NAPPFAST

NCSU APHIS Plant Pest Forecasting System

WEATHER-BASED PEST RISK MAPPING PROJECT

2009

FINAL REPORT

This report fulfills the deliverable of requirement for a final project report to be submitted on or before February 15, 2009.

Edited By Roger Magarey, Dan Borchert, and Glenn Fowler

January 1 to December 31, 2009

Cooperative agreement 09-8100-0862-CA between NCSU and USDA-APHIS-PPQ-CPHST-PERAL/ NCSU

Table of Contents

Executive Summary

This is an annual report for the 2009 NCSU-APHIS Cooperative Agreement **09-8100-0862-CA** 'Weather-based Mapping of Plant Pests,' Turner Sutton, NCSU, Principal Investigator.

Highlights of 2009

- CAPS Top 50 pest risk maps were updated with improved commodity and forest inventory data.
- Development of pathway risk maps.
- Development of Global Analyst tool for improved climate matching.
- Exotic pest targeting tool in NAPPFAST-OBS for risk analysis of highest risk pests and pathways.
- Compilation and cleaning of important phytosanitary and trade data sets.
- Handoff of the NAPPFAST-OBS SCOPE website to Mexico's SINAVEF agency.

Recommendations

- We recommend the continued development of the NAPPFAST-OBS website to support greater risk analysis capabilities and validation of models.

- We recommend that the modeling tools in the old NAPPFAST system be moved to the new NAPPFAST-OBS system. This will allow for superior computational performance and more streamlined system operation.

- We recommend that economic impact maps be added to the suite of CAPS risk maps. This would allow PPQ managers to better target pest detection programs.

- We recommend creation of a graphical modeling building tool. This will allow modelers to create more complex and adaptable models that can predict pest epidemiology beyond that of the existing model templates.

- Add tools to NAPPFAST-OBS to allow PPQ managers to share risk products and model output with the Integrated Plant Health Information System (iPHIS) and with cooperating agencies.

Project Staff and Cooperators

Principal Investigator
Dr. Turner Sutton, Professor of Plant Pathology,
North Carolina State University
Don E Ellis Building 201, Box 7405
Raleigh, North Carolina 27695
Telephone: 919 515-6823, Fax: 919 515-8795

Project Staff at NCSU:
Dr. Roger Magarey, Senior Researcher
Dr. Turner Sutton, NCSU, Principal Investigator
Dr. Jessica Engle, Plant Pathologist, NCSU
Scott Chanelli, Student Assistant
Kathryn Echerd, Student Assistant

Project Staff at APHIS-PPQ-CPHST:
Dr. Daniel Borchert, Risk Analyst
Ms. Peg Margosian, Former GIS Science Fellow
Dr. Glenn Fowler, Risk Analyst,
Mr. Dan Fieselmann, National Science Program Leader for Survey and Detection

Project Staff at ZedX, Inc.:
Dr. Joseph Russo, Agricultural Meteorologist
Dr. Joseph Pietrowicz, Meteorological data manager
Mr. Jay Schlegal, Meteorological specialist
Mr. Brian Holderman, Programmer
Mr. Nathan Gerber, Programmer
Mr. Eric Hudish, Programmer
Ms. Julie Golod, Database Manager (PSU)

Cooperators from Pathways Risk Mapping Group
Dr. Manuel Colunga-Garcia, Michigan State University
Dr. Frank Koch, USFS/NCSU
Dr. Denys Yemshanov, Canadian Forest Service
Dr. Andy Tatem, University of Florida
Ms. Anna Szyniszewska, University of Florida
Ms. Marla Downing, U.S. Forest Service.

1. Introduction

The NCSU/APHIS Plant Pest Forecast System (NAPPFAST) is an internet tool for plant pest modeling using georeferenced climatological weather data. The system is the product of a joint venture between the Animal and Plant Health Inspection Service (APHIS), North Carolina State University (NCSU) and the information technology company ZedX, Inc. The NAPPFAST system was initiated in 2002 based on a 'fill-in-the-blanks' modeling concept developed by Dr. Jack Bailey. The NAPPFAST System is designed to support the predictive pest mapping needs of the Cooperative Agricultural Pest Survey (CAPS) program and the risk assessment activities of the Plant Epidemiology and Risk Assessment Laboratory (PERAL) and is funded by Emergency and Domestic Programs (EDP). NAPPFAST uses "fill in the blank" templates to simplify modeling and mapping of biological events such as the phenological development of insects and the infection periods of plant pathogens. NAPPFAST additionally uses a multi-purpose modeling tool that contains general equations that can be linked to global weather and climate variables.

The following 2009 annual report for the continuing project 'Weather-based Mapping of Plant Pests' (2003-2009), will detail the improvements and upgrades to the NAPPFAST system and related web-tools.

2. NCSU deliverables

2.1 Risk maps for CAPS 2009 and 2010 targets and other targets.

Host Maps

In 2009, the host maps were updated with information from the 2007 agricultural census as reported in NASS. Previously, only the top 28 commodities by sales were used to create the host maps. The updated maps have nearly all commodities surveyed in the agricultural census for a total of 127 possible hosts. The forestry data were also updated from the available data on the USFS Forest Inventory data site (FIDO) site maintained by the forest service. Instead of four classes of wood type (pine, softwoods, soft-hardwoods, hardwoods) the individual genera (*Quercus*, *Populus*, *Ficus*) were grouped to create the 49 forestry layers. The states of Hawaii and Alaska have been added as additional pages of the host map. After consultation with stakeholders, the Puerto Rico territory was also added as an additional page to the maps.

NAPPFAST Maps and Final Risk Maps

The states of Hawaii and Alaska as well as the territory of Puerto Rico will be added to all available NAPPFAST and Final Risk maps. The latter will be updated with new host layers.

Pathway Risk Maps

Pathway Risk maps (importation point maps) were constructed for the CAPS top 50 pests of 2009. Introduction pathways were assessed from a pest's international distribution and from its trade (HS) commodity association. These assessments were made from the Global Pest and Disease (GPDD) and from interception data. Cooperators from the forest service, NCSU (Frank Koch), Michigan State University (Manuel Colunga-Garcia), Kansas State University (Peg Margosian), the Canadian Forest Service (Denys Yemshanov), and PERAL helped develop and program the algorithms to create the maps. The analysis was made using a published methodology (Colunga et al. 2009) [1] which utilizes trade data from the Freight Analysis Framework (FAF) database. Pathways were classified according to 7 international regions of origin and 43 broad commodity classes. Imports entering the United States were distributed to 131 regions including major metropolitan areas, remainder of state and border areas. The maps represent the tonnage of potentially infested commodities (hosts) coming from countries where the pest has been reported.

---

[1] Colunga-Garcia M., Haack R.A., and Adelaja A.O. 2009. Freight Transportation and the Potential for Invasions of Exotic Insects in Urban and Periurban Forests of the United States. Journal of Economic Entomology 102:237-246

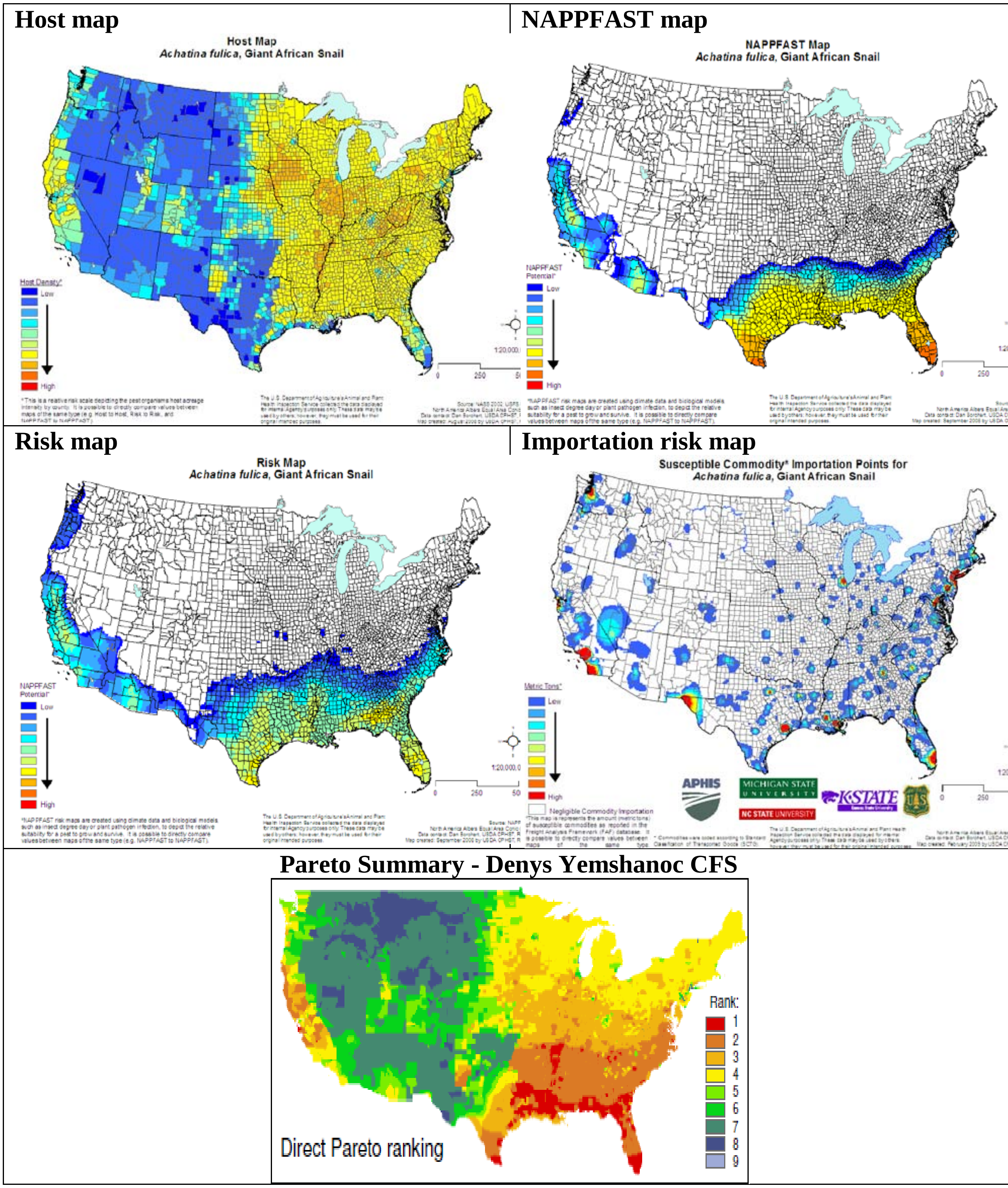

Figure 2.1. *Achatina fulica*, Giant African Snail, host, climatic, risk, importation point, and Pareto summary maps.

Pareto Summary Risk Maps

The host, NAPPFAST, and importation point maps were then combined to create a Pareto Summary risk map. The Pareto map was constructed by Denys Yemshanov CFS with help from Frank Koch (NCSU/CFS) using data provided by NAPPFAST staff. Pareto is a more defensible foundation for incorporating the tradeoffs between multiple risks. Other approaches such as multiplicative or additive weights were tried but did not produce maps that represented the most optimal survey plans. The Pareto approach does not require prior experts' beliefs and the preferences of individual risk components. The final ranks are ordinal and can be used directly to prioritize surveillance and regulatory efforts. They can also be easily rescaled to a desired range (e.g., 0-10). The Pareto maps are currently being constructed but their creation is computationally intensive. They are expected to be completed by mid-2010.

2.2 Training of CAPS cooperators

NAPPFAST staff held a one week workshop for Mexican cooperators on April 6-10, 2009. The workshop included basic training on NAPPFAST, model development exercises for Mexican pests, and a discussion on the Mexican data sharing website. Training power points and cheat sheets in Spanish and English were updated. Dan Borchert conducted a four hour training session on NAPPFAST for pest survey specialists at the Professional Development Center in Maryland.

2.3 Model validations.

NAPPFAST has performed well in two validation studies. The first was a test of the infection models using *Guignardia ctircarpa* (Citrus Black Spot) as a test case. Results showed that *G. citricarpa* was not a pest of concern in citrus districts that had 10 or less days of ascosporic infection during the susceptible period. Observations of pest prevalence were collected for South Africa and Australia. The second study, conducted with 21 insect species was a comparison of the predicted number of generations with observations from the literature. In summary, the primary causes of model failure were day degrees being non-limiting, insufficient biological data for model parameterization and insufficient grid resolution. Overall, 52% of model results were rated as excellent, 24% as moderate, and 24% as poor. These reports will eventually be placed on-line but are available on request.

2.4 International model cooperation.

Roger Magarey and Glenn Fowler participated in the 2009 Pest Risk Mapping Workshop, held in Santa Cruz, CA. The workshop has included researchers from the U.S., Canada, Australia, and New Zealand. NAPPFAST staff made presentations on NAPPFAST-OBS, LBAM, risk mapping of introduction Hot Spots (with Dr. Manuel Colunga-Garcia, MSU), and a proposal for a model shoot out. The workshop focused on a model comparison study but there was controversy over the best methods to conduct such a study.

2.5 Updating user manuals and NAPPFAST.org website.

A NAPPFAST tutorial with Spanish and English power point versions has been placed on the NAPPFAST website. The tutorial includes the following sections: model creation, map creation, climate-matching, map interpretation, and opportunities for collaboration. A spreadsheet with the NAPPFAST infection model has been placed on-line to enable it to be used by researchers.

2.6 Quarterly and annual project reports.

Quarterly reports were delivered on April, July, and October, 2009. This document constitutes the final report.

3 *NAPPFAST*

The following deliverables related to NAPPFAST were completed by the information technology company ZedX. The NAPPFAST-OBS interface provides role-based access and interactive tools for modeling and risk analysis. The NAPPFAST-OBS interface is organized by roles including administrator, modeler, Global Analyst, Program manager, and Survey specialist. Each role contains a number of tools. Individual tools may be shared between roles. In the following two sections, we report on progress to build modeling and risk analysis tools to support PPQ activities.

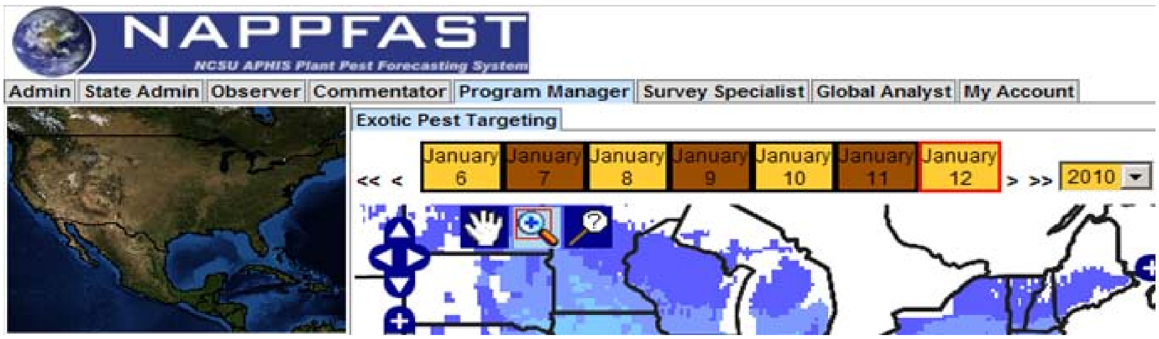

Figure 3.1. Overview of tools and roles in NAPPFAST-OBS

3.1 *Enhance model system infrastructure.*

The administration role is used to organize users and pests. Users including modelers, risk analysts, pest survey specialists, program mangers, and cooperators are organized by role using the user management tool. Users are assigned access to specific pest programs e.g. AHP 2010. The specific pest program contains a list of pests organized by NAPIS code. The code can be added as a tag to all products and data, allowing for correct archival in the system for later viewing and retrieval by the user. Together the user and program management tools allow information in NAPPFAST-OBS to be systematically organized and managed.

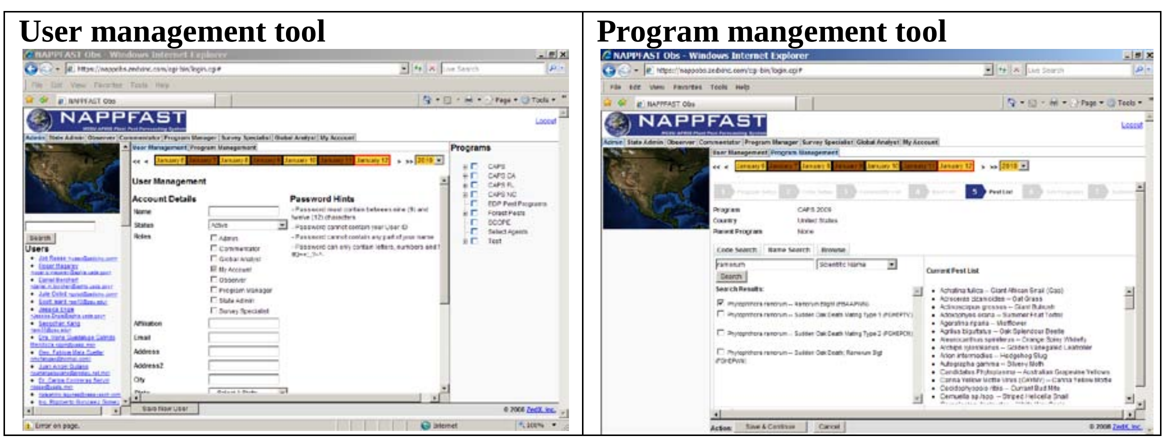

Figure 3.2. Basic infrastructure for NAPPFAST-OBS for program and user management

3.2 *Modeler Role Tools*

In the old version of NAPPFAST, there are three tools, model set-up, history, and climate matching. Design work began in early 2009 to move these tools from the old version of NAPPFAST to the NAPPFAST-OBS modeler role. The task of designing new and improved tools turned out to be more difficult than anticipated because of the new database structure and the modernization of the NAPPFAST-OBS platform. While delayed, the work is expected to be completed in the first quarter of 2010. Consequently, the model location, setup, run, and view tools as part of the modeler role will be included as deliverables for 2010. These tools in their 2009 rendition are described below.

Location – Enables a user as a modeler to create a specific location that can later be selected for later modeling. Users can create a location by clicking on a map or entering a grid location. Tool completed.

Setup – Enables a user as a modeler to parameterize an existing model using an interactive template. Parameters vary by model template but include variable limits e.g. cardinal temperatures, minimum wetness hours, and day degrees by life stage. In addition, users can set-up legend classes for the output variables. This will be an improved version of the existing Edit model tool in the Set-up section of the old version of NAPPFAST.

Run– Enables a user as a modeler to run an existing model by selecting a location, area or continent, model, time period, climatology, input, and output variables and product type graph. A preliminary and simplified version of this tool (Scheduled map request) is available in the History section of the old NAPPFAST system. A new version of this tool in NAPPFAST-OBS will be created in 2010.

View– Enables a user as a modeler to view ad-hoc or automated model products and release them to the administrator for usage in other roles or for sharing with other partner platforms.

3.3 *Climate matching tool*

The role of the Global Analyst (formerly climate analyst) is to construct grid host and pest landscape maps using point data, grid data, and expert opinion as input. The following Global Analyst tools are designed to convert and quality control different third-party data sets into landscape maps and other formats.

- Data Analysis tool converts input data stored in a taxonomic tree into grid landscape maps and conducts either an automatic or manual quality control during the process. The Analyst will be able to dynamically query a data point or grid and edit their stored value(s).
- Drawing tool allows a user to draw and name a bounded area on a user-selected background image.

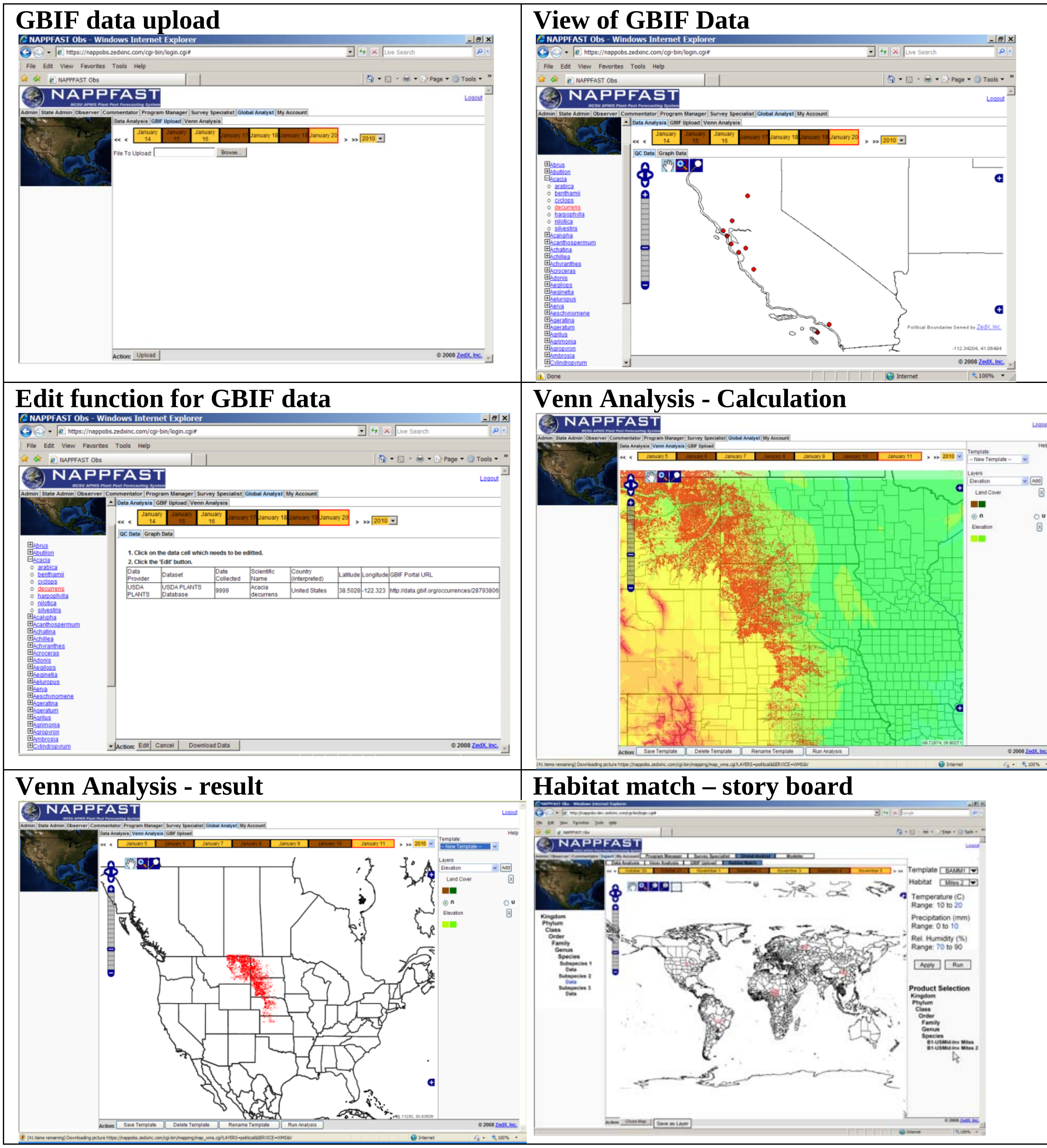

Figure 3.4 Development of the NAPPFAST Global Analyst Tool.

- Venn tool is an algorithm for combining areas by either intersection or union. The Venn tool allows an Analyst to quantify common areas of pest distribution among different geo-referenced data layers, such as political boundaries (nation, state, county), crop, soil, land use, climate, and plant hardiness.
- Habitat Analysis tool is based on the BAMM algorithm, which uses weather data to defined the spatial distribution of a pest. (In progress)
- Publishing tool. Allows analysts to export selected products to program(s).

The Data Analysis, Drawing, and Venn tools have been completed or are in testing phase. The habitat analysis and publishing tools are in progress.

New data sets are in the process of being added to the BAMM climate matching tool. These include 10 km land use/land cover data from the European Space Agency. These data have been upscaled from a 300 m resolution. Other data sets are from the Harmonized World Soil Database was upscaled from a 1 km resolution. A third data set is snow depth from GHCN weather databases. These data are expected to substantially improve climate matching for species that are not climatically driven. In addition, we have been experimenting with a new climate match product which shows users where the predicted distribution is likely to be underestimated based on the climate match for all species in the genera. In addition, research is underway into the feasibility of incorporating a dispersal algorithm into BAMM. The dispersal algorithm could be used to help indicate areas that are both suitable and have most likely been invaded but as yet the pest remains undetected.

3.4 Weather Database and Documentation

Information Technology Company will make available for use comprehensive North American and Global databases as part of the New NAPPFAST platform. The continuing project will cover the following components.

- Sources of weather data, which have been interpolated to 32 km^2 globally, 12 km^2 in North America, 5 km^2 in the conterminous United States (U.S.), and 1 km^2 in small, selected areas within the U.S.
- Online visualization of data quality and record integrity.
- Online documentation of data sources and data processing.

3.4.1 Sources of Weather Data

The weather sources for the NAPPFAST system consist of observed and derived data sets that are stored as either station (point) values or grid values. All data are processed through quality control and quality assurance programs and organized (domains, time frames, historical and real-time) for efficient use in analysis and modeling. Descriptions of data sets currently in NAPPFAST are given in Tables A2.1 – A2.6 (Appendix 2). Included in each table is the period of record, description of data, selected spatial resolution, time step, and variables, and generation procedure if applicable.

The grid data are defined as Phase I (interpolated native values) and Phase II (interpolated native values enhanced with station observations). In Phase I, global grid data are provided by the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Prediction (NCEP) at a native 1 degree spatial resolution. It is then interpolated by the ZedX weather team to a 0.25 degree or 32 km resolution. In Phase II, the native global grid data are combined with station observations that provide more detail in critical geographies. The addition of station observations in the interpolation improves the accuracy of the field representation of a weather variable. The Phase II interpolated data set is also at a 32 km spatial resolution.

NAPPFAST users currently have the option of using either station observations or grid data. If a user chooses station sources, observations are interpolated dynamically (on-the-fly) to a 32 km spatial resolution using either a two-dimensional (2-D) or three-dimensional (3-D) scheme. If a user chooses the grid option, the Phase I and II values have already been interpolated to a 32 km resolution. It should be noted that the Phase II "enhancement" with station observations does not apply to the precipitation variable due to the discrete, discontinuous nature of moisture fields.

In the case of derived weather variables, such as open water evaporation, an algorithm combines multiple "basic" variables, such as temperature, precipitation, wind speed, etc., to derive the desired values. These algorithms will use different weather data sources in their derivation of variables.

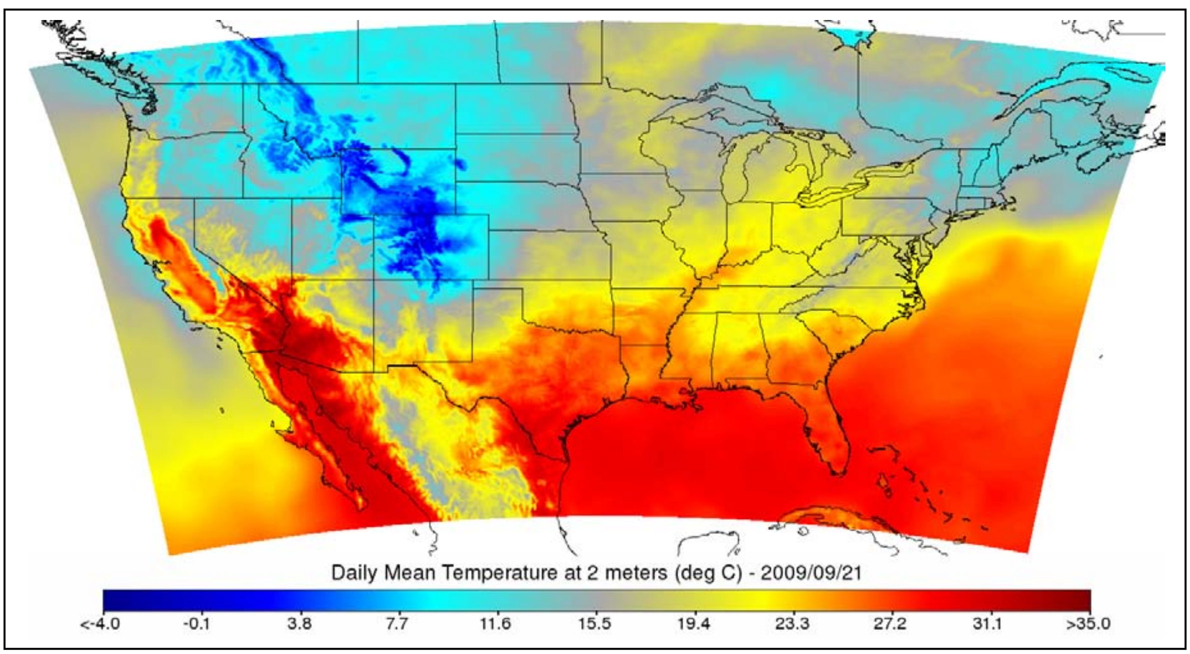

Figure 3.5 Daily mean temperature field on September 9, 2009 for a U.S. centered domain.

During the 2009 development year, considerable work was done by the ZedX weather team to prepare new data sources for NAPPFAST in 2010. An example of a new data source is the Real-Time Mesoscale Analysis (RTMA), which is produced by NOAA-NCEP. A sample of a daily mean temperature field for September 9, 2009 is depicted in Figure 3.5. The RTMA data source consists of grid values at a 5 km resolution for the United States (U.S.) and border portions of Canada and Mexico as shown in Figure 3.5. This new data source along with others will be incorporated in NAPPFAST in the first half of 2010.

3.4.2 Weather Data Quality Control/Quality Assurance

Quality control (QC) and quality assurance (QA) are a normal part of weather data processing. Quality control consists of a number of tests and schemes to ensure that a stored value used in a NAPPFAST application is representative of the "true" weather environment at the reported spatial and temporal resolutions. One internal quality control scheme is called the Percent Absolute Difference or PAD analysis. The PAD analysis scheme identifies biases and uncertainty ranges at fixed levels of risk (e.g. 50, 75, 90, 95, and 99%) in station records. An example of PAD analysis results for a single station and season (spring) is shown in Table 3.1.

A quality control (QC) interface and tests for missing or suspect station observations are depicted in Figure 3.6. The QC tests include plausibility, difference to neighboring (co-located) station(s), missing expected value, climate, linear regression between neighboring (co-located) station(s) and the candidate station, comparisons to grid points, spatial interpolation, and temporal tendency. As shown in figure, a given station on a given day can pass, fail, or not complete a test. The end result of all the tests is a representative "true" value for missing or suspect record.

Table 3.1. Quality control Percent Absolute Difference (PAD) test for station data.

A1 Station

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Station	Season	Hour	Test	SampleSize	Average Obs Value	Average Replacement Value	Avg. Difference	Avg. Absolute Diff.	RMSD	Pct50	Pct75	Pct90	Pct95	Pct9
2	EGLL	Spring	0	Climatology	91	8.58	8.49	-0.09	1.62	1.9209	2	3	3	3	
3	EGLL	Spring	0	Linear Regression	91	8.58	8.48	-0.11	0.74	0.8944	1	1	1	2	
4	EGLL	Spring	0	Grid Point LR	91	8.58	8.21	-0.38	0.84	1.077	1	1	2	2	
5	EGLL	Spring	0	Spatial Interp.	87	8.69	7.26	-1.43	1.43	1.6	1	2	2	3	
6	EGLL	Spring	0	Temporal LR	91	8.58	8.47	-0.12	0.58	0.7348	0	1	1	2	
7	EGLL	Spring	1	Climatology	91	8.23	8.2	-0.03	1.72	2.0518	2	3	3	3	
8	EGLL	Spring	1	Linear Regression	91	8.23	8.1	-0.13	0.87	1.1269	1	1	2	2	
9	EGLL	Spring	1	Grid Point LR	91	8.23	7.81	-0.42	0.8	1.0724	1	1	2	2	
10	EGLL	Spring	1	Spatial Interp.	84	8.25	6.83	-1.42	1.43	1.6432	1	2	3	3	
11	EGLL	Spring	1	Temporal LR	91	8.23	8.2	-0.04	0.48	0.5745	0	1	1	1	
12	EGLL	Spring	2	Climatology	90	7.96	8.21	0.26	1.84	2.1795	2	3	3	4	
13	EGLL	Spring	2	Linear Regression	90	7.96	7.95	-0.01	0.74	0.9644	1	1	2	2	
14	EGLL	Spring	2	Grid Point LR	90	7.96	7.61	-0.34	0.82	1.0344	1	1	2	2	
15	EGLL	Spring	2	Spatial Interp.	82	8.15	6.71	-1.44	1.48	1.6432	1	2	2	3	
16	EGLL	Spring	2	Temporal LR	90	7.96	8.03	0.07	0.53	0.6708	0	1	1	1	
17	EGLL	Spring	3	Climatology	91	7.77	8.17	0.4	1.79	2.1494	2	3	3	3	
18	EGLL	Spring	3	Linear Regression	91	7.77	7.56	-0.21	0.75	0.9274	1	1	1	2	
19	EGLL	Spring	3	Grid Point LR	90	7.71	7.24	-0.48	0.94	1.1662	1	1	2	2	
20	EGLL	Spring	3	Spatial Interp.	90	7.78	6.4	-1.38	1.42	1.5684	1	2	2	2	
21	EGLL	Spring	3	Temporal LR	91	7.77	7.68	-0.09	0.49	0.6164	0	1	1	1	
22	EGLL	Spring	4	Climatology	91	7.52	7.95	0.43	1.89	2.2494	2	3	3	4	
23	EGLL	Spring	4	Linear Regression	91	7.52	7.26	-0.26	0.73	0.8944	1	1	1	2	
24	EGLL	Spring	4	Grid Point LR	91	7.52	7.27	-0.25	0.73	0.9539	1	1	2	2	
25	EGLL	Spring	4	Spatial Interp.	82	7.29	5.96	-1.34	1.36	1.5556	1	2	2	3	
26	EGLL	Spring	4	Temporal LR	91	7.52	7.52	0	0.51	0.6245	0	1	1	1	
27	EGLL	Spring	5	Climatology	91	7.36	7.96	0.59	1.87	2.2405	2	3	3	4	
28	EGLL	Spring	5	Linear Regression	91	7.36	7.25	-0.11	0.88	1.0863	1	1	2	2	
29	EGLL	Spring	5	Grid Point LR	91	7.36	7.25	-0.11	0.85	1.077	1	1	2	2	
30	EGLL	Spring	5	Spatial Interp.	81	7.47	6.15	-1.32	1.39	1.5716	1	2	2	3	
31	EGLL	Spring	5	Temporal LR	91	7.36	7.34	-0.03	0.52	0.6782	0	1	1	1	
32	EGLL	Spring	6	Climatology	91	7.76	8.07	0.31	1.81	2.1517	2	3	3	3	
33	EGLL	Spring	6	Linear Regression	91	7.76	7.68	-0.08	0.85	1.0536	1	1	2	2	
34	EGLL	Spring	6	Grid Point LR	90	7.69	7.62	-0.07	0.73	0.9381	1	1	2	2	
35	EGLL	Spring	6	Spatial Interp.	89	7.82	6.65	-1.17	1.21	1.3528	1	2	2	2	
36	EGLL	Spring	6	Temporal LR	91	7.76	7.58	-0.18	0.57	0.7746	1	1	1	2	
37	EGLL	Spring	7	Climatology	91	8.52	8.44	-0.07	1.77	2.0347	2	3	3	3	
38	EGLL	Spring	7	Linear Regression	91	8.52	8.44	-0.08	0.58	0.7616	1	1	1	1	
39	EGLL	Spring	7	Grid Point LR	91	8.52	8.38	-0.13	0.69	0.8602	1	1	1	2	
40	EGLL	Spring	7	Spatial Interp.	91	8.52	7.43	-1.09	1.09	1.2329	1	1	2	2	
41	EGLL	Spring	7	Temporal LR	91	8.52	8.22	-0.3	0.66	0.8718	1	1	1	2	
42	EGLL	Spring	8	Climatology	91	9.57	9.48	-0.09	1.93	2.2136	2	3	3	3	
43	EGLL	Spring	8	Linear Regression	91	9.57	9.54	-0.03	0.66	0.8602	1	1	1	2	
44	EGLL	Spring	8	Grid Point LR	91	9.57	9.4	-0.17	0.8	0.9849	1	1	2	2	
45	EGLL	Spring	8	Spatial Interp.	91	9.57	8.7	-0.87	0.93	1.0536	1	1	2	2	
46	EGLL	Spring	8	Temporal LR	91	9.57	9.36	-0.21	0.69	0.8307	1	1	1	2	
47	EGLL	Spring	9	Climatology	91	10.77	10.23	-0.54	2.06	2.5	2	3	4	4	
48	EGLL	Spring	9	Linear Regression	91	10.77	10.8	0.03	0.7	0.8832	1	1	1	2	

KFFT 2008-01-12 21:00:00

Dew Point

Raw Value: -1.1

	Test Results	Statistic	Value	Replacement Value
Plausibility	Pass	Percent		
Co-located Stations	Cannot Complete	Difference		
Missing Expected Value	Cannot Complete			
Daily Climate	Pass	Stdev	0.288	-3.1
Spatial-Temporal Co-Point Linear Regression	Fail	stderr_res Z-score	0.18 -3.764	-5.6
Grid Point	Fail	stderr_res Z-score	0.042 -3.097	-2.0
Spatial Interpolation	Fail	Difference	4.72	-5.8
Temporal LR Test	Pass	stderr_res Z-score	0.087 1.677	-0.1
Final Result	Fail			
Value Used				

Figure 3.6. Quality control interface and tests for station data.

Another quality control interface for global stations (METAR and Synoptic networks) is depicted in Figure 3.7. In the figure, METAR stations with suspect data are displayed against a geographic background. A user can click on a station and select a day from the calendar and then view QC results. If there is an anomaly in the QC results, the user can manually override the automatic process by substituting a value. This manual replacement of a value is termed "expert inspection" and is a valid QC procedure.

Using the same QC interface as in Figure 3.7, a user can choose a station and year and compare the daily progression of values against climatological ranges (Figure 3.8). These ranges are at fixed risk levels similar to the PAD analysis. The example variables in Figure 3.8 are the daily maximum (top graph) and minimum (bottom graph) temperatures for the second half of 1960.

Quality assurance is concerned with the reliability and consistency of data delivery. The ZedX weather team has programmed a number of check points that alert the Systems team if data were not delivered or if files were corrupted. The same team also keeps tabs on changes in formats or delivery times by third-party providers, such as the federal government.

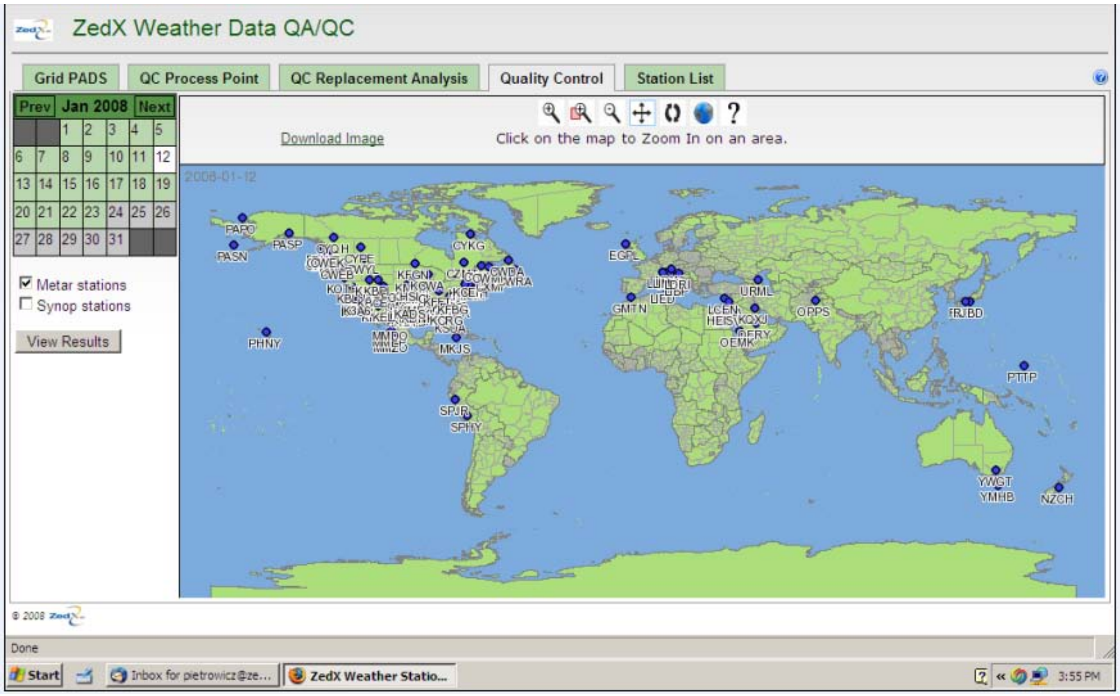

Figure 3.7. Quality control interface with display of METAR stations.

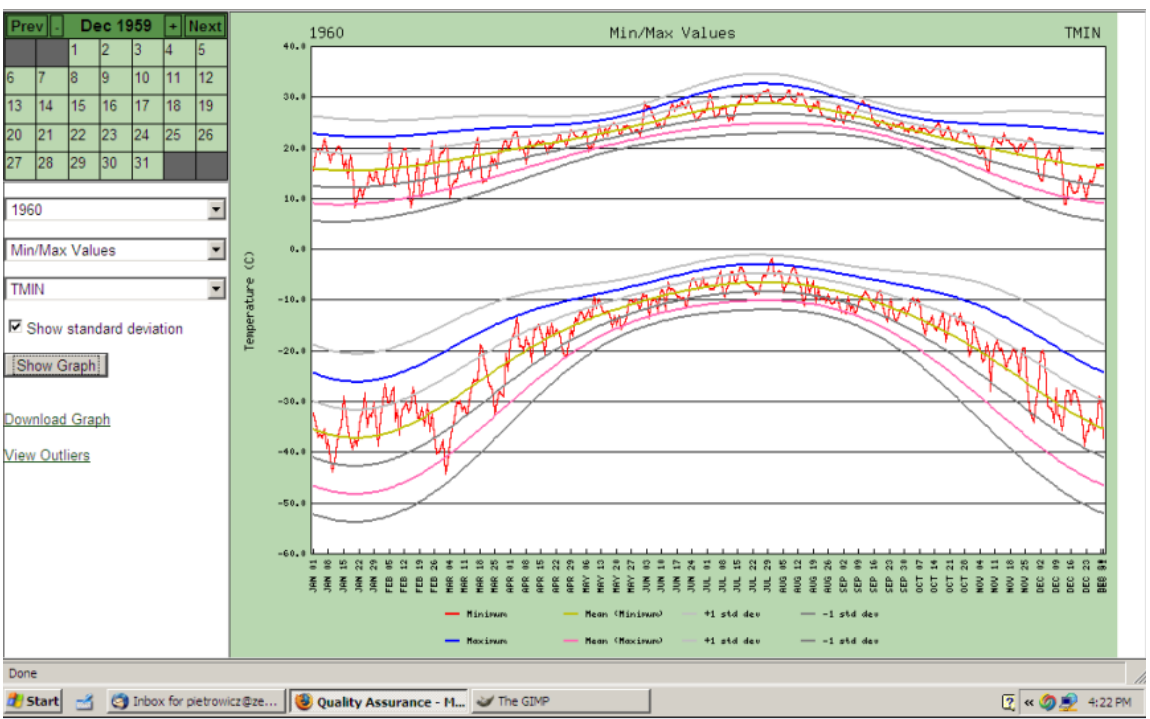

Figure 3.8. Quality control interface with standard deviations for station data.

3.5 *NAPPFAST data flow across the Government Zone Firewall*

The “government zone” firewall was completed during 2009 to meet the Federal Information Processing Standard (FIPS) security requirements for data storage and usage. Considerable time was spent to develop the proper network configuration that allowed public, third-party weather data to pass through the government zone firewall for NAPPFAST applications. A simple flow chart of this configuration and data flow is given in Figure 3.9. The government zone firewall requires continual maintenance as computer equipment and network communications are updated on a regular basis. Also, the firewall protocols must be revisited each time a new weather data source is added to the system.

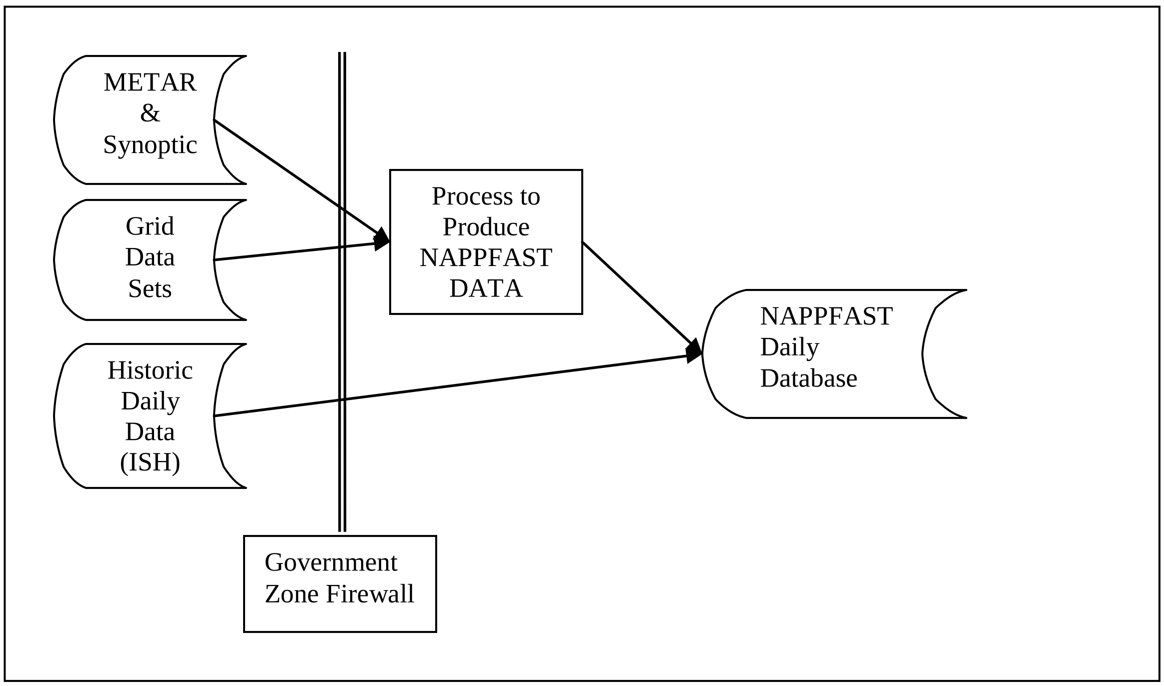

Figure 3.9. Flow chart of government zone firewall configuration and weather data flow.

4 NAPPFAST Exotic Pest Targeting Tool

The exotic pest targeting system will allow APHIS personnel and cooperators to conduct high resolution risk analyses for the highest risk pests, pathways, and points of entry or distribution. The exotic pest targeting system was built by expanding upon the existing NAPPFAST-OBS platform. The NAPPFAST-OBS platform provides role-based access and information technology (IT) tools for risk analysis, viewing, and manipulating pest observations or risk data, GIS data layers, and model output. The NAPPFAST exotic pest targeting system includes tool development for the Program Planner, Survey Specialist, and Risk Analyst roles. In addition, there is the development of cross-role tools to clean critical data sets, upload data, and link third party data sets.

4.1 Program Planner and Survey Specialist Role

The Program planner and Survey specialist role includes access to the Exotic Pest Targeting Tool (EPTT). The tool allows users to see products and data associated with given pests in pest programs either national or state that have been assigned to a user. The following components in the build of the EPTT have been completed

- Organization of pests by under pest program e.g. AHP 2009,
- Interactive maps and ability to view pest observation data across years,
- Static host, climate, and final risk maps available as overlays for selected pests,
- Dynamic risk maps (e.g. weekly phenologies) provided as backgrounds. Weekly phenology maps have been generated for some of the lepidopteran EDP program pests, and
- Aerial imagery at 1-2 m resolution from the National Agricultural Imagery program.

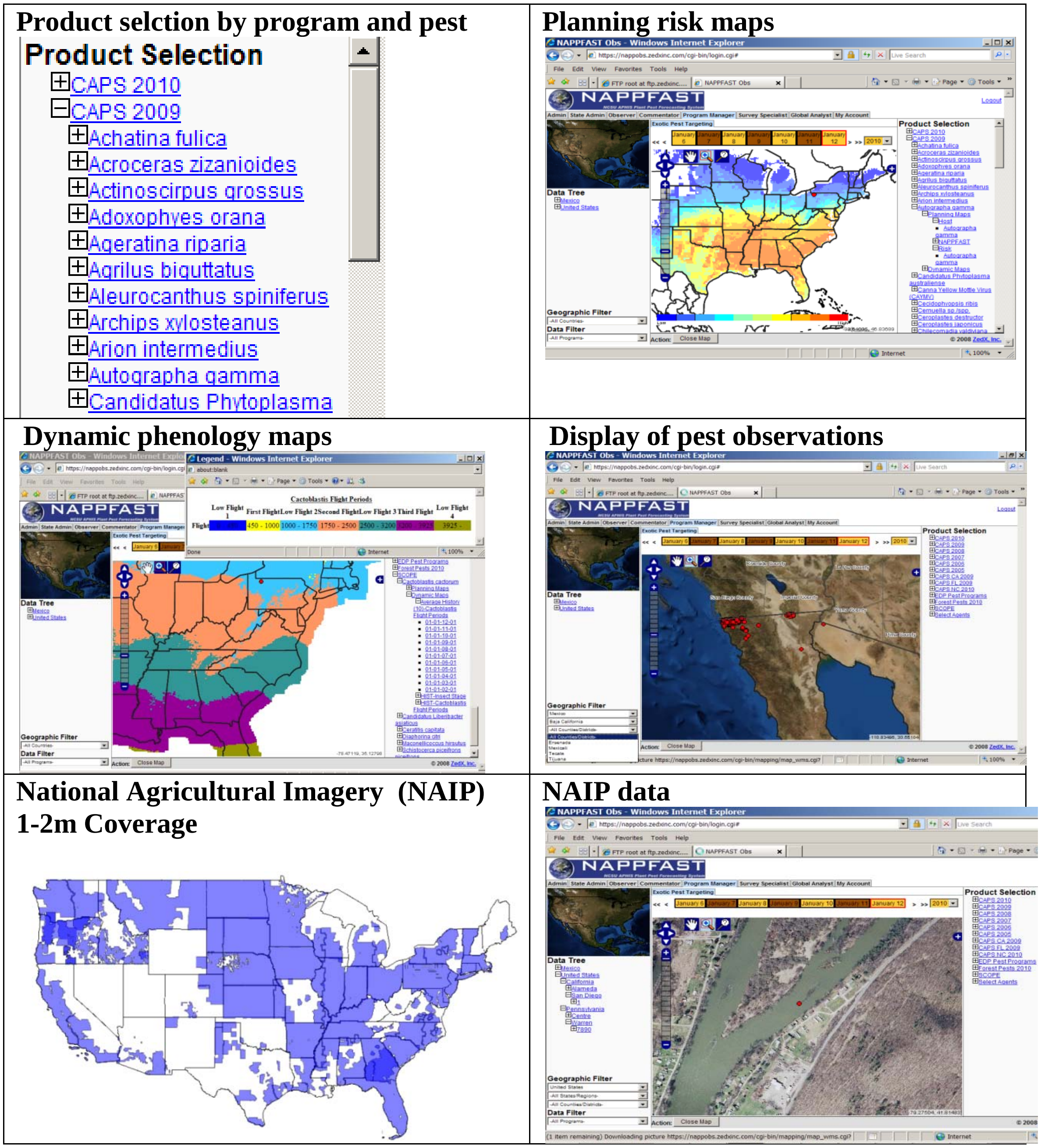

Figure 4.1. Example on-line views from the Exotic Pest Targeting tool.

The following components of the Exotic Pest Targeting tool await completion.

- Display of phytosanitary data. The phytosanitary data will allow users to highlight zip codes or port locations with an elevated phytosanitary risk. Map filters will limit the display of risk data by data set (EAN, PIN 309, 264, 280), country of origin, pest family, pest name, host or commodity name, and/or Harmonized System (HS) code. Details on the cleaned data sets are included below.
- Display of commercial sales data to guide survey specialists to business locations or city blocks with sales of specific commodities e.g. tiles, lumber. The commercial sales data will be uploaded with the pathways upload form. The sales data will be viewed by filtering the data by geography (state or zip), sales volume and NAICS code description, or business type.

- Display of state and county risk maps. These risk maps include an overall risk map, introduction potential, economic, suitability, and natural resource impact. The state risk maps are completed and the county risk maps are under construction by CPHST. Once completed the risk maps will be available in the product menu of NAPPFAST-OBS. Users will be able to select a particular risk map and zoom in on their state or county as well as overlaying the phytosanitary risk data.

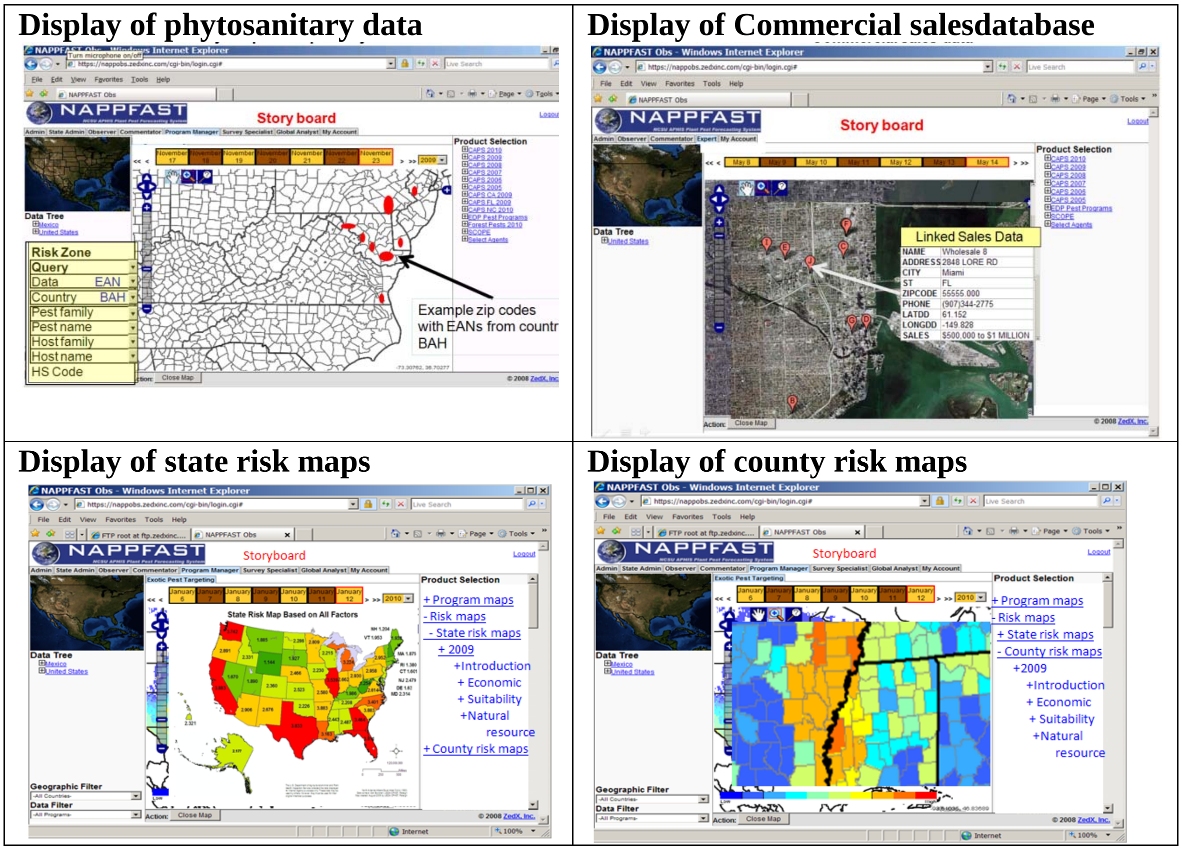

Figure 4.2. Storyboard mockups for remaining tasks in the Exotic Pest Targeting tool.

These additional components are expected to be completed in the Spring of 2010.

4.2 Risk Analyst Role

The role of the Risk Analyst is to create national scale risk analysis products which identify what geographies, hosts, industries, and periods at risk for an invasive species based on knowledge of past or present exotic pest distribution, movement, and behavior.

The Risk Analyst Role will include the following global analyst tools; i) Data QAQC, ii) Venn. In addition, the Risk Analyst role will also include the pathway data upload tool described below and the Pathways Analysis tool.

- Pathway Analysis tool is based on the Pathway algorithm that creates pathway risk maps using a combination of country, pest, host, and harmonized system (HS) code data. The Pathway algorithm will include:
  - i. Estimate distribution of commodities to 131 FAF regions based on FAF and USA Trade data.

ii. The Urban Density function distributes the volume of commodities reaching a FAF region area according to population density or commercial/wholesale land-use NLCD.

The risk analyst role is expected to be ready for testing at the end of February.

4.3 Tool to Upload Risk Analysis Data

A tool for uploading risk analysis data has been developed and is currently available in test form from the administrator role. The tool allows data to be uploaded at different spatial scales, including state, county, zip code, and point. The data fields in an uploaded risk analysis data file supports pathways data including data on point of origin and commodity types (Figure 4.3).

Data source	Observation Date	Data type	Data resolution	Commodity host family	commodity host	Pest family	Pest scientific name	Trade (HS Code)	Quantity code type	Quantity	Quantity un
PPQ	2/12/1992	PIN-309	Country		Rubus idaeus	RHOPALIDAE.		8	Shipment	0	BOX/CARTO
PPQ	2/12/1992	PIN-309	Country		Rubus idaeus	RHOPALIDAE.		8	Shipment	1	BOX/CARTO
PPQ	3/7/1992	PIN-309	Country		Rubus idaeus	RHOPALIDAE.		8	Shipment	0	
PPQ	3/9/1992	PIN-309	Country		Rubus idaeus	RHOPALIDAE.		8	Shipment	0	
PPQ	8/29/1992	PIN-309	Country		Asparagus officinalis	NOCTUIDAE.		8	Shipment	10808	PLANT UNIT
PPQ	8/29/1992	PIN-309	Country		Asparagus officinalis	NOCTUIDAE.		8	Shipment	1000	BOX/CARTO
PPQ	8/29/1992	PIN-309	Country		Asparagus officinalis	NOCTUIDAE.		8	Shipment	420	BOX/CARTO
PPQ	8/29/1992	PIN-309	Country		Asparagus officinalis	NOCTUIDAE.		8	Shipment	1	EACH
PPQ	8/31/1992	PIN-309	Country		Asparagus officinalis	NOCTUIDAE.		8	Shipment	500	BOX/CARTO
PPQ	9/1/1992	PIN-309	Country		Asparagus officinalis	NOCTUIDAE.		8	Shipment	0	
PPQ	9/24/1992	PIN-309	Country		Asparagus officinalis	NOCTUIDAE.		8	Shipment	1500	BOX/CARTO

Country	State	County	zip	Longitude	Latitude	site ID	Location type	Entry port
USA	Colorado		80110					
USA	Colorado		80229					
USA	Colorado		80229					
USA	Colorado		80229					
USA	Colorado		80229					

Shipment date	Origin Country	Origin State	Origin County	Origin Zip	Origin Longitude	Origin Latitude	Orgin site ID	Origin location type	Comments
	Chile								Fruit
	Chile								Fruit
	Chile								Fruit
	Chile								Fruit
	Chile								Fruit
	Chile								Fruit
	Chile								Fruit
	Chile								Fruit
	Chile								Fruit
	Chile								Fruit
	Chile								Fruit

Figure 4.3. Pathway upload form showing example data only (PIN 309 and EAN).

4.4 Tool to Link Third Party Data Sets

During the course of the project it was decided to manually upload the data from GBIF, rather than to connect directly. Consequently, this tool was not built.

4.5 Clean Critical Data Sets

The critical trade, phytosanitary, host, and country distribution data sets required for the Risk Analyst tool will be cleaned. The cleaning process will involve the downloading of data from source databases, data integration, quality control, and summarization. The cleaning process will remove any sensitive or personally identifiable information. The cleaned data would be uploaded to NAPPFAST-OBS using the data upload tool.

Phytosanitary data

The following phytosanitary data sources (Table 4.1), are in the process of being cleaned and will be added to the platform using the Pathways Upload tool.

Table 4.1. Critical phytosanitary data sets for exotic pest targeting

Database	Years	# of records	Entry point*	Destination	Host/commodity*
PPQ 309	03 to 08	383,951	Port	Destination (ST, many missing)	Inspected Host
PPQ 264	03 to 08	559,840	Port	Consignee Zip Code	Genus Sp. Name
PPQ 280	03 to 08	4,566,652	Port	Export Destination Country outside USA	Commodity
PPQ 523	03 to 08	131,099	Various	Destination State/Zip Code	Commodity
AQIM Cargo	03 to 08	39,602	Border Point	Destination city, state, zip code	Commodity
AQIM Vehicles	06 to 08	472,404	Border Point	Destination	
Park Visitors***	N/A	N/A	Park Location	Visitor Zip Code	

*Lat/long or zipcode centroid
** Train, truck, airplane, boat or mail
*** Source is National Park Reservation service. APHIS-WR is cleaning the data.

Commercial sales data

The Salesgenie database was purchased by ZedX. CPHST makes a specific business request e.g. wholesale nurseries in North Carolina and ZedX downloads the sales data and imports it into NAPPFAST-OBS using the pathways upload form. Data will be displayed at the spatial scale of a zip code. Once uploaded, it can be viewed in the Exotic Pest Targeting tool and filtered by geography, sales volume, and NAICS code or business type.

Concordance Tables

The following concordance tables have been or are in the process of being developed for the project.

- The following data sets are supplemental data to be provided by NCSU/APHIS.
  - NAICS to HS concordance table
    - Available at http://www.tradeinfo.net/itc.html
    - For crosswalking sales databases (NAICS) to trade (HS) databases. For use in the exotic pest targeting tool zooming.
  - FAF (SCTG) to HS concordance table
    - Constructed by Manuel Colunga-Gacria (MSU)
    - For crosswalking FAF (STCG) to USA trade (HS) database. For use in the risk analyst tool.
  - Table of Port Lat longs
    - Constructed by Manuel Colunga-Garcia (MSU)
    - For displaying port locations in map format.
  - Commodity list
    - Constructed by Jessica Engle (NCSU)
    - List of FIDO and NASS commodities and trees for use in the pathway analyst tool.
  - Zip code to County FIPS
    - Constructed by Jessica Engle (NCSU).
    - For determining county FIPS code from postal zip code or vice versa.
  - Pest genus to family
    - Constructed by Jessica Engle (NCSU).
    - Converts pest genus name to a family. For use in standardizing PIN 309 data.
  - Host genus to family

- Constructed by Jessica Engle (NCSU).
- Converts pest genus name to a family. For use in standardizing 264 and 280 data.

Other tables to convert shipment values, e.g. boxes, stems etc to tones, are under construction.

5 Mexican Data and Model Sharing Website

The Mexican data and model sharing website has been incorporated into NAPPFAST-OBS as the SCOPE program. The platform includes a spreadsheet for uploading observations and dynamic and planning risk map products for six pests of concern (Table 5.1). The website was rolled out to Mexican cooperators during a meeting at ZedX in State College PA, in October.

The completed SCOPE project included:

- Creation of a restricted access interface,
- Identification of the data formats and transfer protocols for collecting field data,
- Online data uploading form and spreadsheet,
- Layouts, including the placement of IT tools, for the “group and role” interfaces,
- Integration of the appropriate model products from the NAPPFAST system, and
- Training course to be taught at San Luis Potosi in March 2010.

Table 5.1. Pests included in the initial SCOPE program.

Pest Common name	Pest Scientific name	Model products
Cactus Moth	*Cactoblastis cactorum*	Weekly phenology maps
Citrus Greening (Hlb) (Asian)	*Candidatus Liberibacter asciaticus*	Host map
Asiatic Citrus Psyllid	*Diaphorina citri*	Risk map*
Pink Hibiscus Mealybug	*Maconellicoccus hirsutus*	Risk maps available but not yet incorporated
Rice Panicle Mite	*Steneotarsonemus spinki*	
Mediterranean Fruit Fly (medfly)	*Ceratitis capitata*	Weekly phenology models can be created
Central American Locust	*Schistocerca piceifrons piceifrons*	

* Incorporation in progress

The initial development included a translation of the database, interface¸ data upload forms, queries, and help instructions into Spanish. The translation is aided by a data translation table that can be expanded to other languages. Users can select a default language from their ‘My Account’ page.

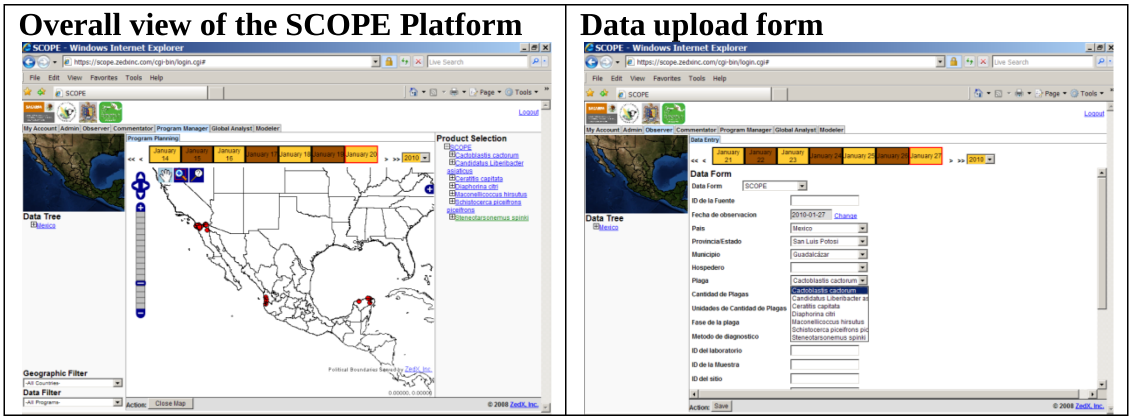

Figure 5.1. View of the SCOPE system

SINAVEF, a Mexican agency charged with developing the Mexican SCOPE system for modeling and data collection, is now funding the ongoing development of the independent SCOPE platform. The SCOPE platform will continue to be developed in parallel with NAPPFAST-OBS and eventually will include tools to securely share selected data sets between the two systems.

Appendix 1. NAPPFAST Presentations and training sessions in 2009

Reports

Chanelli, S., Magarey, R.D., Borchert, D.M., and Engle, J.S. 2009. NAPPFAST Generation Potential Validation. United States Department of Agriculture, Animal and Plant Health Inspection Service, Plant Protection and Quarantine, Center for Plant Health Science and Technology, Plant Epidemiology and Risk Analysis Laboratory (PERAL), Raleigh, NC. 11 p. http:/www.nappfast.org/pest%20reports/

Engle, J.S., Borchert, D.M., and Magarey, R.D. 2009. Pathway Risk Maps for Exotic Plant pests. United States Department of Agriculture, Animal and Plant Health Inspection Service, Plant Protection and Quarantine, Center for Plant Health Science and Technology, Plant Epidemiology and Risk Analysis Laboratory (PERAL), Raleigh, NC. 8 p. http:www.nappfast.org

Magarey, R.D.,Chanelli, S., and Holtz, T. 2009. Validation study and risk assessment *Guignardia citricarpa* (citrus black spot). United States Department of Agriculture, Animal and Plant Health Inspection Service, Plant Protection and Quarantine, Center for Plant Health Science and Technology, Plant Epidemiology and Risk Analysis Laboratory (PERAL), Raleigh, NC. 19 p. http:/www.nappfast.org/pest%20reports/guignardia_citricarpa.pdf.

Published Papers

Magarey, R.D., Colunga-Garcia, M., and Fieselmann, D. (2009) Plant Biosecurity in the United States: Roles, Responsibilities and Information Needs. Bioscience. 59:875-883.

Magarey, R.D., Borchert, D.M., and Schlegel, J. (2009). Global plant hardiness zones for phytosanitary risk analysis. Scientia Agricola. 64:54-59.

Presentations, Posters and training sessions

Borchert, D.M., Magarey, R.D., and Engle, J.S. 2009. NAPPFAST/GIS - US/Mexico Harmonization Workshop. April 6-10, 2009. Raleigh, NC.

Fowler, G.A. 2009. Updated Economic Analysis: Risk to U.S. Apple, Grape, Orange and Pear Production from the Light Brown Apple Moth, *Epiphyas postvittana*. Third Annual Pest Risk Mapping Workshop, August 11-13, 2009, Pescadero, CA.

Fowler, G.A., Garrett, L., Neeley, A., Magarey, R., Borchert, D., and Spears, B. 2009. Light brown apple moth (*Epiphyas postivittana*) economic analysis using GIS and quantitative modeling. Entomological Society of America Annual Meeting. December 13-16, 2009. Indianapolis, Indiana.

Magarey, R.D. 2009. NAPPFAST-OBS: Recent developments in Cyberinfrastructure for pest risk mapping. Third Annual Pest Risk Mapping Workshop, August 11-13, 2009, Pescadero, CA.

Magarey, R.D. and Fieselmann, D.A. 2009. Plant Biosecurity in the United States. Roles Responsibilities and information needs. Plant Pathology Society of North Carolina, March 18, 2009, Raleigh, N.C.

Magarey, R.D. and Russo, J.M. 2009. Model shoot-out. Third Annual Pest Risk Mapping Workshop, August 11-13, 2009, Pescadero, CA.

Magarey, R.D. and Borchert, D.M. 2009. New risk mapping tools for CAPS. National CAPS Committee Meeting. Januray 22, 2009. Raleigh, NC.

Appendix 2. NAPPFAST Weather Data sources

Table A2.1. Global historical weather data available in NAPPFAST.

Period	Description	Selected Resolution	Selected Time step	Selected Variables	Generation Procedure
1979-2009	Station daily values	NA	daily	tmpx tmpm tmpa pcpt lwht eowt wspa rlha	Climate test used to find outliers, and replace with Phase II grid value.
1979- near current	Global Reanalysis II Phase II grid daily values	32 km	daily	tmpx tmpm tmpa lwht eowt wspa rlha	Phase II QA/ QC process
1979- near current	Global Reanalysis II Phase I grid daily values	32 km	daily	pcpt	Phase I QA/ QC process

Table A2.2. Global real-time weather data available in NAPPFAST.

Period	Description	Selected Resolution	Selected Time step	Selected Variables	Generation Procedure
2009-present	Station daily values	NA	daily	tmpx tmpm tmpa pcpt lwht eowt wspa rlha	Climate test Temporal test Spatial test Source comparison test
About one week delay	Global Reanalysis II Phase II grid daily values	32 km	daily	tmpx tmpm tmpa lwht eowt wspa rlha	Phase II QA/ QC process
About 1 week delay	Global Reanalysis II Phase I grid daily values	32 km	daily	pcpt	Phase I QA/ QC process

Table A2.3. North American historical weather data available in NAPPFAST.

Period	Description	Selected Resolution	Selected Time step	Selected Variables	Generation Procedure
1979-2009	Station daily values	NA	daily	tmpx tmpm tmpa pcpt lwht eowt wspa rlha	Climate test Temporal test Spatial test Source comparison test
1979- near current	Global Reanalysis Phase II grid daily values	32 km	daily	tmpx tmpm tmpa lwht eowt wspa rlha	Phase II QA/ QC process
1979- near current	Global Reanalysis II Phase I grid daily values	32 km	daily	pcpt	Phase I QA/ QC process

Table A2.4. North American real-time weather data available in NAPPFAST.

Period	Description	Selected Resolution	Selected Time step	Selected Variables	Generation Procedure
2009 - present	Station daily values	NA	daily	tmpx tmpm tmpa pcpt lwht eowt wspa rlha	Climate test Temporal test Spatial test Source comparison test
About 1 week delay	Global Reanalysis II Phase II grid daily values	32 km	daily	tmpx tmpm tmpa lwht eowt wspa rlha	Phase II QA/ QC process
About 1 week delay	Global Reanalysis II Phase I grid daily values	32 km	daily	pcpt	Phase I QA/ QC process

Table A2.5. Conterminous United States historical weather data available in NAPPFAST.

Period	Description	Selected Resolution	Selected Time step	Selected Variables	Generation Procedures
1979-2009	Station daily values	NA	daily	tmpx tmpm tmpa pcpt lwht eowt wspa rlha	Climate test Temporal test Spatial test Source comparison test
1979- near current	Global Reanalysis II Phase II grid daily values	32 km	daily	tmpx tmpm tmpa lwht eowt wspa rlha	Phase II QA/ QC process
1979- near current	Global Reanalysis II Phase I grid daily values	32 km	daily	pcpt	Phase I QA/ QC process

Table A2.6. Conterminous United States real-time weather data available in NAPPFAST.

Period	Description	Selected Resolution	Selected Time step	Selected Variables	Generation Procedure
2009 – present	Station daily values	NA	daily	tmpx tmpm tmpa pcpt lwht eowt wspa rlha	Climate test Temporal test Spatial test Source comparison test
About 1 week delay	Global Reanalysis II Phase II grid daily values	32 km	daily	tmpx tmpm tmpa lwht eowt wspa rlha	Phase II QA/ QC process
About 1 week delay	Global Reanalysis II Phase I grid daily values	32 km	daily	pcpt	Phase I QA/ QC process