Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

Geoscientific Model Development Discussions

A regional coupled ocean–atmosphere modeling framework (MITgcm–WRF) using ESMF/NUOPC: description and preliminary results for the Red Sea

Rui Sun[1], Aneesh Subramanian[1], Art Miller[1], Matthew Mazloff[1], Ibrahim Hoteit[2], and Bruce Cornuelle[1]

[1]Scripps Institution of Oceanography, La Jolla, California, USA
[2]Physical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

Correspondence: Rui Sun (rus043@ucsd.edu); Aneesh Subramanian (acsubram@ucsd.edu)

Abstract. A new regional coupled ocean–atmosphere model is developed to study air–sea feedbacks. The coupled model is based on two open-source community model components: (1) MITgcm ocean model; (2) Weather Research and Forecasting (WRF) atmosphere model. The coupling between these components is performed using ESMF (Earth System Modeling Framework) and implemented according to National United Operational Prediction Capability (NUOPC) consortium. The regional coupled model allows affordable simulation where oceanic mixed layer heat and momentum interact with atmospheric boundary layer dynamics at mesoscale and higher resolution. This can capture the feedbacks which are otherwise not well-resolved in coarse resolution global coupled models and are absent in regional uncoupled models. To test the regional coupled model, we focus on a series of heat wave events that occurred on the eastern shore of the Red Sea region in June 2012 using a 30-day simulation. The results obtained using the coupled model, along with those in forced uncoupled ocean or atmosphere model simulations, are compared with observational and reanalysis data. All configurations of coupled and uncoupled models have good skill in modeling variables of interest in the region. The coupled model shows improved skill in temperature and circulation evaluation metrics. In addition, a scalability test is performed to investigate the parallelization of the coupled model. The results indicate that the coupled model scales linearly for up to 128 CPUs and sublinearly for more processors. In the coupled simulation, the ESMF/NUOPC interface also scales well and accounts for less than 10% of the total computational resources compared with uncoupled models. Hence this newly developed regional model scales efficiently for a large number of processors and can be applied for high-resolution coupled regional modeling studies.

1 Introduction

Accurate and efficient forecasting of oceanic and atmospheric circulations is essential for a wide variety of high-impact societal needs, including: extreme weather and climate events (Kharin and Zwiers, 2000; Chen et al., 2007); environmental protection and coastal management (Warner et al., 2010); management of fisheries (Roessig et al., 2004), marine conservation (Harley et al., 2006); water resources (Fowler and Ekström, 2009); and renewable energy (Barbariol et al., 2013). Effective forecasting relies on high model fidelity and accurate initialization of the models with observed state of the ocean–atmosphere coupled

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

system. Although global coupled models are being implemented with increased resolution, high-resolution regional coupled models, if properly driven by the boundary conditions, can contribute additional air–sea feedback information to the study of small-scale processes.

A number of regional coupled ocean–atmosphere models have been developed for various goals in the past decades. An early example of building a regional coupled model for realistic simulations focused on accurate weather forecasting in the Baltic Sea (Gustafsson et al., 1998; Hagedorn et al., 2000; Doscher et al., 2002), and suggested that the coupled model improved the SST (Sea Surface Temperature) and atmospheric circulation forecast. Enhanced numerical stability in the coupled simulation was also observed. These early attempts were followed by other practitioners in basin-scale climate simulations (e.g. Huang et al., 2004; Aldrian et al., 2005; Xie et al., 2007; Seo et al., 2007; Somot et al., 2008; Fang et al., 2010; Boé et al., 2011; Gualdi et al., 2013; Chen and Curcic, 2016; Seo, 2017). For instance, Huang et al. (2004) implemented a regional coupled model to study three major important patterns contributing to the variability and predictability of the Atlantic climate. The study suggested that these patterns originate from air–sea coupling within the Atlantic Ocean or by the oceanic responses to atmospheric internal forcing. Seo et al. (2007) studied the nature of ocean–atmosphere feedbacks in the presence of oceanic mesoscale eddy fields in the eastern Pacific Ocean sector. The evolving SST fronts were shown to drive an unambiguous response of the atmospheric boundary layer in the coupled model, and lead to model anomalies of wind stress curl, wind stress divergence, surface heat flux, and precipitation that resemble observations. This study helped substantiate the importance of ocean–atmosphere feedbacks involving oceanic mesoscale variability features.

On top of basin-scale climate simulations, regional coupled models are also used to study weather extremes. For example, the COAMPS (Coupled Ocean/Atmosphere Mesoscale Prediction System) was applied to investigate idealized tropical cyclone events (Hodur, 1997). This work was then followed by other realistic extreme weather studies. For example, extreme bora wind events in the Adriatic Sea were investigated using different regional coupled models (Loglisci et al., 2004; Pullen et al., 2006; Ricchi et al., 2016). The coupled simulation results demonstrated improvements in describing the air–sea interaction processes by taking into account ocean heat fluxes and wind surface wave effects (Loglisci et al., 2004; Ricchi et al., 2016). It was also found in model simulations that SST after bora wind events had a stabilizing effect on the atmosphere, reducing the atmospheric boundary layer mixing and yielding stronger near-surface wind (Pullen et al., 2006). The regional coupled models was also used by Bender and Ginis (2000); Chen et al. (2007); Warner et al. (2010) for improving the tracking of the hurricane path and intensity, predicting SST variation, and forecasting wind speeds.

Regional coupled modeling systems also play important roles in studying the effect of surface variables (e.g., surface evaporation, precipitation, surface roughness) in the coupling process of ocean or lakes. One example is the study conducted by Powers and Stoelinga (2000), who developed a coupled model and investigated the frontal passage over the Lake Erie region. Sensitivity analysis was performed to demonstrate that taking into account lake surface roughness parameterization in the atmosphere model can improve the calculation of wind stress and heat flux. Another example is the investigation by Turuncoglu et al. (2013), who compared a regional coupled model with uncoupled models, and demonstrated the improvement of the coupled model in capturing the response of the Caspian Sea level to climate variability.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

Despite the existing regional coupled ocean–atmosphere models, it is still desirable to develop a new coupled regional ocean–atmosphere model with 'state-of-the-art' physics and using modern coupling framework. The goal of this work is to (1) introduce the design of a newly developed regional coupled ocean–atmosphere modeling system, (2) describe the implementation of the modern coupling framework, (3) present preliminary simulation results in the Red Sea region, and (4) demonstrate and discuss the parallelization of the coupled model. In the coupled system, the ocean model component is the MIT general circulation model (MITgcm) (Marshall et al., 1997); the atmosphere model component is the Weather Research and Forecasting (WRF) model (Skamarock et al., 2005). To couple the model components in the present work, the Earth System Modeling Framework (ESMF) (Hill et al., 2004) is used because of its advantages in conservative re-gridding capability, calendar management, logging and error handling, and parallel communications. The National United Operational Prediction Capability (NUOPC) layer in ESMF is also used (Sitz et al., 2017). The additional NUOPC wrapper layer between coupled model and ESMF simplifies the implementations of component synchronization, execution, and other common tasks in the coupling. To test the coupled model, we focus on a series of heat wave events that occurred on the eastern shore of the Red Sea region in June 2012. The simulated surface variables of the Red Sea (e.g., sea surface temperature, 2-m temperature, and surface heat fluxes) are examined and validated against available observational and reanalysis data. To assess the improvements gained from the coupled simulation, the results are compared with those obtained using stand-alone ocean or atmosphere model. In addition, a scalability test of the coupled model is performed to investigate its parallel capability.

The rest of this paper is organized as follows: the description of the individual modeling components and the design of the coupled modeling system are detailed in Section 2. Section 3 introduces the experiment design, observational data, and analysis methodology. Section 4 discusses the results obtained from the coupled model. Section 5 details the parallelization test of the coupled model. The last section concludes the paper and presents an outlook for future work.

2 Model Description

The newly developed regional coupled modeling system is introduced in this section. The general design of the coupled model, descriptions of individual components, and ESMF/NUOPC coupling framework are presented below.

2.1 General design

The schematic description of the coupled model is shown in Fig. 1(a). The coupled model is comprised of five components: ocean component MITgcm, atmosphere component WRF, MITgcm–ESMF interface, WRF–ESMF interface, and ESMF/NUOPC coupler. They are to be detailed in the following sections.

The coupler component runs in both directions: (1) from WRF to MITgcm, and (2) from MITgcm to WRF. From WRF to MITgcm, the coupler collects the surface atmospheric variables (i.e., surface temperature, pressure, mixing ratio, wind velocity components, precipitation, longwave and shortwave radiations) from the atmosphere component and updates the surface forcing variables (heat flux, wind stress, freshwater flux) to drive the ocean component. From MITgcm to WRF, the coupler collects SST and ocean surface velocity from the ocean component and uses them as the surface boundary condition

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

in the atmosphere component. Re-gridding the data from either model component will be performed by the coupler, in which various coupling intervals and schemes can be specified (Hill et al., 2004).

2.2 MITgcm Ocean Model

The MITgcm (Marshall et al., 1997) is a 3-D, finite-volume, general circulation model used by a broad community of researchers for a wide range of applications at various spatial and temporal scales. The model code and documentation, which are under continuous development, are available on the MITgcm webpage http://mitgcm.org/. The 'Checkpoint 66h' (June 2017) version of MITgcm is used in the present work.

The MITgcm is designed to run on high-performance computing (HPC) platforms and can run in non-hydrostatic and hydrostatic modes. It integrates the primitive (Navier-Stokes) equations, under the Boussinesq approximation, using finite volume method on a staggered 'Arakawa C-grid'. The MITgcm uses modern physical parameterization schemes for subgrid-scale horizontal and vertical mixing and tracer properties. The code configuration includes build-time C pre-processor (CPP) options and run-time switches, which allow for great computational modularity in MITgcm to study a variety of oceanic phenomena (Evangelinos and Hill, 2007).

To implement the MITgcm–ESMF interface, we separated the MITgcm main program into three subroutines that handle initialization, running, and finalization, shown in Fig. 1(b). These subroutines are used by the ESMF/NUOPC coupler that controls the ocean component in the coupled run. The surface boundary fields on the ocean surface is exchanged online[1] via the MITgcm–ESMF interface during the simulation. The MITgcm SST and ocean surface velocity are the export boundary fields, and the atmospheric surface forcing variables are the import boundary fields (see Fig. 1(b)). These boundary fields are registered in the coupler following NUOPC consortium and timestamps[2] are added to them for the coupling. In addition, MITgcm grid information is also provided for online re-gridding of the exchanged boundary fields. To carry out the high-resolution simulation, the MITgcm–ESMF interface runs in parallel via MPI communications. The implementations of the present MITgcm–ESMF interface is based on the baseline MITgcm–ESMF coupler (Hill, 2005), but we updated it to couple the modern version ESMF/NUOPC with MITgcm. We also modified the baseline coupler to receive atmosphere surface fluxes and send ocean surface variables (i.e., SST and ocean surface velocity).

2.3 WRF Atmospheric Model

The Weather Research and Forecasting (WRF) Model (Skamarock et al., 2005) is developed by NCAR/MMM (Mesoscale and Microscale Meteorology Division). It is a 3-D, finite-difference atmospheric model with a variety of physical parameterizations of sub-grid scale processes for predicting a broad spectrum of applications. WRF is used extensively for operational forecasts (http://www.wrf-model.org/plots/wrfrealtime.php) as well as realistic and idealized dynamical studies.

[1] In this manuscript, 'online' means the manipulations are performed via subroutine calls during the execution of the simulations; 'offline' means the manipulations are performed when the simulations are not executing.

[2] In ESMF, 'timestamp' is a sequence of number, usually based on the time, to identify the ESMF fields. Only the ESMF fields having the correct timestamp will be transferred in the coupling.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

Geoscientific Model Development Discussions

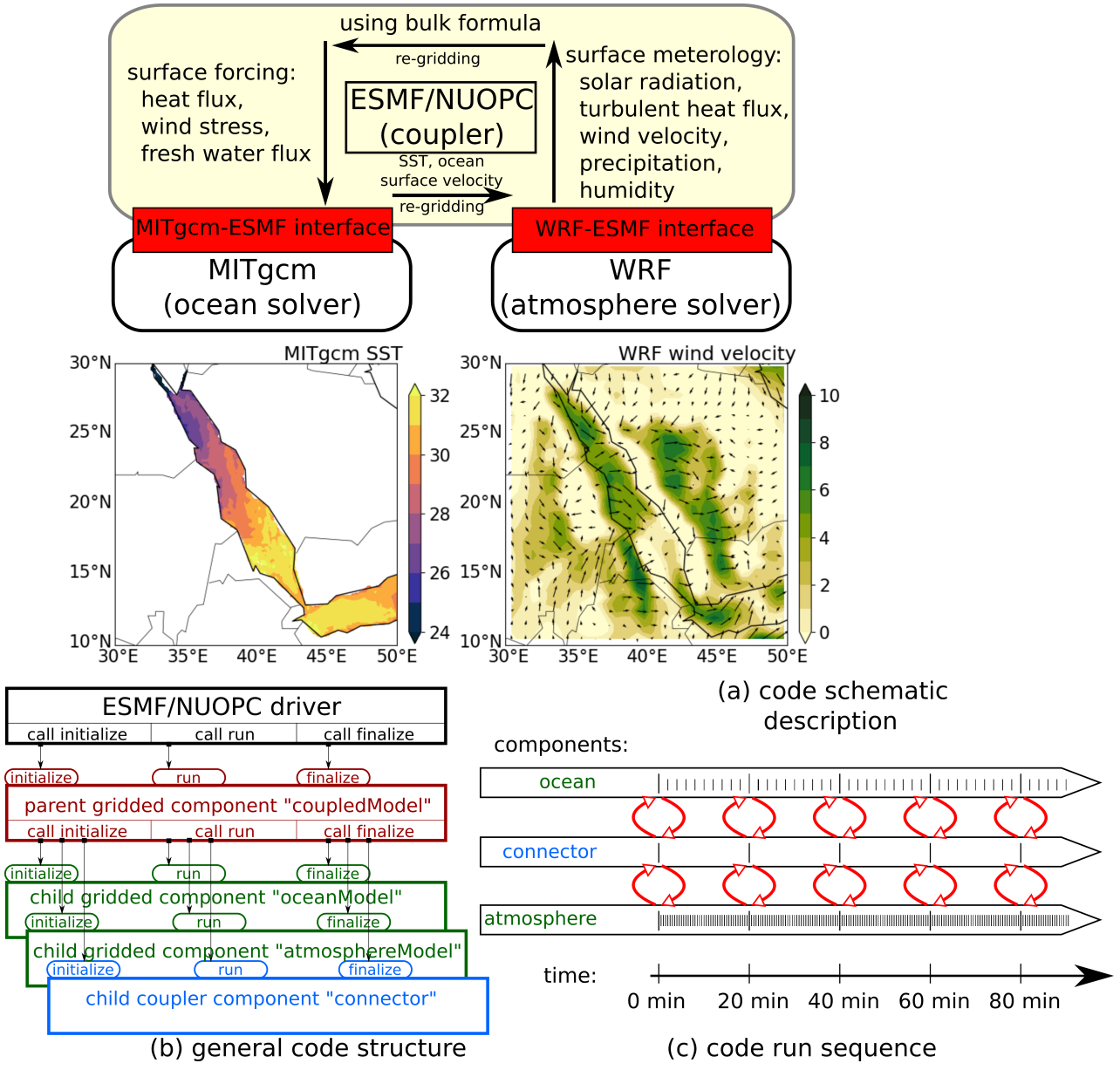


Figure 1. The schematic description, general code structure and run sequence of the coupled ocean–atmosphere model. In panel (a), the white blocks are the ocean and atmosphere components; the red blocks are the implemented MITgcm–ESMF and WRF–ESMF interfaces; the yellow block is the ESMF/NUOPC coupler. In panel (b), the black block is the *application driver*; the red block is the *parent gridded component* called by the *application driver*; the green/blue blocks are the *child gridded/coupler components* called by the *parent gridded component*. In panel (c), each horizontal arrow indicates the time axis of each component; the ticks on the time axis indicate the time step; the boundary condition fields are updated at each coupling interval in the connector.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

In the present work, the Advanced Research WRF dynamic version (WRF-ARW, version 3.9.1.1) is used. It solves the compressible Euler non-hydrostatic equations, and also includes a run-time hydrostatic option. The WRF-ARW uses a terrain-following hydrostatic pressure coordinate system in the vertical direction and utilizes the 'Arakawa-C grid'. WRF incorporates various physical processes including microphysics, cumulus parameterization, planetary boundary layer, surface layer, land surface, and longwave and shortwave radiations, with several options available for each process.

Similar with the implementations in MITgcm, WRF is also separated into initialization, run, and finalization subroutines to enable the WRF–ESMF interface to control the atmosphere model during the coupled simulation, shown in Fig. 1(b). The implementation of the present WRF–ESMF interface is based on the prototype interface (Henderson and Michalakes, 2005). In the present work, the prototype WRF–ESMF interface is updated to a modern version of WRF-ARW and a modern version of ESMF, based on the NUOPC layer. This prototype interface is also expanded to interact with the ESMF/NUOPC coupler to receive the ocean surface variables and send the atmosphere surface fluxes. The surface boundary condition fields are registered in the coupler following the NUOPC consortium with timestamps. The WRF grid information is also provided for online re-gridding by ESMF. To carry out the high-resolution simulation, the WRF–ESMF interface also runs in parallel via MPI communications.

2.4 ESMF/NUOPC Coupler

The coupler is implemented using ESMF version 7.0.0. The ESMF is selected because of its high-performance and flexibility for building and coupling weather, climate, and related Earth science applications (Collins et al., 2005; Turuncoglu et al., 2013; Chen and Curcic, 2016; Turuncoglu and Sannino, 2017). It has a superstructure for representing the model and coupler components and an infrastructure of commonly used utilities, including conservative grid remapping, time management, error handling, and data communications.

The general code structure of the coupler is shown in Fig. 1(b). In the main program, an ESMF/NUOPC driver is created and controls the ESMF component 'coupledModel', which is also the parent component. When the main program calls the parent component, the parent component cascades the calls to the child components and the coupler components. In the present work, the ESMF child gridded ocean and atmosphere components are connected via the ESMF gridded coupler component. The two-way interpolation and data transfer are performed using the coupler component, and ESMF supports different re-gridding approaches or unit conventions (Hill et al., 2004). The ESMF gridded and coupler components can be run in parallel as a group of Persistent Execution Threads (PETs), which are single processing units (i.e. CPU, GPU) defined by ESMF, and the PETs can be created by the user in a flexible way for parallelization. In the present work, PETs are created according to the grid decomposition, and each PET is associated with an MPI process running on a separate processor. ESMF also allows the PETs running in sequential mode, concurrent mode, or a mixed mode. In sequential mode, a set of ESMF gridded/coupler components runs in sequence on the same set of PETs; in concurrent mode, the gridded components are created and run on mutually exclusive sets of PETs, and are coupled by a coupler component. We selected the sequential mode in the implementations.

In ESMF, the gridded components are used to represent models and coupler components are used to connect these models. The interfaces and data structures in ESMF have few constraints, providing the flexibility to be adapted to many modeling

systems. However, the flexibility of the gridded components can limit the interoperability across different modeling systems. To address this issue, the NUOPC layer is developed to provide the coupling conventions and the generic representation of the model components (e.g. drivers, models, connectors, mediators). The NUOPC layer in the present coupled model is implemented according to the documentations (Hill et al., 2004; Theurich et al., 2016), and the ocean/atmosphere component each has:

1. Prescribed variables for NUOPC to link the components;
2. The entry point for registration of the components;
3. An *InitializePhaseMap* which describes a sequence of standard initialization phases, including advertising the fields that a component can provide, checking and mapping the fields to each other, and initializing the fields that will be used;
4. A *RunPhaseMap* that checks the incoming clock of the driver, examines the timestamps of incoming fields, and runs the component;
5. Timestamps on exported fields consistent with the internal clock of the component;
6. The *finalization method* to clean up all allocations.

The subroutines that handle initialization, running, and finalization in MITgcm and WRF will be included in the *InitializePhaseMap*, *RunPhaseMap*, and *finalization method* in the NUOPC layer, respectively.

3 Experiment Design and Observational Datasets

To test the coupled model, we applied it to study a series of heat wave events in the Red Sea region. The simulation of the Red Sea extends from 0000 UTC 01 June 2012 to 0000 UTC 01 July 2012. We select this month because of the record-high surface air temperature observed in the Makkah region, located 70 km inland from the eastern shore of the Red Sea (Abdou, 2014).

The computational domain and bathymetry are shown in Fig. 2. The model domain is centered at 20° N and 40° E, and the bathymetry is from the 2-minute Gridded Global Relief Data (ETOPO2) (National Geophysical Data Center, 2006). WRF is implemented using a horizontal grid of 256×256 points and grid spacing of 8 km, with 40 terrain-following vertical levels, more closely spaced in the atmospheric boundary layer. The time step for atmosphere simulation is 30 seconds. The updated version of the Kain–Fritsch convection scheme (Kain, 2004) is used with the modifications to include the updraft formulation, downdraft formulation, and closure assumption. The Yonsei University (YSU) scheme (Hong et al., 2006) is used for the planetary boundary layer (PBL), and the Rapid Radiation Transfer Model for GCMs (RRTMG; Iacono et al. (2008)) is used for longwave and shortwave radiation transfer through the atmosphere. The Rapid Update Cycle (RUC) land surface model is used for the land surface processes (Benjamin et al., 2004). The MITgcm uses the same horizontal grid spacing as WRF, with 40 vertical z-levels that are more closely spaced near the surface. The time step of the ocean model is 120 seconds. The horizontal sub-grid mixing is parameterized using nonlinear Smagorinsky viscosities, and the K-profile parameterization (KPP) (Large et al., 1994) is used for vertical mixing processes.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

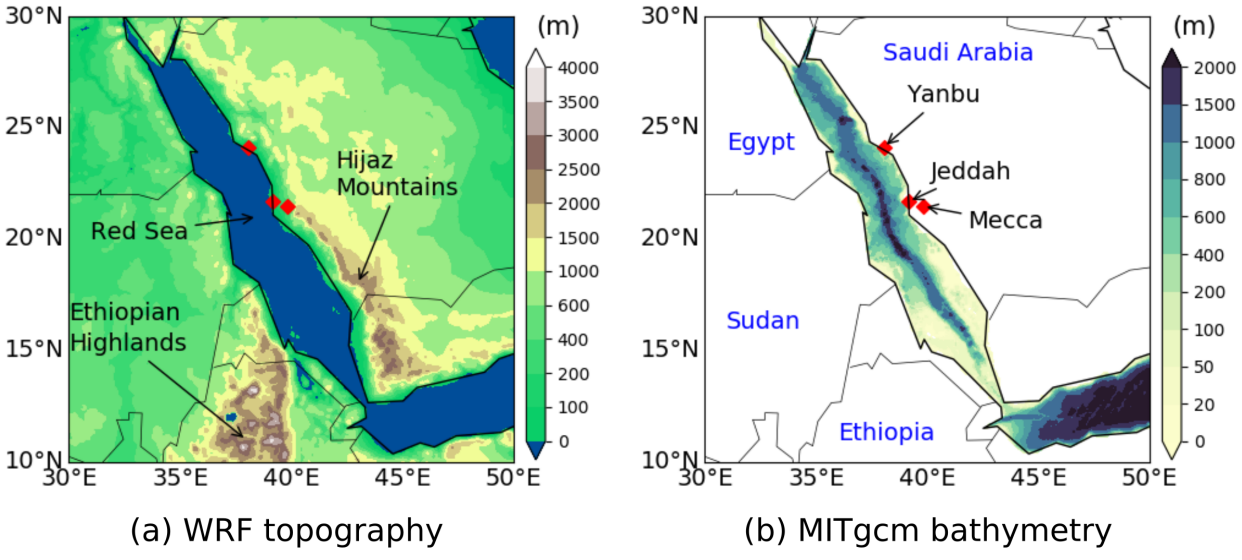


Figure 2. The WRF topography and MITgcm bathymetry in the simulations. Three major cities near the eastern shore of Red Sea are highlighted.

To study the air–sea interactions, the following sets of simulations using different surface forcings are performed:

1. Run CPL: a two-way coupled MITgcm–WRF simulation. The coupling interval is 20 minutes. This run tests the performance of the high-resolution two-way coupled ocean–atmosphere model. The atmosphere is initialized using the ECMWF ERA5 reanalysis dataset, with a grid resolution of 30 km (Hersbach, 2016). The same data also provide the boundary conditions for air temperature, wind speed, and air humidity every six hours. The ocean model uses the assimilated HYCOM/NCODA $1/12^\circ$ global analysis data (http://hycom.org/data-server/glb-analysis) as initial and boundary conditions for ocean temperature, salinity, and horizontal velocities. The boundary conditions for the ocean is updated on a daily basis. The initial condition, boundary condition, and forcing terms of this run are summarized in Table 1.

2. Run ATM.STA: a stand-alone WRF simulation with its initial HYCOM/NCODA SST kept constant throughout the simulation. This run allows to access the WRF model behavior with realistic, but static SST, and serves as a benchmark.

3. Run ATM.DYN: a stand-alone WRF simulation with the SST forcing prescribed using daily HYCOM/NCODA SST. The HYCOM/NCODA SST dataset is selected because it also provides the oceanic boundary condition in the CPL run. This allows accessing the WRF model behavior with updated sea surface temperature.

4. Run OCN.DYN: a stand-alone MITgcm simulation forced by the ERA5 dataset. The bulk formula in MITgcm is used to derive the turbulent heat fluxes. The ERA5 dataset is used because it also provides the atmospheric boundary condition in the CPL run. This run accesses the MITgcm model behavior with prescribed lower-resolution atmospheric surface forcing.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

Geoscientific Model Development Discussions
Open Access
EGU

Table 1. The initial condition, boundary condition and forcing terms used in present simulations.

run	initial and boundary conditions	ocean surface conditions	atmospheric forcings
CPL	ERA5 (atmosphere) HYCOM/NCODA (ocean)	from MITgcm	from WRF
ATM.STA	ERA5	HYCOM/NCODA initial condition kept constant	N.A.
ATM.STA	ERA5	HYCOM/NCODA updated every 24 hours	N.A.
OCN.DYN	HYCOM/NCODA	N.A.	ERA5 + MITgcm bulk formula

The analysis of the results focuses on temperature, heat flux, surface wind, and evaporation. The simulated SST data are validated against the OSTIA (Operational Sea Surface Temperature and Sea Ice Analysis) system in GHRSST (Group for High Resolution Sea Surface Temperature) (Donlon et al., 2012; Martin et al., 2012), and the simulated 2-meter air temperature (T2) is validated against the ECMWF ERA5 dataset. To evaluate the modeling of the heat wave event in three major cities near the eastern shore of Red Sea, the diurnal temperature variation is compared with observed daily maximum and minimum temperatures from NOAA National Climate Data Center (NCDC climate data online at http://cdo.ncdc.noaa.gov/CDO/georegion). Surface heat fluxes (e.g., latent heat, sensible heat, longwave and shortwave radiations), which are important for ocean–atmosphere interactions, are compared with MERRA-2 (Modern-Era Retrospective analysis for Research and Applications, version 2) datasets (Gelaro et al., 2017). The MERRA-2 dataset is selected because it is an independent data compared to the initial and boundary conditions used in the simulations. The data used for validation are outlined in Table 2.

Table 2. The dataset used to validate the simulation results.

variable	validation data
sea surface temperature (SST)	GHRSST and HYCOM
2-meter air temperature (T2)	ERA5 and NCDC climate data
turbulent heat fluxes	MERRA-2
solar radiations	MERRA-2
surface wind	MERRA-2
surface evaporation	MERRA-2

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

4 Results and Discussions

The Red Sea is an elongated basin covering the area between 12-30°N and 32-43°E. The basin is 2250 km long, extending from the Suez and Aqaba gulfs in the north to the strait of Bal el-Mandeb in the south, which connects the Red Sea and the Indian Ocean. Since the global models with coarse resolution cannot properly resolve local features in the narrow basin of the Red Sea (Yao et al., 2014b, a; Zhan et al., 2014), regional models with relatively higher resolutions can be used as dynamical downscaling tools for extreme temperature studies (Li et al., 2018). In this section, results of high-resolution simulations using different model configurations will be presented and examined to assess the performance of the coupled model in simulating the heat wave events in the Red Sea region.

4.1 2-meter Air Temperature (T2)

We begin our analysis by examining the simulated T2 from various experiments. The simulation results obtained from coupled (CPL) run, the ERA5 data, and their associated difference are shown in Fig. 3 after 36 hours and 48 hours. It can be seen in Fig. 3(I) that the CPL run captures the heat wave event in the Red Sea region on June 2nd, compared with the ERA5 dataset in Fig. 3(II). Since ERA5 air temperature data are in good agreement with the NCDC ground observation data in the Red Sea region (comparison not shown), we use ERA5 data to validate the simulation results. The difference between the CPL run and ERA5 dataset is shown in Fig. 3(III). The ATM.STA and ATM.DYN simulation results have consistent patterns with the CPL run results and thus are not shown, but their differences with respect to the ERA5 data are shown in Fig. 3(IV) and 3(V), respectively. Fig. 3(VI) to 3(X) show the same results after 48 hours. It can be seen in Fig. 3 that all simulations reproduce the T2 patterns over the Red Sea region reasonably well compared with the ERA5 data. The mean T2 differences over the sea are -1.55 °C (CPL), -1.66 °C (ATM.STA), and -1.70 °C (ATM.DYN) after 36 hours, and -0.99 °C (CPL), -1.10 °C (ATM.STA), and -1.12 °C (ATM.DYN) after 48 hours. The T2 over the Hijaz Mountains (see Fig. 2(a)) is under-estimated by more than 4 °C in all simulations after 48 hours. This is likely because the simulations on T2 suffer from the mismatches between the model terrain and the actual terrain, especially over complex mountains (Zhang et al., 2013a). The diurnal T2 variation in the simulations is also shown in the snapshots. All simulations can capture the diurnal variation of the T2 in the Red Sea region, and this will be further discussed later in this section.

The simulation results for the heat wave events on June 10th and 24th are shown in Fig. 4 to demonstrate the performance of the coupled model over longer periods of time. It can be seen in Fig. 4(III) and 4(VIII) that the T2 patterns simulated by the coupled run are consistent with the ERA5 dataset. The differences between ATM.STA and ATM.DYN simulation results with respect to the ERA5 data are shown in Fig. 4(IV), 4(V), 4(IX), and 4(X), respectively. It can be seen that the T2 over the sea in CPL simulation has a much smaller difference with the validation ERA5 data (10th: -1.02 °C; 24th: -0.84 °C) compared with the ATM.STA run (10th: -1.56 °C; 24th: -2.13 °C). The CPL run results are closer to the ERA5 dataset because the ocean component (MITgcm) is providing updated SST, which warms the T2; the ATM.STA run uses a constant cooler SST from June 1st, and the T2 is determined by the constant cooler SST. On the other hand, when comparing the CPL run with the ATM.DYN

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

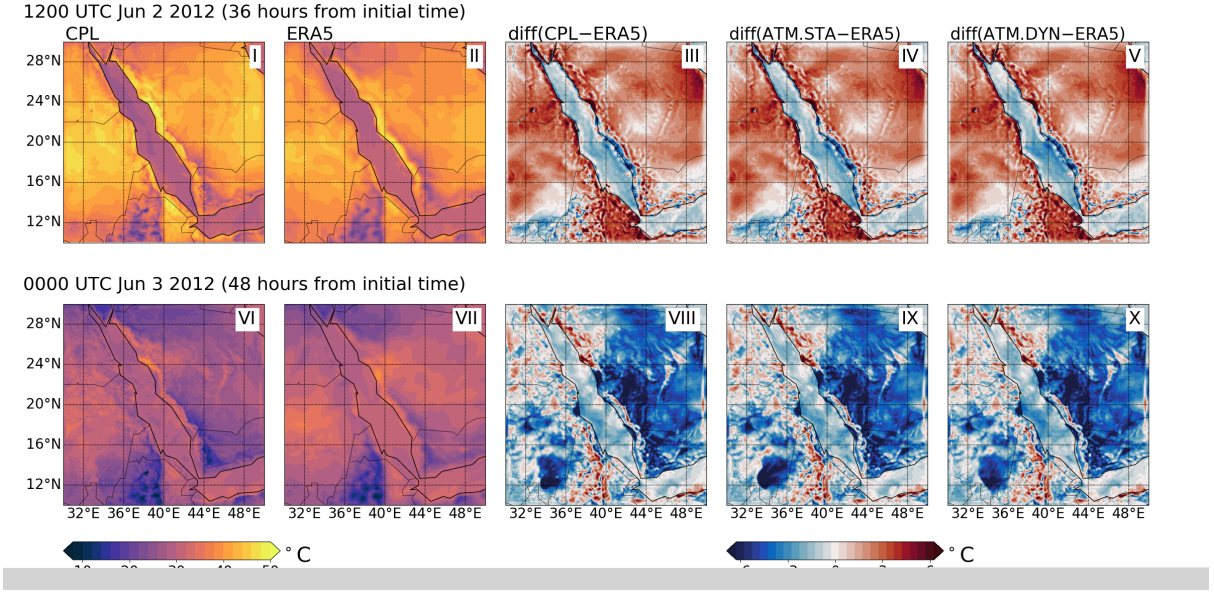


Figure 3. The surface air temperature as obtained from the CPL run, the ERA5 data, and their difference (CPL−ERA5). The difference between ATM.STA and ATM.DYN with the ERA5 data (i.e., ATM.STA−ERA5, ATM.DYN−ERA5) are also presented. The simulation initial time is 0000 UTC Jun 01 2012 for both snapshots. Two snapshots are selected: (1) 1200 UTC Jun 02 2012 (36 hours from initial time); (2) 0000 UTC Jun 03 2012 (48 hours from initial time).

run on June 24th, the difference is very small (-0.10 °C on June 24th). This is because the SST fields from CPL and ATM.DYN runs are similar, which means that the SST in CPL run is tending to be similar to the realistic.

To investigate the diurnal T2 variation in Fig. 3, the time series of T2 in three major cities as simulated in CPL and ATM.STA runs are plotted in Fig. 5, starting from June 1st. The ATM.DYN run results are similar with the CPL run results and thus are not shown. To validate the simulation results, the time series in ERA5 data and the daily observed high/low temperature data from NOAA National Climate Data Center are also plotted. It can be seen that four major heat waves (i.e., June 2nd, 10th, 17th, and 24th) and the T2 variations during the 30-day simulation are all captured by the simulations. Before June 17th (lead time $<$ 16 days), the CPL and WRF.STA runs results are in good agreement with the ground observation and ERA5 dataset. The root mean square error (RMSE) between the simulations and ground observation are 2.79 °C and 2.83 °C for CPL and WRF.STA runs, respectively. However, the error after June 18th (simulation lead time $>$ 17 days) is larger for both CPL (3.42 °C) and WRF.STA (3.94 °C) runs. It can be also seen that the CPL run better captures the daily high temperatures in Yanbu (RMSE difference: 2.77 °C) than ERA5 dataset (RMSE: 5.59 °C), which is probably because ERA5 uses a lower resolution grid and is unable to capture the T2 in the coastal city. This is one of the advantages when employing high-resolution regional simulations. It should be mentioned that both the present simulations and ERA5 dataset reported a T2 that is 4.5 °C lower than observed T2 in Mecca on June 2nd, though the heat wave events in the other cities are still captured. This may be due to the errors in initial conditions, or WRF micro-physics models (e.g., land surface model, the PBL model) are unable to parameterize this extreme

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

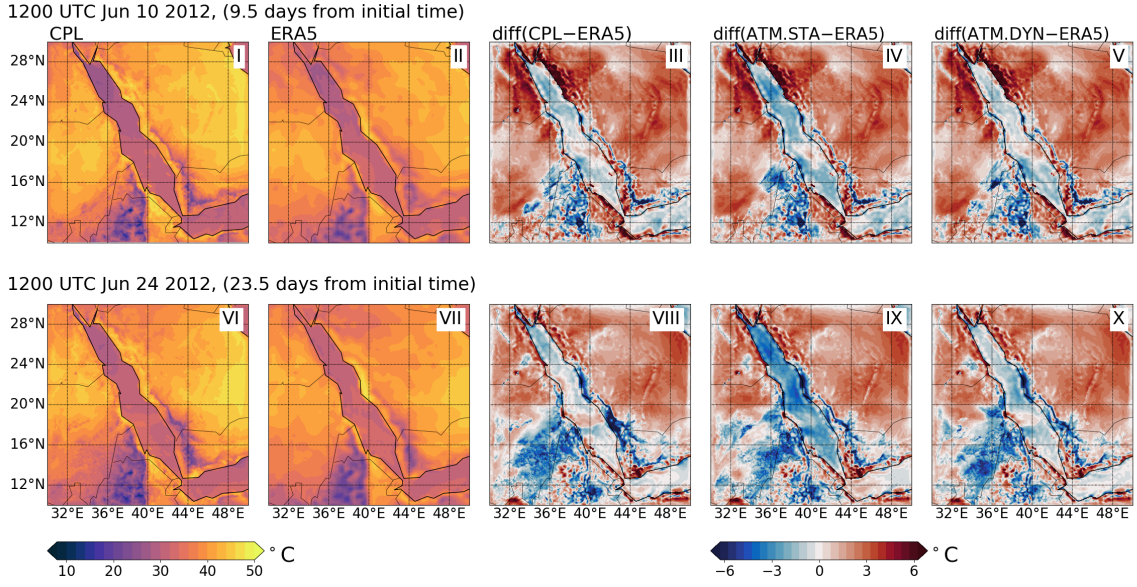


Figure 4. The surface air temperature as obtained from the CPL run, the ERA5 data, and their difference (CPL−ERA5). The difference between ATM.STA and ATM.DYN with the ERA5 data (i.e., ATM.STA−ERA5, ATM.DYN−ERA5) are also presented. The simulation initial time is 0000 UTC Jun 01 2012 for both snapshots. Two snapshots are selected: (1) 1200 UTC Jun 10 2012 (9.5 days from initial time); (2) 1200 UTC Jun 24 2012 (23.5 days from initial time).

event. It can be also seen in the results that taking into account ocean–atmosphere coupling can improve the simulation of T2 in the CPL run. In Fig. 5, the CPL run can better reproduce the evolution of the T2 compare to ATM.STA run during the 30-day simulation: the CPL run better captures the daily high/low temperature in Yanbu and Jeddah (RMSE: 2.69 and 2.81 °C) than ATM.STA run (RMSE: 3.04 and 3.28 °C). However, the difference of T2 in Mecca is negligible (0.05 °C) between CPL and ATM.STA runs. We hypothesize that Mecca is much further away from the Red Sea than Yanbu and Jeddah, which indicates that the influence of air–sea coupling is strong near the coast.

The simulation error of T2 also oscillates diurnally in the present simulations. To demonstrate the diurnal variation of the simulation error quantitatively, the mean deviation and RMSE of T2 between the simulations (i.e., ATM.STA, ATM.DYN, and CPL) and ERA5 data are shown in Fig. 6. To highlight the air–sea interactions in the simulations, only the temperature over the Red Sea is compared. It can be seen in Fig. 6 that the ATM.STA run using the static SST can still capture the T2 patterns in the first week, but it under-predicts T2 by 2.5 °C because of ignoring the SST evolution. On the other hand, CPL run has much smaller mean deviation (-0.49 °C) and root mean square error (1.46 °C) compared with those in ATM.STA run (deviation: -1.34 °C; RMSE 2.04 °C) during the 30-day simulation as the SST evolution is considered. The ATM.DYN run uses the prescribed SST and its results are consistent with those in CPL run (deviation: -0.58 °C; RMSE 1.40 °C), indicating that the coupled model captures the SST revolution. The mean deviation and RMSE of T2 in the present work are similar to those in the benchmark WRF-ARW simulations (Xu et al., 2009; Zhang et al., 2013a). The differences between the present

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

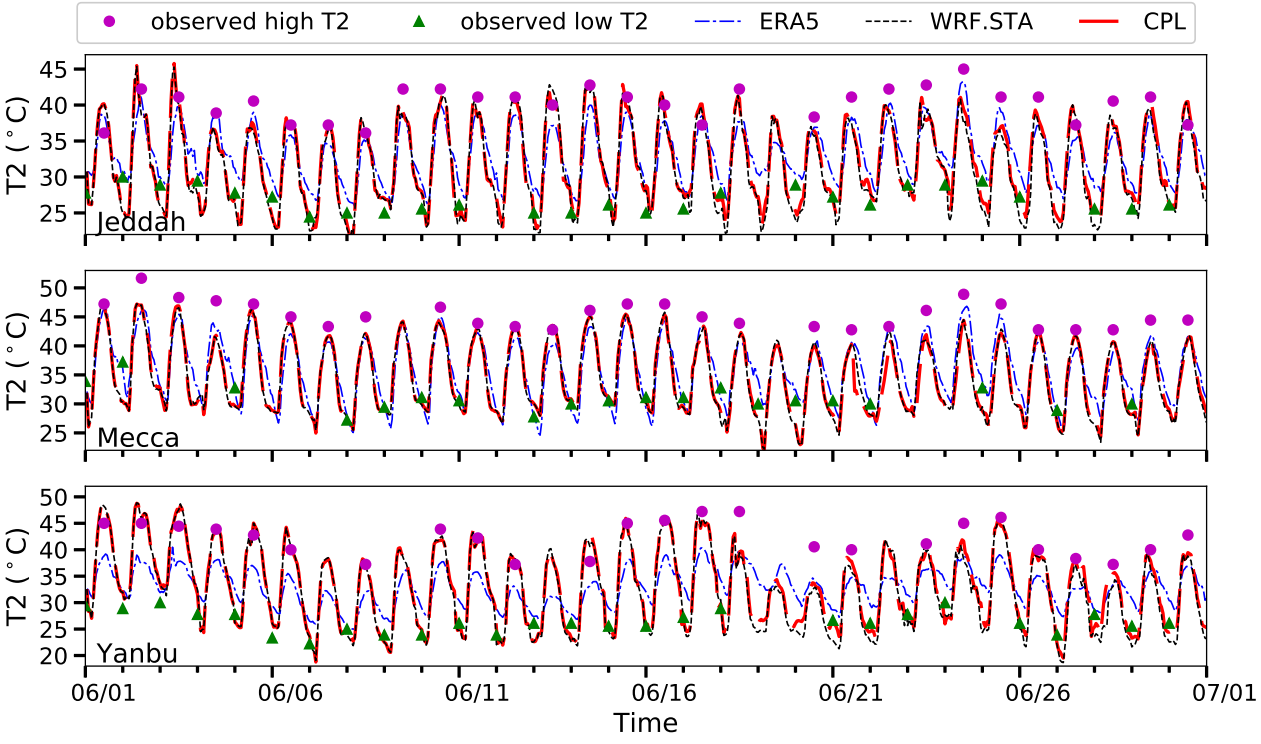


Figure 5. Comparison of the surface air temperature at three major cities near the eastern shore of Red Sea (Jeddah, Mecca, Yanbu) as resulting from CPL and WRF.STA runs. The ATM.DYN run results are similar with the CPL run results and thus are not shown. The temperature data are compared with the time series in ERA5 dataset and daily high/low temperature in the NOAA national data center dataset.

simulations and ERA5 data are plotted below the mean deviation and RMSE to demonstrate the improvement of the CPL run over ATM.STA and ATM.DYN runs. It can be seen that the CPL run captures improved T2 patterns in both mean deviation and RMSE than the ATM.STA run throughout the entire simulation. The deviation and RMSE between CPL run and ATM.DYN runs are consistent within 0.5 ∘C. This demonstrates the capability of the coupled model in performing realistic regional ocean–atmosphere simulations.

4.2 Sea Surface Temperature

The simulated SST patterns are compared to the validation data to demonstrate the performance of the coupled model in capturing the ocean surface state. The daily SST fields from CPL run on June 2nd and 24th are shown in Fig. 7(I) and Fig. 7(VI). To validate the CPL run results, the SST fields obtained in OCN.DYN runs are shown in Fig. 7(II) and 7(VII) and the GHRSST fields are shown in Fig. 7(III) and 7(VIII). It can be seen that both OCN.DYN and CPL runs are able to reproduce the SST patterns reasonably well in comparison with the satellite observations. Though the CPL run uses the surface forcing fields with

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

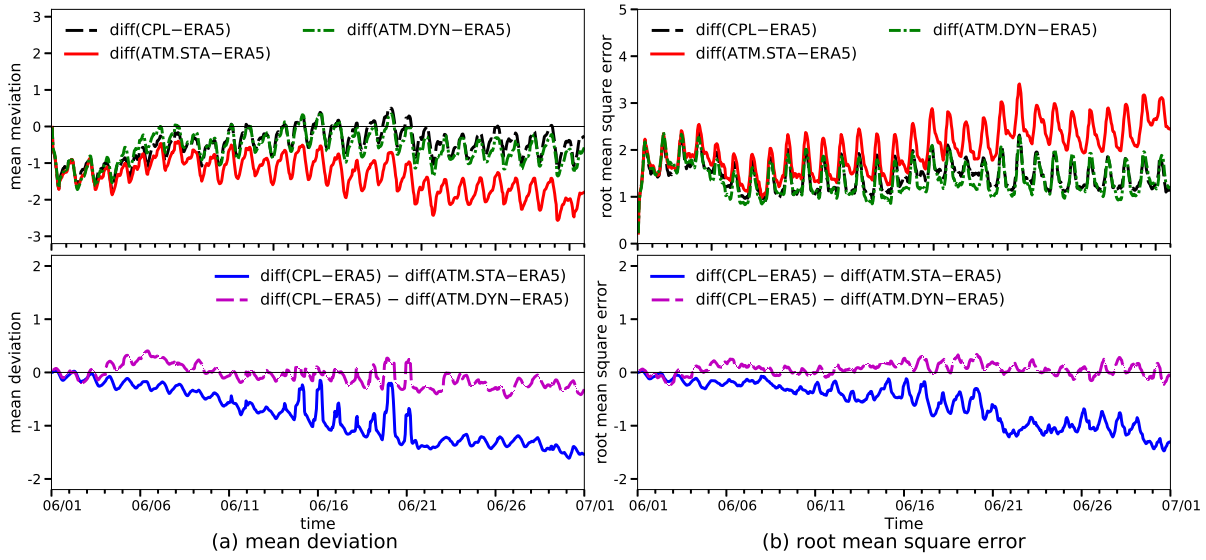


Figure 6. The mean deviation and root mean square error between the surface temperature obtained by the simulations (i.e., ATM.STA, ATM.CPL, and CPL) in comparison with ERA5 data. Only the errors over the Red Sea are considered. The differences between the simulation errors from CPL run and stand-alone WRF simulations are presented below the mean deviation and the root mean square error. The initial time is 0000 UTC Jun 01 2012 for all simulations.

a higher resolution, the SST patterns obtained in both simulations are very similar after two days. On June 24th, the SST patterns in both runs are less similar, but both simulation results are still consistent with GHRSST (RMSE $< 1°$C). Both simulations under-estimate the SST in the northern Red Sea. The CPL run over-estimates the SST in the central and southern Red Sea on June 24th, while the OCN.DYN run under-estimates the SST in the central Red Sea. The difference may be because the CPL run uses the cloud information from the atmosphere component when calculating the surface radiation fluxes, although the uncertainty in the cloud modeling can be significant.

To quantitatively compare the errors in SST results, the time history of the SST in the simulations (i.e., OCN.DYN and CPL) and validation datasets (i.e., GHRSST and HYCOM data) are shown in Fig. 8. The mean deviation and RMSE between simulation results and validation datasets are also plotted. Again, only the errors between daily SST fields are presented because both observational datasets only provide daily data. It can be seen in Fig. 8 that the mean deviation and RMSE of SST in CPL run (deviation: -0.26 °C; RMSE: 0.74 °C) is smaller than that of T2 (deviation: -0.47 °C; RMSE: 1.42 °C) shown in Fig. 6. Generally, the OCN.DYN and CPL runs have a similar range of error compared to both validation datasets, as both simulations are driven by realistic atmospheric forcing. Compared with the HYCOM dataset, the mean deviations of CPL and OCN.DYN runs are small (CPL: -0.12 °C; OCN.DYN: -0.04 °C) before June 10th. After June 11th, the CPL run slightly over-estimated the SST (0.37 °C), but the OCN.DYN run slightly under-estimated it (-0.05 °C). In addition, the RMSEs of both simulations increase in the first 10 days, but the increase is not significant after that. On the other hand, when comparing with the GHRSST, the initial SST patterns in both runs are cooler by 0.8 °C. This is because the HYCOM data is cooler than GHRSST at the

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

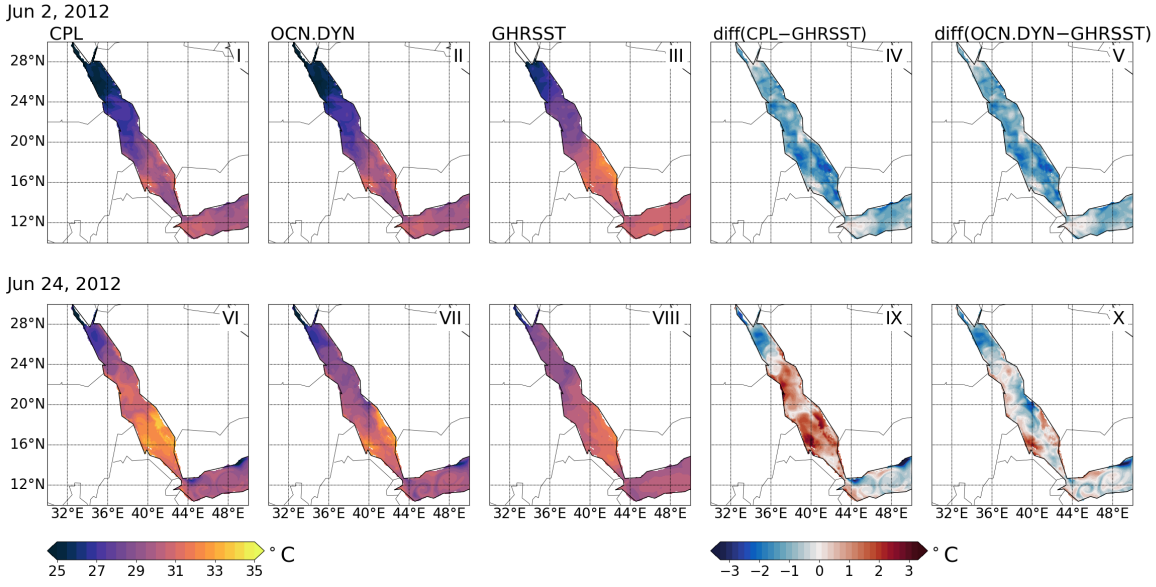


Figure 7. The daily SST patterns obtained by OCN.DYN and CPL runs, and GHRSST dataset. The corresponding differences between the simulations and the GHRSST dataset are also plotted. Two snapshots are selected: (1) Jun 02 2012; (2) Jun 24 2012. The simulation initial time is 0000 UTC Jun 01 2012 for both snapshots.

start of the simulation. After the first 10 days, the difference between GHRSST data and HYCOM decreases, and likewise the difference between the simulation results and GHRSST also decreases. Before June 10th, both CPL and ATM.STA runs under-estimated the SST (CPL: -0.73 °C; OCN.DYN: -0.66 °C). It should be noted that the mean SST in CPL run (-0.01 °C) is closer to GHRSST than OCN.DYN (-0.34 °C) after June 11th.

4.3 Surface Heat Fluxes

The surface heat budget strongly influences the forecast of the surface temperature fields in the simulations. Here we evaluate the performance of the coupled model in capturing the heat fluxes, as compared to the stand-alone simulations. The results are also compared to the MERRA-2 dataset and their differences are plotted.

The turbulent heat fluxes (THF), including the latent heat and sensible heat, and their differences with the validation dataset are shown in Fig. 9. The snapshots of the turbulent fluxes in the heat wave events on June 2nd and 24th are presented. It can be seen that all simulations reproduce the turbulent heat fluxes reasonably well in comparison with the MERRA-2 dataset. On June 2nd, all simulations exhibit similar THF patterns since they have the same initial conditions and air–sea interactions do not significantly impact the THF within two days. On the other hand, for the heat wave event on June 24th, CPL and ATM.DYN runs exhibit more latent heat fluxes coming out of the ocean (157 and 131 W/m^2) than that in ATM.STA run (115 W/m^2). This is because the ATM.STA run is forced by a cooler SST pattern. When forced by cooler SST, the evaporation decreases

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

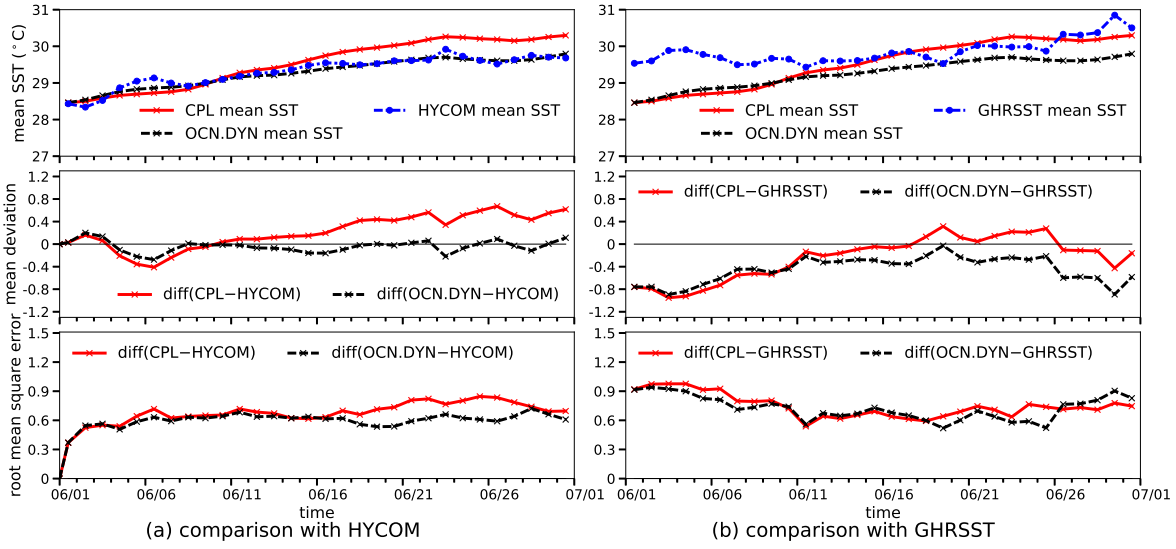


Figure 8. The mean deviation and mean-root-square-error between the daily SST as resulting from the simulations (i.e., OCN.DYN and CPL) in comparison with the observational dataset. Panel (a) shows the comparison with HYCOM dataset and Panel (b) shows the comparison with GHRSST dataset. The initial time is 0000 UTC Jun 01 2012 for all simulations.

and thus the latent heat is smaller. Compared with the latent heat, the sensible heat in the Red Sea region is much smaller in all simulations (10 W/m^2). It should be noted that the MERRA-2 dataset has unrealistically large sensible heat in the coastal regions because its resolution is not adequate to resolve the coastline in the Red Sea region (Gelaro et al., 2017).

The net downward shortwave and longwave heat fluxes are shown in Fig. 10. Again, all simulations reproduce the shortwave and longwave radiation fluxes reasonably well. For the shortwave heat flux, all simulations show similar patterns on both June 2nd and 24th as the air–sea interactions do not significantly impact the solar radiation. However, a small improvement in the CPL (2.19 W/m^2) and ATM.DYN (1.27 W/m^2) runs can be observed in the longwave radiation on June 24th. This is because these two simulations are forced by realistic SST and thus can capture longwave radiation according to the bulk formula. The total downward heat fluxes, which is the sum of the results in Figs. 9 and 10, are shown in Fig. 11. It can be seen that the present simulations over-estimated the total downward heat fluxes (CPL: 646 W/m^2; ATM.STA: 674 W/m^2; ATM.DYN: 663 W/m^2) for both heat wave events compared with MERRA-2 dataset (495 W/m^2), especially in the central Red Sea, the southern Red Sea and the coastal regions. In the central and southern Red Sea, the over-estimation is due to the discrepancies in shortwave solar radiation. To improve the forecast of shortwave radiation, a better understanding of the cloud and aerosol in the Red Sea region is required. In the coastal region, the discrepancy is because MERRA-2 data are only available on a lower resolution grid and do not resolve heat fluxes in the coastal regions. It should be noted that ATM.STA run has the largest discrepancy on June 24th when using a constant SST field. Overall, the present CPL simulations are capable of well capturing all the components of the surface heat fluxes during the heat wave events.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

Geoscientific Model Development Discussions
Open Access
EGU

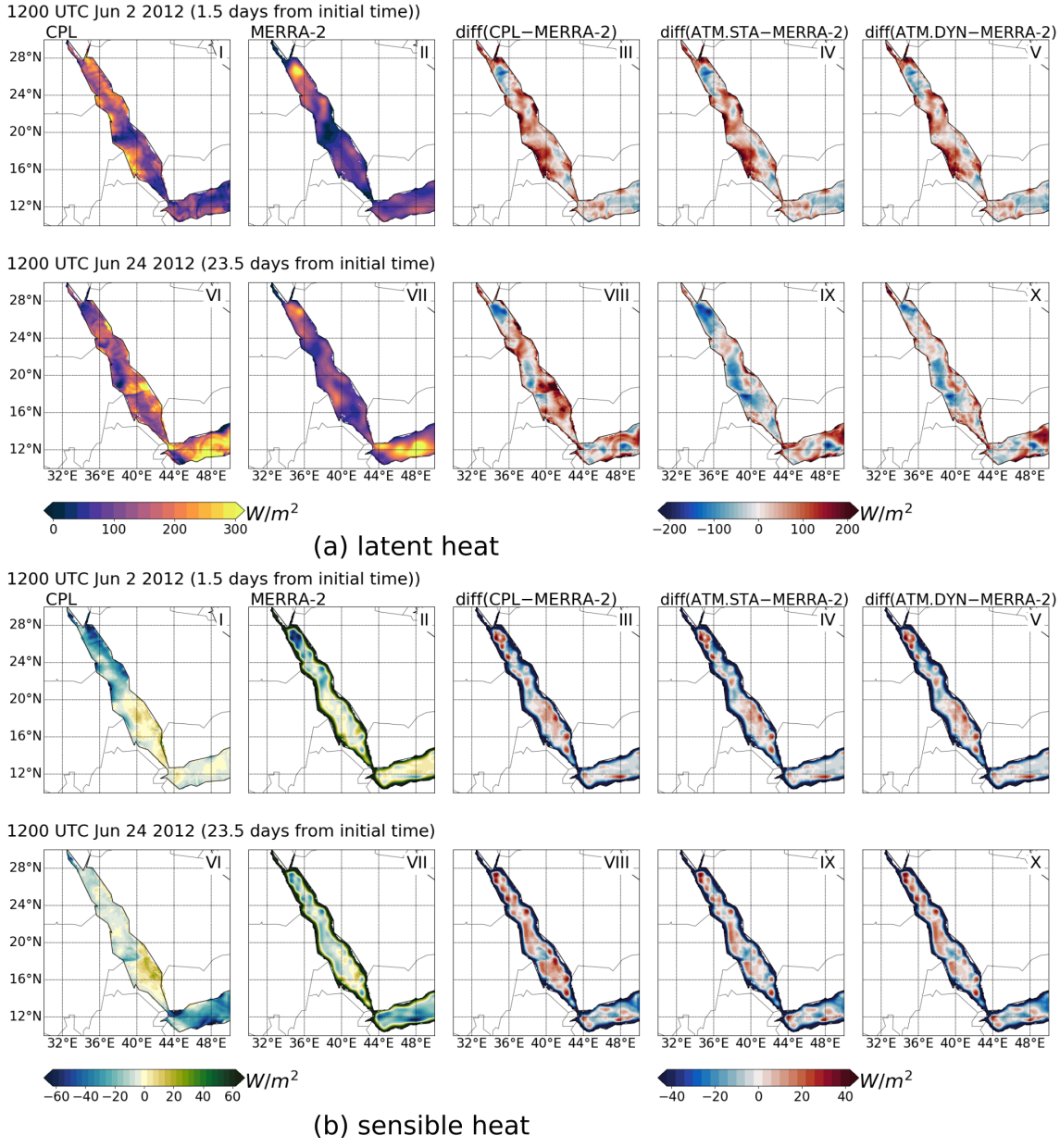
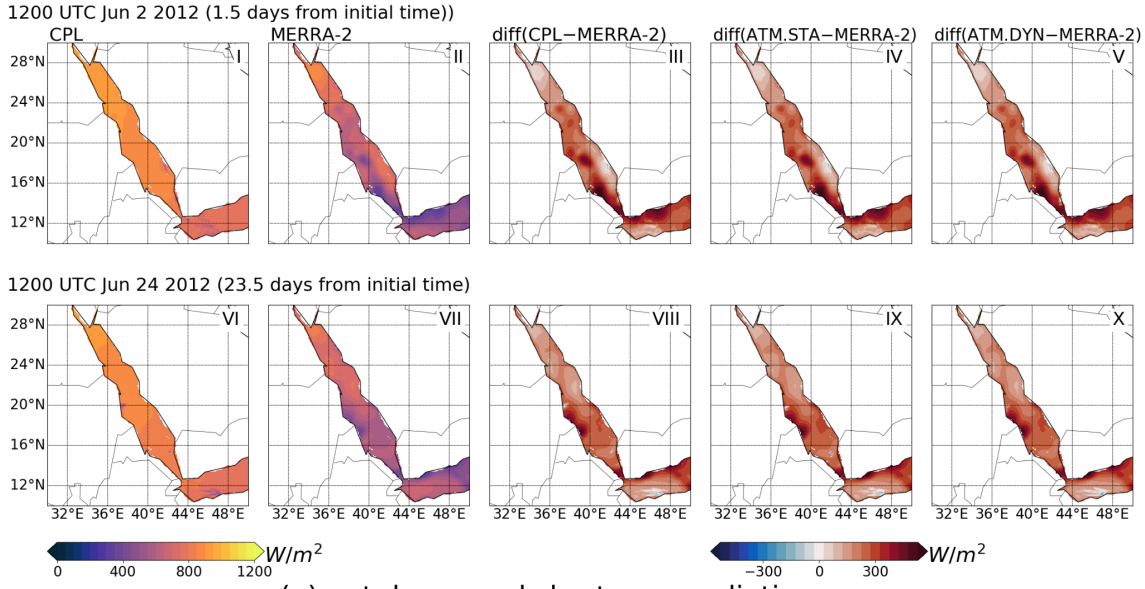


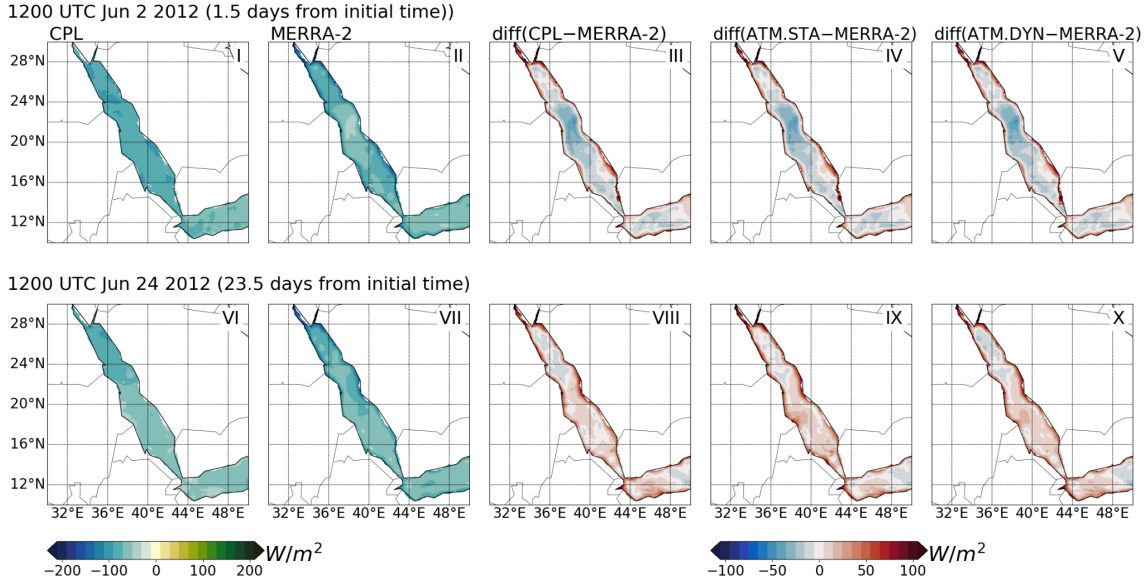
Figure 9. The turbulent heat fluxes out of the sea obtained in CPL run, MERRA-2 data, and their difference (CPL−MERRA-2). The difference between ATM.STA and ATM.DYN with the MERRA-2 data (i.e., ATM.STA−MERRA-2, ATM.DYN−MERRA-2) are also presented. Two snapshots are selected: (1) 1200 UTC Jun 02 2012; (2) 1200 UTC Jun 24 2012. The simulation initial time is 0000 UTC Jun 01 2012 for both snapshots. Only the heat fluxes over the sea is shown to highlight the air–sea interactions.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

Geoscientific Model Development Discussions



(a) net downward shortwave radiation



(b) net downward longwave radiation

Figure 10. The net downward shortwave and longwave heat fluxes obtained in CPL run, MERRA-2 data, and their difference (CPL−MERRA-2). The difference between ATM.STA and ATM.DYN with the MERRA-2 data (i.e., ATM.STA−MERRA-2, ATM.DYN−MERRA-2) are also presented. Two snapshots are selected: (1) 1200 UTC Jun 02 2012; (2) 1200 UTC Jun 24 2012. The simulation initial time is 0000 UTC Jun 01 2012 for both snapshots. Only the heat fluxes over the sea is shown to highlight the air–sea interactions.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

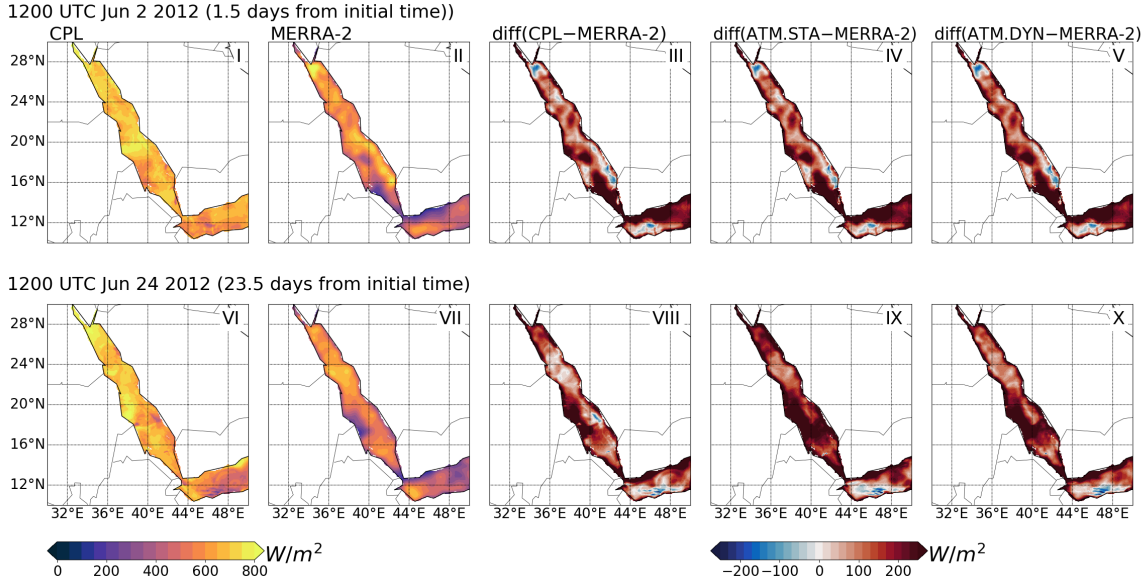


Figure 11. Comparison of the total downward heat fluxes obtained in CPL run, MERRA-2 data, and their difference (CPL−MERRA-2). The difference between ATM.STA and ATM.DYN with the ERA5 data (i.e., ATM.STA−MERRA-2, ATM.DYN−MERRA-2) are also presented. Two snapshots are selected: (1) 1200 UTC Jun 02 2012; (2) 1200 UTC Jun 24 2012. The simulation initial time is 0000 UTC Jun 01 2012 for both snapshots. Only the heat fluxes over the sea is shown to highlight the air–sea interactions.

4.4 Surface Wind and Evaporation

To evaluate the simulation of the surface momentum and freshwater fluxes by the coupled model, the surface wind and evaporation patterns obtained from ATM.STA, ATM.DYN, and CPL runs are presented. The MERRA-2 data is used to validate the simulation results.

The simulated surface wind velocity fields are shown in Fig. 12. They show that the wind velocity magnitude and direction in the CPL run agree well (RMSE: 2.17 m/s) with the MERRA-2 data when using the selected micro-physics schemes. On June 2nd, high-speed wind is observed in the northern and central Red Sea, and the high-resolution CPL run successfully captures the small-scale features of wind speed patterns. On June 24th, the differences between the simulations are larger than those on June 2nd, especially in the central Red Sea and the southern Arabian Peninsula. It should be mentioned that although the SST in the ATM.STA run is lower than the other simulations, the difference in the wind velocity magnitude is small (RMSE, June 2nd: 0.15 m/s; June 24th: 0.74 m/s). This suggests that the ocean–atmosphere coupling does not significantly influence the wind field in the Red Sea region during the heat wave events.

The surface evaporation results are shown in Fig. 13. All simulations reproduce the overall evaporation patterns in the Red Sea. The CPL run is able to capture the relatively high evaporation in the northern Red Sea and the relatively low evaporation in the southern Red Sea in both snapshots, shown in Fig. 13(I) and 13(VI). Again, all simulation results are consistent on

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

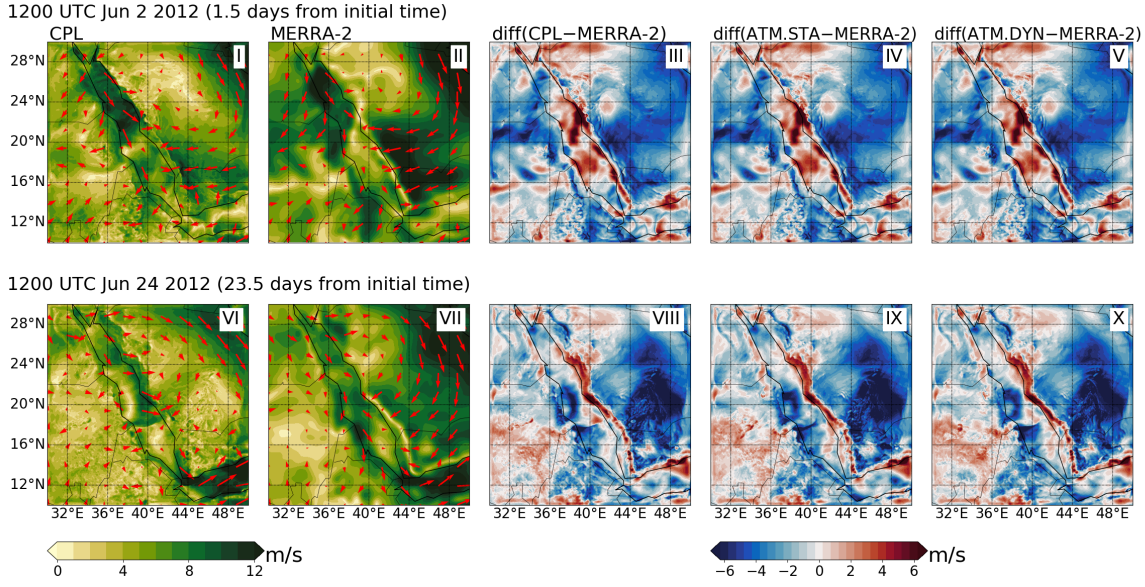


Figure 12. The magnitude and direction of the surface wind obtained in the CPL run, the MERRA-2 data, and their difference (CPL−MERRA-2). The difference between ATM.STA and ATM.DYN with the MERRA-2 data (i.e., ATM.STA−MERRA-2, ATM.DYN−MERRA-2) are also presented. Two snapshots are selected: (1) 1200 UTC Jun 02 2012; (2) 1200 UTC Jun 24 2012. Only the heat fluxes over the sea is shown to highlight the air–sea interactions.

June 2nd because they are driven by the same initial condition and the air–sea interactions do not significantly influence the evaporation fields within two days. However, after 24 days, the CPL run agrees better with MERRA-2 dataset (deviation: 4 cm/year; RMSE: 64 /cm/year) than the ATM.STA run (deviation: -34 cm/year; RMSE: 69 cm/year) by better reproducing the realistic ocean–atmosphere coupling. Although the CPL run results are still consistent with that of the ATM.DYN run, the CPL run over-estimates the evaporation in the southern Red Sea. This is because the CPL run slightly over-estimated the SST, shown in Fig. 7(IX). Since there is no precipitation in three major cities (Mecca, Jeddah, Yanbu) near the eastern shore of the Red Sea during the month according to NCDC climate data, the precipitation results are not shown.

5 Scalability Test

The parallel efficiency is crucial for coupled ocean–atmosphere models for simulating large and complex problems. In this section, the parallel efficiency in the coupled simulations is investigated and presented. We investigate the scalability of the model to evaluate its performance for a constant sized problem simulated using different numbers of processors. Additionally, the CPU time spent on different parts of the coupled model is detailed. The parallel efficiency tests are performed on the COMPAS (Center for Observations, Modeling and Prediction at Scripps) cluster in Scripps Institution of Oceanography (http:

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

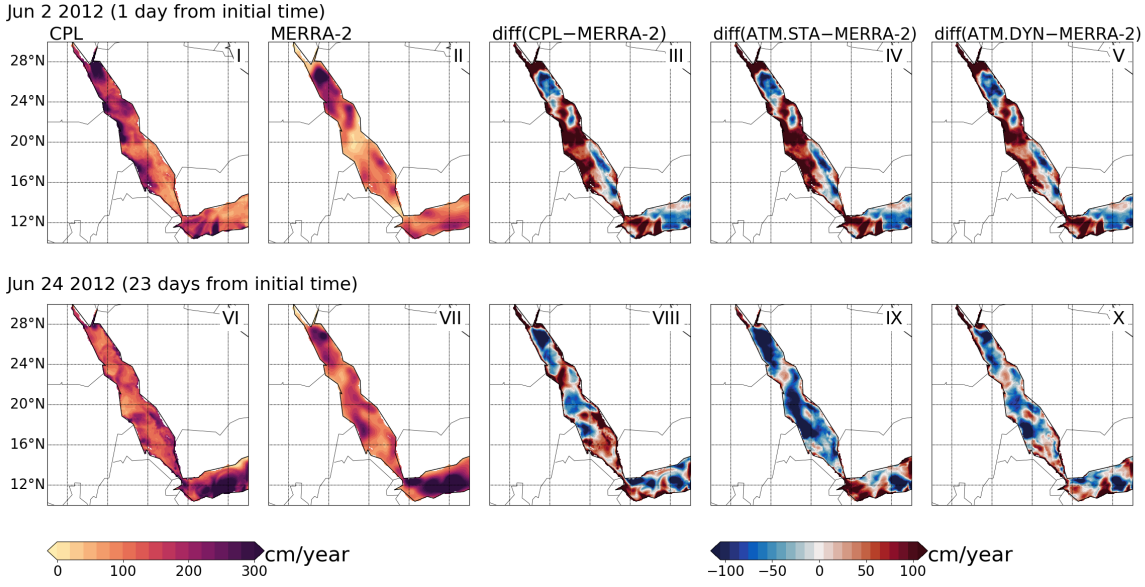


Figure 13. The surface evaporation patterns obtained in the CPL run, the MERRA-2 data, and their difference (CPL−MERRA-2). The difference between ATM.STA and ATM.DYN with the MERRA-2 data (i.e., ATM.STA−MERRA-2, ATM.DYN−MERRA-2) are also presented. Two snapshots are selected: (1) 1200 UTC Jun 02 2012; (2) 1200 UTC Jun 24 2012. Only the evaporations over the sea is shown to highlight the air–sea interactions.

//www.compas.ucsd.edu/). The COMPAS cluster is composed of 1192 Intel 5400 and 5500 series CPUs and has a theoretical peak speed of 12.6 TeraFlops. The cluster uses Myrinet for its high-performance network.

The parallel efficiency of the scalability test is $N_{p0}t_{p0}/N_{pn}t_{pn}$, where N_{p0} and N_{pn} are the number of processors employed in the simulation of the baseline case and the test case, respectively; t_{p0} and t_{pn} are the CPU time. The speed-up is defined as t_{p0}/t_{pn}, which is the relative improvement of the CPU time when solving the problem. The results obtained in the scalability test of the coupled model are shown in Fig. 14. It can be seen that the parallel efficiency is close to 100% when employing less than 64 processors and is still as high as 65% when using 256 processors. The decrease in parallel efficiency is because when using 256 processors, there are only 16×16 grid points in the horizontal plane. The boundary tiles in each processor are 25% of the total, and the parallel communication cost increases significantly. From the results reported in the literature, the parallel efficiency of the coupled model is comparable to other models (Marshall et al., 1997; Zhang et al., 2013b). It is noted in Fig. 14 that the parallel efficiency fluctuates when using 8 to 32 processors. This may be attributed to the fluctuation of the CPU time when solving the systems of linear equations. When using different numbers of processors, the decomposition of the domain leads to different linear equation systems requiring different CPU load and accordingly different convergence time. This fluctuation may also be due to the variation of CPU cache or memory. The fluctuation of the CPU time can also be seen in the speed-up curve, but at smaller magnitude.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

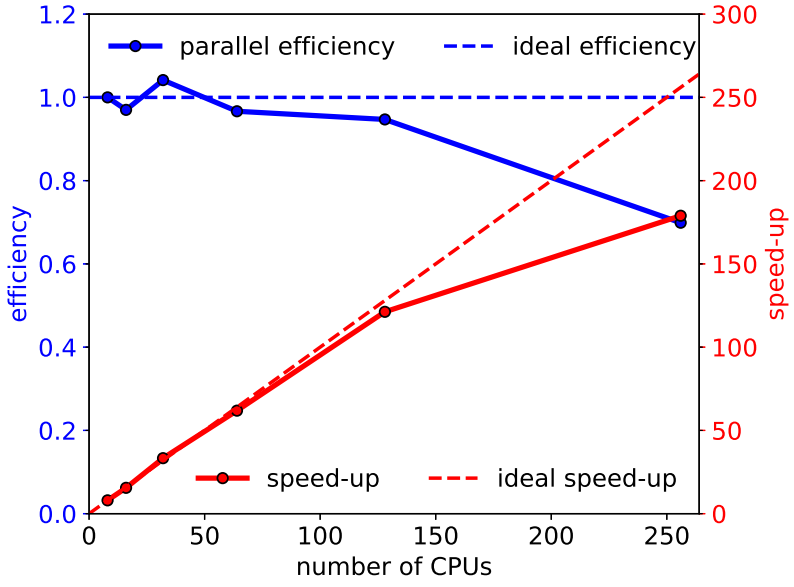


Figure 14. The parallel efficiency test of the coupled model in the Red Sea region. The test cases employ up to 256 CPU cores. The simulation with the smallest case is regarded as base case when computing the speed-up. Tests are performed on the COMPAS cluster in UCSD.

The CPU time spent on coupled run and stand-alone runs is shown in Table. 3. The time spent on the coupler is estimated by subtracting the time spent on stand-alone simulations from the coupled run. The most time-consuming process is the atmospheric model integration, which accounts for 76% to 93% of the total costs. The ocean model integration is the second most time-consuming process, which is 7% to 14% of the total computational costs. The atmospheric model is much more time-consuming than the ocean model because it uses a smaller time step (30 s) than that of the ocean model (120 s) and more complex physics parameterization packages. Moreover, the atmospheric model solves the atmosphere in the entire computational domain, while the ocean model only solves the Red Sea. The coupling process takes less than 5% of the total costs when using fewer than 64 processors. However, when using 256 processors, the proportion of this cost increases to 10% because of the increase of inter-processor communication with more processors. In summary, the scalability test results demonstrate the coupled model can be applied for high-resolution coupled regional modeling studies.

6 Conclusion and Outlook

This study describes the development of a regional coupled ocean–atmosphere numerical framework (MITgcm–WRF) based on the ESMF coupler, with an example of a specific application to simulate the heat wave events in the Red Sea region. Results from the coupled and stand-alone simulations are compared to a wide variety of available observational and reanalysis datasets, aiming to demonstrate the overall performance of the coupled model with respect to stand-alone models. The results obtained from various configurations of coupled and stand-alone model simulations all realistically capture the basic characteristics of

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

Table 3. Comparison of CPU time spent on the coupled run and stand-alone simulations. The CPU times presented here are normalized by the time spent on the coupled run using 256 processors. The CPU time spent on the ESMF/NUOPC coupler is obtained by subtracting two stand-alone simulation time from the CPL run time.

	$N_p = 8$	16	32	64	128	256
CPL run	22.36	11.52	5.37	2.89	1.48	1.00
WRF stand-alone	20.42(91%)	10.41(90%)	4.97(93%)	2.57(89%)	1.27(86%)	0.76(76%)
MITgcm stand-alone	1.76(8%)	0.93(8%)	0.36(7%)	0.20(7%)	0.14(9%)	0.14(14%)
ESMF/NUOPC coupler	0.17(1%)	0.18(2%)	0.03(1%)	0.11(4%)	0.07(5%)	0.10(10%)

the ocean–atmosphere state in the Red Sea region over a 30-day simulation period. The surface air temperature variations in three major cities are consistent with the ground observations and the heat wave events are also well captured in the CPL run. The surface flux fields (e.g., surface air temperature, surface heat fluxes, surface evaporations, surface wind) in the CPL run are consistent with the reanalysis data over the simulation period. The SST fields in CPL run are also consistent with the satellite observation data. Improvements of the coupled model over the stand-alone simulation with static SST forcing are observed in capturing the T2, heat fluxes, evaporation, and wind speed. On the other hand, the difference between coupled simulation and stand-alone simulations with updated forcings is also discussed.

The parallel efficiency of the coupled model is examined by simulating the Red Sea region using increasing number of processors. The coupled model scales linearly for up to 128 CPUs and the parallel efficiency remains about 70% for 256 processors. The CPU time associated with different parts of the coupled simulations is also presented, suggesting good parallel efficiency in both model components and ESMF coupler. Hence the coupled model can be applied for high-resolution coupled regional modeling studies on massively parallel processing supercomputers.

These preliminary results motivate further studies in evaluating and improving this new regional high-resolution coupled ocean–atmosphere model for investigating dynamical processes and forecasting applications in regions around the globe where ocean–atmosphere coupling is important. This regional coupled model can be further improved by developing coupled data assimilation capabilities on initializing coupled forecasts from an assimilated high-resolution analysis state. In addition, the model physics and model uncertainty representation in the coupled system can be enhanced using advanced techniques, such as stochastic physics parameterizations. Future work will involve exploring these and other aspects of developing a high-resolution regional modeling system that is best suited for forecasting and process understanding purposes.

Code and data availability. The coupled code, documentation, and tutorial cases used in this work are maintained using GIT. They are available upon request during the discussion and will be published on GitHub when the manuscript is accepted. ECMWF ERA5 dataset is used as the atmospheric initial and boundary conditions. The ocean model uses the assimilated HYCOM/NCODA $1/12^{\circ}$ global analysis data as initial and boundary conditions. To validate the simulated SST data, we use the OSTIA (Operational Sea Surface Temperature and

Sea Ice Analysis) system in GHRSST (Group for High Resolution Sea Surface Temperature). The simulated 2-meter air temperature (T2) is validated against the ECMWF ERA5 dataset. The observed daily maximum and minimum temperatures from NOAA National Climate Data Center is used to validate the T2 in three major cities. Surface heat fluxes (e.g., latent heat, sensible heat, longwave and shortwave radiations), which are important for ocean–atmosphere interactions, are compared with MERRA-2 (Modern-Era Retrospective analysis for Research and Applications, version 2) datasets.

Author contributions. RS worked on the coding tasks for coupling WRF with MITgcm using ESMF, wrote the code documentation, and performed the simulations for the numerical experiments. RS and ACS worked on the technical details for debugging the model and drafted the initial manuscript. All authors designed the computational framework and the numerical experiments. All authors discussed the results and contributed to the writing of the final manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We appreciate the computational resources provided by COMPAS (Center for Observations, Modeling and Prediction at Scripps) used for this project. We gratefully acknowledge the research funding from KAUST (grant number: OSR-2-16-RPP-3268.02). We are immensely grateful to Caroline Papadopoulos for helping with installing software, testing the coupled code, and using the COMPAS cluster. We appreciate Professor U.U. Turuncoglu sharing part of their ESMF/NUOPC code on GitHub which helped our code development. We wish to thank Dr. Ganesh Gopalakrishnan for setting up the stand-alone MITgcm simulation (OCN.DYN) and providing the external forcings. We also thank Drs. Stephanie Dutkiewicz, Jean-Michel Campin, Chris Hill, Dimitris Menemenlis in ESMF–MITgcm coupling. We also wish to thank Dr. Peng Zhan for discussing the simulations of the Red Sea.

References

Abdou, A. E. A.: Temperature trend on Makkah, Saudi Arabia, Atmospheric and Climate Sciences, 4, 457–481, 2014.

Aldrian, E., Sein, D., Jacob, D., Gates, L. D., and Podzun, R.: Modelling Indonesian rainfall with a coupled regional model, Climate Dynamics, 25, 1–17, 2005.

Barbariol, F., Benetazzo, A., Carniel, S., and Sclavo, M.: Improving the assessment of wave energy resources by means of coupled wave-ocean numerical modeling, Renewable Energy, 60, 462–471, 2013.

Bender, M. A. and Ginis, I.: Real-case simulations of hurricane–ocean interaction using a high-resolution coupled model: effects on hurricane intensity, Monthly Weather Review, 128, 917–946, 2000.

Benjamin, S. G., Grell, G. A., Brown, J. M., Smirnova, T. G., and Bleck, R.: Mesoscale weather prediction with the RUC hybrid isentropic–terrain-following coordinate model, Monthly Weather Review, 132, 473–494, 2004.

Boé, J., Hall, A., Colas, F., McWilliams, J. C., Qu, X., Kurian, J., and Kapnick, S. B.: What shapes mesoscale wind anomalies in coastal upwelling zones?, Climate dynamics, 36, 2037–2049, 2011.

Chen, S. S. and Curcic, M.: Ocean surface waves in Hurricane Ike (2008) and Superstorm Sandy (2012): Coupled model predictions and observations, Ocean Modelling, 103, 161–176, 2016.

Chen, S. S., Price, J. F., Zhao, W., Donelan, M. A., and Walsh, E. J.: The CBLAST-Hurricane program and the next-generation fully coupled atmosphere–wave–ocean models for hurricane research and prediction, Bulletin of the American Meteorological Society, 88, 311–318, 2007.

Collins, N., Theurich, G., Deluca, C., Suarez, M., Trayanov, A., Balaji, V., Li, P., Yang, W., Hill, C., and Da Silva, A.: Design and implementation of components in the Earth System Modeling Framework, The International Journal of High Performance Computing Applications, 19, 341–350, 2005.

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., and Wimmer, W.: The operational sea surface temperature and sea ice analysis (OSTIA) system, Remote Sensing of Environment, 116, 140–158, 2012.

Doscher, R., Willén, U., Jones, C., Rutgersson, A., Meier, H. M., Hansson, U., and Graham, L. P.: The development of the regional coupled ocean-atmosphere model RCAO, Boreal Environment Research, 7, 183–192, 2002.

Evangelinos, C. and Hill, C. N.: A schema based paradigm for facile description and control of a multi-component parallel, coupled atmosphere-ocean model, in: Proceedings of the 2007 symposium on component and framework technology in high-performance and scientific computing, pp. 83–92, ACM, 2007.

Fang, Y., Zhang, Y., Tang, J., and Ren, X.: A regional air-sea coupled model and its application over East Asia in the summer of 2000, Advances in Atmospheric Sciences, 27, 583–593, 2010.

Fowler, H. and Ekström, M.: Multi-model ensemble estimates of climate change impacts on UK seasonal precipitation extremes, International Journal of Climatology, 29, 385–416, 2009.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., et al.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), Journal of Climate, 30, 5419–5454, 2017.

Gualdi, S., Somot, S., Li, L., Artale, V., Adani, M., Bellucci, A., Braun, A., Calmanti, S., Carillo, A., Dell'Aquila, A., et al.: The CIRCE simulations: regional climate change projections with realistic representation of the Mediterranean Sea, Bulletin of the American Meteorological Society, 94, 65–81, 2013.

Gustafsson, N., Nyberg, L., and Omstedt, A.: Coupling of a high-resolution atmospheric model and an ocean model for the Baltic Sea, Monthly Weather Review, 126, 2822–2846, 1998.

Hagedorn, R., Lehmann, A., and Jacob, D.: A coupled high resolution atmosphere-ocean model for the BALTEX region, Meteorologische Zeitschrift, 9, 7–20, 2000.

Harley, C. D., Randall Hughes, A., Hultgren, K. M., Miner, B. G., Sorte, C. J., Thornber, C. S., Rodriguez, L. F., Tomanek, L., and Williams, S. L.: The impacts of climate change in coastal marine systems, Ecology letters, 9, 228–241, 2006.

Henderson, T. and Michalakes, J.: WRF ESMF Development, in: 4th ESMF Community Meeting, Cambridge, USA, Jul 21, 2005.

Hersbach, H.: The ERA5 Atmospheric Reanalysis., in: AGU Fall Meeting Abstracts, San Francisco, USA, Dec 12-16, 2016.

Hill, C., DeLuca, C., Balaji, Suarez, M., and Silva, A.: The architecture of the Earth system modeling framework, Computing in Science & Engineering, 6, 18–28, 2004.

Hill, C. N.: Adoption and field tests of M.I.T General Circulation Model (MITgcm) with ESMF, in: 4th Annual ESMF Community Meeting, Cambridge, USA, Jul 20-21, 2005.

Hodur, R. M.: The Naval Research Laboratory's coupled ocean/atmosphere mesoscale prediction system (COAMPS), Monthly weather review, 125, 1414–1430, 1997.

Hong, S.-Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes, Monthly weather review, 134, 2318–2341, 2006.

Huang, B., Schopf, P. S., and Shukla, J.: Intrinsic ocean–atmosphere variability of the tropical Atlantic Ocean, Journal of Climate, 17, 2058–2077, 2004.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models, Journal of Geophysical Research: Atmospheres, 113, 2008.

Kain, J. S.: The Kain–Fritsch convective parameterization: an update, Journal of applied meteorology, 43, 170–181, 2004.

Kharin, V. V. and Zwiers, F. W.: Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere–ocean GCM, Journal of Climate, 13, 3760–3788, 2000.

Large, W. G., McWilliams, J. C., and Doney, S. C.: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization, Reviews of Geophysics, 32, 363–403, 1994.

Li, D., Zou, L., and Zhou, T.: Regional air–sea coupled model simulation for two types of extreme heat in North China, Climate Dynamics, 50, 2107–2120, 2018.

Loglisci, N., Qian, M., Rachev, N., Cassardo, C., Longhetto, A., Purini, R., Trivero, P., Ferrarese, S., and Giraud, C.: Development of an atmosphere-ocean coupled model and its application over the Adriatic Sea during a severe weather event of Bora wind, Journal of Geophysical Research: Atmospheres, 109, 2004.

Marshall, J., Adcroft, A., Hill, C., Perelman, L., and Heisey, C.: A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers, Journal of Geophysical Research: Oceans, 102, 5753–5766, 1997.

Martin, M., Dash, P., Ignatov, A., Banzon, V., Beggs, H., Brasnett, B., Cayula, J.-F., Cummings, J., Donlon, C., Gentemann, C., et al.: Group for High Resolution Sea Surface Temperature (GHRSST) analysis fields inter-comparisons. Part 1: A GHRSST multi-product ensemble (GMPE), Deep Sea Research Part II: Topical Studies in Oceanography, 77, 21–30, 2012.

National Geophysical Data Center: 2-minute Gridded Global Relief Data (ETOPO2) v2, 2006.

Powers, J. G. and Stoelinga, M. T.: A coupled air–sea mesoscale model: Experiments in atmospheric sensitivity to marine roughness, Monthly weather review, 128, 208–228, 2000.

Pullen, J., Doyle, J. D., and Signell, R. P.: Two-way air–sea coupling: A study of the Adriatic, Monthly weather review, 134, 1465–1483, 2006.

Ricchi, A., Miglietta, M. M., Falco, P. P., Benetazzo, A., Bonaldo, D., Bergamasco, A., Sclavo, M., and Carniel, S.: On the use of a coupled ocean–atmosphere–wave model during an extreme cold air outbreak over the Adriatic Sea, Atmospheric Research, 172, 48–65, 2016.

Roessig, J. M., Woodley, C. M., Cech, J. J., and Hansen, L. J.: Effects of global climate change on marine and estuarine fishes and fisheries, Reviews in fish biology and fisheries, 14, 251–275, 2004.

Seo, H.: Distinct influence of air–sea interactions mediated by mesoscale sea surface temperature and surface current in the Arabian Sea, Journal of Climate, 30, 8061–8080, 2017.

Seo, H., Miller, A. J., and Roads, J. O.: The Scripps Coupled Ocean–Atmosphere Regional (SCOAR) model, with applications in the eastern Pacific sector, Journal of Climate, 20, 381–402, 2007.

Sitz, L., Di Sante, F., Farneti, R., Fuentes-Franco, R., Coppola, E., Mariotti, L., Reale, M., Sannino, G., Barreiro, M., Nogherotto, R., et al.: Description and evaluation of the Earth System Regional Climate Model (RegCM-ES), Journal of Advances in Modeling Earth Systems, 2017.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G.: A description of the advanced research WRF version 2, Tech. rep., National Center For Atmospheric Research Boulder Co Mesoscale and Microscale Meteorology Div, 2005.

Somot, S., Sevault, F., Déqué, M., and Crépon, M.: 21st century climate change scenario for the Mediterranean using a coupled atmosphere–ocean regional climate model, Global and Planetary Change, 63, 112–126, 2008.

Theurich, G., DeLuca, C., Campbell, T., Liu, F., Saint, K., Vertenstein, M., Chen, J., Oehmke, R., Doyle, J., Whitcomb, T., et al.: The earth system prediction suite: toward a coordinated US modeling capability, Bulletin of the American Meteorological Society, 97, 1229–1247, 2016.

Turuncoglu, U., Giuliani, G., Elguindi, N., and Giorgi, F.: Modelling the Caspian Sea and its catchment area using a coupled regional atmosphere-ocean model (RegCM4-ROMS): model design and preliminary results, Geoscientific Model Development, 6, 283, 2013.

Turuncoglu, U. U. and Sannino, G.: Validation of newly designed regional earth system model (RegESM) for Mediterranean Basin, Climate dynamics, 48, 2919–2947, 2017.

Warner, J. C., Armstrong, B., He, R., and Zambon, J. B.: Development of a coupled ocean–atmosphere–wave–sediment transport (COAWST) modeling system, Ocean modelling, 35, 230–244, 2010.

Xie, S.-P., Miyama, T., Wang, Y., Xu, H., De Szoeke, S. P., Small, R. J. O., Richards, K. J., Mochizuki, T., and Awaji, T.: A regional ocean–atmosphere model for eastern Pacific climate: toward reducing tropical biases, Journal of Climate, 20, 1504–1522, 2007.

Xu, J., Rugg, S., Byerle, L., and Liu, Z.: Weather forecasts by the WRF-ARW model with the GSI data assimilation system in the complex terrain areas of southwest Asia, Weather and Forecasting, 24, 987–1008, 2009.

Yao, F., Hoteit, I., Pratt, L. J., Bower, A. S., Köhl, A., Gopalakrishnan, G., and Rivas, D.: Seasonal overturning circulation in the Red Sea: 2. Winter circulation, Journal of Geophysical Research: Oceans, 119, 2263–2289, 2014a.

Yao, F., Hoteit, I., Pratt, L. J., Bower, A. S., Zhai, P., Köhl, A., and Gopalakrishnan, G.: Seasonal overturning circulation in the Red Sea: 1. Model validation and summer circulation, Journal of Geophysical Research: Oceans, 119, 2238–2262, 2014b.

Zhan, P., Subramanian, A. C., Yao, F., and Hoteit, I.: Eddies in the Red Sea: A statistical and dynamical study, Journal of Geophysical Research: Oceans, 119, 3909–3925, 2014.

Zhang, H., Pu, Z., and Zhang, X.: Examination of errors in near-surface temperature and wind from WRF numerical simulations in regions of complex terrain, Weather and Forecasting, 28, 893–914, 2013a.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-252
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 29 November 2018

Zhang, X., Huang, X.-Y., and Pan, N.: Development of the upgraded tangent linear and adjoint of the Weather Research and Forecasting (WRF) Model, Journal of Atmospheric and Oceanic Technology, 30, 1180–1188, 2013b.