



**HORIZON 2020
THEME SC5-2017**



European Climate Prediction system

(GRANT AGREEMENT 776613)

European Climate Prediction system (EUCP)

Deliverable D2.1

***Report on the evaluation and combination of potential
observational/emergent constraints relevant to European climate
projections***

Deliverable Title	<i>Report on the evaluation and combination of potential observational/emergent constraints relevant to European climate projections</i>	
Brief Description	<i>Evaluation and combination of potential observational/emergent constraints relevant to European climate projections</i>	
WP number	2	
Lead Beneficiary	<i>Reto Knutti, ETH Zurich</i>	
Contributors	<i>Lukas Brunner, ETHZ Andrew Ballinger, UEDIN Ben Booth, Met Office Glen Harris, Met Office Gabrielle Heger, UEDIN Reto Knutti, ETHZ Ruth Lorenz, ETHZ Tamzin Palmer, Met Office Said Qasmi, CNRM Aurelian Ribes, CNRM Carl McSweeney, Met Office Kasia Tokarska, ETHZ Sabine Undorf, UEDIN</i>	
Creation Date	<i>03/10/2019</i>	
Version Number	<i>4.0</i>	
Version Date	<i>25/11/2020</i>	
Deliverable Due Date	<i>30/11/2019</i>	
Actual Delivery Date	<i>25/11/2020</i>	
Nature of the Deliverable	<input checked="" type="checkbox"/> x <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>R – Report P - Prototype D - Demonstrator O - Other</i>
Dissemination Level/ Audience	<input checked="" type="checkbox"/> x <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>PU - Public PP - Restricted to other programme participants, including the Commission services RE - Restricted to a group specified by the consortium, including the Commission services CO - Confidential, only for members of the consortium, including the Commission services</i>

Version	Date	Modified by	Comments
0.1	17/10/2019	B. Booth	Included introductions and reworked 3.2
0.2	22/10/2019	L. Brunner	Freeze final draft (subject to minor changes only)
1.0	24/10/2019	L. Brunner	Final draft for review
2.0	27/11/2019	B Booth and L. Brunner	Review and update
3.0	28/11/2019	P Newton	Final review
4.0	25/11/2020	B Booth	Updated to reflect published paper

Contents

1. Executive Summary.....	5
2. Project Objectives	6
3. Overview	7
3.1. Deliverable Outline.....	7
3.2. Science overview	7
4. Core Task 2.1 activities	8
4.1. Selection of observational constraints from the literature	8
4.1.1. Locating analysis in D2.2 deliverable	8
4.1.2. Summary of CNRM activities	9
4.1.3. Emergent constraints in the upcoming CMIP6	11
4.1.4. Reassessing weighting diagnostics for CMIP6	12
4.2. Application of detection and attribution methods to derive constraints	14
4.2.1. Additional information	14
4.3. Impacts of the selected constraints on European climate	17
4.3.1. Additional information.....	17
4.4. Provision of information for task 2.2	21
5. Lessons Learnt, links built and challenges faced	23
6. References	24

1. Executive Summary

The purpose of D2.1 is for the management and funders of EUCP to have the ability to track completion of the tasks originally outlined in Task 2.1, but which are now reported within D2.2, along with the output of Task 2.2. It offers reassurance that the merging of these two tasks, which had good scientific grounds, has not led to any of the earlier tasks not being adequately addressed.

This deliverable was originally intended to report on the application and evaluation of constraints on a range of future climate projections using observations of historical climate (based on Task 2.1 activity). However, many of the results from these activities are now documented and discussed within a joint deliverable, D2.2, with Task 2.2.

The work performed and the main results can be found in the extended D2.2 deliverable. A summary of Task 2.1 activities, based around four themes, is given here:

- **Selection of observational constraints:** Observational constraints have been selected for all methods. They are detailed in the method description of each method (and briefly summarized in section 4.1). In this deliverable we also:
 - Summarize CNRS/CNRM's ongoing evaluation of further potential observational constraints (beyond current implementations)
 - Provide an early assessment of how CMIP6 may change the constraint picture
- **Application of detection and attribution methods to derive constraints:** Two approaches to isolate the human influence on climate have been implemented and are detailed in the extended D2.2 deliverable. Here we briefly note ongoing development of these methods.
- **Impact of selected observations on European climate:** Most of the important progress has been in this area and this is documented in the D2.2 deliverable
- **Provision of information to Task 2.2:** This activity has been completed for most groups and all methods have produced projections based on the selected constraints. CNRM has the ambition to further develop potential observational constraints, beyond these initial deliverables.

The motivation for the merger with D2.2 is outlined in more detail in the deliverable outline (Section 3.1) and lessons learnt (Section 5). To summarize: in practice it was not possible to isolate the impact of observational constraints without first implementing these observations within the climate projection methods and insights on the value of the observations were often method dependent. The extended deliverable avoids duplication of common information between the two deliverables D2.1 and D2.2 that would have resulted if they had not been merged.

2. Project Objectives

WITH THIS DELIVERABLE, EUCP HAS CONTRIBUTED TO THE ACHIEVEMENT OF THE FOLLOWING OBJECTIVES (DESCRIPTION OF ACTION, SECTION 1.1):

No.	Objective	Yes	No
1	Develop an ensembles climate prediction system based on high-resolution climate models for the European region for the near-term (~1-40 years)		no
2	Use the climate prediction system to produce consistent, authoritative and actionable climate information	yes	
3	Demonstrate the value of this climate prediction system through high impact extreme weather events in the near past and near future		no
4	Develop, and publish, methodologies, good practice and guidance for producing and using EUCP's authoritative climate predictions for 1-40 year timescales	yes	

3. Overview

3.1. Deliverable Outline

This deliverable focuses on the evaluation and combination of potential observational and/or emergent constraints relevant to European climate projections explored as part of the development and implementation of the climate projections (described in Deliverable D2.2).

As has been agreed with the European Commission, the core of the analysis that contributed to this activity will form part of an extended D2.2 deliverable. This is because much of the evaluation of observational constraints required their implementation within the projections (mainly to assess what impact or value a given observation or approach had in narrowing spread in future projected changes). This report is structured around the four main activities of Task 2.1 (which informs deliverable D2.1) as follows:

- Selection of constraints from the literature (Section 4.1)
- Application of detection and attribution methods to derive constraints (Section 4.2)
- Impact of constraints on the European regions (Section 4.3)
- Provision of constraints to Task 2.2 (Section 4.4)

Section 4.3 represents the activity that took most of the focus across the participating groups. The impacts of the constraints are at the core of the paper submitted (Brunner *et al.*, 2020).

3.2. Science overview

Deliverable D2.1 documents the progress towards the Task 2.1 objectives. Selections of observational constraints have been made for all the methods that contribute to WP2 (see method description document) and these methods have now been applied. For most groups, the emphasis is now shifting on to their application though there is some work to do by all groups in further understanding the role of the constraints that have been selected. The work at CNRM has the ambition to select and integrate further observational constraints beyond those that are current implemented in their method (method description document, Section B.4). For CNRM therefore, a continued active development in Task 2.1 is anticipated, to identify further observational constraints (we outline some of this activity in section 4.1.2).

The highlight of the work so far has been establishing the role of observations in narrowing the spread of future projected changes. The impact of these observational constraints can be seen by comparing the constrained and unconstrained distributions (see Brunner *et al.*, 2020). Some important findings have emerged from this work. In particular, observations tend to suggest that larger warmings responses are less likely, regardless of the method or choice of observation used across the different climate projection techniques. These are key results in the core paper that will be submitted in time to inform the IPCC AR6 report (deadline December 2019). These results are discussed and detailed in the D2.2 deliverable.

4. Core Task 2.1 activities

4.1. Selection of observational constraints from the literature

Most constraints have been selected and applied within their respective methods. These are briefly summarized here with details in the D2.2 deliverable. CNRM is the one group that is continuing to actively develop their method with a view of selecting further observational constraints. As such, we detail their progress in Section 4.1.2.

4.1.1. Locating analysis in D2.2 deliverable

The selected observational constraints are detailed in the method description document for each of the methods. Here we briefly summarize the nature of the observations selected by each Task 2.1 partner.

ETHZ. Observations used to constrain future projections in the context of model weighting are here referred to as diagnostics following, e.g., Brunner *et al.* (2019). An overview of weighting diagnostics in general and diagnostics used in the ETHZ method (ClimWIP) for Europe in particular, are discussed in Section B.1 of the method description document. As a short summary we here list the diagnostics selected and applied within deliverables 2.1 and 2.2 by ClimWIP: temperature (climatology), precipitation (climatology), shortwave downward radiation (climatology), shortwave upward radiation (climatology and variance), longwave downward radiation (variance). The selection process and related issues are also discussed in Brunner *et al.* (2019) and Lorenz *et al.* (2018).

UEDIN. The Detection and Attribution approach taken by Edinburgh, and explored as potential input in the CNRM method, relies on identifying a human fingerprint on climate using parallel historical experiments that isolate the responses to given forcings. Once identified, it is then possible to quantify what role these fingerprints had in historical observed climate records – which can be used to scale the human contributions to future changes. The approach adopted by Edinburgh is to use optimal fingerprinting, which is already well established in the literature (see method description document, Section B.3). This relies on identifying the European scale human pattern in historical temperature and precipitation (for projections of temperature and precipitation, respectively) and scaling the projected changes so that the future scaled response is consistent with the optimal fingerprint found in the observed historical changes. A detailed description can be found in section B.3 of the method description document.

CNRM. The CNRM method, as currently implemented, makes use of the observed warming during the historical period. This uses both global temperature responses and the temperature changes in the target region. They are continuing to explore the potential to combine these with process-based emergent constraints. The observed warming constraints have shown clear value in narrowing future climate projection spread (see Brunner *et al.*, 2020). The process-based emergent constraint aspect is still in development, and is documented below and in section B.4 in the method description document.

Met Office. The Met Office probabilistic method (section B.5 of method description document) makes use of a broad basket of observations. These fall into two categories: (i) observed estimates of a broad range of climatological variables (such as temperature, humidity, satellite radiation budget, ISCCP cloud amounts) and (ii) observed estimates of historical trends in surface temperature (Briganza estimates); ocean heat content and observed atmospheric CO₂.

4.1.2. Summary of CNRM activities

CNRM worked on providing a constrained probabilistic projection of future changes of the European climate by combining two sources of information: (i) the observed warming during the historical period and (ii) process-based emergent constraints, both using CMIP5. The main output of this technique is a constrained probability density function describing uncertainty on future changes, contributing to Tasks 2.1 and 2.2.

A constraint based on observed warming to date is implemented by considering both global and regional mean temperature. We (i) estimate the forced response of each CMIP model over the historical period, (ii) construct a multi-model distribution which characterizes the model uncertainty in this forced response, and then (iii) sub-select those trajectories which are consistent with available observations, given internal variability. Results obtained using this procedure are illustrated in Figure 1, where it is shown that this constraint leads to a clear reduction of projected future changes in European summer mean temperatures.

Several process-based metrics have been proposed in the literature to constrain future climate changes. This approach consists in identifying a relationship among the models between (i) a metric characterizing a physical process in present-day climate and (ii) the response of the variable of interest in projections. This relationship is based on the assumption that the physical processes existing in the present climate are also involved in driving the future response. Hence, adding the information provided by the observed metric allows the selection of the most realistic models in the present climate, which are likely to provide the most reliable climate projections.

Four metrics from literature have been investigated to constrain projections of European summer temperature. First, Boé and Terray (2014) found that models characterized by a limitation of evapotranspiration by soil moisture tend to simulate larger evapotranspiration decreases and consequently larger surface warming. Second, they also point out that models with a large present-day interannual anti-correlation between cloud cover and temperature over land tend to simulate larger future decrease in cloud cover and therefore a larger surface warming. Third, Cattiaux, Douville and Peings (2013)

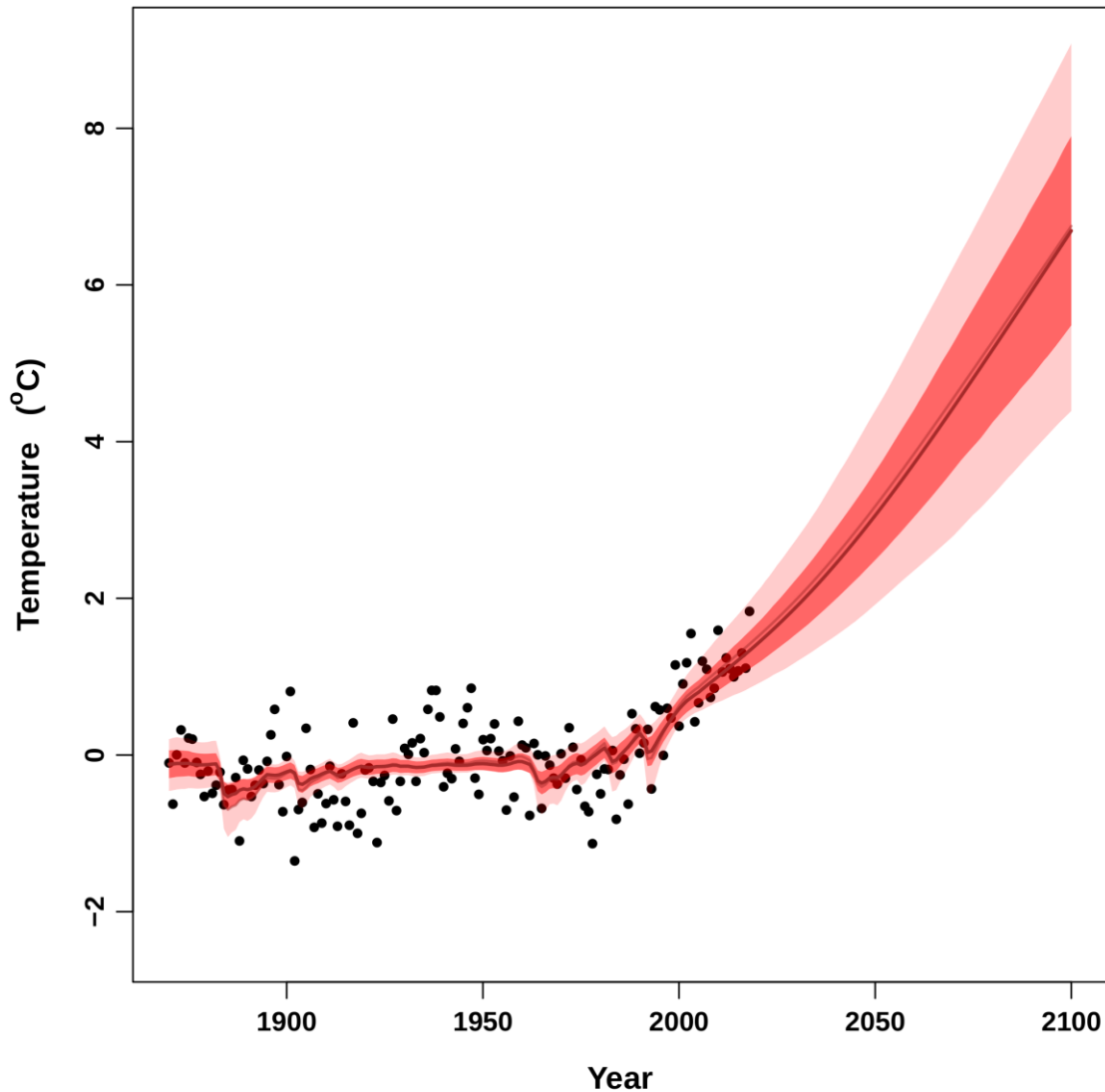


Figure 1 Observations of European summer mean temperature since 1850 (black points), compared to the multi-model distribution of the forced response of this variable to historical forcings and RCP8.5 scenario (multi-model mean in solid line, and 5-95% confidence region in light pink). Uncertainty in European summer mean temperature changes decreases after applying an observational constraint combining observations in European summertime temperature and those of global mean temperature since 1850 (constrained 5-95% confidence region in dark pink). Best estimates before (light brown) and after (brown) applying the observational constraints are almost indistinguishable in this case during the historical period, as observations are well consistent with the multi-model mean estimate. All values are temperature anomalies with respect to the 1870-2018 period.

suggest that CMIP5 models overestimate summer temperatures in Central Europe over the historical period and this may impact the forced response as well. Finally, we investigate the relationship between the temperature trend over the recent period (1979-2014) and temperature changes in projections, following Douville and Plazzotta (2017), who suggest that the projected midlatitude summer drying is underestimated by most CMIP5 models.

In order to quantify the respective contributions of the above-mentioned metrics in the temperature changes, a general multiple regression between temperature changes and the listed metrics has been calculated for 17 CMIP5 models. For each spatial grid point, we consider the “most relevant” metric (or predictor) as the one minimizing the Bayesian

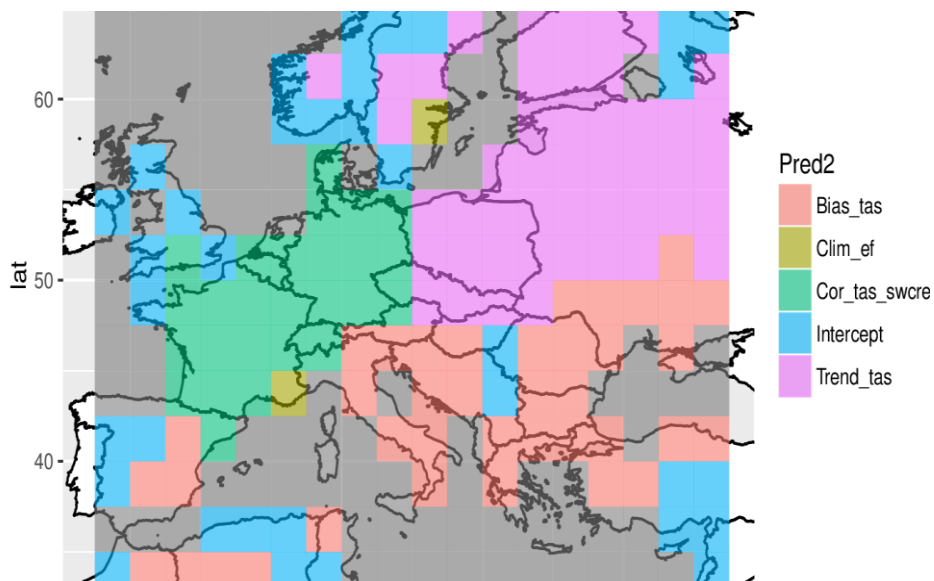


Figure 2 Regression between multi-model temperature changes (2041-2060 relative to 1995-2014) and different metrics (i.e. potential observational constraints) over the reference period: temperature bias (*bias_tas*), evaporative fraction (*Clim_ef*), correlation between shortwave cloud radiative effect and temperature (*Cor_tas_swcre*), temperature trend (*Trend_tas*). The most relevant predictor is shaded for each grid point. If no predictor is relevant, the associated grid point is shaded in blue (*Intercept*).

information criterion among all the possible statistical models (i.e., using all combinations of predictors). Results suggest that Europe is not uniformly sensitive to a single metric in terms of temperature changes (Figure 2). Several sub-regions are noticeable, e.g., western Europe more is sensitive to the role of clouds, while the Mediterranean basin is more sensitive to temperature bias.

Investigations are currently being done in order to effectively combine temperature time series (Figure 1) with the physically-based constraints (Figure 2), and assess the sensitivity of the statistical method to several intrinsic parameters: model error matrices (internal variability), number and strength of the physically-based constraints, estimation of the relationships between constraints and historical temperature/precipitations.

4.1.3. Emergent constraints in the upcoming CMIP6

First studies are being carried out based on the already available CMIP6 models. Motivated by findings that several of the new CMIP6 models have considerably higher climate sensitivity (both in terms of Equilibrium Climate Sensitivity and Transient Climate Response), hence showing more warming compared to CMIP5 (Gettelman *et al.*, 2019; Voltaire *et al.*, 2019; Voosen, 2019; K.B. Tokarska *et al.*, 2020b) look into the recent warming trend for the period 1981-2014 in CMIP5 and CMIP6. They find it to be strongly correlated with the transient climate response (TCR). This strong correlation (Figure 3a-c) can serve as an emergent constraint on the TCR. High TCR models have difficulties reproducing the observed warming trend. The observationally-constrained likely ranges of TCR estimate based on CMIP6, CMIP5 or both combined (Figure 3a-c blue rectangle; Figure 3d blue boxes) are consistent but substantially narrower than those reported by IPCC AR5 of 1.0 - 2.5 °C, regardless of the set

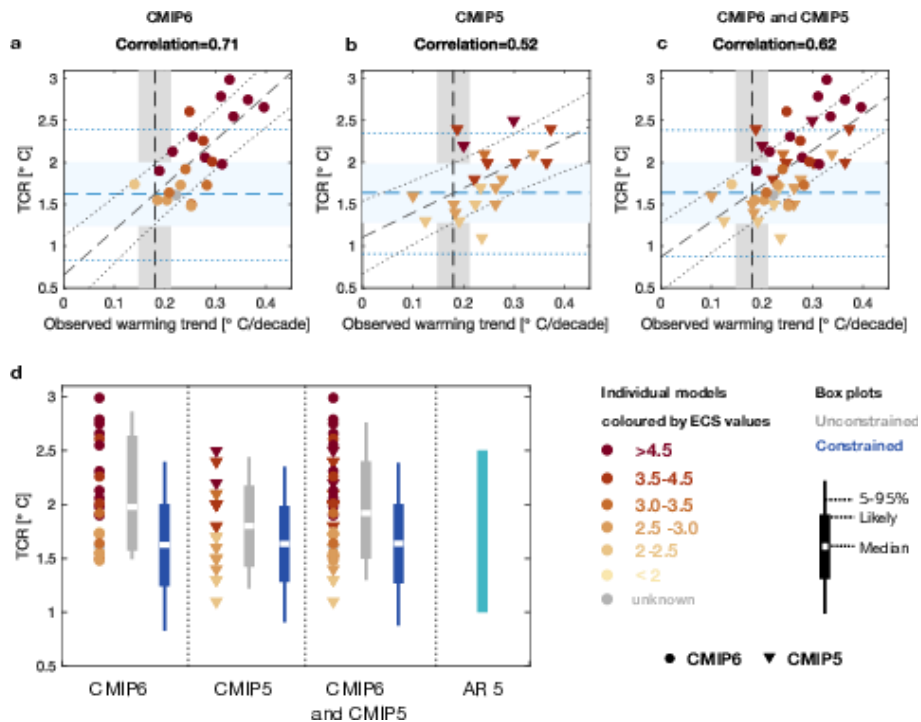


Figure 3: Correlation of the simulated warming trend for the period 1981-2014 with TCR. (a) based on CMIP6 models; (b) based on CMIP5 models; (c) based on the joint distribution of CMIP6 models (circles) and CMIP5 models (triangles). (d) Constrained and unconstrained ranges of TCR based on CMIP6 and CMIP5 models (following from panels a-c), compared with the IPCC AR5 likely range. Unconstrained ranges (gray box plots) are based on raw CMIP models, shown to the left of each box plot by individual dots. Constrained ranges (blue box plots) are based on the emergent constraint (as in top panels). The last box plot in panel (d) shows the IPCC AR5 likely (>66% probability; equivalent to 17-83% range) range. Each box plot shows 5-95% range, likely range, and median value, as illustrated in the legend. Note that the AR5 likely range is an assessment across multiple lines of evidence and studies and thus not fully comparable. Reprinted from K.B. Tokarska et al. (submitted-b) with friendly permission from the authors.

of models used. The two likely ranges (derived here and the IPCC one) are. Reduction in projected ranges indicate that more of the original range of projected TCRs can be excluded due to this constraint, however, these reduced ranges may not be fully comparable as different lines of evidence were combined in AR5 leading to a broader uncertainty range in that case. The observationally-constrained TCR likely range based on CMIP6 models alone, of 1.24 to 2.00 °C (“likely”, 17-83%) with a median of 1.63 °C, is narrower and lower than the raw CMIP6 likely range (of 1.57 to 2.64 °C, median 1.98 °C; Figure 2d grey CMIP6 bar). These results from observationally-constrained TCR range in CMIP6 are consistent with a recent median TCR estimate of 1.67 °C derived from CMIP5 models and a slightly earlier observational period (Jiménez-de-la-Cuesta and Mauritsen, 2019).

4.1.4. Reassessing weighting diagnostics for CMIP6

Due to the delay in the provision of many of the CMIP6 models all diagnostics used in the ETH method (referred to as ClimWIP in D2.1) are based on CMIP5 and have not yet been tested for CMIP6 in detail. However, an analysis of all available CMIP6 models is currently ongoing. This analysis is currently limited to only three temperature-based diagnostics (climatology,

variance and trend) to maximize the number of available models. It uses 40 CMIP5 models and the 13 CMIP6 models which currently provide temperature projections based on the new Socioeconomic Pathway 5-8.5 (SSP585) (Gidden *et al.*, 2019), which is comparable to the Representative Concentration Pathway 8.5 (RCP85) (van Vuuren *et al.*, 2011) used in CMIP5. The rest of the setup is equivalent to the one used in the work for D2.1 and D2.2 and which is detailed in Brunner *et al.* (2019).

Figure 4 shows the unconstrained and constrained time evolution of European summer (JJA) temperature for CMIP5 and CMIP6. It is evident that the unconstrained CMIP6 models available so far reach, on average, considerably higher temperatures than most CMIP5 simulations. Part of stronger warming is related to differences in experimental design (the radiative forcing is expected to differ between SSP585 used in CMIP6 and RCP85 used in CMIP5), but it has also been found that CMIP6 contains several models with considerably higher climate sensitivity compared to CMIP5 (Gettelman *et al.*, 2019; Voltaire *et al.*, 2019; Voosen, 2019). It will therefore be essential to establish if these high climate sensitivity models are consistent with observed changes in the real world using methods such as model weighting or emergent constraints (e.g., K.B. Tokarska *et al.*, 2020b). The preliminary weighting (based on only three temperature-based diagnostics) indeed leads to a downward shift of the unconstrained CMIP6 distribution for Europe indicating that confidence in the upper end of the distribution is lower.

As more and more CMIP6 models become available, results from these assessments are expected to feed into the next WP2 deliverable, D2.3. Planned future work includes comparing the effect of diagnostics between CMIP5 and CMIP6 as well as an analysis of diagnostics with most predictive skill in different regions and seasons.

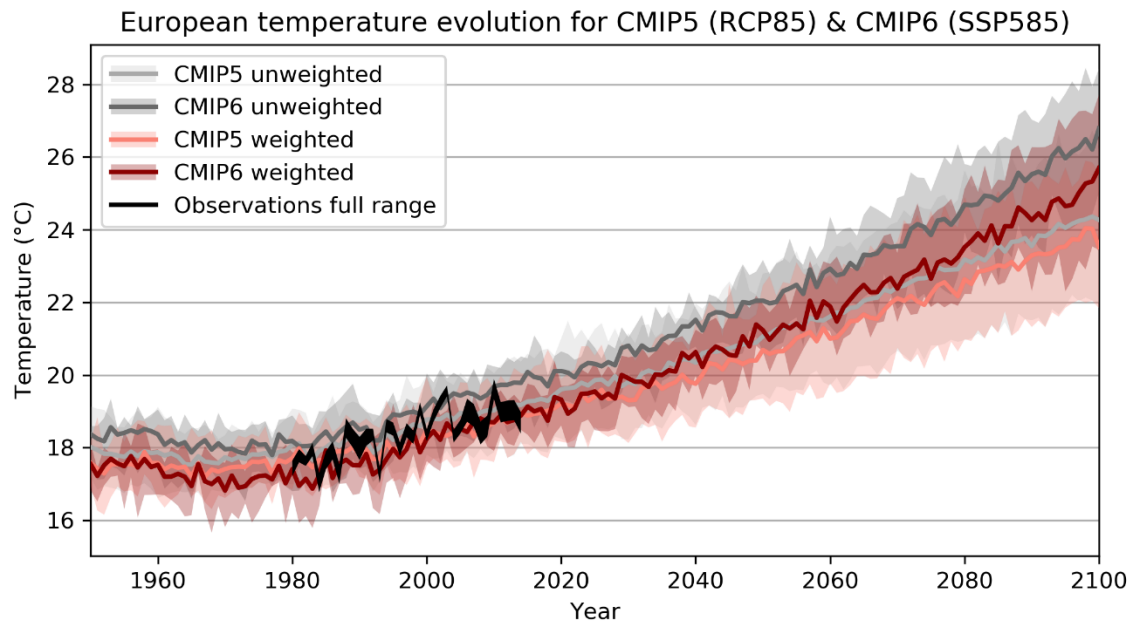


Figure 4: Temperature evolution in Europe for 40 CMIP5 models using RCP85 and 13 CMIP6 models using SSP585. Shown are the unweighted mean and interquartile (gray) and the weighted mean and interquartile (red) as well as an observational estimate based on E-OBS, ERA-Interim, and MERRA2. For more details on the calculation see Brunner *et al.* (2019).

4.2. Application of detection and attribution methods to derive constraints

Detection and attribution techniques have been applied to derive constraints from observed climate change by UEDIN. A summary of literature and a description of the methods can be found in section 2.2 of Brunner *et al.* (2020), as well as in the method description document (section B.3).

CNRM are also pursuing further observational constraints that could be combined with their methods. These include use of idealized single forcing simulations that isolate the human impact on past climate by taking into account each external forcing separately (natural and anthropogenic aerosols, greenhouse gases, stratospheric ozone, which is at the core of the Detection and Attribution approach), but this has not yet arrived at information relevant for the European scale. As such, much of their contribution to date has focused on other observations and historical simulations, which include the effect of all combined external forcings, that have so far shown promise with their method.

4.2.1. Additional information

The UEDIN methodology is being further developed, beyond that presented in the D2.2 deliverable. Here we summarize progress to (i) extend this to wintertime precipitation, (ii) the challenge of how to account for internal variability, and (iii) the rationale behind the choice of separating either the greenhouse gas signal from other forcings or the combined anthropogenic signal from the other forcings (which is used in the projections

in Brunner *et al.* (2020)), rather than separating the three main drivers of climate (the response to natural forcing, the response to greenhouse gases, and the response to other anthropogenic forcing).

- i. UEDIN methodology provides constraints on temperature as well as precipitation for summer. Focus is switching, now, to wintertime constraints where, for precipitation in particular, attention is needed to separate out the NAO role. The ASK contribution is important here, as some publications (e.g. Zhang *et al.*, 2007; Polson *et al.*, 2016) indicate larger changes in observed rainfall than simulated by the model mean, hence the ASK method that allows fitting the magnitude of response to observations even outside the climate model range is very valuable here.
- ii. An aspect of the UEDIN methodology that is still under discussion is the how projections from this method could usefully be compared with other methods. Projected climate will both be a reflection of uncertainty in the projection by the climate change signal uncertainty only, which is the target of the ASK method, and a reflection of uncertainty due to internal climate variability in the future. Currently this discussion is how to include estimates of internal variability so that results from the UEDIN ASK methodology can be compared directly with other methods. The challenge is to assess to what extent internal variability is already partly or wholly aliased within the results from the other methodologies, already, and how to best add it to the overall constraints.
- iii. There are a number of defensible choices that can be made in identifying human signal in past climate. Therefore, a challenging subject is the choice of signals to fit in the ASK method. The most robust method fits three signals to the observations: the response to natural forcing, the response to greenhouse gases, and the response to other anthropogenic forcing (see e.g. Stott and Kettleborough, 2002). Then, in the future, the constrained response to greenhouse gases and other anthropogenic forcing can be combined to a constraint (see also Shiogama *et al.*, 2016). However, presently CMIP6 future single forcing runs are only slowly becoming available and fitting three signals to European changes is too noisy for a useable constraint, as it is even challenging on the global scale (Schurer *et al.*, 2018; Tokarska *et al.*, 2019). Hence the two approaches chosen (see section 2.2 of Brunner *et al.* (2020)). The first one where greenhouse gas responses are separated from the response to all other forcings, with the assumption that the combination in models of aerosols, and natural forcings is approximately of the right size relative to each other. This approach has been explored further recently in Tokarska *et al.* (2019) to estimate greenhouse gas contributions to ocean and global surface temperature warming, with sensitivity tests to using anthropogenic against greenhouse gas only signals, supporting the method chosen here. Another assumption is that of linearity, which has been found to be approximately appropriate (Gillett *et al.*, 2004) and nonlinearity is likely indistinguishable due to high levels of internal climate

variability. The second approach is where anthropogenic responses are separated from the response of all other forcings, with the assumption that the combination of greenhouse gases and aerosols remains approximately the same in the future as they did in the past. Whilst this assumption becomes more questionable towards the end of this century, it may hold reasonably well in the near term, 10-40 years, time horizons of EUCP. Commonalities and differences emerge from these two different approaches (see Brunner *et al.*, 2020).

4.3. Impacts of the selected constraints on European climate

A large part of the Task 2.1 activity has focused on assessing the value and impact of the selected constraints. The presentation of the projections (see Brunner *et al.*, 2020) presents both unconstrained and constrained projections. Differences between these two presentations illustrate the impact of constraints that have been selected for each method. The experimental design adopted to isolate these impacts is motivated and detailed in section 3 of Brunner *et al.*, 2020). The design includes common variables, regions, and time periods as well as a coordinated processing order, in order to facilitate a clean assessment of the role of constraints in the respective methods. Figures 2-6 in Brunner *et al.* (2020) show the effect of constraining compared to the unconstrained distributions and a detailed analysis and discussion is given in section 4 and 5.

4.3.1. Additional information

In addition to the work done to isolate the role of observations in the respective climate projection methods, the **Met Office** have undertaken work to assess the impact of removing simulations that are unable to capture climate processes that are thought to be a key driver of present-day climate. This screening of simulations based on whether they can replicate observed climate processes can be seen as complementary to the probabilistic approaches presented in D2.2 but, as we will come on to discuss, insights from this approach can reach similar conclusions to the probabilistic estimates.

This method is based on the assumption that if a model is unable to reproduce the key factors important for determining the regional climate, the projections from this model are not considered reliable. Each model in the CMIP5 ensemble (where data is available), is firstly assessed against these key performance indicators and poor performers eliminated from the selection. Several models also share large portions of code and therefore have similar errors and projections, Sanderson *et al.* 2015a and 2015b quantifies these similarities. Here we use these two papers to identify and remove 'near-neighbours' and further reduce the selection. This assessment has been carried out previously focusing on processes and performance relevant for UK climate specifically (McSweeney *et al.*, 2018). Here the applicability of this sub-selection of models is assessed for the wider European area.

For a set of UK climate projections created in 2018 (UKCP18) a subset of realizations was selected from a larger ensemble using performance criteria to determine the capability of the models in representing the main physical processes that determine the climate of the UK. The method is described in full in McSweeney *et al.* (2018), following similar principles to McSweeney *et al.* (2015). 31 CMIP5 models with sufficient data availability were screened for

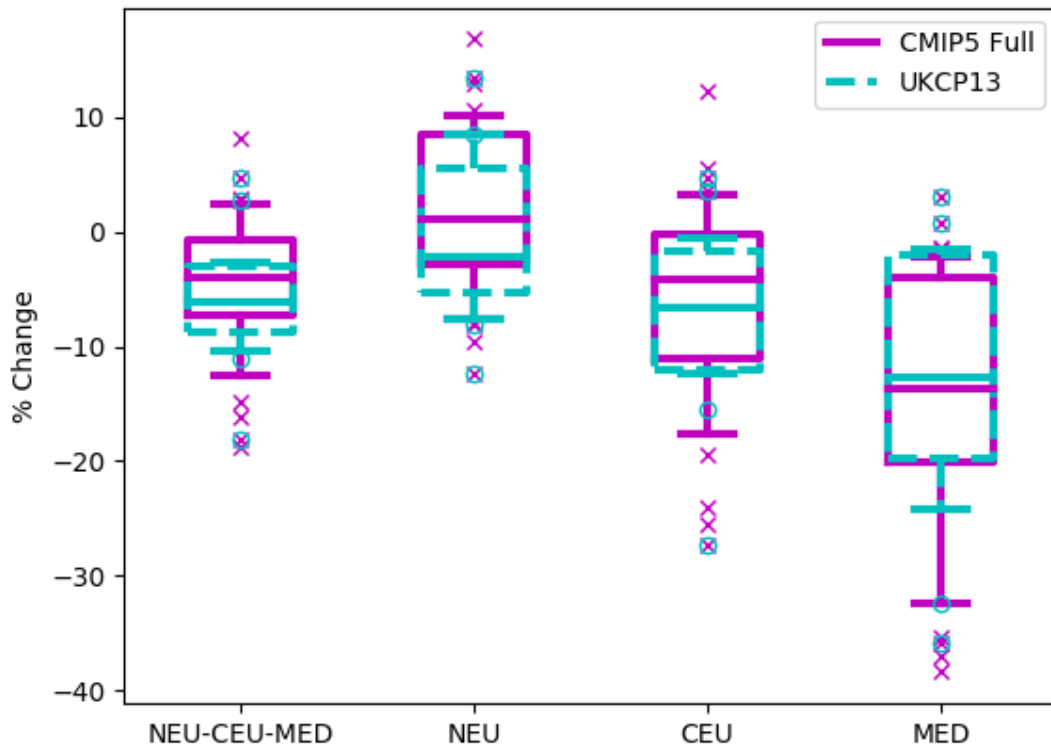


Figure 5: Relative change in JJA average precipitation for the time periods 1995-2014 and 2041-2061. Whiskers show 10th and 90th percentiles.

their performance against relevant regional and global performance criteria and near neighbours in order to identify a subset of 12 -13 members.

Firstly, all models are screened qualitatively for significant performance shortcomings at global and regional scale that are judged to significantly compromise the usefulness of those model projections for regional climate change studies relevant to the UK. The use of global information reflects our assumption that global performance provides evidence relating to the general plausibility of the physical assumptions built into a climate model, while regional analysis indicates the extent to which a combination of remote and local drivers of error might compromise credibility specifically for UK applications.

The regional criteria that the models were assessed against included the climatological circulation patterns of the N. Atlantic/European sector, distribution of daily storm track latitudes, mean temperature biases, frequencies of daily weather types, blocking frequency and the realism of the AMOC. Global criteria included global mean climate variability and drift, and SST errors. A summary of the assessment criteria and a qualitative analysis of each model's performance is shown in Table 1. Three models were found to be too unrealistic in their representation of key characteristics of the regional climate to provide useful projections for the UK. These were IPSL-CM5B-LR (UK affected by a cool bias of 8-9°C and very unrealistic summer circulation), FGOALS-g2 (very unrealistic circulation patterns in both summer and winter) and MIROC5 (unrealistic summer circulation, lacking westerly flow

direction over UK). A further six models were rejected due to very poor performance across multiple regional and/or global criteria.

Table 1: (taken from McSweeney et al., 2018) shows models excluded for poor performance in red. From remaining models (black and green), 13 were selected (green). The subset of 13 models used (green) is could equally be substituted with models indicated in black from the same box. Here we make use of the subset used in McSweeney et al. 2018 (green) to illustrate the impact on wider Europe: ACCESS1-3, bcc-csm1-1, CCSM4, CESM1-BGC, CanESM2, CMCC-CM, CNRM-CM5, EC-Earth, GFDL-ESM2G, HadGEM2-ES, IPSL-CM5A-MR, MPI-ESM-MR, MRI-CGCM3.

<p>bcc-csm1-1 bcc-csm1-1-m</p>	<p>BNU-ESM CCSM4 CESM1-BGC CESM1-CAM5 Nor-ESM1-M</p>	<p>CanESM2</p>	<p>CMCC-CESM CMCC-CM CMCC-CMS</p>	<p>CNRM-CM5</p>
<p>IPSL-CM5A-LR IPSL-CM5A-MR IPSL-CM5B-LR</p>	<p>MIROC5 MIROC-ESM MIROC-ESM-CHEM</p>	<p>FGOALS-g2</p>	<p>MPI-ESM-LR MPI-ESM-MR</p>	<p>MRI-CGCM3</p>
<p>EC-EARTH</p>	<p>inmcm4</p>	<p>CSIRO-Mk3-6-0</p>	<p>GFDL-CM3 GFDL-ESM2G GFDL-ESM2M</p>	<p>ACCESS1-0 ACCESS1-3 HadGEM2-CC HadGEM2-ES</p>

The remaining models were screened for ‘near-neighbours’ which share significant portions of code and are known to generate simulations with similar error and projection characteristics. This clustering by model similarity is intended to aid sub-selection of a single model from each cluster (or two, where four or more non-rejected models exist within a cluster). Some pairs/groups of models within the CMIP5 ensemble are known to share significant parts of their code, and therefore also share error and projection characteristics (Knutti and Sedláček, 2013; Sanderson, Knutti and Caldwell, 2015a, 2015b). Sanderson, Knutti and Caldwell (2015b, 2015a) use a multivariate metric of present-day climatology to quantify similarity between pairs of models from the CMIP5 (and CMIP3) ensemble showing clear relationships between some groups of models. These groups of ‘near neighbours’ include sets of models from particular centres, but also highlighting some groups of models that come from different centres but still share significant portions of code, which may not be so easy to identify without this type of analysis. Sanderson, Knutti and Caldwell (2015b) extend this analysis to future projections, demonstrating that the degree of dependence between models in their present-day climatology applies similarly to projection characteristics.

We would expect that the physical processes important for the UK climate (which McSweeney et al., 2018 focused on) to also be relevant for wider European climatic regions that are also impacted by these processes. Here we assess the potential wider applicability of this method by comparing the impact of filtering and sub-selection on the SREX regions used in the EUCP projections.

The sub-selection of CMIP5 ensemble members selected for UKCP18 (hereafter called CMIP13) was applied to the EUCP regions to see what effect the criteria used for selection would have on the projection range of precipitation and temperature for Europe. The change in average near surface temperature was calculated for each model out of 34 in a standard CMIP5 ensemble, for the time periods 1995-2014 and 2041-2061. The relative change in the

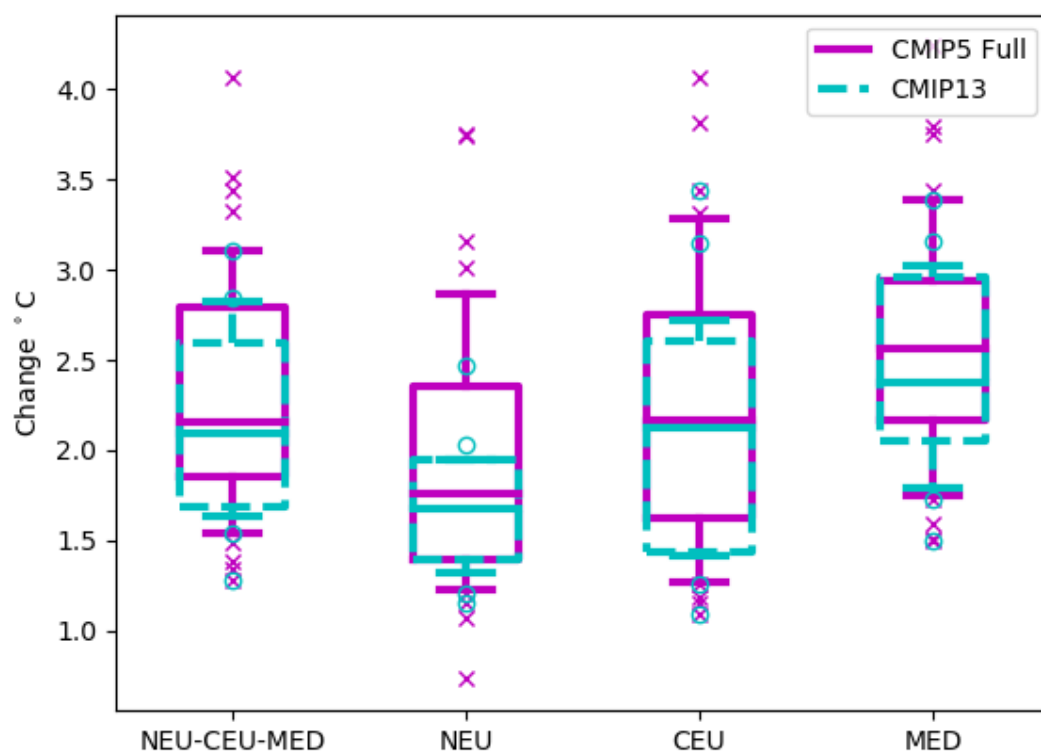


Figure 6: Change in JJA average near surface temperature for the time periods 1995-2014 and 2041. Whiskers show 10th and 90th percentiles.

precipitation was also calculated using 32 models, the EC-EARTH and NorESMI-M models were not included in this ensemble due to lack of data availability.

Figure 5 shows the inter-quartile and 10th to 90th percentile range for the CMIP5 ensemble and the CMIP13 ensemble for each of the EUCP regions along with the combined regions. The projection range is substantially reduced in the inter-quartile range for the CMIP13 ensemble compared to the standard CMIP5 ensemble, showing a decrease in the projected precipitation. The results for the three separate regions show that this reduction in the projected range and decrease in precipitation relative to the standard ensemble occurs in the Northern European region. In the Central and Mediterranean regions there is some reduction in the range of the 10th and 90th percentile but the inter-quartile range is largely unchanged and there is no reduction in the Mediterranean.

Figure 6 shows the projection range of average near surface temperature change for the standard CMIP5 ensemble and CMIP13. The main result of the constrained ensemble is to substantially reduce the projected temperature in the inter-quartile range for the Northern European region. The reduction for the Central and Mediterranean regions is not large and there is only a small decrease in the 90th percentile range in the Mediterranean. It can be noted that the change in the median projection is not substantially changed by CMIP13 for any of the results.

Processes, such as Atlantic circulation, North Atlantic storm tracks, and the AMOC impact the UK and Northern parts of the European climate but are less critical factors in determining more central areas and the Mediterranean. The results show that a sub-selection of the ensemble has the effect of constraining regional projections of climate parameters provided the appropriate climate processes are selected to assess the models for a given region. These results suggest that the existing assessment of poor physical drivers (McSweeney *et al.*, 2018) has wider applicability to Northern Europe, but different physical drivers need to be assessed for the Central European and Mediterranean regions. This lack of constraints in Central Europe and the Mediterranean motivate the need to repeat this form of analysis with a focus on key climate mechanisms in these regions. Further work will involve identifying what these physical drivers are for Central Europe and the Mediterranean and an assessment of how well these are represented in the CMIP6 ensemble.

Perhaps the most surprising outcome from this approach is that the CMIP13 constrained range down weights the warmer and wetter end of the climate projection range, particularly in Northern Europe. This is similar to the effect of observational weighting (see D2.2) in the formal probabilistic approaches, despite starting from very different premises, observations and methodologies. Further work is needed to fully understand this, but it does lend supporting evidence to the probabilistic approaches suggesting that both poor model screening and probabilistic approaches are down weighting models due to poor underlying processes.

4.4. Provision of information for task 2.2

All participating groups have made good progress in implementing their observations within climate projection frameworks produced within Task 2.2. For most groups, this process has largely been completed. CNRM are involved in a more ambitious assessment of potential observations (see section 4.1.2) and we expect this to further influence the implementation by Task 2.2, beyond the timing of this deliverable.

Table 2: Table of qualitative performance flags: Red - not fit to provide useful projections, Orange - significant bias/error, Yellow - relatively poor performance, Grey - no data (modified from McSweeney et al.,2018)

	Regional Criteria								Global Criteria		
	NE Atlantic Weather Types	Annual mean blocking frequency: IPCC	Mean DJF circulation	Mean JJA circulation	Mean temp bias	Atlantic SST bias	AMOC	Storm Track	Global temperature variability/ drift	Remote SST biases	High latitude temperature variability
bcc-csm1-1			Yellow				Grey	Orange			
bcc-csm1-1-m	Yellow	Orange	Yellow	Yellow		Yellow	Grey	Orange			Grey
BNU-ESM	Yellow	Red	Orange				Grey	Orange			
CanESM2											
CESM1-BGC		Grey									Grey
CMCC-CESM				Yellow			Orange	Grey			Grey
CMCC-CM		Grey					Orange	Grey			Grey
CMCC-CMS							Orange	Grey			Grey
CNRM-CM5				Yellow		Yellow				Yellow	
ACCESS1-0		Grey									
ACCESS1-3		Grey									Grey
CSIRO-Mk3-6-0		Grey	Orange				Grey			(CT)	
EC-EARTH							Grey				Grey
inmcm4	Yellow						Orange		Orange	(CT & SO)	
IPSL-CM5A-LR	Yellow			Yellow			Orange				
IPSL-CM5A-MR							Orange				
IPSL-CM5B-LR	Orange			Red	Red		Orange		Orange		
FGOALS-g2	Orange	Orange	Red	Red	Red		Orange	Red			Orange
MIROC5	Yellow		Orange	Red						Orange	
MIROC-ESM	Yellow	Orange	Orange	Yellow				Orange	Orange		
MIROC-ESM-CHEM	Yellow		Orange	Yellow				Orange	Orange		
HadGEM2-CC										Yellow	
HadGEM2-ES		Grey								Yellow	
MPI-ESM-LR											Grey
MPI-ESM-MR											Grey
MRI-CGCM3			Yellow				Orange			Yellow	
CCSM4			Yellow								
Nor-ESM1-M							Orange				
GFDL-CM3		Grey							Orange	Yellow	
GFDL-ESM2G		Grey	Grey							Yellow	
GFDL-ESM2M										Orange	
CESM1-CAM5		Grey						Grey	Grey		Grey

5. Lessons Learnt, links built and challenges faced

For the main discussion see corresponding section in Deliverable 2.2. We will only duplicate discussion on such lesson learned, here, as it relates to the motivation for merging the deliverables from the two tasks (Task 2.1 and Task 2.2) and address the situation at KNMI which directly affecting Task 2.1.

The original proposal had the expectation that the activities focused on identifying useful observational constraints could be a distinct and self-contained activity (Task 2.1). Work was anticipated to start earlier and then expected to inform the development and production of observational constrained estimates of future change within particular methodologies (in Task 2.2). However, we quickly found that these two activities are much more closely coupled than we had anticipated at the proposal stage. Particularly as research moved beyond questions of whether particular model biases were correlated with the range of future changes, we found that assessment of the impact of particular observations required the observations to be implemented within a particular methodology.

As a consequence, we have combined much of the Task 2.1 analysis into an extended D2.2 deliverable. We have also learnt lessons about the limits of how transferable insights are on observational constraints between methods. Philosophical and practical choices made by individual projection methodologies often closely tied insights of the value of particular observations to that methodology. For example, the detection and attribution methodology, adopted by UEDIN, relies on climate simulations which also had parallel single forcing historical simulations to estimate the human fingerprint in observed climate. Insights from the scaling parameters in that methodology (see section B.3 in the method description document) are not directly applicable to other methods. Instead much of the focus has been on quantifying and understanding where different sets of methodologies and observational constraints contribute to robust impacts on projected changes. The choice of observational constraints has been shown to often be tied intrinsically to the methodologies employed. The valuable outcome, therefore, has not been two distinct assessments (firstly on the most useful observational constraints and secondly on their implementation) but rather the comparison of methods and their observational constraints in terms of their combined impact on future projected changes. This is reflected in the merged nature of the deliverable, D2.2.

KNMI originally planned to contribute to Task 2.1 with a method assessing the credibility of the summertime temperature response in the Mediterranean area from present-day climate variability, in relation to the strong soil drying, cloud and precipitation feedbacks occurring in climate models. During the reporting period, however, this work has been put on hold due to long-term illness of the person responsible, and no replacement could be found at KNMI. The delay of this work has only small consequences for the overall WP2 progress as the other methods also deal with the same problem, yet with different approaches. Instead, within WP2, KNMI targeted their efforts on Task 2.2, with a more elaborate analysis of internal variability in the context of the PDF/UQs provided. In addition, work on building dry summer scenario within Task 2.3 has been progressed. Also, KNMI has built strong links with the eScience centre to port part of the KNMI scenario procedure to a common analysis platform.

6. References

- Boé, J. and Terray, L. (2014) 'Land-sea contrast, soil-atmosphere and cloud-temperature interactions: Interplays and roles in future summer European climate change', *Climate Dynamics*, 42(3–4), pp. 683–699. doi: 10.1007/s00382-013-1868-8.
- Brunner, L. *et al.* (2019) 'Quantifying uncertainty in European climate projections using combined performance-independence weighting', *Environmental Research Letters*. doi: 10.1088/1748-9326/ab492f.
- Brunner, L. *et al.* (2020) 'Comparing methods to constrain future European climate projections using a consistent framework'. *Journal of Climate*. 33 (20): 8671–8692. doi: 10.1175/JCLI-D-19-0953.1
- Cattiaux, J., Douville, H. and Peings, Y. (2013) 'European temperatures in CMIP5: origins of present-day biases and future uncertainties', *Climate Dynamics*, 41(11–12), pp. 2889–2907. doi: 10.1007/s00382-013-1731-y.
- Douville, H. and Plazzotta, M. (2017) 'Midlatitude Summer Drying: An Underestimated Threat in CMIP5 Models?', *Geophysical Research Letters*, 44(19), pp. 9967–9975. doi: 10.1002/2017GL075353.
- Gettelman, A. *et al.* (2019) 'High Climate Sensitivity in the Community Earth System Model Version 2 (CESM2)', *Geophysical Research Letters*, 46(14), pp. 8329–8337. doi: 10.1029/2019gl083978.
- Gidden, M. J. *et al.* (2019) 'Global emissions pathways under different socioeconomic scenarios for use in CMIP6: A dataset of harmonized emissions trajectories through the end of the century', *Geoscientific Model Development*, 12(4), pp. 1443–1475. doi: 10.5194/gmd-12-1443-2019.
- Gillett, N. P. *et al.* (2004) 'Testing the linearity of the response to combined greenhouse gas and sulfate aerosol forcing', *Geophysical Research Letters*, 31(14), pp. 2–5. doi: 10.1029/2004GL020111.
- Jiménez-de-la-Cuesta, D. and Mauritsen, T. (2019) 'Emergent constraints on Earth's transient and equilibrium response to doubled CO₂ from post-1970s global warming', *Nature Geoscience*, 2015. doi: 10.1038/s41561-019-0463-y.
- Knutti, R. and Sedláček, J. (2013) 'Robustness and uncertainties in the new CMIP5 climate model projections', *Nature Climate Change*. Nature Publishing Group, 3(4), pp. 369–373. doi: 10.1038/nclimate1716.
- Lorenz, R. *et al.* (2018) 'Prospects and caveats of weighting climate models for summer maximum temperature projections over North America', *Journal of Geophysical Research: Atmospheres*. doi: 10.1029/2017JD027992.
- McSweeney, C. *et al.* (2018) 'Selection of CMIP5 members to augment a perturbed-parameter ensemble of global realisations of future climate for the UKCP18 scenarios'. Hadley Centre Technical Note: 102.
- McSweeney, C. F. *et al.* (2015) 'Selecting CMIP5 GCMs for downscaling over multiple regions', *Climate Dynamics*. Springer Berlin Heidelberg, 44(11–12), pp. 3237–3260. doi: 10.1007/s00382-014-2418-8.

- Sanderson, B. M., Knutti, R. and Caldwell, P. (2015a) 'A representative democracy to reduce interdependency in a multimodel ensemble', *Journal of Climate*, 28(13), pp. 5171–5194. doi: 10.1175/JCLI-D-14-00362.1.
- Sanderson, B. M., Knutti, R. and Caldwell, P. (2015b) 'Addressing interdependency in a multimodel ensemble by interpolation of model properties', *Journal of Climate*, 28(13), pp. 5150–5170. doi: 10.1175/JCLI-D-14-00361.1.
- Schurer, A. *et al.* (2018) 'Estimating the transient climate response from observed warming', *Journal of Climate*, 31(20), pp. 8645–8663. doi: 10.1175/JCLI-D-17-0717.1.
- Shiogama, H. *et al.* (2016) 'Predicting future uncertainty constraints on global warming projections', *Scientific Reports*. Nature Publishing Group, 6(January). doi: 10.1038/srep18903.
- Stott, P. A. and Kettleborough, J. A. (2002) 'Origins and estimates of uncertainty in predictions of twenty-first century temperature rise', *Nature*, 416(6882), pp. 723–726. doi: 10.1038/416723a.
- Tokarska, K. B. *et al.* (2019) 'Quantifying human contributions to past and future ocean warming and thermosteric sea level rise', *Environmental Research Letters*. IOP Publishing, 14(7), p. 074020. doi: 10.1088/1748-9326/ab23c1.
- Tokarska, K. B. *et al.* (2020) 'Observational constraints on the effective climate sensitivity due to greenhouse-gas forcing'. *Environmental Research Letters* 15 034043 doi: 10.1088/1748-9326/ab738f
- Tokarska, K.B. *et al.* (2020) 'Past warming trend constrains future warming in CMIP6 models'. *Science Advances*, 6(12), eaaz9549. doi: 10.1126/sciadv.aaz9549
- Voldoire, A. *et al.* (2019) 'Evaluation of CMIP6 DECK Experiments With CNRM-CM6-1', *Journal of Advances in Modeling Earth Systems*. doi: 10.1029/2019MS001683.
- Voosen, P. (2019) 'New climate models predict a warming surge', *Science*. doi: 10.1126/science.aax7217.
- van Vuuren, D. P. *et al.* (2011) 'The representative concentration pathways: An overview', *Climatic Change*, 109(1), pp. 5–31. doi: 10.1007/s10584-011-0148-z.