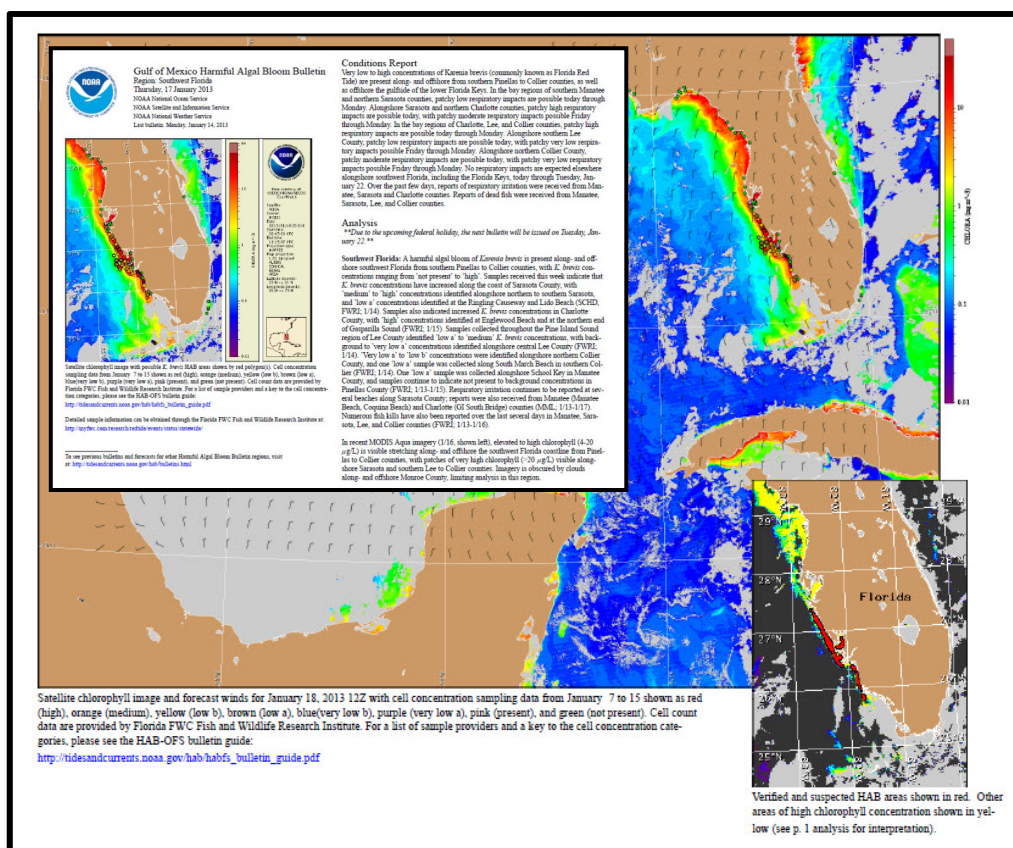


Assessment of the Eastern Gulf of Mexico Harmful Algal Bloom Operational Forecast System (GOMX HAB-OFS):

An Analysis of Forecast Skill and Utilization from May 1, 2008 to April 30, 2014



HAB-OFS bulletin issued for the eastern GOMX on January 17, 2013.

Silver Spring, Maryland
March 2016



noaa

National Oceanic and Atmospheric Administration

U.S. DEPARTMENT OF COMMERCE

National Ocean Service

Center for Operational Oceanographic Products and Services

Center for Operational Oceanographic Products and Services
National Ocean Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

The National Ocean Service (NOS) Center for Operational Oceanographic Products and Services (CO-OPS) provides the National infrastructure, science, and technical expertise to collect and distribute observations and predictions of water levels and currents to ensure safe, efficient and environmentally sound maritime commerce. The Center provides the set of water level and tidal current products required to support NOS' Strategic Plan mission requirements, and to assist in providing operational oceanographic data/products required by NOAA's other Strategic Plan themes. For example, CO-OPS provides data and products required by the National Weather Service to meet its flood and tsunami warning responsibilities. The Center manages the National Water Level Observation Network (NWLON), a national network of Physical Oceanographic Real-Time Systems (PORTS[®]) in major U. S. harbors, and the National Current Observation Program consisting of current surveys in near shore and coastal areas utilizing bottom mounted platforms, subsurface buoys, horizontal sensors and quick response real time buoys. The Center: establishes standards for the collection and processing of water level and current data; collects and documents user requirements, which serve as the foundation for all resulting program activities; designs new and/or improved oceanographic observing systems; designs software to improve CO-OPS' data processing capabilities; maintains and operates oceanographic observing systems; performs operational data analysis/quality control; and produces/disseminates oceanographic products.

Assessment of the Eastern Gulf of Mexico Harmful Algal Bloom Operational Forecast System (GOMX HAB-OFS):

**An Analysis of Forecast Skill and Utilization from
May 1, 2008 to April 30, 2014**

**Karen E. Kavanaugh
Katherine Derner
Edward Davis**

March 2016



**U.S. DEPARTMENT OF COMMERCE
Penny Pritzker, Secretary**

**National Oceanic and Atmospheric Administration
Dr. Kathryn D. Sullivan, NOAA Administrator and Under Secretary of
Commerce for Oceans and Atmosphere**

**National Ocean Service
Dr. Russell Callender, Acting Assistant Administrator**

**Center for Operational Oceanographic Products and Services
Richard Edwing, Director**

NOTICE

Mention of a commercial company or product does not constitute an endorsement by NOAA. Use of information from this publication for publicity or advertising purposes concerning proprietary products or the tests of such products is not authorized.

TABLE OF CONTENTS

TABLE OF CONTENTS.....	iii
LIST OF TABLES.....	vi
LIST OF FIGURES	vii
EXECUTIVE SUMMARY	xi
1. INTRODUCTION	1
1.1 Background.....	1
1.2 Objective.....	2
2. METHODS	3
2.1 Operations.....	3
2.1.1 <i>Modification to HAB-OFS Forecast Models and Skill Assessment Procedures</i>	4
3. RESULTS	7
3.1 Summary of <i>Karenia brevis</i> Events.....	7
3.1.1 <i>Bloom Year: 2008-2009</i>	9
3.1.2 <i>Bloom Year: 2009-2010</i>	11
3.1.3 <i>Bloom Year: 2010-2011</i>	13
3.1.4 <i>Bloom Year: 2011-2012</i>	13
3.1.5 <i>Bloom Year: 2012-2013</i>	16
3.1.6 <i>Bloom Year: 2013-2014</i>	18
3.2 Bulletin Utilization.....	20
3.2.1 <i>Priority Level</i>	21
3.3 Capability of Assessing the Forecasts.....	23

3.4	Accuracy of Forecasts	24
3.4.1	<i>Transport Direction</i>	25
3.4.2	<i>Intensification</i>	26
3.4.3	<i>Bloom Formation at the Coast</i>	27
3.4.4	<i>Respiratory Irritation</i>	29
3.5	Reliability of Forecasts	33
3.5.1	<i>Transport Direction</i>	33
3.5.2	<i>Intensification</i>	33
3.5.3	<i>Bloom Formation at the Coast</i>	33
3.5.4	<i>Respiratory Irritation</i>	35
3.6	Skill of Forecasts.....	37
3.6.1	<i>Transport Direction</i>	37
3.6.2	<i>Intensification</i>	37
3.6.3	<i>Bloom Formation at the Coast</i>	37
3.6.4	<i>Respiratory Irritation</i>	39
3.7	Comparison of New and Old Respiratory Irritation Forecast Model Performance	41
4.	DISCUSSION	43
4.1	Bulletin Utilization.....	43
4.2	Early Warning.....	43
4.3	Forecasts of Transport Direction	44
4.4	Forecasts of Intensification	45
4.5	Forecasts of Potential for Bloom Formation at the Coast.....	46
4.6	Forecasts of Respiratory Irritation	47

4.6.1	<i>Changes to the Respiratory Irritation Forecast Model and Assessment</i>	<i>47</i>
4.6.2	<i>Individual Respiratory Irritation Forecast Levels.....</i>	<i>48</i>
5.	CONCLUSION.....	50
5.1	Modifications Resulting in Improvements.....	50
5.2	Improvement Needs.....	50
5.3	Summary of Recommended Future Actions.....	51
6.	ACKNOWLEDGEMENTS.....	53
7.	REFERENCES	54
8.	APPENDICES	57
	APPENDIX I	58
	APPENDIX II.....	61
A.	Operations.....	61
B.	Definitions of the Forecast Types.....	65
C.	Skill Assessment.....	66
1.	<i>Overview of Procedure</i>	<i>66</i>
2.	<i>Statistical Analysis.....</i>	<i>69</i>
	APPENDIX III.....	73
	APPENDIX IV.....	80

LIST OF TABLES

Table 1. Definitions of nowcast and forecast components.	3
Table 2. The level of respiratory irritation forecasted and the corresponding population potentially affected.....	4
Table 3. Changes to the forecast models from May 1, 2008 through April 30, 2014 that impacted the assessment of bulletin forecasts and utilization.	6
Table 4. Changes to data availability and product dissemination that impacted the assessment of bulletin forecast components from May 1, 2008 through April 30, 2014.	6
Table 5. The number of HAB-OFS products issued during the 2004 to 2014 bloom years.....	8
Table 6. Estimate of the duration (in days) of <i>Karenia brevis</i> bloom features during the 2004 to 2014 bloom years.....	8
Table 7. Priority levels assigned to bulletins indicating the corresponding level of action or response that resource managers might deem necessary based on the status of a harmful algal bloom of <i>Karenia brevis</i>	64
Table 8. The categories assigned to <i>Karenia brevis</i> cell concentrations identified from water samples by state, county, and local organizations in Florida.....	66
Table 9. Data and resources used to assess each forecast type included in a Florida bulletin.	68
Table 10. Definitions of the levels of observed respiratory irritation as assessed by trained beach reporters for the Mote Marine Laboratory Beach Conditions Reporting System for the Gulf Coast of Florida (Kirkpatrick & Currier, 2010).....	68
Table 11. During a <i>Karenia brevis</i> bloom, reports of observed respiratory irritation were used to validate the corresponding level of respiratory irritation forecasted for that region according to this chart.....	69
Table 12. Example of a 2 x 2 contingency table showing the types of correct forecasts (hit and correct rejection) and false forecasts (false alarm and miss), with the letters A through D representing the number of events forecasted and/or observed.	71

LIST OF FIGURES

Figure 1. The number of new <i>Karenia brevis</i> features detected by HAB-OFS satellite imagery or samples collected in the field during the 2004 to 2014 bloom years.....	9
Figure 2. Monthly <i>Karenia brevis</i> samples collected during October through April in the 2008-2009 bloom year.	11
Figure 3. Monthly <i>Karenia brevis</i> samples collected during September through April in the 2009-2010 bloom year.	13
Figure 4. Monthly <i>Karenia brevis</i> samples collected during September through April in the 2011-2012 bloom year.	15
Figure 5. Monthly <i>Karenia brevis</i> samples collected during September of the 2012-2013 bloom year through May of the 2013-2014 bloom year.	18
Figure 6. Monthly <i>Karenia brevis</i> samples collected during October through December of the 2013-2014 bloom year.	19
Figure 7. Number of bulletins that were confirmed utilized and percentage of bulletins confirmed utilized for the 2004 to 2014 bloom years.	21
Figure 8. Average number of bulletins with utilization confirmed for each priority level and average percentage of bulletins utilized over the 2004 to 2014 bloom years. A priority level is assigned to each bulletin based on the need for management response.	22
Figure 9. Percent of bulletins that were confirmed utilized for each priority level for each bloom year from 2004-2014. A priority level is assigned to each bulletin based on the need for management response. No new <i>Karenia brevis</i> bloom developed during BY2010-2011, and therefore no “high” priority bulletins were issued.....	23
Figure 10. Number of assessable and unassessable forecasts for each bloom year from 2004 to 2014. The assessment of forecasts was dependent on the availability of reliable observational data from reputable sources. The label “na” indicates a value was not applicable due to the lack of forecasts issued in the specified bloom year.	24
Figure 11. Accuracy of transport direction forecasts issued during the 2004 to 2014 bloom years. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%.	26
Figure 12. Accuracy of intensification forecasts issued during the 2004 to 2014 bloom years. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation.	27

Figure 13. Accuracy of forecasts of bloom formation at the coast during the 2004 to 2014 bloom years. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation. 28

Figure 14. Accuracy of “low” respiratory irritation forecasts during the 2004 to 2014 bloom years. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation. 30

Figure 15. Accuracy of “moderate” respiratory irritation forecasts during the 2004 to 2014 bloom years. The label “na” indicates a value was not applicable due to no forecast issued in the specified bloom year. The label “0” indicates that the value was 0%. 31

Figure 16. Accuracy of “high” respiratory irritation forecasts during the 2004 to 2014 bloom years. The label “na” indicates a value was not applicable due to no forecast issued in the specified bloom year. The label “0” indicates that the value was 0%. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation. 32

Figure 17. Forecast reliability (bias) in transport, intensification, and bloom formation at the coast forecasts during the 2004 to 2014 bloom years (BY). A score of one indicates no bias, a score greater than one indicates that the forecast system over-forecasted the event, and a score less than one indicates that the forecast system under-forecasted the event. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation. 34

Figure 18. Forecast reliability (bias) in respiratory impact forecasts during the 2004 to 2014 bloom years (BY). A score of one indicates no bias, while a score of greater than one indicates that the forecast system over-forecasted the event. A score of less than one suggests that the forecast system under-forecasted the event. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation. 36

Figure 19. The forecast skill of transport direction, intensification, bloom formation at the coast and all respiratory irritation forecasts during the 2004 to 2014 bloom years (BY). The Heidke skill score is a skill corrected verification measure of categorical forecast performance that references the proportion of correct forecasts relative to the number of correct forecasts that could be made by random chance. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation. 38

Figure 20. The forecast skill of individual levels of respiratory irritation during the 2004 to 2014 bloom years (BY). The Heidke skill score is a skill corrected verification measure of categorical forecast performance that references the proportion of correct forecasts relative to the number of

correct forecasts that could be made by random chance. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation. 40

Figure 21. The forecast skill of the old and new respiratory irritation forecast models for individual levels of respiratory irritation during the 2009 to 2014 bloom years (BY). The Heidke skill score is a skill corrected verification measure of categorical forecast performance that references the proportion of correct forecasts relative to the number of correct forecasts that could be made by random chance. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%. 42

Figure 22. Map of Florida highlighting the three geographic regions for which HAB-OFS bulletins are disseminated. The Northwest region spans the coastal counties from Escambia through Taylor, the Southwest region spans from Dixie through Monroe (including the Florida Keys), and the eastern region spans from Miami-Dade through Nassau..... 63

Figure 23. Distribution of Florida bulletin subscribers among the three geographic regions for which HAB bulletins were disseminated from March 2002 to April 30, 2014. 64

EXECUTIVE SUMMARY

Blooms of the toxic dinoflagellate, *Karenia brevis*, occur nearly every year in coastal regions of the Gulf of Mexico causing potential impacts to public health, ecosystems, and regional economies. To aid early bloom identification and response efforts, NOAA's Gulf of Mexico Harmful Algal Bloom Operational Forecast System (GOMX HAB-OFS) issues semi-weekly bulletins that serve as decision support tools for coastal resource managers, federal and state agencies, public health officials, and academic institutions. In order to continually improve the GOMX HAB-OFS, product utilization and forecast quality (i.e. forecast accuracy, reliability, and skill) are evaluated regularly. This report provides an evaluation of the HAB-OFS products issued for Florida during the bloom years from May 1, 2008 to April 30, 2014 (BY2008-2014). During the assessment period, the HAB-OFS issued a total of 468 bulletins, 5 supplemental bulletins, and 28 conditions updates.

Key Results:

- **Product Utilization:**

- Likely underreported because it was not directly measured via survey or other tool and instead only confirmed when there was evidence available that bulletin content was used by a source such as a state or county agency, research institution, or public media entity.
- Increased from a range of 66.3-83.9% (BY2008-2012) to greater than 93.8% (BY2012-2014) due to greater product awareness in part attributed to the launch of the HAB-OFS Facebook Page in Fall 2012.

- **Early Bloom Detection:**

- Enhancements to satellite imagery products are needed.
 - BY2008-2009: all 3 *K. brevis* features detected first by HAB-OFS.
 - BY2009-2014: all *K. brevis* features detected by sampling first.
 - HAB-OFS decline in detection due to series of Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) sensor outages before its mission was terminated in February of 2011. The replacement, Moderate Resolution Imaging Spectroradiometer (MODIS) Aqua, has not performed as well.
 - Florida partner sampling has also increased, improving their detection capabilities.
- *Next steps:* As of September 2015, HAB-OFS analysts began consulting a newly operational ensemble satellite imagery product, which refines bloom detection (available on bulletin as of 9/8/15). The HAB-OFS is also in the process of evaluating higher resolution chlorophyll products from the Visible Infrared Imaging Radiometer Suite (VIIRS).

- **Respiratory Irritation Forecasts:**

- Improvements implemented to the model in August 2009 resulted in better forecast performance, especially for the “very low” and “low” levels. For example, the Heidke skill score for forecasts issued during BY2013-2014 showed that chance would perform better than the old model at predicting “very low”

respiratory irritation (-17.4%), while the new model showed 100% improvement over chance.

- Forecast performance was variable, improving with increasing bloom severity.
- *Next steps*: Evaluate and integrate new data sets into analysis as they become available, including methods to cost-effectively measure the concentration of brevetoxin in the air and water.

- **Transport Direction Forecasts:**

- BY2012-2014 performance was better than during BY2004-2012 with a 71.7-100% improvement over chance vs. 35.9-84.2%.
- The model did not change, but improvement may have been due to better team training and experience.
- *Next steps*: Explore hydrodynamic models that would enable a prediction of direction and distance of bloom movement.

- **Bloom Formation at the Coast Forecasts**

- Performed 25.0-58.3% better than chance at predicting potential bloom movement to the coast if a bloom had formed.
- Because not enough is known about how blooms form, the forecast predicts potential bloom movement and is prone to false alarms.
- *Next steps*: Evaluate new research findings regarding bloom development to establish a more comprehensive model.

- **Intensification Forecasts**

- Model changes implemented in November 2008 resulted in performance improvements. Forecasts issued during BY2009-2014 exhibited a 46.5-100% improvement over chance vs. a 39.5-51.9% improvement over chance during BY2004-2009.
- *Next steps*: Improvements to bloom detection would also enhance the evaluation of the intensification forecasts.

1. INTRODUCTION

1.1 Background

Blooms of a toxic dinoflagellate, *Karenia brevis*, commonly referred to as “red tide”, occur nearly every year along the eastern Gulf Coast of the United States, typically between August and December, and are reportedly the most common harmful algal bloom (HAB) occurring in the eastern Gulf of Mexico (Stumpf, et al., 2003). *K. brevis* blooms have been linked to fish kills and various marine bird and mammal deaths, and prompt the closure of shellfish beds to prevent Neurotoxic Shellfish Poisoning (NSP) in humans (Tomlinson, et al., 2004). Furthermore, toxins released from *K. brevis* cells can cause respiratory distress, which can include itchy eyes and throat, as well as difficulty breathing, especially for people with chronic respiratory illnesses such as asthma (Kirkpatrick, et al., 2004).

To assist coastal managers in mitigating the impacts of HABs, an ecological forecast system for the Gulf of Mexico was developed by the National Oceanic and Atmospheric Administration (NOAA) through a multi-office effort. In October 2004, this ecological forecast system for the coast of Florida was transitioned from research to operational status within the Center for Operational Oceanographic Products and Services (CO-OPS) establishing the GOMX HAB-OFS. To address the frequent *K. brevis* HABs in the western Gulf of Mexico, the HAB-OFS was also transitioned to operations along the Texas coast in 2010.

Operational GOMX HAB-OFS bulletins are produced twice weekly for a subscriber list of researchers and public service officials during active bloom events (once weekly at times of bloom inactivity). The bulletins provide information concerning the identification of new blooms in addition to monitoring existing blooms by providing forecasts for bloom movement, bloom intensification, and the potential level of respiratory irritation at the coast caused by *K. brevis* aerosols (see APPENDIX I for an example bulletin). These forecasts are publicly available via the Internet at <http://tidesandcurrents.noaa.gov/hab>. Since 2013, the National Weather Service (NWS) has also issued HAB-OFS forecast information as part of a text product called the Beach Hazards Statement when “high” levels of respiratory irritation are forecasted within the area of responsibility of the Weather Forecast Offices of Tampa Bay, Miami, or Key West. The Beach Hazards Statements are disseminated via NWS websites, NOAA Weather Radio and NOAA Weather Wire Service.

As a result of the forecasts in the bulletins, advance cautionary notice is issued to protect beachgoers from respiratory illness and guide mitigation activities related to shellfish harvesting, marine mammal rescues, and fish safety. The bulletins also use satellite imagery to identify potential HABs of *K. brevis*, providing advance notice to appropriate state, county, and local agricultural and health service departments to help the coordination of monitoring efforts. Once a bloom has been identified, the bulletins continue to provide updates on monitoring efforts, indicating the potential geographic extent of the confirmed bloom to allow for more effective and targeted field sampling. This, in turn, assists in confirming the specific location, extent, and severity of a toxic bloom, aids technological development of forecasting methods, and enhances scientific knowledge of the HAB species, *K. brevis*.

1.2 Objective

This report provides an evaluation of the HAB-OFS products issued for Florida during the bloom years from May 1, 2008 to April 30, 2014, with a comparative analysis of previously published data for October 1, 2004 to April 30, 2008 to relate the results across all years (Fisher, et al., 2006; Kavanaugh, et al., 2013). A bloom year (BY) refers to the time period from May 1, XXXX to April 30, YYYY, where BY2008-2009 spans the period from May 1, 2008 to April 30, 2009 and so on. This time period was selected to capture the typical initiation and termination period of *K. brevis* blooms in the Gulf of Mexico, enabling interannual comparisons. The analysis includes an assessment of bulletin utilization, early warning capability and forecast quality (i.e. accuracy, reliability, and skill). Previous publications have detailed the technology, models, and procedures that underlie the Florida HAB-OFS (Stumpf, et al., 2003; Tomlinson, et al., 2004; Stumpf, Litaker, Lanerolle, & Tester, 2008; Stumpf R. , et al., 2009). The results of this assessment will be used to guide enhancements to the operational forecast system with the goals of improving forecast quality through increased scientific understanding and the refinement of forecast models. Some of the recommendations may also be applicable to the HAB-OFS in the western Gulf of Mexico (Texas).

2. METHODS

2.1 Operations

Operations discussed in this report are relevant to the years from BY2004-2014. As shown in Table 1, the Florida HAB-OFS provides forecasts for four different bloom components: transport direction, intensification, potential for bloom formation at the coast, and potential level of associated respiratory irritation (see Table 2). For complete details regarding operations, forecast types, skill assessment, and statistical analysis, see APPENDIX II.

Table 1. Definitions of nowcast and forecast components.

FORECAST COMPONENT	DEFINITION	CATEGORIES	FORECAST BASED ON	EXAMPLE STATEMENT
Transport Direction	Direction a bloom is likely to migrate in relation to the coast	<ul style="list-style-type: none"> • North • South • East • West • No Change 	<ul style="list-style-type: none"> • NWS forecasted winds • Ekman transport • Geostrophic flow 	<i>“Southerly transport of K. brevis concentrations is possible today through August 15”</i>
Intensification	Expected increase in existing K. brevis concentration	<ul style="list-style-type: none"> • Increase • No Change 	<ul style="list-style-type: none"> • NWS forecasted winds • Upwelling/ downwelling favorable conditions 	<i>“Northeast winds forecast today through August 15 may increase the potential for intensification of K. brevis concentrations at the coast”</i>
Potential for Bloom Formation	Conditions are favorable for K. brevis to reach bloom-level concentrations at the coast	<ul style="list-style-type: none"> • Favorable • Unfavorable 	<ul style="list-style-type: none"> • NWS forecasted winds • Upwelling/ downwelling favorable conditions 	<i>“Upwelling favorable conditions today through August 15 may promote bloom formation at the coast”</i>
Respiratory Irritation	Highest potential level of respiratory irritation caused by the bloom (forecast by region for the next 3-4 days)	<ul style="list-style-type: none"> • Very low • Low • Moderate • High • None 	<ul style="list-style-type: none"> • NWS forecasted wind speed and direction • Highest K. brevis concentration within the most recent 7-10 days • Bloom proximity to shore • Validated reports of respiratory irritation at the coast associated with a bloom 	<p>Northern Sarasota: <i>Moderate (Th-M)</i></p> <p>Central Lee, bay regions: <i>Moderate (Th-F), High (Sa-M)</i></p>

Table 2. The level of respiratory irritation forecasted and the corresponding population potentially affected.

RESPIRATORY IMPACT LEVEL	AFFECTED POPULATION			
	None	Chronic Respiratory Conditions	Sensitive	General Public
None	X			
Very Low		X		
Low		X	X	
Moderate		X	X	X
High		X	X	X

2.1.1 Modification to HAB-OFS Forecast Models and Skill Assessment Procedures

Since the previous assessment period from BY2004-2008, the HAB-OFS has made several procedural modifications to improve upon the accuracy of HAB forecast models (see Table 3). There were also changes made to the methods used to assess the quality of forecasts (see Table 4). For a list of changes made to the HAB-OFS from BY2004-2008, please refer to (Kavanaugh, et al., 2013).

2.1.1.1 Forecast Updates

Extent

From BY2004-2010, the HAB team provided forecasts for changes in bloom extent, when applicable. During BY2010-2011, the HAB team decided to discontinue the forecast when analysis revealed that the extent forecast could not be reliably forecasted or assessed and an evaluation for extent forecasts was not included in this report.

Bloom Formation at the Coast

The additional forecast of bloom formation at the coast was added to the bulletin in October, 2008. This forecast was introduced to the HAB-OFS when the forecasts of intensification and transport direction were determined to be insufficient at addressing the formation of a new *K. brevis* bloom at the coast from a location offshore and at depth (Stumpf, Litaker, Lanerolle, & Tester, 2008).

In BY2010-2011, to ensure that assessment criteria were the same as those used by the Florida Fish and Wildlife Conservation Commission/Fish and Wildlife Research Institute (FWRI) and Mote Marine Laboratory (MML), the HAB-OFS changed the definition of a newly formed bloom by raising the threshold for *K. brevis* cell concentrations from 5,000 cells/L to greater than 10,000 cells/L. Since cell concentrations between 5,000-10,000 cells/L did not always precede the development of a bloom along the coast, this revised definition of a bloom refined the bloom formation at the coast forecast and assessment.

Respiratory Irritation

Several aspects of the respiratory irritation forecast model were refined in an effort to improve both the forecast model and assessment procedures. This included modifying the respiratory irritation matrices by creating a distinction between newly developing blooms and established blooms, redefining the threshold for “low b,” “medium,” and “high” concentrations, increasing the duration of forecasts to include the next bulletin day, and including sample data in the respiratory irritation model that was up to 10 days old instead of 7 days old. Additionally, a new map with defined half-county boundaries every 30-60 km along the coast was created (APPENDIX III). The map eliminated the need to reference large bays, or specific coastal landmarks when making a forecast, thereby creating a standard terminology for citing sampling locations and improving the ability of the analysts to assess previous forecasts consistently.

2.1.1.2 Data Updates

Satellite Imagery

In September 2009, a research version of a HAB ensemble imagery product became available for comparison to the operational imagery products with the purpose of refining HAB detection. In February, 2011, SeaWiFS, originally the primary source for ocean chlorophyll imagery was retired following a period of intermittent service outages. The operational backup sensor, MODIS Aqua, succeeded SeaWiFS as the primary operational sensor. In BY2012-2013, MODIS Terra was transitioned for use as an operational backup to the Aqua sensor.

Respiratory Irritation

In July of 2011, stations in Collier County were removed from the MML Beach Conditions Reporting System. Since respiratory irritation was no longer routinely monitored and reported, this reduced the ability of the HAB-OFS to assess respiratory irritation forecasts issued for the region.

2.1.1.3 Product Dissemination Updates

Facebook

In an effort to increase bulletin utilization, reach a larger audience, educate the general public about HABs, and increase the ability to verify the HAB-OFS forecasts, the NOAA HAB-OFS Facebook Page was created in September 2012. Product utilization was confirmed if there was evidence of interactions (“likes,” “shares,” “comments,” and “post clicks”) with conditions reports posted from the bulletins. When Facebook users reported experiencing the symptoms of respiratory irritation associated with a bloom of *K. brevis*, the observations were used to assess the respiratory irritation forecasts for the region. For more information on the assessment procedures, see Appendix I.C.1.

Table 3. Changes to the forecast models from May 1, 2008 through April 30, 2014 that impacted the assessment of bulletin forecasts and utilization.

Bloom Year	Effective Date	Description of Changes from BY2008-2014
<i>Extent Forecasts</i>		
BY2009-2010	12/28/2009	“Extent” forecasts were discontinued due to evaluation of forecast quality (Kavanaugh, et al., 2013).
<i>Bloom Formation at the Coast</i>		
BY2008-2009	06/04/2009	HAB-OFS began forecasting and assessing <i>K. brevis</i> bloom formation at the coast for Southwest Florida.
BY2010-2011	06/28/2010	Threshold for classifying <i>K. brevis</i> cell concentrations as a new bloom increased from 5,000 cells/L to >10,000 cells/L.
<i>Respiratory Irritation Forecasts</i>		
BY2009-2010	08/10/2009	Respiratory irritation table modified to include different criteria for “bloom initiation” and “developed blooms” (Kirkpatrick et al., 2004).
BY2009-2010	10/19/2009	Age of samples used for respiratory irritation forecast increased from 7 days to 10 days old.
BY2010-2010	06/14/2010	<i>K. brevis</i> concentration scale changed for “low b,” “medium,” and “high” concentrations.
BY2010-2010	03/30/2011	Introduced half-county boundary map for respiratory irritation.
BY2012-2013	10/15/2012	Duration of respiratory irritation forecasts were extended through the end of the next bulletin day (<i>Monday-Thursday; Thursday-Monday</i> rather than <i>Monday-Wednesday; Thursday-Sunday</i>).
BY2012-2013	January 2013	More detailed forecast region map was developed. Specific bodies of water and coastal locations described by their half-county location based on the HAB-OFS forecasting regions (i.e. central Lee County instead of Sanibel Island).
BY2012-2013	02/21/2013	NWS National Digital Forecast Database (NDFD) graphical forecasts used instead of marine text forecasts for respiratory irritation forecasts.
<i>Satellite Imagery</i>		
BY2009-2010	09/03/2009	A research version of a HAB ensemble imagery product became available for comparison to the operational imagery products and aid in HAB detection and assessment.
BY2010-2011	February 2011	SeaWiFS mission terminated. MODIS-Aqua replaced SeaWiFS as primary source for ocean color imagery.
BY2012-2013	06/29/2012	MODIS-Terra sensor transitioned for use as an operational backup.

Table 4. Changes to data availability and product dissemination that impacted the assessment of bulletin forecast components from May 1, 2008 through April 30, 2014.

Bloom Year	Effective Date	Description of Changes from BY2008-2014
<i>Bulletin Utilization</i>		
BY2012-2013	09/07/2012	Facebook was used to assess forecasts for bulletin utilization.
<i>Respiratory Irritation Forecasts</i>		
BY2011-2012	07/01/2011	Observations from Collier County no longer reported by MML Beach Conditions Reporting System
BY2012-2013	09/07/2012	Facebook was used to assess forecasts for respiratory irritation.

3. RESULTS

3.1 Summary of *Karenia brevis* Events

From the time the HAB-OFS was transitioned to operations on October 1, 2004 to the end of the most recent bloom year on April 30, 2014, a total of 867 bulletins, 5 supplemental bulletins, and 28 conditions updates were issued, containing 1540 forecasts (see Table 5). There was at least one *K. brevis* bloom in each year with the exception of BY2010-2011.

It is not always possible to verify that discontinuous *K. brevis* bloom features originated from the same bloom due to factors such as clouds in imagery that obscure bloom movement or fluctuations in cell concentrations. Therefore, although the HAB-OFS monitored multiple *K. brevis* bloom features within most years, for the purposes of this evaluation the features identified in each year may have represented either a continuous bloom that formed during the bloom year or the remaining fragments of a continuous bloom that formed earlier in the bloom year. Since BY2004, 10 of the 26 *K. brevis* bloom features were first identified in ocean color satellite imagery by the HAB-OFS. However, for the BY2008-2014 evaluation period, only the bloom during BY2008-2009 was identified by the HAB-OFS team using MODIS Aqua imagery (see Figure 1). The SeaWiFS mission was terminated in February 2011 following intermittent outages including during the initiation periods of the blooms that occurred during BY2008-2009 and BY2009-2010. All 9 of the *K. brevis* bloom features that were monitored during BY2009-2014 were identified first by water samples collected in the field by organizations in Florida (see APPENDIX IV).

From BY2008-2014, *K. brevis* blooms varied both in geographic extent and duration. The BY2012-2013 bloom was the longest lasting bloom of the evaluation period with an estimated 238 days of bloom activity that year (see Table 6). Maps of the monthly *K. brevis* samples collected during BY2008-2014 are shown in Figure 2 through Figure 6. Individual summaries for bloom years 2008 through 2014 can be found below. Kavanaugh et al. (2013) summarized the bloom years from 2004 through 2008.

Table 5. The number of HAB-OFS products issued during the 2004 to 2014 bloom years.

Bloom Year	# of HAB-OFS Products Issued	
	# of Scheduled Bulletins	# of Supplemental/ Conditions Updates
BY2004-2005 (10/1/04 to 4/30/05)	61	2
BY2005-2006 (5/1/05 to 4/30/06)	131	3
BY2006-2007 (5/1/06 to 4/30/07)	96	7
BY2007-2008 (5/1/07 to 4/30/08)	111	18
BY2008-2009 (5/1/08 to 4/30/09)	71	4
BY2009-2010 (5/1/09 to 4/30/10)	83	5
BY2010-2011 (5/1/10 to 4/30/11)	55	1
BY2011-2012 (5/1/11 to 4/30/12)	96	10
BY2012-2013 (5/1/12 to 4/30/13)	99	13
BY2013-2014 (5/1/13 to 4/30/14)	64	2

Table 6. Estimate of the duration (in days) of *Karenia brevis* bloom features during the 2004 to 2014 bloom years.

Bloom Year	# of <i>K. brevis</i> Bloom Features	Estimated Days w/ Bloom
BY2004-2005 (10/1/04 to 4/30/05)	2	178
BY2005-2006 (5/1/05 to 4/30/06)	5	309
BY2006-2007 (5/1/06 to 4/30/07)	1	293
BY2007-2008 (5/1/07 to 4/30/08)	6	215
BY2008-2009 (5/1/08 to 4/30/09)	3	121
BY2009-2010 (5/1/09 to 4/30/10)	2	199
BY2010-2011 (5/1/10 to 4/30/11)	1*	23
BY2011-2012 (5/1/11 to 4/30/12)	2	215
BY2012-2013 (5/1/12 to 4/30/13)	2	238
BY2013-2014 (5/1/13 to 4/30/14)	2	71

* Note: A lingering *Karenia brevis* bloom that initiated during BY2009-2010 lasted until 5/24/2010, an estimated 23 days into the BY2010-2011. However, no new bloom formed in BY2010-2011.

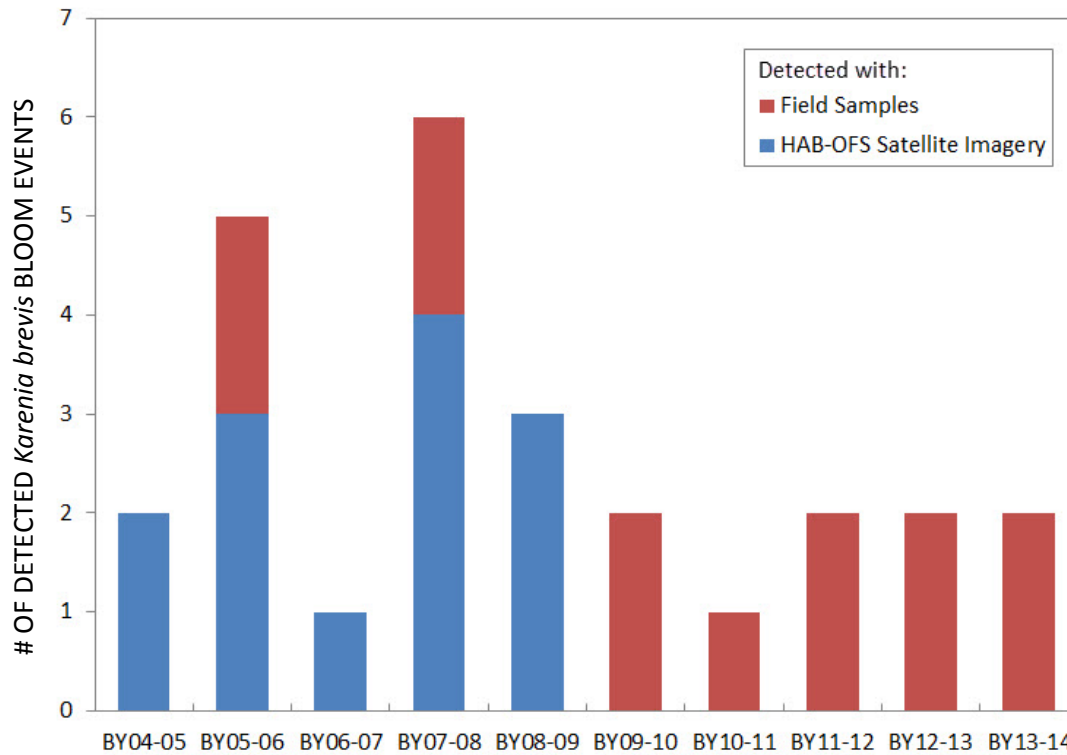
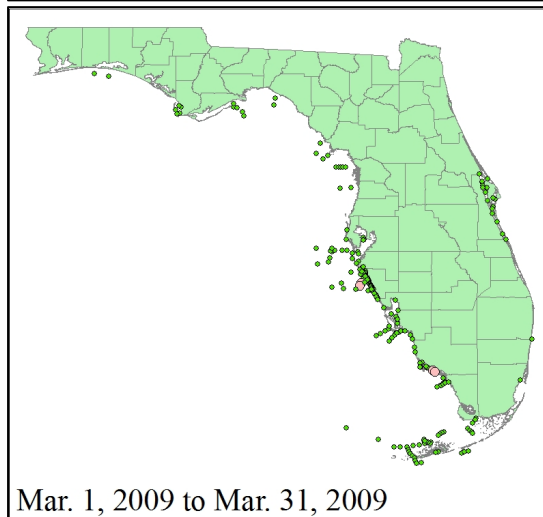
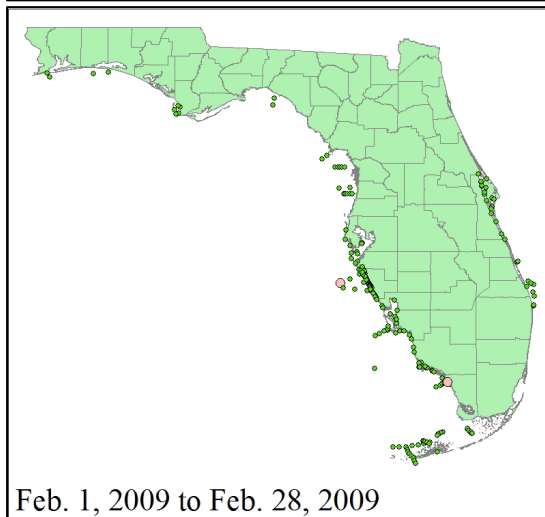
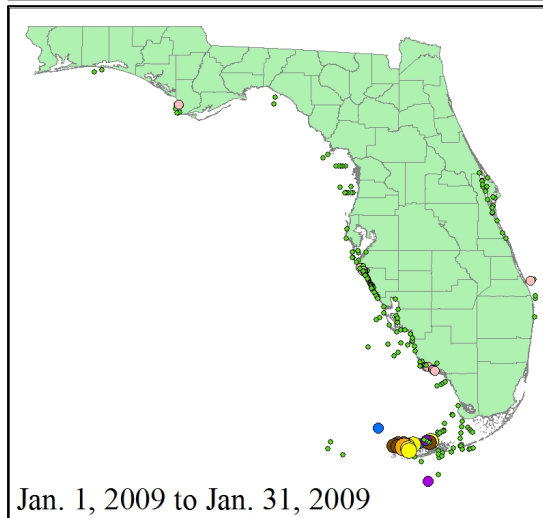
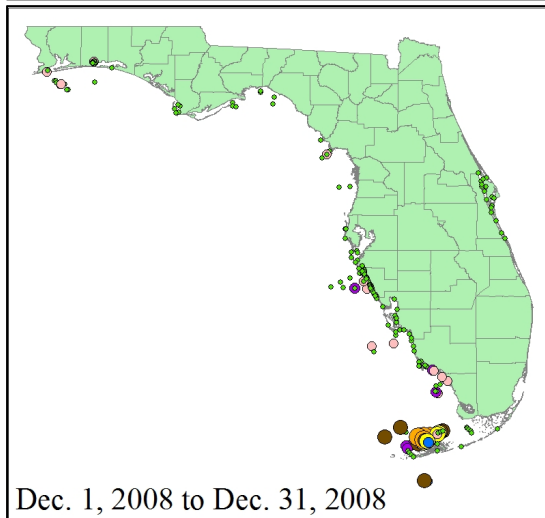
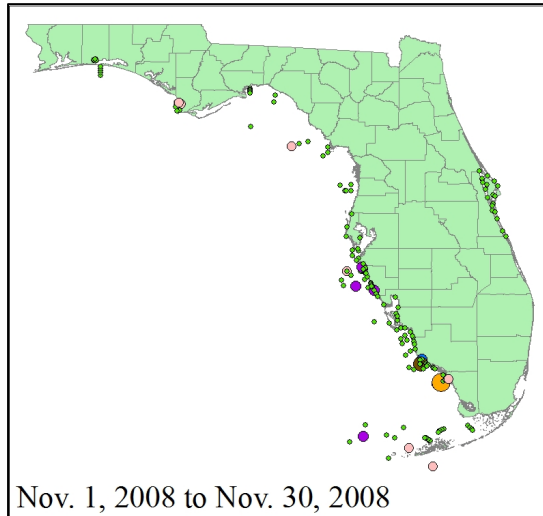
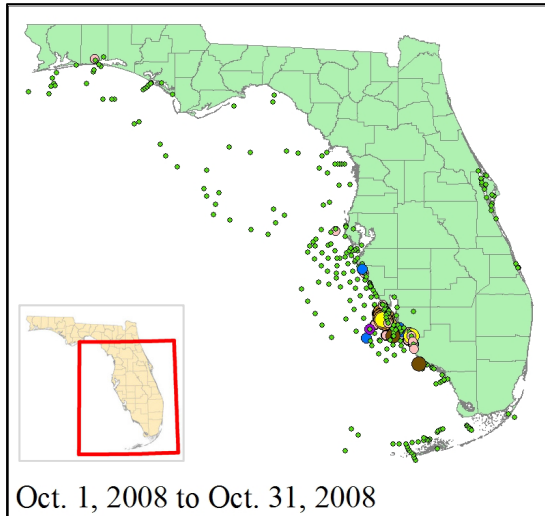


Figure 1. The number of new *Karenia brevis* features detected by HAB-OFS satellite imagery or samples collected in the field during the 2004 to 2014 bloom years.

3.1.1 Bloom Year: 2008-2009

- **Initial Feature Identification:**
 - *Feature #1:*
 - Identified by: HAB-OFS in satellite imagery
 - Timing: October 1, 2008
 - Location: alongshore Sanibel Island in central Lee County
 - *Feature #2:*
 - Identified by: HAB-OFS in satellite imagery
 - Timing: mid-November
 - Location: offshore Monroe County and the Florida Keys
 - *Feature #3:*
 - Identified by: HAB-OFS in satellite imagery
 - Timing: mid-December
 - Location: north of the Florida Keys
- **Duration:** October 2008 to January 2009
- **Extent:** Central Lee County to the lower Florida Keys
- **Transport:**
 - Feature #1 transported south from October through November, extending to the Marco Island region of Collier County.
 - Features #2 and 3 were likely fragments of the bloom that continued to transport south to the lower Keys.

- **Intensity:** Primarily “low” to “medium” cell concentrations (Figure 2).
- **Impacts:** No reports of respiratory irritation could be confirmed during BY2008-2009.
- **Reporting during bloom:** 28 bulletins, 2 supplemental bulletins, 1 conditions update.



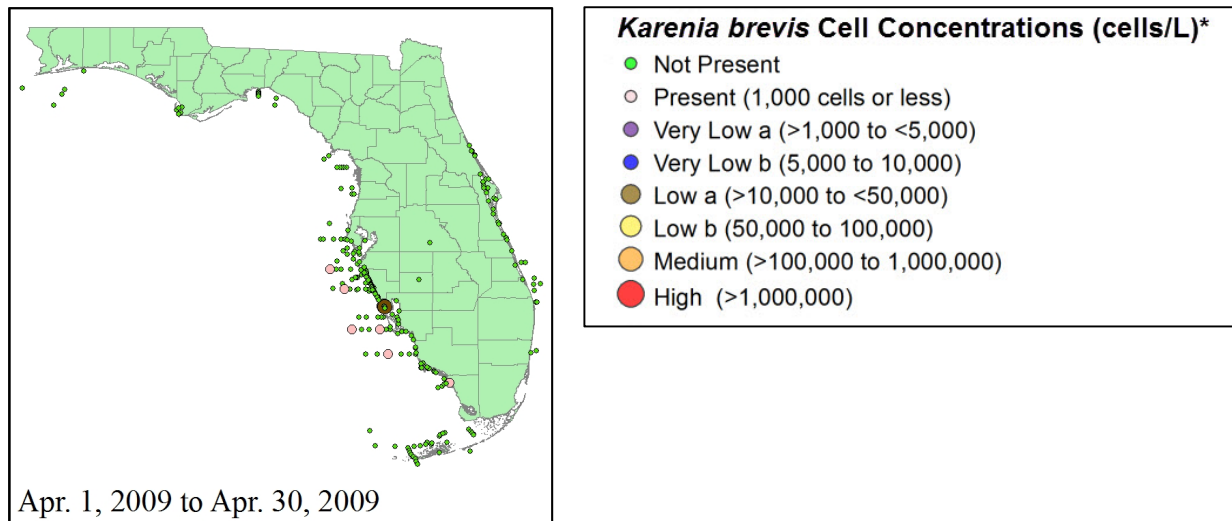
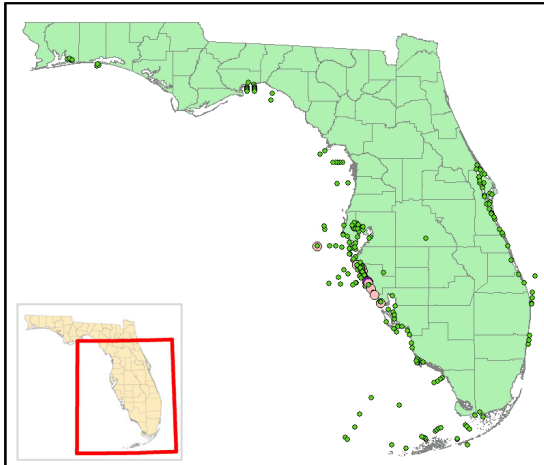


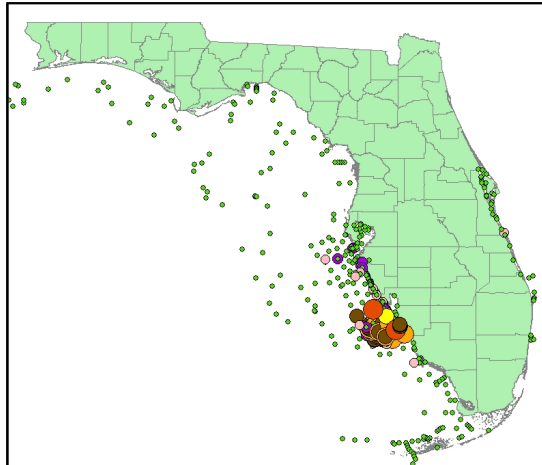
Figure 2. Monthly *Karenia brevis* samples collected during October through April in the 2008-2009 bloom year.

3.1.2 Bloom Year: 2009-2010

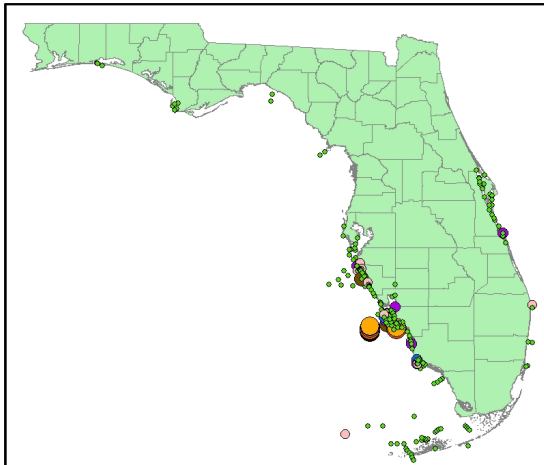
- **Initial Feature Identification:**
 - *Feature #1:*
 - Identified by: Field samples
 - Timing: October 10, 2009
 - Location: offshore Sanibel Island and Estero Island in central and southern Lee County
 - *Feature #2:*
 - Identified by: Field samples
 - Timing: January 19, 2010
 - Location: offshore gulf side of the Florida Keys
- **Duration:** October 2009 to April 2010
- **Extent:** Central Lee County to the lower Florida Keys
- **Transport:**
 - Feature #1 transported south from central and southern Lee County throughout the fall and into the winter, making it past the Marco Island region of central Lee County and into northern Monroe County by early December.
 - Features #2 was likely a fragment of the bloom that continued to transport south to the lower Keys, eventually transporting to the Atlantic side of the Florida Keys in March, where it dissipated.
- **Intensity:** “Very low” to “high” cell concentrations alongshore Southwest Florida, “very low” to “medium” cell concentrations along- and offshore the Florida Keys (Figure 3).
- **Impacts:**
 - Up to “high” levels of respiratory irritation reported from Lee and Collier counties, but no confirmed respiratory irritation in the Florida Keys.
 - Fish kills in Manatee, Lee, and Collier counties.
- **Reporting during bloom:** 52 bulletins, 1 supplemental bulletin, 4 conditions updates.



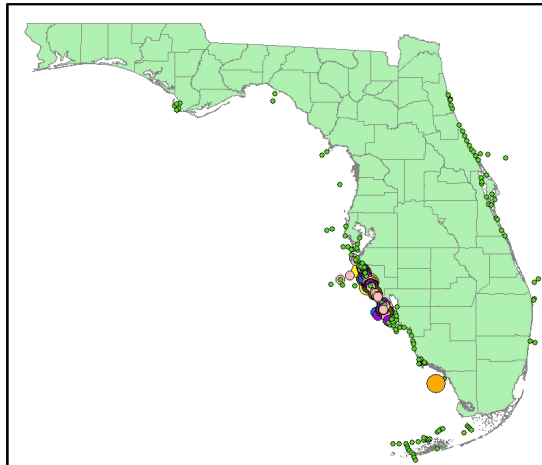
Sep. 1, 2009 to Sep. 30, 2009



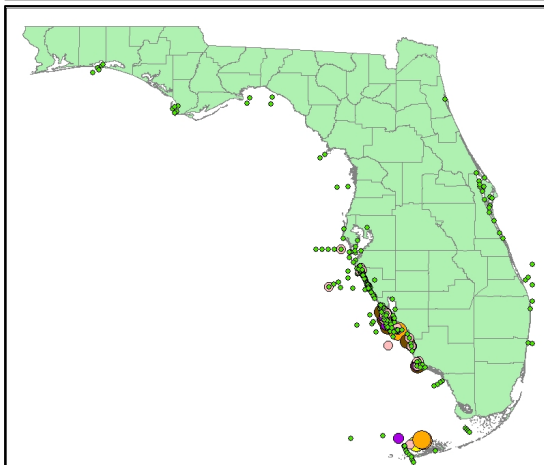
Oct. 1, 2009 to Oct. 31, 2009



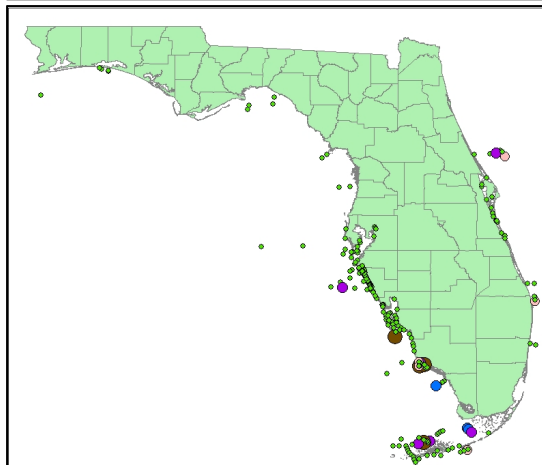
Nov. 1, 2009 to Nov. 30, 2009



Dec. 1, 2009 to Dec. 31, 2009



Jan. 1, 2010 to Jan. 31, 2010



Feb. 1, 2010 to Feb. 28, 2010

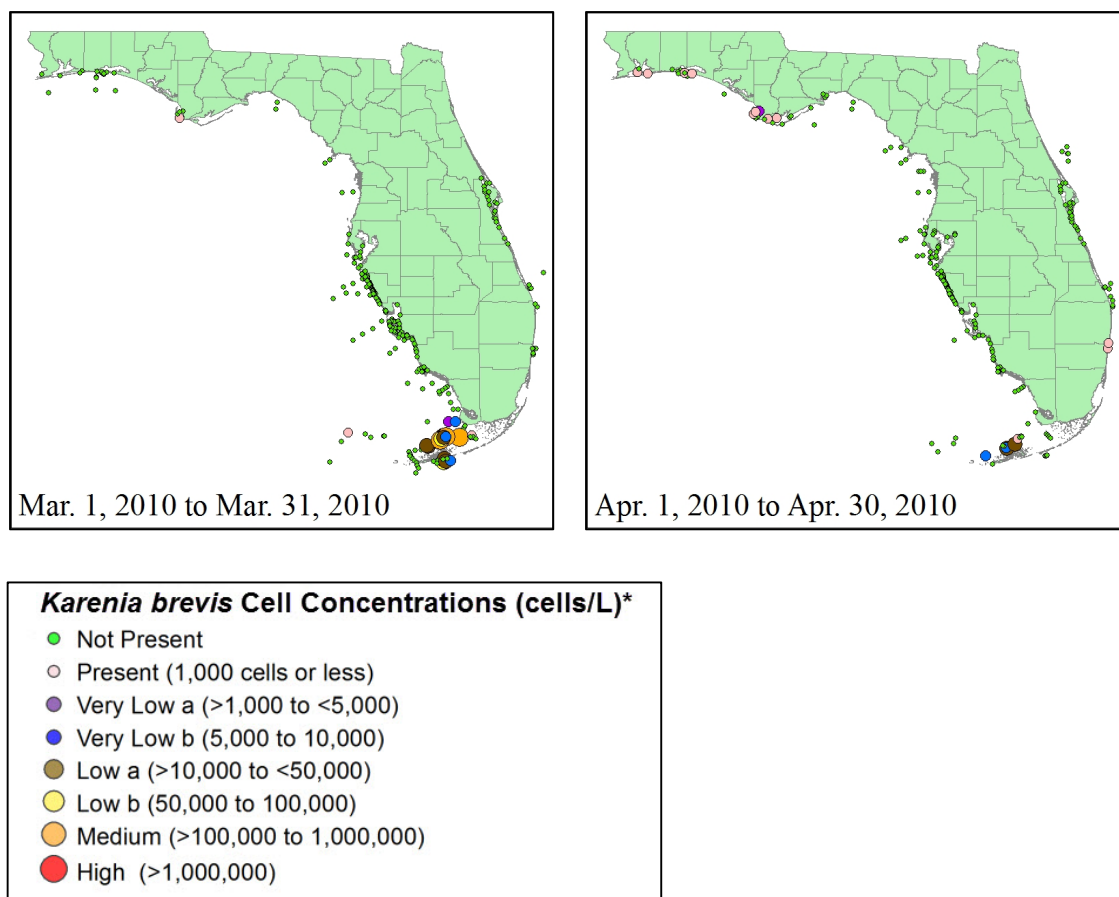


Figure 3. Monthly *Karenia brevis* samples collected during September through April in the 2009-2010 bloom year.

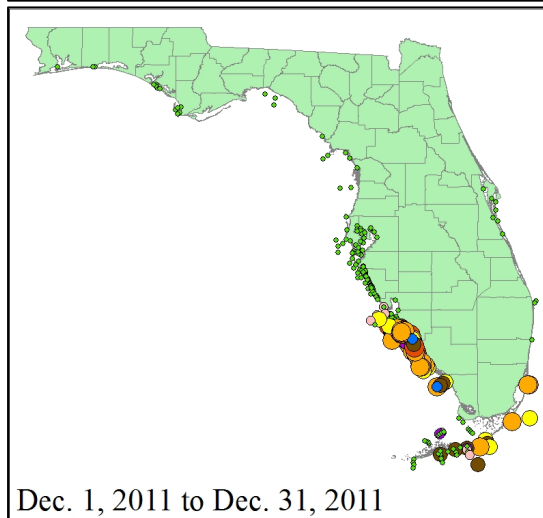
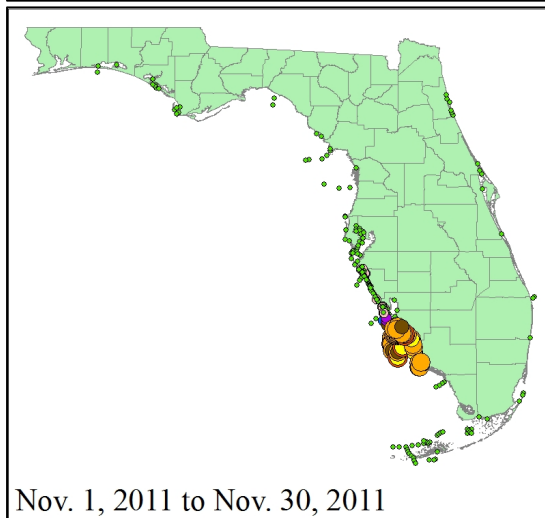
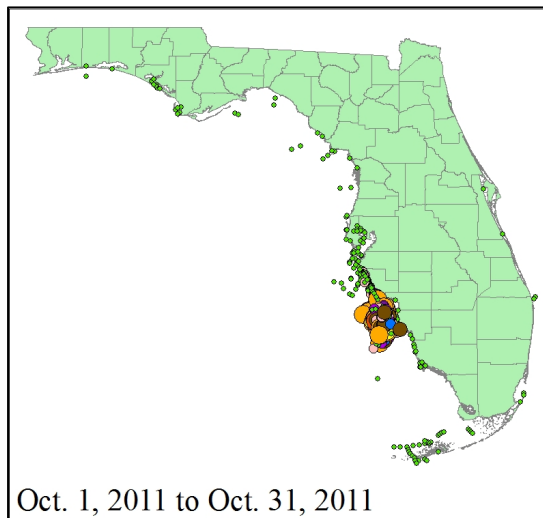
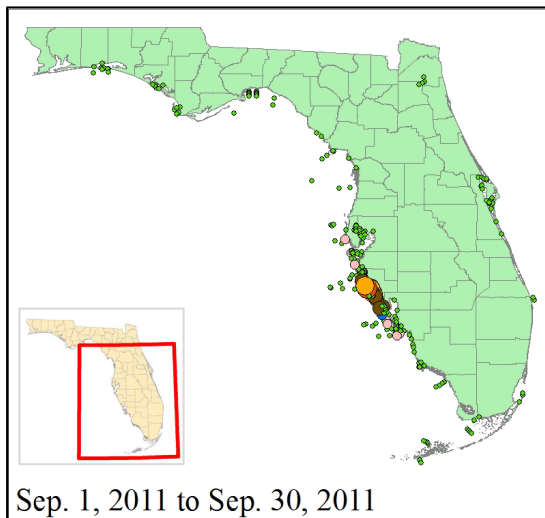
3.1.3 Bloom Year: 2010-2011

No new *K. brevis* bloom developed during the course of BY2010-2011. However, the bloom which initiated on January 19 during the BY2009-2010 interval persisted for 125 days, lingering for 23 days during BY2010-2011 until May 24. In addition, in October HAB-OFS analysts identified a high chlorophyll feature exhibiting the characteristics of a *K. brevis* bloom offshore southern Lee and northern Collier counties and co-located with “very low” concentrations of *K. brevis*. The feature was monitored through November, but was not confirmed as a bloom.

3.1.4 Bloom Year: 2011-2012

- **Initial Feature Identification:**
 - *Feature #1:*
 - Identified by: Field samples
 - Timing: September 28, 2011
 - Location: alongshore Manasota Key in southern Sarasota County
 - *Feature #2:*
 - Identified by: Field samples
 - Timing: December 11, 2011
 - Location: offshore the Upper Keys and Miami-Dade County

- **Duration:** September 2011 to May 2012
- **Extent:** Southern Sarasota County to the Florida Keys
- **Transport:**
 - Feature #1 transported south from southern Sarasota County, lingering alongshore Lee and Collier counties in October and November before extending further southward where it was detected along- and offshore Monroe County and the northern side of the Florida Keys, dissipating in mid-May 2012.
 - Feature #2 was likely a fragment of the Gulf Coast bloom that was transported in mid-December to the Atlantic side of the middle and upper Florida Keys, where it dissipated in early January 2012.
- **Intensity:** “Very low” to “medium” cell concentrations (Figure 4).
- **Impacts:**
 - Up to “high” levels of respiratory irritation reported from Sarasota and Collier counties, and up to “moderate” levels reported from Lee County.
 - Fish kills were reported from northern Sarasota to central Collier counties.
- **Reporting during bloom:** 70 bulletins, 10 conditions updates.



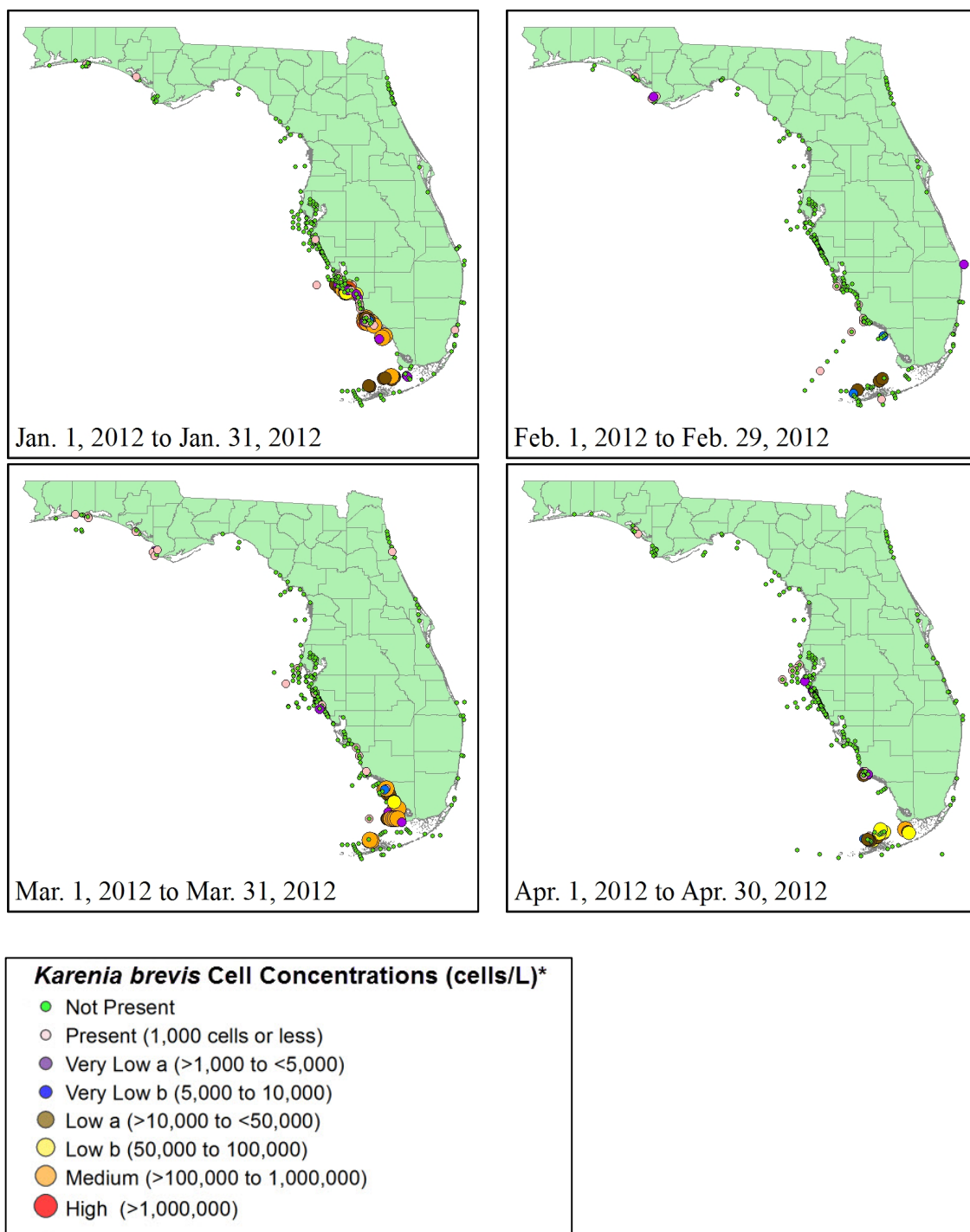
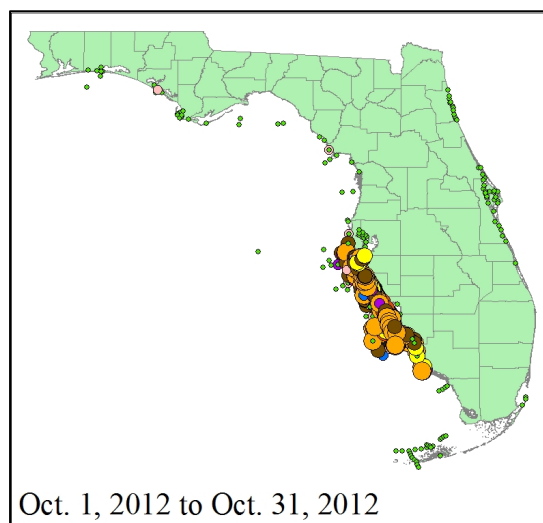
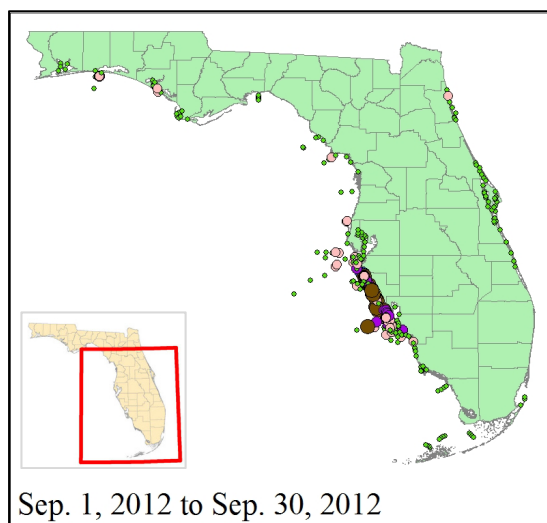
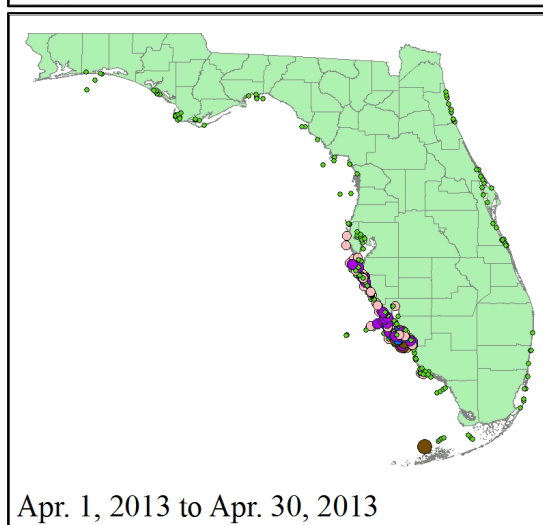
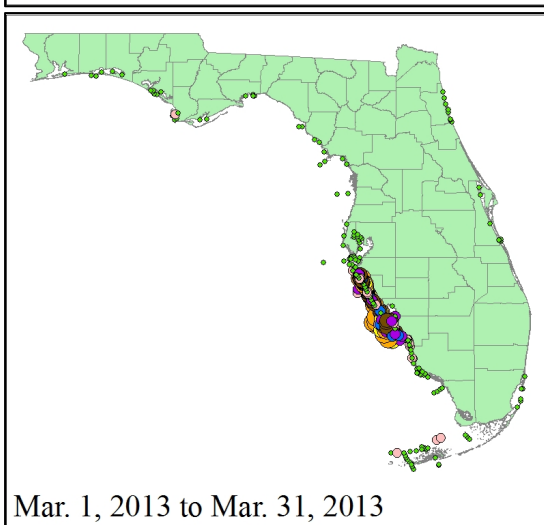
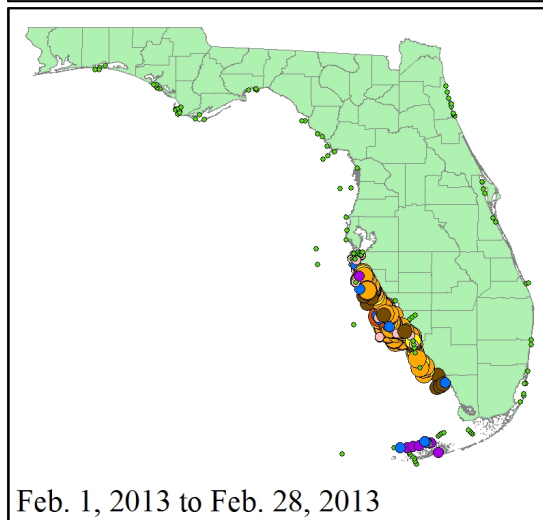
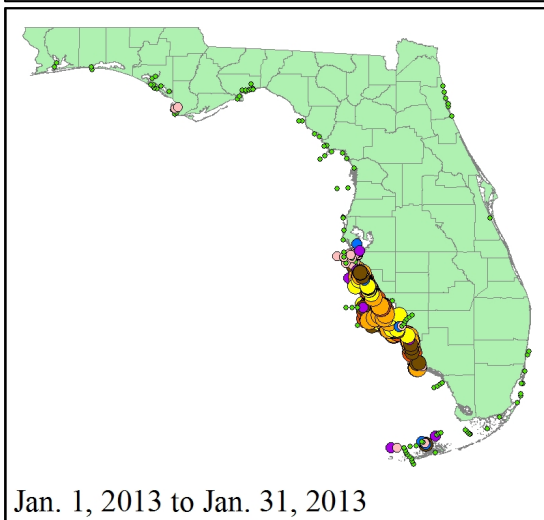
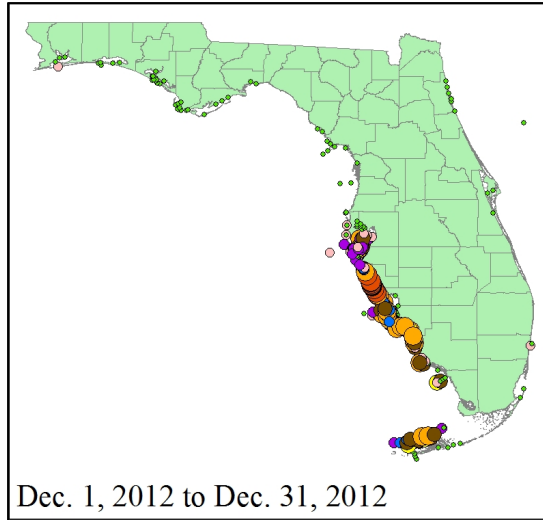
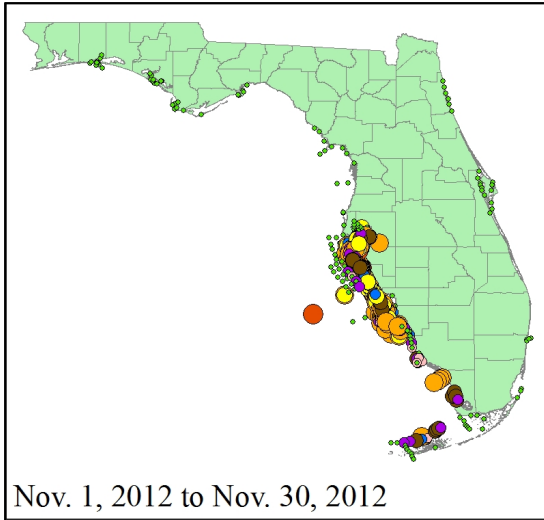


Figure 4. Monthly *Karenia brevis* samples collected during September through April in the 2011-2012 bloom year.

3.1.5 Bloom Year: 2012-2013

- **Initial Feature Identification:**
 - *Feature #1*: Continued from Feature #1 during Bloom Year: 2011-2012
 - *Feature #2*:
 - Identified by: Field samples
 - Timing: September 24, 2012
 - Location: alongshore southern Sarasota County and offshore northern Charlotte County
- **Duration:** September 2012 to May 2013
- **Extent:** Southern Sarasota County to the Florida Keys
- **Transport:**
 - Feature #2 gradually extended southward from southern Sarasota County, lingering in southern Sarasota to northern Monroe counties through the winter and in the bay regions of Lee County through the spring, with the bloom completely dissipating in late May 2013.
 - “Background” to “medium” concentrations were detected along the gulf side of the Florida Keys from November through February.
- **Intensity:** “Low” to “high” cell concentrations (Figure 5).
- **Impacts:**
 - Up to “high” levels of respiratory irritation reported along the coast of Southwest Florida from Sarasota to Collier counties.
 - Fish kills were reported from Pinellas to Collier counties.
 - Deadliest bloom on record for the endangered Florida manatee. As of May 2013, the Florida Fish and Wildlife Conservation Commission (FWC) recorded 277 manatee deaths as a result of brevetoxin exposure, a number that far surpassed the previous record of 151 manatees claimed by *K. brevis* in 1996 (Florida Fish and Wildlife Conservation Commission, 2015).
- **Reporting during bloom:** 79 bulletins, 1 supplemental bulletin, 11 conditions updates.





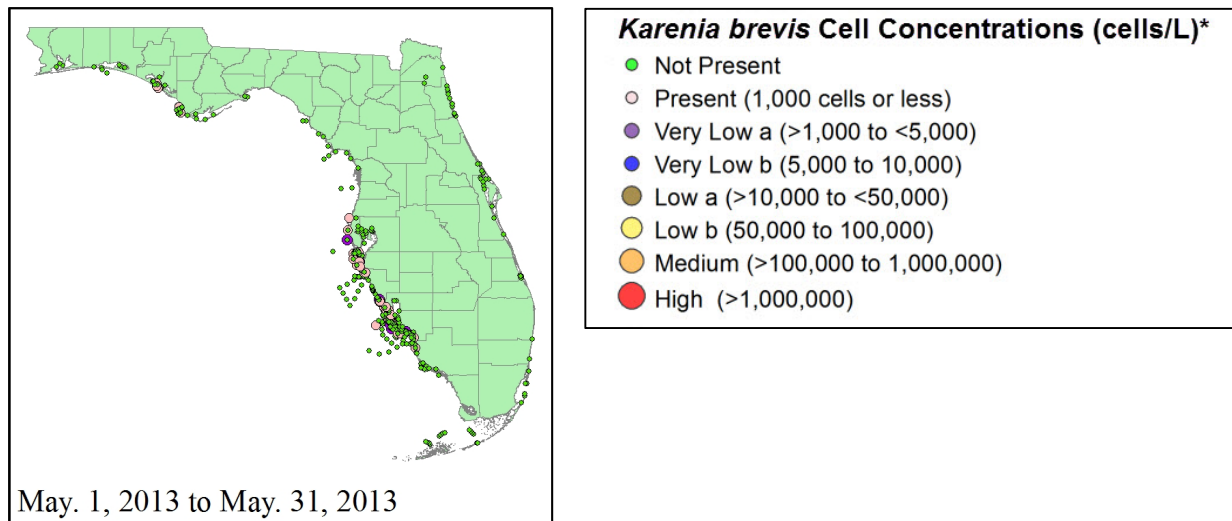


Figure 5. Monthly *Karenia brevis* samples collected during September of the 2012-2013 bloom year through May of the 2013-2014 bloom year.

3.1.6 Bloom Year: 2013-2014

- **Initial Feature Identification:**
 - *Feature #1:* Continued from Feature #1 during Bloom Year: 2012-2013
 - *Feature #2:*
 - Identified by: Field samples
 - Timing: October 28, 2013
 - Location: alongshore southern Sarasota County
- **Duration:** October to December 2013
- **Extent:** Southern Sarasota County to Monroe County
- **Transport:**
 - Feature #2 transported south from southern Sarasota County, slowly expanding and drifting southward into Charlotte, Lee, and northern Collier counties.
 - The bloom remained offshore throughout much of its duration.
- **Intensity:** “Low” to “medium” cell concentrations (Figure 6)
- **Impacts:**
 - Up to “moderate” respiratory irritation reported from Charlotte, Lee, and Collier counties, but few reports received.
 - Fish kills reported from Charlotte, Lee, and Collier counties.
- **Reporting during bloom:** 19 bulletins, 1 supplemental bulletin, 1 conditions update.

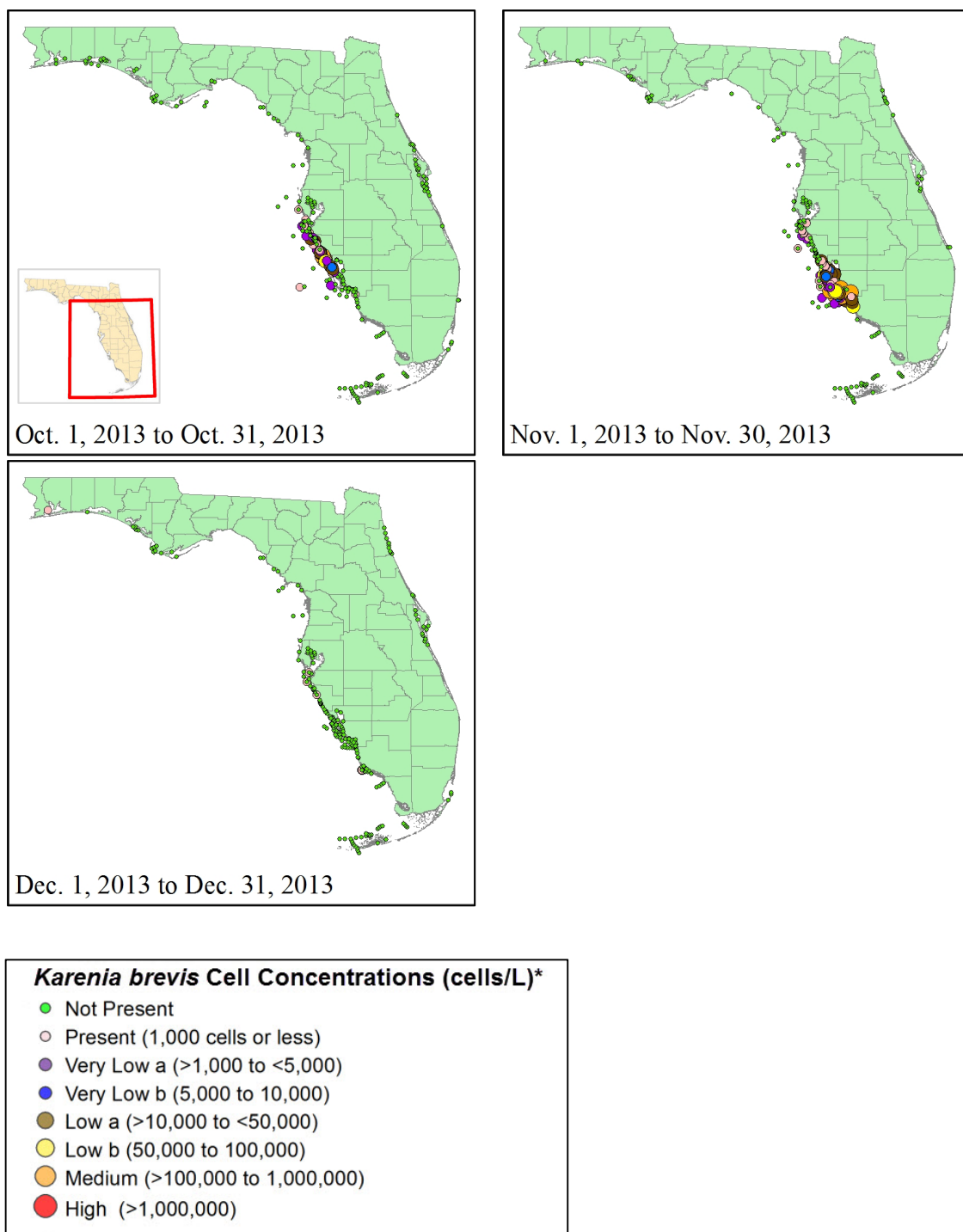
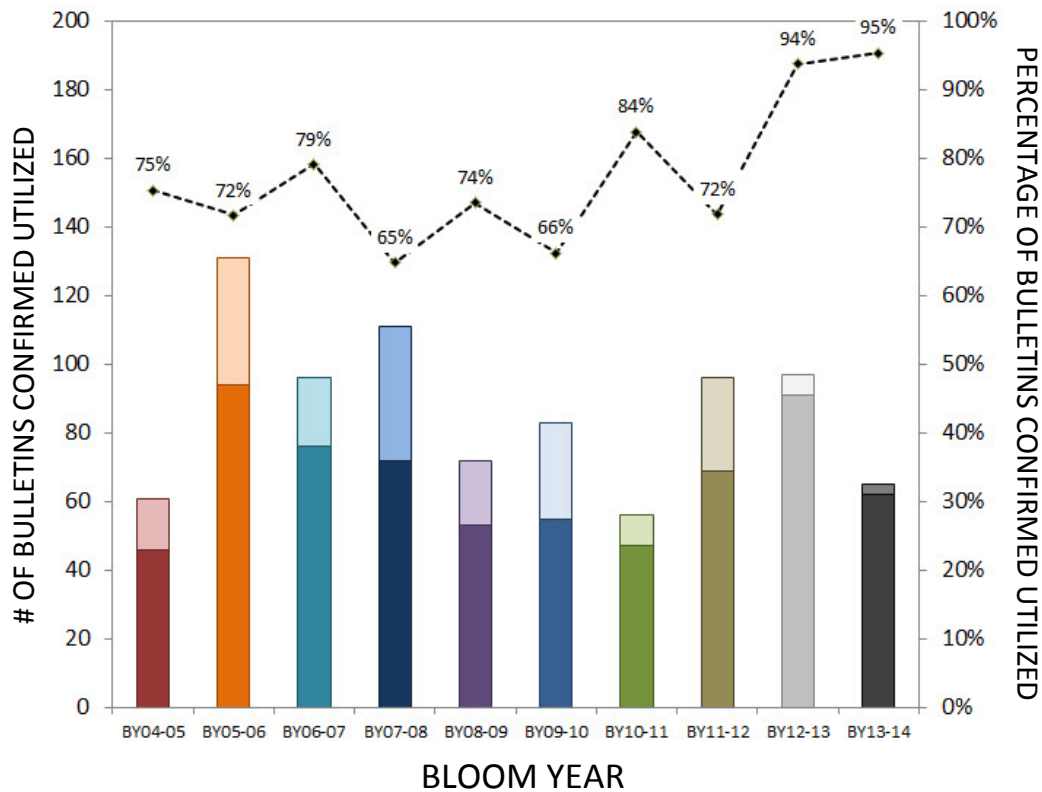


Figure 6. Monthly *Karenia brevis* samples collected during October through December of the 2013-2014 bloom year.

3.2 Bulletin Utilization

There were two categories of usage that counted toward total product utilization, viewing the product and applying its content to bloom response. Confirmation of use was dependent upon the availability of supporting evidence indicating that bulletin content was used by another source such as a state or county agency, research institution, or public media entity. After the HAB-OFS Facebook Page was launched in September 2012, Facebook metrics of active viewership and interactions (“likes,” “shares,” “comments,” and “post clicks”) with the posted conditions reports were used to determine product usage (see Appendix I.C.1). Overall, the proportion of total bulletins with confirmed product utilization increased over the BY2008-2014 assessment period, with 66.3-83.9% confirmed utilized during BY2008-2012 and 93.8-95.4% confirmed utilized in BY2012-2014 (see Figure 7). The increase to greater than 90.0% confirmed utilized in BY2012-2014 was largely due to the launch of the HAB-OFS Facebook Page and the usage of Facebook metrics to assess product utilization. Prior to the launch of the Facebook Page, the average confirmed bulletin utilization during BY2008-2012 was 73.9%, similar to the average during BY2004-2008, 72.8%.

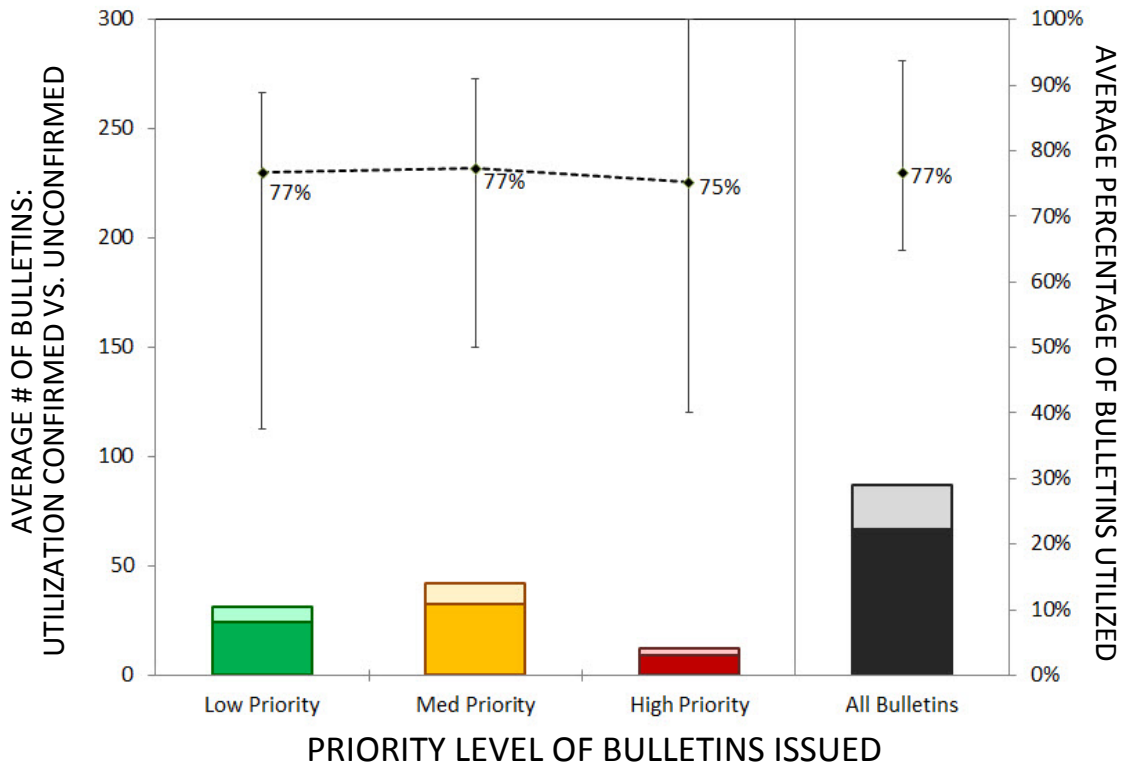


Unconfirmed Product Utilization	04-05	05-06	06-07	07-08	08-09
	09-10	10-11	11-12	12-13	13-14
Confirmed Product Utilization	04-05	05-06	06-07	07-08	08-09
	09-10	10-11	11-12	12-13	13-14

Figure 7. Number of bulletins that were confirmed utilized and percentage of bulletins confirmed utilized for the 2004 to 2014 bloom years.

3.2.1 Priority Level

A priority level (low, medium, or high) was assigned to each bulletin based on bloom activity and the corresponding level of action or response that resource managers might deem necessary (see Table 7). Overall, 76.6% of all bulletins were confirmed utilized during BY2004-2014 (see Figure 8). Bulletin utilization for each of the priority levels individually was similar to the overall utilization, with 76.8% of low, 77.3% of medium, and 75.2% of high priority bulletins confirmed utilized. While the differences in overall utilization between the priority levels are negligible, there has been an increase in utilization since social media began to be used to disseminate the public conditions reports (see Figure 9). Other than perfect utilization for high priority bulletins during the first operational year (BY2004-2005), all confirmed utilization rates greater than 90.0% have occurred from BY2012-2014, during seasons when Facebook was used to disseminate public conditions reports.



Unconfirmed Product Utilization	Low Priority	Medium Priority	High Priority	Total Bulletins
Confirmed Product Utilization	Low Priority	Medium Priority	High Priority	Total Bulletins

Figure 8. Average number of bulletins with utilization confirmed for each priority level and average percentage of bulletins utilized over the 2004 to 2014 bloom years. A priority level is assigned to each bulletin based on the need for management response.

While it might be expected that bulletin utilization should increase as the priority level increases, analysis reveals that this is not always the case (see Figure 9). In fact, from BY2008-2012, the inverse was true, with the proportion of low priority bulletins confirmed utilized greater than that of the medium and high priority bulletins. While utilization of low priority bulletins has varied between individual years, there has been an overall steady increase in low priority bulletin utilization from the beginning of operations (BY2004) to present (BY2014). As shown in Figure 9, the spread between the utilization of each priority level was reduced during several years of the BY2008-2014 assessment period, including BY2009-2010, BY2012-2013, and BY2013-2014. For the latter two years, this can be attributed in part to general increases in utilization resulting from using Facebook to disseminate bulletin information, which in turn increased the ability to confirm product utilization.

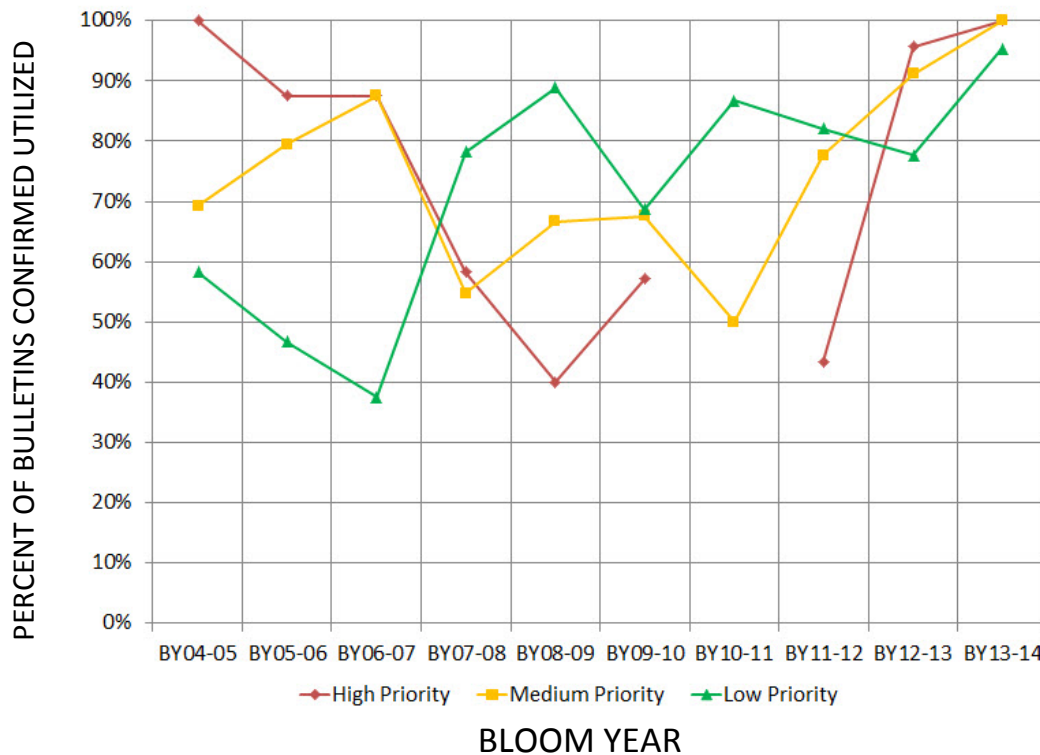


Figure 9. Percent of bulletins that were confirmed utilized for each priority level for each bloom year from 2004-2014. A priority level is assigned to each bulletin based on the need for management response. No new *Karenia brevis* bloom developed during BY2010-2011, and therefore no “high” priority bulletins were issued.

3.3 Capability of Assessing the Forecasts

The assessment of forecasts was dependent on the availability of reliable observational data from reputable government, scientific and academic sources. Due to the patchy nature of blooms, respiratory irritation typically did not affect an entire forecast region (see APPENDIX III) and reports of observations were limited. The majority of respiratory irritation reports were provided by the MML Beach Conditions Reporting System, a network of trained lifeguards who report levels of respiratory irritation in real-time. However, this network does not cover the entire Florida coast and observational evidence is limited elsewhere. Thus, the method of assessing the respiratory irritation forecast could not verify that no irritation was observed throughout an entire forecast area during the forecast period. When the necessary observational evidence was not available from the MML Beach Conditions Reporting System or other sources, the forecast was categorized as “unconfirmed” and forecast quality could not be assessed. Thus, our method of assessment did not allow us to verify that. Bloom duration and intensity varied from year to year, which also influenced the number of forecasts that were issued and could be assessed (see Figure 10).

No new bloom developed along- or offshore the Florida coast during BY2010-2011. As the bloom from the previous year dissipated, forecasts were issued, but only bloom formation at the coast and intensification could be assessed.

For each year during BY2008-2014, the most assessable forecasts were the potential for bloom formation at the coast (87.0-100%). The least assessable were forecasts for respiratory irritation at the coast (0-17.3%), which were also typically the most frequently issued because the conditions could be predicted throughout the year; even when there was no bloom, a forecast for “no” respiratory irritation was issued, which still counted as a prediction that was assessed.

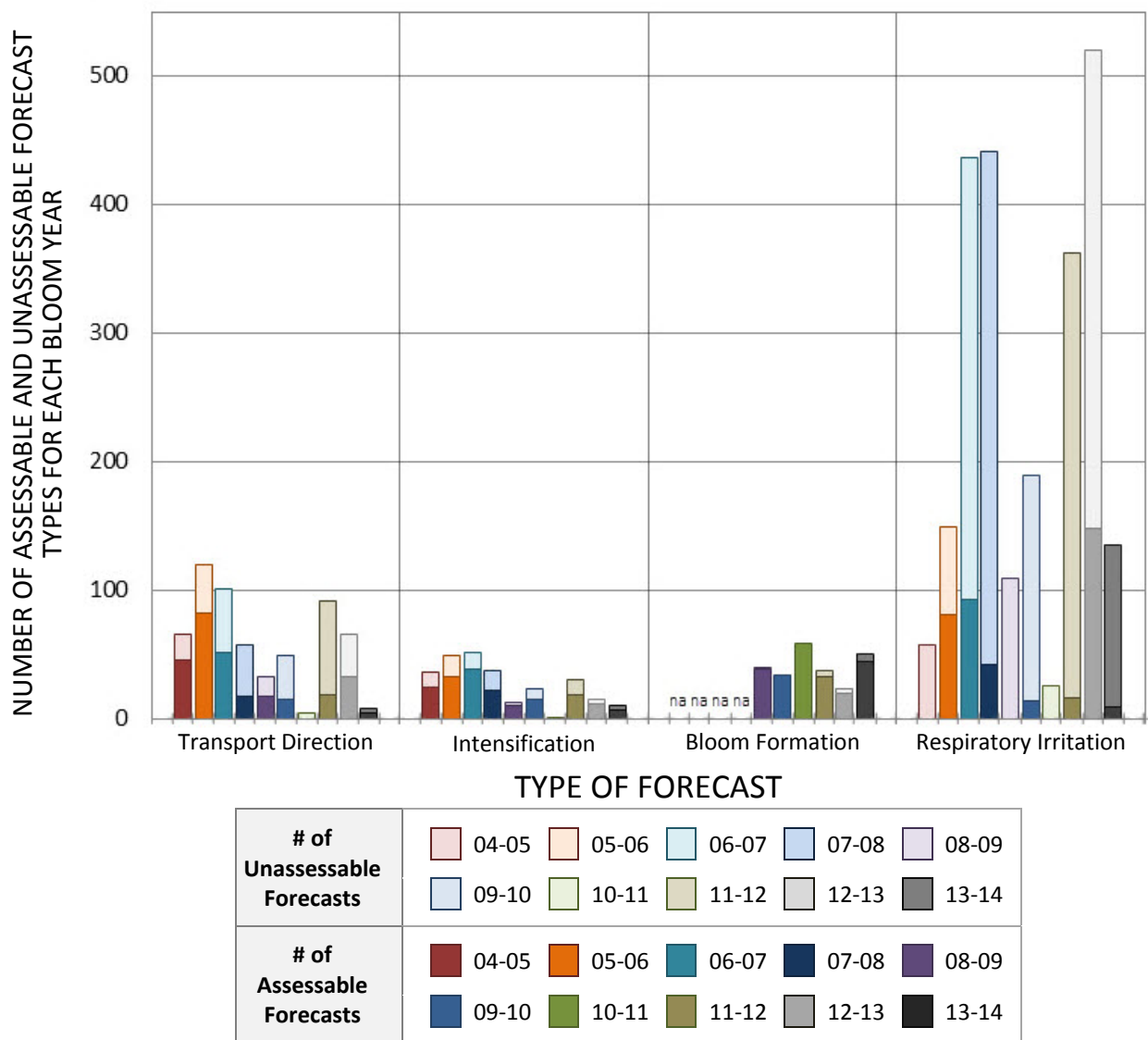


Figure 10. Number of assessable and unassessable forecasts for each bloom year from 2004 to 2014. The assessment of forecasts was dependent on the availability of reliable observational data from reputable sources. The label “na” indicates a value was not applicable due to the lack of forecasts issued in the specified bloom year.

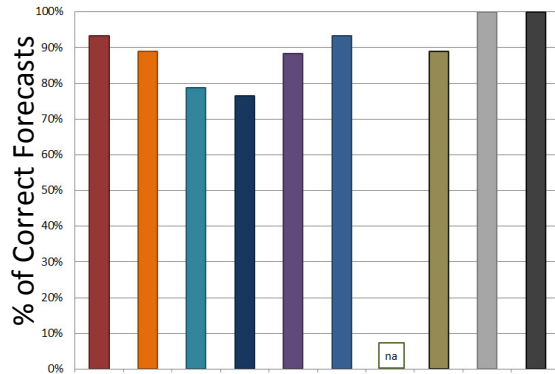
3.4 Accuracy of Forecasts

Forecast accuracy was estimated for each of the forecasts: transport direction, intensification, potential for bloom formation at the coast, and all levels of respiratory irritation. Accuracy was also estimated for the individual respiratory irritation levels: “no impact,” “very low,” “low,”

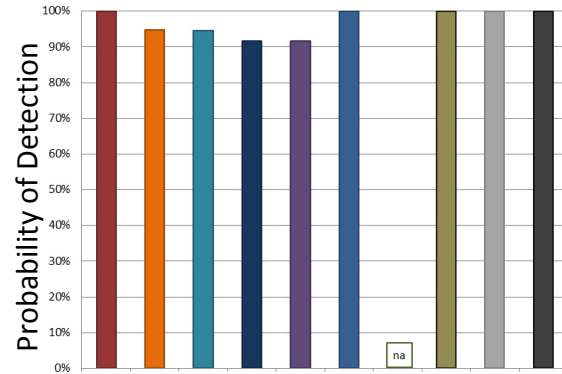
“medium,” and “high”. The four different statistics used to estimate forecast accuracy were proportion correct, probability of detection, threat score, and false alarm ratio (see APPENDIX II.I.B for definitions).

3.4.1 Transport Direction

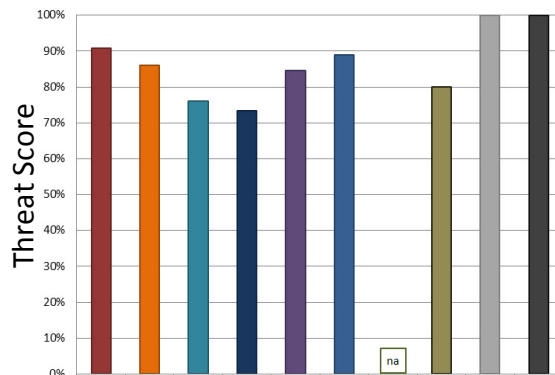
Figure 11 shows that of the assessable forecasts issued during BY2008-2014 transport direction forecasts were consistently accurate, performing better than forecasts issued during BY2004-2008, with a high proportion correct during each bloom year (88.2-100%), high probability of detection (0.917-1.00), high threat score (0.800-1.00), and low false alarm ratio (0.00-0.200). Though the transport direction forecasts issued during BY2013-2014 had perfect scores for each measure of accuracy, the sample size was small with only 4 assessable forecasts during the bloom year. However, the 33 assessable forecasts issued during BY2012-2013 also had perfect scores for accuracy. The accuracy of the transport direction forecasts could not be calculated for BY2010-2011 because poor imagery and limited sampling prevented assessment of the four forecasts issued for the bloom offshore the Florida Keys during the end of BY2009-2010 and the beginning of BY2010-2011.



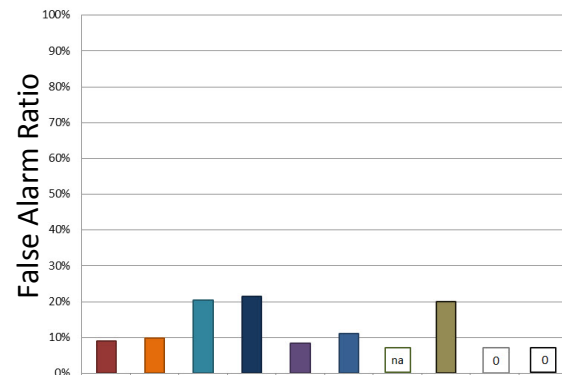
a) Proportion Correct: the number of correct transport direction forecasts and correct rejections compared to the total number of transport direction forecasts issued.



b) Probability of detection: the proportion of observed transport direction events that were correctly forecast.



c) Threat score: measure of forecasted transport direction events, after removing the correct rejections.



d) False alarm ratio: the number of false alarms compared to the total number of transport direction forecasts issued.











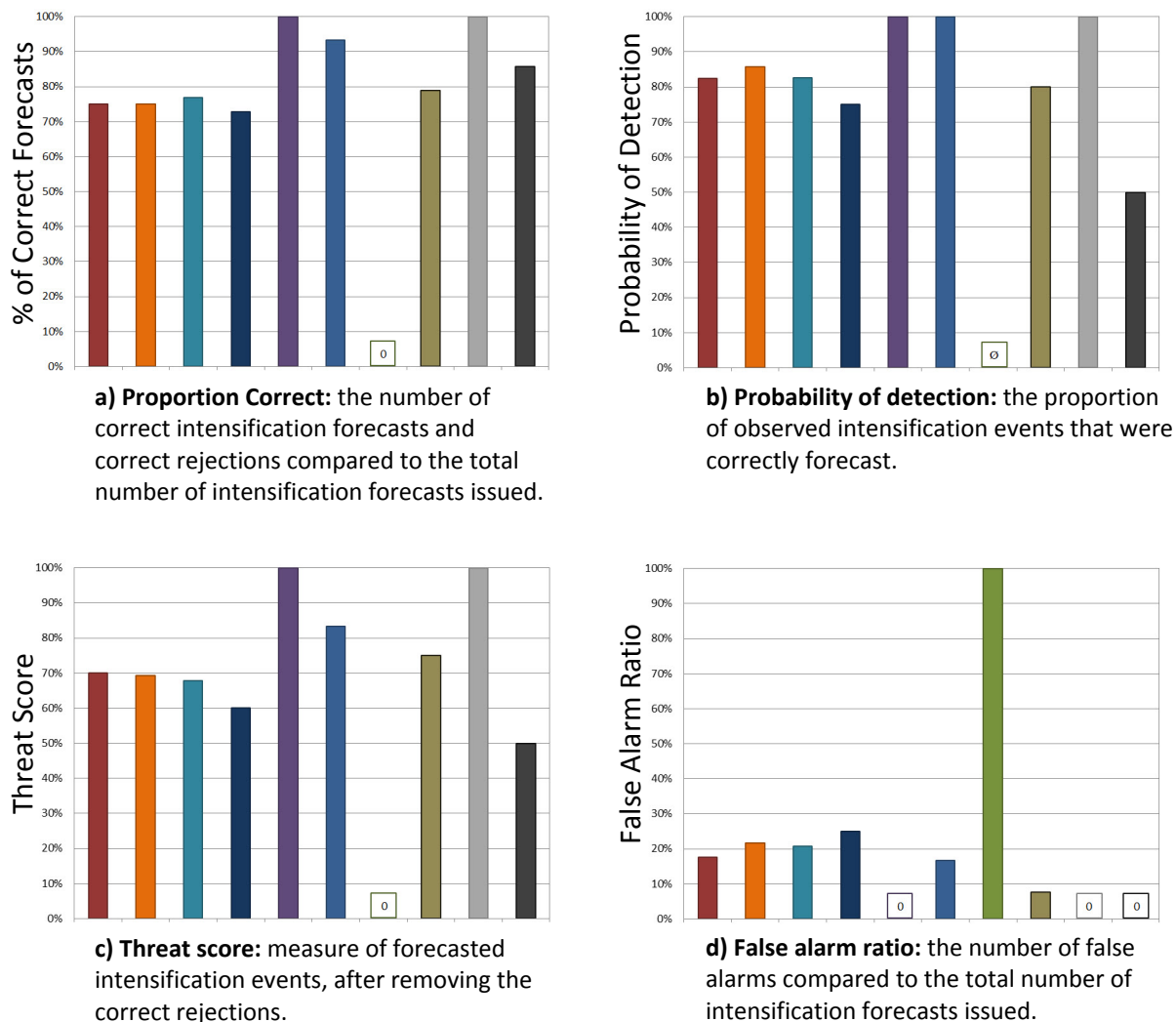
Bloom Years	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	13-14
Colors										
n	45	82	52	17	17	15	0	18	33	4

Figure 11. Accuracy of transport direction forecasts issued during the 2004 to 2014 bloom years. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%.

3.4.2 Intensification

Figure 12 shows that with the exceptions of BY2010-2011 and BY2013-2014, the assessable intensification forecasts issued during BY2008-2014 were consistently more accurate than those issued during BY2004-2008, with a high proportion correct during each bloom year (79.0-100%), high probability of detection (0.800-1.00), a high threat score (0.750-1.00), and a low false alarm ratio (0.00-0.167). During BY2013-2014, only seven assessable intensification forecasts were issued with one observation of intensification when no intensification event was forecasted. Thus, the proportion correct was high (85.7%) and the false alarm ratio was low

(0.00), but the probability of detection and threat score were also low (0.500 each). In BY2010-2011, no new bloom developed, leading to only one assessable forecast for intensification, which was determined to be false. This resulted in a proportion correct of 0.00%, a false alarm ratio of 1.00, and a threat score of 0.00.













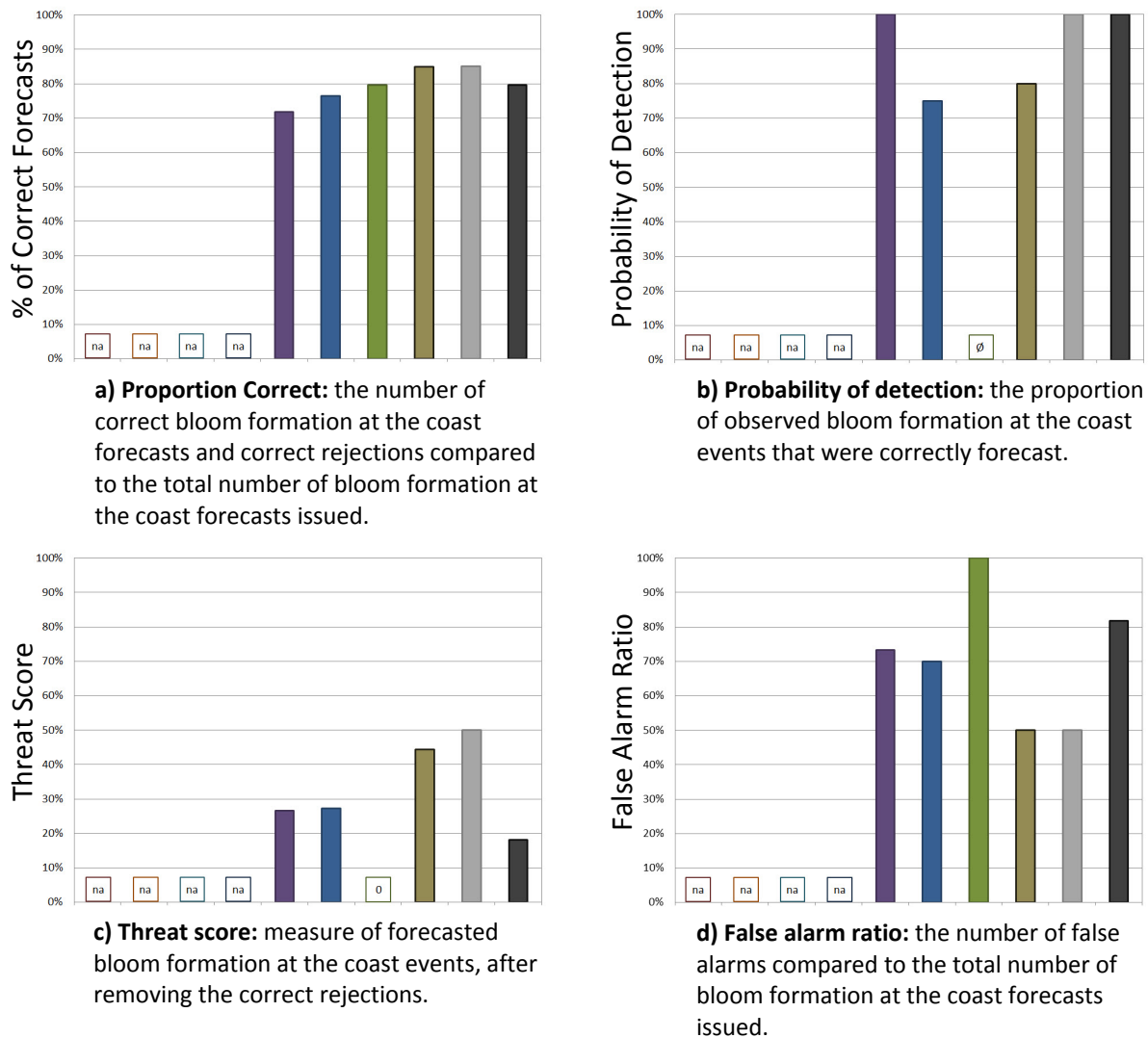
Bloom Years	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	13-14
Colors										
n	24	32	39	22	10	15	1	19	11	7

Figure 12. Accuracy of intensification forecasts issued during the 2004 to 2014 bloom years. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation.

3.4.3 Bloom Formation at the Coast

Forecasts for bloom formation at the coast were not consistently issued until BY2008-2009. During BY2008-2014 forecast accuracy was relatively consistent (see

Figure 13). During BY2008-2014, bloom formation forecasts had a relatively high proportion correct (71.8-85.0%) and probability of detection (0.750-1.00). However, there was also a low threat score (0.182-0.500) and high false alarm ratio (0.500-0.818). In BY2010-2011, no new bloom developed leading to an undefined number when calculating probability of detection.













Bloom Years	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	13-14
Colors										
n	N/A	N/A	N/A	N/A	39	34	59	33	20	44

Figure 13. Accuracy of forecasts of bloom formation at the coast during the 2004 to 2014 bloom years. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation.

3.4.4 Respiratory Irritation

3.4.4.1 All Levels of Respiratory Irritation

While a consistently high proportion correct (81.7-100%) was calculated for respiratory irritation forecasts issued during BY2004-2008, the results for BY2008-2014 varied. Respiratory irritation forecasts issued during BY2011-2013 had the highest proportion correct.

3.4.4.2 No Level of Respiratory Irritation (None)

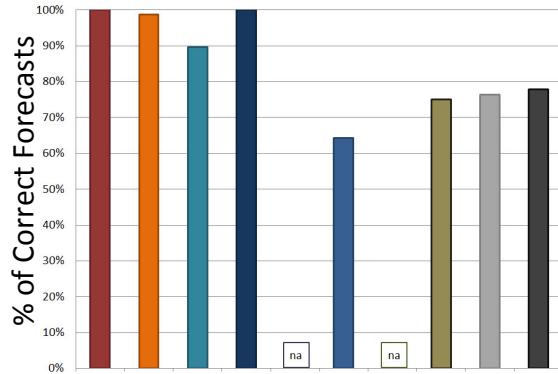
In BY2009-2010 and BY2011-2014 there were forecasts of “none” that were assessed as “false” (i.e. another level of respiratory irritation was observed). This resulted in an undefined probability of detection, a low threat score (0.00), and a high false alarm ratio (1.00). The proportion of correct respiratory irritation forecasts issued during those bloom years varied (64.3-93.8%) based on the occurrence of correct rejections.

3.4.4.3 Very Low Levels of Respiratory Irritation

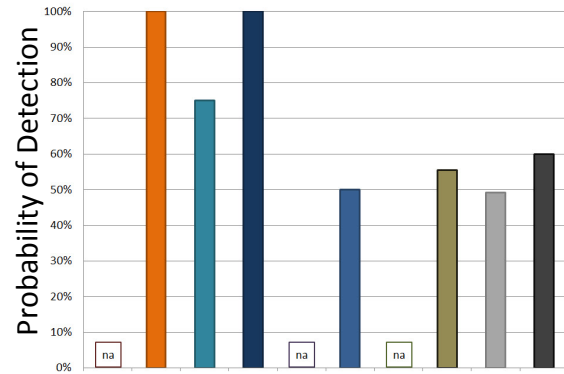
During BY2013-2014, “very low” respiratory irritation forecasts were only assessable twice but were confirmed as being correct both times, yielding perfect scores for proportion correct, probability of detection, threat score, and false alarm ratio. For BY2009-2010, and BY2011-2013, the proportion correct was high (92.9-93.8%), but the threat score was 0.00 and the false alarm ratio was 1.00. Probability of detection could not be calculated for these bloom years.

3.4.4.4 Low Levels of Respiratory Irritation

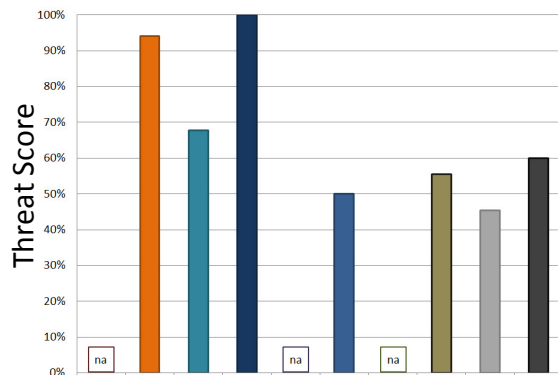
As shown in Figure 14, of the “low” respiratory irritation forecasts issued during BY2009-2010 and BY2011-2014, there was a relatively high proportion correct (64.3-77.8%) with the lowest value for the false alarm ratio (0.00). However, the values for probability of detection ranged between 0.492-0.600 and the threat score values ranged between 0.453-0.600 because in each of those years some observations of “low” respiratory irritation were forecast as other levels of respiratory irritation.



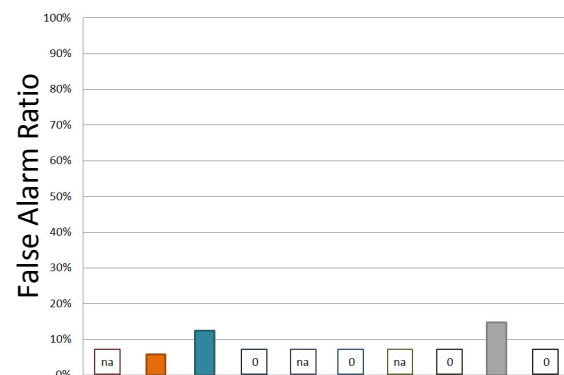
a) Proportion Correct: the number of correct “low” respiratory irritation forecasts and correct rejections compared to the total number of “low” respiratory irritation forecasts issued.



b) Probability of detection: the proportion of observed “low” respiratory irritation events that were correctly forecast.



c) Threat score: measure of forecasted “low” respiratory irritation events, after removing the correct rejections.



d) False alarm ratio: the number of false alarms compared to the total number of “low” respiratory irritation forecasts issued.











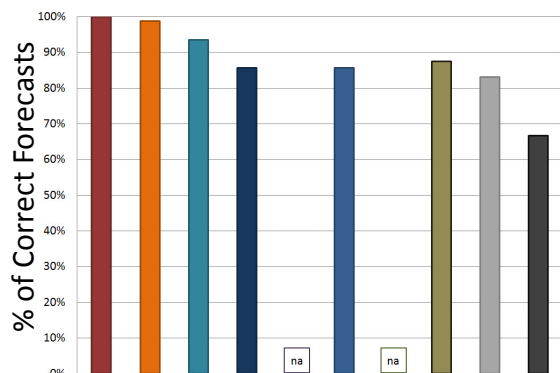
Bloom Years	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	13-14
Colors										
n	0	17	24	9	0	5	0	5	34	3

Figure 14. Accuracy of “low” respiratory irritation forecasts during the 2004 to 2014 bloom years. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation.

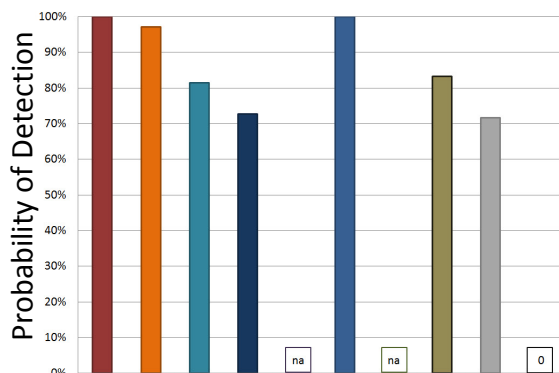
3.4.4.5 Moderate Levels of Respiratory Irritation

Figure 15 shows that “moderate” respiratory irritation forecasts issued during BY2011-2012 and BY2012-2013 had a high proportion correct (87.5% and 83.1%, respectively) and probability of detection (0.833 and 0.717, respectively), relatively high threat scores (0.714 and 0.603, respectively), and low false alarm ratio (0.167 and 0.208, respectively). During BY2009-2010, only one out of three “moderate” respiratory irritation forecasts matched the observations; in the other two instances “low” respiratory irritation was observed instead. Thus, although the proportion correct (85.7%) and probability of detection (1.00) were high, the threat score was

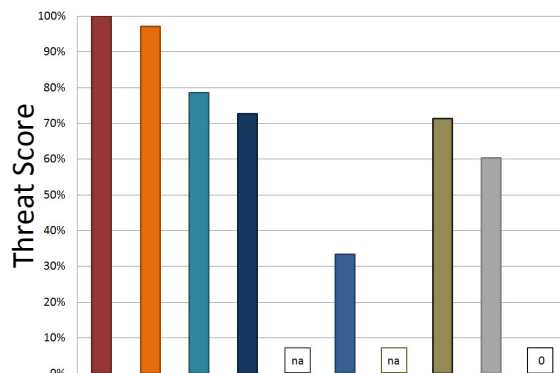
low (0.333) and the false alarm ratio was high (0.667). During BY2013-2014, “moderate” respiratory irritation was observed twice, but another forecast level was predicted, resulting in a low proportion correct (66.7%), probability of detection (0.00), threat score (0.00), and a high false alarm ratio (1.00).



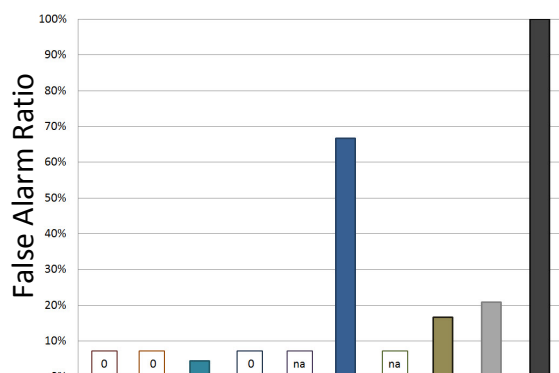
a) Proportion Correct: the number of correct “moderate” respiratory irritation forecasts and correct rejections compared to the total number of “moderate” respiratory irritation forecasts issued.



b) Probability of detection: the proportion of observed “moderate” respiratory irritation events that were correctly



c) Threat score: measure of forecasted “moderate” respiratory irritation events, after removing the correct rejections.



d) False alarm ratio: the number of false alarms compared to the total number of “moderate” respiratory irritation forecasts issued.











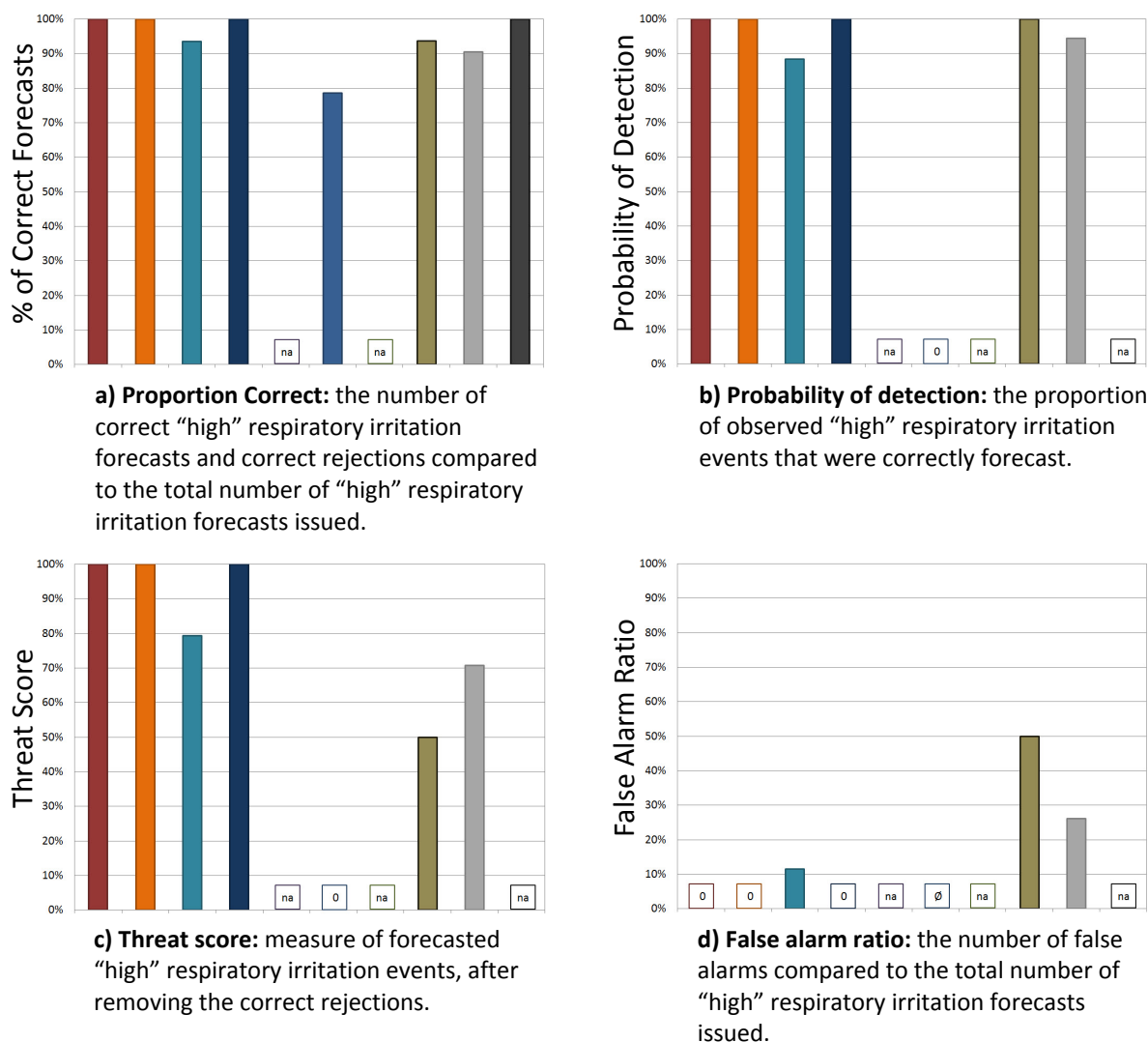
Bloom Years	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	13-14
Colors										
n	3	35	23	16	0	3	0	6	48	1

Figure 15. Accuracy of “moderate” respiratory irritation forecasts during the 2004 to 2014 bloom years. The label “na” indicates a value was not applicable due to no forecast issued in the specified bloom year. The label “0” indicates that the value was 0%.

3.4.4.6 High Levels of Respiratory Irritation

As shown in Figure 16, “high” respiratory irritation forecasts issued during BY2011-2012 and BY2012-2013 had a high proportion correct (93.4-100%) and probability of detection (94.4-

1.00). However, “high” respiratory irritation forecasts matched observations more frequently during BY2012-2013 compared to BY2011-2012, resulting in a higher threat score (0.708 vs. 0.500) and lower false alarm ratio (0.261 vs. 0.500). During BY2009-2010, although no “high” respiratory irritation forecasts were issued, “high” respiratory irritation was observed in three instances resulting in a probability of detection and threat score of 0.00.



Bloom Years	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	13-14
Colors										
n	11	26	26	11	0	0	0	2	46	0

Figure 16. Accuracy of “high” respiratory irritation forecasts during the 2004 to 2014 bloom years. The label “na” indicates a value was not applicable due to no forecast issued in the specified bloom year. The label “0” indicates that the value was 0%. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation.

3.5 Reliability of Forecasts

Forecast reliability was estimated by calculating the bias, a statistic that indicates whether the forecast system consistently over-forecasted or under-forecasted events. Over-forecasting means that an event was forecast more often than it was observed, while under-forecasting means that an event was observed more often than it was forecast. Bias was calculated for each of the forecasts: transport direction, intensification, and bloom formation at the coast (see Figure 17). It was also calculated for the individual respiratory irritation levels ranging from “no” respiratory irritation to “high” respiratory irritation (see Figure 18).

3.5.1 *Transport Direction*

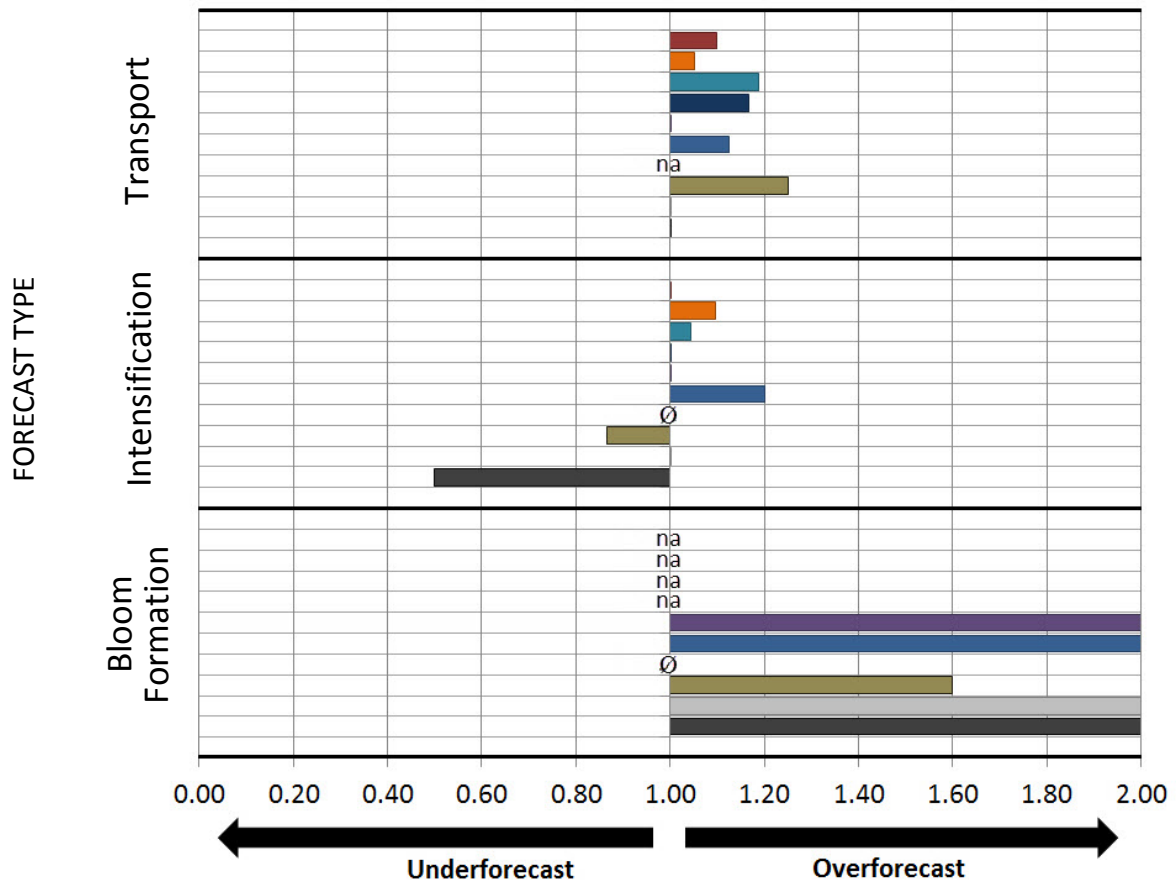
Transport direction forecasts yielded varied bias results during BY2008-2014. As shown in Figure 17, there was no bias (1.00) for transport direction forecasts issued during BY2008-2009, BY2012-2013, and BY2013-2014. Transport direction was slightly over-forecast in BY2009-2010 and BY2011-2012 (1.13 and 1.25, respectively).

3.5.2 *Intensification*

As shown in Figure 17, bias varied in the intensification forecasts issued during BY2008-2014. Intensification was slightly over-forecast in BY2009-2010 (1.20) and under-forecast in BY2011-2012 (0.867) and BY2013-2014 (0.500). However, there was no bias (1.00) in the forecasts issued during BY2008-2009 and BY2012-2013.

3.5.3 *Bloom Formation at the Coast*

As shown in Figure 17, results of the BY2008-2014 assessment period revealed that bloom formation at the coast tended to be over-forecast (1.60-5.50). Bias was undefined for BY2010-2011 because bloom formation at the coast was forecast but not observed, leaving a zero in the denominator.



Number of Assessable Forecasts (n) for each Bloom Year				
Bloom Year		Transport Direction	Intensification	Bloom Formation
	04-05	45	24	N/A
	05-06	82	32	N/A
	06-07	52	39	N/A
	07-08	17	22	N/A
	08-09	17	10	39
	09-10	15	15	34
	10-11	0	1	59
	11-12	18	19	33
	12-13	33	11	20
	13-14	4	7	44

Figure 17. Forecast reliability (bias) in transport, intensification, and bloom formation at the coast forecasts during the 2004 to 2014 bloom years (BY). A score of one indicates no bias, a score greater than one indicates that the forecast system over-forecasted the event, and a score less than one indicates that the forecast system under-forecasted the event. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation.

3.5.4 Respiratory Irritation

3.5.4.1 No Levels of Respiratory Irritation (None)

Forecast bias could not be estimated for the “no” respiratory irritation level because the method of assessment could not verify that no respiratory irritation was observed throughout the entire forecast area during the forecast period (see Section 3.3).

3.5.4.2 Very Low Levels of Respiratory Irritation

As shown in Figure 18, there was no bias (1.00) in the “very low” respiratory irritation forecasts issued during BY2013-2014. Bias could not be calculated for BY2009-2010, BY2011-2012, or BY2012-2013, during which “very low” respiratory irritation was forecast, but not observed, leaving a zero in the denominator.

3.5.4.3 Low Levels of Respiratory Irritation

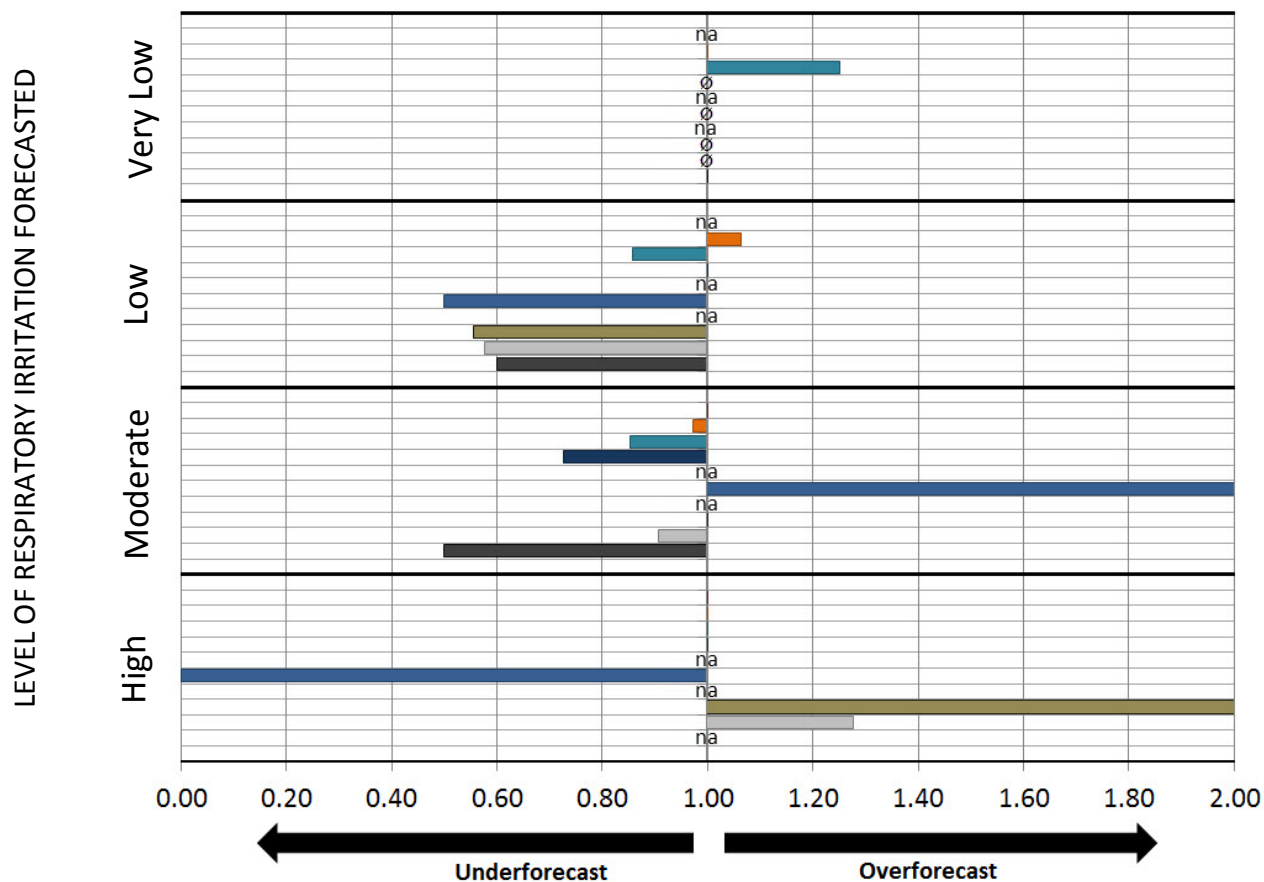
As shown in Figure 18, during BY2008-2014 “low” respiratory irritation was consistently under-forecast (0.500-0.600), meaning that it was observed more often than it was forecast. It was the only respiratory irritation level to be consistently under-forecast.

3.5.4.4 Moderate Levels of Respiratory Irritation

As shown in Figure 18, results varied for “moderate” respiratory irritation forecasts issued during BY2008-2014. During BY2009-2010, “moderate” respiratory irritation was over-forecast, with a bias of 3.00. There was no bias (1.00) for the “moderate” respiratory irritation forecasts issued in BY2011-2012. During BY2012-2014, respiratory irritation was under-forecast (0.906 and 0.500, respectively).

3.5.4.5 High Levels of Respiratory Irritation

While each of the first four years of operation from BY2004-2008 yielded unbiased results (1.00) for “high” respiratory irritation forecasts, results from BY2008-2014 varied (see Figure 18). During BY2011-2013, “high” respiratory irritation was over-forecast (2.00 and 1.28, respectively). “High” respiratory irritation was under-forecast (0.00) in BY2009-2010 when no assessable “high” respiratory irritation forecasts were issued, but “high” respiratory irritation was observed three times.



Number of Assessable Forecasts (n) for each Bloom Year					
Bloom Year		Very Low	Low	Moderate	High
	04-05	0	0	3	11
	05-06	3	17	35	26
	06-07	15	24	23	26
	07-08	3	9	16	11
	08-09	0	0	0	0
	09-10	1	5	3	0
	10-11	0	0	0	0
	11-12	1	5	6	2
	12-13	10	34	48	46
	13-14	2	3	1	0

Figure 18. Forecast reliability (bias) in respiratory impact forecasts during the 2004 to 2014 bloom years (BY). A score of one indicates no bias, while a score of greater than one indicates that the forecast system over-forecasted the event. A score of less than one suggests that the forecast system under-forecasted the event. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation.

3.6 Skill of Forecasts

Forecast skill was estimated by calculating the Heidke skill score, a statistic that represents accuracy relative to chance. It compares the proportion of correct forecasts with an estimate of the correct forecasts that could be due solely to random chance. A score of zero indicates that the forecast is no better than random chance at predicting the event (i.e. no forecast skill), a negative score indicates that the forecast performs worse than chance, and a perfect score is one or 100%. The Heidke skill score was calculated for transport direction, intensification, bloom formation at the coast, and overall respiratory irritation forecasts (see Figure 19). It was also calculated for each level of respiratory irritation forecasted, ranging from “no” respiratory irritation to “high” respiratory irritation (see Figure 20).

3.6.1 Transport Direction

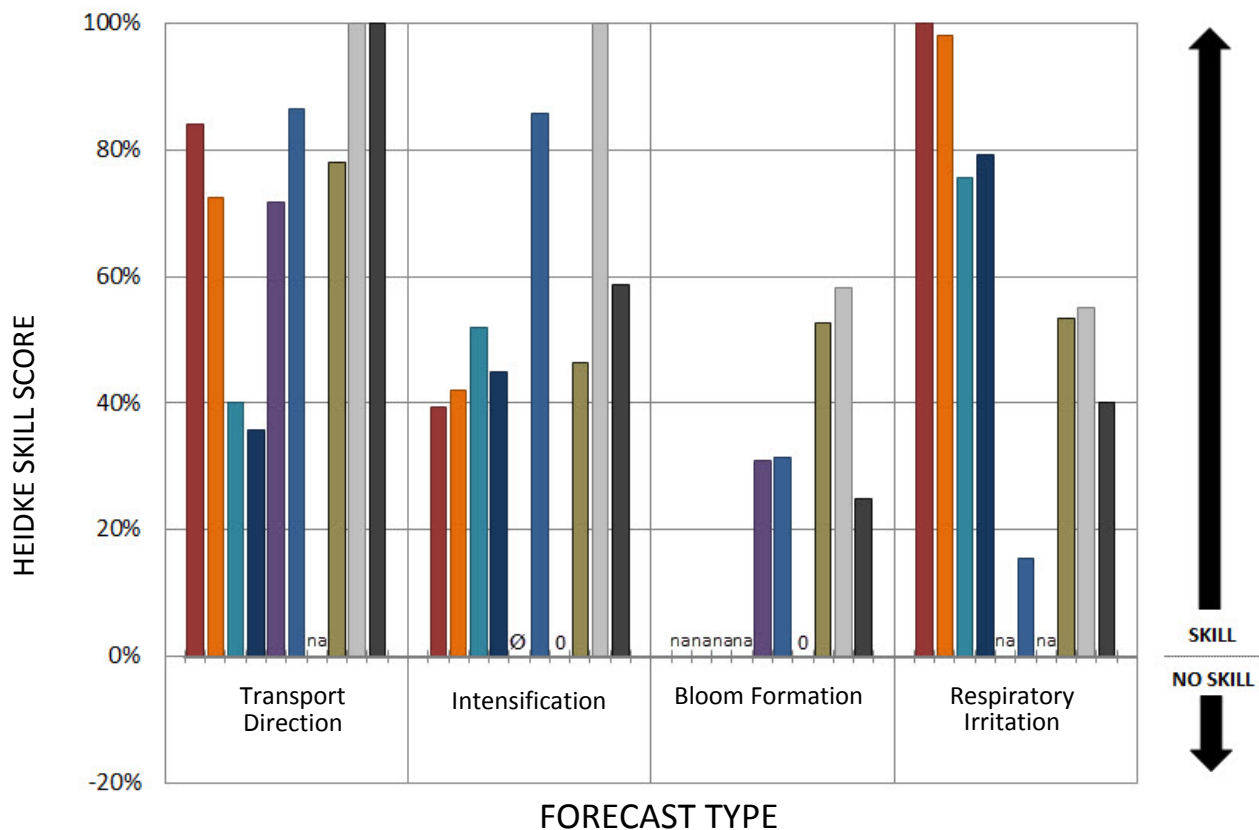
Figure 19 shows that all transport direction forecasts issued during BY2008-2014 performed much better than chance, with Heidke skill scores ranging from 71.7-100% improvement over chance. These results were higher than those for BY2004-2008, in which Heidke skill scores still performed better than chance, but ranged from 35.9-84.2%.

3.6.2 Intensification

Intensification forecasts issued during BY2009-2014 consistently performed better than chance in all but one year (BY2010-2011), with Heidke skill scores ranging from 46.5%-100% improvement over chance (see Figure 19). This showed a higher skill overall than from BY2004-2008 when scores ranged from 39.5-51.9% improvement over chance. During BY2010-2011, one intensification forecast was issued for the dissipating bloom that originated during BY2009-2010, but it was assessed as false resulting in a Heidke skill score of 0.00%. A Heidke skill score could not be calculated for BY2008-2009 because, although intensification events were forecast and observed, there were no correct rejections, leaving a zero in the denominator of the equation.

3.6.3 Bloom Formation at the Coast

The majority of bloom formation at the coast forecasts issued during BY2008-2014 performed better than chance, with Heidke skill scores ranging from 25.0-58.3% improvement over chance (see Figure 19). For BY2010-2011 the Heidke skill score was 0.00% because bloom formation at the coast was forecast to occur 12 times when conditions were favorable although ultimately no bloom formed during that bloom year.



Number of Assessable Forecasts (n) for each Bloom Year					
		Transport Direction	Intensification	Bloom Formation	All Respiratory Irritation
Bloom Year	04-05	45	24	0	14
	05-06	82	32	0	81
	06-07	52	39	0	93
	07-08	17	22	0	42
	08-09	17	10	39	0
	09-10	15	15	34	14
	10-11	0	1	59	0
	11-12	18	19	33	16
	12-13	33	11	20	148
	13-14	4	7	44	9

Figure 19. The forecast skill of transport direction, intensification, bloom formation at the coast and all respiratory irritation forecasts during the 2004 to 2014 bloom years (BY). The Heidke skill score is a skill corrected verification measure of categorical forecast performance that references the proportion of correct forecasts relative to the number of correct forecasts that could be made by random chance. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation.

3.6.4 Respiratory Irritation

3.6.4.1 All Respiratory Irritation Levels

When analyzed together the overall assessable respiratory irritation forecasts issued during BY2008-2014 performed better than chance (15.4-55.0%), though the Heidke skill scores were lower than for BY2004-2008 (75.7-100%).

3.6.4.2 No Levels of Respiratory Irritation (None)

Respiratory irritation forecasts of “none” issued during BY2009-2010 and BY2011-2014 showed no improvement over chance (0.00%) in each case because less than 11 respiratory irritation forecasts of “none” were assessed as false based on other respiratory irritation levels observed within the forecast region.

3.6.4.3 Very Low Levels of Respiratory Irritation

As shown in Figure 20, “very low” respiratory irritation forecasts issued during BY2009-2010 and BY2011-2013 performed no better than chance, with Heidke skill scores of 0.00%. BY2009-2010 and BY2011-2012 included low sample sizes of only n=4 and n=1, respectively. “Very low” respiratory irritation forecasts issued during BY2013-2014 performed much better than chance, with a Heidke skill score of 100%, though this was also based on a small sample size (n=2).

3.6.4.4 Low Levels of Respiratory Irritation

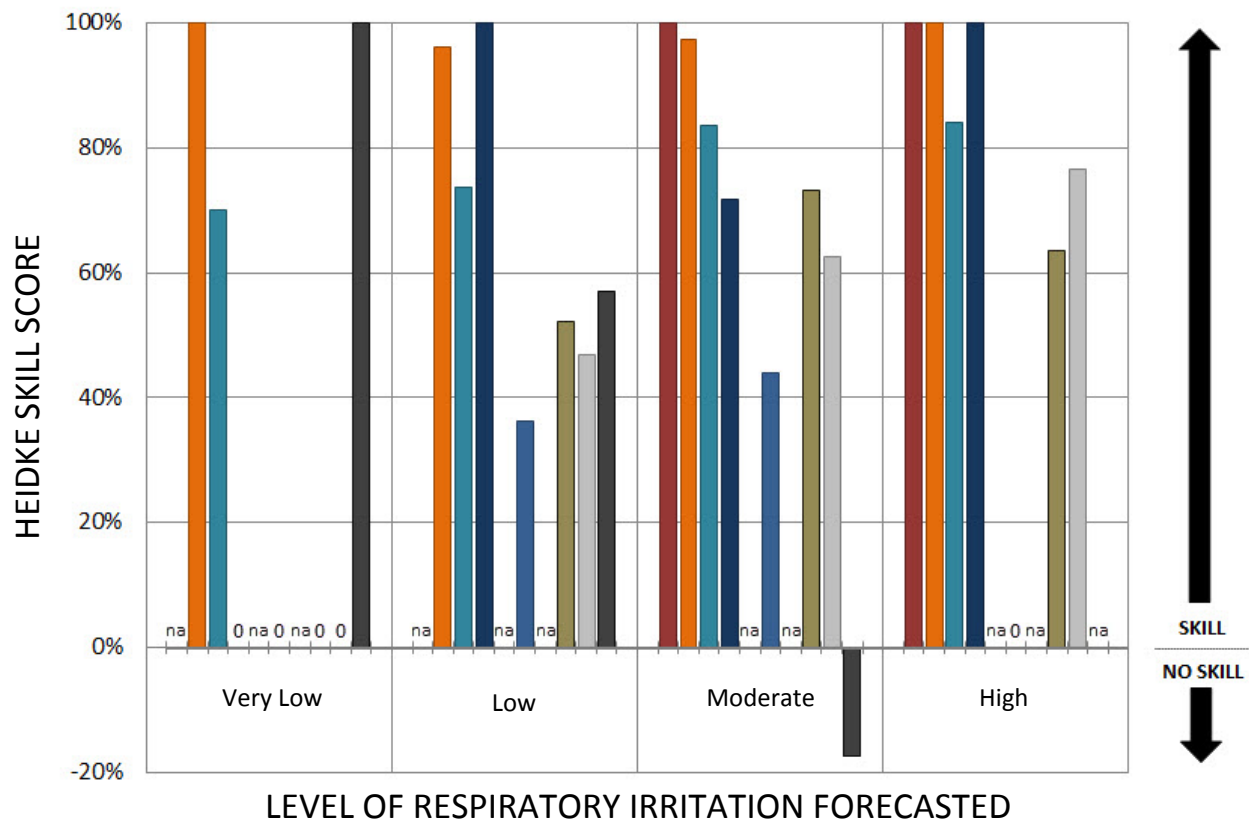
As shown in Figure 20, “low” respiratory irritation forecasts performed much better than chance (36.4-57.1%) in all bloom years that could be assessed from BY2009-2014.

3.6.4.5 Moderate Levels of Respiratory Irritation

Figure 20 shows that “moderate” respiratory irritation forecasts performed much better than chance during BY2009-2010 and BY2011-2013 (44.0-73.3%). By contrast, “moderate” respiratory irritation forecasts issued during BY2013-2014 performed worse than chance (-17.4%) because “moderate” respiratory irritation was forecast only once and “low” respiratory irritation was observed instead. Additionally, in both cases that “moderate” respiratory irritation was observed, the forecast was for “no” respiratory irritation level based on the available water sample data.

3.6.4.6 High Levels of Respiratory Irritation

As shown in Figure 20, Heidke skill scores for “high” respiratory irritation forecasts issued during BY2008-2014 were more variable than those for BY2004-2008. “High” respiratory irritation forecasts issued during BY2011-2012 and BY2012-2013 performed better than chance (63.6% and 76.5%, respectively). During BY2009-2010, “high” respiratory irritation was not forecast to occur, but it was observed three times when another level of respiratory irritation had been forecast based on available water sample data, resulting in a Heidke skill score indicating no improvement over chance (0.00%).



Number of Assessable Forecasts (n) for each Bloom Year					
		Very Low	Low	Moderate	High
Bloom Year	04-05	0	0	3	11
	05-06	3	17	35	26
	06-07	15	24	23	26
	07-08	3	9	16	11
	08-09	0	0	0	0
	09-10	1	5	3	0
	10-11	0	0	0	0
	11-12	1	5	6	2
	12-13	10	34	48	46
	13-14	2	3	1	0

Figure 20. The forecast skill of individual levels of respiratory irritation during the 2004 to 2014 bloom years (BY). The Heidke skill score is a skill corrected verification measure of categorical forecast performance that references the proportion of correct forecasts relative to the number of correct forecasts that could be made by random chance. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%. The label “Ø” indicates a value was undefined due to a zero in the denominator of the equation.

3.7 Comparison of New and Old Respiratory Irritation Forecast Model Performance

On August 10, 2009, a modification to the respiratory irritation forecast model was implemented in order to create a distinction between newly-developing blooms and established blooms. This was done by creating separate respiratory irritation forecast models for newly-developing blooms (<50,000 cells/L) versus already-developed blooms based on the underlying assumption that developing blooms at lower concentrations would be associated with lower levels of respiratory irritation than developed blooms. The new model improved the overall accuracy, reliability, and skill during BY2012-2014, but model performance remained the same when compared to the BY2004-2009 model during BY2009-2012 (see Figure 21). The new model improved the skill of the “very low” and “low” level respiratory irritation forecasts the most.

During BY2009-2014, the new model changed the performance of the “very low” respiratory irritation forecasts during BY2013-2014, increasing the proportion correct from 66.7% to 100%, probability of detection from 0.00 to 1.00, threat score from 0.00 to 1.00, and decreasing the false alarm ratio from 1.00 to 0.00. The reliability improved as well (0.500 to 1.00). The Heidke skill score improved from -17.4% to 100% improvement over chance.

The new model improved the performance of the “low” respiratory irritation forecasts in both BY2012-2013 and BY2013-2014. The greatest improvement in accuracy occurred during BY2013-2014, with an increase in the proportion correct from 55.6% to 77.8% and the threat score from 0.429 to 0.600. The false alarm ratio decreased from 0.400 to 0.00. The probability of detection remained the same (0.600). However, in both BY2012-2013 and BY2013-2014, the new model resulted in “low” respiratory irritation being increasingly under-forecast (0.729 vs. 0.576 and 1.00 vs. 0.600, respectively). The Heidke skill scores in both years demonstrate that the new model showed a greater improvement over chance (38.0% vs. 46.9% and 10.0% vs. 57.1%, respectively).

The performance of the “moderate” respiratory irritation forecast improved during BY2012-2013, but decreased during BY2013-2014. Accuracy improved during BY2012-2013 with an increase in the proportion correct from 78.4% to 83.1%, probability of detection from 56.6% to 71.7%, the threat score from 0.484 to 0.603, and a decrease in the false alarm ratio from 0.231 to 0.208. Reliability improved in both BY2012-2013 and BY2013-2014 (0.736 vs. 0.906 and 0.00 vs. 0.500), but “moderate” respiratory irritation remained under-forecast regardless of the model. The Heidke skill score increased during BY2012-2013 from 50.1% to 62.5%, but decreased during BY2013-2014 from 0.00% to -17.4%.

The performance metrics for the “high” respiratory irritation forecasts remained unchanged using the new model due to the assessment procedures. The new model called for forecasting “high” respiratory irritation instead of “moderate” respiratory irritation in certain scenarios, but during assessment, observations of “moderate” or “high” respiratory irritation both confirm “high” respiratory irritation forecasts.

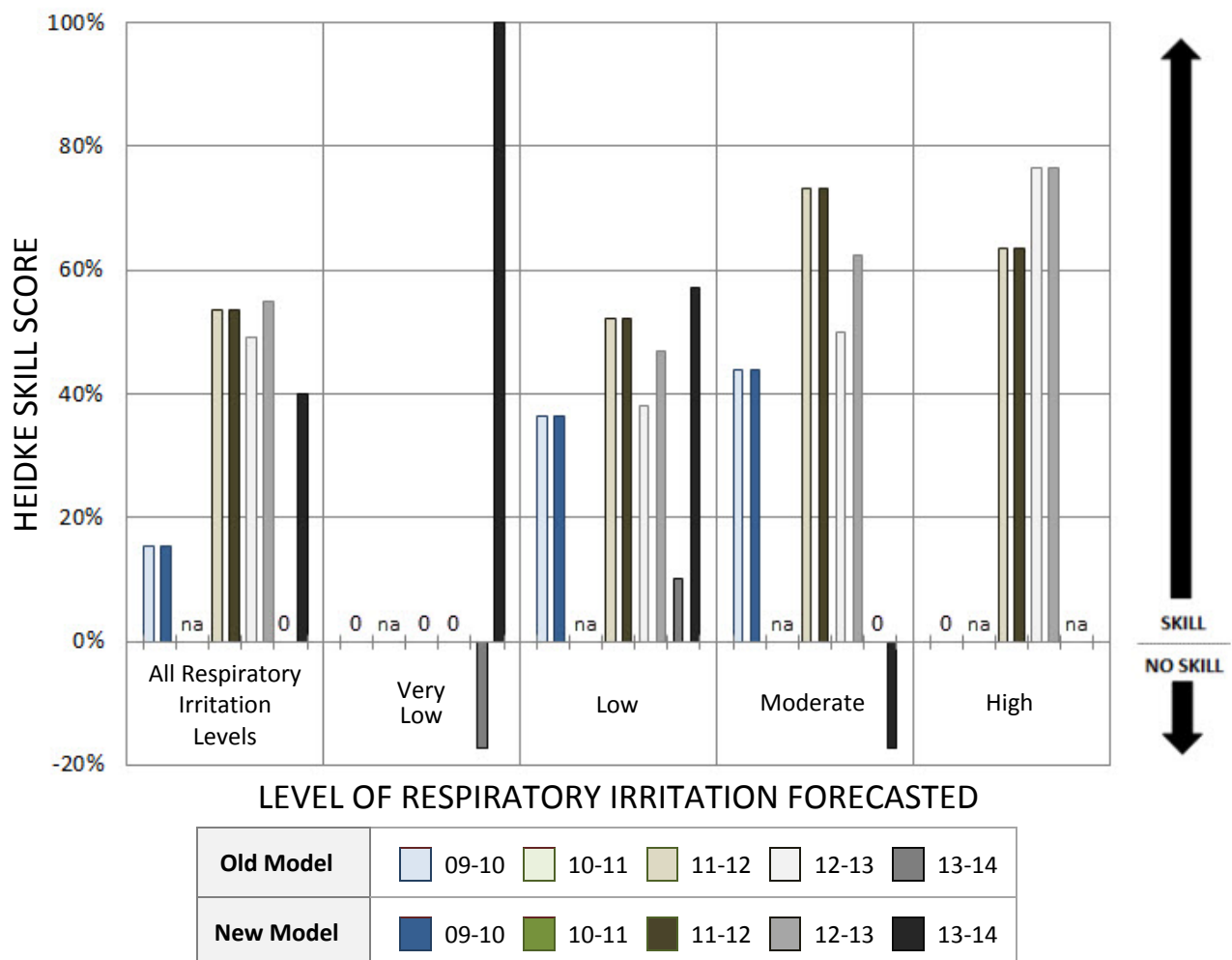


Figure 21. The forecast skill of the old and new respiratory irritation forecast models for individual levels of respiratory irritation during the 2009 to 2014 bloom years (BY). The Heidke skill score is a skill corrected verification measure of categorical forecast performance that references the proportion of correct forecasts relative to the number of correct forecasts that could be made by random chance. The label “na” indicates a value was not applicable due to the lack of assessable forecasts issued in the specified bloom year. The label “0” indicates that the value was 0%.

4. DISCUSSION

4.1 Bulletin Utilization

During BY2008-2012, confirmed product utilization ranged between 66.3-83.9%, similar to the percentage of confirmed utilization during BY2004-2008 (Kavanaugh, et al., 2013). Product utilization increased even more during BY2012-2013 and BY2013-2014, with 93.8% and 95.4% utilization, respectively. This notable increase was due in part to the launch of the HAB-OFS Facebook Page in September 2012.

There were two categories of usage that counted toward total product utilization: viewing the product and applying its content to bloom response. Overall, during BY2008-2012, utilization was most likely underestimated because it was only confirmed when there was evidence available that bulletin content was used by a reputable source such as a state or county agency, research institution, or public media entity. The HAB-OFS Facebook Page helped to more easily measure viewership of the products. Facebook records metrics that detail the ways in which people interact with each post: “likes,” “shares,” “comments,” and “post clicks”; the HAB-OFS was able to use those metrics to better quantify the number of people actively viewing the conditions report section of the Florida HAB bulletin during BY2012-2014 (see I.C.1). However, based on the assessment of Facebook metrics indicating active views of posted bulletin content during the evaluation period, evidence that the public was viewing the product was much easier to determine than confirmation that the product content was being applied to bloom response by Florida partners and stakeholders.

The HAB-OFS Facebook Page increased access to the HAB-OFS forecasts, but additional methods for evaluating utilization and usefulness should also be explored. For instance, tracking receipt of the HAB bulletins by subscribers would help directly assess product viewership. The implementation of routine surveys should also be considered in order to determine if and how the products are being applied, while regularly identifying needs for improvement.

4.2 Early Warning

Advance cautionary notice of the formation of a *K. brevis* bloom can help those involved in bloom response to plan necessary actions, such as closing shellfish beds before a bloom becomes a coastal hazard or minimizing mass marine animal casualties through early coordination of animal rescue, rehabilitation, and release efforts. In BY2008-2009, all three of the *K. brevis* features were detected first by the HAB-OFS using satellite imagery and then confirmed by *in situ* sampling collected by organizations in Florida (see APPENDIX IV). These results are comparable to the success demonstrated from BY2004-2008 when nine of the thirteen *K. brevis* events were identified first by the HAB-OFS (Kavanaugh, et al., 2013). In contrast, during BY2009-2014 all nine of the *K. brevis* events that occurred were first identified by *in situ* sampling.

The decline of the bloom detection capabilities of the HAB-OFS during BY2009-2010 coincides with the last two years of the SeaWiFS mission, which was terminated in February 2011. Outages impacted the availability and quality of data during the initiation of the blooms that occurred in both BY2009-2010 and BY2010-2011, which impacted the early warning capabilities of the HAB-OFS as well as the abilities to assess forecasts of transport and intensification. MODIS Aqua was used as an alternative to SeaWiFS, but blooms may have been less discernible in MODIS imagery because when chlorophyll values were estimated to be greater than 5µg/L, the values estimated by MODIS Aqua were typically higher than SeaWiFS (Stumpf, Tomlinson, & Fisher, MODIS and SeaWiFS Q&A, 2009). An internal comparison between MODIS Aqua and SeaWiFS chlorophyll products indicated that SeaWiFS was more accurate, although the chlorophyll anomalies that flagged the likely bloom location were similar for both products because the anomalies were derived from the average of multiple days of chlorophyll data so as not to be influenced by variation on one particular day (Stumpf, Tomlinson, & Fisher, MODIS and SeaWiFS Q&A, 2009).

Since the *K. brevis* features identified during BY2009-2014 were all detected first by collecting water samples with concentrations too low to detect with satellite imagery, it is also possible that the results are due to improvements in the sampling collection procedures made by organizations in Florida (see APPENDIX IV) during that time. Sampling procedures were improved in part through the Center for Prediction of Red Tides, formed in 2007 as a collaboration between the University of South Florida (USF) and FWRI. FWRI now uses satellite imagery products generated by USF for the Center for Prediction of Red Tides, which are based on different algorithms than the HAB-OFS (Hu, et al., 2011; Hu, Barnes, Murch, & Carlson, 2013). FWRI also uses the USF trajectory model for visualizing forecasted bloom movement, which has enabled FWRI to sample offshore more frequently and effectively, in addition to their routine sampling efforts alongshore (Zheng & Weisberg, 2012).

Recognizing a need to improve the bloom detection capabilities of the HAB-OFS, an ensemble satellite imagery product was developed that incorporates three algorithms: the chlorophyll anomaly, spectral shape at 490nm, and backscatter ratio (b_{bp}) (Derner & Kavanaugh, 2014). There was a reduction in false positives and over-prediction in the ensemble imagery that was evaluated, especially in the typical bloom initiation area from Pinellas County to Cape Romano in central Collier County. The ensemble satellite imagery product was transitioned to operations in September 2015 and will improve bloom detection during BY2015-2016. In addition, the HAB-OFS is currently evaluating the imagery from a new satellite sensor, the Variable Infrared Imaging Radiometer Suite (VIIRS) which has higher resolution imagery, and may improve *K. brevis* bloom detection throughout Florida.

4.3 Forecasts of Transport Direction

Similar to what was found during BY2004-2008 by Kavanaugh et al. (2013), transport direction forecasts were difficult to assess during BY2008-2014 because they required a series of unobscured satellite images where the entire bloom location was consistently distinguishable. Stumpf et al. (2009) determined that only large HABs, covering >10-30 km of coast could be reliably located and validated by imagery. The ability to assess transport direction forecasts may decrease during small, patchy blooms in Southwest Florida that render the forecasts more

difficult to confirm. Additionally, Southwest Florida blooms during BY2008-2014 were shorter in duration. Therefore, with the exception of BY2011-2012, less than half as many forecasts were issued during each year during BY2008-2014 than during BY2004-2007. Assessability may have also decreased due to the transition from SeaWiFS to MODIS Aqua. As noted in section 4.2 Early Warning, SeaWiFS products were found to more accurately detect blooms (Stumpf, Tomlinson, & Fisher, 2009).

During BY2008-2014, the assessable bloom transport direction forecasts performed with comparable accuracy, reliability, and skill to those issued during BY2004-2008. The forecasts were highly accurate and consistently performed better than chance at predicting bloom movement, with the highest performance during BY2012-2014. Although there was no modification to the model made during this time, new training procedures implemented during BY2012-2013 may have contributed to the improvements observed during both the widespread bloom in that year and the much patchier bloom during BY2013-2014.

Similar to transport direction forecasts issued during BY2004-2008, the forecasts were slightly over-forecast during BY2009-2012, but had perfect scores for reliability during BY2008-2009 and BY2012-2014. This means the model was occasionally biased towards predicting the direction of bloom movement more often than bloom transport was observed. The slight tendency towards over-forecasting may in part be due to limitations in the forecast model when attempting to predict the potential speed and distance of the movement (Kavanaugh, et al., 2013). Transport direction forecasts are made for 3-4 days into the future based on the forecasted direction and speed of the wind (Stumpf, Litaker, Lanerolle, & Tester, 2008). The forecast would be assessed as biased even if a bloom was moving south as predicted, but at a slower pace because the transport of a slow moving bloom might not be detectable within the duration of the forecast issued due to the limitations of the resolution of the HAB-OFS satellite imagery products.

The utility of providing only a transport direction forecast for Florida should be evaluated. The Gulf of Mexico Coastal Ocean Observing System Regional Association (GCOOS-RA) identified in the HAB Integrated Observing System (HABIOS) Plan a need to improve the HAB-OFS by using hydrodynamic models to produce transport forecasts that include direction, distance, and the associated uncertainties (Nowlin, Jochens, & Kirkpatrick, 2015). The Texas HAB-OFS uses a particle trajectory model to estimate bloom direction and distance, and a similar procedure should be explored for use in Florida that would complement the efforts of organizations in Florida. FWRI already uses the output of a trajectory model operated by the Center for Prediction of Red Tides to visualize forecasted bloom movement to plan sampling efforts and mitigation strategies. An additional long-term benefit could be that more accurately forecasting transport direction and including transport distance in Florida could eventually lead to the development of more accurate, higher resolution respiratory irritation forecasts.

4.4 Forecasts of Intensification

The model for forecasting bloom intensification is based on evidence that upwelling favorable conditions contribute to the hydrodynamic accumulation of developing blooms and nutrients at the coast (Stumpf, Litaker, Lanerolle, & Tester, 2008). There was a general decrease in the

number of intensification forecasts issued during BY2008-2014 ($n=1-32$) when compared to BY2004-2008 ($n=14-51$). Besides issuing fewer bulletins from BY2008-2014, which could have translated to fewer intensification forecasts being made due to a refinement of the intensification forecast model in November 2008, the forecasts were issued more selectively. The model change discontinued the issuance of intensification forecasts in regions where a bloom had reached “medium” to “high” concentrations. The intensification forecast model change contributed to the forecast performing better during BY2008-2014 than those issued during BY2004-2008.

Assessing the forecasts remained a challenge. With the exception of BY2010-2011, the percentage of assessable intensification forecasts during BY2008-2014 (63.3-83.3%) was similar to BY2004-2008 (59.5-76.5%). During BY2010-2011, the one intensification forecast issued was assessable. The switch from SeaWiFS to MODIS Aqua data did not seem to affect the assessment of intensification forecasts, as it did the assessment of transport direction. Although MODIS Aqua generally appeared to elevate chlorophyll values above those of SeaWiFS, intensification could still be evaluated based on the comparative difference between the chlorophyll values in the images as long as image products from the same sensor were used for each forecast assessment. In addition, during a developing bloom, organizations in Florida routinely collected samples from the same location every few days, enabling the HAB-OFS to verify intensification forecasts even if data from satellite imagery was unavailable. In the future, using higher resolution VIIRS imagery may help to improve assessability through enabling a finer scale comparison of the chlorophyll values in a bloom area. Increasing the frequency of sampling at the routine set of locations would also aid the assessment of intensification forecasts.

4.5 Forecasts of Potential for Bloom Formation at the Coast

Of the types of forecasts issued, forecasts of the potential for bloom formation at the coast had the highest percent assessable (87.0-100%). Due to routine sampling by organizations in Florida (see APPENDIX IV), the HAB-OFS could assess the binary forecast for bloom formation at the coast by determining whether or not *K. brevis* cells were present alongshore within the forecast timeframe.

The model for forecasting the potential for bloom formation at the coast was based on the seasonality of *K. brevis* blooms and the observation that upwelling favorable conditions seem to contribute to the hydrodynamic accumulation of developing blooms at the coast, similar to the basis of the intensification forecast model (Stumpf, Litaker, Lanerolle, & Tester, 2008). Since not enough is currently known about the conditions that cause *K. brevis* blooms to form offshore and at-depth, the HAB-OFS was not able to predict whether a bloom would form. Rather, the forecast predicted whether conditions were favorable to move a newly formed bloom to the coast if a bloom had indeed formed offshore. This meant that the forecast was prone to false alarms, since upwelling favorable conditions during the *K. brevis* season would only move a bloom to the coast if the forecast was preceded by the formation of an offshore bloom in the first place.

During BY2008-2014, coastal bloom formation forecasts had a high proportion correct (71.8-85.0%) and a high probability of detection (0.750-1.00). The forecasts also consistently performed better than chance (25.0-58.3%), with the exception of during BY2010-2011, which did not experience a new bloom. However, as expected with a forecast model prone to false

alarms, the forecasts had a low threat score (0.182-0.500) and high false alarm ratio (0.500-1.00), and were over-forecast (1.60-5.50). Despite its predisposition to false alarms, the coastal bloom formation forecast performed well at identifying when a bloom was unlikely to form at the coast. This information is useful because a model to predict bloom formation has not yet been demonstrated. The GCOOS HABIOS Plan identified the need for a true bloom formation forecast, combined with routine offshore sampling of conditions, and a heuristic biological model (2015). Once the environmental factors that contribute to *K. brevis* bloom development are better understood, it will be possible to model bloom formation and couple it with the bloom formation at the coast forecast model.

4.6 Forecasts of Respiratory Irritation

As was determined in the assessment of BY2004-2008, respiratory irritation forecasts continued to be the most difficult to assess during BY2008-2014, while also being the most frequently forecast (Kavanaugh, et al., 2013). However, the percentage of assessable forecasts decreased between BY2008-2014 when compared to BY2004-2008, with the exception of BY2012-2013. No forecasts were assessable in either BY2008-2009 or BY2010-2011 due to the low severity of the blooms during those years.

The decrease in assessability may be due in part to changes in the forecast regions. Originally, respiratory irritation forecasts were made for entire counties or even multiple counties. Half-county regions were introduced in 2006, but the boundaries were not standardized until 2010. This means that the number of forecasts issued increased because of procedural changes that increased the resolution, while the resolution of available observations, including those from MML, have remained similar since 2009.

The proportion correct and Heidke skill scores during BY2008-2014 also declined compared to forecasts issued during BY2004-2008. As the previous report indicated, the toxicity of each bloom varies. The HAB-OFS used cell concentrations as a proxy for how much brevetoxin aerosol might be present, but the forecast model performed better during large, intense blooms. In the future, direct measurements of the concentration of brevetoxin both in the water and the air should be explored. Quicker analysis of field data may also improve forecast accuracy. Since blooms are variable, the current 3-4 day lag time to collect and analyze samples may result in a forecast being developed based on the inaccurate assumption that the bloom remained in the same location and concentration level indicated by the samples. Regional differences in performance should also be evaluated because not all forecast regions may experience the surf conditions that promote the incorporation of brevetoxins into the marine aerosol (Pierce, et al., 2003). If so, it may be necessary to refine the respiratory irritation forecast model for the characteristics of each forecast region.

4.6.1 Changes to the Respiratory Irritation Forecast Model and Assessment

The respiratory irritation forecast model was modified in August 2009, resulting in an improvement in the overall accuracy, reliability, and skill during BY2012-2014 when compared to the old model and similar performance during BY2009-2012. The new model added a distinction between the level of respiratory irritation associated with a newly developing bloom (<50,000 cells/L) versus a bloom that had already developed with the assumption that developing

blooms would be associated with lower levels of respiratory irritation. “Very low” and “low” level forecasts showed the greatest improvements in performance when the new model was used in place of the old one, in part because “moderate” and “high” level forecasts were assessed based on observations of the general public being affected. Performance of the “moderate” level forecasts was also improved during the severe bloom in BY2012-2013, but not during BY2013-2014 because fewer “moderate” level forecasts were issued during the smaller-scale, lower intensity bloom.

4.6.2 Individual Respiratory Irritation Forecast Levels

4.6.2.1 No Respiratory Irritation Level

Forecasts of “no” respiratory irritation could not be adequately assessed in most cases due to under-reporting. Even with the daily monitoring conducted at select beaches as part of the MML Beach Conditions Reporting System, there are large areas of the coast that are not monitored and blooms often go undetected because their impacts vary throughout the day and are isolated to a few miles of coastline. Therefore, forecasts of “none” could not be assessed unless higher levels of respiratory irritation were observed as was the case during BY2009-2010, BY2011-2012, BY2012-2013, and BY2013-2014.

4.6.2.2 Very Low Levels of Respiratory Irritation

In both Florida and Texas, forecasts of “very low” levels of respiratory irritation could only be confirmed in the rare event that reports of observed respiratory irritation associated with the presence of a *K. brevis* bloom was limited to those suffering from chronic respiratory conditions (such as asthma). The observers of the MML Beach Conditions Reporting System could not report observations that correspond with the “very low” level so the HAB-OFS relied on reports from other reputable sources for assessment. The “very low” level forecasts could only be assessed due to higher levels of respiratory irritation observed in the forecast region, with the exception of BY2013-2014 when the assessable “very low” level forecasts were confirmed yielding perfect accuracy and a performance that was better than chance.

As with each forecast level, direct measures of brevetoxin in the air or water may improve the forecast performance. In addition, Kavanaugh et al. (2013) suggested that the “very low” and “low” forecast levels might be combined creating one level that would impact those with chronic respiratory conditions and otherwise healthy individuals who may be more sensitive to brevetoxin aerosols. This could increase the number of assessable forecasts because there would be more observations available. This option should be explored further.

4.6.2.3 Low and Moderate Levels of Respiratory Irritation

Similar to forecasts of “moderate” level respiratory irritation issued during BY2004-2008, those issued during BY2009-2014 overall performed strongly, with exceptions during BY2009-2010 and BY2011-2014 when there were several instances of “low” levels of respiratory irritation being observed instead. The timing corresponds with when the MML Beach Conditions Reporting System respiratory irritation observations began to be incorporated into the HAB-OFS assessment procedures, a main source of these observations. This therefore may indicate a mismatch between the observation and forecast levels used (see Table 11). MML’s “slight” category was defined as the observation of a “few coughs/sneezes heard” in a thirty second sample period (Mote Marine Laboratory, 2013). According to HAB-OFS procedures, “slight”

was thought to be equivalent to the “low” respiratory irritation forecast level used by the HAB-OFS which indicates that “only sensitive individuals and those with chronic respiratory conditions” would be affected (see Table 11). However, there was a decrease in the number of assessable “low” level forecasts during BY2008-2014 compared to previous years, and those that were assessable were under-forecast because in some cases “slight” respiratory irritation was observed and “moderate” respiratory irritation was forecast. A “moderate” respiratory irritation forecast indicates that “people at the beach may notice mild symptoms”, which, contrary to HAB-OFS procedures, may be more similar to MML’s definition of “slight” (see Table 11). Further investigation is needed to make sure the respiratory irritation forecasts are being appropriately assessed based on the representative observed respiratory conditions.

4.6.2.4 High Levels of Respiratory Irritation

During 2008-2014, forecasts for “high” levels of respiratory irritation were issued much less than during 2004-2008 due in part to interannual variation in the extent, intensity, and toxicity of the blooms. Assessable “high” level forecasts were issued most frequently during the widespread bloom that occurred during BY2012-2013. The majority of those forecasts were evaluated as correct. For 13 of the 47 forecasts, “slight” respiratory irritation was observed instead of “high” which may have been partially due to regional differences. Although the “high” respiratory irritation forecast performed better than chance and had strong scores for accuracy, it was over-forecast when compared to forecasts issued during BY2004-2008 which showed no bias.

Future improvements to the model should focus on reducing the occurrences of over-forecasting to ensure that public health officials continue to be informed when there is a risk of “high” respiratory irritation, while minimizing false alarms that might cause users to lose confidence in the forecast and therefore decide not to respond appropriately when there is a real risk to public health. This again demonstrates the need to continue to improve the forecast model to incorporate regional differences in surf conditions and observations of brevetoxin instead of using cell concentrations as a proxy.

The “high” respiratory irritation forecasts are especially vital for directly protecting public health because the general public is noticeably impacted at that level. The strong performance of the forecast has become even more significant because, since 2013, the NWS has issued the HAB-OFS forecast as part of a text product called the Beach Hazards Statement when “high” levels of respiratory irritation are forecasted by the HAB-OFS within the area of responsibility of the Weather Forecast Offices of Tampa Bay, Miami, or Key West. Beach Hazards Statements issued during the severe bloom in BY2012-2013 successfully increased public awareness of the bloom conditions and enhanced the public’s ability to mitigate the associated impacts.

5. CONCLUSION

From the time the HAB-OFS was transitioned to operations in the eastern Gulf of Mexico on October 1, 2004 to the end of the tenth bloom year on April 30, 2014, a total of 867 bulletins, 25 supplemental bulletins, and 38 conditions updates were issued. During BY2008-2012, confirmed product utilization ranged between 66.3-83.9%, a similar range to the estimate for products issued during BY2004-2008 (Kavanaugh, et al., 2013). The launch of the HAB-OFS Facebook Page in the fall of 2012 resulted in a notable increase in product utilization during BY2012-2013 and BY2013-2014 to 93.4% and 95.4%, respectively.

5.1 Modifications Resulting in Improvements

Several enhancements were implemented during the evaluation period from BY2008-2014. The respiratory irritation forecast model was modified, improving the forecast performance, especially for “very low” and “low” level forecasts. The intensification forecast model was improved, resulting in the forecasts being issued more selectively than before and contributing to general improvements in forecast performance during BY2008-2014. Although there were no changes to the transport direction forecast model during the BY2008-2014 evaluation period the transport direction forecasts issued during BY2012-2014 performed even better than those issued during BY2004-2012. The improvement in performance could have been due to the introduction of new training and review methods for analysts during that period.

During the BY2008-2014 evaluation period, a new forecast was successfully added. In October 2008, the potential for bloom formation at the coast forecast was transitioned to the forecast system. The forecast model was prone to false alarms, but consistently performed better than chance at predicting whether or not a bloom might form at the coast.

5.2 Improvement Needs

During this assessment period (BY2008-2014), the HAB-OFS was unable to consistently meet the needs identified by the GCOOS HABIOS Plan for early detection of HABs that would enable advanced preparation before the blooms reached coastal areas (Nowlin, Jochens, & Kirkpatrick, 2015). Most HABs were detected at the coast without being identified in HAB-OFS satellite imagery products because the currently-used satellite imagery derived from MODIS Aqua did not perform as well as the SeaWiFS sensor did before its mission was terminated in 2012. The first step toward improving satellite imagery was made in August 2015; a HAB ensemble satellite imagery product was transitioned to operations that will reduce false positives and over-prediction in future forecasts and enhance forecast assessment capabilities. The ensemble combines the currently-used chlorophyll anomaly with two additional detection algorithms specific to the optical characteristics of *K. brevis*: a backscatter ratio product (b_{bp}) and spectral shape at 490 nm (Derner & Kavanaugh, 2014).

The GCOOS HABIOS Plan also indicated that providing HAB location data at the current HAB-OFS resolution for the coast (10-50 km/day) is insufficient. In order to adequately support the

protection of public health, HAB location data is needed at a 1-4 km/day resolution, especially in critical areas of public health concern such as shellfish beds and popular recreation sites (Nowlin, Jochens, & Kirkpatrick, 2015). Operational ocean color products are expected to become available from VIIRS in the near future, and the higher resolution imagery may enhance HAB detection capabilities. In addition, the higher resolution satellite imagery combined with the ensemble approach may improve the ability to assess transport direction through enabling the HAB-OFS to detect bloom movement at finer scales.

Evaluations of respiratory irritation forecasts issued during BY2004-2014 demonstrate that the performance varies with the severity of the bloom. There may also be regional variability. One of the objectives identified by the GCOOS HABIOS Plan was to determine the level of toxins associated with blooms, and methods of directly measuring the concentrations of brevetoxin in the air and/or water are currently being developed. This may improve the respiratory irritation forecasts because cell concentrations are only a proxy for toxin data. The methods currently being explored would enable the processing of samples in the field to estimate toxicity, which would also reduce the lag time that exists between sample collection and estimating cell concentrations using microscopy in a laboratory. This in turn would support the HAB-OFS in efforts to better meet the target identified by GCOOS to be able to provide forecasts for critical beaches (Nowlin, Jochens, & Kirkpatrick, 2015).

The transport direction model currently only forecasts the general direction the bloom might move. Since the forecast model was first implemented operationally, more advanced hydrodynamic models have been developed such as the model operated by the University of South Florida. Although the HAB-OFS transport direction forecast performs well, the GCOOS HABIOS Plan indicated a need for 3D hydrodynamic models that provide both the transport direction and distance with associated uncertainties. The HABIOS Plan also recommended providing a graphical representation of the particle trajectories along with coastal GIS maps that include points of interest for human recreation and occupancy in order to support the prediction of animal and human health effects. Improving the transport forecast is also necessary to support the development of a respiratory irritation forecast model capable of meeting the GCOOS target of delivering forecasts at a finer resolution, especially for critical beaches (Nowlin, Jochens, & Kirkpatrick, 2015).

5.3 Summary of Recommended Future Actions

Additional actions with the potential to enhance the operational forecast system should be considered as follows:

- Complete the evaluation of VIIRS chlorophyll products and transition for operational use.
- Enhancement of respiratory irritation forecasts through:
 - creating methods to integrate direct measurements of the concentration of brevetoxin in the air and/or water currently being developed
 - completing an investigation of the HAB-OFS respiratory irritation forecast levels and the levels used by the MML Beach Conditions Reporting System to classify observational data (Mote Marine Laboratory, 2013)
 - exploring regional variation in forecast performance

- Continue the investigation of the use of a hydrodynamic model to forecast bloom transport in order to:
 - increase utility to HAB-OFS forecasters and users
 - improve the performance of the respiratory irritation forecast model
- Examine additional methods of assessing product utilization by subscribers that better capture evidence of the product that is being viewed, how its contents are being applied to bloom response, and whether the products meet user needs.

These enhancements are proposed with the Florida HAB bulletin forecast components in mind. However, some of the recommendations may also be applicable to the western Gulf of Mexico (Texas) HAB Forecast System, which was transitioned to the HAB-OFS in 2010, because the user needs are similar. On a broader scale, the assessment results may also be relevant to addressing stakeholder requirements in other forecast regions in the United States where forecast systems are still being developed.

6. ACKNOWLEDGEMENTS

The success of the Gulf of Mexico HAB-OFS is due to the contributions of numerous individuals and organizations. We would like to acknowledge the personnel in NOAA's Center for Operational Oceanographic Products and Services that served on the HAB-OFS analyst team during this assessment period: Robert Burrows, Edward Davis, Katherine Derner, Lori Fenstermacher, Kathleen Fisher, Karen Kavanaugh, Cristina Urizar, and Hua Yang. Other key NOAA individuals that assisted during the assessment period included Richard Stumpf and Michelle Tomlinson (National Centers for Coastal Ocean Science); Banghua Yan, Michael Soracco, and Kent Hughes (National Environment Satellite Data and Information Service); and Allison Allen, John Cassidy, Zhong Li, and Adria Schneck (Center for Operational Oceanographic Products and Services). Additional organizations that contributed to the HAB-OFS during this period included the Florida Fish and Wildlife Conservation Commission, Mote Marine Laboratory, Florida Department of Health, and NOAA's National Weather Service.

7. REFERENCES

- Derner, K., & Kavanaugh, K. (2014). Assessment of ensemble imagery products for HAB-OFS operational use. Unpublished internal document, NOAA.
- Doswell, C., Davies-Jones, R., & Keller, D. (1990). On summary measures of skill in rare event forecasting based on contingency tables. *Weather and Forecasting*, 576-585.
- Fisher, K., Allen, A., Keller, H., Bronder, Z., Fenstermacher, L., & Vincent, M. (2006). *Annual report of the Gulf of Mexico Harmful Algal Bloom Operational Forecast System (GOM HAB-OFS)*. NOAA Technical Report. NOS CO-OPS 047.
- Florida Fish and Wildlife Conservation Commission. (2015). *Red Tide Manatee Mortalities*. Retrieved August 27, 2015, from Florida Fish and Wildlife Conservation Commission: <http://myfwc.com/research/manatee/rescue-mortality-response/mortality-statistics/red-tide/>
- Hu, C., Barnes, B. B., Murch, B., & Carlson, P. (2013). Satellite-based virtual buoy system to monitor coastal water quality. *Optical Engineering*. Retrieved April 22, 2015, from <http://opticalengineering.spiedigitallibrary.org/article.aspx?articleid=1787513>
- Hu, C., Cannizzaro, J., Carder, K. L., Lee, Z., Muller-Karger, F. E., & Soto, I. (2011). Case Study 7: Red Tide Detection in the Eastern Gulf of Mexico Using MODIS Imagery. In J. Morales, V. Stuart, T. Platt, & S. Sathyendranath, *Handbook of Satellite Remote Sensing Image Interpretation: Applications for Marine Living Resources Conservation and Management* (pp. 95-110). Dartmouth, Canada: EU PRESPO and IOCCG. Retrieved April 22, 2015, from http://www.ioccg.org/handbook/casestudy7_hu_etal.pdf
- Kavanaugh, K., Derner, K., Fisher, K. M., Davis, E., Urizar, C., & Merlini, R. (2013). *Assessment of the eastern Gulf of Mexico Harmful Algal Bloom Operational Forecast System (GOMX HAB-OFS): A comparative analysis of forecast skill and utilization from October 1, 2004 to April 30, 2008*. NOAA Technical Report. NOS CO-OPS 066.
- Kirkpatrick, B., & Currier, R. (2010, September 14). *Patent No. US 7,797,109 B2*. U.S.
- Kirkpatrick, B., Fleming, L., Squicciarini, D., Backer, L., Clark, R., Abraham, W., . . . Baden, D. (2004). Literature review of Florida red tide: Implications for human health effects. *Harmful Algae*, 99-115.
- Mote Marine Laboratory. (2013). *Beach Conditions Reporting System for the Gulf Coast of Florida*. Retrieved February 8, 2013, from <http://coolgate.mote.org/beachconditions/>
- NOAA. (2013). *NOAA Harmful Algal Bloom Operational Forecast System (HAB-OFS)*. Retrieved February 8, 2013, from <http://tidesandcurrents.noaa.gov/hab/index.html>
- NOAA/Space Weather Prediction Center. (2007, October 1). *Forecast Verification Glossary*. Retrieved August 9, 2011, from http://www.swpc.NOAA.gov/forecast_verification/Glossary.html
- Nowlin, W., Jochens, A., & Kirkpatrick, B. (2015). *Harmful Algal Bloom Integrated Observing System Plan*. Gulf of Mexico Coastal Ocean Observing System. Retrieved from <http://gcoos.tamu.edu/documents/HABIOSPlan-Sept2015.pdf>
- Nurmi, P. (2005). General guide to the verification of local weather forecasts. *NOMEK Training*. Oslo.
- Pierce, R., Henry, M., Blum, P., Lyons, J., Cheng, Y., Yazzie, D., & Zhou, Y. (2003, January). Brevetoxin concentrations in marine aerosol: Human exposure levels during a *Karenia*

- brevis harmful algal bloom. *Bulletin of Environmental Contamination and Toxicology*, 70(1), 161-165.
- Stumpf, R., Culver, M., Tester, P., Tomlinson, M., Kirkpatrick, G., Pederson, B., . . . Soracco, M. (2003). Monitoring *Karenia brevis* blooms in the Gulf of Mexico using satellite ocean color imagery and other data. *Harmful Algae*, 147-160.
- Stumpf, R., Litaker, R., Lanerolle, L., & Tester, P. (2008). Hydrodynamic accumulation of *Karenia* off the west coast of Florida. *Continental Shelf Research*, 189-213.
- Stumpf, R., Tomlinson, M., & Fisher, K. (2009, February 5). MODIS and SeaWiFS Q&A. Unpublished internal document, NOAA.
- Stumpf, R., Tomlinson, M., Calkins, J., Kirkpatrick, B., Fisher, K., Nierenberg, K., . . . Wynne, T. (2009). Skill assessment for an operational algal bloom forecast system. *Journal of Marine Systems*, 151-161.
- Thornes, J., & Stephenson, D. (2001). How to judge the quality and value of weather forecast products. *Meteorological Applications*, 307-314.
- Tomlinson, M., Stumpf, R., Ransibrahmanakul, V., Truby, E., Kirkpatrick, G., Pederson, B., . . . Heil, C. (2004). "Evaluation of the use of SeaWiFS imagery for detecting *Karenia brevis* harmful algal blooms in the eastern Gulf of Mexico. *Remote Sensing of Environment*, 293-303.
- Zheng, L., & Weisberg, R. (2012). Modeling the west Florida coastal ocean by downscaling from the deep ocean, across the continental shelf and into the estuaries. *Ocean Modelling*, 48, 10-29.

8. APPENDICES

APPENDIX I

Example of a HAB bulletin for the Southwest Florida region.

APPENDIX II

Methods.

APPENDIX III

Forecast regions defined for Southwest Florida HAB bulletins.

APPENDIX IV

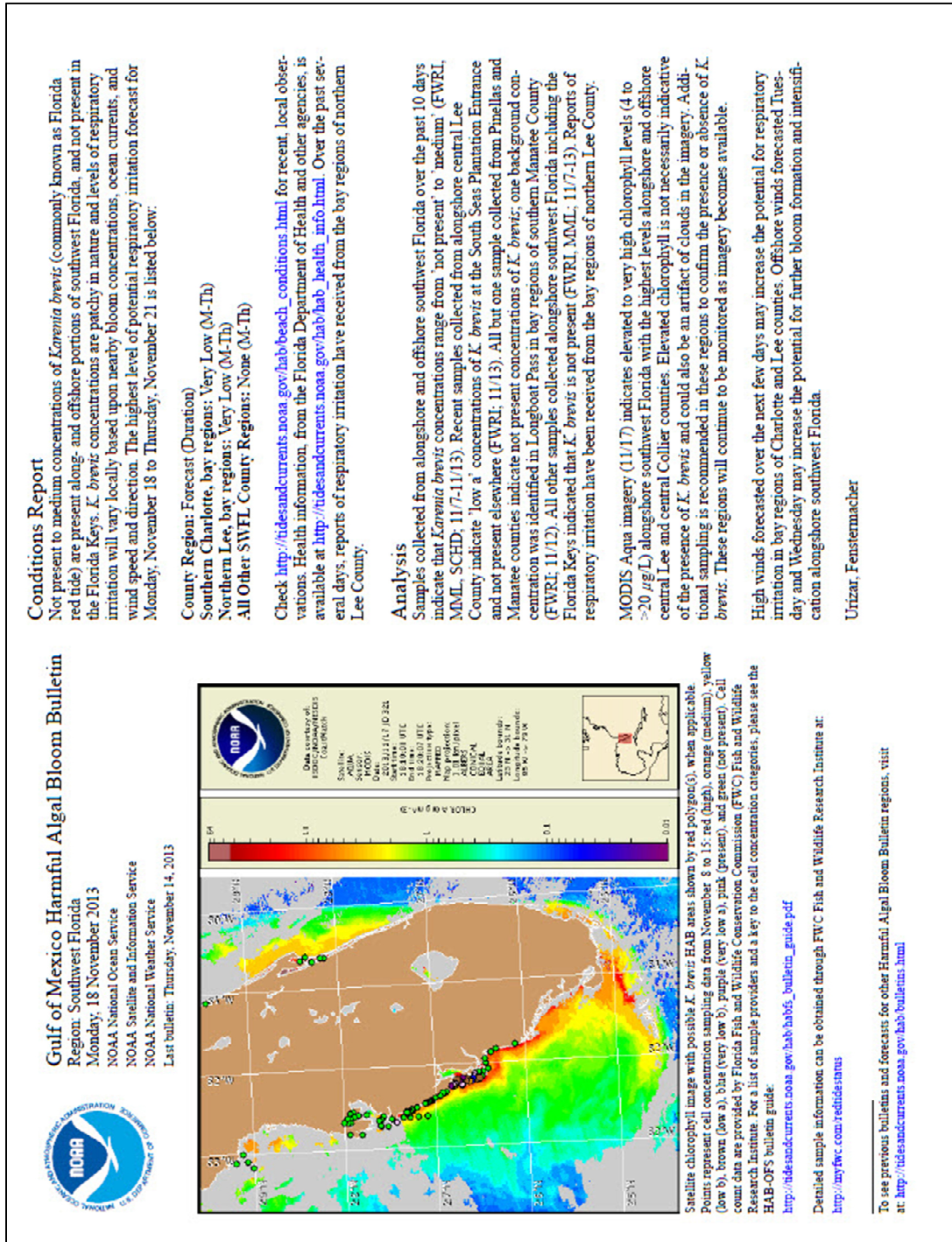
List of organizations that contributed to the 2008-2014 HAB-OFS bulletins for Southwest Florida.

APPENDIX I

Example of a HAB bulletin for the Southwest Florida region. The HAB-OFS Bulletin Guide provides further information on the data that are integrated, components of the bulletin, and how it is used:

http://tidesandcurrents.noaa.gov/hab/habofs_bulletin_guide.pdf.

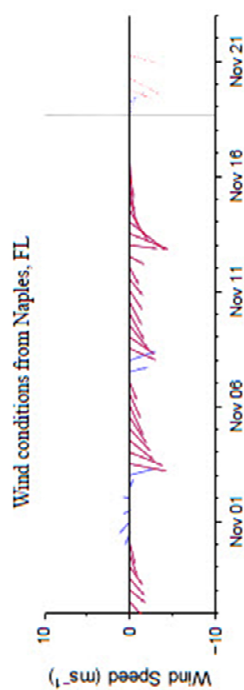
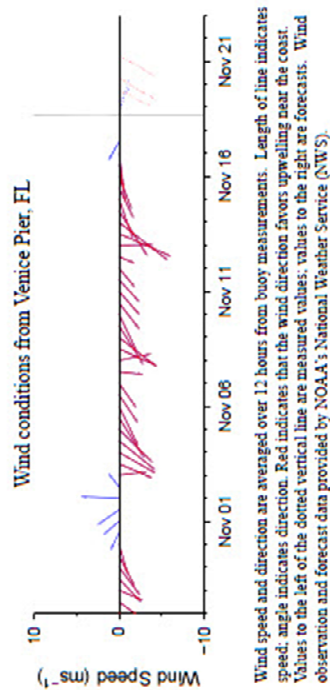
Example of a HAB bulletin for the southwest Florida region (page 1).



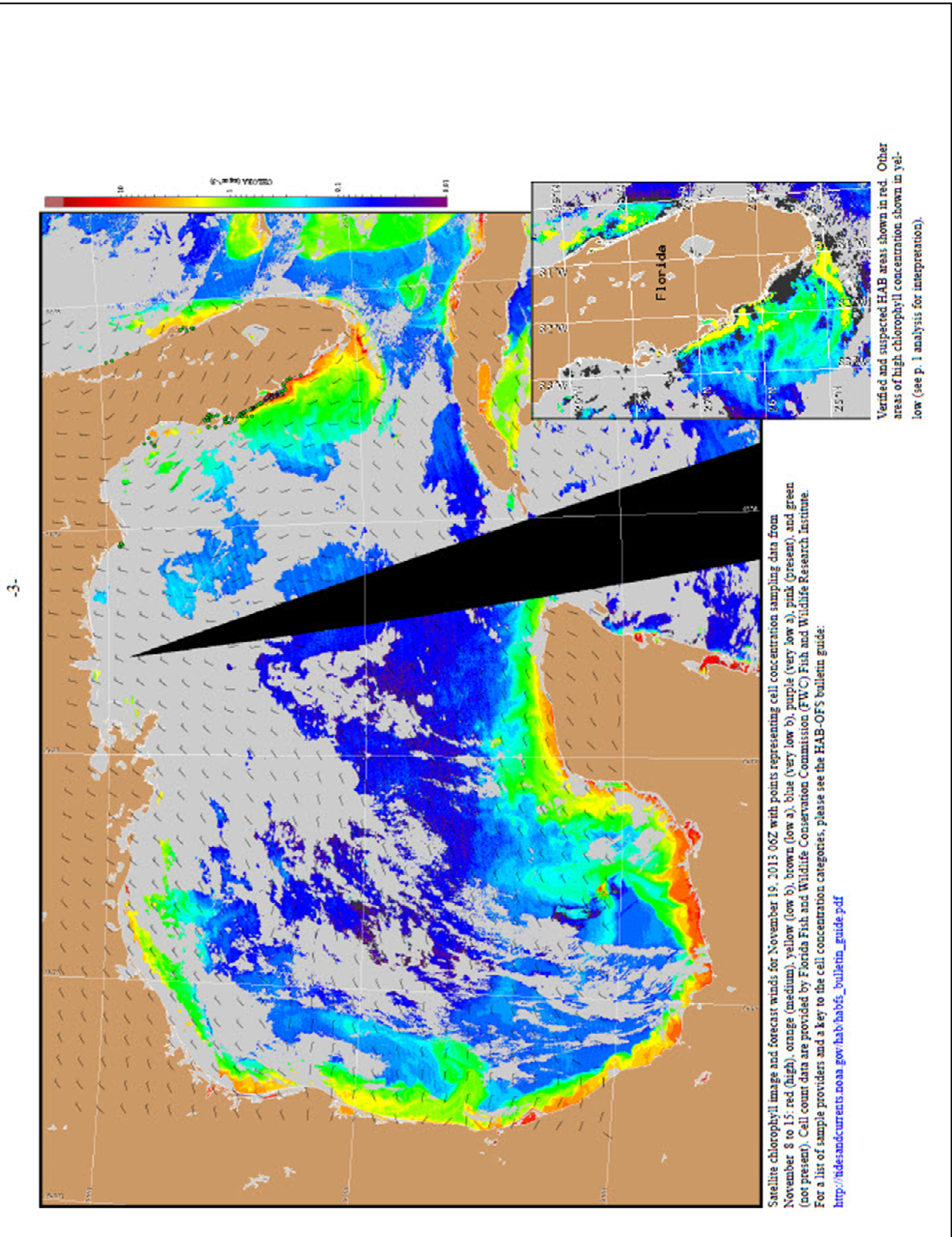
-2-

Wind Analysis

SW Florida: Today: Southerly winds (5 kn, 3 m/s) today becoming westerly (10 kn, 5 m/s) in the afternoon. Westerly winds (5 kn) tonight. Northeasterly winds (10-20 kn, 5-10 m/s) Tuesday and Wednesday. Easterly winds (10-20 kn) Wednesday night and Thursday.



Example of a HAB bulletin for the Southwest Florida region (page 3).



APPENDIX II

Current methods used by the HAB-OFS for forecasting and product assessment.

A. Operations

On October 1, 2004, the Center for Operational Oceanographic Products and Services (CO-OPS) transitioned a new ecological forecast system for HABs in the Gulf of Mexico, known as the GOMX HAB-OFS, from research to operational status. This was part of a NOAA collaborative effort with the National Centers for Coastal Ocean Science (NCCOS: science and research), the Coastal Services Center (CSC: technology development and public outreach), and the National Environmental Satellite, Data and Information Service (NESDIS/CoastWatch Program: satellite ocean color imagery), as well as a NOAA-wide effort to increase and enhance ecological forecasting products and services. Under the system's previous research status, bulletins were issued only as NCCOS resources allowed and bloom occurrence dictated. The operational status enabled regular dissemination of forecast products to accommodate user requirements. In 2008, all remaining technological and outreach activities formerly conducted by CSC were transferred to CO-OPS. Operations discussed in this report are relevant to the years from BY2004-2014.

The GOMX HAB-OFS employs a combination of automated processing and manual analyses using a web-based interface. From BY2004-2011, SeaWiFS satellite ocean color imagery (provided by NOAA's CoastWatch Program) was processed using a chlorophyll algorithm. Daily chlorophyll images were analyzed in conjunction with chlorophyll anomaly imagery highlighting regions of above-average elevated chlorophyll (as determined through a 60-day running mean) to determine the potential boundaries of harmful algal blooms containing the species *K. brevis* (Stumpf, et al., 2003). MODIS ocean color imagery was processed from the Aqua sensor using an identical procedure and served as a backup imagery source from 2004-2011 when SeaWiFS imagery was unavailable due to technical issues. Following the termination of the SeaWiFS satellite in 2011, MODIS Aqua imagery was used as the primary imagery source with MODIS Terra imagery as a back-up source.

The forecast system also incorporated the following data for bloom analysis and confirmation: hindcast and forecast winds available through the National Weather Service's (NWS) National Data Buoy Center (NDBC) and the NWS's North American Mesoscale (NAM) model; a wind transport model developed by NCCOS; and *in situ* *K. brevis* cell count data from several organizations including the FWRI and MML. In 2006, daily respiratory irritation, dead fish, and discolored water reports became available at many beaches in Southwest Florida through the establishment of the MML Beach Conditions Reporting System and were incorporated in subsequent bloom analyses and assessments. Daily respiratory irritation forecasts were developed by forecast regions (see APPENDIX III), based on the highest *K. brevis* concentrations and prevailing wind direction in the region. These resources, coupled with scientific expertise, were synthesized to analyze data and forecast potential for *K. brevis* bloom formation at the coast, transport, intensification, and associated level of respiratory irritation. To produce these forecasts, the HAB-OFS analysts relied mainly upon applying established scientific rules and heuristic and numerical models that NCCOS scientists developed and tested (Stumpf, et al., 2003; Tomlinson, et al., 2004; Stumpf, Litaker, Lanerolle, & Tester, 2008). To

ensure quality control, each bulletin was written by a primary analyst and reviewed by a second analyst. Additional information about the HAB-OFS bulletin contributors and the data they provide is available in APPENDIX IV, the HAB Bulletin Guide at http://tidesandcurrents.noaa.gov/hab/habfs_bulletin_guide.pdf, and at <http://tidesandcurrents.noaa.gov/hab/contributors.html>.

Operational HAB forecasts were communicated through two main products that served as decision support tools.

- 1) The **HAB bulletin** provided a detailed scientific analysis of satellite ocean color imagery, water samples and health reports, and meteorological and oceanographic data, and included all relevant forecasts. The bulletin was disseminated via email to registered coastal resource managers, academics, and public health officials with an email subject line indicating the relevant geographic region (see Figure 22 for a map of the regions and Figure 23 for the geographic distribution of subscribers). The subject line also indicated the priority level of the bulletin for consideration by managers: low, medium, or high (see Table 7).
- 2) The **public conditions reports**, a subset of the HAB bulletins, provided information about the presence or absence of *K. brevis* including a general description of the geographic region affected, forecasted levels of associated respiratory irritation, and any recent observations of respiratory irritation, dead fish or discolored water. The conditions reports were made available on the HAB-OFS website at <http://tidesandcurrents.noaa.gov/hab> immediately following bulletin dissemination. These reports were also made available through weekly updates disseminated by the Florida Department of Health. Beginning in 2012, these reports were also made available through the NOAA HAB-OFS Facebook Page at <https://www.facebook.com/Habredtidewatchnoaagov>.

Both products were routinely updated for the Southwest Florida region twice weekly on Mondays and Thursdays (or the day following a federal holiday) during *K. brevis* bloom events and once weekly during inactive periods. Products for Northwest and East Florida were only updated when *K. brevis* bloom events occurred in those regions (see Figure 22 for a map of the regions). The dissemination of unscheduled supplemental bulletins or conditions updates was also necessary when new data was received that indicated new bloom development or an increase in bloom extent, intensity, or the forecasted level of associated respiratory irritation.

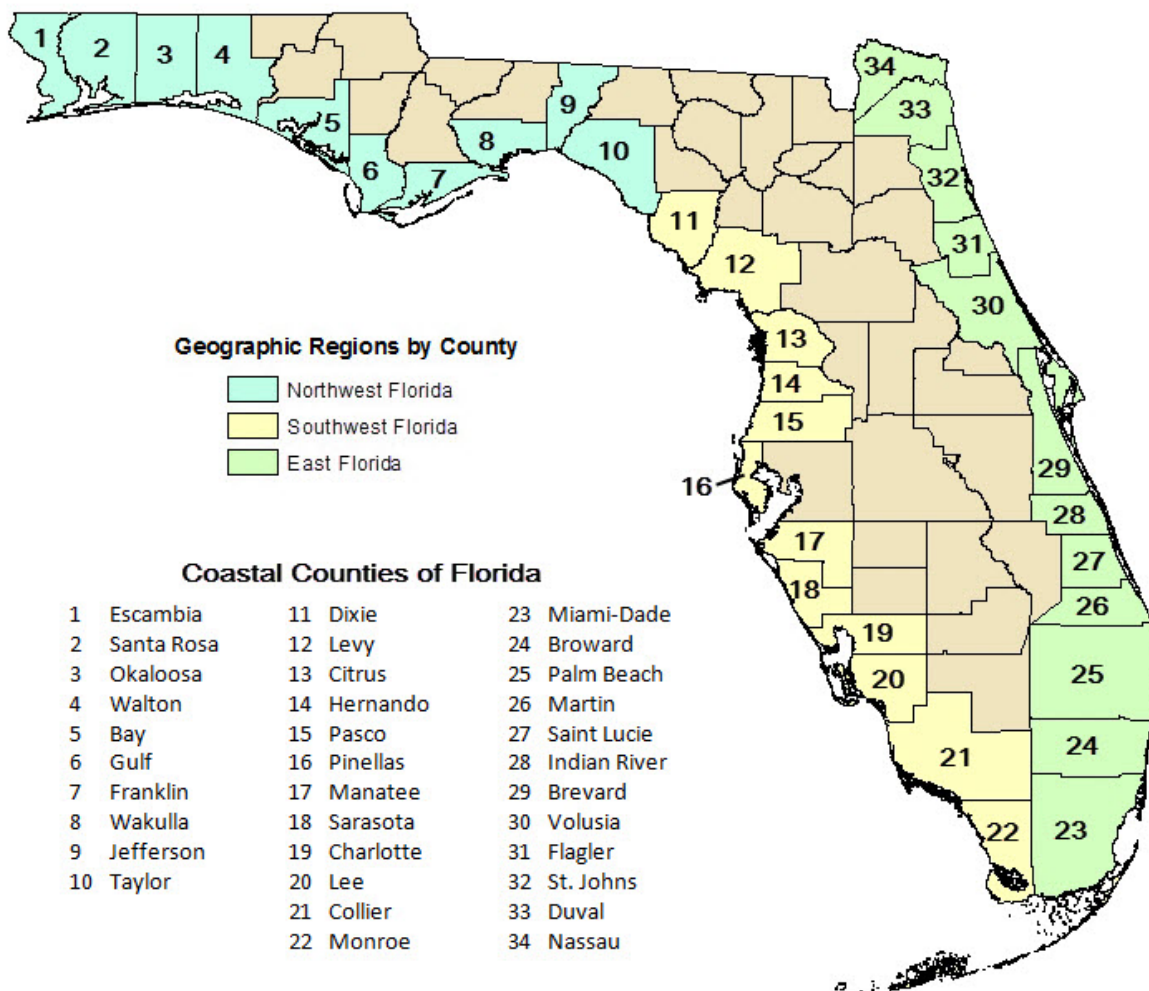


Figure 22. Map of Florida highlighting the three geographic regions for which HAB-OFS bulletins are disseminated. The Northwest region spans the coastal counties from Escambia through Taylor, the Southwest region spans from Dixie through Monroe (including the Florida Keys), and the eastern region spans from Miami-Dade through Nassau.

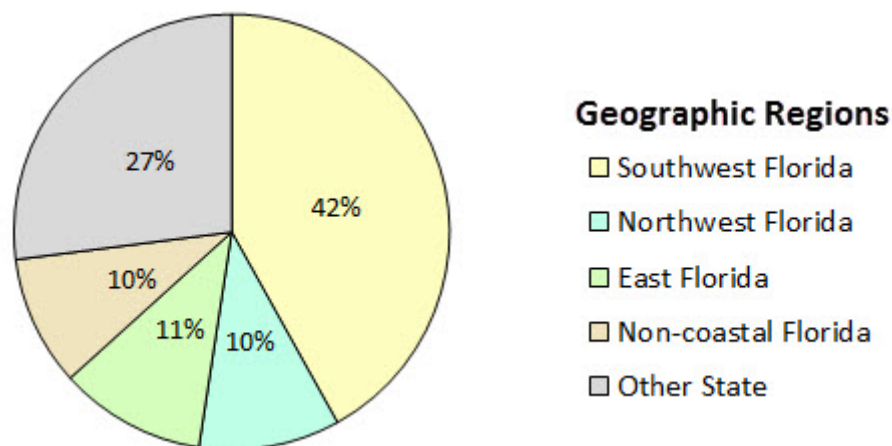


Figure 23. Distribution of Florida bulletin subscribers among the three geographic regions for which HAB bulletins were disseminated from March 2002 to April 30, 2014.

Note: Although the HAB-OFS did not issue operational bulletins until October 2004, users subscribed to the demonstration bulletins disseminated by NCCOS as early as March 2002.

Table 7. Priority levels assigned to bulletins indicating the corresponding level of action or response that resource managers might deem necessary based on the status of a harmful algal bloom of *Karenia brevis*.

PRIORITY LEVEL	DESCRIPTION
Low	<ul style="list-style-type: none"> • Inactive bloom • Resource managers may decide that no new action is necessary
Medium (No change)	<ul style="list-style-type: none"> • Active bloom, but no change in bloom conditions since previous bulletin • Resource managers may or may not decide that new action is necessary
High (Bloom Change)	<ul style="list-style-type: none"> • Active bloom, with recent changes in bloom conditions. Examples: <ul style="list-style-type: none"> ○ New bloom identified ○ Change in bloom extent (i.e. new or increase in coastal area impacted) ○ Bloom intensification (i.e. higher bloom concentrations detected) ○ Increases in the forecasted respiratory irritation levels • Resource managers may decide that immediate action is necessary

Operational status also requires on-call analyst response to public inquiries and bulletin subscription requests. The GOMX HAB-OFS utilized one central telephone number and email distribution address for responding to information requests from the general public and bulletin subscribers, in addition to fielding information requests and comments made through the NOAA HAB-OFS Facebook Page. Most often, inquiries received by the HAB-OFS pertained to present and future bloom conditions, potential impacts at specific locations and times for the purposes of event planning, general background information regarding *K. brevis* blooms and their occurrence, and requests to be added to the bulletin distribution list. Occasionally, the HAB-OFS also received inquiries from members of the public who were experiencing symptoms that might be associated with exposure to *K. brevis*. These inquiries from the public were used in the verification of forecasts and bulletin utilization where applicable.

Maintaining and improving upon past successes required sustained operational status during the 2008 to 2014 bloom years, including the ongoing support of a team of highly trained analysts,

continued adherence to standard operating procedures, maintenance of consistent analytical methods, and the perpetual refinement of tools and methods made possible by a continuing research to operations collaboration.

B. Definitions of the Forecast Types

As shown in Table 1, the Florida HAB-OFS provides forecasts for four different bloom components: transport direction, intensification, potential for bloom formation at the coast, and potential level of associated respiratory irritation (see Table 2). When there are gaps in available data, such as satellite imagery, nowcasts for transport direction, intensification, and potential for bloom formation at the coast are also provided, indicating the potential conditions up to three days prior to the bulletin day. For the purposes of the evaluation, nowcasts were assessed with forecasts. Transport direction is defined as the direction a bloom is likely to migrate, provided as the cardinal direction parallel to the coast, onshore, or no change, and is determined from winds forecasted by the NWS, Ekman transport, and geostrophic flow. Intensification is defined as an expected increase in *K. brevis* concentrations within an existing bloom area at the coast. This is a result of upwelling favorable winds in a developing bloom and downwelling favorable conditions in an established bloom after a period of relaxation. Potential for bloom formation at the coast is defined as the potential for a *K. brevis* bloom developing offshore and at depth to reach the coast. A bloom formation forecast is made from August 1 through December 1 when climatological conditions have been correlated with bloom formation along the Southwest Florida coast (Stumpf, Litaker, Lanerolle, & Tester, 2008). During this time, analysts produce a forecast for bloom formation at the Southwest Florida coast when *K. brevis* concentrations are below 10,000 cells/L and upwelling favorable conditions are present. Analysts also forecast when bloom formation at the coast is unlikely year round.

Although impacts from a bloom include adverse coastal conditions like the presence of dead fish and discolored water, the only impact associated with *K. brevis* blooms that is currently forecasted by the HAB-OFS is the potential level of respiratory irritation at the coast. Respiratory irritation is forecast in levels ranging from “very low” to “high” (in addition to “none” or “not expected”) based on NWS forecasted wind speed and wind direction, water sample location, and *K. brevis* cell concentrations (see Table 8 for cell concentration categories). The levels of respiratory irritation that are forecasted by the HAB-OFS correspond with the part of the population most likely to be affected. The “very low” respiratory irritation level affects only people with severe or chronic respiratory conditions such as cystic fibrosis and asthma. Similarly, the “low” respiratory irritation level affects people who are otherwise healthy, but are more sensitive to *K. brevis* aerosols. The “moderate” respiratory irritation level indicates that the general public may potentially experience mild respiratory symptoms, while the “high” respiratory irritation level affects the general public with adverse respiratory symptoms (NOAA, 2013). Symptoms associated with *K. brevis* include eye and respiratory irritation (coughing, sneezing, tearing, and itching). Refer to Table 2 for more information about the respiratory irritation levels. Due to limited spatial and temporal observations, these forecasts are made for each half-county region (see APPENDIX III for defined forecast regions). The respiratory irritation forecast is only made for coastal and bay regions because respiratory irritation levels are not well understood in open water regions (Stumpf R. , et al., 2009).

Environmental variations in geographic regions influence the forecasts that can be made and the analytical methods employed to develop the forecasts. An example of the variation in regional forecast capabilities is the inability to forecast bloom intensification in the Florida Keys region, as it is done in mainland areas of Southwest Florida and Northwest Florida. Another example is that although blooms have formed in the Northwest Florida region as well as Southwest Florida, based on the seasonal trends that the forecast is based on, bloom formation at the coast is only forecasted after August 1 for the region from Pinellas to Collier counties within Southwest Florida.

Table 8. The categories assigned to *Karenia brevis* cell concentrations identified from water samples by state, county, and local organizations in Florida.

CATEGORY	CELL CONCENTRATION (CELLS/L)
Not Present	0
Present (or Background)	1000 cells or less
Very Low a	>1000 to <5000
Very Low b	5000 to 10,000
Low a	>10,000 to <50,000
Low b	50,000 to 100,000
Medium	>100,000 to 1,000,000
High	>1,000,000

C. Skill Assessment

1. Overview of Procedure

Bulletin forecasts were recorded and evaluated by the primary analyst each week. Bulletin utilization and the forecast quality (i.e. accuracy, reliability, and skill) were assessed using the observational evidence available following the dissemination of each bulletin. All bulletin forecasts and assessments were subsequently reviewed and verified by additional analysts prior to the production of this report.

Product utilization was recorded as “confirmed” in the database when there was reliable evidence that the product was used. Evidence of usage came from sources such as the media and public health reports that referenced bulletin information, indications that sample collection was completed in an area specifically identified in the bulletin to contain a possible or confirmed bloom, and responses or inquiries from both partners and the general public referencing bulletin content. Interactions (“likes,” “shares,” “comments,” and “post clicks”) with conditions update posts issued on the NOAA HAB-OFS Facebook Page, added on September 7, 2012, also counted as confirmation of product utilization. Product utilization was recorded as “unconfirmed” when there was insufficient evidence.

Similarly, bulletin forecasts were evaluated for accuracy using evidence from a variety of sources (see Table 9). Transport direction forecasts were verified based on clear evidence of bloom movement in satellite imagery and/or a geographic shift in the position of *in situ* *K. brevis* concentration data over the specified time period. Intensification forecasts were verified based on evidence that chlorophyll levels in imagery or *in situ* *K. brevis* concentrations had increased,

decreased, or remained stable in the forecasted region. Forecasts of respiratory irritation and bloom formation at the coast were verified based on observational data reported during the specified time period and disseminated by state agencies and research institutions. Beginning in BY2006-2007, the HAB-OFS team was able to utilize the MML Beach Condition Reporting System for the Gulf Coast of Florida, which provides twice daily reports of the level of respiratory irritation associated with *K. brevis* estimated by a network of trained beach observers stationed along the Southwest Florida coast. The observations are classified into four categories of respiratory irritation as outlined in Table 10. (Kirkpatrick & Currier, 2010). The beach conditions reports account for a majority of the respiratory irritation observations provided to the HAB-OFS. Observed respiratory irritation was categorized and forecasts were then assessed using Table 11. Reports of respiratory irritation from additional sources, such as the HAB-OFS Facebook Page established in 2012, were also used if the reports were credible. If the person who experienced respiratory irritation did not have a pre-existing respiratory condition, then the level of irritation was assessed as “moderate”, which affects the general public.

Bulletin forecasts were considered “confirmed” when reliable evidence indicated that the forecasted conditions/events had been observed during the specified forecast period. When evidence indicated that the observed conditions/events were different from those that were forecasted, the forecast was recorded as “false” in the database. Forecast quality could not be analyzed further, and it was categorized as “unconfirmed” when the necessary observational evidence was not available. With regards to respiratory irritation, when reports provided by the MML Beach Conditions Reporting System, FWRI, and other trusted sources did not record respiratory irritation or reported a respiratory irritation of “none,” the observation could not definitively confirm that no respiratory irritation was experienced throughout the entire forecast region, due to the patchy nature of blooms. Therefore, forecasts were assessed as “unconfirmed” when a respiratory irritation level of “none” was reported from alongshore and in the bay regions of Florida.

The assessment data was then grouped together by both U.S. government fiscal year and bloom year. Fiscal year (October 1, XXXX to September 30, YYYY) was used to compare changes that may have occurred from one budget year to the next. However, *K. brevis* blooms more frequently develop between August and December, sometimes spanning two or more fiscal years, potentially skewing the results of statistical analyses. Thus, to avoid this issue, a time span was chosen that would best represent the 365-day HAB cycle (bloom year, BY). The time period from May 1, XXXX to April 30, YYYY was selected to best capture the typical seasonal cycle of *K. brevis* blooms in the Gulf of Mexico, from the initiation phase through termination. This minimized the bias in the evaluation results that might have been due to variations in cell concentrations over the course of a bloom’s life cycle, enabling a more meaningful comparison between years. Assessment statistics and graphs for bloom year are detailed throughout this report.

Table 9. Data and resources used to assess each forecast type included in a Florida bulletin.

FORECAST TYPE	CATEGORIES	ASSESSMENT BASED ON
Transport Direction	<ul style="list-style-type: none"> • North • South • East • West • Onshore • No Change 	<ul style="list-style-type: none"> • Visible movement of feature in satellite imagery • <i>In situ</i> samples confirm cell concentrations in new location • Reports of <i>K. brevis</i> induced respiratory irritation in a new location
Intensification	<ul style="list-style-type: none"> • Increase • No Change 	<ul style="list-style-type: none"> • Localized change in chlorophyll levels visible in satellite imagery • <i>In situ</i> samples confirm change in cell concentrations within the forecast region
Potential for Bloom Formation	<ul style="list-style-type: none"> • Favorable • Unfavorable 	<ul style="list-style-type: none"> • <i>In situ</i> sampling confirms new bloom-level concentrations of <i>K. brevis</i> within the forecasted region • Respiratory irritation reported within the forecasted region
Respiratory Irritation	<ul style="list-style-type: none"> • Very low • Low • Moderate • High • None 	<ul style="list-style-type: none"> • Reports of observed respiratory irritation (see Table 11) received from data partners and trusted sources

Table 10. Definitions of the levels of observed respiratory irritation as assessed by trained beach reporters for the Mote Marine Laboratory Beach Conditions Reporting System for the Gulf Coast of Florida (Kirkpatrick & Currier, 2010).

Level of Respiratory Irritation	Observations during 30 second Sample
None	No coughing/sneezing heard
Slight	A few coughs/sneezes heard
Moderate	A cough/sneeze heard every ~5 seconds
High	Coughing/sneezing heard almost continuously

Table 11. During a *Karenia brevis* bloom, reports of observed respiratory irritation were used to validate the corresponding level of respiratory irritation forecasted for that region according to this chart.

Highest Level of Respiratory Irritation Observed	Highest Level of Respiratory Irritation Forecasted					
	No forecast and/or no bloom	None	Very low	Low	Moderate	High
No reports (no data received)	N/A	Unconfirmed	Unconfirmed	Unconfirmed	Unconfirmed	Unconfirmed
None (no symptoms observed in region)	N/A	Unconfirmed	Unconfirmed	Unconfirmed	Unconfirmed	Unconfirmed
Very Low (only individuals with chronic respiratory conditions)	False	False	Confirmed	False	False	False
Slight (only sensitive individuals & those with chronic respiratory conditions)	False	False	Unconfirmed	Confirmed	False	False
Moderate (general public may notice mild symptoms)	False	False	False	False	Confirmed	Confirmed
High (general public may notice adverse symptoms)	False	False	False	False	Confirmed	Confirmed

2. Statistical Analysis

In order to assess the level of success, verify the forecasts, and continually improve the HAB-OFS, forecast quality and bulletin utilization were evaluated regularly.

a) Capability of Assessing Bulletin Utilization and Forecasts

Before beginning a more extensive evaluation of forecast quality and utilization, the number of bulletins that were capable of being assessed was examined and compared to the number that could not be assessed. As described earlier in the Skill Assessment section of this appendix, the assessment of bulletin utilization and forecasts were limited by the availability of post-bulletin observational evidence. Entries were recorded as unconfirmed when there was insufficient evidence for further assessment. Assessment capability varied, especially between the types of

forecasts (i.e. transport direction, intensification, bloom formation at the coast, and level of respiratory irritation). Reliance on reports of field observations made assessment difficult in some cases. In order to evaluate the assessment capability, we calculated the percent of assessable bulletins for each forecast type as well as utilization.

b) Forecast Frequency

Although all bulletins included at least one forecast, some components were forecasted more often than others. This is a direct result of the bloom conditions during the forecast period. For example, the development of an intensification forecast relied upon the presence of a coastal bloom, whereas a transport direction forecast could be developed for either an active bloom or an unconfirmed feature appearing in the satellite imagery. It could also indicate that some components were easier to forecast than others using established forecast system rules guided by existing scientific knowledge. The frequency that each component was forecast was determined by calculating the proportion of bulletins that included each of the individual components.

c) Forecast Verification and Skill Assessment

Forecast quality was estimated for each of the following forecast types: bloom transport direction, bloom intensification, bloom formation at the coast, and the daily potential level of respiratory irritation at the coast. Statistics were compared between bloom years (5/1/XXXX to 4/30/YYYY) and geographic regions.

Since there is no single measure of the quality of a forecast, several different verification measures were calculated (Doswell, Davies-Jones, & Keller, 1990). All of the forecasts included in the HAB bulletins were binary, i.e. the predicted event was observed to either occur or not occur. Contingency tables were created showing the frequency of “yes” and “no” matched forecasts and observations (see Table 12). In reference to Table 12, there are two types of correct forecasts, indicated by the letters A and D, and two types of false forecasts, indicated by the letters B and C. The letter A represents the number of “hits” or the number of events that were forecasted and also observed. The letter D represents the number of “correct rejections” or the number of times an event was correctly forecast to not occur. The letter B represents the number of “false alarms” or the number of events that were forecasted, but not observed. The letter C represents the number of “misses” or the number of events that were not forecasted, but were observed. The total number of forecasts is represented by N.

Table 12. Example of a 2 x 2 contingency table showing the types of correct forecasts (hit and correct rejection) and false forecasts (false alarm and miss), with the letters A through D representing the number of events forecasted and/or observed.

		EVENT OBSERVED?		
		Yes	No	Marginal Total
EVENT FORECAST?	Yes	<i>Hit</i> (A)	<i>False Alarm</i> (B)	<i>Forecast</i> (A+B)
	No	<i>Miss</i> (C)	<i>Correct Rejection</i> (D)	<i>Not Forecast</i> (C+D)
	Marginal Total	<i>Observed</i> (A+C)	<i>Not Observed</i> (B+D)	<i>Sum Total</i> (A+B+C+D)

There are numerous categorical statistics that can be used to assess forecast quality. The statistics selected for this report include those commonly used for the verification of binary meteorological forecasts and are appropriate for the verification of rare events like harmful algal blooms. Three basic attributes of forecasts were measured: accuracy, reliability, and skill.

Forecast accuracy was measured through the use of four different statistics: proportion correct, probability of detection (or hit rate), false alarm ratio, and threat score (or critical success index). Proportion correct (PC) is measured by the number of correct forecasts compared to the total number of forecasts. With respect to the 2 x 2 contingency table (Table 12):

$$PC = (A+D)/N \quad [\text{range: 0 to 1}] \quad (1)$$

where a perfect score equals one or 100% (Nurmi, 2005). Probability of detection (POD), or hit rate, measures the proportion of observed events that were correctly forecast. With respect to the 2 x 2 contingency table (Table 12):

$$POD = A/(A+C) \quad [\text{range: 0 to 1}] \quad (2)$$

where one is a perfect score (Nurmi, 2005). Since the POD could be artificially inflated by producing excessive “no” forecasts, it should be considered along with a statistic sensitive to the number of false alarms generated by the forecast system. The false alarm ratio (FAR) is a verification measure of categorical forecast performance that compares the number of false alarms to the total number of forecasts. With respect to the 2 x 2 contingency table (Table 12):

$$FAR = B/(A+B) \quad [\text{range: 1 to 0}] \quad (3)$$

where zero is a perfect score (Nurmi, 2005). The threat score (TS) is commonly used to measure the performance of rare event forecasts. It is a measure for the event being forecast after removing the number of times the event was correctly forecasted to not occur. With respect to the 2 x 2 contingency table (Table 12):

$$TS = A/(A+B+C) \quad [\text{range: 0 to 1}] \quad (4)$$

where a perfect score is one (Nurmi, 2005).

The reliability of binary forecasts is often measured by calculating the bias, a statistic that demonstrates whether there are consistent differences between the frequency of observed events and the frequency of event forecasts which would indicate a tendency towards over- or under-forecasting. When events are often predicted, but not observed they are said to be over-forecast. The term under-forecasting describes when forecasts are consistently not issued for events that are observed (Thornes & Stephenson, 2001). The frequency of event forecasts are compared to the frequency of observed events. With respect to the 2 x 2 contingency table (Table 12):

$$BIAS = (A+B)/(A+C) \quad [\text{range: } 0 \text{ to } \infty] \quad (5)$$

where a score of one indicates no bias, and a score greater than one indicates that the forecast system over-forecasts the event. A score of less than one suggests that the forecast system under-forecasts the event (Nurmi, 2005).

Forecast skill is often estimated using a skill score that compares the variation in the accuracy of a forecast with an estimate of the forecast results that could be due solely to chance, climatology, or persistence. The Heidke skill score (HSS) was selected for this assessment because it is commonly used to assess rare event forecasts, such as tornadoes and flash floods (Doswell, Davies-Jones, & Keller, 1990). It is a skill corrected verification measure of categorical forecast performance that references the proportion of correct forecasts relative to the number of correct forecasts that could be made by random chance (NOAA/Space Weather Prediction Center, 2007). With respect to the 2 x 2 contingency table (Table 12), the Heidke skill score is calculated as:

$$HSS = 2(AD-BC) / \{(A+C)(C+D) + (A+B)(B+D)\} \quad [\text{range: } -\infty \text{ to } 1] \quad (6)$$

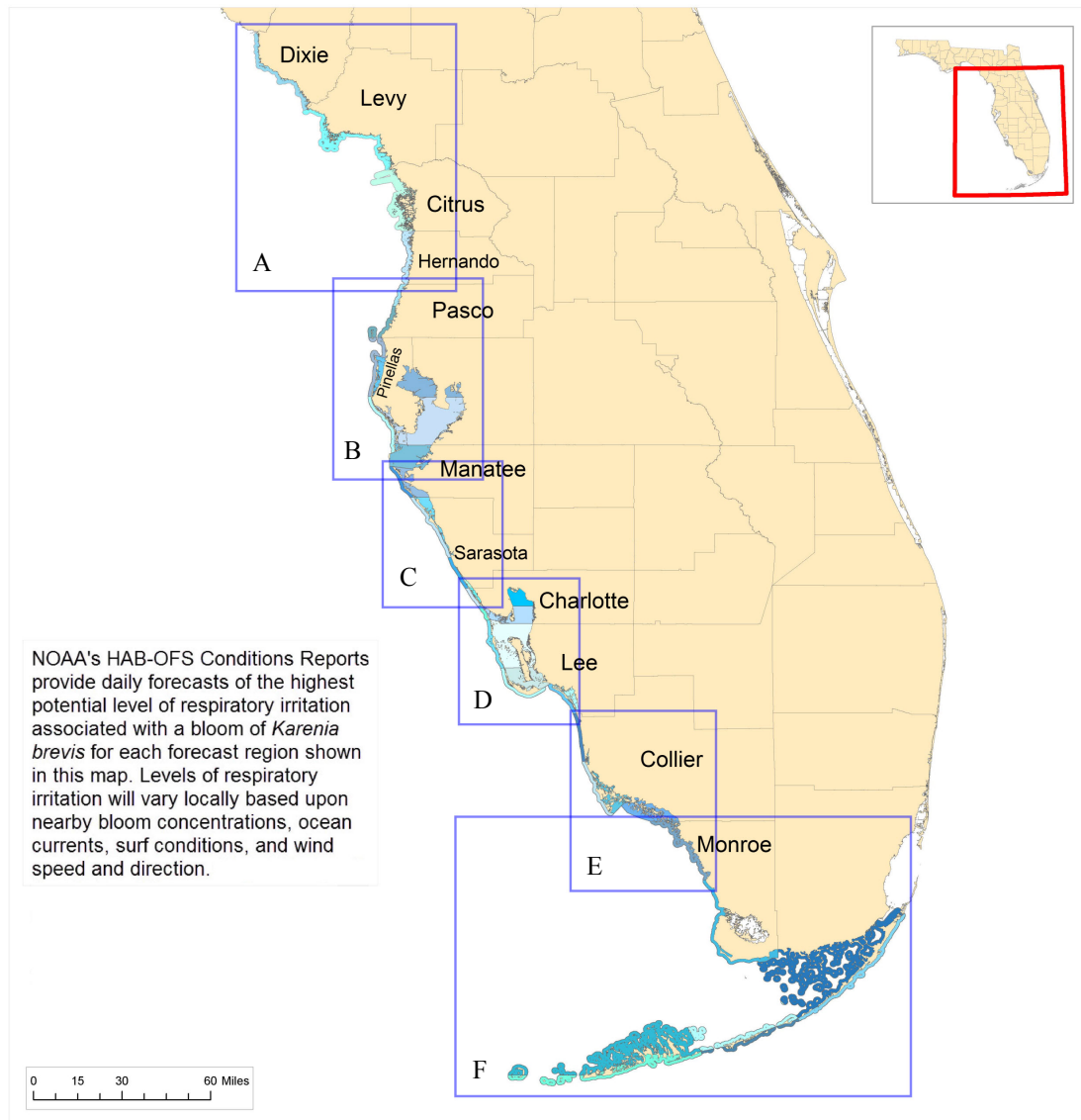
where a perfect score is one or 100%. A score of zero indicates that the forecast is no better than random chance at predicting the event (i.e. no forecast skill) (Nurmi, 2005).

d) Bulletin Utilization

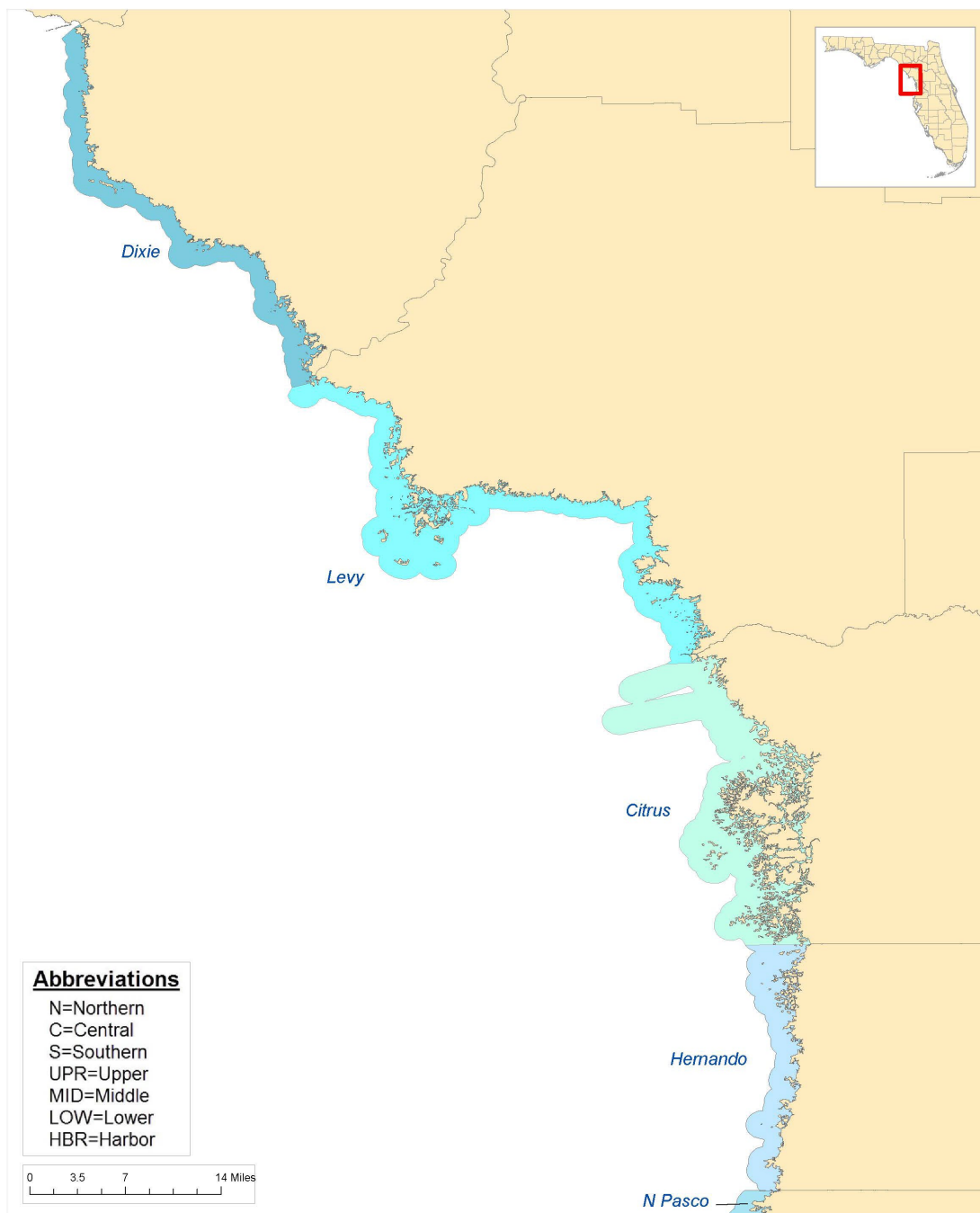
A successful forecast system is one that not only produces accurate forecasts, but also one that is well-used by its intended audience(s). Bulletin utilization was confirmed based on evidence from sources that included sampling response to cited bloom regions, media or public health reports identifying bulletin information, and written or phoned inquiries and responses that were based on bulletin analyses. In BY2012-2013, the NOAA HAB-OFS Facebook Page was created to better disseminate public conditions reports and engage with the general public and bulletin subscribers. Interaction (likes/shares/comments) with Florida conditions report posts on the Facebook Page were counted as bulletin utilization. The proportion of bulletins that were confirmed as utilized was then calculated for each fiscal year, bloom year, and priority level.

APPENDIX III

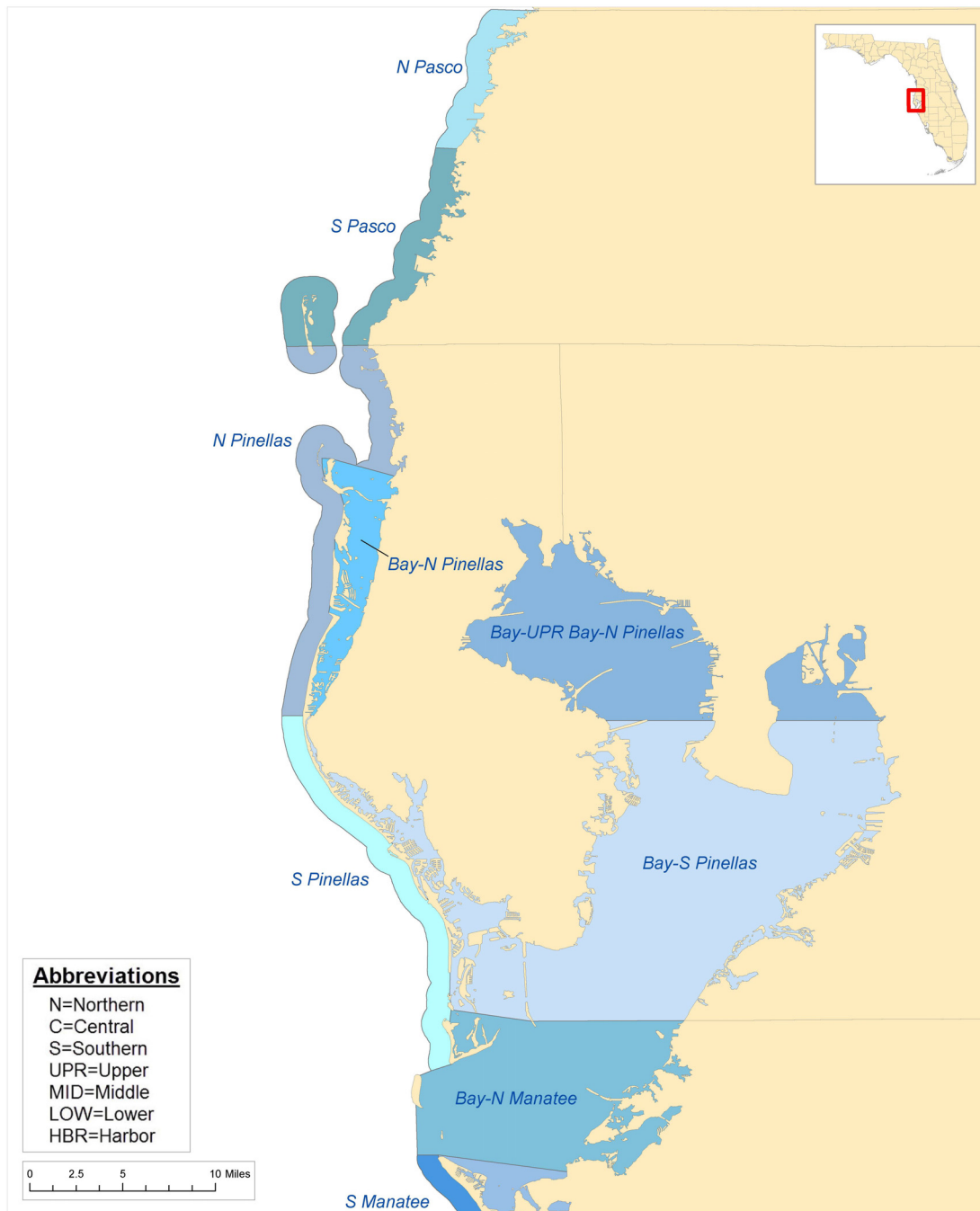
Forecast regions defined for Southwest Florida HAB bulletins.



A. Forecast regions defined for Southwest Florida HAB bulletins: Dixie to Pasco Counties



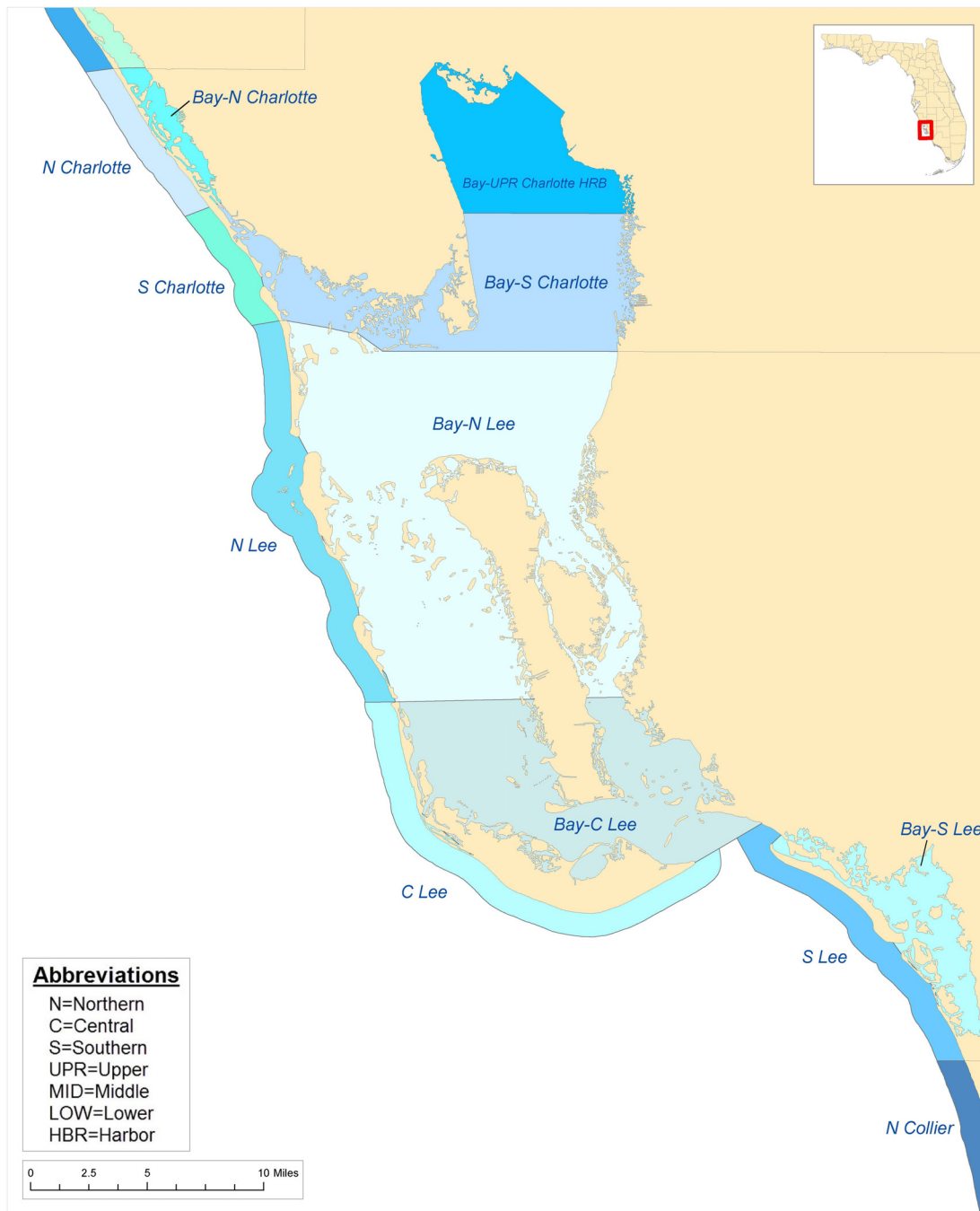
B. Forecast regions defined for Southwest Florida HAB bulletins: Pasco to Manatee Counties



C. Forecast regions defined for Southwest Florida HAB bulletins: Manatee and Charlotte Counties



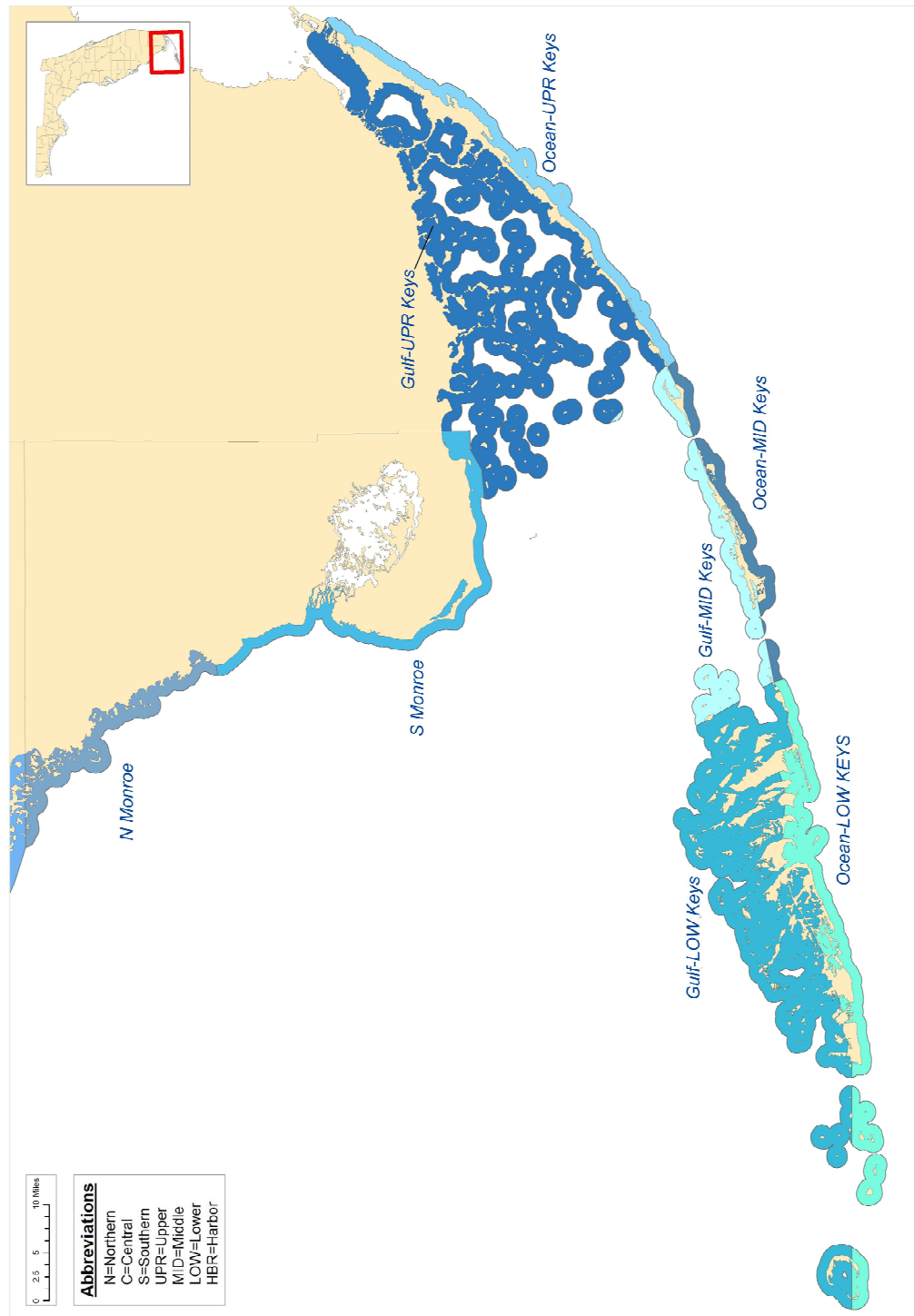
D. Forecast regions defined for Southwest Florida HAB bulletins: Charlotte and Collier Counties



E. Forecast regions defined for Southwest Florida HAB bulletins: Collier and Monroe Counties



F. Forecast regions defined for Southwest Florida HAB bulletins: Monroe County to the Florida Keys



APPENDIX IV

List of organizations that contributed to the 2008-2014 HAB-OFS bulletins for Southwest Florida. The HAB-OFS Bulletin Guide provides further information on the data that are integrated, components of the bulletin, and how data is used:
http://tidesandcurrents.noaa.gov/hab/habfs_bulletin_guide.pdf.

List of organizations that contributed to the 2008-2014 HAB-OFS bulletins for Southwest Florida

Organization	HAB-OFS Contributions	Website
NOAA Center for Operational Oceanographic Products & Services	<ul style="list-style-type: none"> • Forecast analysis • Operations 	http://tidesandcurrents.noaa.gov
NOAA National Centers for Coastal Ocean Science	<ul style="list-style-type: none"> • Research & Development 	http://coastalscience.noaa.gov
NOAA National Weather Service	<ul style="list-style-type: none"> • Wind data 	http://www.weather.gov
NOAA National Data Buoy Center	<ul style="list-style-type: none"> • Wind data 	http://www.ndbc.noaa.gov
NOAA CoastWatch	<ul style="list-style-type: none"> • Remote sensing data 	http://coastwatch.noaa.gov/cwn
NASA SeaWiFS Project	<ul style="list-style-type: none"> • Remote sensing data 	http://oceancolor.gsfc.nasa.gov/SeaWiFS/
NASA MODIS Aqua	<ul style="list-style-type: none"> • Remote sensing data 	http://modis.gsfc.nasa.gov/
Florida Fish and Wildlife Conservation Commission's Fish and Wildlife Research Institute	<ul style="list-style-type: none"> • <i>In situ</i> cell count data • Water sample results • Fish kill database • Other reports of health impacts (i.e. respiratory irritation or discolored water) 	http://myfwc.com/research
Mote Marine Laboratory	<ul style="list-style-type: none"> • <i>In situ</i> cell count data • Water sample results • Beach Conditions Reporting System (including observations of respiratory irritation, dead fish, discolored water, and wind direction) 	http://www.mote.org
Florida Department of Health	<ul style="list-style-type: none"> • Reports of health impacts 	http://www.floridahealth.gov/index.html
Sarasota County Department of Health	<ul style="list-style-type: none"> • Water sample results 	http://www.ourgulfenvironment.net/HomePage.aspx
Collier County Engineering and Natural Resources Division	<ul style="list-style-type: none"> • Water sample results • Other reports of health impacts (i.e. respiratory irritation or discolored water) 	http://www.colliergov.net/Index.aspx?page=113