

# Improving WRF-Hydro runoff simulations of heavy floods through the sea surface temperature fields with higher spatio-temporal resolution

Berina Mina Kilicarslan<sup>1</sup>  | Ismail Yucel<sup>1</sup>  | Heves Pilatin<sup>2</sup> | Eren Duzenli<sup>1</sup> | Mustafa Tugrul Yilmaz<sup>1</sup>

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

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

#### Correspondence

Ismail Yucel, Department of Civil Engineering, Middle East Technical University, Ankara, Turkey.

Email: iyucel@metu.edu.tr

#### Funding information

The Scientific and Technological Research Council of Turkey (TUBITAK), Grant/Award Number: 116Y193

## Abstract

This study investigates the impact of the spatio-temporal accuracy of four different sea surface temperature (SST) datasets on the accuracy of the Weather Research and Forecasting (WRF)-Hydro system to simulate hydrological response during two catastrophic flood events over the Eastern Black Sea (EBS) and the Mediterranean (MED) regions of Turkey. Three time-variant and high spatial resolution external SST products (GHRSST, Medspiration and NCEP-SST) and one coarse-resolution and time-invariant SST product (ERA5- and GFS-SST for EBS and MED regions, respectively) already embedded in the initial and the boundary conditions datasets of WRF model are used in deriving near-surface atmospheric variables through WRF. After the proper event-based calibration is performed to the WRF-Hydro system using hourly and daily streamflow data in both regions, uncoupled model simulations for independent SST events are conducted to assess the impact of SST-triggered precipitation on simulated extreme runoff. Some localized and temporal differences in the occurrence of the flood events with respect to observations depending on the SST representation are noticeable. SST products represented with higher cross-correlations (GHRSST and Medspiration) revealed significant improvement in flood hydrographs for both regions. The GHRSST dataset shows a substantial improvement in NSE (~70%), RMSE reduction up to 20%, and an increase in correlation from 0.3 to 0.8 with respect to the invariable SST (ERA5) in simulated runoffs over the EBS region. The use of both GHRSST and Medspiration SST data characterized with high spatio-temporal correlation resulted in runoff simulations exactly matching the observed runoff peak of 300 m<sup>3</sup>/s by reducing the overestimation seen in invariable SST (GFS) in the MED region. Improved precipitation simulation skills of the WRF model with the detailed SST representation show that the hydrographs of GHRSST and Medspiration simulations show better performance compared to the simulated hydrographs by observed precipitation.

#### KEY WORDS

calibration, GHRSST, Medspiration, sea surface temperature, WRF, WRF-hydro

## 1 | INTRODUCTION

Warming climate results in increased water vapour input into the atmosphere; consequently, triggering the intensity of rainfall events (Allen & Ingram, 2002; Trenberth, 1999). 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.

Accurate forecasts of the spatial distributions of heavy precipitation events and their hydrological responses are among the most significant elements of an accurate flood forecasting system (Ryu et al., 2017; Shih et al., 2014; Yucel & Onen, 2014). In this context, the applications of a hydrometeorological modelling 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 a well-calibrated accurate hydrological model and accurate short-term predictions of atmospheric variables (e.g., precipitation and temperature) via numerical weather prediction (NWP) models driving the hydrological models.

Selection of the NWP model and the input datasets driving the boundary and the initial conditions have a profound effect on the accuracy of the short-term predictions; hence, better operational flood forecasts clearly require improved NWP-based forecasts (Done et al., 2004; Yucel et al., 2015). Such NWP simulations are particularly influenced by the sea surface temperature (SST) state, as oceans/seas supply a significant amount of both energy and water so that the state of the atmospheric forcing variables is impacted (Cassola et al., 2016; Ferrari et al., 2020). 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 the variability in all elements of the water and the energy balances as well as an accurate parameterization of the land-surface to better translate the input atmospheric forcing datasets into hydrological states and fluxes.

SST primarily affects the heat and the water fluxes at the lower boundary of the atmosphere; hence there is a significant relationship between the variabilities of SST and convective extremes. In general, a higher SST state results in increased moisture content in the air and warming of the lower levels of the atmosphere (Lebeaupin et al., 2006). This often results in stronger convection and higher precipitation totals (Trenberth & Shea, 2005; Vecchi & Harrison, 2002). Overall, lower variability of SST in the order of  $\pm 1$  K may dramatically and nonlinearly change the intensity of the development of supercells over the seas (Miglietta et al., 2017). For example, Roxy (2014) shows that  $1^{\circ}\text{C}$  increase in SST causes  $\sim 2$  mm/day increase in precipitation over monsoon basins. Even if SST effects on long-term simulations are identified as small, it may still significantly affect the individual heavy precipitation events (Senatore et al., 2014; Takahashi & Idenaga, 2013). Accordingly, improved representation of SST fields in the Weather Research and Forecasting (WRF) system has considerable improvements in the simulation of the atmospheric boundary

layer processes and flow dynamics (Senatore, Furnari, et al., 2020), in particular the precipitation forecasts (LaCasse et al., 2008; Zhang et al., 2015).

Given that lower atmospheric boundary conditions often drive the precipitation falling on the land surface, SST variations play a key role in the magnitude of the heavy precipitation intensity (Baltaci, 2017; Bozkurt & Sen, 2011; Turuncoglu, 2015). For example, a gradual increase in SST may cause a sudden amplification of convective precipitation extremes over the coastal regions (Meredith et al., 2015). Also, the rapid and abnormal fluctuations in SSTs contribute to the developments of some mesoscale thunderstorms and extreme convective events (Lebeaupin et al., 2006; Senatore et al., 2014). Accordingly, providing high accuracy SST input to the lower boundary is crucial for accurate precipitation modelling, hence for better flood forecasts through NWP models. Despite its significance and impact on the accuracy of the runoff forecasts, the number of studies inter-comparing the impact of spatio-temporal resolution of different SST input datasets over the accuracy of the predicted runoff has remained limited so far (Chen et al., 2009; McCabe & Wolock, 2008; Senatore, Davolio, et al., 2020; Senatore, Furnari, et al., 2020).

A fully distributed, physical-based, multi-scale hydrometeorological modelling system, the WRF-Hydro is developed by the U.S. National Center for Atmospheric Research (NCAR) to investigate critical water issues, including flash flood forecasting applications. Allowing to run both in uncoupled (one-way from the atmosphere to land) mode and fully-coupled (two-way) mode (Gochis et al., 2020), this modelling system links the atmospheric and the hydrological processes. Overall, WRF-Hydro is designed as a framework to couple WRF with a hydrological extension that enables simulation of land surface states and fluxes, including surface overland flow, saturated subsurface flow, channel routing, and vertical energy fluxes between land and atmosphere through physics-based and conceptual approaches. The modelling system has been used in many studies investigating the model performance and application worldwide, including studies focusing on flood simulation, water balance and water management (Kerandi et al., 2018; Li et al., 2017; Naabil et al., 2017; Wehbe et al., 2019).

Despite the studies compare the impact of different sources of precipitation input (i.e., comparing observed and simulated) on runoff simulation and agree that further improvement in the precipitation simulation skills is still needed (Givati et al., 2016; Senatore et al., 2015; Sun et al., 2020), not many studies have investigated the impact of the spatio-temporal resolution of various SST sources over the runoff predictions of WRF-Hydro modelling system via the improvements in the simulated precipitation. Among them, studies that utilized high-resolution SST inputs and implemented parameter calibration in the prediction of runoff have particularly remained limited with the study of Senatore, Furnari, et al. (2020) and Senatore, Davolio, 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 induced by the meteorological,

hydrological and topographical effects. The Eastern Black Sea (EBS) and the Mediterranean (MED) regions of Turkey are among the most vulnerable regions in terms of flood risk in the Anatolian peninsula (Duzenli et al., 2020; Gurer, 1998; Gurer & Ucar, 2009). Forecasting the floods through high-resolution NWP models in EBS and MED regions is critical (Camera et al., 2020), where a gradual increase in SST may cause sudden amplification of convective precipitation extremes over the coastal regions (Meredith et al., 2015) and SST variations play a key role in heavy precipitation events in the Anatolian Peninsula (Baltaci, 2017; Bozkurt & Sen, 2011; Turuncoglu, 2015). The SST forcing is usually considered as a crucial component for the heavy precipitation incidents, particularly over MED region (Robinson et al., 2012; Rebora et al., 2013; Senatore et al., 2014; Cisneros et al., 2016; Jee & Kim, 2017; Pastor et al., 2018; Furnari et al., 2018; Senatore, Furnari, et al. (2020)). Volosciuk et al. (2016) found that Mediterranean SST also impacts the extreme precipitations and floods in Central Europe.

Accordingly, the main goal of this study is to (1) evaluate the impact of the SST products of different resolutions on the simulations of extreme precipitation causing significant floods over the catchments of MED and EBS coastal regions, (2) investigate the impact of the SST products represented at different resolutions on the accuracy of the storm hydrographs derived using WRF-Hydro after a proper calibration, (3) investigate the effect of SST sensitivity over different geographic regions as EBS and MED represented with humid and semi-arid climates, respectively.

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 (GHRSST, Medspiration, NCEP and ERA5/GFS), while WRF-Hydro parameters responsible for hydrological processes are calibrated by using twenty-one events over MED and EBS regions. Four sub-basins in EBS and three sub-basins in MED regions are selected, while three events for each sub-basin are used to perform the calibration. Both SST-driven WRF precipitation and WRF-Hydro runoff simulations are validated against ground station-based observations for two catastrophic flood events that occurred on 24 August 2015 over EBS and 16 December 2018 over the MED regions. Thereby, the impact of the sensitivity of SST products over the accuracy of the WRF-Hydro model predictability is assessed through generated both precipitation and runoff estimates.

## 2 | DATA AND METHODS

### 2.1 | Study area and event description

This study focuses on the floods generated by two SST-triggered heavy precipitation events occurred in the catchments of EBS and MED regions that are characterized by different climate. Nested 3-km WRF domains (d02) covering the EBS and MED regions, selected basins together with their channel networks, location of

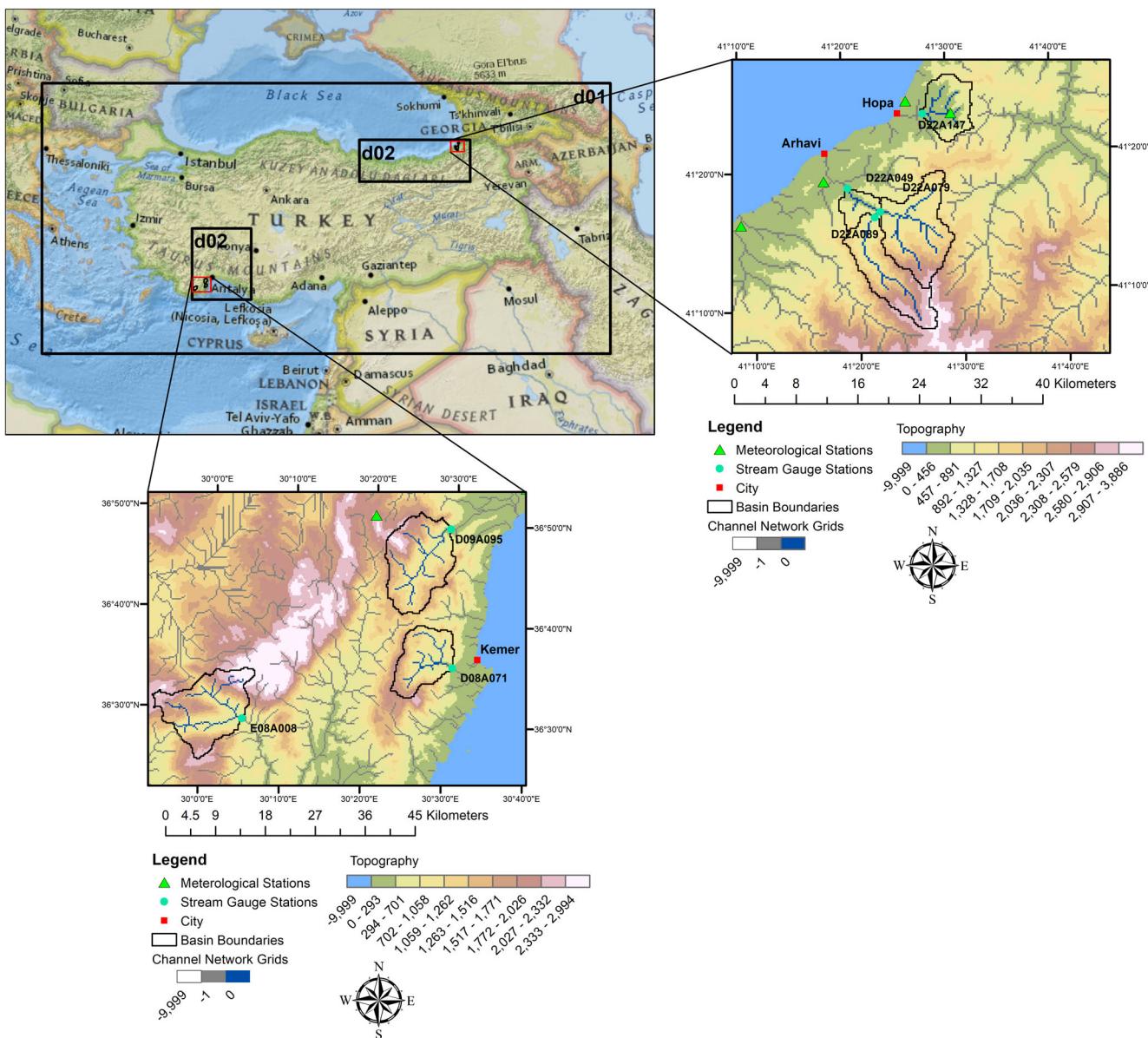
both meteorological and stream gauge stations are shown in Figure 1.

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. Due to small basin structures and steep rocky characteristics, river systems react quickly to moderate precipitation events and may cause flash floods (Eris & Agiralioglu, 2018; Gurer & Ucar, 2009). 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 over Turkey (Baltaci, 2017).

MED region has a typical Mediterranean climate prevailing humid and semi-humid subtropical characteristics with a rainy winter/spring and a severe hot-dry summer (Turkes, 1996). 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 points out that the strong orographic lifting dependency and elevated heat sources for convective initiation exist over both regions. Since high SST increases the moisture content in the air, it has a critical role in the occurrence of flood events over such regions located in coastal areas with complex topography.

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 a total of 135 mm of precipitation accumulated within 24-h. On the other hand, for the MED event that occurred on 16 December 2018, the peak hourly precipitation was recorded as 53.1 mm at Antalya-Ovacik station, while it received a total daily precipitation amount of 651.7 mm. This event was registered as the highest precipitation record measured in Turkey (Kaya et al., 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 a typical mesoscale convective signature, whereas the frontal system is dominant for the event that occurred over the MED region during the winter season.

Runoff observations obtained at stations D22A049, D22A079, D22A089 (all three in Arhavi province) and D22A147 (in Hopa province) over the EBS region and stations D08A071, D09A095 and E08A008 over MED region are utilized for calibration and validation in this study (Figure 1 and Table 1). The model simulations are performed for 10-days for each calibration event in Table 1 to allow the model to spin-up. Average streamflow observations ( $m^3/s$ ) collected by the State Hydraulic Works (SHW) of Turkey are provided at daily timesteps at five of the stream gauge stations (D22A079, D22A089 and D22A147 over EBS region; D09A095 and E08A008 over MED region) and hourly timesteps at two of the stations (D22A049 over EBS and D08A071 over MED). However, the hourly streamflow data at these two stations are limited and only available after 2016.



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

## 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), as shown in Figure 1. The outer domain extends between longitudes 23.5°E and 47.5°E and latitudes 34.5°N and 43.5°N, where this outer domain contains 232 × 111 grid points. The inner domain over the MED region is placed between latitudes 47.5°N and 32.4°N and longitudes 34.5°E

and 36.4°E, where this inner domain contains 73 × 88 grid points. The inner domain over the EBS region is placed between latitudes 47.5°N and 41.6°N and longitudes 23.5°E and 36.9°E, where this inner domain contains 136 × 52 grid points.

In a sensitivity study of simulating extreme precipitation events using the WRF model, Duzenli et al. (2020) found that the Global Forecasting System (GFS) initial and boundary datasets over the MED region was the most representative, while The European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 (ECMWF, 2020; NOAA, 2015) provide the most representative dataset over the EBS region. Therefore, these two datasets are also used in this current study for each region.

**TABLE 1** Drainage areas and calibrated event periods of each selected basin over EBS and 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	164.5	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

**TABLE 2** SST products and initial boundaries included as meteorological forcings in the scope 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 NCEP	ERA5 Reanalysis Medspiration NCEP	08/17/2015	08/27/2015
MED	GFS GHR Medspiration NCEP	GFS Forecast Medspiration NCEP	12/10/2018	12/20/2018

Two different types of simulations are performed: one using the SST fields constant as extracted from ERA5 and GFS (for EBS and MED regions, respectively) defined at the initial time step and remained unchanged throughout the model simulation (i.e., time-invariant SST option in the WRF model setup) and another simulation that enables the daily update of ERA5 and GFS model SST values with higher spatial resolution products during the simulations (i.e., time-variant SST option).

In addition to time-invariant ERA5 and GFS SSTs, three other time-variant 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 (GHRSST) (Team GHRSST, 2010); (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 (Medspiration, GHRSST and NCEP) have high spatial resolutions (0.022°, 0.01° and 0.083°, respectively) and are provided on daily basis. From here on, the SST products used in this study will be referred as Medspiration, GHRSST, 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.

Given there are no buoy observations over the study regions to validate the accuracy of SST products, here it is suggested to use the average cross-correlations as an indicator of true signal (Yilmaz et al., 2014), assuming there is no other common spatial and temporal signal between the products (i.e., higher average cross-correlations imply a better product).

In addition, for the evaluation of the precipitation simulated by different SST products, the root mean squared error (RMSE) and the mean absolute error (MAE) are calculated for the simulated precipitation values for each SST product corresponding to the point observations over rain gauges.

## 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). For the LSM, land cover classifications are provided by the Moderate Resolution Imaging Spectroradiometer (MODIS) Modified IGBP 20-category land cover product, and the soil classifications are derived based on the 1-km STATSGO database (details in the Supporting Information). 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 the 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 et al. (2008) using the WRF-Hydro Pre-Processing toolbox in the GIS environment (Sampson & Gochis, 2015).

In calibration simulations of the WRF-Hydro model, WRF-derived hourly precipitation forcing input is replaced with the interpolated observed precipitation. Observed precipitation maps are created by inverse distance weighting (IDW) method using the point observations. Evaluation of the interpolated precipitation is performed based on the leave-one-out approach. At each step, one rain gauge is excluded from the interpolation, and the interpolated value over the location of excluded rain gauge is determined. This process is repeated for all rain gauges. Statistical measures, including RMSE and MAE, were calculated between the interpolated precipitation values and the point observations over rain gauges.

Based on streamflow data availability in hourly or daily temporal resolution, 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. The calibration is performed using hourly streamflow data for two events in 2016 and 2017

hydrologic years in D22A049 of EBS and one event in 2017 hydrologic year in D08A071 of MED (see Table 1). Daily calibration is applied for all other events in Table 1. Also, for validating SST events, hourly streamflow data is used in MED, while daily streamflow data is used in EBS. Calibration of the model is manually employed with a step-wise approach as described in Yucel et al. (2015). In the first step, parameters controlling the hydrograph volume called infiltration factor (REFKDT), surface retention depth factor (RETDEPRTFAC) and deep drainage coefficient (SLOPE) are calibrated. Surface roughness coefficient factor (OVROUGHRTFAC), channel Manning roughness scaling factor (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. A similar procedure is commonly adopted for the calibration of WRF-Hydro in terms of water balance and its distribution (Liu et al., 2020; Naabil et al., 2017; Senatore et al., 2015; Yang et al., 2018; Yucel et al., 2015). Some parameters (REFKDT, SLOPE and MANN) are defined in tabular value format considered as global values over the domain. Others are defined as pixel specific (RETDEPRTFAC, OVROUGHRT and LKSATFAC) that enables changing parameter value for each pixel over the model domain. In this study, pixel-specific parameters are tuned at the basin level.

Calibration of the REFKDT parameter (default value of 3.0), which impacts the amount of water infiltrating through the soil column, is performed with parameter values between 0.5 and 5.0 with 0.5 increments. RETDEPRTFAC is the multiplier of the retention depth before the flow is routed as overland flow and influences the overland flow amount calculated after ponded water exceeds the predefined water depth. Initially, surface retention depth equals 1 mm within the whole domain, and the parameter (default value of 1.0) is calibrated within the range of 0.0–10.0 with a 1.0 increment. The SLOPE parameter (default value of 0.1) controlling the openness of the bottom soil column to the conceptual bucket is calibrated between the values of 0.1 and 1.0 with 0.3 increments. Similar to Wang et al. (2019), only the first class of the nine SLOPE\_DATA categories represented in GENPARM.TBL is subjected to tuning. OVROUGHRTFAC (default value of 1.0) is calibrated for values ranging from 0.1 to 1.0 with 0.3 increments. It impacts the speed of the infiltration excess water transmitted over the surface until it reaches the channel network grids. Manning's Roughness coefficients for all stream orders (MannN column of the channel parameters table in the Supporting information) are calibrated with a scaling factor (MANN) within a range from 0.5 to 2.0 with 0.5 increments rather than changing coefficients individually. It controls the conveyance time of the flow through the channel network, which can be interpreted as the higher MANN values create a slower peak time and lower hydrograph volume at peak time. Lastly, the LKSATFAC parameter, which affects the lateral redistribution of infiltrated water, is calibrated for the values of 10, 100, 1000 (default) and 10 000.

Statistical measures are implemented between observed and simulated discharge for the model accuracy evaluation, namely bias, root mean square error (RMSE), correlation coefficient (RR), Nash–Sutcliffe efficiency (NSE) and Kling–Gupta efficiency (KGE) 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, NSE and KGE are sensitive to both the shape and the volume of the hydrograph and reflect the overall fit of the hydrograph (Gupta et al., 2009; Moriasi et al., 2007). 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.

After performing the calibration, the calibrated WRF-Hydro model is validated for SST event simulations over EBS and MED regions. At this stage, the WRF-Hydro model is driven by the meteorological forcing datasets generated utilizing the time-invariant ERA5 and GFS SSTs, and three other time-variant external SST products as Medspiration, GHSST and NCEP. In addition, these SST events are also simulated with meteorological forcing data in which WRF-derived precipitation is replaced with observed precipitation.

### 3 | RESULTS

#### 3.1 | Calibration of the WRF-Hydro model

Results for the hourly calibration of selected parameters within the WRF-Hydro model are shown in Figure 2. In this figure, the first column (a-f) represents the calibration results of the event that occurred between 19 October 2016 and 29 October 2016 over the D22A049 basin of the EBS region, while the second column (g-l) belongs to the event that occurred between 07 March 2017 and 17 March 2017 at basin D08A071 located over MED region. Two more additional events belonging to each catchment are also used in the calibration process (Table 1). Table 3 shows the average statistical measures calculated for the WRF-Hydro model set up with default parameter set and for the simulation of a selected parameter value of each catchment considered over EBS region and over MED region, respectively. Since the calibration procedure is carried out in a step-wise approach, hydrographs with 3.0 of REFKDT value represents the default simulation of corresponding basins (Figure 2(a) and (b)).

Figure 2(a) and (g) show the calibration results of D22A049 and D08A071 basins, respectively. It can be inferred as the lower the REFKDT value lower the infiltration capacity of the soil column, in turn, the higher the hydrograph volume. According to the statistics and comparison with the calibration hydrographs based on the other two events, it is decided on to the lowest value (0.5) of REFKDT as optimum for both basins. However, there is still an underestimation observed in the D22A049 hydrograph volume in Figure 2(a). The simulated first peak on day-8 is lowered, and the simulated hydrograph is fed through the observed peak that occurred between day-7 and day-8. On the contrary, when the average bias is calculated for three events, bias turns into  $3.72 \text{ m}^3/\text{s}$  in Table 3. A similar contrast is also observed in the D08A071 station. Negative bias is observed for the average of three events (Table 3), while an overestimation is

observed for the represented event in Figure 2(b). A pronounced increase is noticed in the KGE value (from  $-0.16$  to  $0.47$ ) in D22A147 (Table 3). The overall statistic shows that the REFKDT parameter is strongly sensitive in both regions.

Figure 2(b) and (h) show the calibration results of the RETDEPRTFAC parameter with the range of 0.0–10.0 with 1.0 increment. Simulated hydrographs of both basins are not showing an apparent response to the RETDEPRTFAC parameter (Table 3). Since EBS and MED regions have steep topography, little water accumulation over the terrain is expected to be observed. Therefore, the optimum RETDEPRTFAC parameter value is selected as 0.0 for both basins.

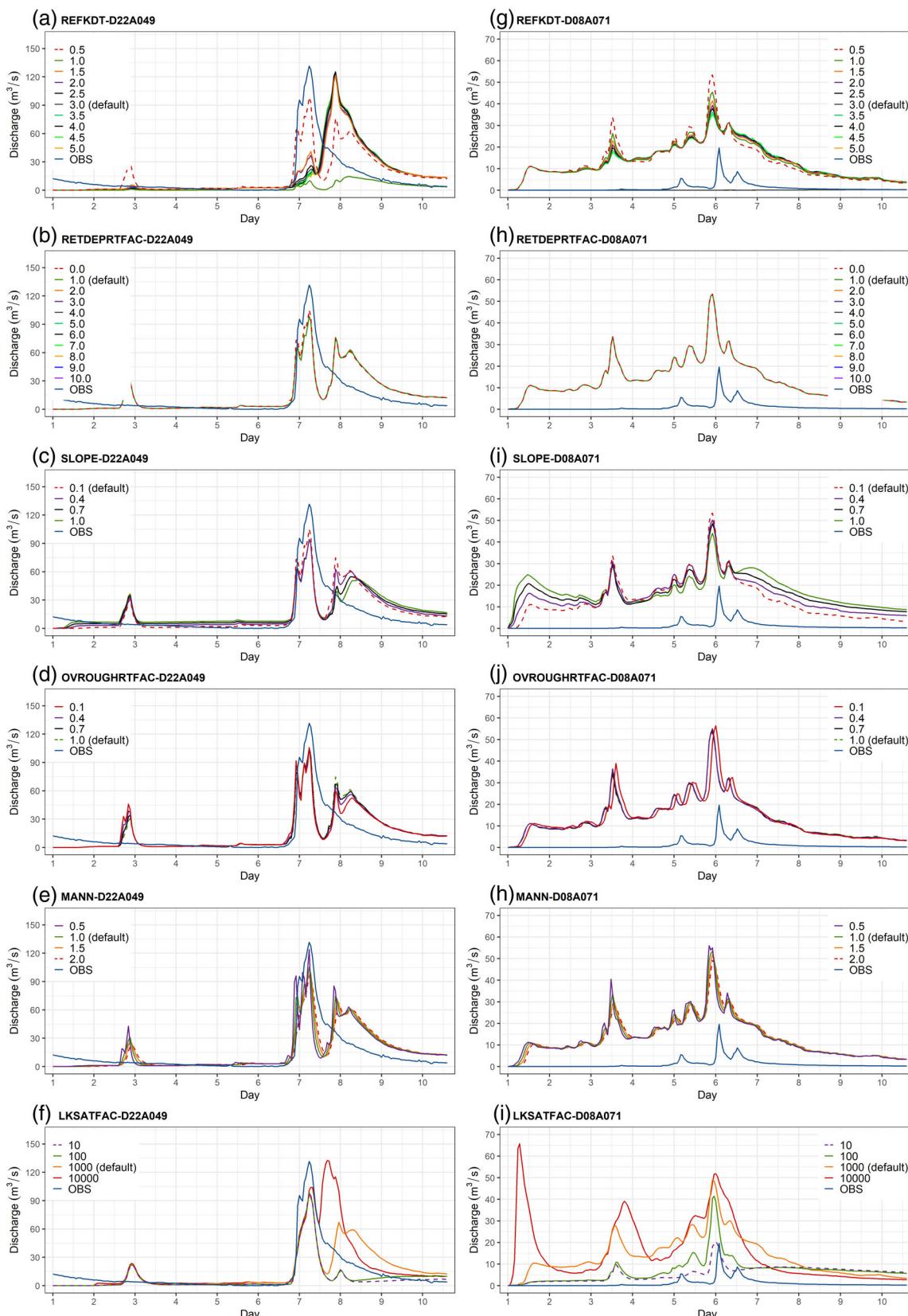
In Figure 2(c) and (g), the SLOPE parameter shows little influence on simulated hydrographs in terms of statistics. The default value (0.1) is selected as an appropriate SLOPE value for the model in both basins. However, it is observed that the other calibrated events in D22A147 and D09A095 basins show improvement in RMSE and RR with the SLOPE parameter (Table 3).

For the second step, parameters controlling hydrograph shape and timing are considered for the calibration process. Figure 2(d) and (j) show the results from the calibration of the OVROUGHRTFAC parameter. According to statistical measures, the default value of OVROUGHRTFAC (1.0) is found to be the optimum for all basins except the value of 0.1, selected for basin D09A095 (Table 3).

Figure 2(e) and (k) show the MANN parameter tuning results. The highest correlation is observed for the value of 2.0 that the KGE value is improved by  $\sim 15\%$  in D22A147. In general, RMSE improvement is observed in all basins for a value of 2.0 (Table 3). Also, a similar improvement is observed for a value of 0.5 in E08A008. Thus, the scaling factor (MANN) is selected as 0.5 for E08A008, while 2.0 is selected for others.

Lastly, the calibration of the LKSATFAC parameter is shown in Figure 2(f) and (l). It appears that LKSATFAC is the most sensitive parameter in both regions, particularly for the MED region. It influences peak timing and its magnitude with a significant decrease. Over both regions, the value of 10 is determined as the optimum value for LKSATFAC.

In Table 3, progressive improvement in RMSE and RR is observed from the first simulation (with default parameter set) to the simulation of LKSATFAC in a step-wise manner for both basins. For each streamflow station in Table 3, three events are used for calibration. The calculated statistics in Table 3 is the average of the results acquired from three events. As a result of manual step-wise calibration, an optimum parameter value depending on the calculated statistics for each event among the three can be found as a different value. Therefore, an optimum representative parameter value is selected qualitatively, not following a systematic way. With the calibration, the RR increases from 0.13 to 0.56, while KGE increases from 0.13 to 0.40 for D22A49. On the other hand, bias switches to the negative value, which is likely resulted from the effect of the sharp decrease in the recession stage in Figure 2(f). In D22A147, significant improvement is observed in the RR (from 0.38 to 0.71), NSE (from  $-0.58$  to 0.72) and KGE ( $-0.16$  to 0.61) at the end of the calibration process.



**FIGURE 2** Calibration results of the selected WRF-hydro model parameters, namely REFKDT, RETDEPRTFAC, SLOPE, OVROUGHRTFAC, MANN and LKSATFAC: (a)–(f) left column for the event occurred between 19 October 2019 and 29 October 2016 and basin D22A049 located over EBS region; (g)–(l) right column for the event occurred between 07 March 2017 and 17 March 2017 and basin D08A071 located over MED region. The dashed line shows the modelled hydrograph for the selected optimum parameter value. Calibration is performed with hourly streamflow data in these sub-basins

**TABLE 3** Average statistics of (bias, root mean square error (RMSE), correlation coefficient (RR), Nash–Sutcliffe efficiency (NSE) and Kling–Gupta efficiency (KGE) calibrated parameters for three events compare to default parameter set for D22A049 and D22A147 basins over EBS region and D08A071, D09A095 and E08A008 basins over MED region. Default parameter set is defined as 3.0, 1.0, 0.1, 1.0, 1.0 and 1000 for the parameters: REFKDT, RETDEPRTFAC, SLOPE, OVROUGHRTFAC, MANN and LKSATFAC, respectively

D22A049							
Region		Parm. Values	Bias	RMSE	RR	NSE	KGE
EBS	Default parameter set		4.24	40.55	0.13	-0.39	0.13
	REFKDT	0.5	3.72	40.48	0.38	-0.33	0.30
	RETDEPRTFAC	0.0	4.00	40.45	0.39	-0.35	0.31
	SLOPE	0.1	4.00	40.45	0.39	-0.35	0.31
	OVROUGHRTFAC	1.0	4.00	40.45	0.39	-0.35	0.31
	MANN	2.0	3.69	37.54	0.39	-0.14	0.40
	LKSATFAC	10	-2.34	32.16	0.56	0.16	0.40
D22A147							
		Parm. Values	Bias	RMSE	RR	NSE	KGE
MED	Default parameter set		0.48	5.75	0.38	-0.58	-0.16
	REFKDT	0.5	0.58	3.20	0.63	0.51	0.47
	RETDEPRTFAC	0.0	0.97	3.16	0.63	0.52	0.47
	SLOPE	1.0	1.01	2.88	0.67	0.58	0.47
	OVROUGHRTFAC	1.0	1.01	2.88	0.67	0.58	0.47
	MANN	2.0	0.78	2.65	0.68	0.63	0.54
	LKSATFAC	10	0.47	2.22	0.75	0.72	0.61
D08A071							
		Parm. Values	Bias	RMSE	RR	NSE	KGE
MED	Default parameter set		-5.28	16.68	0.43	-0.03	-0.12
	REFKDT	0.5	-1.02	30.12	0.44	-13.53	-3.18
	RETDEPRTFAC	0.0	-0.47	30.53	0.44	-13.54	-3.17
	SLOPE	0.1	-0.47	30.53	0.44	-13.54	-3.17
	OVROUGHRTFAC	1.0	-0.48	29.84	0.46	-13.52	-3.16
	MANN	2.0	-0.50	29.85	0.49	-13.09	-3.14
	LKSATFAC	10	-5.57	26.30	0.46	-1.26	-0.75
D09A095							
		Parm. Values	Bias	RMSE	RR	NSE	KGE
MED	Default parameter set		-0.10	12.40	0.59	-1.69	-0.55
	REFKDT	0.5	1.31	9.67	0.73	-1.21	-0.29
	RETDEPRTFAC	0.0	1.59	9.51	0.65	-1.16	-0.25
	SLOPE	1.0	1.66	8.77	0.70	-0.78	-0.14
	OVROUGHRTFAC	0.1	1.69	8.55	0.70	-0.59	-0.06
	MANN	2.0	1.70	8.35	0.81	-0.54	-0.04
	LKSATFAC	10	2.29	9.02	0.77	-0.93	-0.14
E08A008							
		Parm. Values	Bias	RMSE	RR	NSE	KGE
MED	Default parameter set		19	23.45	0.35	-283.11	-13.75
	REFKDT	0.5	18.93	23.17	0.51	-289.21	-13.89
	RETDEPRTFAC	0.0	18.95	23.14	0.51	-288.88	-13.87
	SLOPE	0.1	18.95	23.14	0.51	-288.88	-13.87
	OVROUGHRTFAC	1.0	18.92	23.14	0.51	-288.89	-13.88
	MANN	0.5	18.97	22.99	0.50	-289.34	-13.86
	LKSATFAC	10	6.22	8.07	0.40	-54.98	-5.14

For D08A071, an improvement is observed only in the RR, while bias and RMSE increase after the calibration of the MANN (Table 3). In D09A095 and E08A008, statistics at the end of the calibration process show an improvement compared to the model performed with default parameters (Table 3). E08A008 exhibits nearly no response to the RETDEPRTFAC, 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 D22A049 and D22A147 in the EBS region and D08A071, D09A095 and E08A008 in the MED region and default parameter set are given in Table 3. In addition to Table 3, for D22A079 and D22A089 in the EBS region, the same parameter values (0.5 and 0.0) with the other basins are determined for REFKDT and RETDEPRTFAC. For the other parameters, the value of 0.1 and 1.0 for SLOPE, the value of 1.0 and 0.1 for OVROUGHRTFAC, and the value of 10 000 and 1000 for LKSATFAC are selected for D22A079 and D22A089, respectively. For the MANN parameter, the value of 2.0 is selected for both basins.

### 3.2 | Spatio-temporal evaluation of SST products

GFS and ERA5 products have coarser spatial ( $0.25^\circ$ ) resolution than GHRSST (1.1-km), Medspiration (2.2-km) and NCEP products (9-km). In this study, GHRSST, 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 3, while spatially averaged time series are given in Figure 4. The eastern part of the EBS region is depicted by warmer temperatures ( $\sim 301$  K) than the western part ( $\sim 297$  K) consistently by all products (except for NCEP). Over the MED region, inter-product consistency is much smaller than EBS region, that spatial variability of average temperature is the largest (Figure 3). Overall, all products, except for constant ERA5 and GFS, are temporally consistent with each other, particularly over the MED region (Figure 4).

For any product, spatial and temporal cross-correlations are calculated with other products, and then these cross-correlations are averaged. In general, the average temporal cross-correlations are higher (0.79–0.84) over MED region than EBS (0.24–0.60), while vice versa for spatial cross-correlations (between 0.48–0.83 for EBS and 0.11–0.39 for MED region). Overall, average spatio-temporal cross-correlation for GHRSST (0.61) is higher than Medspiration (0.54), which is higher than NCEP (0.36); this order is also valid for average spatial and temporal cross-correlations as well as EBS and MED regions; this implies, among the time-variant SST products, GHRSST is the best and NCEP is the least performing products for the events and regions focused in this study. These results also resonate with the spatial resolution of the products that GHRSST, having highest accuracy, also have highest spatial resolution, while NCEP having the last accuracy also have the coarsest spatial resolution compared against GHRSST and Medspiration; these results imply overall there is a consistency between the spatial resolution of the SST products and their

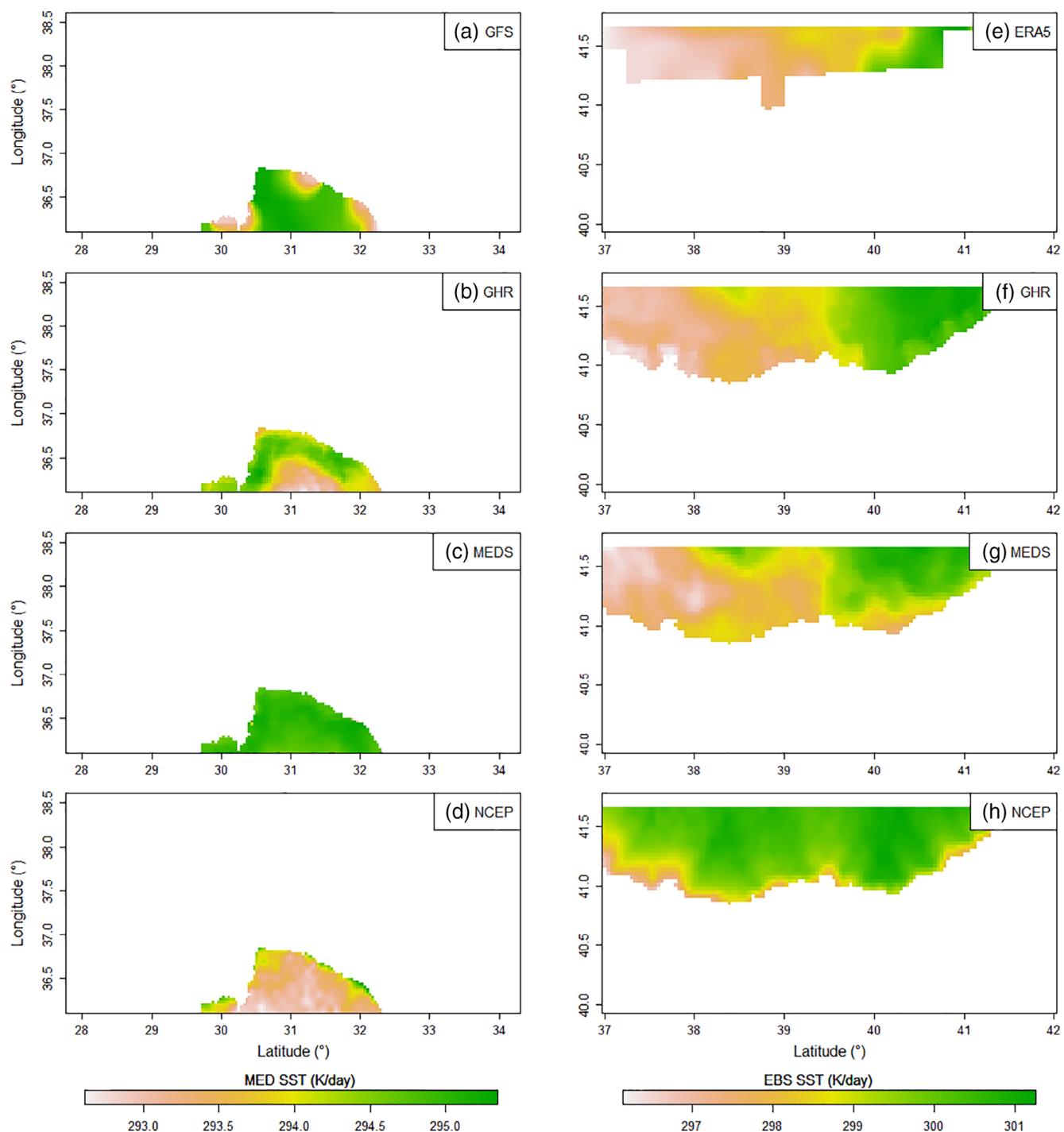
accuracy over both regions (i.e., higher the spatial resolution higher the accuracy).

### 3.3 | Precipitation evaluation for SST cases

RMSE and MAE values are calculated for interpolated precipitation and simulated precipitation for each SST corresponding to the point observations available in each model domain (Table 4). It is found that both RMSE and MAE values for the interpolation method are smaller than modelled precipitation obtained using each SST. For example, GHR SST simulation in the MED region shows higher RMSE and MAE values ( $\sim 1.89$  and  $\sim 0.81$ , respectively) than those values of interpolation ( $\sim 1.17$  and  $\sim 0.47$ , respectively), yet it shows relatively the lowest RMSE and MAE values among the other modelled precipitations. The mean RMSE of modelled precipitation in EBS is 29% greater than the RMSE of interpolated precipitation, while it is 41% higher than the interpolated precipitation RMSE in the MED region. Results imply that the error introduced by the WRF model to the precipitation is higher than that of interpolated precipitation.

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 5 shows the statistical measures calculated for each SST case in both basins. In Figure 5(a), the precipitation time series is represented from 08/17/2015 00:00:00 UTC to 27 August 2015 00:00:00 UTC (241-h). The maximum precipitation amount for D22A147 is recorded as 26.3 mm for the 178th hour, which corresponds to 24 August 2015 09:00:00 UTC. However, the maximum precipitation for the EBS region for this event was recorded as 32.4 mm at 24 August 2015 00:00:00 UTC. This maximum precipitation amount is observed within the boundary of the D22A049 basin. Nevertheless, as shown in Figure 5(a), the basin-average precipitation of the D22A147 provides 16.1 mm at the 169th 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 5, it can be depicted that using an external high-resolution SST dataset still improves the accuracy of the simulated precipitation in comparison with observed precipitation, especially for Medspiration. Besides, GHRSST simulation overestimates the observed peak precipitation. Other simulated peaks are lower than the GHRSST simulation, but they are closer to the observed peak.

In Figure 5(b), the basin-averaged precipitation time series is represented from 10 December 2018 00:00:00 UTC to 20 December 2018 00:00:00 UTC (241-h). Peak time and precipitation magnitude for the whole MED region are recorded as 53.1 mm at 162th hour (10 December 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 a similar

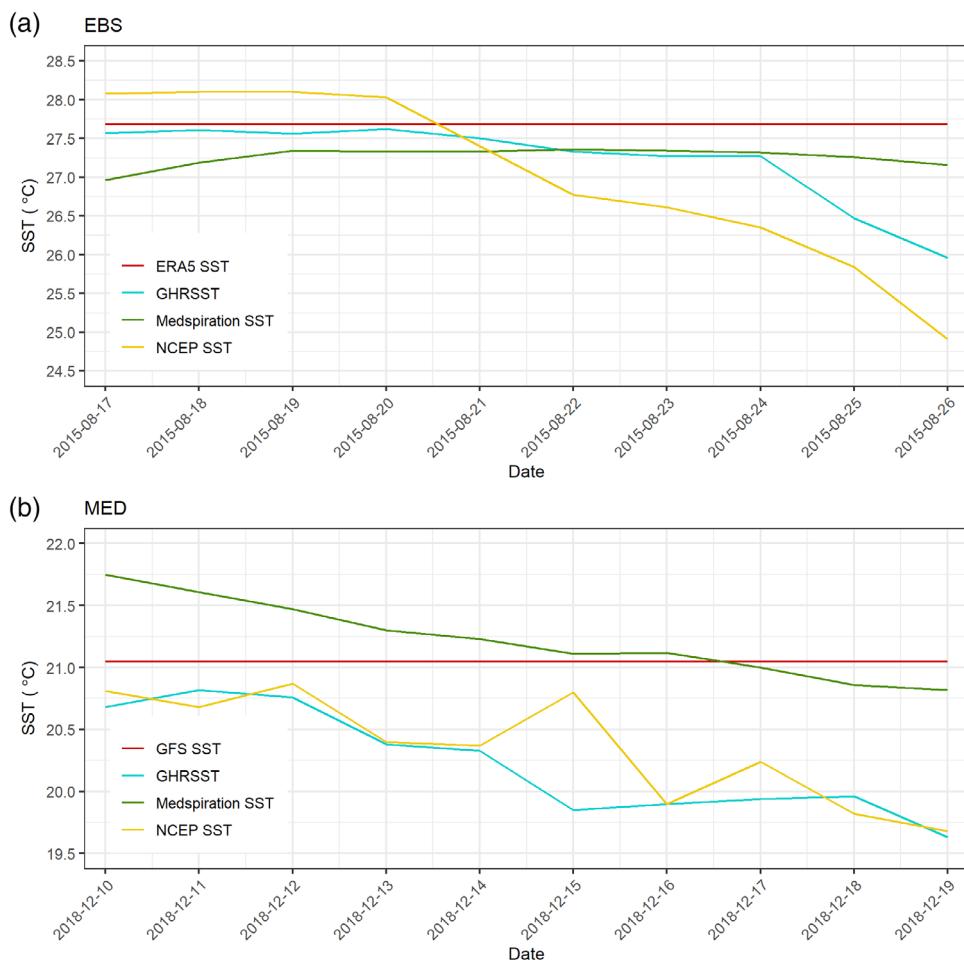


**FIGURE 3** Temporally averaged spatial distribution of SST products of ERA5/GFS, GHSST, Medspiration and NCEP: (a)–(d) over MED and (e)–(h) over EBS regions

trend as the observation with minor overestimations. Nonetheless, it appears that external SST simulations are able to improve the precipitation volume with reduced bias corresponding to the observed precipitation. Modest delays in peak time (1–2 h) are observed for GFS SST, GHSST and NCEP SST 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  $0.56 \text{ m}^3/\text{s}$  (Table 5). Medspiration shows the best model performance in terms of all statistics calculated with respect to the observed precipitation compared to the rest (Table 5).

Figure 6 shows the spatial distribution of observed precipitation and simulated precipitations from the WRF model created by different SST datasets on peak day (24 August 2015) over the EBS region. It is



**FIGURE 4** Spatially-averaged temporal distribution of SST products of ERA5/GFS, GHRSSST, Medspiration and NCEP: (a) over EBS and (b) MED regions

**TABLE 4** Comparison of mean absolute error (MAE) and RMSE values calculated for precipitation observations from rain gauges, interpolated and modelled (with ERA5/GFS, GHR, MED and NCEP SSTs) precipitation data extracted from the same rain gauge locations

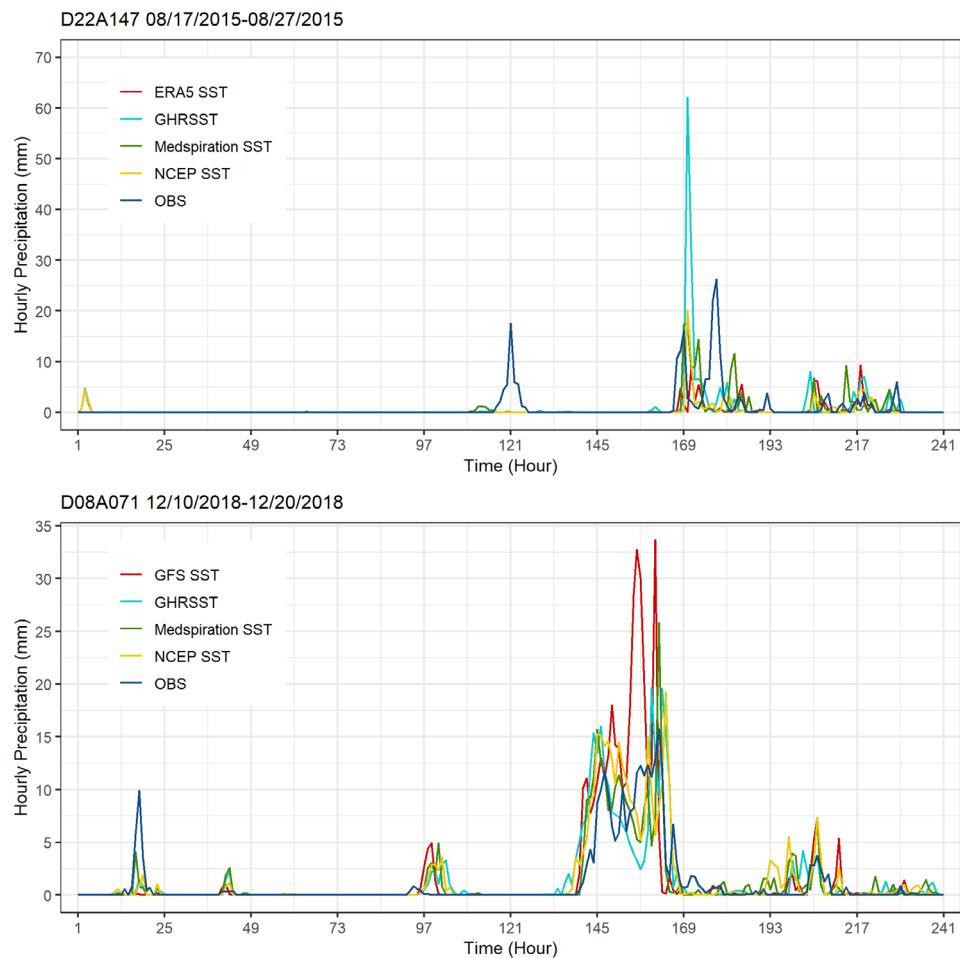
Region	Precipitation source	MAE	RMSE
EBS	Interpolated observation	0.240	0.954
	ERA5-SST	0.334	1.297
	GHR-SST	0.347	1.369
	MED-SST	0.341	1.370
	NCEP-SST	0.327	1.302
MED	Interpolated observation	0.466	1.165
	GFS-SST	0.868	2.129
	GHR-SST	0.808	1.885
	MED-SST	0.825	1.943
	NCEP-SST	0.812	1.901

noteworthy that in Figure 6, the GHRSSST simulation shows an overestimation in the 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)).

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

**FIGURE 5** Time series of hourly precipitation that (a) D22A147 basin over EBS region receives during the event occurred in 17 August 2015–27 August 2015 and (b) D08A071 basin over MED region receives during the event occurred in 10 December 2018–20 December 2018 for 10 days. Outputs are generated from the WRF model with the native SST field from ERA5 reanalysis data (ERA5-SST) for EBS region and GFS forecast data (GFS-SST) for MED region with different SST products: GHRSSST, Medspiration and NCEP

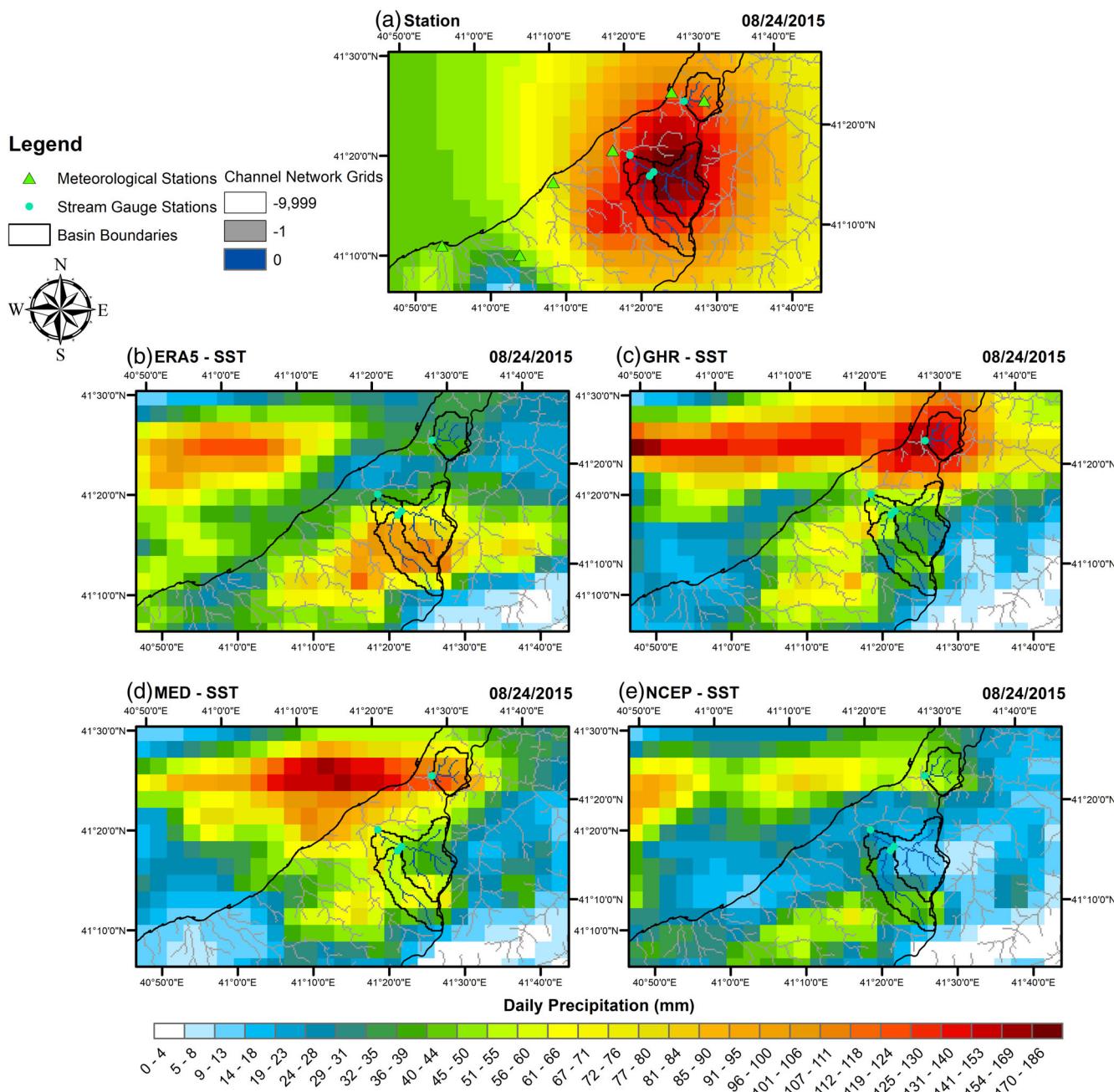


**TABLE 5** Statistics of bias, RMSE and RR between observed and modelled precipitations with different SST datasets of ERA5, GHRSSST, Medspiration and NCEP for D22A147 over EBS and GFS, GHRSSST, Medspiration and NCEP for D08A071 over MED regions. Statistics of bias, RMSE, RR, NSE and KGE between observed and modelled hydrographs with observed precipitation and different SST simulations for the corresponding basins

Station		WRF precipitations			WRF-hydro streamflows (calibrated parameter set)				
		Bias	RMSE	RR	Bias	RMSE	RR	NSE	KGE
D22A147	Observed Precip.	-	-	-	-7.11	14.25	0.83	0.37	0.063
	ERA5-SST	-0.54	3.19	0.03	-9.92	20.13	0.42	-0.26	-0.39
	GHR-SST	-0.06	5.30	0.01	-6.16	10.83	0.83	0.63	0.30
	MED-SST	-0.24	3.55	0.03	-8.32	15.86	0.86	0.22	-0.08
	NCEP-SST	-0.54	3.38	0.01	-9.60	18.98	0.82	-0.12	-0.29
D08A071	Observed Precip.	-	-	-	-57.48	84.53	0.40	-0.13	0.11
	GFS-SST	0.56	3.45	0.60	-24.98	128.81	0.18	-1.62	0.02
	GHR-SST	0.18	2.35	0.52	-42.73	83.25	0.30	-0.09	0.31
	MED-SST	0.13	1.86	0.67	-43.63	83.57	0.31	-0.10	0.30
	NCEP-SST	0.33	2.23	0.60	-40.76	81.50	0.32	-0.05	0.33

predictions to the observation in terms of precipitation depth (Figure 7[d and e]). Especially, the 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 the Medspiration simulation over D08A071



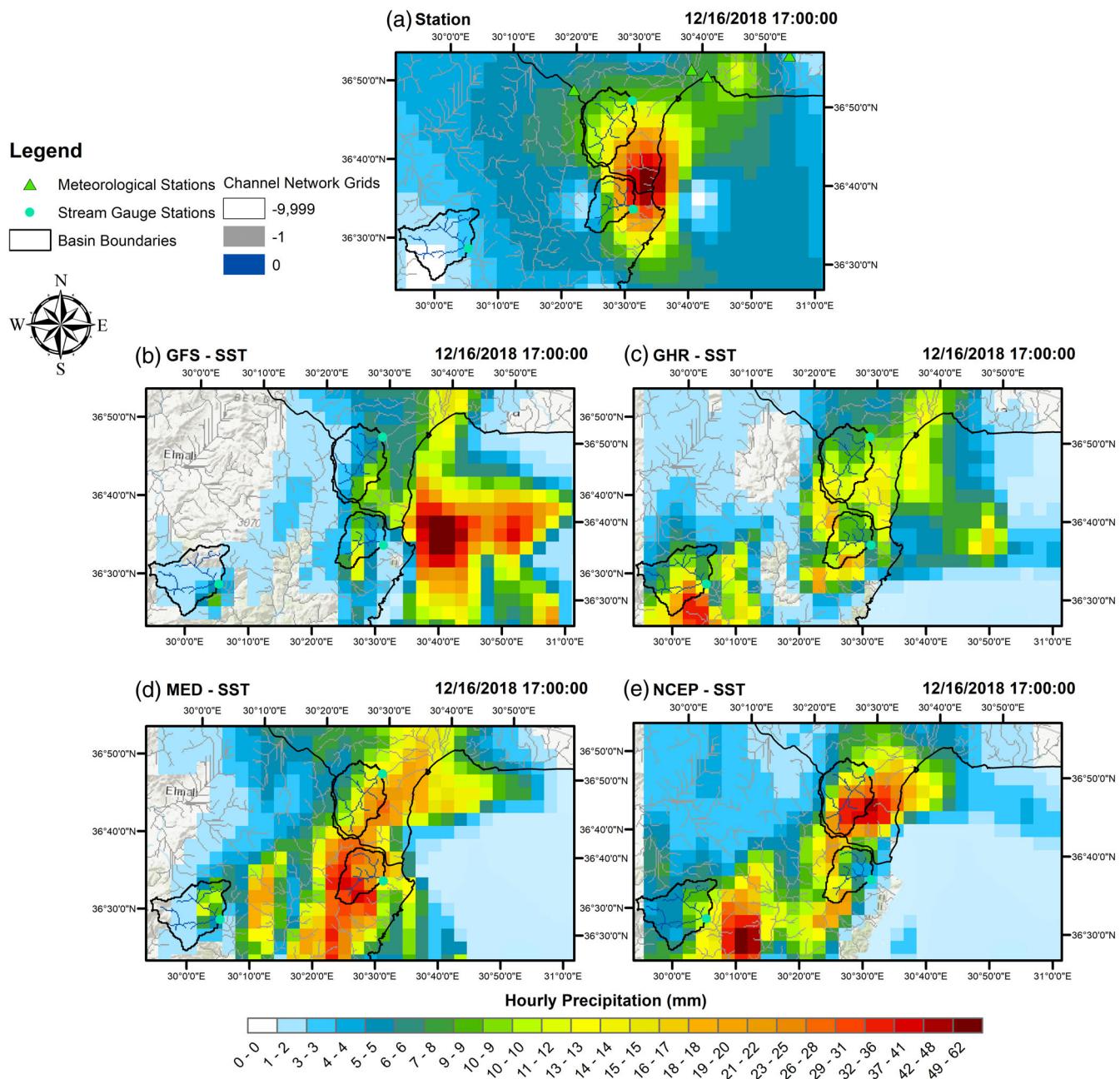
**FIGURE 6** Spatial distribution of daily precipitation at the peak day (24 August 2015) for the run period of 17 August 2017–27 August 2017 over the EBS region. (a) the map at the top shows the interpolated observed precipitation map obtained from meteorological station data (green triangles). Black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

is 25.8 mm, which corresponds to the darker orange colouring of the basin grids (Figure 7(d)).

### 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. Figure 8

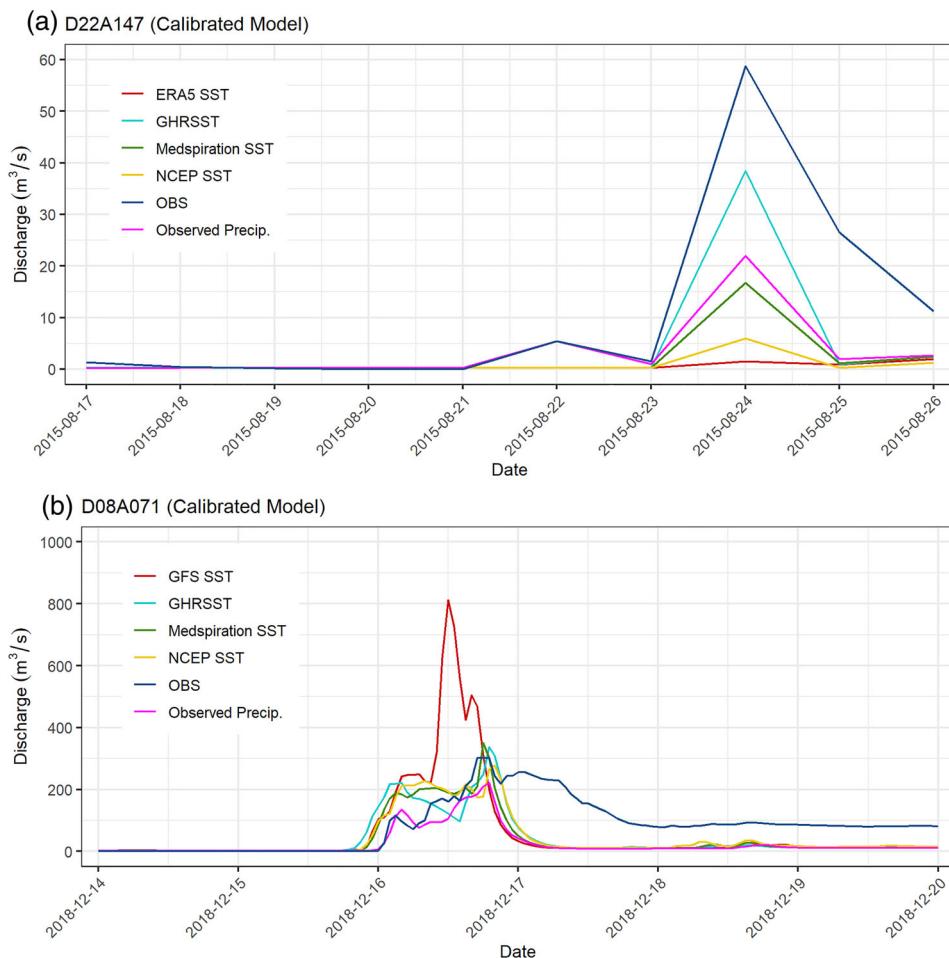
shows the simulated hydrographs by calibrated models forced by different SST simulations and observed precipitation in D22A147 and D08A071. ERA5 and NCEP hydrographs show substantial underestimation for the peak volume (Figure 8(a)). This may be due to the negative bias observed in precipitation in Figure 5(a) for ERA5 and NCEP simulations (They are the ones with the highest negative bias among other SST simulations and show negative NSE and KGE). Medspiration simulation and simulation with observed precipitation



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

create slightly better hydrograph volume and statistics than ERA5 and NCEP simulations. Though the GHRSSST generates overestimation in precipitation and misses the event peak time for D22A147 as discussed in the previous section (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 24 August 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 NSE value of 0.63 and the KGE value of 0.30 is achieved through GHRSSST simulation (Table 5). On average, RRs increase from 0.42 (for ERA5) to 0.8 for the simulated hydrographs with high-resolution SST datasets. Overall, the statistical



**FIGURE 8** Comparison of observed hydrograph with the simulated hydrographs generated using precipitation inputs derived with native SST field (ERA5/GFS), GHRSSST, Medspiration, NCEP and interpolated point observations  
 (a) for event 17 August 2015–27 August 2015 in D22A0147 and  
 (b) for event 10 December 2018–20 December 2018 in D08A071

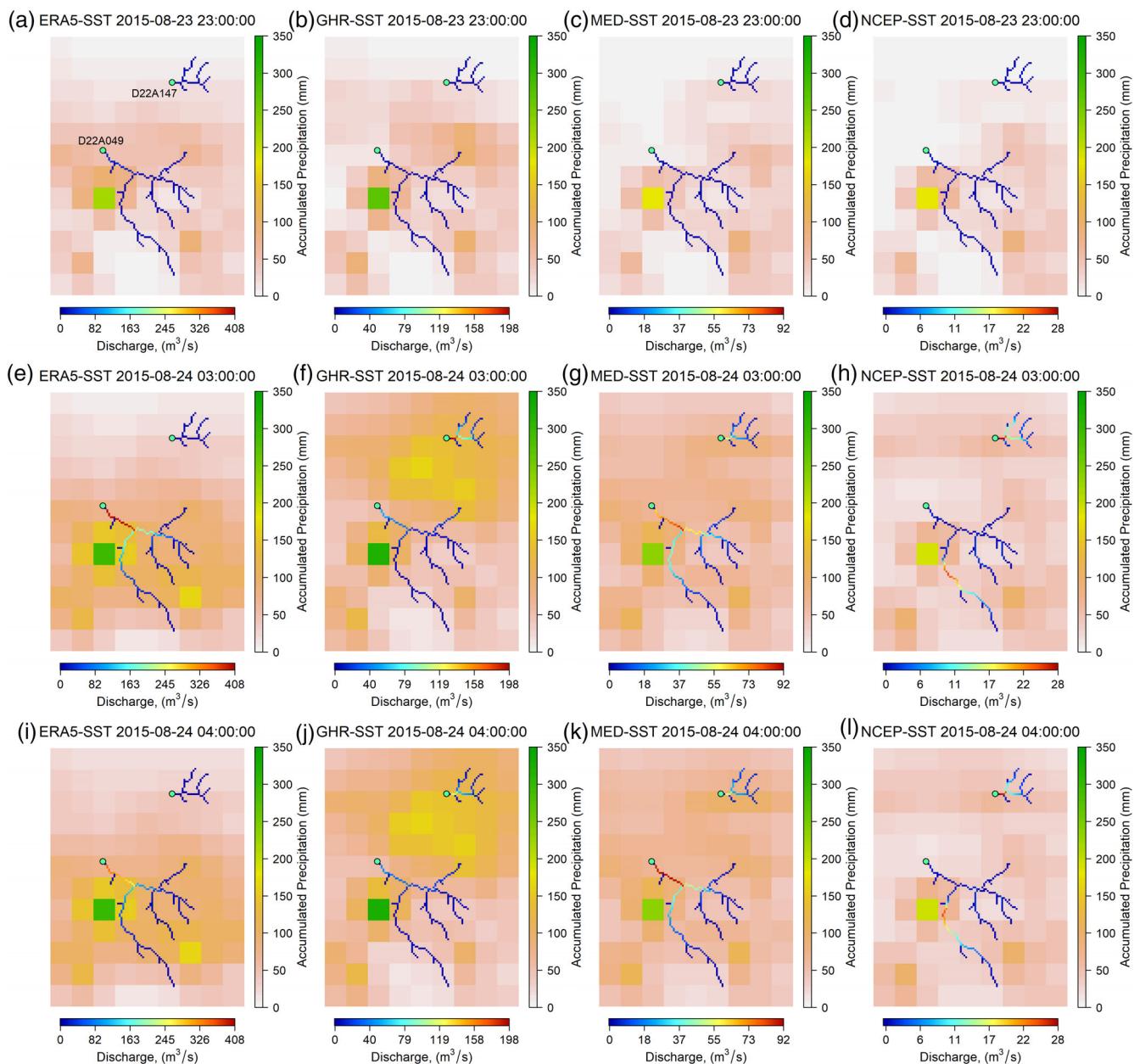
measures in Table 5 show that the GHRSSST simulation yields improved performance for all statistics with respect to the observed precipitation simulation. Also, the RR is increased by  $\sim 4\%$  for Medspiration simulation. GHRSSST simulation especially creates a substantial improvement ( $\sim 70\%$  increment in NSE) in the performance of the hydrograph volume compared to observed precipitation simulation. These results indicate that the GHRSSST is the most representative SST dataset for D22A147 among the other SST datasets.

Comparison of hourly observed hydrograph and simulated hydrographs forced by four different SST events and observed precipitation with the set of calibrated parameters for the D08A071 is represented in Figure 8(b) (plotted for the last 6 days of the model run period). Minor delays in the primary hydrograph peak time are observed for the observed precipitation simulation and the SST simulations except for the GFS simulation. They overestimate the observed discharge until peak time, yet the underestimation in the falling limb stage causes negative bias between  $-24.98$  and  $-57.48 \text{ m}^3/\text{s}$ , as shown in Table 5. The simulated hydrograph of GFS SST produces an anomalously higher peak of  $812.9 \text{ m}^3/\text{s}$  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 ( $-24.98 \text{ m}^3/\text{s}$ ), it produces the highest RMSE (128.81) and lowest RR (0.18) (Table 5). 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). Simulated hourly hydrographs with the calibrated parameter set in Figure 8(b) represent a reasonable behaviour in the rising limb part till their peak values, but their falling limb parts decrease more sharply after the calibration. Compared with the high-resolution SST simulation, observation precipitation simulation shows a significant underestimation in hydrograph volume and causes a negative bias of  $-53.48 \text{ m}^3/\text{s}$  while achieving a better correlation value (0.40). Overall, external high-resolution SST products significantly capture the observed peak value with an improvement in the ERA5 hydrograph by the reduction of  $\sim 550 \text{ m}^3/\text{s}$ .

### 3.5 | Evaluation of rainfall-runoff representations

Figure 9 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

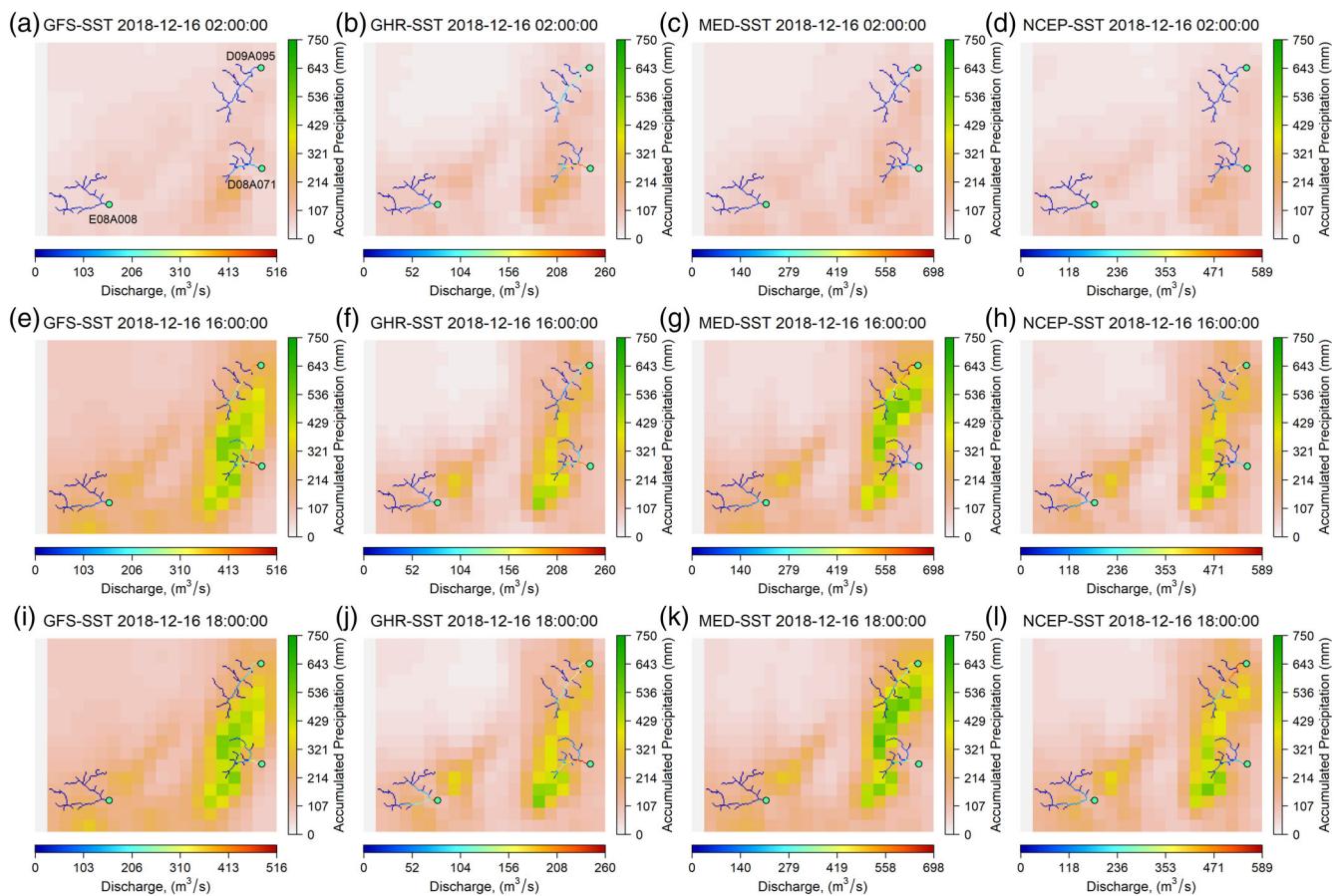


**FIGURE 9** Overlapped dynamic maps of accumulated precipitation simulated by WRF model (3-km) operated with four different SST datasets (ERA5, GHR-SST, Medspiration and NCEP) and discharge simulated by WRF-hydro model (250-m) over EBS region at 23 August 2015 23:00:00, 24 August 2015 03:00:00, and 08/24 August 2015 04:00:00. Stream gauges are denoted as blue dots

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 9(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 9(f), at the second time step, the D22A147 basin receives the highest precipitation compared to others; this result is consistent with the simulated precipitation with GHR-SST generating 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 h) to the river network and collected to the outlet point and lead to flooding. This is clearly seen in Figure 9(e) for the D22A049, in Figure 9(f) for the D22A147, and in Figure 9(g and 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 9[i-l]).

Figure 10 shows overlapped dynamic maps of accumulated precipitation simulated by the WRF model through using four different SST datasets (GFS, GHR-SST, Medspiration and NCEP) and simulated discharges on gridded river networks corresponding to these four precipitation estimates over the MED region for D08A071, D09A095



**FIGURE 10** Overlapped dynamic maps of accumulated precipitation simulated by WRF model (3-km) operated with four different SST datasets (GFS, GHRSST, Medspiration and NCEP) and discharge simulated by WRF-hydro model (250-m) over MED region at 16 December 2018 02:00:00, 16 December 2018 16:00:00, and 16 December 2018 19:00:00. Stream gauges are denoted as blue dots

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 10(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  $m^3/s$  as a response to the accumulated precipitation for 14 h, especially over the upper basin (Figure 10[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  $m^3/s$  (Figure 10(g)). In Figure 10(l), due to the minor delays in primary peak time discharge in hydrographs of GHRSST and NCEP, the channel grid network is still on the rising limb stage with respect to the simulated hydrographs in Figure 8(b).

## 4 | DISCUSSION

### 4.1 | Discussion of results

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 significantly affect the accuracy of runoff simulation in terms of timing and magnitude of the peak value and overall volume. Duzenli et al. (2020) found that among uncertainties related to the error sources resulted from physics parameterization and initial condition, the parameterization affects the precipitation variability more than the initial/boundary data source for the simulated extreme events in MED and EBS regions. However, the changes introduced by GFS and ERA5 at the initial state and through time-varying boundary conditions may propagate differently through WRF model simulations. Therefore, the SST impacts on the results are further influenced by the use of these different forcing datasets. It is worth to mention to understand the uncertainties of the results. As it is expected, the uncertainty raised by the error sources of the lower boundary representation through different SST datasets is of primary importance in weather prediction efforts. The outcomes achieved from this study highlight that studying the time-variant, high-resolution SST on the complex coastal topographies like MED and EBS regions (Ferrari et al., 2020; Papanastasiou et al., 2010; Senatore et al., 2014) can help to improve the precipitation prediction accuracy and reduce the uncertainties of mesoscale modelling. Water and energy fluxes between the land/sea surface and the atmospheric boundary layer entail complex interactions. Therefore, in making a

hydrometeorological prediction for floods/flash-floods across various space and time scales, the sensitivity analysis focusing on better representation of lower SST boundary layer within the atmospheric-hydrological processes is imperative. Similarly, updating the lower boundary with more appropriate SST helps to reduce the overall error in storm hydrographs derived using the one-way coupled system. Overall, the GHRSST product yields the highest spatio-temporal correlation, while NCEP yields the least among the temporally variable SST products. Consistent with this result, GHRSST-based runoff simulations yield the highest accuracy, while NCEP shows the lowest among the temporally variable SST products. These results clearly show the significance of using higher spatio-temporal resolution SST products in the simulation of heavy rainfall and extreme runoff.

In this study, cross-correlations are used as a validation tool, where dense buoy observations are non-existent over the study regions. Accordingly, the consistency between the cross-correlation-based accuracy estimates and the runoff simulations shows that 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. According to Jee and Kim (2017), detailed SST representation in the WRF model shows a negligible influence on the simulated precipitation for short-term events. However, this result is consistent with the findings of previous studies agreeing that the high-resolution SST representation improves the performance of short-term simulations (Cassola et al., 2016; Iivatek-Şahdan et al., 2018; Lebeaupin et al., 2006).

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 time-variant SST products from GHRSST, Medspiration and NCEP over both study regions revealed significant changes in heavy precipitation amounts for the selected short-term events with respect to invariant 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. Conversely, GFS SST creates the highest overestimation in basin-average precipitation, around 17 mm for the D08A071 basin, and produces precipitation over the sea (agreeing with the study of Cassola et al., 2016 conducted in MED region). The high spatio-temporal resolution SSTs (GHRSST and Medspiration) help resolve high variability in rainfall, and its hydrologic response resulted from a mesoscale convective system that occurred in the ESB region. As a result, the Medspiration over the MED region and GHRSST over the EBS region revealed the best basin-averaged precipitation representation that directly translates into an improvement in surface runoff prediction in small catchments of both study regions.

Results also present an agreement with the studies of Senatore et al. (2015), Givati et al. (2016) and Sun et al. (2020) that the models forced with the observation precipitation show better performance in simulated hydrographs compare to the hydrographs of coarse SSTs

(ERA5 and GFS) in both regions. On the other hand, results show that it may occur that SST-driven simulations result in a hydrologic better performance than those run with observations. The observed precipitation errors resulted from the interpolation method in both regions are 30%–40% smaller than the modelled precipitation. The density of the observation network representing complex terrain is critical in reducing this error percentage further. Error resulting from the observed precipitation simulations may be sourced from the reason that most of the meteorological observation stations are located at the valley floors; thus, they could not represent orographic precipitation characteristics of the region (Eris & Agiralioglu, 2018). This situation may be attributed to our case that the observed precipitation data could be underestimating the observed hydrograph.

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. In the model, infiltration-excess, interflow and channel routing mechanisms influence the discharge simulation. Non-homogenized distribution of precipitation rate in the semi-arid region dominated by infiltration excess mechanism results in underestimating or neglecting some infiltration excess (e.g., basins in MED region). Moreover, the interflow pattern should be improved in the model. The weakness in the interflow mechanism is responsible for the discharge overestimations of falling limbs, which has been presented in this study and other literature involving humid (Ryu et al., 2017; Senatore et al., 2015), semi-arid (Lahmers et al., 2019; Silver et al., 2017), and arid basins (Silver et al., 2017). There are four soil layers in the model, but saturated interflow is still calculated in one layer, with its thickness as the sum of the four layers (Gochis et al., 2020). This consideration makes the outflow of interflow more homogenized and concentrated. However, it is worth mentioning that the calibration method showed better performance in the humid EBS region than that in the semi-arid MED region. The statistics particularly for NSE and KGE, are represented by high positive values (0.2–0.72) in the EBS region. Moreover, 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 a 1-h period. The high-resolution gridded rainfall-runoff coupling greatly benefits from monitoring the water excess condition for a given storm over topographically complex and steeply small watersheds.

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 (Senatore et al., 2015; Silver et al., 2017; Yucel et al., 2015). Among the parameters, REFKDT, SLOPE, MANN and LKSATFAC revealed an important impact on making reliable runoff predictions in both regions, but especially the saturated hydraulic conductivity parameter factor (LKSATFAC) became substantially critical over the MED region.

## 4.2 | Added utility compared against earlier studies

Hydrometeorological modelling systems, together with the local topography, drive the hydrological responses seen during extreme

events. However, so far, only a few studies (Senatore, Davolio, et al., 2020; Senatore, Furnari, et al., 2020) have investigated the added utility of different SST products driving the WRF forecasts coupled with hydrological models during extreme events. These two studies analysed the impact of the SST products, while they went beyond the investigation of the impact of the SST products over the acquired precipitation forecasts and evaluated the impact of SST products over the forecasts of hydrological responses during different extreme events. In these two studies, SST sensitivity was investigated in the context of overall mesoscale modelling uncertainty related with initial and boundary conditions for the first time using WRF/WRF-Hydro modelling system, with and without data assimilation (Senatore, Furnari, et al., 2020). Similarly, in this current study, the impact of the accuracy of the SST products over the hydrological responses is also investigated using the same modelling system. The additional contributions and the advances acquired in this study compared with the studies of Senatore, Furnari, et al. (2020) and Senatore, Davolio, et al. (2020) can be listed as follows:

- In the studies of Senatore, Furnari, et al. (2020) and Senatore, Davolio, et al. (2020), the impact of high-resolution SST is not found significant as it can be hidden due to other more relevant sources of uncertainty (initial and boundary conditions). On the other hand, here in this study, significant improvements in the runoff simulations are obtained through the update of the SST field with the high-resolution SST products in the WRF model for the investigated extreme events.
- The studies of Senatore, Furnari, et al. (2020) and Senatore, Davolio, et al. (2020) have been performed over the basins situated along the Mediterranean Sea coastal region, while this current study is implemented over basins situated along both Mediterranean and Black Sea coastal regions (i.e., MED and EBS regions). Accordingly, the sensitivity of the hydrological responses to high-resolution SST products is investigated over a different region with significantly different climatic properties and has not been explored before (i.e., EBS). The results show that the hydrological response improvement acquired using high-resolution SST products is greater over the EBS region than the MED region.
- The studies of Senatore, Furnari, et al. (2020) and Senatore, Davolio, et al. (2020) have utilized only Medspiration SST products with  $0.022^\circ$  spatial resolution, while in this study, Medspiration, GHR, NCEP SST products with  $0.022^\circ$ ,  $0.01^\circ$  and  $0.083^\circ$  spatial resolutions, respectively, are all used. Accordingly, the sensitivity of hydrological response of additional SST products is investigated, expanding the high-resolution SST input product choices to be able to obtain more general conclusions.
- Temporal accuracy assessments have been performed in the study of Senatore, Furnari, et al. (2020) utilizing collected buoy observations. Here in this current study, both the temporal and the spatial accuracy assessments have been done utilizing SST cross-correlation information of four SST products, where dense buoy observations are absent in the study area. Accordingly, this current

study advances the study of Senatore, Furnari, et al. (2020) via additional spatial accuracy assessments.

## 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 time-variant and time-invariant 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 the WRF 3-km domain to the Hydro model. GFS over the MED region and ERA5 data over the EBS region include their own SST values (considered as coarse resolution and time-invariant data sets), whereas GHRSST, Medspiration and NCEP products are described as high-resolution and time-variant 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:

- The manual step-wise calibration showed better performance in the humid EBS region than that in the semi-arid MED region.
- Using higher spatio-temporal resolution SST products (Medspiration and GHRSST), prevailing higher average cross-correlation is also highly influential in capturing the temporal and spatial variability of precipitation in small catchments. This effect is variable from region to region.
- The impact of high spatio-temporal resolution SST update on simulated hydrographs over both regions is highlighted in terms of more accurate predictions of peak discharges. Overall, high-resolution GHRSST and Medspiration show more significant improvement compared to other SST datasets to capture peak discharge timing and magnitude for storm hydrographs simulated over both regions.
- Depending on the spatial and temporal distribution of precipitation, runoff simulations performed by time-invariant SSTs resulted in poor performance. While there is no hydrologic response in the EBS, the flood peak magnitude is simulated 2.5 times higher than the observed peak magnitude in MED.
- With the calibrated model, it may occur that SST-driven simulations result in a hydrologic better performance than those run with observations, particularly for higher resolution SSTs.

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 for operational forecasts of extreme events. Under the consideration of abnormal SST changes exacerbated by changing climate, time-variant SST features characterized

with high spatio-temporal resolution should be accounted for extreme weather event evaluations in complex coastal topographical regions. Also, the discharge simulation mechanism of WRF-Hydro contributed from interflow pattern should be improved to account for overestimations of falling limb issued in this study and literature. A more systematic manual calibration approach instead of a qualitative manner needs to be developed.

## ACKNOWLEDGEMENTS

We thank the Turkish State Meteorological Service for the ground-based meteorological observations, The State Hydraulics Works of Turkey for streamflow observations, The Group for High-Resolution Sea Surface Temperature for GHRSST datasets, CERSAT for Medspiration datasets, NCEP for RTG\_SST datasets and GFS products, ECMWF for ERA5 products, and NCAR/NOAA for serving and improving the WRF model. This project is funded by The Scientific and Technological Research Council of Turkey (TUBITAK) [Grant Number 116Y193]. We also thank to the constructive comments of two anonymous reviewers and the editor.

## DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## ORCID

Berina Mina Kilicarslan  <https://orcid.org/0000-0001-5299-1334>

Ismail Yucel  <https://orcid.org/0000-0001-9073-9324>

## REFERENCES

- Allen, M. R., & Ingram, W. J. (2002). Constraints on future changes in climate and the hydrologic cycle. *Nature Communications*, 419 (September), 225–232. <https://www.nature.com/articles/nature01092>
- Baltaci, H. (2017). Meteorological analysis of flash floods in Artvin (NE Turkey) on 24 August 2015. *Natural Hazards and Earth System Sciences*, 17(7), 1221–1230. <https://doi.org/10.5194/nhess-17-1221-2017>
- Bozkurt, D., & Sen, O. L. (2011). Precipitation in the Anatolian peninsula: Sensitivity to increased SSTs in the surrounding seas. *Climate Dynamics*, 36(3), 711–726. <https://doi.org/10.1007/s00382-009-0651-3>
- Camera, C., Bruggeman, A., Zittis, G., Sofokleous, I., & Arnault, J. (2020). Simulation of extreme rainfall and streamflow events in small Mediterranean watersheds with a one-way-coupled atmospheric-hydrologic modelling system. *Natural Hazards and Earth System Sciences*, 20(10), 2791–2810. <https://doi.org/10.5194/nhess-20-2791-2020>
- Cassola, F., Ferrari, F., Mazzino, A., & Miglietta, M. M. (2016). The role of the sea on the flash floods events over Liguria (northwestern Italy). *Geophysical Research Letters*, 43(7), 3534–3542. <https://doi.org/10.1002/2016GL068265>
- CERSAT. (2012). Medspiration. <http://cersat.ifremer.fr/thematic-portals/projects/medspiration>
- Chen, W., Wang, L., Xue, Y., & Sun, S. (2009). Variabilities of the spring river runoff system in East China and their relations to precipitation and sea surface temperature. *International Journal of Climatology*, 29(10), 1381–1394. <https://doi.org/10.1002/joc.1785>
- Cisneros, M., Cacho, I., Frigola, J., Canals, M., Masqué, P., Martrat, B., Casado, M., Grimalt, J. O., Pena, L. D., Margaritelli, G., & Lirer, F. (2016). Sea surface temperature variability in the Central-Western Mediterranean Sea during the last 2700 years: A multi-proxy and multi-record approach. *Climate of the Past*, 12(4), 849–869. <https://doi.org/10.5194/cp-12-849-2016>
- Done, J., Davis, C. A., & Weisman, M. (2004). The next generation of NWP: Explicit forecasts of convection using the weather research and forecasting (WRF) model. *Atmospheric Science Letters*, 5(6), 110–117. <https://doi.org/10.1002/asl.72>
- Duzenli, E., Yucel, I., Pilatin, H., & Yilmaz, M. T. (2020). Evaluating the performance of a WRF initial and physics ensemble over eastern Black Sea and Mediterranean regions in Turkey. *Atmospheric Research*, 248 (August), 105184. <https://doi.org/10.1016/j.atmosres.2020.105184>
- ECMWF. (2020). ERA5. <https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>
- Eris, E., & Agiralioglu, N. (2018). Streamflow map of the Eastern Black Sea region, Turkey. *Scientia Iranica*, 25(3A), 1048–1056. <https://doi.org/10.24200/sci.2017.4261>
- Ferrari, F., Cassola, F., Tuju, P. E., Stocchino, A., Brotto, P., & Mazzino, A. (2020). Impact of model resolution and initial/boundary conditions in forecasting flood-causing precipitations. *Atmosphere*, 11(6), 1–20. <https://doi.org/10.3390/atmos11060592>
- Furnari, L., Senatore, A., & Mendicino, G. (2018). Impact of sea surface temperature representation uncertainty on operational hydrometeorological forecasts over coastal Mediterranean catchments impact of sea surface temperature representation uncertainty on operational hydrometeorological forecasts over. (August 2019), 19–20.
- Givati, A., Gochis, D., Rummler, T., & Kunstmann, H. (2016). Comparing one-way and two-way coupled hydrometeorological forecasting systems for flood forecasting in the Mediterranean region. *Hydrology*, 3(2), 19. <https://doi.org/10.3390/hydrology3020019>
- Gochis, D. J., Barlage, M., Cabell, R., Casali, M., Dugger, A., FitzGerald, K., M. McAllister, J. McCreight, A. Rafiee Nasab, L. Read, K. Sampson, D. Yates & Zhang, Y. (2020). The WRF-Hydro modeling system technical description, (version 5.1.1). NCAR Technical Note, 107. [https://ral.ucar.edu/sites/default/files/public/projects/wrf\\_hydro/technical-description-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)
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1–2), 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
- Gurer, I. (1998). Flood disasters and preventative measures in Turkey. *Journal of Natural Disaster Science*, 20(1), 1–9.
- Gurer, I., & Ucar, I. (2009). Flood disasters' inventory in Turkey in 2009. Eleventh International Symposium on Water Management and Hydraulic Engineering, (September), 371–380.
- Hirabayashi, Y., Mahendran, R., Koitala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H., & Kanae, S. (2013). Global flood risk under climate change. *Nature Climate Change*, 3(9), 816–821. <https://doi.org/10.1038/nclimate1911>
- Ivatek-Šahdan, S., Stanešić, A., Tudor, M., Odak Plenković, I., & Janečković, I. (2018). Impact of SST on heavy rainfall events on eastern Adriatic during SOP1 of HyMeX. *Atmospheric Research*, 200 (September), 36–59. <https://doi.org/10.1016/j.atmosres.2017.09.019>
- Jee, J. B., & Kim, S. (2017). Sensitivity study on high-resolution WRF precipitation forecast for a heavy rainfall event. *Atmosphere*, 8(6), 96. <https://doi.org/10.3390/atmos8060096>
- Kaya, O. F., Guler, M., Altan, A., & Yorgancı, I. (2019). 15-16-17 Aralık 2018 Tarihlerinde Antalya İli Kemer/Ovacık Mevkiinde Görülen Aşırı Yağış Hadisesinin Aktüel Hava Haritaları, Uzaktan Algılama ve Sayısal Hava Tahmin Ürünleri ile Sinoptik ve Hidrolojik Analizi. 10. Ulusal Hidroloji Kongresi 9–12.
- Kerandi, N., Arnault, J., Laux, P., Wagner, S., Kitheka, J., & Kunstmann, H. (2018). Joint atmospheric-terrestrial water balances for East Africa: A WRF-hydro case study for the upper Tana River basin. *Theoretical and Applied Climatology*, 131(3–4), 1337–1355. <https://doi.org/10.1007/s00704-017-2050-8>

- Kunstmann, H., & Stadler, C. (2005). High resolution distributed atmospheric-hydrological modelling for alpine catchments. *Journal of Hydrology*, 314(1–4), 105–124. <https://doi.org/10.1016/j.jhydrol.2005.03.033>
- LaCasse, K. M., Splitter, M. E., Lazarus, S. M., & Lapenta, W. M. (2008). The impact of high-resolution sea surface temperatures on the simulated nocturnal Florida marine boundary layer. *Monthly Weather Review*, 136(4), 1349–1372. <https://doi.org/10.1175/2007MWR2167.1>
- Lahmers, T. M., Gupta, H., Castro, C. L., Gochis, D. J., Yates, D., Dugger, A., Goodrich, D., & Hazenberg, P. (2019). Enhancing the structure of the WRF-hydro hydrologic model for semiarid environments. *Journal of Hydrometeorology*, 20(4), 691–714. <https://doi.org/10.1175/JHM-D-18-0064.1>
- Lebeaupin, C., Ducrocq, V., & Giordani, H. (2006). Sensitivity of torrential rain events to the sea surface temperature based on high-resolution numerical forecasts. *Journal of Geophysical Research Atmospheres*, 111(12), 1–19. <https://doi.org/10.1029/2005JD006541>
- Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos*, 89(10), 93–94. <https://doi.org/10.1029/2008EO100001>
- Li, L., Gochis, D. J., Sobolowski, S., & Mesquita, M. D. S. (2017). Evaluating the present annual water budget of a Himalayan headwater river basin using a high-resolution atmosphere-hydrology model. *Journal of Geophysical Research*, 122(9), 4786–4807. <https://doi.org/10.1002/2016JD026279>
- Liu, Y., Liu, J., Li, C., Yu, F., Wang, W., & Qiu, Q. (2020). Parameter sensitivity analysis of the WRF-hydro modeling system for streamflow simulation: A case study in semi-humid and semi-arid catchments of northern China. *Asia-Pacific Journal of Atmospheric Sciences*, 57, 451–466. <https://doi.org/10.1007/s13143-020-00205-2>
- McCabe, G. J., & Wolock, D. M. (2008). Joint variability of global runoff and global sea surface temperatures. *Journal of Hydrometeorology*, 9(4), 816–824. <https://doi.org/10.1175/2008JHM943.1>
- Meredith, E. P., Maraun, D., Semenov, V. A., & Park, W. (2015). Evidence for added value of convection-permitting models for studying changes in extreme precipitation. *Nature*, 175(4449), 238. <https://doi.org/10.1038/175238c0>
- Miglietta, M. M., Mazon, J., Motola, V., & Pasini, A. (2017). Effect of a positive sea surface temperature anomaly on a Mediterranean tornadic supercell. *Scientific Reports*, 7(1), 1–8. <https://doi.org/10.1038/s41598-017-13170-0>
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulation. *Columbia Médica*, 50(3), 885–900. <https://doi.org/10.1234/590>
- Naabil, E., Lamptey, B. L., Arnault, J., Kunstmann, H., & Olufayo, A. (2017). Water resources management using the WRF-hydro modelling system: Case-study of the Tono dam in West Africa. *Journal of Hydrology: Regional Studies*, 12(June), 196–209. <https://doi.org/10.1016/j.ejrh.2017.05.010>
- NCEP, & NOAA. (2014). NCEP SST analysis. <https://polar.ncep.noaa.gov/sst/ophi/verification.shtml>
- NOAA. (2015). NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive. Retrieved from Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory website. <https://doi.org/10.5065/D65D8PWK>
- Papanastasiou, D. K., Melas, D., & Lissaridis, I. (2010). Study of wind field under sea breeze conditions: An application of WRF model. *Atmospheric Research*, 98(1), 102–117. <https://doi.org/10.1016/j.atmosres.2010.06.005>
- Pastor, F., Valiente, J. A., & Palau, J. L. (2018). Sea surface temperature in the Mediterranean: Trends and spatial patterns (1982–2016). *Pure and Applied Geophysics*, 175(11), 4017–4029. <https://doi.org/10.1007/s00024-017-1739-z>
- Pilatin, H. (2020). Sensitivity and sea surface temperature analyses of WRF model for predicting heavy rainfall events observed in eastern Black Sea and Mediterranean regions of Turkey. *Middle East Technical University*.
- Rebora, N., Molini, L., Casella, E., Comellas, A., Fiori, E., Pignone, F., Siccardi, F., Silvestro, F., Tanelli, S., & Parodi, A. (2013). Extreme rainfall in the mediterranean: What can we learn from observations. *Journal of Hydrometeorology*, 14(3), 906–922. <https://doi.org/10.1175/JHM-D-12-083.1>
- Robinson, I., Piollé, J. F., LeBorgne, P., Poulter, D., Donlon, C., & Arino, O. (2012). Widening the application of AATSR SST data to operational tasks through the Medspiration service. *Remote Sensing of Environment*, 116, 126–139. <https://doi.org/10.1016/j.rse.2010.12.019>
- Roxy, M. (2014). Sensitivity of precipitation to sea surface temperature over the tropical summer monsoon region-and its quantification. *Climate Dynamics*, 43(5–6), 1159–1169. <https://doi.org/10.1007/s00382-013-1881-y>
- Ryu, Y., Lim, Y. J., Ji, H. S., Park, H. H., Chang, E. C., & Kim, B. J. (2017). Applying a coupled hydrometeorological simulation system to flash flood forecasting over the Korean peninsula. *Asia-Pacific Journal of Atmospheric Sciences*, 53(4), 421–430. <https://doi.org/10.1007/s13143-017-0045-0>
- Sampson, K., & Gochis, D. (2015). WRF hydro GIS pre-processing tools, Version 2. 2 Documentation. 1–39.
- Senatore, A., Davolio, S., Furnari, L., & Mendicino, G. (2020). Reconstructing flood events in mediterranean coastal areas using different reanalyses and high-resolution meteorological models. *Journal of Hydrometeorology*, 21(8), 1865–1887. <https://doi.org/10.1175/JHM-D-19-0270.1>
- Senatore, A., Furnari, L., & Mendicino, G. (2020). Impact of high-resolution sea surface temperature representation on the forecast of small Mediterranean catchments' hydrological responses to heavy precipitation. *Hydrology and Earth System Sciences*, 24(1), 269–291. <https://doi.org/10.5194/hess-24-269-2020>
- Senatore, A., Mendicino, G., Gochis, D. J., Yu, W., Yates, D. N., & Kunstmann, H. (2015). Fully coupled atmosphere-hydrology simulations for the Central Mediterranean: Impact of enhanced hydrological parameterization for short and long time scales. *Journal of Advances in Modeling Earth Systems*, 7(4), 1693–1715. <https://doi.org/10.1002/2015MS000510>
- Senatore, A., Mendicino, G., Knoche, H. R., & Kunstmann, H. (2014). Sensitivity of modeled precipitation to sea surface temperature in regions with complex topography and coastlines: A case study for the Mediterranean. *Journal of Hydrometeorology*, 15(6), 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., Ziquan, L., Berner, J., Wei Wang, Jordan G. Powers, Michael G. Duda, Dale M. Barker & Huang, X.-Y. (2019). A description of the advanced research WRF model version 4. NCAR technical note NCAR/TN-475+STR. <http://library.ucar.edu/research/publish-technote>
- Sun, M., Li, Z., Yao, C., Liu, Z., Wang, J., Hou, A., Zhang, K., Huo, W., & Liu, M. (2020). Evaluation of flood prediction capability of the WRF-hydro model based on multiple forcing scenarios. *Water (Switzerland)*, 12(3), 874. <https://doi.org/10.3390/w12030874>
- Takahashi, H. G., & Idenaga, T. (2013). Impact of sst on precipitation and snowfall on the sea of Japan side in the winter monsoon season: Time-scale dependency. *Journal of the Meteorological Society of Japan*, 91(5), 639–653. <https://doi.org/10.2151/jmsj.2013-506>

- Team GHRSST. (2010). Black Sea Ultra High Resolution SST L4 Analysis 0.01 deg Resolution|PO.DAAC/JPL/NASA. [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)
- Trenberth, K. E. (1999). Conceptual framework for changes of extremes of the hydrological cycle with climate change. *Climatic Change*, 42, 327–339.
- Trenberth, K. E., & Shea, D. J. (2005). Relationships between precipitation and surface temperature. *Geophysical Research Letters*, 32(14), 1–4. <https://doi.org/10.1029/2005GL022760>
- 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), 363–380.
- Turuncoglu, U. U. (2015). Identifying the sensitivity of precipitation of Anatolian peninsula to Mediterranean and Black Sea surface temperature. *Climate Dynamics*, 44(7–8), 1993–2015. <https://doi.org/10.1007/s00382-014-2346-7>
- Vecchi, G. A., & Harrison, D. E. (2002). Monsoon breaks and subseasonal sea surface temperature variability in the Bay of Bengal. *Journal of Climate*, 15(12), 1485–1493. [https://doi.org/10.1175/1520-0442\(2002\)015<1485:MBASSS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1485:MBASSS>2.0.CO;2)
- Volosciuk, C., Maraun, D., Semenov, V. A., Tilinina, N., Gulev, S. K., & Latif, M. (2016). Rising mediterranean sea surface temperatures amplify extreme summer precipitation in central europe. *Scientific Reports*, 6(February), 1–7. <https://doi.org/10.1038/srep32450>
- Wang, J., Wang, C., Rao, V., Orr, A., Yan, E., & Kotamarthi, R. (2019). A parallel workflow implementation for PEST version 13.6 in high-performance computing for WRF-hydro version 5.0: A case study over the midwestern United States. *Geoscientific Model Development*, 12(8), 3523–3539. <https://doi.org/10.5194/gmd-12-3523-2019>
- Wehbe, Y., Temimi, M., Weston, M., Chaouch, N., Branch, O., Schwitalla, T., Wulfmeyer, V., Zhan, X., Liu, J., & Al Mandous, A. (2019). Analysis of an extreme weather event in a hyper-arid region using WRF-Hydro coupling, station, and satellite data. *Natural Hazards and Earth System Sciences*, 19(6), 1129–1149. <https://doi.org/10.5194/nhess-19-1129-2019>
- Yang, Y., Yuan, H., & Yu, W. (2018). Uncertainties of 3D soil hydraulic parameters in streamflow simulations using a distributed hydrological model system. *Journal of Hydrology*, 567(June), 12–24. <https://doi.org/10.1016/j.jhydrol.2018.09.042>
- Yilmaz, M. T., Anderson, M. C., Zaitchik, B., Hain, C. R., Crow, W. T., Ozdogan, M., Chun, J. A., & Evans, J. (2014). Comparison of prognostic and diagnostic surface flux modeling approaches over the Nile River basin. *Water Resources Research*, 50(1), 386–408. <https://doi.org/10.1002/2013WR014194>
- Yucel, I., & Onen, A. (2014). Evaluating a mesoscale atmosphere model and a satellite-based algorithm in estimating extreme rainfall events in northwestern Turkey. *Natural Hazards and Earth System Sciences*, 14(3), 611–624. <https://doi.org/10.5194/nhess-14-611-2014>
- Yucel, I., Onen, A., Yilmaz, K. K., & Gochis, D. J. (2015). Calibration and evaluation of a flood forecasting system: Utility of numerical weather prediction model, data assimilation and satellite-based rainfall. *Journal of Hydrology*, 523, 49–66. <https://doi.org/10.1016/j.jhydrol.2015.01.042>
- Zhang, X., Cornuelle, B., Martin, A., Ralph, M. & Weihs, R. (2015). West-WRF sensitivity to sea surface temperature boundary condition in California precipitation forecasts of AR related events A33H-2496 model forecast RMS errors conclusions. AGU E-Poster, 10.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Kilicarslan, B. M., Yucel, I., Pilatin, H., Duzenli, E., & Yilmaz, M. T. (2021). Improving WRF-Hydro runoff simulations of heavy floods through the sea surface temperature fields with higher spatio-temporal resolution. *Hydrological Processes*, 35(9), e14338. <https://doi.org/10.1002/hyp.14338>