EVALUATION OF THE *FORETELL*™ CONSORTIUM OPERATIONAL TEST: WEATHER INFORMATION FOR SURFACE TRANSPORTATION

Evaluation Strategy

July 1, 1998

Prepared for:

U.S. Department of Transportation
Joint Program Office for ITS, HVH-1
Room 3400
400 Seventh Street, SW
Washington, DC 20590

Prepared by:

Battelle
505 King Avenue
Columbus, OH 43201

TABLE OF CONTENTS

Page

LIST OF TABLES

TABLE OF CONTENTS
(Continued)

Page

LIST OF FIGURES

EXECUTIVE SUMMARY

FHWA has awarded funding to the *FORETELL*™ Consortium (Castle Rock Services, Iowa, Wisconsin and Missouri DOTs) for the development and operational testing of a multi-regional road and weather forecasting/dissemination system, in partnership with the National Weather Service (NWS) and Environment Canada (EC). An important component of this initiative is an independent evaluation of the effectiveness of the services.

FORETELL™ plans to establish an Intelligent Transportation System (ITS) Service Center to disseminate weather and pavement condition information to users. The fundamental functions of the service center will be to: use NWS and EC data sources and models to provide current conditions and forecasts; use transfer energy balance models and solar gain and snow drift algorithms to develop pavement condition forecasts; adjust weather forecast and pavement condition predictions using real time field sensor information from stationary and mobile road weather information systems (RWIS), and observations; disseminate value-added tailored information to travelers, DOT maintenance personnel, and others using available/emerging commercial and ITS traveler information media.

As with all ITS Field Operational Tests, the FHWA also will conduct an independent evaluation of the project. Battelle was selected to perform the evaluation. The first step in understanding the project and the primary evaluation issues is to develop an overall strategy for conducting the evaluation. This document defines that strategy.

FORETELL™ Consortium has recently published a System Design Concept (SDC) document. This document, dated March 1998, defines the goals and objectives of the Program, the deficiencies in weather information within the transportation system, the *FORETELL*™ approaches to address these deficiencies and the system configuration being designed to achieve their goals. The SDC is an important resource to help understand the evaluation strategy defined within this document.

Evaluation Strategy

The purpose of the independent evaluation is to assess the effectiveness of the *FORETELL*™ Program in achieving certain ARTS goals and objectives. Independent evaluations of ITS Operational Tests require a well documented structured approach to ensure meaningful results. One of the goals of the evaluation process is to determine the feasibility of the *FORETELL*™ Program and the possibility of widespread deployment.

The following fundamental principals will guide the Evaluation Team's conduct of the project evaluation:

- Extensive integration with Project Team to ensure continuity and consistency
- Strategy consistent with and supportive of ARTS Strategic Plan goals
- Focus on user decisions and operational improvements
- Utilize sound technical evaluation approaches (simple, meaningful, and achievable)
- Comprehensive in scope, but selective in practice (consistent with budget allocations).

A successful evaluation must answer some fundamental questions:

- Is the *FORETELL*™ information adding value to users beyond what they can obtain from existing sources?
- Is the new information changing users' behavior? How?
- What impact will this program have on ARTS goals and objectives (outcomes)?

The evaluation measures of success will be of two types: outputs and outcomes. The output measures evaluate the *FORETELL*™ Program system performance. The outcome measures evaluate the operational improvements achieved through deployment of the Program. Both types of measures are valid and important to the success of the evaluation. Figure E -1 illustrates the relationship between the *FORETELL*™ Program process and the evaluation goals. The project begins with improved roadway and weather information. Users of this information make decisions that affect results. The evaluation goals are shown at the bottom of the figure and will be assessed at each step in the process.

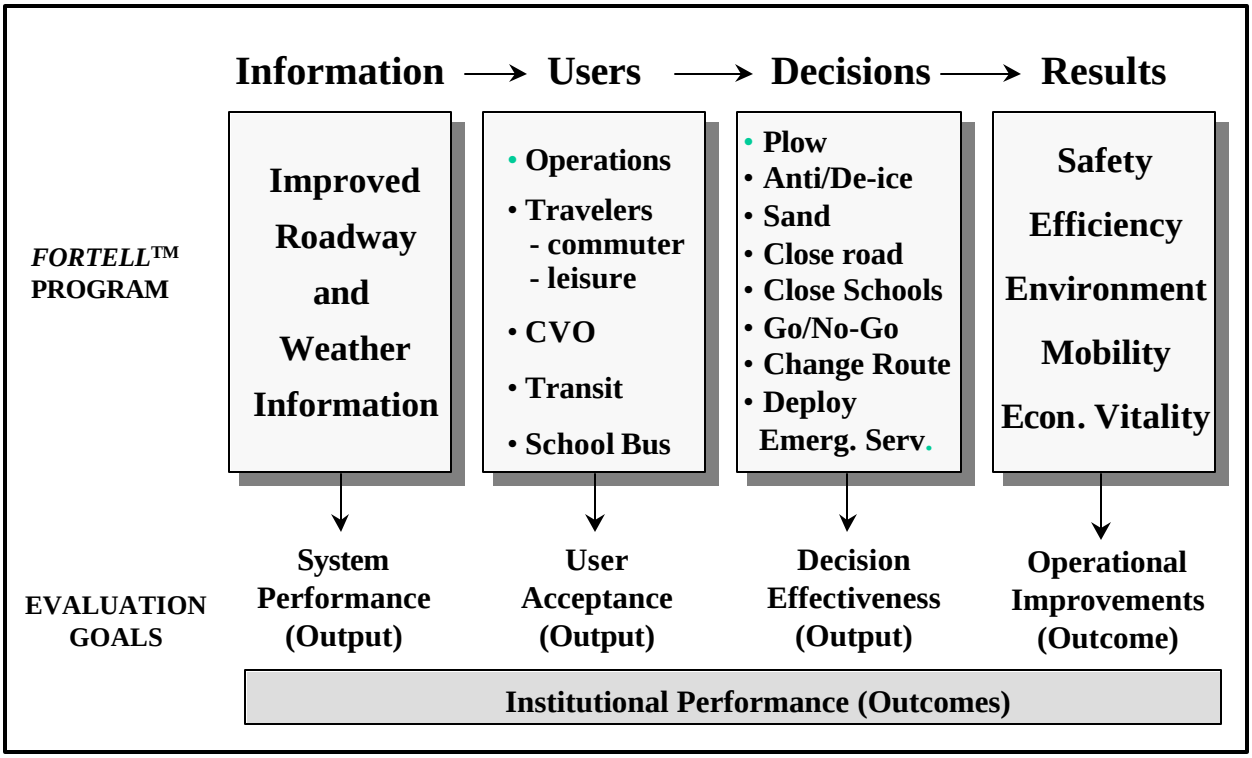


Figure ES-1. *FORETELL*™ Information Link to User Decisions and Evaluation Outcomes

The top-level evaluation goals are therefore to evaluate system performance, user acceptance, decision effectiveness, operational improvements and institutional issues. This strategy document identifies the lower-level objectives within each goal and the appropriate measures that will be used to evaluate the *FORETELL*™ Program.

1.0 INTRODUCTION

Accurate weather information is a critical element in the daily lives of most Americans. In many cases, weather information helps us determine when to take a trip, which route, or whether to go at all. It guides the actions of state department's of transportation that maintain our interstates and state highways. It also affects how and when our commerce is transported.

When weather turns wintry with snow and ice it can not only change our daily habits, it can be deadly. Over 17% of all fatal crashes occur during winter weather conditions. Of those, 60% happen in rural areas (most on non-interstate roadways). The Federal Highway Administration (FHWA) Weather Team believes more accurate and accessible weather information is the solution to these issues. FHWA recently awarded a rural ITS Operational Test to the *FORETELL*™ Consortium to demonstrate approaches to deliver accurate weather information to all who need/want it.

As with all ITS Field Operational Tests, the FHWA also will conduct an independent evaluation of the project. FHWA selected Battelle to perform the evaluation. The first step in the evaluation project is to develop an overall strategy for conducting the evaluation. This report defines that strategy.

1.1 *FORETELL*™ PROGRAM

FORETELL™ is a multi-state initiative bringing ITS together with advanced weather prediction systems to create operational highway maintenance management and traveler information systems throughout North America. *FORETELL*™ participants envision:

- developing a self-sustaining road and weather information system fully integrated within a wider basket of ITS services;
- reducing winter-condition related road deaths by at least 15%; and
- creating a viable road and weather information network across the continent.

The *FORETELL*™ 's mission is to deliver the benefits of advanced weather prediction systems and ITS technologies to travelers, shippers, and transportation system operators. The Program envisions a widely accessible real time road and weather information system that will support seamless

information sharing for travelers and highway maintenance managers. Major partners in *FORETELL*™ include state governments, private entities, Canadian agencies, and the U.S. Department of Transportation (U.S. DOT).

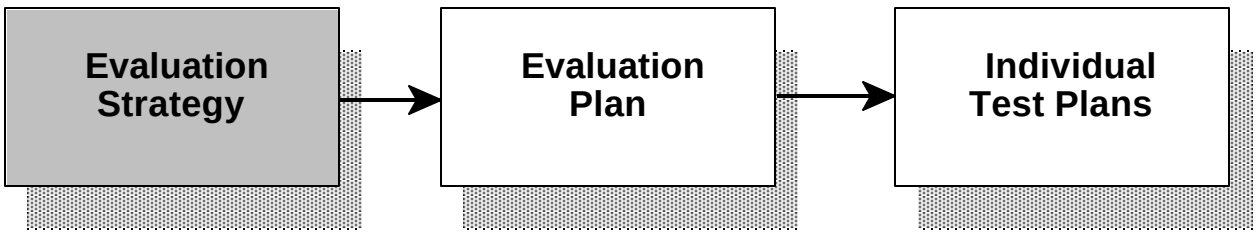
FORETELL™ Consortium has recently published a System Design Concept (SDC) document. This document, dated March 1998, defines the goals and objectives of the Program, the deficiencies in weather information within the transportation system, the *FORETELL*™ approaches to address these deficiencies and the system configuration being designed to achieve their goals. The SDC is an important resource to help understand the evaluation strategy defined within this document.

1.2 EVALUATION OVERVIEW

The purpose of the independent evaluation is to assess the effectiveness of the *FORETELL*™ Program in achieving its goals and objectives. Independent evaluations of ITS Field Operational Tests require a well documented, structured approach to ensure meaningful results.

Typically, multiple planning documents are developed that define the project evaluation. This extensive evaluation planning must be completed before the project testing can begin and is coordinated with the Project Team, obtaining their input and involvement in the process. In this case, the first step in the planning process is to develop the evaluation strategy.

This Strategy was developed by the Evaluation Team with extensive involvement and cooperation of the *FORETELL* ™ Project Team. It provides the foundation for developing the evaluation plan and individual test plans, which will define the details of each area having evaluated. The strategy will define the overall approach to the independent evaluation of the project.



1.2.1 EVALUATION KICK-OFF WORKSHOP

The Evaluation Kick-off Workshop was held in March 1998. In attendance were Gary Nelson (Mitretek), Peter Davies (Castle Rock), Fred Kitchener (Battelle), John Whited (IA DOT), Rich Naistat (National Weather Service), Dean Deeter (Castle Rock), Bill Stone (MO DOT), Bradley Skarpness (Battelle) and Ed Boselly (Battelle). Absent from the meeting were representatives from Forecast Systems Laboratory (FSL) and Environment Canada (EC).

During the meeting, the participants reviewed the *FORETELL*™ Program, brainstormed the potential evaluation goals and objectives, and prioritized (by voting) the evaluation goals and objectives. The evaluation strategy defined in this document reflects the input, comments, and suggestions of the meeting participants.

1.2.2 GUIDING PRINCIPALS

The following fundamental principles are used to guide the Evaluation Team's conduct of the project evaluation:

- Extensive integration with the Project Team to ensure continuity and consistency
- Strategy consistent with and supportive of Advanced Rural Transportation System (ARTS) Strategic Plan goals
- Focus on user decisions and operational improvements
- Utilize sound technical evaluation approaches (simple, meaningful, and achievable)
- Comprehensive in scope, but selective in practice (consistent with budget allocations)
- Provide frequent impartial feedback to *FORETELL*™ Team to enable continuous improvements.

1.3 DOCUMENT CONTENT

The following chapter describes the *FORETELL*™ program goals and their relationship to the ARTS Strategic Plan goals. Next, the evaluation strategy provides a discussion of the complexities of the activities to be implemented, the proposed questions to be answered, importance of the outputs and outcomes, proposed evaluation goals, and the user decision hierarchy. The evaluation strategy goals and objectives, responsibilities, and evaluation tools are defined. The evaluation management, future

activities, schedule, and deliverables are discussed in the final chapter.

2.0 UNDERSTANDING THE *FORETELL*™ PROGRAM

FORETELL™ 's long term plan is to provide accurate weather and road condition information to travelers, shippers and transportation system operators across North America. This ITS Field Operational Test is being implemented as a starting point and focuses on three mid-western states (Iowa, Wisconsin, and Missouri) and most of Minnesota and Illinois. A parallel program is being developed in Ontario, Canada. The *FORETELL*™ SDC document defines in detail the information to be provided and how it will be generated. Those details will not be duplicated here. Instead, a program summary is discussed below to help understand the basis for the evaluation strategy formulation.

The market analysis conducted by *FORETELL*™ identifies significant deficiencies with current approaches to weather and road condition information development, production, and dissemination. These deficiencies include:

- Lack of information and geographic coverage
- Insufficient timeliness
- Inaccuracies that result in lack of confidence in making decisions
- Lack of necessary detail
- Difficulties in acquiring information and the high cost of acquiring it.

In response to these apparent deficiencies in the current system, *FORETELL*™ plans to provide information on current conditions ("nowcasts") and forecasts of weather information and road conditions to operations staff, transit operators, commercial vehicle operators, school management, and commuter and leisure travelers. The specific information is identified in the SDC and varies by user group and need.

FORETELL™ plans to provide a "one-stop-shopping" so that users will not have to integrate information from multiple sources. The information will be disseminated primarily through the following media:

- Internet/World Wide Web
- E-mail
- Fax
- Phone/Cell Phone
- Digital Messaging
- Pagers.

The Evaluation Team is specifically interested in understanding how this improved information will address the goals and objectives of the ARTS Strategic Plan. The SDC document addresses the benefits expected to accrue from dissemination of *FORETELL*™ 's information and relates those benefits to the ARTS goals and objectives. How *FORETELL*™ expects to meet these goals and objectives is important to the Evaluation Team because these expectations of the Project Team need to be measured as part of the evaluation. Not all of these project expectations are measurable within the scope, schedule, and budget of this evaluation. Additional effort is required to prioritize areas that can be evaluated within resource and time constraints. Table 1 maps the links between ARTS goals and objectives to *FORETELL*™ project expectations.

3.0 EVALUATION CONTEXT (Understanding the Complexities)

The Evaluation Team must understand the complexities of this project to create a foundation for evaluation activities. A successful evaluation must answer some fundamental questions:

- Is the *FORETELL*™ information adding value beyond what is available to users from existing sources?
- Is the new information changing users' behavior and how?
- What impact will this program have on ARTS goals and objectives (outcomes)?

This chapter provides the background, framework, and approach for the evaluation strategy.

Table 1. Mapping of ARTS Goals/Objectives to *FORETELL*™ 's Project Expectations

ARTS Goal	ARTS Objectives	*FORETELL*™ 's Expectations
Safety & Security	• Reduce frequency of crashes • Reduce rate of crashes • Reduce severity of fatal crashes • Reduce exposure to unsafe conditions	• Reduce winter-related road deaths by 15% • Help reduce frequency, rate and severity of crashes with more detailed, accurate and timely weather and road condition information to: Maintenance crews Highway patrol Travelers • Help alleviate exposure to unsafe conditions through dissemination of information
Efficiency	• Reduce congestion and delay • Improve vehicle routing and diversion • Improve operations and maintenance resource management	• Reduce O&M cost for road maintenance Optimize labor call out Efficient fleet deployment Attention to critical times and places • Help reduce delay with more accurate weather and road condition information
Environmental Conservation	• Reduce vehicle miles traveled • Reduce emissions • Improve hazardous material response	• Improve water and air quality Efficient allocation of chemicals (less runoff into rivers/lakes) Efficient allocation of sand (less particulates in the air)
Mobility/ Convenience	• Increase percent of population with available and convenient transportation services • Improve access to services/tourist areas • Improve communications within rural areas	• Enhance travelers choices and availability • Increase number of people with access • Expand information services and communication connectivity • Assist with school closure decisions • Improve timing/scheduling of transportation services • Improve transportation service efficiency and maintenance programs
Economic Vitality and Productivity	• Improve access to and from rural communities for travel, goods, services, and information • Improve knowledge of goods, services, opportunities through en-route information and transportation service information • Improve transportation and communication facilities in rural communities	• Facilitate open and safe travel and transport to and from communities Weather information Road condition information • Support tourists transportation decisions • Support successful business and tourism activities in rural communities

3.1 IMPORTANCE OF EVALUATION OUTPUTS AND OUTCOMES

The evaluation measures of success will be of two types: outputs and outcomes. The output measures evaluate the *FORETELL*™ Program system performance. The outcome measures evaluate the operational improvements achieved through deployment of the Program. Both types of measures are valid and important to the success of the evaluation. Figure 1 illustrates the relationship between the *FORETELL*™ information, decision links and the Evaluation Goals. The five goals of this evaluation are to assess the system performance, user acceptance, decision effectiveness, operational improvements, and institutional performance The evaluation outcomes relate directly to the ARTS goals.

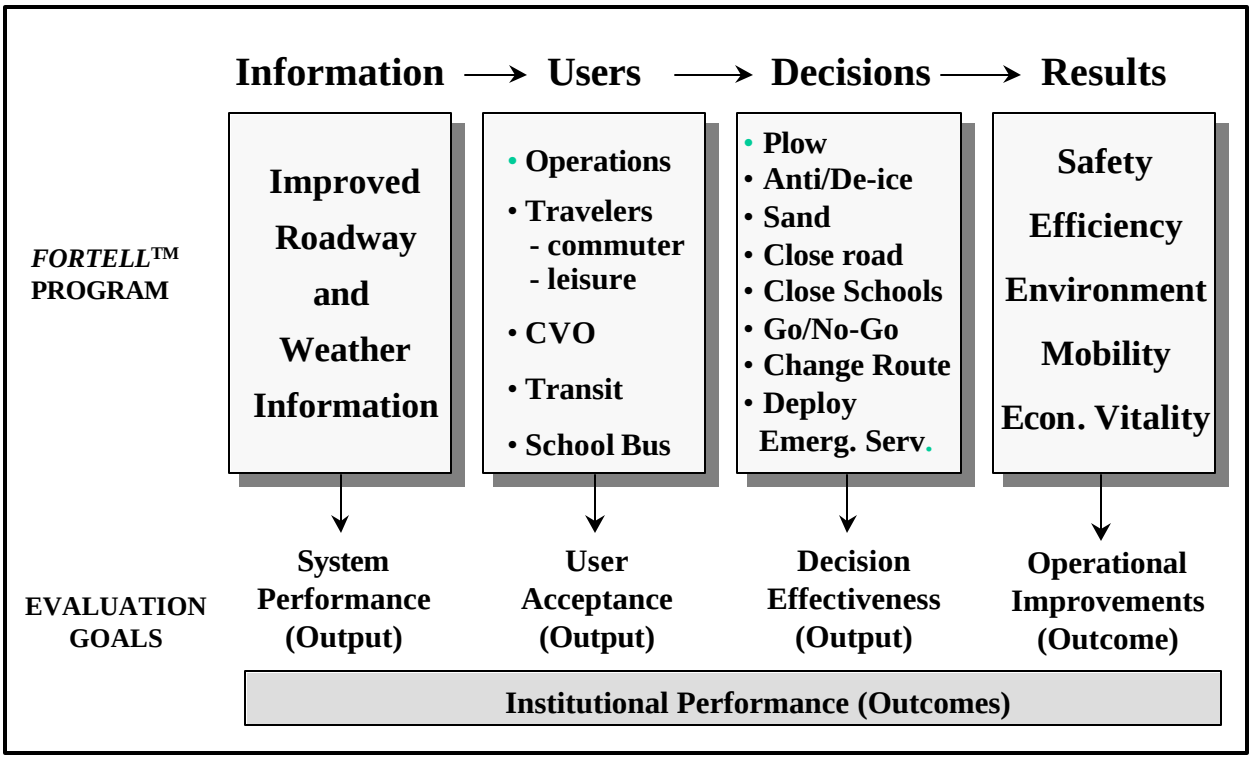


Figure 1. *FORETELL*™ Information Link to User Decisions and Evaluation Outcomes

The process begins with improved weather information from *FORETELL*™. As documented in the SDC, many processes and links with National Weather Service (NWS) data are required to achieve the desired improvement in accuracy, detail, and content of the information. There are significant system performance issues that require evaluation and documentation throughout the development of the program. Although this will not be a major emphasis of the evaluation, it is important to measure and characterize the overall performance of the *FORETELL*™ system as they relate to issues that may affect the data being disseminated.

The information is obtained by a set of users (identified in Figure 2) whose functions and information needs differ. As discussed later in this chapter, the state operations (maintenance) programs are a key factor in the program's success. A primary evaluation goal is to measure user acceptance among maintenance personnel. The evaluation will focus on assessing how information was received, how it was used, its perceived value, and the effect of the information on personnel behavior or operations. Questions to be answered in this category are defined in the next chapter.

The effectiveness of decisions made by users on the basis of program information is the focus of the evaluation at this stage in the process. Many of these decision effectiveness measures may be surrogates for evaluating the operational improvement outcomes (which are in general very difficult to measure).

Better user decisions made with improved information are expected to result in operational improvements in the system. The following operational improvements are the goals of the ARTS Strategy Plan:

- improved safety;
- enhanced efficiency;
- environmental conservation;
- enhancing mobility and convenience; and
- encouraging economic vitality.

It is the goal of the evaluation process to measure the operational improvements, since they are

the ultimate goals of the project. Measuring many of these effects accurately will be challenging, since they depend on conditions that are not under our control. It will be necessary to use output measurements as surrogates for outcomes. In addition, the evaluation will document the institutional issues that may arise during the execution of the program.

3.2 PROPOSED EVALUATION GOALS

Based on the previous discussion, the following top-level evaluation goals have been established:

- System performance
- User acceptance
- Decision effectiveness
- Institutional issues
- Operational improvements

These evaluation goals will be further defined in the next chapter.

3.3 DECISION HIERARCHY

Further analysis of user decisions reveals a two-tiered process. Figure 2 illustrates the differences between the tiers in terms of their decision level. This is an important aspect of the project dynamics and has significant implications for the evaluation.

Tier 1 users will receive the full complement of value-added weather information. The operations staff, (traffic managers, state maintenance forces, state patrol, and emergency services) will use this information to make proactive decisions to improve the safety of the roadways (Tier 1). The actual road condition and status is fed back to *FORETELL*™ or NWS where new predictions are made and disseminated. The actions of the operations staff affect the status of the road, and therefore, the other users. Other users make reactive decisions to weather and roadway information.

Road condition information obtained from the network systems is important to other users (travelers, transit and paratransmit operators, commercial vehicle operators, and school officials). The

decisions these users make may affect other road users, but generally do not affect the conditions of the roadways (thus, the Tier 2 designation). Some Tier 2 users such as school bus and paratransmit operators will also be providing feedback to *Foretell*™ or NWS if their vehicles are equipped with mobile sensors.

The distinction between proactive (Tier 1) users and reactive (Tier 2) users has a significant effect on the evaluation strategy. Although all the users are important, the greatest impact of improved weather information will be realized by operations staff 's use, acceptance, and resulting actions (behavior or operational changes). Consequently, this evaluation will focus on how this information impacts and changes their procedures. In terms of outcomes, the actions of Tier 1 users have the greatest potential of realizing beneficial improvements in safety, efficiency, environmental conservation, mobility, and economic vitality. The decisions by the other users will have a secondary impact on the potential outcomes of the project.

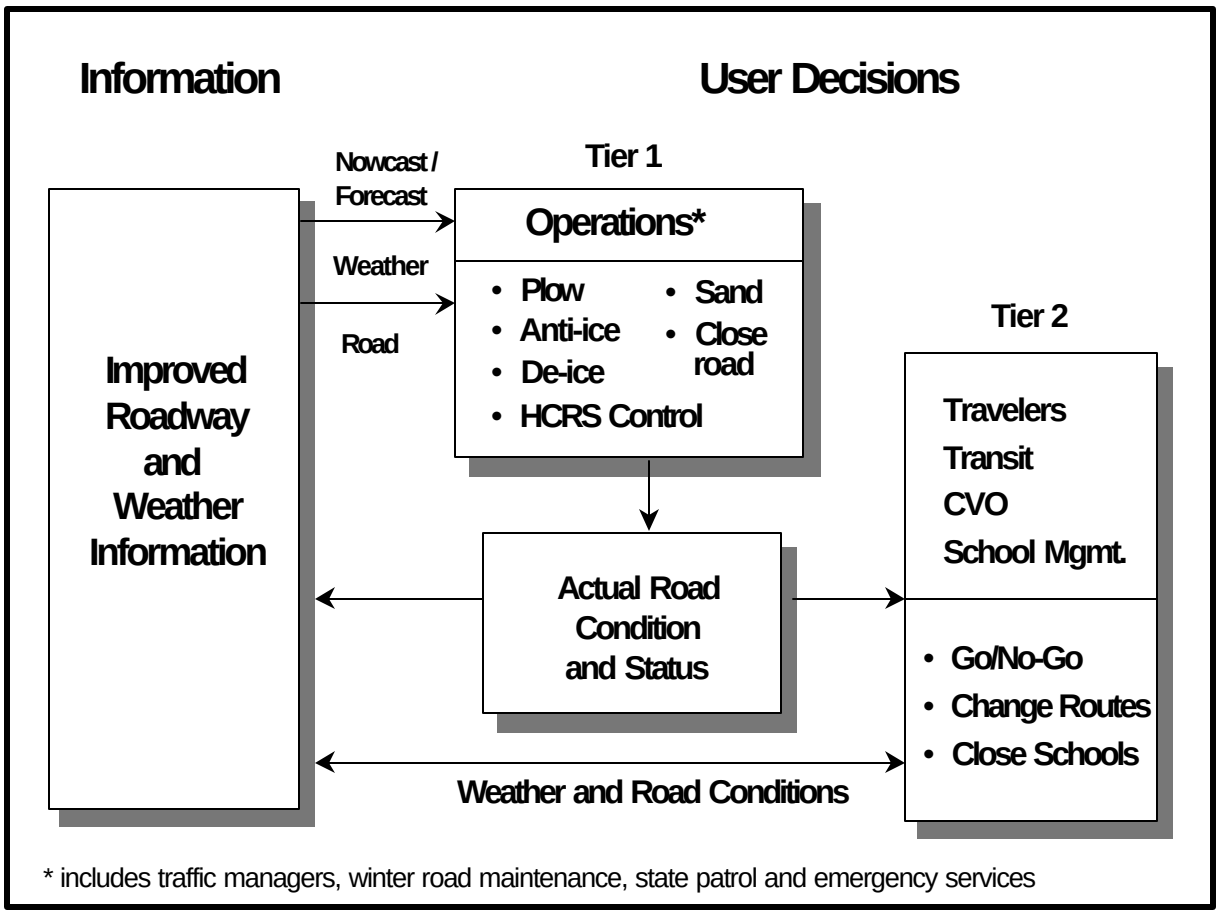


Figure 2. Two Tier Decision Process Identifies Operations Staff as Key

4.0 EVALUATION STRATEGY

As mentioned in Section 3.0, this evaluation must address three fundamental questions:

- Is the *FORETELL*™ information adding value beyond what the users can obtain from existing sources?
- Is the new information changing user's behavior and how?
- What impact will this program have on ARTS goals and objectives (outcomes)?

In this section, these three questions are broken down into specific questions directed toward the five evaluation goals identified in Figure 1. These five evaluation goals have specific questions along with lower-level objectives that were developed from the questions during a meeting of the Battelle Evaluation Team and *FORETELL*™ Project Team. The specific questions will later be used to develop quantifiable hypotheses during the formulation of the evaluation plan. The objectives within each evaluation goal and the proposed measures or measurement method that will be used to evaluate the *FORETELL*™ Program are presented in Section 4.1- 4.7.

Responsibility for collecting data for an evaluation goal will be divided between the Battelle Evaluation Team and the *FORETELL*™ Project Team. The *FORETELL*™ Project Team will be primarily responsible for collecting data associated with the system performance and institutional issues while the Battelle Team will be primarily responsible for assessing user acceptance, decision effectiveness, and operational improvements.

The following sections within this chapter more clearly define each of the evaluation goals. Each section expands on the overall evaluation system description in Figure 1 and identifies the evaluation objectives, the measurements and data sources, and responsibilities of the Evaluation and *FORETELL*™ Project Teams. At this time, the level of detail is commensurate with a "strategy" and will be expanded during the development of the Evaluation Plan.

4.1 SYSTEM PERFORMANCE

Weather information from the NWS will be collected, modeled, and disseminated from the ITS Service Center (SC). System performance of the SC involves the data accuracy, and operational

availability and effectiveness of the center. The primary questions associated with the evaluation of the SC focus on four areas:

- System Performance:
  How accurate are the forecasts prepared by the SC?
  How precise and timely is the road condition (HCRS) information provided by the SC?
- System Reliability:
  Is the system producing reliable output (i.e., given the same inputs, producing the same outputs)?
- Operational Availability:
  Is the SC up and running when it is supposed to be?
- Operational Effectiveness:
  Is the SC delivering the products it is supposed to?

A summary of the evaluation strategy for the system performance goal is illustrated in Figure 3.

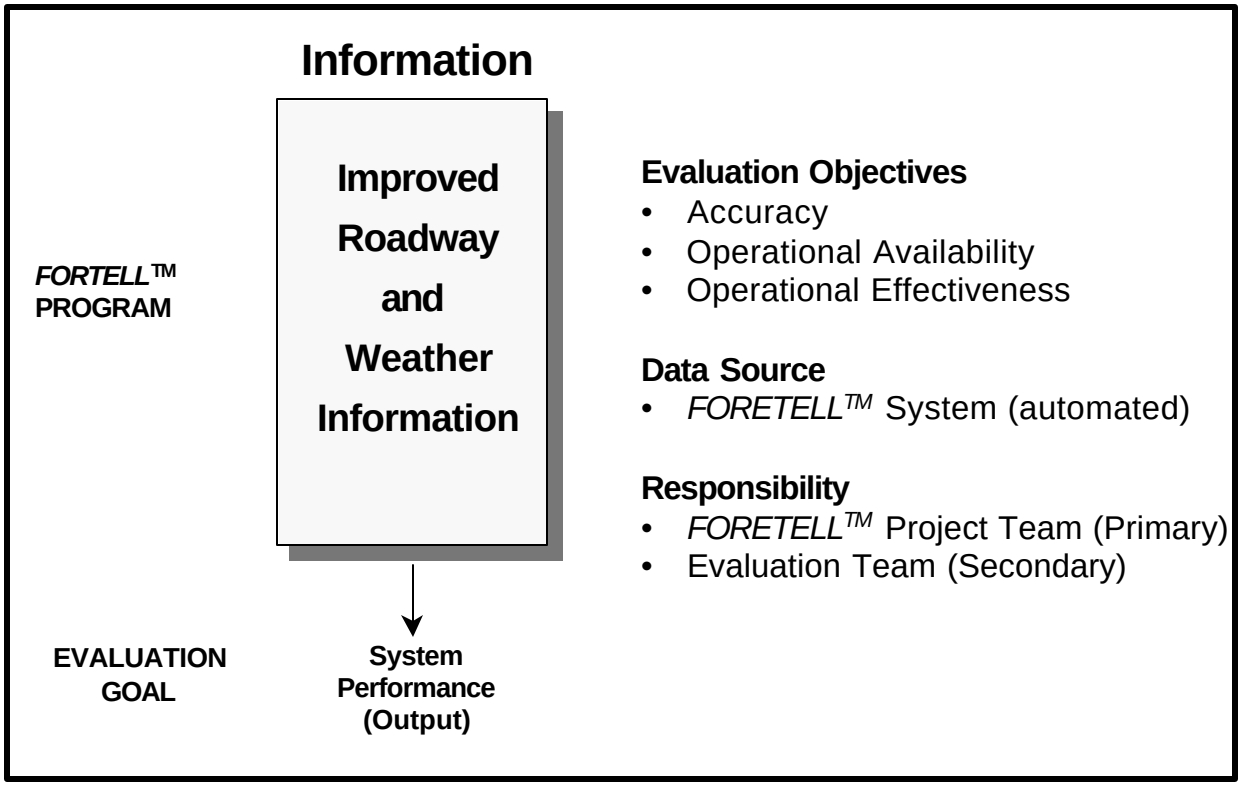


Figure 3. An Example of System Performance of SC Goals and Objectives

A detailed list of objectives and measures to evaluate the accuracy, operational availability, and operational effectiveness with respect to the system performance of the SC is provided in Table 2.

Table 2. System Performance, Objectives, and Measures

Evaluation Goal	Evaluation Objectives	Measures[1]
System Performance	**Accuracy**	
	• Atmospheric data forecast vs. actual[2]	• Precipitation type, rate, amount, start time and finish time • Minimum, maximum air temp • Minimum, maximum dew point temp; • Minimum visibility • Wind speed and direction (vector error)
	• Road condition data forecast vs. actual[3]	• Minimum, maximum pavement surface temperature (0-6 hr and 24 hr forecasts) • Pavement surface condition (dry, wet, freezing, frozen)[3] • Snow or ice amount, type, start time and finish time
	• Mobile platform observation vs. final site observation[4]	• Air temperature • Dew point temperature • Pavement surface temperature • Surface condition
	Operational Availability	
	• Atmospheric data updated on time?	Time data are available vs. time needed (or percentage of time data are not available);
	• Road conditions updated on time?	Time data are available vs. time needed (or percentage of time data are not available);
	• HCRS status updated on time?	Time data are available vs. time needed (or percentage of time data are not available);
	• Weather forecast preparation occurring as planned?	Percentage of time forecasts are issued when required;
	• Road condition forecast preparation (SC models) occurring as planned?	Percentage of time forecasts are issued when required.
	Operational Effectiveness	
	• Information dissemination subsystems updated on time?	Time products are delivered for dissemination vs. time required.
	• Service Center products delivered on time?	• Percentage of time weather forecasts are delivered on time? • Percentage of time road condition forecasts are delivered on time? • Percentage of time HCRS information is available on time?

[1] Information will be gathered and supplied by *FORETELL*™.
[2] 0-6 hr and 24 hr forecasts to be compared with measurements at selected, agreed-to RWIS sites.
[3] This forecast may not be measurable because it requires a prediction of maintenance actions.
[4] Mobile observations to be compared with fixed observations at selected, agreed-to RWIS sites.

The level of detail presented in Table 2 has been prepared by the Evacuation Team in order to assist the *FORETELL*™ Project Team develop software for its data collection and archiving schemes. *FORETELL*™ will assume responsibility to generate and analyze this data.

4.2 USER ACCEPTANCE

The primary question that needs to be answered with respect users acceptance of the SC weather information is: Are the *FORETELL*™ products disseminated in a manner (form and function) that assists users in making timely and efficient decisions? To address this primary question, four fundamental issues have to be evaluated:

- Did the user receive the information?
- Did they understand it?
- Was it useful?
- Did it change their behavior?

These specific questions translate into four objectives: receipt of information, use of information, perceived value, and behavior change summarized in Figure 4.

As shown in Table 3, within each of these objective areas are additional questions that must be addressed to evaluate the objectives. Survey methods and interviews will be used to assess the user acceptance objectives.

Table 3. User Acceptance, Objectives, and Measurement Methods

Evaluation Goal	Evaluation Objectives	Measurement Methods
User Acceptance	**Receipt of Information** • What media was it received through? • Was it received on time? • Was it what user wanted? • Was it in the form the user wanted?	Survey and/or Interview Users[1]
	Use of Information • Did user understand information? • Did user know how to use the info.? • Did user take action based on info.? • Was info. there when you needed it?	Survey and/or Interview Users
	Perceived Value • Did user like the information? • Did user think it was correct? • Was information worth it? • Reaction to sponsorship?	Survey and/or Interview Users

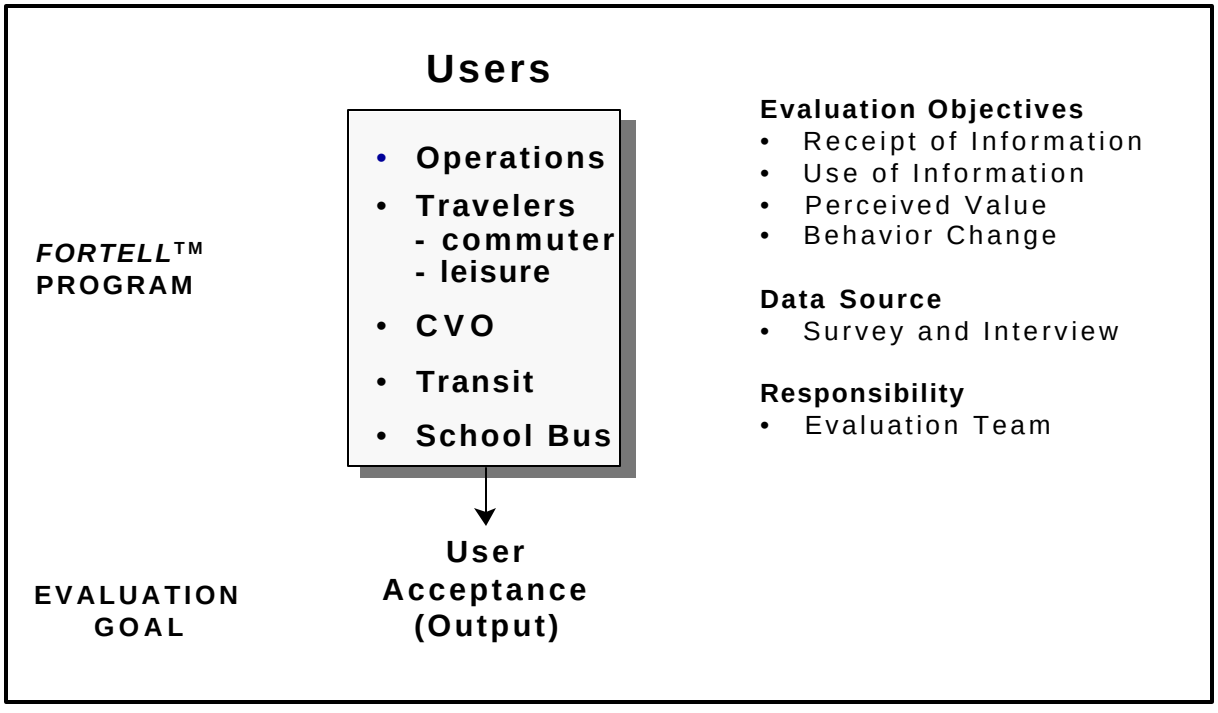


Figure 4. User Acceptance Evaluation Goal and Objective

	Behavior Change • Did info. affect users' decision? • Did info. increase user comfort level of decision?	Survey and/or Interview Users

[1] Survey will focus on highway maintenance and operations personnel.

The Battelle Team will have primary responsibility to develop and conduct these surveys, and analyze the results.

4.3 DECISION EFFECTIVENESS

The primary question to be addressed with respect to the effectiveness of decisions made with SC weather information is: Are the *FORETELL*™ products providing the appropriate information for making timely and efficient decisions compared to similar decisions made before the SC weather information was available? This question needs to be addressed by operations personnel, travelers, CVO's, transit/para-transit operators, and school management. Our evaluation objectives are summarized in Figure 5.

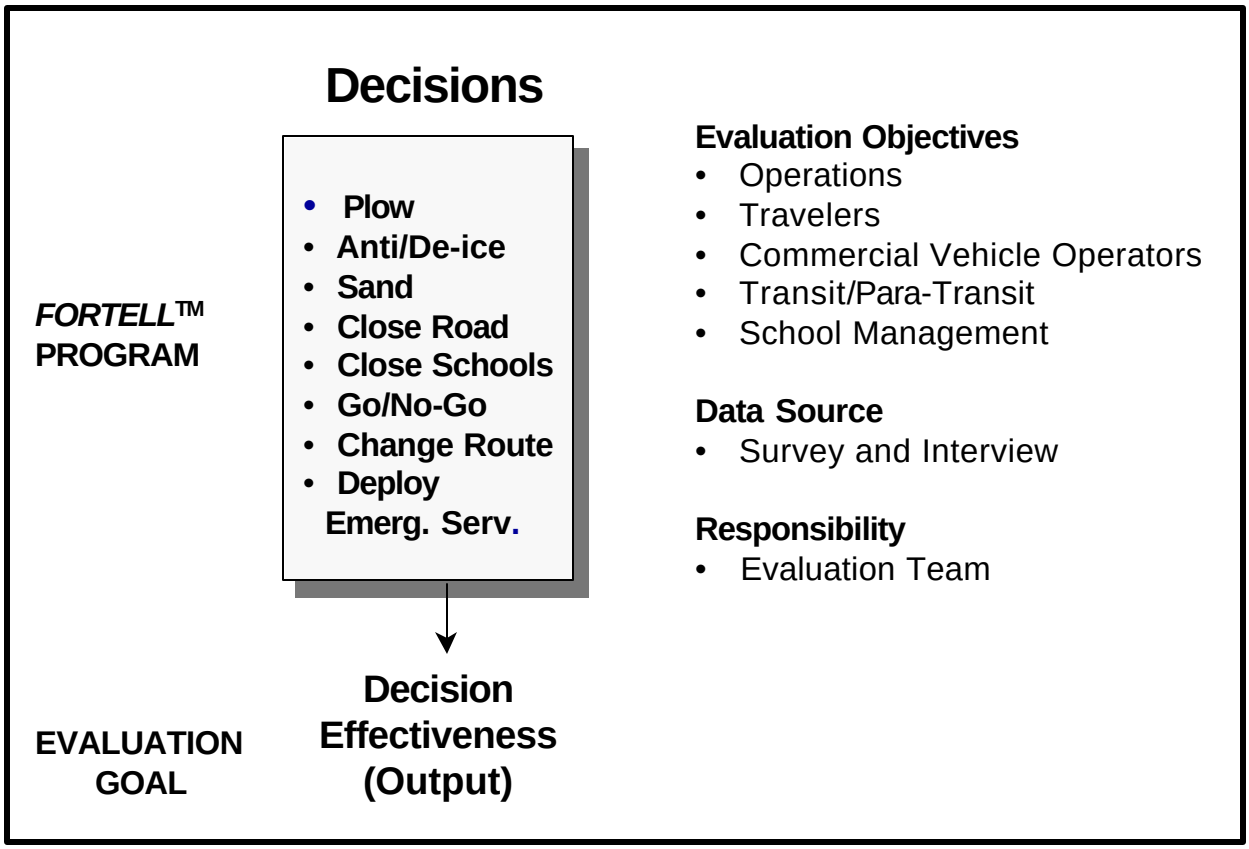


Figure 5. Decision Effectiveness Goals and Objectives

The decision effectiveness of weather information for most users will be assessed using simple 2x2 contingency tables or decision matrices. For example, a maintenance supervisor has to make a decision to call in a crew or not because of inclement weather forecast. The supervisors call/no-call decision will be based on weather information from the SC. Under this scenario four results are possible:

SF/SO – Snow forecasted and snow observed (crew called in and deployed)

SF/NSO – Snow forecasted and no snow observed (crew called in and deployed unnecessarily)

NSF/SO – No snow forecast and snow observed (crew called in after snow has fallen)

NSF/NSO – No snow forecasted and no snow observed (crew not called in and crew not needed)

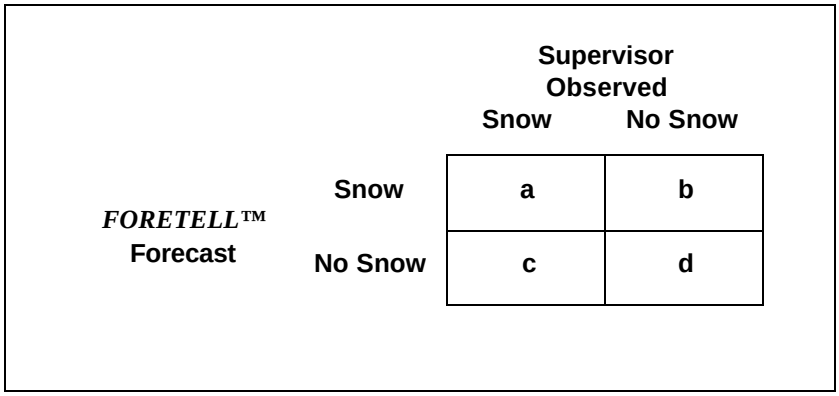


Figure 6. A Decision Matrix Example

The number of occurrences in cell b are associated with the false alarm rate and unnecessary expenditure of resources. The occurrence in cell c can result in untreated conditions, which may require more resources to correct, or increased accident rates. The occurrences in cells a and d are correct forecasts. A similar matrix will be developed for locations in close proximity where the SC weather information will not be available. The with and without SC weather information matrices will then be compared.

A more detailed list of important components that should be evaluated for each objective is provided in Table 4.

Table 4. Decision Effectiveness, Objectives, and Measurement Method

Evaluation Goal	Evaluation Objectives	Measurement Method
Decision Effectiveness	**Operations** • Staff efficiency • Route/Location of work • Pavement treatment type • Road Closure • Incident response • HCRS • Traffic/Advisories/Control • Dispatch • Road condition/Other information	Survey and/or Interview Users
	Travelers - Commuters • Go/No-go • Route selection • Vehicle/Equipment • Timing • Mode	Potentially survey commuters, but limited capability to do so
	Travelers - Leisure • Go/No-go • Route selection • Vehicle/Equipment • Timing • Mode • Itinerary • Weather at destination	Survey Interview, or Intercept Users
	Commercial Vehicle Operators • Go/No-go • Route selection • Load configuration • Schedule • Itinerary • Fuel type	Survey and/or Interview Users
	Transit/Paratransit • Crew/Vehicle preparation • Route selection • Trip cancellations • Go/No-go • Communications with patrons • Scheduling/Timing	Survey and/or Interview Users
	School Management • Start delay • Closures • Bus route selection	Survey and/or Interview Users

An analysis of these types of tables for other users will be utilized to evaluate decision effectiveness along with surveys and interviews. The Battelle Team will take primary responsibility to develop these decision matrices, and surveys to evaluate the effectiveness of decisions.

4.4 OPERATIONAL IMPROVEMENTS

Operational improvements are the primary focus of the evaluation, and are linked to the ARTS Strategic Goals (Figure 7). Directly evaluating safety and security issues, for example, may be difficult. Thus, surrogate measures will be sought to assess any trends. Similarly, it may be very difficult to measure improvements to water quality or air quality associated with improved surface weather

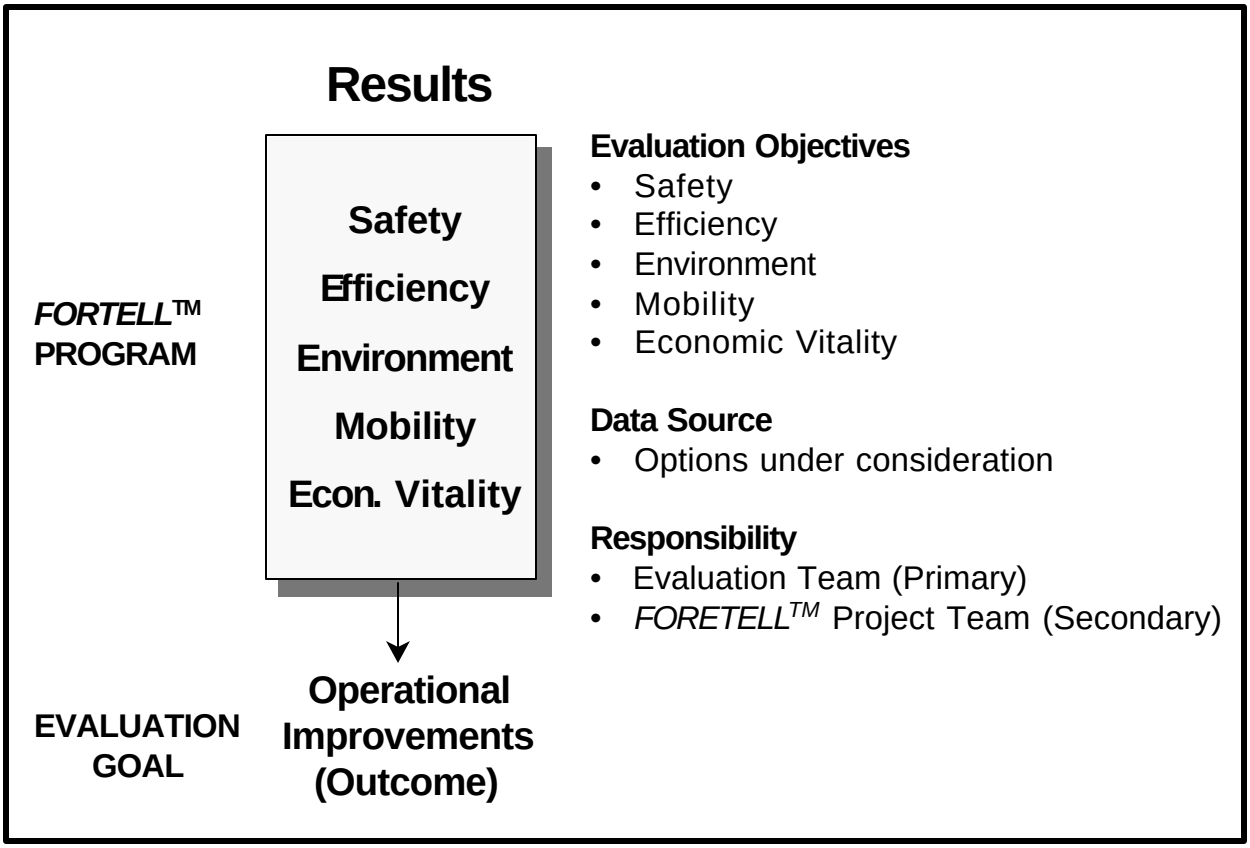


Figure 7. Operational Improvements and Evaluation Objectives

information. On the other hand, some effects of SC information on maintenance operations may be measurable.

Table 5 contains a more detailed list of areas that could be assessed to evaluate operational improvements associated with improved surface weather information.

Table 5. Operational Improvements, Objectives, and Measurement Methods

Evaluation Goal	Evaluation Objectives	Measurement Methods
Operational Improvements	**Safety/Security** • Reduce winter-related road deaths by 15% • Reduce frequency, rate, and severity of crashes • Alleviate exposure to unsafe conditions	Law enforcement – safer driving behavior User surveys

	Efficiency
	• Reduce O&M costs (road maint.):
	  • Labor call out
	  • Fleet deployment
	• Reduce delay/congestion
		Highway agencies documents
		
		Highway agencies verify

	Environmental Conservation • Improve water quality • Improve air quality	 Depts of ecology/environment measure Depts of ecology/environment measure
	Mobility/Convenience • Enhance travelers choices and availability • Increase access for people • Expand information services and communication connectivity • Improve timing of trans. services • Improve trans. service efficiency and maintenance programs	 System Operations Data Surveys and Interviews Traffic Data Collection
	Economic Vitality • Facilitate open and safe travel to and from rural communities • Facilitate open and safe transport to and from rural communities • Support tourists trans. decisions • Support successful business and tourism in rural communities	 Historical and Existing Data Surveys and Interviews

The Battelle Evaluation Team will be primarily responsible for evaluating operational improvements.

4.5 INSTITUTIONAL PERFORMANCE

There are many questions associated with institutional issues that will be considered. For example:

- Is there sufficient interest in and use of the products to indicate *FORETELL*™ is sustainable and has potential for expansion?

- What institutional barriers were apparent in conducting this operational test that would prevent full-scale deployment?
- How well did the *FORETELL*™ Team work with the recipients of the information?
- How well did the *FORETELL*™ Team work with the Canadian participants?
- What organizational/procedural changes might improve overall system performance?

These questions and others listed in Table 6 will be addressed by interviewing NWS personnel, stakeholders, state officials, etc.

Table 6. Institutional Performance, Objectives, and Measurement Methods

Evaluation Goal	Evaluation Objectives	Measurement Methods
Institutional Issues	NWS Local Weather Office Resources	Interview *FORETELL*™ managers, NWS, and local weather officials
	Liability with respect to forecasted condition	Interview Highway Agency representatives
	Contracting Process	Interview *FORETELL*™ managers
	Stakeholder Participation/Commitment	Interview Stakeholders
	Stakeholder Cooperation/Coordination	Interview Stakeholders
	Expandability/Sustainability	Interview *FORETELL*™ manager and users
	New Commercial Ventures	Interview *FORETELL*™ manager
	FORETELL™ Business Model	Interview *FORETELL*™ managers
	Ownership of Information/Products	Interview *FORETELL*™ and DOT managers
	Cost	Interview *FORETELL*™ and DOT managers
	Standards/ITS Architecture Compliance	Interview *FORETELL*™ managers and NTCIP
	HCRS Implementation Demands	Interview *FORETELL*™ managers
	Meterological involvement	Interview Stakeholders
	Compatibility of SC Output	Interview DOT managers

FORETELL™ and Battelle will share responsibilities to acquire information from stakeholders

to assess institutional issues. The Battelle Evaluation Team will be responsible for the evaluation of institutional performance.

4.6 RESPONSIBILITIES

In addition to the Evaluation Team funds, the *FORETELL* ™ Project Team also has funds allocated to support the evaluation activities. The *FORETELL* ™ Project Team evaluation support activities will primarily focus on the efforts to automatically collect data within the information generation systems. An example of this is the forecast to actual comparisons. This data will be collected and stored for subsequent evaluation. The Evaluation Team is responsible for all analysis of the data collected.

4.7 EVALUATION TOOLS

Battelle will be conducting data collection before, during and after implementation of the *FORETELL* ™ Program. Obtaining these multiple measures at different points in time will improve the team's assessment of the effectiveness of the ITS system. The data collection will take the following forms.

- Mail and intercept surveys
- Focus groups and personal interviews
- Traffic data collection activities
- Systems operational data
- Historical and existing data

5.0 EVALUATION TASKS, MANAGEMENT, AND FUTURE PLANNING ACTIVITIES

The Evaluation Team's required tasks are as follows:

- Task 1: Develop Evaluation Strategy
- Task 2: Prepare Evaluation Plan
- Task 3: Prepare Individual Test Plans
- Task 4: Collect Baseline Data
- Task 5: Implement and Perform Individual Tests
- Task 6: Report Findings and Results
- Task 7: Archive Data

The first three tasks (development of evaluation plans) define and guide the evaluation

activities. The product of Task 1 is this Evaluation Strategy. The next two tasks further define the evaluation activities and are briefly discussed below.

5.1 EVALUATION PLAN

The evaluation strategy defined within this document (evaluation Task 1) will be used as a foundation to develop the more detailed evaluation plan. Once FHWA approves the strategy, work will begin to create the project Evaluation Plan (evaluation Task 2).

The evaluation plan will provide the exact criteria, measurements, data required, and data sources for each evaluation goal and objective. The elements of the evaluation plan will include an introduction, background, objectives of the evaluation, a system description, description of the evaluation management structure, evaluation approach, estimated level of effort, work breakdown structure, evaluation schedule, data management plan, and deliverables reports. It will also define specific "tests" that will require the development of an individual test plan.

5.2 INDIVIDUAL TEST PLANS

Following the approval of the evaluation plan, specific individual test plans will be developed for each "test" defined in the Evaluation Plan (evaluation Task 3). A separate individual test plan will be prepared for each full-scale field test. This document will provide the final details to conduct a test of specific hypotheses and collect the necessary data for subsequent evaluation. An example of a required test might be: determining the accuracy of the mobile RWIS sensors. It is anticipated that test plans will have to be specifically developed to assess the equipment's ability to accurately measure the friction coefficient, pavement temperature, freezing point, chemical content and concentration of the surface material, and the depth of the material on the roadway. Other specifics provided in individual test plans include, as appropriate: copies of surveys and interview forms, and the number of phone queries, and web site hits planned to determine if the information disseminated is useful to travelers, maintenance operators, traffic operational centers, and other users.

The elements of the individual test plans will include: test plan objective, approach (strategy, data collection methods, analysis methods, and key supporting conditions), test schedule, pre-test

activities, test activities, post-test activities, data requirements, data analysis plan, report format and expected contents, and estimated resources required to complete all test activities.

5.3 ORGANIZATION AND RESPONSIBILITIES OF THE EVALUATION PROJECT TEAM

Battelle Memorial Institute (BMI), under contract to FHWA, is leading the Evaluation Team. Dr. Bradley Skarpness is Battelle's Project Manager. Battelle has two subcontractors supporting its efforts on this project (Figure 8): Management Solutions (Fred Kitchener, Principal) and Weather Solutions Group (Ed Boselly, President). This team brings extensive experience in ITS operational test implementation and technical evaluations, the applications of weather-related systems, meteorology, and project management. For questions or comments related to this evaluation strategy, please contact Bradley Skarpness at Battelle.

Dr. Skarpness will work closely with the Chair of the FHWA Weather Team and *FORETELL*™ Consortium to ensure that the evaluation plan is integrally related to the project scope and design as specified in the system design concept. He will provide the overall direction to the team and maintain routine communication with the COTR. The day-to-day responsibilities will be assigned to key members of the project team with Mr. Kitchener taking a lead role as the team's Principal Investigator. He will work with the rest of the project team, under Dr. Skarpness' direction, to develop the evaluation strategy plan and assist in the coordination of activities to implement test plans. Mr. Boselly, who has worked on the SHRP and other weather-related projects, will be responsible for assessing the reliability and accuracy of mesoscale predictions. Ms. Burkom and Mr. Williams will be responsible for managing the survey operations and developing the survey questionnaires.

Table 7 provides the allocation of hours by task for key personnel. The indicated effort is limited which requires a focused and well-managed process to achieve the evaluation goals within these constraints.

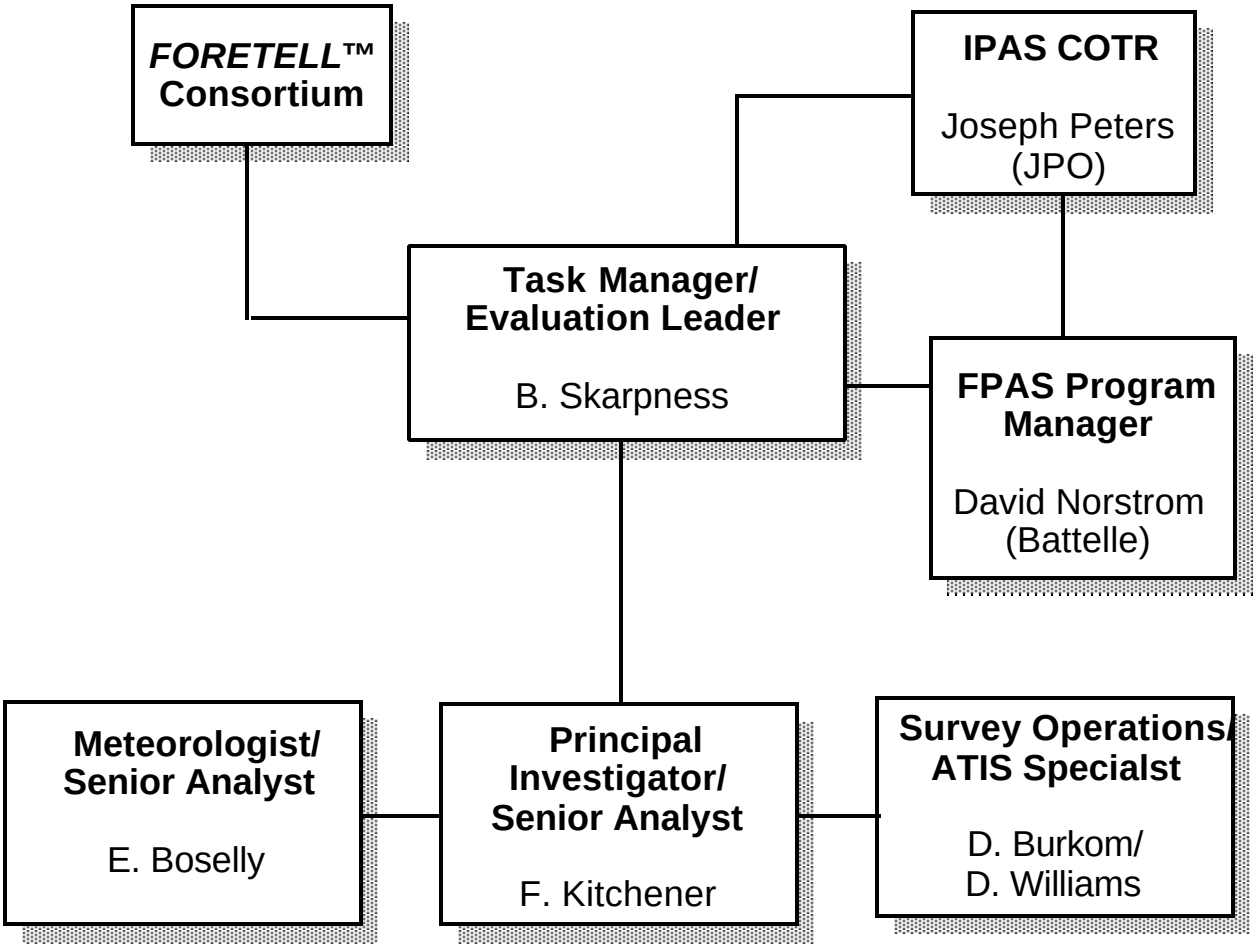


Figure 8. ***FORETELL*™ Consortium Operational Test: Weather Information for Surface Transportation Evaluation Team**

Table 7. Allocation of Hours by Task for Proposed Personnel

Name	Project Role	Task							Total
		1	2	3	4	5	6	7	
Key Personnel									
B. Skarpness	Task Manager and Evaluation Leader	40	20	20	20	10	50		160
F. Kitchener	Principal Investigator and Senior Analyst	125	75	75	110	75	160	80	700
E. Boselly	Senior Analyst and Meteorologist	100	80	60	80		80		400
D. Williams	ATIS Specialist	20	20						40
D. Burkom	Survey Operations Manager	40	20	15		20	15		110
Support Staff									
M. Greene	Data Collection Supervisor				40	60			100
B. Herman	Data Preparation Supervisor				20	20			40
J. Holdcraft	Data Analyst/Statistician		15	15			180	100	310
J. Hayes	Junior Analyst and Meteorologist		14	20	20	20	20		94
Data Collection Crew					280	560			840
Secretarial Support		20	8	40	8	8	40		124
Total									2918

5.4 PROPOSED EVALUATION SCHEDULE

The current evaluation schedule includes four data collection periods (Fall 1998, 1998-99 Winter, 1999 Summer, 1999-2000 Winter), with corresponding evaluations, and a final evaluation report due in the summer of 2000. The Evaluation Team recognizes from their experience on similar projects the possibility that contingencies exist that may delay the evaluation of a given planned component of the project. Some contributing factors are:

- System performance and integration issues that prevent the system from going online as scheduled (Fall 1998)
- Required market penetration necessary to evaluate the behavior changes and impacts of user decisions

- Weather conditions necessary for appropriate data collection (e.g., Mother Nature may not provide winter weather conditions with sufficient frequency or severity to test system effectiveness.)
- Assessing longitudinal issues, such as safety, will take a minimum of 3 years of data after implementation.

The Evaluation Team recommends that no changes in the schedule be made at this time. However, because of possible delays in development and the prospect of additional funding the evaluation schedule should be re-evaluated during the spring each year. At that time the Evaluation Team will recommend an evaluation and testing schedule consistent with the current status of the project.

The proposed schedule is shown in Figure 9 with a possible optional two-year data collection period, which would allow for two additional winter and summer testing seasons. It would also allow *FORETELL*™ the ability to enhance their data accuracy and availability to the full complement of users.

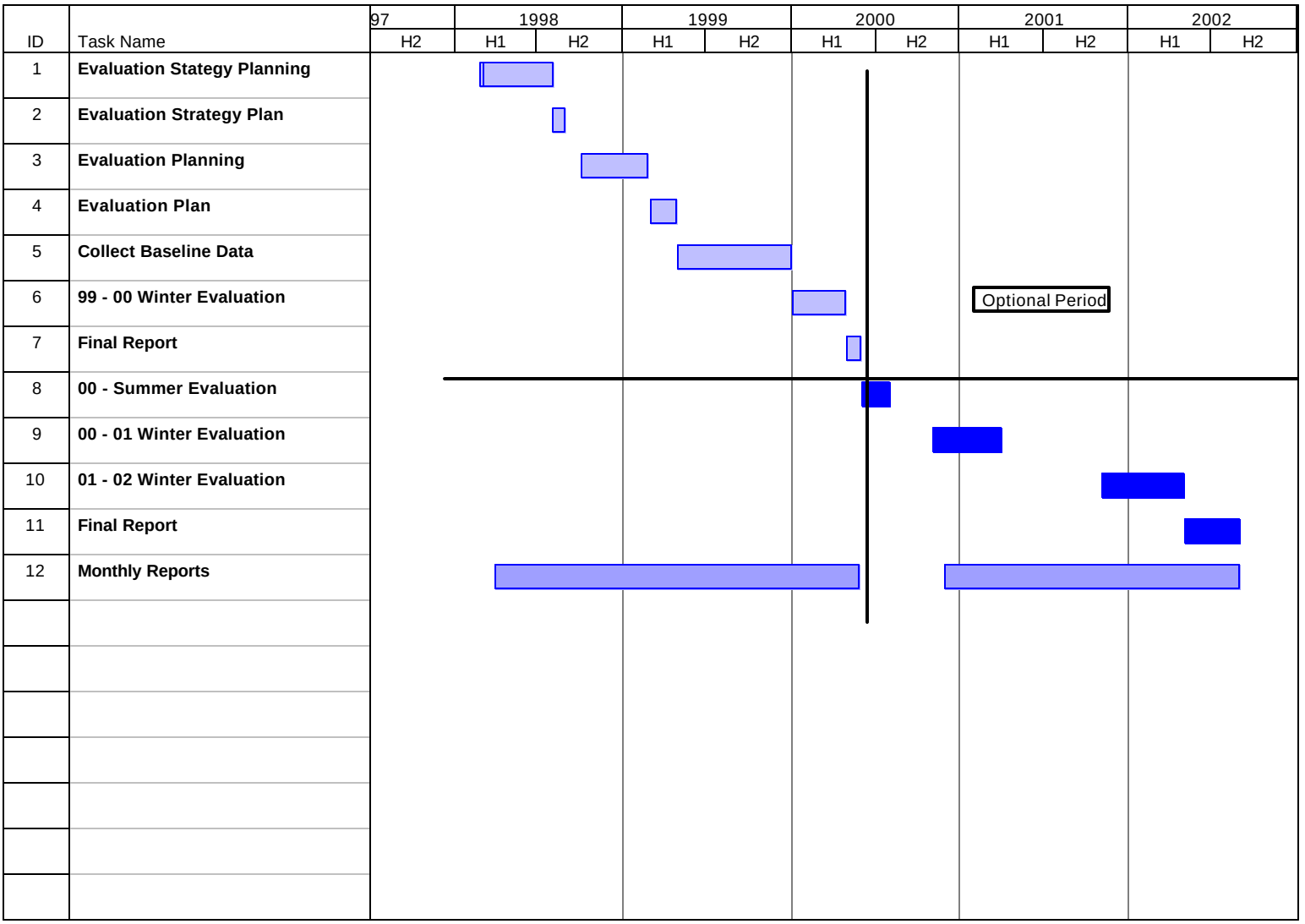


Figure 9. Proposed Evaluation Schedule