

# 1 Improving WRF-Hydro Runoff Predictions of Heavy Floods Through 2 Higher Spatio-Temporal Sea Surface Temperature Products

3 **Running Head:** Flood prediction with WRF-Hydro: Influences of Sea Surface Temperature

4 Berina Mina Kilicarslan <sup>1</sup>, Ismail Yucel <sup>1</sup>, Heves Pilatin <sup>2</sup>, Eren Duzenli <sup>1</sup>, and Mustafa Tugrul  
5 Yilmaz <sup>1</sup>

6 <sup>1</sup>Department of Civil Engineering, Middle East Technical University, Ankara, Turkey

7 <sup>2</sup> Earth System Science Department, Graduate School of Natural and Applied Sciences,  
8 Middle East Technical University, Ankara, Turkey

9

## 10 **Abstract**

11 In this study, the impact of spatio-temporal accuracy of four different sea surface temperature  
12 (SST) datasets on the accuracy of the Weather Research and Forecasting (WRF)-Hydro  
13 system to simulate hydrological response during two catastrophic flood events over Eastern  
14 Black Sea (EBS) and Mediterranean (MED) regions of Turkey is investigated. Three time-  
15 varying and high spatial resolution external SST products (GHRSSST, Medspiration, and  
16 NCEP-SST) and one coarse-resolution and invariable SST product (ERA5- and GFS-SST for  
17 EBS and MED regions, respectively) already embedded in the initial and boundary condition  
18 dataset of WRF model are used in deriving near-surface weather variables through WRF.  
19 After the proper event-based calibration performed to the WRF-Hydro using hourly and daily  
20 streamflow data of small catchments in both regions, uncoupled model simulations for  
21 independent SST events are conducted to assess the impact of SST-triggered precipitation on  
22 simulated extreme runoff. Some localized and temporal differences in the occurrence of the  
23 flood events with respect to observations depending on the SST representation are noticeable.  
24 SST products represented with higher temporal and spatial correlation revealed significant  
25 improvement in flood hydrographs for both regions. The higher spatial and temporal  
26 correlations of GHRSSST dataset show RMSE reduction up to 20% and increase in correlation  
27 from 0.3 to 0.8 with respect to the invariable SST (ERA5) in simulated runoffs over the EBS  
28 region. The error reduction with GHRSSST reached 35% after the calibration of hydrological  
29 model parameters compared to not calibrated model. The use of both GHRSSST and

30 Medspiration SST data characterized with high spatio-temporal correlation resulted in runoff  
31 simulations exactly matching the observed runoff peak of 300 m<sup>3</sup>/s by reducing the  
32 overestimation seen in not calibrated runs over the MED region.

33 **Keywords:** WRF-Hydro, WRF, Calibration, Sea Surface Temperature, GHRST,  
34 Medspiration

35

## 1. INTRODUCTION

Warming climate results in increased water vapor input into the atmosphere; consequently, triggering the intensity of rainfall events. (Trenberth, 1999; Allen & Ingram, 2002). The impact of the flood events might be exacerbated in time with the changing climate (Hirabayashi et al., 2013). Accordingly, accurate flood forecasting is important for many operational applications.

The forecast of heavy precipitation events with their spatial distributions and the forecast of their hydrological response are among the most significant elements of an accurate flood forecast (Shih, Chen, & Yeh, 2014; Yucel & Onen, 2014; Ryu et al., 2017). In this context, the application of a hydrometeorological modeling framework that can integrate atmospheric and hydrological models are started to be used commonly in practice for flood forecasting (Kunstmann & Stadler, 2005). Accordingly, accurate short-term predictions of runoff inherently require well-calibrated accurate hydrological model and accurate short-term predictions of atmospheric variables (e.g., precipitation and temperature) driving this hydrological model.

Selection of the numerical weather prediction (NWP) model and the datasets driving its boundary and initial conditions have profound effect over the accuracy of the short-term predictions of the atmospheric forcing datasets; hence, better operational flood forecasts clearly require improved NWP simulations. Such NWP simulations are particularly impacted from the SST state, as oceans/seas supply significant amount of both energy and water that the state of the atmospheric forcing variables are heavily impacted. Studies focusing on improvement of the accuracy of the existing operational flood forecasts, particularly near the coastal regions with complex topography, require an ocean-land-atmosphere coupled system to better reflect variability in all elements of the water and the energy balances as well as for accurate parameterization of the land-surface to better benefit from the input atmospheric forcing dataset.

SST primarily affects the heat and the water fluxes at the lower boundary of the atmosphere, hence there is a significant relationship between SST variations and convective extremes. In general, increasing SST state increases the moisture content in the air and warms the low level of the atmosphere (Lebeaupin, Ducrocq, & Giordani, 2006). This often results in stronger convection and higher precipitation totals over coastal regions. Overall, even

67 variations of SST in order of  $\pm 1$  °K may dramatically and nonlinearly change the intensity of  
68 the development of supercells over the seas (Miglietta, Mazon, Motola, & Pasini, 2017).  
69 Even if SST effects on long-term simulations are identified as small, it may still significantly  
70 affect the individual heavy precipitation events (Senatore, Mendicino, Knoche, &  
71 Kunstmann, 2014). Accordingly, improved representation of SST fields has a not negligible  
72 impact on simulation of the atmospheric boundary layer processes and flow dynamics  
73 (Senatore, Furnari, & Mendicino, 2020).

74 Given lower atmospheric boundary conditions often drive the precipitating water on the land  
75 surface, SST variations play a key role in heavy precipitation events (Bozkurt & Sen, 2011;  
76 Turuncoglu, 2015; Baltaci, 2017). A gradual increase in SST may cause a sudden  
77 amplification of convective precipitation extremes over the coastal regions (Meredith,  
78 Maraun, Semenov, and Park, 2015). Accordingly, providing higher accuracy SST input is  
79 crucial for accurate modeling of precipitation, hence for accurate flood forecasts through  
80 NWP models. Despite its significance and impact over the accuracy of the runoff forecasts,  
81 the number of studies inter-comparing the impact of spatio-temporal accuracy of different  
82 SST input datasets over the accuracy of the predicted runoff has remained limited so far  
83 (McCabe & Wolock, 2008; Chen, Wang, Xue, & Sun, 2009; Senatore et al., 2020).

84 A fully distributed, physical-based, multi-scale hydrometeorological modeling system, the  
85 WRF-Hydro is developed by the U.S. National Center for Atmospheric Research (NCAR) to  
86 investigate critical water issues, including flash flood forecasting applications. Allowing to  
87 run both in uncoupled (one-way from the atmosphere to land) mode and fully-coupled (two-  
88 way) mode (Gochis et al., 2020), this modeling system links the atmospheric and the  
89 hydrological processes. Overall, WRF-Hydro is designed as a framework to couple WRF  
90 (i.e., a NWP model) with a hydrological extension that enables simulation of land surface  
91 states and fluxes, including surface overland flow, saturated subsurface flow, and channel  
92 routing and vertical energy fluxes between land and atmosphere through physics-based and  
93 conceptual approaches. Despite many studies have been performed so far investigating the  
94 performance and application of the WRF-Hydro model (Kerandi et al., 2018; Wehbe et al.,  
95 2019; Varlas et al., 2019; Sun et al., 2020), not many studies have investigated the impact of  
96 the spatio-temporal accuracy of various SST sources over the predictions of runoff using  
97 WRF-Hydro modelling system. Among them, studies utilized high-resolution SST inputs and

implemented parameter calibration in prediction of runoff have particularly remained limited with the study of Senatore et al. (2020).

Surrounded by sea from three sides and having one of the most complex topography in the region, Turkey has many locations living with significant potential flood threats produced by the meteorological, hydrological, and topographical differences. EBS and MED regions of Turkey are among the most vulnerable regions in terms of flood risk in the Anatolian peninsula (Gurer, 1998; Gurer & Ucar, 2009; Duzenli, Yucel, Pilatin, & Yilmaz, 2020). Forecasting the floods through high resolution NWP models in MED region is critical (Camera, Bruggeman, Zittis, Sofokleous, and Arnault, 2020), where a gradual increase in SST may cause sudden amplification of convective precipitation extremes over the Black Sea coastal regions (Meredith, Maraun, Semenov, and Park (2015) and SST variations play a key role in heavy precipitation events in the Anatolian Peninsula (Bozkurt & Sen, 2011; Turuncoglu, 2015; Baltaci, 2017). On the other hand, the number of studies investigating the impact of utilizing various spatio-temporal accuracy SST products over the formation of heavy precipitation that may cause floods over EBS and MED regions remains lacking.

Accordingly, the main goal of this study is to 1) evaluate the impact of the spatio-temporal accuracy of SST products on the accuracy of the modelled hydrological response over the small catchments located in MED and EBS coastal regions with different climatic characteristics, 2) investigate the impact of calibration of WRF-Hydro parameters over the benefit obtained from the use of high spatio-temporal accuracy SST products, 3) investigate the consistency of the sensitivity analysis to different geographic regions with vastly diverse climate.

In this study, the uncoupled WRF-Hydro simulations are forced by the WRF model meteorological forcing data created via initial and lower boundary conditions updated with different SST products (GHRSSST, Medspiration, NCEP, ERA5/GFS), while WRF-Hydro parameters responsible from hydrological processes are calibrated. WRF precipitation forecasts and WRF-Hydro runoff simulations are independently validated using ground observations collected during three different heavy precipitation events for each basin over MED and EBS regions. Thereby, the accuracy of the WRF-Hydro model predictability is assessed not only with SST product sensitivity but also with model parameter calibration.

## 2. DATA AND METHODS

### 2.1 Study Area and Event Description

Two significant SST-related heavy precipitation events (Pilatin, 2020). generated flash flood over catchments located in the EBS and MED regions with different climatic characteristics are considered for analysis. Nested 3-km WRF domains (d02) covering the EBS and MED regions, selected basins together with their channel networks, location of both meteorological and stream gauge stations are shown in Figure 1.

EBS region is located in the North-Eastern part of Turkey, where mountains lie parallel to the shore and act as a barrier to humid air currents. The mountains rise above 3000 m and result in complex topography and steep-sloped characteristics (Eris & Agiralioglu, 2018). Due to small basin structures and steep rocky characteristics, river systems can react quickly to moderate precipitation events and cause flash floods (Gurer & Ucar, 2009; Eris & Agiralioglu, 2018). The region exhibits a humid climate and receives rainfall throughout the year (Turkes, 1996). It has the highest mean annual recorded precipitation exceeding 2200 mm (Baltaci, 2017).

MED region has typical Mediterranean climate prevailing humid and semi-humid subtropical characteristics with a rainy winter/spring and a severe hot-dry summer (Turkes, 1996). The precipitation amount of the region is more than 1000 mm, and in many points, it exceeds 2000 mm (Turkes, 1996; Eris & Agiralioglu, 2018). Mean annual precipitation is 800 mm over the MED coasts, and it increases up to 1500 mm over the Taurus Mountains (Turkes, 1996; Turkes, 1999). Details of air masses affecting the regions are described by Duzenli et al. (2020). Typical topographic characteristics and sea effect point out that the strong orographic lifting dependency and elevated heat sources for convective initiation exist in both regions. Since high SST increases the moisture content in the air, it has a critical role in the occurrence of flood events in such regions located in coastal areas with complex topography.

[Insert Figure 1]

The peak hourly precipitation amount that occurred on 24 August 2015 over the EBS region is recorded as 32.4 mm at Artvin-Arhavi, while total 135 mm of precipitation accumulated within 24-hours. On the other hand, for the MED event occurred on 16 December 2018, the peak hourly precipitation was recorded as 53.1 mm at Antalya-Ovacik station, while it

received the total daily precipitation amount of 651.7 mm. This event was registered as the highest precipitation record measured in Turkey (Kaya, Guler Altan, & Yorganci, 2019). This value is almost three times higher than the monthly average precipitation in December (265.3 mm) for Antalya city. The precipitation system for the event that occurred during the summer season over the EBS region shows typical mesoscale convective signature, whereas the frontal system is dominant for the event occurred over the MED region during the winter season.

Over the EBS region, the drainage area of the D22A049 stream gauge and its sub-basins (D22A079 and D22A089) located in Arhavi province and the drainage areas of the D22A147 stream gauge in Hopa province are selected as study basins while the drainage areas of D08A071, D09A095, and E08A008 stream gauges are selected over MED region for WRF-Hydro Model (Figure 1 and Table 1). The streamflow observations from 7 stream gauge stations are provided by the State Hydraulic Works (SHW) of Turkey. Streamflow is provided as an average daily record in m<sup>3</sup>/s for selected stations and event periods except for D22A049 and D08A071; it is provided as an hourly record for the events that occurred after 2016 (Table 1).

[Insert Table 1]

## 2.2 WRF Model

In this study, the Advanced Research WRF model version 4.0 (Skamarock et al., 2019) developed by NCAR is used to reproduce the meteorological forcing data of the WRF-Hydro model for the selected heavy precipitation events. Two-way nesting model configuration is applied with spatial resolution specified at 9-km for the outer domain (d01) and 3-km for the inner domain (d02). The outer domain as shown in Figure 1 extends 23.5°E-47.5°E;34.5°N-43.5°N, and contains 232 × 111 grid points. Also, the inner domain over the MED region is placed between 47.5°N – 32.4°N, 34.5°E – 36.4°E coordinates with 73 × 88 grid points, while over the EBS region, it is placed between 47.5°N – 41.6°N, 23.5°E – 36.9°E coordinates with 136 × 52 grid points.

In this study, two different Global Circulation Models (GCMs) are selected as initial and boundary conditions to be used in the WRF model. Following the studies of Duzenli et al. (2020) and Pilatin (2020), the Global Forecasting System (GFS) forecast dataset is used over the MED region, while The European Centre for Medium-Range Weather Forecasts

(ECMWF) ERA5 Re-analysis dataset (ECMWF, 2020; NOAA, 2015) is used over EBS region simulations as not updated.

In addition to ERA5 and GFS, three other external SST datasets are used for the sensitivity analysis in this study: 1) Medspiration Level 4 Ultra-High-Resolution Foundation Sea Surface Temperature (CERSAT, 2012); 2) The Group for High-Resolution Sea Surface Temperature Level 4 Ultra-High Resolution (GHRSSST) (Team GHRSSST, 2010a; Team GHRSSST, 2010b); 3) Real-Time, Global, Sea Surface Temperature (RTG\_SST\_HR) represented by the National Centers for Environmental Prediction (NCEP), National Oceanic and Atmospheric Administration (NOAA) (NCEP & NOAA, 2014). These products have high spatial resolutions ( $0.01^\circ$ ,  $0.022^\circ$ , and  $0.083^\circ$ , respectively) and are provided on daily basis. From here on, the SST products used in this study will be referred as Medspiration, GHRSSST, NCEP, ERA5, and GFS. Information about simulation periods of the WRF model runs using these SST products over each study region are given in Table 2.

[Insert Table 2]

### **2.3 WRF-Hydro Model**

This study operates the WRF-Hydro model version 5.1.1. configured in an uncoupled way over the 3-km nested domain (d02) of the WRF model. Noah–Multi Parameterization (Noah-MP) is selected for the model configuration as the land surface model (LSM). In model physics options, surface overland and subsurface routing modules are activated for the whole domains, whereas the channel routing module is only activated within the study basins. The baseflow bucket model is also activated with the pass-through option. Detailed descriptions of WRF-Hydro model structure and routing modules are available in (Gochis et al., 2020). After the moisture states are calculated for the land surface column, the LSM grid disaggregates into the high-resolution routing grids of 250-m resolution for both study regions. High-resolution routing layers are produced from a hydrologically conditioned digital elevation model (DEM) from the HydroSHEDS of Lehner, Verdin, and Jarvis (2008) using the WRF-Hydro Pre-Processing toolbox in the GIS environment.

In calibration simulations of the WRF-Hydro model, among meteorological inputs derived from WRF model the hourly precipitation field is updated by the observed precipitation. Based on streamflow data availability, model calibration is performed for three events for each basin (7 basins in total, see Table 1), and the SST events are used independently to



validate the calibrated parameter set in terms of the performance of the WRF-Hydro model. Calibration of the model is manually employed with a step-wise approach as described in Yucel, Onen, Yilmaz, and Gochis (2015). In the first step, parameters controlling the hydrograph volume called infiltration factor (REFKDT), surface retention depth (RETDEPRT), and deep drainage coefficient (SLOPE) are calibrated. Surface roughness coefficient (OVROUGHRT), channel Manning roughness coefficient (MANN), and saturated hydraulic conductivity factor (LKSATFAC) being considered as parameters controlling hydrograph shape (temporal distribution and peak timing) are calibrated in the second step. Similar procedure is commonly adopted for the calibration of WRF-Hydro in terms of water balance and its distribution (Yucel et al., 2015; Senatore et al., 2015; Naabil, Lamptey, Arnault, Kunstmann, & Olufayo, 2017; Yang, Yuan, & Yu 2018; Liu et al., 2020). Some parameters (REFKDT, SLOPE, MANN) are defined in tabular value format considered as global values over the domain. Others are defined as pixel specific (RETDEPRT, OVROUGHRT, LKSATFAC) that enables to change parameter value only for each basin.

Statistical measures are implemented between observed and simulated discharge for the model accuracy evaluation, namely bias, root mean square error (RMSE), and correlation coefficient (RR) to find the best parameter value among the different events for each basin. Bias represents the degree of overestimation and underestimation in hydrograph volume. RR reflects the linear relationship between observed and modelled flow and calculates the capturing performance of the timing and shape of the hydrograph. Besides, RMSE is sensitive to both the shape and the volume of the hydrograph (Moriasi et al., 2007; Gupta, Kling, Yilmaz, & Martinez, 2009). This statistical evaluation is performed based on hourly or daily time steps depending on the available temporal resolution of streamflow data of selected stream gauges.

### **3. RESULTS**

#### **3.1 Spatio-Temporal Accuracy Evaluation of SST Products**

GFS and ERA5 products have coarser spatial ( $0.25^\circ$ ) resolution than GHRSSST (1.1-km), Medspiration (2.2km), and NCEP products (9-km). In this study, GHRSSST, Medspiration, and NCEP products are selected to have daily temporal resolutions, while GFS and ERA5 SST products temporally remained constant during event simulations. Temporally averaged (10-days) spatial distribution of these products are shown in Figure 2, while spatially

252 averaged time series are given in Figure 3. Eastern part of the EBS region is depicted by  
253 warmer temperatures ( $\sim 301\text{K}$ ) than western part ( $\sim 297\text{K}$ ) consistently by all products (except  
254 for NCEP). Over MED region, inter-product consistency is much smaller than EBS region  
255 that spatial variability of average temperature is largest (Figure 2). Overall, all products,  
256 except for constant ERA5 and GFS, are temporally consistent with each other particularly  
257 over MED region (Figure 3).

258 For any product, spatial and temporal cross-correlations are calculated with other products,  
259 and then these cross-correlations are averaged (Table 3). Given there are no buoy  
260 observations over the study regions to validate the accuracy of SST products, here average  
261 cross-correlations are used as an indicator of true signal assuming there is no other common  
262 spatial and temporal signal between the products (i.e., higher average cross-correlations  
263 imply a better product). In general, the average temporal cross-correlations are higher over  
264 MED region than EBS, while vice versa for spatial cross-correlations (Table 3). Overall,  
265 average spatio-temporal cross-correlation for GHRSSST (0.61) is higher than Medspiration  
266 (0.54), which is higher than NCEP (0.36); this order is also valid for average spatial and  
267 temporal cross-correlations as well as EBS and MED regions; this implies, among the time  
268 varying SST products, GHRSSST is the best and NCEP is the least performing products for the  
269 events and regions focused in this study.

270 [Insert Figure 2]

271 [Insert Figure 3]

272 [Insert Table 3]

### 273 3.2 Calibration of the WRF-Hydro Model

274 Results for the hourly calibration of selected parameters within the WRF-Hydro model is  
275 shown in Figure 4. In this figure, first column (a-f) represents the calibration results of the  
276 event occurred between 10/19/2016 to 10/29/2016 at D22A049 basin located over EBS  
277 region while the second column (g-l) belongs to the event occurred between 03/07/2017 to  
278 03/17/2017 at basin D08A071 located over MED region. Two more additional events belong  
279 to each catchment are also used in the calibration process (Table 1). Table 4 and Table 5  
280 show the average statistical measures calculated for the WRF-Hydro model set up with

281 default parameter set and for the simulation of selected parameter value of each catchment  
282 considered over EBS region and over MED region, respectively.

283 [Insert Figure 4]

284 [Insert Table 4]

285 [Insert Table 5]

286 Depending on the step-wise approach, the calibration procedure starts with the group of  
287 parameters controlling the hydrograph volume. Initially, calibration of the REFKDT  
288 parameter (default value of 3.0) is performed with the parameter values between 0.5 and 5.0  
289 with 0.5 increments. Figure 4(a) and Figure 4 (g) show the results of D22A049 and D08A071  
290 basins, respectively. It can be inferred as the higher the REFKDT value lower the infiltration  
291 capacity of the soil column, in turn, the higher the hydrograph volume. According to the  
292 statistics and comparison with the calibration hydrographs based on the other two events, it is  
293 decided on to the lowest value (0.5) of REFKDT as optimum for both basins. However, there  
294 is still an underestimation observed in the D022A049 hydrograph volume in Figure 4 (a). The  
295 simulated first peak in day-8 is lowered, and the simulated hydrograph is fed through the  
296 observed peak that occurred in between day-7 and day-8. On the contrary, when the average  
297 bias is calculated for three events, bias turns into 3.72 in Table 4. Similar contrast is also  
298 observed in the D08A071 station. Negative bias is observed for the average of three events  
299 (Table 5), while an overestimation is observed for the represented event in Figure 4 (b).  
300 Overall statistic shows that REFKDT parameter strongly sensitive in both regions.

301 Figure 4 (b) and Figure 4 (h) shows the calibration results of the RETDEPRTFAC parameter  
302 with the range of 0.0-10.0 with 1.0 increment. Simulated hydrographs of both basins are not  
303 showing an apparent response to the RETDEPRTFAC parameter (Table 4 and Table 5).  
304 Since EBS and MED regions have steep topography, little water accumulation over the  
305 terrain is expected to be observed. Therefore, the optimum RETDEPRTFAC parameter value  
306 is selected as 0.0 for both basins.

307 The SLOPE is considered for the model calibration using values between 0.1 and 1.0 with 0.3  
308 increments. Similar to Wang et al. (2019), only the first class of the nine SLOPE\_DATA  
309 categories represented in GENPARM.TBL is subjected to tuning. This parameter controls the  
310 openness of the bottom soil column to the conceptual bucket. It shows little influence on

311 simulated hydrographs in terms of statistics. The default value is selected as an appropriate  
312 SLOPE parameter value for the model in both basins. However, it is observed that the other  
313 calibrated events in D22A147 and D09A095 basins show improvement in RMSE and  
314 correlation coefficient with the SLOPE parameter (Table 4 and Table 5).

315 For the second step, parameters controlling hydrograph shape and timing are considered for  
316 the calibration process. Figure 4 (d) and Figure 4 (j) show the results from the calibration of  
317 the OVROUGHRTFAC parameter with values ranging from 0.1 to 1.0 with 0.3 increments.  
318 OVROUGHRTFAC has an impact on the speed of the infiltration excess water transmitted  
319 through the channel network grids. According to statistical measures, the default value of  
320 OVROUGHRTFAC is found to be the optimum for all basins except the value of 0.1,  
321 selected for basin D09A095 (Table 5).

322 Manning's Roughness scaling factor for all stream orders is calibrated with a scaling factor  
323 (MANN) within a range from 0.5 to 2.0 with 0.5 increments. MANN controls the conveyance  
324 time of the flow through the channel network, which can be interpreted as the higher MANN  
325 values creates a slower peak and lower hydrograph volume. Figure 4 (e) and Figure 4 (k)  
326 show that the highest correlation is seen for the value of 2.0. In addition, RMSE improvement  
327 is observed in all basins for value of 2.0. Also, similar improvement is observed for value of  
328 0.5 in E08A008 (Table 4 and Table 5). Thus, scaling factor (MANN) is selected as 0.5 for  
329 E08A008, while 2.0 is selected for others.

330 Lastly, the LKSATFAC parameter, which affects the lateral redistribution of infiltrated water,  
331 is calibrated for the values of 10, 100, 1000 (default), and 10000, as it is shown in Figure 4 (f)  
332 and Figure 4 (l). It appears that LKSATFAC is the most sensitive parameter in both regions  
333 particularly for the MED region. It influences peak timing and its magnitude with a  
334 significant decrease. Over both regions, the value of 10 is determined as the optimum value  
335 for LKSATFAC.

336 In Table 4, progressive improvement in RMSE and correlation coefficient is observed from  
337 the first simulation (with default parameter set) to the simulation of LKSATFAC in step wise  
338 manner for both basins. With the calibration, correlation coefficient is increased from 0.13 to  
339 0.56, while RMSE is reduced from 40.55 to 32.16 for D22A49. On the other hand, bias  
340 switches to negative value which is likely resulted from the effect of sharp decrease in the  
341 recession stage in Figure 4(f). In D22A147, significant improvement is observed in

correlation coefficient (from 0.38 to 0.71) at the end of the calibration process. For D08A071, an improvement is observed only in correlation coefficient, while bias and RMSE increase after the calibration of the MANN (Table 5). In D09A095 and E08A008, statistics at the end of the calibration process show an improvement compared to the model performed with default parameters (Table 5). E08A008 exhibits no response to the RETDEPRT, SLOPE and OVROUGHRTFAC. As a result, it appears that the WRF-Hydro model is considerably sensitive to the LKSATFAC parameter especially in the MED region. Calibrated parameters for each basin with their default values are shown in Table 6.

[Insert Table 6]

### 3.3 Precipitation evaluation for each SST case

Figure 5 (a) and (b) show the comparison between observed and WRF-derived basin-averaged precipitation time series of each SST case for D22A147 and D08A071 basins, respectively. On the other hand, Table 7 shows the statistical measures calculated for each SST case in both basins. In Figure 5 (a), the precipitation time series are represented from 08/17/2015 00:00:00 UTC to 08/27/2015 00:00:00 UTC (241-hours). The maximum precipitation amount for D22A147 is recorded as 26.3 mm for the 178<sup>th</sup> hour, which corresponds to 08/24/2015 09:00:00 UTC. However, the maximum precipitation for the EBS region for this event was recorded as 32.4 mm at 08/24/2015 00:00:00 UTC. The spatial patterns of this precipitation amount measured in the meteorological station towards the D22A049, not in the range of D22A147 boundaries. Nevertheless, as shown in Figure 5 (a), the effect of event center on the basin-average precipitation of the D22A147 is still observed, and it is recorded as 16.1 mm at the 169<sup>th</sup> hour, which corresponds to the event peak time for the EBS region. Also, it can be interpreted that simulations performed with different SST datasets are able to catch the general pattern of the observation, except they generate the primary peak couple of hours earlier than the observation peak. However, notwithstanding the poor statistical measures (low correlation of 0.01-0.03 and high RMSE of 3.19-5.30) in Table 7, it can be depicted that using an external high-resolution SST dataset still improves the accuracy of the simulated precipitation, especially for Medspiration. Besides, GHR SST simulation overestimates the observed peak precipitation. Other simulated peaks are lower than the GHR SST simulation, but they are closer to the observed peak.

In Figure 5 (b), the basin-averaged precipitation time series are represented from 12/10/2018 00:00:00 UTC to 12/20/2018 00:00:00 UTC (241-hours). Peak time and precipitation magnitude for the whole MED region is recorded as 53.1 mm at 162<sup>th</sup> hour (at 12/10/2018 17:00:00 UTC). The maximum basin-average precipitation value of 15.7 mm is calculated at the same time step for the D08A071. Overall, simulated precipitations show nearly the similar trend as the observation with minor overestimations. Nonetheless, it appears that external SST simulations are able to improve the precipitation volume with reduced bias. Modest delays in peak time (1-2 hours) are observed for GFS SST, GHRSSST, and NCEP simulations, while Medspiration catches the exact peak time. Comparing with the observed peak precipitation amount, the GFS SST creates the highest overestimation around 17 mm, and in terms of model run period, it creates a positive bias value of 0.56 (Table 7). Medspiration shows the best model performance in terms of all statistics calculated with respect to the observed precipitation compared to the rest (Table 7).

[Insert Figure 5]

[Insert Table 7]

Figure 6 shows the spatial distribution of observed precipitation and simulated precipitations from the WRF model created by different SST datasets in peak day (08/24/2015) over the EBS region. Observed precipitation map is created by IDW method using the point observations, as shown in Figure 6 (a). It is noteworthy that in Figure 6 GHRSSST simulation shows an overestimation in spatial distribution of precipitation over the D22A147 compared to observed precipitation (Figure 6 (c)). Medspiration generates the closest precipitation distribution to the observation over the D22A147, consistent with the previously mentioned remark that Medspiration improves the accuracy of precipitation estimates compared to native coarse-resolution SST dataset (ERA5) in Figure 6 (d). Medspiration and GHRSSST simulations also overestimate the precipitation towards the coastline, where they produce more than 140 mm of daily precipitation (Figure 6 (c and d)). Besides, NCEP simulation leads to the underestimation of the simulated precipitation as shown in Figure 6 (e). On the other hand, GHRSSST catches the observed event location compare to other simulations considerably (Figure 6 (c)).

[Insert Figure 6]

For the MED region, Figure 7 shows the spatial distribution of simulated precipitation (GFS, GHRSSST, Medspiration, and NCEP) and observed precipitation with a maximum depth of 53.1 mm at the peak hour (Figure 7 (a)). Simulation performed with GFS SST shows overestimation in terms of precipitation amount. It also misses the event location and creates the event over the sea near the coastline instead of over the land (Figure 7 (b)). Besides, simulations performed with external high-resolution SST datasets are reasonably well represented compared to GFS simulations to catch the event location over the land. Figure 7 (c) shows that GHRSSST simulation can capture the observed event location yet, it cannot generate enough precipitation and causes underestimation with a depth of 16-18 mm, which is due to the modest delay in peak time mentioned earlier. Medspiration and NCEP simulations reveal much closer precipitation predictions to the observation in terms of precipitation depth (Figure 7 (d-e)). Especially, Medspiration simulation steps forward in generating similar precipitation depth and catching the similar hotspot of the observed event in Figure 5 (b). However, it overestimates the observed precipitation by ~8 mm (the highest hourly precipitation for Medspiration simulation over D08A071 is 25.8 mm which corresponds to the darker orange coloring of the basin grids (Figure 7 (d)).

[Insert Figure 7]

### 3.4 Evaluation of the WRF-Hydro for SST events

The performance of the calibrated WRF-Hydro model is evaluated using each SST case in D22A147 and D08A071 basins. In Figure 8 shows the simulated hydrographs by uncalibrated and calibrated models in D22A147. ERA5 and NCEP hydrographs show substantial underestimation for the peak volume (Figure 8 (a)). This may due to the negative bias observed in precipitation in Figure 5 (a) for ERA5 and NCEP simulations (They are the ones with the highest negative bias among other SST simulations). Medspiration simulation creates slightly better hydrograph volume and statistics compare to ERA5 and NCEP simulation. Though the GHRSSST generates overestimation in precipitation and misses the event peak time for D22A147 as discussed in the previous session (Figure 5 (a)), the daily mean discharge of the GHRSSST simulation makes the best improvement in the discharge estimation. This is due to the fact that the WRF simulation of the GHRSSST generated the most realistic amount of water volume that the D22A147 received on peak day (in 08/24/2015) as shown in Figure 6 (c). Therefore, the daily mean of the total water conveyed to the channel network after the water balance calculations resulted in the closest simulated

discharge volume to the observed one with the lowest negative bias and RMSE values among the other simulations (Figure 8 (a)). The bias value of the simulated hydrographs with GHR SST precipitation is reduced by -1.8 (from -10.5 to -8.7) while RMSE is reduced by 4.2 (from 20.7 to 16.5) as compared to hydrograph simulated with ERA5 precipitation (Table 8). On average, correlation coefficients increase from 0.3 (for ERA5) to 0.8 for the simulated hydrographs with high-resolution SST datasets. A sharp decrease in the recession stage in the hydrographs of all simulations is observed as different from the observed hydrograph. Overall, from the statistical measures in Table 8, it can be seen that simulated hydrographs obtained from WRF model forcings derived by high-resolution SST datasets show better performance in terms of both peak timing and hydrograph volume corresponding to the observed hydrograph.

[Insert Figure 8]

[Insert Table 8]

In Figure 8 (b), the realistic volume increase is observed in the simulated hydrographs through the calibrated set of parameters in the D22A147. The correlation coefficients of simulated hydrographs are similar to those before calibration, except the correlation coefficient of ERA5 simulation increases from 0.3 to 0.4. Medspiration and NCEP hydrographs volumes are improved, and they are way closer to the volume of observed hydrograph, but their underestimation is still higher compare to GHR SST hydrograph. The calibrated parameter set also substantially increases the GHR SST hydrograph volume and makes it closer to the observation compare to other simulations. For GHR SST simulated hydrograph, bias and RMSE is reduced by -2.5 (from -8.7 to -6.2) and 5.7 (from 16.5 to 10.8), respectively (Table 8). These results indicate that the GHR SST is the most representative SST dataset for D22A147 among the other SST datasets in the way of its positive effect on simulated hydrograph and the calibration of the WRF-Hydro model is also essential to further improve the model simulation, especially in terms of hydrograph volume.

Comparison of hourly observed hydrograph and simulated hydrographs forced by four different SST events in the D08A071 basin is represented in Error: Reference source not found (a) (plotted for the last six days of the model run period). Figure 9 (b) shows the equivalent plots with the set of calibrated parameters for the D08A071. In Figure 9 (a), hourly simulated discharge patterns are well matched with the observation for high-resolution



SST datasets (GHRSSST, Medspiration, and NCEP) simulations with the correlation coefficient values of  $\sim 0.6$  (Table 8). Minor delays in the primary hydrograph peak time are observed for the simulated hydrographs with GHRSSST and NCEP. They overestimate the observed discharge until peak time, yet the underestimation in the falling limb stage causes negative bias between -18.85 and -26.24 as shown in the Table 8. The simulated hydrograph of GFS SST produces a substantially higher peak of  $877.4 \text{ m}^3/\text{s}$  compared to the observed hydrograph and mismatches the hydrograph timing trend. The overestimation in a peak discharge of this hydrograph is likely due to the positive bias in the peak time of hourly precipitation time series of GFS SST in Figure 5 (b). Though the GFS SST hydrograph has the lowest bias value (-7.1), it produces the highest RMSE (125.9) and lowest correlation coefficient (0.3) (Table 8). Therefore, the simulated hydrograph shows better performance in terms of hydrograph peak timing and magnitude with the WRF forcing updated by external high-resolution SSTs, consistent with that they show the closer spatial distribution of precipitation to observation in peak time over the D08A071 (Figure 7).

[Insert Figure 9]

Simulated hourly hydrographs with the calibrated parameter set in Figure 9 (b) represent better behaviour in rising limb part till their peak values, but their falling limb parts decreases more sharply after the calibration. It can be interpreted that model is trying to adapt to extraordinarily high observed peak discharges ( $301.4 \text{ m}^3/\text{s}$ ) via calibration. This is likely the evidence for the discrepancy in statistical measures in Table 8, are getting worse after the calibration of the model. The observed peak value is greatly captured by external high-resolution SST products with a reduction of  $\sim 100 \text{ m}^3/\text{s}$ .

### 3.5 Evaluation of Rainfall-Runoff Representations

Figure 10 shows overlapped dynamic maps of accumulated precipitation simulated by the WRF model using four different SST datasets (ERA5, GHRSSST, Medspiration and NCEP) and simulated discharges on the gridded river networks corresponding to these four precipitation estimates over the EBS region for D22A49 and D22A147 basins. Blue dots over the maps highlight the location of outlet points (stream gauge station from Figure 1) of the basins. The first-time step in Figure 10 (a-d) shows the accumulated precipitation shortly before the start of the precipitation event and the state of the river networks of the D22A49 and D22A147 having the discharge at the baseflow level. In Figure 10 (f), at the second time

step, the D22A147 basin receives the highest precipitation compared to others; this result is consistent with that the simulated precipitation with GHRSSST generates the highest overestimation stated in the previous section. Due to the steep slope characteristics of the basins over the EBS region, it can be seen that the precipitation is immediately conveyed (less than 1 hour) to the river network and collected to the outlet point and lead to flooding. This is clearly seen in Figure 10 (e) for the D22A049, in Figure 10 (f) for the D22A147, and in Figure 10 (g-h) for both basins. For the third time step, the river network responds with lowered discharge values and lastly returns to the baseflow since there is no significant precipitation observed at the previous time step (Figure 10 (i-l)).

[Insert Figure 10]

Figure 11 shows overlapped dynamic maps of accumulated precipitation simulated by the WRF model through using four different SST datasets (GFS, GHRSSST, Medspiration, and NCEP) and simulated discharges on gridded river networks corresponding to these four precipitation estimates over the MED region for D08A071, D09A095, and E08A008 basins. The first-time step (02:00:00 UTC) demonstrates the precipitation event start over the basins located towards the east at which channel grids of mentioned basins are started to be filled with water (Figure 11 (a-d)). At 16:00:00 UTC, the simulated discharge amount with GFS-SST at the outlet of the D08A071 reaches from 142 to 516 m<sup>3</sup>/s as a response to the accumulated precipitation for 14 hours, especially over the upper basin (Figure 11 (a and e)). The precipitation event takes place towards the D09A095 for MED-SST simulation, and it appears that precipitated water is collected from the upper basin and conveyed to the outlet point and reaches the discharge value of 698 m<sup>3</sup>/s (Figure 11 (g)). In Figure 11 (l), due to the minor delays in primary peak time discharge in hydrographs of GHRSSST and NCEP, the channel grid network still on the rising limb stage with respect to the simulated hydrographs in Figure 9 (b)

[Insert Figure 11]

#### 4. DISCUSSION

Various SST products indeed resulted in different precipitation variability both in space and time over both regions, while the spatial and the temporal differences in precipitation greatly affect the accuracy of runoff simulation in terms of timing and magnitude of the peak value, and overall volume (Yucel et al., 2015; Senatore et al., 2020). Overall, GHRSSST product

yield the highest spatio-temporal accuracy, while NCEP yield the least among the temporally variable SST products. Consistent with this result, GHRSSST-based runoff simulations yield the highest accuracy, while NCEP the lowest among the temporally variable SST products. These results clearly show the significance of using higher spatio-temporal accuracy SST products in the simulation of heavy rainfall and extreme runoff.

In this study, cross-correlations are used as a validation tool, where buoy observations are non-existent over the study regions. Accordingly, the consistency between the cross-correlation based accuracy estimates and the runoff simulations show such cross-correlation-based methodology can be used over other remote locations that do not have buoy observations to validate SST products.

In general, simulated hydrographs show strong sensitivity to simulated precipitation inputs based on different SST products as well as significant variability from event to event. It is indicated that WRF Hydro and its calibration process function reasonably well in that calibration tends to improve model simulations when appropriate precipitation inputs are used. With the hourly calibration procedure, simulated hydrographs shapes over both regions are significantly improved. With sharp and steep small catchments over the EBS, the hydrologic response is very fast and overland flow is quickly joined to the river networks and pours to the outlets within 1-h period. The high-resolution gridded rainfall-runoff coupling greatly benefits to monitor the water excess condition for a given storm over topographically complex and steeply small watersheds.

In event simulations by WRF, the updates in SST through model integration are usually not activated because the variability of SST is small during a short event period. However, it is expected that changing climate causes abnormal SST changes that trigger the formation of the occurrence of heavy precipitation events (Pilatin, 2020). The daily updated SST products from GHRSSST, Medspiration, and NCEP over both study regions revealed significant changes in heavy precipitation amounts with respect to the not-updated (native) SST products from GFS over MED and ERA5 over EBS. They improved the accuracy of predictions in terms of storm location, timing, and extent particularly over the MED region. As a result, the Medspiration over the MED region and GHRSSST over the EBS region revealed the best basin-averaged precipitation representation that directly translates into improvement in surface runoff prediction in small catchments of both study regions.

The high spatio-temporal resolution SSTs (GHRSSST and Medspiration) help resolve high variability in rainfall and its hydrologic response resulted from a mesoscale convective system occurred in the ESB region. The calibrated WRF-Hydro model significantly highlighted the improvement provided by these two SST products over the EBS region. Even though the statistics show some degradation in runoff results after model calibration, the calibrated model indeed improved the rising limb parts of the storm hydrographs till their peak occurrence particularly for Medspiration- and GHRSSST-based simulations over the MED region. Since the MED SST event produced an observed peak around 300 m<sup>3</sup>/s, the calibration became highly sensitive to this peak value and therefore it showed a poor performance in describing the falling limb parts of the hydrographs. The effective parameter sets controlling the volume and shape of the hydrograph need to be identified prior to the operational runoff forecast to perform more accurate forecasts (Yucel et al., 2015; Senatore et al., 2015; Silver, Karnieli, Ginat, Meiri, & Fredj, 2017). Among the parameters, REFKDT, SLOPE, MANN and LKSATFAC revealed an important impact on making reliable runoff prediction in both regions but especially the saturated hydraulic conductivity parameter factor (LKSATFAC) became substantially critical over the MED region.

## 5. CONCLUSIONS

This study investigated the hydrologic response of the small catchments characterized by complex coastal orography and diverse climate to the heavy precipitation events simulated by various SST products featured as coarse- and high-resolution, and daily updated and not updated within the WRF model. The flood hydrographs of the heavy rainfall events are simulated using the physical-based and fully-distributed WRF-Hydro model configured with one-way coupling from WRF 3-km domain to the Hydro model. GFS over the MED region and ERA5 data over EBS region include their own SST values (considered as coarse resolution and not updated data sets), whereas GHRSSST, Medspiration, and NCEP products are described as high-resolution and updated external products used in both study regions. Calibration of the WRF-Hydro model is carried out for two different groups of parameters controlling hydrograph volume and shape in a step-wise approach to improve the performance of the WRF-Hydro model further. The main findings of this study are listed as follows:

- Using higher spatio-temporal SST products (Medspiration and GHRSSST) is highly influential in capturing the temporal and spatial variability of precipitation in small catchments. This effect is variable from region to region.
- For operational forecasts of extreme events, higher spatio-temporal products should be used to improve the accuracy of the runoff.
- High spatio-temporal resolution SST update impact on simulated hydrograph over both regions is highlighted in terms of predicting peak discharge values more accurately by their effect of changing precipitation spatial distribution, and intensity.
- Calibration of the model further improved the model statistical measures for simulated hydrographs over the EBS region, and it was observed that the hydrographs simulated over the MED region are way more sensitive to the calibration, especially in terms of peak timing and magnitude, though the statistical measures were degraded in the falling limb part of the hydrographs.
- The effect of calibrated parameters on statistics improvement was found slightly better than the SST effect over the EBS region, while over the MED region, both SST and calibration effects were found prominent in terms of hydrograph improvement capacity.
- Improvements acquired from different SST products with various spatio-temporal resolution vary. Overall, high-resolution GHRSSST and Medspiration show more significant improvement compared to other SST datasets to capture peak discharge timing and magnitude for hydrographs simulated over both regions.

Overall, the findings of this study from the precipitation and hydrograph simulations demonstrate the potential benefit of using high-resolution SST datasets in initial and lower boundary conditions of the WRF model simulations. Under the consideration of abnormal SST changes exacerbated by changing climate, time-varying SST features characterized with high spatio-temporal accuracy should be accounted for extreme weather event evaluations in complex coastal topographical regions. Additionally, the effect of the WRF-Hydro model calibration on simulated hydrographs displays satisfactory enhancement. Such improvements are considered noteworthy in terms of early warning systems, especially regions under the significant influence of sea effect in atmospheric conditions and have a complex topographical characteristic that poses high flood risk.

## ACKNOWLEDGEMENTS

620 We thank to GDM and SHW of Turkey for the ground-based meteorological observation and  
621 streamflow datasets, The Group for High-Resolution Sea Surface Temperature for GHR SST  
622 datasets, CERSAT for Medspiration datasets, NCEP for RTG\_SST datasets and GFS  
623 products, ECMWF for ERA5 products, and NCAR/NOAA for serving and improving the  
624 WRF model. This project is funded by The Scientific and Technological Research Council of  
625 Turkey (TUBITAK) [Grant Number 116Y193].

## 626 DATA AVAILABILITY

627 All of the datasets used in the current study can be found in the following open-source online  
628 data repositories:

629 GFS dataset from <https://rda.ucar.edu/datasets/ds084.1/> at NCAR. ERA5 dataset from [https://](https://cds.climate.copernicus.eu/#!/search?text=ERA5&type=dataset)  
630 [cds.climate.copernicus.eu/#!/search?text=ERA5&type=dataset](https://cds.climate.copernicus.eu/#!/search?text=ERA5&type=dataset) at ECMWF.

631 GHR SST dataset from [https://podaac.jpl.nasa.gov/dataset/OISST\\_UHR\\_NRT-GOS-L4-](https://podaac.jpl.nasa.gov/dataset/OISST_UHR_NRT-GOS-L4-MED-v2.0?ids=TemporalResolution&values=Daily&search=l4)  
632 [MED-v2.0?ids=TemporalResolution&values=Daily&search=l4](https://podaac.jpl.nasa.gov/dataset/OISST_UHR_NRT-GOS-L4-MED-v2.0?ids=TemporalResolution&values=Daily&search=l4) for Mediterranean and  
633 [https://podaac.jpl.nasa.gov/dataset/OISST\\_UHR\\_NRT-GOS-L4-BLK-v2.0?](https://podaac.jpl.nasa.gov/dataset/OISST_UHR_NRT-GOS-L4-BLK-v2.0?ids=TemporalResolution&values=Daily&search=l4)  
634 [ids=TemporalResolution&values=Daily&search=l4](https://podaac.jpl.nasa.gov/dataset/OISST_UHR_NRT-GOS-L4-BLK-v2.0?ids=TemporalResolution&values=Daily&search=l4) for Eastern Black Sea regions at NASA  
635 Physical Oceanography Distributed Active Archive Center (PODAAC).

636 Medspiration dataset from <http://cersat.ifremer.fr/thematic-portals/projects/medspiration> at  
637 CERSAT.

638 NCEP dataset from [ftp://polar.ncep.noaa.gov/pub/history/sst/rtg\\_high\\_res](ftp://polar.ncep.noaa.gov/pub/history/sst/rtg_high_res) at NCEP.

## 639 REFERENCES

640 Allen, M. R., & Ingram, W. J. (2002). Constraints on future changes in climate and the  
641 hydrologic cycle. *Nature.Com*, 419(September), 225–232. Retrieved from  
642 <https://www.nature.com/articles/nature01092>

643 Baltaci, H. (2017). Meteorological analysis of flash floods in Artvin (NE Turkey) on 24  
644 August 2015. *Natural Hazards and Earth System Sciences*, 17(7), 1221–1230.  
645 <https://doi.org/10.5194/nhess-17-1221-2017>

646 CERSAT. (2012). Medspiration. Retrieved January 7, 2021, from  
647 <http://cersat.ifremer.fr/thematic-portals/projects/medspiration>

648 Chen, W., Wang, L., Xue, Y., & Sun, S. (2009). Variabilities of the spring river runoff  
649 system in East China and their relations to precipitation and sea surface temperature.  
650 *International Journal of Climatology*, 29(10), 1381–1394.  
651 <https://doi.org/10.1002/joc.1785>

652 Duzenli, E., Yucel, I., Pilatin, H., & Yilmaz, M. T. (2020). Evaluating the performance of a  
653 WRF initial and physics ensemble over Eastern Black Sea and Mediterranean regions in  
654 Turkey. *Atmospheric Research*, 248(August 2020), 105184.  
655 <https://doi.org/10.1016/j.atmosres.2020.105184>

656 ECMWF. (2020). ERA5. Retrieved January 7, 2021, from  
657 <https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>

658 Eris, E., & Agiralioglu, N. (2018). Streamflow map of the Eastern Black Sea Region, Turkey.  
659 *Scientia Iranica*, 25(3A), 1048–1056. <https://doi.org/10.24200/sci.2017.4261>

660 Gochis, D. J., Barlage, M., Cabell, R., Casali, M., Dugger, A., FitzGerald, K., ... Zhang, Y.  
661 (2020). The WRF-Hydro modeling system technical description, (Version 5.1.1). *NCAR*  
662 *Technical Note*, 107. Retrieved from  
663 [https://ral.ucar.edu/sites/default/files/public/projects/wrf\\_hydro/technical-description-](https://ral.ucar.edu/sites/default/files/public/projects/wrf_hydro/technical-description-user-guide/wrf-hydro-v5.1.1-technical-description.pdf)  
664 [user-guide/wrf-hydro-v5.1.1-technical-description.pdf](https://ral.ucar.edu/sites/default/files/public/projects/wrf_hydro/technical-description-user-guide/wrf-hydro-v5.1.1-technical-description.pdf)

665 Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the  
666 mean squared error and NSE performance criteria: Implications for improving  
667 hydrological modelling. *Journal of Hydrology*, 377(1–2), 80–91.  
668 <https://doi.org/10.1016/j.jhydrol.2009.08.003>

669 Gurer, I. (1998). Flood Disasters and Preventative Measures in Turkey. *Journal of Natural*  
670 *Disaster Science*, 20(1), 1–9.

671 Gurer, I., & Ucar, I. (2009). Flood Disasters' Inventory in Turkey in 2009. *Eleventh*  
672 *International Symposium on Water Management and Hydraulic Engineering*,  
673 (September), 371–380.

674 Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., ...  
675 Kanae, S. (2013). Global flood risk under climate change. *Nature Climate Change*, 3(9),  
676 816–821. <https://doi.org/10.1038/nclimate1911>

- 677 Kaya, O. F., Guler, M., Altan, A., & Yorganci, I. (2019). 15-16-17 ARALIK 2018  
678 TARİHİNLERİNDE ANTALYA İLİ KEMER/OVACIK MEVKİİNDE GÖRÜLEN  
679 AŞIRI YAĞIŞ HADİSESİNİN AKTÜEL HAVA HARİTALARI, UZAKTAN  
680 ALGILAMA VE SAYISAL HAVA TAHMİN ÜRÜNLERİ İLE SİNOPTİK VE  
681 HİDROLOJİK ANALİZİ. 10. *Ulusal Hidroloji Kongresi 9-12*, (November).
- 682 Kerandi, N., Arnault, J., Laux, P., Wagner, S., Kithaka, J., & Kunstmann, H. (2018). Joint  
683 atmospheric-terrestrial water balances for East Africa: a WRF-Hydro case study for the  
684 upper Tana River basin. *Theoretical and Applied Climatology*, 131(3–4), 1337–1355.  
685 <https://doi.org/10.1007/s00704-017-2050-8>
- 686 Kunstmann, H., & Stadler, C. (2005). High resolution distributed atmospheric-hydrological  
687 modelling for Alpine catchments. *Journal of Hydrology*, 314(1–4), 105–124.  
688 <https://doi.org/10.1016/j.jhydrol.2005.03.033>
- 689 Lebeaupin, C., Ducrocq, V., & Giordani, H. (2006). Sensitivity of torrential rain events to the  
690 sea surface temperature based on high-resolution numerical forecasts. *Journal of*  
691 *Geophysical Research Atmospheres*, 111(12), 1–19.  
692 <https://doi.org/10.1029/2005JD006541>
- 693 Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from  
694 spaceborne elevation data. *Eos*, 89(10), 93–94. <https://doi.org/10.1029/2008EO100001>
- 695 Liu, Y., Liu, J., Li, C., Yu, F., Wang, W., & Qiu, Q. (2020). Parameter Sensitivity Analysis  
696 of the WRF-Hydro Modeling System for Streamflow Simulation: a Case Study in Semi-  
697 Humid and Semi-Arid Catchments of Northern China. *Asia-Pacific Journal of*  
698 *Atmospheric Sciences*. <https://doi.org/10.1007/s13143-020-00205-2>
- 699 McCabe, G. J., & Wolock, D. M. (2008). Joint variability of global runoff and global sea  
700 surface temperatures. *Journal of Hydrometeorology*, 9(4), 816–824.  
701 <https://doi.org/10.1175/2008JHM943.1>
- 702 Miglietta, M. M., Mazon, J., Motola, V., & Pasini, A. (2017). Effect of a positive Sea Surface  
703 Temperature anomaly on a Mediterranean tornadic supercell. *Scientific Reports*, 7(1), 1–  
704 8. <https://doi.org/10.1038/s41598-017-13170-0>
- 705 Moriasi, D. N., Arnold, J. G., Liew, M. W. Van, Bingner, R. L., Harmel, R. D., & Veith, T.



706 L. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in  
 707 Watershed Simulation. *Colombia Medica*, 50(3), 885–900. <https://doi.org/10.1234/590>

708 Naabil, E., Lamptey, B. L., Arnault, J., Kunstmann, H., & Olufayo, A. (2017). Water  
 709 resources management using the WRF-Hydro modelling system: Case-study of the Tono  
 710 dam in West Africa. *Journal of Hydrology: Regional Studies*, 12(June), 196–209.  
 711 <https://doi.org/10.1016/j.ejrh.2017.05.010>

712 NCEP, & NOAA. (2014). NCEP SST Analysis. Retrieved January 7, 2021, from  
 713 <https://polar.ncep.noaa.gov/sst/ophi/verification.shtml>

714 NOAA. (2015). NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive. Retrieved  
 715 from Research Data Archive at the National Center for Atmospheric Research,  
 716 Computational and Information Systems Laboratory website:  
 717 <https://doi.org/10.5065/D65D8PWK>

718 Pilatin, H. (2020). *SENSITIVITY AND SEA SURFACE TEMPERATURE ANALYSES OF*  
 719 *WRF MODEL FOR PREDICTING HEAVY RAINFALL EVENTS OBSERVED IN*  
 720 *EASTERN BLACK SEA AND MEDITERRANEAN REGIONS OF TURKEY*. Middle East  
 721 Technical University.

722 Senatore, A., Furnari, L., & Mendicino, G. (2020). Impact of high-resolution sea surface  
 723 temperature representation on the forecast of small Mediterranean catchments’  
 724 hydrological responses to heavy precipitation. *Hydrology and Earth System Sciences*,  
 725 24(1), 269–291. <https://doi.org/10.5194/hess-24-269-2020>

726 Senatore, A., Mendicino, G., Gochis, D. J., Yu, W., Yates, D. N., & Kunstmann, H. (2015).  
 727 Fully coupled atmosphere-hydrology simulations for the central Mediterranean: Impact  
 728 of enhanced hydrological parameterization for short and long time scales. *Journal of*  
 729 *Advances in Modeling Earth Systems*, 7(4), 1693–1715.  
 730 <https://doi.org/10.1002/2015MS000510>

731 Senatore, A., Mendicino, G., Knoche, H. R., & Kunstmann, H. (2014). Sensitivity of modeled  
 732 precipitation to sea surface temperature in regions with complex topography and  
 733 coastlines: A case study for the Mediterranean. *Journal of Hydrometeorology*, 15(6),  
 734 2370–2396. <https://doi.org/10.1175/JHM-D-13-089.1>

- Shih, D. S., Chen, C. H., & Yeh, G. T. (2014). Improving our understanding of flood forecasting using earlier hydro-meteorological intelligence. *Journal of Hydrology*, 512, 470–481. <https://doi.org/10.1016/j.jhydrol.2014.02.059>
- Silver, M., Karnieli, A., Ginat, H., Meiri, E., & Fredj, E. (2017). An innovative method for determining hydrological calibration parameters for the WRF-Hydro model in arid regions. *Environmental Modelling and Software*, 91, 47–69. <https://doi.org/10.1016/j.envsoft.2017.01.010>
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Zhiquan, L., Berner, J., ... Huang, X.-Y. (2019). A Description of the Advanced Research WRF Model Version 4. *NCAR Technical Note NCAR/TN-475+STR*. Retrieved from <http://library.ucar.edu/research/publish-technote>
- Sun, M., Li, Z., Yao, C., Liu, Z., Wang, J., Hou, A., ... Liu, M. (2020). Evaluation of flood prediction capability of the WRF-hydro model based on multiple forcing scenarios. *Water (Switzerland)*, 12(3). <https://doi.org/10.3390/w12030874>
- Team GHR SST. (2010a). Black Sea Ultra High Resolution SST L4 Analysis 0.01 deg Resolution | PO.DAAC / JPL / NASA. Retrieved January 7, 2021, from [https://podaac.jpl.nasa.gov/dataset/OISST\\_UHR\\_NRT-GOS-L4-BLK-v2.0?ids=TemporalResolution&values=Daily&search=l4](https://podaac.jpl.nasa.gov/dataset/OISST_UHR_NRT-GOS-L4-BLK-v2.0?ids=TemporalResolution&values=Daily&search=l4)
- Team GHR SST. (2010b). Mediterranean Sea Ultra High Resolution SST L4 Analysis 0.01 deg Resolution | PO.DAAC / JPL / NASA. Retrieved January 7, 2021, from [https://podaac.jpl.nasa.gov/dataset/OISST\\_UHR\\_NRT-GOS-L4-MED-v2.0?ids=TemporalResolution&values=Daily&search=l4](https://podaac.jpl.nasa.gov/dataset/OISST_UHR_NRT-GOS-L4-MED-v2.0?ids=TemporalResolution&values=Daily&search=l4)
- Trenberth, K. E. (1999). Conceptual framework for changes of extremes of the hydrological cycle with climate change. *Climatic Change*, 42, 327–339.
- Turkes, M. (1996). Spatial and temporal analysis of annual rainfall variations in Turkey. *International Journal of Climatology*, 16(9), 1057–1076. [https://doi.org/10.1002/\(SICI\)1097-0088\(199609\)16:9<1057::AID-JOC75>3.0.CO;2-D](https://doi.org/10.1002/(SICI)1097-0088(199609)16:9<1057::AID-JOC75>3.0.CO;2-D)
- Turkes, M. (1999). Vulnerability of Turkey to desertification with respect to precipitation and aridity conditions. *Turkish Journal of Engineering and Environmental Sciences*, 23(5),

764 363–380.

765 Varlas, G., Anagnostou, M. N., Spyrou, C., Papadopoulos, A., Kalogiros, J., Mentzafou, A.,  
766 ... Katsafados, P. (2019). A multi-platform hydrometeorological analysis of the flash  
767 flood event of 15 November 2017 in Attica, Greece. *Remote Sensing*, 11(1).  
768 <https://doi.org/10.3390/rs11010045>

769 Wang, J., Wang, C., Rao, V., Orr, A., Yan, E., & Kotamarthi, R. (2019). A parallel workflow  
770 implementation for PEST version 13.6 in high-performance computing for WRF-Hydro  
771 version 5.0: a case study over the midwestern United States. *Geoscientific Model*  
772 *Development*, 12(8), 3523–3539. <https://doi.org/10.5194/gmd-12-3523-2019>

773 Wehbe, Y., Temimi, M., Weston, M., Chaouch, N., Branch, O., Schwitalla, T., ... Al  
774 Mandous, A. (2019). Analysis of an extreme weather event in a hyper-arid region using  
775 WRF-Hydro coupling, station, and satellite data. *Natural Hazards and Earth System*  
776 *Sciences*, 19(6), 1129–1149. <https://doi.org/10.5194/nhess-19-1129-2019>

777 Yang, Y., Yuan, H., & Yu, W. (2018). Uncertainties of 3D soil hydraulic parameters in  
778 streamflow simulations using a distributed hydrological model system. *Journal of*  
779 *Hydrology*, 567(June), 12–24. <https://doi.org/10.1016/j.jhydrol.2018.09.042>

780 Yucel, I., & Onen, A. (2014). Evaluating a mesoscale atmosphere model and a satellite-based  
781 algorithm in estimating extreme rainfall events in northwestern Turkey. *Natural*  
782 *Hazards and Earth System Sciences*, 14(3), 611–624. [https://doi.org/10.5194/nhess-14-](https://doi.org/10.5194/nhess-14-611-2014)  
783 611-2014

784 Yucel, I., Onen, A., Yilmaz, K. K., & Gochis, D. J. (2015). Calibration and evaluation of a  
785 flood forecasting system: Utility of numerical weather prediction model, data  
786 assimilation and satellite-based rainfall. *Journal of Hydrology*, 523, 49–66.  
787 <https://doi.org/10.1016/j.jhydrol.2015.01.042>

788

789 **TABLES**

790 Table 1 Drainage areas and calibrated event periods of each selected basin over EBS and  
791 MED regions.

Region	Station	Drainage Area (km <sup>2</sup> )	Calibration Event Period	
			Start	End
EBS	D22A049	175.8	08/27/2016	09/06/2016
			09/20/2017	09/30/2017
			10/19/2016	10/29/2016
	D22A079	85.8	10/19/2016	10/29/2016
			10/01/2018	01/11/2018
			06/24/2019	07/04/2019
	D22A089	71.5	08/27/2016	09/06/2016
			09/20/2017	09/30/2017
			10/19/2016	10/29/2016
	D22A147	41.9	08/27/2016	09/06/2016
			09/20/2017	09/30/2017
			10/19/2016	10/29/2016
MED	D08A071	98.3	01/09/2015	01/19/2015
			03/07/2017	03/17/2017
			03/23/2015	04/02/2015
	E08A008	164.6	01/09/2015	01/19/2015
			03/07/2017	03/17/2017
			03/23/2015	04/02/2015
	D09A095	164.6	01/21/2014	01/31/2014
			01/09/2015	01/19/2015
			03/23/2015	04/02/2015

792

793

794

795

796 Table 2 SST products and initial boundaries included as meteorological forcings in the scope  
 797 of this study and model run periods corresponding to EBS and MED region.

Region	Meteorological Forcings		Model Run Periods	
	SST Products	Initial and Boundary Conditions	Start Date	End Date
EBS	ERA5			
	GHR Medspiration	ERA5 Reanalysis	08/17/2015	08/27/2015
	NCEP			
MED	GFS			
	GHR Medspiration	GFS Forecast	12/10/2018	12/20/2018
	NCEP			

798 Table 3 Average spatial and temporal cross correlations of SST products over the study  
799 regions and periods

Average Cross Correlations	Spatial		Temporal	
	EBS	MED	EBS	MED
ERA5/GFS	0.75	0.11	-	-
GHRSSST	0.83	0.39	0.60	0.84
NCEP	0.48	0.23	0.24	0.79
MED	0.73	0.35	0.48	0.79

800

801

802

803 Table 4 Average statistics of (Bias, Root Mean Square Error (RMSE), and Correlation  
804 Coefficient (RR)) calibrated parameters for three events compare to default parameter set for  
805 D22A049 and D22A147 basins over EBS region.

	D22A049				D22A147			
	Parm. Values	Bias	RMSE	RR	Parm. Values	Bias	RMSE	RR
Default Parameter Set		4.24	40.55	0.13		0.48	5.75	0.38
REFKDT	0.5	3.72	40.48	0.38	0.5	0.58	3.20	0.63
RETDEPRT	0.0	4.00	40.45	0.39	0.0	0.60	3.18	0.62
SLOPE	0.1	4.00	40.45	0.39	1.0	1.01	2.88	0.67
OVROUGHRTFAC	1.0	4.00	40.45	0.39	1.0	1.01	2.88	0.67
MANN	2.0	3.69	37.54	0.39	2.0	0.85	2.76	0.64
LKSATFAC	10	-2.34	32.16	0.56	10	0.55	2.33	0.71

806

807

Table 5 Average statistics of (Bias, Root Mean Square Error (RMSE), and Correlation Coefficient (RR)) calibrated parameters for three events compare to default parameter set for D08A071, D09A095 and E08A008 basins over MED region.

	D08A071				D09A095			
	Parm. Values	Bias	RMSE	RR	Parm. Values	Bias	RMSE	RR
Default Parameter Set		-5.28	16.67	0.44		2.78	17.02	0.45
REFKDT	0.5	-1.02	30.12	0.44	0.5	1.31	9.67	0.73
RETDEPRT	0.0	-0.47	30.53	0.44	0.0	5.28	16.58	0.42
SLOPE	0.1	-0.47	30.53	0.44	1.0	5.48	15.65	0.48
OVROUGHRTFA C	1.0	-0.47	30.53	0.44	0.1	1.69	8.55	0.70
MANN	2.0	-0.50	29.85	0.49	2.0	1.70	8.35	0.81
LKSATFAC	10	-5.57	26.30	0.46	10	2.29	9.02	0.77
	E08A008							
	Parm. Values	Bias	RMSE	RR				
Default Parameter Set		12.22	15.98	0.25				
REFKDT	0.5	11.81	15.38	0.39				
RETDEPRT	0.0	11.80	15.35	0.39				
SLOPE	0.1	11.80	15.35	0.39				
OVROUGHRTFA C	1.0	11.80	15.35	0.39				
MANN	0.5	11.84	15.19	0.37				
LKSATFAC	10	2.52	4.23	0.31				



813    Table 6 Default and calibrated parameter values for each basin.

Parameter	Default Parameter Value	Calibrated Parameter Value						
		EBS				MED		
		D22A049	D22A079	D22A089	D22A147	D08A071	D09A095	E08A008
REDKT	3.0	0.5	0.5	0.5	0.5	0.5	0.5	0.5
RETDEPRTFAC	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
SLOPE	0.1	0.1	0.1	1.0	1.0	0.1	1.0	0.1
OVROUGHRTFAC	1.0	1.0	1.0	0.1	1.0	1.0	0.1	1.0
MANN	1.0	2.0	2.0	2.0	2.0	2.0	2.0	0.5
LKSATFAC	1000	10	10000	1000	10	10	10	10

815 Table 7 Statistics of Bias, RMSE, and RR between observed and modelled precipitations with  
816 different SST datasets a) ERA5, GHRSSST, Medspiration, and NCEP for D22A147 over EBS  
817 and b) GFS, GHRSSST, Medspiration, and NCEP for D08A071 over MED are shown.

Station	SST WRF Runs	Bias	RMSE	RR
D22A147	ERA5-SST	-0.54	3.19	0.03
	GHR-SST	-0.06	5.30	0.01
	MED-SST	-0.24	3.55	0.03
	NCEP-SST	-0.54	3.38	0.01
D08A071	GFS-SST	0.56	3.45	0.60
	GHR-SST	0.18	2.35	0.52
	MED-SST	0.13	1.86	0.67
	NCEP-SST	0.33	2.23	0.60

818

819

820 Table 8 Statistics of Bias, RMSE, and RR between observed and modelled hydrographs of  
821 D22A147 and D08A071 for SST events over EBS and MED regions.

Station	SST WRF-Hydro Runs	Default Parameter Set			Calibrated Parameter Set		
		Bias	RMSE	RR	Bias	RMSE	RR
D22A147	ERA5-SST	-10.46	20.69	0.29	-9.92	20.13	0.42
	GHR-SST	-8.71	16.49	0.83	-6.16	10.82	0.83
	MED-SST	-10.24	20.13	0.86	-8.32	15.86	0.86
	NCEP-SST	-10.42	20.55	0.83	-9.60	18.98	0.82
D08A071	GFS-SST	-7.07	125.97	0.30	-24.98	128.81	0.18
	GHR-SST	-26.56	57.30	0.62	-42.73	83.25	0.30
	MED-SST	-26.24	59.70	0.59	-43.63	83.57	0.31
	NCEP-SST	-18.85	58.79	0.60	-40.76	81.50	0.32

822

## 823 **FIGURE LEGENDS**

824 Figure 1 The outer and nested domains (d01 and d02) of the WRF model for EBS and MED  
825 regions are displayed in the top-left. Boundaries of the selected basin, their outlet points  
826 (stream gauge stations denoted as blues dots), channel network grids in the WRF-Hydro  
827 model, and the meteorological station (denoted as a green triangle) are shown in the zoomed  
828 maps with the high-resolution topography layer at the background.

829

830 Figure 2 Temporally averaged spatial distribution of SST products over MED (left column)  
831 and EBS (right column) regions.

832

833 Figure 3 Spatially-averaged temporal distribution of SST products over EBS (upper panel)  
834 and MED (lower panel) regions.

835

836 Figure 4 Calibration results of the selected WRF-Hydro model parameters, namely REFKDT,  
837 RETDEPRT, SLOPE, OVROUGHRTFAC, MANN, and LKSATFAC: a-f) left column for  
838 event occurred between 10/19/2019 to 10/29/2016 and basin D22A049 located over EBS  
839 region; g-l) right column for event occurred between 03/07/2017 to 03/17/2017 and basin  
840 D08A071 located over MED region. Dashed line shows the hydrograph for selected optimum  
841 parameter value.

842

843 Figure 5 Time series of hourly precipitation that a) D22A147 basin over EBS region receives  
844 during the event occurred in 08/17/2015-08/27/2015 and b) D08A071 basin over MED region  
845 receives during the event occurred in 12/10/2018-12/20/2018 for 10 days. Outputs are  
846 generated from WRF model with the native SST field from ERA5 Reanalysis data (ERA5-  
847 SST) for EBS region and GFS Forecast data (GFS-SST) for MED region with different SST  
848 products: GHRSSST, Medspiration, and NCEP.

849

850 Figure 6 Spatial distribution of daily precipitation at the peak day (08/24/2015) for run period  
851 of 08/17/2017 – 08/27/2017 over EBS region. a) The map at the top shows the interpolated  
852 observed precipitation map obtained from meteorological stations data (green triangles).  
853 Black line indicates the boundaries of selected basins for this study while blue dots show the  
854 corresponding stream gauge stations. The four maps at the sub-panels refer the simulated  
855 precipitations by WRF model derived by different SST data sources for the peak hour: b)  
856 ERA5, c) GHRSSST, d) Medspiration and e) NCEP, respectively.

857

858 Figure 7 Spatial distribution of hourly precipitation at the peak hour (12/16/2018 17:00:00  
859 UTC) for run period of 12/10/2018–12/20/2018 over MED region. a) The map at the top  
860 shows the interpolated observed precipitation map obtained from meteorological stations data  
861 (green triangles). Black line indicates the boundaries of selected basins for this study while  
862 blue dots show the corresponding stream gauge stations. The four maps at the sub-panels  
863 refer the simulated precipitations by WRF model derived by different SST data sources for  
864 the peak hour: b) GFS, c) GHRSSST, d) Medspiration and e) NCEP, respectively.

865

866 Figure 8 Comparison of observed hydrographs with the simulated hydrographs generated  
867 using precipitation inputs derived with native SST field (ERA5), GHRSSST, Medspiration and  
868 NCEP a) prior to the calibration and b) with the calibrated parameter set of the WRF-Hydro  
869 model for event 08/17/2015-08/27/2015 in D22A147.

870

871 Figure 9 Comparison of observed hydrographs with the simulated hydrographs generated  
872 using precipitation inputs derived with native SST field (GFS), GHRSSST, Medspiration and  
873 NCEP a) prior to the calibration and b) with the calibrated parameter set of the WRF-Hydro  
874 model for event 12/10/2018-12/20/2018 in D08A071.

875

876 Figure 10 Overlapped dynamic maps of accumulated precipitation simulated by WRF model  
877 (3-km) operated with 4 different SST datasets (ERA5, GHRSSST, Medspiration and NCEP)  
878 and discharge simulated by WRF-Hydro model (250-m) over EBS region at 08/23/2015  
879 23:00:00, 08/24/2015 03:00:00, and 08/24/2015 04:00:00. Stream gauges are denoted as  
880 blue dots.

881

882 Figure 11 Overlapped dynamic maps of accumulated precipitation simulated by WRF model  
883 (3-km) operated with 4 different SST datasets (GFS, GHRSSST, Medspiration and NCEP) and  
884 discharge simulated by WRF-Hydro model (250-m) over MED region at 12/16/2018  
885 02:00:00, 12/16/2018 16:00:00, and 12/16/2018 19:00:00. Stream gauges are denoted as  
886 blue dots.