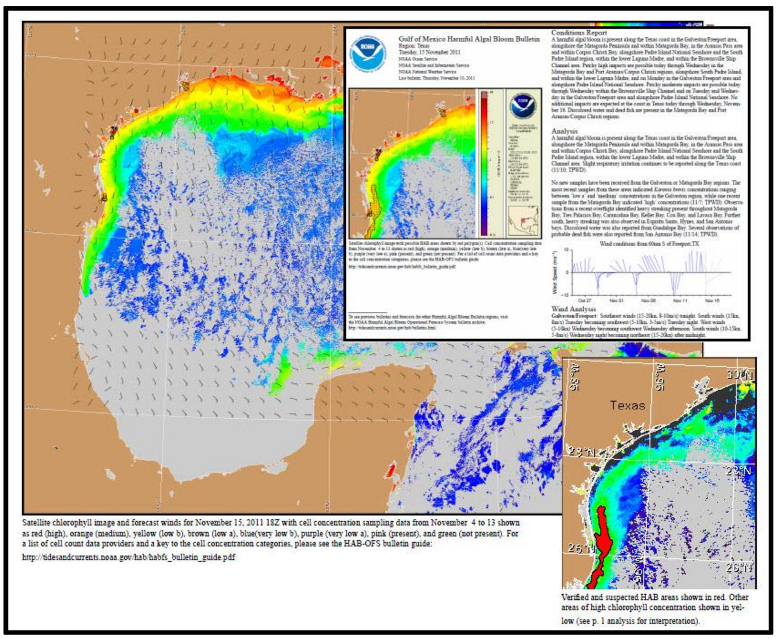
NOAA Technical Report NOS CO-OPS 080

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

An Analysis of Forecast Skill and Utilization from October 1, 2010 to April 30, 2014



HAB-OFS bulletin issued for the Western GOMX on November 15, 2011.

Silver Spring, Maryland
March 2015

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.

NOAA Technical Report NOS CO-OPS 080

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

An Analysis of Forecast Skill and Utilization from October 1, 2010 to April 30, 2014

Karen E. Kavanaugh
Katherine Derner
Edward Davis
Cristina Urizar

March 2015

U.S.DEPARTMENT OF COMMERCE
Penny Pritzker, Secretary

National Oceanic and Atmospheric Administration
Dr. Kathryn 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

LIST OF TABLES

LIST OF FIGURES

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 U.S., typically between August and December. While the western Gulf of Mexico has not typically experienced *K. brevis* blooms annually, analysis of historical records suggests that these blooms may be occurring with increasing frequency along the coast of Texas (Magana, Contreras, & Villareal, 2003). Numerous fish kills and various marine bird and mammal deaths have been linked to *K. brevis* blooms, and even very low concentrations of *K. brevis* from 5,000 to 10,000 cells/L prompt the closure of shellfish beds to prevent Neurotoxic Shellfish Poisoning (NSP) in humans (Tomlinson, et al., 2004). Under certain wind conditions and wave action, the cells of *K. brevis* can lyse releasing a toxin into the water. This toxin is then incorporated into the marine aerosol. Inhaling the toxin causes respiratory irritation 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). Winds can carry the toxic aerosols from nearshore surface blooms to distances at least 4.2 km inland from the beach, prompting necessary advisories at afflicted beaches (Kirkpatrick, et al., 2010).

To assist coastal managers in mitigating the impacts of harmful algal blooms (HABs), an ecological forecast system for the Gulf of Mexico was developed through the efforts of multiple offices within the National Oceanic and Atmospheric Administration (NOAA). In October 2004, this ecological forecast system was transitioned from research to operational status along the coast of Florida, creating the Gulf of Mexico HAB Operational Forecast System (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 during active bloom events (once weekly at times of bloom inactivity) and provide information concerning the possible identification of new blooms, in addition to monitoring existing blooms. Bulletins for the western Gulf of Mexico provide forecasts of bloom movement, including transport direction and distance, and daily coastal respiratory irritation. These forecasts are publicly available via the Internet at http://tidesandcurrents.noaa.gov/hab.

As a result of the forecasts in the bulletins, advance cautionary notice can be issued to protect beachgoers from experiencing respiratory irritation; necessary actions, such as closing shellfish beds, can be initiated before a bloom becomes a coastal hazard; and mass marine animal casualties can be minimized through providing advance information to personnel responsible for animal rescue, rehabilitation and release. The bulletins identify potential areas of HABs using satellite imagery and make use of transport models that project potential bloom movement. By doing so, the bulletins provide advance notice to appropriate state, county, and local agricultural and health service departments to initiate sampling programs to confirm the identity of any anomalously high chlorophyll features present in the imagery as blooms of *K. brevis*. If a feature is found to contain *K. brevis* at a concentration level capable of causing NSP when ingested, shellfish harvesting is prohibited in the region of the bloom and shellfish bed closures are listed on regional hotlines and via the Internet at http://www.dshs.state.tx.us/seafood/redtide.shtm. 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 in the 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 Texas during the bloom years from October 1, 2010 to April 30, 2014, with comparisons to those issued for Florida where possible (Kavanaugh, et al., 2013). A bloom year (BY) refers to the time period from May 1, XXXX to April 30, YYYY, where BY2010-2011 spans the period from May 1, 2010 to April 30, 2011 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 first bloom year only spans from October 1 rather than May 1 because the Texas HAB-OFS was not operational until October 1, 2010. 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 Texas HAB-OFS (Stumpf, et al., 2003; Tomlinson, et al., 2004; Wynne, Stumpf, Tomlinson, Ransibrahmankul, & Villareal, 2005). The results of this assessment will be used to guide enhancements to the operational forecast system with the goal 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 eastern Gulf of Mexico.

METHODS

2.1 Operations

On October 1, 2010, the NOAA Center for Operational Oceanographic Products and Services (CO-OPS) transitioned an operational forecast system for HABs in the Gulf of Mexico from research to operational status for the western Gulf of Mexico (Texas) (see Figure 1). This transition followed several years of successful operations of the GOMX HAB-OFS for the eastern Gulf of Mexico (Florida), which was transitioned to operations in 2004. These forecast systems are a 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, 2004-2008), 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 employee resources allowed and bloom occurrence dictated. The operational status enables regular dissemination of forecast products to accommodate user requirements. This report details the operational Texas bulletins for October 2010 through April 2014 (BY2010-2014).

During the BY2010-2014 assessment period, Moderate Resolution Imaging Spectroradiometer (MODIS) satellite ocean color imagery (provided by NOAA's CoastWatch Program) was processed from the Aqua sensor using a chlorophyll algorithm. Daily chlorophyll images were analyzed in conjunction with chlorophyll anomaly imagery highlighting regions of above-average chlorophyll (as determined through a 60-day running mean) to determine the potential presence or existing boundaries of harmful algal blooms containing the species *K. brevis* (Stumpf, et al., 2003). The surface waters along the Texas coast are prone to a high amount of suspended sediment, especially along the northeast coast, because the fine sediments are easily resuspended (unlike the coarser sediment on the Florida shelf). During resuspension events, benthic chlorophyll and sediment exceed the chlorophyll-a concentration in the water column so a revised chlorophyll algorithm developed by NCCOS is used that subtracts an estimate of the resuspended chlorophyll from the chlorophyll anomaly (Wynne, Stumpf, Tomlinson, Ransibrahmankul, & Villareal, 2005).

The following data was also incorporated during the BY2010-2014 assessment period for bloom analysis and confirmation: observed winds available through the National Data Buoy Center (NDBC) and the North American Mesoscale (NAM) model forecast winds; observed and forecast currents from the Texas General Land Office's Texas Automated Buoy System (TGLO/TABS) and Texas A&M University (TAMU); and *in situ K. brevis* cell count data from the Texas Parks and Wildlife Department (TPWD) and the TAMU Imaging FlowCytobot housed at the University of Texas Marine Sciences Institute in Port Aransas, TX. Using observed and forecasted current data, HAB analysts used the General NOAA Operational Modeling Environment (GNOME) particle trajectory modeling tool to predict the distance and direction of bloom movement. Reports of respiratory irritation, dead fish, and discolored water from TPWD, the Red Tide Rangers, and other organizations were also incorporated in bloom analyses and assessments. These resources, coupled with scientific expertise, were synthesized to analyze the bloom status and forecast *K. brevis* bloom transport direction, transport distance, and associated respiratory irritation. In addition to modeling tools, to produce these forecasts, the HAB-OFS analysts relied upon mental integration methods, applying established scientific rules and heuristic and numerical models that NCCOS scientists developed and tested (Stumpf, et al., 2003; Tomlinson, et al., 2004; Stumpf, et al., 2009). To ensure quality control, each bulletin was written by a primary analyst and reviewed by a second analyst for consensus. Additional information about the HAB-OFS bulletin contributors and the data they provide is available in APPENDIX III, the HAB Bulletin Guide at http://tidesandcurrents.noaa.gov/hab/habfs_bulletin_guide.pdf and at http://tidesandcurrents.noaa.gov/hab/contributors.html.

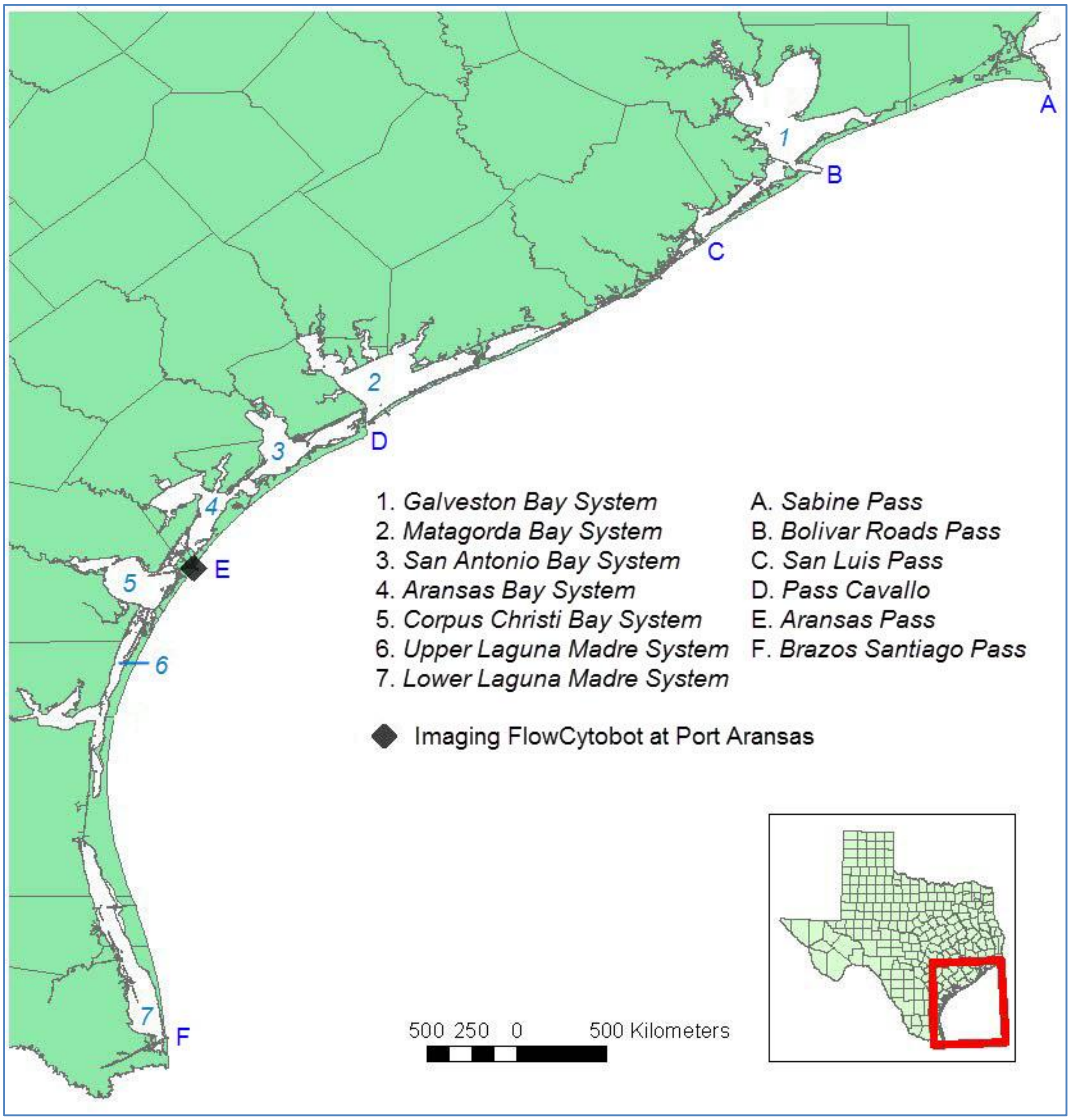


Figure 1. Map of Texas illustrating the major bay systems and coastal areas covered by the HAB-OFS. The bay regions and passes are commonly referenced as landmarks in the bulletins. The map also notes the location of Texas A&M University's Imaging FlowCytobot stationed in Aransas Pass at the University of Texas Marine Science Institute near Port Aransas.

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

1) The **HAB bulletins** provided a detailed scientific analysis of satellite ocean color imagery, water samples and health reports, meteorological and oceanographic data, and included all relevant forecasts. Each bulletin was disseminated via email to registered coastal resource managers, academics, and public health officials with an email subject line indicating the priority level of the bulletin for consideration by managers: low, medium, or high (see Table 1).

2) The **public conditions reports**, a subset of the HAB bulletins, provided information about the presence or absence of a HAB of *K. brevis* including a general description of the geographic region affected, forecasts of associated respiratory irritation, and any recent observations of respiratory irritation, dead fish, or discolored water. The conditions reports were available on the HAB-OFS website at http://tidesandcurrents.noaa.gov/hab immediately following bulletin dissemination. Beginning in 2012, these reports were also made available through the NOAA HAB (Red Tide) Watch Facebook Page at https://www.facebook.com/Habredtidewatchnoaagov.

Both products were routinely updated twice weekly on Mondays and Thursdays (or the day following a federal holiday) during HAB events and once weekly during inactive periods. The dissemination of unscheduled supplemental bulletins or conditions updates was also necessary when new data was received that indicated an increase in bloom extent, intensity, or the forecasted level of associated respiratory irritation.

Table 1. 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	• Inactive bloom • Resource managers may decide that no new action is necessary
Medium (No Change)	• 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)	• Active bloom, with recent changes in bloom conditions. Examples: ○ 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 levels of 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 (Red Tide) Watch Facebook Page. Frequently inquiries pertained to the present and future bloom conditions or potential impacts at specific locations and times to enable event planning. Inquiries received by the HAB-OFS also sought 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*.

Over the course of the assessment period, in order to manage the workload with the backups needed for operations, the number of analysts trained to support the Texas HAB-OFS increased from two to six, similar to the Florida HAB-OFS. In addition, maintaining and improving upon the Texas HAB-OFS required 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.

2.2 Definitions of the Forecast Types

The Texas HAB-OFS provides forecasts for three different bloom components: transport direction, transport distance, and potential level of respiratory irritation (see Table 2). Transport direction and distance are estimated by using the observed currents from TGLO/TABS, forecasted currents from the TGLO/TABS/TAMU ROMS-based hydrodynamic model, as well as the GNOME software. Transport direction is defined as the direction a bloom is likely to migrate (either north or south), and the transport distance is measured and rounded to the nearest 10km. 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 was forecasted by the HAB-OFS during the evaluation period was the potential for coastal respiratory irritation.

Respiratory irritation is forecasted in levels ranging from “very low” to “high” (in addition to “none” or “not expected”) based on wind direction and speed, as well as the nearby *K. brevis* cell concentrations identified in water samples (see Table 3 for cell concentration categories). In addition to manually collected water samples, TAMU’s Imaging FlowCytobot, located at the Port Aransas Ship Channel, was used to identify *K. brevis* cell concentrations (see APPENDIX III). The instrument is programmed to collect water samples from the channel at regular intervals and estimate the number of *K. brevis* cells per sample using automated image processing and an established classification procedure (GCOOS). The HAB-OFS received reports of the cell concentrations from TAMU via TPWD. As of the beginning of BY2013-2014, an hourly time series plot of *K. brevis* cell concentrations was made accessible as a product from the Gulf of Mexico Coastal Ocean Observing System (GCOOS), a regional association within the U.S. Integrated Ocean Observing System.

Symptoms associated with *K. brevis* include eye and respiratory irritation (coughing, sneezing, tearing, and itching) to beachgoers. 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). Refer to Table 4 for more information about the respiratory irritation levels. Due to limited spatial and temporal observations, these forecasts are made for geographic forecast regions approximately 30-60 km in length and only for coastal and bay regions because respiratory irritation levels are not well understood in open water regions (Stumpf, et al., 2009). The forecast regions currently used for Texas are defined in APPENDIX II.

Environmental variations, such as hydrodynamics, influence the forecasts that can be made and the analytical methods employed to develop the forecasts in Texas and Florida. For example, forecasts for intensification and potential for bloom formation at the coast are not made for

Texas because the models developed for Florida fail to reliably predict the accumulation of *K. brevis* at the coast. The factors that cause blooms to develop and intensify along the Texas coast are currently being investigated.

Table 2. Forecast definitions. Additional examples of the forecast statements can be found within the text of the sample bulletin in APPENDIX I.

FORECAST	DEFINITION	CATEGORIES	BASED ON	EXAMPLE STATEMENT
Transport Direction	Direction bloom is likely to migrate in relation to the coast	• North • South • No Change	• Observed, local ocean currents from TABS • Forecasted currents from TGLO/TABS/ TAMU ROMS Current Model • GNOME particle trajectory model	*"Forecast models based on predicted near-surface currents indicate a potential maximum transport of 50km south from the Port Aransas region May 5-9"*
Transport Distance	Distance bloom is likely to migrate in relation to the initial location estimated	• Rounded to the nearest 10km		
Respiratory Irritation	Potential level of respiratory irritation caused by the bloom (*forecast by region for the next 3-4 days*)	• Very low • Low • Moderate • High • None	• Forecasted wind speed and direction • Highest *K. brevis* concentration within most recent 7-10 days • Bloom proximity to shore • Validated reports of respiratory irritation at the coast associated with a bloom	***San Jose Island region**: Moderate (Th-M)*

Table 3. The categories assigned to *Karenia brevis* cell concentrations identified from water samples by state, county, and local organizations in Texas.

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

Table 4. The level of respiratory irritation forecasted and the corresponding population potentially affected.

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

2.3 Skill Assessment

2.3.1 Overview of Procedure

Bulletin forecasts were recorded and evaluated by HAB-OFS analysts 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. There were two categories of usage that counted toward total product utilization, viewing the product and applying its content to bloom response. 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,” and “comments”) on conditions report posts made on the NOAA HAB (Red Tide) Watch Facebook page, added on September 7, 2012, also counted as confirmation of product utilization. Bulletin utilization was recorded as “unconfirmed” when there was insufficient evidence. The utilization assessment was conducted for each bulletin issued.

Similarly, bulletin forecasts were evaluated using evidence from a variety of sources (see Table 5). Transport 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.

Table 5. Data and resources used to assess each forecast type included in a Texas bulletin.

FORECAST TYPE	CATEGORIES	ASSESSED BASED ON
Transport Direction	• North • South • No Change	• Visible movement of feature in satellite imagery • *In situ* samples confirm cell concentrations in new location • Reports of *K. brevis* induced respiratory irritation in a new location • GNOME particle trajectory model
Transport Distance	• Rounded to the nearest 10km	• Visible movement of feature in satellite imagery • *In situ* samples confirm cell concentrations in new location • Reports of *K. brevis* induced respiratory irritation in a new location • GNOME particle trajectory model
Respiratory Irritation	• Very low • Low • Moderate • High • None	• Reports of observed respiratory irritation (see Table 6)

Forecasts of respiratory irritation were verified based on observational data reported during the specified time period and disseminated by state agencies and research institutions. Sources of observed respiratory irritation data used for verification included public health reports and emails from reputable sources. Observed respiratory irritation was categorized by TPWD using a scale ranging from “very mild aerosols” to “high aerosols” as outlined in Table 6. In addition, TPWD may also use terms like “diminishing aerosols” or “increasing aerosols” when comparing the current conditions to those previously reported. Table 6 was then used to assess the forecasts based on the reports of observed respiratory irritation.

Bulletin forecasts for respiratory irritation and transport 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. Bulletin forecasts were categorized as “unconfirmed” when the necessary observational evidence was not available and forecast quality could not be analyzed further. With regards to respiratory irritation, when reports provided by TPWD did not record respiratory irritation, 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 respiratory irritation level of “none” were reported from alongshore and in the bay regions of Texas.

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 bloom year or 365-day HAB cycle. 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 6. During a *Karenia brevis* bloom, Texas Parks and Wildlife Department reports of observed respiratory irritation were used to validate the corresponding level of respiratory irritation forecasted for that region according to this chart. Due to the patchy nature of blooms, when respiratory irritation levels of "none" were reported, the observations could not definitively confirm that no respiratory irritation was experienced throughout the 30-60km forecast region. Therefore, forecasts were assessed as "unconfirmed" when respiratory irritation levels of "none" were reported in the forecast region.

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 mild aerosols *(only individuals with chronic respiratory conditions)*	False	False	Confirmed	False	False	False
Mild aerosols *(only sensitive individuals & those with chronic respiratory conditions)*	False	False	Unconfirmed	Confirmed	False	False
Aerosols *(general public may notice mild symptoms)*	False	False	False	False	Confirmed	Confirmed
High aerosols *(general public may notice adverse symptoms)*	False	False	False	False	Confirmed	Confirmed

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

Modifications to skill assessment procedures during the assessment period from BY2010-2014 were primarily made to the methods for forecasting and assessing respiratory irritation (see Table 7). Unlike in Florida, where sampling alongshore and offshore most of the coastline is frequent, certain regions of the Texas coast are sampled significantly less often than other regions though both may be subject to an equal number of respiratory irritation forecasts over a given HAB season. As a result, beginning during BY2011-2012, reports of respiratory irritation were reported as "respiratory irritation is possible" when infrequent sampling restricted the ability to specify an actual respiratory irritation level. These forecasts were assessed as binary events where respiratory irritation either was or was not observed rather than following the matrix in Table 6. Though this scenario continued to be possible during bulletins after BY2011-2012, analysts did not make forecasts without a corresponding irritation level after that bloom year.

In addition, respiratory irritation forecast region boundaries were defined on a map in BY2013-2014 (see APPENDIX II). Before this, the Texas HAB-OFS referenced geographically-identifiable landmarks to broadly indicate where respiratory irritation was forecast (i.e. Aransas Pass, Matagorda Peninsula). In BY2013-2014, a set of boundaries was created to define specific geographical forecast regions to maintain consistency from bulletin to bulletin.

Table 7. Changes during the evaluation period that impacted the assessment of bulletin forecasts and utilization.

Bloom Year	Effective Date	Description of Change
Respiratory Irritation Forecast		
BY2011-2012	11/15/11	Information from the Texas Parks and Wildlife Department and Red Tide Rangers (Texas Coastal Naturalist) Facebook Pages were used to assess forecasts of respiratory irritation.
BY2011-2012	11/28/11	Due to infrequent sampling, at times respiratory irritation was forecast without specifying a level.
BY2013-2014	9/16/13	Defined set of geographic boundaries, 30-60km in length, were created to standardize respiratory irritation forecast regions.
Bulletin Utilization		
BY2012-2013	9/13/12	HAB-OFS Facebook Page was used to assess bulletin utilization.

2.3.3 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.

2.3.3.1 Capability of Assessing Bulletin Utilization and Forecasts

Before beginning a more extensive evaluation of forecast quality and bulletin 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 in the Skill Assessment section (2.3) and Table 5, the assessment of bulletin utilization and forecasts was 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, transport distance, and respiratory irritation). Reliance on reports of field observations for forecasts made along sparsely populated or undeveloped stretches of coastline made assessment difficult in some cases. In order to evaluate the assessment capability, we calculated the percent of bulletins where each forecast type and utilization could be assessed.

2.3.3.2 Forecast Verification and Skill Assessment of Categorical Variables

Forecast quality was estimated for each of the following forecast types: bloom transport direction 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 statistics were calculated (Doswell, Davies-Jones, & Keller, 1990). Excluding the distance component of the transport forecasts, all of the forecasts included in the Texas HAB bulletins were binary, i.e. the predicted event was observed to either occur or not occur. For these forecasts, contingency tables were created showing the frequency of "yes" and "no" matched forecasts and observations (see Table 8). In reference to Table 8, 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. D represents the number of "correct rejections" or the number of times an event was correctly forecast to not occur. B represents the number of "false alarms" or the number of events that were forecasted, but not observed. 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 8. 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**	*Hit (A)*	*False Alarm (B)*	*Forecast (A+B)*
	No	*Miss (C)*	*Correct Rejection (D)*	*Not Forecast (C+D)*
	Marginal Total	*Observed (A+C)*	*Not Observed (B+D)*	*Sum Total (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 8):

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

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 8):

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

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 8):

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

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 8):

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

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 8):

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

where a score of one indicates no bias, while 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 8), the Heidke skill score is calculated as:

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

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).

2.3.3.3 Forecast Verification and Assessment of Continuous Variables

Transport distance is a continuous variable so different statistics were used to evaluate transport distance forecasts than for forecasts of categorical variables. The transport distance forecasted was compared to the distance observed.

Forecast reliability, or bias, was estimated by calculating the Mean Error (ME) as follows:

$$ME=(1/n)\ \Sigma\ (F_i-O_i) \qquad \text{[range: } -\infty \text{ to } \infty\text{]} \qquad (7)$$

where 0 is a perfect score, n is the sample size, F_i is the forecasted value and O_i is the observed value. A negative value indicated that transport distance was under-forecast, while a positive value indicated that it was over-forecast.

To estimate the accuracy of the set of transport distance forecasts, the average magnitude of the errors was determined by calculating the Mean Absolute Error (MAE) as follows:

$$MAE=(1/n)\ \Sigma\ |F_i-O_i| \qquad \text{[range: 0 to } \infty\text{]} \qquad (8)$$

where 0 is a perfect score, n is the sample size, F_i is the forecasted value and O_i is the observed value. Smaller values of MAE are more accurate. Since the MAE does not distinguish between positive and negative magnitudes, it was compared to the ME.

The Root Mean Square Error, an estimate of the average magnitude of errors weighted according to the square of the error was also calculated as follows:

$$RMSE=\sqrt{[(1/n)\ \Sigma\ (F_i-O_i)^2]} \qquad \text{[range: 0 to } \infty\text{]} \qquad (9)$$

where 0 is a perfect score, n is the sample size, F_i is the forecasted value and O_i is the observed value. The RMSE is very sensitive to small sample sizes and outliers so that was considered when interpreting the results.

The variation in the errors in the set of transport distance forecasts, or the spread of values around the average, was estimated by comparing the MAE and RMSE. The greater the difference between MAE and RMSE, the greater the variance in the individual errors in the set.

2.3.3.4 Bulletin Utilization

A successful forecast system is one that not only produces high quality forecasts, but also one that is well-used by its intended audience. 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 BY 2012-2013, the NOAA HAB (Red Tide) Watch Facebook Page was created to better disseminate public conditions reports and engage with the general public and bulletin subscribers. Interaction (likes/shares/comments) with Texas 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 bloom year and priority level.

RESULTS

3.1 Summary of *Karenia brevis* Events

From the time the HAB-OFS was transitioned to operations in the western Gulf of Mexico on October 1, 2010 to the end of the fourth BY on April 30, 2014, a total of 219 bulletins and 6 supplemental bulletins and/or conditions updates were issued, containing 305 forecasts (see Table 9). There were three separate *K. brevis* events during this time, one each bloom year from BY2011 to BY2014. There were no *K. brevis* events during BY2010-2011. All three *K. brevis* events were first identified by water samples collected in the field under the coordination of TPWD. These *K. brevis* events varied both in geographic extent and duration. The BY2011-2012 bloom was both the longest lasting bloom (151 days) and the largest bloom in terms of geographic extent (affecting over 75% of the Texas coastline). The BY2012-2013 and BY2013-2014 blooms lasted for approximately 32 and 35 days respectively (see Table 10). The BY2012-2013 bloom was the smallest in terms of geographic extent; while the BY2013-2014 bloom affected approximately 40-45% of the Texas coastline. All three blooms were patchy in nature covering multiple disconnected portions of the Texas coastline. Maps of the monthly *K. brevis* samples collected during BY2011-2014 are shown in Figures Figure 2-Figure 4.

Table 9. The number of HAB-OFS products issued during the 2010 to 2014 bloom years.

Bloom Year	# of HAB-OFS Products Issued	
	# of Scheduled Bulletins	*# of Supplemental/ Conditions Updates*
10/1/10 to 4/30/11	30	0
5/1/11 to 4/30/12	74	3
5/1/12 to 4/30/13	57	2
5/1/13 to 4/30/14	58	1

Table 10. Estimate of the duration (in days) of *Karenia brevis* bloom events detected during the 2010 to 2014 bloom years.

Bloom Year	# of *K. brevis* Events Detected	Bloom Duration (in days)
10/1/10 to 4/30/11	0	0
5/1/11 to 4/30/12	1	151
5/1/12 to 4/30/13	1	32
5/1/13 to 4/30/14	1	35

3.1.1 Bloom Year: 2010-2011

The HAB-OFS was transitioned to operations in the western Gulf of Mexico on October 1, 2010. As a result, the analysis for BY2010-2011 spans from October 1, 2010 to April 30, 2011 rather than from May 1, 2010. No *K. brevis* blooms occurred during BY2010-2011. A total of 30 bulletins were disseminated during this bloom year.

3.1.2 Bloom Year: 2011-2012

The first Texas bloom monitored by the HAB-OFS occurred in the 2011-2012 bloom year and it was detected by samples collected in the field by TPWD (see APPENDIX III). A total of 41 bulletins, two supplemental bulletins and one conditions update were disseminated during the bloom. Maps of the monthly *K. brevis* samples collected during this period are shown in Figure 2 along with a key to cell concentration categories.

The BY2011-2012 bloom was first identified in the Brownsville Ship Channel region at the southern end of the Texas coastline from samples collected on September 14, 2011. TPWD had received reports of stressed and dead fish in the region. During their investigation, they encountered dead fish and discolored water and experienced respiratory irritation. The water samples indicated “high” cell concentrations (>1,000,000 cells/L) and further investigation revealed that the fish kill extended as far as 7 miles along the ship channel. One week later, multiple “high” cell concentrations (>1,000,000 cells/L) were also identified in the San Luis Pass region of southern Galveston Island.

The BY2011-2012 bloom was one of the state’s largest blooms in terms of geographic extent. By late October, bloom-level *K. brevis* concentrations had been identified throughout the Texas coast including the Galveston/Freeport area, alongshore the Matagorda Peninsula and within Matagorda Bay, in the Aransas Pass area and within Corpus Christi Bay, alongshore the Padre Island National Seashore and the South Padre Island area, within the lower Laguna Madre, and within the Brownsville Ship Channel. *K. brevis* concentrations ranging from “medium” to “high” (>100,000 cells/L) persisted at least through December in most regions and began dissipating in late December 2011 and early January 2012.

Of the three blooms discussed in this report, the BY2011-2012 bloom was by far the longest lasting (see Table 10) and it was the longest lasting bloom on record for the state of Texas (NOAA, 2012; Sherman, 2011). The bloom lingered through early February and completely dissipated by February 13, 2012. Affecting over 75% of the Texas coastline and bay areas, the BY2011-2012 bloom resulted in the deaths of over 4 million fish and over 100 marine birds, numerous respiratory irritation reports, and the closure of all oyster harvesting along the Texas coast for several months (Red tide toxin found in dead ducks, 2012). According to TPWD, preliminary costs to the shellfish industry alone, due to bloom-related closures, amounted to at least 7 million dollars (NOAA, 2012; Pack, 2012).

3.1.3 Bloom Year: 2012-2013

The only bloom of BY2012-2013 was first detected via sampling alongshore Galveston Island and from the mouth of Galveston Bay at the Bolivar Roads Pass region by TPWD on August 12, 2012. As in the BY2011-2012 bloom, the sampling was conducted during an investigation of reports of dead fish and respiratory irritation in the region. Samples indicated up to “medium” concentrations (>100,000 to 1,000,000 cells/L) of *K. brevis*.

The BY2012-2013 bloom was initially confined to Galveston Bay and along- and offshore Galveston Island and the Bolivar Peninsula. Approximately 11 days later, on August 23, 2012, new samples indicated “low a” concentrations (>10,000 to < 50,000 cells/L) of *K. brevis* near the Texas/Mexico border at Boca Chica Beach. In both regions, subsequent samples indicated that the patchy bloom was dissipating. No reports of dead fish were received after the initial report that led to the identification of the bloom. By September 13, 2012, the short-lived (32 days; see Table 10) and patchy bloom had dissipated. A total of nine bulletins were disseminated during the bloom. Maps of the monthly *K. brevis* samples collected during this period are shown in Figure 3, along with a key to cell concentration categories.

3.1.4 Bloom Year: 2013-2014

As in the previous two blooms in Texas, TPWD was prompted to collect samples that identified the existence of a new bloom after receiving reports of respiratory irritation. On August 27 and 28, 2013, one “medium” concentration (>100,000 to 1,000,000 cells/L) was found at Surfside Beach, near Freeport, TX, and multiple “low a” and “low b” concentrations (>10,000 to < 50,000 cells/L and 50,000 to 100,000 cells/L, respectively) were found within and at the mouth of Galveston Bay. While performing the sampling, TPWD personnel also reported discolored water at various locations. Analysts working on the first bulletin for the BY2013-2014 event highlighted an anomalous patch of elevated to high levels of chlorophyll (2 to >20 µg/L) extending from Galveston Island westward to East Matagorda Bay. Although sampling had not yet confirmed the presence of the bloom that far east, the satellite imagery was clear.

Two days later, a “low a” concentration (>10,000 to < 50,000 cells/L) was identified in Sargent Beach located near the East Matagorda Bay. Additionally, the highest sample concentration identified for this event, a “high” concentration of approximately 1.3 million cells/L, was identified at the northeast end of the Galveston Yacht Basin on August 29, 2013. Multiple “very low a” and “low a” concentrations (>1,000 to <5,000 cells/L and >10,000 to <50,000 cells/L, respectively) were also identified in Galveston Bay.

By early September, samples confirmed that the bloom was already dissipating in the Galveston Island, Galveston Bay and Bolivar Peninsula region. Cell concentrations ranged from not present to “low b” (from 0 to 100,000 cells/L) in samples collected on September 2 and 3, 2013. At the same time, samples indicated *K. brevis* cell concentrations ranging from not present to “medium” (from 0 to 1,000,000 cells/L) along the Padre Island National Seashore. While not as widespread as the BY2011-2012 *K. brevis* event, the presence of the bloom along the Padre Island National Seashore meant that this event was affecting approximately 40-45% of the Texas coastline.

In mid-September, cell concentrations continued to confirm that the bloom was dissipating in the Galveston Island, Galveston Bay and Bolivar Peninsula region. At the same time, samples were indicating that *K. brevis* was no longer present in the Padre Island National Seashore. The only increase in cell concentrations was seen alongshore the Port Aransas/Mustang Island region.

By October 3, 2013, the BY2013-2014 event had completely dissipated. This bloom was just 3 days longer than that from the previous year (see Table 10). Other than the first week of the bloom, no additional reports of impacts, including respiratory irritation, dead fish, or discolored water, were received throughout the remainder of the bloom. A total of 10 bulletins and one conditions update were disseminated during the bloom. Maps of the monthly *K. brevis* samples collected during this period are shown in Figure 4, along with a key to cell concentration categories.

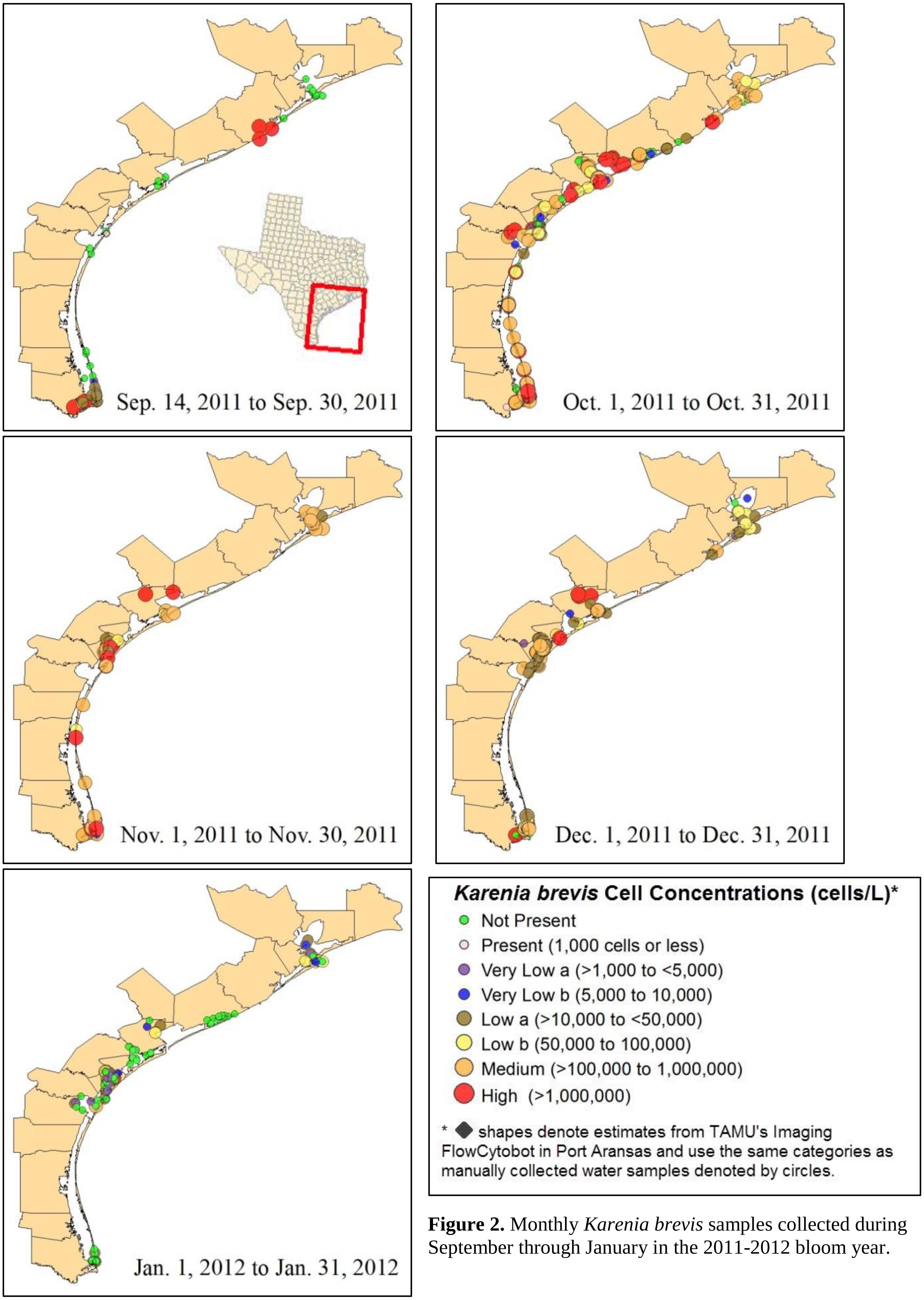


Figure 2. Monthly *Karenia brevis* samples collected during September through January in the 2011-2012 bloom year.

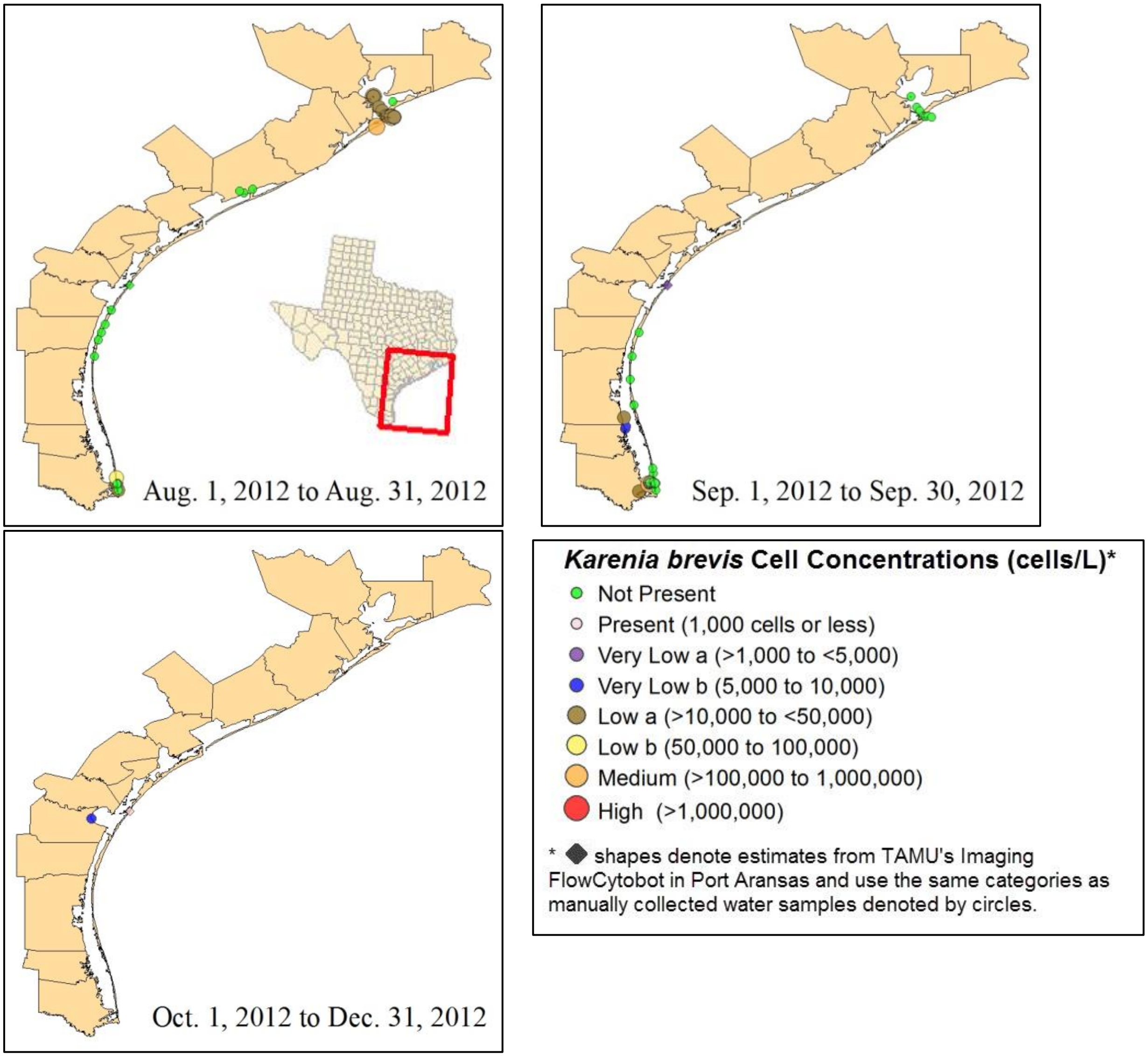


Figure 3. Monthly *Karenia brevis* samples collected during August through December in the 2012-2013 bloom year.

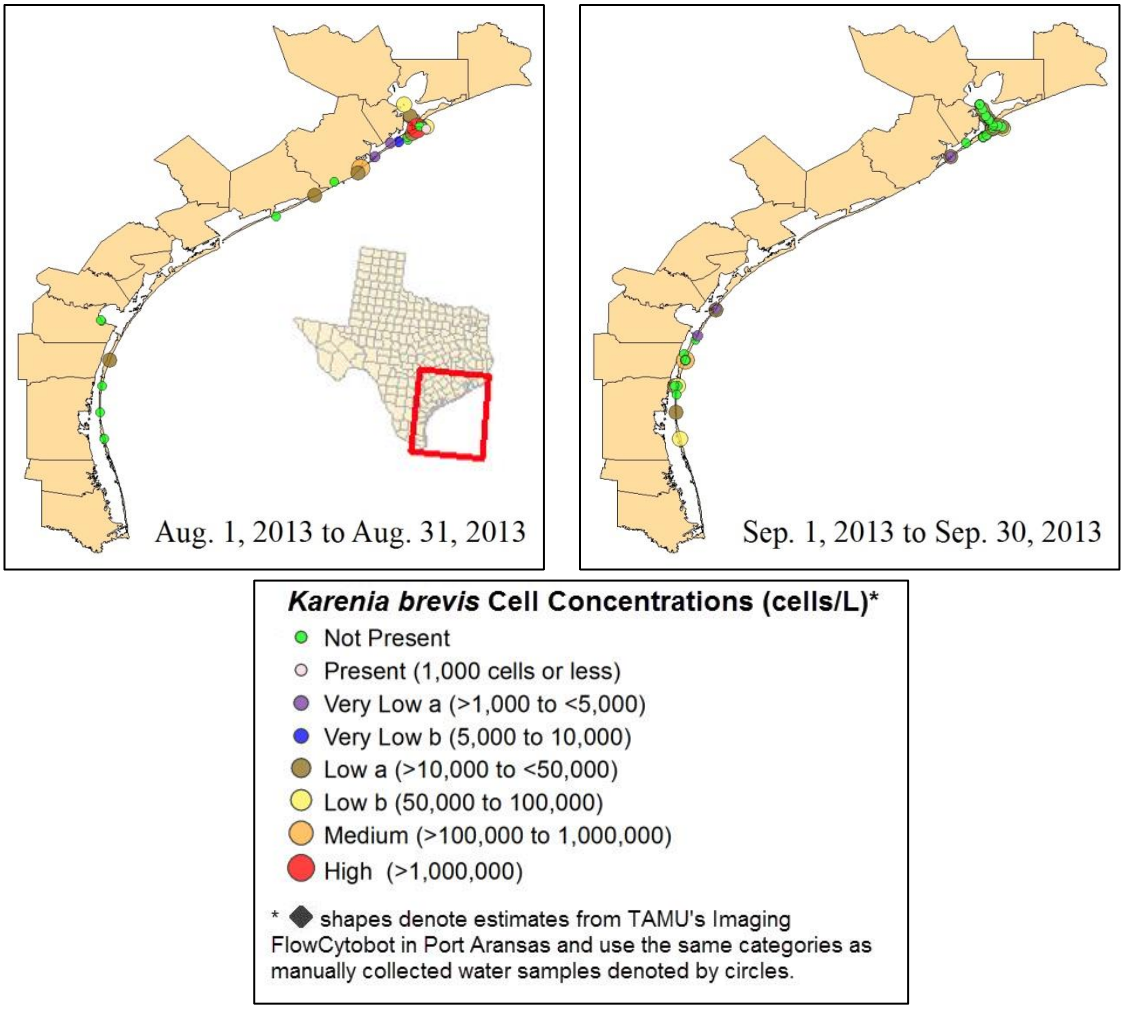
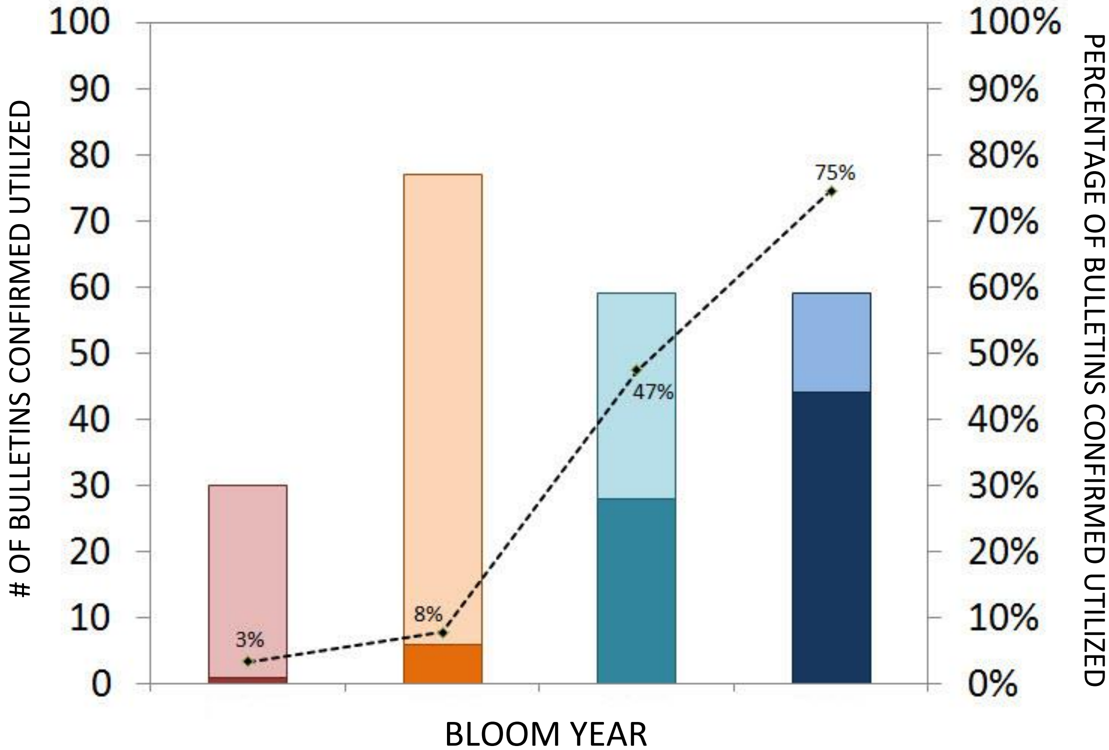


Figure 4. Monthly *Karenia brevis* samples collected during August and September in the 2012-2013 bloom year.

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 measurements of viewership and interactions were used to determine if the posted conditions reports were used. Overall the proportion of total bulletins with confirmed utilization increased significantly over the four bloom years covered in this assessment, with 3.33% confirmed utilized during BY2010-2011 and 74.6% confirmed utilized during BY2013-2014 (see Figure 5). The difference was due to the launch of the HAB-OFS Facebook Page and the usage of Facebook metrics to assess utilization. In fact, during BY2012-2014, all cases of confirmed bulletin utilization were verified using Facebook metrics.



Unconfirmed Bulletin Utilization	10/1/10 to 4/30/11	5/1/11 to 4/30/12	5/1/12 to 4/30/13	5/1/13 to 4/30/14
Confirmed Bulletin Utilization	10/1/10 to 4/30/11	5/1/11 to 4/30/12	5/1/12 to 4/30/13	5/1/13 to 4/30/14

Figure 5. Number of bulletins that were confirmed utilized and percentage of bulletins confirmed utilized out of the total number of bulletins for the 2010 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 (Table 1). Utilization of each bulletin varied according to the priority level assigned to the bulletin. Overall, 35.1% of all bulletins were confirmed utilized during BY2010-2014 and, as with overall utilization by year, utilization of each priority level of bulletin has also shown an increase since bulletins were first issued in October 2010 (see Figure 6). Low priority bulletins were the most frequently issued and confirmed utilized with 43.8% confirmed utilized during BY2010-2014. The utilization of low priority bulletins increased each year from 3.00% confirmed utilized during BY2010-2011 to 81.3% confirmed utilized in BY2013-2014. No high or medium priority bulletins were disseminated during BY2010-2011 because there was no bloom during that time. High priority bulletins were the next most frequently confirmed utilized with 20.0% confirmed utilized on average from BY2010-2014. High priority bulletins showed fairly consistent utilization in BY2011-2012 and BY2013-2014, with 21.4% and 25.0% confirmed utilized during these years, respectively. Only two high priority bulletins were issued in BY2012-2013, neither of which were confirmed utilized. Medium priority bulletins were the least frequently confirmed utilized with 11.1% confirmed utilized on average during BY2010-2014. While the greatest number of medium priority bulletins were disseminated in BY2011-2012 (31 bulletins), only 3.23% were confirmed utilized. This may be due to underreporting during that year. Seven medium priority bulletins were issued in each of the subsequent bloom years, with 0.00% confirmed utilized in BY2012-2013 and 57.1% confirmed utilized during BY2013-2014.

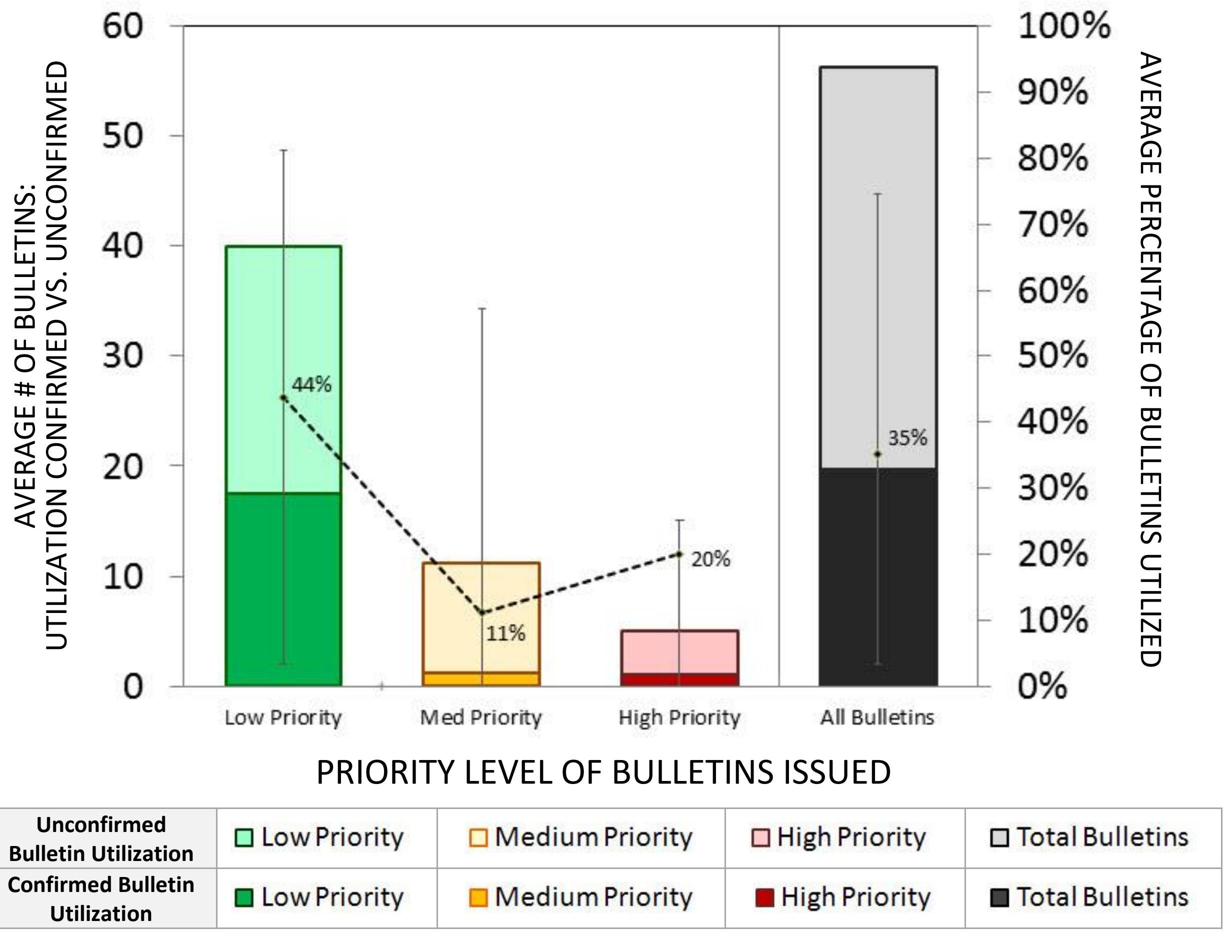
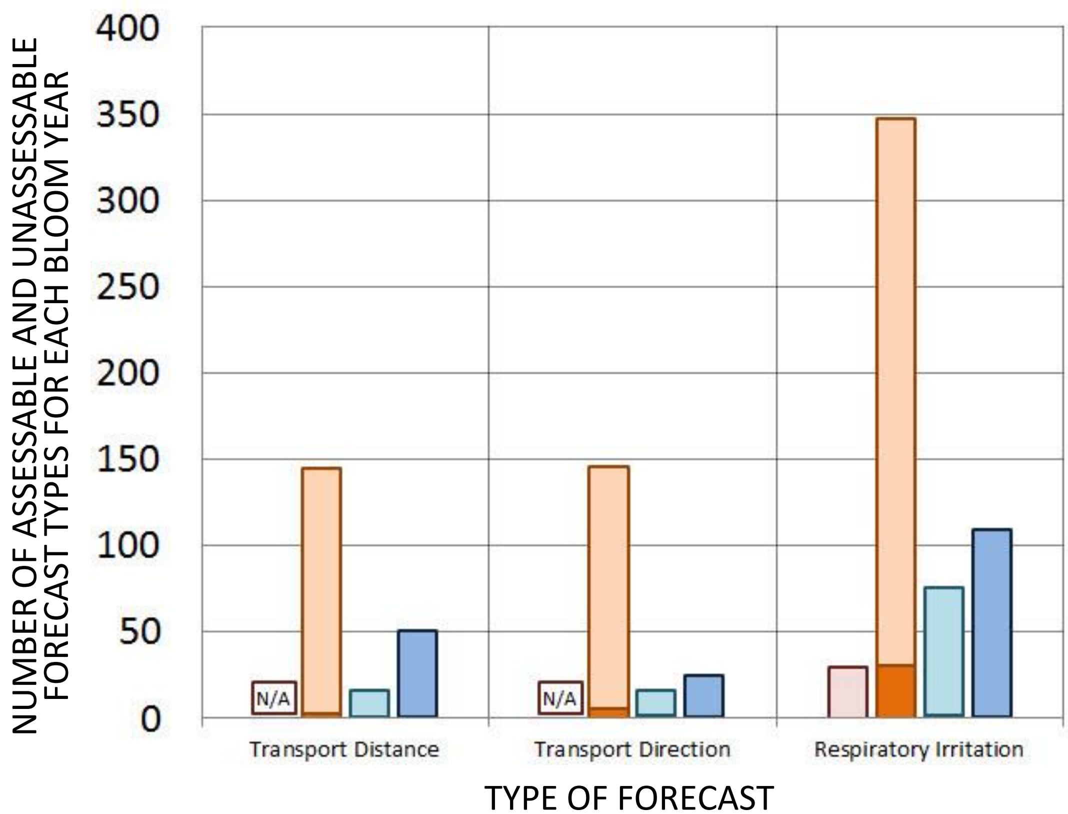


Figure 6. Average number of bulletins with utilization confirmed for each priority level and average percentage of bulletins utilized over the 2010 to 2014 bloom years. A priority level is assigned to each bulletin based on the need for management response.

Following the launch of the HAB-OFS Facebook Page on September 16, 2012, posts from low priority bulletins were the most highly utilized. Of the 98 low priority bulletin posts, 92 were confirmed utilized (93.9%). By comparison, 8 of the 15 medium and high priority bulletin posts were confirmed utilized (53.3%). Facebook users also responded to low priority bulletins with more active interactions, “liking,” “sharing,” or “commenting” on 91 of the 98 posts (92.9%). Only one of the interactions that confirmed utilization of a low priority bulletin post was through clicking on the post without liking, sharing, or commenting. Medium and high priority bulletin posts were liked, shared, or commented on 6 times out of the 15 total posts (40.0%). In two cases, medium and high priority posts were confirmed utilized based on clicking on the post without liking, sharing, or commenting.

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. The forecast was categorized as unconfirmed when the necessary observational evidence was not available and forecast quality could not be assessed. Since large stretches of the Texas coast are inaccessible or not frequently visited, observational evidence required for validation was not always available, leading to a large variation in assessment capability between bloom years. TPWD also does not sample routinely, which leads to data gaps. Furthermore, bloom duration and intensity varied from year to year, influencing the number of forecasts that could be assessed (see Figure 7).



# of Unassessable Forecasts	10/1/10 to 4/30/11	5/1/11 to 4/30/12	5/1/12 to 4/30/13	5/1/13 to 4/30/14
# of Assessable Forecasts	10/1/10 to 4/30/11	5/1/11 to 4/30/12	5/1/12 to 4/30/13	5/1/13 to 4/30/14

Figure 7. Number of assessable and unassessable forecasts during bloom years from 2010 to 2014. The assessment of forecasts was dependent on the availability of reliable observational data from reputable sources.

Note: There was no *Karenia brevis* bloom during BY2010-2011 so no transport distance or transport direction forecasts were issued.

There was no bloom along- or offshore the Texas coast during the BY2010-2011 interval and therefore, no assessable forecasts were made, although 30 forecasts for the “none” level of respiratory irritation were issued. During the BY2011-2014 interval, a total of 886 individual forecasts were made for the three forecast types (transport direction, transport distance, respiratory irritation).

During BY2011-2012, the highest percentage of assessable forecasts was for respiratory irritation (8.93%, 31 of 347 total forecasts) followed by transport direction (4.11%, 6 of 146 total forecasts) and transport distance (2.07%, 3 of 145 total forecasts). During BY2012-2013 and BY2013-2014, the highest percentages of assessable forecasts were for transport direction (12.5%, 2 of 16 total forecasts in BY2012-2013 and 4.00%, 1 of 25 total forecasts in BY2013-2014). This was followed by transport distance (6.25%, 1 of 16 total forecasts in BY2012-2013 and 1.96%, 1 of 51 total forecasts in BY2013-2014). In both years, the lowest percentages of assessable forecasts were for respiratory irritation (2.63%, 2 of 76 total forecasts in BY2012-2013 and 0.920%, 1 of 109 total forecasts in BY2013-2014).

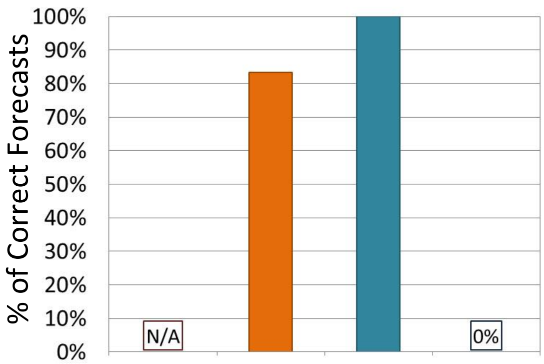
From BY2011-2014, the proportion of all assessable forecasts were distributed unevenly (3.80-90.9%) between bloom years. Of all of the forecasts issued for blooms, 70.3% were made during BY2011-2012, but 90.9% of all assessable forecasts were issued in that year. Comparatively, only 5.20% and 3.90% of all assessable forecasts were issued for BY2012-2013 and BY2013-2014, respectively.

3.4 Accuracy of Categorical Forecasts

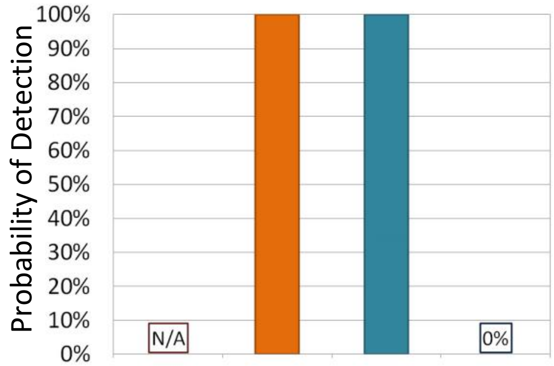
Forecast accuracy was estimated for each of the categorical forecasts: transport direction and respiratory irritation. Accuracy was also estimated for the individual respiratory irritation levels: “none,” “very low,” “low,” “medium,” and “high.” The four statistics used to estimate forecast accuracy were proportion correct, probability of detection (POD), threat score (TS), and false alarm ratio (FAR) (see Section 2.3 for definitions).

3.4.1 Transport Direction

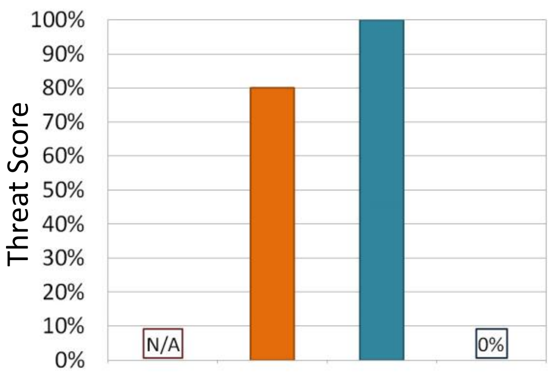
Figure 8 shows that the transport direction forecast accuracy varied from BY2011-2014. Because there was no bloom activity in BY2010-2011, no transport direction forecasts were issued. Of the assessable forecasts for BY2011-2012, transport direction forecasts were consistently accurate, with a high proportion correct (83.3%), high probability of detection (1.00), high threat score (0.800), and relatively low false alarm ratio (0.200). The assessable forecasts for BY2012-2013 also showed high accuracy, with perfect results for each of the accuracy statistics (proportion correct, probability of detection, threat score, and false alarm ratio); however, these results were estimated from a very small sample size (n=2) so the results may or may not be representative of the overall performance of the forecasts issued. Statistics calculated for BY2013-2014 were based on only one assessable forecast (n=1) of “no change” which was verified to be false (or incorrect), producing a proportion correct, probability of detection, and threat score of 0.00, and an undefined false alarm ratio.



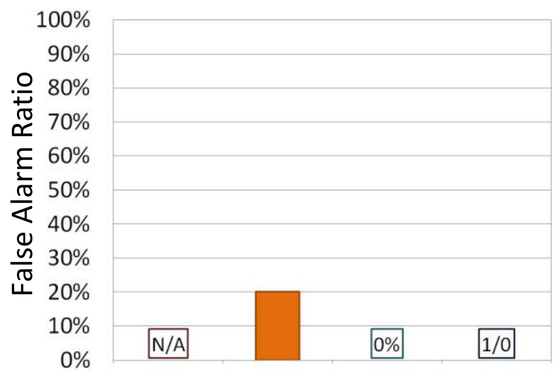
a) Proportion Correct: the number of correct transport direction forecasts and correct rejections compared to the total number of transport directions forecasts issued.



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.

Number of Assessable Forecasts (n) for each Bloom Year				
Bloom Years	■ 10/1/10 to 4/30/11	■ 5/1/11 to 4/30/12	■ 5/1/12 to 4/30/13	■ 5/1/13 to 4/30/14
n	0	6	2	1

Figure 8. Accuracy of transport direction forecasts issued during the 2010 to 2014 bloom years. There was no *Karenia brevis* bloom during BY2010-2011 and no transport direction forecasts were issued as indicated by values of N/A.

Note: Values of 1/0 indicate that the denominator of the calculation was zero (see Section 2.3 for an explanation of the statistical analyses used).

3.4.2 Respiratory Irritation

3.4.2.1 All Levels of Respiratory Irritation

During BY2010-2014, the accuracy of respiratory irritation forecasts was highly variable. While respiratory irritation forecasts issued during BY2011-2012 were consistently accurate with a high proportion correct (92.9%), during BY2013-2014 only one assessable forecast was issued (n=1), which was confirmed false (incorrect), resulting in a proportion correct of 0.00%. Likewise, in BY2012-2013, the proportion correct was relatively low (50.0%), but with a low number of assessable forecasts (n=2). There was no bloom activity during BY2010-2011, so no respiratory irritation forecasts were issued during that time.

3.4.2.2 No Level of Respiratory Irritation (None)

Respiratory irritation forecasts were validated based on reports of coastal field observations. Due to the patchy nature of blooms, respiratory irritation typically does not affect an entire forecast area and observation reports are limited. Thus, our method of assessment did not allow us to verify that no irritation was observed throughout an entire forecast area during the forecast period. In both BY2012-2013 and BY2013-2014, there was one “none” forecast that was confirmed false (i.e., a level of respiratory irritation other than “none” was observed) during each bloom year. Each of these years had a low number of assessable forecasts. Out of two assessable respiratory irritation forecasts issued during BY2012-2013, this resulted in a proportion correct of 50.0%, undefined probability of detection, low threat score (0.00), and high false alarm ratio (1.00). Similarly, because this was the only assessable forecast issued in BY2013-2014 and it was confirmed false (incorrect), this resulted in a proportion correct of 0%, undefined probability of detection, low threat score (0.00), and high false alarm ratio (1.00).

3.4.2.3 Very Low Levels of Respiratory Irritation

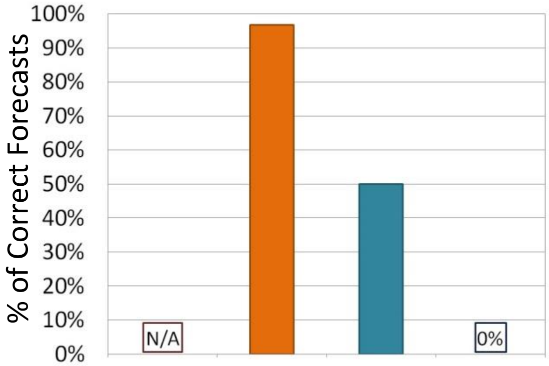
There were no assessable “very low” respiratory irritation forecasts issued or confirmed observations of “very low” levels of respiratory irritation during BY2011-2014.

3.4.2.4 Low Levels of Respiratory Irritation

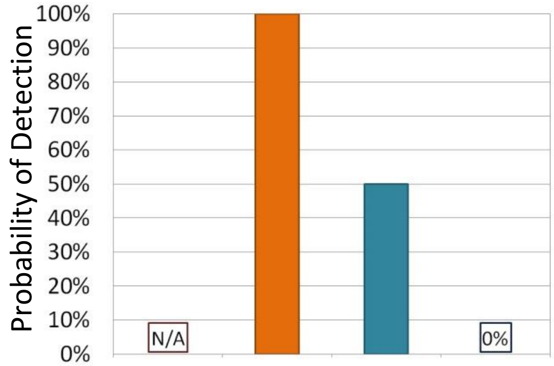
During BY2011-2012, only three “low” respiratory irritation forecasts were issued with variable accuracy; while the forecasts resulted in a high proportion correct (93.6%) and perfect false alarm ratio (0.00), the probability of detection and threat scores were relatively low (33.3% for each). This is the result of a combination of confirmed correct “low” respiratory irritation forecasts and observed “low” respiratory irritation that was not correctly forecast. There were no assessable “low” respiratory irritation forecasts issued or confirmed observed during BY2012-2014.

3.4.2.5 Moderate Levels of Respiratory Irritation

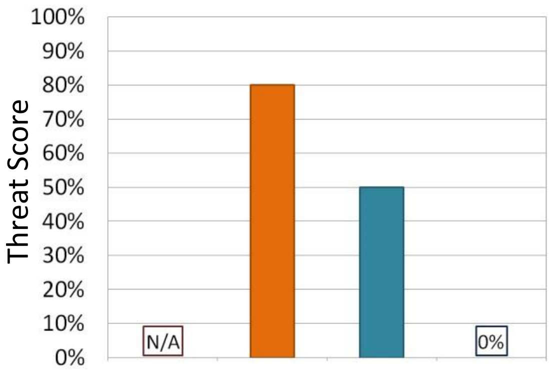
Figure 9 shows that “moderate” respiratory irritation forecasts were the most accurate during BY2011-2012, with a relatively high proportion correct (96.8%), probability of detection (1.00), and threat score (0.800), and low false alarm ratio (0.200). During BY2012-2013 only two “moderate” respiratory irritation forecasts were issued, showing less accuracy with a low proportion correct (50.0%), probability of detection (0.500), and threat score (0.500), but also a low false alarm ratio (0.00). While no “moderate” respiratory irritation forecasts were issued during BY2013-2014, there was one report of observed “moderate” respiratory irritation, resulting in a low proportion correct (0.00%), probability of detection (0.00), and threat score (0.00). The false alarm ratio was undefined since there were no assessable “moderate” forecasts issued.



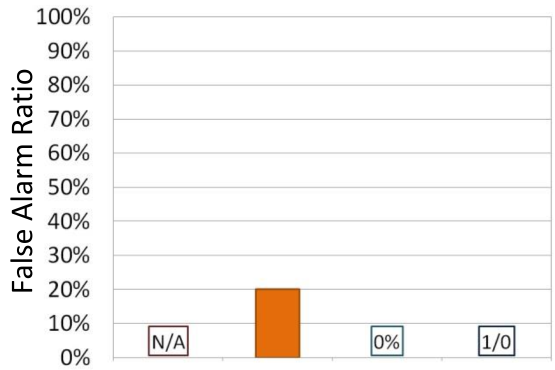
a) Proportion Correct: the number of correct moderate respiratory irritation forecasts and correct rejections compared to the total number of moderate respiratory irritation forecasts issued.



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



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.

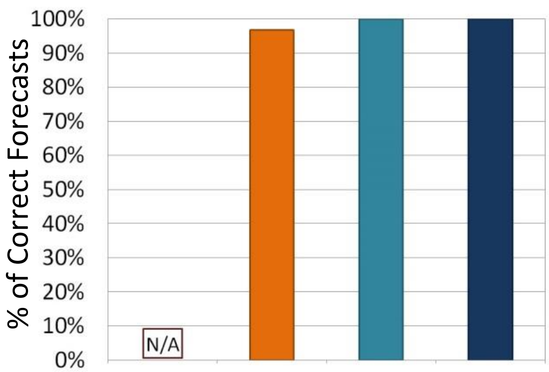
Number of Assessable Forecasts (n) for each Bloom Year				
Bloom Years	10/1/10 to 4/30/11	5/1/11 to 4/30/12	5/1/12 to 4/30/13	5/1/13 to 4/30/14
n	0	5	1	0

Figure 9. Accuracy of moderate respiratory irritation forecasts issued during the 2010 to 2014 bloom years (BY). There was no *Karenia brevis* bloom during BY2010-2011 and no moderate respiratory irritation forecasts were issued as indicated by values of N/A.

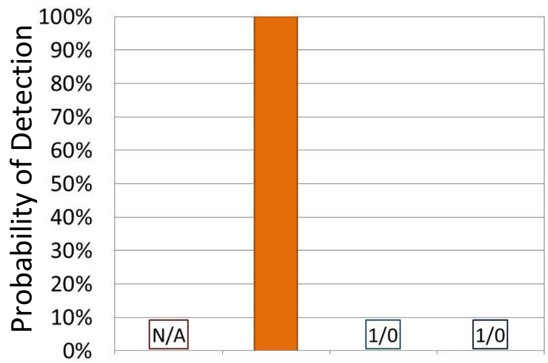
Note: Values of 1/0 indicate that the denominator of the calculation was zero (see Section 2.3 for an explanation of the statistical analyses used).

3.4.2.6 High Levels of Respiratory Irritation

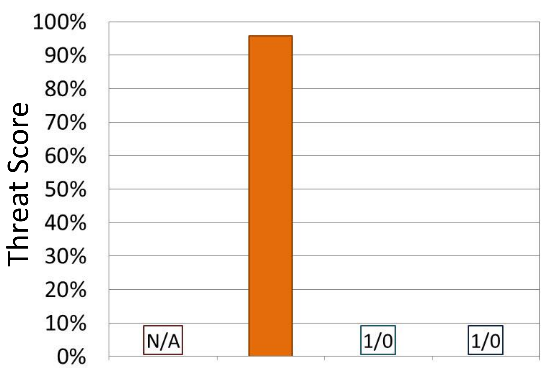
As demonstrated in Figure 10, the 22 assessable “high” respiratory irritation forecasts issued during BY2011-2012 were very accurate, with a high proportion correct (96.7%), probability of detection (1.00), threat score (0.957), and very low false alarm rate (0.043). There were no assessable “high” respiratory irritation forecasts issued or confirmed observed during BY2012-2014.



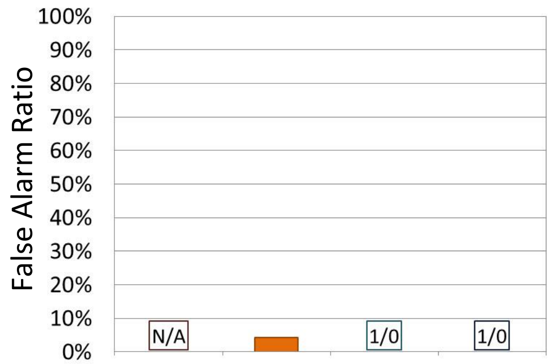
a) Proportion Correct: the number of correct high respiratory irritation forecasts and correct rejections compared to the total number of high respiratory irritation forecasts issued.



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



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



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

Number of Assessable Forecasts (n) for each Bloom Year				
Bloom Years	10/1/10 to 4/30/11	5/1/11 to 4/30/12	5/1/12 to 4/30/13	5/1/13 to 4/30/14
n	0	23	0	0

Figure 10. Accuracy of "high" respiratory irritation forecasts during the 2010 to 2014 bloom years (BY). There was no *Karenia brevis* bloom during BY2010-2011 and no high respiratory irritation forecasts were issued as indicated by values of N/A.

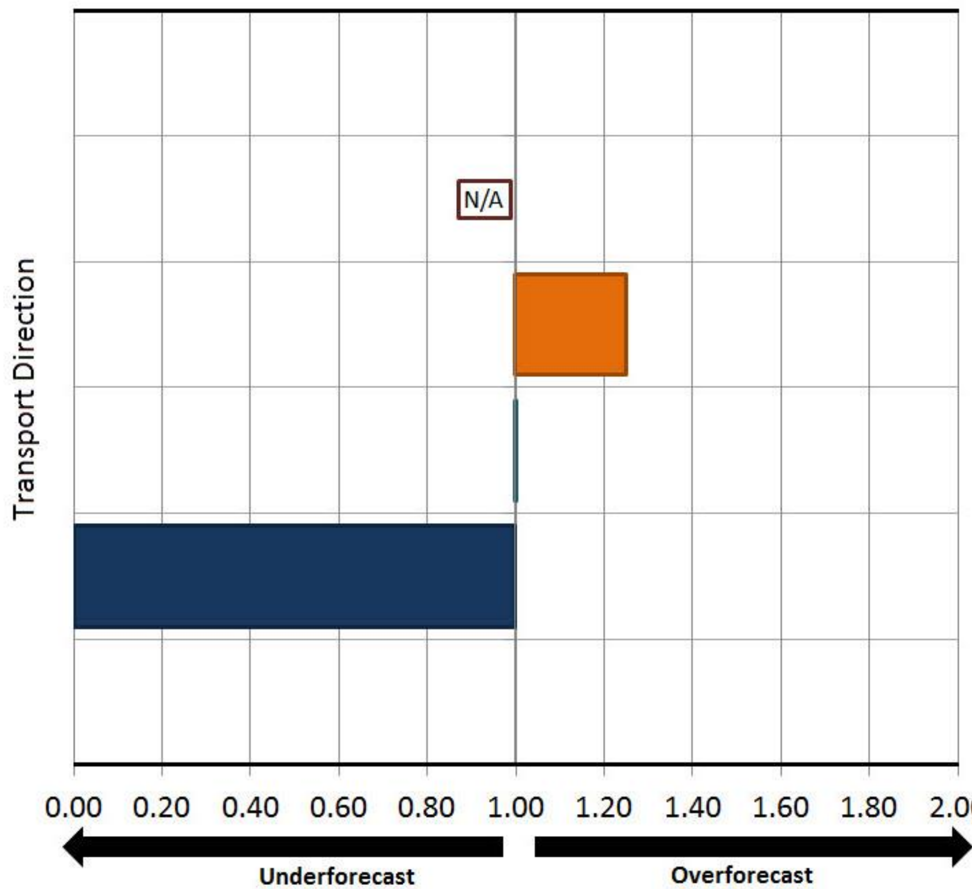
Note: Values of 1/0 indicate that the denominator of the calculation was zero (see Section 2.3 for an explanation of the statistical analyses used).

3.5 Reliability of Categorical Forecasts

Forecast reliability for categorical forecasts 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 categorical forecasts: transport direction (see Figure 11) and each of the individual respiratory irritation levels ranging from "no" respiratory irritation to "high" (see Figure 12).

3.5.1 Transport Direction

Figure 11 shows that there was high variability in forecast reliability for transport direction forecasts issued during BY2011-2014. Because there was no bloom activity in BY2010-2011, no transport direction forecasts were issued (N/A). There was no bias (1.00) in transport direction forecasts issued in BY2012-2013. Transport was slightly over-forecast during BY2011-2012, with a bias of 1.25, meaning bloom movement was forecasted slightly more often than it was observed. Conversely, transport was under-forecast (0.00) during BY2013-2014 meaning bloom movement was often observed when no transport was forecasted.



Number of Assessable Forecasts (n) for each Bloom Year				
Bloom Years	10/1/10 to 4/30/11	5/1/11 to 4/30/12	5/1/12 to 4/30/13	5/1/13 to 4/30/14
n	0	6	2	1

Figure 11. Forecast reliability (bias) in transport direction forecasts during the 2010 to 2014 bloom years (BY). A score of one indicates no bias, while a score 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. There was no *Karenia brevis* bloom during BY2010-2011 and no transport direction forecasts were issued as indicated by values of N/A.

3.5.2 Respiratory Irritation

3.5.2.1 No Levels of Respiratory Irritation (None)

Respiratory irritation forecasts were validated based on reports of coastal field observations. Due to the patchy nature of blooms, respiratory irritation typically does not affect an entire forecast area and observation reports are limited. Thus, our method of assessment did not allow us to verify that no respiratory irritation was observed throughout an entire forecast area during the forecast period. During the blooms in both BY2012-2013 and BY2013-2014, a “moderate” level of respiratory irritation was observed when “none” was forecasted. However, no observations of “none” were reported resulting in the bias being undefined for this forecast level in both years.

3.5.2.2 Very Low Levels of Respiratory Irritation

With no bloom during BY2010-2011 and no assessable “very low” respiratory irritation forecasts issued or observed during BY2011-2014, no bias statistics could be measured for “very low” respiratory irritation forecasts (N/A) (see Figure 12).

3.5.2.3 Low Levels of Respiratory Irritation

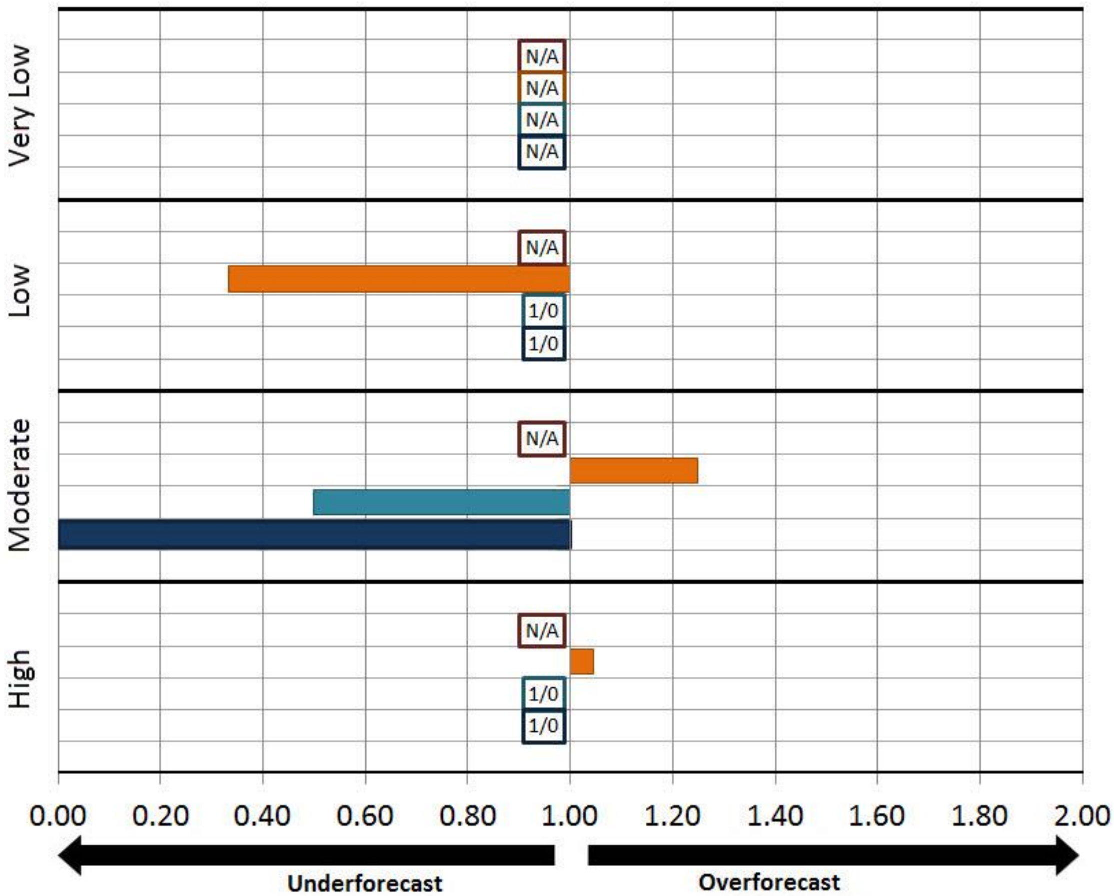
From BY2010-2014, the only assessable “low” respiratory irritation forecasts were issued during BY2011-2012 (see Figure 12). During BY2011-2012, only one assessable “low” respiratory irritation forecast was issued, but “low” respiratory irritation was observed three times, resulting in a bias of 0.330, meaning that “low” respiratory irritation forecasts were slightly under-forecast during this time.

3.5.2.4 Moderate Levels of Respiratory Irritation

As shown in Figure 12, bias results varied for “moderate” respiratory irritation forecasts issued during BY2011-2014. During BY2011-2012, the four assessable “moderate” respiratory irritation forecasts were slightly over-forecast, with a bias of 1.25, meaning forecasts for “moderate” levels of respiratory irritation were more often issued than observed. The following bloom year, BY2012-2013, respiratory irritation was under-forecast, with a bias of 0.500, meaning “moderate” levels of respiratory irritation were observed more often than forecast. This was the result of one “moderate” respiratory irritation forecast being issued and confirmed correct, and one “moderate” respiratory irritation observed, but not forecast. “Moderate” respiratory irritation was also under-forecast during BY2013-2014, with a bias of 0.00. No assessable “moderate” respiratory irritation forecasts were issued, however there was one observation of “moderate” respiratory irritation.

3.5.2.5 High Levels of Respiratory Irritation

“High” respiratory irritation was only very slightly over-forecast in BY2011-2012, with observations of “high” respiratory irritation reported for 22 out of 23 times “high” respiratory irritation was forecasted, resulting in a bias of 1.05 (see Figure 12). With no bloom during BY2010-2011 and no “high” respiratory irritation forecasts issued or confirmed observed in BY2012-2014, bias statistics were undefined for “high” respiratory irritation forecasts for these years.



Number of Assessable Forecasts (n) for each Bloom Year					
Bloom Years		10/1/10 to 4/30/11	5/1/11 to 4/30/12	5/1/12 to 4/30/13	5/1/13 to 4/30/14
Level of Respiratory Irritation	***Very Low***	0	0	0	0
	Low	0	1	0	0
	Moderate	0	5	1	0
	High	0	23	0	0

Figure 12. Forecast reliability (bias) in respiratory irritation forecasts during the 2010 to 2014 bloom years (BY). A score of one indicates no bias, while a score 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. There was no *Karenia brevis* bloom during BY2010-2011 and no respiratory irritation forecasts were issued.

Note: Values of 1/0 indicate that the denominator of the calculation was zero (see Section 2.3 for an explanation of the statistical analyses used). Values of N/A indicate that no assessable respiratory irritation forecasts were issued.

3.6 Skill of Categorical Forecasts

Forecast skill for the categorical forecasts 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, for the overall forecast of respiratory irritation (see Figure 13), and for individual respiratory irritation forecasts, ranging from "no" respiratory irritation to "high" (see Figure 14). Due to a lack of verified forecasts, no Heidke skill scores could be calculated for BY2010-2011.

3.6.1 Transport Direction

Figure 13 shows that the transport direction forecasts issued during BY2011-2012 performed much better than chance, with a Heidke skill score of 57.1%. During BY2013-2014, transport direction forecasts performed no better than random chance, with a Heidke skill score of 0.00%. However, this Heidke score was obtained during an interval where only two forecasts could be assessed so the score calculated from such a small sample size may not have been representative of the performance of all of the transport direction forecasts.

3.6.2 All Respiratory Irritation Levels

Figure 14 shows that all respiratory irritation forecasts issued during BY2011-2012 performed much better than chance, with a Heidke skill scores indicating 91.4% improvement over chance. During BY2012-2013 and BY2013-2014, respiratory irritation forecasts performed no better than random chance, with a Heidke skill score of 0.00%. However, this Heidke score was obtained during years when no more than two forecasts could be assessed.

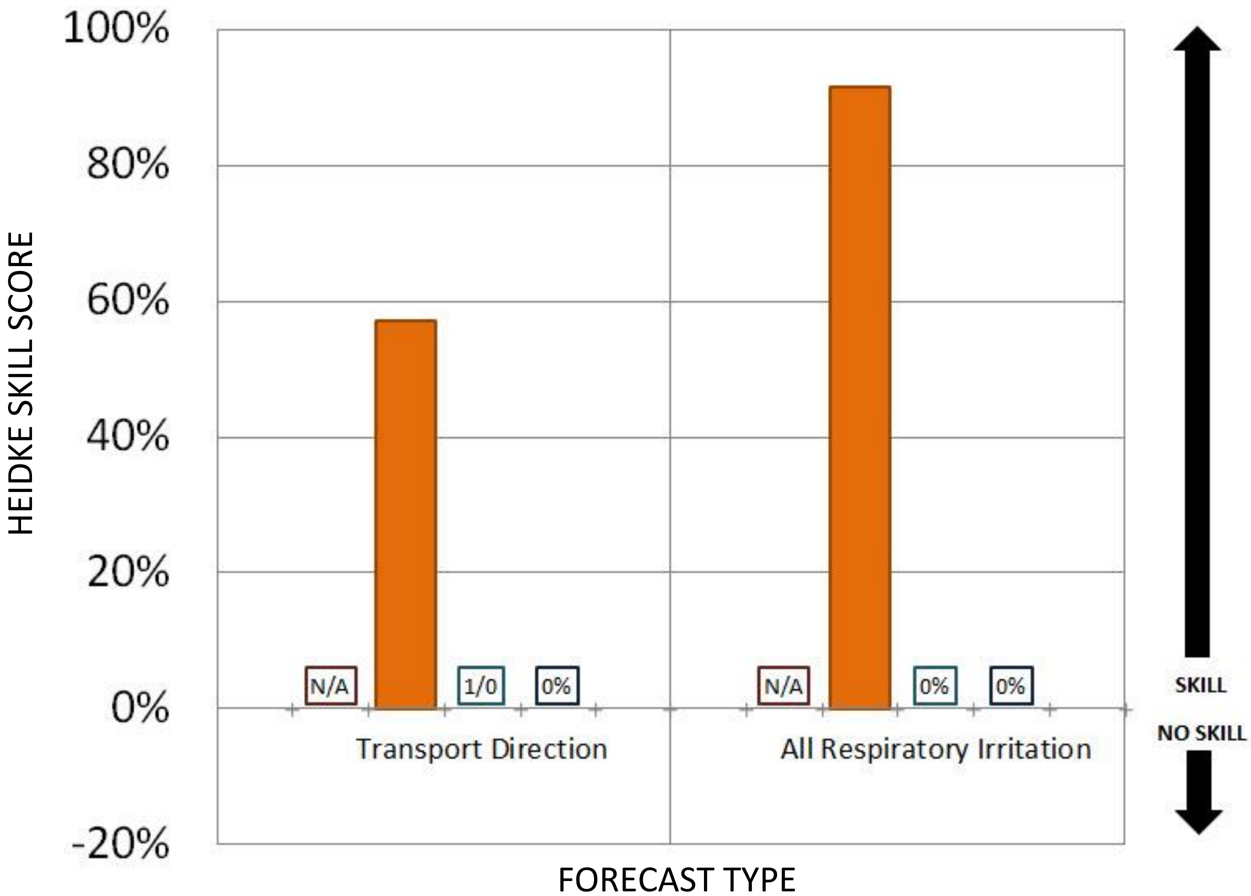
3.6.3 Individual Respiratory Irritation Levels

3.6.3.1 No Levels of Respiratory Irritation (None)
The Heidke skill scores could not be calculated for forecasted irritation levels of "none" issued during BY2010-2012 because there were no forecasts of "none" during that period that could be assessed as confirmed or false. During BY2012-2014, two forecasts of the "none" level of respiratory irritation were assessed as false because of observations of "moderate" levels of respiratory irritation confirmed in the forecast region. Therefore, for both BY2012-2013 and BY2013-2014, the Heidke skill score indicated no improvement over chance (0.00%).

3.6.3.2 Very Low Levels of Respiratory Irritation
There were no assessable "very low" respiratory irritation forecasts issued or observed during the entire BY2010-2014 interval so the Heidke skill score could not be calculated.

3.6.3.3 Low Levels of Respiratory Irritation
Only BY2011-2012 had assessable "low" respiratory irritation forecasts or observations, which allowed a Heidke skill score to be calculated. During BY2011-2012, "low" respiratory irritation forecasts performed better than chance, with a Heidke skill score showing 47.5% improvement over chance. However, this was the lowest Heidke skill score of any assessable forecasted level

of respiratory irritation during BY2011-2012. The Heidke skill scores could not be calculated for "low" respiratory irritation forecasts issued for Texas during BY2010-2011 because there was no bloom, and no assessable "low" respiratory irritation forecasts were issued during BY2012-2014.



Number of Assessable Forecasts (n) for each Bloom Year					
Bloom Years		10/1/10 to 4/30/11	5/1/11 to 4/30/12	5/1/12 to 4/30/13	5/1/13 to 4/30/14
Forecast Type	***Transport Direction***	0	6	2	1
	All Respiratory Irritation	0	31	2	1

Figure 13. The forecast skill of transport direction and all respiratory irritation during the 2010 to 2014 bloom years (BY). There was no *Karenia brevis* bloom during BY2010-2011 and no forecasts were issued as indicated by values of N/A. 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.

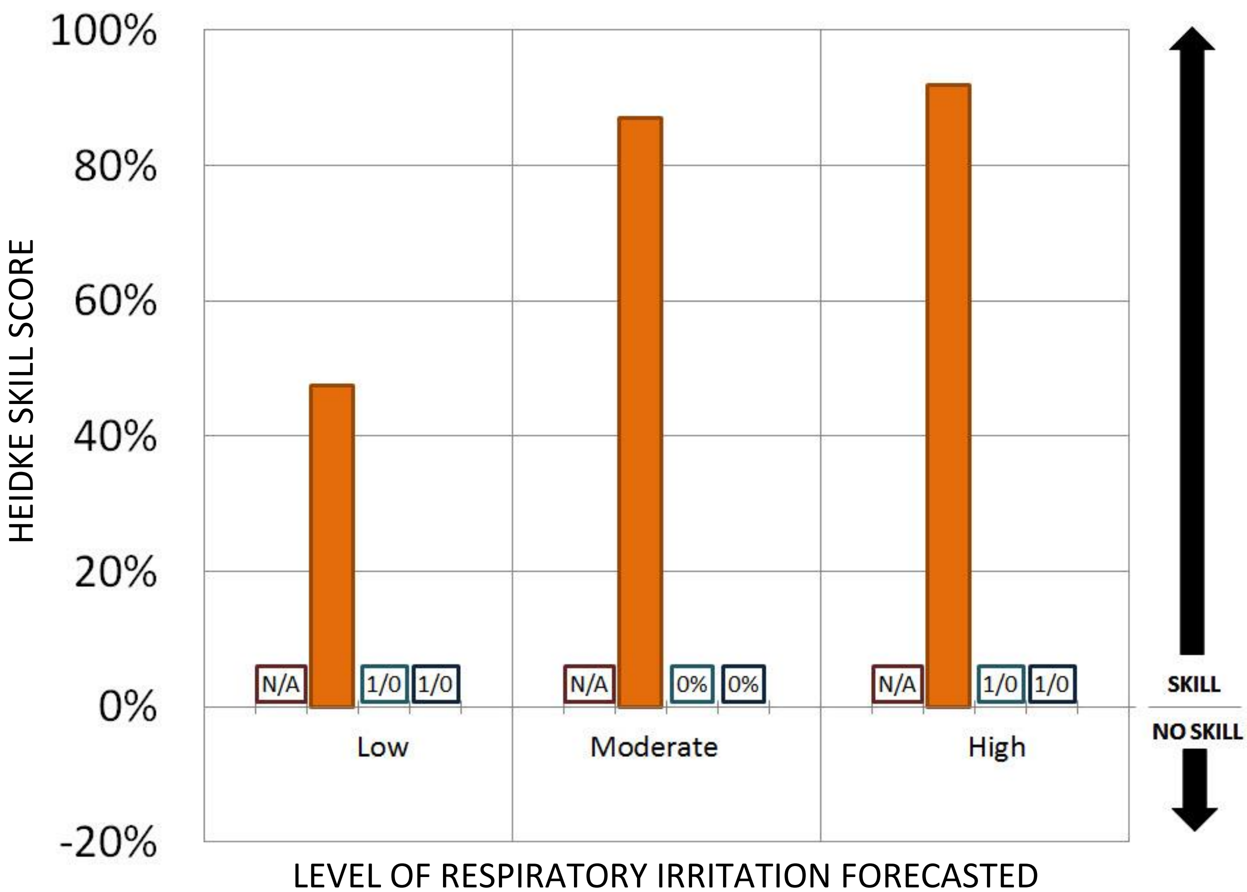
Note: Values of 1/0 indicate that the denominator of the calculation was zero (see Section 2.3 for an explanation of the statistical analyses used).

3.6.3.4 Moderate Levels of Respiratory Irritation

Figure 14 shows that during BY2011-2012, “moderate” respiratory irritation forecasts performed much better than chance, with Heidke skill scores indicating an 87.0% improvement over chance. During BY2012-2013, the Heidke skill score indicated no improvement over chance (0.00%). Although the only assessable “moderate” forecast issued was confirmed correct, there were no correct rejections-cases where a level of respiratory irritation other than “moderate” was both forecasted and confirmed to occur. During BY2013-2014, no assessable forecasts for “moderate” respiratory irritation were issued and again there were no correct rejections assessed. As a result, the Heidke skill score for “moderate” respiratory irritation during BY2013-2014 was 0.00%.

3.6.3.5 High Levels of Respiratory Irritation

Only BY2011-2012 had assessable “high” respiratory irritation forecasts and observations allowing a Heidke skill score to be calculated. Figure 14 shows that during BY2011-2012, “high” respiratory irritation forecasts performed much better than chance, with a Heidke skill score showing 91.9% improvement over chance. This was the highest Heidke score of any assessable forecast for respiratory irritation during BY2011-2012.



Number of Assessable Forecasts (n) for each Bloom Year					
Bloom Years		10/1/10 to 4/30/11	5/1/11 to 4/30/12	5/1/12 to 4/30/13	5/1/13 to 4/30/14
Level of Respiratory Irritation	***Low***	0	1	0	0
	Moderate	0	5	1	0
	High	0	23	0	0

Figure 14. The forecast skill of individual levels of respiratory irritation during the 2010 to 2014 bloom years (BY). There was no Karenia brevis bloom during BY2010-2011 and no respiratory irritation forecasts were issued as indicated by values of N/A. 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.

Note: Values of 1/0 indicate that the denominator of the calculation was zero (see Section 2.3 for an explanation of the statistical analyses used).

3.7 Transport Distance Forecast Verification

There were no blooms during BY2010-2011 and only 5 assessable forecasts of transport distance issued during BY2011-2014. Figure 15 shows a scatter plot of the forecasted and observed transport distances. The sample size of assessable forecasts was too small to estimate a trend in the data, but 3 of the 5 forecasts appeared to be close to the observed distances. The residuals (the difference between forecasted distance and observed distance) were also plotted against the observed distances as shown in Figure 16. Again, there was a very small difference between the forecasted and observed distances for 3 of the 5 assessable forecasts. In both Figures Figure 15 and Figure 16, one of the forecasts issued during BY2011-2012 stood out with a residual of 200 km, but without a larger sample size it was unclear if it was an outlier.

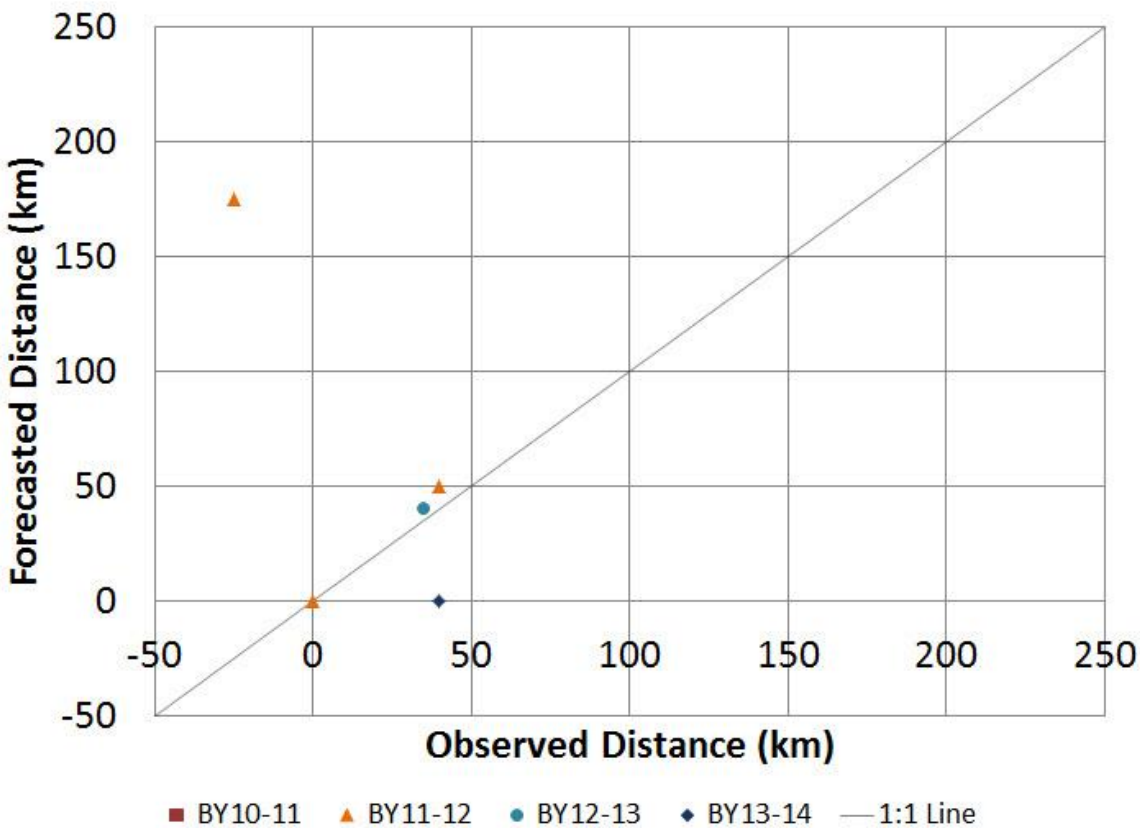


Figure 15. Scatter plot for assessable forecasts and observations of *Karenia brevis* bloom transport distance during the 2010 to 2014 bloom years (BY). There was no bloom during BY2010-2011. A 1:1 line is added to facilitate interpretation.

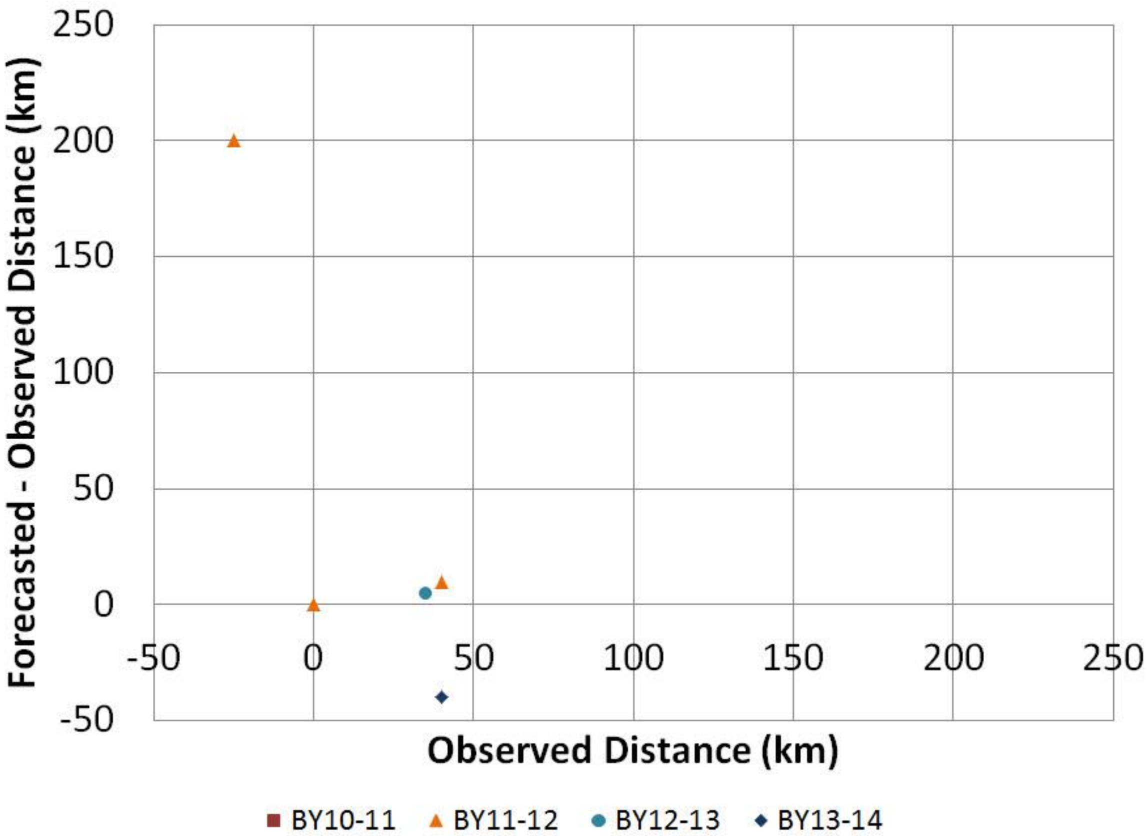


Figure 16. Scatter plot for the residuals (assessable forecasts-observed distance) and observations of *Karenia brevis* bloom transport distance during the 2010 to 2014 bloom years (BY). There was no bloom during BY2010-2011.

3.7.1 Reliability

Transport distance is a continuous variable so in order to assess the bias in the forecasts of transport distance the mean error (ME) was estimated (see Table 11). Overall, during BY2011-2014, transport distance was over-forecast with an ME of 35.0 km (n=5). Very few forecasts of transport distance were assessable so there was a wider range in the bias results. Transport distance was over-forecast during BY2011-2012 with an ME of 70.0 km (n=3). As shown in Figure 16, there was large deviation with a residual of 200 km. When this was excluded, the ME changed to 5.00 km (n=2). During BY2012-2013, transport distance was only slightly over-forecast with an ME of 5.00 km (n=1). Transport distance was under-forecast during BY2013-2014 with an ME of -40.0. (n=1).

3.7.2 Accuracy

Transport distance is a continuous variable so in order to assess the accuracy of the forecasts of transport distance the average magnitude of the error was estimated by determining the mean absolute error (MAE) and root mean square error (RMSE). The error variance was then calculated as the difference between the RMSE and MAE (see Table 11). Overall, during BY2011-2014, the mean absolute error (MAE) was estimated to be 35.0 km (n=5) and the RMSE was 78.3 km. The variance was 43.3 km. During BY2011-2012 the MAE was estimated to be 70.0 km with an RMSE of 121 km (n=3) and a variance of 51.2 km. This was a large variance in error, but RMSE is sensitive to small sample sizes and outliers. As shown in Figure 16, there was an outlier with a residual of 200 km. When the outlier was excluded, the MAE changed to 5.00 km with a RMSE of 7.10 km (n=2) and a variance of 2.10 km. During BY2012-2013, the estimated MAE was 5.00 km with a RMSE of 5.00 km (n=1). During BY2013-2014 the MAE was 40.0 and the RMSE was 40.0 (n=1). Since there was only one assessable forecast of transport distance during both BY2012-2013 and BY2013-2014, there was no variance in error to calculate.

Table 11. The Mean Error, Mean Absolute Error, Root Mean Square Error and Variance in Error calculated for the transport distance forecasts issued during the 2010 to 2014 bloom years (BY). There were no *Karenia brevis* blooms during BY2010-2011.

	BY 2010-2011	BY 2011-2012	BY 2012-2013	BY 2013-2014	ALL: BY 2010-2014
# of Assessable Forecasts (n)	N/A	3	1	1	5
Mean Error (ME)	N/A	70.0	5.00	-40.0	35.0
Mean Absolute Error (MAE)	N/A	70.0	5.00	40.0	35.0
Root Mean Square Error (RMSE)	N/A	121	5.00	40.0	78.3
Variance in Error (RMSE-MAE)	N/A	51.2	N/A	N/A	43.3

DISCUSSION

4.1 Early Warning

Advance cautionary notice of the formation of a *K. brevis* bloom can help those involved in bloom response 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. All three of the *K. brevis* events during BY2010-2014 were first identified by *in situ* sampling collected by organizations in Texas (see APPENDIX III). This stands in sharp contrast to the success the HAB-OFS has demonstrated providing advance notice of HABs in Florida, where nine of the thirteen *K. brevis* events during BY2004-2008 were first identified by the HAB-OFS team using satellite imagery (Kavanaugh, et al., 2013).

The reason for the difference in performance is the distinct optical characteristics of the Texas coast. The chlorophyll anomaly product used by the HAB-OFS is not specific to *K. brevis* and also highlights blooms of other species of algae as well as benthic algae and sediments that are resuspended in the water column by wind and wave action, resulting in false positives (Tomlinson, Wynne, & Stumpf, 2009). Since the surface waters along the Texas coast are prone to resuspension events, a revised chlorophyll anomaly product was used that subtracts an estimate of the resuspended chlorophyll from the chlorophyll anomaly (Wynne, Stumpf, Tomlinson, Ransibrahmankul, & Villareal, 2005). Despite the use of the revised product, during BY2010-2014, the HAB-OFS was unable to reliably detect and track HABs along- and offshore the Texas coast using satellite imagery alone. The revised chlorophyll anomaly highlighted large areas of the coast, whether or not a *K. brevis* bloom had been confirmed at the coast. The anomaly was often so large in extent that *K. brevis* blooms were not discernible within the highlighted area. Due to the limited availability of the water sample data required to validate satellite imagery, it was not possible to determine if the anomalies were correctly flagging blooms or if the anomalies were false positives. Most samples were collected from locations within bays where the chlorophyll anomaly product is known to perform poorly, and no samples were collected offshore where *K. brevis* blooms are known to develop. For that reason, most Texas bulletins did not identify a *K. brevis* bloom location from satellite imagery, and early warning of the presence of an offshore bloom was not possible.

Enhancements to the satellite imagery products and increased sampling are needed. It is clear that the methods used to estimate the resuspended chlorophyll must be modified to improve the performance of the revised chlorophyll anomaly product. Doing so should reduce the frequency and extent of the false positives in imagery. The HAB-OFS is currently evaluating the imagery from a new sensor, NOAA’s Variable Infrared Imaging Radiometer Suite (VIIRS). The higher resolution may help improve *K. brevis* bloom detection. Since the majority of the *in situ* samples are collected from the bay regions, especially in shellfish harvesting areas, research should also be conducted to develop algorithms that enable the detection of *K. brevis* blooms in the bays. The Ocean and Land Colour Instrument (OLCI) aboard European Space Agency’s Sentinel-3 is anticipated to provide higher resolution chlorophyll products that may potentially improve detection of blooms within the bays.

Additionally, *in situ* data needs to be collected offshore and along the coast so that the revised chlorophyll anomaly product can be validated and the movement of identified blooms can be monitored consistently. Verifying the satellite imagery products with *in situ* sample data will inform future evaluations, aiding in the enhancement of the satellite imagery product algorithms. The manual collection of offshore water samples and aerial overflights can be expensive so the use of autonomous underwater vehicles (AUVs), additional Imaging FlowCytobots moored in other locations along the Texas coast, and other data collection methods should be investigated. An expansion of volunteer networks, like the Red Tide Rangers, would also inexpensively increase both coastal and offshore water sample collection. Data collection and sharing should be supported and integrated through GCOOS, as is currently done with the Imaging FlowCytobot data, TGLO/TABS/TAMU modeled currents, and other products relevant to HAB monitoring and forecasting, with the objective of forming a more extensive and effective monitoring system for the Gulf of Mexico.

4.2 Bulletin Utilization

During BY2010-2014, bulletin utilization increased drastically from 3.33% in BY2010-2011 to 74.6% in BY2013-2014. Some of this increase may be due to an increase in the awareness of the product; the number of subscribers increased by 32.7% during the same time period (NOAA, 2014). Overall, utilization was most likely underestimated from BY2010-2012 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. Unlike those in Florida, Texas organizations did not regularly acknowledge the HAB bulletins as a resource in their public communications. The HAB-OFS Facebook Page, launched in September of 2012, simplified the process of evaluating utilization because Facebook records metrics that detail the ways that people interact with each post. Using Facebook, the HAB-OFS was able to better quantify the number of people reading the conditions report section of the Texas HAB bulletin during BY2012-2014. It is important to note that Facebook, as well as some of the other sources of bulletin utilization, provided evidence of product viewership. However, evidence of applying product content to bloom response was less common from the methods employed during the evaluation period.

Of bulletins issued in Florida during BY2004-2008, the majority of utilized bulletins were categorized as high priority, followed by medium priority. A similar trend was expected for Texas since the priority levels indicate to resource managers the status of a bloom event. Bulletins are categorized as medium or high priority if a *K. brevis* bloom is present that may require response. On the other hand, low priority bulletins are issued when there is no bloom and action from resource managers is likely unnecessary. However, high priority bulletins issued in Texas were only the highest proportion of utilized bulletins during BY2011-2012. Since there was no bloom during BY2010-2011, only low priority bulletins were issued, and only one was confirmed utilized. During BY2012-2014, the low priority bulletins had the highest proportion utilized. During this timeframe, the HAB-OFS Facebook Page raised the visibility of all bulletin excerpts posted. However, because of the algorithms used by Facebook, not all posts reach a follower’s Newsfeed. Since low priority bulletins were more frequently issued, they were more likely to be seen on Facebook, especially if a follower had interacted with a similar post previously. Furthermore, Facebook users may have actively interacted with low priority bulletin posts, responding with likes, shares, and comments because of the favorable nature of the forecast statement that "no respiratory irritation is expected." This suggests that members of the general public may use and respond to HAB-OFS products in different ways than subscribers of the HAB bulletin such as resource managers and researchers.

The HAB-OFS Facebook Page increased access to the HAB-OFS forecasts, but overall, there were few confirmations of utilization received through direct correspondence from bulletin subscribers. Since data on utilization of the product is extremely important for guiding improvements, additional methods for evaluating utilization and usefulness should also be explored, including tracking receipt of the HAB bulletin by subscribers and implementing routine surveys.

4.3 Forecasts of Transport Direction and Transport Distance

The percentage of transport forecasts issued that could be assessed was much lower for Texas than for Florida. During BY2010-2014, only 4.11-12.5% of forecasts of transport direction issued for Texas were assessable compared to 29.8-69.0% issued for Florida during BY2004-2008. Similarly, only 2.07-6.25% of the transport distance forecasts issued for Texas were assessable forecasts, but no comparison exists for Florida since transport distance was not forecasted in that region. HAB-OFS transport forecasts issued for either Florida or Texas can be difficult to assess because they require a series of satellite images or water samples where the bloom location is consistently distinguishable. Clouds in satellite imagery render bloom boundaries indiscernible. Stumpf et al. (2009) determined that only large HABs, covering more than 10 to 30 km of coast could be reliably located and validated by sampling and imagery. That may explain why the majority of assessable transport forecasts were issued for the BY2011-2012, which was one of the most geographically extensive and longest lasting *K. brevis* blooms recorded in Texas history (NOAA, 2012; Sherman, 2011). In fact, 6 out of 9 assessable forecasts of transport direction and 3 out of 5 of the assessable forecasts of transport distance were issued during the BY2011-2012 bloom.

The size of the bloom was not the only factor that affected the assessment of Texas transport forecasts. As previously mentioned, *K. brevis* blooms were often indistinguishable from resuspended benthic chlorophyll and sediments, even though the revised chlorophyll anomaly product was used. Water samples were not collected offshore and were rarely collected in several areas of the coast, creating data gaps that also prevented transport forecasts from being validated.

Only nine of the 187 forecasts (4.81%) of transport direction issued for Texas were assessable during BY2010-2014 making it difficult to draw meaningful conclusions about the overall forecast quality of either the transport direction or transport distance forecasts. Most of the transport direction forecasts that could be validated were highly accurate and had Heidke scores indicating a performance much better than chance when predicting the direction of bloom movement. Overall, measures of the accuracy, reliability, and skill of assessable transport direction forecasts were comparable to the evaluation results for forecasts issued for Florida during BY2004-2008 despite different methods being used to forecast transport direction in each region (Kavanaugh, et al., 2013). Just five of the 208 forecasts (2.40%) of transport distance were assessable, too few to evaluate the statistical significance of the forecast quality. On average, the assessable forecasts were over-forecast with a mean error of 35.0 km, but three forecasts had very low residuals and were within 10 km of the observed transport distance. One of the assessable forecasts issued during BY2011-2012 seemed to be an outlier, yielding over-forecasting with a mean error of 70.0 km. Despite this, the assessable forecasts for transport distance were similar to the resolution of the extent forecasts issued in Florida which was estimated to be approximately 30.0 km (Stumpf, et al., 2009). However, the sample size was too small to thoroughly evaluate the forecast accuracy and determine the significance of the results relative to chance.

Due to the limited assessment capability, this evaluation could not use the results of the statistical analyses to definitively identify aspects of the forecast methods that require refinement, but there may still be some modifications that should be considered in the future. The TGLO/TABS/TAMU ROMS-based hydrodynamic model and GNOME oil spill model used to nowcast and forecast surface currents for the HAB transport forecasts have been evaluated and used operationally for tracking oil spills and for researching *K. brevis* bloom movement in the Gulf of Mexico (Hetland & Campbell, 2007; Martin, et al., 2005; Walpert, Guinasso, Lee, & Martin, 2011). However, the procedure used by the HAB-OFS to forecast bloom transport requires evaluation. Inaccuracy may have been introduced into several of the steps. For transport forecasts, the starting location of the bloom was often estimated from cell counts in a nearby bay. The TGLO/TABS/TAMU ROMS-based hydrodynamic model does not include the bay regions where many of the samples were collected, which might also have introduced error to the forecast. Additionally, the forecasted distance of bloom movement may be inexact because the estimated starting locations for both the nowcast and the forecast were entered by hand into GNOME. The distance from the estimated starting location to the nowcast or forecast location was then manually measured. Furthermore, only the modeled particles closest to the coast were considered, excluding the predicted offshore movement of particles, even though there was often notable variation between the two.

Improvements to the revised chlorophyll anomaly as well as the frequency and extent of water sample collection are imperative in order to thoroughly evaluate the transport forecast methods and recommend changes. Although a more thorough evaluation could not be completed for the transport forecasts at this time, there are two modifications that might be considered in the future. Firstly, since the majority of the water sample data is collected from the bay regions, the CO-OPS Northern Gulf of Mexico Operational Forecast System (NGOFS) hydrodynamic model should be explored as an alternative to the TGLO/TABS/TAMU ROMS-based model because it nests a high resolution model for the northwestern Gulf of Mexico (NWGOFS) which includes Matagorda Bay, Galveston Bay, and the Sabine Neches region (NOAA, 2014). This would potentially allow the HAB-OFS to forecast the movement of *K. brevis* blooms in and out of Galveston Bay, a region where blooms frequently occur. Secondly, the procedures for determining the starting location of the nowcast should be revised. Rather than inputting the location by hand, the satellite imagery or water samples should be converted to particles and ingested into GNOME similar to the procedure used for the Lake Erie HAB Demonstration Forecast System (Wynne, et al., 2011). Before this modification is possible though, the revised chlorophyll anomaly product must be refined to enable better HAB detection. Once chlorophyll products become available from Sentinel-3, an evaluation should determine if the higher resolution imagery could enhance detection of HABs within the bays.

4.4 Respiratory Irritation

As was found to be the case for the evaluation of the HAB-OFS in Florida, respiratory irritation forecasts proved to be the most difficult forecast to assess because observational data from the field is required for validation (Kavanaugh, et al., 2013). Blooms are patchy by nature and their associated impacts are sporadic. Respiratory irritation is most likely underreported because even if some forecast areas have brevetoxin aerosols blowing onshore as forecasted, without experienced observers present, the respiratory irritation associated with the bloom might not be correctly reported, if reported at all.

It was more difficult to assess forecasts of respiratory irritation issued for Texas than for Florida; only 0.917-8.93% of respiratory irritation forecasts issued for Texas were assessable compared to 10.0-54.0% issued for Florida (Kavanaugh, et al., 2013). In part, this was due to the fact that unlike Florida, which hosts Mote Marine Laboratory's Beach Conditions Reporting System, Texas does not have an established network for collecting observational data specific to respiratory irritation from beaches along the coast. Instead, observations of respiratory irritation were summarized by TPWD from intermittent reports received from those collecting samples in the field and the general public. Consequently, the level of observed respiratory irritation was not always categorized systematically and observations were not collected on a routine basis. Past evaluations of respiratory irritation forecasts for Florida suggested that reliance on anecdotal information resulted in a bias toward reports of more severe respiratory irritation levels, which made it difficult to assess the impacts of small, patchy blooms (Kavanaugh, et al., 2013; Stumpf, et al., 2009). This explains why 31 of the 34 assessable respiratory irritation forecasts issued between 2010 and 2014 were issued during the BY2011-2012 Texas bloom, one of the most geographically extensive and longest lasting blooms that Texas has on record (NOAA, 2012; Sherman, 2011).

The respiratory irritation forecast model was originally designed to predict conditions along the coast of Florida. In Texas however, most of the water samples were collected from bay regions, meaning the respiratory irritation forecasts may also have varied in accuracy, but without systematically collected observational evidence, this could not be determined. Consistently using defined categories for the levels of respiratory irritation and increasing the frequency and extent of the observations would help validate forecasts, as well as enable improvements to the forecasting procedures.

The inability to assess many of the respiratory irritation forecasts limits the evaluation of forecast quality. There were not enough assessable forecasts issued during BY2012-2014 for the statistical analyses to yield meaningful results so only the forecasts issued during the BY2011-2012 are discussed in depth.

Overall, the assessable respiratory irritation forecasts issued in Texas during the BY2011-2012 bloom were highly accurate and consistently performed much better than chance. The proportion correct and Heidke skill scores were similar to those calculated for forecasts issued in Florida during BY2004-2008.

4.4.1 No Respiratory Irritation Level

When a bloom is not present, a forecast of “no” respiratory irritation along the Texas coast is issued, and therefore the majority of bulletins forecasted “no” respiratory irritation and could not be confirmed. Forecasts of “no” respiratory irritation could not be adequately assessed in most cases. The exceptions were at the beginning of blooms in both BY2012-2013 and BY2013-2014 when reports of observed impacts prompted TPWD to coordinate the collection of water sample data. The forecasts of respiratory irritation are based on observational data. Since respiratory irritation was observed before other observational data was collected, the HAB-OFS forecasts at the beginning of these blooms were shown to be inaccurate. The routine collection of observational data, as well as increased water samples and enhancements to satellite imagery, would all help prevent such forecast errors from recurring.

4.4.2 Very Low Levels of Respiratory Irritation

Forecasts of “very low” levels of respiratory irritation are difficult to assess in both Florida and Texas because the respiratory irritation level is limited to members of the population who suffer from chronic respiratory conditions (like asthma) and are especially sensitive to brevetoxin aerosols. The “very low” level forecast can only be confirmed in the rare event when reports of observed respiratory irritation indicate that only someone suffering from chronic respiratory issues had experienced symptoms associated with the presence of a *K. brevis* bloom. Additionally, if a bloom is known to be in the region, the area may be avoided. Such reports were never received for blooms in Texas during BY2011-2014.

4.4.3 Low Levels of Respiratory Irritation

Observations of “low” levels of respiratory irritation may have been underreported. Of the 31 assessable forecasts issued during BY2011-2012, only one was an assessable forecast for “low” levels of respiratory irritation. Observations of “low” levels of respiratory irritation were reported two other times during that bloom year, but the forecasts were for “moderate” or “high” levels of respiratory irritation. Therefore, although the forecast performed better than chance, it was not as accurate as forecasts for “moderate” or “high” levels of respiratory irritation and it was slightly under-forecast.

Observations of less severe levels of respiratory irritation, such as “very low” and “low” levels, are often underreported, especially without a systematic way of collecting the observations. In Florida, before the HAB-OFS began using data from Mote Marine Laboratory’s Beach Conditions Reporting System in August 2006, no observations of “low” respiratory irritation levels were reported during the bloom in BY2004-2005 and only 15 observations were reported during the two blooms in BY2005-2006 (Kavanaugh, et al., 2013). In contrast, over 350 observations of “low” respiratory irritation levels were reported during the bloom in BY2006-2007 by Mote Marine Laboratory alone (Kirkpatrick, et al., 2008; Mote Marine Laboratory, 2007). This clearly indicates that a comprehensive network of observers providing routine observations enable the “low” respiratory irritation level to be assessed.

4.4.4 *Moderate and High Levels of Respiratory Irritation*

The assessable forecasts for "moderate" and "high" levels of respiratory irritation were highly accurate during BY2011-2012. Both sets of forecasts were also very reliable, although "moderate" respiratory irritation forecasts were slightly more over-forecasted than "high" respiratory irritation forecasts. Both "moderate" and "high" respiratory irritation forecasts performed much better than chance at predicting the observed conditions.

These findings are noteworthy because the "moderate" and "high" respiratory irritation forecasts are arguably the most vital forecasts for directly protecting public health so it is essential that they are the best performing forecasts issued by the HAB-OFS. Not only were the forecasts issued for Texas highly accurate, but over-forecasting was minimal. Over-forecasting the "moderate" or "high" respiratory irritation forecasts could have potentially undermined the credibility of the forecasts themselves and jeopardized tourism by unnecessarily discouraging people from visiting the forecast regions. However, it is also important that the forecasts were not under-forecast. The forecasts erred on the side of caution meaning the public was properly warned about possible health risks.

The same methods were used to forecast respiratory irritation in both Texas and Florida, but there were differences between the two regions that could affect the forecast performance. For instance, in Texas, most of the water sample data was collected in the bay regions. The same level of respiratory irritation might not result in the bays as it would in an adjacent coastal region, in part because of differences in the winds and wave action. Water samples and observations of respiratory irritation were also not collected and reported as extensively or frequently in Texas as in Florida, and the limitation of that critical input data could have affected the forecasts as well. Despite these potential factors that could cause differences, during BY2011-2012 the measures of accuracy and skill of the "moderate" and "high" forecasts issued in Texas were comparable to those of forecasts issued in Florida during BY2004-2008. In fact, the forecasts performed better in some cases (Kavanaugh, et al., 2013).

There were limited observations of "moderate" respiratory irritation and no reports of "high" respiratory irritation during the BY2012-2014 Texas blooms. Unfortunately, meaningful evaluations of the forecasts issued during the blooms BY2012-2014 were not possible. Potentially the blooms in both years were so patchy and small in scale that either no existing respiratory irritation levels were observed or only low levels of brevetoxin were produced, similar to what may have occurred during the BY2007-2008 bloom in southwest Florida with comparable results (Kavanaugh, et al., 2013). Since systematic observations are not collected routinely in Texas, it is difficult to determine if the forecasts were inaccurate or if the conditions were underreported. Methods to collect and incorporate observations beyond cell concentrations should be investigated. In Florida, Mote Marine Laboratory's Beach Conditions Reports have been an excellent tool for providing daily estimates of the level of respiratory irritation at chosen sites. Since cell concentrations are only a proxy for how much brevetoxin aerosol might be present and the actual amount produced by a *K. brevis* bloom varies, direct measures of the concentration of brevetoxin both in the water and the air should be investigated. Research is currently being conducted to create an inexpensive test kit for brevetoxin with a concept similar to the domoic acid dip stick test kit developed with support through the Ecology and

Oceanography of HABs (ECOHAB) and Monitoring and Event Response for HABs (MERHAB) research programs (NOAA, 2014). Even if a test kit is developed for brevetoxin, an expansion of trained personnel and volunteer networks, like the Red Tide Rangers, will be required in order to collect daily observations along the Texas coast. Data collection and sharing should also be integrated with other products relevant to HAB monitoring and forecasting available through GCOOS.

CONCLUSION

Since October 1, 2010, the HAB-OFS has provided the western Gulf of Mexico with operational forecasts for *K. brevis*, the species commonly known in the region as red tide. HAB-OFS forecasts and analyses were disseminated to subscribers through the HAB bulletin product on a biweekly basis during an active bloom and once a week when no bloom was present in Texas. This report provides an evaluation of the HAB-OFS products issued for Texas during the bloom years from October 1, 2010 to April 30, 2014, with comparisons to those issued for Florida where possible (Kavanaugh, et al., 2013). The analysis includes an assessment of bulletin utilization, early warning capability, and forecast quality.

From the time the HAB-OFS was transitioned to operations in the western Gulf of Mexico on October 1, 2010 to the end of the fourth bloom year on April 30, 2014, a total of 219 bulletins and 6 supplemental bulletins and/or conditions updates were issued. During BY2010-2012, two analysts, specially trained to utilize established standard operating procedures and analytical methods, were responsible for creating and disseminating the bulletins. Since then, additional resources were required and four additional analysts were trained and incorporated into the rotation.

From BY2010-2014, the percentage of confirmed bulletin utilization increased from 3.33% in BY2010-2011 to 74.6% in BY2013-2014 due to an increase in awareness of the product over the course of four years and the launch of the HAB-OFS Facebook Page in the fall of 2012. Confirmed bulletin utilization during BY2013-2014 in Texas was comparable to the average confirmed utilization of all bulletins issued for Florida during BY2004-2008 (72.0%) (Kavanaugh, et al., 2013).

In Florida, during BY2004-2008, the HAB-OFS played an important role in planning bloom response efforts in advance of impacts reaching the coast by providing advance detection of blooms via satellite imagery. Nine out of thirteen *K. brevis* blooms were identified first by the HAB-OFS in satellite imagery during that period. In contrast, no early warning was given by the HAB-OFS for the three *K. brevis* blooms that occurred during BY2011-2014 along the Texas coast. Instead, organizations in Texas were prompted to collect *in situ* water sample data that confirmed the presence of *K. brevis* blooms following receipt of reports of fish kills and respiratory irritation along the coast. The surface water along the Texas coast is more prone to resuspension events than the southwest Florida coast which inhibits bloom detection via satellite imagery using the chlorophyll anomaly product developed for Florida. The HAB-OFS used a revised chlorophyll anomaly product for HAB monitoring along the Texas coast, in the attempt to separate the effect of resuspended chlorophyll from the regular chlorophyll anomaly, but *K. brevis* blooms were often still indiscernible within the area highlighted by the revised anomaly. Furthermore, if the anomaly flagged a feature offshore, ground truthing of the satellite imagery was hindered since limited *in situ* data was collected along the coast and none was collected offshore.

The limited availability of necessary observational data also inhibited the assessment of the forecasts. During BY2010-2014, a total of 961 forecasts were issued indicating the potential direction and distance of bloom transport and the potential level of associated respiratory irritation. However, only a small percentage of each type of forecast could be verified with the available observational data. A large proportion of assessable forecasts were issued for the Texas bloom during BY2011-2012, one of the most geographically extensive and longest lasting ever recorded in Texas (NOAA, 2012; Sherman, 2011). Large HABs, covering greater than 10-30 km of coast are more reliably located and validated by sampling and imagery, and the verification of the respiratory irritation forecasts is biased towards reports of more severe levels of irritation (Kavanaugh, et al., 2013; Stumpf, et al., 2009). The majority of assessable forecasts for transport direction, transport distance, and respiratory irritation were accurate, but due to the small sample size of assessable forecasts the results may or may not represent the overall quality of forecasts issued during BY2010-2014.

Respiratory irritation forecasts were the most frequently issued, but they were also among the most difficult to assess because of reliance on reports of observations from the field, which were rare. Although conclusions could not be drawn from the small sample size of assessable forecasts during BY2012-2014, the assessable respiratory irritation forecasts issued in Texas during the BY2011-2012 bloom were highly accurate and consistently performed much better than chance. The proportion correct and Heidke skill scores were similar to those calculated for forecasts issued in Florida during BY2004-2008. The "moderate" and "high" level respiratory irritation forecasts had the greatest accuracy, reliability, and skill of all forecast types issued by the HAB-OFS. This is especially significant because these forecasts have the greatest potential to directly protect public health.

The number of assessable transport direction and transport distance forecasts was even smaller than for respiratory irritation. Nine out of 187 forecasts of transport direction were assessable. Most of those nine assessable forecasts were highly accurate and performed better than chance at predicting the direction of bloom movement. Only five of the 208 forecasts of transport distance were assessable. Of those, three forecasts had very low residuals when compared to the observed bloom movement, but the sample size of assessable forecasts was too small to thoroughly evaluate the forecast accuracy, reliability, and the overall significance of the results relative to chance.

This assessment clearly demonstrates that the revised chlorophyll anomaly product needs to be modified to enable improved detection of blooms from the satellite imagery. Since the majority of *in situ* sample data is collected in the bay regions, research should also be conducted to develop an algorithm that enables *K. brevis* bloom detection in the bays. Increased frequency and extent of *in situ* data would also aid in ground truthing the bloom location, transport, and respiratory irritation.

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

- Investigate the use of the CO-OPS nested NGOFS model to forecast bloom transport direction and distance in order to:

    - enable more accurate, higher resolution spatial and temporal predictions of bloom movement
    - forecast bloom movement in and out of Galveston Bay, where blooms frequently persist and where many water samples are collected
- Explore a procedure to directly input the nowcast starting location into GNOME using satellite imagery and/or water sample data similar to the methods used for the Lake Erie Demonstration Forecast System (Wynne, et al., 2011).
- Once chlorophyll products become available from Sentinel-3, evaluate the higher resolution imagery to determine a method to enhance HAB detection within the bays.
- Upgrade the respiratory irritation forecasts through:
    - exploring ways to support GCOOS and increase the frequency and extent of observational data
    - developing methods to integrate direct measurements of the concentration of brevetoxin in the air and water
- Examine additional methods of assessing product utilization by subscribers that better capture evidence of both the product being viewed and its contents being applied to bloom response.

These enhancements are proposed with the Texas HAB-OFS in mind. However, some of the recommendations may also be applicable to the Florida HAB-OFS. On a broader scale, the assessment results may also be relevant to the potential expansion of the HAB-OFS to include new forecast regions in the United States.

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: Edward Davis, Katherine Derner, 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 John Cassidy (Center for Operational Oceanographic Products and Services). Additional organizations that contributed to the HAB-OFS during this period included the Texas Parks and Wildlife Department, Texas A&M University, Texas Sea Grant, Texas General Land Office, and NOAA's National Weather Service.

REFERENCES

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.

GCOOS. (n.d.). *Harmful Algal Blooms (HAB): Hourly time series plots*. Retrieved February 2, 2015, from Gulf of Mexico Coastal Ocean Observing System: http://gcoos.org/products/index.php/bio/hab/time-series/

Hetland, R. D., & Campbell, L. (2007). Convergent blooms of Karenia brevis along the Texas coast. *Geophysical Research Letters*. doi:10.1029/2007GL030474

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., Nierenberg, K., Reich, A., Backer, L. C., Stumpf, R., . . . Kirkpatrick, G. (2008). Florida Red Tide and Human Health: A Pilot Beach Conditions Reporting System to Minimize Human Exposure. *Science of the Total Environment*, 1-8.

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.

Kirkpatrick, B., Pierce, R., Cheng, Y., Henry, M., Blum, P., Osborn, S., . . . Baden, D. (2010). Inland transport of aerosolized Florida red tide toxins. *Harmful Algae*, 186-189.

Magana, H. A., Contreras, C., & Villareal, T. A. (2003). A historical assessment of Karenia brevis in the western Gulf of Mexico. *Harmful Algae*, 163-171.

Martin, R. D., Guinasso, N. L., Lee, L. L., Walpert, J. N., Bender, L. C., Hetland, R. D., . . . Howard, M. K. (2005). Ten years of realtime, near-surface current observations supporting oil spill response. *International Oil Spill Conference Proceedings: May 2005. 2005*, pp. 541-545. Miami Beach, Florida: IOSC. doi:http://dx.doi.org/10.7901/2169-3358-2005-1-541

Mote Marine Laboratory. (2007). Beach Conditions Reporting System data set, 2006 to 2007. [Data set].

NOAA. (2012, September/October). Red Tide Rangers Wrangle Data During Texas Harmful Algal Blooms. *Coastal Services, 15*(4), pp. 6-7. Retrieved January 5, 2014, from http://coast.noaa.gov/digitalcoast/sites/default/files/files//digitalcoast/node/304/edit/issue%20%281%29.pdf#page=5

NOAA. (2013). *NOAA Harmful Algal Bloom Operational Forecast System (HAB-OFS)*. Retrieved February 8, 2013, from http://tidesandcurrents.noaa.gov/hab/index.html

NOAA. (2014). [HAB-OFS Subscriber List and User Metrics]. Unpublished raw data.

NOAA. (2014, October 31). *Harmful Algal Blooms, Hypoxia & Climate Change: West Coast Region.* Retrieved February 2, 2015, from NOAA's Ecological Forecasting Services: http://www.cop.noaa.gov/stressors/extremeevents/hab/current/project_details.aspx

NOAA. (2014, September 16). *Northern Gulf of Mexico Operational Forecast System (NGOFS)*. Retrieved from NOAA Tides & Currents: http://tidesandcurrents.noaa.gov/ofs/ngofs/ngofs.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

Nurmi, P. (2005). General guide to the verification of local weather forecasts. *NOMEK Training.* Oslo.

Pack, W. (2012, February 23). Oyster eaters can soon sate their hunger. *Houston Chronicle*. Retrieved March 1, 2012, from http://www.chron.com/business/article/Oyster-eaters-can-soon-sate-their-hunger-3357533.php

Red tide toxin found in dead ducks. (2012, February 23). *The Brownsville Herald*. Retrieved February 29, 2012, from http://www.brownsvilleherald.com/news/valley/article_4741c9f6-64da-5860-914e-a776320f74f7.html

Sherman, C. (2011, October 17). Texas Gulf Coast Sees Largest Algae Bloom in Years. *The Boston Globe*. Retrieved January 5, 2015, from http://www.boston.com/news/nation/articles/2011/10/17/gulf_algae_bloom_affects_much_of_texas_gulf_coast/

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., 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.

Tomlinson, M., Wynne, T., & Stumpf, R. (2009). An evaluation of remote sensing techniques for enhanced detection of the toxic dinoflagellate, Karenia brevis. *Remote Sensing of Environment*, 598-609.

Walpert, J., Guinasso, N., Lee, L., & Martin, R. (2011). Texas Automated Buoy System-sustainable ocean observations to help protect the environment. *OCEANS 2011* (pp. 1-7). Waikoloa, HI: IEEE.

Wynne, T. T., Stumpf, R. P., Tomlinson, M. C., Ransibrahmankul, V., & Villareal, T. A. (2005). Detecting Karenia brevis blooms and algal resuspension in the western Gulf of Mexico with satellite ocean color imagery. *Harmful Algae*, 992-1003.

Wynne, T. T., Stumpf, R. P., Tomlinson, M. C., Schwab, D. J., Watabayashi, G. Y., & Christensen, J. D. (2011). Estimating cyanobacterial bloom transport by coupling remotely sensed imagery and a hydrodynamic model. *Ecological Applications*, 2709-2721.

APPENDICES

APPENDIX I
Example of a HAB bulletin for the Texas region.

APPENDIX II
Forecast regions defined for Texas HAB bulletins.

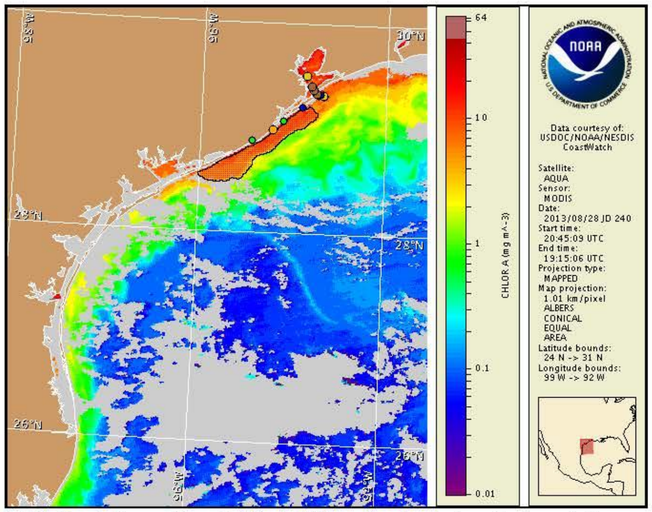
APPENDIX III
List of organizations that contributed to the 2010-2014 HAB-OFS bulletins for Texas.

APPENDIX I

Example of a HAB bulletin for the Texas 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/habfs_bulletin_guide.pdf.

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

Gulf of Mexico Harmful Algal Bloom Bulletin
Region: Texas
Thursday, 29 August 2013
NOAA National Ocean Service
NOAA Satellite and Information Service
NOAA National Weather Service
Last bulletin: Monday, August 26, 2013



Satellite chlorophyll image with possible *K. brevis* HAB areas shown by red polygon(s), when applicable. Points represent cell concentration sampling data from August 19 to 28: red (high), orange (medium), yellow (low b), brown (low a), blue (very low b), purple (very low a), pink (present), and green (not present). Cell count data are provided by Texas Parks and Wildlife Department. For a list of sample providers and a key to the cell concentration categories, please see the HAB-OFS bulletin guide: http://tidesandcurrents.noaa.gov/hab/habfs_bulletin_guide.pdf

Detailed sample information can be obtained through the Texas Parks and Wildlife Department at: http://www.tpwd.state.tx.us/landwater/water/environconcerns/hab/redtide/status.phtml

To see previous bulletins and forecasts for other Harmful Algal Bloom Bulletin regions, visit the NOAA Harmful Algal Bloom Operational Forecast System bulletin archive: http://tidesandcurrents.noaa.gov/hab/bulletins.html

Conditions Report

Not present to medium concentrations of *Karenia brevis* (commonly known as Texas red tide) are present along- and offshore Texas. *K. brevis* concentrations are patchy in nature and levels of respiratory irritation will vary locally based upon nearby bloom concentrations, ocean currents, and wind speed and direction. The highest level of potential respiratory irritation forecast for Thursday, August 29 to Tuesday, September 3 is listed below:

Region: Forecast (Duration)
Gulf Coast- Bolivar Peninsula: Moderate (Th-M), Very low (Tu)
Galveston Bay Region: Moderate (Th-Tu)
Gulf Coast- Galveston Island: Moderate (Th-M), Very low (Tu)
Gulf Coast- San Luis Pass to East Matagorda Bay Peninsula: Moderate (Th-M), Very low (Tu)

Check http://tidesandcurrents.noaa.gov/hab/beach_conditions.html for recent, local observations. Health information, from the Texas Department of State Health Services and other agencies, is available at http://tidesandcurrents.noaa.gov/hab/hab_health_info.html. Reports of respiratory irritation were received from the Galveston Island region and Surfside Beach yesterday.

There are currently patches of a bloom of the algae *Aureoumbra lagunensis* in the upper Laguna Madre region. This algae species does not produce the respiratory irritation associated with the Texas red tide caused by *Karenia brevis*, but it may cause discolored water and fish kills.

Analysis

As of today, August 29, Texas bulletins will be issued twice weekly on Mondays & Thursdays due to current harmful algal bloom activity. Due to the upcoming Federal Holiday, the next bulletin will be issued on Tuesday, September 3.

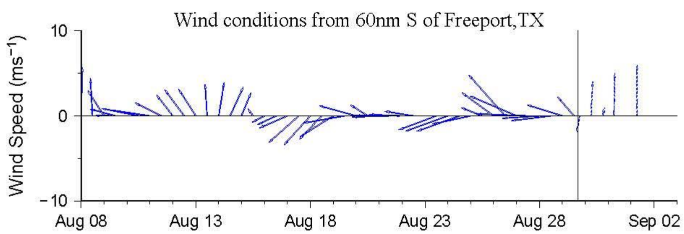
A harmful algal bloom of *Karenia brevis* has been identified in the Bolivar Peninsula, Galveston, and San Luis Pass to East Matagorda Bay regions of Texas. In the San Luis Pass to East Matagorda Bay region, *K. brevis* concentrations range between not present and 'medium' with the highest concentration found at Surfside Beach (TPWD, 8/27-28). In the Galveston Island, Galveston Bay, and Bolivar Peninsula regions, *K. brevis* concentrations range between not present and 'low b' with the highest concentrations found at Houston Ship Channel Marker 47 and in Bolivar Roads Pass adjacent to the Galveston Yacht Basin (TPWD, 8/27-28).

Recent MODIS Aqua imagery from 8/28 (shown left) is partially obscured by clouds along- and offshore the Texas coast. Elevated to very high chlorophyll concentrations (2->20 μg/L) are visible in patches alongshore extending up to 20km offshore from the Sabine Pass to Matagorda Pass regions, with patches of very high chlorophyll (>20 μg/L) extending south from the Surfside Beach region. Elevated chlorophyll is not necessarily indicative of the presence of *K. brevis* and could also be due to the resuspension of benthic chlorophyll and sediments along the coast. In situ sampling is necessary to confirm the presence of *K. brevis*.

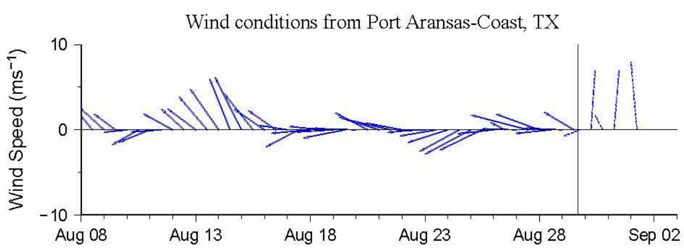
Example of a HAB bulletin for the Texas region (page 2).

Forecast models based on predicted near-surface currents indicate that the maximum bloom transport may be negligible (<10km) from coastal sample locations in the Surfside Beach region and the potential transport from the Port Aransas region would also be negligible (<10km) from August 28 to September 1.

Davis, Kavanaugh



Wind speed and direction are averaged over 12 hours from buoy measurements. Length of line indicates speed; angle indicates direction. Red indicates that the wind direction favors upwelling near the coast. Values to the left of the dotted vertical line are measured values; values to the right are forecasts. Wind observation and forecast data provided by NOAA's National Weather Service (NWS).

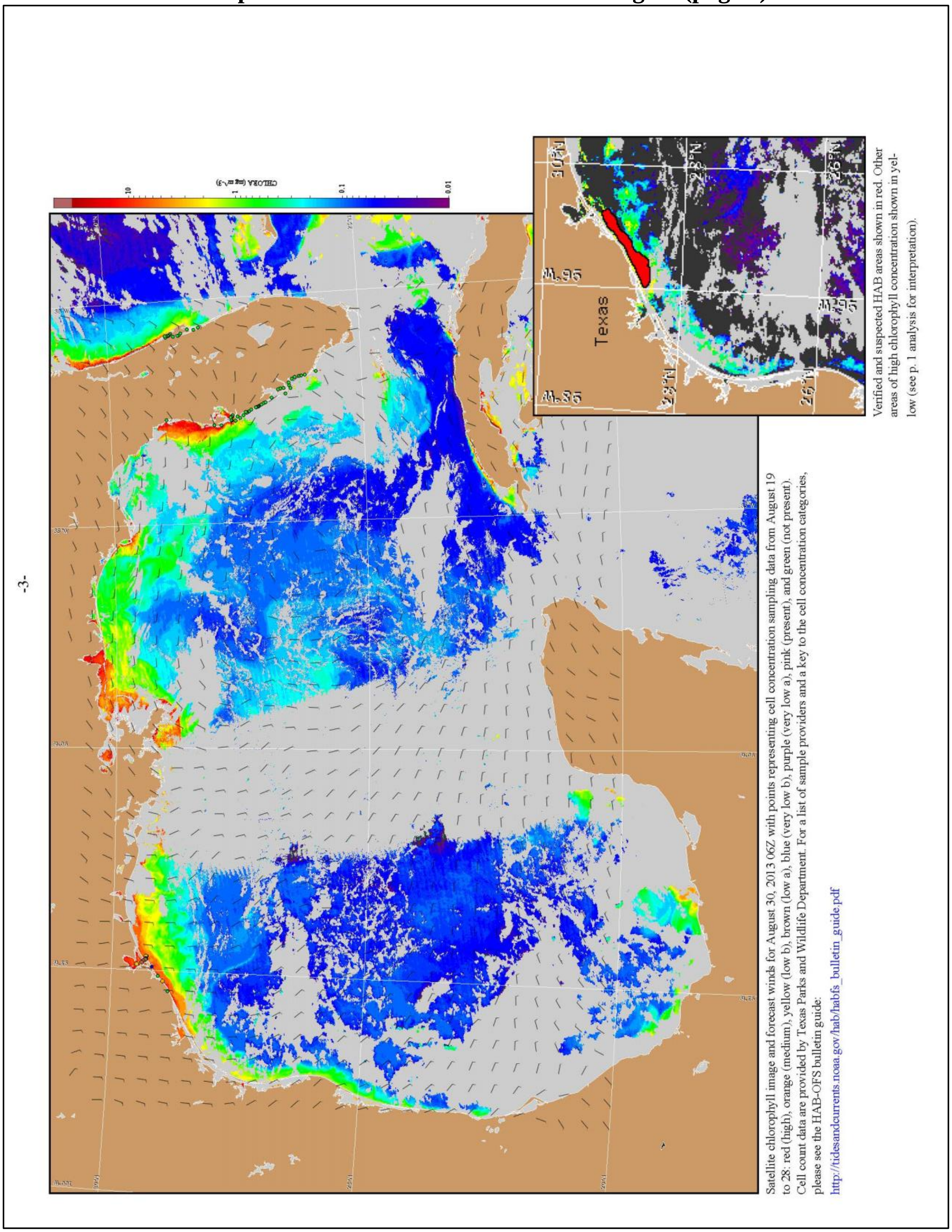


Wind Analysis

Galveston Island: Northwest winds (5kn, 3m/s) today becoming south in the afternoon. South winds (10kn, 5m/s) tonight. West winds (5kn) Friday becoming south in the afternoon. South winds (5-15kn, 3-8m/s) Friday night through Monday.

Port Aransas: Southeast winds (5-10kn, 3-5m/s) today becoming south winds (10-15kn, 5-8m/s) tonight. Southwest winds (5-10kn) Friday shifting to southeast winds in the afternoon. South winds (10-15kn) Friday night through Saturday. Southeast winds (10-15kn) Sunday through Monday.

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


-3-

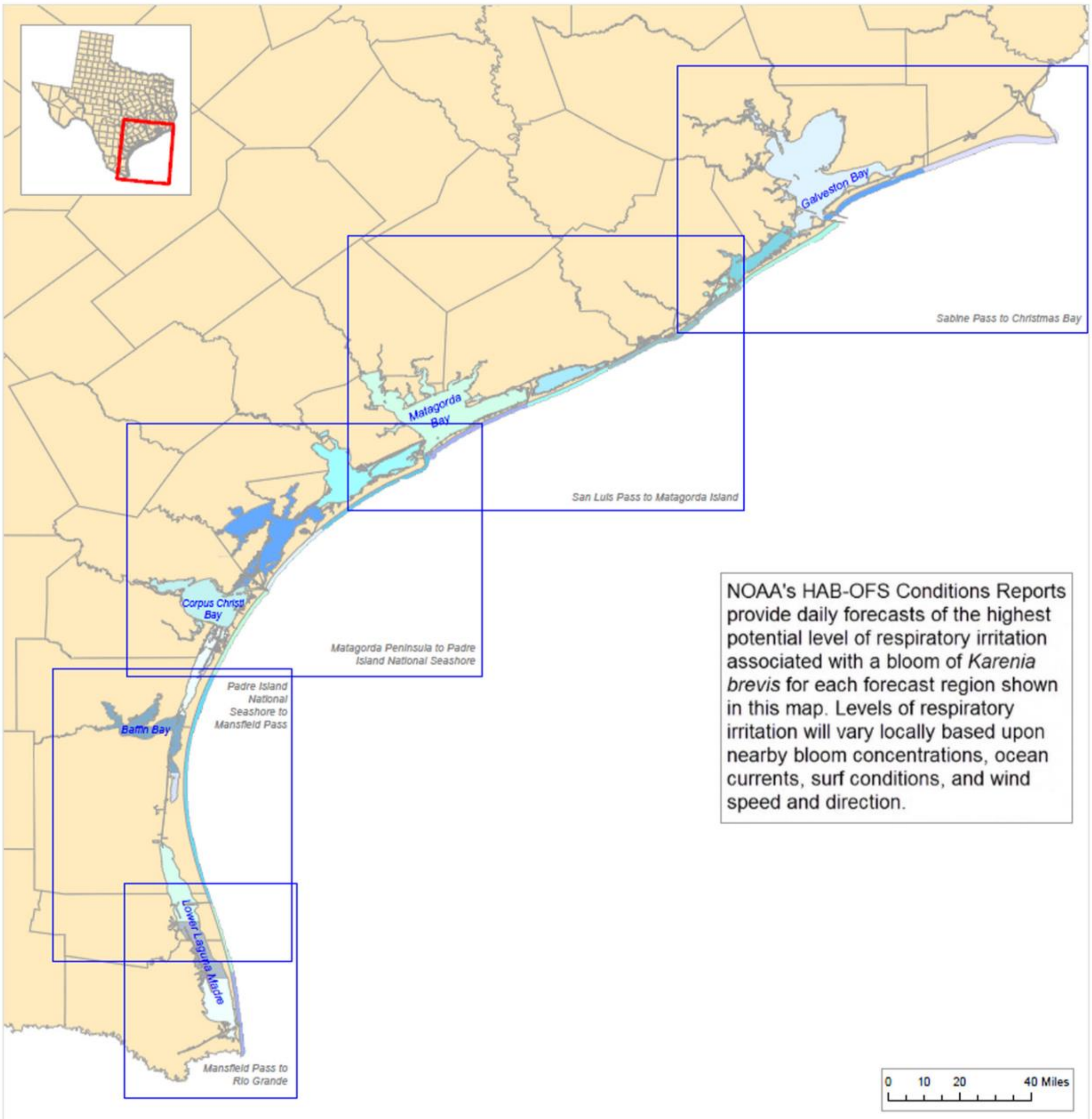
Satellite chlorophyll image and forecast winds for August 30, 2013 06Z with points representing cell concentration sampling data from August 19 to 28: red (high), orange (medium), yellow (low b), brown (low a), blue (very low b), purple (very low a), pink (present), and green (not present). Cell count data are provided by Texas Parks and Wildlife Department. For a list of sample providers and a key to the cell concentration categories, please see the HAB-OFS bulletin guide:

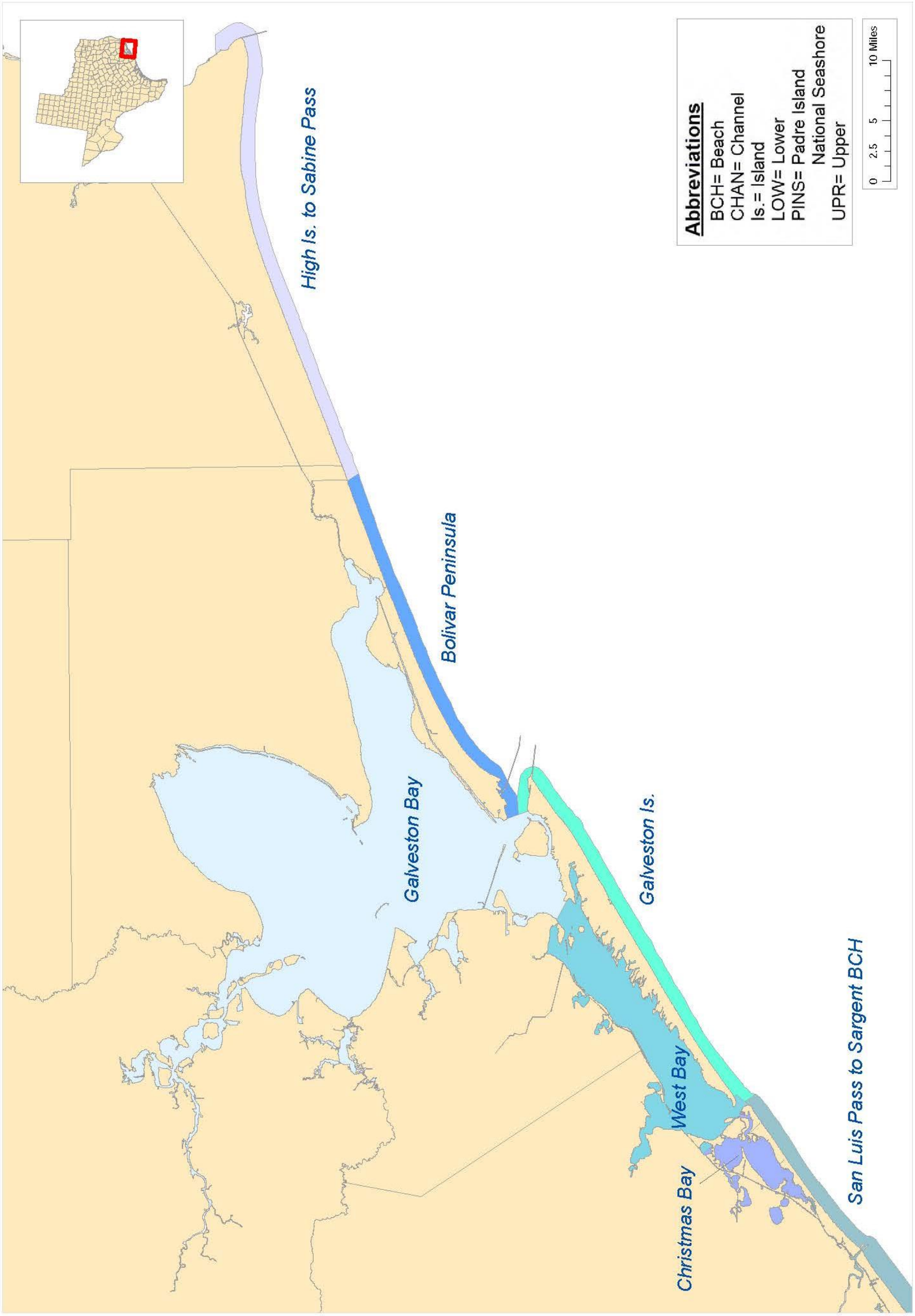
http://tidesandcurrents.noaa.gov/hab/habfs_bulletin_guide.pdf

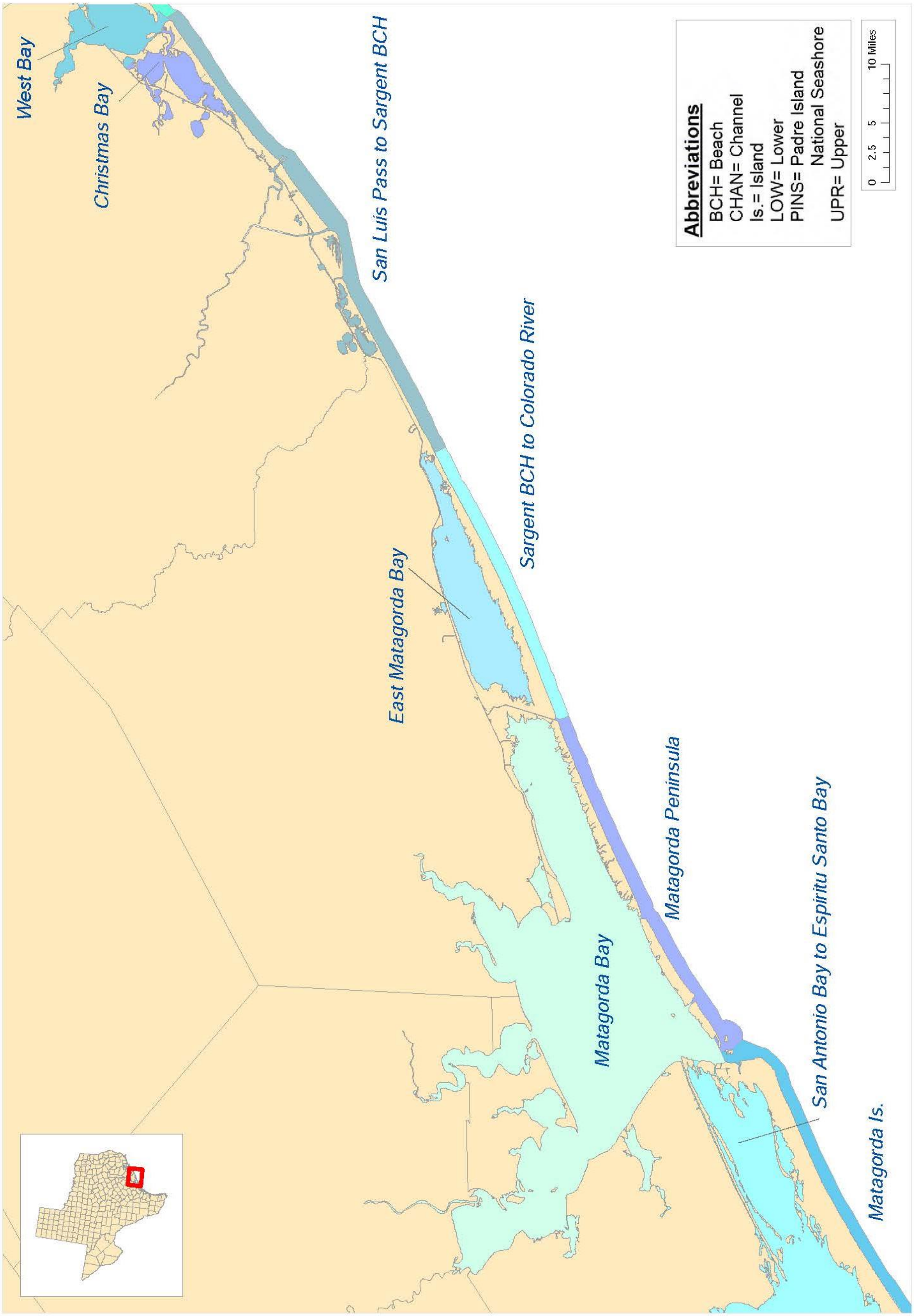
Verified and suspected HAB areas shown in red. Other areas of high chlorophyll concentration shown in yellow (see p. 1 analysis for interpretation).

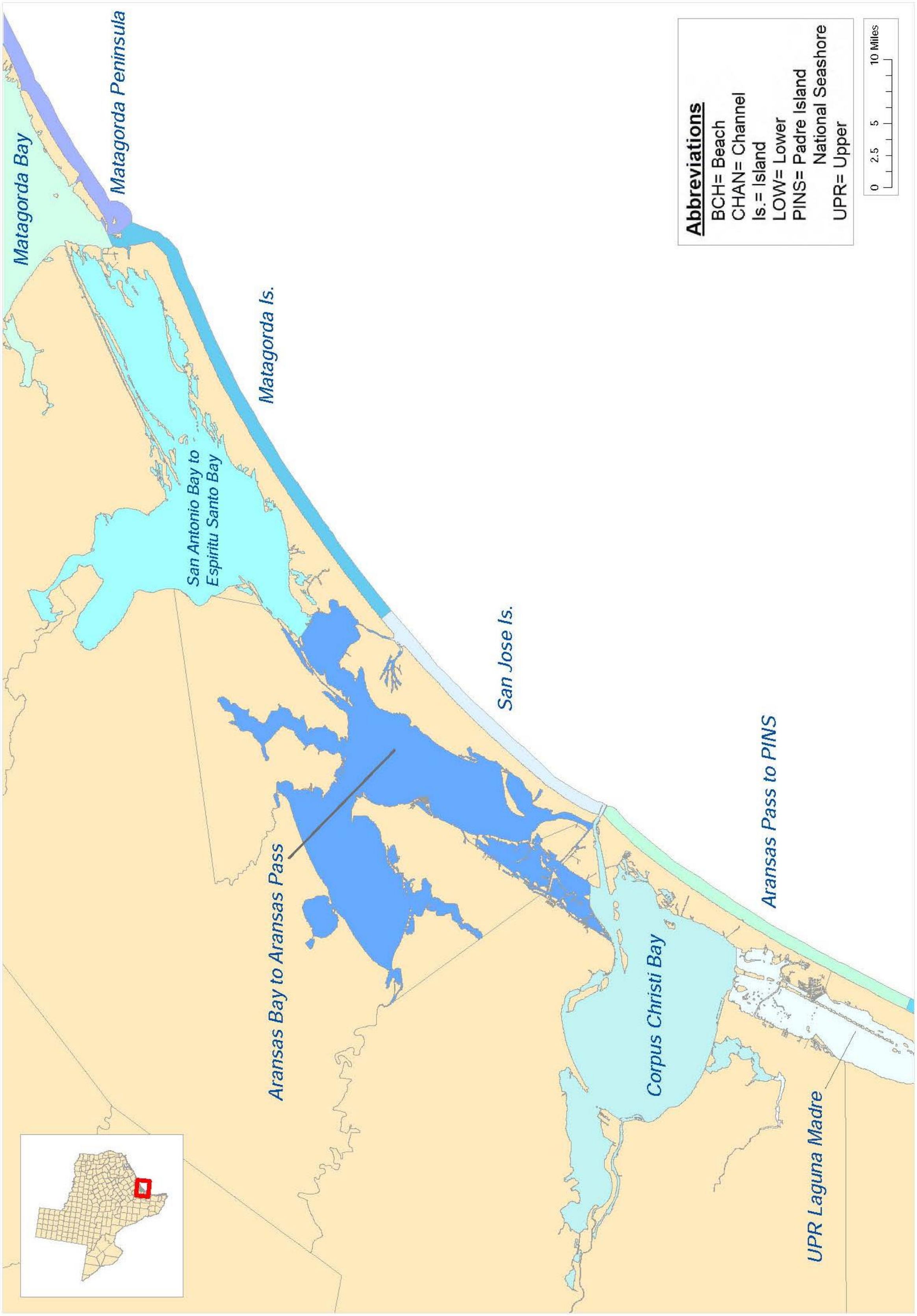
APPENDIX II

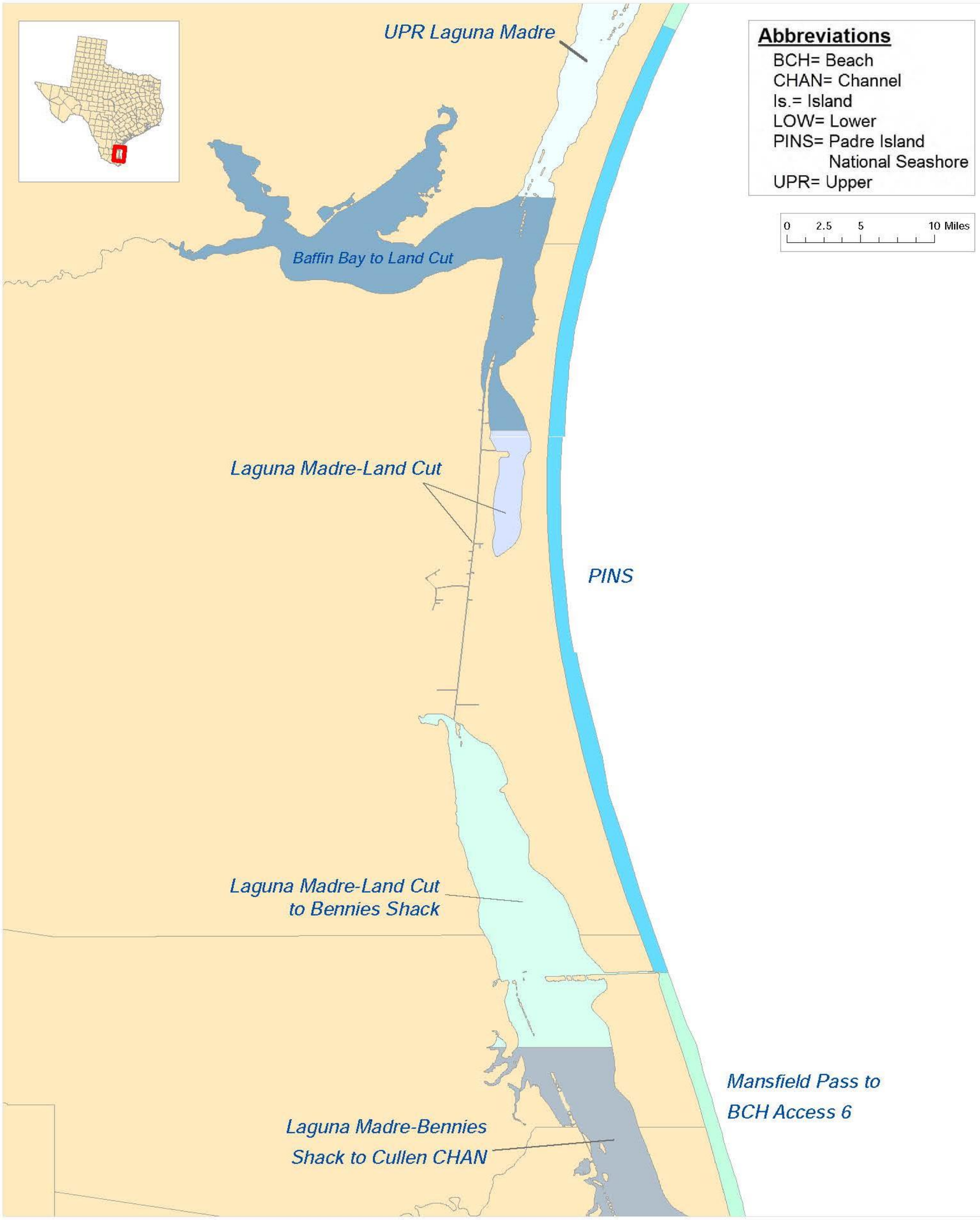
Forecast regions defined for Texas HAB bulletins.

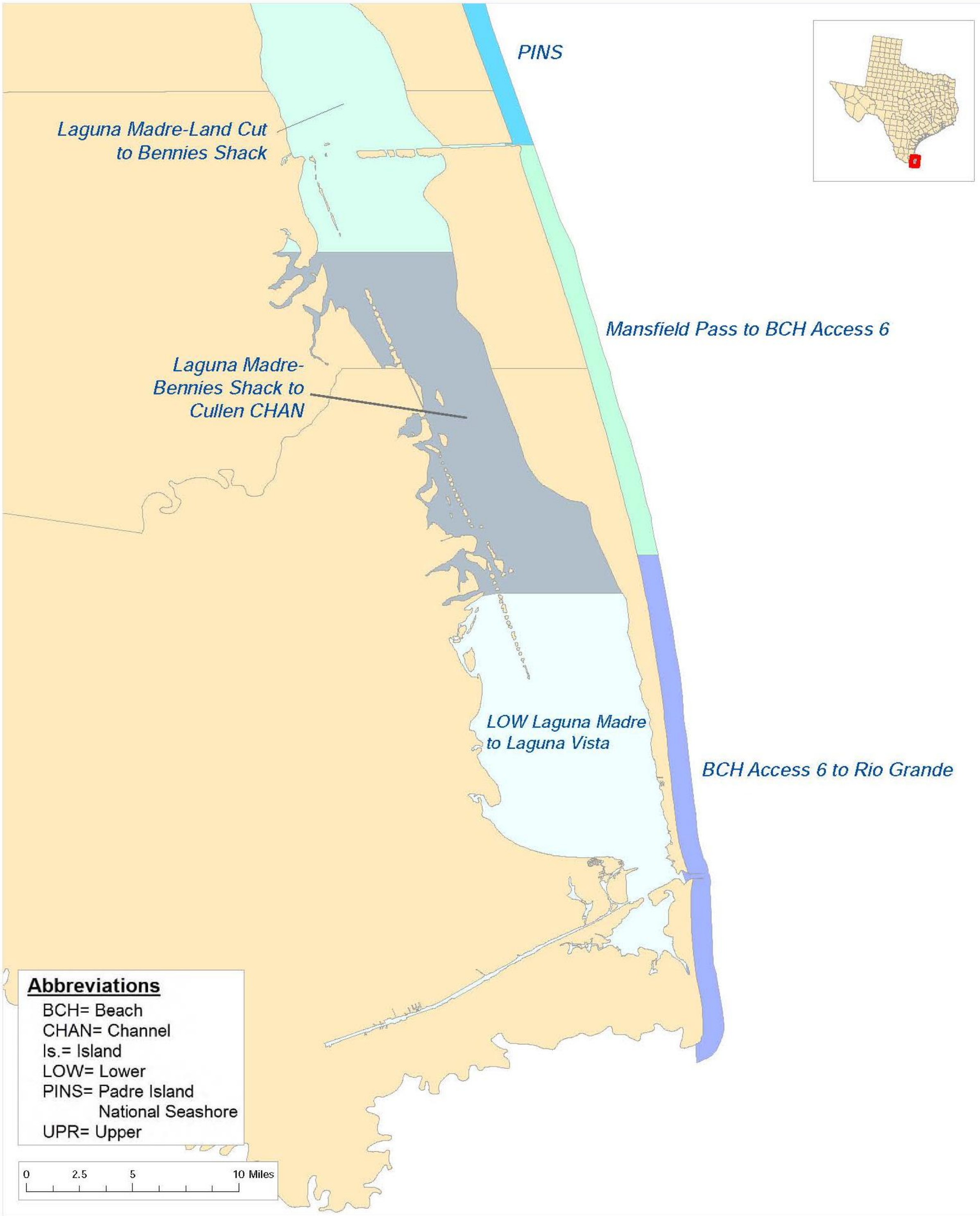



High Is. to Sabine Pass
Bolivar Peninsula
Galveston Bay
Galveston Is.
West Bay
Christmas Bay
San Luis Pass to Sargent BCH
Abbreviations
BCH= Beach
CHAN= Channel
Is.= Island
LOW= Lower
PINS= Padre Island
National Seashore
UPR= Upper
0 2.5 5 10 Miles


West Bay
Christmas Bay
San Luis Pass to Sargent BCH
Sargent BCH to Colorado River
East Matagorda Bay
Matagorda Peninsula
Matagorda Bay
San Antonio Bay to Espiritu Santo Bay
Matagorda Is.
Abbreviations
BCH= Beach
CHAN= Channel
Is.= Island
LOW= Lower
PINS= Padre Island National Seashore
UPR= Upper
0 2.5 5 10 Miles


Matagorda Bay
Matagorda Peninsula
Matagorda Is.
San Antonio Bay to Espiritu Santo Bay
San Jose Is.
Aransas Bay to Aransas Pass
Aransas Pass to PINS
Corpus Christi Bay
UPR Laguna Madre
Abbreviations
BCH= Beach
CHAN= Channel
Is.= Island
LOW= Lower
PINS= Padre Island National Seashore
UPR= Upper
0 2.5 5 10 Miles


Abbreviations
BCH= Beach
CHAN= Channel
Is.= Island
LOW= Lower
PINS= Padre Island National Seashore
UPR= Upper
0 2.5 5 10 Miles
UPR Laguna Madre
Baffin Bay to Land Cut
Laguna Madre-Land Cut
PINS
Laguna Madre-Land Cut to Bennies Shack
Mansfield Pass to BCH Access 6
Laguna Madre-Bennies Shack to Cullen CHAN


PINS
Laguna Madre-Land Cut
to Bennies Shack
Mansfield Pass to BCH Access 6
Laguna Madre-
Bennies Shack to
Cullen CHAN
LOW Laguna Madre
to Laguna Vista
BCH Access 6 to Rio Grande
Abbreviations
BCH= Beach
CHAN= Channel
Is.= Island
LOW= Lower
PINS= Padre Island
National Seashore
UPR= Upper
0 2.5 5 10 Miles

APPENDIX III

List of organizations that contributed to the 2010-2014 HAB-OFS bulletins for Texas. 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 2010-2014 HAB-OFS bulletins for Texas

Organization	HAB-OFS Contributions	Website
NOAA Center for Operational Oceanographic Products & Services	• Forecast analysis • Operations	http://tidesandcurrents.noaa.gov
NOAA National Centers for Coastal Ocean Science	• Research & Development	http://coastalscience.noaa.gov
NOAA National Weather Service	• Wind data	http://www.weather.gov
NOAA CoastWatch	• Remote sensing data	http://coastwatch.noaa.gov/cwn
NASA MODIS Aqua	• Remote sensing data	http://modis.gsfc.nasa.gov/
Texas Parks and Wildlife Department	• *In situ* cell count data • Water sample results including those collected by other organizations such as the Texas Red Tide Rangers, a Texas Sea Grant partner • Other reports of health impacts (i.e. respiratory irritation, dead fish or discolored water)	http://tpwd.texas.gov/
Texas A&M University	• Cell counts estimated by the Imaging FlowCytobot deployed in the pier laboratory at the University of Texas-Marine Sciences Institute located on the Port Aransas ship channel.	http://gcoos.org/products/index.php/bio/hab/
Texas General Land Office	• Observed surface wind current data from Texas Automated Buoy System (TABS)	http://tabs.gerg.tamu.edu/
Texas General Land Office/ Texas Automatic Buoy System/ Texas A&M University	• ROMS-based forecast model for surface currents in the Gulf of Mexico	http://seawater.tamu.edu/tglo/