



NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

AUTOMATED THUNDERSTORM FORECASTING VALIDATION

by

Joshua D. Carter

December 2020

Thesis Advisor:

Joel W. Feldmeier

Co-Advisor:

Daniel Tyndall,

Naval Research Laboratory

Second Reader:

John M. Peters

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			<i>Form Approved OMB No. 0704-0188</i>	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE December 2020		3. REPORT TYPE AND DATES COVERED Master's thesis
4. TITLE AND SUBTITLE AUTOMATED THUNDERSTORM FORECASTING VALIDATION			5. FUNDING NUMBERS	
6. AUTHOR(S) Joshua D. Carter				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.			12b. DISTRIBUTION CODE A	
13. ABSTRACT (maximum 200 words) Increased use of long-duration unmanned aerial vehicles (UAVs), as well as other factors, has led to a growing demand on the naval meteorology and oceanography (METOC) community for mission-tailored meteorological support. To creatively meet this need, the Naval Research Laboratory (NRL) Marine Meteorology Division has developed methodologies to process output from the Navy Global Environmental Model (NAVGEN). Automated products replicate human weather forecasters' written and unwritten rules to produce guidance for icing, turbulence, and thunderstorms. This is a force multiplier, as it takes the skilled, but few in number, human forecasters from being "in the loop" to being "over the loop." Much like the National Weather Service Aviation Weather Center's World Area Forecast System, utilizing Global Forecast System (GFS) output, NRL's existing automated thunderstorm algorithm uses convective precipitation and cloud thickness variable output to compute areas of expected thunderstorms. This thesis examines the current NRL method, and a few alternatives, in the Geostationary Satellite East coverage area and in two operationally relevant locations. Partial validation of forecasts is conducted using statistical analyses and comparison with satellite lightning observations. Results indicate that NRL's current methodology performs the best overall; new methodologies slightly lag behind.				
14. SUBJECT TERMS thunderstorm forecasting, NAVGEN, automation			15. NUMBER OF PAGES 87	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UU	

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

AUTOMATED THUNDERSTORM FORECASTING VALIDATION

Joshua D. Carter
Lieutenant Commander, United States Navy
BS, Embry-Riddle Aeronautical University, 2008

Submitted in partial fulfillment of the
requirements for the degree of

**MASTER OF SCIENCE IN METEOROLOGY AND PHYSICAL
OCEANOGRAPHY**

from the

**NAVAL POSTGRADUATE SCHOOL
December 2020**

Approved by: Joel W. Feldmeier
Advisor

Daniel Tyndall
Co-Advisor

John M. Peters
Second Reader

Wendell A. Nuss
Chair, Department of Meteorology

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Increased use of long-duration unmanned aerial vehicles (UAVs), as well as other factors, has led to a growing demand on the naval meteorology and oceanography (METOC) community for mission-tailored meteorological support. To creatively meet this need, the Naval Research Laboratory (NRL) Marine Meteorology Division has developed methodologies to process output from the Navy Global Environmental Model (NAVGEM). Automated products replicate human weather forecasters' written and unwritten rules to produce guidance for icing, turbulence, and thunderstorms. This is a force multiplier, as it takes the skilled, but few in number, human forecasters from being "in the loop" to being "over the loop." Much like the National Weather Service Aviation Weather Center's World Area Forecast System, utilizing Global Forecast System (GFS) output, NRL's existing automated thunderstorm algorithm uses convective precipitation and cloud thickness variable output to compute areas of expected thunderstorms. This thesis examines the current NRL method, and a few alternatives, in the Geostationary Satellite East coverage area and in two operationally relevant locations. Partial validation of forecasts is conducted using statistical analyses and comparison with satellite lightning observations. Results indicate that NRL's current methodology performs the best overall; new methodologies slightly lag behind.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I.	INTRODUCTION.....	1
II.	OVERVIEW	3
A.	TRADITIONAL THUNDERSTORM FORECAST METHODS	3
1.	K-index.....	3
2.	CAPE and CIN.....	4
III.	METHODS	7
A.	AUTOMATED THUNDERSTORM FORECAST METHODS	7
1.	Numerical Weather Model Based Thunderstorm Prediction.....	7
2.	Automated Forecasts	8
B.	FORECASTING ALGORITHMS	9
1.	NOAA/Cumulonimbus Algorithm	9
2.	FAST Algorithm.....	10
3.	K Index Algorithm.....	11
4.	CAPE Algorithm	11
5.	CIN Algorithm	11
6.	Modified CAPE Algorithm	11
C.	VERIFICATION METHOD	12
D.	TESTING REGIONS	13
1.	Contingency Tables.....	16
2.	Probability of Detection.....	16
3.	Proportional Correct	17
4.	False Alarm Rate.....	17
5.	Critical Success Index/Threat Score	17
IV.	ANALYSIS	19
A.	OVERALL RESULTS.....	19
B.	GLM EAST COVERAGE RESULTS	21
C.	VIRGINIA REGION RESULTS.....	24
D.	FLORIDA REGION RESULTS.....	28
E.	QUALITATIVE ANALYSIS.....	33
F.	ADDITIONAL NAVGEM VERIFICATION	42
V.	CONCLUSION	47
A.	GENERAL CONCLUSIONS	47

B.	CONTINUATION IDEAS	49
APPENDIX A. STATISTICAL DATA.....		51
A.	PROBABILITY OF DETECTION FOR ALL REGIONS.....	51
B.	PROPORTIONAL CORRECT FOR ALL REGIONS.....	51
C.	FALSE ALARM RATE FOR ALL REGIONS	52
D.	CRITICAL SUCCESS INDEX FOR ALL REGIONS.....	52
APPENDIX B. ADDITIONAL FORECASTS		53
A.	SEPTEMBER 07 00Z	53
B.	SEPTEMBER 07 18Z	55
C.	SEPTEMBER 11 18Z	57
D.	SEPTEMBER 13 18Z	59
E.	SEPTEMBER 22 00Z	61
LIST OF REFERENCES		63
INITIAL DISTRIBUTION LIST		67

LIST OF FIGURES

Figure 1.	Area of CAPE. Source: Stull (2000).....	5
Figure 2.	Area of CIN. Source: Stull (2000).	6
Figure 3.	GLM East coverage area.....	14
Figure 4.	Virginia region	15
Figure 5.	Florida region.....	15
Figure 6.	Forecast Contingency Table. Source: Wilks (2011).	16
Figure 7.	GLM East statistical averages.....	21
Figure 8.	GLM East probability of detection by forecast lead time (TAU).....	22
Figure 9.	GLM East proportional correct rates over time	23
Figure 10.	GLM East false alarm rates over time	23
Figure 11.	GLM East critical success index over time.....	24
Figure 12.	Virginia region statistical averages	25
Figure 13.	Virginia region probability of detection over time	26
Figure 14.	Virginia region proportional correct rates over time	26
Figure 15.	Virginia region false alarm rates over time.....	27
Figure 16.	Virginia region critical skill index over time.....	28
Figure 17.	Florida region statistical averages.....	29
Figure 18.	Florida region probability of detection over time	30
Figure 19.	Florida region proportional correct rates over time	31
Figure 20.	Florida region false alarm rates over time	32
Figure 21.	Florida region critical success index over time.....	33
Figure 22.	September 19, 2020 06Z Surface analysis for the eastern seaboard and western Atlantic	34

Figure 23.	September 19, 2020 06Z U.S. surface analysis with radar mosaic overlay.....	35
Figure 24.	September 19, 2020 06Z U.S. surface analysis with satellite overlay	36
Figure 25.	September 20, 2020 06Z GLM thunderstorm data	37
Figure 26.	September 19, 2020 06Z FAST thunderstorm forecast	38
Figure 27.	September 19, 2020 06Z K index thunderstorm forecast	39
Figure 28.	September 19, 2020 06Z CAPE thunderstorm forecast.....	39
Figure 29.	September 19, 2020 06Z CIN thunderstorm forecast	40
Figure 30.	September 19, 2020 06Z Modified CAPE thunderstorm forecast.....	40
Figure 31.	September 09, 2020 00Z FAST thunderstorm forecast	41
Figure 32.	September 09, 2020 00Z GLM thunderstorm data	42
Figure 33.	NAVGEM air temperature verification for September 2020	43
Figure 34.	NAVGEM relative humidity verification for September 2020	44
Figure 35.	NAVGEM mean sea level pressure verification for September 2020	45
Figure 36.	September 07, 2020 00Z Surface analysis with radar mosaic overlay	53
Figure 37.	September 07, 2020 00Z FAST thunderstorm forecast	54
Figure 38.	September 07, 2020 00Z GLM thunderstorm data	54
Figure 39.	September 07, 2020 18Z U.S. surface analysis with radar mosaic overlay.....	55
Figure 40.	September 07, 2020 18Z Modified CAPE forecast	56
Figure 41.	September 07, 2020 18Z GLM thunderstorm data	56
Figure 42.	September 11, 2020 18Z U.S. surface analysis with satellite overlay Source: Weather Prediction Center.....	57
Figure 43.	September 11, 2020 18Z FAST thunderstorm forecast	58
Figure 44.	September 11, 2020 18Z GLM thunderstorm data	58

Figure 45.	September 13, 2020, 18Z U.S. surface analysis with satellite overlay Source: Weather Prediction Center.....	59
Figure 46.	September 13, 2020 18Z FAST thunderstorm forecast	60
Figure 47.	September 13, 2020 18Z GLM thunderstorm data	60
Figure 48.	September 22, 2020 00Z U.S. surface analysis with satellite overlay Source: Weather Prediction Center.....	61
Figure 49.	September 22, 2020 00Z FAST thunderstorm forecast	62
Figure 50.	September 22, 2020 00Z GLM thunderstorm data 092200	62

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

Table 1.	K index values.....	4
Table 2.	GLM EAST.....	20
Table 3.	Virginia region.....	20
Table 4.	Florida region.....	20

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

CAPE	Convective Available Potential Energy
CB	Cumulonimbus Clouds
CIN	Convective Inhibition
CSI	Critical Success Index
DSRC	Department of Defense Super Computer Resource Center
FAR	False Alarm Rate
FAST	Forecast Automation System for Triton
GFS	Global Forecast System
GLM	Global Lightning Mapper
GOES	Geostationary Satellite
HPC	High Performance Computer
ISR	Intelligence, Surveillance, and Reconnaissance
LFC	Level of Free Convection
METOC	Meteorology and Oceanography
NAVGEN	Navy Global Environment Model
NCEP	National Center for Environmental Prediction
NRL	Naval Research Laboratory
NOAA	National Oceanic and Atmospheric Administration
PC	Proportional Correct
POD	Probability of Detection
SAS	Simplified Arakawa Schubert
UAV	Unmanned Aerial Vehicle

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to, first and foremost, acknowledge and thank Dr. Dan Tyndall and Dr. Joel Feldmeier, CDR, USN. Their support, guidance and mentorship were absolutely key to completing this research. Their knowledge and patient leadership were defining qualities that helped motivate me from the beginning of this project. I was thankful every day that they served as my advisors.

In addition to my advisors, I would like to thank Dr. John Peters for serving as my second reader. Your expertise and guidance greatly enhanced my research. I was always absolutely astonished with your depth of knowledge.

I would also like to thank the faculty and staff at the Naval Postgraduate School. Their knowledge and help were overwhelming during my period of study. There is not a doubt in my mind that they are among the best in the world.

My cohort also deserves to be acknowledged. The comradery we shared will never be forgotten. I will always be thankful for the times you picked me up when I was down and the laughs we shared together. I hope to consider you good friends for life.

Last, but certainly not least, I would like to thank my wife, Deborah, and my two sons, William and Benjamin. Your love and support during this time of “new normal” was always comforting and much needed. I always look forward to the next chapter in our lives with you all by my side.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

Thunderstorms pose a great risk to naval operations, whether they occur over land, in the air, or out to sea. Destructive winds, heavy precipitation, hail, lightning, and a myriad of other effects from thunderstorms can inhibit a mission, or even cause damage to equipment and personnel. One study of Navy and Marine Corps aviation mishaps involving aircrew found that 19 percent of aviation mishaps were weather related, and two-thirds of those mishaps were judged to be preventable by aircrew with a perfect forecast (Cantu 2001). Long-range unmanned aerial vehicles (UAVs) have become more prominent in naval operations. As stated in the Department of the Navy Strategic Roadmap for Unmanned Systems (Secretary of the Navy 2018), there is a growing chance for weather-related incidents “that have the potential to cause loss of control, loss of communication, diminished aerodynamic performance, and may negatively affect the operator, such as wind, turbulence, rain, solar storms, temperature extremes, humidity, snow and ice” (Ranquist 2017). Although termed unmanned, most UAV operations have remote human pilots and mission planners who require useful weather forecasts.

Therefore, it is paramount for naval weather forecasters to accurately predict destructive weather events in a timely manner to ensure operational success and safety. However, forecasts of thunderstorms for localized areas (hereafter “point forecasts”) can be problematic for the novice and experienced forecaster alike. Bouttier and Marchal (2020) even noted “point thunderstorm forecasts tend to have little predictive value beyond a few hours.” Naval forecasters face an even more complicated problem with thunderstorm forecasts as they must account for wide swaths of the globe, in compressed mission time frames, supporting a variety of platforms, to include submarines, ships, aircraft, and UAVs sometimes traversing long distances.

Navy leaders recognize that forecasting can be a time and manpower intensive process. Further, the Navy has indicated that it will use UAVs such as the Triton more and more extensively, in missions such as intelligence, surveillance, and reconnaissance (ISR). As Rear Admiral John Okon, Oceanographer of the Navy stated at the 100th Annual meeting of American Meteorological Society:

One key advancement in the Navy this year is that we brought our first Triton ISR tail into operations. With that comes a significant challenge with the forecast for that platform. We're really pressing automation with Triton on how we're producing those forecasts and working to move the forecaster from "in the loop" to "on the loop." The way the Navy is moving with unmanned ISR we will quickly outgrow the forecast capability, the human organic forecast capability, if we don't figure this out. (Okon 2020)

Facing this dilemma, the Naval Research Laboratory (NRL) in Monterey, California, has developed a system called the Forecast Automation System for Triton (FAST). FAST helps forecasters brief mission planners on the many weather hazards that are faced during UAV operations. FAST output, including pre-production of (modifiable) briefing and planning products, is a force multiplier. A forecaster can now add context and value to an already meteorologically sound product, rather than constructing a forecast brief from scratch. This automated system uses Navy Global Environmental Model (NAVGEM) data fields to produce products. Some of the products are easily accessed, as they are direct outputs from the model, such as temperature and winds. Others, such as thunderstorms, aircraft icing, and turbulence related forecast parameters, must be post-processed from directly output model variables. As these algorithms come into operational use, a thorough evaluation of their accuracy is important considering the platforms they will be supporting.

While automated thunderstorm forecasting is a new technique in the Navy, not all automated forecasting software is new. The Automated Tropical Cyclone Forecast (ATCF) system, also developed at NRL, has been used by the Joint Typhoon Warning Center for decades with great success (Sampson and Schrader 1999). This system provides many automated features to a tropical cyclone forecaster, including the ability to construct consensus model tracks by blending multiple types of models, which has greatly improved forecasts throughout the years. This also reduces the workload of the forecaster, who must process a lot of data in a short amount of time. FAST is expected to have similar long-term benefits.

II. OVERVIEW

A. TRADITIONAL THUNDERSTORM FORECAST METHODS

As previously stated, thunderstorm forecasting can be difficult, owing to the rapid nonlinear error growth in numerical models that occurs in the presence of moist convection (Zhang et al. 2003). As Stull (2011) states, “Forecasting thunderstorms is not easy. Thunderstorm processes are very nonlinear (e.g., thunderstorms grow explosively), and are extremely dependent on initial conditions such as triggering, shear, and static stability. Individual storms are relatively short lived (15 to 30 min) and are constantly changing in intensity and movement during their lifetimes.” Another reason for this difficulty is that there is no one clear-cut method for predicting these events. Every forecaster has a preferred method at their location. These techniques can range from the calculating various indices, to evaluating upper air soundings, or interpreting model output. While each forecasting technique has its strengths and weaknesses, this paper will be focused on techniques that can be automated and are limited to the current FAST NRL thunderstorm prediction algorithm, CAPE and CIN output from NAVGEM model fields, a modified CAPE evaluation, and the K index.

1. K-index

The K-index has been used by forecasters to produce thunderstorm forecasts for decades. It has been noted in forecasters’ handbooks as having the best accuracy in maritime environments, making it of high interest for naval forecasters (Mireles et al. 2007). The K Index is a measurement of thunderstorm potential based on the vertical temperature lapse rate between 850, 700 and 500 hPa. It also encompasses the amount, and the vertical extent of, low-level moisture in the atmosphere by using the temperature dewpoints from 850 and 700 hPa. It is calculated by using the formula:

$$K = T(850 \text{ hPa}) + Td(850\text{hPa}) - T(500 \text{ hPa}) - DD(700 \text{ hPa})$$

where T represents temperature, Td represents dewpoint temperature, and DD represents dewpoint depression (i.e., $T - Td$) at the indicated pressure levels. In general, the higher the

K index value, the greater the potential for thunderstorms with heavy rain. Typical threshold values are listed in Table 1 (NOAA 2015).

Table 1. K index values

K below 30:	Thunderstorms with heavy rain or severe weather possible.
K over 30:	Better potential for thunderstorms with heavy rain.
K $\geq$ 40:	Best potential for thunderstorms with very heavy rain.

K index values are typically large when there is cold air aloft (500hPa) and warm air at low-levels (850hPa), which is consistent with small lower-tropospheric static stability (not shown). Additionally, the presence of dry air aloft (700 hPa) will result in lower K index values than when there is comparatively moist air aloft (NOAA 2015). This is consistent with the idea that the entrainment of dry air aloft by cumulus congestus clouds reduces their chances of growing into thunderstorm-producing cumulonimbus clouds (Morrison 2017; Morrison et al. 2020; Peters et al. 2020c). The temperature and dew point that are required to calculate the K index are typically collected by rawinsondes or produced by model output, which can be advantageous in that these values are usually easy to obtain and calculate.

2. CAPE and CIN

When examining upper air soundings, one indicator that forecasters use for predicting thunderstorms is Convective Available Potential Energy (CAPE). CAPE represents the vertical integral of the positive buoyancy an un-diluted air parcel would achieve if were lifted to its equilibrium level (EL; Stull 2000). Nonzero CAPE is a necessary, but not sufficient, condition for thunderstorms (Doswell et al. 2006). To reach the area of positive buoyancy, a lifted air parcel must overcome the negatively buoyant region of convective inhibition (CIN) below the LFC that usually extends to the surface. See Figures 1 and 2 for a visual depiction of CAPE and CIN. On these diagrams, the pink and green shaded areas represent the respective amount of CAPE or CIN. Without an

upward oriented pressure gradient acceleration, an air parcel will not be able to overcome the CIN and thus, a thunderstorm will not form. Upward pressure gradient accelerations that allow a parcel to surpass the LFC and continue out of the CIN region could come from a multitude of sources such as orographic lifting, fronts (Stull 2000), and outflow boundaries (e.g., Rotunno et al. 1988). Furthermore, localized surface heating may locally erode CIN and allow parcels to reach their LFCs uninhibited. However, CAPE and CIN calculations solely do not account for these sources of upward acceleration; this is usually taken into consideration by the forecaster using other diagnostics. There are other factors to consider when evaluating CAPE and CIN for thunderstorm potential as well. If a parcel is lifted past the LFC into the positively buoyant region, in order to support thunderstorm development, the upper portion of the atmosphere must be moist enough, otherwise entrainment will dry out the parcel and cease the convective processes. Finally, strong vertical wind shear may initially inhibit thunderstorm growth (Peters et al. 2019a), but once convection matures, shear results in stronger and more sustained thunderstorms (Peters et al. 2019b). Calculations of CAPE and CIN require more information (i.e., a vertical profile of the atmosphere) and require a more complicated calculation than K index.

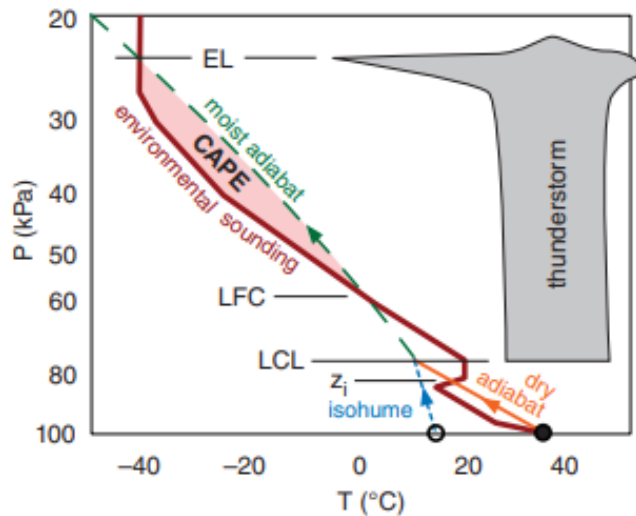


Figure 1. Area of CAPE. Source: Stull (2000).

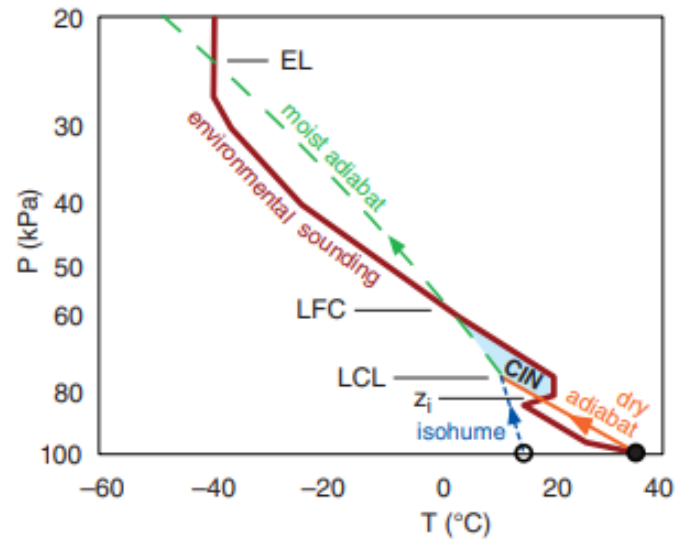


Figure 2. Area of CIN. Source: Stull (2000).

III. METHODS

A. AUTOMATED THUNDERSTORM FORECAST METHODS

1. Numerical Weather Model Based Thunderstorm Prediction

The above referenced traditional thunderstorm techniques typically require observations at short time leads on very local spatial scales, or large quantities of output variables from numerical weather prediction models. For decades, the resolution of global numerical weather prediction models was too coarse for the models to explicitly resolve thunderstorms in a realistic way. This made direct model output of thunderstorm potential or thunderstorm related variables have questionable use for operational naval forecasting. However, years of model improvements have led to global models resolving features on a scale that would have been unthinkable to most naval forecasters just a few years ago, making the application of human forecaster rules to model output more feasible (CNMOC 2020).

With these improvements, forecasting thunderstorm activity using model output has become common practice for forecasters. With a plethora of model fields to choose from, forecasters have their choice of what to base their forecast off of. Vorticity, humidity levels, precipitation values, and even forecast simulated radar reflections are some examples of these fields and are usually combined to make a more accurate forecast. However, doing so still requires a thorough knowledge of the area that the forecast is being issued for. Currently, the highest horizontal resolution of numerical weather prediction models, which is on the order of 1 km, is still insufficient for the proper resolution of processes that are internal to thunderstorms (e.g., Bryan and Fritsch 2003). Models run with a horizontal grid spacing greater than 4 km must parameterize convection (e.g., Weisman et al. 1997). Additionally, knowledge of local geographic features by forecasters is important. This is particularly true for features that are not explicitly resolved by forecast models. These small features (e.g., hills, mountains, lakes, and others) can easily be overlooked to the detriment of a forecast. This is especially likely when a forecaster is concerned with large areas of coastline (or even forecasts on a continental scale) as naval

forecasters typically must do for varied mission support. It is anticipated that a human forecaster “over-the-loop” could still add value above straight model output in these cases.

2. Automated Forecasts

Automated forecasting is a new method that takes model data fields and makes more detailed forecasts in what is commonly termed as postprocessing. During this postprocessing, output from the numerical models is combined with simple equations or decision rules to give categorical or numerical forecasts. These forecasts do not require any human input, other than the initial programming required to make the simple equations or decision rules that the forecasts are based on. Human forecasters can review the output, and accept, reject, or modify the forecast.

Automated thunderstorm forecasts can also be a blend of traditional thunderstorm forecasting methods with model data and computer calculations. That is, the insights gained from decades of aviation related thunderstorm forecasting can be partially distilled into “rules of thumb,” and those rules can be programmed into the selected post-processing techniques. For example, as noted in Table 1, K indices greater than 40 indicate the highest potential for thunderstorms. That K index value comes from human forecaster experience with thousands of storms and can be made an empirical decision rule in automated forecasting. A simple procedure for doing this could consist of model output being used to calculate K index, and the decision to forecast thunderstorms being applied across a geographic area where the calculated K index exceeds 40. Such automation can be advantageous because it can be done quickly across large areas and requires no human intervention. Note, all the disadvantages of using the traditional methods and model output still exist, which include insufficient spatial and temporal resolution. Human intervention and flexible rules on how to do thunderstorm forecasts may be necessary for the foreseeable future.

To compare the results of automated thunderstorm forecasts using different techniques in postprocessing, and to be consistent with current methodology at NRL, NAVGEM model data fields were used with all forecasting methods considered here. NAVGEM is the U.S. Navy’s global numerical weather prediction model that is run

operationally at the Fleet Numerical Meteorology and Oceanography Center in Monterey, California. It has a three-time level semi-Lagrangian and semi-implicit dynamical core employing spherical Legendre polynomials for computation of horizontal derivatives, and fourth order diffusion (FNMOC 2015). The model is run every six hours at synoptic times (00, 06, 12, and 18 UTC) with output at 0.25 and 0.5 degree resolution. The approximate grid spacing of the model is 19 km (CNMOC 2020). NAVGEM's convective parametrization scheme is a simplified Arakawa-Schubert (SAS) scheme for large scale precipitation and uses the National Center for Environmental Prediction Global Forecast System (GFS) parameterization for shallow cumulus (FNMOC 2015). The convective parametrization and shallow cumulus scheme estimate the transport and microphysical properties of convection that are not explicitly resolved in the dynamical core of the model.

B. FORECASTING ALGORITHMS

1. NOAA/Cumulonimbus Algorithm

The National Weather Service's Aviation Weather Center has developed an automated thunderstorm forecast system that focuses on producing fields of thunderstorms and thunderstorm-caused cloud cover. The method uses the Global Forecast System (GFS) forecast fields with a resolution of 1.25 degrees which provide 3-hour precipitation rates, cloud top pressure, and cloud bottom pressure from the cumulus parameterization scheme. They calculate convective cloud cover (C_c) from the aforementioned variables using the formula:

$$C_c = 0 \text{ for } P < 1.62$$

$$C_c = a + b \ln P \text{ for } 1.62 < P < 984$$

$$C_c = 0.8 \text{ for } P > 984$$

where P is precipitation rate from the cumulus parameterization scheme [$\text{kg m}^{-2} \text{s}^{-1} 10^{-6}$] and the constants are $a = 0.06$ and $b = 0.125$. To ensure the calculated convective cloud is deep enough to be a cumulonimbus cloud (CB), these additional conditions were imposed:

1. The top of the convective cloud must be above 400 mb
2. The depth of the cloud must be more than 300 mb.

The heights of cloud bases and tops are calculated using a standard atmosphere definition, e.g., the NOAA Standard Atmosphere (COESA 1976). The output of the algorithm is: CB cover, height and pressure at the base of CB, height and pressure at the top of CB, height and pressure at the bottom of embedded CB, height and pressure at the top of embedded CB (Trojan 2007). While this algorithm is a good example of what automated forecast are available, it is considered an experimental product meant for visual comparison to commercial flight planning software, and is not operationally approved (B. Pettegrew, NOAA, 2020, personal communication). Therefore, it was not an algorithm that was compared in this study.

2. FAST Algorithm

The current thunderstorm algorithm within FAST uses existing NAVGEM convective precipitation, cloud base, and cloud top output from the SAS scheme to determine the presence of thunderstorms. SAS precipitation fields are operationally produced by FNMOC; cloud base and cloud top fields are currently supplied by a development run of NAVGEM. The currently implemented FAST algorithm diagnoses thunderstorms when the 3-hour SAS precipitation total is greater than 2.5 mm and when the distance between the cloud top and the cloud base is at least 3,000 m at a NAVGEM grid point. The relationship used by this algorithm is similar to the one used for aviation forecasts at the World Area Forecast Center (WAFC) Washington. However, the WAFC imposes an additional constraint in which cloud tops must exceed 24,000 ft. Thunderstorm tops within the FAST application are set to the NAVGEM cloud tops (D. Tyndall, NRL, 2020, personal communication). While NAVGEM does have the capability to output quarter-degree spatial resolution fields, the currently implemented FAST algorithm uses the half-degree resolution for the initial operating capability of the software. The other forecast algorithms examined here utilized the increased resolution model data.

3. K Index Algorithm

The K Index algorithm calculates the K Index value at each grid point of the NAVGEM model using the forecasted temperatures, dewpoint temperatures, and dewpoint depressions at the applicable pressure levels. After the K Index is computed, a grid point is identified as having a thunderstorm when the K index value is greater than 35. This threshold value was determined by using standard values of K Index for determining thunderstorms and upon examination of K Index values compared to satellite observations (NOAA 2015).

4. CAPE Algorithm

The CAPE algorithm uses CAPE fields that are produced as part of the operational NAVGEM production cycle. Any grid point exceeding 2,000 J/kg of CAPE is identified as having a thunderstorm. The 2,000 J/kg threshold value was determined by standard forecasting values for thunderstorm prediction (Rauber et al. 2002) and was also compared to satellite and surface analyses observations for consistency. Although not undertaken in this study, a future examination of algorithm performance with lower thresholds of CAPE may be useful.

5. CIN Algorithm

The CIN algorithm uses CIN fields that are produced as part of the operational NAVGEM production cycle. Any grid point which has CIN values greater than -250 J/kg is labeled as a storm. The -250 J/kg threshold value was based upon Thompson et al. (2007) and Colby (1984). Further constraints may have been needed for this algorithm. See the analysis chapter for more discussion.

6. Modified CAPE Algorithm

The modified CAPE algorithm was based off of a formula from Peters et al. (2020b) that calculates “entrainment CAPE.” This calculation accounts for the entrainment effects of dry air at mid-levels of the atmosphere by using a relationship with vertical velocity, updraft width in a thunderstorm, and environmental wind shear (Peters et al. 2020b). In environments with strong shear, updrafts are wide and the parcels rising through the center

of updrafts are largely undiluted (e.g., Peters et al. 2019b, 2020a, 2020b). In these instances, modified CAPE will be very close to standard CAPE calculations. In contrast, when shear is weak, updrafts are narrow and parcels that are rising through these updrafts experience considerable entrainment-driven dilutions. In these instances, modified CAPE will be much less than standard CAPE calculations reflecting the deleterious effects of entrainment on updraft buoyancy. Finally, modified CAPE will be closer to standard CAPE when middle tropospheric relative humidity is large, and smaller than standard CAPE when middle tropospheric relative humidity is comparatively small. This is because entrainment-driven dilution of updraft buoyancy is larger when relative humidity is smaller.

To calculate this formula, the algorithm constructs an atmospheric sounding at each grid point by using the NAVGEM model fields to obtain the environmental variables needed. After these variables are acquired, the calculation is conducted. The threshold for this algorithm was set at 1,000 J/kg based on a comparison of satellite data and observations made by a few sample forecasts prior to the conducting the forecast used in this research.

C. VERIFICATION METHOD

Thunderstorm forecasts made by the algorithms were generated from archived NAVGEM 0–72 hour forecast fields between 0000 UTC 31 August to 1800 UTC 20 September 2020. A determination of a thunderstorm was made based on the threshold values set for each algorithm and were mapped on the NAVGEM forecast grid. The FAST and K index algorithm were able to be run every 3 hours as the data for these algorithms was available at this temporal resolution. The CAPE and CIN NAVGEM fields were only available every 6 hours and the modified CAPE algorithm was only run every 6 hours as well due to the computational cost of the algorithm.

Data from the Global Lightning Mapper (GLM) sensor on GOES-East was used to validate the preceding forecast algorithms using lightning as confirmation of a thunderstorm. The GLM validation data consists of observed lightning flashes accumulated over a ± 60 -minute period about the validation time for forecast data. Accumulated GLM flashes were mapped on to the NAVGEM forecast grid (either the half degree or quarter degree resolution grid, depending on the resolution of the input data used by each

algorithm) as the GLM flashes are discrete latitude/longitude points. If any lightning flashes were observed at a grid point, that grid point was marked as an inferred thunderstorm during the time period. The ± 60 -minute period was picked to account for the temporal spacing in the model data, which has a 3 to 6-hour step between forecast times.

To validate the results of the forecast data, the forecast grids from the algorithms were compared to the inferred thunderstorm grids from the GLM data grid point by grid point. To account for phasing errors, hits were scored if an inferred thunderstorm event occurred within 2 grid points of thunderstorms forecasted using the FAST method, and within 4 grid points of thunderstorms forecasted using all other methods. The difference between these values is due to the resolution of the NAVGEM data files between the different methods; a distance of 2 grid points in the FAST forecast grids is approximately the same as 4 grid points in the other methods due to the coarser resolution input fields used by the FAST method.

D. TESTING REGIONS

The algorithms were evaluated over the entire coverage area of the GLM sensor (seen in Figure 3) as well as two smaller regions, chosen for their proximity to military aviation bases, greater expected availability of surface weather reports and analysis, and for their climatologically different weather regimes. Both of the smaller regions are located on the East Coast of the United States and henceforth will be referred to as the Virginia region (Figure 4) and the Florida region (Figure 5) while the coverage area of the GLM sensor will be referred to as GLM East. The entirety of GLM East area was used due to the availability of the sensor to observe lightning across the broad area and due to the similarity of the area to large area forecast that naval forecasters frequently make. The GLM coverage area was set to the same coverage area as the advanced baseline imager of the GOES as the exact coverage of the GLM sensor could not be ascertained from the data files. The Virginia region was expected to experience more mid-latitude type weather patterns while the Florida region was expected to have more tropical/sub-tropical weather. The GLM East area has all weather patterns within its hemispherical coverage. As currently written, the modified CAPE algorithm is not optimized for a large number of grid point forecast that

would be required for the GLM East area. This makes modified CAPE calculations quite lengthy and thus that particular algorithm was only run in the smaller regions.

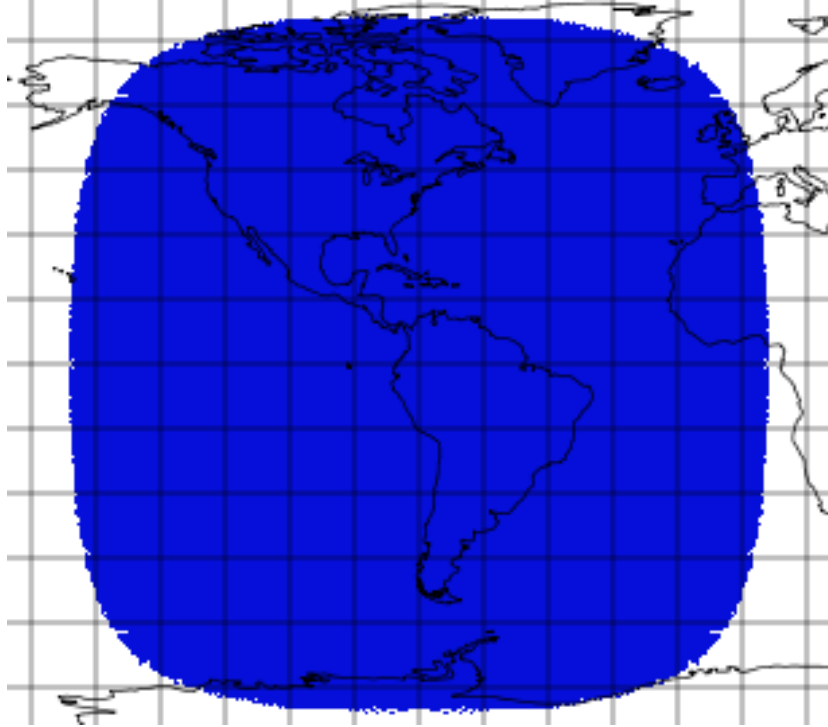


Figure 3. GLM East coverage area

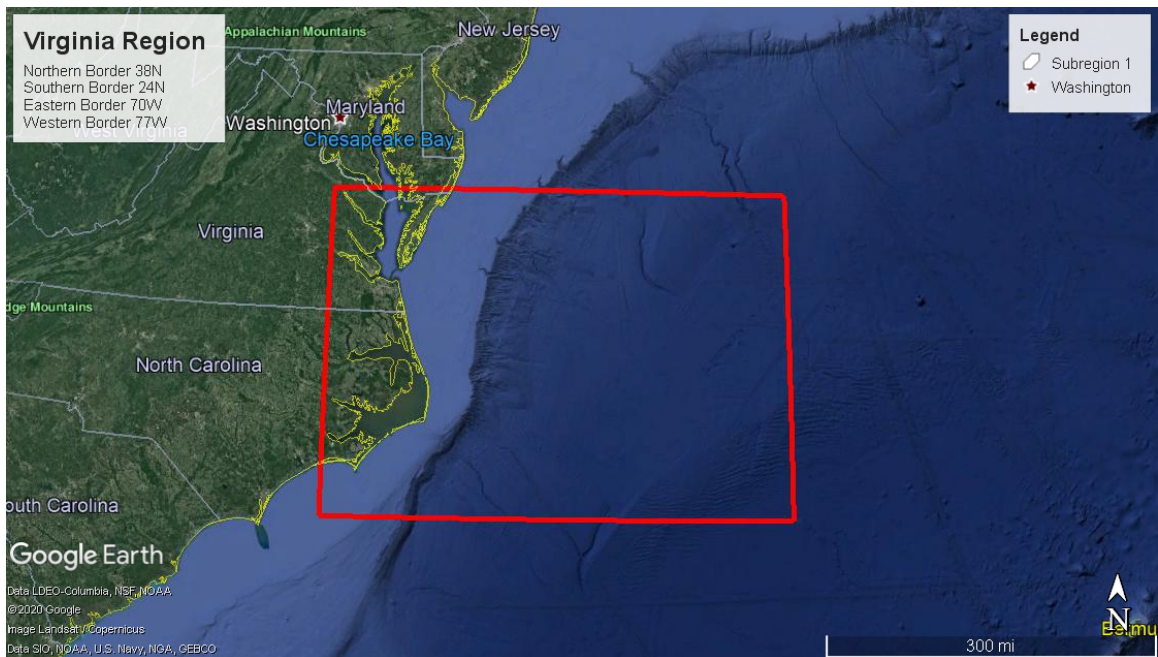


Figure 4. Virginia region

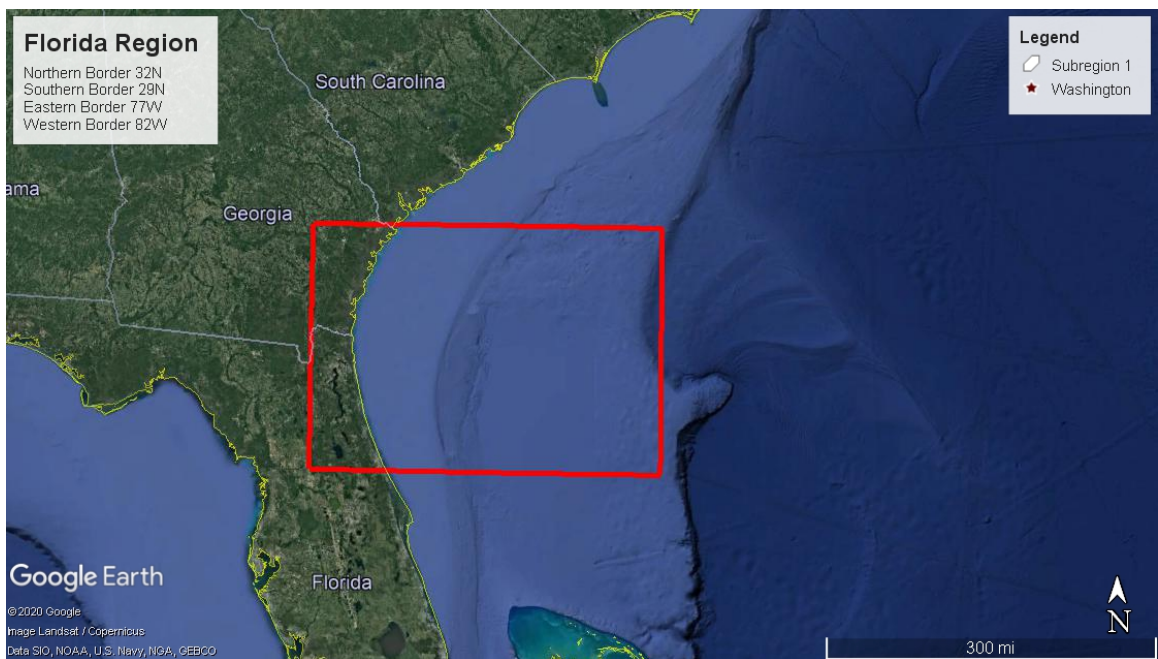


Figure 5. Florida region

E. VERIFICATION STATISTICS

1. Contingency Tables

To compute the statistics for the evaluation of the algorithms' performance, a simple two by two contingency table method was used (Figure 6). The left column of the table represents when the algorithms forecasted a thunderstorm, while the top row is when a thunderstorm was inferred from GLM data. Position a is when the algorithm was true (i.e., made a positive categorical thunderstorm forecast) and corresponded with an inferred thunderstorm from the GLM observations. This is a correct forecast and is referred to as a hit. Position b is when the algorithm predicted a thunderstorm, but one was not observed; this is known as a false alarm. Position c is when a thunderstorm was observed but was not forecast and is called a miss. Position d is when no thunderstorm was forecast and none were observed, also known as a non-event or a correct negative (Wilks 2011).

		Observation	
		YES	NO
Fore- cast	YES	a correct event forecast	b false alarms
	NO	c surprise events	d non-events

Figure 6. Forecast Contingency Table. Source: Wilks (2011).

2. Probability of Detection

Probability of detection (POD) is a ratio of correctly forecasted lightning events, to the total number of observed lightning events. The highest POD is 1 and lowest is zero. POD can be calculated by using the formula (Wilks 2011):

$$\text{POD} = a/(a+c).$$

3. Proportional Correct

Proportional correct (PC) is the ratio of correct forecasts to the total number of forecasted or observed lightning events. The best PC is one, while the worst is zero and can be found by using the formula (Wilks 2011):

$$PC = (a+d)/(a+b+c+d).$$

4. False Alarm Rate

The false alarm rate (FAR) is the ratio of falsely predicted lightning events to the total number of events without observed lightning. Generally, a small FAR is to be desired, but a low FAR can sometimes come at the expense of POD. Thus, the acceptable threshold for FAR is situationally dependent. FAR can be calculated by the formula (Wilks 2011):

$$FAR = b/(b+d).$$

5. Critical Success Index/Threat Score

The critical success index (CSI), also known as the Threat Score, is the number of correct event forecasts divided by the total number of occasions on which the event was forecast and observed. It is the proportional correct forecast for the number of forecasts with the correctly forecast non-event removed. The CSI is particularly beneficial when the hit forecast occurs less often than the non-events, as it implicitly minimizes FAR while maximizing the POD. The worst CSI is zero, while a perfect score is one, and can be found by using the formula (Wilks 2011):

$$CSI = a/(a+b+c).$$

THIS PAGE INTENTIONALLY LEFT BLANK

IV. ANALYSIS

A. OVERALL RESULTS

After comparing the algorithms of interest, certain patterns become evident. The FAST algorithm is the best overall performer, with scores at or above average compared to the other methods, in all contingency table based metrics (explained above). FAST also has the lowest FAR and highest CSI in each area tested. CIN had the highest POD in each region, but also had the highest FAR. We thus infer that the threshold for CIN was set incorrectly and thunderstorms were highly overpredicted. Additionally, it was found CIN should be used in areas where CAPE is present, which additionally accounted for the erroneous results. While, still included in this analysis, the results of this algorithm should be considered with caution. Using K index yielded mixed results. The K index method showed above average results in all categories in the GLM East coverage area, an above average PC score for the Virginia region, and above average for all metrics, except POD in the Florida region. Modified CAPE had the closest performance to the FAST algorithm and scored the highest PC in the Florida region. Tables 2–4 summarize the main results.

Table 2. GLM EAST

GLM East Averages							
	CAPE	CIN	FAST	K Index	Mod Cape	Average	
POD	0.70	1.00	0.87	0.91		0.87	Above Average(Below for FAR)
PC	0.98	0.08	0.97	0.96		0.75	Best
FAR	0.54	0.94	0.52	0.61		0.65	
CSI	0.39	0.06	0.45	0.38		0.32	

Table 3. Virginia region

Virginia Region Averages							
	CAPE	CIN	FAST	K Index	Mod Cape	Average	
POD	0.85	1.00	0.94	0.76	0.90	0.89	Above Average(Below for FAR)
PC	0.92	0.34	0.93	0.92	0.92	0.81	Best
FAR	0.39	0.66	0.23	0.48	0.36	0.42	
CSI	0.54	0.34	0.74	0.46	0.60	0.53	

Table 4. Florida region

Florida Region Averages							
	CAPE	CIN	FAST	K Index	Mod Cape	Average	
POD	0.81	1.00	0.94	0.86	0.89	0.90	Above Average(Below for FAR)
PC	0.91	0.58	0.89	0.92	0.92	0.84	Best
FAR	0.30	0.44	0.21	0.24	0.22	0.28	
CSI	0.60	0.56	0.75	0.67	0.71	0.66	

B. GLM EAST COVERAGE RESULTS

Within the GLM East coverage area, most of the algorithms performed well, with the exception of CIN, which continued to be over predictive. The FAST algorithm performed best overall by having both the lowest FAR and highest CSI. CAPE had a surprisingly high PC rate, the best noted for all of the algorithms, but had the lowest POD. K index was very consistent, having above average results in all categories.

Examination of how each algorithm performed over forecast lead time was also carried out (e.g., see Figures 8–11). In the global-scale region of the GLM coverage area, the algorithms yielded similar values to one another. It was noted that the CAPE algorithm POD declined during the first 24 hours of forecast time, but slightly increased and then maintained a roughly consistent level afterwards. The FAST algorithm produced an interesting result, with an alternating increase and decrease (i.e., sawtooth) pattern in scores emerging that repeats itself over a 12 to 15-hour period. Initially it was thought this could be related to some diurnal effect. However, since the forecasts statistics were aggregated from every 6-hourly model run from each day, diurnal effects would not be expected to be consistent across all output.

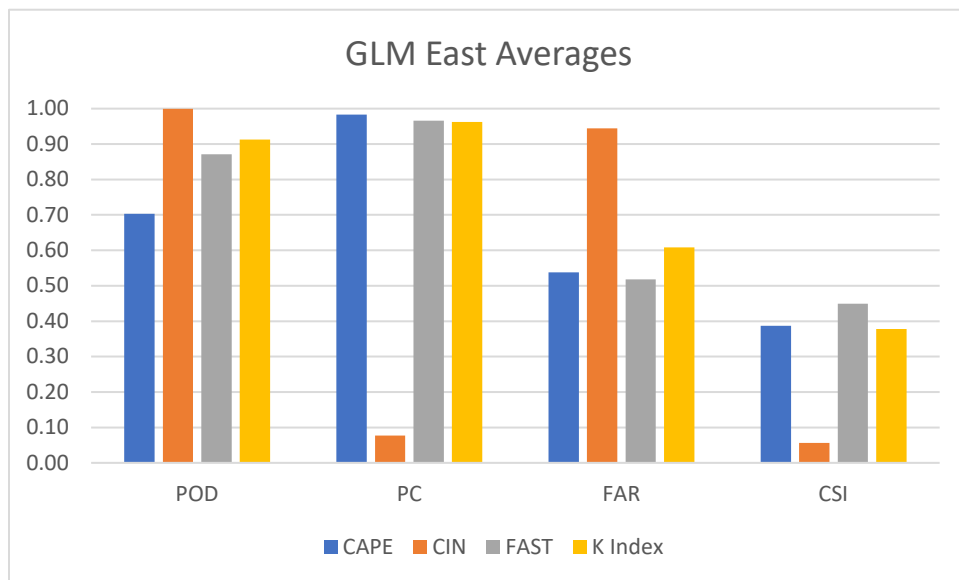


Figure 7. GLM East statistical averages

Examining POD in particular, the GLM East results shows the aforementioned sawtooth pattern with the FAST algorithm, accompanied by a slight degradation in performance over the 72 hours. K index algorithm output displayed a similar degradation of POD score with time. The CAPE algorithm, as previously discussed, notably showed declining performance in the first 24 hours before becoming approximately stable around the 0.7 range for the remaining forecast times.

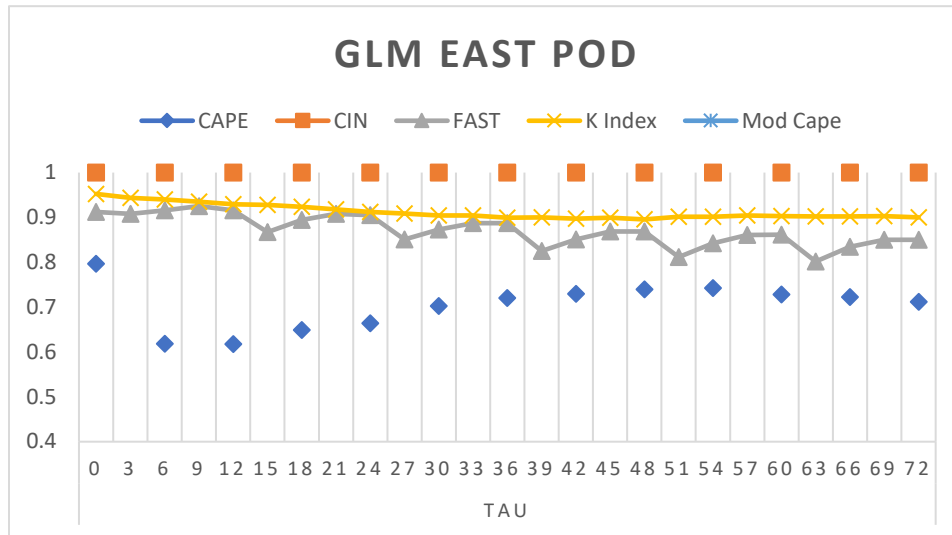
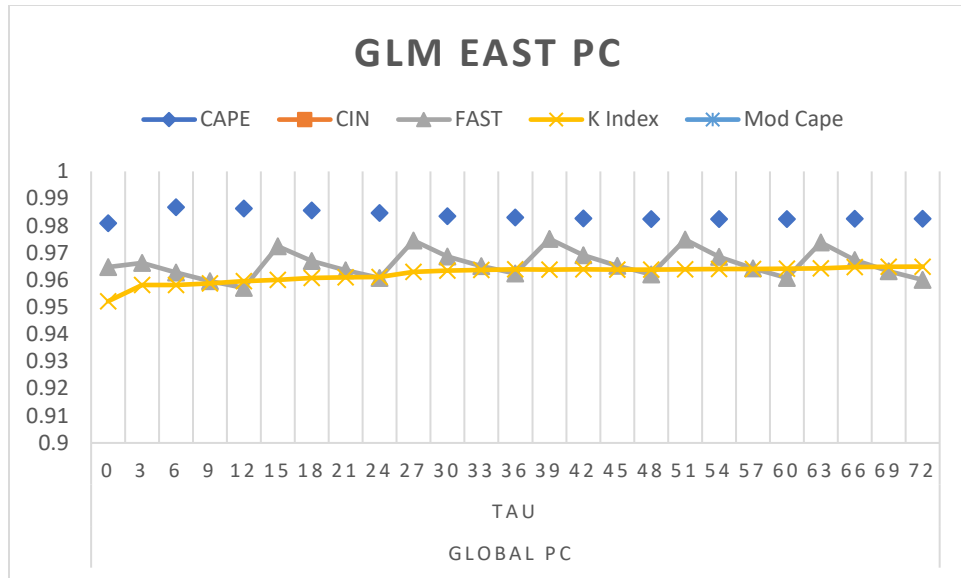


Figure 8. GLM East probability of detection by forecast lead time (TAU)

The GLM East PC scores revealed that the CAPE and K index algorithms showed little variation in performance of this metric with forecast time. The FAST algorithm was also relatively stable, in terms of PC, over forecast time, with the exception of the aforementioned saw tooth pattern. PC results for the CIN algorithm (not pictured) were very low, generally around 0.1, consistent with the assessment that the algorithm was over predictive.



CIN was not pictured on this graph as it was too low to display properly.

Figure 9. GLM East proportional correct rates over time

FAR for the various algorithms in the GLM East area over time were similar to the PC results in that each algorithm's score varied little over 72 hours of forecasts. The saw tooth pattern was still observed in the FAST algorithm. CIN FAR was relatively high, as expected with its over predictive results.

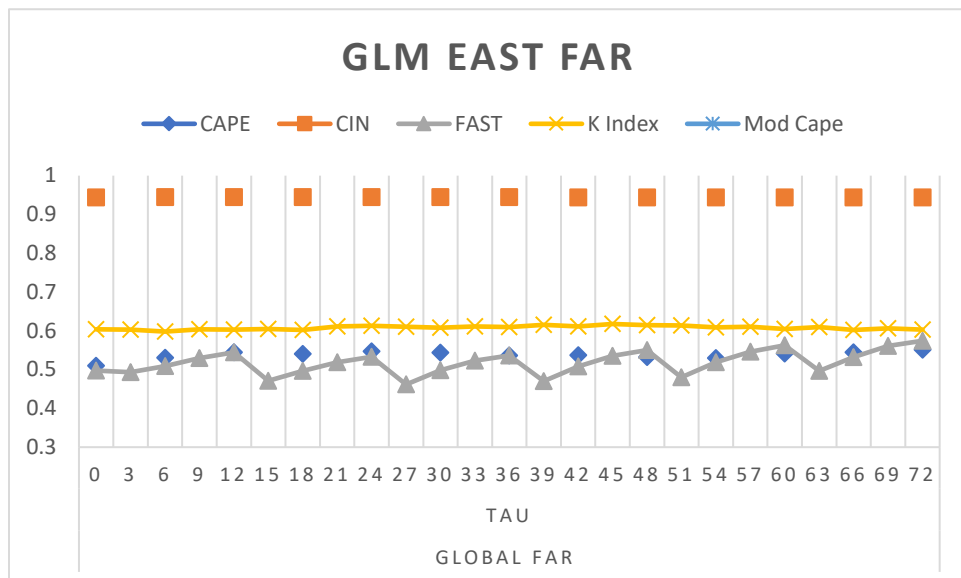


Figure 10. GLM East false alarm rates over time

The FAST algorithm had the highest CSI of all the algorithms over 72 hours of forecasts, although the saw tooth pattern remains. CAPE and K Index performed similarly to each other, both having lower CSI than FAST. More variation was noted in the CAPE algorithm performance. CIN, again, was consistently the poorest performing over the forecast period.

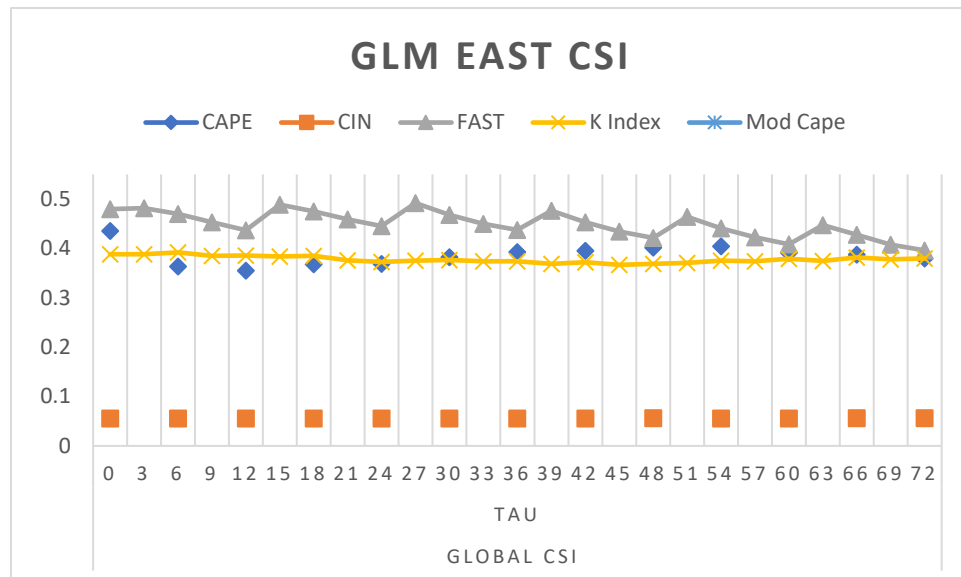


Figure 11. GLM East critical success index over time

C. VIRGINIA REGION RESULTS

Algorithmic performance in the Virginia region is displayed in Figures 12–16. Similar to the GLM East results, the FAST algorithm showed best overall performance with the highest POD (aside from CIN), highest PC rate, highest CSI and lowest FAR. The modified CAPE algorithm also performed well and was above average in all categories. CAPE and K index performed overall worse than FAST and modified CAPE, but better generally than CIN which was over predictive. More variability in results was seen in this region than in the GLM East area, this is tentatively attributed to a smaller number of thunderstorms for sampling.

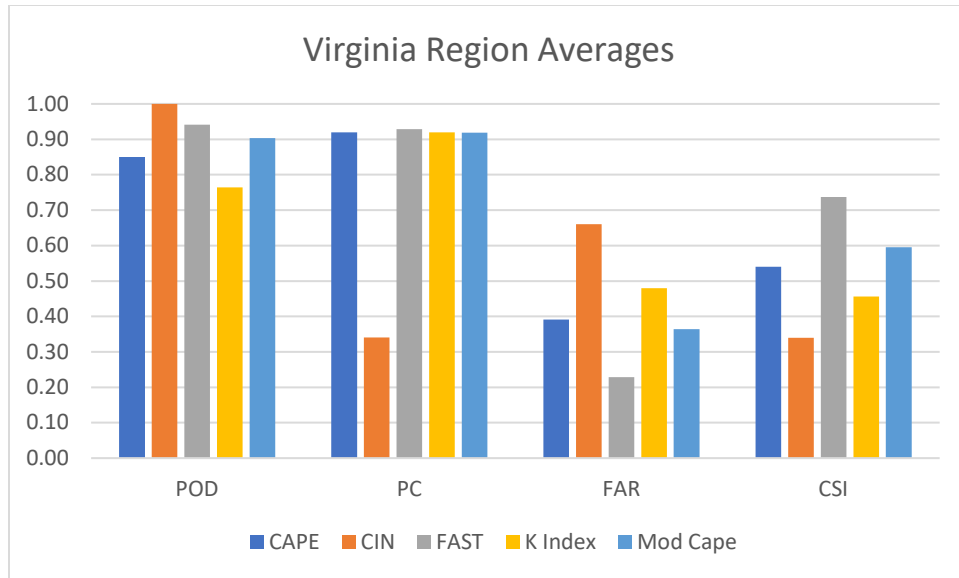


Figure 12. Virginia region statistical averages

The POD results over 72 hours of forecast time in Virginia region yielded interesting results. The CIN algorithm continued to have relatively high POD, consistent with being over predictive. Neglecting CIN, the FAST algorithm was seen to have highest POD, with performance from the modified CAPE algorithm being relatively close. K index related scores declined quickly with forecast time and became erratic after approximately 30 hours. CAPE in the Virginia region had a similar performance as in the GLM East region with a decline in POD noted in early forecast lead times and then better performance after 24 hours.

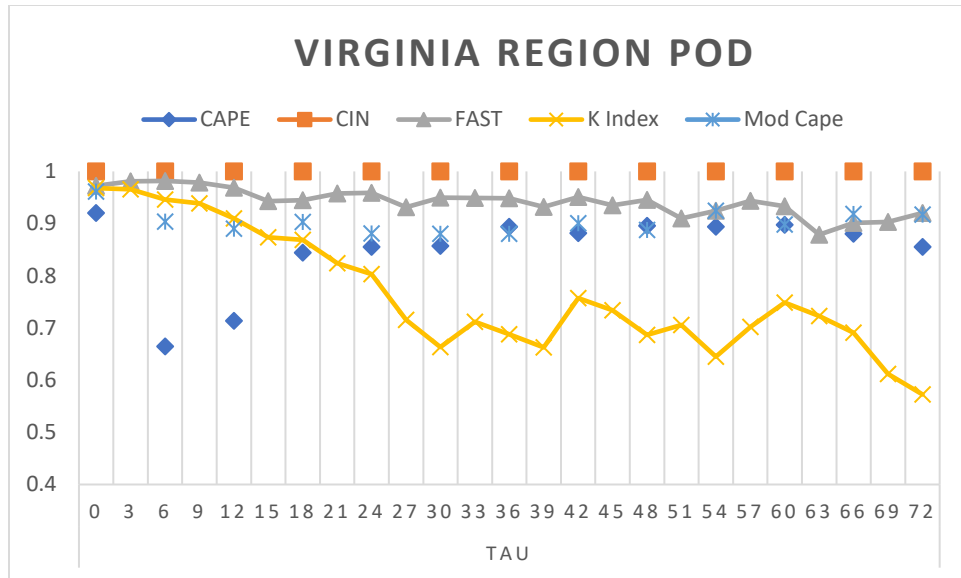
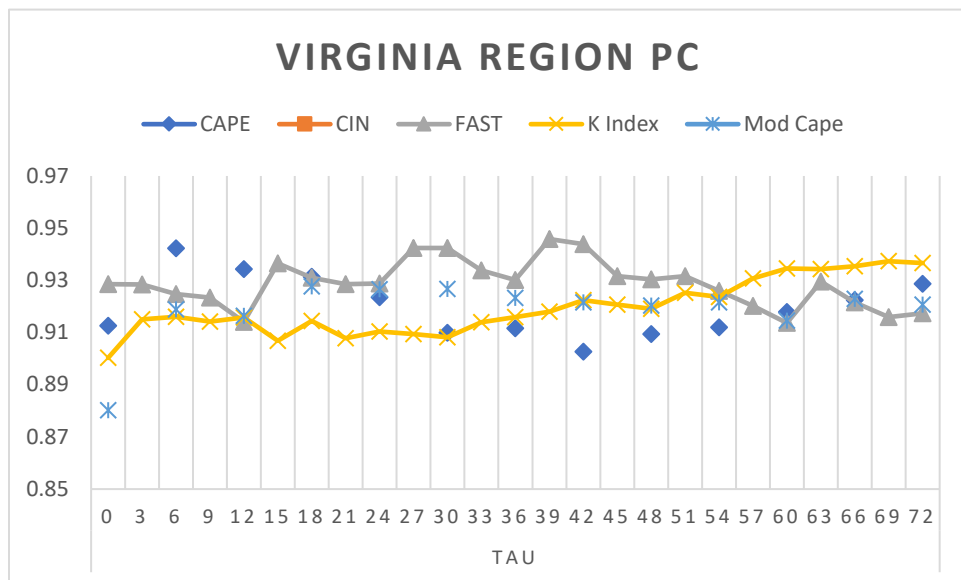


Figure 13. Virginia region probability of detection over time

PC rates over time in the Virginia region were slightly erratic. However, neglecting CIN which performed relatively poorly, the remaining algorithms varied from 0.88 to 0.94 over the forecast period. For this region, K index PC scores over 72 hours of forecasts was the most consistent and displayed slight increase with time.



CIN was not pictured as it was too low to display

Figure 14. Virginia region proportional correct rates over time

The FARs for the Virginia region were considerably lower than for the GLM East area. However, the relative performance was similar with FAST overall best. The CAPE and modified CAPE FAR results were very close to each other, with modified CAPE performing slightly better. The K index algorithm FAR generally increased over time but otherwise did not show consistent trends. CIN remained consistently high.

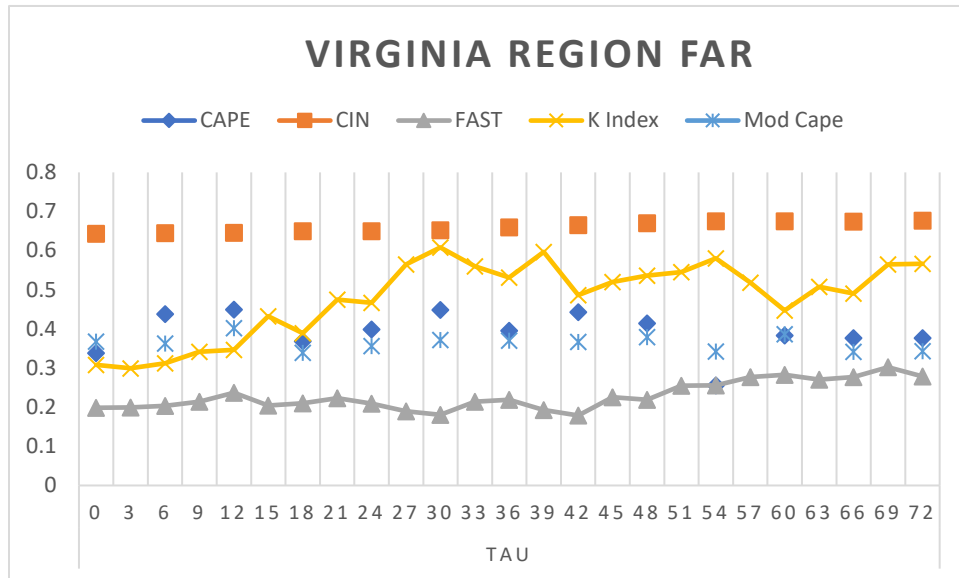


Figure 15. Virginia region false alarm rates over time

CSI results over time yields a useful depiction of the performance of the algorithms for this region. The FAST algorithm was the top performer with a higher CSI than all algorithms through all forecast hours. K index performance quickly faded after 12 hours. Modified CAPE slightly bested conventional CAPE over the forecast period while CIN continued to perform relatively poorly.

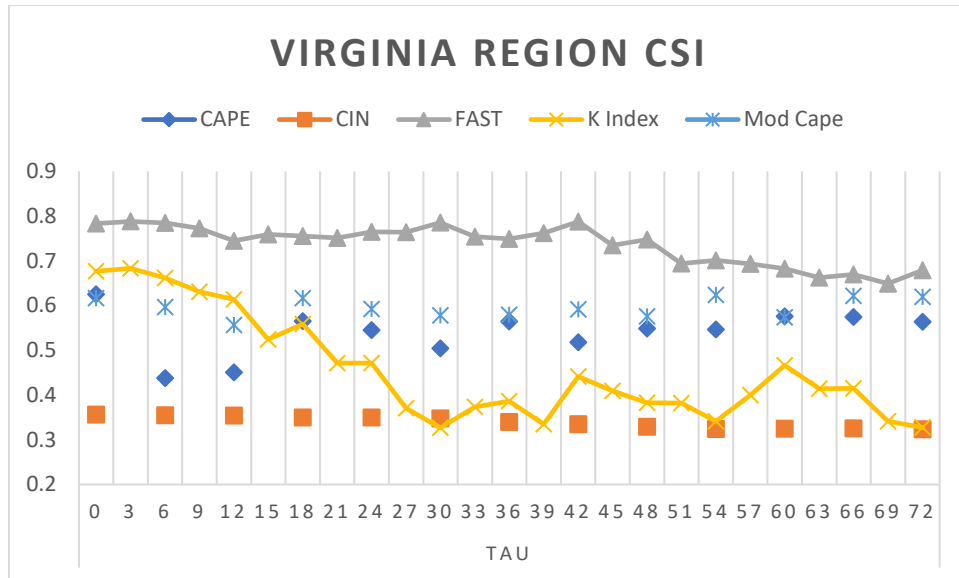


Figure 16. Virginia region critical skill index over time

D. FLORIDA REGION RESULTS

Figures 17 through 21 display summary results for the Florida region. Overall performance, as displayed in Figure 17, suggests FAST remains the most useful algorithm, showing the highest CSI and lowest FAR. However, the modified CAPE algorithm was close to the FAST in all categories and even had a higher PC rate. K index and CAPE overall performance was lower, but relatively close. The CIN algorithm displayed similar results as in the other two evaluations above. For the period of this research, more thunderstorms were observed in the Florida region than in the Virginia region, which may make the Florida results more robust.

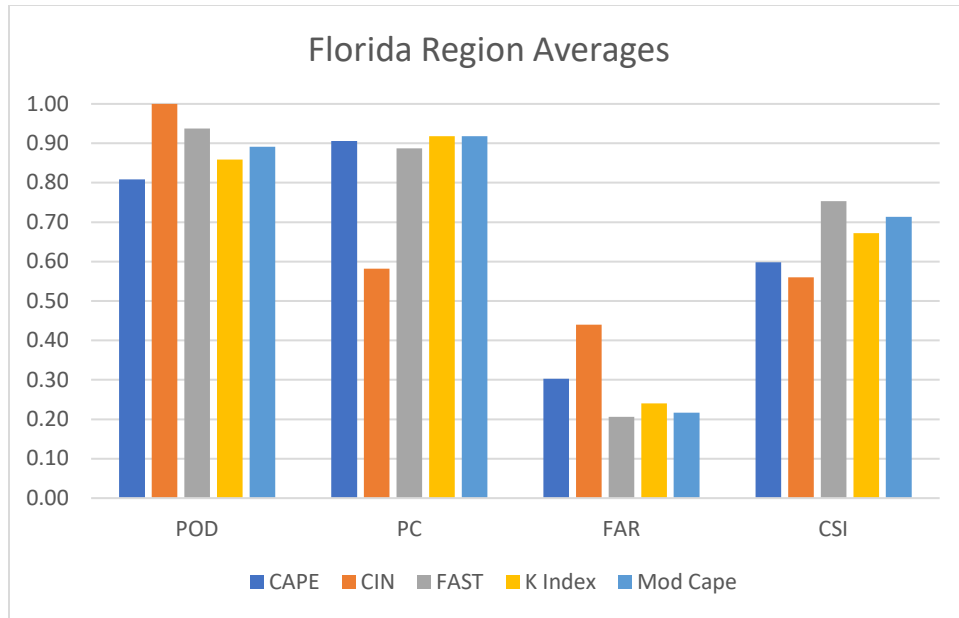


Figure 17. Florida region statistical averages

The POD results over time for the Florida region revealed slightly more consistent results than Virginia region. CIN remained high across the forecast period. FAST continued to have the highest POD of the remaining algorithms, with a saw tooth pattern becoming apparent again. K index degraded over time similarly to the Virginia region. CAPE again showed the pattern of quickly degrading performance after the initial forecast hour and then slightly improvement over time. Modified CAPE had a similar result as CAPE but did not degrade as much in the first six hours and became more consistent around 0.9 for the remainder of the forecast range.

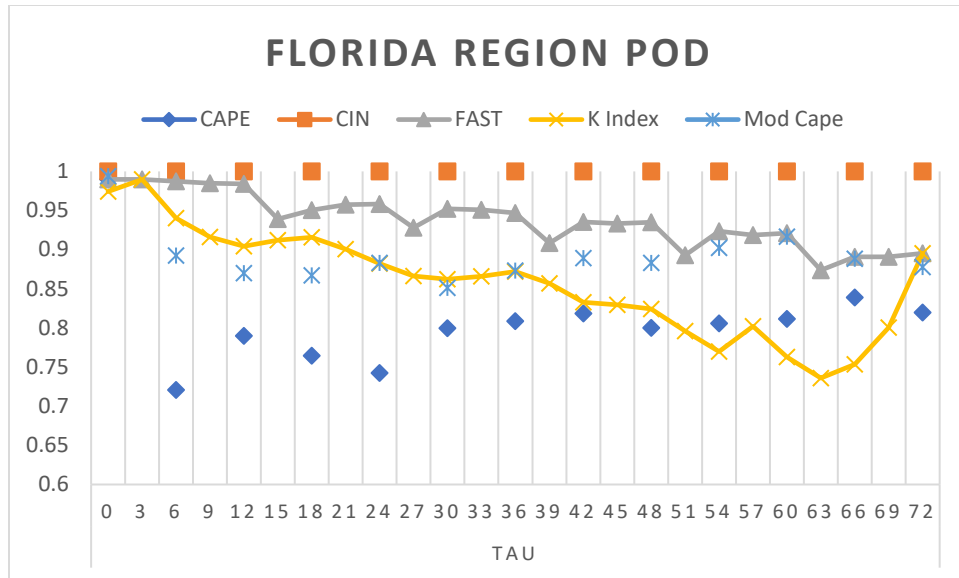


Figure 18. Florida region probability of detection over time

PC rates for algorithms in this region were relatively close, fairly consistent around 0.9 mark. Surprisingly, this is the only category in which the FAST algorithm was not one of the better performing algorithms. CAPE, K index and modified CAPE generally performed better across the forecast periods. CIN remained at the bottom of the group but was higher than in the Virginia region or in the overall GLM East areas.

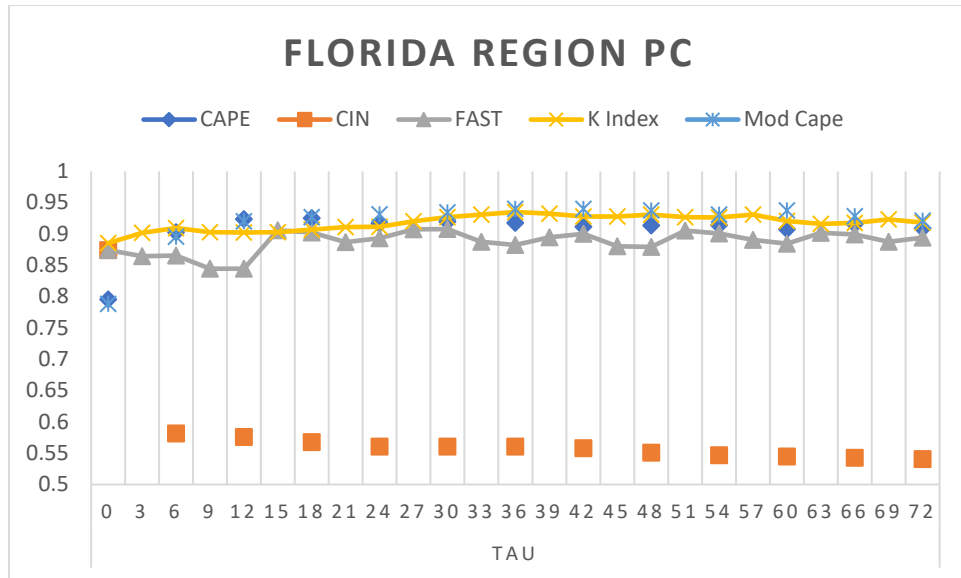


Figure 19. Florida region proportional correct rates over time

False alarm rates in region were the lowest observed across the three areas examined. K index, modified CAPE, and the FAST all performed well with FAST generally lower over time. CAPE FAR was slightly higher than FAST, modified CAPE, or K Index. CIN still displayed a high FAR, but it was noted to be lower than the other regions examined.

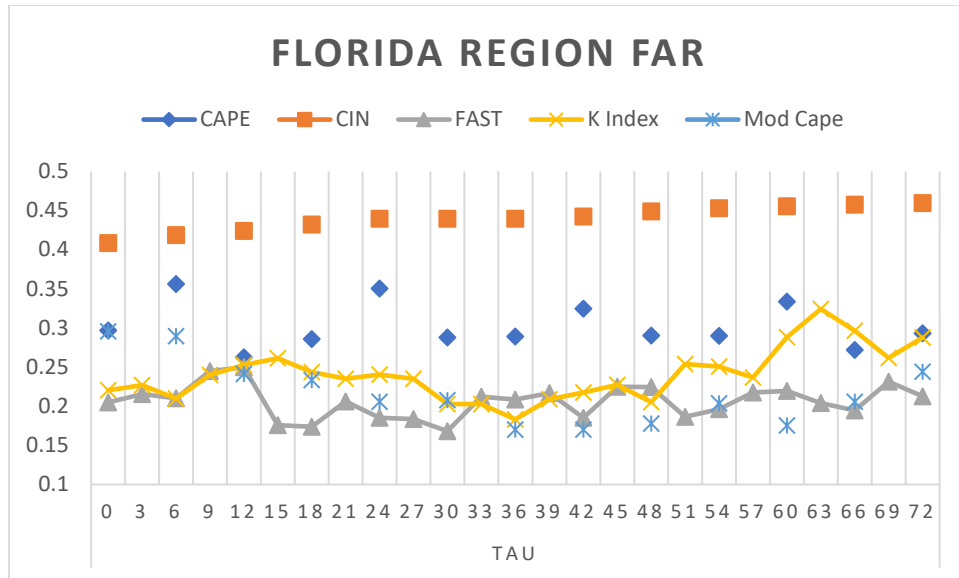


Figure 20. Florida region false alarm rates over time

CSI results for this region over time continued to show the FAST algorithm outperforming the rest. Interestingly, the modified CAPE algorithm, while performing well below FAST in the first 36 hours, improved over time and was in line with FAST through the remainder of the forecast period. K index generally declined over time while CAPE was near the bottom, almost in line with CIN, and was slightly erratic. CIN remained the worst overall performer but still was better than in Virginia and GLM East regions.

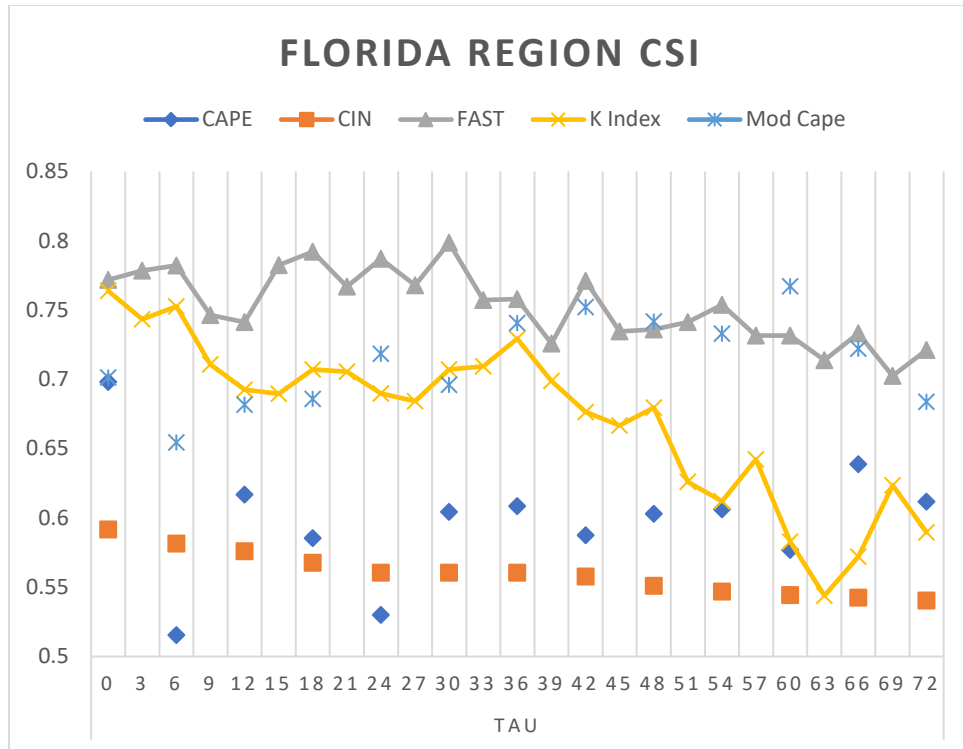


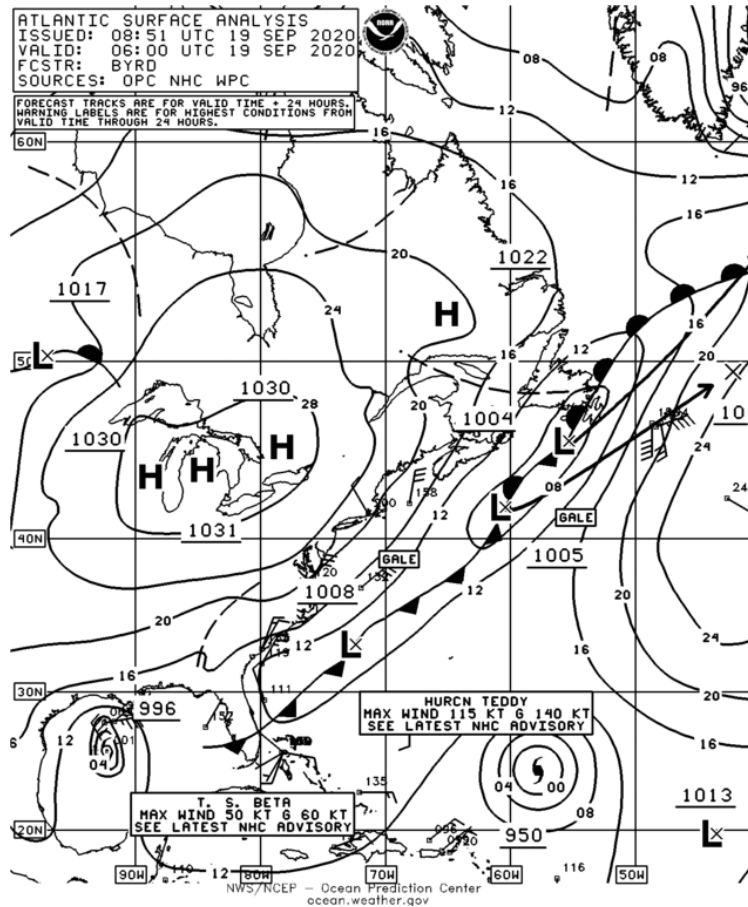
Figure 21. Florida region critical success index over time

E. QUALITATIVE ANALYSIS

In addition to the above analysis, a qualitative inspection of a select few forecasts was performed. Here we concentrate on the 72-hour forecast from 16 September 2020 at 06Z with a verification time of 19 September 2020 at 06Z. This day was chosen due to the performance of the algorithms on that day and the active weather patterns present. Several other selected forecasts are also available in Appendix B.

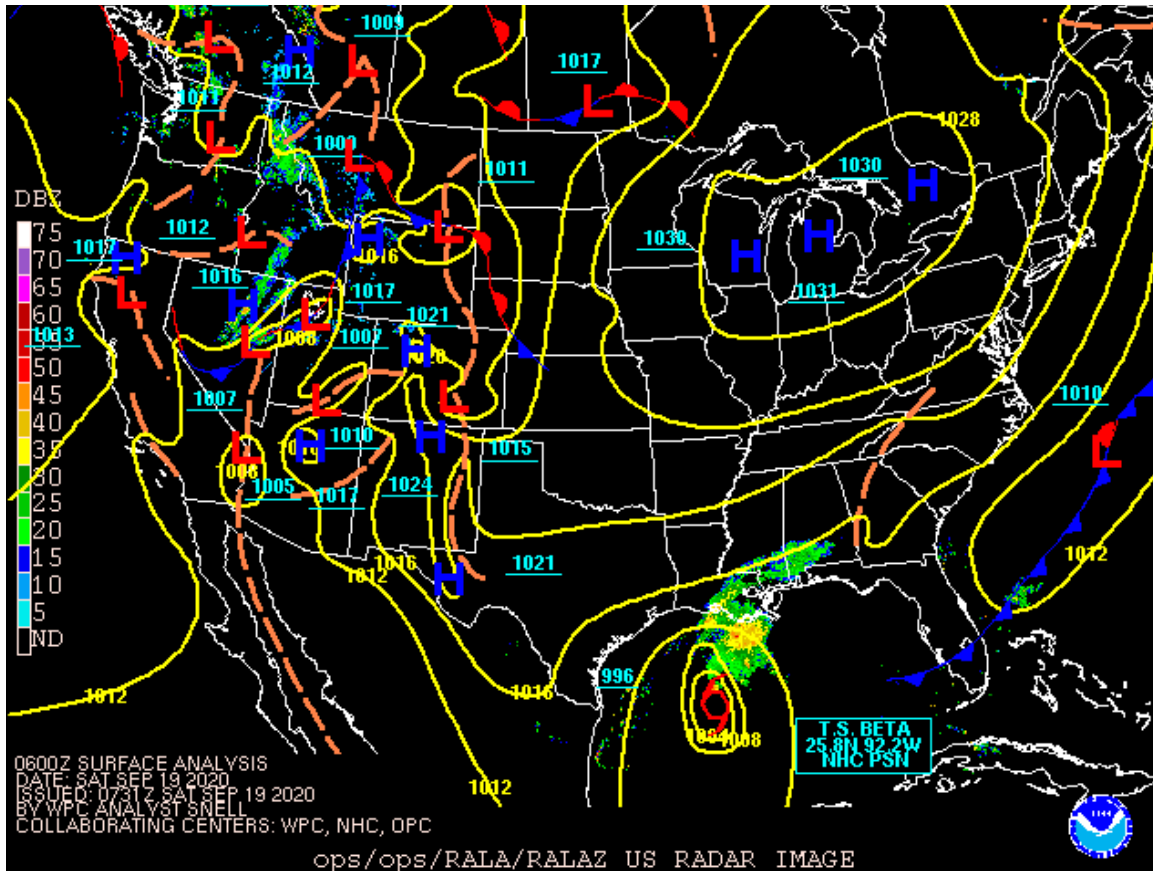
Three surface analysis charts (Figures 22–24) are provided along with the forecasts. One surface analysis additionally includes a weather radar mosaic (Figure 23) from the valid time, and another includes a satellite overlay (Figure 24) to highlight the existent weather. From these charts a large cold front is observed offshore of the Eastern Seaboard of the United States. Also, two tropical systems are observed. Hurricane Teddy is northeast of Puerto Rico and Tropical Storm Beta is south of Louisiana. Additionally, a complicated weather system is draped across the northern mountain west region of the U.S. With the satellite and radar overlays, several areas of thunderstorms can be observed along the

Mexican coast. Thunderstorms can be inferred by the brighter colored areas in the overlays, which represent colder brightness temperatures in the satellite overlay and higher dBz levels in the radar overlay.



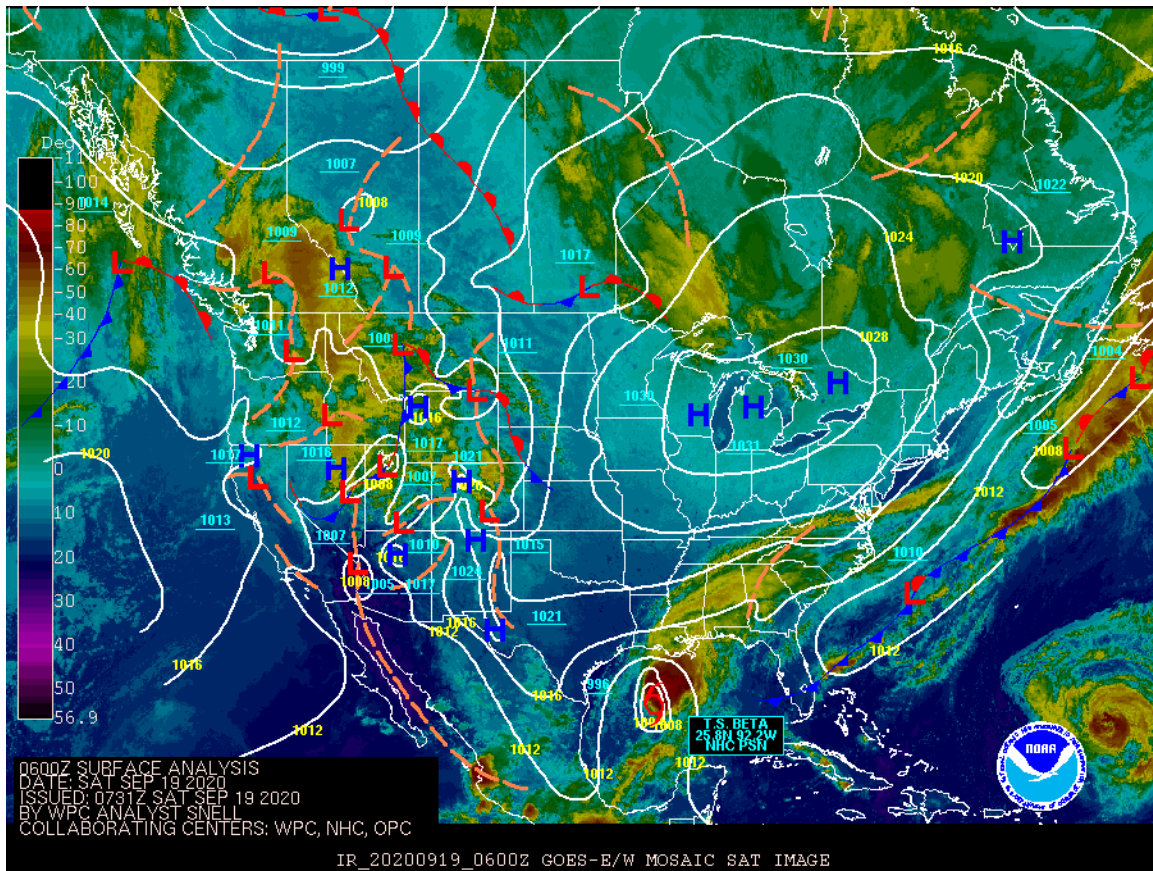
Map retrieved from the Ocean Prediction Center, November 13, 2020, https://www.ncei.noaa.gov/data/ncep-charts/access/2020/09/19/atlw_sfcanal_06.2020091908.gif_06.2020091908.gif

Figure 22. September 19, 2020 06Z Surface analysis for the eastern seaboard and western Atlantic



Map retrieved from the Weather Prediction Center, November 13, 2020,
https://www.wpc.ncep.noaa.gov/archives/web_pages/sfc/sfc_archive_maps.php?arcdte=09/19/2020&selmap=2020091906&maptype=radsfcus_exp

Figure 23. September 19, 2020 06Z U.S. surface analysis with radar mosaic overlay



Map retrieved from Weather Prediction Center, November 13, 2020,
https://www.wpc.ncep.noaa.gov/archives/web_pages/sfc/sfc_archive_maps.php?arcdte=09/19/2020&selmap=2020091906&maptype=satsfcnp

Figure 24. September 19, 2020 06Z U.S. surface analysis with satellite overlay

Comparing the surface analysis to the GLM data (Figure 25) reveals some intriguing results. While most of the convection along the Mexican coast seems to be easily identifiable, the two tropical system show a particularly small signature, especially when compared to the radar and satellite data.

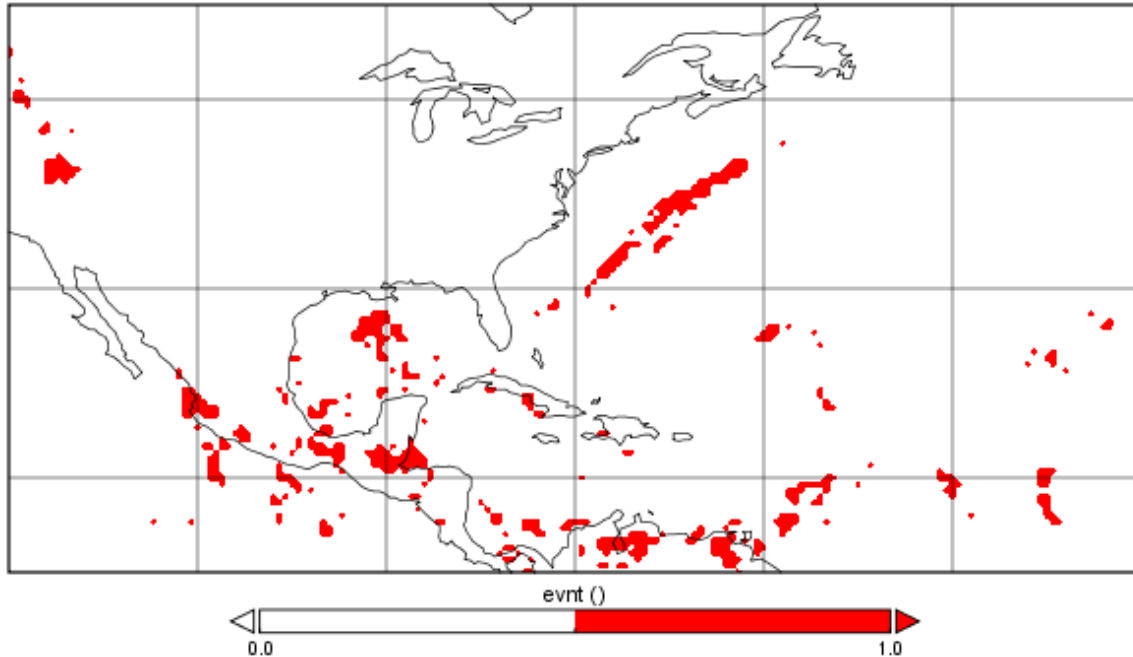


Figure 25. September 20, 2020 06Z GLM thunderstorm data

The forecasts from that time frame are contained in Figures 26 through 30. In these figures, areas colored red are areas of predicted thunderstorms determined by each algorithm. The FAST forecast (Figure 26) seems to paint the most accurate picture when compared to observations and analysis. The over predictiveness that was suggested by the statistical analysis was observed across Georgia with a large area of thunderstorms predicted but none observed. Additionally, FAST predicted a smaller area of thunderstorms in the central plains with none observed. The cold front is depicted as being larger than what was observed in the GLM data but seems to fit better with the satellite overlay in that region. However, the two tropical system's signatures in the FAST forecast appears more realistically than in the GLM data.

The K index (Figure 27) forecast, like the FAST forecast, more accurately predicts the signature of the tropical systems than the observed GLM data. However, the K index forecast seems to be even more over predictive than the FAST with multiple areas of predicted storms with none observed in the GLM data or the analysis. This was most notable, similar to FAST, in the area around Georgia and in the central US.

CAPE (Figure 28) actively over forecasted large areas of thunderstorms across the tropics but had little to no thunderstorm forecasts north of 30N latitude, including completely under forecasting the cold front. CIN (Figure 29) as previously noted, was highly over predictive with a majority of the whole forecast being marked as thunderstorm, indicating that the threshold and results were incorrect. Modified CAPE (Figure 30) was contained to the two smaller regions so there were not as many observations for this algorithm's performance. While the algorithm predicted an area of thunderstorms off the coast of Northern Florida, none were observed in the GLM data. However, the radar overlay in that area suggest there may be some small thunderstorms active in that vicinity.

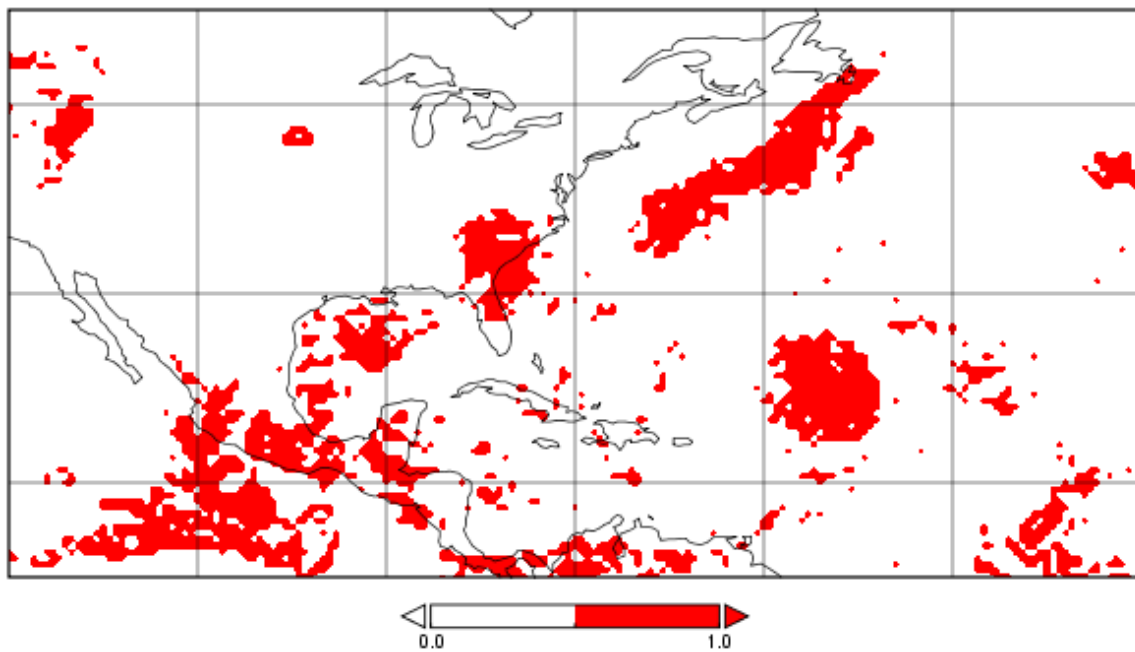


Figure 26. September 19, 2020 06Z FAST thunderstorm forecast

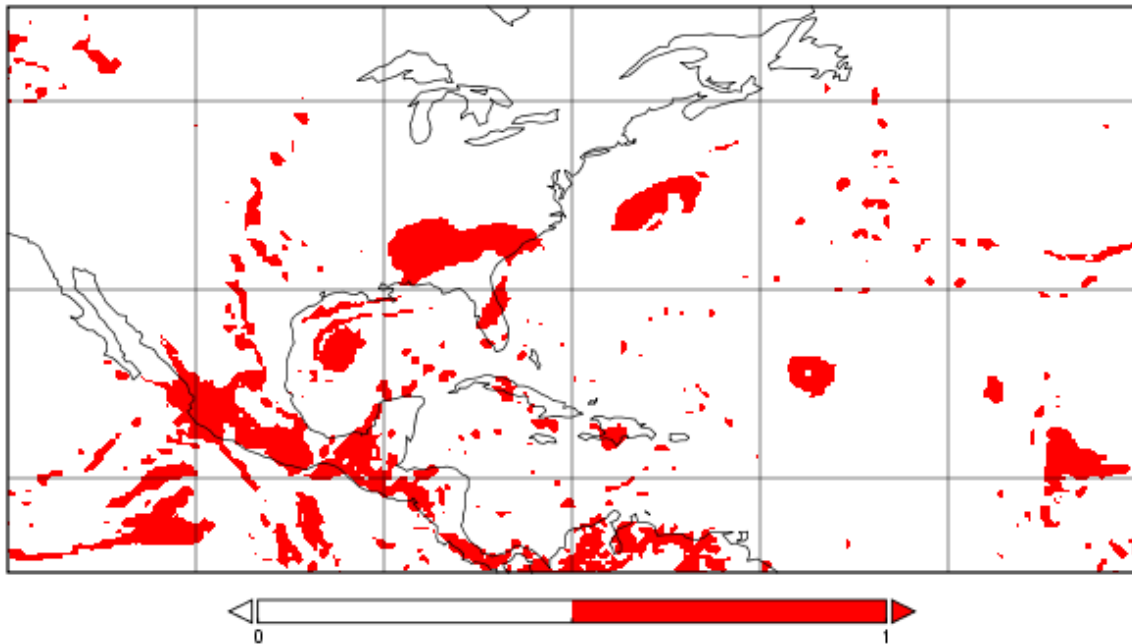


Figure 27. September 19, 2020 06Z K index thunderstorm forecast

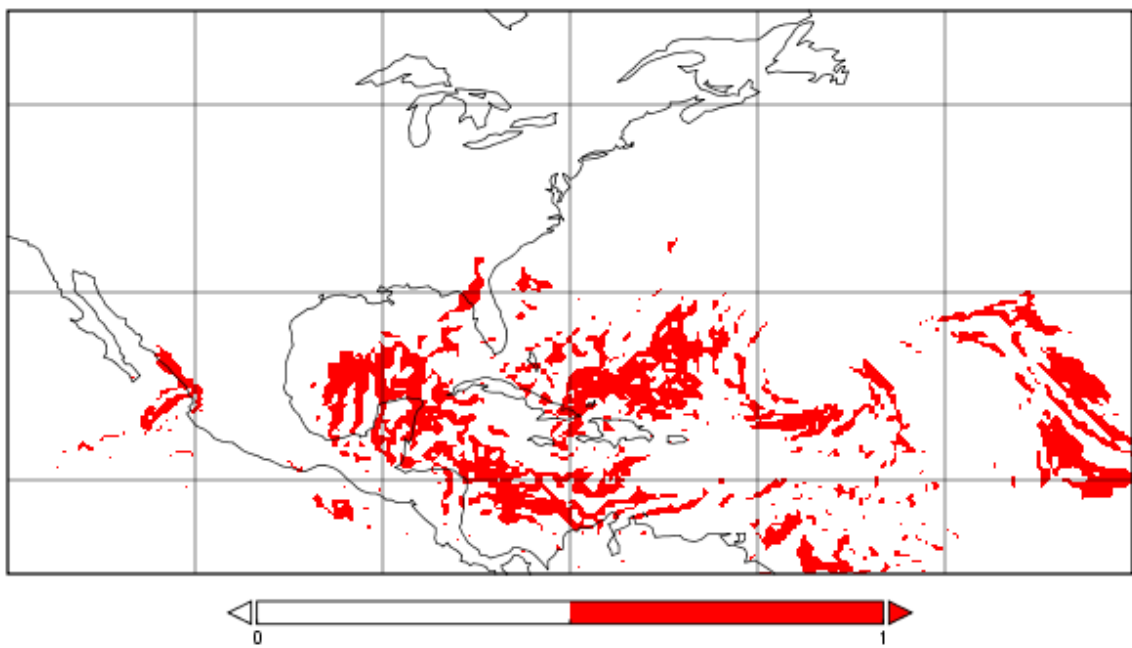


Figure 28. September 19, 2020 06Z CAPE thunderstorm forecast

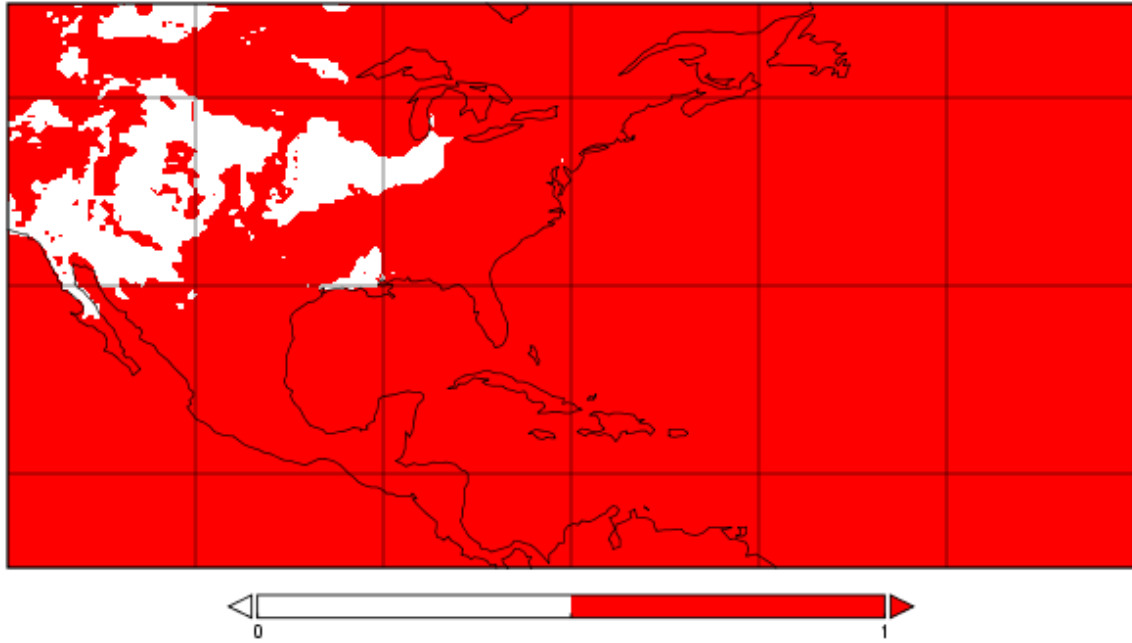


Figure 29. September 19, 2020 06Z CIN thunderstorm forecast

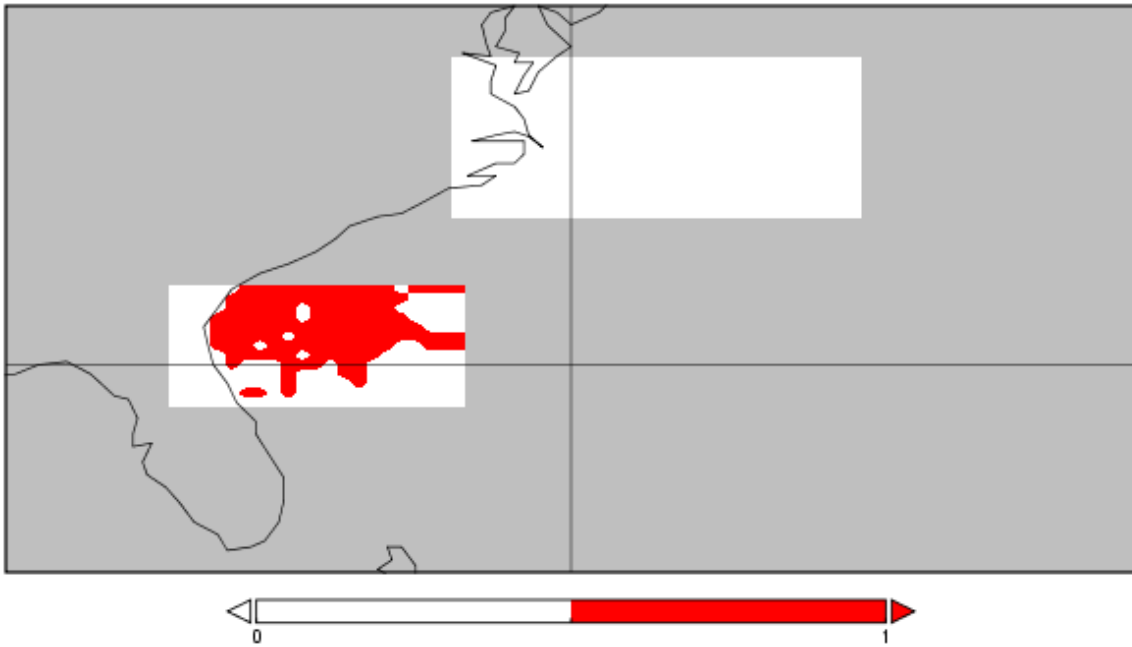


Figure 30. September 19, 2020 06Z Modified CAPE thunderstorm forecast

One explanation for why the GLM data shows less inferred thunderstorms than the forecast data is that lightning in oceanic thunderstorms has been noted to be less than continental thunderstorms (Williams and Stanfill 2004). Overall, as seen in Figure 31 and Figure 32 which are from September 9th, thunderstorms were forecast along intertropical convergence zone, but few were observed, and this pattern was also observed in several other forecasts that were analyzed. This is most likely due this oceanic effect, however, these storms are also towards the edge of coverage for the GLM sensor. It is not known at this time if the lack of GLM observations in these regions is due to the oceanic thunderstorm effect, the GLM coverage area being set incorrectly, or if these storms were just too far of nadir for the sensor to be observed. This is a possible reason for the higher false alarm rates seen in the GLM East region comparatively to the two smaller regions which lie well within the GLM coverage area and closer to nadir of the sensor.

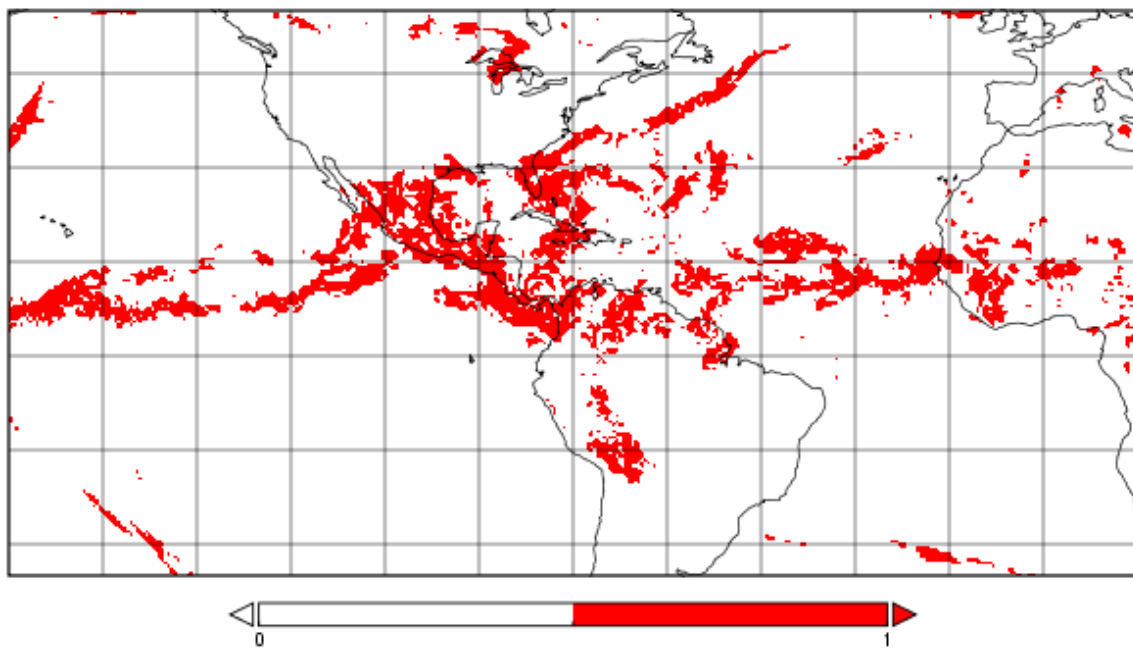


Figure 31. September 09, 2020 00Z FAST thunderstorm forecast

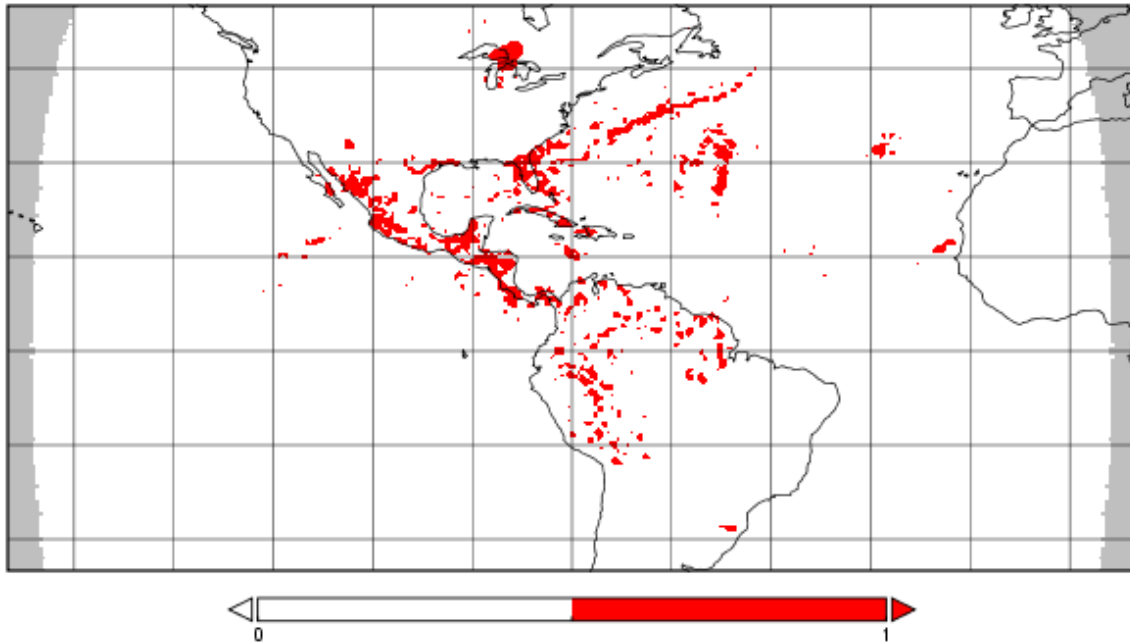


Figure 32. September 09, 2020 00Z GLM thunderstorm data

F. ADDITIONAL NAVGEM VERIFICATION

In order to possibly explain the saw tooth patterns noticed with FAST performance, an examination of general NAVGEM performance over September 2020 was performed utilizing the FNMOC model verification page. Here again, a slight sawtooth pattern was observed for air temperature (Figure 33), relative humidity levels (Figure 34), and mean sea level pressure (Figure 35). These patterns suggest the pattern of alternating increasing and decreasing levels of performance are present across many different fields of NAVGEM output and are not solely a feature of FAST algorithm performance. It was also noted that this sawtooth effect was seen in some variables in several of the other global models that were included in the verification.



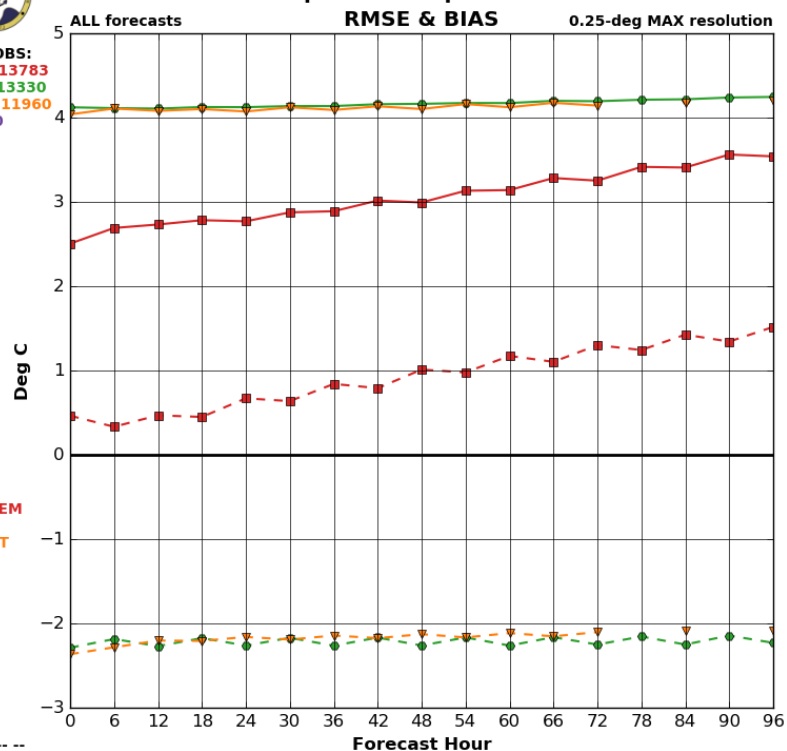
AVG OBS:
NAV: 13783
GFS: 13330
UKM: 11960
JMA: 0

NAVGEN
GFS
UKMET
JMA

BIAS -- --
RMSE ____

GLOBAL Air Temperature Sep 2020

UNCLASSIFIED



Approved for Public Release; Distribution is Unlimited

Retrieved from FNMOC, 15 November, 2020,
https://www.fnmoc.navy.mil/verify.cgi/cgi-bin/verif_all.cgi?area=global&dtg=sep2020&prod=sairtsa&set=standard&type=regional&hr=a#sairtsa

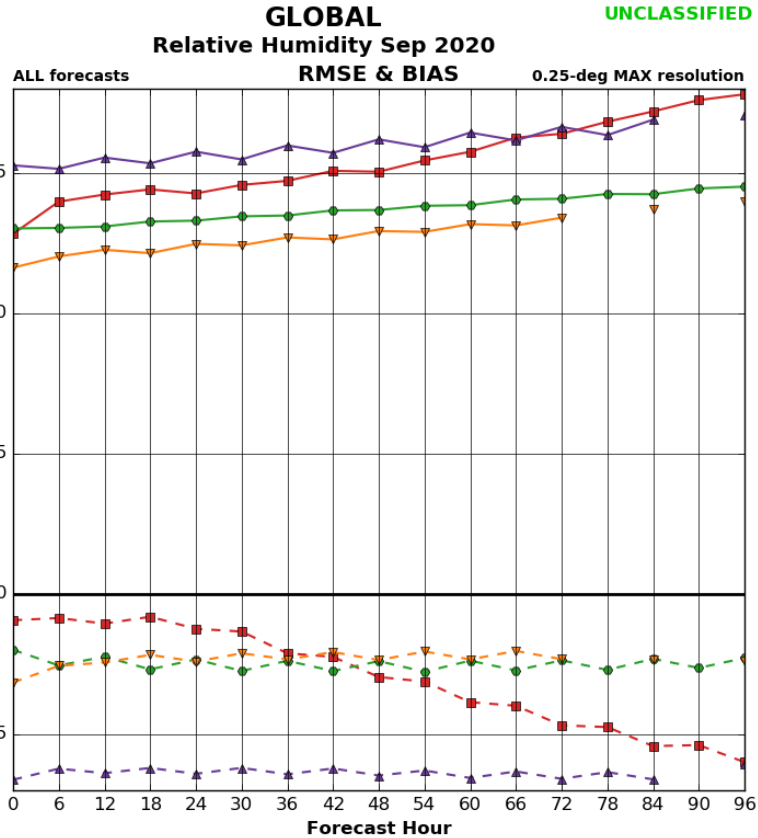
Figure 33. NAVGEN air temperature verification for September 2020



AVG OBS:
NAV: 13212
GFS: 12992
UKM: 11669
JMA: 12063

NAVGEN
GFS
UKMET
JMA

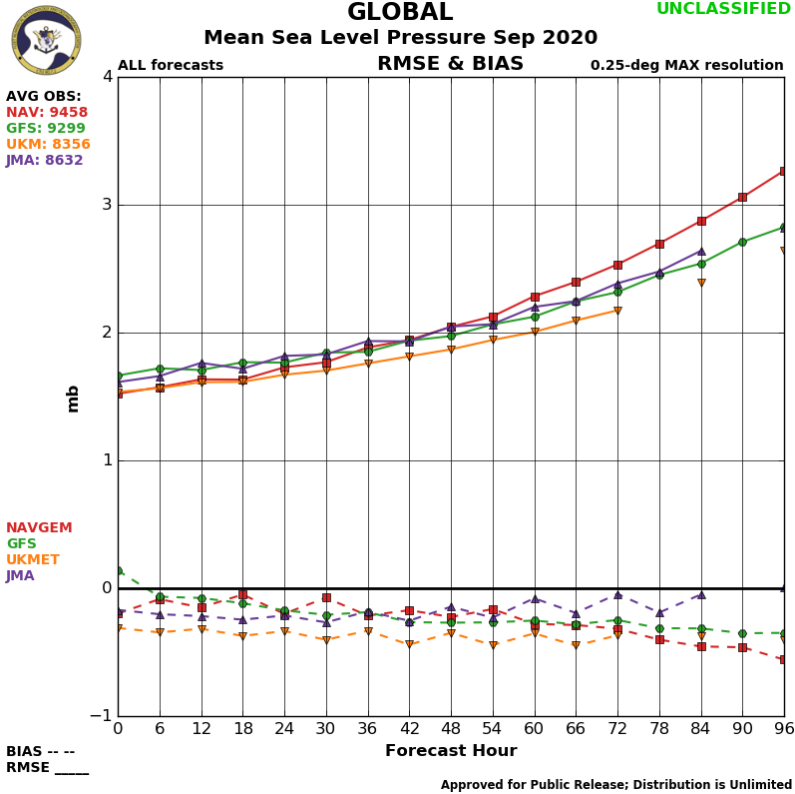
BIAS ---
RMSE —



Approved for Public Release; Distribution is Unlimited

Retrieved from FNMOC, 15 November, 2020,
https://www.fnmoc.navy.mil/verify.cgi/cgi-bin/verif_all.cgi?area=global&dtg=sep2020&prod=sairtsa&set=standard&type=regional&hr=a#sairtsa

Figure 34. NAVGEN relative humidity verification for September 2020



Retrieved from FNMOC, 15 November, 2020,
https://www.fnmoc.navy.mil/verify.cgi/cgi-bin/verif_all.cgi?area=global&dtg=sep2020&prod=sairtsa&set=standard&type=regional&hr=a#sairtsa

Figure 35. NAVGEM mean sea level pressure verification for September 2020

THIS PAGE INTENTIONALLY LEFT BLANK

V. CONCLUSION

A. GENERAL CONCLUSIONS

In conclusion, the NRL FAST algorithm was overall the top performing algorithm for predicting thunderstorm activity in the regions examined. In the GLM East region, it had a 0.87 POD, 0.97 PC rate, 0.52 FAR, and a 0.45 CSI. Additionally, the FAST CSI and FAR were the best of all the algorithms. The FAST algorithm also performed even better than the GLM East averages in the smaller regions. All the algorithms displayed operationally significant FAR. In naval forecasting and planning, a FAR tolerance will vary based on the specifics of individual mission type. However, a FAR above 0.50 may seriously degrade customer trust in a forecast. Regardless, the FAST algorithm was the least over predictive. Also, considering the observations made during the qualitative analysis section of this paper, the FAR could be skewed high due to the lack of lightning observed in an especially busy hurricane season. Considering FAST is to be used as an operational product that will support long endurance UAVs, some over predictiveness is acceptable as safety of flight is key. It is typically better to unnecessarily divert or delay a flight, than to lose an airframe to a weather-related accident. Additionally, while the algorithm will be used as the primary forecasting technique, a skilled forecaster over-the-loop, overlooking the results, should add some skill.

The other algorithms did show skill, sometimes even outperforming FAST in various measures over time, but never consistently across all categories. Another point of interest for these results is the resolution that the algorithms were run at. FAST was run at half-degree resolution to be consistent with the initial operating capability of the software, whereas the other algorithms were run at the recently increased output resolution of NAVGEM. Despite this, FAST was still able to outperform the other algorithms. One may anticipate this better performance will continue when FAST is run with higher resolution fields. This will need to be verified.

While FAST did outperform the other algorithms, it was not perfect in its temporal or spatial predictions. This suggests that these algorithms are better at generally assessing

areas of thunderstorms and are not an exact point forecast tool that would be optimum for every specific operation. However, forecasting an accurate broad area of thunderstorms at a fast speed is better than forecasting from a blank page. Operationally, Navy mission planners looking 72 hours into the future may be able to tolerate less time fidelity as a tradeoff for situational awareness of thunderstorms overall. The famous quote from George Box comes to mind, “All models are wrong, but some are useful.”

Physically, it makes the most sense that the FAST algorithm would perform the best. The necessary ingredients for moist convection are moisture, instability, and a “trigger” that lifts parcels to their LFC (e.g., Doswell 2006). Convective parameterizations like the SAS implicitly account for all of these ingredients. For instance, these schemes are only triggered if moisture and CAPE are present, and if a convective trigger condition is met. Thus, by using convective precipitation from the SAS, the FAST algorithm accounts for all of the necessary ingredients for convection. In contrast, CAPE and modified CAPE account for moisture and conditional instability, but there is no necessary trigger mechanism in the calculation, limiting its ability to be a completely automated thunderstorm predictor. While modified CAPE would better account for entrainment effects than the traditional CAPE calculations, it would still not account for a trigger mechanism. Note that cumulus parametrizations also account for the role of entrainment in modulating updraft buoyancy, and thus the FAST algorithm implicitly accounts for entrainment. K index, while simple and efficient, does not take into all the necessary ingredients, just some accounting for moisture and instability in the lower and mid-levels that a skilled forecaster may find useful, but most likely limited if automated. Furthermore, CIN does not perform as well as it does not account for any of the necessary ingredients, as it predicts areas where convection is highly unlikely. However, it could be combined with CAPE and some form of triggering mechanism and be useful in the future.

An intriguing part of the results is the slight saw tooth pattern that was observed several times. Initially thought to be some type of diurnal effect, but because the algorithms were initialized at different date time groups (spanning 00, 06, 12, and 18 UTC model runs), this would not explain the results. One possibility is that the region forecasted (Eastern Americas and Atlantic Ocean) sees an approximate diurnal convective maximum

period that typically occurs over land in the early evening and over water in the early morning. This would fall between the synoptic forecast times. However, that would not fully explain the dips in performance. The NAVGEM verification data in September 2020 also showed a slight saw tooth pattern for some variables. Although the variation is unexplained at this time, the variables for the FAST algorithm may be more strongly linked to the NAVGEM metric variability than the other algorithms. This effect may be worth further investigation.

B. CONTINUATION IDEAS

Although reinforcing the validity of the current FAST thunderstorm forecast methodology, the results of this research indicate areas for further needed research. While all the algorithms were based on the NAVGEM global model because of the operational need for large area forecasts, it would be worth seeing how performance could be improved by retooling the algorithms to mesoscale models. The finer resolution of these models may be better at capturing the smaller scale thunderstorm activity. Also, NAVGEM was the only global model used for this research, again based on operational need, but it would be worth examining how other global models perform with these algorithms, or if some sort of consensus approach would perform better. Further studies could also focus on seasonal variability among other regions of the globe.

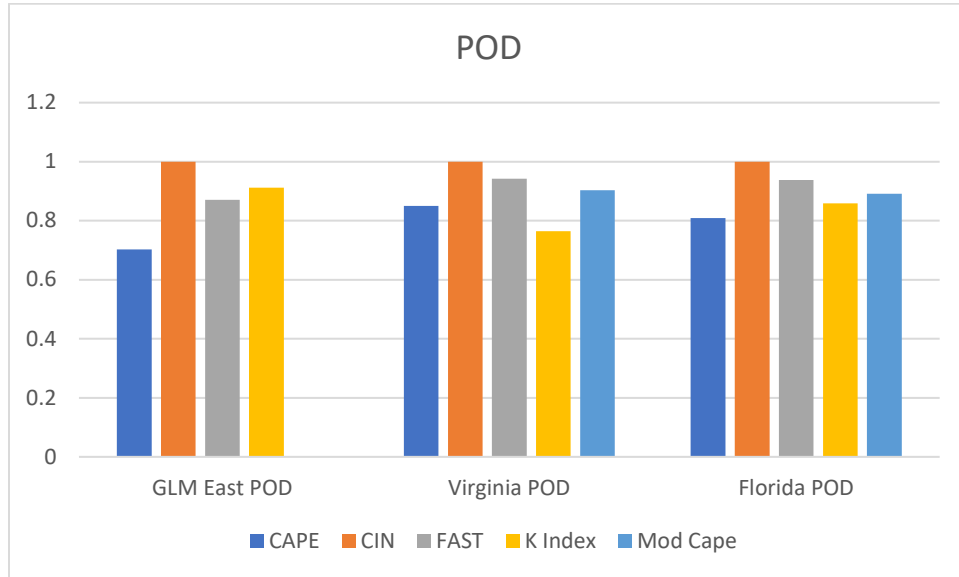
Additionally, as GLM data was used as a verification metric for this paper, another verification metric may better quantify the effects of thunderstorms. Considering the observed low amount of lightning typically seen in tropical systems and that thunderstorm impacts to operations are not just limited to the area of lightning, using another metric such as radar or other satellite observations may be useful.

While beyond the scope of this paper, using some sort of machine learning approach to developing better algorithms to predict thunderstorm activity may be the way of the future. While some patterns occur to the human eye readily, and can be observed in this paper, there may be things occurring that a machine learning approach would be better qualified to observe and predict.

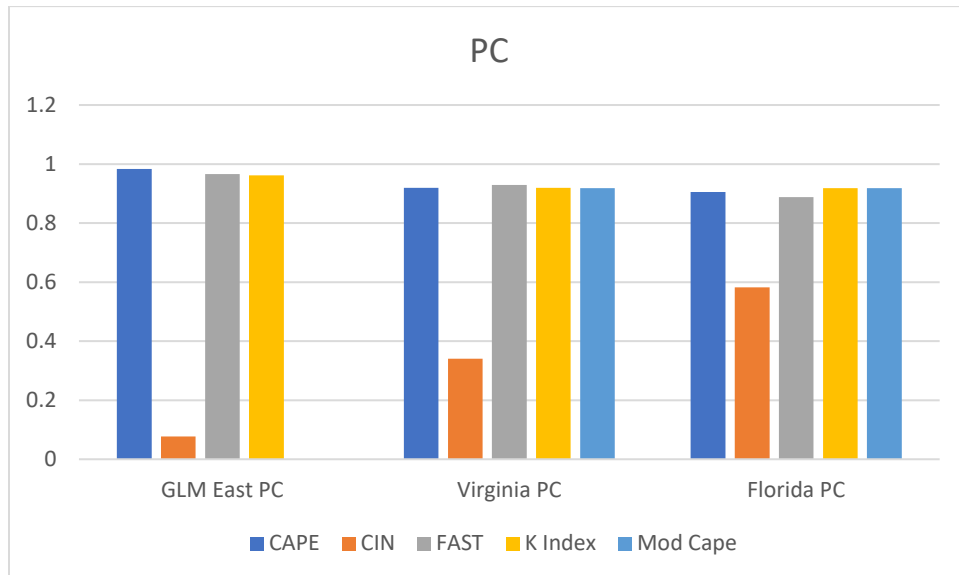
As automated approaches to weather forecasting become more common other validations should occur. While this paper was concentrated on automated thunderstorm forecasting, the FAST software also has other automated forecasts, most notably icing and turbulence. These forecasts should undergo a validation process as well considering the important operations they will support. We recommend the Naval METOC community continue to use the current NAVGEM thunderstorm forecast for mission planning and situational awareness at leads of 72 hours or less and anticipate shorter lead forecasts from COAMPS to be developed.

APPENDIX A. STATISTICAL DATA

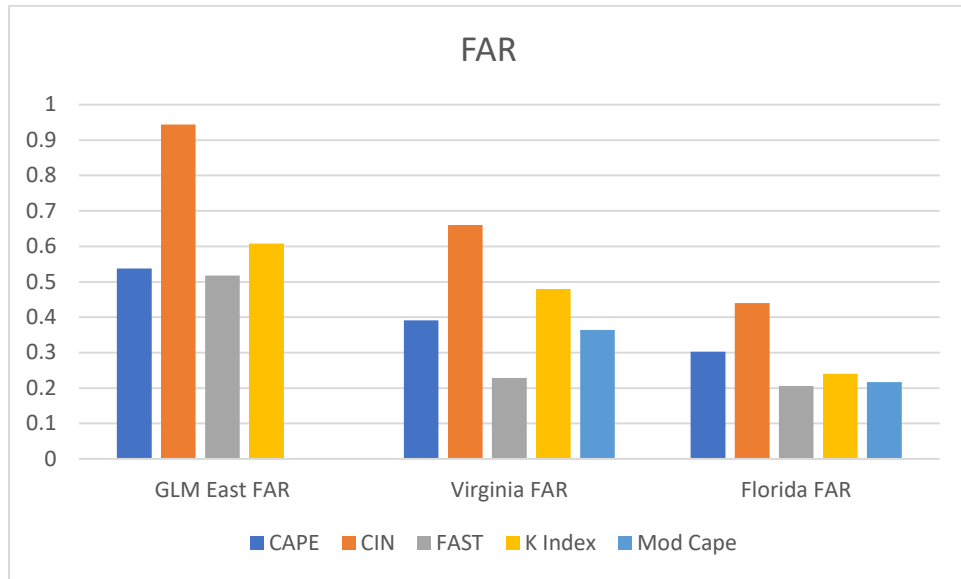
A. PROBABILITY OF DETECTION FOR ALL REGIONS



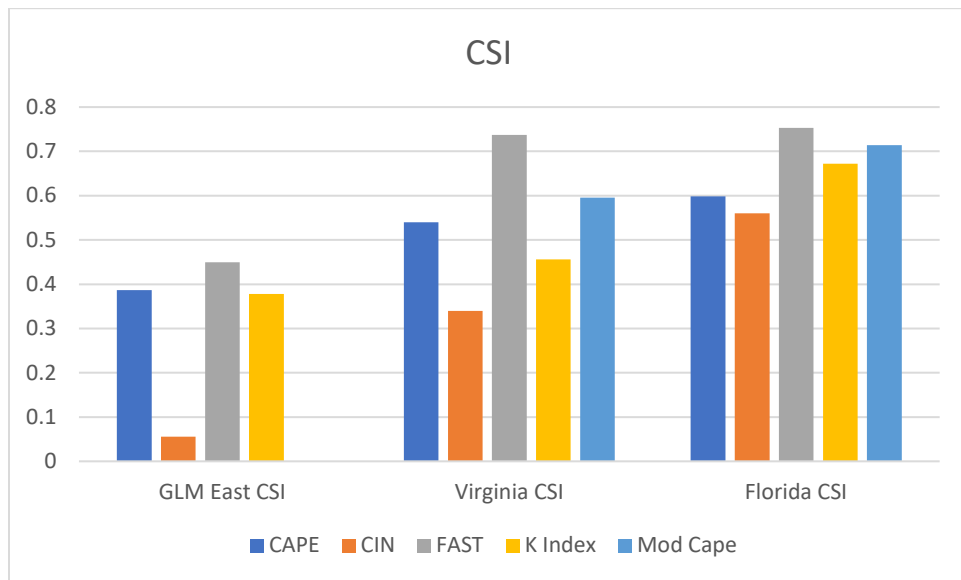
B. PROPORTIONAL CORRECT FOR ALL REGIONS



C. FALSE ALARM RATE FOR ALL REGIONS

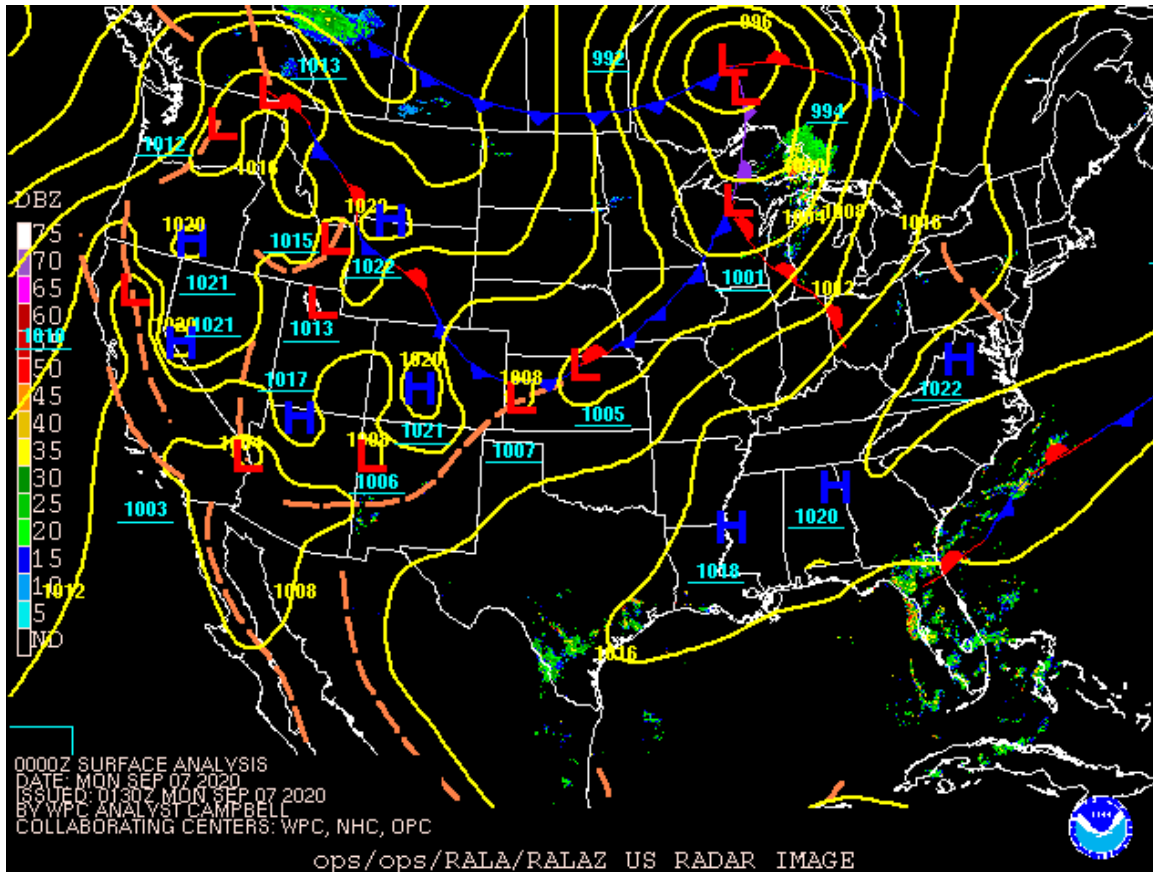


D. CRITICAL SUCCESS INDEX FOR ALL REGIONS



APPENDIX B. ADDITIONAL FORECASTS

A. SEPTEMBER 07 00Z



Map retrieved from the Weather Prediction Center, November 20, 2020,
https://www.wpc.ncep.noaa.gov/archives/web_pages/sfc/sfc_archive_maps.php?arcdte=09/07/2020&selmap=2020090700&maptype=radsfcus_exp

Figure 36. September 07, 2020 00Z Surface analysis with radar mosaic overlay

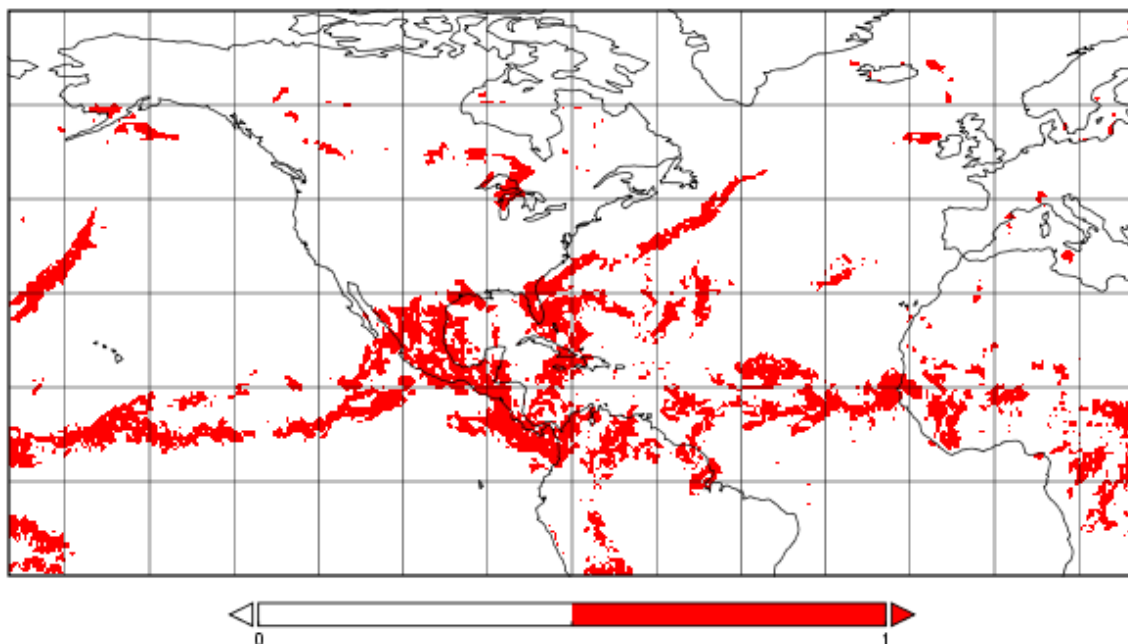


Figure 37. September 07, 2020 00Z FAST thunderstorm forecast

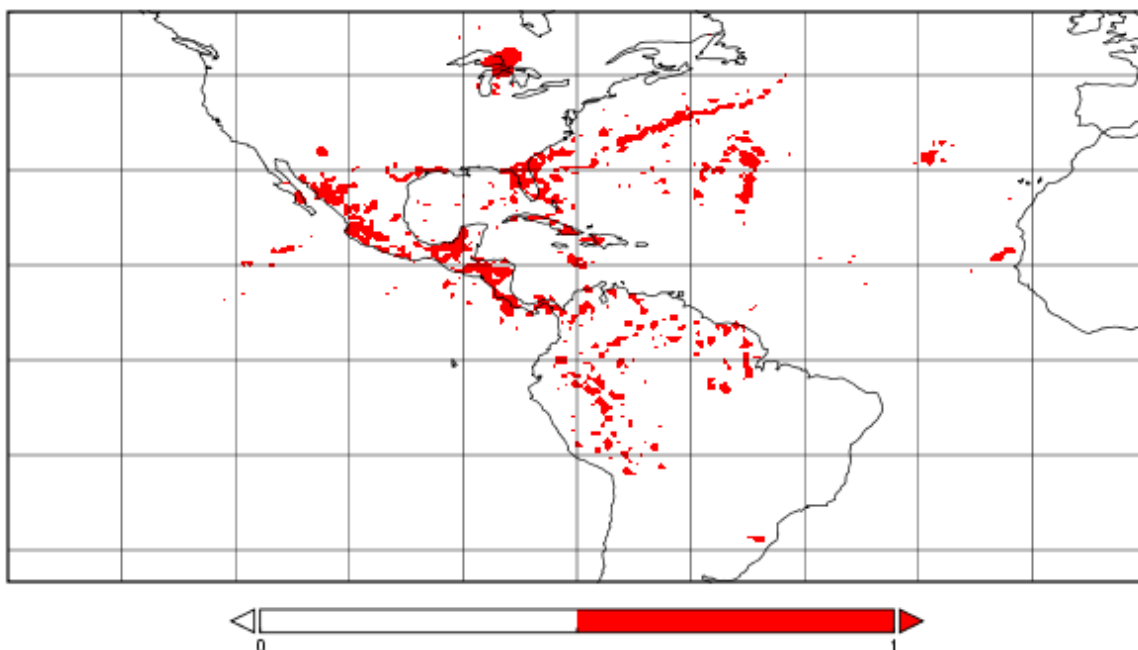
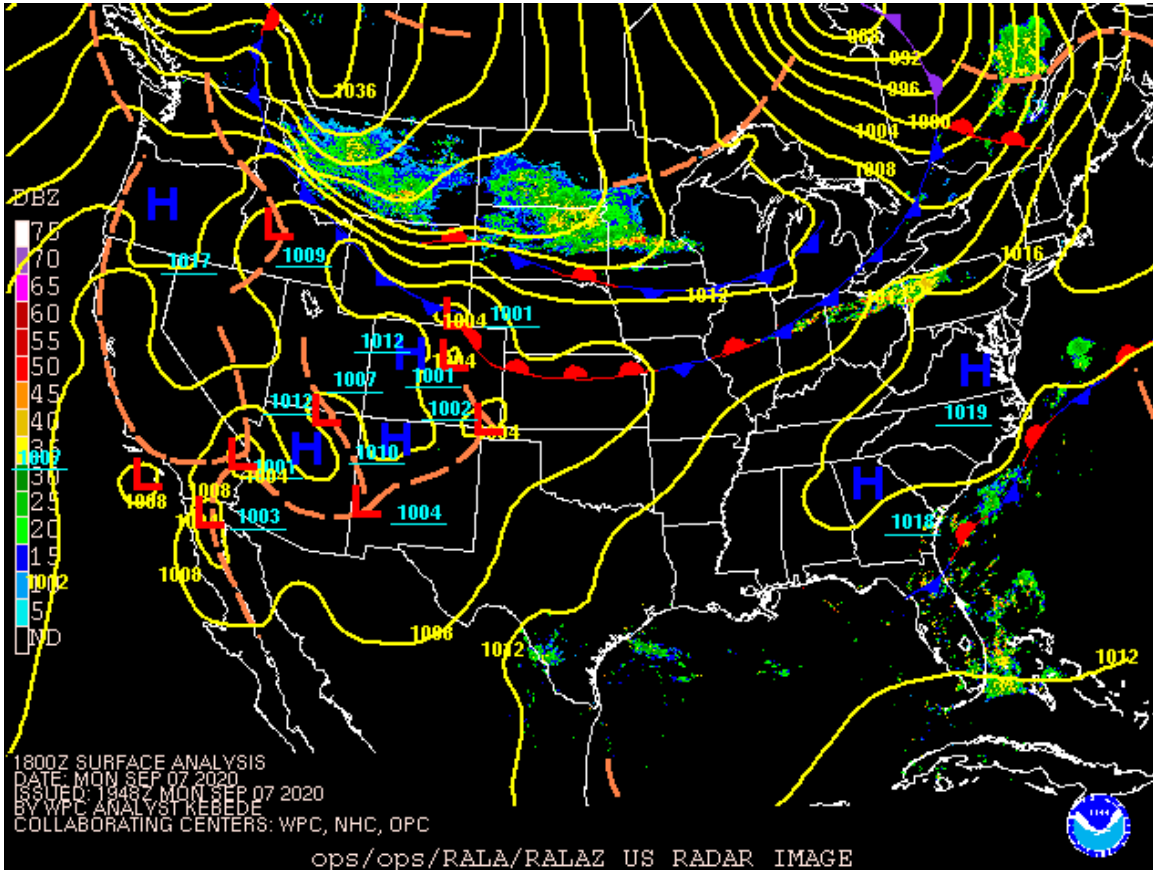


Figure 38. September 07, 2020 00Z GLM thunderstorm data

B. SEPTEMBER 07 18Z



Map retrieved from the Weather Prediction Center November 20, 2020,
https://www.wpc.ncep.noaa.gov/archives/web_pages/sfc/sfc_archive_maps.php?arcdte=09/07/2020&selmap=2020090718&matype=radsfcus_exp

Figure 39. September 07, 2020 18Z U.S. surface analysis with radar mosaic overlay

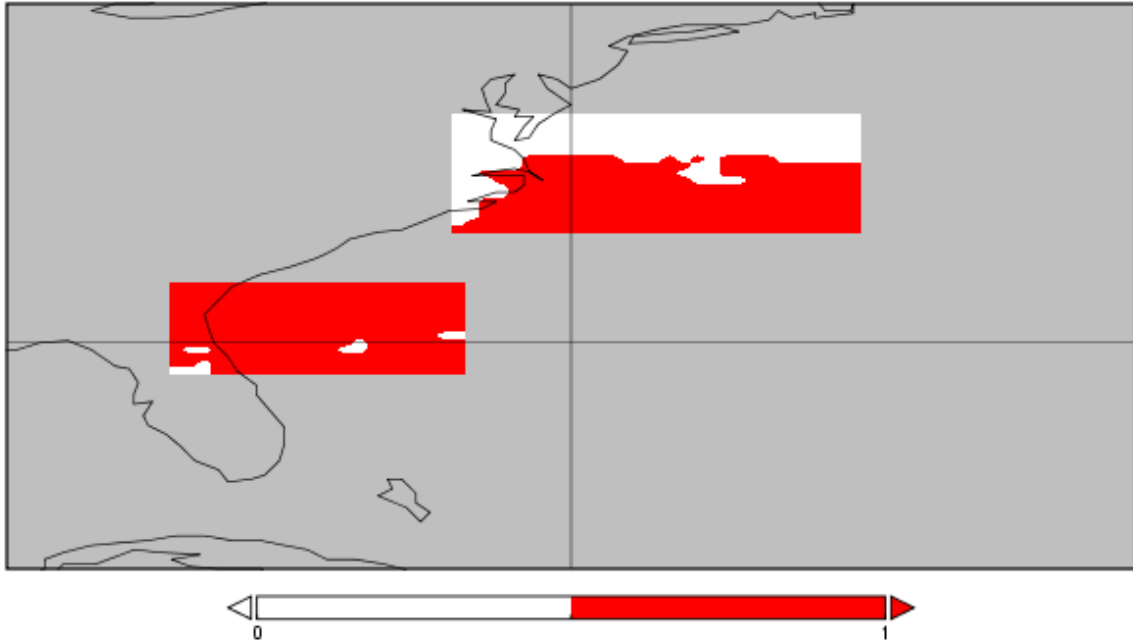


Figure 40. September 07, 2020 18Z Modified CAPE forecast

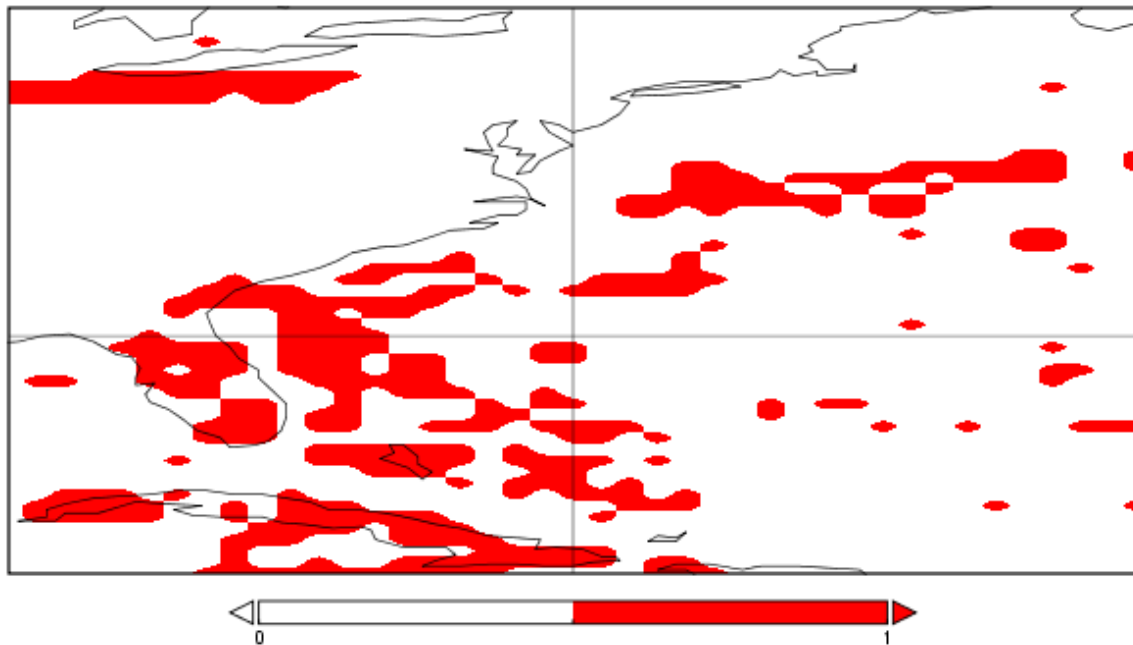
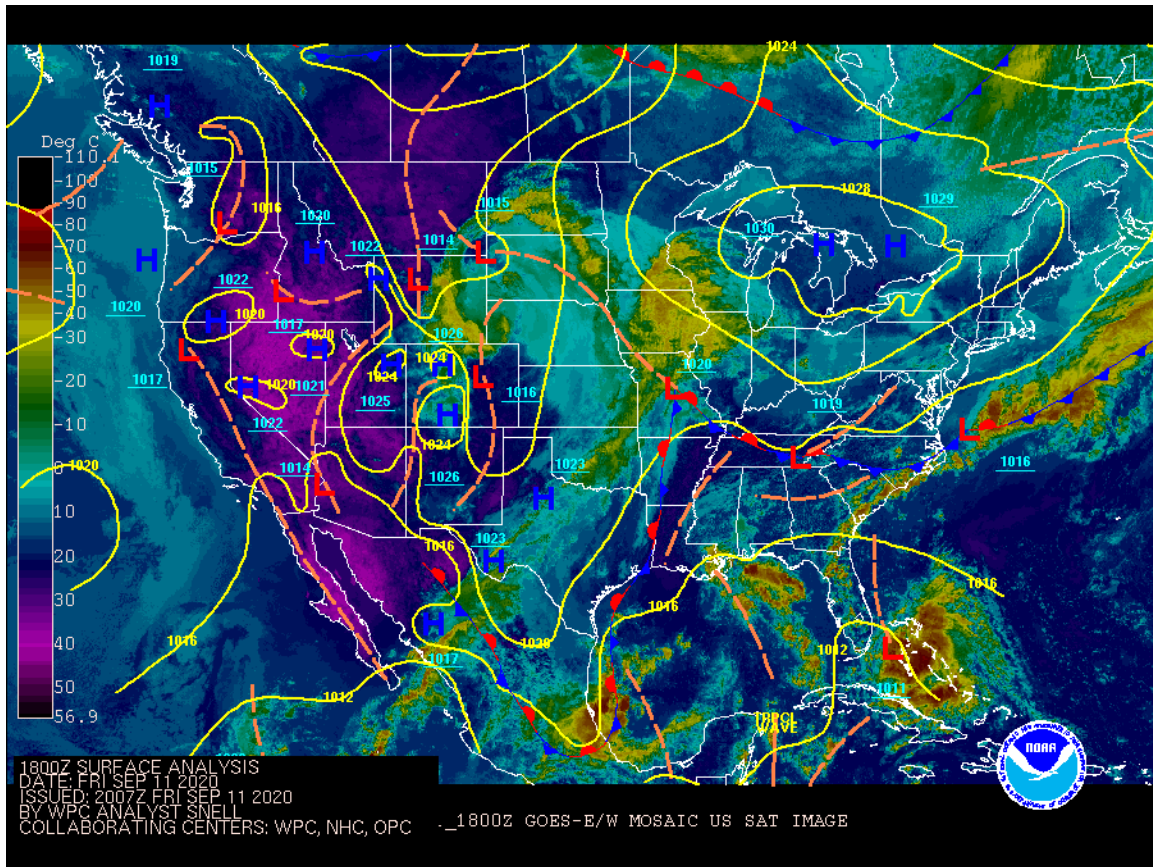


Figure 41. September 07, 2020 18Z GLM thunderstorm data

C. SEPTEMBER 11 18Z



Map retrieved from the Weather Prediction Center, November 20, 2020,
https://www.wpc.ncep.noaa.gov/archives/web_pages/sfc/sfc_archive_maps.php?arcdte=09/11/2020&selmap=2020091118&maptype=ussatsfc

Figure 42. September 11, 2020 18Z U.S. surface analysis with satellite overlay Source: Weather Prediction Center

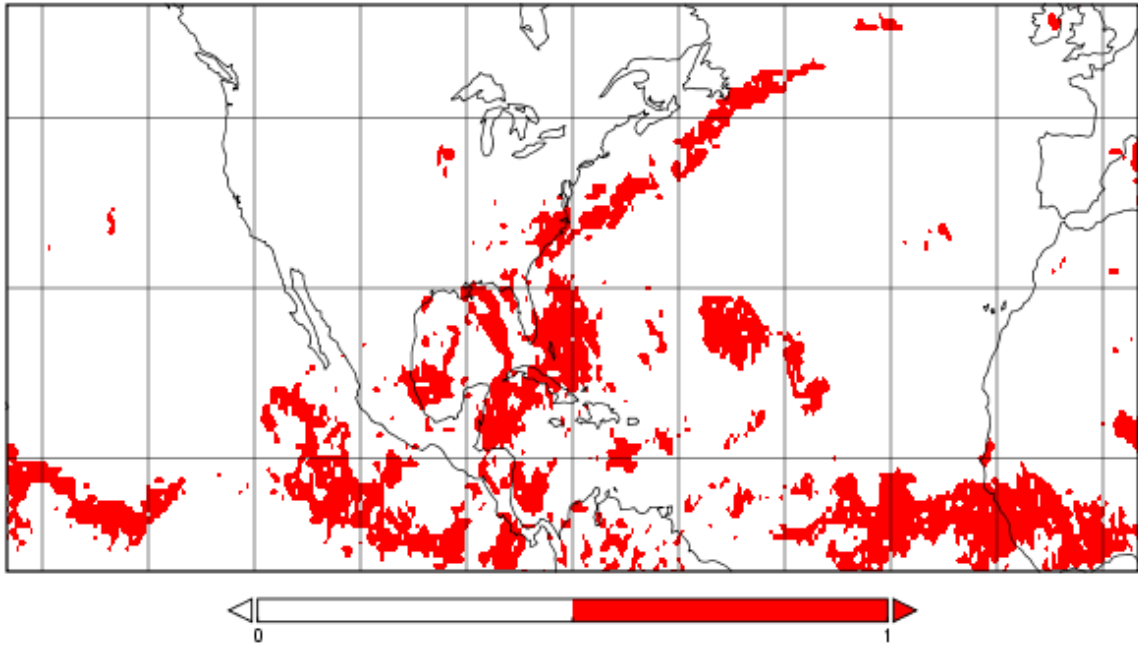


Figure 43. September 11, 2020 18Z FAST thunderstorm forecast

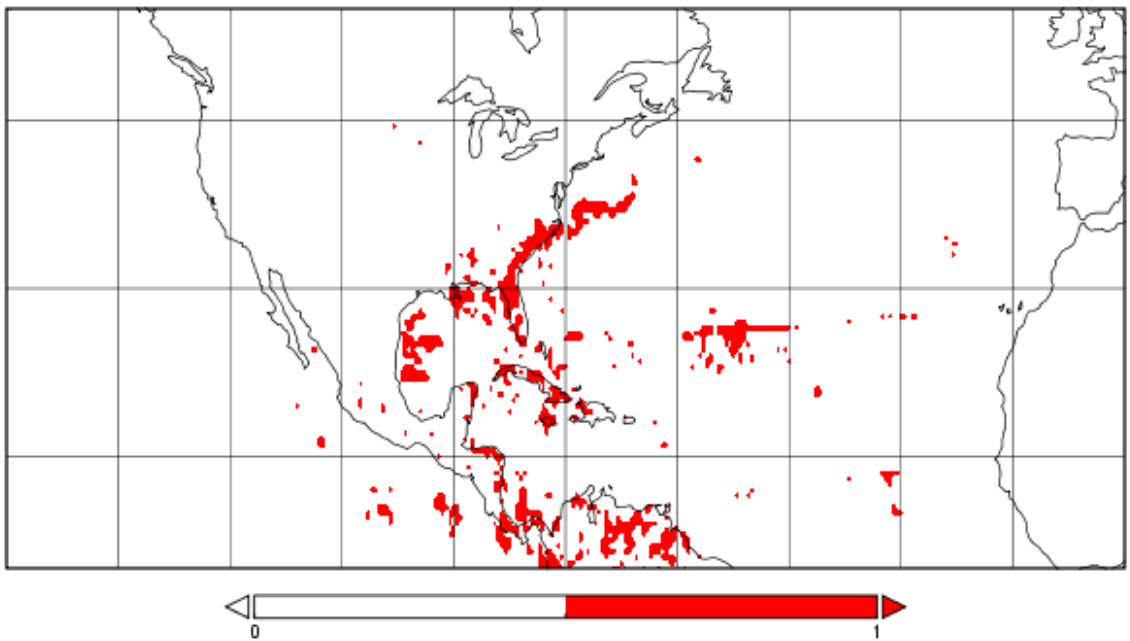
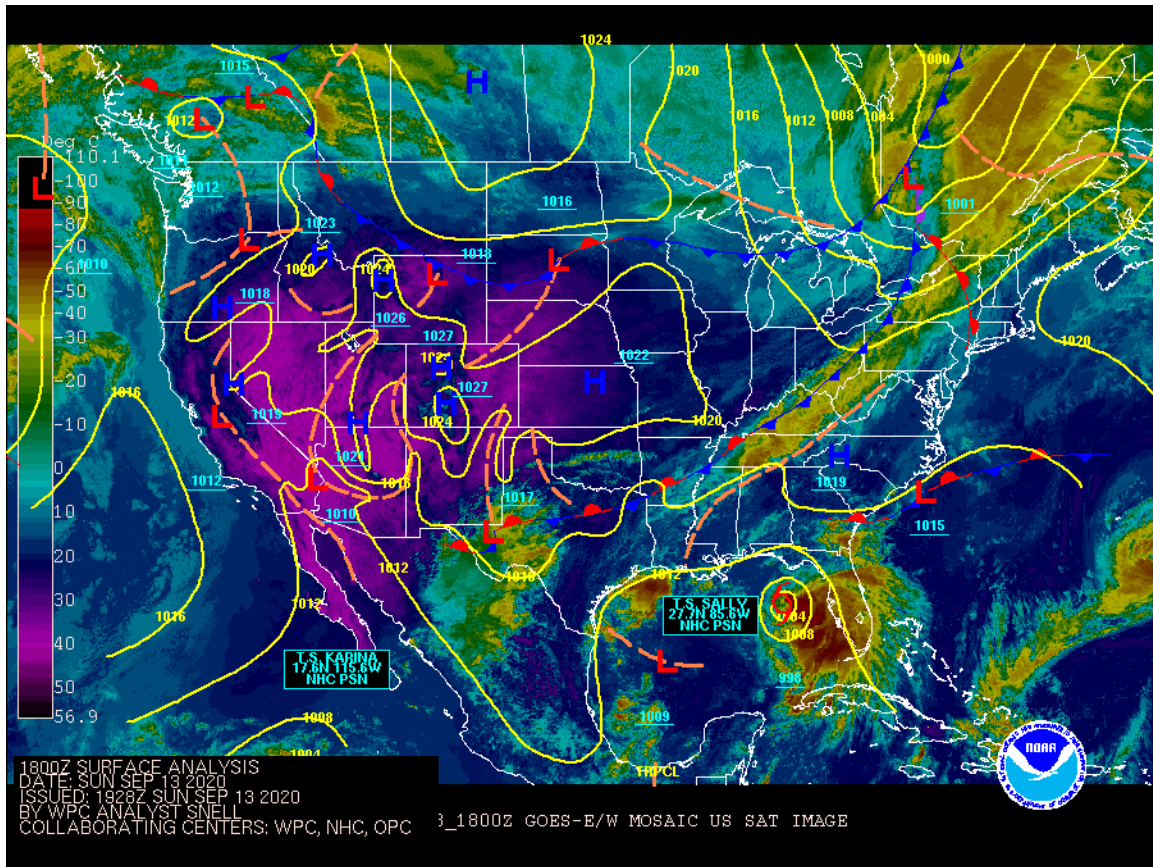


Figure 44. September 11, 2020 18Z GLM thunderstorm data

D. SEPTEMBER 13 18Z



Map retrieved from the Weather Prediction Center, November 20, 2020,
https://www.wpc.ncep.noaa.gov/archives/web_pages/sfc/sfc_archive_maps.php?arcdte=09/13/2020&selmap=2020091318&maptype=ussatsfc

Figure 45. September 13, 2020, 18Z U.S. surface analysis with satellite overlay Source: Weather Prediction Center

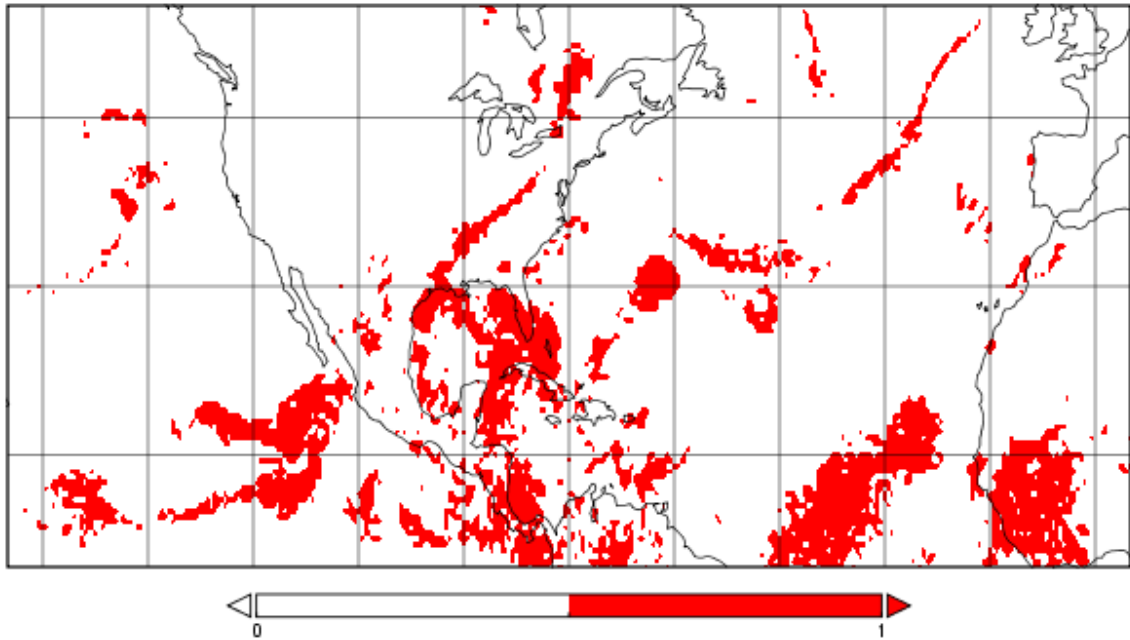


Figure 46. September 13, 2020 18Z FAST thunderstorm forecast

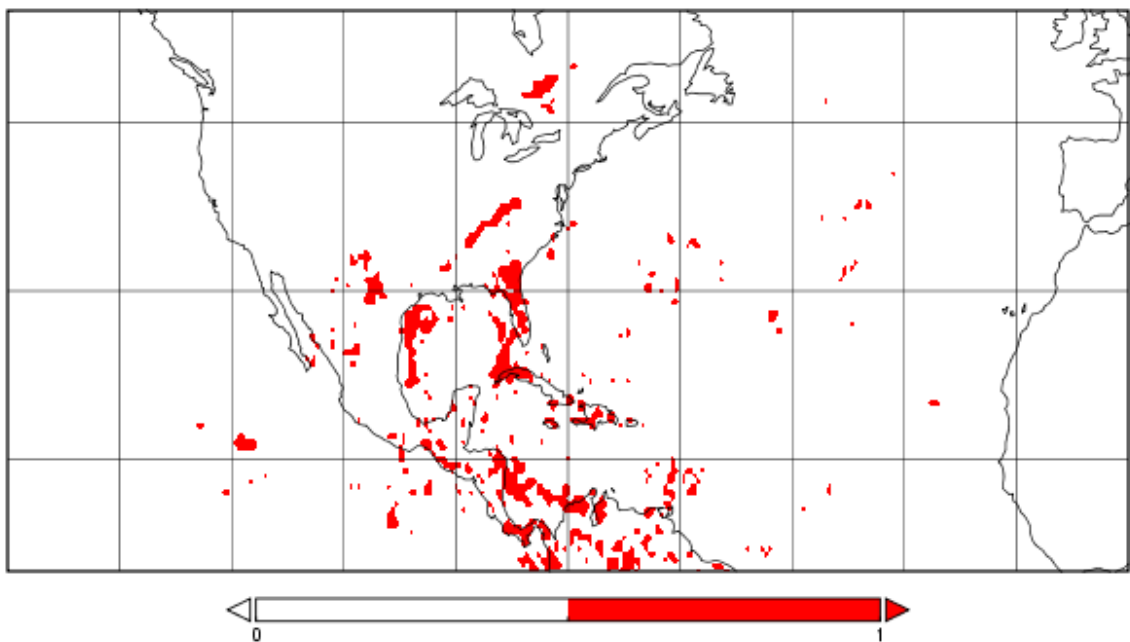
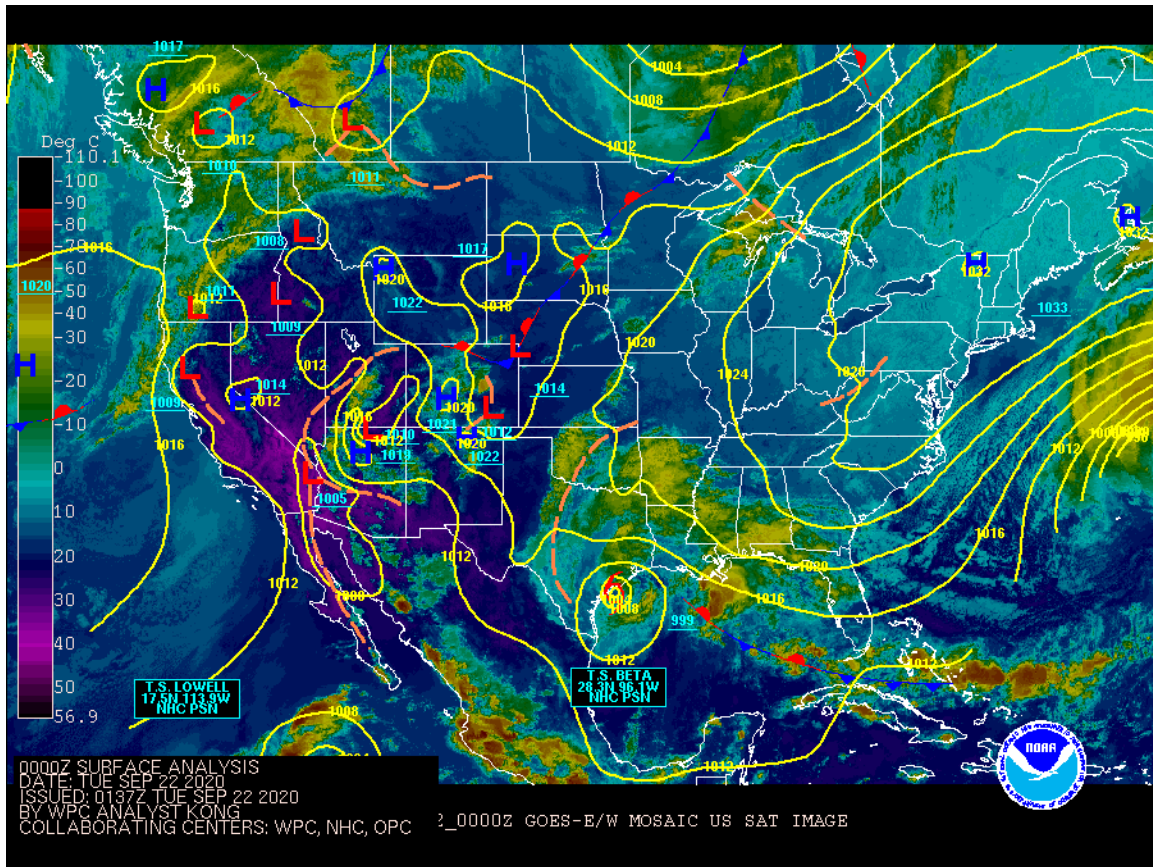


Figure 47. September 13, 2020 18Z GLM thunderstorm data

E. SEPTEMBER 22 00Z



Map retrieved from the Weather Prediction Center, November 20, 2020,
https://www.wpc.ncep.noaa.gov/archives/web_pages/sfc/sfc_archive_maps.php?arcdte=09/22/2020&selmap=2020092200&maptype=ussatsfc

Figure 48. September 22, 2020 00Z U.S. surface analysis with satellite overlay Source: Weather Prediction Center

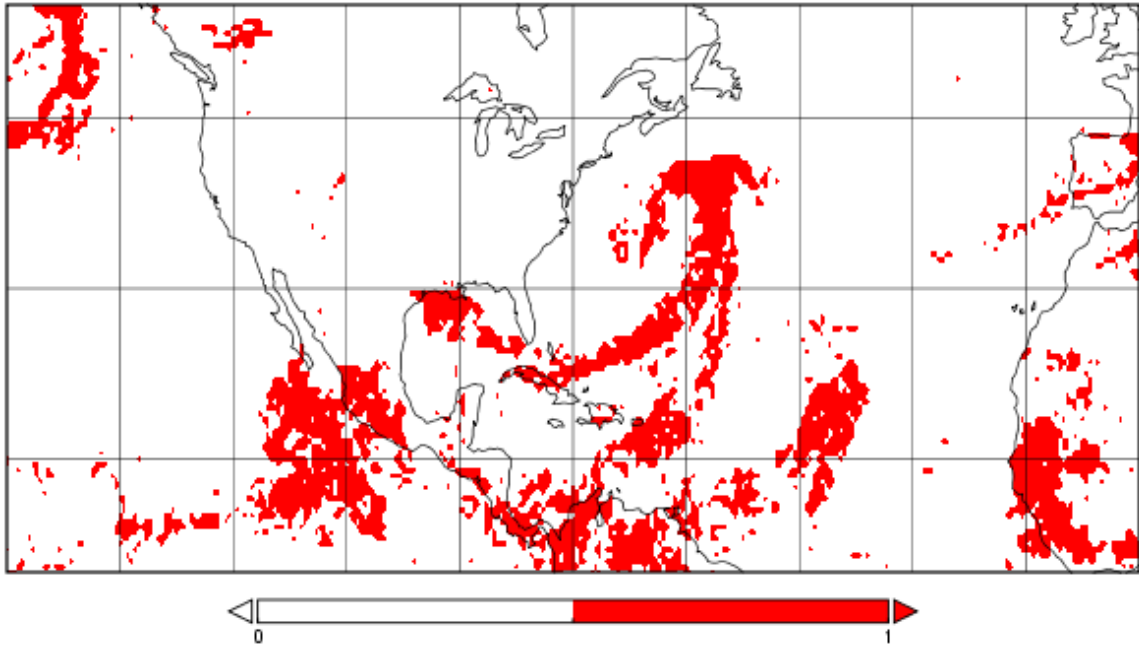


Figure 49. September 22, 2020 00Z FAST thunderstorm forecast

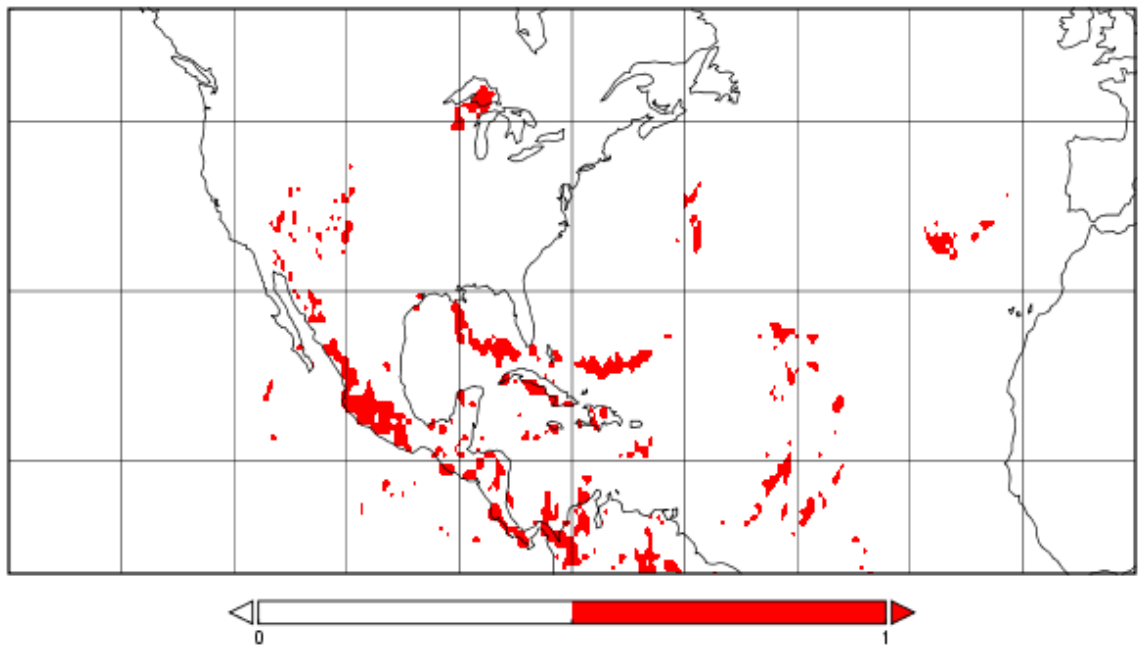


Figure 50. September 22, 2020 00Z GLM thunderstorm data 092200

LIST OF REFERENCES

- Blaylock, B. K., and J. D. Horel, 2020: Comparison of lightning forecasts from the high-resolution rapid refresh model to geostationary lightning mapper observations. *Wea. Forecasting*, **35**, 401–416, doi:10.1175/WAF-D-19-0141.1.
- Bouttier, F., and H. Marchal, 2020: Probabilistic thunderstorm forecasting by blending multiple ensembles, *Dynamic Meteor. and Oceanography*, **72**, 1–19, doi:10.1080/16000870.2019.1696142
- Bryan, G. H., J. C. Wyngaard, and J. M. Fritsch, 2003: Resolution requirements for the simulation of deep moist convection. *Monthly Weather Review*, **131**, 2394–2416.
- Cantu, A., 2001: The role of weather in Class A naval aviation mishaps. M.S. thesis, Dept. of Meteorology, Naval Postgraduate School, 106 pp.
- CNMOC, 2020: Naval oceanography improves the fleet’s asymmetric advantage. Accessed 27 October 2020, <https://www.doncio.navy.mil/CHIPS/ArticleDetails.aspx?ID=13427>
- Colby, F. P., 1984: Convective inhibition as a predictor of convection during AVE-SESAME II. *Mon..Wea.. Rev.*, **112**, 2239–2252, doi:10.1175/1520-0493(1984)112<2239:CIAAPO>2.0.CO;2.
- COESA, 1976: U.S. Standard Atmosphere, 1976. NOAA 227 pp.
- Doswell III, C., and D. Schultz, 2006: On the Use of Indices and Parameters in Forecasting Severe Storms. *E-Journal Of Severe Storms Meteorology*, **1**. Accessed December 2, 2020, <https://ejssm.org/ojs/index.php/ejssm/article/view/11/12>
- FNMOCC, 2015: Model information. Accessed 4 March 2020, <https://portal-alpha.fnmoc.navy.mil/modelinfo/ModelInfo.html>
- Mireles, M. R., K. L. Pederson, and C. H. Elford, 2007: *Meteorological Techniques*. AFWA, 233 pp.
- Morrison, H., 2017: An analytic description of the structure and evolution of growing deep cumulus updrafts. *Journal of the Atmospheric Sciences*, **74**, 809–834, doi:10.1175/JAS-D-16-0234.1
- Morrison, H., J. M. Peters, A. C. Varble, W. M. Hannah, and S. E. Giangrande, 2020: Thermal chains and entrainment in cumulus updrafts, Part 1: Theoretical description. *Journal of the Atmospheric Sciences*, **77**, doi:10.1175/jas-d-19-0243.1.

- NOAA, 2015: Environmental parameters and indices. Accessed 4 March 2020, <https://www.weather.gov/lmk/indices>
- Okon, J., 2020: Naval meteorology and oceanography 2020 American meteorological society update. Accessed 28 February 2020, <https://ams.confex.com/ams/2020Annual/videogateway.cgi/id/520512?recordingid=520512>
- Peters, J. M., W. Hannah, and H. Morrison, 2019a: The influence of vertical wind shear on moist thermals. *Journal of the Atmospheric Sciences*, **76**, 1645–1659, doi:10.1175/JAS-D-18-0296.1
- Peters, J. M., C. J. Nowotarski, and H. Morrison, 2019b: The role of vertical wind shear in modulating maximum supercell updraft velocities. *Journal of The Atmospheric Sciences*, **76**, 3169–3189, doi:10.1175/JAS-D-19-0096.1
- Peters, J.M., C. J. Nowotarski, J. P. Mulholland, and R. L. Thompson, 2020a: The Influences of Effective Inflow Layer Streamwise Vorticity and Storm-Relative Flow on Supercell Updraft Properties. *Journal of the Atmospheric Sciences*, **77**, doi:10.1175/JAS-D-19-0355.1.
- Peters, J. M., H. Morrison, C. J. Nowotarski, J. P. Mullholland, and R. L. Thompson, 2020b: A formula for the maximum vertical velocity of supercell updrafts, *J. Atmos. Sci.* **77**, 3747–3757. doi:10.1175/JAS-D-20-0103.1
- Peters, J. M., H. Morrison, A. C. Varble, W. M. Hannah., and S. E. Giangrande, 2020c: Thermal chains and entrainment in cumulus updrafts, Part 2: Analysis of idealized simulations. *Journal of the Atmospheric Sciences*, **77**, doi:10.1175/jas-d-19-0244.1
- Ranquist, E. A., M. Steiner M., and B. Argrow, 2017: Exploring the range of weather impacts on UAS operations *18th Conference on Aviation, Range, and Aerospace Meteorology*, Seattle, WA, Amer. Meteor. Soc., JP3.1, <https://ams.confex.com/ams/97Annual/webprogram/Paper309274.html>
- Rauber R.M., J. E. Walsh, and D. J. Charlevoix, D.J., 2002: *Severe and Hazardous Weather*, Kendall/Hunt, 558 pp.
- Rotunno, R., J. B. Klemp, and M. L. Weisman, 1988: A theory for strong, long-lived squall lines. *Journal of the Atmospheric Sciences*, **45**, 463–485, doi:10.1175/1520-0469(1988)045<0463:ATFSSL>2.0.CO;2
- Sampson, C. R., A. J. Schrader, 1999: The automated tropical cyclone forecasting system (Version 3.2). *Bull. Amer. Meteor. Soc.*, **81**, 1231–1240, doi:10.1175/1520-0477(2000)081%3C1231:TATCFS%3E2.3.CO;2

- Secretary of the Navy, 2018: DON strategic roadmap for unmanned systems. Accessed 7 November 2020, <https://www.secnav.navy.mil/rda/Documents/DON-Strategic-Roadmap-for-Unmanned-Systems.docx>
- Stull, R. B., 2000: *Meteorology for Scientist and Engineers*, Brooks/Cole, 502 pp.
- Trojan, G, 2008: *GRIB Aviation Products, WAFC Washington Progress Report*, Science International Applications Corporation and Environmental Modeling Center/NCEP/NWS/NOAA 16 pp.
- Thompson, R., C. Mead, and R. Edwards, 2007: Effective Storm-Relative Helicity and Bulk Shear in Supercell Thunderstorm Environments. *Weather and Forecasting*, **22**, doi:10.1175/WAF969.1.
- Weisman, M. L., W. Skamarock, and J. B. Klemp, 1997: The resolution dependence of explicitly modeled convective systems. *Monthly Weather Review*, **125**, 527–548, doi:10.1175/1520-0493(1997)125<0527:TRDOEM>2.0.CO;2
- Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*, Academic Press, 676 pp.
- Williams E., and S. Stanfill, 2003: The physical origin of the land–ocean contrast in lightning activity. *Comptes Rendus Physique*, **3**, 1277–1292, doi:10.1016/S1631-0705(02)01407-X.
- Zhang, F., C. Snyder, and R. Rotunno, 2003: Effects of moist convection on mesoscale predictability. *Journal Of The Atmospheric Sciences*, **60**, 1173–1185, doi:10.1175/1520-0469(2003)060<1173:EOMCOM>2.0.CO;2

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
Ft. Belvoir, Virginia
2. Dudley Knox Library
Naval Postgraduate School
Monterey, California