Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development Discussions
Open Access

Towards in-situ visualization integrated earth system models: RegESM 1.1 regional modelling system

Ufuk Utku Turuncoglu [1]

[1]Informatics Institute, Istanbul Technical University, 34469, Istanbul, Turkey

Correspondence: Ufuk Utku Turuncoglu (ufuk.turuncoglu@itu.edu.tr)

Abstract. The data volume being produced by regional and global multi-component earth system models are rapidly increasing due to the improved spatial and temporal resolution of the model components, sophistication of the used numerical models in terms of represented physical processes and their non-linear complex interactions. In particular, very short time steps have to be defined in multi-component and multi-scale non-hydrostatic modelling systems to represent the evolution of the fast-moving processes such as turbulence, extra-tropical cyclones, convective lines, jet streams, internal waves, vertical turbulent mixing and surface gravity waves. Consequently, the used small time steps cause extra computation and disk I/O overhead in the used modelling system even if today's most powerful high-performance computing and data storage systems are being considered. Analysis of the high volume of data from multiple earth system model components at different temporal and spatial resolution also poses a challenging problem to efficiently perform integrated data analysis of the massive amounts of data by relying on the conventional post-processing methods available today. This study basically aims to explore the feasibility and added value of integrating existing in-situ visualization and data analysis methods with the model coupling framework (ESMF) to increase interoperability between multi-component simulation code and data processing pipelines by providing easy to use, efficient, generic and standardized modeling environment for earth system science applications. The new data analysis approach enables simultaneous analysis of the vast amount of data produced by multi-component regional earth system models (atmosphere, ocean etc.) during the run process. The methodology aims to create an integrated modeling environment for analyzing fast-moving processes and their evolution in both time and space to support better understanding of the underplaying physical mechanisms. The state-of-art approach can also be used to solve common problems in earth system model development workflow such as designing new sub-grid scale parametrizations (convection, air-sea interaction etc.) that requires inspecting the integrated model behavior in a higher temporal and spatial scale during the run or supporting visual debugging of the multi-component modeling systems, which usually are not facilitated by existing model coupling libraries and modeling systems.

1 Introduction

The multi-scale and inherently coupled (atmosphere, ocean, land etc.) earth system models make them challenging to study and understand. Rapid developments in earth system science, as well as in high performance computing and data storage, have enabled fully coupled regional or global Earth System Models (ESMs) to better represent relevant processes and complex cli-

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

mate feedbacks and interactions among coupled components. Regional ESMs are generally used when the spatial and temporal resolution of the global climate models are not sufficient to resolve local features such as complex topography, land-sea gradients and the influence of human activities in a smaller spatial scale. The complexity of the fully coupled state-of-art regional ESMs have made them computationally more demanding compared to conventional standalone models, and in many cases, to global climate models. To that end, there are obvious needs to develop novel modelling tools, model coupling frameworks (at the required level of sophistication in terms of both physics and computational methods) as well as innovative and efficient modelling systems to enable truthful predictions based on accurate and realistic representation of earth systems.

A modeling framework is defined as an environment for coupling model components through a common calling interface. The modeling framework aims to reduce complexity of regular tasks such as spatial interpolation across different computational grids and to transfer data among model components, in order to increase the efficiency and interoperability of multi-component coupled model systems by providing standardized calling interfaces. Moreover, the synchronization of the execution of individual model components and the exchange of metadata among them are also simplified with model coupling frameworks. The Earth System Modeling Framework (ESMF) is one of the most popular example for this approach (Hill et al., 2004a, b; Collins et al., 2005). The ESMF consists of a standardized superstructure for coupling components of Earth system applications through a robust infrastructure of high-performance utilities and data structures that ensure consistent component behavior (Hill et al., 2004). In addition to the ESMF model coupling framework, the Model Coupling Toolkit (MCT; Jacob et al., 2005; Larson et al., 2005), Model Coupling Environmental Library (MCEL; Bettencourt, 2002), OASIS (Redler et al., 2010; Valcke, 2013) and C-Coupler (Liu et al., 2014) can be given as other examples of this approach.

The ESMs make use of state-of-art modeling frameworks, model components and libraries to have better representation of the earth system processes. The key component of coupled modeling systems is often the driver, which is mainly responsible for data exchange, spatial interpolation and synchronization between model components, subject to their individual computational grids. In this approach, direct interaction do not have to occur between individual model components, since the data transfer is managed by the specialized driver component. Coupling between individual earth system components adds extra overhead and complexity to the overall modeling system in terms of network bandwidth usage, computation, disk I/O and data storage. Due to the high volume of data produced by multiple model components, the results are generally stored in a limited temporal resolution such as six hourly or daily averages, which are processed after numerical simulations are finished. This conventional approach is commonly known as post-processing. A typical configuration, e.g. a thirty years simulation of a fully coupled regional atmosphere-ocean model might produce some Terabyte of output (in four dimensions; time, height or depth, latitude and longitude), which strongly depend on horizontal and vertical resolution of model components, the number of variables and the time interval defined to store the simulation results. In this case, the post-processing step commonly involves creating temporal (i.e. seasonal climatology, daily and/or monthly averages) and spatial averages, time series, cross-sections in a defined axis and/or custom pathways, added value information produced by set of statistical methods and the visualization of the model results. In general, the high volume of data produced by the numerical modeling systems may not enable to store all the critical and valuable information to use later, despite recent advances in disk systems. As a result, the fast-moving processes such

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

as extreme precipitation events, convection, turbulence and non-linear interactions among the model components cannot be analyzed in high temporal and spatial scales with the conventional post-processing approach.

The analysis of leading high-performance computing systems reveals that the rate of disk I/O performance is not growing at the same speed as the peak computational power of the systems (Ahern, 2012; Ahrens, 2015). The recent report of U.S. Department of Energy (DOE) also indicates that the expected rate of increase in I/O bandwidth (100 times) will be slower than the peak system performance (500 times) of the new generations of exascale computers (Ashby et al., 2010). In addition, the movement of large volumes of data across relatively slow network bandwidth servers fail to match the ultimate demands of data processing and archiving tasks of the present high resolution multi-component earth system models. In short, the conventional post-processing approach has become a bottleneck in monitoring and analysis of fast-moving processes that require very high spatial resolution, due to the present technological limitations in high-performance computing and storage (Ahrens et al., 2014). In the upcoming exascale computing era, brilliant new data analysis and visualization methods are needed to evocatively overcome the above limitations.

Besides to the conventional data analysis approach, the so-called in-situ visualization and co-processing approaches allow researchers to simultaneously analyze the output while running the numerical simulations. The coupling of computation and data analysis helps to facilitate efficient and optimized data analysis and visualization pipelines and boosts the data analysis workflow. Recently, a number of in-situ visualization systems for analyzing numerical simulations of earth system processes have been implemented. For instance, the ocean component of Model for Prediction Across Scales (MPAS) has been integrated with an image-based in-situ visualization tool to examine the important elements of the simulations and reduce the data needed to preserve those elements by creating flexible work environment for data analysis and visualization (Ahrens et al., 2014; O'Leary et al., 2016). Additionally, the same modeling system (MPAS-Ocean) has been used to study eddies in large scale, high-resolution simulations. In this case, the in-situ visualization workflow is designed to perform eddy analysis at higher spatial and temporal resolutions than available with conventional post-processing facing storage size and I/O bandwidth constraints (Woodring et al., 2016). Moreover, a regional weather forecast model (Weather Research and Forecasting Model; WRF) has been integrated with in-situ visualization tool to track cyclones based on an adaptive algorithm to perform efficient online visualization (Malakar et al., 2012). In spite of the lack of generic and standardized implementation for integrating model components with in-situ visualization tools, the previous studies have shown that in-situ visualization is able to produce analyses of simulation results, revealing many details in an efficient and optimized way. It is obvious that more generic implementations could facilitate easy integration of the existing standalone and coupled earth system models with available in-situ visualization tools (Ahrens et al., 2005; Ayachit, 2015; Childs et al., 2012) and improve interoperability between such tools and non-standardized numerical simulation codes.

The main aim of this paper is to explore the added value of integrating in-situ analysis and visualization methods with a model coupling framework (ESMF) to provide in-situ visualization for easy to use, generic, standardized and robust scientific applications of earth system modeling. The implementation allows existing earth system models coupled with the ESMF library to take advantage of in-situ visualization capabilities without extensive code restructuring and development. Moreover, the integrated model coupling environment allows sophisticated analysis and visualization pipelines by combining information

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

coming from multiple earth system model components (i.e. atmosphere, ocean, wave, land-surface) in various spatial and temporal resolutions. Detailed studies of key physical processes and interactions among model components are vital to the understanding of complex physical processes and could potentially open up new possibilities for the development of earth system models.

The next section (Sect. 2) describes the design of a fully coupled regional earth system model used in this study. Section 3 provides information about the implementation and design of the in-situ visualization integrated modelling framework employed. Then, the initial results from newly designed modeling framework will be demonstrated in Section 4, along with some benchmark results (Sect. 5). Finally, a summary and discussion are given in the last section (Sect. 6).

2 The Design of the Modeling System

The RegESM (Regional Earth System Model; 1.1) modeling system is able to couple four different model components (atmosphere, ocean, wave and river routing) to support many different modeling applications that might require detailed representation of the interactions among different earth system processes (Fig. 1a-b). The state-of-art driver that is responsible for the orchestration of the overall modeling system resides in the middle and acts as a translator among model components. In the design of the coupled modeling system, the coupling interfaces and driver are mainly developed by Istanbul Technical University (ITU) while MITgcm ocean model component is modified by the Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA) to work as a part of RegESM modeling system (Turuncoglu and Sannino, 2017).

The following sections basically aim to give brief information about the individual model components, which are used to construct the modeling system, as well as the design details of the driver to combine independently developed and maintained model components in an efficient and standardized way.

2.1 Atmosphere Models

The flexible design of RegESM modeling system allows to choose different atmospheric model component (ATM) in the configuration of the coupled model for various type of application. Currently, two different atmospheric model are compatible with RegESM modeling system: **1)** RegCM4 (Giorgi et al., 2012), which is developed by the Abdus Salam International Centre for Theoretical Physics (ICTP) and **2)** the Advanced Research Weather Research and Forecasting (WRF) Model (ARW; Skamarock et al., 2005), which is developed and sourced from National Center for Atmospheric Research (NCAR). The details of the individual atmospheric model can be found in following section but RegCM4 is selected as an atmospheric model component in this study because it is well tested with verified in different model domains such as Caspian Sea (Turuncoglu et al., 2013), Mediterranean Basin (Surenkok and Turuncoglu, 2015; Turuncoglu and Sannino, 2017) and Black Sea. Additionally, the current implementation of RegESM WRF interface is still experimental and does not support online nesting yet but this will be included in the next version of the modeling system (RegESM 2.0).

2.1.1 RegCM

The dynamical core of the RegCM4 is based on the primitive equation, hydrostatic version of the National Centre for Atmospheric Research (NCAR) and Pennsylvania State University mesoscale model MM5 (Grell, 1995). The model includes two different land surface models: **1)** Biosphere-Atmosphere Transfer Scheme (BATS; Dickinson et. al., 1989) and **2)** Community Land Model (CLM), version 4.5 (Tawfik and Steiner, 2011). The model also includes specialized physical parameterizations to define air-sea interaction over the sea and lake (one-dimensional lake model; Hostetler et al., 1993). The Zeng Ocean Air-Sea Parameterization (Zeng et al., 1998) is basically extended to introduce the atmosphere model as a component of the coupled modeling system. In this way, the atmospheric model is able to exchange both two and three-dimensional fields (i.e. sea surface temperature, surface roughness, runoff, surface wind speed, cloud liquid water content) with other model components such as ocean, wave and river routing components that are active in an area inside of the atmospheric model domain as well as in-situ visualization component. In this design, the driver and coupling interface are developed by ITU to adapt easily to the newer version of RegCM model without major code change in the model component and the driver. In this study, specifically RegCM 4.6 is used to perform in-situ visualization integrated simulations.

2.1.2 WRF

The WRF model is suitable for a broad range of applications and has variety of options to choose parameterization schemes for convection, planetary boundary layer (PBL), explicit moisture, radiation and soil processes to make available investigation of different earth system processes. The model consists of fully compressible non-hydrostatic equations and the prognostic variables include the three-dimensional wind, perturbation quantities of pressure, potential temperature, geo-potential, surface pressure, turbulent kinetic energy and scalars (water vapor mixing ratio, cloud water etc). A few modifications are done in WRF (version 3.8.1) model to couple it with RegESM modeling system. These modifications include rearranging of WRF time related subroutines, which are inherited from older version of ESMF Time Manager API that was available in 2009, to compile model with newer version of ESMF library (version 7.1.0) together with older version that requires mapping of time manager data types between old and new versions. In addition, the model is modified to exchange data with driver component to support different coupled model applications such as atmosphere-ocean interaction.

2.2 Ocean Models

The current version of the coupled modeling system supports two different ocean model components (OCN): **1)** Regional Ocean Modeling System (ROMS revision 809; Shchepetkin and McWilliams, 2005; Haidvogel et al., 2008), which is developed and distributed by Rutgers University and **2)** MIT General Circulation Model (MITgcm version c63s; Marshall et al., 1997a, b). In this case, ROMS and MITgcm models are selected due to their large user communities and different vertical grid representations. Although the selection of ocean model components depend on user experience and application, often the choice of vertical grid system has a determining role in some specific applications. For example, the ROMS ocean model uses terrain following (namely s-coordinates) vertical grid system that allows better representation of the coastal processes but

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

MITgcm uses z levels and generally used for the applications that involves open oceans and seas. Similar to the atmospheric model component, both ocean models (ROMS and MITgcm) are slightly modified to allow data exchange with the other model components. In the current version of the coupled modeling system, there is no interaction between wave and ocean model components, which could be crucial for some applications (surface ocean circulation and wave interaction etc.) that need to consider the two-way interaction between waves and ocean currents. Fortunately, there is a plan to implement ocean-wave interface in the future release of the coupled modeling system (RegESM 2.0). In general, the exchange fields between ocean and atmosphere closely depend on the application and the studied problem. In some studies, the ocean model requires heat, freshwater and momentum fluxes to be provided by the atmospheric component, while in others, the ocean component retrieves atmospheric data (i.e. surface temperature, humidity, surface pressure, wind components, precipitation) to calculate fluxes internally, by using bulk formulas. In the current design of the coupled modeling system, the driver allows to select the desired exchange fields from the predefined list of the available fields. In this way, the coupled modeling system can be adapted to different applications without any code customizations in both the driver and individual model components.

2.2.1 ROMS

The ROMS is a three-dimensional, free-surface, terrain-following numerical ocean model that solves the Reynolds-averaged Navier-Stokes equations using the hydrostatic and Boussinesq assumptions. The governing equations are in flux form and the model uses Cartesian horizontal coordinates and sigma vertical coordinates with three different stretching functions. The model also supports second, third and fourth order horizontal and vertical advection schemes for momentum and tracers via its preprocessor flags. To allow coupling with RegESM modeling system, a set of preprocessor flags such as *REGCM_COUPLING* are included to the model.

2.2.2 MITgcm

The MIT general circulation model (MITgcm) is a generic and widely used ocean model that solves Boussinesq form of Navier-Stokes equations for an incompressible fluid. It supports both hydrostatic and non-hydrostatic applications with a spatial finite-volume discretization on a curvilinear computational grid. The model has implicit free surface in the surface and partial step topography formulation to define vertical depth layers. The model used in the coupled modeling system is slightly modified by ENEA to allow data exchange with other model components. The detailed information about the regional applications of the MITgcm ocean model was initially described in the study of Artale et al. (2010) using PROTEUS modeling system, which is also developed by ENEA.

2.3 Wave Model

Surface waves play a crucial role in the dynamics of the planetary boundary layer (PBL) in the atmosphere and the currents in the ocean. To have better representation of atmospheric PBL and the surface conditions (i.e. surface roughness, friction velocity, wind speed), the wave component is included to the coupled modeling system. In this case, the wave component is based on WAM Cycle-4 (4.5.3-MPI). The WAM is a third-generation model, based on pure physics without any assumption on the spectral shape (Monbaliu et al., 2000). It considers all the main processes that control the evolution of a wave field in deep water, namely the generation by wind, the nonlinear wave-wave interactions, and also white-capping. To include WAM model (provided by ENEA) as a component of coupled modeling system, the model again is slightly modified to exchange fields with other model components. In this case, RegCM4 atmospheric model component provides surface atmospheric conditions (i.e. wind speed components or friction velocity and wind direction) to wave model and receives surface roughness to calculate air-sea transfer coefficients and fluxes over sea using Zeng ocean air-sea parameterization (Zeng et al., 1998). In this implementation, it is also possible to define a threshold for maximum roughness length (the default value is 0.02) and friction velocity in the configuration file of atmospheric model component to ensure stability of the overall modeling system. The initial results to investigate the added value of atmosphere-wave coupling in the Mediterranean Sea can be found in Surenkok and Turuncoglu (2015).

2.4 River Routing Model

To simulate the lateral freshwater fluxes (river discharges) at the land surface and to provide river discharge to ocean model component, the modeling system uses Hydrological Discharge (HD, version 1.0.2) model developed by Max Planck Institute (Hagemann and Dumenil, 1998; Hagemann and Lydia, 2001). The model is basically designed to run in a fixed global regular grid with 0.5° horizontal resolution. In this case, model uses pre-computed river channel network to simulate the horizontal transport of the runoff within model watersheds using different flow processes such as overland flow, baseflow and riverflow. In this case, the model integrated with driver uses a daily time step and requires daily time series of surface runoff and drainage as input fields. The river routing model (RTM) plays important role in the freshwater budget of the ocean model by closing the water cycle between atmosphere and ocean model components. The original version of the model is slightly modified to support interaction with the coupled model components. To close water cycle between land and ocean, model retrieves surface and sub-surface runoff from atmospheric component and provides estimated river discharge to the selected ocean model component (ROMS or MITgcm). In the current design of the driver, rivers can be represented in two different ways: **1)** individual point sources that are vertically distributed to model layers, and **2)** imposed as freshwater surface boundary condition like precipitation (P) or evaporation minus precipitation (E-P). The model is also modified to write restart and output files directly in NetCDF (Network Common Data Form) format rather than using Max Planck Institute's SERVICE data type.

2.5 The Driver: RegESM

The RegESM (version 1.1) is completely redesigned and improved version of the previously used and validated coupled atmosphere-ocean model (RegCM-ROMS) to study regional climate of Caspian Sea and its catchment area (Turuncoglu et al., 2013). To simplify the design and to create more generic, extensible and flexible modeling system that aims to support easy integration of multiple model components and applications, the RegESM uses driver based model coupling approach. In this case, all the model components are combined using ESMF (version 7.1.0) framework to structure coupled modelling system. The ESMF framework is selected because of its unique online re-gridding capability, which allows the driver to readily

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

perform different interpolation types (bilinear, conservative etc.) over the exchange fields (i.e. sea surface temperature, heat and momentum fluxes) and the National Unified Operational Prediction Capability (NUOPC) layer. The NUOPC layer basically simplifies component synchronization and run sequence by providing additional programming interface between coupled model and ESMF framework. It also provides the capability of transferring computational grids among the model components, which has critical importance in the integration of the modeling system with co-processing environment (see also Sect. 3). The RegESM modelling system also uses NUOPC layer to support various configuration of component interactions such as defining multiple coupling time steps (fast and slow time steps; Fig. 2) among the model components and explicit and semi-implicit type of model interaction (Fig. 3). An example configuration of the four-component (ATM, OCN, RTM and WAV) coupled modeling system can be seen in Fig. 2. In this case, the RTM component runs in a daily time step (slow) and interacts with ATM and OCN components but ATM and OCN components can interact more frequently (i.e. 3-hours, fast).

The interaction among the model components and driver are facilitated by the connector components provided by NUOPC layer. Connector components are mainly used to create link between individual model components and driver. The NUOPC layer also specifies an interface to define the model and connector components and their interactions (also called as run sequences). To support different applications, The RegESM modeling system can be configured with two different type of time-integration scheme (coupling type) between atmosphere and ocean components: **1)** explicit and **2)** semi-implicit (Fig. 3). In explicit type of coupling, two connector components (ATM-OCN and OCN-ATM direction) are executed at every coupling time step and model components start and stop at the same model time (Fig. 2a). Unlike, explicit coupling, a leap-frog style interaction is supported between atmosphere and ocean components using semi-implicit type of model coupling. In this case, at every coupling time step, the ocean model receives surface boundary conditions from the atmospheric model at one coupling time step ahead of the current ocean model time (Fig. 2b). The main advantage of using of the semi-implicit coupling is that it ensures the stability of the overall modeling system.

As described earlier, the execution of the model components is mainly controlled by the driver. In this case, both sequential and concurrent execution of the model components are allowed in the current version of the modeling system. If the model components and the driver are configured to run in sequence on the same set of PETs (Persistent Execution Threads), then the modelling system executes in a sequential mode. This is much more efficient way to run the modeling system in case of limited computing resources. In concurrent type of execution, the model components run in mutually exclusive sets of PETs but the driver uses union of available computational resources (or PETs) of interacted model components. In this way, the modelling system can support a variety of computing systems ranging from local servers to large computing systems that could include high-speed performance networks, accelerators (i.e. Graphics Processing Unit or GPU) and parallel I/O capabilities. The main drawback of concurrent execution approach is to assign correct amount of computing resource to individual model components, which is not a easy task and might require a extensive performance benchmark of specific configuration of the model components, to achive best available computational performance.

In general, the design and development of the coupled modeling systems involve a set of technical difficulties that arise due to the usage of the different computational grids in the model components. One of the most common examples is the mismatch between the land-sea masks of the model components (i.e. atmosphere and ocean models). In this case, the unaligned land-sea masks might produce artificial and/or unrealistic surface heat and momentum fluxes around the coastlines, narrow bays, straits and seas. The simplest solution of this issue is to modify the land-sea masks of the individual model components manually to align them. However, the main disadvantage of this solution is the required time and the difficulty to fix the land-sea masks of the different model components (especially when the grid resolution is high). In addition, same procedure must be repeated every time when the model domain (i.e. shift or change in model domain) or horizontal grid resolution is changed. As a result, this approach is considered as application specific and very time-consuming. Unlike manual editing of the land-sea masks, customized interpolation techniques that also include extrapolation support might help to create more generic and automatized solutions. The RegESM modeling system uses extrapolation approach to overcome the mismatched land-sea mask problem for the interaction between atmosphere, ocean and wave components. To perform extrapolation, the driver uses a specialized algorithm (Fig. 4) to find the mapped and unmapped ocean grid points in the interpolation stage for every coupling direction (i.e. atmosphere-ocean, ocean-atmosphere, atmosphere-wave). According to the algorithm, the mapped grid points have same land-sea mask type in both model components (i.e. both are sea or land). On the other hand, the land-sea mask type does not match completely in the case of unmapped grid points and they cannot be filled with the standard interpolation method such as bilinear interpolation. In this case, the two-step interpolation is performed to fill unmapped grid points.

In the first step, exchange field is interpolated from source to destination grid (i.e. from atmosphere to ocean) using grid points just over the sea (Fig. 4). Then, result field is used to fill unmapped grid points using nearest-neighbor type interpolation that is performed in the destination grid (from mapped grid points to unmapped grid points). One of the main drawback of this method is that the result field might include unrealistic values and/or sharp gradients in the areas of complex land-sea mask structure (i.e. channels and/or straits) but this might be fixed by applying light smoothing after interpolation or using more sophisticated extrapolation techniques such as the sea-over-land approach (Kara et al., 2007; Dominicis et. al., 2014), which are not included in the current version of the modeling system. In addition, the usage of mosaic grid along with second-order conservative interpolation method, which gives smoother results when the ratio between horizontal grid resolutions of source and destination grids are high, might also help to solve the unaligned land-sea mask problem. Fortunately, the next public release of ESMF framework will supports higher-order conservative interpolation as well as extrapolation to fill unmapped grid points.

3 Integration of RegESM modeling system with co-processing component

The newly designed modeling framework is defined as a combination of the ParaView co-processing plugin - which is called Catalyst (Fabian et. al., 2011) - and ESMF library that is specially designed to couple different earth system models to create more complex regional and global modelling systems.

In conventional co-processing enabled simulation systems, the Catalyst is used to integrate ParaView visualization pipeline (defined as a Python script) with the simulation code to support in-situ visualization through the use of application specific custom adaptor code. The adaptor code basically acts as a wrapper layer and transform information coming from simulation code to the ParaView in a compatible format that is defined using VTK (Visualization Toolkit) API. Moreover, the adaptor

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

code is responsible for defining underlying computational grid and associating them with the multi-dimensional fields. After defining computational grids and fields, the ParaView processes the received data to perform co-processing to create desired products such as rendered visualizations, added value information (i.e. spatial and temporal averages, derived fields) and/or writing raw model data to the storage system (Fig. 5a).

On the other hand, the implemented novel approach aims to create more generic and standardized co-processing environment specifically designed for earth system science (Fig. 5b). By this approach, the existing earth system models, which are coupled with ESMF library using NUOPC interface, might benefit to use integrated modeling framework to analyze the data flowing from multi-component and multi-scale modeling system without extensive code development and restructuring. In this design, the adaptor code interacts directly with the driver that is developed by using ESMF and provides an abstraction layer for the co-processing component. As discussed previously, the ESMF framework basically uses standardized interface (initialization, run and finalize routines) to plug new model components into exiting modeling system such as RegESM in an efficient and optimized way. To that end, the new approach will benefit from the standardization of common tasks in the model components to integrate co-processing component with the existing modeling system. In this case, all the information (grid, field and their associated metadata information) required by ParaView, Catalyst plugin is received from the driver and direct interaction between individual model components and the adaptor code is not allowed (Fig. 5b). The implementation logic of the adaptor code is very similar to the conventional approach (Fig. 5a) but in this case, it uses standardized interface of ESMF framework and NUOPC layer to define computational grid and associated two and/or three-dimensional fields of model components. The adaptor basically maps the ESMF field (i.e. *ESMF_Field*) and grid (i.e. *ESMF_Grid*) objects to their VTK equivalents through the use of VTK API, which is provided by ParaView co-processing plugin. Along with the usage of the new approach, the interoperability between simulation code and in-situ visualization system are enhanced and standardized. The new design also ensures easy to develop, extensible and flexible integrated modeling environment for earth system science.

The development of the adaptor component plays important role in the overall design and performance of the integrated modeling environment. The adaptor code mainly includes set of functions for the initialization (defining computational grids and associated input ports), run and finalize the co-processing environment. Similarly, ESMF framework also uses same approach to plug new model components into the modeling system as ESMF components. In ESMF framework, the simulation code is separated into three basic components (initialization, run and finalize) and calling interfaces are triggered by driver to control the simulation codes (i.e. atmosphere and ocean models). In this case, the initialization phase includes definition and initialization of the exchange variables, reading input (initial and boundary conditions) and configuration files and defining the underlying computational grid (step 1 in Fig. 6). The run phase includes a time stepping loop to run the model component in a defined period and continues until simulation ends (step 4 in Fig. 6). The time interval to exchange data between model and co-processing component can be defined using coupling time step just like the interaction among other model components. According to the ESMF convention, the model and co-processing components are defined as a gridded component while the driver is a coupler component. In each coupling loop, the coupler component prepares exchange fields according to the interaction among components by applying re-gridding, performing unit conversion and common operations over the fields (i.e. rotation of wind field).

To allow interaction with the co-processing component, the driver is extended to redistribute two and three-dimensional fields from physical model components. In the initialization phase, the numerical grid of ESMF components are transformed into their VTK equivalents using adaptor code (step 3 in Fig. 6). In this case, *ESMF_Grid* object can be used to create *vtkStructuredGrid* and/or *vtkUnstructuredGrid* along with their modified parallel two-dimensional decomposition configuration, which is supported by ESMF/NUOPC grid transfer capability. The simplified diagram of the ESMF/NUOPC grid transfer feature is also shown in Fig. 7. In this case, each model component transfers their numerical grid representation to co-processing component at the beginning of the simulation (step 1 in Fig. 6) while assigning independent two-dimensional decomposition ratio to the retrieved grid definitions. In the example configuration, the atmosphere model with 2x3 decomposition ratio (in x and y direction) is mapped to 2x2 in co-processing component (Fig. 7). Similarly, the ocean model transfers its numerical grid with 4x4 decomposition ratio to co-processing component with 2x2 (Fig. 7). The main advantage of the generic implementation of the driver component is to assign different computational resources to both model and co-processing components. By this design, the computational resource with accelerator support (GPU etc.) can be independently used by co-processing component to process (iso-surface extraction, volume rendering, texture mapping etc.) the high volume of data in an efficient and optimized way. The initialization phase is also responsible to define exchange fields that will be exchanged among the model components and maps *ESMF_Field* representations defined in physical model components to *vtkMultiPieceDataSet* objects in co-processing component. Due to the modified two-dimensional domain decomposition structure of the numerical grids of the simulation codes, the adaptor code also modifies the definition of ghost regions that are defined as a small subset of global domain that are used to perform numerical operations around edges of the decomposition elements. In this case, the ghost regions (or halo regions in ESMF convention) are updated by using specialized calls and after that the simulation data are passed (as *vtkMultiPieceDataSet*) to the co-processing component. During the simulation, the co-processing component of the modeling system also synchronizes with the simulation code and retrieves updated data (step 5 in Fig. 6) to process and analyze the results (step 6 in Fig. 6). The interaction between driver and the adaptor continues until the simulation ends (step 4, 5 and 6 in Fig. 6) and the driver continues to redistribute the exchange fields using *ESMF_FieldRedist* calls. The driver also supports vertical interpolation of the three-dimensional exchange fields to height (from s-coordinates of ROMS ocean model) or depth coordinate (from sigma coordinates of RegCM atmosphere model) before passing information to the co-processing component. Then, finalizing routines of model and co-processing components are called to stop the model simulations and the data analysis pipeline that destroy the defined data structure/s and free the memory(step 7-8 in Fig. 6).

4 Use Case and Performance Benchmark

To test the capability of newly designed integrated modeling system that is described briefly in the previous section, the three components (atmosphere, ocean and co-processing) configuration of RegESM 1.1 modeling system is implemented to analyze category 5 Hurricane Katrina. Hurricane Katrina was the costliest natural disaster and has been named one of the five deadliest hurricanes in the history of the United States and the storm is currently ranked as the third most intense United States landfalling tropical cyclone. After established in the southern Florida coast as a weak category 1 storm near 22:30 UTC 25 August

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

2005, it strengthened to a category 5 storm by 12:00 UTC 28 August as the storm entered the central Gulf of Mexico (GoM). To observe the evolution of the Hurricane Katrina and understand importance of air-sea interaction in terms of its development and predictability, the model simulations are performed between 27-30 Aug. 2005, which is the most intense period of the cyclone, for three days. The next section mainly includes detailed information of three components configuration of the modeling system as well as used computing environment, the preliminary benchmark results that are done in limited computing resource and analysis of the evolution of Hurricane Katrina.

4.1 Working Environment

The model simulations and performance benchmarks are done on a cluster (SARIYER) provided by National Center for High Performance Computing (UHeM) in Istanbul, Turkey. The CentOS 7.2 operating system installed in compute nodes are configured with a two Intel Xeon CPU E5-2680 v4 (2.40GHz) processor (total 28 cores) and 128 GB RAM. In addition to the compute nodes, the cluster is connected to a high-performance parallel disk system (Lustre) with 349 TB storage capacity. The performance network, which is based on Infiniband FDR (56 Gbps) is designed to give highest performance for the communication among the compute servers and the disk system. Due to the lack of GPU accelerators in the entire system, the in-situ visualization integrated performance benchmarks are done with support of software rendering provided by Mesa library. Mesa is an open source OpenGL implementation that supports a wide range of graphics hardwares each with its own back-end called a renderer. Mesa also provides several software-based renderers for use on systems without graphics hardware. In this case, ParaView is installed with Mesa support to render information without using hardware-based accelerators.

4.2 Domain and Model Configurations

The regional earth system model (RegESM 1.1) is configured to couple atmosphere (ATM; RegCM) and ocean (OCN; ROMS) models with newly introduced novel in-situ visualization component (COP; ParaView Catalyst version 5.4.1) to analyze evolution of Hurricane Katrina and to assess the overall performance of the modeling system. In this case, two atmospheric model domains were designed for RegCM simulations using one-way nesting approach, as shown in Fig. 8. The outer atmospheric model domain (low-resolution; LR) with a resolution of 27-km is centered at 77.5°W, 25.0°N and covers almost entire United States, western part of Atlantic Ocean and north-eastern part of Pacific Ocean for better representation of the large scale atmospheric circulation systems and minimize the effect of lateral boundaries of the atmospheric model in the simulation results of inner model domain (high-resolution; HR). The horizontal grid spacing of second domain is 3-km and covers the entire GoM and western Atlantic Ocean to provide high resolution atmospheric forcing for coupled atmosphere-ocean model simulations and perform cloud resolving simulations. Unlike the outer domain, the model for inner domain is configured to use non-hydrostatic dynamical core to allow better representation local scale vertical acceleration and important pressure features.

The lateral boundary condition for outer domain is obtained from European Centre for Medium-Range Weather Forecasts (ECMWF) latest global atmospheric reanalysis (ERA-Interim project; Dee et. al., 2011), which is available at 6-h intervals at a resolution of 0.75°x0.75° in the horizontal and 37 pressure levels in the vertical. On the other hand, the lateral boundary condition of the HR domain, which is used in the fully coupled model simulations, is specified by the results of the LR

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development Discussions
Open Access

domain. Concerning cumulus convection, Massachusetts Institute of Technology-Emanuel convective parameterization scheme (MIT-EMAN; Emanuel, 1991; Emanuel and Zivkovic-Rothman, 1999) is used in outer model simulations. Along with selected cumulus convection parameterization, sub-grid explicit moisture (SUBEX; Pal et al., 2000) scheme is used to represent large-scale precipitation for LR domain.

As it also shown in Fig. 8, the ROMS ocean model is configured to cover entire the GoM to allow better tracking of the Hurricane Katrina. In this case, the used ocean model configuration is very similar to the configuration used by Physical Oceanography Numerical Group (PONG), Texas A&M University (TAMU), in which the original model configuration can be accessible from their THREDDS server. The ocean model has a spatial resolution of 1/36°, which corresponds to a non-uniform resolution of around 3 km (655 x 489 grid points) with highest grid resolution in the northern part of the domain. The model has 60 vertical sigma layer ($\theta_s = 10.0$, $\theta_b = 2.0$) to provide detailed representation of the main circulation pattern of the region and vertical tracer gradients. The bottom topography data of the GoM is constructed using the ETOPO1 dataset (Amante and Eakins, 2009) and minimum depth (h_c) is set to 400 m. The bathymetry data are also modified so that the ratio of depths of any two adjacent grids does not exceed 0.25 to enhance stability of the model and ensure hydrostatic consistency creation that prevents pressure gradient error. The Mellor-Yamada level 2.5 turbulent closure (MY; Mellor and Yamada, 1982) is used for vertical mixing, while rotated tensors of harmonic formulation are used for horizontal mixing. The lateral boundary conditions for ROMS ocean model are provided by Naval Oceanographic Office Global Navy Coastal Ocean Model (NCOM) during 27-30 August 2005.

The model coupling time step between atmosphere and ocean model component is set to 1 hour but 6 minutes coupling time step is used to provide one-way interaction with co-processing component to study Hurricane Katrina in a very high temporal resolution. In the coupled model simulations that use same configuration of the ocean model but different resolution of the atmospheric model component, the SST data provided to the atmospheric model in the region where their numerical grids overlap. In the rest of the domain, atmospheric model uses SST data provided by ERA-Interim dataset. The results of performance benchmark also include additional tests with smaller coupling time step such as 3 minutes for the interaction with the co-processing component. In this case, the model simulations for the analysis of the Hurricane Katrina runs over three days, but only one day of simulation length is chosen in the performance benchmarks to reduce amount of used computing resources.

4.3 Performance Benchmark

To assess the overall performance of the coupled modeling system by focusing overhead of the newly introduced co-processing component, a set of simulations are performed with different model configurations (Table 1). The performance benchmarks include analysis of the extra overhead provided by the co-processing component, coupling interval between physical models and co-processing component under different rendering load such as various visualization pipelines (Table 1). To scale up to large number of processors, two different atmospheric model configurations are defined (a low-resolution, LR and high-resolution HR). The LR model domain includes around 900.000 grid points in atmospheric model while HR domain contains 25 million grid points. In both case, the ocean model configuration is same and it has around 19 million grid points. Beside the change of the dynamical core of atmospheric model component in HR case (non-hydrostatic), the rest of the model configurations are preserved. To isolate the overhead of the driver from the overhead of the co-processing component, first individual model components (ATM and OCN) are run in standalone mode and then, the best scaled model configurations in terms of two-dimensional decomposition configuration are used in the coupled model simulations (CPL and COP). Due to the current limitation in the integration of the co-processing component, the coupled model only supports sequential type execution (see section 2.5 for more information) when co-processing component is activated but this limitation will be removed in the future version of the modeling system (RegESM 2.0). As mentioned in the previous section, the length of the simulations are kept relatively short (1 day) in the benchmark analysis to perform many simulations with different model configurations (coupling interval etc.), visualization pipelines and domain decomposition parameters.

The benchmark results of standalone model components (ATM and OCN) can be seen in Fig. 9. In this case, two different atmospheric model configurations are considered to see the effect of the domain size and non-hydrostatic dynamical core in the benchmark results (LR and HR; Fig. 8). The results show that the model scales pretty well and it is clear that HR case shows better scaling results than LR configuration of the atmospheric component (ATM) as expected. It is also shown that around 588 processors, which is the highest available compute resource, the communication among the processors dominate the benchmark results and even HR case does not gain further performance (Fig. 9a). Similar to atmospheric model component, the ocean model (OCN) is also tested to find the best two-dimensional domain decomposition configuration (tiles in x and y direction). As it can be seen from the Fig. 9b, the selection of the tile configuration affects the overall performance of the ocean model. In general, model scales better if tile in x direction is bigger than tile in y direction but this is more evident in the small number of processors. This is mainly due to the memory management of Fortran programming language (column-major order) as well as the total number of active grid points (not masked as land) placed in each tile. On the other hand, the tile options must be selected carefully while considering the dimension of the model domain in each direction. In some tile configuration, it is not possible to run the model due to the used underlying numerical solver and the required minimum ghost points. To summarize, the ocean model scales well until 588 cores with the best tile configurations shown in Fig. 9b.

Using the benchmark results of standalone atmosphere and ocean models, the performance of the two-component modeling system (CPL) can be investigated. In this case, the best two-dimensional decomposition parameters of the standalone ocean model simulations are used in the coupled model simulations (Fig. 9b). The comparison of the standalone and coupled model simulations show that the driver component introduces additional 5-10% (average is 5% for LR and 6% for HR cases) overhead in the total execution time, which slight increases along with the used total number of processors, which is acceptable when increased number of MPI communication between the components are considered (Fig. 9 and 10a-b). The extra overhead is mainly due to the interpolation (sparse matrix multiply performed by ESMF) and extrapolation along the coastlines to match land-sea masks of the atmosphere and ocean models and fill the unmapped grid points to exchange data (Fig. 4).

To investigate the overhead introduced by the newly designed co-processing component, the three-component modeling system (COP) is tested with three different visualization pipelines (P1, P2 and P3; Table 1) using two different atmospheric model configurations (LR and HR) and coupling interval (3 and 6 minutes with co-processing). In this case, the measured total execution time during the COP benchmark results also includes vertical interpolation to map data from sigma coordinates to height coordinates for both physical model components (ATM and OCN). As shown in Fig. 10a-b, the co-processing components require 10-40% extra execution time for both LR and HR cases depending on used visualization pipeline when it is compared with CPL simulations. The results also reveal that the fastest visualization pipeline is P3 and the slowest one is P1 for the HR case (Fig. 10b). Table 1 also includes the execution time of the single visualization pipeline (measured by using *MPI_Wtime* call) isolated from the rest of the tasks. In this case, each rendering task gets 2-4 seconds for P1 and P2 cases and 7-15 seconds for P3 case in LR atmospheric model configuration. For HR case, P1 and P2 takes around 17-80 seconds and P3 case is rendered in around 8-10 seconds. These results show that the time spent in co-processing component (basically sending data to ParaView, Catalyst and rendering to create output) fluctuates too much and do not show predictable stable behavior. This might be due to the special configuration of the ParaView, which is configured to use software-based rendering to process data in CPUs and load in the used high-performance computing system (UHeM) even if the benchmark tests are repeated multiple times.

In addition to the testing modelling system with various data processing load, a benchmark with increased coupling time step is also performed (see P23M in Fig. 10b). In this case, the coupling time step between physical model components and co-processing component is increased (from 6 minutes to 3 minutes) to produce output in doubled frame rate but coupling interval between physical model components (ATM and OCN) are kept same (1 hours). The benchmark results show that increased coupling time step also rises overhead due to the co-processing from 45% to 60% for HR case and pipeline P2 when it is compared with the results of two-component simulations (CPL; Fig. 10b). It is also shown that the execution time of co-processing enabled coupled simulations increase but the difference between P2 and P23M cases are reduced from 66% to 37% when number of processor increased from 140 to 588.

Besides the minor fluctuations in the benchmark results, the modelling system with co-processing component scales pretty well to higher number of processors (or cores) without any major performance pitfalls in the current configuration. On the other hand, the usage of accelerator enabled ParaView configuration (i.e. using NVIDIA EGL library) and ParaView plugins with accelerator support such as NVIDIA IndeX volume rendering plugin to process data on GPU will definitely improve the benchmark result. The NVIDIA IndeX for ParaView Plugin basically enables large-scale and high-quality volume data visualization capabilities of the NVIDIA IndeX library inside the ParaView and might help to reduce time to process high-resolution spatial data (HR case). This will be investigated future when NVIDIA IndeX plugin supports in-situ visualization under ParaView. In addition, the model configurations used in the benchmark simulations also write simulation results to the disk in netCDF format. In case of disabling of writing data to disk or configure the models to write data with large time intervals (i.e. monthly), the simulations with active co-processing component will run much faster and make analysis of the model results in real time efficiently especially in live mode (see Section 5.1).

5 Results

As it indicated in the previous sections, the newly designed modelling system is able to analyze numerical simulation results in both in-situ (or live) and co-processing (or post-processing) mode and this section aims to give more detailed information about two different approaches by evaluating numerical simulation of Hurricane Katrina in both mode to reveal the designed modelling system capability and its limitations.

5.1 Live Visualization Mode

While the live visualization designed to examine the simulation state at a specific point in time, the temporal filters such as ParticlePath, ParticleTracer, TemporalStatistics that are designed to process data using multiple time steps cannot be used in this mode. However, live visualization mode allows to connect to the running simulation anytime through the ParaView GUI in order to make detailed analysis by modifying existing visualization pipelines defined by Python script. In this case, numerical simulation can be paused while visualization pipeline is modified and continue to run with the revised one. It is oblivious that the live visualization capability gives a full control to the user to make further investigation about the simulation results and facilitate better insight into underlying physical process and its evolution in time.

The current version of the co-processing enabled modeling system is able to process data of multiple model components by using multi-channel input port feature of ParaView Catalyst. In this case, each model has two input channels based on the rank of exchange fields. For example, atmospheric model component has *atm_input2d* and *atm_input3d* input channels to make available processing both two and three-dimensional exchange fields. The underlaying ESMF adaptor resides in the driver side and provides two grid definitions (2d and 3d) for each model components for further analysis. In this design, ParaView Co-processing Plugin is used to generate Python co-processing scripts and user need to map data sources to input channels by using predefined names such as *atm_input2d* and *ocn_input3d*. Then, adaptor provides required data to co-processing component through each channel to perform rendering and data analysis in real time. The fields that are used in the co-processing component are defined by generic ASCII formatted driver configuration file (*exfield.tbl*), which is also used to exchange data among physical model components such as atmosphere and ocean models. Fig. 11 shows a screenshot of live visualization of three-dimensional relative humidity field provided by the low-resolution atmospheric model component, underlying topography information and vorticity of ocean surface that is provided by ocean model component.

5.2 Co-processing Mode

In addition to live visualization model that is described briefly in the previous section, ParaView Catalyst also allows to process and store data using predefined co-processing pipeline (in Python) for further analysis. Co-processing mode can be used for two purposes: **1)** the simulation output can be directed to the co-processing component to calculate added value information such as vorticity from wind components or eddy kinetic energy from ocean current and stored in a disk for further analysis and **2)** storing simulation output in a higher temporal resolution to process it later (post-processing) or create a representative dataset that can be used to create visualization pipeline for co-processing or live visualization modes. In this case, the newly designed modelling system is able to apply multiple visualization and data processing pipeline to the simulation results at each coupling time step to make different set of analysis in the same numerical simulation for more efficient data analysis. The modelling system also facilitates multiple input ports to process data flowing from multiple earth system model components. In this design, input ports are defined automatically by co-processing component based on activated physical model components

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

(ATM, OCN, etc.) and each model components have two ports to handle two and three-dimensional grids (and fields) separately such as *atm_input2d*, *atm_input3d*, *ocn_input2d* and *ocn_input3d*.

To test the capability of the co-processing component, the evolution of Hurricane Katrina is investigated by using two different configuration of coupled model (COP_LR and COP_HR) that are also used to analyze overall computational performance of the modelling system (see Section 4.3). In this case, both model configuration uses same configuration of OCN model component but different horizontal resolution of ATM model is considered (27 km for LR and 3 km for HR cases).

Figure 12 shows 3-hourly snapshots of the model simulated clouds that are generated by processing three-dimensional relative humidity field calculated by low-resolution version of coupled model (COP_LR) using NVIDIA IndeX volume rendering plugin as well as stream lines of Hurricane Katrina, which is calculated using three-dimensional wind field. The visualization pipeline also includes sea surface height and surface current from ocean model component to make integrated analysis of the model results. Figure 12a-b shows the streamlines that are produced by extracting the hurricane using ParaView *Threshold* filter. In this case, extracted region is used as a seed to calculate backward and forward streamlines. In Figure 12c-e, sea surface height, sea surface current and surface wind vectors (10-meters) are shown together to give insight about interaction of ocean related variables with atmospheric wind. Lastly, the hurricane reaches to the land and start to disappear due to increased surface roughness and lack of energy source (Fig. 12f). While, low-resolution of atmosphere model configuration is used, the information produced by the new modeling system enabled to investigate the evolution of the hurricane in a very high temporal resolution, which was impossible before. A day-long animation that is also used to create Figure 12 can be found as a supplemental video.

In addition to the analysis of low resolution model results to reveal evolution of the hurricane in a very high temporal resolution, low and high-resolution model results are also compared to see the added value of the increased horizontal resolution of the atmospheric model component in terms of representation of the hurricane and its structure. To that end, a set of visualization pipelines are designed to investigate the vertical updraft in the hurricane, simulated track, precipitation pattern and ocean state. In this case, two time snapshots are considered: **1)** 28 August 2005 0000 UTC, when it is the early stage of the hurricane in Category 5 and **2)** 29 August 2005 0000 UTC when it is just before Katrina makes its third and final landfall near the Louisiana–Mississippi border, where the surface wind is very strong and surface currents had a strong onshore component (McTaggart-Cowan et al., 2007a, b). In the analysis of vertical structure, the hurricane is isolated based on the criteria of surface wind speed that exceeds 20 m/s and the seed (basically set of points in *vtkPoints* data type) input for ParaView *StreamTracerWithCustomSource* filter are defined dynamically using *ProgrammableFilter* as a circular plane with a radius of 1.2° and points distributed with 0.2° interval in both direction (x and y) around the center of mass of the isolated region. Then, forward and backward streamlines of vorticity are computed separately to see inflow at low and mid levels and outflow at upper levels for both low (COP_LR; Fig. 13a, b, d and e) and high-resolution (COP_HR; Fig. 14a, b, d and e) cases. The analysis of simulations reveal that the vertical air movement shows higher spatial variability in high-resolution simulation (COP_HR) case even if the overall structure of the hurricane is similar in both cases. As it is expected, the strongest winds occur in a region formed as a ring around the eyewall of the hurricane, which is the lowest surface pressure occurs. In addition, the analysis of cloud liquid water content also shows that low and mid-levels of the hurricane have higher water content in a decreasing trend

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development Discussions Open Access EGU

with height and spatial distribution of precipitation is better represented in high resolution case (Fig. 14a-b and d-e), which is consistent with the previous modelling study of Trenberth et al. (2007). It is also seen that the realistic principal and secondary precipitation bands around eye of the hurricane are more apparent and well structured in high-resolution simulation while low-resolution case does not show those small scale features (Fig. 13a-b and d-e). In addition to the analysis of the atmospheric model component, the loop current, which is a warm ocean current that flows northward between Cuba and the Yucatan Peninsula and moves north into the Gulf of Mexico, loops east and south before exiting to the east through the Florida Straits and joining the Gulf Stream, is well defined by the ocean model component in both cases (Fig. 13c and f; Fig. 14c and f). The track of the hurricane is also compared with the HURDAT2 second-generation North Atlantic (NATL) hurricane database, which is the longest and most complete record of tropical cyclone (TC) activity in any of the world's oceans (Landsea and Franklin, 2013). In this case, the eye of the hurricane is extracted as a region that surface pressure anomaly is greater than 15 mb (shown as a circular region near the best track). As it can be seen from figures, Katrina move over in the central Gulf, which is mainly associated with the loop current and persistent warm and cold eddies, and intensified as it passed over the region due to the high ocean heat content in both simulation (Fig. 13c and f and Fig. 14c and f). The comparison of the low and high-resolution simulations also indicate that the diameter of hurricane-force winds at peak intensity is bigger in high-resolution simulation case at 29 August 2005 0000 UTC (Fig. 13f and Fig. 14f). An animation that shows the comparison of low and high-resolution model results can be found as a supplemental video.

While the main aim of this paper is to give design details of the new in-situ visualization integrated modeling system and show its capability, the performance of the coupled modeling system to represent one of the most destructive hurricane is very satisfactory especially for high-resolution case (COP_HR). Nonetheless, the individual components (atmosphere and ocean) of the modeling system can be tuned to have better agreement with the available observations and previous studies. Specifically for the analysis of the hurricane, a better storm tracking algorithm need to be implemented using ParaView *Programmable Filter* by porting existing legacy Fortran codes for more accurate storm tracking in both live and co-processing mode.

6 Summary and Conclusions

In this study, the newly developed state-of-art in-situ visualization integrated modeling system (RegESM) is used to demonstrate feasibility and added value of the integrated modeling environment to analyze high volume of data coming from multi-component earth system model in an integrated way, which was not possible before. In this case, ParaView, Catalyst plugin is used as a co-processing component to process and render data. The results of the selected use case (Hurricane Katrina) show that the co-processing component provides easy to use and generic modeling and data analysis environment, which is independent from the underlying physical model components used. Moreover, it promotes the usage of co-processing capability with the existing earth system models, which is coupled using ESMF framework and NUOPC layer, without major code restructuring and development and help to increase the interoperability between earth system models and ParaView, Catalyst in-situ visualization plugin. In the current implementation, the prototype version of adaptor code basically acts as a wrapper or abstraction layer to simplify and standardize the regular tasks to integrate the simulation code with in-situ visualization and analysis environment. The driver is also responsible to redistribute the data to co-processing component while preserving its numerical grid along with the support of vertical interpolation. The coupling of co-processing component with the generic driver facilitate to define custom data processing pipelines (defined by special Python scripts) easily and allows integrated analysis of data originated from different components (i.e. atmosphere and ocean models) of the RegESM modeling system in a very high temporal resolution. In this way, RegESM modeling system can be used to study various physical processes (i.e. extreme precipitation events, air-sea interaction, convection and turbulence) that could not be analyzed with the conventional post-processing approaches. While the results of the in-situ visualization integrated modeling system are encouraging, the co-processing component will be extended to support different regional and global computational grid representations supported by ESMF library such as unstructured meshes for having a generic adaptor for various model applications. Additionally, we are currently exploring: **1)** the way to optimize the grid transfer feature and mapping exchange fields to enhance the overall performance of the modelling environment in terms of memory usage and computational efficiency especially for very high resolution applications (< 3 km), **2)** possibility of automatic detection of accelerators (GPUs) through the use of driver component and assigning available GPU resources automatically to the co-processing component for rendering, **3)** improving modelling system and co-processing component to allow nested applications (both atmosphere and ocean), **4)** developing more application of the integrated modeling environment to analyze different physical processes such as air-sea interactions in upwelling regions under extreme atmospheric forcing conditions.

Code availability. The RegESM modeling system is open source and available under the MIT License, making it suitable for the community usage. The license allows to modification, distribution, private and commercial uses. The source code for all versions of RegESM driver including 1.1 is distributed through the public code repository hosted by GitHub (https://github.com/uturuncoglu/RegESM). The user guide and detailed information about the modeling system is also distributed along with the source code using same code repository. The RegESM source code includes the required code patches for the individual model components to use them as a component in the modelling system. On the other hand, the source code of individual model components such as ocean, wave and river routing components and co-processing tool (ParaView, Catalyst) used in the modelling system are distributed mainly by their home institutes and might apply different licensing types. The reader that might want to get more information about the individual model components and their license type could refer to the web sites of them. The release version 1.1 is permanently archived on Zenodo and accessible under the digital object identifier doi:10.5281/zenodo.1307212.

Competing interests. The author declare that he has no conflict of interest.

Acknowledgements. This study has been supported by a research grant (116Y136) provided by The Scientific and Technological Research Council of Turkey (TUBITAK). The computing resources used in this work were provided by the National Center for High Performance Computing of Turkey (UHeM) under grant number 5003082013 and 5004782017. The Quadro K5200 used for the development of the prototype version of the modeling system was donated by the NVIDIA Corporation as part of the Hardware Donation Program. The author extends his grateful thank to Rocky Dunlap and Robert Oehmke from NOAA/ESRL, and CIRES, Boulder, Colorado, Gerhard Theurich from Science Applications International Corporation, McLean, Virginia and Andy C. Bauer from Kitware Inc., USA and Mahendra Roopa from NVIDIA for their very useful suggestions and comments.

References

Ahrens, J.: Increasing scientific data insights about exascale class simulations under power and storage constraints. IEEE Computer Graphics and Applications, 35, 2, pp. 8–11, 2015.

Ahrens, J., Jourdain, S., O'Leary, P., Patchett, J., Rogers, D.H., Fasel, P., Bauer, A., Petersen, M., Samsel, F.: In Situ MPAS-Ocean Image-Based Visualization. Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis SC14, 16–21 Nov. 2014, New Orleans, LA, USA, 2014.

Ahrens, J., Geveci, B., Law, C.: ParaView: An End-User Tool for Large Data Visualization, Visualization Handbook, Elsevier, ISBN-13: 978-0123875822, 2005.

Ahern, S.: High Performance Visualization: Enabling Extreme- Scale Scientific Insight. CRC Press/Francis–Taylor Group, 2012, ch. The Path to Exascale, pp. 331–353. ISBN 978-1439875728

Amante, C. and Eakins, B. W.: ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis, NOAA Technical Memorandum NESDIS NGDC-24, 19 pp, 2009.

Artale, V., Calmanti, S., Carillo, A., Dell'Aquila, A., Hermann, M., Pisacane, G., Ruti, P.M., Sannino, G., Struglia, M.V., Giorgi, F., Bi, X., Pal, J.S., Rauscher, S.: An atmosphere ocean regional climate model for the mediterranean area: assessment of a present climate simulation. Climate Dynamics, 35, 721–740, 2010.

Ashby, S., Beckman, P., Chen, J., Colella, P., Collins, B., Crawford, D., Dongarra, J., Kothe, D., Lusk, R., Messina, P., Mezzacappa, T., Moin, P., Norman, M., Rosner, R., Sarkar, V., Siegel, A., Streitz, F., White, A., Wright, M.: The opportunities and challenges of exascale computing: summary report of the advanced scientific computing advisory committee (ASCAC) subcommittee at the US Department of Energy Office of Science. pp. 71, 2010.

Ayachit, U.: The ParaView Guide: A Parallel Visualization Application, Kitware, ISBN 978-1930934306, 2010.

Bettencourt, M. T.: Distributed model coupling framework, in: Proceedings of the 11th IEEE Symposium on High Performance Distributed Computing, Edinburgh, UK, 23–26 July 2002, 284–290, 2002.

Childs, H., Brugger, E., Whitlock, B., Meredith, J., Ahern, S., Pugmire, D., Biagas, K., Miller, M., Harrison, C., Weber, G.H., Krishnan, H., Fogal, T., Sanderson, A., Garth, C., Bethel, E.W., -Camp, D., Rubel, O., Durant, M., Favre, J.M., Navr'atil, P.: VisIt: An End-User Tool For Visualizing and Analyzing Very Large Data. High Performance Visualization-Enabling Extreme-Scale Scientific Insight, Chapman and Hall/CRC, pp. 520, 2012.

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, International Journal of High Performance Computing Applications, 19, 341–350, 2005.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bid- lot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Holm, E. V., Isaksen, L., Kallberg, P., Kohler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, 2011.

Dickinson, R. E., Errico, R. M., Giorgi, F., and Bates, G. T.: A regional climate model for the western United States, Climatic Change, 15, 383–422, 1989.

Dominicis, M.D., Falchetti, S., Trotta, F., Pinardi, N., Giacomelli, L., Napolitano, E., Fazioli, L., Sorgente, R., Haley, P., Lermusiaux, P., Martins, F., Cocco, M.: A relocatable ocean model in support of environmental emergencies. The Costa Concordia emergency case. Ocean Dyn., 64, pp. 667–688, 2014.

Emanuel, K. A.: A scheme for representing cumulus convection in large-scale models, J. Atmos. Sci., 48(21), 2313–2335, 2011.

Emanuel, K. A., and Zivkovic-Rothman, M.: Development and evaluation of a convection scheme for use in climate models, J. Atmos. Sci., 56, 1766–1782, 1999.

Fabian, N., Moreland, K., Thompson, D., Bauer, A., Marion, P., Geveci, B., Rasquin, M., Jansen, K.: The paraview coprocessing library: A scalable, general purpose in situ visualization library. In Large Data Analysis and Visualization (LDAV), 2011 IEEE Symposium on (Oct 2011), pp. 89–96, 2011.

Giorgi, F., Coppola, E., Solmon, F., Mariotti, L., Sylla, M. B., Bi, X., Elguindi, N., Diro, G. T., Nair, V., Giuliani, G., Turuncoglu, U. U., Cozzini, S., Guttler, I., O'Brien, T. A., Tawfik, A. B., Shalaby, A., Zakey, A. S., Steiner, A. L., Stordal, F., Sloan, L. C., and Brankovic, C.: RegCM4: model description and preliminary tests over multiple CORDEX domains, Climate Research, 52, 7–29, 2012.

Grell, G., Dudhia, J., and Stauffer, D. R.: A description of the fifth-generation Penn State/NCAR mesoscale model (MM5), Technical Note NCAR/TN-398+STR, NCAR, 117 pp.,1995.

Haidvogel, D. B., Arango, H. G., Budgell, W. P., Cornuelle, B. D., Curchitser, E., DiLorenzo, E., Fennel, K., Geyer, W. R., Hermann, A. J., Lanerolle, L., Levin, J., McWilliams, J. C., Miller, A. J., Moore, A. M., Powell, T. M., Shchepetkin, A. F., Sherwood, C. R., Signell, R. P., Warner, J. C., and Wilkin, J.: Ocean forecasting in terrain-following coordinates: formulation and skill assessment of the Regional Ocean Modeling System, J. Comput. Phys., 227, 3595–3624, 2008.

Hagemann, S., Dumenil, L.: A parameterization of the lateral waterflow for the global scale. Clim Dyn, 14, pp. 17–31, 1998.

Hagemann, S., Lydia, D.G.: Validation of the hydrological cycle of ECMWF and NCEP reanalyses using the MPI hydrological discharge model. J Geophys Res, 106, pp. 1503–1510, 2001.

Hill, C., DeLuca, C., Balaji, V., Suarez, M., and Da Silva, A.: The architecture of the Earth System Modeling Framework, Comput. Sci. Eng., 6, 18–28, 2004.

Hill, C., DeLuca, C., Balaji, V., Suarez, M., Da Silva, A., Sawyer, W., Cruz, C., Trayanov, A., Zaslavsky, L., Hallberg, R., Boville, B. A., Craig, A., Collins, N., Kluzek, E., Michalakes, J., Neckels, D., Schwab, E., Smithline, S., Wolfe, J., Iredell, M., Yang, W., Jacob, L. R., and Larson, J. W.: Implementing Applications with the Earth System Modeling Framework, Lecture Notes in Computer Science, 3732, 563–572, 2004.

Hostetler, S. W., Bates, G. T., and Giorgi, F.: Interactive Coupling of a Lake Thermal Model with a Regional Climate Model, J. Geophys. Res, 98, 5045–5057, 1993.

Jacob, R., Larson, J., and Ong, E.: M X N Communication and Parallel Interpolation in Community Climate System Model Version, International Journal of High Performance Computing Applications, 19, 293–307, 2005.

Kara, B.A., Wallcraft, A.J., and Hurlburt, H.E.: A Correction for Land Contamination of Atmospheric Variables near Land?Sea Boundaries. J. Phys. Oceanogr., 37, pp. 803–818, 2007.

Landsea, C.W. and Franklin, J.L.: Atlantic Hurricane Database Uncertainty and Presentation of a New Database Format. Mon. Wea. Rev., 141, 3576–3592, 2013.

Larson, J., Jacob, R., and Ong, E.: The Model Coupling Toolkit: a new Fortran90 toolkit for building multiphysics parallel coupled models, International Journal of High Performance Computing Applications, 19, 277–292, 2005.

Liu, L., Yang, G., Wang, B., Zhang, C., Li, R., Zhang, Z., Ji, Y., and Wang, L.: C-Coupler1: a Chinese community coupler for Earth system modeling, Geosci. Model Dev., 7, 2281–2302, 2014.

McTaggart-Cowan, R., Bosart, L.F., Gyakum, J.R., and Atallah, E.H.: Hurricane Katrina (2005). Part I: Complex Life Cycle of an Intense Tropical Cyclone. Mon. Wea. Rev., 135, 3905–3926, 2007.

McTaggart-Cowan, R., Bosart, L.F., Gyakum, J.R., and Atallah, E.H.: Hurricane Katrina (2005). Part II: Evolution and Hemispheric Impacts of a Diabatically Generated Warm Pool. Mon. Wea. Rev., 135, 3927–3949, 2007.

Malakar, P., Natarajan, V., Vadhiyar, S.: Integrated parallelization of computations and visualization for large-scale applications. IEEE 26th International Parallel and Distributed Processing Symposium Workshops PhD Forum (IPDPSW), 21-25 May 2012, Shanghai, China, 2012.

Marshall, J., Adcroft, A., Hill, C., Perelman, L., Heisey, C.: A finite-volume, incompressible Navier Stokes model for, studies of the ocean on parallel computers. J Geophys Res C Oceans, 102, C3, 5753–5766, 1997.

Marshall, J., Hill, C., Perelman, L., Adcroft, A.: Hydrostatic, quasi-hydrostatic, and non-hydrostatic ocean modeling. J Geophys Res C Oceans,102, C3, 5733–5752, 1997.

Mellor, G. L., and Yamada, T.: Development of a turbulence closure model for geophysical fluid problems, Rev. Geophys., 20, 851–875, 1982.

Monbaliu, J., Hargreaves, R., Albiach, J., Luo, W., Sclavo, M., and Gunther, H.: The spectral 5 wave model, WAM, adapted for applications with high spatial resolution, Coast. Eng., 41, 41–62, 2000.

O'Leary, P., Ahrens, J., Jourdain, S., Wittenburg, S., Rogers, D.H., Petersen, M.: Cinema image-based in situ analysis and visualization of MPAS-ocean simulations. Parallel Computing, 55, pp. 43–48, 2016.

Pal, J. S., Small, E.E, and Eltahir, E.A.B.: Simulation of regional-scale water and energy budgets: Representation of subgrid cloud and precipitation processes within RegCM, J. Geophys. Res.-Atmospheres, 105(D24), 29579–29594, 2000.

Redler, R., Valcke, S., and Ritzdorf, H.: OASIS4 ? a coupling software for next generation earth system modelling, Geosci. Model Dev., 3, 87–104, 2010.

Shchepetkin, A. F. and McWilliams, J. C.: The Regional Ocean Modeling following coordinates ocean model, Ocean Modell., 9, 347–404, 2005.

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. NCAR Tech. Note TN-468+STR, 88 pp., 2005.

Surenkok, G. and Turuncoglu, U.U.: Investigating the Role of Atmosphere–Wave Interaction in the Mediterranean Sea using coupled climate model (RegESM). EGU General Assembly 2015, 12–17 April 2015, 17, EGU2015-3644, Vienna, Austria, 2015.

Tawfik, A.B. and Steiner, A.L.: The role of soil ice in land–atmosphere coupling over the United States: a soil moisture precipitation winter feedback mechanism. J Geophys Res. 116, D02113, 2011.

Trenberth, K.E., Davis, C.A., and Fasullo, J.: Water and energy budgets of hurricanes: Case studies of Ivan and Katrina, J. Geophys. Res., 112, D23106, 2007.

Turuncoglu, U.U., Giuliani, G., Elguindi, N., Giorgi, F.: Modelling the Caspian Sea and its catchment area using a coupled regional atmosphere-ocean model (RegCM4-ROMS): model design and preliminary results, Geosci. Model Dev., 6, pp. 283–299, 2013.

Turuncoglu, U.U., and Sannino, G.: Validation of newly designed regional earth system model (RegESM) for Mediterranean Basin. Climate Dynamics, 48, 9–10, pp. 2919–2947, 2017.

Valcke, S.: The OASIS3 coupler: a European climate modelling community software, Geosci. Model Dev., 6, 373–388, 2013.

Woodring, J., Petersen, M., Schmeisser, A., Patch-Ett, J., Ahrens, J., Hagen H.: In situ eddy analysis in a high-resolution ocean climate model. IEEE Transactions on Visualization and Computer Graphics, 22, 1, pp. 857–866, 2016.

Zeng, X., Zhao, M., and Dickinson, R. E.: Intercomparison of bulk aerodynamic algorithms for computation of sea surface fluxes using TOGA COARE and TAO data, J. Climate, 11, 2628–2644, 1998.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

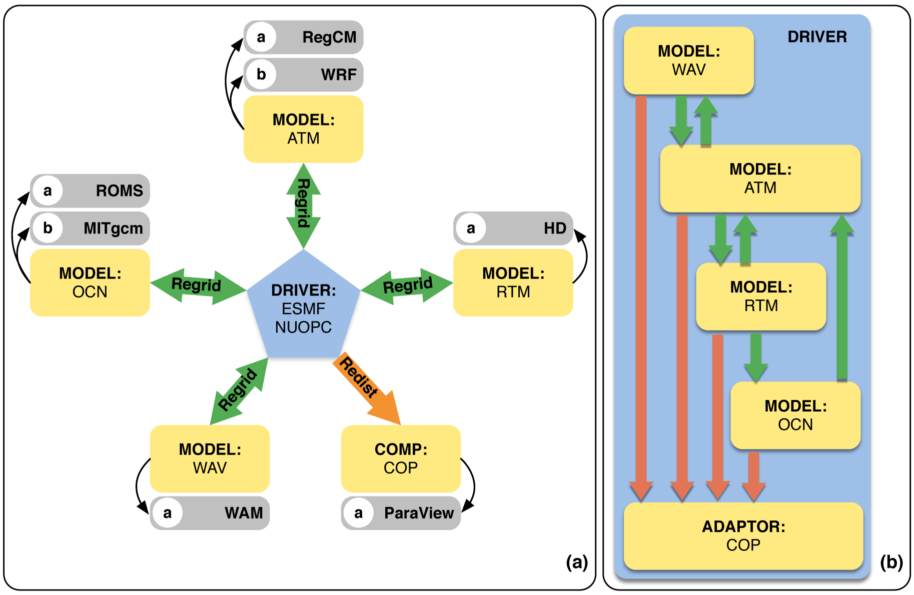


Figure 1. Design of the RegESM coupled modeling system: **(a)** model components including co-processing component, **(b)** their interactions (orange arrows represent the redistribution and green arrows shows regridding).

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development Discussions
Open Access

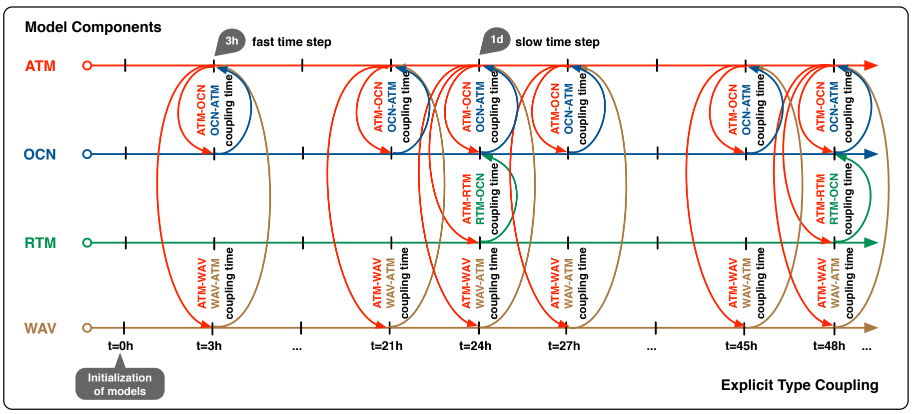


Figure 2. The run sequence of model components in case of explicit type coupling. In this case, the fast coupling time step is used for the interaction between atmosphere, ocean and wave components. The slow coupling time step is only used to interact with river routing component.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development Discussions
Open Access

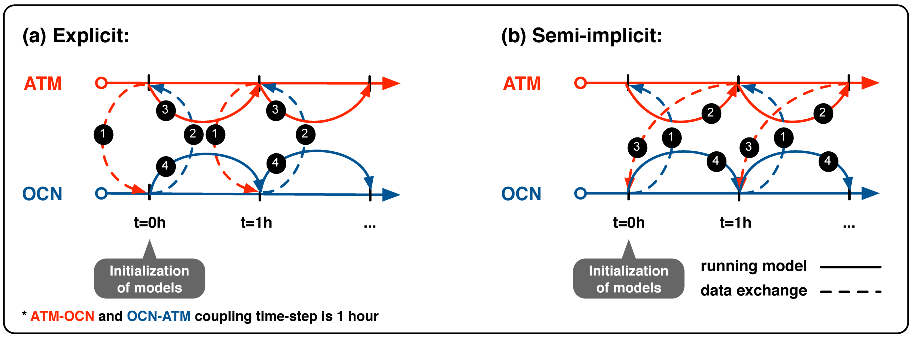


Figure 3. Schematic representation of **(a)** explicit and **(b)** semi-implicit model coupling between two model components (atmosphere and ocean). The numbers indicate the execution orders, which is initialized in each coupling interval.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development
Discussions
Open Access

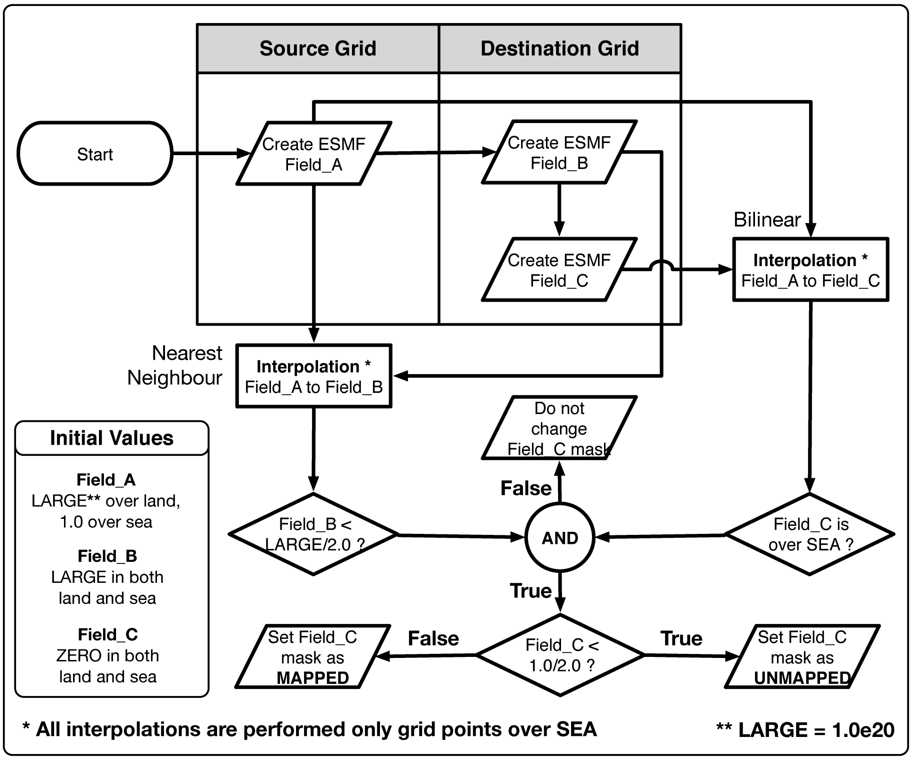


Figure 4. Processing flow chart of algorithm to find mapped and unmapped grid points for two-step interpolation.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development Discussions
Open Access

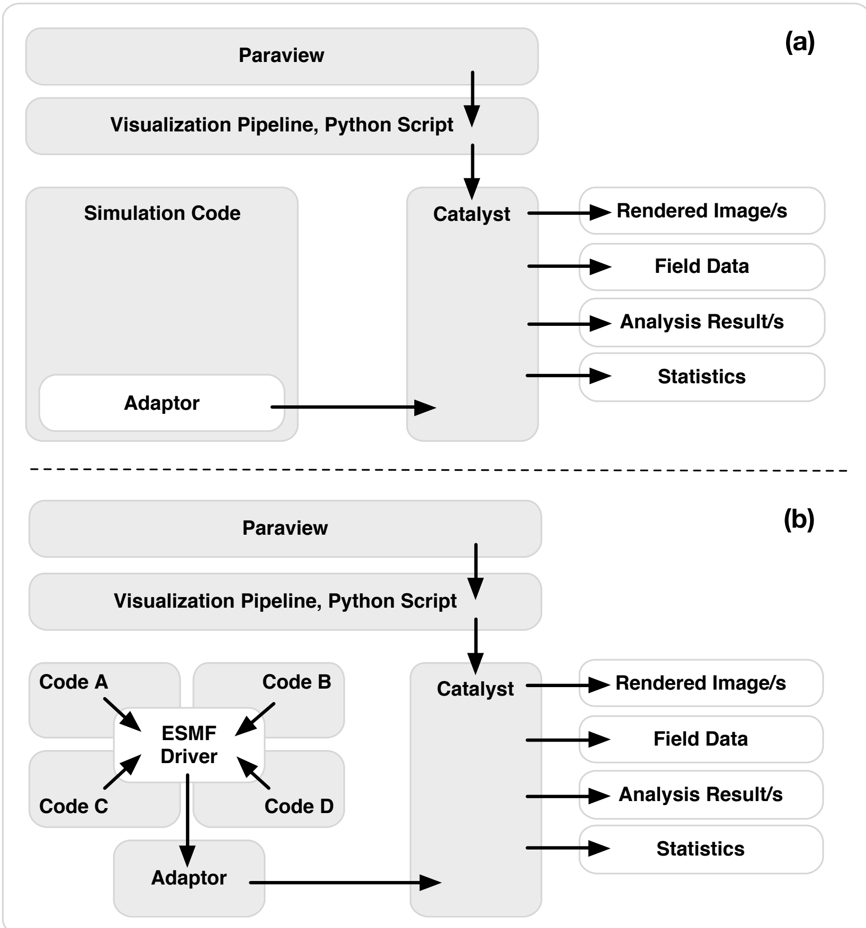


Figure 5. Comparison of the **(a)** conventional and **(b)** ESMF integrated in-situ visualization system

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development Discussions
Open Access

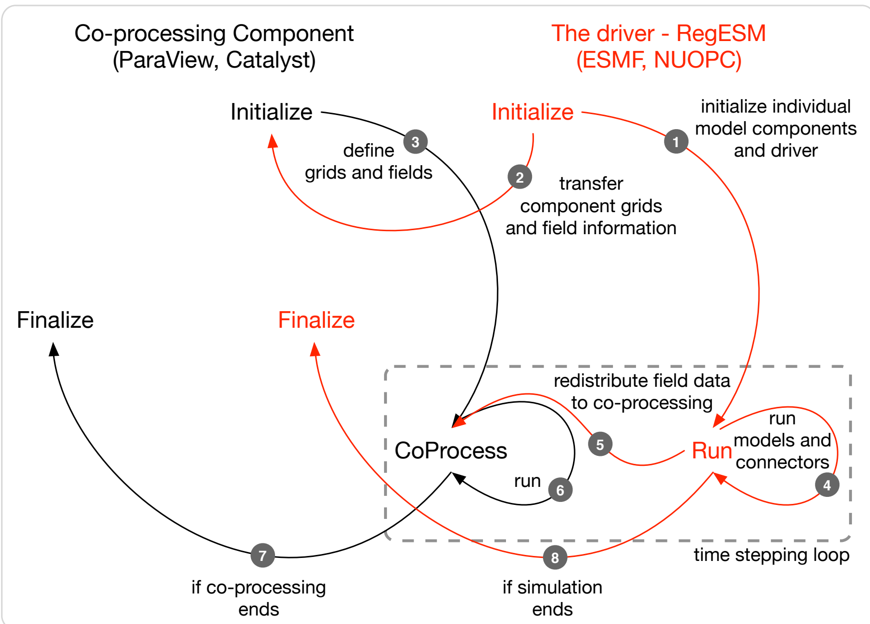


Figure 6. The interaction between driver defined by ESMF, NUOPC and co-processing component (Paraview, Catalyst).

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

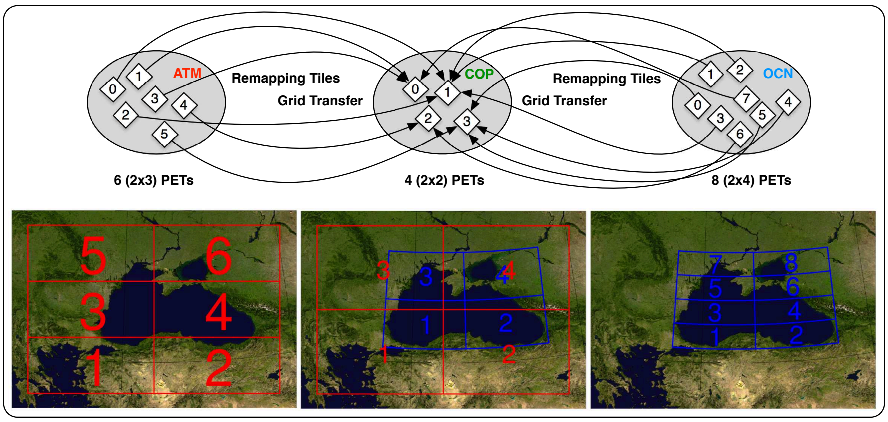


Figure 7. Two component (atmosphere and ocean) representation of grid transfer and remapping feature of ESMF/NUOPC interface.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development Discussions
Open Access

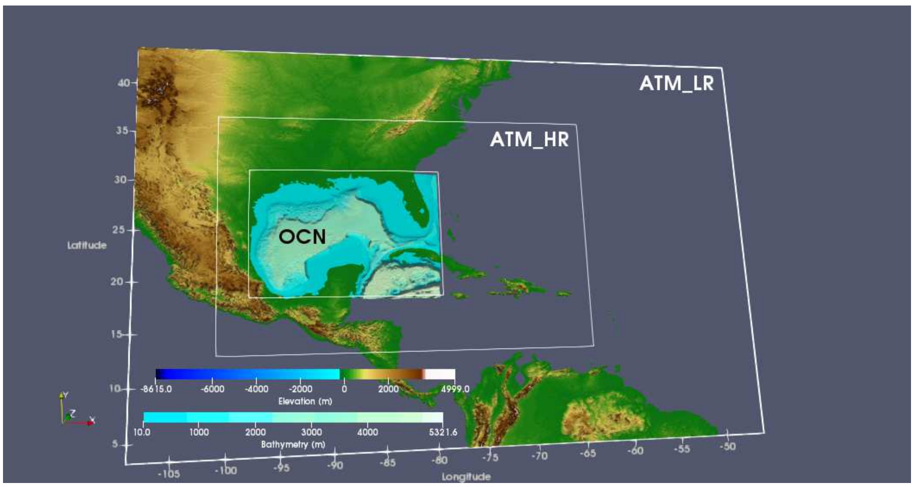
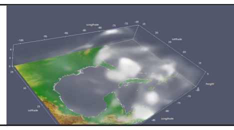
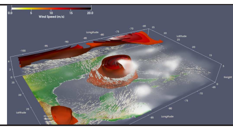
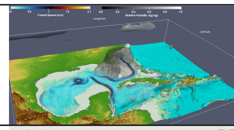
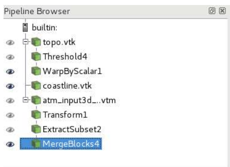
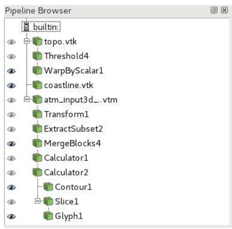
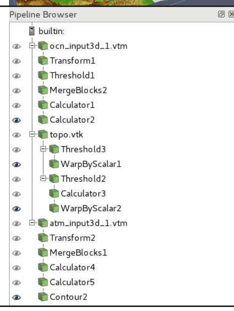


Figure 8. Domain for the RegESM simulations with topography and bathymetry of the region. The solid white boxes represent boundaries of atmosphere (both outer and inner) and ocean model domains.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Table 1. Tested model configurations for benchmark simulations. Note that the dimension of vertical coordinates of ATM and OCN components are shown in here after vertical interpolation from sigma to height and s-coordinates to depth.

	P1: Case I	P2: Case II	P3: Case III
Visualization			
Pipeline			
Primitives	**ATM:** Contour for topography polyline for coastline and direct volume rendering for clouds	**ATM:** same with previous case but it includes iso-surface for wind speed and glyph for wind at specified level	**ATM:** Contour for topography, iso-surface for wind speed colored by relative humidity **OCN:** Contour for bathymetry, direct volume rendering for current
Domain Size	**ATM** **LR:** 170 x 235 x 27 **HR:** 880 x 1240 x 27	**ATM** Same with **Case I**	**ATM** Same with **Case I** **OCN** 653 x 487 x 21
Number of Fields	**1 x 3D ATM** Relative Humidity	**4 x 3D ATM** Relative Humidity Wind (u, v, w)	**4 x 3D ATM** Relative Humidity Wind (u, v, w) **4 x 3D OCN** Ocean Current (u, v, w) Land-Sea Mask
Data Size **ATM+OCN (MB)**	**LR:** 8.3 **HR:** 224.0	**LR:** 33.2 **HR:** 896.0	**LR:** 33.2+25.4 = 58.6 **HR:** 896.0+25.4 = 921.4
Time (s)	**LR:** 2.3 – 3.7 **HR:** 17.7 – 65.0	**LR:** 2.3 – 3.8 **HR:** 18.4 – 79.3	**LR:** 6.8 – 14.6 **HR:** 7.8 – 10.1

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development
Open Access
Discussions

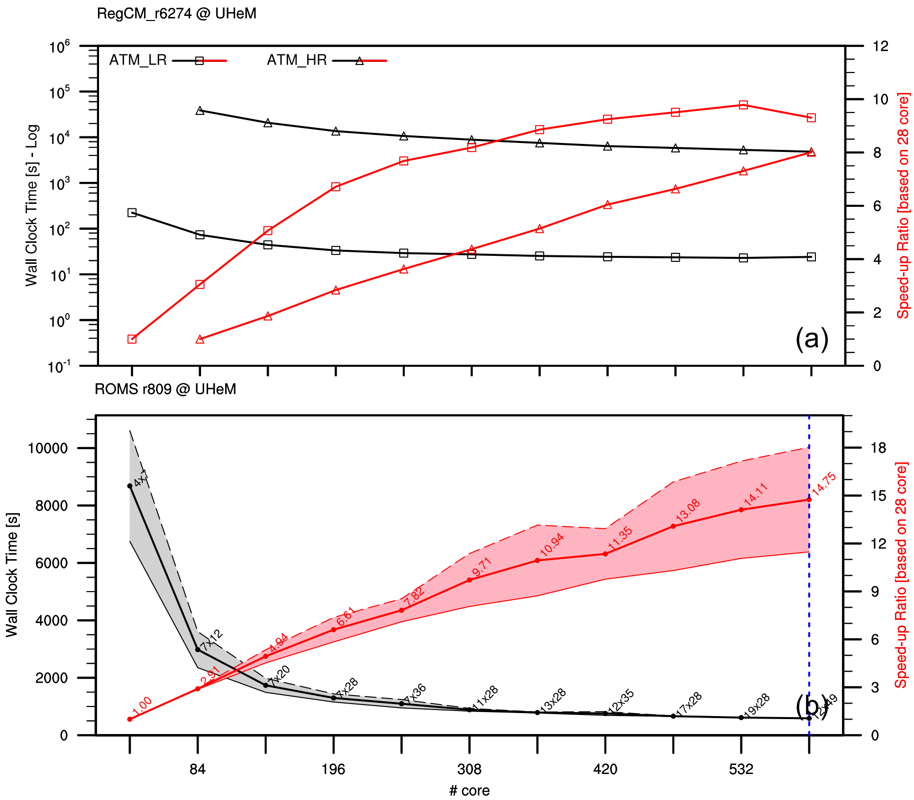


Figure 9. Benchmark results of standalone **(a)** atmosphere (ATM; both LR and HR) and **(b)** ocean (OCN) models. Note that timing results of atmosphere model is in log axes to show both LR and HR cases in the same figure. The black lines represent measured wall clock times in second and red lines (and shaded envelope) show speed-up. The best two-dimensional decomposition parameters, timing results and speed-up are shown as line for the ocean model case.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development Discussions
Open Access

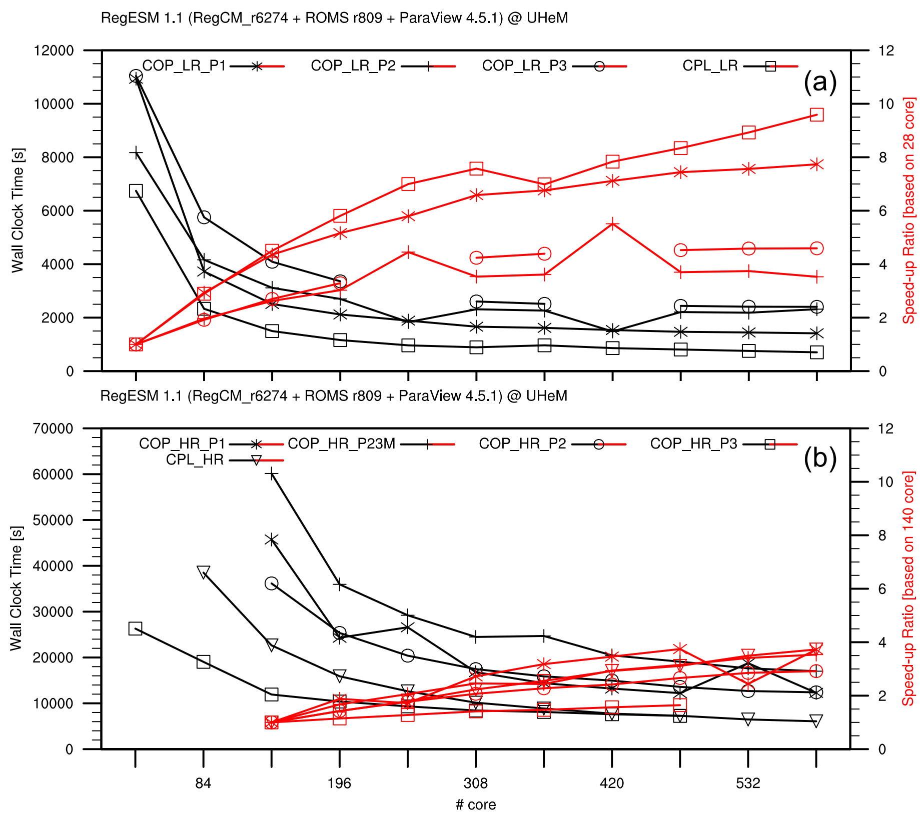


Figure 10. Benchmark results of coupled model **(a)** LR and **(b)** HR cases. CPL represents two-component case and COP shows three-component case including co-processing component. Note that the HR case requires at least 140 cores to run and the speed-up results are given based on 140 cores.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development Discussions
Open Access

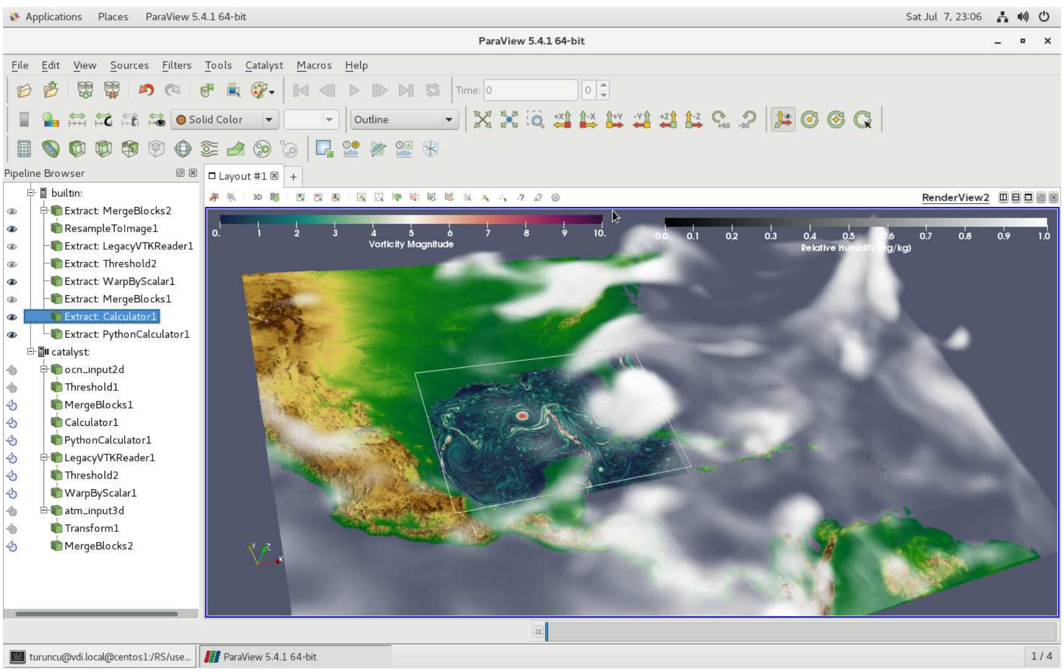


Figure 11. Volume rendering of atmospheric relative humidity field (*atm_input3d*) as well as vorticity field in the ocean surface (*ocn_input2d*) from COP_LR simulation using ParaView Catalyst in live mode.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

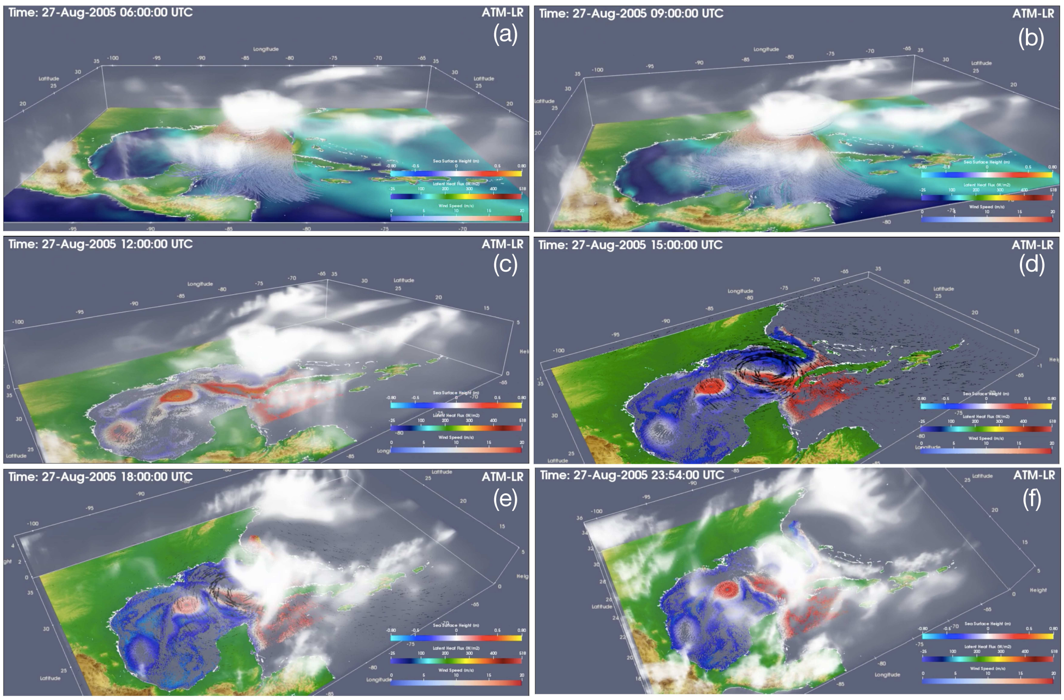


Figure 12. Rendering of multi-component (ATM-OCN-COP) fully coupled simulation using ParaView. The temporal interval for the processed data is defined as 6-minutes. The evolution of the hurricane is shown in supplemental video.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development Discussions
Open Access

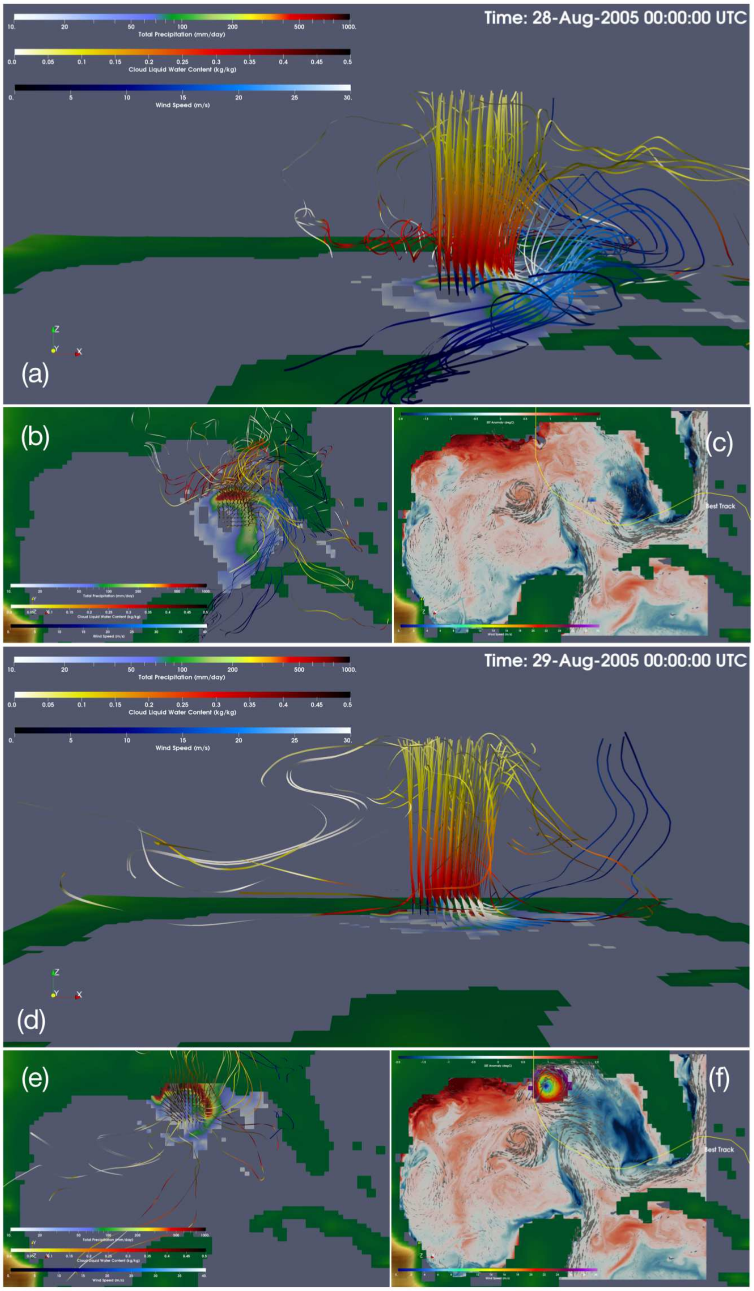


Figure 13. Rendering of three-dimensional vorticity streamlines (1/s), total precipitation (mm/day) and sea surface temperature anomaly (degC) of COP_LR simulation for 28-Aug-2005 00:00 UTC **(a-c)** and 29-Aug-2005 00:00 UTC **(d-f)**. Streamlines are calculated only from the eye of the hurricane. In this case, red and yellow colored forward streamlines represents cloud liquid water content (kg/kg) and blue colored backward streamlines indicates wind speed (m/s). The yellow solid line represents the best track of Hurricane Katrina, which is extracted from HURDAT2 database.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-179
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 27 July 2018

Geoscientific Model Development Discussions
Open Access

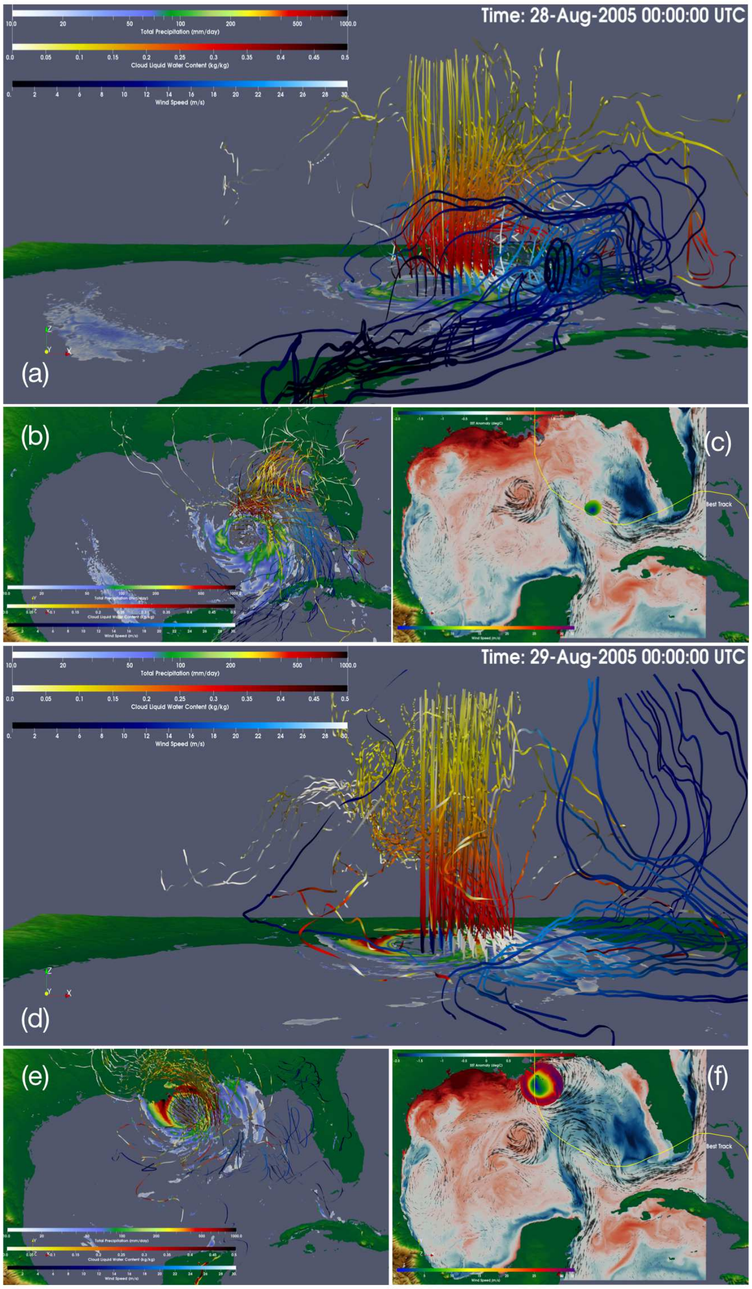


Figure 14. Same with Fig. 13 but for COP_HR simulation. The comparison of low and high resolution model results is shown in supplemental video.