3.8 LATEST DEVELOPMENT IN THE NWS' EXTRA-TROPICAL STORM SURGE MODEL, AND PROBABILISTIC EXTRA-TROPICAL STORM SURGE MODEL

Huiqing Liu[1*], Arthur Taylor[2] and Kwangmin Kang[3]
1. Ace Info Solutions, Reston, VA
2. NOAA / NWS / STI / Meteorological Development Laboratory, Silver Spring, MD
3. KBR Wyle, 601 Jefferson St., Houston, TX 77002, 301-427-9445

1. INTRODUCTION

The National Weather Service's (NWS) Meteorological Development Laboratory (MDL) developed the Extra-Tropical Storm Surge (ETSS) model in 1995 (Kim et al. 2006) to predict storm surge to the coastline, which was later modified to provide coastal inundation based on storm surge and tide (Liu et al. 2015 and Liu et al. 2016). The resulting coastal inundation model provided the basis for MDL's development of the Probabilistic Extra-Tropical Storm Surge (P-ETSS) model in 2017 (Liu and Taylor 2018). When P-ETSS was implemented in 2017, MDL had established the capability to operationally provide both deterministic and probabilistic inundation guidance based on storm surge and tide from extra-tropical storms four times a day along the United States' Eastern and Western seaboards, the Gulf of Mexico and Alaska.

The deterministic model (ETSS) can be improved as higher resolution (13 km vs 55 km or 0.5 degree) wind and pressure input are now available from the Global Forecast System (GFS) (Taylor et al. 2015). Additionally, the East Coast basin hasn't been updated in 9 years, so updating it with new bathymetry and topography data will improve its accuracy. This will also allow the coverage to be expanded to cover Puerto Rico and Virgin Islands in anticipation of wave coupling. Also, the Gulf of Mexico basin hasn't been updated in 7 years, so updating it with new bathymetry and topography data will improve its accuracy. This will also allow the coverage to be expanded to cover the entire Gulf of Mexico further improving its accuracy and potentially assisting Mexico.

As for the probabilistic guidance, the number of members used in the extra-tropical ensemble is significantly less than the number used in the tropical ensemble. P-ETSS uses the 21 member Global Ensemble Forecast System (GEFS) 55 km wind and pressure fields whereas the Probabilistic tropical-cyclone storm Surge (P-Surge) (Taylor and Glahn 2008) uses an approximately 630 member parametric tropical cyclone wind ensemble. This difference in number of ensemble members leads to concerns about P-ETSS' reliability and spread, so MDL has worked to increase the number of ensembles in P-ETSS by using the 42 member North American Ensemble Forecast System (NAEFS). Unfortunately it is only available at 0 and 12Z, so at 6 and 18Z P-ETSS would still need to use the 21 GEFS members. The result would be a system with alternating characteristics of ensemble spread potentially causing user confusion. So while the validation results for the NAEFS based products are presented here, MDL is still developing a method to smooth out the cycle variations.

Additionally, ETSS does not account for model bias, nor water level components such as sea level rise, waves, and river flooding. MDL addressed this in 2000 by developing a station based post-processing methodology for ETSS which statistically accounts for those components based on recent observations. This has not been done for P-ETSS, so another improvement is to add ETSS' statistical post processing methodology at stations to enhance the overall guidance.

This paper describes the details of these efforts and provides validation using historic events. Section 2 describes the improvements of ETSS and P-ETSS. Section 3 lists the historic storms along with observations used to validate the ETSS and P-ETSS results. Section 4 presents the results. Section 5 discusses the post processing adjustment. The paper concludes with a summary and discussion in Section 6.

* Corresponding Author: Huiqing Liu, 1325 East-West Hwy, Silver Spring, MD 20910-3280; e-mail: huiqing.liu@noaa.gov

2. IMPROVEMENTS OF ETSS and P-ETSS

Table 1a. The status of current operational ETSS 2.2 and P-ETSS 1.0. Red text indicates a feature that is planned to be enhanced in ETSS 2.3 and P-ETSS 1.1.

	ETSS	P-ETSS
Model Type	Deterministic	Ensemble
Phenomena	Overland inundation based on surge + tide	
Areas: no nesting	Gulf of Alaska (Created - Apr 2008) Bering Beaufort Chuckchi Seas (Created - Nov 2015) West Coast (Created - Feb 2017)	
Areas: fine res nesting	East Coast (Created - Feb 2009) Gulf of Mexico (Created - Jan 2011)	
Forcing Resolution	3-hourly 0.5 degree (55-km)	
Forcing Frequency	4x a day - GFS wind/pressure	4x a day - 21 member GEFS wind/pressure
Post-Processing	Station based bias adjustment	None
Images	Created on WCOSS' development machine, which is not as stable as WCOSS' operational machine.	

Table 1b. The updates for ETSS 2.3 and P-ETSS 1.1 with green text indicating proposed improvements.

	ETSS	P-ETSS
Areas: fine resolution nesting	Upgrade East Coast (Created - Feb 2018) Upgrade Gulf of Mexico (Created - Jun 2018)	No Change
Forcing Resolution	1-hourly 13-km GFS winds	No Change
Forcing Frequency	No Change	2x a day - 42 member NAEFS 2x a day - 21 member GEFS
Post-Processing	No change	Add Station based bias adjustment
Images	Created on WCOSS' operational machine, which is very stable.	

Table 1a describes NWS' current operational ETSS (version 2.2) and P-ETSS (version 1.0) models. ETSS is a deterministic coastal inundation model forced by the 3-hourly 0.5 degree (55 km) Global Forecast System (GFS) and P-ETSS is a coastal inundation ensemble model forced by the 3-hourly 0.5 degree (55 km) GEFS. Both systems provide storm surge and overland inundation guidance 4 times a day based on surge and tide for all U.S coastal areas. They provide finer resolution guidance along the East Coast and the Gulf of Mexico by nesting high resolution basins. ETSS has a station based bias correction post-processing whereas P-ETSS does not. The images displaying ETSS and P-ETSS results are created on NOAA's Weather and

Climate Operational Supercomputing System (WCOSS) development machine.

The proposed improvements for both models are listed in Table 1b. For ETSS (version 2.3), the first improvement is to expand the East Coast basin to cover Puerto Rico and the Virgin Islands in anticipation of future wave coupling. The second improvement is to expand the Gulf of Mexico basin to cover the whole Gulf of Mexico and parts of the Yucatan Peninsula (Fig. 1), which should allow ETSS to better model phenomena that come from outside our coastal area of interest (e.g. forerunner surge or reflection of waves off the Mexican coastline). The third improvement is to upgrade from the 3-hourly 0.5 degree GFS winds to 1-hourly Semi-Lagrangian 13-km GFS winds for forcing.

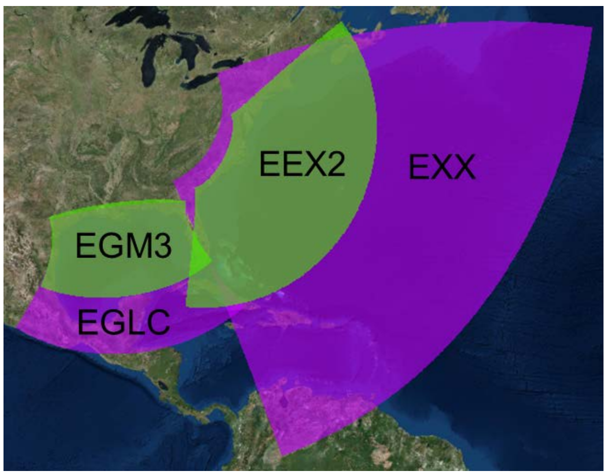

Fig. 1 Old and New East Coast basins and Gulf of Mexico basins. EXX (purple) replaces EEX2 (green) and EGLC (purple) replaces EGM3 (green).

For P-ETSS (version 1.1), the first improvement is to increase the number of ensemble members via the 42 member NAEFS. Since the NAEFS is only available at 0 and 12Z, P-ETSS currently uses the 21 member GEFS at 6 and 18Z. As mentioned earlier, this will not be part of P-ETSS until a method is developed to smooth out the cycle variations; however the results of using the 0 and 12Z NAEFS are presented later in the paper. The second improvement to P-ETSS is to add the same bias correction at stations that is done for ETSS.

The last improvement for both models is to transition the image creation software from WCOSS' development machine to WCOSS' production machine. This should improve the stability of the ETSS and P-ETSS guidance websites. While not a scientific improvement, this is important as the websites have proven to be very useful for local forecast offices when they are issuing extra-tropical storm surge watch and warning forecasts. Getting stable model guidance to the forecaster is critical for them to issue timely and accurate forecasts.

3. HISTORICAL STORMS

To evaluate the improvements of ETSS 2.3 and P-ETSS 1.1 over the current operational versions, a quantitative analysis using retrospective model runs was made over the past two years. The two-year limitation was due to National Centers for Environmental Prediction (NCEP) not having an official archive of the 13 km resolution GFS wind data. MDL's archive contains data from September 2017 to February 2019 with some gaps due to WCOSS development and production machine switches. For NAEFS, there is a similar limitation. NCEP didn't have an official archive of the 0.5 degree resolution winds from the 21 Canadian members (half of the NAEFS ensemble members). The GEFS group had their own archived data between July 2017 and May 2018 with some gaps.

Of the interesting storms over the last two years, we chose to evaluate ETSS model performance with 17 storm events and P-ETSS model performance with 9 storm events. Eight storm events were dropped from the P-ETSS evaluation due to gaps in the archived NAEFS data. Table 2a lists the storm events used to evaluate the ETSS model, which includes both tropical and extra-tropical storms impacting the East Coast, Gulf of Mexico and Alaska over the past two years. The storm events used to evaluate the P-ETSS model are listed in Table 2b.

Table 2a. The storms used to validate ETSS 2.3

Storm #	Storm-name	Year	Forecast Start Time
1	Irma	2017	SEP 08 18Z – 13 06Z
2	Jose	2017	SEP 15 18Z – 23 06Z
3	Maria	2017	SEP 22 18Z – 30 06Z
4	NYC	2017	OCT 26 18Z – NOV 03 06Z
5	AK	2017	DEC 17 18Z – 25 06Z
6	EST	2018	DEC 31 18Z – JAN 07 06Z
7	EST	2018	JAN 13 18Z – 21 06Z
8	AK	2018	JAN 13 18Z – 21 06Z
9	AK	2018	FEB 16 18Z – 24 06Z
10	EST	2018	FEB 26 18Z – MAR 11 06Z
11	AK	2018	FEB 26 18Z – MAR 11 06Z
12	EST	2018	MAR 09 18Z – 16 06Z
13	EST	2018	MAR 17 18Z – 25 06Z
14	EST	2018	APR 11 18Z – 20 06Z
15	GOM	2018	APR 11 18Z – 20 06Z
16	AK	2018	APR 21 18Z – 29 06Z
17	Alberto	2018	MAY 21 18Z – 31 06Z

Table 2b. The storms used to validate P-ETSS 1.1

Storm-#	Storm-name	Year	Forecast Start Time
1	NYC	2017	OCT 26 18Z – NOV 03 06Z
2	AK	2017	NOV 18 18Z – 26 06Z
3	AK	2017	DEC 17 18Z – 25 06Z
4	EST	2018	DEC 31 18Z – JAN 07 06Z
5	EST	2018	FEB 26 18Z – MAR 11 06Z
6	EST	2018	APR 11 18Z – 20 06Z
7	GOM	2018	APR 11 18Z – 20 06Z
8	AK	2018	APR 21 18Z – 29 06Z
9	Alberto	2018	MAY 21 18Z – 31 06Z

4. RESULTS

4.1 ETSS RESULTS

ETSS skill scores for 12, 24, 36, 48, 60, 72, 84 and 96 hour projection windows were evaluated against the tide gauge observation time frames based on statistical scores calculated from a 96 hour time series. The reason a 96 hour time series was selected was to have a consistent time window for all events that focused on when the water levels were most significantly impacted. The 96 hour time series was created by splicing together 6 hour slices from consecutive model runs. For example, the 24-hr projection window spliced hours 19 to 24 from one model run to hours 19 to 24 from the next consecutive model run. This results in a relatively constant projection thereby reducing the impact of errors within different projections on the assessment. Model performance was then assessed based on the average of the following scores over the various tide gauge observation time frames:

1) Root Mean Squared Error (RMSE),

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}}$$

2) Peak Absolute Error (PAE),

$$PAE = abs(\max(X_{obs,i}) - \max(X_{model,i}))$$

The 17 storm events were divided into 3 groups based on region of impact: East Coast (EST Region), Gulf of Mexico (GOM Region) and Alaska (AK Region). The average RMSE and PAE for the tide gauge observation time frames for the 12, 24, 36, 48, 60, 72, 84 and 96 hour projection windows are shown in Fig. 2. The results show a reduction of the average RMSE and PAE from the updated ETSS model for all projection hours in all locations.

It is worth noting that the improvement for the Alaska region is the smallest among the three regions. This is likely because the wind forcing for Alaska was improved but not the basin. For the East Coast and the Gulf of Mexico, we improved both the wind forcing and updated the basins.

It is also worth noting that the RMSE and PAE values in the Alaska region are larger than in the other two regions because no high resolution smaller basins are available for nesting. For the East Coast and Gulf of Mexico, ETSS nests smaller high resolution basins within larger coarse basins (Liu et al. 2015). The results suggest that developing high resolution grids in the Alaska region will improve ETSS model guidance there.

The storm specific reduction (ETSS 2.2 – ETSS 2.3) to RMSE per projection for the 17 storm events are listed in Table 3. The positive numbers are green as the newer version reduced the RMSE score, whereas the negative values are red since the newer version increased the RMSE. The result shows that ETSS 2.3 performs better in most scenarios, with just a few cases of small negative values in the later projections.

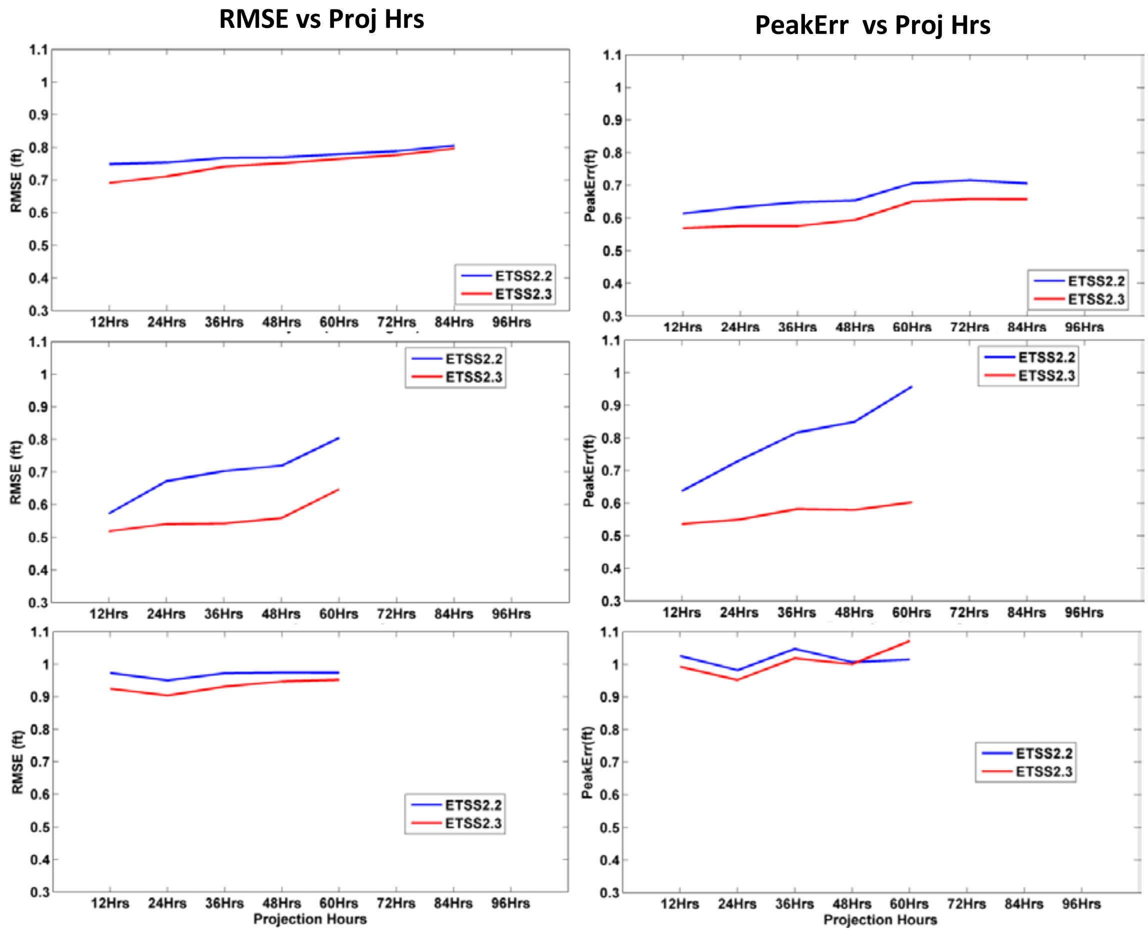

Fig. 2 The left column is average RMSE versus projection hours. The right column is average Peak Error versus projection hours. The top row is for the East Coast region, the middle is for the Gulf of Mexico region and the bottom is for the Alaska region.

Table 3 Reductions to ETSS RMSE (ETSS 2.2– ETSS 2.3) in feet

Storm	12-hr	24-hr	36-hr	48-hr	60-hr	72-hr	84-hr	96-hr
1-Irma-17	0.07	0.23	0.26	0.21	0.03	*	*	*
2-Jose-17	0.03	0.03	0.01	0.01	-0.03	-0.04	-0.05	-0.04
3-Maria-17	0.00	0.01	0.02	0.02	0.01	0.01	0.01	0.00
4-NYC-17	0.07	0.06	0.03	-0.02	0.01	0.01	0.00	*
5-AK-17	0.06	0.06	0.03	-0.01	0.00	0.01	0.00	-0.01
6-EST-18	0.06	0.02	-0.02	-0.04	-0.01	0.01	0.01	0.04
7-EST-18	0.11	0.10	0.09	0.08	0.07	*	*	*
8-AK-18	0.06	0.06	0.06	0.05	0.06	*	*	*
9-AK-18	0.04	0.04	0.03	0.02	0.00	0.00	-0.03	-0.04
10-EST-18	0.14	0.12	0.10	0.09	0.08	0.08	0.07	0.05
11-AK-18	0.04	0.02	0.04	0.03	0.02	0.02	0.05	0.05
12-EST-18	0.03	0.05	0.03	0.04	0.03	0.01	0.00	*
13-EST-18	0.03	0.04	0.08	0.09	0.10	0.08	0.05	0.04
14-EST-18	0.11	0.10	0.11	0.13	0.13	0.12	0.12	*
15-GOM-18	0.10	0.11	0.12	0.10	0.11	0.12	*	*
16-AK-18	0.05	0.05	0.05	0.05	0.03	0.05	0.04	0.05
17-Aberto-18	0.20	0.23	0.27	0.21	0.14	0.15	0.06	0.03

***** - Insufficient model results due to missing GFS wind data.

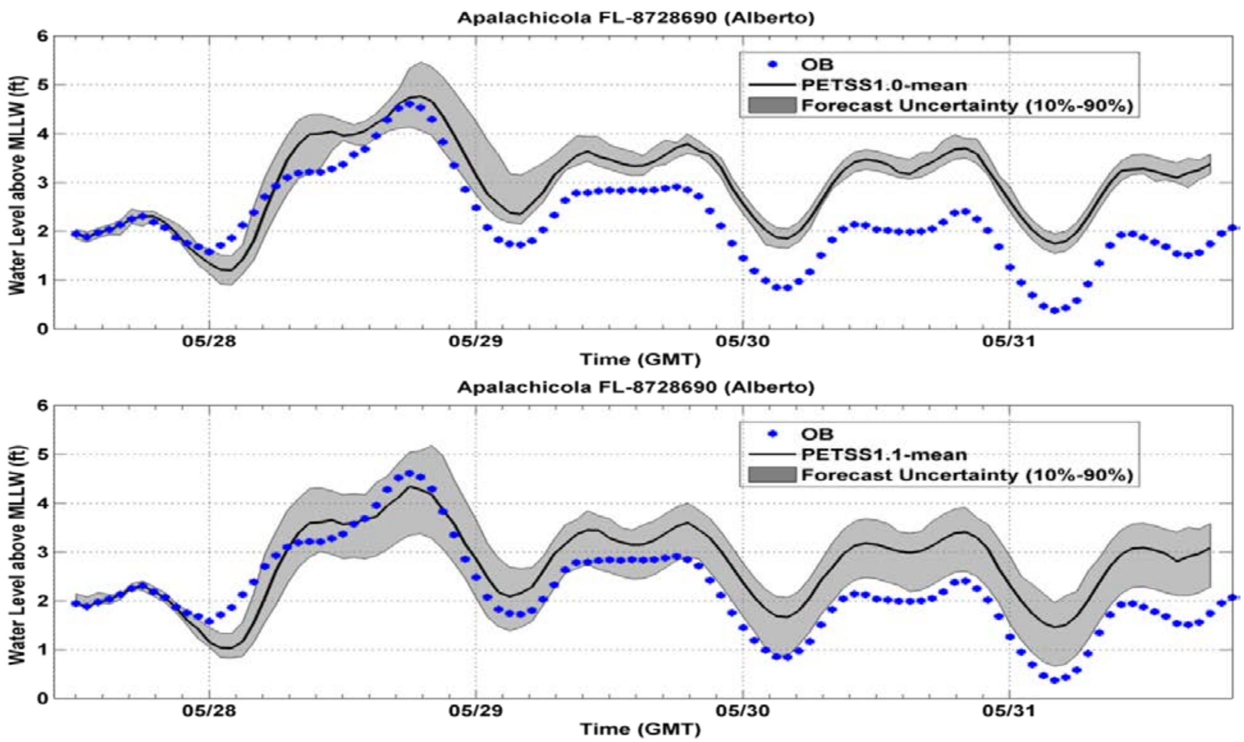

Fig. 3. Model results on May 27th 2018 at 12Z from P-ETSS1.0 (top) and P-ETSS1.1 (bottom).

4.2 P-ETSS RESULTS

To start, we consider a hydrograph of P-ETSS 1.0 and 1.1 model guidance created at 12Z on May 27th, 2018 at the Apalachicola, FL tide gauge for Hurricane Alberto (Fig. 3). Ideally for ensemble model results, the observations will either match the ensemble mean, or fall within the uncertainty range. While the ensemble mean for P-ETSS 1.1 shows some slight improvement over P-ETSS 1.0 in the first 36 hours, after that both are under forecast and it is hard to tell which performs better. However, the observation falls inside the model uncertainty area in P-ETSS 1.1 much more than in P-ETSS 1.0, so P-ETSS 1.1 is better capturing the uncertainty of the forecast. To assess how well it captures the uncertainty of the forecast, we add a new skill score:

3) Percentage of Observations (POU) that fall inside the area of forecast uncertainty,

$$POU = \frac{n}{96} * 100$$

Where n is the total number of hourly observation that fall inside the area of P-ETSS model uncertainty during the 96-hr time frame.

We evaluated the performance of P-ETSS 1.1 for all 9 storm events. To do so we calculated a similar RMSE and PAE as we did for ETSS, except using the ensemble mean, and we'll calculate the POU. The average RMSE and PAE over the observation time frames, in the 3 regions, for the 12, 24, 36, 48, 60, 72, 84 and 96 hour projections are shown in Fig. 4. Fig. 5 shows the POU in the 3 regions for different projection hours. The RMSE and PAE shows that the overall performance of the P-ETSS 1.1 ensemble mean is comparable to the P-ETSS 1.0 ensemble mean in all three regions. However, the POU shows significant improvements in P-ETSS1.1 for all 3 regions.

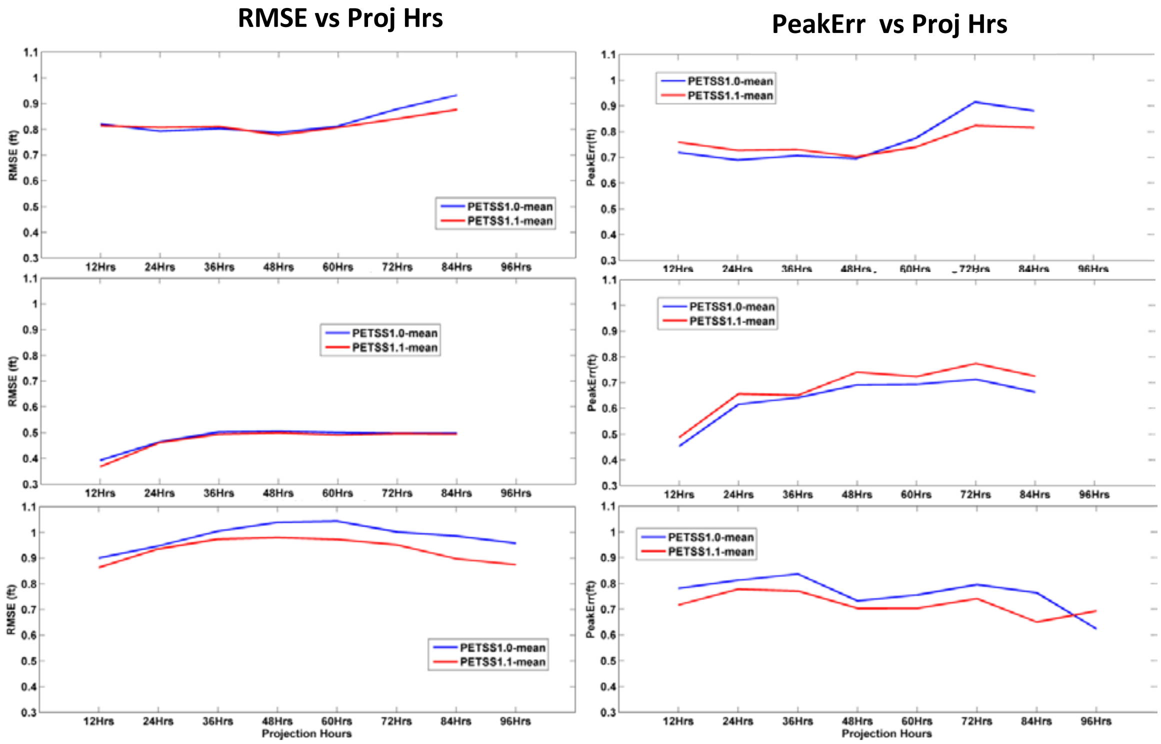

Fig. 4 Same as Fig. 2 but for P-ETSS1.0 and P-ETSS1.1

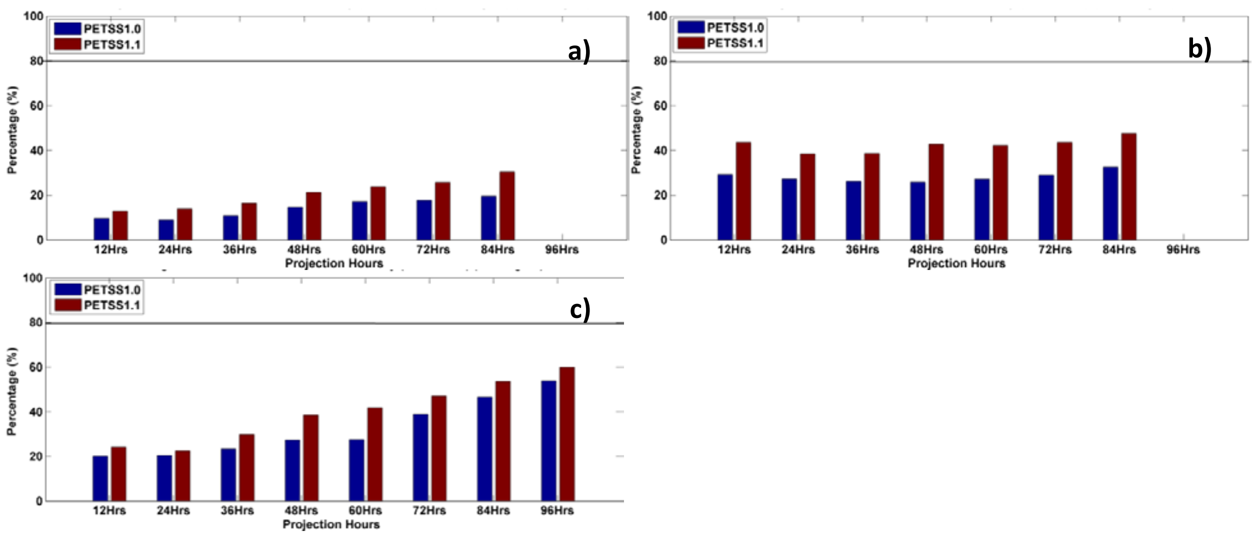

Fig. 5. POU for PETSS 1.0 and PETSS 1.1 per projection hour for a) the East Coast; b) the Gulf of Mexico and c) Alaska. The 96 hour results are omitted for (a) and (b) due to gaps in the NAEFS archive.

The storm specific improvements of the POU score per projection for the 9 storm events are listed in Table 4. It displays the POU score for P-ETSS 1.1 and then, in parentheses, the improvement over P-ETSS 1.0. The POU of P-ETSS 1.1 improved significantly for almost all projections for the 9 storm events. However, the POU value is still far from our goal of 80%, which means there are still many times that the observations are not within the model area of uncertainty. This implies that we need to add more ensemble members to P-ETSS to better capture the uncertainty.

Table 4 Percentage of observations that fall between 10 and 90% exceedance (uncertainty areas). A (B), where A is the P-ETSS1.1 score B is the improvement over Version 1.0 (e.g. Version 1.1 – Version 1.0)

Storm	12-hr	24-hr	36-hr	48-hr	60-hr	72-hr	84-hr	96-hr
1-NYC-17	12 (+1)	12 (+1)	14 (+7)	19 (+7)	22 (+5)	21 (+6)	27 (+8)	*
2-AK-17	16 (+5)	15 (+5)	23 (+10)	29 (+12)	34 (+17)	40 (+10)	47 (+3)	57 (+5)
3-AK-17	35 (+5)	32 (+2)	33 (+5)	41 (+9)	40 (+14)	54 (+5)	69 (+16)	80 (+18)
4-EST-18	19 (+7)	20 (+8)	23 (+7)	29 (+12)	30 (+10)	34 (+13)	36 (+14)	*
5-EST-18	13 (+1)	16 (+8)	19 (+7)	26 (+5)	32 (+5)	36 (+10)	47 (+17)	*
6-EST-18	4 (+1)	5 (+1)	5 (+1)	6 (+1)	6 (+0)	5 (+1)	5 (+1)	*
7-GOM-18	29 (+18)	31 (+20)	30 (+20)	30 (+20)	30 (+14)	30 (+10)	35 (+10)	*
8-AK-18	22 (+3)	20 (+0)	34 (+4)	46 (+13)	52 (+13)	47 (+10)	46 (+3)	43 (-4)
9-Alberto-18	89 (+21)	73 (+10)	74 (+14)	82 (+28)	83 (+29)	82 (+30)	86 (+31)	*

***** - Insufficient model results due to missing GEFS wind data.

5. POST PROCESSING ADJUSTMENT

As mentioned previously, both ETSS and P-ETSS models omit wave setup and run-up, flooding from rain, and sea level rise. Additionally they include errors from both the wind models (GFS, GEFS) and storm surge model. To address these omissions and errors, the ETSS model incorporated a simple and efficient post-processing methodology at stations (Schuster and Taylor 2015). A similar technique is now used within P-ETSS to account for the same thing. Specifically the P-ETSS post-processing uses observations, tides, and the P-ETSS ensemble mean storm surge guidance to calculate the average anomaly over the past 5 days. It then adds, for the first 12 hours, a linearly interpolated anomaly (from the instantaneous anomaly value to the 5-day average anomaly value) to the ensemble mean, 10% exceedance and 90% exceedance water levels at a particular station. After 12 hours, it adds the 5-day average anomaly to those same products.

The impact of adding station-based post-processing to P-ETSS model results are demonstrated in detail in Liu and Taylor, 2018, so we skip that here and simply incorporate that post-processing into our model package which will be run on the WCOSS production machine.

6. SUMMARY AND DISCUSSION

In all of the storm events, both ETSS 2.2 and P-ETSS 1.1 provided better guidance for most of the 12 to 96 hour forecast projections. Furthermore, ETSS 2.2 performs better in the East Coast and the Gulf of Mexico regions than in the Alaska region indicating a need for nesting smaller high resolution basins within the larger coarser Alaska basin. Though the P-ETSS 1.1 ensemble mean does not show significant improvements over the current version, the POU for P-ETSS 1.1 does. However, the specific POU score per projection for the 9 storm events indicates we're still far from our goal of 80%. More members need to be added to P-ETSS to expand its estimate of uncertainty, which will not only improve the POU, but also improve the performance of the ensemble mean.

In addition to the model improvements, this implementation includes station-based post-processing to the P-ETSS model, which provides an efficient way to account for omitted physical terms and model errors. Additionally, by migrating the image production to the WCOSS production machine, MDL is able to provide a more stable web service for displaying ETSS and P-ETSS model guidance to local forecast offices.

However, there are a number of actions that can be taken to improve ETSS and P-ETSS. The station-based post-processing adjustment is an efficient way to account for various biases; however since it is done after the inundation calculations, it only improves water level guidance at stations and doesn't help the inundation guidance. Therefore, in the short term, MDL plans to add an initial water level estimate to both ETSS and P-ETSS before the models run. Doing this will improve the inundation calculation as it will be done within the actual model run. We'll still do post-processing at stations to adjust to local water level observations, but the adjustment will be smaller. Additionally, MDL plans to develop a method to smooth out the P-ETSS cycle variations caused by the lack of NAEFS at 6 and 18Z. MDL also plans to upgrade P-ETSS by using the new East Coast and Gulf of Mexico basins that have been added to ETSS. Finally we plan to upgrade ETSS to use the new south Florida basin.

In the longer term, MDL plans to incorporate a fast wave model when one becomes available, and improve P-ETSS by incorporating rainfall model output along the river boundary. Finally, we would like to add other fast storm surge models into the P-ETSS scheme to create multi-storm surge model ensemble products.

7. ACKNOWLEDGEMENTS

The authors would like to thank Yuejian Zhu, Bo Cui and Richard Wobus from the GEFS modeling group for support in preparing the historical wind data.

8. REFERENCES

Kim, S. C., J. Chen, and W. A. Shaffer, 1996: An Operational Forecast Model for Extratropical Storm Surges along the U.S. East Coast. *Preprints, Conference on Coastal Oceanic and Atmospheric Prediction, 76th Amer. Meteor. Soc. Annual Meeting*, Atlanta, GA, 281-286.

Liu, H., A. Taylor and R. Schuster, 2015: Creating Inundation Guidance from NWS's Extra-Tropical Storm Surge Model, *Preprints, 13th Symposium on the Coastal Environment, 95th Amer. Meteor. Soc. Annual Meeting*, Phoenix, AZ, 3.2, 7 pp.

Liu, H., A. Taylor and C. Belton, 2016: Creating Storm Tide Inundation Guidance and Bridging the Bering Strait with NWS' ETSS Model, *Preprints, 14th Symposium on the Coastal Environment, 96th Amer. Meteor. Soc. Annual Meeting*, New Orleans, LA, 4.4, 8 pp.

Liu, H. and A. Taylor, 2018: Development of the NWS' Probabilistic Extra Tropical Storm Surge Model and Post-Processing Methodology. *Preprints, 16th Symposium on the Coastal Environment, 98th Amer. Metor. Soc. Annual Meeting*, Austin, TX, 1.2, 8 pp.

Taylor, A., and Glahn, B., 2008: Probabilistic Guidance for Hurricane Storm Surge. *Preprints, 19th Conference on Probability and Statistics, 88th Amer. Metor. Soc. Annual Meeting,* New Orleans, LA, 7.4, 8 pp.

Taylor, A., H. Liu, and R. Schuster, 2015: Finer Wind Field Resolution for NWS's Extra-Tropical Storm Surge Model. *Preprints, 13th Symposium on the Coastal Environment, 95th Amer. Meteor. Soc. Annual Meeting*, Phoenix, AZ, 3.1, 6 pp.