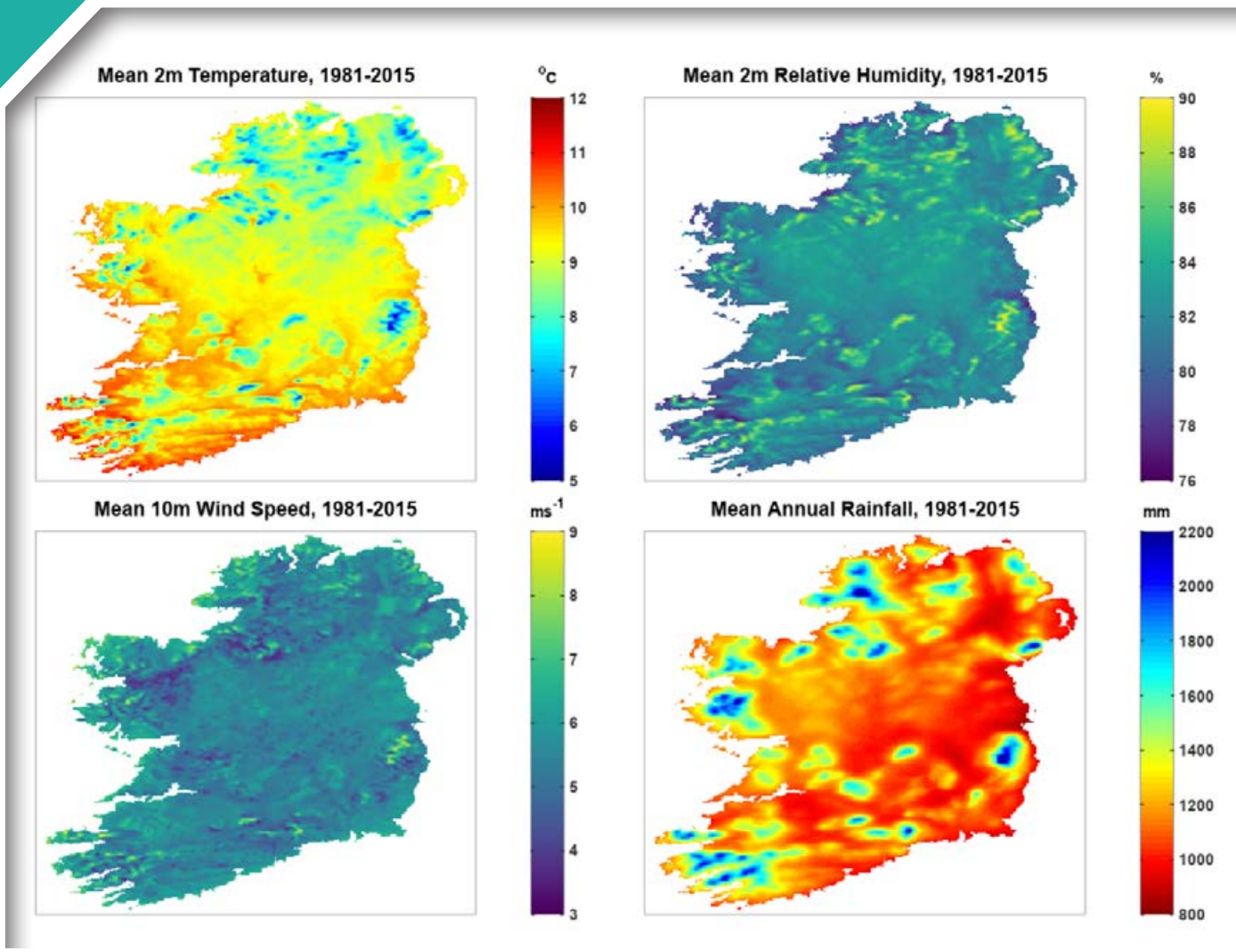
epaResearch
Climate - Water - Sustainability
Identifying pressures • Informing policy • Developing solutions

Report No.350

Towards a Definitive Historical High-resolution Climate Dataset for Ireland – Promoting Climate Research in Ireland

Authors: Jason Flanagan and Paul Nolan



www.epa.ie

Rialtas na hÉireann
Government of Ireland

ENVIRONMENTAL PROTECTION AGENCY

The Environmental Protection Agency (EPA) is responsible for protecting and improving the environment as a valuable asset for the people of Ireland. We are committed to protecting people and the environment from the harmful effects of radiation and pollution.

The work of the EPA can be divided into three main areas:

Regulation: *We implement effective regulation and environmental compliance systems to deliver good environmental outcomes and target those who don't comply.*

Knowledge: *We provide high quality, targeted and timely environmental data, information and assessment to inform decision making at all levels.*

Advocacy: *We work with others to advocate for a clean, productive and well protected environment and for sustainable environmental behaviour.*

Our Responsibilities

Licensing

We regulate the following activities so that they do not endanger human health or harm the environment:

- waste facilities *(e.g. landfills, incinerators, waste transfer stations)*;
- large scale industrial activities (*e.g. pharmaceutical, cement manufacturing, power plants)*;
- intensive agriculture *(e.g. pigs, poultry)*;
- the contained use and controlled release of Genetically Modified Organisms *(GMOs)*;
- sources of ionising radiation *(e.g. x-ray and radiotherapy equipment, industrial sources)*;
- large petrol storage facilities;
- waste water discharges;
- dumping at sea activities.

National Environmental Enforcement

- Conducting an annual programme of audits and inspections of EPA licensed facilities.
- Overseeing local authorities' environmental protection responsibilities.
- Supervising the supply of drinking water by public water suppliers.
- Working with local authorities and other agencies to tackle environmental crime by co-ordinating a national enforcement network, targeting offenders and overseeing remediation.
- Enforcing Regulations such as Waste Electrical and Electronic Equipment (WEEE), Restriction of Hazardous Substances (RoHS) and substances that deplete the ozone layer.
- Prosecuting those who flout environmental law and damage the environment.

Water Management

- Monitoring and reporting on the quality of rivers, lakes, transitional and coastal waters of Ireland and groundwaters; measuring water levels and river flows.
- National coordination and oversight of the Water Framework Directive.
- Monitoring and reporting on Bathing Water Quality.

Monitoring, Analysing and Reporting on the Environment

- Monitoring air quality and implementing the EU Clean Air for Europe (CAFÉ) Directive.
- Independent reporting to inform decision making by national and local government *(e.g. periodic reporting on the State of Ireland's Environment and Indicator Reports)*.

Regulating Ireland's Greenhouse Gas Emissions

- Preparing Ireland's greenhouse gas inventories and projections.
- Implementing the Emissions Trading Directive, for over 100 of the largest producers of carbon dioxide in Ireland.

Environmental Research and Development

- Funding environmental research to identify pressures, inform policy and provide solutions in the areas of climate, water and sustainability.

Strategic Environmental Assessment

- Assessing the impact of proposed plans and programmes on the Irish environment *(e.g. major development plans)*.

Radiological Protection

- Monitoring radiation levels, assessing exposure of people in Ireland to ionising radiation.
- Assisting in developing national plans for emergencies arising from nuclear accidents.
- Monitoring developments abroad relating to nuclear installations and radiological safety.
- Providing, or overseeing the provision of, specialist radiation protection services.

Guidance, Accessible Information and Education

- Providing advice and guidance to industry and the public on environmental and radiological protection topics.
- Providing timely and easily accessible environmental information to encourage public participation in environmental decision-making *(e.g. My Local Environment, Radon Maps)*.
- Advising Government on matters relating to radiological safety and emergency response.
- Developing a National Hazardous Waste Management Plan to prevent and manage hazardous waste.

Awareness Raising and Behavioural Change

- Generating greater environmental awareness and influencing positive behavioural change by supporting businesses, communities and householders to become more resource efficient.
- Promoting radon testing in homes and workplaces and encouraging remediation where necessary.

Management and structure of the EPA

The EPA is managed by a full time Board, consisting of a Director General and five Directors. The work is carried out across five Offices:

- Office of Environmental Sustainability
- Office of Environmental Enforcement
- Office of Evidence and Assessment
- Office of Radiation Protection and Environmental Monitoring
- Office of Communications and Corporate Services

The EPA is assisted by an Advisory Committee of twelve members who meet regularly to discuss issues of concern and provide advice to the Board.

EPA RESEARCH PROGRAMME 2014–2020

Towards a Definitive Historical High-resolution Climate Dataset for Ireland – Promoting Climate Research in Ireland

(2016-CCRP-FS.28)

EPA Research Report

Prepared for the Environmental Protection Agency

by

Irish Centre for High-End Computing

Authors:

Jason Flanagan and Paul Nolan

ENVIRONMENTAL PROTECTION AGENCY
An Ghníomhaireacht um Chaomhnú Comhshaoil
PO Box 3000, Johnstown Castle, Co. Wexford, Ireland

Telephone: +353 53 916 0600 Fax: +353 53 916 0699
Email: info@epa.ie Website: www.epa.ie

ACKNOWLEDGEMENTS

This report is published as part of the EPA Research Programme 2014–2020. The EPA Research Programme is a Government of Ireland initiative funded by the Department of the Environment, Climate and Communications. It is administered by the Environmental Protection Agency, which has the statutory function of co-ordinating and promoting environmental research.

The authors would like to acknowledge the members of the project steering committee, namely Seamus Walsh (Met Éireann), Ray McGrath (Meteorology and Climate Centre, University College Dublin), Saji Varghese (Met Éireann), Frank McGovern (EPA), Phillip O'Brien (EPA) and Anne Mason (Project Manager on behalf of EPA Research).

DISCLAIMER

This report is based on research carried out/data from January 2017 to January 2019. More recent data may have become available since the research was completed.

The EPA Research Programme addresses the need for research in Ireland to inform policymakers and other stakeholders on a range of questions in relation to environmental protection. These reports are intended as contributions to the necessary debate on the protection of the environment.

EPA RESEARCH PROGRAMME 2014–2020
Published by the Environmental Protection Agency, Ireland

ISBN: 978-1-84095-952-9 October 2020

Price: Free Online version

Project Partners

Jason Flanagan
Irish Centre for High-End Computing
NUI Galway
Co. Galway
Ireland
Tel.: + 353 1 524 1608 (ext. 24)
Email: jason.flanagan@ichec.ie

Paul Nolan
Irish Centre for High-End Computing
Technology & Enterprise Campus
Grand Canal Quay
Dublin 2
Ireland
Tel.: + 353 1 524 1608 (ext. 24)
Email: paul.nolan@ichec.ie

Alastair McKinstry
Irish Centre for High-End Computing
NUI Galway
Co. Galway
Ireland
Tel.: + 353 1 524 1608 (ext. 68)
Email: alastair.mckinstry@ichec.ie

Contents

List of Figures

List of Tables

Executive Summary

There is strong and constant demand from various sectors (research, industry and government) for long-term, high-resolution (both temporal and spatial) gridded climate datasets. To help address this demand, in 2017 the Irish Centre for High-End Computing (ICHEC) completed two high-resolution simulations of the Irish climate, utilising the regional climate models COSMO-CLM5 (1.5 km) and the Weather Research and Forecasting (WRF) model v3.7.1 (2 km). In addition, in 2017 the Irish meteorological service Met Éireann completed a 36-year reanalysis with data assimilation (MÉRA) at 2.5 km resolution for the period 1981–2016. All three datasets have been archived at hourly intervals and contain hourly outputs for an array of subsurface, surface and atmospheric fields for the entire 36-year period from 1981 to 2016. Together, the three datasets represent the most comprehensive, up-to-date and accurate picture of the historical (1981–present) Irish climate.

It should be expected that different models will exhibit different errors as per the climate field in question. A priori knowledge of data errors helps end-users tailor their needs accordingly. For instance, an application may require fine resolution without too much regard to minor errors. Or an application may be highly dependent on the accuracy of the underlying dataset.

In this work, a full list of the climate variables that have been produced by these simulations is provided. Uncertainty estimates and skill scores for several of the more in-demand variables (essential climate variables; ECVs), as determined from potential end-users, have been produced utilising a variety of observations (station, satellite, turbine and radiosonde) and are presented here. Gridded datasets that are of interest to the renewable energy sector (winds at turbine heights and solar fields) have been validated and are described.

A total of 36 climate change indices (Table ES.1), 27 of which have been recommended by the Expert Team on Climate Change Detection and Indices, have been produced and analysed. These indices will be of interest to climate change researchers as well as public health and planning authorities, among others (e.g. the Office of Public Works, Teagasc and the National Parks and Wildlife Service).

A total of 23 climate indices that are of interest to the agricultural sector have also been produced for the period 1981–2000 and are presented. The list includes average daily temperature and temperature range; average January/July temperature values; average monthly rainfall; mean number of wet (1 mm precipitation) days; mean number of 5, 10, 15, 20, 25 and 30 mm precipitation days; mean January, July and annual relative humidity; average annual wind speed; air frost indices; 5-year return period rainfall amounts (at several timescales); and a driving rain index. Some of the more commonly used climate variables and all the climate indices have been made available online and their location is provided.

Several recommendations based on the research undertaken during the lifetime of this project are made. These include the following:

- The use of fifth generation European Centre for Medium-Range Weather Forecasts atmospheric reanalysis of the global climate (ERA5) data to extend the reanalysis/downscalings described here would result in even greater accuracy and temporal coverage.
- The validated gridded climate variables that are now available can provide the data resources for numerous new studies that will be of benefit to Ireland – for example, they can motivate an improved understanding of climate change at local levels and extreme weather events.
- MÉRA shows the least errors for most of the climate variables examined (hourly/daily precipitation, hourly/daily 2 m temperature, 10 m wind speed and direction, relative humidity, sea level pressure, global irradiance, upper-air wind speed and direction) and should be considered the primary source. However, the finer resolution COSMO-CLM and/or WRF data may be preferred for variables for which there is little error difference (e.g. WRF for 10 m and upper-air winds) or for which the corresponding MÉRA variable is

Table ES.1. The 36 climate change indices with descriptions produced from temperature and precipitation fields

Climate index	Description
FD, ID and SD	Annual number of frost, icing and summer days
TN	Annual number of tropical nights
GSL and GSS	Growing season length and start
TX_X and TN_X	Monthly maxima of daily maximum and minimum temperatures
TX_N and TN_N	Monthly minima of daily maximum and minimum temperatures
DTR	Daily temperature range
TN10p and TX10p	Percentage of days when the minimum/maximum temperature is less than the 10th percentile (1961–1990) calendar day minimum/maximum
TN90p and TX90p	Percentage of days when the minimum/maximum temperature is greater than the 90th percentile (1961–1990) calendar day minimum/maximum
WSDI and CSDI	Warm spell and cold spell duration index
NWSP and NCSP	Number of warm and cold spells per period
SDII	Simple precipitation intensity index
R*N*mm (*N* = 1, 5, 10, 15, 20, 25, 30)	Number of days with precipitation greater than 1, 5, 10, 15, 20, 25 and 30 mm
CDD and CWD	Number of consecutive dry days and consecutive wet days
CDDP and CWDP	Number of periods of more than 5 CDDs or CWDs
PRCPTOT	Total annual precipitation from wet days
Rx1day and Rx5day	Monthly maximum 1-day and consecutive 5-day precipitation amounts
R95pTOT and R99pTOT	Percentage of annual precipitation from wet days that exceeds the 95th and 99th percentile of (wet day) precipitation in the period 1961–1990

unavailable (e.g. COSMO-CLM5 for convective available potential energy 3 km, Showalter index and surface lifted index).

- Some of the rainfall and temperature indices examined suggest changes in rainfall and temperature patterns. A more robust statistical approach (e.g. non-parametric trend analysis, further regional and/or monthly/seasonal analysis) would shed further light on the perceived changes.
- The solar fields examined suggest a strong north–south gradient with an along-coastline resource that could be exploited for future installations.

1 Introduction

There is a constant demand from industry, research and governmental sectors for high-quality, long-term gridded climate datasets with high spatial and temporal resolution for conducting climate research. Such datasets have the potential to be utilised in a wide range of applications, including agricultural (Collins *et al*., 1996), hydrology (Duethmann *et al*., 2013), public health (Ziese *et al*., 2014), energy (Šúri *et al*., 2007) and fundamental studies (Becker *et al*., 2013) of observed climate change trends and variability.

In Ireland, station observations have traditionally been used to describe the Irish climate and produce gridded datasets. For instance, daily and monthly gridded datasets (at 1 km resolution) of precipitation have been created for Ireland (Walsh, 2012, 2016) and are based on station data from Met Éireann's rainfall network. The identification of changes in Irish precipitation patterns, whether they are driven by natural variability or man-made climate change, is particularly important to the country, with recent projections pointing to an increased likelihood of summer droughts and winter flooding (Nolan *et al.*, 2013a,b). Unfortunately, gridded datasets based on station observations come with numerous caveats, as detailed by Prein and Gobiet (2017): they may not be particularly representative in regions with few stations and station densities may change over time; and station data are prone to error and/or missing values, precipitation under-catch and excessive smoothing. Furthermore, some observational datasets may have inhomogeneities, caused by changes in instrumentation, location or methods of measurement over time.

Gridded observational climate datasets for a wide range of parameters are not readily available for climate research applications in Ireland. Climatologically important variables, such as wind speed and direction, humidity and radiation, are measured at a limited number of weather stations.

The reanalysis outputs from numerical weather models represent an alternative to observations for the production of gridded datasets. These outputs provide a consistent physical analysis over time. The European Centre for Medium-Range Weather Forecasts (ECMWF) has initiated several global reanalysis datasets beginning with ERA-15 (1979–1993; 190 km resolution; Gibson *et al*., 1995). Model improvements over the intervening years have led to higher resolution datasets: ERA-40 (1957–2002; 125 km; Uppala *et al*., 2005); ERA-Interim (1979–present; 80 km; Dee *et al*., 2011) and, more recently, ERA5 (31 km; 1950–present, with data for 1979–present currently available from the Copernicus Climate Data Store, https://cds.climate.copernicus.eu, and the remainder to be released throughout 2020).

Today, even with the most up-to-date global climate models (GCMs), long climate simulations are computationally feasible only with horizontal grid spacing of ~30 km or coarser. Such resolutions are inadequate to simulate the detail and pattern of Ireland's climate that are required at regional and local levels; for instance, climate fields such as precipitation, wind speed and direction are strongly influenced by the local topography. Fortunately, the computational obstacles can be overcome through the application of numerical weather prediction (NWP) or regional climate models (RCMs) to achieve higher resolutions than extant global reanalysis datasets.

Regional reanalysis and dynamical downscaling are two methods often used to achieve a higher resolution (and overcome the associated computational cost). Both methods include forcing (with global reanalysis data) at the boundaries, but differ in other ways. Reanalyses simulate past weather/climate utilising a consistent NWP and data assimilation scheme. Assimilation ensures that historical observations are included so as to produce the best representation of climate at any given time. Dynamical downscaling makes use of an RCM, typically with nested domains and without data assimilation. The computational cost of running the RCM for a given (high) resolution is considerably less than that of a global model as the simulated RCM domain is considerably smaller. High-resolution RCMs demonstrate an improved ability to simulate precipitation (Kendon *et al*., 2012; Lucas-Picher *et al*., 2012) and improve the simulation of topography-influenced phenomena and extremes

with a relatively small spatial or short temporal character (Flato *et al*., 2013). An additional advantage is that the physically based RCMs explicitly resolve mesoscale atmospheric features and may provide a better representation of convective (Rauscher *et al*., 2010) and extreme (Kanada *et al*., 2008) precipitation. Other examples of the added value of RCMs include improved simulation of strong mesoscale cyclones (Cavicchia and von Storch, 2011), North Atlantic tropical cyclone tracks (Daloz *et al*., 2015) and near-surface wind speeds (Kanamaru and Kanamitsu, 2007), particularly in coastal areas with complex topography (Winterfeldt *et al*., 2011). 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 … and for mesoscale phenomena and extremes" (Flato *et al*., 2013).

In recent years, numerous regional reanalyses and downscaled products for Europe have been published; two examples of the former are the High Resolution Limited Area Model (HIRLAM; 1979–2014; 22 km; Dahlgreen *et al*., 2016) and COSMO-REA6 (1997–2004; 6 km; Bollmeyer *et al*., 2015), while two examples of the latter are described by Lucas-Picher *et al.* (2012) and Dasari and Challa (2015). Although there are numerous high-resolution regional reanalysis datasets available, until recently (2017) none has covered Ireland at spatial resolutions higher than 6 km.

Prior to the current research, two high-resolution Irish climate datasets that cover the period 1981–2016 were produced by researchers at the Irish Centre for High-End Computing (ICHEC). The datasets were created by downscaling ERA-Interim data using the Weather Research and Forecasting (WRF) model v3.7.1 (Skamarock *et al.*, 2008) and COSMO-CLM5 (Rockel *et al.*, 2008). These RCMs were run at 2 km and 1.5 km spatial resolution, respectively, with two additional 6 km and 18 km simulations run for both models. The data output from each model was archived at 1-hour intervals.

In addition, in 2017 the Irish meteorological service, Met Éireann, completed a 36-year reanalysis (MÉRA) at 2.5 km resolution for the period 1981–2016 (Gleeson *et al.*, 2017). Although the MÉRA resolution is coarser than those of the two ICHEC simulations, it has the advantage of data assimilation that utilised time series of surface observations (Whelan *et al*., 2016). As such, the MÉRA dataset is expected to show better skill than the ICHEC datasets for assimilated surface variables (e.g. temperature, pressure, 10 m wind speeds). The MÉRA datasets, which are stored as a series of 3-hour and 33-hour forecasts, are archived by Met Éireann at 1-hour intervals.

The overall aim of this research project was the production and dissemination of validated, long-term, high-resolution (spatial and temporal) gridded datasets of climate variables and derived products that are of use to researchers, planners and policymakers from various diverse sectors (e.g. climate science, renewable energy and agriculture).

Descriptions of the ICHEC model set-ups are provided in Chapter 2 (the MÉRA set-up has been described elsewhere, e.g. Gleeson *et al*., 2017), as well as a description of the model outputs that are available. In Chapter 3, descriptions of the data validations conducted for several commonly used climate parameters (e.g. precipitation, 2 m temperature, sea level pressure, relative humidity, 10 m winds) and other more "exotic" climate parameters [e.g. convective available potential energy (CAPE) 3 km] are provided. This is followed in Chapter 4 by a description of validated wind and solar radiation fields, for use in renewable energy applications. In Chapter 5, we give a detailed description of 36 (19 temperature and 17 precipitation) gridded climate change indices that were produced for each of the models (temperature: COSMO-CLM and MÉRA only). In Chapter 6, we describe a selection of products (24 in total) that are of interest to the agricultural sector and that are based on the outputs of MÉRA. In Chapter 7, we give details of the data and their accessibility. Finally, recommendations for future work are given in Chapter 8.

2 Methods and Models

Both the ICHEC WRF and COSMO-CLM RCM simulations were performed utilising nested domains with 18 km, 6 km and 2 km (WRF) or 18 km, 6 km and 1.5 km (COSMO-CLM) resolutions. Figure 2.1 illustrates the spatial coverage and topography of the three COSMO-CLM and WRF domains. The 18 km simulations were driven at the boundaries by ERA-Interim reanalysis data, produced by ECMWF at 80 km resolution, with all outputs from each individual nested domain (for both COSMO-CLM and WRF) archived at hourly intervals.

The WRF model used (v3.7.1) provides topography data at four resolutions (approximately 11, 6, 2 and 0.6 km at Irish latitudes) that can be used to construct terrain data for the model grid. Given that some climate variables (e.g. winds) are affected by nearby

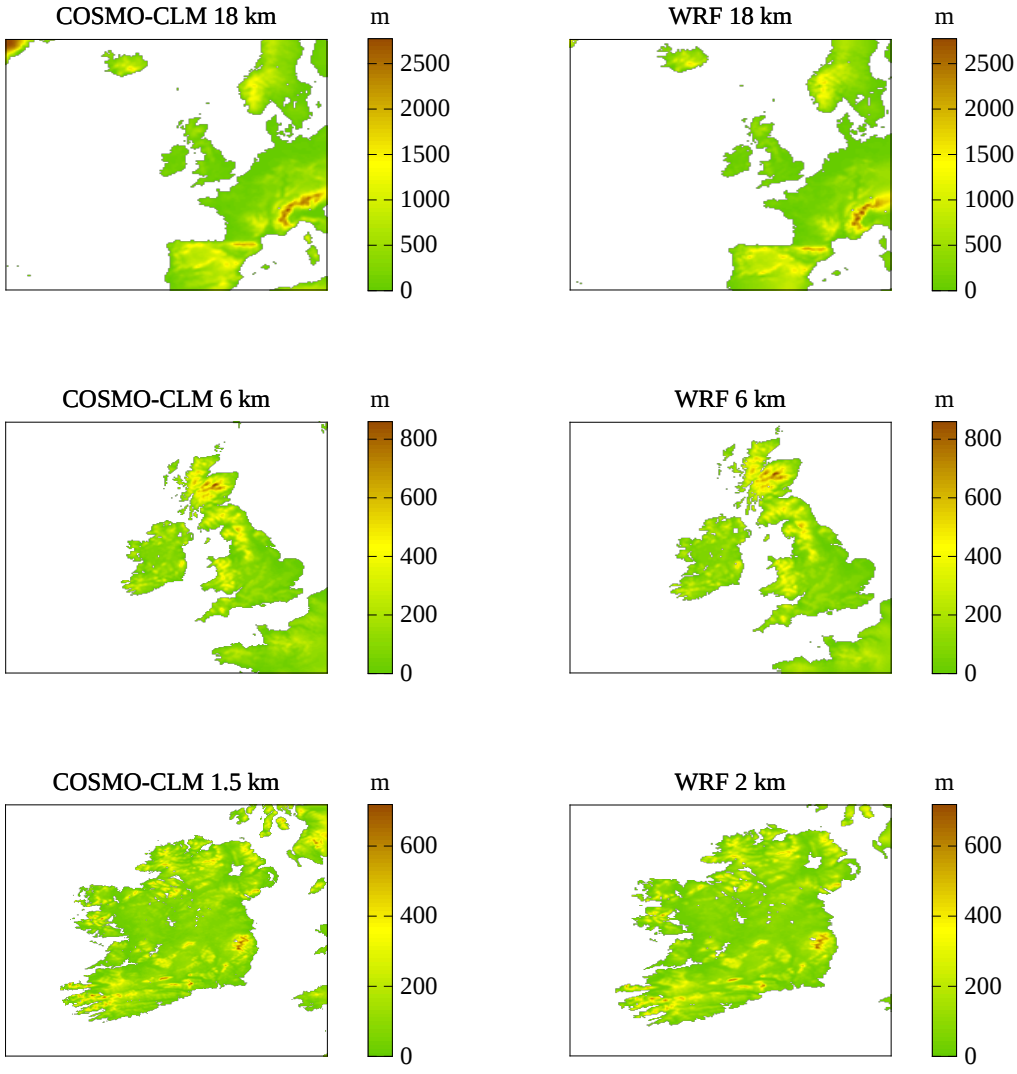


Figure 2.1. The nested domains used for the COSMO-CLM (left column) and WRF (right column) model runs showing model topography at three spatial resolutions: 18 km (top row), 6 km (middle row) and 1.5/2 km (bottom row). Source: Flanagan *et al.* (2019). This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/

topography, it was realised that underlying data with much finer resolution were required. Therefore, a 1 arc second (approximately 20 m) topography dataset [the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM)] was obtained and incorporated into the WRF simulations using the WRF Preprocessing System (WPS). The COSMO-CLM model already includes the high-resolution ASTER topography dataset as part of the preprocessing stage (ExtPAR), while MÉRA utilises Met Éireann's operational HARMONIE domain (Whelan *et al*., 2016; Gleeson *et al*., 2017).

Both COSMO-CLM and WRF have numerous parameter schemes that can potentially affect outputs. For instance, it is known that the choice of (WRF) subgrid orographic, flow blocking and gravity wave drag schemes can influence bias in 10 m wind speeds, 2 m temperature and surface pressure (Koo *et al*., 2018). To ensure that the most accurate options were employed, the results from several 1-month validation simulations previously performed at ICHEC (Nolan *et al*., 2017) were utilised. Summaries of the individual model settings, where different from the default option (or between different resolutions), are given in Tables 2.1 and 2.2.

Table 2.1. Namelist options used for each of the three COSMO-CLM simulations

COSMO-CLM option	Namelist	1.5 km	6 km	18 km
Time step	dt	12	50	120
Number of soil levels	ke_soil	8	8	8
Number of vertical levels	ke_tot	40	40	40
Radiation scheme call interval (hours)	hincrad	0.125	0.25	0.33
Convection scheme call interval (number of time steps)	ninconv	4	2	1
Moist convection scheme	itype_conv	3	0	0
Microphysics parameterisation scheme	itype_gscp	4	4	3
Subgrid scale orography	lsso	False	True	True
Aerosol option	itype_aerosol	2	2	2
Solar surface albedo	itype_albedo	2	2	2

Table 2.2. Namelist options used for each of the three WRF simulations

WRF option	Namelist	Physics scheme	2 km	6 km	18 km
Adaptive time step	max_time_step	N/A	19	57	171
Number of soil levels	num_soil_layers	N/A	4	4	4
Number of vertical levels	e_vert	N/A	50	50	50
Microphysics	mp_physics	WSM6	6	6	6
PBL scheme	bl_pbl_physics	YSU	1	1	1
Convective scheme	cu_physics	Kain–Fritsch	0	1	1
Short-wave radiation	ra_sw_physics	RRTMG	4	4	4
Long-wave radiation	ra_lw_physics	RRTMG	4	4	4
Land surface model	sf_surface_physics	Noah	2	2	2

N/A, not applicable; PBL, planetary boundary layer; RRTMG, rapid radiative transfer model for global climate models; WSM6, WRF single moment microphysics class 6; YSU, Yonsei University.

The models were each run for the period 1981–2016. The improvements in resolution gained are illustrated in Figure 2.2, where mean temperature data over Ireland for the period 1981–2014 from the original ERA-Interim dataset are shown alongside the COSMO-CLM 1.5 km downscaled data.

All data from the two ICHEC simulations and MÉRA were archived at 1-hour intervals (where appropriate). A full list of the datasets output by the three models is given in Tables 2.3–2.5 (COSMO-CLM), Tables 2.6–2.8 (WRF) and Tables 2.9–2.15 (MÉRA).

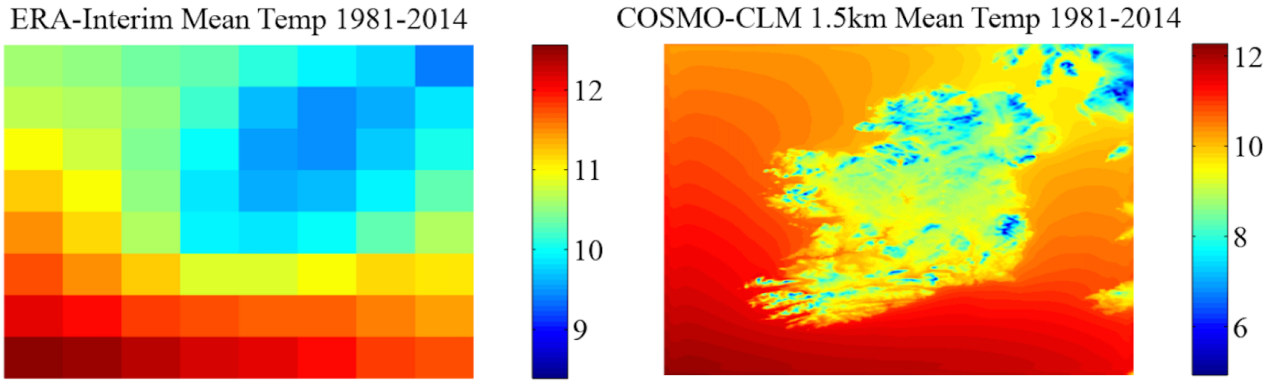


Figure 2.2. A comparison of ERA-Interim 80 km resolution data (left) with downscaled COSMO-CLM 1.5 km data (right). The mean 2 m temperature for the period 1981–2014 is shown.

Table 2.3. COSMO surface/near-surface parameters archived by ICHEC at 1-hour intervals

Variable	Unit
Precipitation rate	$kg\,m^{-2}\,s^{-1}$
Large-scale rainfall, convective rainfall, large-scale snowfall, convective snowfall, large-scale graupel, total precipitation amount, surface run-off, surface evaporation, subsurface run-off, vertical integrated water vapour, vertical integrated cloud ice, vertical integrated cloud water	$kg\,m^{-2}$
Total cloud cover, low cloud cover, medium cloud cover, high cloud cover, surface albedo	0–1
Surface temperature, 2 m temperature, 2 m dew point temperature, snow surface temperature	K
Surface pressure, mean sea level pressure	Pa
U (eastward) and *V* (northward) components of 10 m wind, maximum 10 m wind speed	$m\,s^{-1}$
Surface net downward short-wave (SW) radiation, average surface net downward SW radiation, direct surface downward SW radiation, averaged direct surface downward SW radiation, averaged surface diffuse downward SW radiation, averaged surface diffuse upward SW radiation, averaged downward long-wave (LW) radiation at the surface, averaged upward LW radiation at the surface, averaged surface net downward LW radiation, averaged surface photosynthetic active radiation, surface latent heat flux, surface sensible heat flux	$W\,m^{-2}$
Surface roughness length, thickness of snow, height of freezing level	m
Surface specific humidity, 2 m specific humidity	$kg\,kg^{-1}$
2 m relative humidity	%
CAPE 3 km	$J\,kg^{-1}$
Surface lifted index, Showalter index	None

Table 2.4. COSMO subsurface parameters at eight levels archived by ICHEC at 1-hour intervals

Variable	Unit
Soil temperature	K
Thickness of moisture content of soil layer	m

Table 2.5. COSMO upper-air parameters at 20, 40, …, 200 m archived by ICHEC at 1-hour intervals

Variable	Unit
U (eastward) and *V* (northward) components of wind	$m\,s^{-1}$
Air density	$kg\,m^{-3}$

Table 2.6. WRF surface or near-surface parameters archived by ICHEC at 1-hour intervals

Variable	Unit
Total precipitation, accumulated snowfall	mm
Total cloud fraction	0–1
Surface temperature	K
Surface pressure, sea level pressure	Pa
2 m temperature	°C
Time-varying roughness height, physical snow depth	m
Water vapour mixing ratio at 2 m	$kg\,kg^{-1}$
Relative humidity at 2 m	%
U (eastward) and *V* (northward) components of wind at 10 m, maximum 10 m wind speed at previous output time, friction velocity	$m\,s^{-1}$
Air density at lowest model level	$kg\,m^{-3}$
Short-wave (SW) downward surface flux, SW downward accumulated surface flux, bucket SW flux downward at surface accumulated, ground heat flux	$W\,m^{-2}$
Liquid path water, ice path water, water evaporation flux at surface	$kg\,m^{-2}$

Table 2.7. Hourly WRF subsurface parameters (5, 25, 75 and 150 cm below surface) archived by ICHEC

Variable	Unit
Soil temperature	K
Soil moisture	$m^3\,m^{-3}$

Table 2.8. WRF upper-air parameters at 40, 60, …, 120 m and archived by ICHEC at 1-hour intervals

Variable	Unit
U (eastward) and *V* (northward) components of wind	$m\,s^{-1}$

Table 2.9. Hourly MÉRA surface and near-surface parameters archived by Met Éireann

Variable	Unit
Surface temperature, 2 m temperature, maximum temperature, minimum temperature	K
Surface pressure, mean sea level pressure	Pa
2 m relative humidity	%
U (eastward) and *V* components (northward) of 10 m wind, gust *U* and *V* components	$m\,s^{-1}$
Total cloud cover, low, medium and high cloud cover, total cloud cover (fog)	0–1
Sensible heat flux, latent heat flux of evaporation, latent heat flux of sublimation, direct short-wave (SW) irradiance, net SW irradiance, long-wave (LW) irradiance, net LW irradiance, global irradiance, direct normal irradiance	$J\,m^{-2}$
Cloud base, cloud top, height of $T'_w = 0°$ isotherm, height of 0° isotherm, mixed layer depth, visibility	m
Momentum flux (*U* and *V* components)	$N\,m^{-2}$
Precipitation total, rain, graupel, snow, water evaporation, snow sublimation, snow depth, hail diagnostic	$kg\,m^{-2}$
Icing index, precipitation type	–
Lightning	m^{-3}

Table 2.10. MÉRA top-of-the-atmosphere parameters archived by Met Éireann at 1-hour intervals

Variable	Unit
Direct short-wave (SW) irradiance, net SW irradiance, net SW irradiance accumulated, net long-wave (LW) irradiance, net LW irradiance accumulated	$J\,m^{-2}$

Table 2.11. Hourly upper-air MÉRA parameters available at 12 heights (from 30 m to 400 m)

Variable	Unit
Temperature	K
U (eastward) and *V* (northward) components of wind	$m\,s^{-1}$
Relative humidity	%

Table 2.12. MÉRA upper-air parameters available at 100, 200, 300, 400, 500, 600, 700, 800, 850, 900, 925, 950 and 1000 hPa and archived by Met Éireann at 1-hour intervals

Variable	Unit
Geopotential	$m^2\,s^{-2}$
U (eastward) and *V* (northward) component of wind, vertical velocity	$m\,s^{-1}$
Relative humidity	%
Cloud ice, cloud water	$kg\,m^{-2}$
Temperature	K

Table 2.13. MÉRA pseudo-satellite parameters archived by Met Éireann at 1-hour intervals

Variable	Unit
Cloud-top temperature (IR), water vapour brightness temperature T_b (WV), T_b (WV) + cloud correction	K
Cloud water reflectivity	–

IR, infrared.

Table 2.14. Hourly vertically integrated MÉRA parameters archived by Met Éireann

Variable	Unit
Precipitable water, rain, snow, graupel, cloud ice, cloud water	$kg\,m^{-2}$

Table 2.15. Hourly and subsurface (root and deep) MÉRA parameters archived by Met Éireann

Variable	Unit
Soil temperature (surface and deep)	K
Soil moisture content (surface, root and deep)	$kg\,m^{-3}$
Surface soil ice (surface, root and deep)	$m^3\,m^{-3}$

3 Data Validations

The three new datasets discussed here contain many climate variables for which observations have previously been unavailable either at such temporal/spatial coverage and/or resolution or, indeed, at all. Given the datasets' novelty and potential, it is important that some indicator of quality be attached. To this end, uncertainty estimates [bias, absolute error, standard deviation (STD) and root mean square error (RMSE)] and several skill scores (where appropriate) were calculated for average annual and daily precipitation and 2 m temperature utilising gridded datasets of observations made available by Met Éireann and the UK Met Office. The results of these analyses are presented in section 3.1. Uncertainty estimates for hourly precipitation, 2 m temperature, 10 m winds, relative humidity and mean sea level pressure were calculated utilising station observations and are presented in section 3.2.

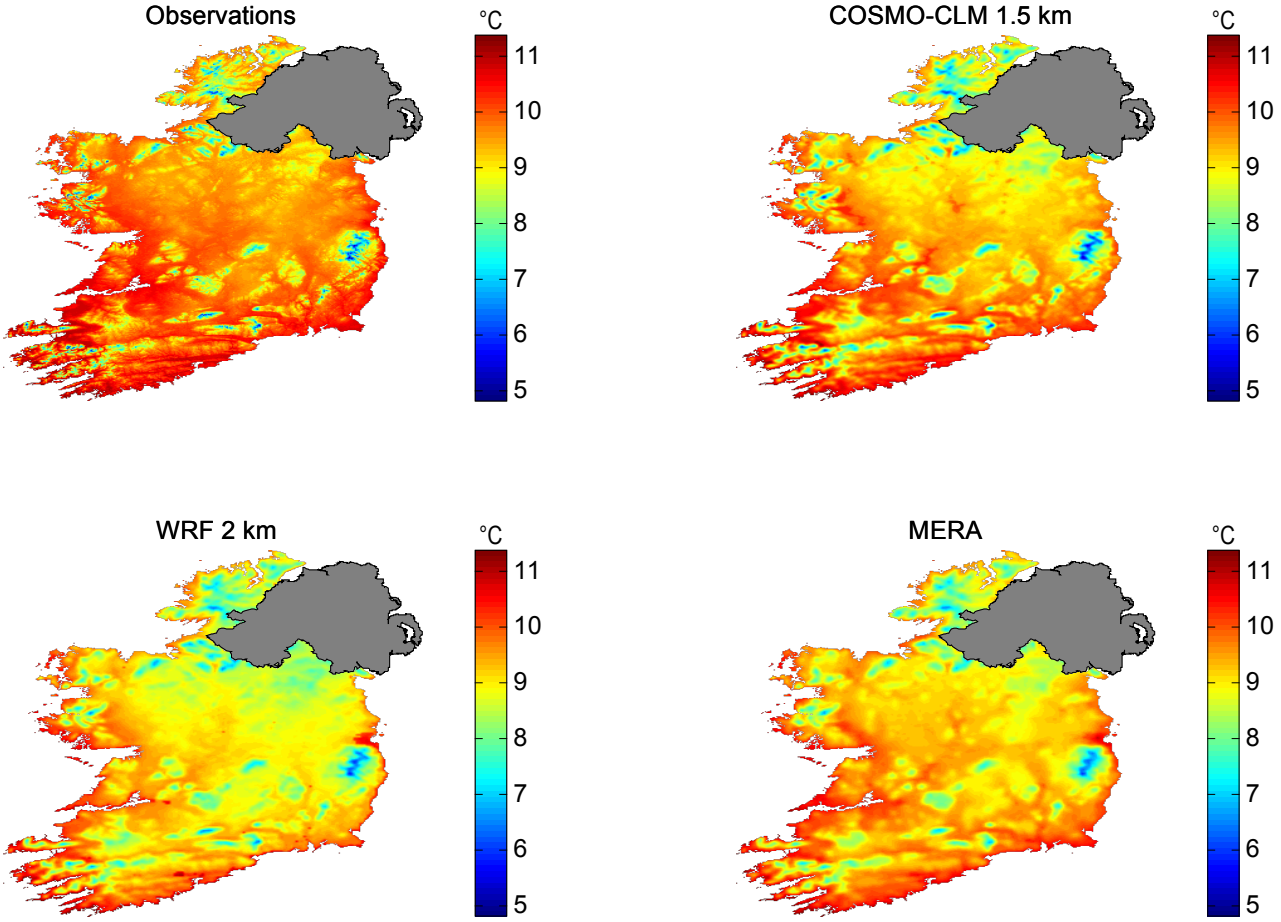
3.1 Annual and Daily Precipitation and 2 m Temperature

Gridded datasets of observed daily temperature and accumulated precipitation at 1 km resolution covering Ireland (Walsh, 2012, 2016, 2017) for the period 1981–2015 were obtained from Met Éireann. UK Met Office accumulated precipitation datasets with equal resolutions and time ranges were acquired for Northern Ireland from the Centre for Ecology and Hydrology (Tanguy *et al*., 2016). Since they are based on observations, the 1 km daily precipitation fields are the most authoritative source of information at daily (and longer) timescales and provide a strong benchmark against which the model data can be measured. However, they do not provide sub-daily information, which can come only from a limited number of stations or from models. The gridded datasets are available in monthly comma-separated values (CSV) files and require a level of processing before they can be used in any later analyses: the precipitation files contain some erroneous negative values that must be masked; easting and northing co-ordinates must be transformed to longitude–latitude pairs; and the gridded datasets are at 1 km resolution, whereas the COSMO, WRF and MÉRA datasets are at 1.5, 2 and 2.5 km, respectively. This last point is relatively straightforward to deal with for temperature data – (model grid) adjusted values can be easily obtained through interpolation (we have chosen bilinear interpolation for ease of implementation). However, the calculation of precipitation values requires a degree of care, as there are differences in how the observed values and the model values have been calculated. The observed values are recorded at a point location, whereas the model values represent accumulations averaged over a grid cell. To deal with this difference, a routine has been developed that effectively overlays the observed grid with the model grid. For each cell on the model grid, an average precipitation amount is calculated from those observed values that fall within the cell. To prevent the possibility of double-counting, any observed value that is used in a calculation is then marked as "missing". This routine has been applied to the gridded observations for each of the three model grids.

Daily and (subsequently) annual records of 2 m temperature and precipitation were built from each of the three models. For COSMO and WRF, each daily record is simply a sum of hourly values over the given time range 00:00–00:00 UTC (Coordinated Universal Time) for temperature and 09:00–09:00 UTC for precipitation – the latter to match climate station observational practice. For MÉRA, daily records of precipitation were built utilising 33-hour forecast files (thereby avoiding any negative impact from errors linked to model spin-up through use of the 3-hour forecast files) and consecutive subtractions of the 09:00 UTC forecast from the 33-hour forecast for each day. For 2 m temperature, hourly values were first obtained from the 3-hour files, followed by daily averaging. Annual records are then easily obtained through summation (precipitation) and averaging (2 m temperature) and comparisons with gridded observations are made.

3.1.1 Uncertainty estimates

The comparison for average annual 2 m temperature is shown in Figure 3.1. All three models capture the spatial distribution seen in the observations – cooler



Average annual 2 m temperature (°C), 1981–2015

Figure 3.1. Average annual temperature (°C) for the period 1981–2015 from Met Éireann gridded observations (top left), COSMO-CLM (top right), WRF (bottom left) and MÉRA (bottom right). Source: Flanagan *et al.* (2019). This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/

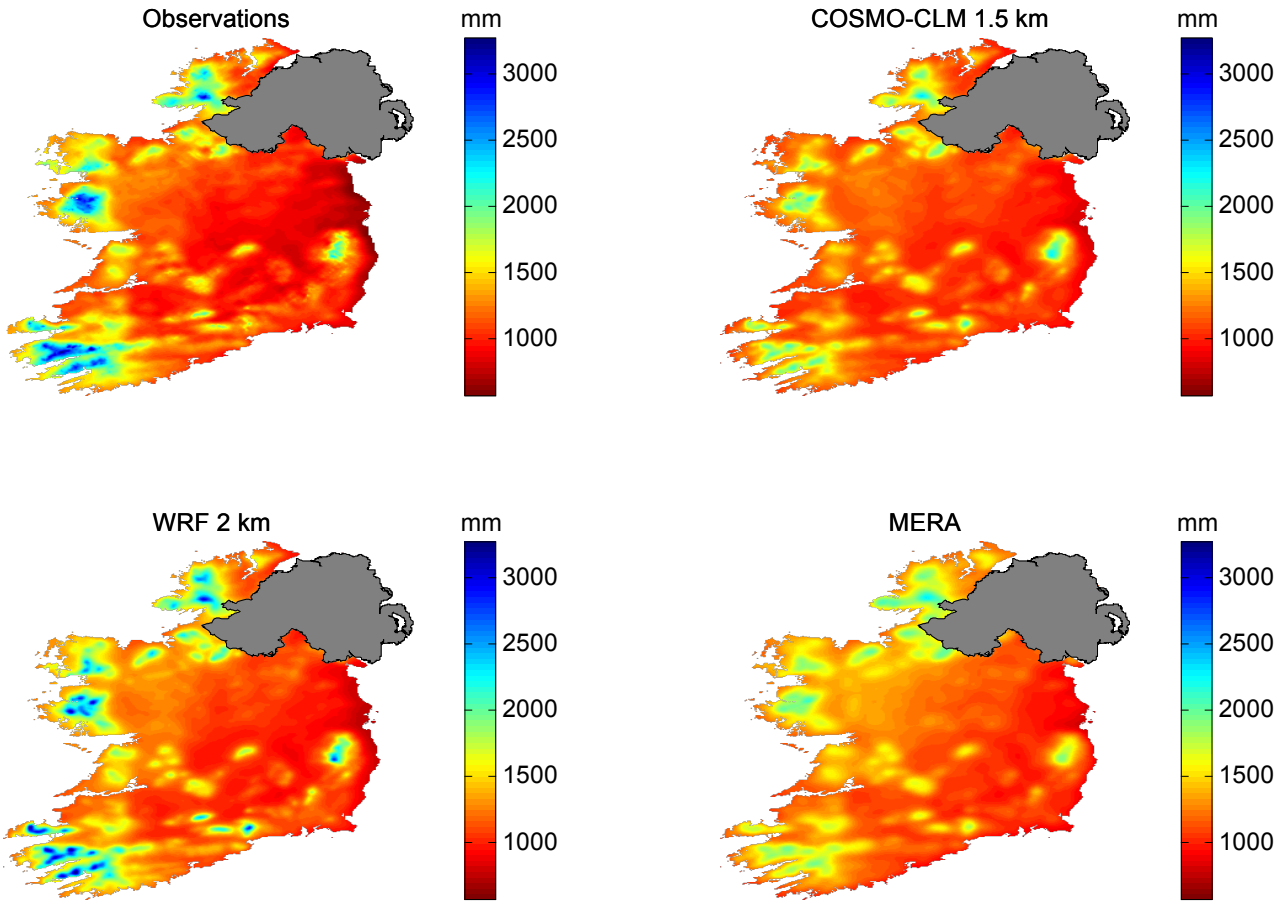
temperatures in the north and over the mountains – with WRF performing less well than either MÉRA or COSMO. This is clarified in Table 3.1, where WRF is found to have a much larger bias (–0.626°C) and mean absolute error (MAE; 0.641°C) than either MÉRA (bias = −0.382°C, MAE = 0.414°C) or COSMO (bias = −0.358°C, MAE = 0.372°C). All three models display similar variance – STDs are all in the range 0.2–0.26°C.

The comparison for average annual precipitation is shown in Figure 3.2. Again, all three models capture the spatial distribution seen in the observations – higher rainfall amounts in the west and over the mountains. In Table 3.2, WRF is shown to have the lowest MAE and STD (7.68% and 7.3%, respectively). Corresponding values for MÉRA and COSMO are 11.58% and 12.55%, respectively, and 10.02% and 12.41%, respectively.

A clearer picture emerges as the timescale for comparison is reduced from annual to daily. The results for daily 2 m temperature are given in Table 3.3. At this timescale, MÉRA has both a lower MAE (0.889°C) and a lower STD (1.042°C) than either COSMO (1.19°C, 1.494°C) or WRF (1.181°C, 1.327°C). The results for daily precipitation using both the Met Éireann (Ireland) and the UK Met Office (Northern Ireland) gridded observations are given in Table 3.4. For the Met Éireann dataset, MÉRA again

Table 3.1. Annual average 2 m temperature error found for the three models when utilising Irish gridded observation datasets provided by Met Éireann

Model	Bias (°C)	MAE (°C)	STD (°C)	RMSE (°C)
COSMO (1.5 km)	–0.36	0.37	0.20	0.41
WRF (2 km)	–0.63	0.64	0.26	0.68
MÉRA	–0.38	0.41	0.26	0.46



Average annual precipitation (mm), 1981–2015

Figure 3.2. Average annual rainfall (mm) for the period 1981–2015 from Met Éireann gridded observations (top left), COSMO-CLM (top right), WRF (bottom left) and MÉRA (bottom right). Source: Flanagan *et al.* (2019). This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. This figure has been modified.

Table 3.2. Annual average precipitation error found for the three models when utilising Irish gridded observation datasets provided by Met Éireann

Model	Bias (%)	MAE (%)	STD (%)	RMSE (%)
COSMO (1.5 km)	0.25	10.0	12.4	12.42
WRF (2 km)	5.98	7.68	7.30	9.44
MÉRA	5.20	11.58	12.55	13.83

Table 3.3. Mean daily 2 m temperature error found for the three models when utilising Irish gridded observation datasets provided by Met Éireann

Statistic	COSMO	WRF	MÉRA
Bias (°C)	–0.36	–0.63	–0.38
MAE (°C)	1.19	1.18	0.89
STD (°C)	1.49	1.33	1.04
RMSE (°C)	1.54	1.47	1.11

Table 3.4. Twenty-four-hour rainfall accumulation error found for the three models when utilising Irish (ROI) gridded observation datasets provided by Met Éireann and Northern Ireland (NI) gridded datasets sourced from the UK Met Office

Statistic	COSMO: ROI (NI)	WRF: ROI (NI)	MÉRA: ROI (NI)
Bias (mm)	0.18 (0.37)	0.24 (0.22)	0.09 (0.42)
MAE (mm)	4.07 (3.89)	3.24 (3.34)	2.60 (2.66)
STD (mm)	6.67 (6.33)	5.67 (5.66)	4.55 (4.49)
RMSE (mm)	6.67 (6.33)	5.68 (5.67)	4.55 (4.51)

displays the lowest MAE (2.6 mm) and STD (4.55 mm), followed by WRF (3.24 mm, 3.34 mm) and then COSMO (4.07 mm, 6.67 mm). A similar pattern is seen for the UK Met Office dataset comparison.

In addition to the analysis utilising gridded observations, a second analysis of daily precipitation amounts utilising data from 484 Met Éireann stations was performed. The MAE values obtained were 2.91 mm (COSMO), 2.63 mm (WRF) and 2.23 mm (MÉRA), with average RMSE values of 5.71 mm (COSMO), 5.19 mm (WRF) and 4.36 mm (MÉRA).

3.1.2 Skill scores

Numerous skill scores (Borsche *et al.*, 2015, Dahlgren *et al.*, 2016) based on the Met Éireann daily gridded temperature and Met Éireann (Ireland) and UK Met Office (Northern Ireland) precipitation datasets were calculated for each of the three models. These include (but are not limited to) accuracy (section 3.1.2.1), frequency bias (section 3.1.2.2), hit rate (section 3.1.2.3), false alarm rate (section 3.1.2.4), Hanssen–Kuiper skill score (KSS; section 3.1.2.5) and equitable threat score (section 3.1.2.6).

As a first step in this analysis, several representative thresholds were chosen. For precipitation, zeros were filtered out of the data and the thresholds chosen were 0.1, 1, 5, 10, 20, 30 and 50 mm. For temperature, the thresholds chosen were −5, 0, 5, 10,15 and 20°C. Appropriate contingency tables of the form shown in Table 3.5 were then calculated for each threshold to facilitate further analysis.

Figure 3.3 shows the 24-hour precipitation frequency distribution for the gridded observations and each of the three models. The values are given as percentages, as the differing model grid resolutions mean different total numbers of observations and tend to distort the visual comparison. Additionally, the observed values were taken directly from the original data files, after removal of negatives and transformation of co-ordinates but before grid spacing was dealt with. This was done purely for visualisation purposes. From Figure 3.3, it is clear that the three models capture the overall trend of observed rainfall quite well. A similar trend (not shown) was found for 2 m temperature.

3.1.2.1 Accuracy

The accuracy score was calculated using the formula:

$$\text{Accuracy} = \frac{a+d}{T}, \tag{3.1}$$

where a, d and T are as described in Table 3.5. This equation answers the simple question of what fraction of the model forecasts were correct. The results for each of the three models are shown in Table 3.6 (2 m temperature), Table 3.7 (Ireland 24-hour precipitation) and Table 3.8 (Northern Ireland 24-hour precipitation). For 2 m temperature, MÉRA has better scores than

Table 3.5. Layout of contingency tables used for threshold analysis

	Observed		
Modelled	**Yes**	**No**	**Total**
Yes	Hits (*a*)	False alarms (*b*)	Modelled yes
No	Misses (*c*)	Correct negatives (*d*)	Modelled no
Total	Observed yes	Observed no	Total (*T*)

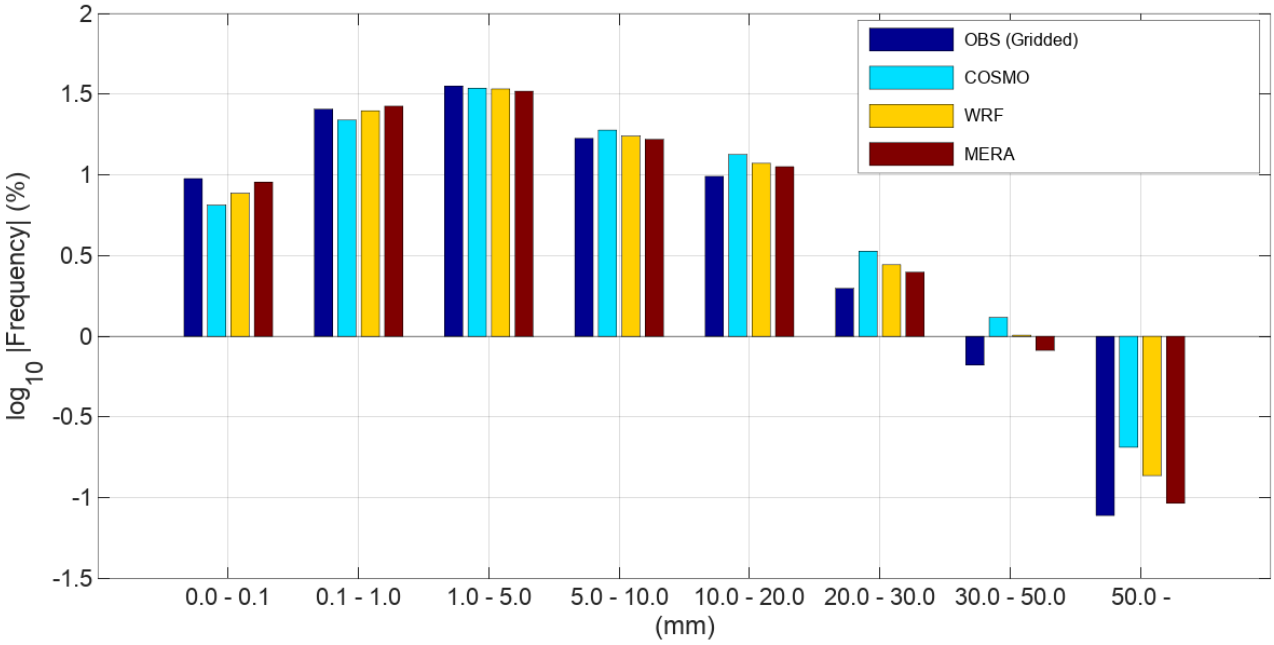


Figure 3.3. The 24-hour precipitation frequency distribution for gridded observations and each model.

Table 3.6. A selection of daily mean 2 m temperature skill scores found for each model for Ireland

		Threshold (°C)					
Skill score	**Model**	**–5**	**0**	**5**	**10**	**15**	**20**
Accuracy	COSMO	0.999	0.982	0.922	0.927	0.940	0.994
	WRF	0.999	0.986	0.918	0.923	0.943	0.997
	MÉRA	0.999	0.991	0.946	0.937	0.953	0.998
Frequency bias	COSMO	1	0.985	0.946	0.963	1.006	2.655
	WRF	1	0.990	0.938	0.924	0.786	0.836
	MÉRA	1	0.994	0.977	0.950	0.767	1.082
Hit rate	COSMO	1	0.984	0.926	0.904	0.762	0.736
	WRF	1	0.988	0.920	0.880	0.662	0.483
	MÉRA	1	0.993	0.956	0.908	0.692	0.659
False alarm rate	COSMO	0.682	0.149	0.099	0.053	0.034	0.006
	WRF	0.716	0.188	0.092	0.039	0.018	0.001
	MÉRA	0.417	0.150	0.107	0.037	0.011	0.001
KSS	COSMO	0.318	0.835	0.827	0.851	0.727	0.731
	WRF	0.284	0.799	0.828	0.840	0.644	0.482
	MÉRA	0.583	0.843	0.849	0.870	0.681	0.658
Equitable threat score	COSMO	0.156	0.350	0.592	0.743	0.569	0.250
	WRF	0.161	0.391	0.582	0.730	0.550	0.356
	MÉRA	0.375	0.519	0.681	0.774	0.609	0.462

Table 3.7. A selection of 24-hour rainfall accumulation skill scores found for the three models when utilising Irish gridded observation datasets provided by Met Éireann

		Threshold (mm)						
Skill score	**Model**	**0.1**	**1.0**	**5.0**	**10.0**	**20.0**	**30.0**	**50.0**
Accuracy	COSMO	0.91	0.74	0.73	0.82	0.94	0.98	0.997
	WRF	0.89	0.77	0.78	0.86	0.96	0.99	0.998
	MÉRA	0.89	0.80	0.83	0.89	0.96	0.99	0.998
Frequency bias	COSMO	0.98	0.93	0.99	1.11	1.35	1.53	1.98
	WRF	0.99	0.96	1.03	1.14	1.30	1.39	1.58
	MÉRA	0.98	0.95	1.02	1.11	1.18	1.14	1.11
Hit rate	COSMO	0.94	0.80	0.63	0.52	0.38	0.30	0.20
	WRF	0.94	0.82	0.68	0.57	0.42	0.33	0.21
	MÉRA	0.93	0.83	0.73	0.64	0.48	0.39	0.27
False alarm rate	COSMO	0.82	0.45	0.21	0.12	0.04	0.01	0.002
	WRF	0.73	0.34	0.17	0.09	0.03	0.01	0.001
	MÉRA	0.61	0.25	0.13	0.07	0.02	0.01	0.001
KSS	COSMO	0.12	0.34	0.42	0.41	0.35	0.29	0.19
	WRF	0.21	0.48	0.51	0.48	0.39	0.32	0.21
	MÉRA	0.32	0.58	0.60	0.57	0.46	0.38	0.27
Equitable threat score	COSMO	0.05	0.19	0.26	0.24	0.17	0.13	0.07
	WRF	0.11	0.30	0.34	0.29	0.21	0.16	0.09
	MÉRA	0.17	0.39	0.43	0.37	0.27	0.22	0.15

Table 3.8. A selection of 24-hour rainfall accumulation skill scores found for the three models when utilising Northern Ireland gridded datasets sourced from the UK Met Office

		Threshold (mm)						
Skill score	**Model**	**0.1**	**1.0**	**5.0**	**10.0**	**20.0**	**30.0**	**50.0**
Accuracy	COSMO	0.91	0.74	0.72	0.83	0.95	0.99	0.998
	WRF	0.90	0.75	0.76	0.86	0.96	0.99	0.999
	MÉRA	0.91	0.80	0.81	0.88	0.97	0.99	0.999
Frequency bias	COSMO	0.99	0.97	1.05	1.18	1.35	1.56	2.17
	WRF	0.99	0.97	1.03	1.13	1.23	1.31	1.72
	MÉRA	1.00	1.02	1.10	1.21	1.24	1.18	1.16
Hit rate	COSMO	0.95	0.81	0.63	0.50	0.35	0.36	0.15
	WRF	0.94	0.81	0.64	0.51	0.33	0.24	0.13
	MÉRA	0.95	0.86	0.74	0.63	0.46	0.35	0.15
False alarm rate	COSMO	0.84	0.50	0.23	0.12	0.03	0.01	0.001
	WRF	0.77	0.40	0.18	0.09	0.02	0.01	0.001
	MÉRA	0.69	0.35	0.16	0.08	0.02	0.005	0.001
KSS	COSMO	0.11	0.32	0.40	0.38	0.32	0.25	0.14
	WRF	0.17	0.41	0.46	0.41	0.31	0.24	0.13
	MÉRA	0.26	0.52	0.59	0.55	0.44	0.34	0.15
Equitable threat score	COSMO	0.05	0.18	0.25	0.22	0.16	0.11	0.05
	WRF	0.09	0.25	0.30	0.24	0.16	0.11	0.05
	MÉRA	0.15	0.35	0.40	0.34	0.25	0.19	0.07

either COSMO or WRF at all thresholds, with COSMO outperforming WRF at moderate thresholds (5°C and 10°C) only. For 24-hour precipitation, MÉRA again displays marginally better scores than WRF (at all thresholds, both Ireland and Northern Ireland), with WRF, in turn, outperforming COSMO.

3.1.2.2 Frequency bias

The frequency bias was calculated using the formula:

$$\text{Bias} = \frac{a+b}{a+c}, \tag{3.2}$$

where *a*, *b* and *c* are as described in Table 3.5. The frequency bias answers the question of how the model frequency of "yes" events compares with the observed frequency of "yes" events, with 1 being a perfect score. As shown in Table 3.6 (2 m temperature), all three models tend to underpredict at most thresholds (apart from MÉRA and COSMO at high thresholds), with MÉRA outperforming both WRF and COSMO at lower (0°C and 5°C) and higher (20°C) thresholds. At moderate thresholds (10°C and 15°C) COSMO is the better performer. In Table 3.7 (Ireland precipitation) and Table 3.8 (Northern Ireland precipitation), for thresholds below 5 mm the models perform similarly, tending to slightly underpredict. Above this threshold, each model overpredicts, with MÉRA outperforming WRF, which in turn outperforms COSMO.

3.1.2.3 Hit rate

The hit rate (HR) was calculated using the formula:

$$\text{HR} = \frac{a}{a+c}, \tag{3.3}$$

where *a* and *c* are as described in Table 3.5. The HR answers the question of what fraction of the observed "yes" events are correctly predicted by the models, with 1 being a perfect score. For 2 m temperature (Table 3.6) all three models display a decreasing ability with increasing threshold, with MÉRA being the best performer at lower thresholds (0°C, 5°C and 10°C) and COSMO outperforming both WRF and MÉRA at higher thresholds (15°C and 20°C). For 24-hour precipitation [Table 3.7 (Ireland) and Table 3.8 (Northern Ireland)], values again decrease at increasing thresholds, with MÉRA outperforming WRF [at all but the lowest (0.1 mm) threshold], which in turn outperforms COSMO.

3.1.2.4 False alarm rate

The false alarm rate (FAR) was calculated using the formula:

$$\text{FAR} = \frac{b}{b+d}, \tag{3.4}$$

where *b* and *d* are as described in Table 3.5. The FAR answers the question of what fraction of the observed "no" events are incorrectly predicted by the models as a "yes" event, with 0 being a perfect score. For 2 m temperature (Table 3.6) all three models display an increasing ability with increasing threshold, with MÉRA being the best performer at the lowest (–5°C) and higher (10°C, 15°C and 20°C) thresholds. For the remaining (0°C and 5°C) thresholds, COSMO is marginally best. For 24-hour precipitation [Table 3.7 (Ireland) and Table 3.8 (Northern Ireland)], values again decrease (i.e. improve) for increasing thresholds, with MÉRA consistently outperforming WRF, which, in turn, outperforms COSMO.

3.1.2.5 Hanssen–Kuiper skill score

The KSS (Dahlgreen *et al.*, 2016) is simply the difference between the HR (equation 3.3) and the FAR (equation 3.4). It is used to indicate how well each model separated the "yes" events from the "no" events, with 0 indicating no skill and 1 being a perfect score. For 2 m temperature (Table 3.6) all three models display the highest skill at moderate thresholds (0°C, 5°C and 10°C), with MÉRA being the best performer. At the lowest (–5°C) and higher (15°C and 20°C) thresholds, skill scores decrease for all three models, with MÉRA showing higher skill at the lowest threshold and COSMO showing higher skill otherwise. WRF does not display the highest skill at any threshold. For 24-hour precipitation (Tables 3.7 and 3.8), the picture is somewhat simpler – values are again highest at moderate thresholds, with MÉRA consistently outperforming WRF, which, in turn, outperforms COSMO.

3.1.2.6 Equitable threat score

The equitable threat score (ETS) (Gandin and Murphy, 1992) was calculated using equations 3.5 and 3.6:

$$\text{ETS} = \frac{a-a^*}{a+b+c-a^*}, \tag{3.5}$$

where

$$a^* = \frac{(a+c)(a+b)}{T} \quad (3.6)$$

and *a, b, c* and *T* are as described in Table 3.5. The ETS measures how well the modelled "yes" predictions correspond to the observed "yes" events when adjusted for hits due to random chance (*a**). Here, 0 indicates no skill and 1 indicates perfect skill. For 2 m temperature (Table 3.6) each model displays the highest skill at moderate thresholds (5°C, 10°C and 15°C) and decreased skill at lower (–5°C, 0°C) and the highest (20°C) thresholds. At all thresholds MÉRA is the best performer of the three models, with COSMO the second-best performer at moderate thresholds (WRF otherwise). For 24-hour precipitation (Tables 3.7 and 3.8) the trend is the same as that seen for the KSS – values are highest at moderate thresholds, with MÉRA consistently outperforming WRF, which, in turn, outperforms COSMO.

3.1.2.7 Fractions skill score

An early question that needed to be addressed was what, if any, added value an increase in spatial resolution brings. The fractions skill score (FSS) (Roberts and Lean, 2008) provides a measure of forecast skill against spatial scale for selected thresholds and is described as follows:

- **Step 1:** Convert to binary fields:
  - thresholds (q) are chosen (either characteristic values or percentiles);
  - $I_O = \begin{cases} 1, O_r \geq q \\ 0, O_r < q \end{cases}$ and $I_M = \begin{cases} 1, M_r \geq q \\ 0, M_r < q \end{cases}$ (3.7)
- **Step 2:** Generate fractions for neighbourhoods of length n:
  - $O(n)(i,j) = \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n} I_O$

$$\left[i+k-1-\frac{n-1}{2}, j+l-1-\frac{n-1}{2}\right] \quad (3.8)$$

  - $M(n)(i,j) = \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n} I_M$

$$\left[i+k-1-\frac{n-1}{2}, j+l-1-\frac{n-1}{2}\right] \quad (3.9)$$

- **Step 3:** Compute FSSs:
  - $MSE_{(n)} = \frac{1}{N_x N_y}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\left[O_{(n)i,j} - M_{(n)i,j}\right]^2$ (3.10)
  - $MSE_{(n)ref} = \frac{1}{N_x N_y}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y} O^2_{(n)i,j} + \sum_{i=1}^{N_x}\sum_{j=1}^{N_y} M^2_{(n)i,j}$ (3.11)
  - $FSS_{(n)} = 1 - \frac{MSE_{(n)}}{MSE_{(n)ref}}$. (3.12)

A script to implement the FSS method was written and run for several thresholds (50th, 75th and 95th percentiles) for each of COSMO 1.5 km, WRF 2 km and MÉRA 2.5 km for neighbourhoods of length 6 km and 18 km. Example results for the three high-resolution precipitation datasets (95th percentile, 18 km length scale) are given in Table 3.9 (columns 2–4), alongside results from the COSMO and WRF 18 km datasets (columns 5 and 6) only (there are no lower resolution MÉRA datasets). The patterns in Table 3.9 (for high-resolution datasets MÉRA marginally outperforms WRF and WRF outperforms COSMO, while for lower resolution datasets the skill scores are lower) were seen for all thresholds and length scales tested.

3.2 Hourly Estimates

Hourly synoptic station observations of 2 m temperature, precipitation, 10 m winds, relative humidity and mean sea level pressure were obtained from Met Éireann. There are 25 stations in total,

Table 3.9. Mean FSSs and STDs for 24-hour precipitation for each of the three high-resolution models at length scales of 18 km (columns 2–4) and for the COSMO and WRF 18 km datasets (columns 5 and 6)

	18 km length scale from:				
Statistic	**COSMO 1.5 km**	**WRF 2 km**	**MÉRA 2.5 km**	**COSMO 18 km**	**WRF 18 km**
Mean (FSS)	0.335	0.404	0.408	0.229	0.239
STD (FSS)	0.285	0.279	0.269	0.249	0.229

The threshold used in the analysis is the 95th percentile.

with varying record lengths available. Each of the three model datasets was preprocessed so that a comparison with these observations could be made.

3.2.1 2 m temperature (°C)

A summary of the results from the analysis of 2 m temperature is given in Figure 3.4. The (mean) error values are −0.11°C (COSMO), −0.24°C (WRF) and -0.05°C (MÉRA), while the STDs are 1.93°C (COSMO), 1.85°C (WRF) and 1.31°C (MÉRA).

3.2.2 Precipitation (mm)

For hourly precipitation amounts, COSMO-CLM and WRF show remarkably similar errors (Figure 3.5) and STDs (Figure 3.6) – overall error values are less than 0.01 mm and overall STDs are 0.63 mm for both models. Additionally, the MAEs are 0.18 mm for both models. By comparison, the analysis performed for hourly MÉRA precipitation gave the following values: < 0.01 mm (error), 0.55 mm (STD) and 0.16 mm (MAE).

3.2.3 10 m winds (m s^{-1})

An analysis of 10 m wind speed and direction was performed utilising hourly station data obtained from Met Éireann. In total, there are 23 stations where such data are recorded. The data are provided in CSV format and a latitude–longitude location is provided with each file. A degree of caution is required when handling these datasets – unfortunately, they are not always continuous (i.e. missing data) or of equal duration, which can cause erroneous results if not accounted for. Octave scripts were developed to correctly process these datasets and produce hourly time series at each station location that can then be compared with the three model outputs. The three model datasets contain hourly 10 m U (eastward) and V (northward) wind components for the period 1981–2015. A script that utilises Climate Data Operators (CDO) commands has been developed to extract 10 m wind speeds and direction based on the following simple formulae: speed = sqrt ($U^2 + V^2$) and direction = 180 + (180/pi)*atan^2 (U, V). This CDO-based script first interpolates both (U, V) wind components to

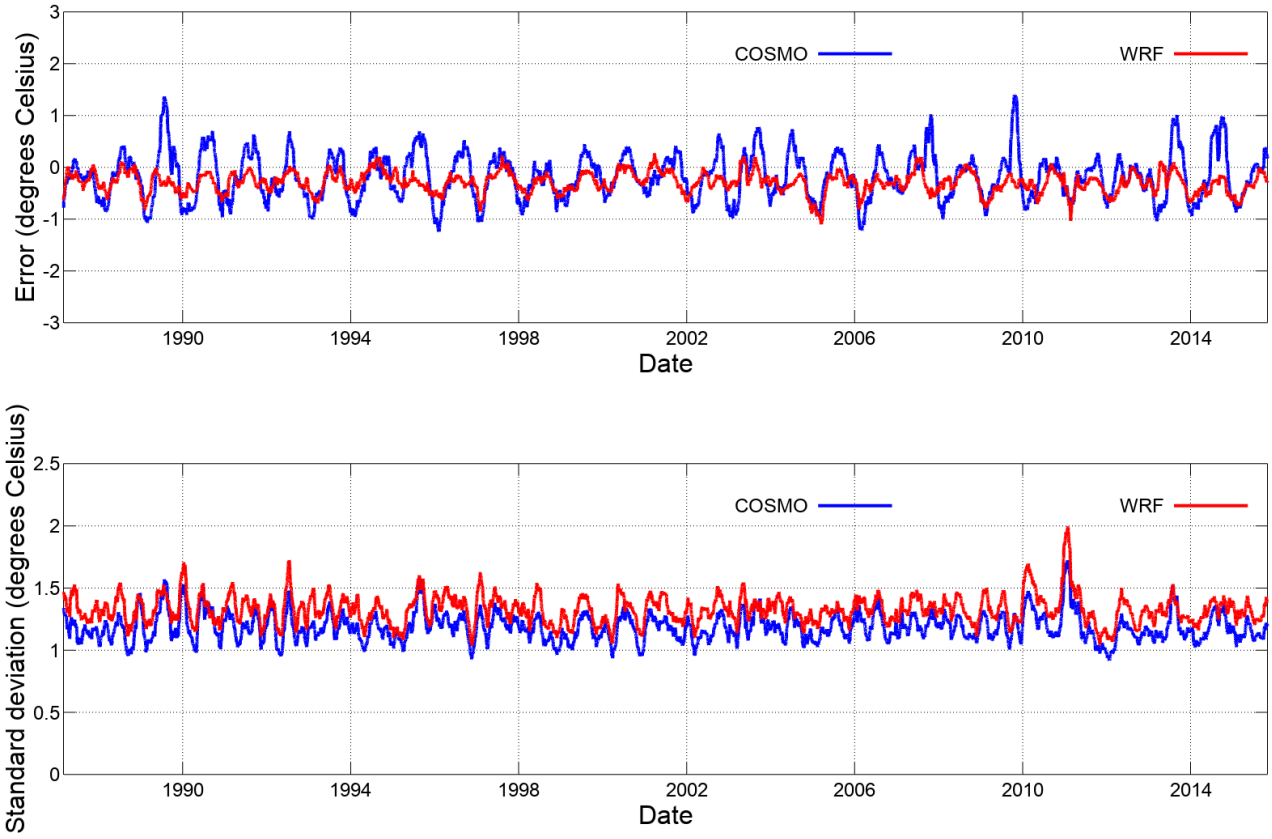


Figure 3.4. Sixty-day running average of the hourly (mean) temperature errors (top panel) and STDs (bottom panel) for COSMO-CLM and WRF, from comparisons with available station observations.

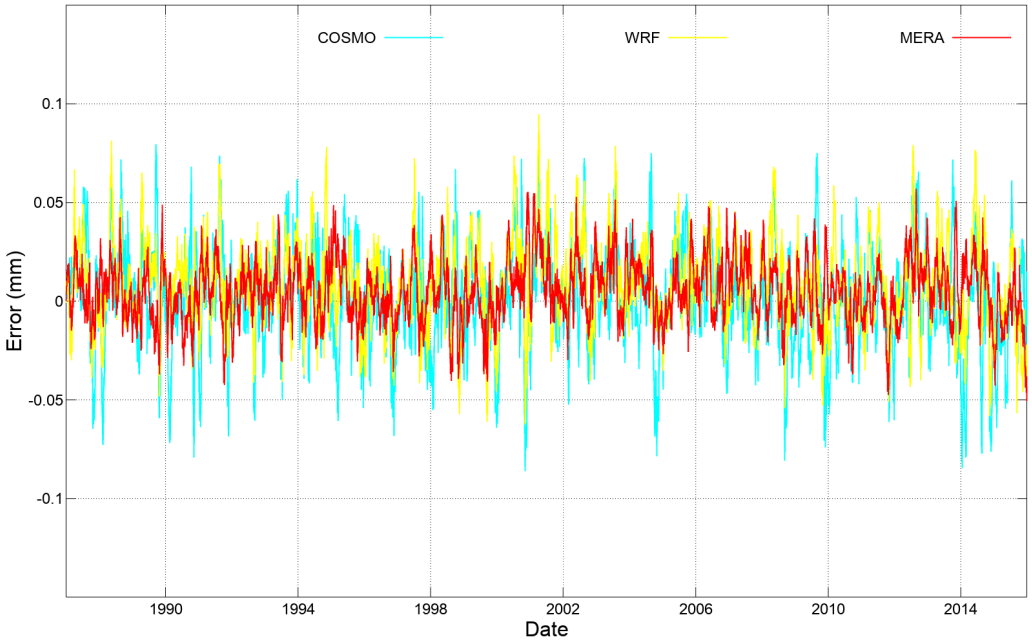


Figure 3.5. Thirty-day running average of the hourly rainfall (mean) error for the three models and available station observations.

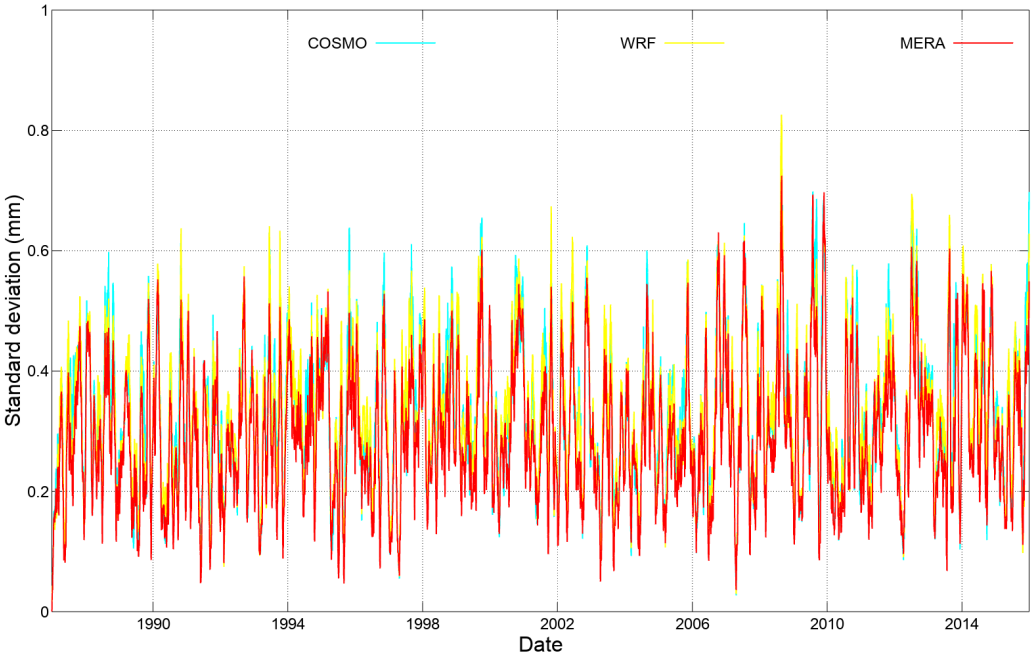


Figure 3.6. Thirty-day running average of the hourly rainfall (mean) STD for the three models and available station observations.

each station location for comparison with observations. Another Octave script was developed to correctly match each fixed-length, continuous model time series with the relevant variable-length, non-continuous observed time series. This script also produces a wind rose at each location for the observations and each model. An example from Casement Aerodrome, Dublin, for the period from1 January 1987 to 31 December 2015 is shown in Figure 3.7, in which each of COSMO, WRF and MÉRA have captured the fundamental characteristics of the observed wind profile. The overall bias, MAE, STD and RMSE were calculated at each location. The wind speed results are given in Table 3.10.

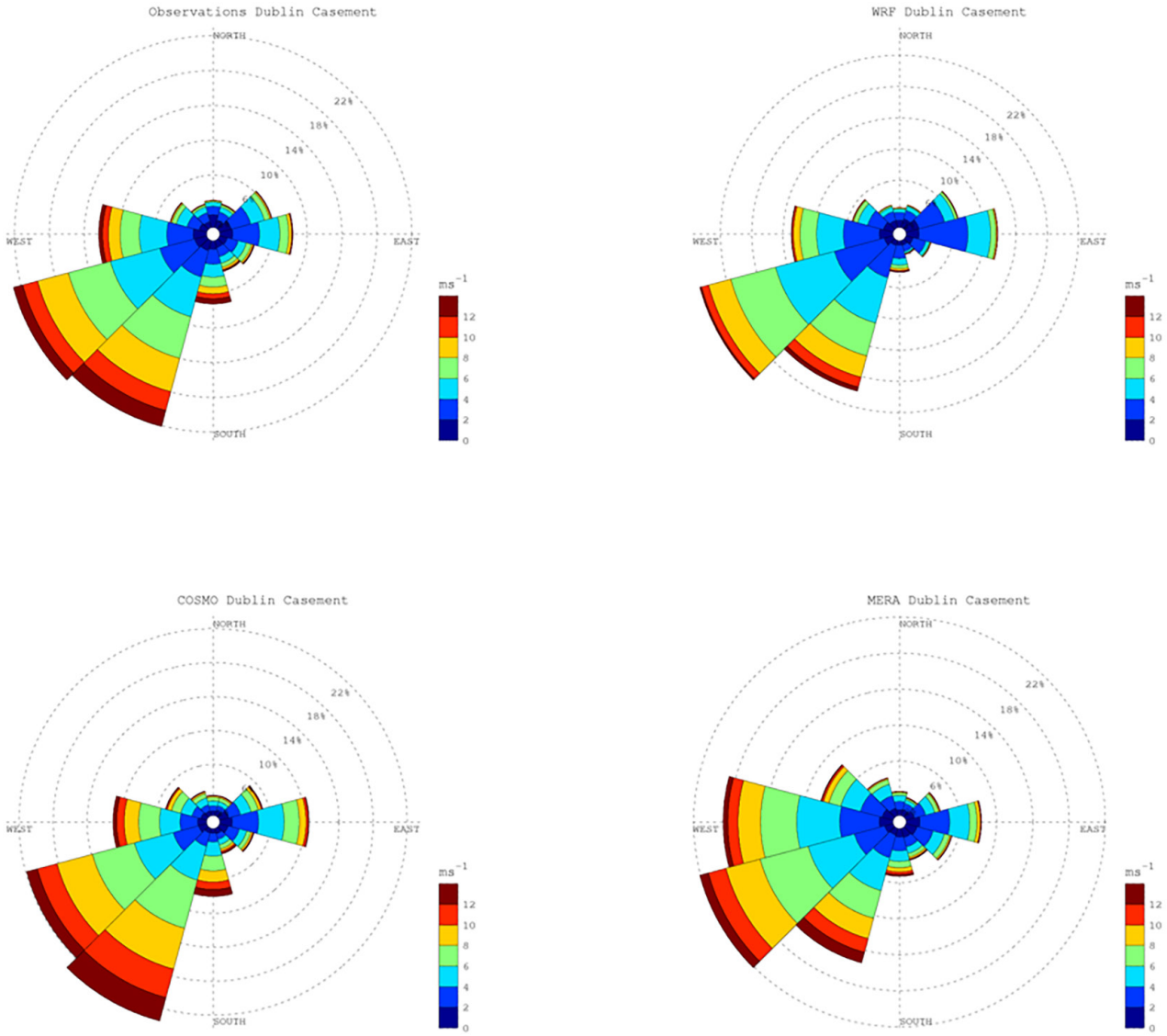


Figure 3.7. 10 m wind roses generated for the Casement Aerodrome station in Dublin over the period from 1 January 1987 to 31 December 2015. Clockwise from top left: observations-, WRF-, MÉRA- and COSMO-based wind roses.

3.2.4 Relative humidity (%)

Relative humidity bias, MAE, percentage error, absolute percentage error, correlation and associated STDs were calculated at each station location for the three models. The overall results of this analysis, in which all station data were amalgamated to form one single dataset, are given in Table 3.11.

3.2.5 Sea level pressure (hPa)

Hourly sea level pressure bias, MAE, percentage error, absolute percentage error, correlation and associated STDs were calculated at each station location for each model. The results of this analysis, in which all station data were amalgamated to form one single dataset, are given in Table 3.12.

Table 3.10. Hourly 10 m wind speed MAEs ($m\,s^{-1}$) and STDs ($m\,s^{-1}$) for COSMO (C), WRF (W) and MÉRA (M) determined using observations from 23 Met Éireann weather stations

Station	MAE (W)	STD (W)	MAE (C)	STD (C)	MAE (M)	STD (M)
Oak Park, Carlow	1.30	1.71	1.77	1.95	0.99	1.20
Ballyhaise, Cavan	1.35	1.38	2.28	1.79	1.82	1.40
Shannon Airport	2.20	2.38	2.13	2.27	1.17	1.40
Cork Airport	1.43	1.83	1.72	2.04	1.04	1.36
Moore Park, Cork	1.21	1.46	2.08	1.92	1.66	1.45
Roches Point, Cork	2.10	2.50	2.01	2.59	1.30	1.70
Sherkin Island, Cork	2.15	2.28	1.97	2.42	1.38	1.53
Finner, Donegal	1.67	2.04	1.81	2.27	1.15	1.47
Malin Head, Donegal	2.59	2.69	2.17	2.74	1.72	1.83
Casement, Dublin	1.65	2.12	1.91	2.34	1.06	1.41
Dublin Airport	1.76	2.09	1.64	2.13	1.07	1.30
Athenry, Galway	1.28	1.53	1.78	1.76	1.61	1.47
Mace Head, Galway	1.83	2.39	1.94	2.61	1.29	1.70
Valentia, Kerry	1.59	2.01	2.04	2.22	1.11	1.42
Belmullet, Mayo	2.04	2.40	1.90	2.43	1.29	1.55
Claremorris, Mayo	1.28	1.69	1.33	1.76	1.26	1.39
Knock Airport	1.26	1.67	1.57	1.92	0.97	1.17
Newport, Mayo	1.81	2.30	1.79	2.34	1.45	1.88
Dunsany, Meath	1.18	1.48	1.96	1.81	1.25	1.20
Mt Dillon, Roscommon	1.18	1.46	2.04	1.85	0.89	1.10
Gurteen, Tipperary	1.15	1.52	1.52	1.74	1.06	1.15
Mullingar, Westmeath	1.33	1.51	1.99	1.77	1.52	1.26
Johnstown, Wexford	1.55	1.68	2.49	2.06	2.03	1.57

Table 3.11. Overall relative humidity hourly uncertainty estimates for each of the three models determined through comparison with data from 25 Met Éireann synoptic stations

Statistics	COSMO-CLM	WRF	MÉRA
Bias (STD)	–2.86 (5.85)	0.95 (6.61)	–0.42 (3.59)
MAE (STD)	7.58 (3.53)	7.67 (3.93)	5.47 (2.19)
% error (STD)	–2.75 (7.64)	2.16 (8.92)	–0.04 (4.72)
Absolute % error (STD)	9.39 (4.94)	9.83 (6.08)	6.89 (3.33)
Correlation	0.66	0.57	0.80

Table 3.12. Overall sea level hourly uncertainty estimates for each of the three models determined through comparison with data from 25 Met Éireann synoptic stations

Statistics	COSMO-CLM	WRF	MÉRA
Bias (STD)	–0.87 (2.56)	–0.20 (2.39)	0.03 (0.51)
MAE (STD)	1.96 (1.87)	1.69 (1.70)	0.37 (0.36)
% error (STD)	–0.09 (0.25)	–0.02 (0.24)	< 0.01 (0.05)
Absolute % error (STD)	0.19 (0.19)	0.17 (0.17)	0.04 (0.04)
Correlation	0.98	0.98	1.0

4 Renewable Energy Fields

4.1 Upper-air Winds

There is great interest in long-term time series of wind speed and direction, particularly at the 80 m level, the standard turbine height. Numerous attempts to obtain observational data were made, with varying degrees of success. Wind speed and direction data were obtained from 10 and 7 Midlands turbines, respectively (section 4.1.1), and radiosonde data were obtained from two locations (section 4.1.2).

4.1.1 Turbine data

The observational data files required a degree of processing before they could be used for model validation: the values are comma separated but with many different data elements and durations; there are various conventions for handling errors and missing values; the data are 10-minute resolution averages, maximums and minimums for wind speed (10 series) and vector averages for wind direction (seven series); and the data are from numerous (non-modelled) heights – 48, 68, 69, 70, 80 and 80.5 m. Octave scripts were written to process the observational data into 1-hour resolution time series, with missing/error values being flagged for future masking. Additional scripts were written that generate time series of appropriate duration from the three models. The model *U* and *V* components, recorded at various heights (e.g. 20, 40, 60, 80 m, …), were first interpolated to turbine heights using a simple log power law:

$$(U_{new}, V_{new}) = (U_{old}, V_{old}) * \log_e[h_{new}/R]/\log_e[h_{old}/R] \quad (4.1)$$

where *h* is height, *R* is the roughness length and the subscripts old/new denote known/interpolated values. Once the new velocity components are found, they are interpolated to the latitude/longitude of the turbine and wind speeds and directions are then calculated. The results of the subsequent turbine height wind speed and direction validations are given in Tables 4.1 and 4.2, respectively. At each location, MÉRA consistently showed the lowest mean MAE and STD, with WRF outperforming COSMO.

4.1.2 Radiosonde data

Radiosonde data were acquired from two locations: Castor Bay, County Down, and Valentia, County Kerry. The data files contain numerous upper-air measurements, including wind speed and direction, from ground level up to approximately 25 km and were processed for comparison with the three models. An example of wind speed data processed for comparison with COSMO-CLM at Castor Bay is shown in Figures 4.1 and 4.2. The uncertainty estimates found for wind speed are given in Tables 4.3 and 4.4, while those for wind direction are given in Tables 4.5 and 4.6. The results of other, “more exotic” parameter

Table 4.1. MÉRA (M), WRF (W) and COSMO (C) wind speed (m s^{-1}) error analysis for 10 turbines at various heights

	MAE			STD			Bias		
Height (m)	M	W	C	M	W	C	M	W	C
70	1.18	1.65	2.27	1.44	2.14	2.23	–0.50	–0.37	–1.68
80.5	1.21	1.75	1.99	1.46	2.29	2.35	–0.53	–0.05	–0.91
80.5	1.21	1.76	1.99	1.47	2.30	2.37	–0.52	–0.05	–0.90
70	1.17	1.59	2.12	1.45	2.07	2.22	0.43	–0.08	–1.47
80	1.19	1.74	2.34	1.47	2.24	2.26	0.43	0.27	–1.83
80	1.20	1.75	2.29	1.47	2.23	2.24	0.51	0.34	–1.76
70	1.12	1.64	1.72	1.47	2.14	2.27	0.37	0.25	–0.10
80	1.13	1.82	1.91	1.50	2.30	2.50	0.32	0.51	0.51
80	1.12	1.82	1.91	1.51	2.31	2.49	0.30	0.49	0.49
69	1.22	1.66	2.09	1.58	2.18	2.31	–0.02	–0.11	–1.21

Table 4.2. MÉRA (M), WRF (W) and COSMO (C) wind direction (degrees) error analysis for seven turbines at various heights

	MAE			STD			Bias		
Height (m)	M	W	C	M	W	C	M	W	C
68	14.1	23.9	29.5	22.5	37.1	44.9	1.85	–0.98	–7.31
78	12.8	23.1	28.6	21.0	36.4	44.1	2.59	0.69	–6.09
70	14.2	24.7	30.8	22.4	36.9	44.2	–2.79	–5.91	–11.1
48	15.9	24.7	27.4	25.6	38.3	41.2	–1.89	–3.60	–5.30
78	15.8	25.0	28.2	25.5	38.5	41.5	–1.19	–3.78	–9.00
48	16.3	25.6	30.2	25.1	38.3	45.1	–4.66	–4.24	–5.79
68	16.0	25.5	31.9	24.8	37.7	45.2	–3.95	–7.00	–11.24

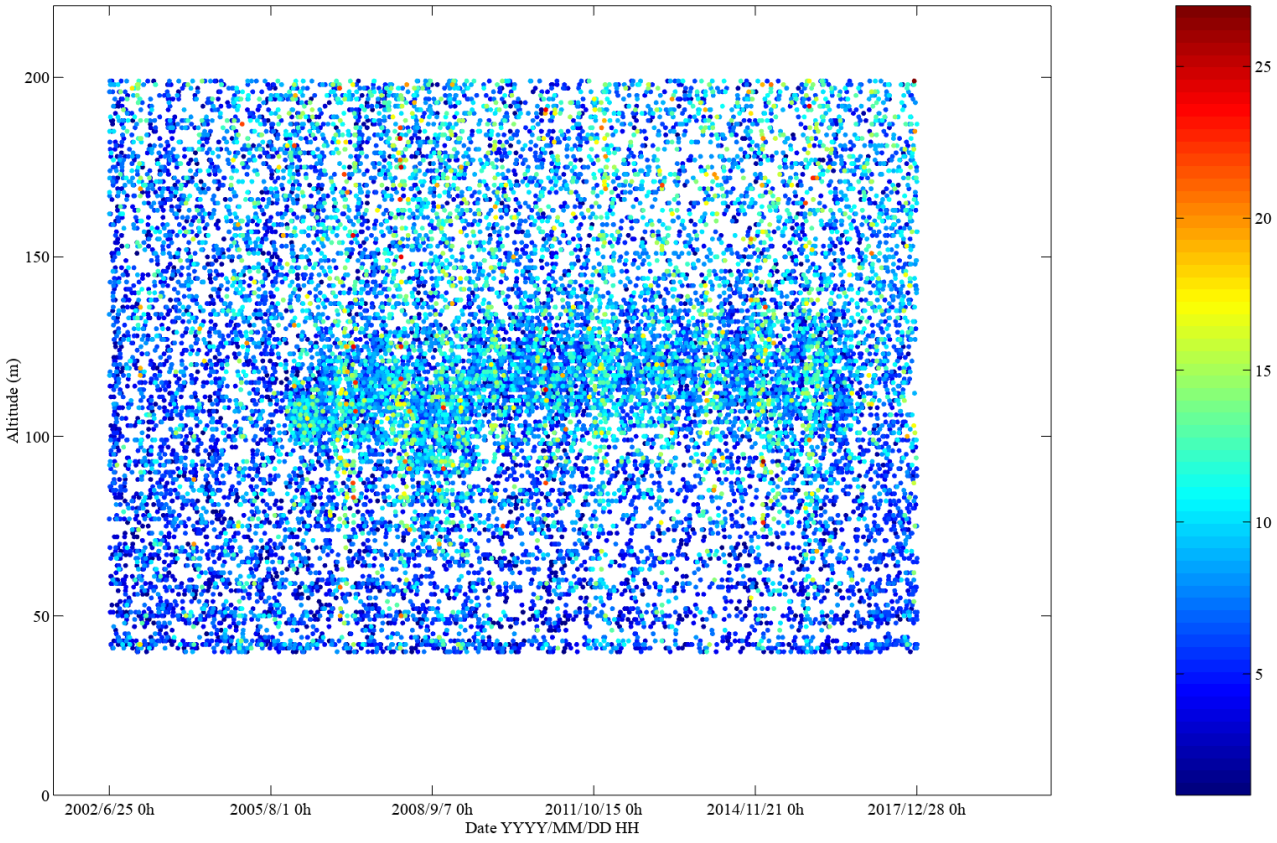


Figure 4.1. Processed Castor Bay radiosonde wind speeds (m s^{-1}).

validations (CAPE 3 km, Showalter index and surface lifted index) facilitated by radiosonde data acquisition are shown in Table 4.7.

4.2 Solar Fields

4.2.1 Satellite data

Satellite-based radiation data were obtained from two different sources, the Photovoltaic Geographical Information System (PVGIS) and the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), and were used to validate the three model outputs. The PVGIS data are at 0.025° (approximately 1.5 km) resolution, have an overall mean bias of 2% (Huld *et al.*, 2012) and contain average annual and seasonal sums for the period 1998–2011. The EUMETSAT data are at 0.05° (approximately 3 km) resolution and contain hourly sums over the period 5 November 2011 to 31 December 2015, with a typical annual bias of 0.2% (EUMETSAT, 2018). The data required several

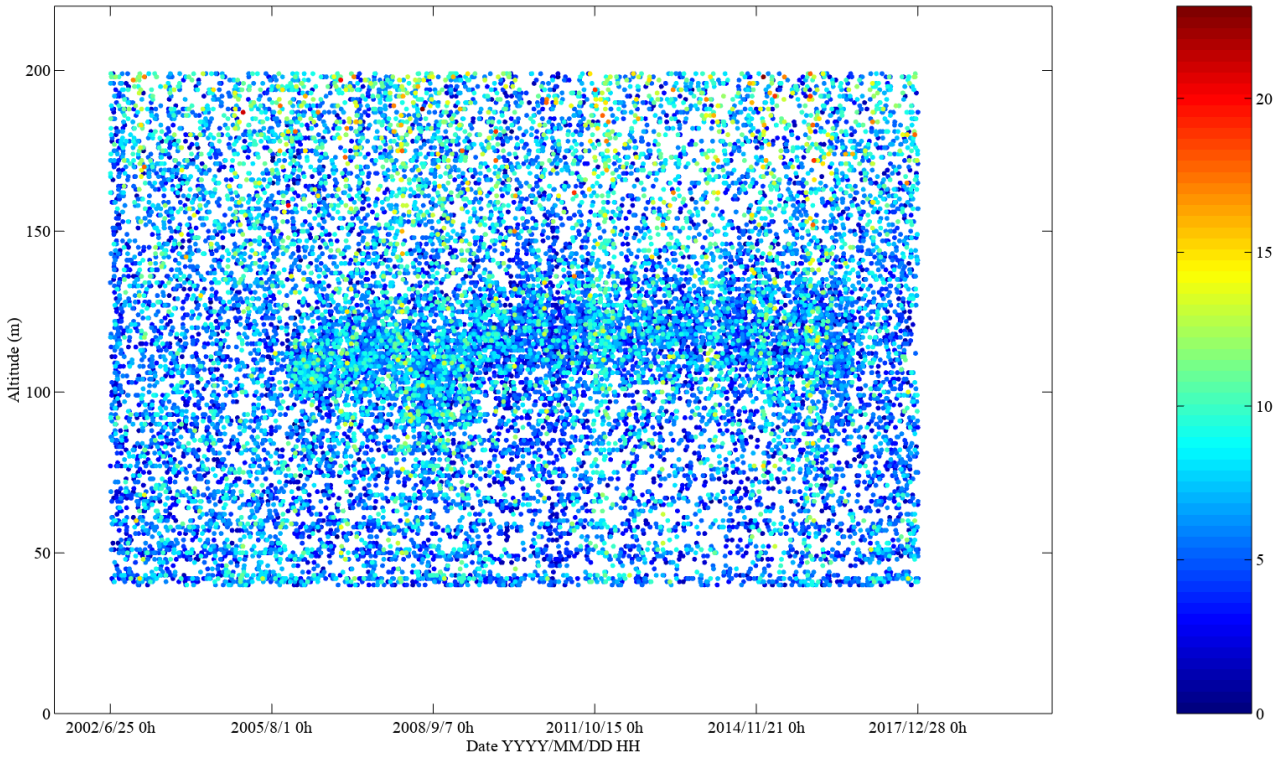


Figure 4.2. Processed Castor Bay COSMO-CLM wind speeds (m s⁻¹).

Table 4.3. Castor Bay wind speed uncertainty estimates (m s⁻¹)

	Model		
Statistic	COSMO	WRF	MÉRA
Bias	–0.80	1.32	1.20
STD	2.45	2.55	1.85
MAE	1.95	2.28	1.74

Table 4.4. Valentia wind speed uncertainty estimates (m s⁻¹)

	Model		
Statistic	COSMO	WRF	MÉRA
Bias	–0.12	0.62	0.02
STD	3.29	3.24	2.38
MAE	2.51	2.50	1.70

Table 4.5. Castor Bay wind direction uncertainty estimates (degrees)

	Model		
Statistic	COSMO	WRF	MÉRA
Bias	–6.46	1.43	1.17
STD	50.5	48.9	37.1
MAE	35.9	33.9	24.99

Table 4.6. Valentia wind direction uncertainty estimates

	Model		
Statistic	COSMO	WRF	MÉRA
Bias	–0.02	0.96	3.52
STD	48.9	44.6	35.0
MAE	32.4	29.1	21.0

Table 4.7. A selection of parameter validations performed for COSMO-CLM data using radiosonde data from two locations (Valentia and Castor Bay)

	Valentia			Castor Bay		
Parameter	Bias	MAE	STD	Bias	MAE	STD
CAPE 3 km ($J kg^{-1}$)	2.31	2.31	47.5	–0.68	10.33	34.9
Showalter index	–0.05	2.58	3.51	0.43	2.52	3.36
Surface lifted index	0.06	2.27	3.03	1.08	2.58	3.09

processing steps before use: the satellite datasets are on much larger domains and at different resolutions (0.05°) from the three models; the MÉRA 3-hour forecasts contain 1-, 2- and 3-hour accumulations; and the COSMO data are split into averaged direct and averaged diffuse surface downward short-wave radiation. Several CDO-based scripts were developed to prepare both the observational and the model data for error analysis. Results from the PVGIS analysis are shown in Figure 4.3; the WRF data display greater errors than either the COSMO or the MÉRA data. The PVGIS results (for all three models) and EUMETSAT results (for MÉRA and COSMO only) are given in Table 4.8. In the PVGIS analysis, COSMO and MÉRA show similar error trends.

4.2.2 Station data

Yearly global irradiance totals were derived from 18 Met Éireann daily station observations and checked against the satellite observations (Figure 4.4). The length and range of each observation time series are station dependent but typically cover multiple years over the period 1987–2015. However, error values (not shown) are similar to those found during the satellite data analysis.

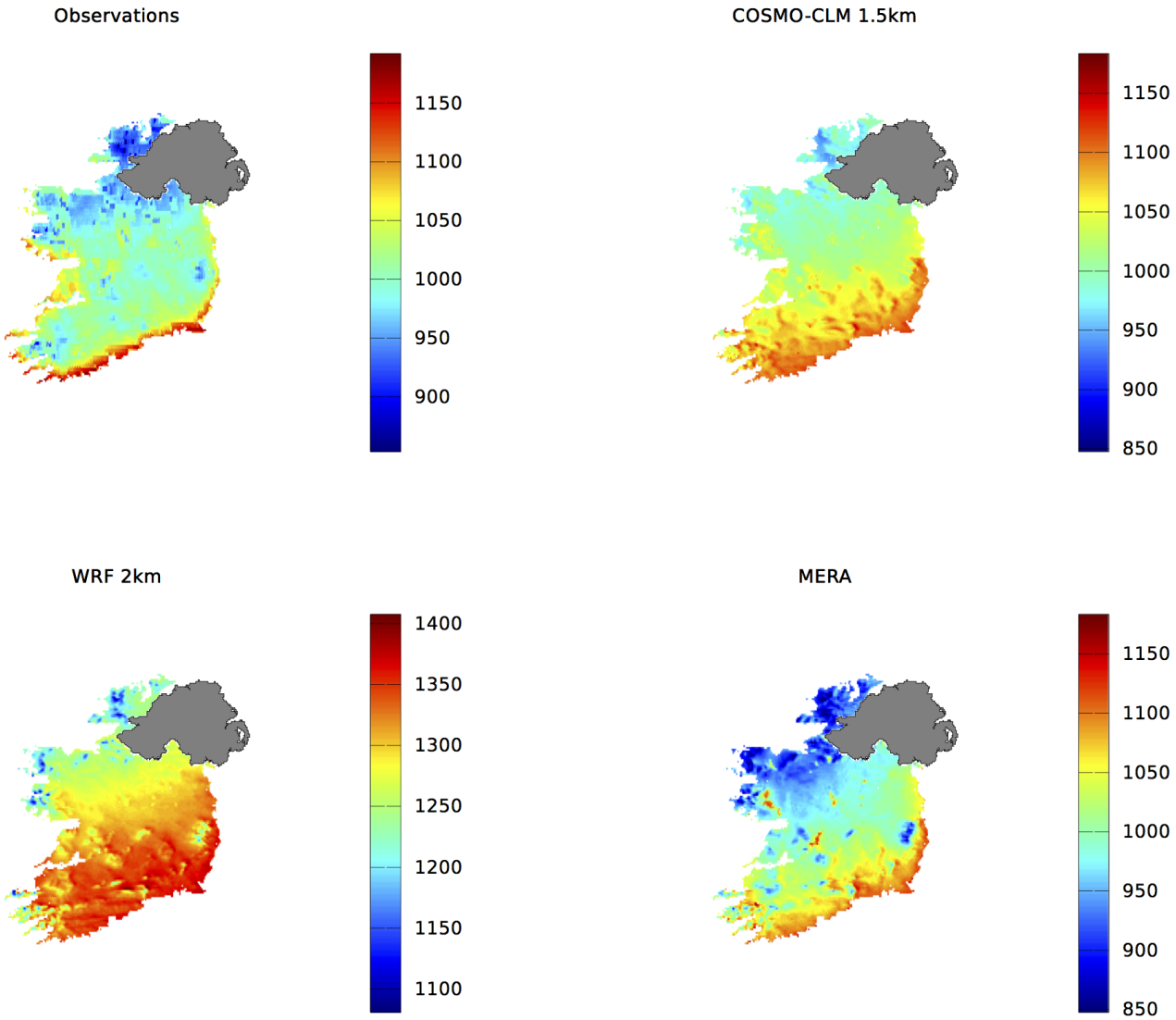


Figure 4.3. Average annual global irradiance (kW h m^{-2}) for the period 1998–2011 from the PVGIS dataset (top left). Both COSMO (top right) and MÉRA (bottom right) capture the overall range and spatial trends, with WRF (bottom left) displaying greater errors.

Table 4.8. Error analysis of global irradiance based on the PVGIS (1998–2011) annual average dataset and preliminary use of the EUMETSAT hourly dataset (for 2012 only)

Dataset	Model	Mean % error	Mean absolute % error	STD (%)	RMS (%)
PVGIS (1998–2011)	COSMO	2.83	3.58	3.47	4.48
	WRF	29.17	29.17	4.76	29.56
	MÉRA	–0.81	3.48	4.46	4.54
EUMETSAT (2012 only)	COSMO	6.3	6.49	5.4	8.40
	MÉRA	0.15	3.42	4.6	4.64

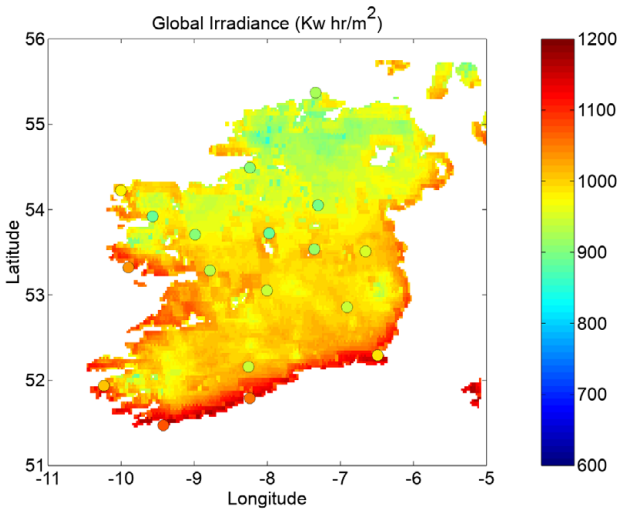


Figure 4.4. PVGIS 1998–2011 annual global irradiance (kW h m^{-2}) superimposed with average annual values (small circles) derived from 18 Met Éireann daily stations.

5 European Climate Assessment and Dataset Indices

A major output from this project are gridded datasets of climate change indices. The Expert Team on Climate Change Detection and Indices (ETCCDI) produced a list of 27 core indices [used in the European Climate Assessment and Dataset (ECA&D)] that are relevant to climate change detection and that are based on daily temperature and precipitation; of the 27, 16 are temperature based. The CDO package provides a way to calculate many of these indices. Thus, numerous bash scripts that take advantage of task farming and embedded CDO commands have been written to allow their calculation, as well as nine other sundry indices. Some indices require analysis of the model (mean, maximum and minimum) 2 m temperature and precipitation (where appropriate) data at annual timescales, while others require analysis at monthly timescales. For both categories, the indices are calculated directly from the model datasets. Once the model data have been preprocessed into appropriate subsets (based on the relevant timescale), the algorithms/scripts required to produce each individual index are both similar and relatively straightforward. For other indices, the calculations are more involved as they require comparison with the base period 1961–1990. A high level of care is required when calculating these indices for "in-base" model data (i.e. model data from 1981 to 1990). The reasoning, reference period data and methods applied to overcome this problem are discussed below for the calculation of the percentage of days when the minimum temperature is below the 10th percentile (TN10p).

5.1 Annual Temperature Indices

These indices require yearly comparison with various temperature thresholds and include the number of frost days, icing days, summer days and tropical nights and the length (and start) of the growing season. The indices were calculated as follows.

5.1.1 Number of frost days

This index is defined as the annual count of days when the daily minimum temperature (TN) is less than 0°C. More precisely, for each year *j* in the period 1981–2017, the minimum temperature TN_{ij} is found for every *i* day. The number of days when $TN_{ij} < 0°C$ is then counted. This index was calculated using the CDO command *eca_fd* (i.e. "CDO *eca_fd infile outfile"*), where *infile* is an annual file containing TN at each model grid point and *outfile* is the number of frost days per grid point. Some example results using COSMO-CLM 1.5 km and MÉRA temperature data (for the years 1981, 1991, 2001 and 2011) are given in Figures 5.1 and 5.2, respectively. The spatially averaged values over the period 1981–2017 are shown in Figure 5.3 for both COSMO-CLM and MÉRA. In Figure 5.3, a downward trend is discernible, with especially cold years in 1985/1986 and 2010. Simple linear regressions of the form $y = a + bx$ are also shown for both COSMO-CLM and MÉRA. For COSMO-CLM, $b = -0.227$ [95% confidence interval (CI) −0.61 to 0.16], while for MÉRA, $b = -0.228$ (95% CI −0.62 to 0.16); the slopes (b) are similar and both suggest a downward trend. Exclusion of the 2010 outlier results in a strengthening of this trend: $b = -0.357$ for COSMO-CLM and $b = -0.363$ for MÉRA.

5.1.2 Number of icing days

The number of icing days is defined as the annual count of days when the daily maximum temperature (TX) is less than 0°C. For each year *j* in the period 1981–2017, the maximum temperature TX_{ij} is found for every *i* day. The number of days when $TX_{ij} < 0°C$ is then counted. This index was calculated using the CDO command *eca_id*. Some example results using COSMO-CLM 1.5 km and MÉRA data (for the years 1981, 1991, 2001 and 2011) are given in Figures 5.4 and 5.5, respectively. Spatially averaged values over the period 1981–2017 are shown in Figure 5.6 for both COSMO-CLM and MÉRA. Since Ireland's climate is relatively mild, there is a paucity of data points around which to base a regression. Consequently, the regression slopes are more susceptible to outliers and have not been calculated here.

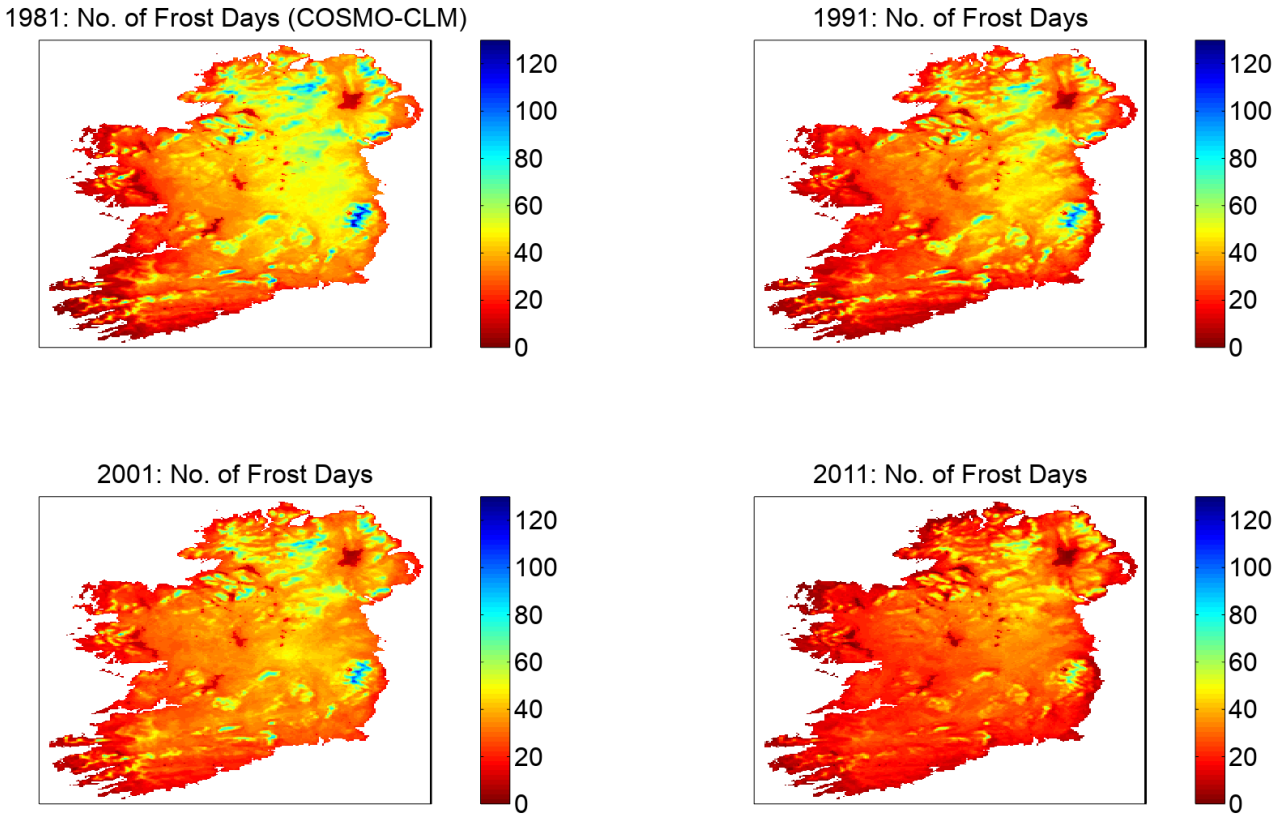


Figure 5.1. Number of annual frost days found using COSMO-CLM 1.5 km data for the years 1981, 1991, 2001 and 2011, showing a noticeable decrease in frost days in more recent years.

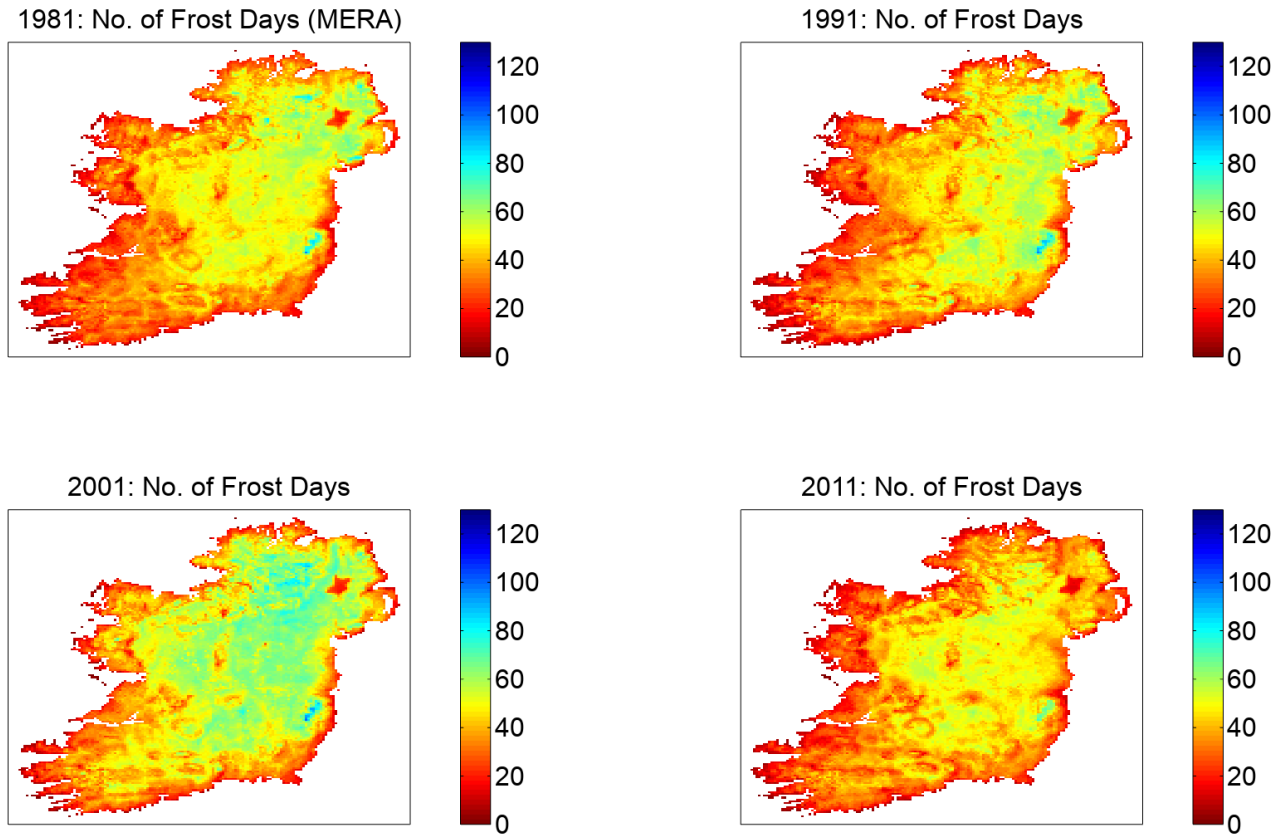


Figure 5.2. Number of annual frost days found using MÉRA data for the years 1981, 1991, 2001 and 2011.

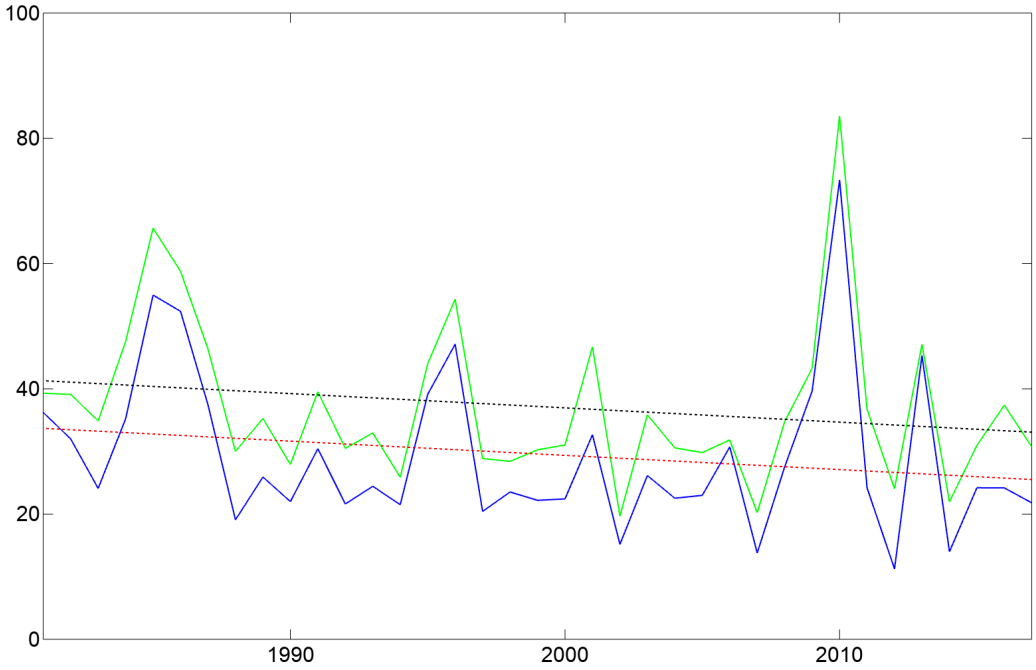


Figure 5.3. Spatially averaged number of frost days (FDs) per annum for the period 1981–2017 found using COSMO-CLM 1.5 km (blue line) and MÉRA (green line) temperature data. Also shown are two simple linear fits of the form FD = *a* + *bx*. For COSMO-CLM (red dotted line) *b* = −0.227, while for MÉRA (black dotted line) *b* = −0.228.

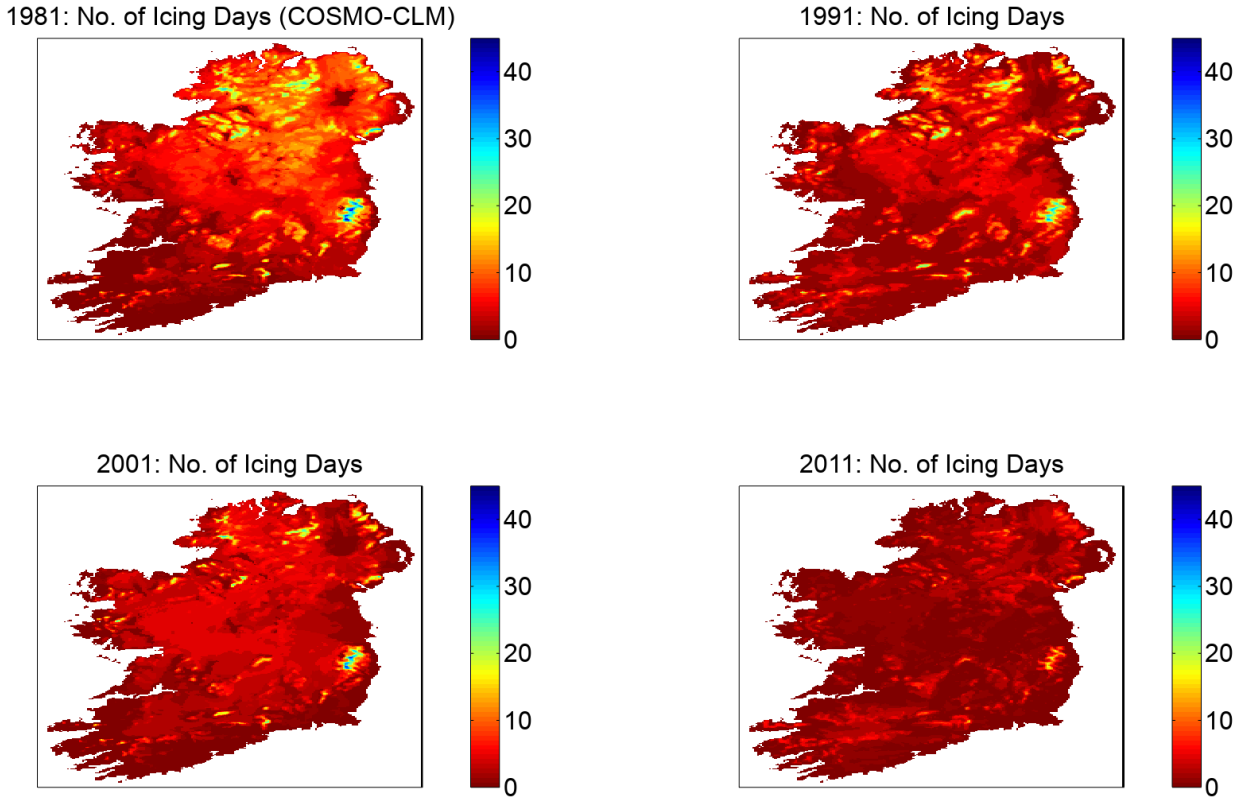


Figure 5.4. Number of annual icing days found using COSMO-CLM 1.5 km data for the years 1981, 1991, 2001 and 2011.

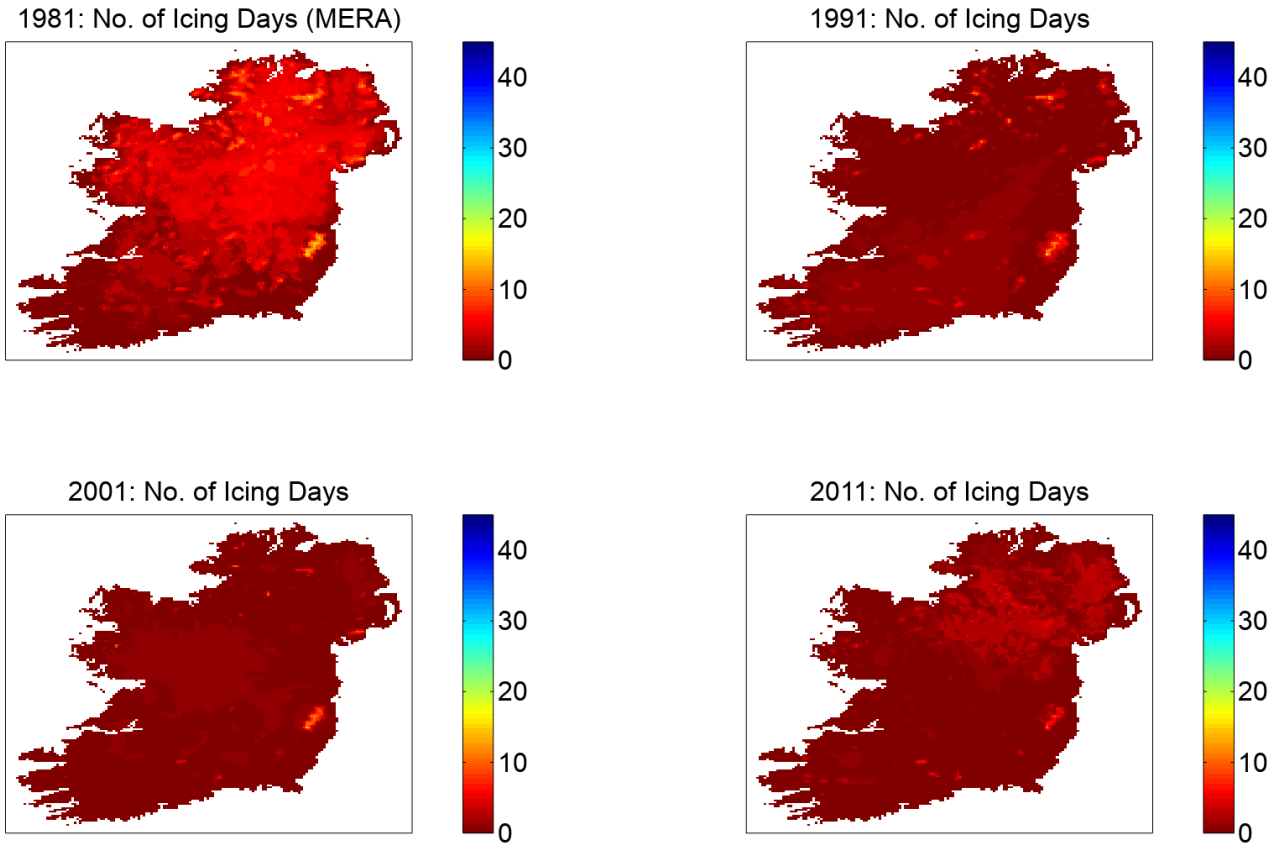


Figure 5.5. Number of annual icing days found using MÉRA data for the years 1981, 1991, 2001 and 2011.

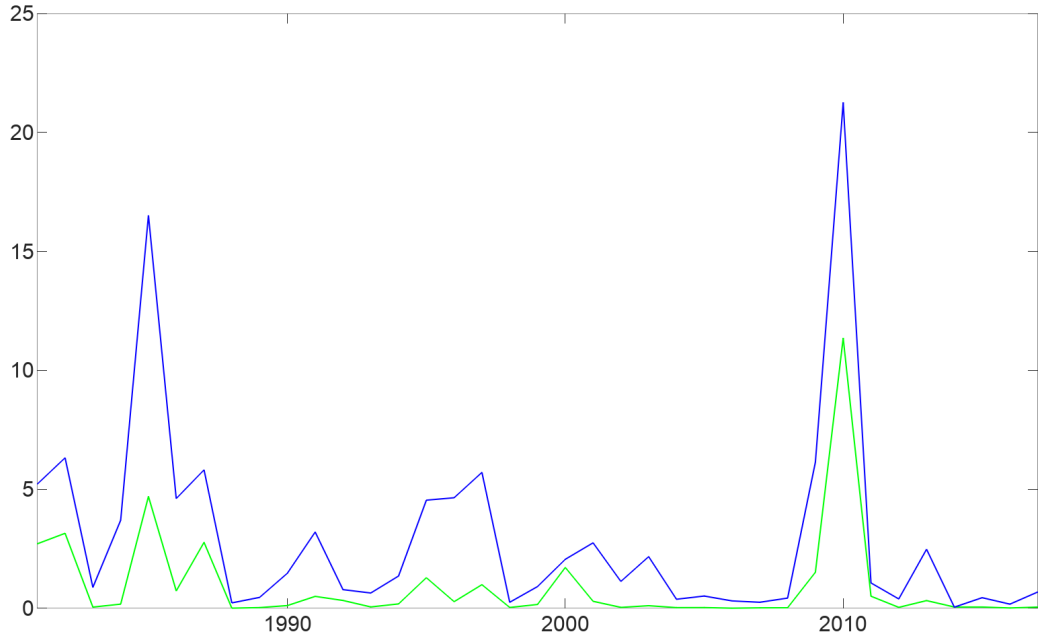


Figure 5.6. Spatially averaged number of icing days per annum for the period 1981–2017 found using COSMO-CLM 1.5 km (blue line) and MÉRA (green line) data.

5.1.3 Number of summer days and tropical nights

The number of summer days is the annual count of days when TX is greater than 25°C, while the number of tropical nights is the annual count of days when TN is greater than 20°C. These indices have been calculated for the period 1981–2017 using the CDO commands *eca_su* and *eca_tr*, respectively. Given Ireland's relatively mild climate, it can be expected that these indices are low in value, particularly the number of tropical nights. Figures 5.7 and 5.8 show some example outputs for number of summer days for individual years using COSMO-CLM and MÉRA data, respectively, while Figure 5.9 shows the spatially averaged number of summer days per year over the entire 1981–2017 period for COSMO-CLM and MÉRA. Figure 5.10 shows the equivalent data for the number of tropical nights. As expected, the values for each index are low, with no immediate trend recognisable from the data.

5.1.4 Growing season length

For the northern hemisphere, the growing season length (GSL) is defined as the annual count of days between the first span of 6 consecutive days with a daily mean temperature (TG) greater than 5°C and the first span (after 1 July) of 6 consecutive days with a TG of less than 5°C. The CDO command *eca_gsl* was used to output the growing season start (GSS) and GSL for each year. Example spatial outputs of GSS and GSL for the years 1981 and 2017 are given in Figures 5.11 and 5.12 for COSMO-CLM and MÉRA, respectively. The models show similar results for GSL, with noticeably increasing values towards the south-west and an almost year-round growing season. The spatially averaged annual GSL for the period 1981–2017 is shown in Figure 5.13 for COSMO-CLM and MÉRA. Linear regressions ($y = a + bx$) for both models are also shown. Both models suggest an increasing GSL: $b = 0.43$ (95% CI −0.19 to 1.05) for COSMO-CLM and $b = 0.62$ (95% CI −0.03 to 1.27) for MÉRA. Linear regressions for GSS were also tested (not shown). As expected, both regressions have a

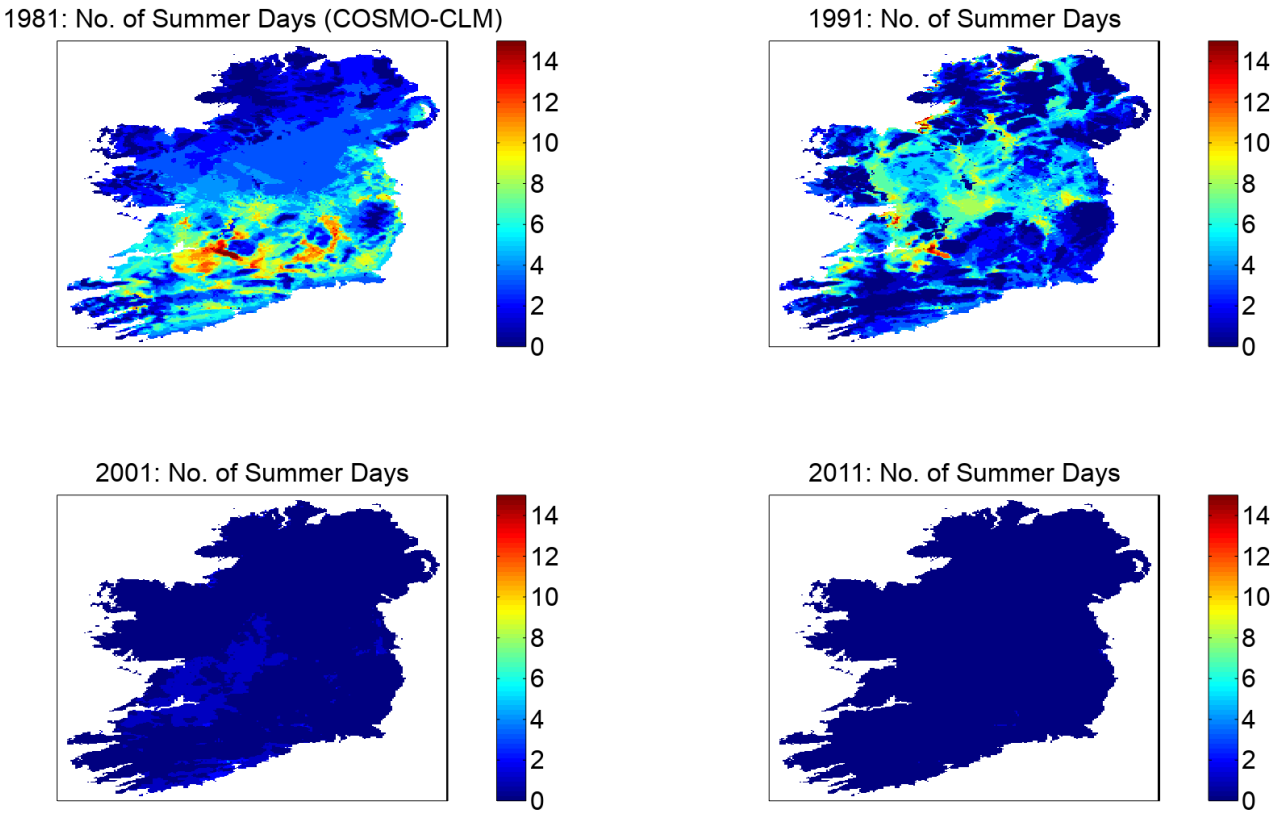


Figure 5.7. Number of annual summer days found using COSMO-CLM 1.5 km data for the years 1981, 1991, 2001 and 2011.

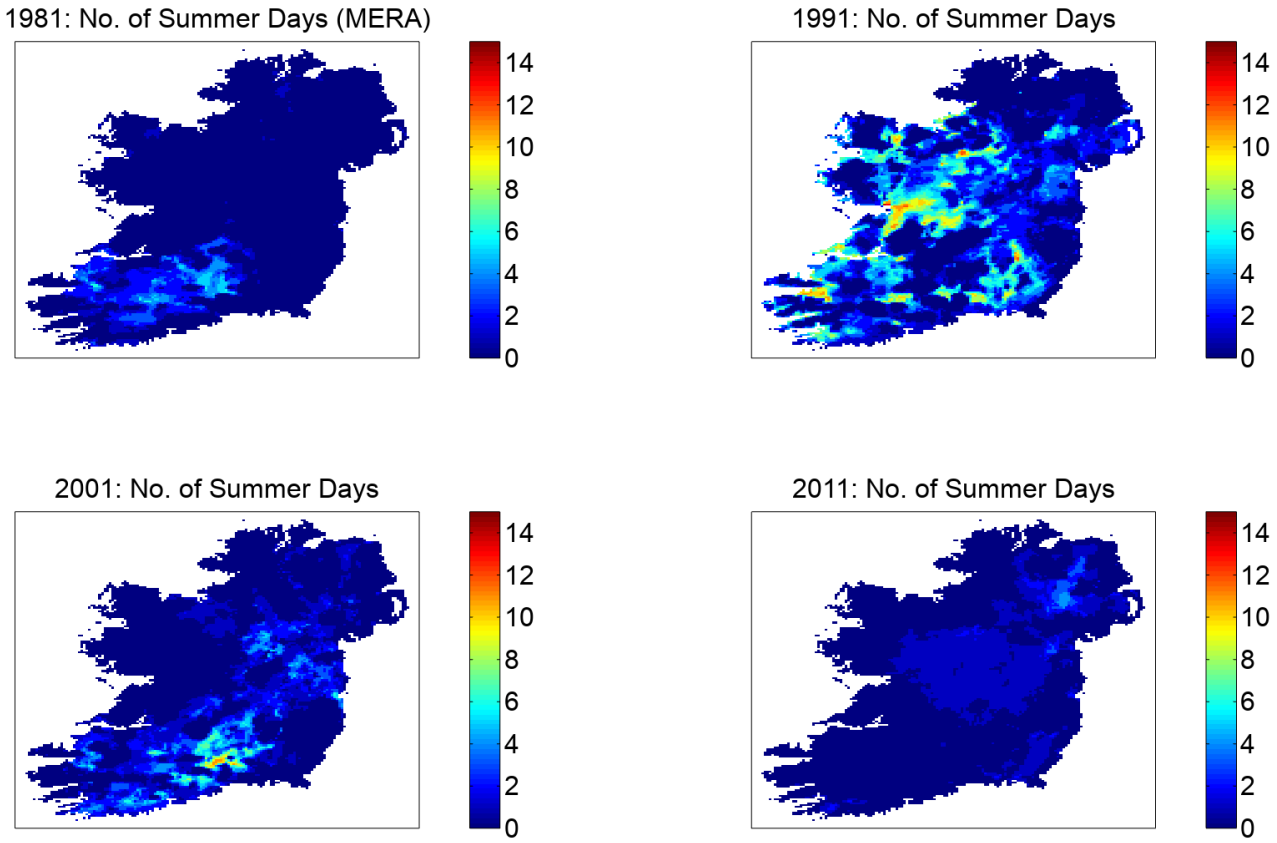


Figure 5.8. Number of annual summer days found using MÉRA data for the years 1981, 1991, 2001 and 2011.

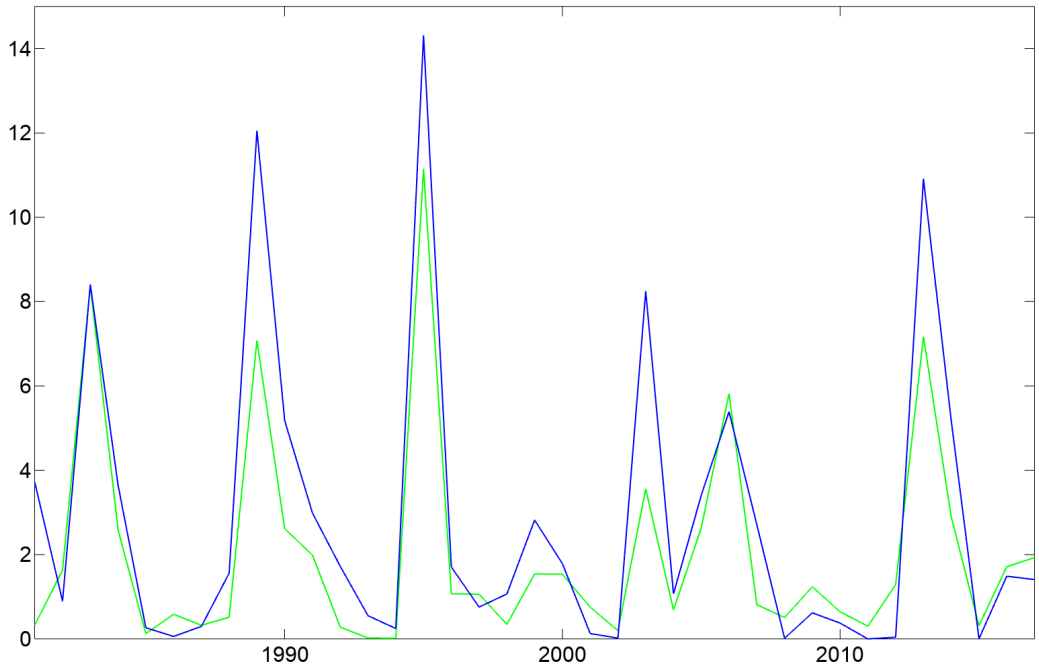


Figure 5.9. Spatially averaged number of summer days per annum for the period 1981–2017 found using COSMO-CLM 1.5 km (blue line) and MÉRA (green line) data.

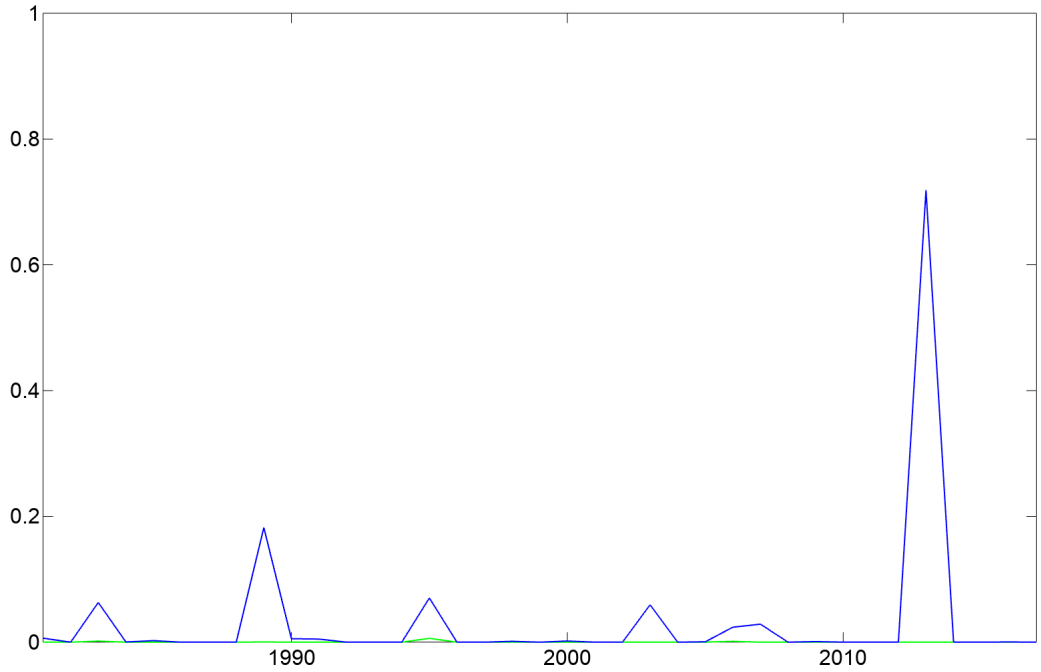


Figure 5.10. Spatially averaged number of tropical nights per annum for the period 1981–2017 found using COSMO-CLM 1.5 km (blue line) and MÉRA (green line) data.

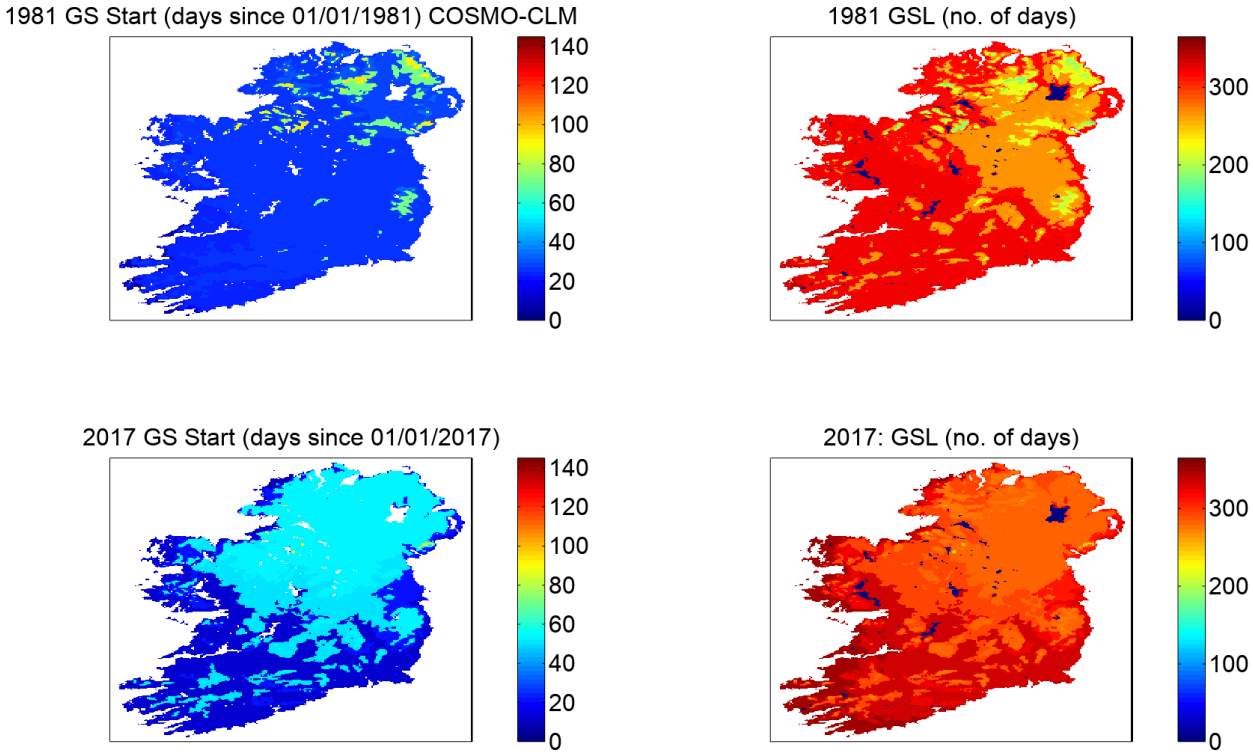


Figure 5.11. COSMO-CLM GSS (left column) and GSL (right column) for the years 1981 (top row) and 2017 (bottom row).

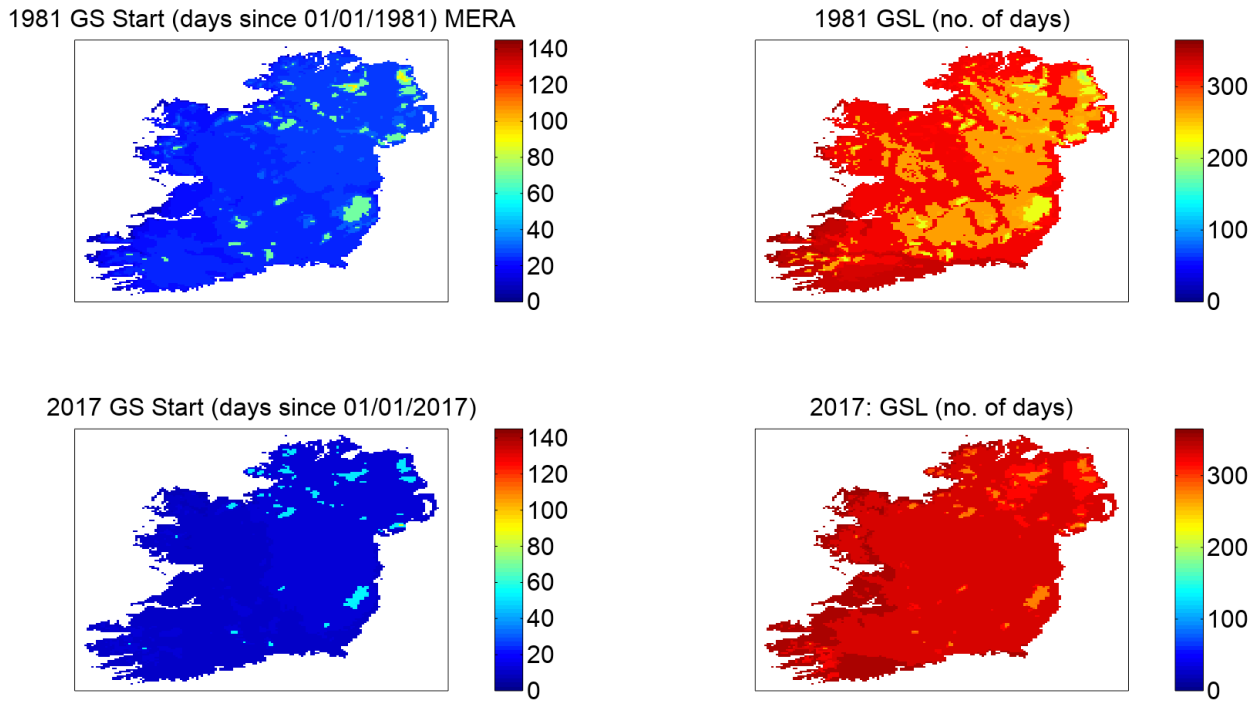


Figure 5.12. MÉRA GSS (left column) and GSL (right column) for the years 1981 (top row) and 2017 (bottom row).

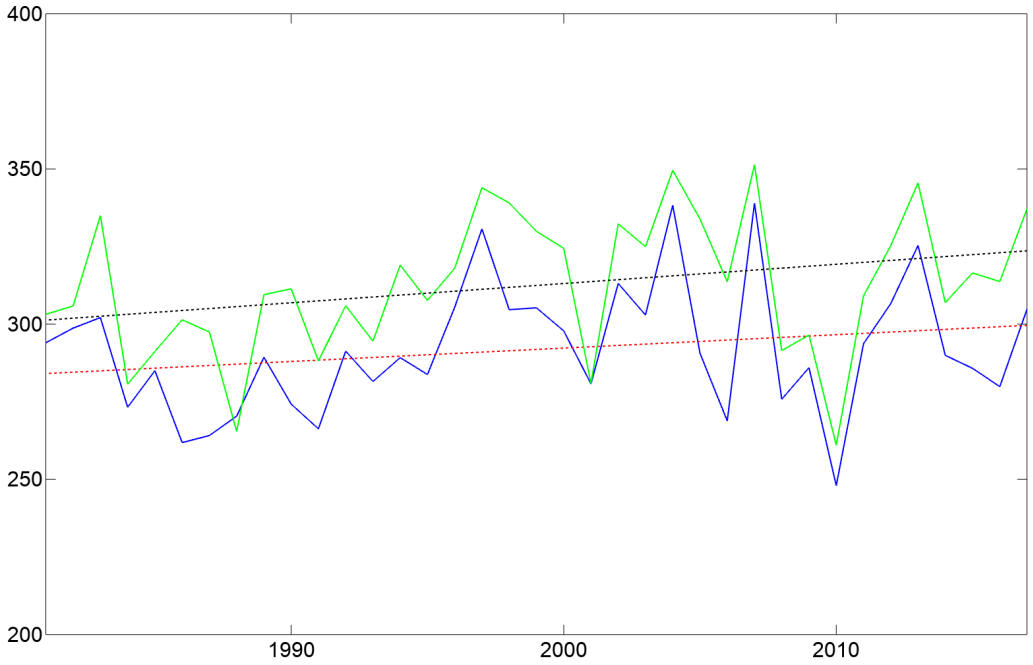


Figure 5.13. Spatially averaged annual GSL for the period 1981–2017 found using COSMO-CLM 1.5 km (blue line) and MÉRA (green line) data. Also shown are two simple linear fits of the form GSL = *a* + *bx*. For COSMO-CLM (red dotted line) *b* = 0.431, while for MÉRA (black dotted line) *b* = 0.620.

negative slope ($b = -0.12$, 95% CI −0.72 to 0.48, for COSMO-CLM; $b = -0.31$, 95% C.I −0.85 to 0.23, for MÉRA), suggesting increasingly earlier start dates.

5.2 Monthly Temperature Indices

Five monthly indices for each month (k) and year (j) were calculated (with units of °C) for the period 1981–2017. These are the monthly maxima of daily maximum (TX_X) and minimum (TN_X) temperatures (TX_{Xkj} and TN_{Xkj}) and monthly minima of daily maximum (TX_N) and minimum (TN_N) temperatures (TX_{Nkj} and TN_{Nkj}), and the monthly daily temperature range (DTR), which is simply the monthly mean difference between TX and TN. Examples of COSMO-CLM TX_X and MÉRA TN_X spatial maps (for the months January, April, July and October 1981) are shown in Figures 5.14 and 5.15, respectively. The spatially averaged values for COSMO-CLM TX_N and MÉRA TN_N for each month (by season) over the period 1981–2017 are given in Figures 5.16 and 5.17, respectively. An example of the spatially averaged DTR calculated for MÉRA for the period 1981–2017 is shown in Figure 5.18.

5.3 Base Period (1961–1990) Temperature Indices

Several climate change indices that require comparison with the base/reference period 1961–1990 were calculated. Gridded observations of TG, TX and TN at 1 km spatial resolution were previously acquired from Met Éireann (during the first 6 months of this project) and have been utilised here to construct the base period. For the indices calculated here, a potential problem can arise as a result of inhomogeneity across the "in-base" (1981–1990) and "out-of-base" (1991–2017) periods. The ETCCDI recommends that the bootstrapping algorithm outlined in Zhang *et al*. (2005) be followed. This procedure was implemented for the following indices: percentage of days when TN < 10th percentile (TN10p), percentage of days when TX < 10th percentile (TX10p), percentage of days when TN > 90th percentile (TN90p), percentage of days when TX > 90th percentile (TX90p), warm spell duration index (WSDI) and cold spell durarion index (CSDI) and is explained below in the context of the TN10p calculations.

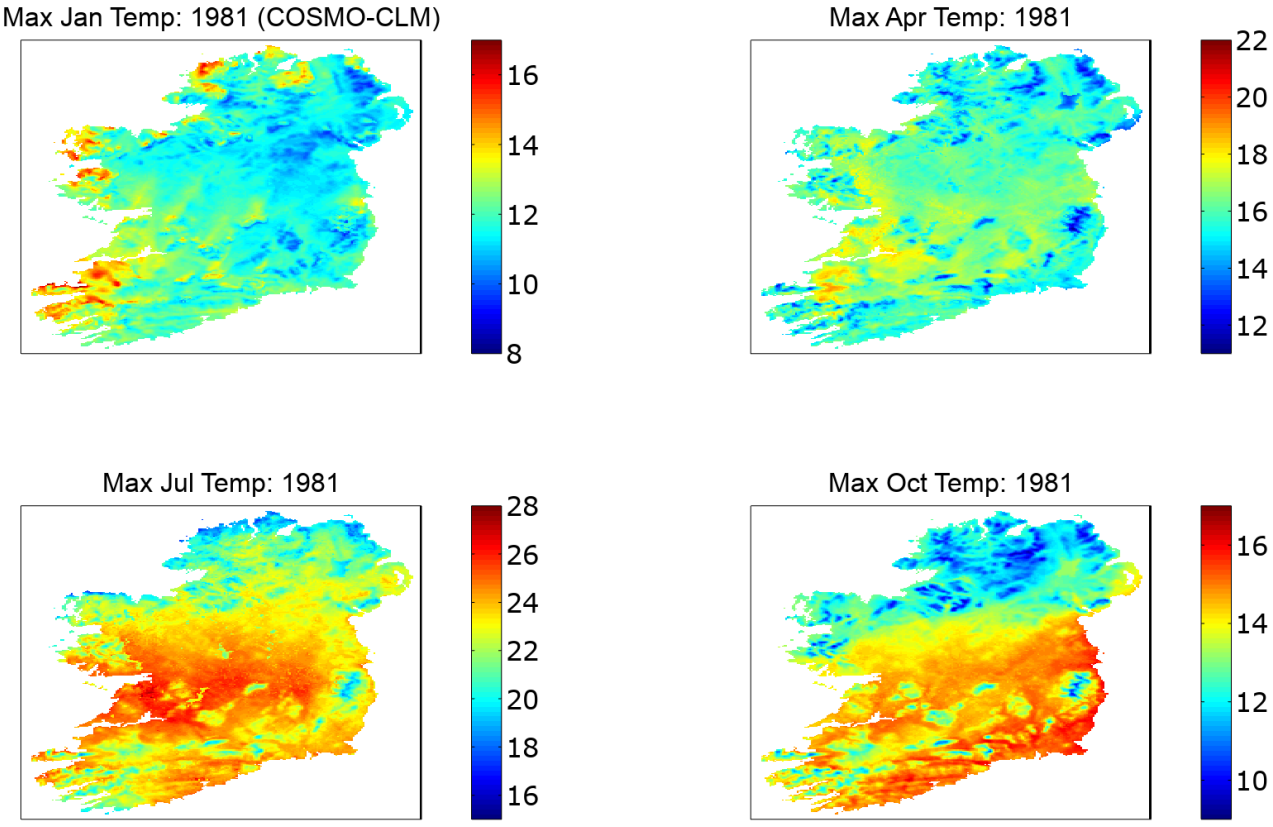


Figure 5.14. COSMO-CLM TX_X (°C) for the months January (top left), April (top right), July (bottom left) and October (bottom right) 1981.

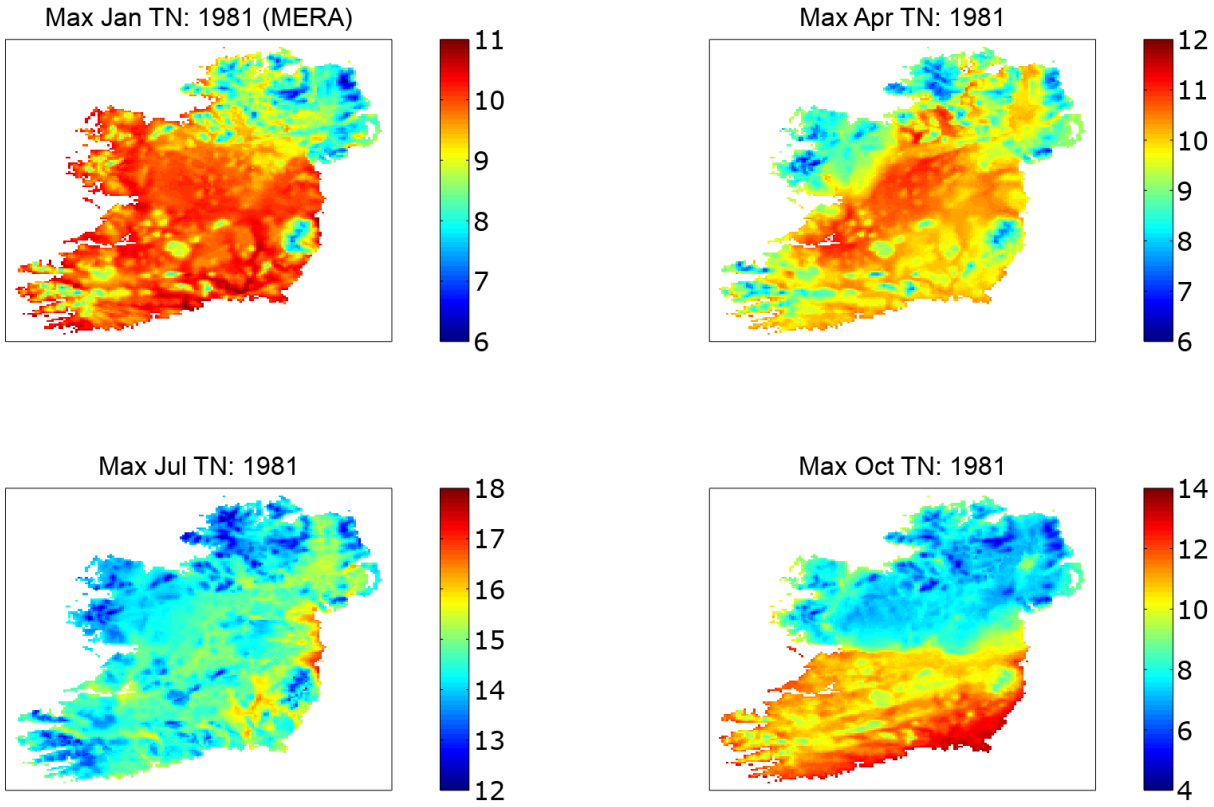


Figure 5.15. MÉRA TN_x (°C) for the months January (top left), April (top right), July (bottom left) and October (bottom right) 1981.

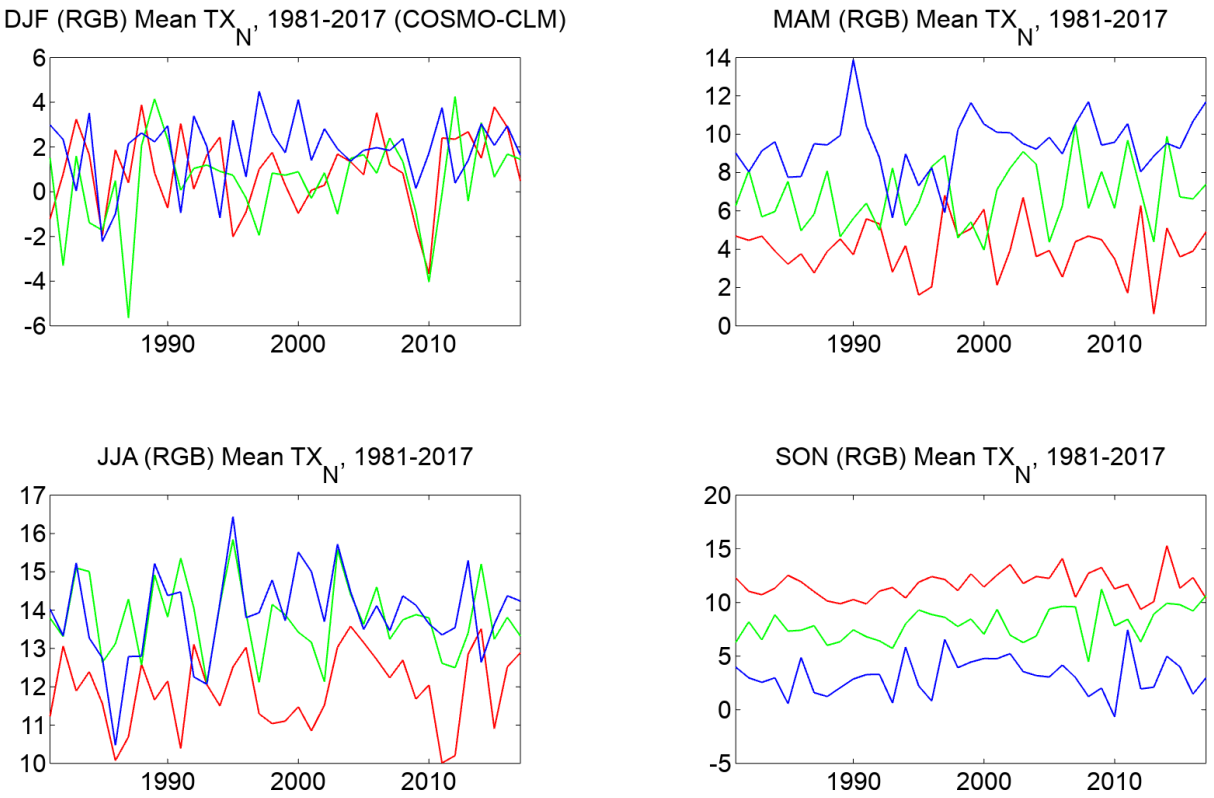


Figure 5.16. Spatially averaged COSMO-CLM monthly TX_N (°C) for the period 1981–2017, grouped according to season: winter (top left), spring (top right), summer (bottom left) and autumn (bottom right). For each season, the months are represented by the colours red, green and blue, in that order, e.g. December (red), January (green) and February (blue).

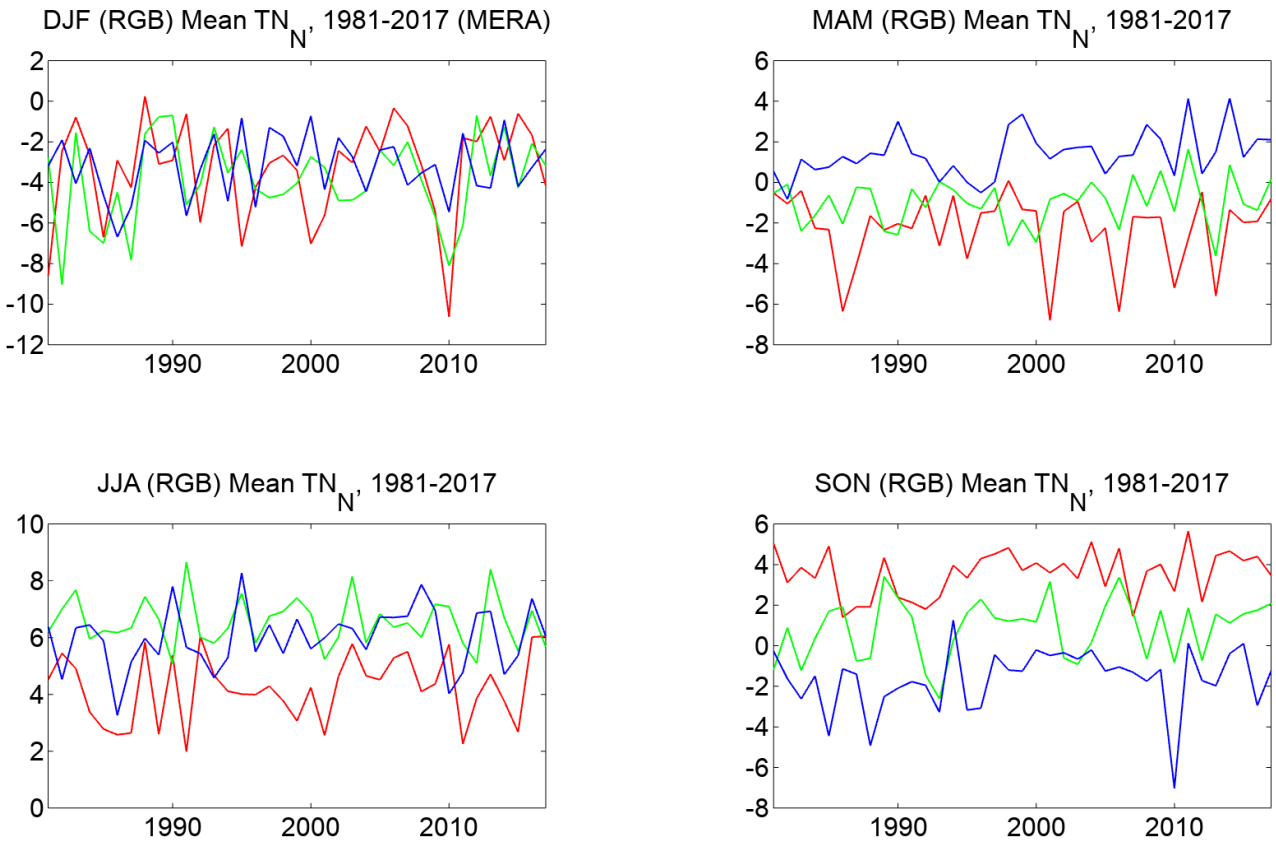


Figure 5.17. Spatially averaged MÉRA monthly TN_N (°C) for the period 1981–2017, grouped by season: winter (top left), spring (top right), summer (bottom left) and autumn (bottom right). For each season, the months are represented by the colours red, green and blue, in that order, e.g. December (red), January (green) and February (blue).

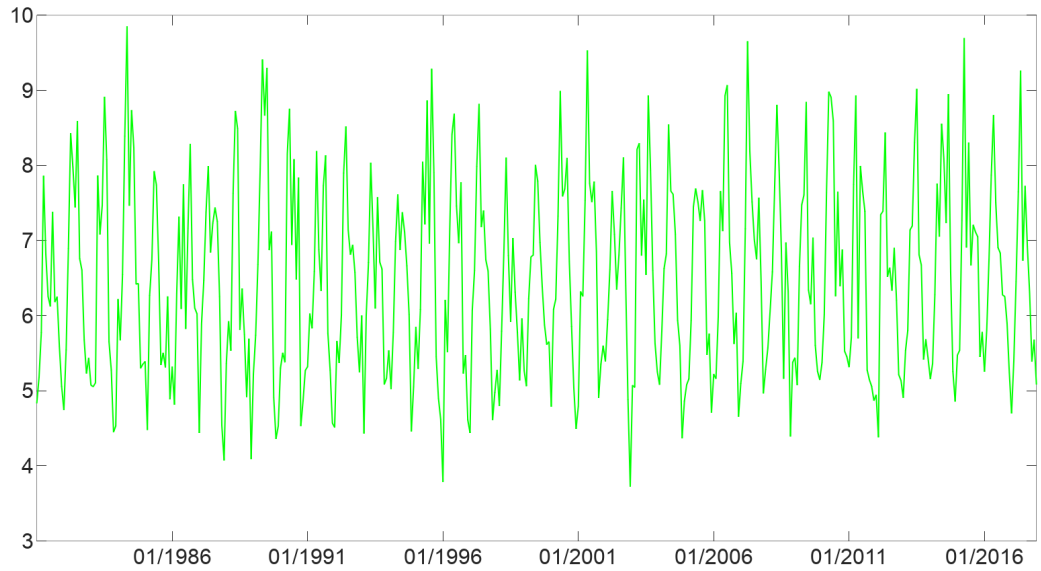


Figure 5.18. Spatially averaged MÉRA monthly average DTR (°C) for the period 1981–2017.

5.3.1 Percentage of days when TN<10th percentile

The daily TN for each day *i* over a period *j* was calculated; here, *j* is a calendar year from the period 1981–2017. These TN_{ij} were then compared with $TN_{in}10$, which is the calendar day 10th percentile centred on a 5-day window for the base period. The percentage of days when $TN_{ij} < TN_{in}10$ was then calculated. For those years that are "in-base" years (1981–1990), the following algorithm (Zhang *et al*., 2005) was implemented:

1. The period 1961–1990 was divided into a single-year "in-base" dataset that contains the year for which the exceedance was to be calculated and a "29-year base period" from which the percentiles were to be calculated.
2. A selection of 29 new "temporary 30-year base periods" were constructed from the "29-year base period" by iteratively making a copy of one individual year within the "29-year base period".
3. Twenty-nine different TN10p values were then calculated by comparison of the "in-base" year from step 1 with each of the new "temporary 30-year base periods" constructed in step 2.
4. An average was then taken across all 29 TN10p values to form the TN10p value for this "in-base" year.

The procedure outlined above was then repeated for each of the years that are "in base", i.e. 1981–1990, generating a large volume of data. The processing was achieved through a combination of bash, CDO (e.g. the *eca_tn10p* command) and task farm scripts. A summary of the (annual, spatially averaged) results is given in Figure 5.19, where a potential downward trend was quantified by simple linear fits with slope $b = -0.064$ (95% CI −0.13 to 0) for COSMO-CLM and $b = -0.079$ (95% CI −0.17 to 0.01) for MÉRA. This downward trend suggests warming, as fewer daily minimum temperatures are below the 1961–1990 threshold.

5.3.2 Percentage of days when TX<10th percentile

The daily TX for each day *i* over a period *j* was calculated; here, *j* is a calendar year from the period 1981–2017. These TX_{ij} were then compared with $TX_{in}10$, which is the calendar day 10th percentile centred on a 5-day window for the base period. The

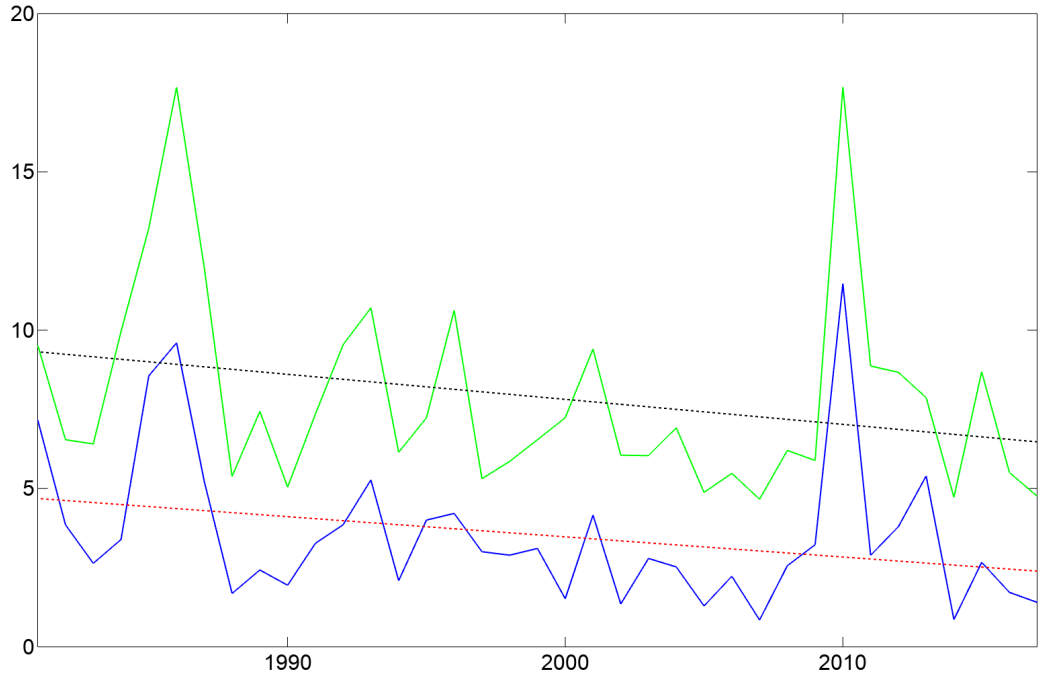


Figure 5.19. Spatially averaged COSMO-CLM (blue line) and MÉRA (green line) (annual) TN10p (%) found from comparison of the years 1981–2017 with the base period 1961–1990 and using the algorithm outlined in Zhang *et al.* (2005) for "in-base" years (1981–1990). Also shown are linear fits for both COSMO-CLM (red dotted line) and MÉRA (black dotted line).

percentage of days when $TX_{ij} < TX_{in}10$ was then calculated (CDO command *eca_tx10p*). In Figure 5.20, the annual spatially averaged TX10p values found for COSMO-CLM and MÉRA are shown. As with TN10p, there is a noticeable downward trend, indicating warming, as fewer daily maximum temperatures are lower than the historical threshold. Also shown in Figure 5.20 are linear fits with slope $b = -0.20$ (95% CI −0.36 to −0.03) for COSMO-CLM and $b = -0.27$ (95% CI −0.42 to −0.12) for MÉRA, which quantify this trend.

5.3.3 *Percentage of days when TX > 90th percentile*

Values for TX_{ij} (calculated earlier) were compared with $TX_{in}90$, which is the calendar day 90th percentile centred on a 5-day window for the base period. The percentage of days when $TX_{ij} > TX_{in}90$ was then calculated (CDO command *eca_tx90p*). In Figure 5.21, the annual spatially averaged TX90p values found for COSMO-CLM and MÉRA are shown. There is a noticeable upward trend for MÉRA TX90p and a much weaker (practically non-existent) trend for COSMO-CLM TX90p. The MÉRA trend suggests warming, as increasing numbers of daily maximum temperatures are higher than the historical threshold. These trends are quantified in Figure 5.21 by linear fits with slope $b = 0.004$ (95% CI −0.11 to 0.12) for COSMO-CLM and $b = 0.08$ (95% CI −0.01 to 0.16) for MÉRA.

5.3.4 *Percentage of days when TN > 90th percentile*

Values for TN_{ij} (calculated earlier) were compared with $TN_{in}90$, which is the calendar day 90th percentile centred on a 5-day window for the base period. The percentage of days when $TN_{ij} > TN_{in}90$ was then calculated (CDO command *eca_tn90p*). In Figure 5.22, the annual spatially averaged TN90p values found for COSMO-CLM and MÉRA are shown. As with TX90p, there is a much stronger upward trend for MÉRA TN90p than for COSMO-CLM TN90p. The MÉRA upward trend suggests warming, as increasing numbers of daily minimum temperatures are higher than the historical threshold. Both trends are quantified in Figure 5.22 by linear fits with slope $b = 0.003$ (95% CI −0.12 to 0.13) for COSMO-CLM and $b = 0.04$ (–0.05 to 0.13) for MÉRA.

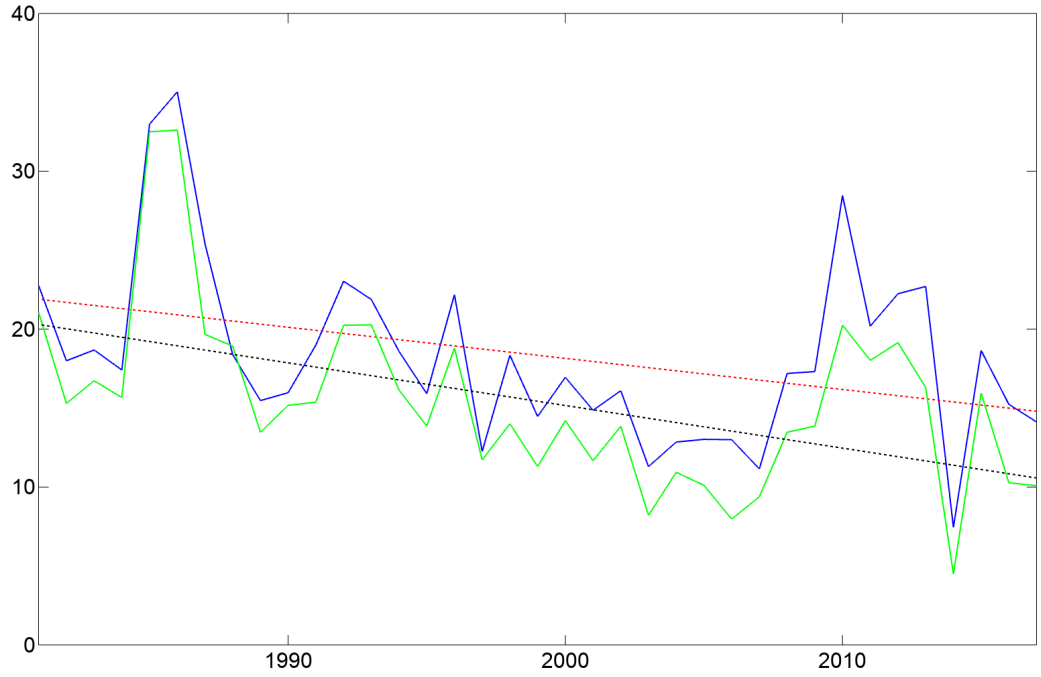


Figure 5.20. Spatially averaged COSMO-CLM (blue line) and MÉRA (green line) (annual) TX10p (%) found from comparison of the years 1981–2017 with the base period 1961–1990 and using the algorithm outlined in Zhang *et al.* (2005) for "in-base" years (1981–1990). Also shown are linear fits for both COSMO-CLM (red dotted line) and MÉRA (black dotted line).

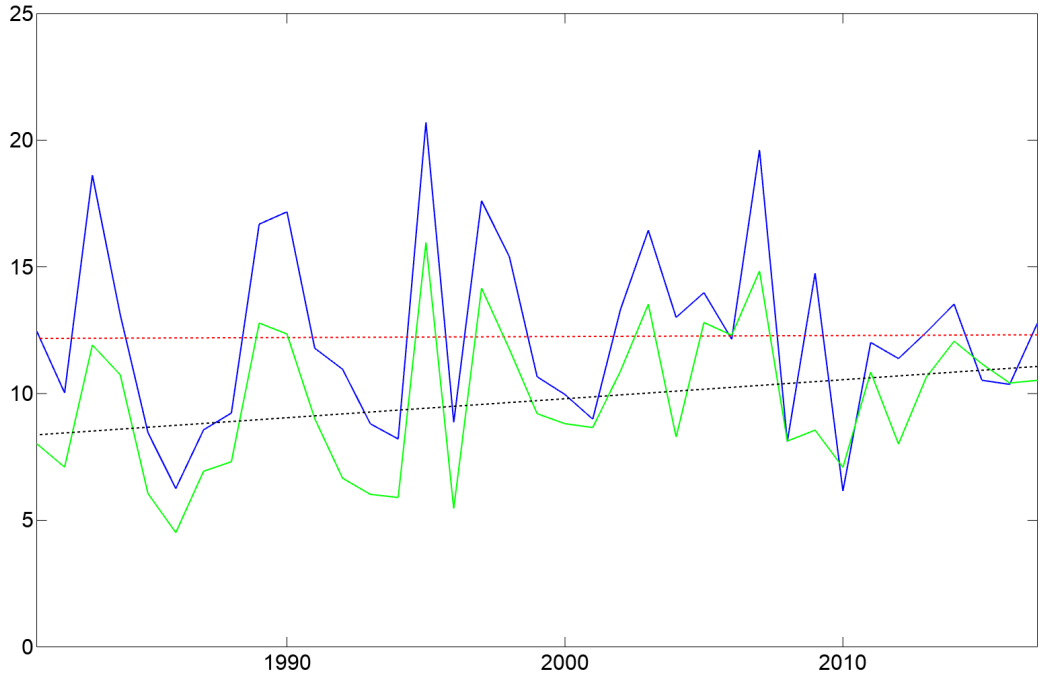


Figure 5.21. Spatially averaged COSMO-CLM (blue line) and MÉRA (green line) (annual) TX90p (%) found from comparison of the years 1981–2017 with the base period 1961–1990 and using the algorithm outlined in Zhang *et al.* (2005) for "in-base" years (1981–1990). Also shown are linear fits for both COSMO-CLM (red dotted line) and MÉRA (black dotted line).

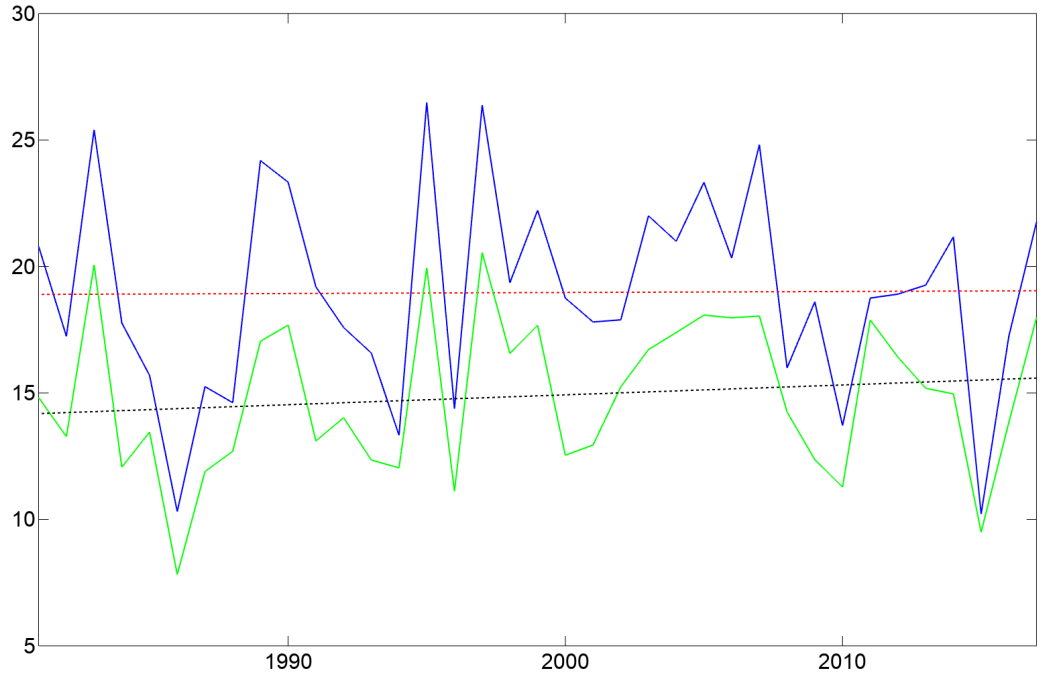


Figure 5.22. Spatially averaged COSMO-CLM (blue line) and MÉRA (green line) (annual) TN90p (%) found from comparison of the years 1981–2017 with the base period 1961–1990 and using the algorithm outlined in Zhang *et al.* (2005) for "in-base" years (1981–1990). Also shown are linear fits for both COSMO-CLM (red dotted line) and MÉRA (black dotted line).

5.3.5 Warm spell duration index

The 1981–2017 annual count of warm spell days has been calculated for both COSMO-CLM and MÉRA. A warm spell is defined as a period of at least 6 consecutive days when TX > 90th percentile (from the reference period 1961–1990). As part of the calculation, the CDO command *eca_hwfi* was invoked, which has two outputs: WSDI and the number of warm spell periods (NWSPs) per time period (i.e. year). In Figure 5.23, the annual spatially averaged WSDI values for COSMO-CLM and MÉRA are presented.

5.3.6 Cold spell duration index

The 1981–2017 annual count of cold spell days has been calculated for both COSMO-CLM and MÉRA. A cold spell is defined as a period of at least 6 consecutive days when TN < 10th percentile (from the reference period 1961–1990). The CDO command *eca_cwfi* was utilised as part of the processing, which, like *eca_hwfi*, has two outputs: CSDI and the number of cold spell periods (NCSPs) per period (i.e. year). Figure 5.24 shows the annual spatially averaged CSDI values for COSMO-CLM and MÉRA.

5.4 Annual Precipitation Indices

A total of 13 annual precipitation indices were calculated for the period 1981–2015 utilising data from COSMO-CLM (1.5 km), WRF (2 km) and MÉRA (2.5 km). Each index requires different levels of processing and, in most cases, extant CDO commands were used. The datasets created are as follows: the simple precipitation intensity index (SDII); the annual count of days when precipitation is ≥ 1, 5, 10, 15, 20, 25 and 30 mm (R*N*mm); the maximum length of dry and wet spells [number of consecutive dry days (CDDs) and consecutive wet days (CWDs)]; the number of periods with more than 5 consecutive dry or wet days (CDDP/CWDP); and the annual total precipitation for wet days (PRCPTOT).

5.4.1 Simple precipitation intensity index

The SDII is based on the amount of precipitation (RR_{wj}) that occurs on wet days *w* (RR ≥ 1 mm) in period *j* and is given by the following formula:

$$SDII_j = \sum_{w=1}^{N} RR_{wj} / N \qquad (5.1)$$

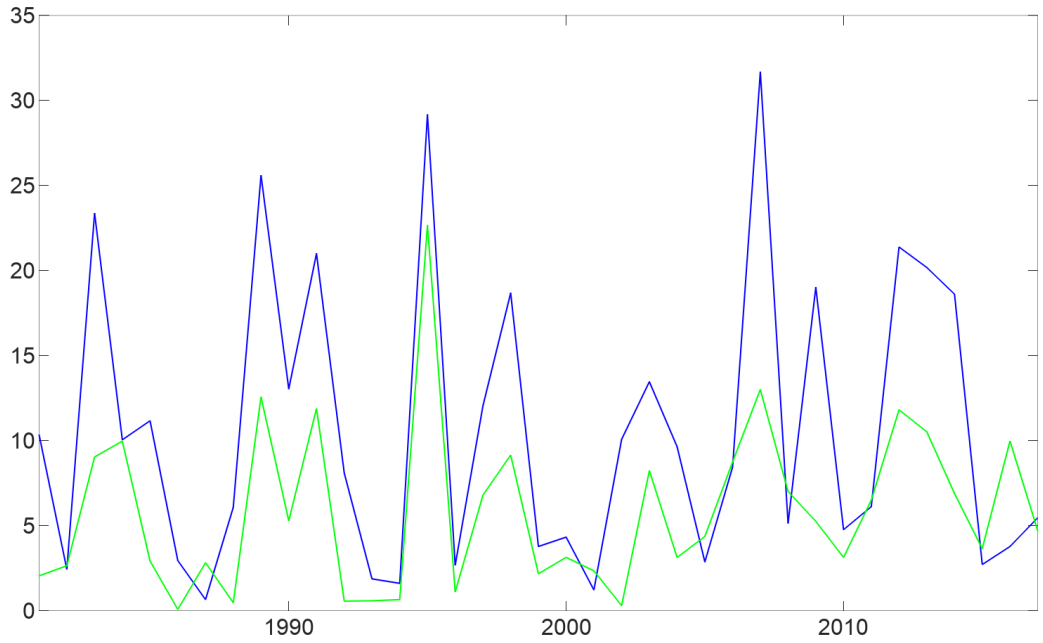


Figure 5.23. Spatially averaged COSMO-CLM (blue line) and MÉRA (green line) (annual) WSDI (days) found from comparison of the years 1981–2017 with the base period 1961–1990 and using the algorithm outlined in Zhang *et al.* (2005) for "in-base" years (1981–1990).

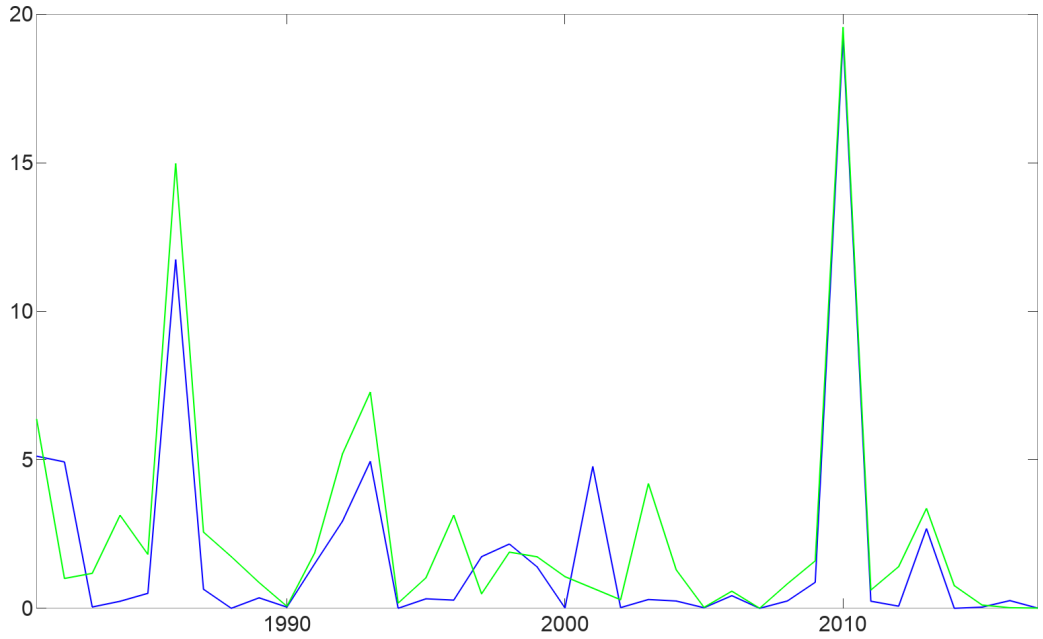


Figure 5.24. Spatially averaged COSMO-CLM (blue line) and MÉRA (green line) (annual) CSDI (days) found from comparison of the years 1981–2017 with the base period 1961–1990 and using the algorithm outlined in Zhang *et al.* (2005) for "in-base" years (1981–1990).

where N represents the number of wet days in period j. The SDII was calculated at monthly, seasonal and annual timescales for each of the three models (as well as from 1 km daily gridded observations from Met Éireann, for comparison) after preprocessing (i.e. the generation of monthly, seasonal and annual precipitation files) and running the CDO command *eca_sdii*. Examples of the (temporal and spatial) mean summer and winter SDII over the period 1981–2015 are given in Figures 5.25 and 5.26, respectively. In general, all three models showed a consistently higher summer SDII than the observations, with COSMO-CLM showing the highest SDII and WRF and MÉRA showing comparable SDIIs. For the winter SDII, the modelled and observed values were similar.

5.4.2 RNmm

The index R*N*mm describes the annual count of days when precipitation is greater than or equal to N mm, with N determined by meteorologically significant precipitation amounts. For instance, the values $N = 1$, 10 and 30 were used to determine the number of wet, heavy precipitation and extremely heavy precipitation days, respectively, and were calculated for each of the three models over the period 1981–2015 utilising the CDO command *eca_RR1*. Other values used were $N = 5$, 15, 20 and 25 mm. Figure 5.27 illustrates the R10mm values (temporal and spatial means) found for each of the three models. There is a clear east–west divide visible in each spatial map, with the west showing higher R10mm values than the east. R10mm values are also noticeably higher in mountainous regions. The three models display similar temporal patterns (Figure 5.27, bottom right panel) with no obvious overall trend; linear fits give slopes of $b = -0.024$ for COSMO-CLM, 0.041 for WRF and −0.005 for MÉRA. The results for R20mm (Figure 5.28) and R30mm (Figure 5.29) show a decrease in the number of days that fall into these categories (when compared with R10mm). The east–west pattern is no longer obvious, with mountainous areas accounting for the majority of the R20mm and R30mm results. Although weak, there appears to be a positive temporal trend for both R20mm (Figure 5.28, bottom right panel) and R30mm (Figure 5.29, bottom right panel). Linear fits show agreement in sign and (to a lesser extent) magnitude of this trend among all three models. The slopes found for R20mm were $b = 0.018$ for COSMO-CLM, 0.019 for WRF and 0.016 for MÉRA, while the values found for R30mm were $b = 0.013$ for COSMO-CLM, 0.008 for WRF and 0.011