm tracks

The radar loop components 1 and 2 both received importance scores ranking in quartile 2, while the future radar and storm tracks had importance scores in quartiles 3 and 4. The top of the ranking scale responses for the Currency/Latency attribute suggests that radar is perceived as the most readily available source of current precipitation information. The components derived from the primary radar images (items 3 and 4) were not accepted as well. Future radar and storm tracks had quality ratings within each of the attributes that were moderate to low. Future radar in particular had attribute scores that fell into quartile 4.

GENERAL STATEMENT: The radar information scores indicate users see the fundamental radar images as easy to use, timely, and reasonably accurate. Future radar and storm track derived services need to improve to gain an equivalent acceptance to the observed radar products.

3.4 Comparison of Element Results from the 2008 and 2010 Surveys

Product Components have direct relationships to the *Products* and *Elements* used in the 2008 survey. Section 2.1 discussed the relationships between these three basic entities. Each of the *Elements* evaluated in the 2008 survey could be attributed to one or more *Product Component* formats. It was a simple process to take the list of all the *Product Component* forms for each of the *Elements* used in 2008, gather all responses from each of the related *Product Components*, and compute quality averages for each of the six attributes, the composite average, and the importance. Since *Element* importance was not one of the questions asked in the 2008 survey, *Element* importance became a new measurement tool starting with the 2010 survey.

To compare the derived *Element* statistics from 2010 with those in 2008, the 2010 averages had to be separated by primary management strategy class. All the *Product Component* responses for individuals classified as Advisory, Control, or Treatment members were consolidated separately. These values were then used to create a table of quality averages for each of the attributes. The responses from all six attributes were then used to form the composite average. These computations created a set of derived averages for each quality attribute, the composite average, and the importance for all of the *Elements* based upon the input into the 2010 survey. Similar values were then extracted from the results of the 2008 survey. Table 14, Table 15, and Table 16 list the composite average values from both phases of the survey for Advisory, Control, and Treatment respondents, respectively.

Since participants in the 2010 survey could respond to questions on any of the *Product Components*, it became possible to compute an estimated response for all *Elements*. However, the data from the 2008 Survey was limited to only those *Elements* that were in the survey for each of the primary management strategy groups. Therefore, Table 14 through Table 16 contain composite averages for all *Elements* (except where there were no responses) from the 2010 survey and composite averages from the *Elements* defined for that strategy group in the 2008 survey. The list of *Elements* was ordered in each of the tables based upon the composite averages derived from the 2010 survey computations. Comparison of the lists of *Elements* from the three tables shows there is a different ranking scheme for each of the primary management strategy groups, illustrating each group's uniquely different evaluation of road weather information quality.

It would be possible to compare the responses in each of the quality attribute classes, but the number of responses involved in those averages was often too small to support a statistically significant comparison. There were two reasons for the small number of responses. First, the number of respondents who entered quality attribute scores for a number of the *Product Components* was small, sometimes limited to one or two respondents. Second, respondents jumped around in the 2010 survey and chose to only answer questions for certain *Products*. Since the computation of the derived 2010 *Element* averages often came from *Product Components* from 5 or more *Products*, the lack of answers in one or more of these from a given respondent not only decreased the number in the computation of the mean, but also changed the average to reflect the influence of those *Product Components* from which there were scores. The most obvious examples of the limited sample set can be seen in the Control strategy group results

(Table 15) where a number of Elements had no input source or numbers in the single digits. Since the means in Table 15 are composite averages, one respondent would create six valid entries if that individual entered scores for all six attributes. This suggests a number of the averages in the *Element* list on the 2010 side of the comparison came from only one or two respondents. The composite average buffers part of the limited number of respondents. However, the computation of the individual attributes would be more severely affected by the small number of valid scores. Therefore, the analysis was limited to just the composite averages.

The differences in the rightmost column of each comparison table were computed by subtracting the means from the 2008 survey from the means from the 2010 survey. Therefore, positive differences indicate an increase in the average quality values over the two years and negative differences a decrease. The average difference at the bottom of each table was derived from the mean response quality rating results from both the 2008 and the 2010 studies using all quality attribute responses. It is not derived by averaging the differences in the column of figures above it. Likewise, the standard deviation is also computed using all responses. The average differences in each of the three strategies provide the best source for determining whether the changes are statistically significant because of the larger size of the information set compared to the individual *Element* comparisons. The average mean differences for the Advisory, Control, and Treatment strategy results were 0.54, 0.54, and 0.08, respectively. An analysis of variance T-test on these increases indicates they are statistically significant increases at the 95% confidence level.

Although all three tests indicate a significant improvement in the quality scores between 2008 and 2010, the Control statistics remain problematic as an indicator. The seven large positive increases in quality scores at the top of the Control list are all attributable to a single respondent who submitted a very high (Likert score of 5) to all attribute scores for six of the seven entries. If these values were removed there would still be a positive increase in the quality scores, but the magnitude would be closer to 0.3. The Advisory score increases from 2008 to 2010 are consistently greater. This is substantiated by the relatively small standard deviation value associated with the mean difference value. The results from the Treatment strategy support the overall trend of an increase in the *Element* scores; however, this increase was composed of nearly equivalent individual *Element* differences having both positive and negative values. Much of the increase in the quality measure in the Treatment group is attributable to precipitation-related *Elements* while the decrease in quality appears in those *Elements* dealing with chemical concentration and flood-related *Elements*.

Table 14. Comparison of combined Quality attribute average results for *Elements* from Advisory respondents in 2008 and 2010 surveys

ELEMENT	2010 SURVEY		2008 SURVEY		DIFFERENCE
	N	MEAN	MEAN	N	
Camera view of weather	78	4.19			
Camera view of traffic	78	4.19			
Visibility	24	4.17	3.65	51	0.52
Road closure	42	4.17	4.12	41	0.05
Severe weather advisory	144	4.13	3.26	46	0.87
Snow rate	23	4.13			
Treatment recommendation	18	4.11			
Air temperature	246	4.10			
Winter weather advisory	114	4.10	3.38	48	0.72
Dense fog advisory	114	4.10	3.46	46	0.64
Camera view of road	78	4.09			
Cloud cover	23	4.04			
Pavement temperature	150	4.04			
Wind advisory	66	4.03	3.08	50	0.95
Wind speed	293	4.02	3.42	67	0.60
Wind gust	89	4.00			
Precipitation rate	24	4.00			
Precipitation type probability	26	4.00			
Wind direction	293	3.98	3.28	65	0.70
Precipitation start time	40	3.98			
Radar images	168	3.96			
Dew point temperature	204	3.96			
Precipitation end time	40	3.95			
Relative humidity	220	3.91			
Precipitation type	226	3.89			
Minimum temperature	41	3.88	3.46	68	0.42
Weather type	148	3.87	3.22	58	0.65
Flood stage	42	3.86			
Precipitation probability	67	3.85	3.04	50	0.81
Maximum temperature	41	3.83	3.42	62	0.41
Pavement condition	147	3.82	3.74	47	0.08
Freeze point temperature	83	3.81			
Snow accumulation	110	3.73			
Flood advisory	42	3.69	3.46	35	0.23
Precipitation accumulation	117	3.67	3.18	51	0.49
Chemical concentration	85	3.36			
Frost probability	27	3.22			
AVERAGE DIFFERENCE					**0.54**
STANDARD DEVIATION					**0.27**

Table 15. Comparison of combined Quality attribute average results for *Elements* from Control respondents in 2008 and 2010 surveys

ELEMENT	2010 SURVEY		2008 SURVEY		DIFFERENCE
	N	MEAN	MEAN	N	
Precipitation rate	6	5.00	3.65	23	1.35
Snow rate	6	5.00	3.67	18	1.33
Precipitation probability	6	5.00	3.85	20	1.15
Precipitation type probability	6	5.00	3.85	20	1.15
Visibility	6	5.00	3.80	20	1.20
Frost probability	6	5.00	3.85	20	1.15
Cloud cover	6	4.83	2.36	22	2.47
Precipitation accumulation	24	4.67	3.68	25	0.99
Snow accumulation	24	4.67	3.78	18	0.89
Weather type	18	4.56	3.74	19	0.82
Precipitation start time	7	4.43			
Precipitation end time	7	4.43			
Air temperature	52	4.37	3.96	25	0.41
Wind gust	18	4.33	4.00	25	0.33
Relative humidity	52	4.33	3.72	25	0.61
Wind direction	52	4.33	4.00	25	0.33
Wind speed	52	4.33	4.00	25	0.33
Dew point temperature	52	4.27	3.68	19	0.59
Camera view of weather	12	4.25	4.33	24	-0.08
Camera view of traffic	12	4.25	3.96	24	0.29
Precipitation type	41	4.20			
Wind advisory	9	4.11	3.94	18	0.17
Pavement temperature	34	4.00	3.84	25	0.16
Winter weather advisory	15	3.87	3.87	15	0.00
Dense fog advisory	15	3.87	4.05	19	-0.18
Severe weather advisory	24	3.83	4.06	16	-0.23
Camera view of road	6	3.83	4.36	25	-0.53
Pavement condition	31	3.81	3.89	19	-0.08
Road closure	6	3.50	4.06	17	-0.56
Freeze point temperature	27	3.26			
Radar images	18	3.17			
Chemical concentration	28	3.04			
Maximum temperature	0	0.00	3.29	21	
Minimum temperature	0	0.00	3.57	21	
Treatment recommendation	0	0.00			
Flood advisory	0	0.00	3.88	16	
Flood stage	0	0.00			
AVERAGE DIFFERENCE					**0.54**
STANDARD DEVIATION					**0.693**

Table 16. Comparison of combined Quality attribute average results for *Elements* from *Treatment* respondents in the 2008 and 2010 surveys

ELEMENT	2010 SURVEY		2008 SURVEY		DIFFERENCE
	N	MEAN	MEAN	N	
Precipitation type probability	56	4.23	3.85	59	0.38
Precipitation rate	56	4.20	3.75	56	0.45
Snow rate	56	4.20	3.84	58	0.36
Severe weather advisory	96	4.19	3.69	51	0.50
Wind advisory	47	4.15	3.78	60	0.37
Wind gust	129	4.15	4.08	66	0.07
Precipitation start time	86	4.12	3.85	59	0.27
Precipitation end time	86	4.12	3.88	59	0.24
Cloud cover	50	4.10	3.45	53	0.65
Frost probability	56	4.09	3.76	67	0.33
Radar images	282	4.05			
Precipitation probability	86	4.03	3.88	60	0.15
Wind direction	393	4.02	4.04	67	-0.02
Pavement temperature	222	4.02	4.28	68	-0.26
Wind speed	393	4.01	4.08	65	-0.07
Air temperature	363	4.01	4.24	68	-0.23
Camera view of weather	77	3.95			
Camera view of traffic	77	3.95			
Dew point temperature	357	3.95	4.06	68	-0.11
Weather type	175	3.93	3.84	58	0.09
Visibility	51	3.92	3.78	59	0.14
Camera view of road	77	3.91			
Relative humidity	355	3.90	4.03	64	-0.13
Winter weather advisory	71	3.89	3.83	54	0.06
Dense fog advisory	71	3.89	3.62	55	0.27
Precipitation type	314	3.86	3.98	63	-0.12
Precipitation accumulation	192	3.76	3.88	60	-0.12
Flood stage	12	3.75	3.61	31	0.14
Treatment recommendation	42	3.74	3.73	44	0.01
Maximum temperature	30	3.73	3.92	61	-0.19
Snow accumulation	173	3.74	3.92	64	-0.18
Pavement condition	214	3.72	3.76	63	-0.04
Minimum temperature	30	3.70	3.95	61	-0.25
Flood advisory	12	3.67	4.00	36	-0.33
Freeze point temperature	135	3.35	3.72	58	-0.37
Chemical concentration	129	3.09	2.89	58	0.20
Road closure	6	2.83	4.16	19	-1.33
AVERAGE DIFFERENCE					**0.08**
STANDARD DEVIATION					**0.882**

3.5 Comparison of the Product Results from the 2008 and 2010 Surveys

Each of the *Products* derived in the 2010 survey was composed of one or more *Product Components*. The components within each of the *Products* are shown in Table 2. To facilitate a comparison between the *Product* results from 2008 and 2010 surveys it was necessary to compute an estimate of the quality metrics for the 2010 *Product* groupings. To accomplish this, the scores from all of the *Product Components* within each *Product* classification were used to create averages for the six attributes, the composite average, and the importance for each *Product*. These derived values are representations of the score that might have occurred had the respondents answered questions about *Product* quality rather than *Product Component* quality. As with the *Element* analysis, the *Product* quality measures were computed separately from responses attributable to survey participants in the Advisory, Control, and Treatment strategy groups. Because of sample size considerations, the comparison of results between the two survey years was limited to the composite average metrics.

The comparisons of the results from each of the three management strategy classes collected during the 2008 and 2010 studies are provided in Table 17, Table 18, and Table 19. As with the *Element* analysis, the results from the Control group are questionable due to the small sample size and the fact a significant portion of the quality results are primarily from a single survey respondent. The results from the Advisory and Treatment comparisons have a higher statistical significance based upon the T-test and seem more reasonable based upon the consistency of the difference values as indicated by the relatively small standard deviation measures. Both of the Advisory and Treatment differences are negative but small. The T-test indicates both of these differences are statistically significant at the 95% confidence level, but would not be significant at the 99% confidence level.

The results of the *Product* comparison based upon the Advisory and Treatment comparisons suggest the measure of quality of road weather information based upon *Product* classifications has decreased slightly between 2008 and 2010. The decrease in the average quality for all *Products* combined in the Advisory strategy class is primarily due to decreased quality scores for the Pavement Forecast. The Road Condition *Product* also exhibited a significant decrease in the quality score, but the number of responses is 1/10 of the number for the Pavement Forecast *Product*. Part of the issue with the difference computation for Pavement Forecast is that the *Product* in the 2008 Advisory survey was actually a Route Specific Forecast typically used with 511 or related traveler information services. The road weather parameters in that 2008 *Product* were limited to Air Temperature, Wind Direction, Wind Speed, Weather, and Type of Precipitation. This set of components is quite different from the components in the Pavement Forecast *Product* computed in the 2010 survey. In addition, because of the uniqueness of the Route Specific Forecast, only four (4) survey participants responded to questions about this *Product*. Since there was limited response to the Route Specific Forecast in 2008 and there was a need to reduce the size of the survey, the Route Specific Forecast was replaced by the Pavement Forecast. This makes the comparison of the Route Specific Forecast with the Pavement Forecast of questionable value. If the Pavement Forecast *Product* is removed from the computation of the difference, the result changes to -0.11 instead of -0.19.

For the Treatment strategy class, the decrease is predominantly associated with decreases in the quality measures for the ESS History, MDSS, and Road Condition *Products*. Although the MDSS and Road Condition *Products* had substantial decreases in their quality averages from 2008 to 2010, the number of valid scores used to compute the 2010 *Product* average was small compared to the number used to derive the ESS History average. A closer review of the process that computes the ESS History Product indicates the derived average would be close to 4.00 if the average were computed without the Pavement Condition, Chemical Concentration, and Freeze Point numbers. This is still considerably different from the average calculated with the 2008 responses. There is definitely a decrease in the quality metric from 2008 to 2010 for this *Product*.

Table 17. Comparison of combined Quality attribute average results for *Products* from *Advisory* respondents in the 2008 and 2010 surveys. The DIFFERENCE values are 2010 survey mean values minus the 2008 survey mean values

PRODUCT	2010 SURVEY		2008 SURVEY		DIFFERENCE
	N	MEAN	MEAN	N	
Road Weather Alerts	90	4.18			
Camera Images	156	4.14			
Regional Map	398	4.11			
Maintenance Decision Support System	18	4.11			
Watches and Warnings	210	4.10	4.08	48	0.02
Pavement Weather Forecast	537	4.04	4.61	24	-0.57
Radar	168	3.96			
Weather Summary	668	3.87			
ESS Current Conditions	675	3.84			
ESS Histories	168	3.82			
Zone Forecast	246	3.79	3.85	48	-0.06
Flood Warning	84	3.77			
Weather History	112	3.64			
Road Condition Report	54	3.57	4.17	54	-0.60
AVERAGE DIFFERENCE					**-0.19**
STANDARD DEVIATION					**0.33**

Table 18. Comparison of combined Quality attribute average results for *Products* from *Control* respondents in the 2008 and 2010 surveys. The DIFFERENCE values are the 2010 survey mean values minus the 2008 survey mean values

PRODUCT	2010 SURVEY		2008 SURVEY		DIFFERENCE
	N	MEAN	MEAN	N	
Weather History	42	5.00			
Road Condition Report	6	5.00	4.25	18	0.75
Pavement Weather Forecast	132	4.99	3.16	12	1.83
Weather Summary	108	4.36			
Camera Images	18	4.11	4.31	24	-0.20
ESS Current Conditions	127	3.96		12	
Watches and Warnings	33	3.91	3.82		0.09
Regional Map	120	3.83			
Road Weather Alerts	12	3.50			
Radar	18	3.17			
ESS Histories	38	2.79			
Zone Forecast	0		3.33	18	
Maintenance Decision Support Systems	0				
Flood Warning	0				
AVERAGE DIFFERENCE					**0.34**
STANDARD DEVIATION					**0.902**

Table 19. Comparison of combined Quality attribute average results for *Products* from *Treatment* respondents in the 2008 and 2010 surveys. The DIFFERENCE values are the 2010 survey mean values minus the 2008 survey mean values

PRODUCT	2010 SURVEY		2008 SURVEY		DIFFERENCE
	N	MEAN	MEAN	N	
Pavement Weather Forecast	1199	4.19	4.27	42	-0.08
Watches and Warnings	143	4.17	3.79	24	0.38
Radar	282	4.05			
Camera Images	154	3.93			
Regional Map	625	3.90	4.24	42	-0.34
ESS Histories	348	3.84	4.27	36	-0.43
Weather Summary	769	3.84	3.93	42	-0.09
Maintenance Decision Support Systems	42	3.74	4.46	36	-0.72
Zone Forecast	180	3.73	3.48	54	0.25
ESS Current Conditions	768	3.73	4.23	59	-0.50
Flood Warning	24	3.71	3.62	18	0.09
Weather History	328	3.61	3.90	40	-0.29
Road Weather Alerts	30	3.27	3.42	24	-0.15
Road Condition Report	28	3.11	3.90	36	-0.79
AVERAGE DIFFERENCE					**-0.08**
STANDARD DEVIATION					**0.363**

3.6 COMPARISON OF THE PRODUCT RESULTS FOR REPEAT PARTICIPANTS

Changes in the level of the quality metrics from 2008 to 2010 were expected, especially with the assessment of the results for *Products* and *Elements* because of the modification of the survey structure. The analysis of the differences so far has compared results generated from all 2010 participants against the results from all 2008 participants for each of the three primary management strategy groups. This approach provided the largest sample size on which to perform the comparison. In theory, if the sample size was large enough the collective metric from that group should be the most representative assessment of quality. One of the difficulties experienced in the execution of the survey was the level of participation was lower than desired. The limited sample set had the potential to create biased results due to the way specific participants answered survey questions. A small number of individuals could impact the results in any survey year due to the way they approach the survey or an individual interested in influencing or disrupting the survey could conceivably significantly influence the metrics for a given survey year. Therefore, there was interest in determining if the results obtained from repeat participants would generate results similar or different from the composite set analyzed thus far.

Twenty-six (26) survey participants started both the 2008 and 2010 surveys. Eight of the participants indicated a primary management strategy classification of Treatment in both surveys, six classified themselves as Advisory in both surveys, and twelve changed their primary management strategy class or did not define their strategy class in the 2010 survey. Although these 26 surveys were tied to specific email addresses that had to be identified to initiate the survey, an optional question in the demographic section of the 2010 survey requesting the participant's name yielded a name different from the user identified in the email address. It is estimated that surrogate respondents may represent 10 to 20% of the responses. What is not known is whether the 2008 survey also had a significant number of surrogate participants and whether these surrogates were the same for both surveys. Since 45 individuals started the 2010 survey, if 26 were repeat participants then the 2010 survey had 19 new participants.

Since there were definitely repeat participants in the Advisory and Treatment strategy groups, plus a group whose strategy class was not defined, the analysis of repeat participants was done in three separate collections of participants: Advisory, Treatment, and Other. The results from each participant in its respective group are provided in Table 20. It is immediately evident from this table that it is difficult to compare results between the 2008 and 2010 surveys on an individual by individual basis. The greatest limitation to an effective comparison is the fact that many of the repeat participants did not evaluate the quality of the same *Products* in both of the surveys. The 2008 survey limited the *Products* the user could evaluate by design. The 2010 survey allowed participants to submit scores on the *Product Components* for any or even all *Products* in the survey, but participants in the 2010 survey could skip (and did skip) *Products* the users did not use or preferred not to rate. Participants even skipped components within *Product* groups they chose to evaluate. Therefore, the comparisons contained a large number of missing entries and had results from one survey year that did not have a corollary set of results from the other survey year.

One would like to average the differences and establish a single indication of the change in quality means from 2008 to 2010 either by *Product* or for all *Products*. However, it is not statistically possible to sum or average the differences in these results because the number of responses within any *Product* category is not uniform from individual to individual. Further, what is not seen in the results in Table 20 is the results in both the 2008 and 2010 surveys were derived from partial responses within that *Product* category in many cases. That is, in the 2008 survey respondents did not always respond to all attributes and in 2010 respondents did not necessarily respond to all *Product Components* within a *Product* group or even to all attributes within a given *Product Component*. What can be done with the results is to establish a repetitive pattern of trends in the data. For example, the following patterns appear in the data in Table 20: a general tendency for negative differences in the Treatment group, a tendency for positive differences in the Other collective group, and an undefined trend pattern in the Advisory collective group.

The decrease in the *Product* quality scores from 2008 to 2010 calculated for the Treatment management strategy group (Table 19) is also apparent in the results from the Treatment section of Table 20. There were 19 survey participants in the 2010 survey whose dominant management strategy responsibility was classified as Treatment in the demographic question on responsibilities. Results from the 2010 survey indicate that the number of Treatment participants who entered quality scores for the *Product Components* varied from 5 to 15 and that the counts were almost always the same for all components within a given *Product*. The individual results for the Treatment participants in Table 20 suggest that 6 of the 8 repeat participants actively responded in both surveys; therefore, the repeat responders make up a substantial portion of the total survey results. Twenty-three (23) of the 31 difference calculations for the Treatment group in Table 20 are negative. This replicates the pattern seen with the complete set of Treatment users and suggests that the repeat participants were responsible for or had a major influence on the reduction in the overall Product quality scores from 2008 to 2010.

The uncertain trend results in the Advisory group were based on four difference calculations that really do not provide an adequate measure for analysis. What is noticeable is that repeat respondent #4 had a quality score decrease of 1.48 for the Pavement Weather Forecast *Product*. Since the Pavement Weather Forecast comparison in Table 17 was based upon input from 4 participants in the 2010 survey, the -1.48 difference from repeat respondent #4 had a significant influence on the overall -0.57 difference result in Table 17.

Eight of the 11 differences in the Other group were positive or increases in the assessed quality. These repeat participants could not be traced specifically to one of the management strategies in both surveys and thus ended up in the Other category; however, each was placed in one of the strategy groups in the 2010 survey. These repeat participants influenced the increase in the overall *Product* quality scores for the Control group (Table 18) and probably moderated the negative results in the overall Treatment group.

Table 20. Comparison of 2008 and 2010 results from repeat participants

ADVISORY PARTICIPANTS

PRODUCT	RESPONDENT 1			RESPONDENT 2			RESPONDENT 3			RESPONDENT 4		
	2010	2008	DIFF	2010	2008	DIFF	2010	2008	DIFF	2010	2008	DIFF
Weather Summary	4.48									3.47		
Weather History												
ESS Current Conditions										3.29		
ESS Histories												
Regional Map	4.11											
Zone Forecast	4.14	4.00	0.14							3.50	3.00	0.50
Pavement Weather Forecast		4.00								3.52	5.00	-1.48
Road Weather Alerts	4.00									3.33		
Watches and Warnings	3.94									3.50		
MDSS										3.33	3.33	0.00
Road Condition Report		4.00								3.08		
Flood Warning	3.58											
Camera Images												
Radar												

PRODUCT	RESPONDENT 5			RESPONDENT 6		
	2010	2008	DIFF	2010	2008	DIFF
Weather Summary						
Weather History						
ESS Current Conditions						
ESS Histories						
Regional Map						
Zone Forecast		4.00				
Pavement Weather Forecast		4.00				
Road Weather Alerts						
Watches and Warnings						
MDSS						
Road Condition Report		4.00			5.00	
Flood Warning						
Camera Images						
Radar						

TREATMENT PARTICIPANTS

PRODUCT	RESPONDENT 1			RESPONDENT 2			RESPONDENT 3			RESPONDENT 4		
	2010	2008	DIFF	2010	2008	DIFF	2010	2008	DIFF	2010	2008	DIFF
Weather Summary				4.77			3.02	3.17	-0.15	4.21	3.50	0.71
Weather History					3.67		3.43	4.83	-1.40		3.67	
ESS Current Conditions	3.91			3.30	4.00	-0.70	3.91	4.50	-0.59	3.48	4.00	-0.52
ESS Histories					4.00		4.15	5.00	-0.85	3.89	4.00	-0.11
Regional Map	4.72			4.65			3.18	4.00	-0.82	4.12	4.50	-0.38
Zone Forecast					4.00		3.72	3.33	0.39	4.22	3.33	0.89
Pavement Weather Forecast	4.77			4.45			4.43	4.83	-0.40	4.12	3.33	0.79
Road Weather Alerts											3.17	
Watches and Warnings				4.35				4.17		4.50	3.17	1.33
MDSS				4.00	4.50	-0.50		4.33		3.50	4.33	-0.83
Road Condition Report				300			2.00			2.83	3.33	-0.50
Flood Warning								3.00		3.42		
Camera Images												
Radar												

PRODUCT	RESPONDENT 5			RESPONDENT 6			RESPONDENT 7			RESPONDENT 8		
	2010	2008	DIFF	2010	2008	DIFF	2010	2008	DIFF	2010	2008	DIFF
Weather Summary		4.33		3.27			3.96	4.33	-0.37	2.87	5.00	-2.13
Weather History		3.75		3.00			3.95	4.33	-0.38		2.83	
ESS Current Conditions	2.67	3.83	-1.17	4.60				4.67		3.94	5.00	-1.06
ESS Histories	2.69	3.00	-0.31	4.67				5.00				
Regional Map		4.50		3.33			4.70	4.17	0.53	4.00	4.00	0.00
Zone Forecast		4.00					3.72	3.17	0.56	2.83		
Pavement Weather Forecast		4.00		3.83			4.08	4.83	-0.76	4.17	5.00	-0.83
Road Weather Alerts		3.00		3.17				4.17				
Watches and Warnings		4.50		5.00			3.78			3.33		
MDSS	4.00	4.50	-0.50							4.33	5.00	-0.67

Road Condition Report		4.17		4.00							4.50	
Flood Warning		4.00		4.00								
Camera Images												
Radar												
OTHER PARTICIPANTS - UNDEFINED												
PRODUCT	RESPONDENT 1			RESPONDENT 2			RESPONDENT 3			RESPONDENT 4		
	2010	2008	DIFF	2010	2008	DIFF	2010	2008	DIFF	2010	2008	DIFF
Weather Summary	4.07			4.59			3.63					
Weather History				4.67								
ESS Current Conditions	4.45			3.63			3.78					
ESS Histories												
Regional Map	4.00			4.78								
Zone Forecast		3.17		4.64								
Pavement Weather Forecast		3.17		4.63								
Road Weather Alerts				4.67								
Watches and Warnings		3.83		4.67								
MDSS				4.67								
Road Condition Report		3.50										
Flood Warning				4.50								
Camera Images	4.07									4.07		
Radar	4.07									4.07		
PRODUCT	RESPONDENT 5			RESPONDENT 6			RESPONDENT 7			RESPONDENT 8		
	2010	2008	DIFF	2010	2008	DIFF	2010	2008	DIFF	2010	2008	DIFF
Weather Summary				3.85			4.20					
Weather History												
ESS Current Conditions				4.00	4.33	-0.33	3.55			3.93		
ESS Histories					4.00							
Regional Map					4.00		4.08				4.50	
Zone Forecast		3.67			2.83		3.75	3.17	0.58		3.33	
Pavement Weather Forecast					3.50		4.19	3.17	1.02			
Road Weather Alerts											3.33	
Watches and Warnings							4.50	3.83	0.67	4.00	3.17	0.83
MDSS					4.00					4.00		
Road Condition Report				3.17	4.33	-1.17	3.33	3.50	-0.17	4.00	3.00	1.00
Flood Warning							4.50					
Camera Images	4.07			4.07			4.07			4.07		
Radar	4.07			4.07			4.07			4.07		
PRODUCT	RESPONDENT 9			RESPONDENT 10			RESPONDENT 11			RESPONDENT 12		
	2010	2008	DIFF	2010	2008	DIFF	2010	2008	DIFF	2010	2008	DIFF
Weather Summary				3.20			4.31			5.00		
Weather History				3.21						5.00		
ESS Current Conditions	3.21									4.65	4.00	0.65
ESS Histories	2.79											
Regional Map				3.55			3.23			4.25		
Zone Forecast											3.00	
Pavement Weather Forecast							4.08			4.99		
Road Weather Alerts				2.83								
Watches and Warnings												
MDSS							2.83					
Road Condition Report												
Flood Warning												
Camera Images	4.07			4.07			4.07			4.07		
Radar	4.07			4.07			4.07			4.07		

The *Product* quality attribute difference scores seemed to have distinct patterns within each of the management strategy groups. The question then was whether particular attributes were responsible for this quality change pattern. The differences in the scores within the attributes were calculated and are presented in

Table 21, Table 22, and Table 23 for the Advisory, Control, and Treatment strategy classes, respectively. The decreases in quality noted in the Advisory and Treatment groups are notably attributable to decreases in the quality scores in the Accuracy/Precision, Timeliness/Reliability, and Ease of Use attributes. Even the increase in quality for the Control group is suppressed for the Accuracy/Precision and Ease of Use attributes. The intriguing decrease in quality is the change in the quality scores in the Relevance attribute in the Treatment group. One would not expect that DOT personnel would change their assessment of the relevance of their resources. This unexpected decrease may be an indicator that some other factor(s) are influencing the results. Three plausible causes may be the change in the structure of the surveys with the concomitant derivation of *Product* results from *Product Component* responses; the small sample size used in the computations; or it may reflect a change in how the Treatment respondents related to the survey. In reference to the third possible cause, it has already been established that the composite scores reflect considerable influence from repeat participants. A computation of the Relevance differences for the eight repeat participants illustrates that the computed Relevance difference in Table 23 also occurs for repeat participants.

PRODUCT	DIFFERENCE
Weather Summary	-0.07
Weather Histories	-0.21
ESS Observations	-0.01
ESS Histories	-0.65
Regional Map	-0.39
Zone Forecast	-0.19
Pavement Forecast	0.08
Road Weather Alerts	0.00
Watches and Warnings	0.12
MDSS	-0.25
Road Condition Reports	-1.00

Even these numbers cannot be compared directly to one another since they are derived from computations done with considerably different sample sets. MDSS has already been shown to be an inappropriate comparison. Therefore, other than the Road Condition Reports difference, it appears that most of the decrease in Relevance comes from the perceived value of history information and the display of observations.

The limitations in the ability to summarize individual differences into a collective comparison make an argument for the comparison of quality from survey to survey using a composite of all results where the determination of quality is based upon the responses from the entire sample set. This has been the approach used in this report. Further, to get a good representative quality value for the *Product* and *Element* analyses, it will require a significantly larger number of responses in order to counteract the partial responses that were received from individual users.

Table 21. Differences in the quality attribute averages between the 2008 and 2010 surveys for Advisory participants

DIFFERENCES IN ADVISORY QUALITY ATTRIBUTE MEAN VALUES (2010 SURVEY - 2008 SURVEY)						
PRODUCT	Accuracy/Precision	Completeness	Relevance	Currency/Latency	Timeliness/Reliability	Ease of Use
Weather Summary						
Weather History						
ESS Current Conditions						
ESS Histories						
Regional Map						
Zone Forecast	-0.40	0.33	-0.24	0.16	-0.21	0.18
Pavement Weather Forecast						
Road Weather Alerts						
Watches and Warnings	0.09					
MDSS						
Road Condition Report	-0.66	-0.44	-0.11	-0.89	-0.67	-0.44
Flood Warning						
AVERAGE DIFFERENCE	**-0.33**	**0.07**	**-0.01**	**-0.17**	**-0.29**	**-0.22**

Table 22. Differences in the quality attribute averages between the 2008 and 2010 surveys for Control participants

DIFFERENCES IN CONTROL QUALITY ATTRIBUTE MEAN VALUES (2010 SURVEY - 2008 SURVEY)						
PRODUCT	**Accuracy/Precision**	**Completeness**	**Relevance**	**Currency/Latency**	**Timeliness/Reliability**	**Ease of Use**
Weather Summary						
Weather History						
ESS Current Conditions						
ESS Histories						
Regional Map						
Zone Forecast	0.00	0.00	0.00	0.00	0.00	0.00
Pavement Weather Forecast	2.00	2.00	2.00	2.00	2.00	0.95
Road Weather Alerts						
Watches and Warnings	-0.33	0.04	-0.33	0.58	0.33	-0.83
MDSS						
Road Condition Report	1.00	2.00	1.00	1.00	1.00	1.00
Flood Warning						
AVERAGE DIFFERENCE	**0.24**	**0.37**	**0.24**	**0.33**	**0.30**	**0.10**

Table 23. Differences in the quality attribute averages between the 2008 and 2010 surveys for Treatment participants

DIFFERENCES IN TREATMENT QUALITY ATTRIBUTE MEAN VALUES (2010 SURVEY - 2008 SURVEY)						
PRODUCT	Accuracy/Precision	Completeness	Relevance	Currency/Latency	Timeliness/Reliability	Ease of Use
Weather Summary	-0.37	0.09	0.07	0.22	-0.44	-0.11
Weather History	-0.38	-0.11	-0.11	-0.55	-0.79	0.31
ESS Current Conditions	-0.74	-0.33	-0.16	-0.70	-0.93	-0.39
ESS Histories	-0.88	-0.41	-0.55	0.21	0.21	-0.53
Regional Map	-0.71	-0.26	-0.29	-0.43	-0.13	-0.18
Zone Forecast	-0.04	0.80	-0.04	0.47	0.43	0.13
Pavement Weather Forecast	-0.25	0.09	0.07	0.01	-0.54	0.06
Road Weather Alerts	0.15	0.15	-0.70	0.55	0.15	-1.20
Watches and Warnings	-0.01	0.88	0.17	0.25	0.50	0.75
MDSS	-0.71	-0.79	-0.36	-0.96	-0.97	-0.46
Road Condition Report	-0.80	-0.63	-0.80	-0.70	-0.83	-0.92
Flood Warning	0.08	-0.17	-0.17	-0.17	0.33	0.33
AVERAGE DIFFERENCE	**-0.36**	**0.01**	**-0.23**	**-0.08**	**-0.19**	**-0.16**

4.0 Summary and Conclusions

4.1 Summary of the 2010 Survey Results

The 2010 survey was a substantial modification from the first Baselining survey executed in 2008 *Product Components* became the primary resource to acquire the quality of road weather information replacing the questions regarding *Elements* and *Products* that were used in the 2008 survey. The characteristics that make *Product Components*, *Elements*, and *Products* unique entities were covered in depth in Section 2.1. That section also described how the three entities were related to one another. The results of the *Product Component* computations from the 2010 survey were covered in Section 3.2 and this was followed by an interpretation and analysis of the *Product Component* results. Since *Product Components* are constituents of both *Elements* and *Products*, quality scores were derived for both of these classifications from the *Product Component* results. The derivation process is explained in Section 2.3. The results from the derivation of the *Element* average scores in 2010 were compared to the 2008 average scores in Section 3.4. Similarly, the results of the derived *Product* average scores were compared to the 2008 average scores in Section 3.4. The analysis of the *Product* results was expanded in Section 3.5 to look at the responses from known repeat participants. This summary reviews the key points obtained from the structure of the 2010 survey and the analysis of the results. It is organized around the three formats that have been utilized as tools to extract the metrics on road weather information quality: *Product Components*, *Elements*, and *Products*.

<u>4.1.1 2010 Product Component Results</u>

The results from the 2010 survey represent the input from 37 individuals; 15 indicated their primary management responsibilities were Advisory in nature, 5 indicated Control, and 17 selected Treatment.

Survey respondents could submit quality scores for each of 92 *Product Components* by selecting one of five quality rating scores ranging from Very High to Very Low for each of six quality attribute categories and their importance. Respondents could skip Product Components that they did not use; therefore, the results from the 37 respondents reflect quality ratings of those *Product Components* that are actively used by the DOT participants in the survey. The actual responses amounted to 45% of the total possible answers. The average number of responses per *Product Component* attribute was 16.5; however, the actual number of responses per *Product Component* ranged from 4 to 31 with 68% of the number of responses between 10 and 23. The number of responses for a few of *Product Component* attributes meant their average Likert scores were marginally reliable. Nevertheless, the number of responses for most of the responses was adequate to make the comparison of quality attribute means statistically significant. However, if the responses were subdivided into their management strategy groups, the sample size for comparisons between participants in the three groups would have had a predominance of *Product Component* averages that were not statistically significant. Therefore, the *Product Component* analysis was limited to a comparison of responses for the combined set of participants.

The computations for each attribute and the composite attribute included the average, the median, the standard deviation, the maximum, the minimum, and the skewness of the distribution. For comparison purposes, the averages within each of the attributes and the composite attribute were ranked and these rankings were separated into quartiles that were color-coded to permit visual recognition of ranking patterns.

The statistical analysis of the *Product Component* responses (Table 9) indicates the quality scores average near a Likert score of 4 (survey response of HIGH) and the importance of the *Product Components* is rated above HIGH (4.25). The quality attribute distributions are not perfectly normal curves but skewed toward lower scores. In general, most of the responses fell between 3 and 4.5 (MODERATE to midway between HIGH and VERY HIGH).

The ranking information indicates the Pavement Forecast data is the most important road weather resource along with the observed precipitation data from the National Weather Service. Other highly rated road weather information included camera imagery, road weather alerts, and basic radar imagery. The least important of the road weather parameters are the current observations of relative humidity, dew point temperature, and chemical concentration; most of the historical information; the forecasted max/min temperatures; flood information; and future radar.

For quality, the Pavement Forecast data was rated as having the highest quality with the exception of the Probability of Precipitation and Chemical Concentration. This group of components was followed by Watches and Warnings and Radar information. The quality rankings at the low end of the ranking scale included *Product Components* dealing with historical information, any form of observed or forecasted pavement conditions, zone forecasts, and road condition reports.

Specific findings that were derived from the analyses in Section 3.3 were:

- Weather Summaries suffered because the key components of interest were considered lower in accuracy and not as current as desired by the DOT personnel.
- DOT personnel found the NWS History information of less value than the information in other Products across all quality attributes.
- ESS Observations have a lower level of importance to the DOT respondents than many of the other resources in this Baselining Study and are viewed as having marginal quality to the users.
    - Data accuracy, timeliness, and reliability need further attention
    - These limitations impact usage of data by DOT users, NWS, weather service providers, media, and the *Clarus* program
- ESS Histories were considered one of the least important tools and they generally received quality rankings that were below the median values.
- Map displays had a middle of the road *Product Components* both in importance and quality attribute scores.
- Pavement Weather Forecasts are the highest priority road weather information tool and DOT participants find the quality of the forecast information well above average except for pavement condition-related *Product Components*.

- Road Weather Alerts are important to DOT users, but they are disappointed with the limited content of the information and its availability on a timely basis.
- Watches and Warnings are an important decision support tool that rates second to Pavement Forecasts in its value to the DOT participants in this study.
- MDSS Treatment Recommendations are viewed as a relevant and reasonably important tool but are thus far not perceived as one of the more accurate, timely, reliable, and easy to use resources.
- Road Condition Reports are considered of average importance by DOT users; these users see a lot of room for improvement in this resource. Flood Warnings are of lesser value to DOT users and the quality of the information provided is lower than with most other road weather information resources.
- Flood Warnings are of lesser value to DOT users and the quality of the information provided is lower than with most other road weather information resources.
- Camera imagery is an important resource for DOT operations and survey participants suggest that this tool could be even more effective with more or better selection of camera views.
- Radar information scores indicate users see the fundamental radar images as easy to use, timely, and reasonably accurate. Future radar and storm track derived services need to improve to gain an equivalent acceptance of the observed radar products.

4.1.2 2010 Element Results

The *Product Component* results from the 2010 survey were transformed to derived *Element* results using the technique described in Section 2.3 in order to provide a base for comparison of *Element* results between 2008 and 2010. Since the 2008 surveys had separate question sets based on the participant's selected management strategy, the derived *Element* results from the 2010 survey also had to be separated by strategy. The *Product Component* answers for the participants in each strategy were transformed to derived *Element* scores for each strategy and these answers were then compared to the 2008 scores. Because the responses from the 37 participants were reduced to much smaller sample numbers in each of the strategies, only the composite average results were compared between the two survey years. Tables 14 – 16 in Section 3.4 contain the results of the comparisons for the three different strategy groups.

Because all 2010 survey participants could submit quality scores for any *Product Component*, it was possible to derive *Element* scores for nearly all of the *Elements* in each of the three management strategy groups. These derived composite average scores and the number of valid responses were listed in Tables 14 – 16 next to the *Element* names that had been ordered based upon the average scores. To the right of these columns the average composite averages of the quality attribute scores from the 2008 survey and the number of answers used to compute the averages were listed. For all *Elements* where there were results from both surveys, a simple difference was computed representing the 2010 results minus the 2008 results. In this way, the difference indicated an increase in quality if the number was positive.

The Advisory management strategy group had composite average scores for *Elements* in both surveys for 15 of the potential 37 *Elements*. All of the computed differences were positive indicating an increase in the perceived quality for these *Elements* over the two-year period. The average increase was 0.54 or nearly half of one step in the Likert scale. The distribution of the differences was relatively tightly arrayed around the difference mean. The *Elements* showing the most significant increases included the advisory messages (wind, severe weather, winter weather, and dense fog) and precipitation probability.

The ordered ranking of the composite attribute scores for the Advisory group indicated the group rated the *Elements* dealing with traffic issues as having the highest quality. These *Elements* included camera views of weather and traffic, visibility, road closures, snow rate, and severe weather advisories. Each composite average score was 4.1 or higher. At the opposite end of the scale were *Elements* dealing with pavement conditions (frost probability, chemical concentration, freeze point temperature) and accumulation amounts of rain and snow.

The Control management strategy group had composite average scores for *Elements* in both surveys for 26 of the 37 *Elements*. Six of these 26 computed differences were negative. Even so, the average of all individual *Element* differences was 0.54 indicating the same level of increase for the composite group of differences as seen in the Advisory data. However, the distribution of differences for the Control group was considerably more spread out. The top seven *Elements* in the ordered ranking accounted for nearly all of the average increase in the quality scores and all seven of these scores in 2010 came from a single individual who entered a quality rating of 5 (VERY HIGH) for all but a couple of *Product Components*. If those seven *Elements* were thrown out, the average difference would remain positive but drop to 0.22. This type of change is indicative of two issues addressed in Section 4.2: insufficient sample size and the influence of the Likert scoring mechanism.

The ordered ranking is a bit confused because of the top seven scores of 5.00. Ignoring these seven top *Elements* in the ranking, the highest ranked remaining *Elements* are precipitation and snow accumulation, weather type, and precipitation start and end times. These *Elements* had been near the bottom of the Advisory ranking. However, pavement condition, chemical concentration, freeze point temperature, and road closures were at the bottom of the Control rankings. One surprising *Element* also near the bottom was Radar, reflecting the negative influence of the Future Radar and Storm Tracks components.

4.1.3 2010 Product Results

The *Product Component* results from the 2010 survey were transformed to derived *Product* results using the technique described in Section 2.3. The derived *Product* results became the comparison metric to assess the change in quality since the 2008 survey. As with the *Element* comparison the *Product* comparison had to be done in management strategy groups. The *Product Component* answers for the participants in each strategy were transformed to derived *Product* scores for each strategy and these answers were then compared to the 2008 scores. Because the responses from the 37 participants were reduced to much smaller sample numbers in each of the strategies, only the composite average results were compared between the two survey years. Tables 17 – 19 in Section 3.5 contain the results of the comparisons for the three different strategy groups.

As with the derived *Element* scores, there were derived *Product* scores for all *Products* in all three strategies. The only three exceptions were the Zone Forecast, MDSS, and Flood Warning *Products* in the Control strategy for which there were no responses in the 2010 survey. The derived *Product* results were placed in Tables 17 – 19 and the *Product* categories were placed in descending order based upon the derived average Likert score. The averages and the number of responses used to compute the average scores for each *Product* for both survey years were placed in the row following the *Product* name. Where there were results in both years, the differences were computed by subtracting the 2008 average from the 2010 average.

The Advisory management strategy group had duplicate sets of results in 4 of the 14 *Product* categories. Three of the four *Products* differences had negative difference values and the average of all four was -0.19 indicating an overall decrease in assessed quality of the *Products* within the Advisory group. The Watches and Warnings and Zone Forecast *Products* were essentially unchanged; however, the Pavement Forecast and Road Condition Report *Products* both were down roughly 0.5 units on the Likert scale.

Road Weather Alerts and Camera Images had the highest *Product* averages in the Advisory group matching the emphasis on traffic-related concerns seen in the *Element* quality ranking.

The Control management strategy group had results for both survey years in 4 of the 14 *Product* categories. Differences for 3 of the 4 were positive and the average of the four combined also were positive (0.34). However, there is some question regarding the validity of two of the higher positive differences. These two are derived from the answers from one individual who consistently rated the quality of a number of *Product Components* as 5. If these two results are ignored then the average of the other two is closer to zero and even slightly negative. Recall there were only 5 participants defined as members of the Control group. Anomalous results from just one individual when the overall sample size is just five participants would likely disrupt the statistics and seems to be what is going on in the overall Control results.

The Treatment management strategy group had results in 12 of the 14 *Product* categories. The differences between the 2008 and 2010 averages produced negative results in 9 of the 12 comparisons and the average difference for all comparisons was -0.08. The Watches and Warnings and Zone Forecast Products had increases around 0.30 while ESS Current Conditions, ESS Histories, Road Condition Reports, and MDSS had decreases of 0.5 or greater. The average difference value of -0.08 appears incorrect based upon the preponderance of negative values for the individual differences. However, when the overall average difference is computed using the influence of all responses to compute the mean for each survey, the average difference is actually closer to zero. In fact, if the influence of the Treatment Recommendation is dropped from the comparison, the overall decrease drops to -0.04.

Pavement Weather Forecasts, Watches and Warnings, and Radar Products topped the list of Products in quality while Weather Histories, Road Weather Alerts, and Road Condition Reports were at the bottom of the ranked list.

4.1.4 Repeat Participants

By looking at the *Product* results from only those individuals who participated in both surveys, it was anticipated there might be clarification regarding the decreases in the *Product* quality between 2008 and 2010. This assumption was not substantially supported by the results of this confined sample of participants. It proved difficult to compare results between the 2008 and 2010 surveys on an individual-by-individual basis. Many of these repeat participants did not answer all of the questions in both of the surveys. Further, it was not obvious that the tendency for lower quality scores apparent in the larger composite sample was present in the repeat data set, although there seemed to be the tendency indicated within the Treatment group.

A significant challenge was the fact that it was not possible to sum the differences in these results because the number of responses within any *Product* category that generated the differences was not uniform from individual to individual. In addition, the results in both surveys were typically derived from partial responses within each Product category. That is, in 2008 survey respondents did not always respond to all attributes and in the 2010 survey respondents did not necessarily respond to all *Product Components* within a *Product* group or all attributes within a given *Product Component.*

The analysis of the repeat participants illustrates an important point regarding the comparison of results between sequential surveys. Individual participants in sequential surveys will have the option to submit quality assessments of whichever *Product Components* or attributes they choose. At the *Product Component* level that would mean the comparisons must be done at the attribute level since there is no guarantee the repeat participant will submit quality assessments for all six attributes for a given *Product Component* consistently. Consistency becomes even more of an issue for *Products* and *Elements* that are derived from *Product Components* since participants have the option to skip *Product Component* questions in *Products* they don't use or prefer to skip in one survey but not the other. These factors make a strong argument to perform the comparison of quality from survey to survey based upon the composite of all results from a relatively large sample of participants.

4.2 Sources of Errors and Inconsistencies in Results

The 2008 and 2010 surveys yielded subjective assessments of quality and importance. Likert scores were derived from these responses, and subsequently analyzed using standard variance statistical methods and ranking techniques. In addition, the survey format was modified between the 2008 and 2010 surveys, although techniques were employed to permit comparison of the results. The survey design and each of the processing steps have the potential to cause errors or inconsistencies in the results and in the analysis of the results. This section explores the potential sources of errors or inconsistencies in the results.

4.2.1 Potential Sources of Error

Errors in the analysis of a topic such as quality are generally tied to the design of the survey instrument. Surveys must ask the appropriate questions to acquire the metric desired and they must do this without push back from those willing to complete the survey. As the series of surveys has evolved, the design team has gained knowledge regarding the process of acquiring metrics on road weather information quality. After the 2008 survey the question set needed to be changed to eliminate some deficiencies in the original design and more importantly ask questions that addressed the specific tools that DOT users used to make decisions. Therefore, survey design and modification of the design were potential sources of error in the results.

4.2.1.1 2008 Survey Design

The 2008 survey was built around *Elements* and *Products*. In addition, the design team felt the survey should be separated into three distinct survey instruments to fit the specific interests of the three primary management strategies in DOT operations. Finally, the survey was designed to assess the quality of road weather information based upon input from DOT personnel within one of the three management strategies. It turns out these design factors created unforeseen impediments for an ongoing longitudinal assessment of road weather quality.

1. There is uncertainty that the 2008 quality metrics were effective measures of road weather information quality.

The analysis of the results from the 2008 survey coupled with feedback from participants regarding how they should answer questions about *Elements* and *Products* highlighted a fundamental issue with *Elements* and *Products*. Regarding the *Element* questions, survey participants indicated uncertainty about what they should evaluate. DOT users need specific forms of an *Element* to make operational decisions. For example, if a maintenance supervisor needs to make a treatment decision on an upcoming event, that supervisor needs forecasted values of precipitation type, precipitation rates, the precipitation start time, pavement temperature, and related road weather parameters. Current and historical observations of these parameters are not the appropriate resource. The maintenance supervisor knows to turn to a specific Product to get the necessary information to support the decision process. Therefore, specific forms of *Elements* are needed such that when presented with questions on the quality of *Elements*, the survey participants asked the question, “how do I rate the quality of the various forms that the *Elements* take?” Participants were faced with a similar dilemma with the *Product* questions. Participants use specific products to get the components within the *Product* to make a decision. Some of the components were valuable decision support resources and some were not. The question that arose was “how do I rate the quality of a ‘thing’ that is composed of a series of information pieces that have different levels of quality?” *Elements* and *Products* were not the entity that the DOT survey participants used to determine the quality of road weather information. The real resource was the components within specific *Products*. This was the driving force to change the 2010 survey to address *Product Components* rather than *Elements* and *Products*. The ongoing concern in performing comparisons against the 2008 results is that the results may be an abstract or fuzzy estimate of the user’s actual assessment of quality.

2. The fixed set of questions on *Elements* and Products in the 2008 survey for different strategy groups may have artificially constrained user input of quality assessments.

The selection of which *Elements* and *Products* were typical resources used by members of the different management strategy classes represented an artificial characterization of what users in the different management strategies use to make decisions. In the 2010 survey, participants were given the opportunity to select which resources they use from the entire set of possible road weather information tools and then provide their assessment of quality for each of the tools that they use. They could skip resources they did not use, but were not constrained to the set that they could evaluate. The results from 2010 indicate that participants in each strategy do use decision support resources that the design team had not allowed them to address in 2008. The influence this has had on the Baselining process is to create limitations in the 2008 survey data set and thus may limit comparisons of specific *Elements* or *Products* that were legitimate resources.

3. **The limitation of the survey population to DOT participants confines the quality assessment to the perspective of that user community.**

DOT personnel comprise a significant portion of the road weather information user community; however, the stakeholder community interested in road weather information is considerably more extensive. Notable stakeholder groups include the National Weather Service, general weather service providers, surface transportation weather service providers, and travelers (commercial and private). These stakeholder groups have an interest in the quality of road weather information that parallels the interest of the DOT users and in a number of situations likely assess the quality of the resource information more critically than the DOT users. There is no intent to change the survey population; however, it is important to recognize that the Baselining Study reflects the perspective of DOT users.

4.2.1.2 2010 Survey Design

The 2010 survey was redesigned to address *Product Components* instead of *Elements* and *Products*. This meant the metrics provided by survey participants represented a new quality measure that did not have an obvious, direct correlation to the metrics acquired from the 2008 survey. Since *Product Components* were constituents of both *Elements* and *Products* a technique was implemented to transform the *Product Component* responses into derived values that were theoretically equivalent to the *Element* and *Product* answers in the 2008 survey. This permitted one to track the change in the quality of these derived *Element* and *Product* scores going forward from 2010 and compare the scores against the 2008 survey values. Nevertheless, the modification did pose the potential for differences between the 2010 results and those from 2008.

1. ***Element* averages are a composite of the individual components of the *Element* and not a single quality assessment.**

Element scores in 2010 were derived from the Product Component scores from all components of that *Element*. Table 4 lists the source components for each *Element*. The premise behind this technique is that the *Product Components* provide a more reliable quality assessment for each of the *Product Components* containing that road weather parameter; (aka, *Element*) and, the average of all of the component scores will provide a better estimate of the score for that *Element* than the single quality assessment score captured in the 2008 survey (see Section 4.2.1.1, item 1). However, it is unknown which approach provides a better assessment of quality and whether the difference between the 2008 and 2010 *Element* scores is due to an actual change in user perception of quality or due to the change in methodology. The BAH team had not expected to see much of a change in the *Element* scores over the two-year period since the use of road weather information is a mature practice. Therefore, the limited increase of the *Element* scores between 2008 and 2010 for the Treatment group and the Control group (once the influence of the one user with consistent very high scores was dismissed) positively reinforced the design change to *Product Components*. The increase of 0.54 units on the Likert scale for the Advisory group did raise a flag that the derived Element averages may represent different measurement tools and are not as comparable as expected.

2. **Product averages are a composite of the individual components within each *Product* and not a single quality assessment.**

Product scores in 2010 were derived from the *Product Component* scores from all components within that *Product*. Table 2 lists the source components for each *Product*. It was anticipated that the average of the component scores will provide a better estimate of the score for that *Product* than the single quality assessment score captured in the 2008 survey. However, this is not substantiated and will be difficult to prove. Although the direct comparisons of individual *Product* scores between 2008 and 2010 were negative, the composite scores of all comparisons were only slightly negative and close to zero difference. Even so the preponderance of individual negative comparisons, especially in the Treatment group, raises questions, such as, was there really a decrease in the quality assessment and if so is this a real difference or is it attributable to the change in the design?

4.2.1.3 Analytical Computations

The analytical computations used to determine the quality relationships in this report and the report from the 2008 survey require numerous steps. Although the algorithms and the separation of data into strategy groups has been checked and the results seem reasonable, it is possible that mistakes were made in the entry of the computational algorithms or the transfer of data from the 2008 survey for comparison of results. There may be errors in the computed results that were not caught during the review of the computational steps.

4.2.2 Potential Sources of Uncertainty

Uncertainty is always a concern in the assessment of a subjective measure such as quality derived from a survey. The uncertainty derives from human factor influences associated with the community of participants in the survey, the survey testing methodology, and the analysis process. This section considers those factors that seem to have the most potential to cause uncertainty for the 2010 survey and its comparison to the 2008 survey results.

4.2.2.1 User's Assessment of Quality

The objective of the Baselining Study was the measurement of the quality of road weather information. Quality was the primary metric. Importance was a secondary consideration. The results from the 2008 survey raised questions regarding whether the survey actually measured quality or value. In itself, whether the quality attributes measured quality or value was not critical. But, the importance of understanding the definition of "quality" and how that definition impacted its use as a metric influenced two aspects of the study. First, it influenced the interpretation of the study results and impacted how various readers with different backgrounds would evaluate the results. Second, it influenced the organization and structure of the survey questions since the questions needed to be formulated in a manner appropriate to extract the desired metric. It was decided that quality represented the attributes, characteristics, or properties of a thing or phenomenon that can be observed and interpreted and which may be approximated (quantified) but cannot be measured exactly.

The surveys assessed overall quality of road weather information by looking at six quality attributes. The expectation was that these attributes would differentiate various aspects related to quality. A review of the six attribute definitions indicates that three of the six attributes included terms indicating the use of the parameter to meet user needs. These attributes (Completeness, Relevance, and Ease of Use) seem to relate more closely to value judgments rather than strict quality assessments. Conversely, Accuracy/Precision, Currency/Latency, and Timeliness/Reliability were more related to the measurable and verifiable aspects of a thing or phenomenon, expressed in numbers or quantities

One of the significant findings in the results from both surveys was that certain road weather elements delivered to users in *Product* bulletins often had unexpected quality attribute scores. The situation was particularly apparent for scores in the Currency/Latency and Timeliness/Reliability attributes. Components delivered in a given *Product* bulletin should have had essentially the same scores within both the Currency/Latency and Timeliness/Reliability attribute categories since they were all delivered together. However, the results showed significant variability between the attribute scores of components within the same *Product*. What was evident was that if a component was ranked high across all other attributes then the scores within these two attributes were also high; and, if the attribute scores were generally low, then the scores within these two attributes were low. Because of their nature, these two variables should have provided an assessment of quality of the delivery process of these components. The results indicate the DOT responses were not based upon quality but rather a more general value of the component without regard to the specific attribute.

These factors indicate that how one interprets quality can impact decisions regarding how one applies the metric employed in the survey.

4.2.2.2 Survey Composition and Questioning Technique

The issue associated with the design change of the surveys was covered in Sections 4.2.1.1 and 4.2.1.2. The design change creates uncertainty in the ongoing comparison of the 2010 survey results against those from the 2008 survey. The 2010 technique to derive comparable *Product* and *Element* values will produce results that may not be compatible with the 2008 results. It will not be possible to resolve this uncertainty until there is a string of survey results using the 2010 survey structure. If a discernable trend becomes obvious after four or five executions of the survey, then it may be possible to ascertain whether the 2008 survey results are comparable to those designed from the 2010 questions.

The structure and format of the questions also may induce some degree of uncertainty in the results. This uncertainty is tied in with the definition of quality discussed in Section 4.2.2.1. The questions and the answer categories in 2010 were redesigned from 2008 in an attempt to provide participants with a set of questions that would permit input of their quality scores in a reference framework that matched their usage of the road weather information resources. Hopefully, the new question set samples their sense of quality in a manner that best fits the user's perception of quality. This should limit some of the uncertainty potentially caused by the sampling technique.

4.2.2.3 Statistical Methods Used in the Analysis

Much of the analysis from the 2008 and 2010 surveys is based upon the statistical analysis of variance of the combined responses from each set of survey responses. The responses from all respondents may be divided into defined sub-groups (e.g., Advisory, Control, and Treatment groups), but the results from each of these groups then becomes another data set that can be used for statistical analysis and comparison. Thus, statistical computations are involved in the generation of nearly all of the analyses and comparisons of results between surveys. The numbers generated by the statistical operations are of most interest; however, if one does not understand whether the statistical tool yields a valid or reasonable result, the conclusions made from looking at the numbers may be incorrect. There are a few key assumptions about the statistical methods and the data from the surveys that may impact the validity of the statistical results and the subsequent interpretation of the results. This section looks at factors in the computational process that may create uncertainty in the results.

1. The use of the Likert scoring scale influences the distribution of the average scores.

The quality and importance results in this report have been converted from the original selections made by each participant from each of the multiple choice questions associated with each Product Component. The users picked the most appropriate text response from the list of options that defined the quality or importance in a five-step range from Very Low to Very High. These text responses were converted to the Likert scores ranging from 1 = Very Low to 5 = Very High. The individual responses were still discrete numerical values; however, once these responses were grouped and statistically averaged, the average now becomes a value on a continuous scale from 1 to 5. In the mental process of assessing quality to answer a survey question, each participant would rate the quality for each question on a continuous scale within the range from 1 to 5. In order to respond to the survey the participant would need to determine which answer option provided the closest scoring option. This is a common process associated with surveys, multiple-choice tests, or most classifications humans perform. However, the process may tend to bias the results, since respondents will need to shift their true answers up or down to fit one of the multiple choice options. Theoretically, the rounding up and down should compensate for this process but the compensating effect requires a large sample set which was not present in the 2008 and 2010 surveys.

The discrete nature of the responses also comes into play in the comparison of results between survey years. With a small sample size, the effects of a change from one Likert score to another can significantly change the results. For example, if the majority of the participants are wavering between two responses (say, moderate and high) then their mental score on the Likert scale would actually be somewhere near 3.5. Suppose in the first survey year most of these individuals select moderate in the survey and the resulting survey computed average yielded 3.1 because of the predominance of moderate (3) answers. Then, in year two, most of these individuals lean the other direction and select high (4). The resulting average might be closer to 3.9. In reality, their true assessed mental average did not change much (still about 3.5) but the survey results are distinctly different. This example is probably not totally but it illustrates the effect that the use of the Likert scale can potentially have on the results. One might expect that changes up in the Likert scores would be countered by changes down in the Likert scores by other participants. However, results suggest that the ongoing assessment of quality averages around 4 and the distribution is skewed toward lower scores. This implies that a good share of respondents are entering Likert scores of 3 or 4 and are likely at the point where their actual assessment of quality would need to either pushed upward to 4 or downward to 3. The effect of the scoring scheme is not an obvious factor but does offer a source for uncertainty in the computed averages, especially where the number of responses is small.

2. The sample size limits the statistical significance of the results and creates uncertainty in the analysis.

A very rough estimate of the total number of DOT users in the United States who actively use road weather information to support their decisions is 16,000 individuals (1,600 at a management level, 4,400 at a supervisory level, and 10,000 at a field operation level). The number of participants in the 2008 survey was 26 and the number in 2010 was 37. Based upon these numbers the sample size represents 0.2% of the DOT population actively using road weather information. From a statistical significance approach, if the desire is to be 95% confident that all of the answers are accurate within ±4% of findings, the sample size from a population of 16,000 needs to be at least 580 responses. To get a representative sample from the 1600 individuals at the management level, there would need to be at least 437 responses. The word 'responses' was used rather than participants because participants have the option to skip sections of the survey. The response rate for any given *Product Component* attribute in the 2010 survey was just under 50% of the participants. This implies that the sample size necessary to provide the level of confidence desired needs to be in the range of 900 to 1000 participants.

The uncertainty associated with the sample size becomes very evident when statistical significance is applied to the participation in the 2010 survey. On average, there were 18 responses to each of the *Product Component* questions. The potential survey group was the managers at state headquarters who use or are familiar with the use of road weather information within their organization. The group included individuals from all three management strategies. For estimation purposes, this group probably totals around 200. Using the sample size of 18 responses and the population size of the 200, the confidence interval for the scores is ±22%. This confidence interval appears much larger than what the results from the two surveys seem to portray; however, the statistical analysis is a strong indication that the sample size used in the two surveys limits any expectation that the numerical results generated from the survey are absolute representations of the assessed quality of road weather information for the respective survey years. Further, the uncertainty in the computed results for any given year makes the assessment of change in quality a questionable result.

4.2.2.4 Human Factors

Humans participate in surveys all of the time and view any particular survey with a set of preconceived notions. These notions impact how an individual responds to a particular survey and therefore they have the potential to modify the participant's actual assessment of the parameter being measured. The influences of these personal notions would be difficult to measure directly but they must be considered as they relate to the Baselining results.

1. Interpretation of the survey questions

The questions in the 2010 survey were configured to extract the quality and importance of road weather information using a straightforward multiple-choice format. The quality assessment was structured to assess quality by the six attributes. The questions were stated simply and were intended to elicit a response that rated the quality or value of *Product Component* for the specified attribute. The questions were designed *a priori* to acquire a specific quality score for the designated attribute. The design seemed to be free of confusion. However, survey participants viewed certain questions differently than was designed. For example, in the question on the Currency/Latency of Weather History data a survey respondent left this comment, "Kind of a strange question. This is history of weather. Of course it won't be current." The question was actually intended to determine whether the specific components in the Weather History Product were current and not delayed in the process of going from the source (NWS) to the user. The survey participant saw the Currency/Latency attribute relating to the parameter and not the delivery of the parameter. Other instances of interpretations different from the original intent of the survey can be seen in user comments. This fact implies that one participant may be answering a different question than others taking the survey and not answering the question intended. This would make the average of all responses an imperfect representation of the level of quality for that specific question. Possibly, all respondents interpreted the question in the same manner as the participant who left the comment. In that case, the analysis done by the BAH team could provide an incorrect interpretation of the scores.

2. Personal scaling and weighting criteria

Each individual responds to the rating scales in surveys in a different way. Some are critical and tend to input lower scores sensing that the highest scores should be reserved for perfection or near perfection. Others are more affirmative, and tend to respond with higher scores throughout the survey. If the sample size is large, the influence from any single user will be masked by the responses of other users. This may not be the case for small sample sizes such as those involved in the 2010 survey. Of particular note is the effect of the one Control strategy user who input scores of 5 for nearly all attributes for those *Product Components* the respondent chose to answer. Those *Product Components* and the *Products* and *Elements* derived from the *Product Components* rose to the top of the rankings in each of the analyses. The rankings were different if the scores from this one individual were deleted. Therefore, for small sample sizes, the results may reflect an influence from a small number of individuals and not reflect the true quality score that would accrue from a larger sample size.

A related situation relating to a user's selection of a scoring level also appears to affect the results. It deals with the tendency to submit scores reflecting a similar value across all attributes of one *Product Component* or to submit similar scores consistently in a series of *Product Components* within a given *Product*. The evaluation team saw this pattern at several locations in the 2010 survey. The source may be psychological or due to the mechanics of the survey input process. When a survey respondent reached a *Product Component* that he/she valued highly, the entire set of attribute scores tended to be higher. Where a *Product Component* was perceived as having limited value, the scores for all attributes were lower. When this pattern was seen for successive components in a give product, it was suspected that the individual respondent was placing similar value on all components rather than rating each component independently. Figure

3 presents the typical page with which survey participants dealt. Recurrent scores could reflect user bias toward the components of a given *Product*, but it could also reflect the mechanics of entering quality scores. As a measure of expediency, an individual wanting to move through the survey quickly might select a measure of quality for the first component and then select radio buttons in the same column for the remainder of the components. Either way, these factors had an effect on the scores with differing levels of influence in different parts of the survey and create an uncertainty that would be hard to eliminate.

3. Mood of the participant

A respondent's attitude toward involvement in the survey may also affect the quality responses. For an individual who ends up participating in the survey by directive or under duress, the frustration of the situation is likely to translate into the responses that are different from the participant's true sense of quality. In a similar way, participants who start the survey with good intentions to provide an accurate sense of quality can get tired of the survey process. As they tire, their attitude toward the survey changes and the original intent is lost. Thus, attitude is another unknown human factors influence that most likely has some level of influence on the results.

4. Experience of most recent season influencing primary use of road weather information

Each survey is expected to capture the participant's assessment of quality at the time of the survey and it would be anticipated that that assessment would be primarily influenced by the quality of the information received recently or at least in last six months. If the conditions in the last six months had been benign and the weather information had been reasonably good in support of the participant's decision requirements, the survey scores are likely to remain consistent with previous scores or even biased in a positive direction. On the other hand, if conditions had been unusually harsh and the participant experienced difficulty dealing with the conditions or was stressed by budget issues, lack of materials, and numerous incidents, quality scores will likely be biased in a negative direction. Therefore, what happened in the period before the survey may influence the results in that survey year, and therefore impact the comparison with the results of another survey year.

Any one of these four human factors effects may cause slightly inconsistent results that are difficult if not impossible to determine empirically. Nevertheless, their influence is likely to cause enough "noise" in the results that one cannot accept the result without some uncertainty that they are not without error.

4.3 Respondents' Comments

Survey participants were given the opportunity to make comments regarding the content of the question sets at the end of each of the *Product* questions (see Figure 3 for an example). In addition, the participants could answer a final set of questions just prior to the final submission of their survey results. Since there were questions for each of the attributes and for importance, comments were possible for each of the six quality attributes and the importance for each of the *Products*. Comments were submitted throughout the survey with the number of comments per *Product* section varying from 0 to 15% of the total number of respondents. Two or three participants submitted comments to a majority of the question sets. The entire set of comments is provided in Appendix B. The organization of the listing of comments follows the format of the survey and includes at least one row for each *Product* and its attributes. Many of the rows contain no comments. Several attribute/*Product* pairs received multiple comments. Each of the comments was numbered sequentially to permit an easy way to reference the comments.

A number of the comments highlighted user needs or stated issues that indicated specific user needs. Examples include:

#	PRODUCT	PARAPHRASED COMMENT
8	Weather Summary	Need to have all precipitation reports in rate, not Y/N
11	Weather Summary	Need pavement temperature
12	Weather Summary	Need short term precipitation amounts and rates
34	ESS Current Conditions	Accurate precipitation type and rate are a must
40	ESS Current Conditions	Type of precipitation is a necessity
45	ESS Current Conditions	ESS data needs to be updated once every 6 min
53	ESS Current Conditions	Most important parameters are least accurate
73	Pavement Forecast	Need snow forecast amounts in map format
77	Road Condition Report	State input method is time consuming
82	Watches and Warnings	Time step of 15 min may be insufficient for use
87	MDSS	Decision support needs to add contingency planning
95	Cameras	Need good nighttime visibility

Comments also pointed out both benefits and debits of the road weather resources. Some specific comments were:

#	PRODUCT	PARAPHRASED COMMENT
41	ESS Current Conditions	Can't count on chemical concentration information
56	ESS Histories	Critical for litigation
72	Pavement Forecast	Chemical concentration data not used
76	Road Weather Alerts	Not accurate
78	Watches and Warnings	Not accurate enough
91	MDSS	Too complex for many users
94	Cameras	Night views are not accurate enough
100	Radar	Future radar needs additional development effort

Perhaps the most apparent and disconcerting observation in reading the comments was that many of the comments did not relate to the content being addressed in that section of the survey, or they represented a different perspective from that intended by the design team. Of particular note are responses about pavement parameters and MDSS capabilities in the NWS Weather Summary and Weather History sections. This indicated the survey participants did not understand they were answering questions about specific types of road weather bulletins in different sections of the survey. It appears respondents glossed over the descriptions before the question sets and just started answering questions as if the questions were intended to address all of the road weather information possible. When the user got to the end of the question set and it didn't include the parameters that were important to the user, he/she left a comment that a specific parameter missing from the list was really the important criterion in the user's decision process. These users failed to recognize that the parameter of their interest was not provided in this *Product* and was not delivered by the provider of the *Product*.

The question sets were also designed to ask questions regarding the quality of the components in the question list (see Figure 3) relative to one specific attribute. This was done intentionally to make it easy for the survey participant to focus on the quality of a single quality attribute for all the components in one sequence of answers. Several comments indicate the respondents were not aware of this organization, and they referred to some other quality metric rather than the one being tested at the time. Similarly, there are indications that survey participants did not fully understand the definitions of certain attributes. The most notable was the reference to the relevancy of the Currency/Latency and Timeliness/Reliability questions for the Products dealing with historical information (comments 37, 57, and 59). These comments imply that additional consideration is necessary in the structure of the survey to assure survey participants completely comprehend what they are rating.

At the end of the survey, there was a final opportunity to make comments and/or suggestions about the survey. Three of the participants specifically indicated the survey was too long. These written comments about the survey length were also echoed by verbal communications between several of the participants and the BAH team regarding the user's frustration with the length of the survey.

4.4 Summary of Key Findings

The pavement weather forecast parameters, precipitation observations, road closure alerts, and camera images of the roadway are the most important road weather information requirements of the DOT personnel who participated in the 2010 survey. These components reside at the top of the importance list in Table 24. Historical information; observations dealing with relative humidity, dew point temperature, chemical concentration, and wind direction; and future radar were the items that held positions at the low end of the importance list. Although these items were at the lower end of the importance scale, they still had importance scores of 3.5 to 4 or between medium-high to high.

Table 24. Product Component importance ranked from highest mean score to lowest

Rank	Product Component	N	Rank	Product Component	N
1	Pavement Forecast - Precipitation Type	14	**47**	ESS Observations - Freeze Point Temperature	21
2	Pavement Forecast - Precipitation Start Time	14	**48**	Flood Warning - Current flood stage	9
3	Pavement Forecast - Precipitation End Time	14	**49**	Pavement Forecast - Wind gust	14
4	Pavement Forecast - Snow Amount	14	**50**	ESS Histories - Precip Type	10
5	Pavement Forecast - Snow Rate	13	**51**	ESS Histories - Pavement Temperature	10
6	Pavement Forecast - Pavement Temperature	15	**52**	Map - Wind Speed	21
7	Pavement Forecast - Precipitation accumulation	14	**53**	Weather Summary - Air Temperature	30
8	Pavement Forecast - Probability of Precipitation	14	**54**	Zone Forecast - Wind Direction	12
9	Road Weather Alerts - Road closure	9	**55**	Zone Forecast - Wind Speed	12
10	Weather History - Snow Amount	11	**56**	ESS Observations - Wind Speed	25
11	Pavement Forecast - Precipitation rate	14	**57**	Radar - Storm tracks	19
12	Pavement Forecast - Wind Speed	14	**58**	Road Condition Report - Road condition	14
13	Pavement Forecast - Probability of Precipitation Type	14	**59**	ESS Observations - Air temperature	25
14	Weather History - Precip Amount	11	**60**	ESS Histories - Precip Start	10
15	Camera - View of road	27	**61**	ESS Histories - Precip End	10
16	Weather Summary - Precip Amount	29	**62**	Map - Air Temperature	21
17	Pavement Forecast - Weather	13	**63**	Weather History - Dew Point	11
18	Weather Summary - Snow Amount	28	**64**	Pavement Forecast - Chemical Concentration	11
19	Weather Summary - Precip Type	30	**65**	Pavement Forecast - Relative Humidity	14
20	Pavement Forecast - Pavement Condition	15	**66**	Map - Freeze Point Temperature	17
21	Pavement Forecast - Air Temperature	14	**67**	ESS Observations - Dew Point	24
22	Pavement Forecast Dew Point	14	**68**	ESS Observations - Wind Gust	24
23	Pavement Forecast - Wind Direction	14	**69**	Map - Dew Point	19
24	ESS Observations - Precip Type	23	**70**	Map - Wind Direction	21
25	ESS Observations - Pavement Temperature	25	**71**	Flood Warning - Forecasted flood stage	9
26	Road Weather Alerts - Alert parameters	13	**72**	ESS Observations - Wind Direction	25
27	Watches and Warnings - Severe weather warnings	21	**73**	Pavement Forecast - Cloud Cover	14
28	Camera - View of weather	26	**74**	Zone Forecast - Minimum Temperature	12
29	Weather Summary - Weather	30	**75**	ESS Histories - Wind Direction	10
30	Zone Forecast - Weather	12	**76**	ESS Histories - Wind Speed	10
31	ESS Observations - Pavement Condition	25	**77**	Radar - Future radar	20
32	Radar - Radar loop	20	**78**	ESS Histories - Dew Point	9
33	Watches and Warnings - Severe weather watches	21	**79**	Weather Summary - Wind Direction	30
34	Weather History - Wind Direction	11	**80**	ESS Histories - Air Temperature	10
35	Weather History - Wind Speed	11	**81**	ESS Histories - Relative Humidity	10
36	Map - Pavement Temperature	21	**82**	ESS Histories - Pavement Condition	10
37	Zone Forecast - Probability of Precipitation	12	**83**	Zone Forecast - Maximum Temperature	12
38	Pavement Forecast - Percent Probability of Frost	15	**84**	Weather History - Relative Humidity	11
39	Radar - Radar loop with precip type coloring	21	**85**	Weather Summary - Dew Point	27
40	Pavement Forecast - Visibility	13	**86**	Map - Relative Humidity	20
41	Weather Summary - Wind Speed	30	**87**	Map - Chemical Concentration	16
42	Watches and Warnings - Special weather statements	20	**88**	ESS Observations - Chemical Concentration	21
43	MDSS - Treatment recommendations	10	**89**	ESS Histories - Chemical Concentration	8
44	Map - Precip Type	21	**90**	ESS Histories - Freeze Point Temperature	8
45	Map - Pavement Condition	21	**91**	ESS Observations - Relative Humidity	24
46	Weather History - Air Temperature	11	**92**	Weather Summary - Relative Humidity	27

When ranked, the overall quality scores of the *Product Components* in the 2010 survey closely followed the distribution of the ranking of the importance scores. Pavement Forecast components dominated the top of the list along with the Watches and Warnings and Radar loops (see Table 25). The lower end of the overall quality scores was populated with all parameters dealing with pavement condition (pavement condition, chemical concentration, and freeze point), observations of precipitation type and amount, and historical information. The actual scores of the parameters at or near the top end of the scale ranged from 4.0 to 4.35 while the scores at the bottom of the scale ranged from 3.0 to 3.75.

Table 25. *Product Component* overall quality ranked from highest mean score to lowest

Rank	Product Component	N	Rank	Product Component	N
1	Pavement Forecast - Wind Direction	87	47	Radar - Storm tracks	102
2	Pavement Forecast - Wind Speed	87	48	Road Weather Alerts - Road closure	54
3	Pavement Forecast - Weather	87	49	ESS Histories - Air Temperature	50
4	Pavement Forecast - Air Temperature	87	50	ESS Histories - Relative Humidity	50
5	Pavement Forecast - Precipitation Type	88	51	ESS Observations - Pavement Temperature	149
6	Pavement Forecast - Wind gust	87	52	ESS Histories - Pavement Temperature	50
7	Pavement Forecast - Precipitation Start Time	86	53	ESS Histories - Precip Type	48
8	Pavement Forecast - Precipitation accumulation	86	54	Pavement Forecast - Percent Probability of Frost	89
9	Pavement Forecast - Precipitation End Time	86	55	Road Weather Alerts - Alert parameters	78
10	Pavement Forecast - Pavement Temperature	91	56	Weather Summary - Dew Point	152
11	Pavement Forecast - Probability of Precipitation	88	57	Map - Pavement Condition	105
12	Pavement Forecast - Dew Point	85	58	Weather History - Air Temperature	70
13	Pavement Forecast - Snow Rate	85	59	Weather History - Wind Direction	70
14	Pavement Forecast - Probability of Precipitation Type	88	60	Weather Summary - Precip Type	182
15	Pavement Forecast - Snow Amount	86	61	MDSS - Treatment recommendations	60
16	Pavement Forecast - Relative Humidity	85	62	Weather History - Wind Speed	70
17	Pavement Forecast - Precipitation rate	86	63	Map - Freeze Point Temperature	90
18	Watches and Warnings - Severe weather warnings	135	64	Flood Warning - Current flood stage	54
19	Pavement Forecast - Pavement Condition	91	65	Weather Summary - Relative Humidity	151
20	Pavement Forecast - Cloud Cover	79	66	Weather History - Dew Point	69
21	Radar - Radar loop	126	67	Weather Summary - Weather	183
22	Watches and Warnings - Severe weather watches	129	68	Zone Forecast - Minimum Temperature	71
23	Map - Wind Direction	128	69	Zone Forecast - Maximum Temperature	71
24	Weather Summary - Air Temperature	177	70	Zone Forecast - Wind Speed	71
25	Camera - View of road	167	71	ESS Histories - Precip Start	47
26	Watches and Warnings - Special weather statements	122	72	ESS Histories - Precip End	47
27	ESS Observations - Air Temperature	149	73	Zone Forecast - Wind Direction	71
28	Radar - Radar loop with precip type coloring	126	74	Zone Forecast - Weather	71
29	Map - Wind Speed	128	75	Radar - Future radar	114
30	Pavement Forecast - Visibility	81	76	Weather History - Relative Humidity	69
31	Weather Summary - Wind Speed	183	77	Weather Summary - Snow Amount	166
32	ESS Observations - Wind Speed	149	78	Zone Forecast - Probability of Precipitation	71
33	Map - Air Temperature	128	79	Weather Summary - Precip Amount	177
34	ESS Observations - Wind Direction	149	80	Flood Warning - Forecasted flood stage	54
35	Weather Summary - Wind Direction	183	81	ESS Observations - Precip Type	139
36	Map - Pavement Temperature	116	82	Map - Chemical Concentration	79
37	ESS Observations -Wind Gust	149	83	Pavement Forecast - Chemical Concentration	53
38	Map - Dew Point	121	84	ESS Observations - Pavement Condition	147
39	Camera - View of weather	161	85	Road Condition Report - Road condition	88
40	Map - Relative Humidity	124	86	ESS Histories - Pavement Condition	49
41	ESS Observations - Dew Point	142	87	Weather History - Snow Amount	64
42	Map - Precip Type	124	88	Weather History - Precip Amount	70
43	ESS Histories - Wind Direction	50	89	ESS Observations - Freeze Point Temperature	123
44	ESS Histories - Wind Speed	50	90	ESS Histories - Freeze Point Temperature	32
45	ESS Histories - Dew Point	44	91	ESS Observations - Chemical Concentration	126
46	ESS Observations - Relative Humidity	148	92	ESS Histories - Chemical Concentration	37

The Accuracy/Precision quality attribute seemed to be the most discriminating attribute of the six attributes. Accuracy was mentioned several times in the comments and accuracy scores seemed to oppose the importance scores. This fits with the sentiment reflected in comment 53, "With exception of air temerature (sic) the most important information is the least accurate, this kills us!!" Table 26 lists the *Product Components* ranked according to their Accuracy/Precision attribute scores. The *Product Components* with the highest accuracy scores were the non-precipitation ESS and Weather Summary observations, Camera images, and a few of the non-precipitation Pavement Forecast parameters.

Table 27. *Product Component* Accuracy/Precision quality attribute ranked from highest mean score to lowest

Rank	Product Component	N	Rank	Product Component	N
1	ESS Observations - Wind Direction	24	47	Weather Summary - Relative Humidity	23
2	Pavement Forecast - Wind Direction	15	48	Weather Summary - Precip Type	31
3	Camera - View of road	28	49	Watches and Warnings - Severe weather watches	22
4	ESS Observations - Air Temperature	24	50	Road Weather Alerts - Alert parameters	13
5	Pavement Forecast - Air Temperature	15	51	ESS Observations - Pavement Temperature	24
6	ESS Observations - Wind Speed	24	52	Pavement Forecast - Precipitation Start Time	15
7	Pavement Forecast - Cloud Cover	13	53	Pavement Forecast - Snow Amount	15
8	Camera - View of weather	27	54	Pavement Forecast - Visibility	14
9	Pavement Forecast - Wind Speed	15	55	Radar - Radar loop with precip type coloring	21
10	Pavement Forecast - Weather	15	56	Weather Summary - Weather	31
11	ESS Observations - Dew Point	22	57	Map - Precip Type	21
12	ESS Observations - Relative Humidity	23	58	Zone Forecast - Maximum Temperature	12
13	ESS Observations - Wind Gust	24	59	Pavement Forecast - Precipitation End Time	15
14	Weather Summary - Wind Direction	31	60	Flood Warning - Current flood stage	9
15	Weather Summary - Air Temperature	30	61	ESS Histories - Precip Type	8
16	Weather Summary - Wind Speed	31	62	Pavement Forecast - Precipitation rate	15
17	Weather History - Wind Direction	13	63	Pavement Forecast - Pavement Condition	15
18	Pavement Forecast - Dew Point	14	64	Road Condition Report - Road condition	15
19	Pavement Forecast - Relative Humidity	14	65	Radar - Storm tracks	17
20	Pavement Forecast - Precipitation Type	16	66	Zone Forecast - Minimum Temperature	12
21	Watches and Warnings - Severe weather warnings	23	67	Map - Pavement Condition	17
22	Map - Wind Direction	22	68	ESS Histories - Pavement Temperature	8
23	Map - Wind Speed	22	69	Zone Forecast - Wind Speed	12
24	Pavement Forecast - Precipitation accumulation	15	70	MDSS - Treatment recommendations	10
25	Weather History - Dew Point	12	71	Map - Freeze Point Temperature	15
26	Map - Air Temperature	22	72	Pavement Forecast - Percent Probability of Frost	16
27	Radar - Radar loop	21	73	Zone Forecast - Wind Direction	12
28	Watches and Warnings - Special weather statements	20	74	Weather Summary - Precip Amount	30
29	Road Weather Alerts - Road closure	9	75	ESS Observations - Precip Type	21
30	Weather Summary - Dew Point	24	76	Zone Forecast - Weather	12
31	ESS Histories - Air Temperature	8	77	Flood Warning - Forecasted flood stage	9
32	ESS Histories - Relative Humidity	8	78	ESS Histories - Precip Start	8
33	ESS Histories - Wind Direction	8	79	ESS Histories - Precip End	8
34	ESS Histories - Wind Speed	8	80	Zone Forecast - Probability of Precipitation	12
35	Pavement Forecast - Probability of Precipitation	16	81	Radar - Future radar	19
36	Pavement Forecast - Probability of Precipitation Type	16	82	Weather History - Precip Amount	13
37	Pavement Forecast - Pavement Temperature	16	83	Map - Chemical Concentration	13
38	Pavement Forecast - Wind gust	15	84	ESS Observations - Pavement Condition	23
39	ESS Histories - Dew Point	7	85	Weather Summary - Snow Amount	27
40	Map - Dew Point	20	86	Weather History - Snow Amount	12
41	Weather History - Air Temperature	13	87	Pavement Forecast - Chemical Concentration	10
42	Weather History - Wind Speed	13	88	ESS Histories - Pavement Condition	8
43	Weather History - Relative Humidity	12	89	ESS Observations - Freeze Point Temperature	20
44	Map - Relative Humidity	21	90	ESS Observations - Chemical Concentration	20
45	Pavement Forecast - Snow Rate	15	91	ESS Histories - Freeze Point Temperature	5
46	Map - Pavement Temperature	19	92	ESS Histories - Chemical Concentration	6

At the bottom end of the scale were all *Product Components* dealing with Chemical Concentration, Freeze Point Temperature, and Pavement Condition. The scores for those *Product Components* at the top of the list ranged from 4.0 to 4.35, which were similar to the scores for the top components of the overall attribute scores; however, the scores of the lower end of the scale were 3.0 to 3.5. The unique feature of the Accuracy/Precision distribution is the number of scores at the lower end of the range. Statistically, the distribution is strongly skewed toward these negative scores (see Table 9), which caused the mean and median scores for the Accuracy/Precision attribute to be lower than the other attributes. The lower scores for the accuracy of the parameters dealing with pavement condition parameters and precipitation parameters appear to be the biggest challenge for the weather support community.

The comparison of the results between the 2008 and 2010 surveys suggests the *Element* scores for all three strategy classes were slightly higher in 2010 than in 2008, whereas the *Product* scores for 2010 were slightly lower than in 2008. However, the sample size for the two years is not large enough to support any conclusion that these differences are significant.

4.5 RECOMMENDATIONS

The results from the 2008 and 2010 surveys argue that there are three predominant issues with the current Baselining technique: sample size, length of the survey, and clarity of the survey questions. Nearly all errors and inconsistencies addressed in Section 4.2 are tied back to one of these three causes.

4.5.1 Sample Size and Sample Community

The approach used thus far in the establishment of a baseline for the quality of road weather information is to limit the survey to those individuals at the management level, typically within the headquarters facility of the various departments of transportation. This has created a potential sample group of roughly 100 potential participants who indicated they were willing to participate. Approximately a quarter to a third of this group has actually completed the survey. Because the DOT community is comprised of individuals with diverse responsibilities and therefore varied interests in the types of road weather information each uses, the 2010 survey was designed to permit users to skip types of information not used to support operational decisions. This further limits the sample size. From a statistical perspective, the sample size is insufficient to adequately permit a statistically significant analysis of the survey results (see item 2 in Section 4.2.2.3 for further discussion).

The management personnel who have participated in the 2008 and 2010 surveys have a good understanding of the influence of road weather information on the overall operations, but in many cases lack the day-to-day understanding of the influence of road weather information on operations. This survey would provide a better grasp of the value of road weather information within the DOT community if the sample included the input of DOT personnel in the field.

Increasing the sample size is probably the most effective way to minimize the inconsistencies seen in the 2008 and 2010 surveys. The small number of survey participants has allowed those who have participated to greatly influence the results. Personal preferences, or even somewhat disinterested execution of the survey at times, have created apparent biases in the results. These biases would be masked by a larger number of responses.

Therefore, the first recommendation is to increase the size of the survey group.

However, there are a couple of significant challenges to expanding participation in the survey. First, a mechanism is needed to distribute the survey to a much larger set of potential participants while still allowing a way to manage the logistics of monitoring participation in the survey. Second, the survey instrument has to be modified so the broader DOT community is willing to take the time to complete the survey. Possible approaches to this challenge are addressed in

Sections 4.5.2 and 4.5.3. The BAH research team feels it will difficult to increase participation without a modification in the perceived length of the survey and the amount of time required to complete the survey at one sitting. Thus, it is necessary to address the sample format associated with the increase in the sample community.

4.5.2 Length of the Survey

The design and structure of the 2010 survey is appropriate to acquire the type of information necessary to adequately assess the quality of road weather information and the detailed characteristics inherent in the assessment of that quality. However, the survey process needed to extract the information to establish the baseline is too lengthy to suit those willing to participate. Several of those who have provided excellent and relatively thorough responses in the 2008 and 2010 have committed a substantial amount of time to provide the baseline information already acquired. Nevertheless, these participants have commented that the survey is way too long and needs to be shortened if they are asked to participate again. The length will become an even more critical issue with the broader DOT community who are less likely to spend more than a few minutes answering a set of survey questions.

Therefore, modification of the how the survey is organized and administered seems to be the most reasonable resolution to the length issue. However, any modification must consider the structure of the database and how changes will impact the structure of the database. One solution might be to reduce the number of possible responses while not changing the fundamental setup of the survey structure or the quality metrics captured by the survey. A better approach may be to separate the survey into a series of smaller surveys, each addressing different *Products* and each performed during different periods of the survey year. There are 14 different *Products* in the current survey format. If the survey was set up to cover a reasonable number of *Product Components* during one survey session, then the survey could be divided into five sub-surveys with each survey being available for input for up to one month. The sub-survey entry periods could be separated by a two-week hiatus. The entire sequence would take approximately seven months to complete. Because some *Products* contain a large number of *Product Components* and others only one or two, the sub-surveys would cover anywhere from two *Products* and possibly up to six.

The second recommendation is to divide the survey into four or five separate survey components that would be completed over a six to seven month period.

4.5.3 Clarity of the Survey

The comments indicate confusion regarding the intent of several questions. These tangible indications were also inferred by patterns in the results from both the 2008 and 2010 surveys. Therefore, the questions need to be reformatted to assure participants answer the correct questions. This can be accomplished by making two modifications.

Third, it is recommended that each question section begin with an example of the type of bulletin that is being evaluated.

Survey participants are more likely to jump over text that describes the *Product* than pass up an example of the type of data with which they normally deal. If participants see what they are evaluating, then it is likely they will have a reference frame that focuses their efforts on the *Product* under review.

Fourth, it is recommended the questions need to be simplified and put in terms that the respondents understand.

For example, in Figure 3, the question should drop the "1. In your opinion, please indicate the level of Accuracy/Precision for the following Road Weather Components with Current Weather Products" and "Accuracy/Precision – How close is the observed or forecasted condition to the actual condition?" and insert "How close is the observed component to reporting the correct condition?" The questions will need to be adjusted for each of the *Products* to reference the components as advisory, forecast, historical, etc

4.5.4 Coordinated Recommendations

Fifth, it is recommended that all three of the modification types listed above be coordinated to produce a modified format for the survey.

Each modification has the potential to resolve some of the uncertainties noted in the analysis of the results from the 2008 and 2010 surveys. It is believed that, jointly, these modifications will greatly reduce most of the analysis issues encountered in this study, and minimize the frustration associated with the size of the survey and the detail being asked of the participants.

Push back in answering multiple surveys may still occur. That situation is acceptable as long as the number of individuals willing to respond to specific sub-surveys is increased appreciably since statistically the averages derived from different sample sets from the same population should be essentially the same. However, it is hoped that many participants will answer all sub-surveys if each of the surveys takes about 30 minutes or less, and each survey session is separated by at least two weeks or more.

The greatest challenge to this approach will be an effective and active "advertising campaign" to get adequate participation of DOT users. Preparation for the survey will require working with individual states or web services to establish methods to inform potential survey participants about each of the sub-surveys, how they can participate, what they will be asked to evaluate, roughly how long it will take them to do the survey, and what benefit their input will provide back to them. Although the last point is a difficult sell, a good case can be made that if service providers know what users like and don't like and where the providers need to make improvement, then the service provider will take action to resolve weak services. The survey offers an opportunity for users to critically evaluate road weather information and show their dissatisfaction with services that are not meeting their needs. Users may even express their concerns as comments. In the process, these same users will have the opportunity to indicate where road weather information satisfies their operational needs. Hopefully, if the survey is presented to the broad spectrum of DOT personnel as an opportunity to express their feelings about road weather information and the survey is seen as a relatively easy to complete, the participation level will provide the desired sample size.

APPENDIX A

RESPONDENT ENTITY

Draft MySQL Table Creation Syntax for Respondent

```
CREATE TABLE ‘respondent‘ (
	‘id‘ int NOT NULL auto_increment,
	‘participant_id‘ char(64) NOT NULL,
	‘date‘ datetime NOT NULL,
	‘comments‘ text,
	‘state‘ char(2) NOT NULL,
	‘jurisdiction‘ char(64) NOT NULL,
	‘psc‘ enum(’advisory’,’control’,’treatment’) NOT NULL,
	‘job_class‘ char(64) NOT NULL,
	‘experience‘ float unsigned NOT NULL,
	‘name‘ char(128) default NULL,
	‘advisory‘ int NOT NULL,
	‘control‘ int NOT NULL,
	‘treatment‘ int NOT NULL,
	‘phase‘ enum(’baseline’,’survey 2010’) NOT NULL,
	PRIMARY KEY (‘id‘)
)
```

PRODUCT COMPONENT ENTITY

Draft MySQL Table Creation Syntax for PC

```
CREATE TABLE ‘product_component‘ (
	‘id‘ int NOT NULL auto_increment,
	‘name‘ char(64) NOT NULL,
	PRIMARY KEY (‘id‘)
)
```

QUALITY ATTRIBUTE ENTITY

Draft MySQL Table Creation Syntax for Quality Attribute

```
CREATE TABLE ‘quality_attribute‘ (
	‘id‘ int NOT NULL auto_increment,
	‘name‘ char(64) NOT NULL,
	PRIMARY KEY (‘id‘)
)
```

PRODUCT ENTITY

Draft MySQL Table Creation Syntax for Product

```
CREATE TABLE ‘product‘ (
	‘id‘ int NOT NULL auto_increment,
	‘name‘ char(64) NOT NULL,
	PRIMARY KEY (‘id‘)
)
```

USES RELATIONSHIP

Draft MySQL Table Creation Syntax for Uses

```
CREATE TABLE ‘uses‘ (
	‘usage‘ tinyint(1) default NULL,
	‘method‘ char(64) default NULL,
	‘manner_selection‘ enum(’no direct cost’,’fee-based’) default NULL,
	‘manner_comments‘ text,
	‘cost-benefit_selection‘ enum(’great benefit’,’near equal’,’little benefit’) ‘cost-
	benefit_comments‘ text,
	‘respondent_id‘ int default NULL,
	‘product_id‘ int default NULL,
	KEY ‘respondent_id‘ (‘respondent_id‘),
	KEY ‘product_id‘ (‘product_id‘),
	CONSTRAINT FOREIGN KEY (‘respondent_id‘)
	REFERENCES ‘respondent‘ (‘id‘),
	CONSTRAINT FOREIGN KEY (‘product_id’)
	REFERENCES ‘product‘ (‘id‘)
)
```

ELEMENT ENTITY

Draft MySQL Table Creation Syntax for Element

```
CREATE TABLE ‘element‘ (
	‘id‘ int NOT NULL auto_increment,
	‘name‘ char(64) NOT NULL,
	PRIMARY KEY (‘id‘)
)
```

RESPONDS RELATIONSHIP

Draft MySQL Table Creation Syntax for Responds

DRAFT MYSQL TABLE CREATION SYNTAX FOR RESPONDS

```
CREATE TABLE ‘responds‘ (
‘selection‘ int default NULL,
‘comments‘ text,
‘respondent_id‘ int default NULL,
‘product_id‘ int default NULL,
‘quality_attribute_id‘ int(11) default NULL,
‘element_id‘ int default NULL,
KEY ‘respondent_id‘ (‘respondent_id‘),
KEY ‘product_id‘ (‘product_id‘),
KEY ‘quality_attribute_id‘ (‘quality_attribute_id‘),
KEY ‘element_id‘ (‘element_id‘),
CONSTRAINT FOREIGN KEY (‘respondent_id‘)
REFERENCES ‘respondent‘ (‘id‘),
CONSTRAINT FOREIGN KEY (‘product_id‘)
REFERENCES ‘product‘ (‘id‘),
CONSTRAINT FOREIGN KEY (‘quality_attribute_id‘)
REFERENCES ‘quality_attribute‘ (‘id‘),
CONSTRAINT FOREIGN KEY (‘element_id‘)
REFERENCES ‘element‘ (‘id‘)
```

APPENDIX B

COMMENTS FROM THE 2010 SURVEY

#	PRODUCT	ATTRIBUTE	COMMENTS
1	Weather Summary	Accuracy/Precision	The Weather stations around Washington are not all the same as there are SSI and Viasala systems. There are some trouble areas around the state where the weather stations results don't match what is actually occuring with the weather.
2			There is a disconnect between actual pavement temperatures and forecast pavement temperatures.
3			The Lufft sensor has proven to give very good results in precipitation type and amounts. We currently have very few of these units utilized.
4			Sometimes, precipitation is very accurate; other times, it's not
5			Information from our weather services provider has historically been more accurate than the national weather services information for winter weather conditions.
6		Completeness	We don't use everything in the product. Generally we utilize only the items that would affect travel plans..
7			My level of completeness would be higher if the information was more tailored to a more pin pointed level by a user defined specifi area. NWS is too general and zonal.
8			The weather stations don't report the same tye of weather information as some weather stations report the precipitation as either yes/no and others report inches per hour.
9			More detailed data is for a shorter time period usually less than 36 hours. At times it is necessary to have more detailed data for a longer period, up to 5 days.
10		Relevance	It's all relevant for winter operations even beyond snow removal operations and treatment recomendations.
11			Need Pavement Temperature and would rate that as very high
12			Relevance of precipitation and snow amounts are higher when the amounts are high or fall in a short amount of time or when combined with other factors. For instance, blizzard conditions would mean high winds in addition to snow amount.
13			Same comment as number 2. Sometimes necessary to have for a longer 5 day period.
14			The information provided generally is provided for an area such as a city. To accurately report weather that affects travel plans, we need the information at the roadway level
15		Currency/Latency	This depends on the circumstances. During precipitation events, it seems to be more current but when it is not storming, the informatoin seems to be almost static.
16			Pavement Temperature is very high in terms of Currency

#	PRODUCT	ATTRIBUTE	COMMENTS
17			Very satisfied with the currency of the data.
18			As long as the weather condition is still active when the report is received, we can utilize the information.
19		Timeliness/Reliability	Pavement Temperature generally has a high timeliness / Reliablility
20			Washington State maintains and operates more than 100 weather sations and at any given time there is typically 10 percent that are not working due to a system failure of one device or onother.
21			MDSS has improved the timliness of the forecasts. No complaints.
22			I rate this low for the general user. Unless the user has had training, the information is difficult to understand and interpret.
23		Ease of Use	Pavement Temperature would be very high in ease of use for decisions
24			Again, MDSS has packaged this information in an easy to use, electronic format.
25			The information may come in for an area. The operators would have to try to convert that to the street level to determine how that will affect travel plans.
26		Importance	Pavement Temperature is very important is supporting operational decisions
27			All items are critical in winter maintenance decision making.
28		Accuracy/Precision	limited weather reporting stations and microclimates of the west limit accurate historical data we can apply throughout the state.
29			At times the dynamics of the storm change and precip amount are in the moderate category. Some storms are very difficult to forecast.
		Completeness	
	Weather History	Relevance	
30		Currency/Latency	Kind of a strange question. This is history of weather. Of course it won"t be current.
31		Timeliness/Reliability	This depends on the location of the RWIS as not all weather sations communicate the same way. Some cellular modem, microwave, or DSL.
		Ease of Use	
32		Importance	Pavement Temperature
33			I rate this moderate as a view of the statewide accuracy. Some regions would rate higher if they are more aggressive in maintaining the equipment.
34			Accurate precipitation indication is a must.
35			IL DOT provides ESS info to Weather Provider. Chemical and Freeze point are not included in our requirements.
36	ESS Current Conditions	Accuracy/Precision	We have not done a rigorous comparison of all parameters as yet.
37			We found that surface condition was only accurate 20 percent of the time during a winter event.
38			Existing legacy RWIS are in the process of being phased out with new RWIS being deployed within construction activity. No data to rate at this time.
39			Our ESS system has been mostly not working over the past year due to a system

#	PRODUCT	ATTRIBUTE	COMMENTS
			conversion and new vendor. Questions will be answered on the basis of performance prior to conversion.
40	ESS Current Conditions	Completeness	We don't have dew point as a reading on our current system. Not all of our precip sensors tell us the type of precip (I wish they did). Pavement temperature is a falicy, we get a reading of the temperature of the puck in the road. Not all of our sensors give us pavement condition as a reading. Not all of our sensors give us chemical concentration read outs. None of our sensors, except a few bridge spray systems, give us freeze point temperatures
41			We cannot count on the chemical concentration data that comes to us.
42		Relevance	We also use a "grip" value from the non-invasive pavement sensors.
43			We expect the data to be very useful once RWIS is fully deployed.
44			chemical concentration and freeze point tend to be less representative of conditions over an area than the other parameters
45		Currency/Latency	This really depends on how the server is configured and how often the sites are polled. 15 minutes works, but I wish it was every 6 minutes.
46			sometimes there is a delay in data....we are working internally to correct this.
47			When our system is delivering the data to theirs all if fine. Our system is getting some age and it has been know to break down. This stops the flow of data and timeliness is affected. When we deliver, they deliver..
48		Timeliness/Reliability	Some sites have delays due to communications. Others are timely.
49			Most of the time.
50			In the past we have had some problems with our RWIS system and have worked to provided consistently reliable information more often
51		Ease of Use	With our current interface. (SCAN WEB)
52		Importance	Year round.
53			With exception of air temerature the most important information is the least accurate, this kills us!!
54			chemical concentration and freeze point are subject to local (as little as a square inch or less) treatment and may not accurately reflect conditions over a wider area
55	ESS Histories	Accuracy/Precision	I don't use the information in this way
56			Critical information for litigation
		Completeness	
		Relevance	
57		Currency/Latency	Since we are looking at Past Conditions, the latency question is mute.
58		Timeliness/Reliability	not sure how this question differes from the previous question....either current RWIS

			data is there or it isn't
59			Since we are looking at Past Conditions, the timeliness question is mute.
		Ease of Use	
		Importance	
60		Accuracy/Precision	High, because we can drill down to more specific information by route.
61			Currently considering proposals from various vendors for this product.
62		Completeness	And it is all on one map in one location. We can layer the information we need.
63			Would like better snow totals map.
		Relevance	
64	Regional Weather Map	Currency/Latency	It is as good and as current as the information we feed it.
65			Information updated every 3 hours.
66		Timeliness/Reliability	We do not have great confidence in the reliability of our chemical concentration sensors
67		Ease of Use	We are trying to upgrade to a more accurate precipitation sensor that can indicate precipitation type.
		Importance	
		Accuracy/Precision	
68		Completeness	Much like NWS and not detailed to site specific locations.
69			Should include sub surface temp using sub surface probes for load restriction analysis.
70	Zone Forecast	Relevance	It gives us a general sense of what to expect.
		Currency/Latency	
		Timeliness/Reliability	
71		Ease of Use	Text discussions are very good if you can figure out the zone they are describing.
		Importance	
72		Accuracy/Precision	chemical concentration as reflected in the current and future condtions is extremely important, however I don't look at the chemical concentration, only the effect it has on predicted pavement condition (wet, slush, ice, etc)
73		Completeness	Would like better presentation (map view) of total snow through the forecast period and total snow after the storm.
	Pavement Forecast	Relevance	
74		Currency/Latency	Value Added Meteorology is a wonderful thing.
		Timeliness/Reliability	
75		Ease of Use	Providing training has been given.
		Importance	
76	Road Weather Alerts	Accuracy/Precision	Would love to use, on account of innacuracies warnings are disregarded
		Completeness	
77		Relevance	Programming the regional text alerts is time consuming in a statewide application. One has to program county by county to establish a grid. It is more user friendly for localized alerts.

		Currency/Latency	
		Timeliness/Reliability	
		Ease of Use	
		Importance	
78		Accuracy/Precision	We use the NWS CAP XML feed to post Warnings on our 511 map.
			Just not reliably accurate.
79		Completeness	There are some issues with the formatting of their XML feed. It's not formatted in the way they say it should be (mainly issues with the warning times).
80			Sometimes if the statement covers a large area it is more difficult to understand the timing for the condition in a specific location
81	Watches & Warnings	Relevance	Just wish it was more accurate.
82		Currency/Latency	Severe weather can move through very quickly 15 minute updates may not follow the storm quickly enough.
		Timeliness/Reliability	
83		Ease of Use	As mentioned above, there are issues with the formatting of the time.
84		Importance	Operators still have to interpret the information and post to the roadway level.
85		Accuracy/Precision	We are having a hard time getting on top of this issue. Some operators find it hard to break past practices and try something different. Some have recorded what the recommendation is and shown how wrong it was but we haven't seen what was reported to the system that would have driven that forecast, so we are not sure we gave the system proper observations.
86			You need good ground truth coming in from the AVL/MDC in order to get good recommendations back.
		Completeness	
87	MDSS	Relevance	Only offers guidance for most likely scenario. What if the storm shifts start time or phase? How will the treatments and crew scheduling be different under slightly different weather scenarios? Our current MDSSs offer no help in contingency planning.
88		Currency/Latency	If we give it current and accurate date on a timely basis.
89		Timeliness/Reliability	How the product is delivered to the users is very important.
90		Ease of Use	Very intuitive.
91			Some users believe MDSS is too complex and requires a great deal of effort to get the information they are searching for.
92		Importance	Decision support is the most important part of MDSS.
	Road Reports	Accuracy/Precision	
		Completeness	
		Relevance	

		Currency/Latency	
		Timeliness/Reliability	
		Ease of Use	
		Importance	
		Accuracy/Precision	
		Completeness	
	Flood Reports	Relevance	
	Flood Reports	Currency/Latency	
		Timeliness/Reliability	
		Ease of Use	
93		Importance	very important if the water is close to the road, otherwise not very important.
94		Accuracy/Precision	Night views are less accurate
95		Completeness	We are responsible for our own camera images. The cameras provide marginal images of road surface conditions and when weather is bad the lenses can get covered providing a less than desireable image. Also, we don't have a method of obtaining night time images, which are needed for operations. And, we need more cameras to provide better coverage.
	Camera Images	Relevance	
96		Currency/Latency	Depending on the number of preset camera angles and the polling frequency.
		Timeliness/Reliability	
97		Ease of Use	If the camera views are delivered in a format that is easy to access, such as selecting views on the radar map.
98		Importance	Camera images are a great way of getting ground truth to support the sensor data that is coming in.
99		Accuracy/Precision	If you can get the future radar.
100			The Future radar still needs more work.
101		Completeness	If you can get the future radar.
102		Relevance	If you can get the future radar.
103	Radar	Currency/Latency	If you can get the future radar.
104		Timeliness/Reliability	If you can get the future radar.
105			If you can get the future radar.
106		Ease of Use	Currently being integrated into desktop application which eliminates the need to maintain a separate Internet Explorer window for monitoring.
107		Importance	If you can get the future radar.
108	Final Comments		Just beginning to use RWIS in our region
109			On the history questions, I do use them but very seldom and only when asked.
110			Weather Integration is a priority for our 24/7 TMC which serves two state DOTs. New RWIS is being deployed; NWS conponents are being integrated into our mgmt software and vendors are being evaluated for providing additional digital information.

111		In my current management position I don't make the day to day operational decisions I used to as a Highway Superintendent or Supervisor. Therefore I don't need as much current and continuous weather information as possible.
112		I do check the RWIS cameras and associated data, in addition to weather forecasts on occasion during extreme or above average weather events, in case there is a need to mobilize additional resources to respond.
113	Final Comments	Too long.
114		I manage the ATIS CARS program, so most of these didn't really apply to me.
115		Survey too long and duplicative.
116		Survey is tooo long

Federal Highway Administration
U.S. Department of Transportation
1200 New Jersey Avenue, S.W.
Washington, DC 20590
Toll-Free "Help Line" 866-367-7487
www.ops.fhwa.dot.gov