THE STEVENS FLOOD ADVISORY SYSTEM: OPERATIONAL H^{3}E FLOOD FORECASTS FOR THE GREATER NEW YORK / NEW JERSEY METROPOLITAN REGION

N. GEORGAS, A. BLUMBERG, T. HERRINGTON, T. WAKEMAN, F. SALEH, D. RUNNELS, A. JORDI, K. YING, L. YIN, V. RAMASWAMY, A. YAKUBOVSKIY, O. LOPEZ, J. MCNALLY, J. SCHULTE & Y. WANG
Davidson Laboratory, Stevens Institute of Technology, USA.

ABSTRACT

This paper presents the automation, website interface, and verification of the Stevens Flood Advisory System (SFAS, http://stevens.edu/SFAS). The fully-automated, ensemble-based flood advisory system dynamically integrates real-time observations and river and coastal flood models forced by an ensemble of meteorological models at various scales to produce and serve street scale flood forecasts over urban terrain. SFAS is applied to the Greater NY/NJ Metropolitan region, and is used routinely by multiple forecast offices and departments within the US National Weather Service (NWS), regional and municipal Offices of Emergency Management, as well as the general public. Every six hours, the underlying H^{3}E (Hydrologic–Hydraulic–Hydrodynamic Ensemble) modelling framework, prepares, runs, data-assimilates, and integrates results from 375 dynamic model simulations to produce actionable, probabilistic ensemble forecasts of upland and coastal (storm surge) flooding conditions with an 81-h forecast horizon. Meteorological forcing to the H^{3}E models is provided by 125 weather model ensemble members as well as deterministic weather models from major weather agencies (NCEP, ECMWF, CMC) and academia. The state-of-the-art SFAS, a replacement of the well-known, but deterministic, Storm Surge Warning System (SSWS) that was highlighted during Hurricanes Irene and Sandy and more recently extratropical cyclone Jonas, has been operational since the end of 2015.

Keywords: early warning systems, emergency management, ensemble forecast, flood advisory system, integrated flood forecasts, river flooding, storm surge.

1 INTRODUCTION

In the United States of America alone, annualized inland hydrologic flood losses over the 30-year period ending at the end of water year (WY) 2014, stood at $7.96 billion in damages, and 82 fatalities, per year [1]. These hydrologic losses are adjusted for inflation but do not include losses from coastal storm tide flooding due to Extratropical and/or Tropical Cyclones (ETCs and TCs, respectively). Storm tide flooding can be defined as coastal flooding caused by a storm's surge pushing ocean waters to rise above local astronomical tide levels, with the phase of the local astronomical tide having important effects on total water levels above ground [2].

New York Harbor (NYH) and its connected tidal waterways lie at the apex of the New York Bight, and are surrounded by over 21 million people in cities such as New York, Newark, Jersey City and Hoboken. The metropolitan region with the highest risk for storm tide flood-

WIT CONFERENCES
This paper is part of the Proceedings of the 5th International Conference on Flood Risk Management and Response (FRIAR 2016)
www.witconferences.com

ISSN: 2041-9031 (paper format), ISSN: 2041-904X (online), http://www.witpress.com/journals
DOI: 10.2495/SAFE-V6-N3-648-662

ing is the New York City Core Based Statistical Area (CSA), which includes New York City (NYC), Long Island and a portion of New Jersey, with a total 685,152 homes at risk and a reconstruction-cost-value of over $244 Billion [3]. As with Hurricane Sandy in 2012, TCs can cause significant coastal storm tide flooding in the Eastern United States and the New York / New Jersey Metropolitan Region. In New York counties alone, aggregated storm tide losses from Hurricane Sandy were estimated to total $23B of over $50B country-wide [4, 5]. New Jersey was the second coastal state worst hit. Historically, many of the historical records for storm tide flooding in the region have been caused by cold-season ETCs, mid-latitude cyclones locally called Nor'Easters [6]. On January 23 2016 for example, Nor'Easter Jonas caused severe tidal flooding in southern New Jersey: Jonas' storm tide broke the local records previously set by Sandy or the historic December 1992 Nor'Easter, while dumping near 30 inches of snow in the northern part of the State and New York City; an area that did not experience major coastal flooding during the event [7].

Importantly, TCs are not only a coastal flood hazard but an important inland flooding agent in the Eastern United States [8]. As recently as between October 1–5 2015, locally extreme rainfall totals exceeded 20-inches in several eastern US states resulting from the convergence of a powerful low pressure system / frontal boundary and copious moisture from Hurricane Joaquin in the Atlantic [9]. A year before Hurricane Sandy, Hurricane Irene and Tropical Storm Lee in the summer of 2011 brought severe inland flooding into the NY State region and broke records for both storm surge and river stage flooding [10, 11].

To characterize deaths related to Sandy, the US Centers for Disease Control and Prevention (CDC) analyzed data on 117 hurricane-related deaths captured by American Red Cross mortality tracking during October 28–November 30, 2012. 53 of these deaths occurred in New York and 34 in New Jersey, with drowning being the most common cause of death related to Sandy, and 45% of drowning deaths occurring in flooded homes in NYC's Evacuation Zone A. The CDC report states: 'Drowning is a leading cause of hurricane death but is preventable with advance warning systems and evacuation plans' [12].

Seventy-eight percent of emergencies are weather related [13, 14], with flooding now being the second leading cause of death behind heat-related incidents according to the CDC. Driving a car into flood waters is the leading cause of flood-related deaths, making the 'Turn Around, Don't Drown' message of the US National Weather Service (NWS) one of the most ubiquitous and recognizable signs in the US. NWS is working hard to implement its Weather-Ready Nation plan with a goal to minimize preventable losses in life and property from flooding. To effectively minimize flood risks, flood forecasting and warning systems must provide precise flood information to emergency managers and decision-makers. Producing accurate flood forecasts requires numerical models that properly take into account the important physical processes that affect a storm tide or river stage and consequently the flooded areas [11, 15–17]. An accessible interface that is quick to understand and easy to use in decision making is then needed to effectively communicate these forecasts to the public.

Stevens Institute of Technology has built and operates an automated Storm Surge Warning System (SSWS) for New Jersey and New York since 2002 [18]. The system extended its capacity over the years with support from the New York Harbor Observing and Prediction System (NYHOPS) and its models [19, 20] and was highlighted during Hurricane Sandy as being preferred by the NYC Office of Emergency Management for its ease of use [2, 21]. With significant funding through the Port Authority of New York and New Jersey (the PANYNJ), the system has recently been updated and renamed the Stevens Flood Advisory System (SFAS, http://stevens.edu/SFAS). In this paper, the design of the latest generation

SFAS and its website interface is presented, followed by a discussion on the system's verification and improvements upon other current SSWSs in the US, and closing with a discussion of ongoing research and development.

2 DESIGN OF THE STEVENS FLOOD ADVISORY SYSTEM

The system diagram shown in Fig. 1 abstracts the components of SFAS. Real-time water level observations (coastal storm tide and river stage) and an ensemble of water level forecasts produced by operational computer models at stations in and around the NY/NJ Greater Metropolitan area are acquired and stored in relational databases, registered to equivalent vertical datums, and then compared against critical flood levels to interpret whether a station is presently flooding or whether it is expected to flood in the future over the next 81 h. It also indicates how severe the flooding may be. The data are then visualized on the SFAS website. For tidal stations, storm surge is also computed as the deviation between the observed or forecast water level from the astronomical tide prediction for each station.

To provide common context for the severity of occurring or anticipated flooding, we use the high water level terminology of the NWS for the critical flood levels [22]: 'Near' flooding (equivalent to 'Action stage' for inland river stations) and 'Minor', 'Moderate', and 'Major' flood levels. These levels are set and provided through regional NWS Weather Forecast Offices (PHI, OKX, ALY, and GYX) for accuracy and consistency [23]. Station locations displayed on the main SFAS map (Fig. 2) start blinking at Near (Action) flood stage. Minor Flooding represents minimal or no property damage but possibly some public threat (e.g. inundation of roads). Moderate Flooding signifies some inundation of structures and roads near the coastline or stream and possible evacuations of people and/or transfer of property to higher elevations. Major Flooding indicates extensive inundation of structures and roads and significant evacuations of people and/or transfer of property to higher elevations. An SFAS user has the ability to register for automated flood alerts for the stations he/she is most interested in by providing his e-mail and information on the website; an automated e-mail will be sent when water level at a registered station is forecast by the modeling ensemble to rise above Minor Flood stage.

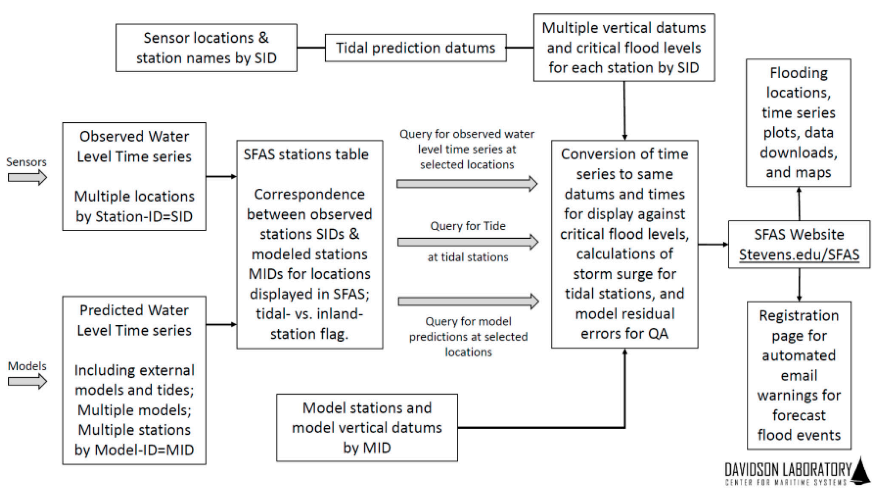

Figure 1: Stevens flood advisory system diagram.

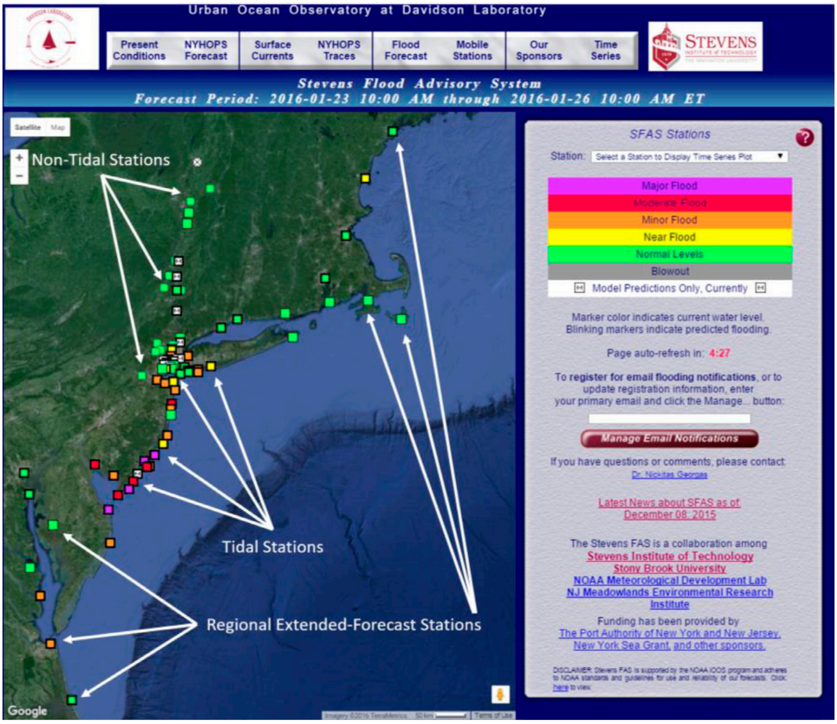

Figure 2: Screenshot taken from the opening page of the SFAS website (http://stevens.edu/SFAS), at 10 am EST (UTC-5) January 23 2016, during Nor'Easter Jonas, showing Major (purple) and Moderate (red) flooding occurring in the Delaware Bay and southern New Jersey coastal regions, with Minor flooding (orange) occurring in New York Harbor and western Long Island. All stations with forecast water levels above Near-flood stage are blinking and outlined with thicker black lines. Arrows point to examples of inland (non-tidal) flood station locations (with circles), and tidal station locations (squares), either within the NY/NJ greater metropolitan area that the NYHOPS model covers, or regional extended-forecast stations forecast by SNAP.

2.1 SFAS Website

Figure 2 shows the opening page of the SFAS website (http://stevens.edu/SFAS). This page is also accessible from the NYHOPS website (http://stevens.edu/NYHOPS) by clicking on the top banner's 'Flood Forecast' button. White arrows are included on Fig. 2 to highlight different types of stations shown on the map: inland stations (rivers, using circles), and tidal stations (ocean coast, estuarine, or tidal rivers, using squares). The latter station are further broken down to indicate those stations for which flood advisories are based on the NYHOPS model for the NY/NJ Greater Metropolitan area or by the SNAP extended regional model (see section 2.2). The user can see if a station is presently experiencing flooding by comparing its marker colour to the colour key found on the right hand side of the webpage. Note that the grey colour is used if a 'blowout tide' condition is presently occurring, a blowout tide being a 'negative storm surge' condition where the waters have dropped too low and safe vessel passage through navigable waters is a concern.

If a user would like to receive email notifications of predicted flooding, he/she can enter their primary email address in the control panel to the right, and click the 'Manage Email Notifications' button. On the page that appears, users may select the stations for which they would like to receive notifications, optionally, provide an alternate email, or deregister previously registered stations.

If a user clicks on a location marker, an 'info window' will appear collectively noting the current observed water level, the most severe forecast level in the forecast period, and the date and time of the expected event. By default, an ensemble average of forecast predictions created at the Davidson Laboratory of Stevens Institute of Technology is used in SFAS. To display a time series plot (Fig. 3), which includes 81 or more hours of forecast water levels, and observed data as available, users should click on the 'Show Time Series Plot' in the info window or use the Station drop down in the control panel to select a station.

On the top-left panel of the SFAS-station time series webpage (Fig. 3), the user can shift between stations through the 'Station' drop-down menu. The best Stevens ensemble forecast for the selected location will be displayed by default. Where available, a user may select to display a different forecast time series from a number of forecast models described in the following section by selecting them from the 'Forecast:' drop down box, which includes three models from the NWS (Fig. 3). The numerical values for the selected forecast time series can be downloaded by clicking on the 'Extended download' button at the bottom of the control box. For consistency, time series of observed (where available) and forecast water levels are

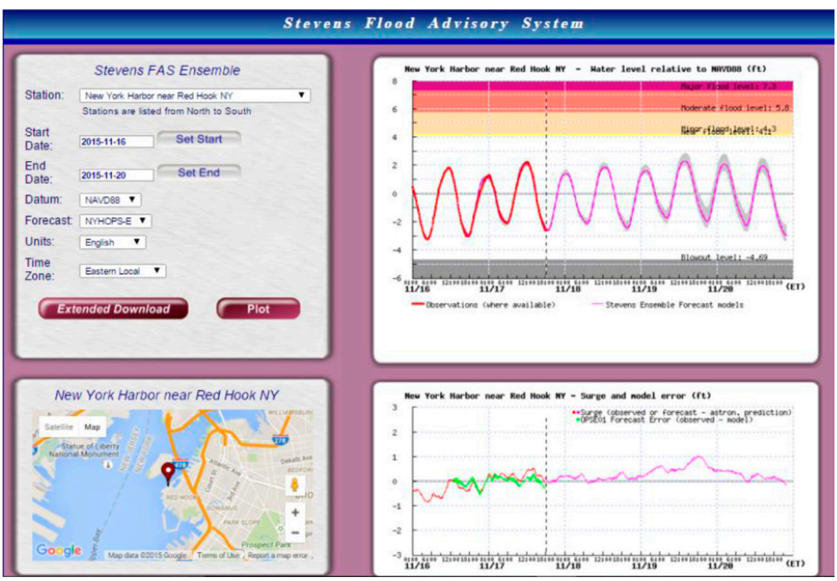

Figure 3: SFAS-station, water level time series page. The selected model prediction (in this case, the NYHOPS Ensemble Forecast, NYHOPS-E; pink line) appears on the top right panel, for a new tidal station (New York Harbor near Red Hook, NY), including the 5th to 95th percentile forecast uncertainty (grey area), along with past and real-time observations (red colour).

shown and can be downloaded referenced to a 0 level being the North American Vertical Datum of 1988 (NAVD88). The user can, however, select among multiple vertical Datums, Units, etc., depending on familiarity or need, as well as look or download past data and predictions using the control box options shown on Fig. 3. The map on the bottom left of the webpage shows the selected station location in anticipation of predicted overland flood depth maps that will be placed there in the future, while the bottom right plot shows the model's deviation from observed data (green line) together with the observed (red) or model-forecast (pink) storm surge above astronomically-predicted tide levels (for tidal stations only).

2.2 Supporting infrastructure: Observations and Models

The Davidson Laboratory has been installing 11 dual (for redundancy) environmental monitoring stations (Aquatrack 5002 & Greenspan Analytical EC250 C/T sensors and power cells) near critical infrastructure sites, such as airports and marine terminals, through a collaboration with the PANYNJ. This environmental data and collected water levels are sent to the Stevens Institute databases in real-time through cellular connections from the station data loggers every 6 minutes. Real-time data from partner institutions and agencies (see Acknowledgements section) are retrieved through the internet. Such regional collaborations have helped greatly in expanding the scope of the system.

Regarding water level forecasts at SFAS stations, most water level forecast data are generated at the Davidson Laboratory. The overall research objective is to predict inundation occurring from a storm surge|rainfall event at the street level scale. The Storm Surge flood forecasts, and automated e-mail flood advisories that previous SSWS users were familiar with are now of improved quality because they are ensemble-model based. This is accomplished by running many numerical hydrodynamic and hydrologic models on the in-house Pharos Hyperscale Computing Facility, a 1,320-core Hewlett-Packard supercomputer built on sheltered high-ground based on HP Proliant microserver architecture, Mellanox FDR InfiniBand network and a 2.2 PB Seagate online storage solution, dual head nodes, Uninterrupted Power Supplies, Cooling Towers, and a Caterpillar generator for storm resiliency.

Figure 4 shows the domains and scope of the presently operational Davidson Laboratory models. Three sets of 125 linked coastal and inland flood models with a forecast horizon of at least 81 h, are reinitialized every 6 h and are based on different atmospheric model predictions of surface meteorological factors, such as near-surface winds, barometric pressure, and rainfall. Meteorological forcing for the 125 members of the 'Stevens Ensemble' is provided through output fields from the ECMWF ensemble and ECMWF High-Resolution [24], SREF Ensemble [25], GEFS Ensemble [26], GFS, including its experimental version GFS-e [27], NAM [28], Rutgers-WRF [29], and the Canadian Meteorological Centre's Ensemble (CMC) [30] models (Fig. 4).

The computational grid of the three-dimensional NYHOPS built using the sECOM hydrodynamic model, consists of 147×452 nodes and its horizontal resolution ranges from approximately 7.5 km at the open ocean boundary to less than 50 m in NY/NJ Harbor [19, 20]. The grid has 10 vertical layers that stretch from the water surface to the local bottom relief. The NYHOPS grid is nested in (derives open ocean boundary conditions from) the regional-scale Stevens Northwest Atlantic Prediction (SNAP) model domain, which is also based on the sECOM code but using a 5-km constant resolution grid [31]. River stages at SFAS stations are forecast by the Stevens HYDRO model, based on the HEC-HMS code [16]. River discharges from HYDRO are coupled into the NYHOPS model. This operational

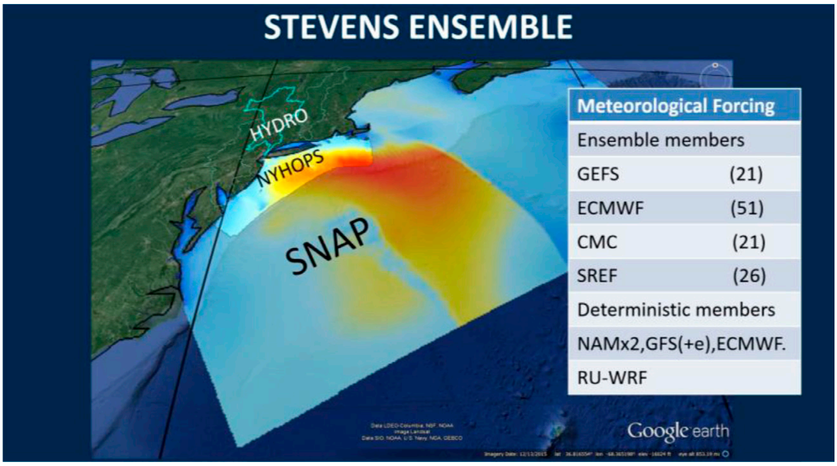

Figure 4: sECOM-NYHOPS three dimensional 125-member hydrodynamic ensemble linked to offshore Stevens Western North Atlantic Prediction (sECOM-SNAP) model and inland Stevens Hydrologic (HMS-HYDRO) ensembles.

framework offers an improvement over AHPS in this particular region [16, 32]. The AHPS streamflow forecasts are at 6-h time intervals using one weather control member for input and with a lead time that is less than 60 h in this region [33]. A new ensemble AHPS is being evaluated presently by the NWS.

The NYHOPS model ensemble members span 81 h from present. Every 6 h, each of 125 NYHOPS ocean model forecasts is run based on a different meteorological forecast. These 125 different water level predictions provide uncertainty estimates and are processed to create the consensus weighted ensemble average prediction that the automated e-mail advisories are based on, called 'NYHOPS-E.' Ensemble-averaged water level predictions are displayed in the SFAS time series and are enveloped by the uncertainty around that ensemble (5th to 95th percentile) prediction to better depict a range of possible futures (Fig. 3). Selecting the most probable of these possible futures, the one that is most likely to be experienced over the next 3–4 days, is the focus of ongoing research in ensemble forecasting.

An extended storm surge forecast ensemble (SNAP-Ex) that spans at least 105 h from present is also available at some regional stations based on 76 of 125 total members of the SNAP regional ensemble that predict further than 81 h out (51 ECMWF, ECMWF-HR, 21 GEFS, GFS, GEFS-e, and RU-WRF). 17 new regional stations, from Maine to North Carolina, including cities like Portland, Boston, Baltimore, and Annapolis, were added to SFAS, with predictions based on that extended regional ensemble.

3 COMPARISONS OF ENSEMBLE WATER LEVEL PREDICTIONS TO EXISTING DETERMINISTIC SYSTEMS

The NWS has developed two deterministic models for ETCs [34]: The Extratropical Storm Surge model (ETSS), based on the SLOSH numerical model, and the Extra Tropical Storm and Tide Operational Forecast System (ESTOFS) based on the 2D ADCIRC model. For Extra-Tropical storms, both models use a nominal 2.5 km, are forced by GFS winds, and are initialized every six hours.

Gridded water level results from these models at specific stations are debiased against observations by an anomaly-correction procedure used by NOAA's Meteorological Development Lab (MDL), and subsequently posted on MDL's Extra-Tropical Surge Guidance website. From there, debiased water levels from the two models at SFAS stations are downloaded each hour, stored at the Stevens databases along with other models, and displayed in SFAS as NOAA guidance together with the Stevens NYHOPS-E, SNAP, and SNAP-Ex ensembles. Di Liberto *et al.* [15] compared NYHOPS water level predictions to ETSS for 75 available days during the 2007-08 and 2008-09 cold season to assess the benefit of creating a multi-model storm surge ensemble using NYHOPS, ETSS, and a Stony Brook University ADCIRC-based ensemble. Forecasters from the NWS OFSs who use SFAS routinely and regularly note on the benefit of having these different guidance sets for the same stations and valid times displayed on one website, SFAS. They routinely do their own multi-model ensemble averaging to devise their coastal flooding forecasts. In the process of verifying that ensemble, Di Liberto *et al.* [15] found that ETSS was significantly biased, but after an imposed debiasing produced quite accurate results; yet, it was still somewhat inferior to a deterministic NYHOPS model based on NAM winds.

Here, as part of the new ensemble verification, we expanded and updated that work by comparing the accuracy of the water level forecasts that debiased ETSS and now ESTOFS provided for 130 days mostly in the cool season of 2015-2016 to that of the 3 new three-dimensional model ensembles (the NYHOPS-E, SNAP, and SNAP-Ex ensembles created at the Davidson Laboratory based on the sECOM model). The method used was the following:

a. Extract the predictions at several stations from the five systems collected hourly at Stevens Institute databases, four times a day, at 0 z, 6 z, 12 z, and 18 z each day beginning on December 30 2015 at 0 z, real time. The process was automated and continued henceforth each day. The dataset considered here ends on April 23 2016 at 18 z, real time, some 135 days later. This real-time operation disallows any posterior data or forecast corrections, but ensures that the downloaded model predictions are truly as they were stored and displayed at SFAS at the time of the scheduled downloads, across all 5 models considered.
b. The forecasts from each model were then binned in six hourly increments as the 0–6 h, 6–12 h, 12–18 h forecast-bin for each model, etc., to the end of each model's forecast horizon.
c. The data were used to compare the models performance at each station, primarily using the same metrics used at [15]: The Mean Error (ME), and the Root-Mean-Square-Error (RMSE). For ease of presentation, model results on Fig. 5 are aggregated over several stations in four station categories, somewhat arbitrarily, and where available from each model, corresponding to SFAS stations on Fig. 2:
    i. Coastal NYHOPS stations: Ocean City Inlet, MD; Lewes, DE; Cape May, NJ; Atlantic City, NJ; Montauk, NY; Newport, RI.
    ii. Estuarine NYHOPS stations: Reedy Point, DE; Sandy Hook, NJ; The Battery, NY; Kings Point, NY; New Haven, CT; Bridgeport, CT; New London, CT.
    iii. SNAP-South: Annapolis, MD; Cambridge, MD; Lewisetta, VA; Baltimore, MD; Chesapeake Bay Bridge Tunnel, VA; Duck, NC.
    iv. SNAP-North: Woods Hole, MA; Nantucket Island, MA; Boston, MA; Fort Point, NH; and Portland, ME.

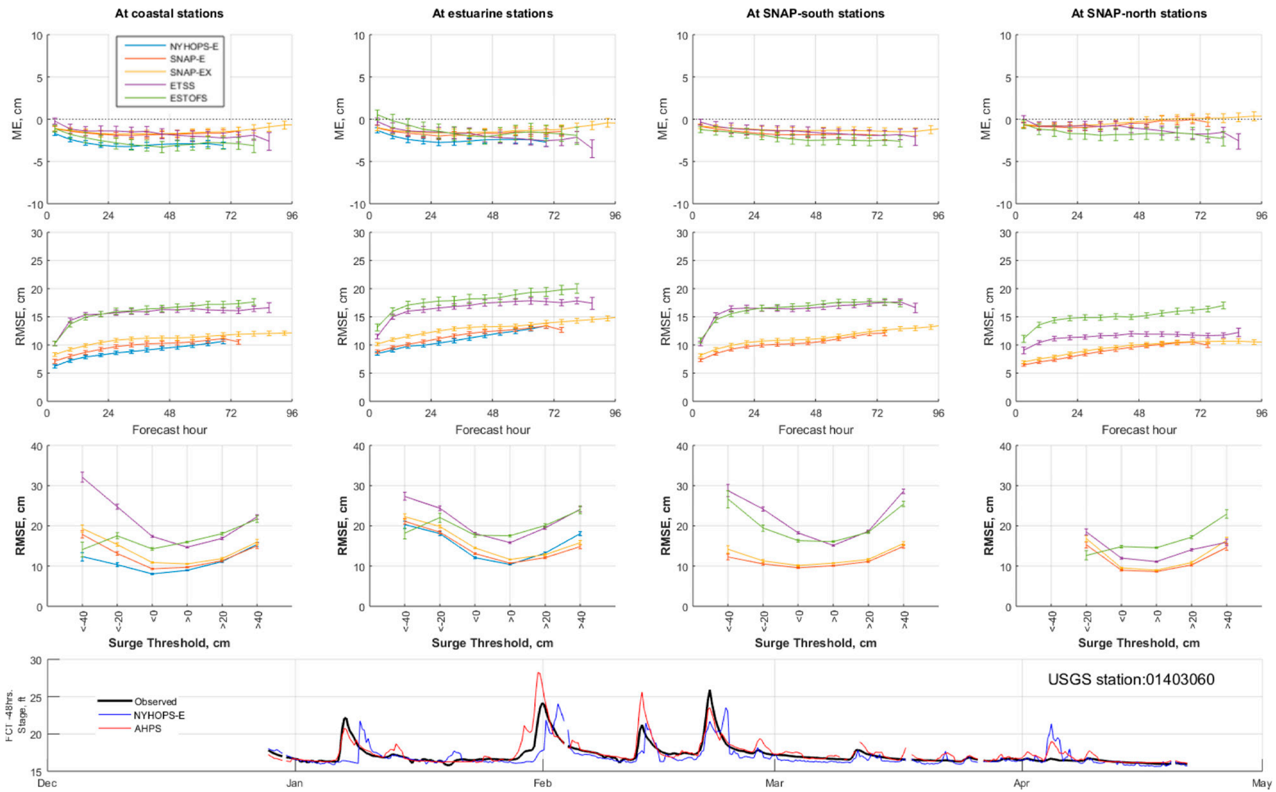

Figure 5: (Previous page). Top three panels: Multi-model comparison for storm surge forecasts against observations across several stations operated by the National Ocean Service (NOS) and included in SFAS. Top row: ME by forecast lead hour. Second row: RMSE by forecast lead hour. Third row: RMSE by surge threshold, similar to [15], across all 0–72 h forecasts. Bottom row: Comparison between the derived HYDRO hydrologic ensemble stage ('NYHOPS-E') time series and the AHPS deterministic model stage against observations collected by USGS at an inland river station: 01403060, Raritan River below Calco Dam at Bound Brook NJ. Stages were derived from 6 h-averaged NYHOPS-E discharge data, through the same rating curves, to match AHPS's six hourly outputs. Only the 48 h forecast lead time time-series is shown.

d. 90% confidence intervals around the calculated means were generated using Matlab's boot-strapping by random replacement that was repeated 1000 times for each interval shown. If the confidence intervals of two particular samples did not overlap, they were considered different at the 90% level.
e. River stages between AHPS and HYDRO were treated similarly, except that the latest rating curve obtained from USGS per station was used to convert river discharge, which is what the hydrologic models predict, to river stage, and is what is used in SFAS for comparison to flood stages. Also, the AHPS forecast horizon was short, near 60 h.

The top three panels of Fig. 5 clearly show that all five models produce reasonable results, with the NYHOPS-E storm surge predictions continuing to have overall the least amount of error residuals at all stations considered. ME's are small, on the order of 2–3 cm for all models and regions, a significant improvement for the NOAA ETSS model compared to 5 years ago. The NYHOPS-E across-station forecast-wide RMSE is ~ 10.5 cm during this 135-day period that included three Nor'Easter storms, including ETC Jonas, which was a record-setting storm for some locations. The NYHOPS-E RMSE is about 50% less than that of the NOAA models (ETSS and ESTOFS). All models have higher errors at estuarine stations compared to coastal stations. Model errors grow with forecast lead time, with ETSS and ESTOFS experiencing a very quick increase in RMSE shortly after forecast start, and then following the gradients of the Stevens models; the net effect being that the Stevens models maintain comparable performance to the NOAA models initial 12 h performance for a much longer lead, up to at least 96 h in the SNAP-Ex case. The outlier is the ETSS model at SNAP-North stations, where that model's error does not appear to grow significantly with lead time. Interestingly, the differences between the higher resolution NYHOPS model performance from the lower resolution SNAP model, both based on the sECOM code, but with SNAP having less comprehensive forcing that NYHOPS, are barely significant at the 90% level. They are about the same magnitude different as SNAP to itself using less Ensemble models (SNAP vs. SNAP-Ex), with the leaner SNAP-Ex ensemble having an increase in error compared to the full 125-member SNAP for similar lead times. As expected, RMSE grows with positive or negative surge threshold on Fig. 5, third row from top. The exception is the constancy or increase of skill for the ESTOFS model at low surges in three of the four regions. Comparing ESTOFS and ETSS alone, no consistently better model emerges for the period considered here.

For river stage, the example in Fig. 5 is largely representative of what happens in other stations too. In the case of inland stations, overall for the period considered, NOAA AHPS did significantly better than the NYHOPS HYDRO ensemble in predicting the observed river stage, by as much as 50% in upper Hudson watersheds (not shown). Resolution and spread of rainfall appear to be key here. In the case of the Stevens inland hydrology ensemble, NYHOPS-E is presently the median of all precipitation-driven hydrologic predictions. Given the high spread in the precipitation data and the way it propagates in the hydrologic model this appears to be a limiting approach; research is ongoing on other ensembling techniques. The AHPS performance in RMSE is of the same order of magnitude as the tidal NOAA models (not shown). The example shown in Fig. 5 is for the 48 h lead forecast time. The error growth with lead time appears to be quasi-linear to lead time, with the two models being roughly equivalent at small lead times, and growing toward the forecast horizon.

4 PERFORMANCE COMPARISONS FOR INDIVIDUAL MEMBERS IN THE NYHOPS ENSEMBLE TO EACH OTHER, AND THE ENSEMBLE MEAN

Figure 6 compares the performance of individual members of the 125-member NYHOPS ensemble to each other and to NYHOPS-E ensemble mean for coastal stations, using the Grand mean RMSE (GM-RMSE) over all runs in the verification period as a performance metric. Figure 6 clearly shows that the NYHOPS-E weighted ensemble used in SFAS has indeed a lower GM-RMSE than any of the 125 NYHOPS members that comprise it. This is both true at The Battery NOS station, but also when all 13 stations are considered (Fig. 6).

As expected, NYHOPS members forced by deterministic members of the same meteorological models show better performance in simulating storm tide than NYHOPS members forced by the lower resolution meteorological ensemble members of the same model (GFS deterministic vs. GEFS ensemble, for instance). Comparing across meteorological ensembles, the

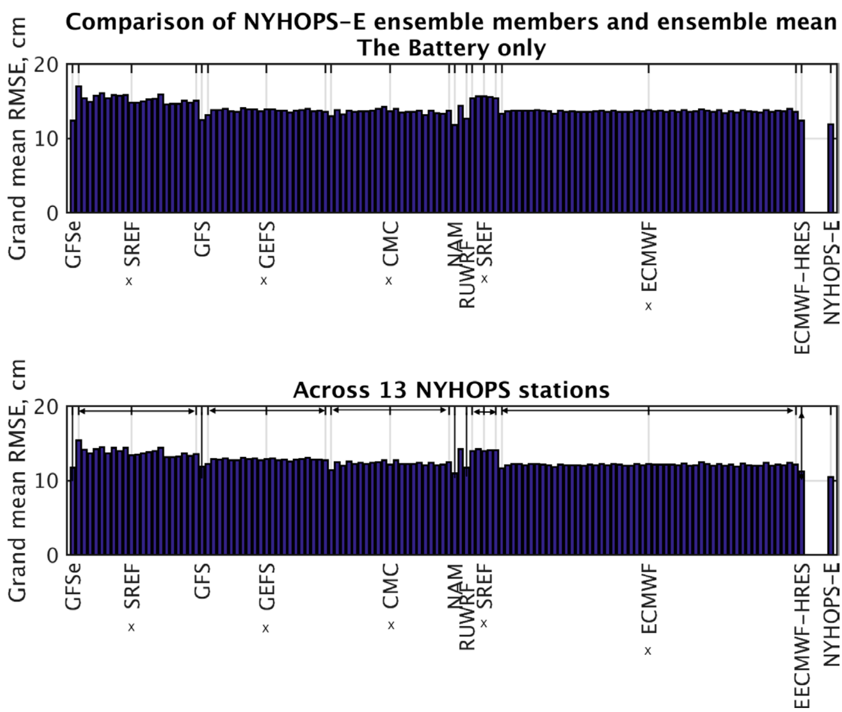

Figure 6: Comparison of 125 NYHOPS Ensemble Members to each other and to the NYHOPS-E weighted ensemble mean (right-most bar): The bar charts depict Grand Mean RMSE across all forecast times, four forecast cycles a day, and 135 days in the verification period at The Battery, NY (top panel), or at all 13 NYHOPS stations considered together (bottom panel). NYHOPS members forced with deterministic high-resolution meteorological models are highlighted with black vertical lines on the bottom panel. Perturbed meteorological ensembles of a single model Horizontal arrows span NYHOPS members forced by a perturbed ensemble of a single meteorological model, highlighted with an x below its name.

SREF-forced ensemble members perform quantitatively worse than average, even though the meteorological fields from SREF have a relatively high 16 km resolution. Also note the big discrepancy in Grand Mean RMSE between individual SREF members. This appears to be a sign of an over-dispersed SREF ensemble; the model is run with two different physical packages and strong negative and positive perturbations. This finding does not appear to necessarily mean that the SREF ensemble is not a good ensemble to use to force an ensemble for storm surge, however. For example, the SNAP-Ex model that does not use SREF, nor the CMC, was slightly inferior to the SNAP ensemble that did (Fig. 5).

On the other end of the spectrum, Fig. 6 shows that the NYHOPS members forced with the ECMWF ensemble have slightly lower GM-RMSE than the ones forced with the CMC and GEFS ensembles, and also that ECMWF is not over-dispersed because the standard deviation in GM-RMSE among its members is the lowest compared to that of other ensembles (its bar-chart profile is flatter). CMC and GEFS also appear to be similarly-dispersed and quite similar in performance here. Further, and as expected, careful consideration of each ensemble internal structure reveals that control members tend to have the smaller GM-RMSE compared to an ensemble's perturbed members.

5 CONCLUSIONS AND FUTURE WORK

The redesigned, and now ensemble-based SFAS, was presented and its flood-predicting performance was compared to other operational flood prediction systems. The work shows that the new ensemble water level forecasts in SFAS provide improved predictions at tidal stations, thus improving the assessment of coastal flood risks when compared with deterministic water level forecasts, and leading the way for similar changes occurring at the NWS level. Section 4 showed that the weighted ensemble average has improved performance (significantly lower storm surge RMSE) at coastal stations compared to even some of the best members of the NYHOPS-E ensemble. Uncertainties in weather inputs may result in false warnings and missed flooding events, reducing the potential to effectively mitigate flood damage, and need to be accounted for through ensemble techniques [35].

The research presented also indicates the value of using increased resolution meteorological forecasts in such an ensemble: with the exception of SREF, the NYHOPS ensemble members that were forced with the higher resolution versions of each meteorological model, as well as the control members of the lower resolution ensembles produced results with the lowest RMSE in forecast water level. Statistical downscaling of meteorological ensembles to their high-resolution equivalents might increase accuracy further, and such downscaling methods should be investigated in the future.

Stevens SFAS forecasts were very accurate during winter storm Jonas, predicting no significant coastal flooding effects at Port Authority facilities well in advance of the storm, unlike the significant flooding predicted for South Jersey (https://twitter.com/PortAuthOEM/status/690897301905227776). SFAS has become part of the NWS information flow: The web site has been entrained into the NWS pages (http://www.weather.gov/phi/tides; Clicking on the New York Harbor map, the link takes a user to the Stevens web site). The NYHOPS ensemble tracked *Jonas* tidal water levels at the 85 SFAS stations for the most part within the 5%–95% uncertainty envelope displayed in the SFAS website. It is of concern that at most locations the observations were skewed high, closer to the 95% of the forecast ensemble envelope for this Nor'easter Storm, as this may be an indication of the weighted ensemble process over-smoothing peaks. There was also large uncertainty around the river water level ensemble prediction, primarily due to the large uncertainties associated with precipitation forecasts. Alternative methods to minimize that uncertainty and produce a more accurate ensemble are currently under active research and development with some promising theoretical results. Work needs to continue on developing better ensemble and debiasing techniques for tidal and inland stations and improving downscaling of the lower-resolution members of the ensemble.

With regard to SFAS website improvements, although inundation maps are now created with bath-tubbing techniques (http://hudson.dl.stevens-tech.edu/njdemo/), dynamic street-scale resolution sECOM models (as in [32]) for the PANYNJ facilities is now a focus activity. For the SFAS region as a whole, the hydrodynamic resolution deficit at the NJ Back Bays became obvious during Nor'Easter Jonas, with observations at a few sites in the Intracoastal Waterway (Sea Isle City NJ, Stone Harbor NJ, and even Ocean City Inlet MD) coming out of the 5%–95% Stevens Ensemble predictions at times.

ACKNOWLEDGMENTS

This work was partially funded by a research task agreement entered between the Trustees of the Stevens Institute of Technology and the PANYNJ, effective August 19, 2014. New York Sea Grant, New Jersey Sea Grant, Connecticut Sea Grant, the Long Island Sound Study, the NJ Coastal Protection Technical Assistance Service, the Borough of Manasquan, NJ, the Monmouth County Office of Emergency Management, the NJ Meadowlands Research Institute of the New Jersey Sports and Exposition Authority, the New Jersey Office of Emergency Management, the Hudson River Environmental Conditions Observing System, the United States Geologic Survey, and the NOAA Integrated Ocean Observing System program and the Mid-Atlantic Regional Association for Coastal Observing Systems (MARACOOS) are acknowledged for their contributions and support to the SFAS. We are thankful for the academic numerical models contributed to SFAS by Rutgers University and Stony Brook University, as well as for the operational models contributed by the US National Centers for Environmental Prediction, the Canadian Meteorological Centre, the European Centre for Medium-Range Weather Forecasts, and the NOAA Meteorological Development Laboratory. Finally, the authors are grateful to Ms. Elizabeth Lamb for help editing this document.

REFERENCES

[1] NWS, United States Flood Loss Report – Water Year 2014, available at: http://www.nws.noaa.gov/hic/summaries/WY2014.pdf (accessed 01 Febuary 2016), 2014.

[2] Georgas, N., Orton, P., Blumberg, A., Cohen, L., Zarrilli, D. & Yin, L., The impact of tidal phase on hurricane sandy's flooding around New York City and Long Island Sound. *Journal of Extreme Events*, **1**(01), p. 1450006, 2014.

[3] Botts, H., Du, W., Jeffery, T., Kolk, S., Pennycook, Z. & Suhr, L., CoreLogic Storm Surge Report 2015, Irvine, CA. p. 44, 2015.

[4] Schubert, C.E., Busciolano, R., Hearn Jr, P.P., Rahav, A.N., Behrens, R., Finkelstein, J., Monti Jr, J. & Simonson, A.E., *Analysis of Storm-Tide Impacts from Hurricane Sandy in New York*, US Geological Survey, 2015.

[5] Blake, E.S., Kimberlain, T.B., Berg, R.J., Cangialosi, J. & Beven II, J.L., Tropical cyclone report: Hurricane Sandy, National Hurricane Center, **12**, pp. 1–10, 2013.

[6] Anderson, M.L., Chen, Z.Q., Kavvas, M.L. & Feldman, A., Coupling HEC-HMS with atmospheric models for prediction of watershed runoff. *Journal of Hydrologic Engineering*, **7**(4), pp. 312–318, 2002.
http://dx.doi.org/10.1061/(ASCE)1084-0699(2002)7:4(312)

[7] USGS, January 22–24, 2016 *Summary of Coastal Flooding in New Jersey Caused by the Nor'easter,* available at: http://nj.usgs.gov/hazards/flood/flood1601/ 2016.
[8] Villarini, G. & Goska, R., Smith, J.A. & Vecchi, G.A., North atlantic tropical cyclones and U.S. flooding. *Bulletin of the American Meteorological Society*, **95**(9), pp. 1381–1388, 2014.
http://dx.doi.org/10.1175/BAMS-D-13-00060.1
[9] Berg, R., National Hurricane Center Tropical Cyclone Report Hurricane Joaquin (AL112015) 28 September – 7 October 2015, available at: http://www.nhc.noaa.gov/data/tcr/AL112015_Joaquin.pdf. 2016.
[10] Lumia, R., Firda, G.D. & Smith, T.L., Floods of 2011 in New York: U.S. geological survey scientific investigations report 2014–5058, p. 236, avaialble at: http://dx.doi.org/10.3133/sir20145058. 2014.
[11] Orton, P., Georgas, N., Blumberg, A. & Pullen, J., Detailed modeling of recent severe storm tides in estuaries of the New York City region. *Journal of Geophysical Research: Oceans*, **117**(C9), 2012.
http://dx.doi.org/10.1029/2012JC008220
[12] Noe, R.M., Casey-Lockyer, M., Mertzlufft, C., Heick, R., Yard, E. & Wolkin, A., Deaths associated with hurricane sandy. *Morbidity and Mortality Weekly Report (MMWR) Centers for Disease Control and Prevention*, **62**(20), pp. 393–397, 2013, available at: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6220a1.html. 2013.
[13] Weaver, J., Harkabus, L.C., Braun, J., Miller, S., Cox, R., Griffith, J. & Mazur, R.J., An overview of a demographic study of united states emergency managers. *Bulletin of the American Meteorological Society*, **95**(2), pp. 199–203, 2014.
http://dx.doi.org/10.1175/BAMS-D-12-00183.1
[14] Hoss, F. & Fischbeck, P., Increasing the value of uncertain weather and river forecasts for emergency managers. *Bulletin of the American Meteorological Society*, **97**(1), pp. 85–97, 2016.
http://dx.doi.org/10.1175/BAMS-D-13-00275.1
[15] Di Liberto, T., Colle, B.A., Georgas, N., Blumberg, A.F. & Taylor, A.A., Verification of a multimodel storm surge ensemble around New York City and Long Island for the cool season. *Weather and Forecasting*, **26**(6), pp. 922–939, 2011.
http://dx.doi.org/10.1175/WAF-D-10-05055.1
[16] Saleh, F., Ramaswamy, V., Georgas, N., Blumberg, A. & Pullen, J., A Retrospective streamflow ensemble forecast for an extreme hydrologic event: a case study of Hurricane Irene and on the Hudson River basin. *Hydrology and Earth System Sciences*, **2016**, pp. 1–29, 2016.
http://dx.doi.org/10.5194/hess-20-2649-2016
[17] Marsooli, R., Orton, P.M., Mellor, G., Blumberg, A.F., Georgas, N.A., Coupled Wave-Circulation Model for Numerical Simulation of Storm Tides and Waves, Ocean Modeling, Submitted.
[18] http://www.weather.gov/media/phi/veteransday2009_pres.pdf.
[19] Georgas, N. & Blumberg, A., Establishing confidence in marine forecast systems: the design and skill assessment of the New York Harbor observation and prediction system, version 3 (NYHOPS v3). Estuarine and Coastal Modeling, 660-685. 2010.
[20] Georgas, N., Blumberg, A.F., Bruno, M.S. & Runnels, D.R., Marine forecasting for the New york urban waters and harbor approaches: the design and automation of NYHOPS. In 3rd International Conference on Experiments/Process/System Modelling/Simulation & Optimization. Demos T. Tsahalis, Ph. D., University of Patras, Greece, 2009.

[21] NWS, Service Assessment: *Hurricane/Post-Tropical Cyclone Sandy. October 22–29, 2012*, Silver Spring, MD, USA: National Oceanic and Atmospheric Administration National Weather Service. 2013.

[22] NWS, *National Weather Service Manual 10-950,* DECEMBER 4, 2012, Operations and Services Hydrologic Services Program, NWSPD 10-9, available at: http://www.nws.noaa.gov/directives/. 2012.

[23] http://hudson.dl.stevens-tech.edu/SFAS/SFAS_Latest_News.pdf.

[24] Molteni, F., Buizza, R., Palmer, T.N. & Petroliagis, T., The ECMWF ensemble prediction system: Methodology and validation. *Quarterly Journal of the Royal Meteorological Society*, **122**(529), pp. 73–119, 1996. http://dx.doi.org/10.1002/qj.49712252905

[25] http://www.spc.noaa.gov/exper/sref/

[26] https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-ensemble-forecast-system-gefs.

[27] http://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forcast-system-gfs.

[28] http://www.ncdc.noaa.gov/data-access/model-data/model-datasets/north-american-mesoscale-forecast-system-nam.

[29] http://rucool.marine.rutgers.edu/COOL-Data/ruwrf-model-output.

[30] https://weather.gc.ca/ensemble/index_e.html.

[31] Blumberg, A.F., Georgas, N., Yin, L., Herrington, T.O. & Orton, P.M., Street-scale modeling of storm surge inundation along the New Jersey Hudson river waterfront. *Journal of Atmospheric and Oceanic Technology*, **32**(8), pp. 1486–1497, 2015.

[32] McEnery, J., Ingram, J., Duan, Q., Adams, T. & Anderson, L., NOAA'S advanced hydrologic prediction service: building pathways for better science in water forecasting. *Bulletin of the American Meteorological Society*, **86**(3), pp. 375–385, 2005. http://dx.doi.org/10.1175/BAMS-86-3-375

[33] Adams, T., Verification of the NOAA/NWS MMEFS Operational Hydrologic Ensemble Forecasting System in the Ohio River Valley. *World Environmental and Water Resources Congress 2015 at Floods, Droughts, and Ecosystems*, pp. 2688–2700, 2015.

[34] http://www.stormsurge.noaa.gov/models_obs_modeling.html.

[35] https://ams.confex.com/ams/96Annual/webprogram/Paper279348.html.