REQUEST FOR A SPECIAL PROJECT 2022–2024

MEMBER STATE: Ireland

Principal Investigator¹: Paul Nolan, paul.nolan@ichec.ie

Affiliation: Irish Centre for High-End Computing, National University of Ireland Galway (NUIG) and Met Éireann

Address: Irish Centre for High-End Computing (ICHEC), IT Building, South Campus, National University of Ireland Galway (NUIG), Galway, Ireland

Other researchers: Dr Jason Flanagan, jason.flanagan@ichec.ie Irish Centre for High-End Computing, NUIG

Project Title: High-Resolution Downscaled RCM-CMIP6 Simulations

<table>
<thead>
<tr>
<th>If this is a continuation of an existing project, please state the computer project account assigned previously.</th>
<th>SP spienola</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting year: (A project can have a duration of up to 3 years, agreed at the beginning of the project.)</td>
<td>2022</td>
</tr>
<tr>
<td>Would you accept support for 1 year only, if necessary?</td>
<td>YES X</td>
</tr>
</tbody>
</table>

Computer resources required for 2022-2024:
(To make changes to an existing project please submit an amended version of the original form.)

<table>
<thead>
<tr>
<th></th>
<th>2022</th>
<th>2023</th>
<th>2024</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Performance Computing Facility (SBU)</td>
<td>60 million</td>
<td>70 million</td>
<td>60 million</td>
</tr>
<tr>
<td>Accumulated data storage (total archive volume)²</td>
<td>30,000</td>
<td>30,000</td>
<td>30,000</td>
</tr>
</tbody>
</table>

¹ The Principal Investigator will act as contact person for this Special Project and, in particular, will be asked to register the project, provide annual progress reports of the project’s activities, etc.
² These figures refer to data archived in ECFS and MARS. If e.g. you archive x GB in year one and y GB in year two and don’t delete anything you need to request x + y GB for the second project year etc.

Continue overleaf
Principal Investigator:  Paul Nolan, paul.nolan@ichec.ie

Project Title:  High-Resolution Downscaled RCM-CMIP6 Simulations

Extended abstract

Note that the proposed project is a continuation of a previous Special Project (spiennola): “High Resolution EC-Earth Simulations”. The goal of this previous research project was two-fold; (i) develop, tune and test the EC-Earth modelling system in preparation for CMIP6, (ii) and simulate the effects of climate change at the global scale. The following EC-Earth CMIP6 contributions were completed:

- 7 x EC-Earth AOGCM Historical/Veg Simulations 1850-2014
- 28 x EC-Earth AOGCM/Veg ScenarioMIP 2015-2100 Simulations; 7 x SSP1-2.6, 7 x SSP2-4.5, 7 x SSP3-7.0 & 7 x SSP5-8.5

The data were fully validated and analysed (e.g., Döscher et al., 2021; Nolan et al., 2020b) and shared with the international research community on the ESGF data hosting system. The resulting datasets comprise Ireland’s contribution to the Coupled Model Intercomparison Project (phase 6) (CMIP6) and are currently being analysed for inclusion in the U.N. Intergovernmental Panel on Climate Change (IPCC) AR6 reports.

The scientific objectives of the proposed project are to dynamically downscale CMIP6 EC-Earth data to provide high resolution regional climate projections for Europe (Euro-CORDEX) and Ireland using both standard atmosphere-only (WRF and COSMO-CLM) and coupled atmosphere-ocean-wave (COAWST) Regional Climate Models. The resulting datasets will add to a larger ensemble of RCM data currently being produced at a national and European scale (EURO-CORDEX).

1. Regional Climate Modelling

The impact of increasing greenhouse gases and changing land use on climate change can be simulated using Global Climate Models (GCMs). However, due to computational constraints, long climate simulations using ensembles of GCMs are currently feasible only with horizontal resolutions of ≈50 km or coarser. Because climate fields such as precipitation, wind speed and temperature are closely correlated to the local topography, this is inadequate to simulate the detail and pattern of climate change and its effects on the future climate at a regional scale. Furthermore, and of particular relevance to Ireland and Western Europe, numerous studies have shown that, even at 50 km grid spacing, GCMs severely under-resolve both the number and intensity of cyclones (e.g., Zhao et al., 2009; Camargo, 2013; Zappa et al., 2013).

To overcome these limitations, the RCM method dynamically downscales the coarse information provided by the global models and provides high-resolution information on a subdomain covering Ireland. The computational cost of running the RCM, for a given resolution, is considerably less than that of a global model. The approach has its flaws: all models have errors, which are cascaded in this technique, and new errors are introduced via the flow of data through the boundaries of the regional model. Nevertheless, numerous studies have demonstrated that high-resolution RCMs improve the simulation of fields, such as precipitation (Lucas-Picher et al., 2012; Kendon et al., 2012, 2014; Bieniek et al., 2015; Nolan et al., 2015, 2017, 2020a) and topography-influenced phenomena and extremes with relatively small spatial or short temporal character (Feser et al., 2011; Feser and Barcikowska, 2012; Shkol’nik et al., 2012; IPCC, 2013a). An additional advantage is that the physics-based RCMs explicitly resolve more small-scale atmospheric features and provide a better representation of convective precipitation (Rauscher et al., 2010) and extreme precipitation (Kanada
et al., 2008; Nolan et al., 2017). Other examples of the added value of RCMs are improved simulations of near-surface temperature (Feser, 2006; Di Luca et al., 2016; Nolan, 2020a), European storm damage (Donat et al., 2010), strong mesoscale cyclones (Cavicchia and Storch, 2011), North Atlantic tropical cyclone tracks (Daloz et al., 2015) and near-surface wind speeds (e.g. Kanamuru and Kanamitsu, 2007; Nolan et al., 2014; Nolan, 2015), particularly in coastal areas with complex topography (Feser et al., 2011; Winterfeldt et al., 2011). The added value of RCMs in the simulation of cyclones is particularly important for the current study, as low pressure systems are the main delivery mechanism for precipitation and wind in Ireland and Western Europe.

Furthermore, numerous studies have demonstrated that increased RCM spatial resolution results in a more accurate representation of the climate system. Low-resolution RCMs use parameterised convection schemes, meaning that the heaviest precipitation events (e.g. convective systems on hot summer days) may not be adequately represented in the simulations (Prein et al., 2013; Kendon et al., 2014). Zängl et al. (2015) investigated heavy rainfall events over the North-Alpine region and found that increasing the mesh size (9, 3 and 1 km) resulted in a stepwise improvement in skill. Similarly, Nolan et al. (2017) found that RCM accuracy increased with higher spatial resolution; however, reducing the horizontal grid spacing below 4 km provided relatively little added value. The Intergovernmental Panel on Climate Change (IPCC) has concluded that there is “high confidence that downscaling adds value to the simulation of spatial climate detail in regions with highly variable topography (e.g., distinct orography, coastlines) and for mesoscale phenomena and extremes” (IPCC, 2013).

1.2 Coupled Atmosphere-Ocean-Wave Regional Modelling

Numerous studies have demonstrated improved skill in the simulations of storms when using regional coupled atmosphere-ocean-wave regional models in place of standalone atmosphere models (e.g., Chen et al. 2007; Warner et al. 2010; Liu et al. 2011; Renault et al. 2012; Zhao et al. 2017; Planezze et al. 2018). Studies such as Stūld et al. (2015), Staneva et al. (2016) and Wahle et al. (2016) have demonstrated that surface wind speeds, waves and storm surges provide a closer match to observations when simulated with a coupled atmosphere-wave model. Other studies demonstrated that regional coupled models better represent air-sea interaction during high-impact weather events such as cyclones and depressions (Ratnam et al. 2009; Samala et al. 2013). The number of convective-resolving coupled atmosphere-ocean-wave models currently in operation is limited. The most widely used coupled model is the COAWST modelling system, developed by the United States Geological Survey (USGS) and comprising of (i) the WRF atmospheric numerical weather model, (ii) the Regional Oceanic Modelling System (ROMS), (iii) the SWAN wave model developed by Delft University, (iv) WAVEWATCH III, (v) the Infragravity wave model and (vi) the National Community Sediment Transport Model (NCSTM) model. These model components are coupled together using the Model Coupling Toolkit (MCT) and the Spherical Coordinate Remapping and Interpolation Package (SCRIP). Numerous international studies have demonstrated the advantages of using COAWST in place of non-coupled models. For example, COAWST was used at high resolutions (up to 3 km for the atmosphere and up to 1 km for the ocean and wave models) in the Mediterranean region (Renault et al. 2012; Carniel et al. 2016; Ricchi et al. 2016). These studies highlight that high-resolution coupling improves the simulation results significantly. Olabarrieta et al. (2012) used COAWST to simulate hurricane Ida and demonstrated substantial improvement over uncoupled model systems. Zambon et al. (2014) used COAWST to simulate hurricane Ivan and compared the results with those from WRF atmosphere-only simulations. The results show a dramatic improvement by the COAWST coupled model in the simulation of ocean and atmosphere parameters during, and after, the hurricane. Olabarrieta et al. (2011) simulated a large storm event affecting Willapa Bay (Washington State) during 22 to 29 October 1998 and found that model results were in very good agreement with observed water elevations, currents and wave measurements.
2. Work Plan

The main component of the proposed research involves simulating regional climate change in Ireland (and Europe via EURO-CORDEX) at fine detail. This is currently being achieved by running a large ensemble of high-resolution downscaled simulations using the most up-to-date Regional Climate Models (RCMs) (both standard and coupled atmosphere-ocean-wave), CMIP6 Global Climate Models and all four “tier-1” SSP emission scenarios (SSP1-2.6, SSP2-4.5, SSP3-7.0 & SSP5-8.5) for the period 1980-2100. Additionally, the accuracy of the model predictions is enhanced by increasing the model resolution (~4km) and using fully coupled atmosphere-ocean-wave RCMs. Furthermore, the RCM work will contribute to the CORDEX project by running the required outer domain of selected RCM simulations on the Euro-CORDEX domains, conforming to the CORDEX standards, and extending the simulation period to 1950-2100 (for outer domain).

As part of the proposed Special Project, one EC-Earth ensemble member will be downscaled using the COSMO-CLM, WRF and COAWST RCMs for the historical period 1980-2014 and future period 2015-2100 under all four ScenarioMIP “tier 1” SSPs; SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5. The resulting datasets will add to a larger ensemble of RCM data currently being produced at a national and European scale (EURO-CORDEX).

Specifically, the below simulations will be run as part of the proposed Special Project:

- 3 historical simulations (1980-2014) RCM-EC-Earth simulations
- 12 future simulations (2015-2100) RCM-EC-Earth simulations; 3 RCMs x 4 SSPs (SSP1-2.6, SSP2-4.5, SSP3-7.0 & SSP5-8.5)

2.1 Regional Climate Models

The future regional climate of Europe and Ireland will be simulated at high-spatial resolution using both standard (COSMO-CLM; v5 and WRF; v4) and coupled atmosphere-ocean-wave (COAWST v3.5) RCMs. The COSMO-CLM RCM is the COSMO weather forecasting model in climate mode (Rockel et al., 2008). The Consortium for Small-scale Modeling–Climate Limited-area Modelling COSMO model is the non-hydrostatic operational weather prediction model used by the German weather service (Deutscher Wetterdienst; DWD). The proposed WRF simulations will adapt the Advanced Research WRF (ARW, v4) dynamical core, with development led by the US National Center for Atmospheric Research (NCAR) (Skamarock et al., 2008; Powers et al., 2017).

The Coupled-Ocean-Atmosphere-Wave-Sediment-Transport (COAWST) Modelling System (Warner et al., 2010) is an agglomeration of open-source modelling components that has been tailored to investigate coupled processes of the atmosphere and ocean in coastal regions. The proposed project team has implemented COAWST (v3.4) on ICHEC’s HPC system Kay. This version of COAWST is composed of the following:

- Coupler: - Model Coupling Toolkit (MCT) v 2.6.0
- Ocean: - Regional Ocean Modeling System (ROMS) svn 934
- Atmosphere: - Weather Research and Forecasting Model (WRF) v 4.0.3
- Wave(s): - Simulating Waves Nearshore (SWAN) v 41.20 and WAVEWATCH III (WW3) v 5.16
- Infragravity wave model (InWave) v 1.1
- Sediment Transport: - Community Sediment Transport Modeling Systems (CSTMS)
- Sea Ice: - Sea Ice model

The RCM configurations were validated by downscaling European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim and ERA5 reanalyses for multi-decadal time periods and comparing the output with observational data. Extensive validations were carried out to test the ability
of the RCMs to accurately model the climate of Ireland. For an in-depth validation of the RCMs, please refer to Nolan (2015) and Nolan et al. (2017, 2020a), Flanagan et al. (2019, 2020) and Werner et al. (2019), whose results confirm that the output of the RCMs exhibit reasonable and realistic features, as documented in the historical data record, and consistently demonstrate improved skill over the GCMs. The results of these validation analyses confirm that the RCM configurations and domain size of the current study are capable of accurately simulating the climate of Ireland.

We have already started downscaling a subset of CMIP6 EC-Earth data using WRF, COSMO-CLM and COAWST on the ICHEC system. Figure 1 presents preliminary projections of 2m temperature under four SSPs emission scenarios.

Figure 1. Preliminary RCM-EC-Earth (CMIP6) projections of mean annual 2-m temperature. All RCM ensemble members were run with 4-km grid spacing. In each case, the future 30-year period is compared with the past period, 1981–2010. Current work is focusing on greatly increasing the ensemble size.

The most up-to-date versions of WRF, COSMO-CLM and COAWST were installed and tested on the ICHEC kay machine using the intel compilers. The WRF RCM domains are presented in Figure 2(a); the outer and inner domains have 20km and 4km grid spacings, respectively. The RCMs are very computationally expensive so a careful scale testing was completed to determine the optimal configuration and ensure an efficient use of compute resources. It was found that use of classic-netcdf (with no compression) leads to an increase in efficiency of ~15% over netcdf4. Furthermore, it was demonstrated that using adaptive time stepping adds an additional increase in efficiency of ~5%. The timing results for a one-year WRF simulation are presented in Figure 2(b). An absence of timing data for certain node counts implies the simulation failed. The timing results show that the WRF
configuration scales well up to 11 nodes (440 cores) with the optimal setup using a hybrid MPI-OpenMP compilation (with 4 or 5 OpenMP threads per node being optimal). The results show that one year of a WRF simulations takes 20.4 hours. When accounting for data pre-processing and queue time, it takes approximately 21 hours to complete one year of a WRF simulation. A careful scale testing of the COSMO-CLM and COAWST simulations was also completed. It was found that COSMO-CLM and COAWST are 0.7 and 1.7 times as expensive compared to WRF, respectively.

3. Justification of Computing and Storage Resources

As the Atos Sequana XH2000 HPCF machine will not be operational until 2022, it is not possible to accurately assess scaling statistics. However, it can be assumed that the new ECMWF system will be significantly more powerful when compared to the ICHEC machine (installed in 2018). To this end, for the purpose of the current proposal, we assume that 11 nodes on the ICHEC machine (440 cores) is comparable to 3 nodes (384 cores) on the new ECMWF system. Any surplus of compute resources will be used for pre-/post-processing of data, model testing and used to contribute to the CORDEX project by running the required outer domains of selected RCM simulations on the Euro-CORDEX domains, conforming to the CORDEX standards, and extending the simulation period to 1950-2100 (for the outer domain).

The above scaling results were used to provide the resource request figures of Table 1 using 3 nodes on the Atos Sequana system and the following SBU calculations:

\[ \text{SBU} = \text{compute time} \times \text{number of physical CPUs} \times 18.84 \]

![Figure 2. (a) WRF Test-domains, (b) timing results on ICHEC kay machine using the intel compilers](image)
<table>
<thead>
<tr>
<th>Description</th>
<th>Simulation time [hr]</th>
<th>Total SBUs = $4 \times 128 \times [\text{simulation_time}_{\text{hr}}] \times 18.84$</th>
<th>Total Archive</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experiment 1: WRF-CMIP6</strong></td>
<td>7731</td>
<td>56 million</td>
<td>20 TB</td>
</tr>
<tr>
<td>35-year WRF historical simulations (1980-2014)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 × 86-year future WRF simulations (2015-2100)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Experiment 2: COAWST-CMIP6</strong></td>
<td>13143</td>
<td>95 million</td>
<td>50TB</td>
</tr>
<tr>
<td>35-year COAWST historical simulations (1980-2014)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 × 86-year future COAWST simulations (2015-2100)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Experiment 3: COSMO-CLM-CMIP6</strong></td>
<td>5412</td>
<td>39 million</td>
<td>20 TB</td>
</tr>
<tr>
<td>35-year COSMO-CLM historical simulations (1980-2014)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 × 86-year future COSMO-CLM simulations (2015-2100)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>26286</td>
<td>190 million</td>
<td>90 TB ††</td>
</tr>
</tbody>
</table>

The WRF, COSMO-CLM and COAWST simulations produce 21, 20 and 40GB of data per simulation year, respectively resulting in a total archive volume of 92 TB. Furthermore, storage resources of approximately 50 GB per simulation year will be required for short-term storage of RCM boundary data. Taking the above into account, we request 90TB of storage over the 3-year project timeframe. This estimate is an absolute upper value as boundary data will be regularly deleted and the RCM data will be transferred to local storage resources as the simulations complete.

**References**


Flanagan, J. and Nolan, P.: Towards a Definitive Historical High-resolution Climate Dataset for Ireland – Promoting Climate Research in Ireland, EPA Research 350, available at: https://www.epa.ie/pubs/reports/research/climate/researchreport350/


Zambon, J. B., R. He, and J. C. Warner (2014), Investigation of hurricane Ivan using the coupled ocean-atmosphere-wave-sediment transport (COAWST) model, *Ocean Dynamics*, 64(11), 1535-1554


