Chapter 7

Probabilistic Projections and Wider Uncertainty

Authors: David Sexton[1], Tamara Janes[1]

Met Office internal reviewers: Glen Harris[1] and James Murphy[1]

1 - Met Office, Exeter, UK

Contents

7.1. Introduction

Society needs to make plans to adapt to the level of climate change that is expected to happen in response to the increased amounts of greenhouse gases and pollution that mankind has put into the atmosphere since pre-industrial times. Any decision-making framework that aims to make adaptation plans resilient to a range of plausible climate changes needs to account for the key uncertainties. One such uncertainty is associated with information about future climate change based on computer simulations. The principal sources of climate change uncertainty are (i) the range of possible emissions scenarios, (ii) externally- and internally-generated climate variability, on a range of time scales, and (iii) modelling uncertainty arising from an incomplete understanding and the approximate representation by climate models of how the concentrations or emissions are converted to radiative forcing, and how the climate system responds to this forcing.

This project provides a new set of plausible realisations for Singapore and SE Asia, which are designed and generated in Chapters 3, 4 and 5. They sample climate variability, two scenarios (RCP4.5 and RCP8.5), and are driven by a small number of CMIP5 GCMs to account for modelling uncertainty. These simulations therefore partially sample the three main sources of uncertainty described above, and prove to be a valuable tool for stakeholders. This is because they provide information that is spatially, temporally and physically consistent on the spatial and time scales that are useful for impacts studies and relevant to adaptation problems. A disadvantage is that limitations in the experimental design, including the small sample size of the simulations, implies that they do not fully represent the range of plausible future climate outcomes consistent with current knowledge, particularly in relation to the implications of modelling uncertainty. For example, this “wider uncertainty” can arise because different ensembles of coupled ocean-atmosphere GCMs can give different envelopes of future change for key variables of interest to users (e.g. Murphy et al. 2014). Another example is that ensembles of earth system models featuring an interactive carbon cycle give different ranges of change to ensembles relying on a single prescribed pathway for future CO_2 concentrations, and representing only uncertainties in physical climate feedbacks (e.g. Booth et al. 2012; Murphy et al. 2014).

It is important that stakeholders are aware of the extent to which the plausible realisations span this wider uncertainty, so that they can make more reliable decisions. In this study this can be done in two ways. Firstly it is possible to quantify key known uncertainties using a probability density function (PDF), which is a rigorous synthesis of multiple lines of evidence (ensembles of climate simulations, observations, and expert choices) based on a Bayesian framework. This was done in 2009 for the UK (UKCP09; Murphy et al 2009) using variants of the climate model, HadCM3, hereafter called HadCM3Q variants. For the Singapore projections presented in the project, a similar method can be applied, with an essential modification to use the regional climate projections generated in Chapter 5, rather than the European regional model simulations used in UKCP09. In practice the method requires assumptions to be checked on a variable by variable basis. If these requirements are satisfied, then the PDFs have the advantage that they can be used to assess the risk associated with various levels of climate change, and this can be useful in decision making. However, if the assumptions are invalidated then a simpler alternative presentation of the wider uncertainty is required. This simplified information, which is typically a plausible range of climate change, should still be included in the decision-making process; it just has to be done in

a simpler way, as the plausible ranges are not supported by the formal estimates of the relative risk of different outcomes that PDFs provide. There is a need then to assess whether the PDFs generated using the UKCP09 method for different Singapore variables are valid, before deciding how to present the uncertainty. For this study, for variables which do not pass the criteria for making PDFs, we present a wider range of responses by pooling the HadCM3Q and CMIP5 ensembles.

The methodology for producing PDFs, its critical assumptions, and the underpinning experimental design are outlined in Section 7.2. Section 7.3 describes three assessment criteria, and the assessment itself. In Section 7.4, estimates of uncertainty are shown for temperature and precipitation changes with some guidance. After a summary of results in Section 7.5, we make some recommendations for the use of the probabilistic information, and discuss some of its limitations.

7.2. Methodology

7.2.1 Outline of methodology

The method used here to make our probabilistic climate projections (Sexton et al 2012; Harris et al 2013) consists of three stages and is based on the seven ensembles outlined in Table 7.2.1. The first stage (Sexton et al 2012) uses a Bayesian framework (Rougier 2007) to predict, at the resolution of the global climate model, the distribution of equilibrium response to doubled CO_2 levels. The method combines information from a) a perturbed parameter ensemble (PPE; ensemble 1 in Table 7.2.1), where ensemble members are based on a standard version of the HadCM3 climate model but differ in the values of the model parameters that control its climate processes, b) a multimodel ensemble of other international climate models (Meehl et al 2007), and c) observations. Expert judgement is also included, for example in specifying prior distributions for uncertain model parameters, and in the choice of observations. The Bayesian method requires a more robust sampling of the plausible set of parameter combinations than is provided by the PPE, because the number of parameters perturbed (31 in this case) defines a space too large to be comprehensively covered by the number of feasible simulations (280 in this case). Fuller sampling is achieved by building an emulator, a statistical model trained upon the emergent properties of the PPE, which can be used to predict the recent mean climate and the equilibrium response to a doubling of CO_2 for any combination of parameter values. The Bayesian framework allows the projections to be constrained by a set of multiannual mean observations by weighting different model variants according to their ability to simulate aspects of historical mean climate. The framework recognises that climate models are imperfect, and combines information from the emulator and ensemble b) to specify and include “structural” modelling uncertainty (explored by climate models other than HadCM3-based ones) in land/atmosphere component of the climate model in the predicted probabilities.

The second stage uses a time scaling approach (Harris et al 2013) to provide probabilistic projections for regional climate change for different time periods during the 21st century by combining information from the probabilistic projections from stage one with GCM ensembles that explore uncertainties in the time-dependent response to historical forcings and projected future emissions (Ensembles 2-6). The time-dependent

regional response is emulated by assuming a linear variation with global annual mean temperature change, the latter being predicted by a Simple Climate Model (SCM). The time scaling is done for each sampled parameter combination; the PDFs of equilibrium response to doubled CO2 concentrations from stage one are sampled jointly to provide i) the climate feedbacks required to drive the SCM and ii) the normalised response per unit degree of global temperature change. The SCM is an energy balance model coupled to a simple carbon cycle model, which also includes aerosol forcing and ocean heat uptake. These Earth system components are sampled from a set of input parameters that have been calibrated to reproduce the response of the transient perturbed parameter (Ensembles 3-6) and multimodel ensemble simulations. The sampled projections are then reweighted, based upon the likelihood that they correctly replicate observed historical changes in surface temperature, and combined to provide time-dependent PDFs to the end of the twenty-first century, for the A1B emissions scenario.

The third stage converts GCM resolution PDFs from the first two stages to PDFs at the resolution of the regional climate simulations, 12km. This stage (Murphy et al 2009) first builds statistical relationships (here a simple linear regression model) pooled over the 9 RCM-GCM pairs from Chapters 4 and 5 linking the climate change in the regional climate simulation to a predictor variable from its driving CMIP5 GCM (note: GFDL-ESM-2G was dropped from the selected set of 10 CMIP5 climate models chosen in Chapter 3; see Chapter 4). The statistical model then provides a transfer function that generates a PDF at the local scale based on the PDF of the GCM predictor variable.

This third stage is the only one that differs from the three stages of UKCP09 in that it uses a different set of GCM and RCM data to define the statistical relationships. Note that, unlike UKCP09, both CMIP3 (for stage one) and CMIP5 (for stage three) data are used to make PDFs here. In the first stage, which is identical to the UKCP09 method, the Bayesian framework is applied to slab models from Ensemble 1 and the corresponding CMIP3 slab ensemble is used to provide the multimodel data for estimating the structural uncertainty. As CMIP3 contributes to stage one, we also used this data in the comparison of HadCM3Q, CMIP3, and CMIP5 projections in section 7.3.3.

For this methodology, to provide PDFs that are reliable for decision making, a number of assumptions have to be satisfied which will be assessed in Section 7.3:

- The climate models used need to have an acceptable level of performance in simulating present-day climate
- In the first stage of the method, there needs to be a good degree of overlap in the projections based on the perturbed parameter ensemble of slab model variants (Ensemble 1), and the other CMIP3 models used to account for structural uncertainty[1]. For example, in UKCP09, PDFs of change in latent heat flux were not provided, due to a lack of overlap between the two ensembles.
- The slab models also need to be good predictors of the coupled model response in Stage 2. The method for deriving time-dependent changes takes account of results from CMIP3 multi-model simulations of globally averaged climate change,

[1] In particular, the first stage of the methodology assumes that a relatively large PPE of simulations of the slab model configuration of HadCM3 provides a reasonable first order estimate of uncertainty associated with physical climate feedbacks, so that the additional contribution provided by a much smaller multi-model ensemble, while important, is not so large as to be render the PDFs excessively dependent on the limited number of alternative models available.

but relies exclusively on a HadCM3 PPE to predict how *spatial patterns* of the equilibrium response to doubled CO_2 are modified by the effects of dynamical ocean processes in transient simulations. This limits the extent to which structural uncertainty involving coupled atmosphere-ocean processes is represented.

- The main modelling uncertainties are explored by the GCM simulations, with the downscaling step adding local information due to the regional effects of orography, coastlines and other processes resolved by the regional climate model. Therefore PDFs should not be expected to capture phenomena that are due to local processes that act at spatial scales below the resolution of the regional climate simulations.
- There needs to be a linear relationship between some GCM predictor for which the above criteria apply and the RCM variable of interest. Statistical transformations can be applied to the GCM and RCM data to satisfy this condition.
- Since the UKCP09 method was developed using HadCM3Q and CMIP3 models, simulations based on improved climate models have been provided for CMIP5. For the method to provide PDFs that remain useful in light of these latest modelling results, both the HadCM3Q and CMIP3 projections should be consistent by showing sufficient overlap with the CMIP5 results.

More generally, there are inevitably limitations on how this or any other information based largely on computer simulations of climate change can be used, given that not all earth system processes capable of affecting future climate are included in contemporary climate models. This needs to be considered in the decision-making process. We have listed the main limitations in Box 7.6.1 and the caveats on their use in Box 7.6.2.

7.2.2 Experimental details

A *slab model*, as used in Ensemble 1, and Stage 1 of the probabilistic projection methodology, is an atmosphere model coupled to a very simple thermodynamic mixed-layer ("slab") ocean rather than a fully dynamic ocean model. It provides a relatively inexpensive configuration with which to sample parametric model uncertainties efficiently, in this case via large ensembles of the equilibrium response to an idealised doubling of CO_2 concentrations.

The ensembles 2-6 are used in Stage 2 of the methodology described in Section 7.2.1. They are simulations of transient climate change from 1860-2100 based on the coupled ocean atmosphere configuration of the climate model, and use *flux adjustment*. This is an empirical technique to limit the development of systematic biases or long term drifts in global and regional sea surface conditions (typically SST and surface salinity) in coupled models. The adjustments, which vary regionally and seasonally but not from year to year, consist of an additional surface flux of heat and moisture that has to be added in order to counter the combined effects of climatological errors in surface exchanges and ocean transports. Flux adjustments were used in all the HadCM3-based coupled ocean-atmosphere ensembles run for UKCP09 in order to avoid the risk that small structural deficiencies in the simulation of planetary energy balance might otherwise have

restricted the sampling of parametric model uncertainties to an artificially small region of the model parameter space (Collins et al 2011).

Observational data sets used for constraining the projections can be found in Sexton et al (2012) and Harris et al (2013). The probabilistic projections are made for the SRES A1B emissions scenario (Nakicenovic et al 2000) as the ensembles have been generated using this scenario. In Section 7.6, we discuss a simple relationship for approximating the PDFs for temperature change for the RCP8.5 emission scenario from those for SRES A1B.

Table 7.2.1. List of ensembles used to make probabilistic climate projections in Chapter 7.

Ensemble Number (and references)	Model configuration	Uncertainty explored by perturbing model parameters	Forcing scenario and nature of response	Ensemble size
1 Murphy et al (2004) Webb et al (2006) Rougier et al (2009) Sexton et al (2012)	Atmosphere coupled to thermodynamic 'slab' ocean	Land/atmosphere/sea-ice	Equilibrium response to doubled CO_2 concentrations	280
2 Harris et al (2013)	Atmosphere coupled to dynamic ocean	Land/atmosphere/sea-ice	Transient response to A1B	17
3 Collins et al (2007) Brierley et al (2010)	Atmosphere coupled to dynamic ocean	Ocean	Transient response to A1B	17
4 Ackerley et al (2011)	Atmosphere coupled to dynamic ocean	Sulphate aerosol chemistry	Transient response to A1B	17
5 Harris et al (2013)	Atmosphere coupled to dynamic ocean	Land/atmosphere/sea-ice	Transient response to A1B but with no changes in greenhouse gas concentrations	17
6 Booth et al (2012)	Atmosphere coupled to dynamic ocean and dynamic vegetation model	Terrestrial carbon cycle	Transient response to A1B	17
7 (see chapters 3, 4 and 5)	Regional climate model driven by subset of CMIP5 GCMs members.	Land/atmosphere/sea-ice/ocean structural uncertainty	Transient response to A1B	5

7.3. Assessment Criteria

The previous section outlined a list of assumptions and choices that need to be tested before PDFs can be passed on to stakeholders for use in their adaptation planning. Chapter 7 only supplies PDFs for variables that meet these criteria. For variables that do not pass them, alternative presentations of the wider uncertainty (defined in the Introduction) are used instead of PDFs. Based on discussions with MSS, the variables assessed here include seasonal mean surface air temperature and precipitation, as these are considered to be in part affected by large scale drivers such as global warming or regional circulation patterns, though this varies with season. In contrast, wind speed was not included, as the critical events that should be accounted for in planning are wind gusts associated with thunderstorms, which are locally driven.

The three assessment criteria were:

a. **Evaluation of HadCM3's performance**: The HadCM3Q perturbed parameter ensemble simulations can provide useful information for Singapore, using the same evaluation techniques as outlined in Chapter 3 for CMIP5 models.

b. **Downscaling relationships**: the required downscaling relationships between GCM- and RCM-scale future changes, derived in this case from CMIP5-driven RCMs, are appropriate for this purpose

c. **Comparison of HadCM3 and CMIP3/CMIP5 projections**: the CMIP5 projections show nothing substantially different to CMIP3/HadCM3Q, thereby questioning the credibility of the older models.

In the following three subsections each assessment criterion is evaluated.

7.3.1 Evaluation of HadCM3's performance

Chapters 4 and 5 are designed to equip stakeholders with local scale information, based on the latest state-of-the-art climate models generated for CMIP5, downscaled with a regional climate model, HadGEM3-RA. The sub-selection process in Chapter 3 provided an efficient way of sampling a range of outcomes using a subset of those CMIP5 climate models with an acceptable level of performance in simulating key climate features that affect South East Asian climate. However, CMIP5 is limited in its ability to explore the wider uncertainty, in two key respects. Firstly, the spread of outcomes simulated by any ensemble of climate models is dependent on the set of models or model variants included (e.g. Murphy et al. 2014), so no single ensemble can necessarily span a full range of plausible outcomes. Secondly, the set of earth system process uncertainties represented depends on how the simulations are designed. In particular, the CMIP5 simulations used for this project are concentration-driven, although Booth et al (2012), Lambert et al (2013), and Murphy et al (2014) show that a key source of uncertainty for regional climate projections is how the land carbon cycle converts emissions in to atmospheric CO_2 concentrations. This then affects the Earth's radiation budget and subsequently the global and then regional temperature changes.

The purpose of Chapter 7 is to provide stakeholders with an assessment of the wider uncertainty, either by using PDFs, or for some variables, an alternative presentation of wider uncertainty if PDFs are assessed to be unreliable. Either approach requires the use of a larger number of plausible models than is used in Chapter 5, such as those described in Section 7.2. Current supercomputing resources prohibit the use of CMIP5 models to generate the large ensembles that are needed, so HadCM3, a CMIP3-generation model, is used instead. This means the ensemble used to explore the wider uncertainty does not benefit from the improvements made to CMIP5 models to better represent regional climate around South East Asia. However, the non flux-adjusted HadCM3 standard model was evaluated in Appendix Chapter 3 A3.1. The main point for South East Asia was that HadCM3 was found to be in the major cluster of CMIP5 models which instead of reproducing the observed sudden northward transition of the ITCZ during boreal spring, demonstrate a widening of the ITCZ during the summer months (i.e. the equatorial region remains too wet during June-August).

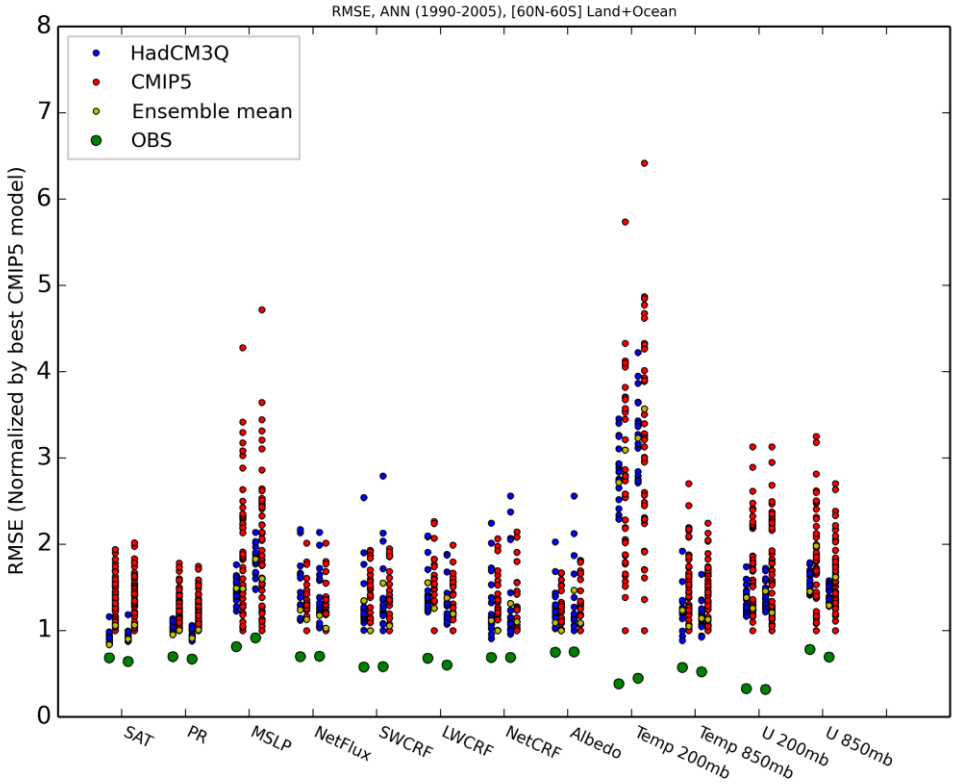

Figure 7.3.1. Comparison between 17 perturbed HadCM3Q variants and the available CMIP5 coupled ocean-atmosphere models of normalised root mean square errors (RMSE) in simulated 15 year averages of annual values of several climate variables (see text) for the period 1990-2005. All model output is regridded to HadCM3 2.5x3.75º latitude-longitude grid prior to the calculation of RMSE values, which are calculated by averaging regional squared errors from 60ºN-60ºS. Polar latitudes are excluded because uncertainties in some of the observational datasets are particularly large in those regions. For each climate variable there are two sets of results, as the analysis is done for two observational data sets. Green dots indicate the RMSE between the two observational data sets. To compare performance across variables, each RMSE value is normalised by dividing it by the RMSE of the best-performing CMIP5 model for each variable and each observational data set. Therefore for each column of red dots, the lowest value is guaranteed to be 1. The olive dots are the RMSE values of the mean of the relevant ensemble.

However, regional climate features are not the only consideration for assessing performance, with many larger scale drivers being important. For instance, some examples where large scale drivers outside the local region are important to local climate change are (i) local temperature changes will be part of a global pattern of warming, (ii) positions of the ITCZ will be controlled in part by inter-hemispheric temperature contrast, sea surface temperature changes outside the local region (see assessment of third criterion below) and other teleconnections to remote regions (Smith et al 2012), and (iii) monsoons are affected by changes in the land-sea contrast. There still remain systematic global errors common to HadCM3, CMIP3 and CMIP5 models, providing an unquantifiable source of uncertainty for projections on regional climate. So in addition to regional climate assessment in Chapter 3, a HadCM3 PPE is evaluated here to assess its performance on a global basis. This is consistent moreover with our approach of using of globally-representative large scale features of the mean climate to constrain the PDFs (see Section 7.2).

In this sub-section we provide some quantitative comparisons between the CMIP5 models and the 17-member perturbed parameter ensemble of variants of the HadCM3Q coupled ocean-atmosphere model (Ensemble 2 in Table 7.2.1). We necessarily limit our analysis of ensemble performance to comparison of this HadCM3Q ensemble against CMIP5 coupled ocean-atmosphere simulations, since the CMIP5 experiment protocol did not include production of simulations based on slab model configurations (Ensemble 1 in Table 7.2.1).

We focus mainly on global performance metrics relating to the simulation of long-term climatological averages of a set of physical variables commonly used to evaluate climate models. This is consistent with the approach used to apply weights to alternative model variants in UKCP09 (Sexton et al 2012), who use a number of metrics designed to estimate model quality on a global basis, and across a broad basket of variables; this is because influences on regional climate change can in general arise from a complex mix of local and remote dynamic and thermodynamic influences. While the analysis presented below is limited in its scope, it is nevertheless sufficient to give a broad first-order picture of how the ensembles of parametrically perturbed HadCM3Q compare in credibility to CMIP5 models.

Figure 7.3.1 shows normalised root-mean squared error (RMSE) values for historical long-term averages of spatial fields for a number of key variables for the region 60ºN-60ºS, including both land and ocean areas. These consist of air temperature at the surface (SAT) and in the lower and upper troposphere (Temp 850mb, Temp 200mb), precipitation (PR), mean sea level pressure (MSLP), the net radiative flux at the top of the atmosphere (NetFlux), shortwave, longwave and net effects of cloud on NetFlux (SWCRF, LWCRF, NetCRF), the fraction of incoming shortwave solar radiation returned to space (Albedo), and the westerly wind component in the lower and upper troposphere (U 850mb, U 200mb). This set of variables covers several basic emergent properties of climate, all influenced to varying degrees by atmospheric dynamics, radiative transfer, the water cycle and surface exchanges. The RMSE values essentially represent the globally-averaged magnitude value of regional errors for the relevant ensemble member and variable.

The results show that the HadCM3Q ensemble generally compares favourably with CMIP5. For most variables the best HadCM3Q performers would rank amongst the top CMIP5 models. For surface air temperature and precipitation, there are many CMIP5 models which score worse than the least skilful of the HadCM3Q variants, whilst a number of the HadCM3Q members give scores slightly better than any of the best

CMIP5 models. For the tropospheric temperature and wind metrics, the HadCM3Q simulations span a narrower range of scores than CMIP5: Temp 200mb is the weakest variable regarding HadCM3Q performance, although the largest biases always come from the CMIP5 ensemble, while several CMIP5 members outperform the best of the HadCM3Q variants (with the exception of Temp 850mb). One or two HadCM3Q members do score worse than all CMIP5 models for some of the radiative flux diagnostics.

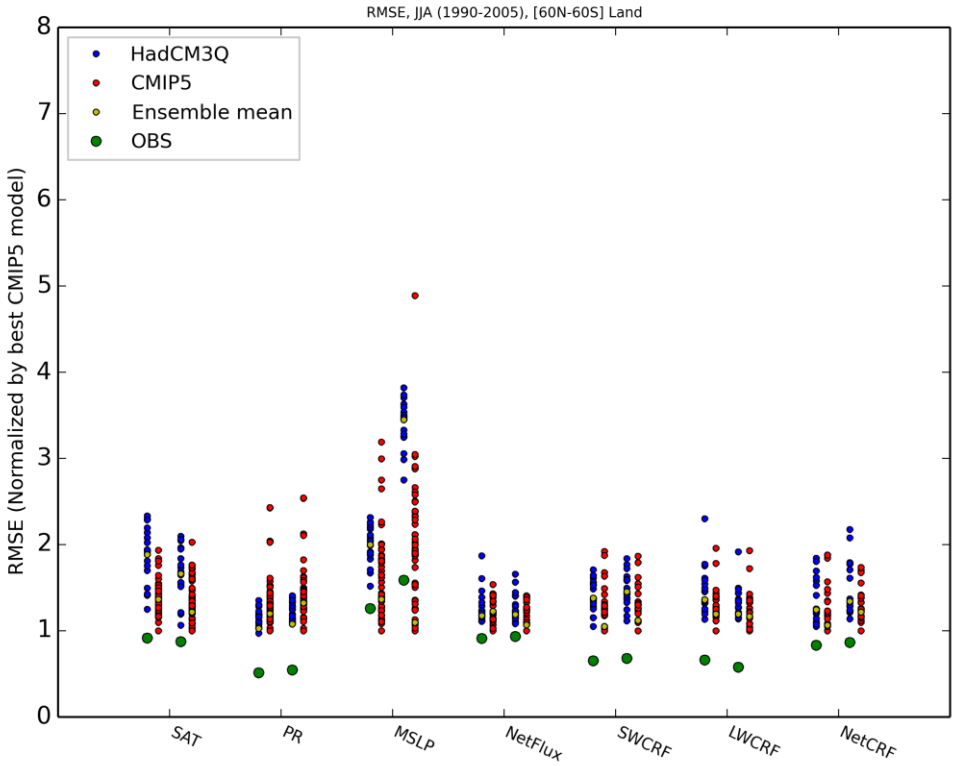

Figure 7.3.2. As Figure 7.3.1 but for JJA over land regions only and for a reduced set of variables.

However even in these cases the two ensembles span similar overall ranges of performance, and several individual members of HadCM3Q are competitive with the better CMIP5 models. Figure 7.3.1 also demonstrates that all contemporary models still contain significant systematic biases, since the best performing models (or model variants) invariably give larger RMSE values than those obtained from the differences between the two observational datasets (green dots in Figure 7.3.1). For RMSE values calculated over land points only, the conclusions above also hold in the case of annual mean climate (not shown). For precipitation, the HadCM3Q variants are better than most of the CMIP5 models but not as good as the best CMIP5 model.

Seasonal comparisons over land (see Figure 7.3.2 for JJA; DJF not shown) of HadCM3Q and CMIP5 climatology also show favourable results. Note for this assessment of model performance, we use DJF as this is generally more relevant to other worldwide regions, rather than NDJ used in other chapters and later on in this chapter, which is relevant for regional changes over South East Asia and Singapore. There are only a few variables (some of the radiative fluxes and JJA surface air temperature) where the worst simulation belongs to the HadCM3Q ensemble rather than CMIP5. In JJA, errors in surface air temperature are typically somewhat larger in

HadCM3Q than in CMIP5. This arises from a warm bias in central regions of the Eurasian and north American continental landmasses (Murphy et al 2009). Analysis of a more recent ensemble of variants of the earth system configuration of HadCM3 (Murphy

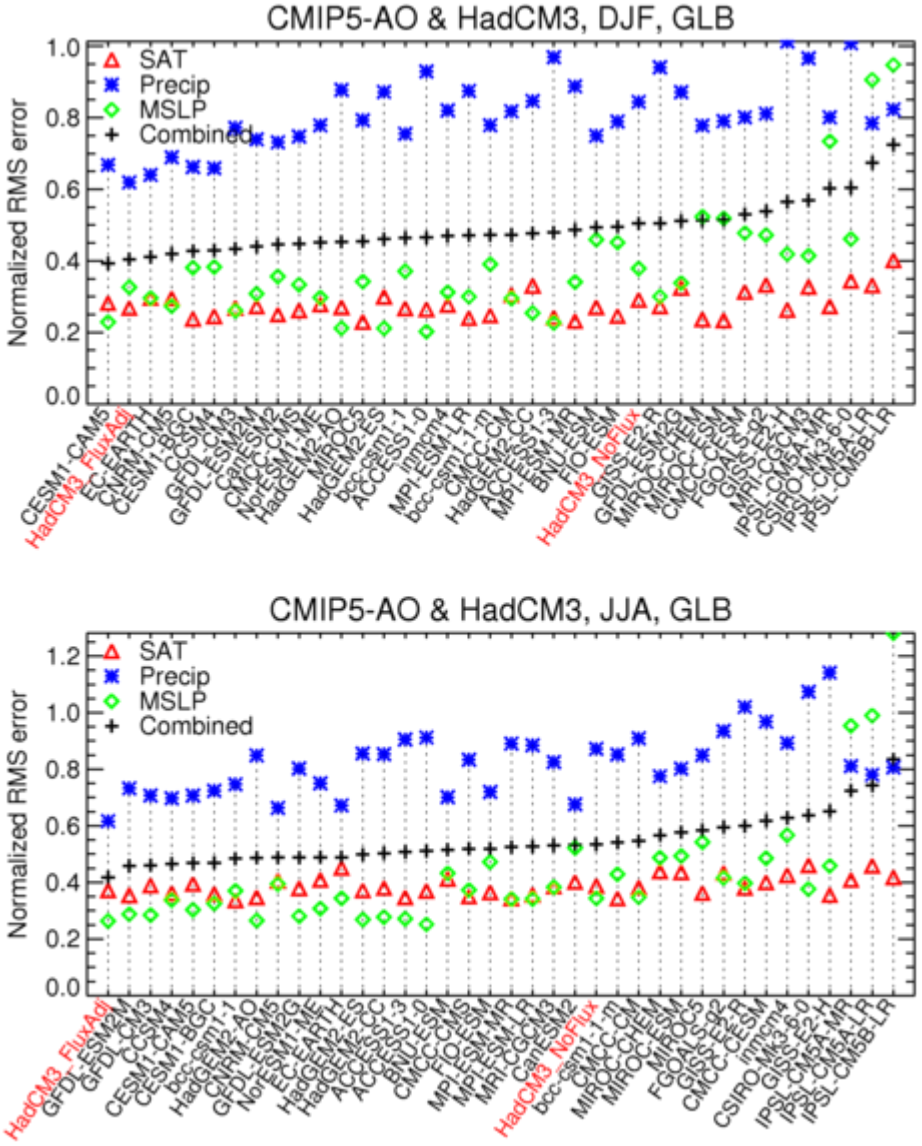

Figure 7.3.3. Normalised root mean square errors of 1.5m temperature, precipitation, and MSLP over the whole globe for HadCM3 and CMIP5 models (from Murphy et al 2014). Normalisation is different to that used in Figures 7.3.1-3; here the mean square error is divided by the spatial gridpoint standard deviation of the observed climatological field. Models are ranked according to a combined score. Top panel shows DJF, bottom panel shows JJA.

et al 2014) shows that warm biases in summer SAT over continental regions in the northern hemisphere are largest in model variants which simulate the lowest values of soil moisture content. For most variables, comparisons between the HadCM3 and CMIP5 scores are consistent across the two observational data sets. However, for mean sea level pressure in JJA, the performance of HadCM3Q relative to CMIP5 appears acceptable when verifying observations from the HadSLP2 dataset (Allan and Ansell 2006) are used (left-hand set of dots in the MSLP column of Figure 7.3.2), but not when ERA-Interim reanalysis data (Dee et al 2011) is used (right-hand set of dots). The different scores reflect large differences between the verification datasets in

mountainous regions (especially over the Tibetan plateau). The RMSE scores of Figure 7.3.2 suggest that several CMIP5 models score better against either observational dataset than the two datasets score against each other. This however more likely reflects limitations of the verification data than a true absence of systematic biases in any of the models. These results should therefore be interpreted with caution. Over ocean regions (not shown), RMSE scores are much more consistent between the two observational datasets, and HadCM3Q performs well with respect to CMIP5.

Overall, Figures 7.3.1 and 7.3.2 suggest that HadCM3Q performs well against CMIP5, with the exception of summer surface air temperature over northern hemisphere continents. However, it is important to note that the HadCM3Q simulations used flux-adjustments in order to reduce systematic biases in regional sea surface temperature (SST) and salinity values, and also reduce the risk of under-sampling the effects of parametric model uncertainties. Some CMIP3 models also used flux adjustments, but no CMIP5 models do. While the use of flux adjustments can be justified in a climate prediction context (as a means of supporting uncertainty quantification and reducing the effects of historical SST biases on simulated future changes), their use does not improve the fundamental quality of a climate model from a process perspective, so it is important to understand to what extent their use may favourably influence the scores of model performance metrics, compared to the scores for models which do not use flux adjustments.

In Figure 7.3.3 we compare RMSE values for the key variables of surface air temperature, precipitation and MSLP, found in flux-adjusted and non-flux adjusted configurations of the version of HadCM3 using standard parameter settings (noting that the non-flux adjusted version also omitted aerosol-cloud interactions). Values for CMIP5 coupled ocean-atmosphere models are also shown. RMSE values are calculated over both land and ocean and are normalised in this case using the spatial standard deviation of the observed field, so that the values represent typical regional errors relative to the observed contrasts between different regions. The black crosses in Figure 7.3.3 show a combined skill metric for each model, obtained by averaging the normalised RMSE value for the three variables. The flux-adjusted HadCM3 simulation gives a better combined score than all CMIP5 models in JJA (lower panel), and for all but one model in DJF (top panel). The HadCM3 simulation without flux adjustments gives a combined score within the body of CMIP5 results, though towards somewhat below the average performance level.

Over land (Figure 7.3.4), the impact of flux adjustments is much smaller: both HadCM3 simulations give scores towards the upper end of the CMIP5 performance range. In Figure 7.3.3, the main impact of using flux adjustments is to improve the precipitation score. This arises mainly from improvements over the tropical oceans. For example, McSweeney et al (2012) note improvements in the tropical Pacific resulting from partial correction of a cool bias in the region of the equatorial cold tongue in SST. Such improvements are also consistent with previous work demonstrating reduction of precipitation biases resulting from the correction of SST errors (e.g. Ashfaq et al 2011). Overall, Figures 7.3.3 and 7.3.4 show that the use of flux adjustments does play a role in improving the performance of HadCM3 in simulating long term mean climate, particularly for precipitation over the oceans, however the performance of HadCM3 remains reasonably competitive with CMIP5 models even when flux adjustments are not used, especially over land. This suggests that the representation of processes determining the climatological average characteristics in this core UKCP09 ensemble remains consistent with the latest performance benchmark set by CMIP5, despite the improvements in CMIP5 performance relative to CMIP3 models (Flato et al 2013). This is because

HadCM3 is one of the better CMIP3 models (e.g. Reichler and Kim 2008), and was therefore an appropriate choice as the basis for the systematic approach to uncertainty quantification achieved via its use in perturbed parameter ensembles for UKCP09.

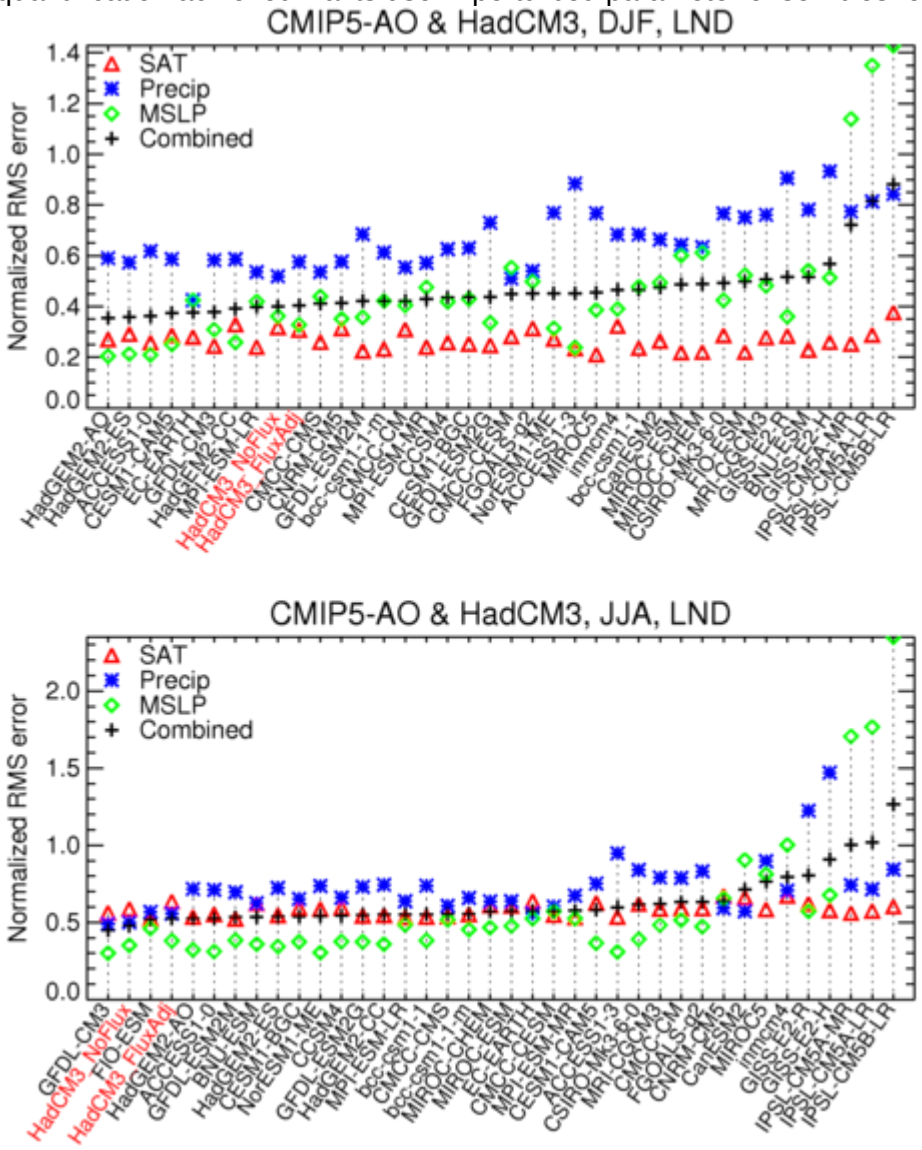

Figure 7.3.4. As Figure 7.3.3 but over all land.

7.3.2 Evaluation of downscaling relationships

The approach taken to make the PDFs at the RCM scale requires that large-scale changes in the GCM bear a clear physical and statistical relationship to the RCM changes. To assess this we show maps of correlations (see Figures 7.3.5 and 7.3.7) between the RCM change at the grid point defined as Macritchie in Figure 7.4.1, and the GCM changes relative to 1961-90. The correlations at a single GCM point (and later, the linear regression estimates used for the downscaling) are based on all the points from the 9 GCM-RCM pairs e.g. just like the scatterplot in Figure 7.3.6 (lower panel).

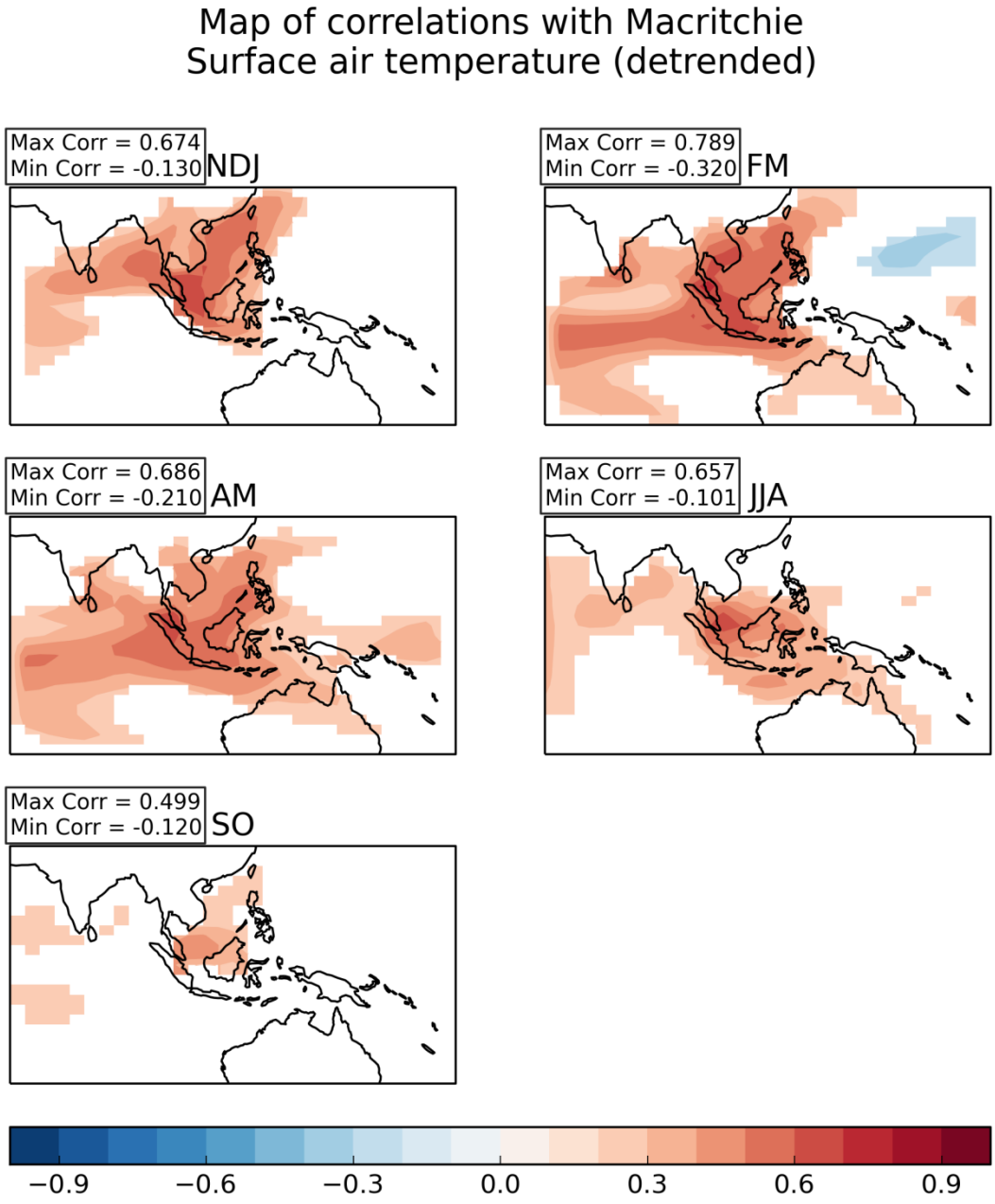

Figure 7.3.5. Maps of the correlation of temperature changes which have been detrended by removing variability using a 50-year low-pass filter. The correlations are between the RCM change at the Macritchie RCM point and the GCM changes relative to 1961-90 based all 9-member ensemble of GCM-RCM pairs. All GCM runs are first regridded to a common grid, here HadCM3, using only land points to regrid to HadCM3 land points, and only ocean points to regrid to HadCM3 ocean points. The title of each panel is the meaning period e.g. SO is September-October mean.

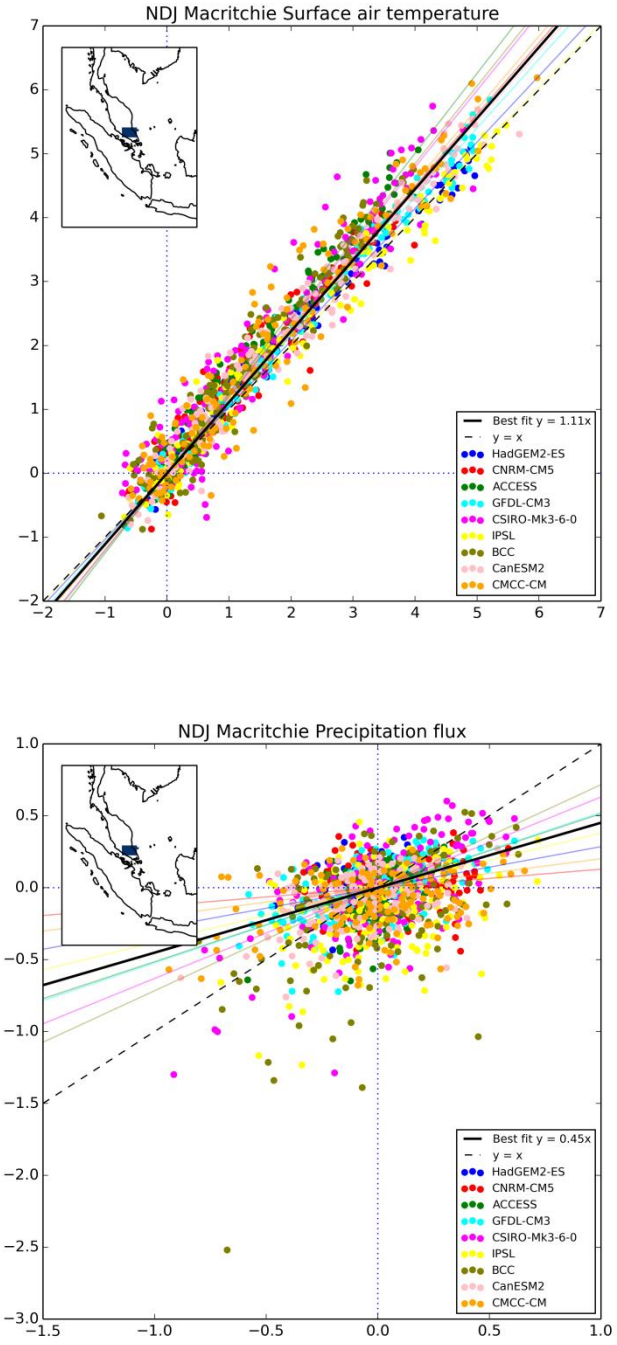

Figure 7.3.6. Scatter plots of GCM anomalies at the grid point highlighted in the inset versus RCM anomalies for the Macritchie RCM grid point for temperature (upper panel) and log (precipitation) (lower panel). The logarithm of precipitation is used so that the residuals from a linear fit are Gaussian. In each panel, both RCM and GCM anomalies are the change in the NDJ value for these variables and each year relative to 1961-90. The different coloured points identify the nine different CMIP5 models used to drive the RCM simulations in Chapters 4 and 5. The CMIP5 GCM data were regridded to the HadCM3 grid before the anomalies were calculated, using only land points to regrid to HadCM3 land points, and only ocean points to regrid to HadCM3 ocean points. The coloured lines show the best fit line for each CMIP5 model, and the black line shows the fit using all points. The dashed line shows y=x for reference.

To make the correlation maps, the analysis is repeated for each HadCM3 grid box. For temperature changes, there is an extra processing step before the correlations are estimated, as variations on time scales of 50 years or more have first of all been removed from the GCM and RCM data. This effectively removes the influence of long term global warming on the correlation maps, which exerts a strong influence on the correlations across the globe. By removing this influence, the choice of predictor is focused on those grid points that have strong relationships with Singapore on time scales of a few years or less. Note, however, that when regression relationships are built for downscaling the probabilistic projections at the GCM scale (see below), no filtering is done. Therefore the long term warming signal is included, as this is important to capture, alongside the effects of variability on shorter time scales. Figure 7.3.5 shows that there is a relationship between Singapore temperatures and the larger region covering much of Indonesia, the South China Sea, and parts of the Indian Ocean. The strength of the pattern varies with season, with the strongest relationship in February-March, the weakest in September-October. Importantly, this indicates that the filtered temperature variations in the regional climate simulations are being driven by the large scale variations provided by the driving GCMs. Therefore, it is appropriate to apply a downscaling step to probabilistic projections estimated at the GCM scale, to generate PDFs at a more relevant, local spatial scale.

The correlation maps serve a secondary purpose as they indicate the GCM land points which have the highest correlations with the local RCM variability, and are therefore the best choices to use as predictors. We also chose to use the same GCM land point to predict the RCM changes for all local Singapore points and seasons. Whilst this choice is not absolutely necessary, it helps to maintain some coherence between PDFs and, data sampled from them (see Section 7.6), across spatial points and seasons. Full coherence as seen in the climate model simulations is not possible as the current method to build an emulator (see Section 7.2) is only 'approximately' multivariate - separate emulators are built for each climate variable (e.g. temperature change at a given GCM grid point) and multivariate relationships are represented by correlations in emulation errors for different variables (Sexton et al 2012).

For temperature changes, we chose the GCM point over South Malaysia as shown in the inset of Figure 7.3.6 (upper panel). The estimate of the regression slope (see legend in Figure 7.3.6) is based on all values from the 9-member ensemble of GCM-RCM pairs and zero intercept is assumed. In contrast to the process used to select GCM predictor grid points (Figure 7.3.5 and discussion above), no detrending is done when the final downscaling relationships are built. This is because it is important in predictive mode to include the effects of both long term trends, and natural variability. The downscaling step in Section 7.2 uses the estimate of the error variance from the regression as well as this slope for each RCM point, variable, and meaning period. For temperature changes in NDJ and other seasons (not shown) there is clearly a strong relationship between the temperature change at the GCM point and the Macritchie RCM point. The slope for NDJ mean temperature change is 1.11 indicating the RCM at Macritchie typically warms at a slightly faster rate than the driving GCM. In this example, the slopes for the individual CMIP5 GCM-RCM pairs are very similar indicating that there is not much modelling uncertainty in this estimate. It is also important to validate the linear regression model. Inspection of standard diagnostic plots for regression (not shown) for each meaning period and Singapore location shows that the linear model is suitable. However, when downscaling annual mean fit line tends to over-predict the RCM changes by about 0.5^{o}C. This seems to arise mainly from the hottest seasons of the year (April-May and September-October), in which the regional model tends to moderate the warming simulated by the driving GCM simulations. The difference of 0.5^{o}C is small compared to

the overall range of the projections (see Section 7.4) but the upper tails in these seasons may well be slight over-estimates.

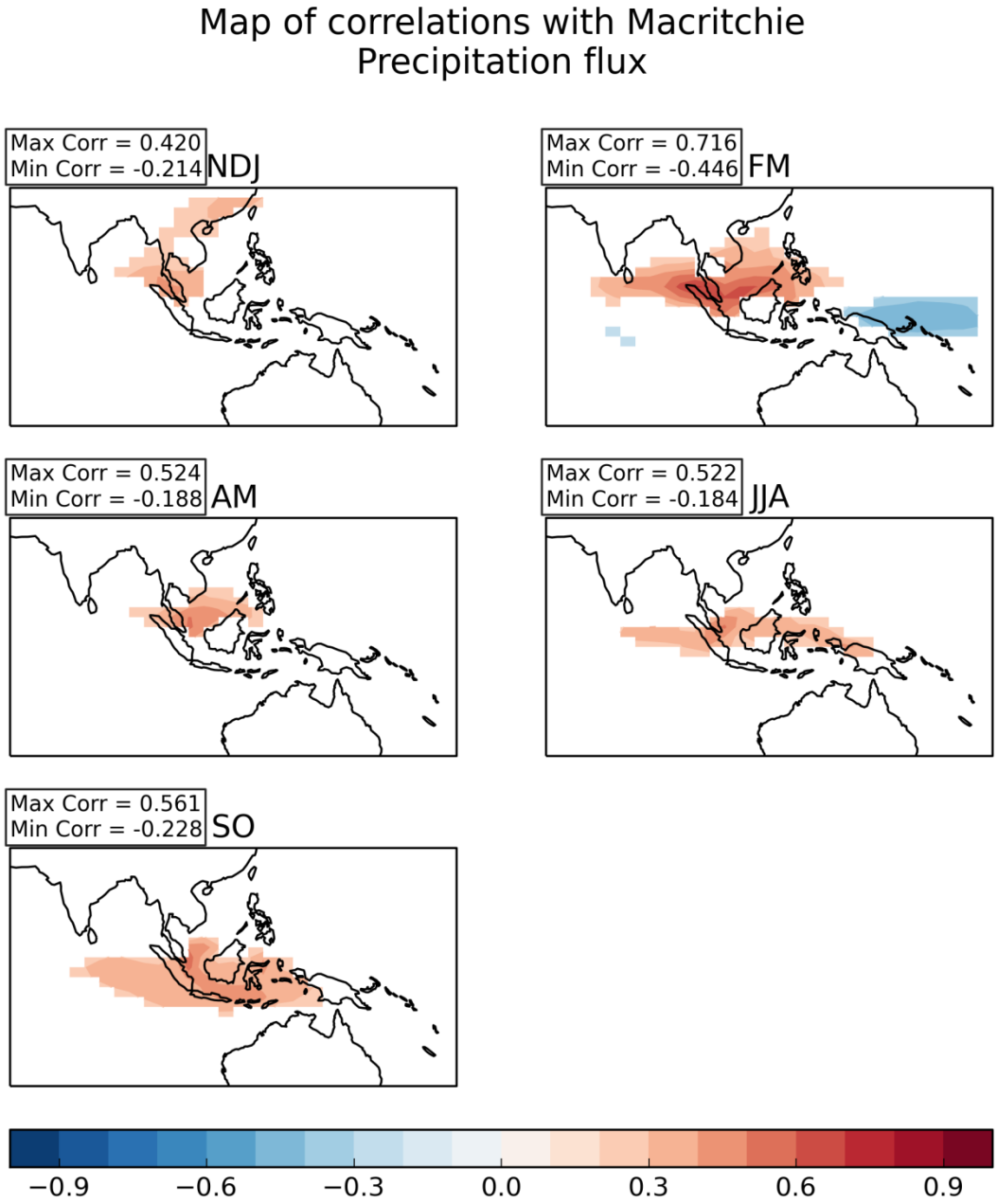

Figure 7.3.7. As for Figure 7.3.6 but for log(precipitation) anomalies.

For precipitation changes, there is a weaker link between the changes at the Macritchie RCM point and the wider regional variations coming from the GCM simulations. Nevertheless, Figure 7.3.7 shows the link between local precipitation changes and the ITCZ and in NDJ, to South East China. This may be that the regional climate simulations are representing the real effect that precipitation extremes in NDJ are linked to cold air intrusions from mid-latitudes, but this needs further study. For a given GCM point, it is clear from Figure 7.3.6 (lower panel) that the slope for individual GCM pairs can be different. One strategy would be to build nine PDFs, each based on a single GCM-RCM pair, and then merge them. However, this could generate modality in the resultant PDFs

providing spurious detail that might confuse users or be over-interpreted. However, there is also a lot of overlap between the scatter from the 9 GCM-RCM pairs. Therefore, it is appropriate to sample regression errors using pooled estimates of error variance. The pooled estimate of the error variance is larger than the nine individual estimates, because it captures the uncertainty from inter-model differences in the downscaling relationships, as well as uncertainty arising within any specific GCM-RCM pair. For precipitation, we find that the same South Malaysian GCM point (see the inset in Figure 7.3.6 lower panel) is the best choice across all seasons. For examples like September-October, where there is a weak slope, the probabilistic projections at the RCM scale are dominated by information from the regional climate simulations, with the GCM scale information being effectively down-weighted. However, we show below that the third assessment criterion is not satisfied, so we stop short of attempting probabilistic projections including downscaling, preferring to illustrate the presence of wider uncertainty by simply showing plausible GCM-scale realizations at the South Malaysian grid point.

7.3.3 Comparison of HadCM3 and CMIP3/CMIP5 projections

7.3.3.1 Temperature

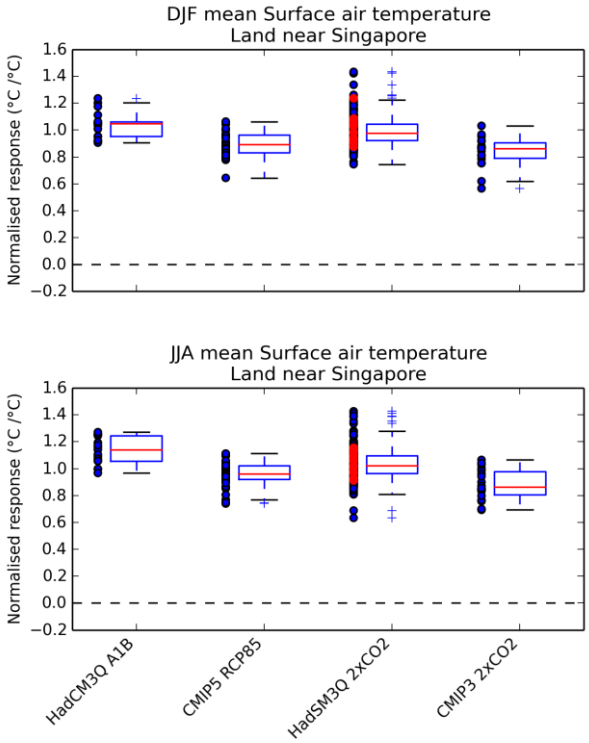

Figure 7.3.8. Comparison of the normalised DJF (upper panel) and JJA (lower panel) mean temperature response per unit global warming for each member (dots) of four perturbed parameter or multimodel ensembles covering three forcing scenarios for an average over 5 GCM land points (two in Malaysia, three covering Sumatra)**. Box and whiskers show the median (red line), the box extends from the 1st to 3rd quartile, whiskers extend the quartiles by 1.5 times the inter-quartile range, and crosses show values outside the whiskers. Red dots show the subset of 17 HadSM3Q ensemble members that correspond to the HadCM3Q A1B experiment.**

The range of future projections for global temperature change in CMIP5 is different from that of either the HadCM3Q ensemble, or the probabilistic range derived from the full

UKCP09 methodology, due to the use of different emissions scenarios regardless of any differences in the range of climate sensitivities explored by either ensemble, or the consideration of carbon cycle feedbacks in the case of UKCP09. Therefore, we would expect to see some divergence in envelopes of regional change, since global mean warming is a key driver of regional uncertainties (e.g. Harris et al 2006), particularly for temperature-related variables. To overcome this complication in comparing raw data from HadCM3Q A1B and CMIP5 RCP6 and RCP8.5 simulations, we instead analyse the normalised response of temperature and other variables, defined as the change in the relevant variable per unit change in global temperature. This allows us to focus specifically on potential differences in the regional patterns of response. Note that CMIP3 is used here as the slab model responses due to doubling CO2 levels are used to provide the estimate of structural uncertainty in the first stage of the method to make PDFs (see section 7.2).

Figure 7.3.8 compares the range of normalised DJF and JJA mean temperature change per unit global warming for four ensembles made up of different combinations of climate models and forcing scenarios. There are two main comparisons to be made, and conclusions from these are similar for both seasons. First, there is reasonable overlap between the normalised temperature response in the HadCM3Q 2xCO2 and CMIP3 2xCO2 slab ensembles, which implies that the first stage of the PDF method (see Section 7.2) will work. That the CMIP3 ensemble explores the lower half of the HadCM3Q range suggests that the term in stage one of the method to account for structural uncertainty will act to lower the normalised response from the emulator. Second, there is good agreement between CMIP3 and CMIP5 suggesting that PDFs of temperature at the GCM scale will be consistent with the newer improved models.

7.3.3.2 Precipitation

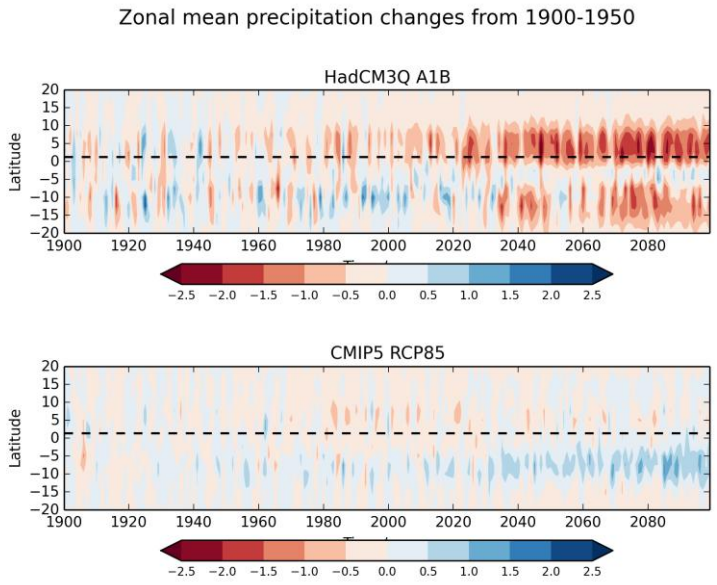

Figure 7.3.9. Hovmöller plots of HadCMQ A1B (top panel) and CMIP5 RCP85 (bottom panel) ensemble mean changes in DJF mean precipitation (mm/day) relative to 1900-1950, zonally averaged in the longitude band 85°E to 115°E.

McSweeney et al (2012) showed that the 17 flux-adjusted HadCM3Q variants give a drying response over the South East Asia region, in contrast to an increase in rainfall indicated by the CMIP3 models. The CMIP3 and CMIP5 ensemble mean responses are similar so we show CMIP5 results to help make this analysis more relevant to the results in Sections 7.4 and 7. 5. Figure 7.3.9 shows that the evolution of rainfall zonally averaged over a longitudinal region from 85^{o}E to 115^{o}E is clearly different between the HadCM3Q and CMIP5 ensembles.

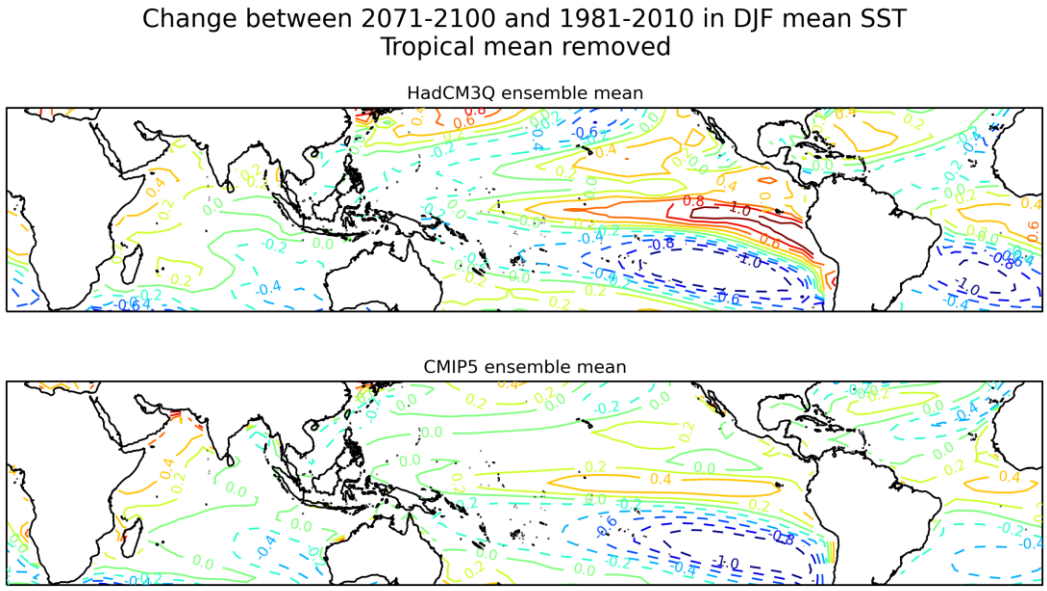

Figure 7.3.10. Ensemble mean changes in local SST relative to changes in tropical mean SST from 2071-2100 relative to 1981-2000 for HadCM3Q A1B (top panel) and CMIP5 RCP85 (bottom panel).

After some investigation, the reason for this difference in response was found to be linked to differences in the response of SSTs over the eastern and central tropical Pacific, and the implications this has for local stability and rainfall throughout the rest of the tropics. Following the approach taken by Chadwick et al (2013), we show the changes in local tropical SST relative to the changes in the tropical mean SST, as this is a proxy for local stability. This is because the boundary layer temperatures are influenced by local tropical SST changes, whereas the free tropical troposphere above the boundary layer is well mixed and its temperature changes are related to changes in the tropical mean SST. The differences between the local boundary layer temperature and temperatures in the free troposphere aloft are related to local stability. Figure 7.3.10 shows that the HadCM3Q ensemble mean SST changes over the eastern and central tropical warm more than over the rest of the tropics, including the warm pool region and ocean around Singapore and Indonesia. Figures 7.3.11 and 7.3.12 overlay the DJF and JJA mean precipitation changes over the change in SSTs relative to the tropical mean, to show the strong relationship that exists, and help explain why the precipitation changes in HadCM3Q over the eastern Indian ocean, in the seas around Indonesia and the warm pool are different to that found in CMIP5.

Both HadCM3Q and CMIP5 patterns of ensemble mean SST change seem plausible, although different. One potential factor is that the use of flux adjustments in HadCM3Q could exert a first order effect. This could either be beneficial, due to a reduction in local

SST biases removing sources of error in the simulated future changes (e.g. Ashfaq et al. 2011), or problematic due to some distortion of the dynamics of atmosphere-ocean coupling (e.g. Spencer et al. 2007). However, a comparison (not shown) of the responses from two HadCM3Q variants which were run both with, and without flux adjustment shows that flux adjustment is not responsible for the difference between HadCM3Q and CMIP5. We therefore assess that the HadCM3Q ensemble is exploring a plausible scenario not covered by the CMIP5 models, and is therefore providing a useful insight into the wider uncertainty. Based on the assessment criteria outlined in Section 7.2.1, we assess that this lack of overlap means that PDFs based on HadCM3Q and CMIP3/CMIP5 for the South East Asian region would be unreliable.

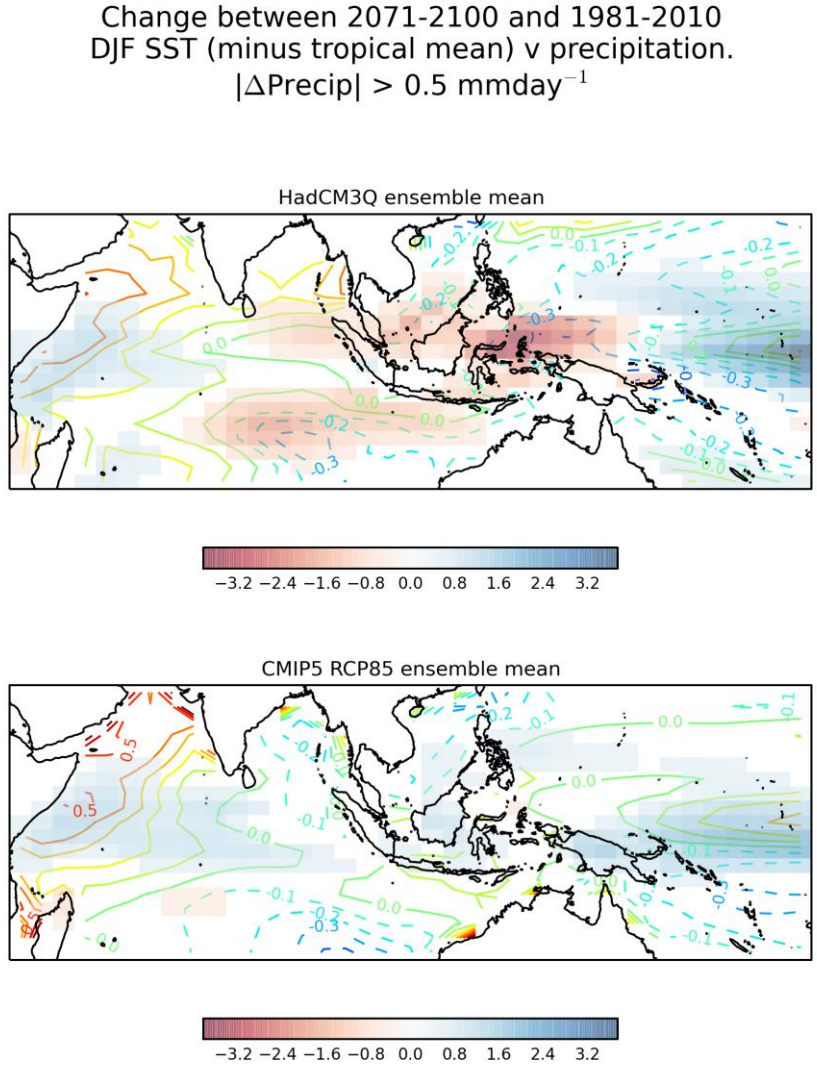

Figure 7.3.11. Ensemble mean changes in DJF mean precipitation (shading) overlaid on local SST relative to changes in DJF mean tropical mean SST (contours) from 2071-2100 relative to 1981-2000 for HadCM3Q A1B (top panel) and CMIP5 RCP85 (bottom panel).

The precipitation response though around Singapore is weaker, suggesting a need to check whether the lack of overlap extends to the more local scale. We compared the normalised response for precipitation over a more local region, that is the average over 5 GCM land points (two in Malaysia, three covering Sumatra). Figure 7.3.13 compares the

normalised responses for DJF and JJA mean precipitation changes. The results show good overlap between the slab model ensembles. The coupled simulations show more overlap than was found for the larger South East Asia region above. However, for JJA the HadCM3Q coupled runs show a drier envelope of responses than the corresponding slab simulations, arising from coupling the atmosphere to a dynamic rather than thermodynamic ocean. In contrast, the CMIP5 transient simulations do not show such a shift, relative to the CMIP3 slab simulations. No such issues stand out for DJF.

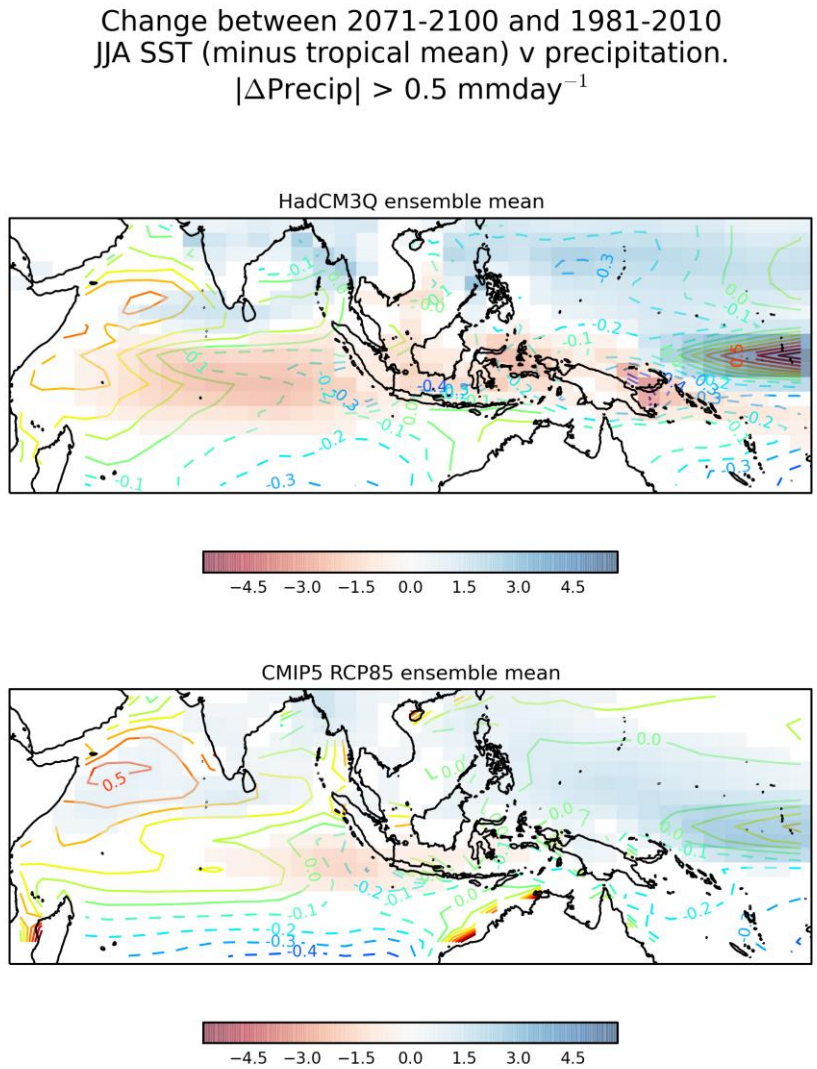

Figure 7.3.12. As Figure 7.3.11 but for JJA.

For DJF, we attempted to actually produce PDFs of transient future changes at the GCM grid box over South Malaysia. The plume of uncertainty (not shown) implies a 70% chance that DJF precipitation will be drier than 1961-90 throughout much of 21st century. Moreover, the 17 HadCM3Q variants cover the 30th-70th percentile range, indicating that the PDFs are being dominated by the response simulated by the HadCM3Q coupled model variants described above. Therefore, we conclude that the PDFs do not present a balanced assessment of the uncertainty that evenly spans HadCM3Q and the CMIP5 models.

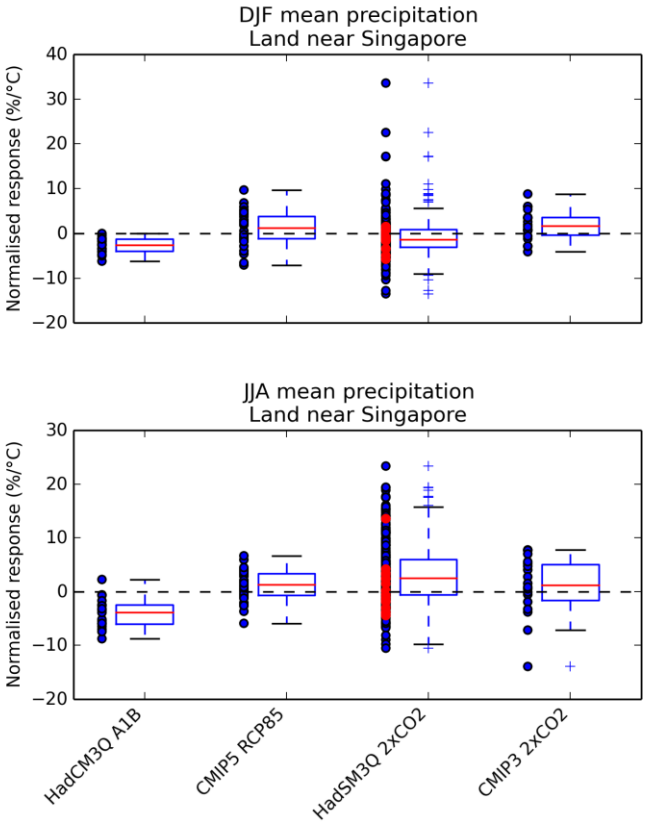

Figure 7.3.13. As Figure 7.3.8 but for normalised DJF and JJA mean percentage precipitation response per unit global warming for each member.

7.3.3.3 Summary of assessment

Whilst the first two criteria (model performance and the ability to find downscaling relationships between a GCM grid box and local Singapore RCM grid boxes) were found to be satisfactory, the third criterion (consistency between CMIP5, CMIP3 and HadCM3Q future responses) was only satisfactory for temperature and not precipitation changes. Therefore, probabilistic projections are produced for temperature change only. The wider uncertainty for precipitation change is assessed using pooled ensembles of CMIP5 and HadCM3Q results.

7.4. Results

7.4.1 Temperature

The method outlined in Section 2 was applied to temperature changes, to make probabilistic projections for individual seasons from 1950-2100. PDFs were made for 9 Singapore locations (see Figure 7.4.1) and 18 meaning periods (each month, annual means and seasonal means for NDJ, FM, AM, JJA, and SO). Note that for regional

projections of South East Asia, it is more appropriate to use NDJ rather than DJF as used for the assessment of model performance in section 7.3.1.

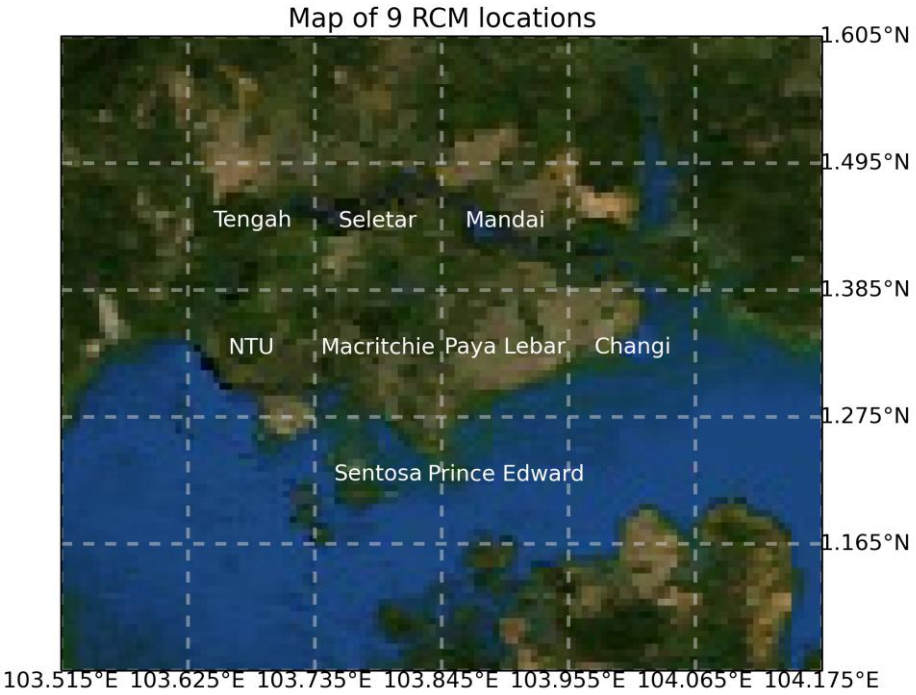

Figure 7.4.1 Map showing the regional climate model grid (dotted lines) and the 9 Singapore locations named after a key station in each grid box.

Figure 7.4.2 is an example of the plots that can be found in the pre-prepared graphics forming one of the outputs from this work package. It shows the evolution of seven percentiles (see legend) for Annual mean temperature change relative to 1961-90 from 1961 to 2100 for the 9 Singapore RCM grid boxes. The projections are similar across all locations, showing an increase in the spread through time due to an increase in the uncertainty arising from the climate change signal. Uncertainty due to natural year-to-year variability is also included in the uncertainty, and is based on the variability as simulated by the HadCM3Q ensemble. The temperature changes around the 50th percentile are relatively more likely than the changes in the tails of the distribution (see Figure 7.6.1 right panel in the section 7.6). The upper tails are very similar for all locations, although by 2100, Sentosa and Prince Edward, which are ocean points in the regional climate simulations, are not as warm as the others by about 1^{o}C. In Section 7.5, Table 7.5.1 summarises how the temperature changes and the uncertainty increases with time.

The lower red line, which is the 5th percentile, crosses zero change at around 2020 indicating that under the A1B SRES scenario there is a 5% chance of a year around this time being as cool as the 1961-90 baseline. Beyond about 2050, what was considered typical in 1961-90 is extremely unlikely. The record annual mean Singapore temperature anomaly (which occurred in 1998) was 1.46^{o}C relative to 1961-90 according to CRUTEM4.2.0.0 (Jones et al 2012) downloaded from http://www.metoffice.gov.uk/hadobs/crutem4/. Such a temperature change is just above the 96th percentile for 1998, could become typical around 2030 (when the 50th percentile line reaches this value), and would be considered a very cold year by the late

2070s (when the 2.5% line reaches this value). The magnitude of natural variability can be gauged from the spread in the plume during the 20th century, and is an appreciable source of uncertainty.

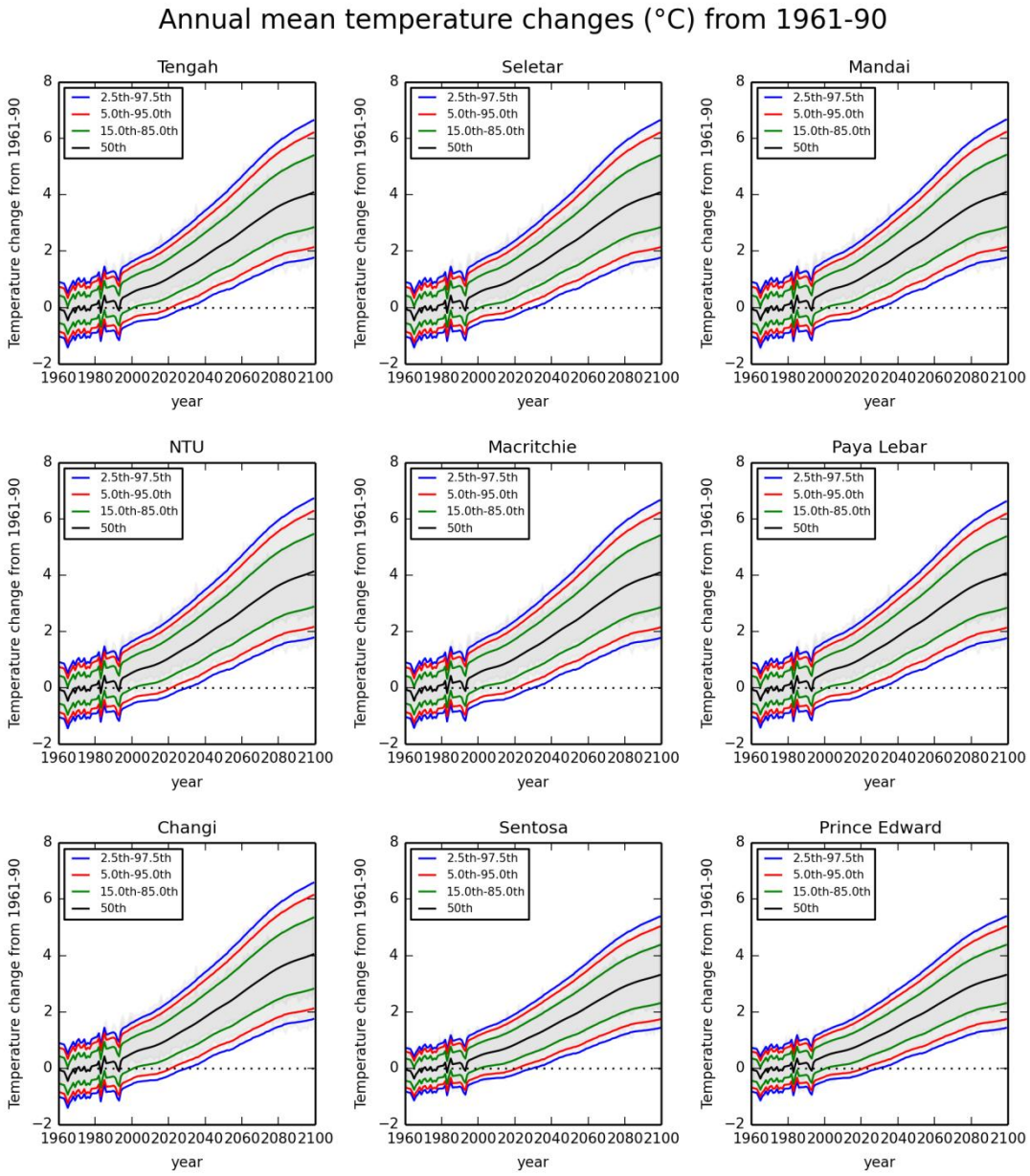

Figure 7.4.2. The time variation of six percentiles (see legend) and the 50th percentile (median) for Annual mean temperature change relative to 1961-90 for each year from 1961 to 2100 under the SRES A1B emissions scenario. The shaded plume shows the same six percentiles estimated from the sampled data (see Section 7.6). Each panel shows one of the 9 Singapore locations. The high frequency variations seen in the late 20th century but not in the 21st century are due to the way the method used to make the PDFs handles the effect of volcanic eruptions over the year of the eruption and the subsequent three years.

7.4.2 Precipitation

Figure 7.4.3 compares the percentage precipitation changes relative to 1902-2009 for 30-year means in 10 year steps for the South Malaysia GCM grid point, for each of the 9 CMIP5 models used to drive the RCMs in Chapters 4 and 5 (solid lines). The envelopes show the ensemble range after the highest and lowest values at each time point have been removed, to reduce the effect of outliers. Percentage change is used here to remove the effects of differences in the 1902-2009 mean absolute precipitation across the HadCM3Q and CMIP5 models. The baseline period 1902-2009 was chosen to correspond with a baseline used for observations in Figures 7.4.4 and 7.4.5. Percentage change in precipitation is not strongly dependent on emission scenario, so it is acceptable, to first order, to combine HadCM3Q output for the SRES A1B scenario with CMIP5 output for RCP8.5.

There is an increase in spread in both the 9 CMIP5 simulations and the CMIP5 and HadCM3Q ensembles throughout the 21st century. Pooling the envelopes from the two ensembles, there is no clear statement that can be made about the sign of precipitation changes for the Singapore region.

The CMIP5 time series show a variety of low frequency characteristics, ranging from quasi-linear trends (e.g. CanESM2 in annual mean panel), to time series which show multi-decadal variability with a single turning point (e.g. CNRM-CM for JJA mean precipitation changes). The latter behaviour may result from shifting rainfall patterns (Hawkins et al 2014) or natural climate variability. Therefore, unlike the steadily increasing signals for temperature change, stakeholders need to consider that the largest precipitation changes for some of the 9 CMIP5-driven RCM runs might occur during the 21st century rather than at the end of it. In addition to the pronounced low frequency variability on multidecadal time scales, it is possible to see modest natural variability on the decadal time scale.

For NDJ, it is clear that the 9 CMIP5 models do not span the wider uncertainty implied by the combination of the two envelopes of uncertainty from CMIP5 and HadCM3Q. Therefore, it will be important that stakeholders appreciate this point before using results from the RCM simulations generated for Chapters 4 and 5. Detailed comparison of the range of response for the 9 GCMs selected in Chapter 5 and the wider range of response for the pooled CMIP5 and HadCM3Q ensembles is presented in the summary below (Section 7.5). For JJA, the range of HadCM3Q and CMIP5 is spanned by the Chapter 5 subset throughout the 21st century with the exception of the dry tail provided by the HadCM3Q results beyond 2050.

Whilst the 30-year averages used so far are suitable for showing that there are possible long term changes not covered by the 9 CMIP5 models used in Chapter 5, they remove much of the year-to-year variability which can lead to extremes that impact on society. Therefore we now show wider uncertainty for 1-year averages and 5-year running means (see Figures 7.4.4 and 7.4.5). For a given time point, the envelopes are based on a 5th-95th confidence interval, estimated by pooling 20-years of data from all the available model simulations, centred on that time point. In particular, the focus here is on relatively dry years as these are often driven by large scale features of the climate system. In contrast, extremely high rainfall events are linked with thunderstorms and local scale convection and are not expected to be captured by our analysis based on GCM-scale output.

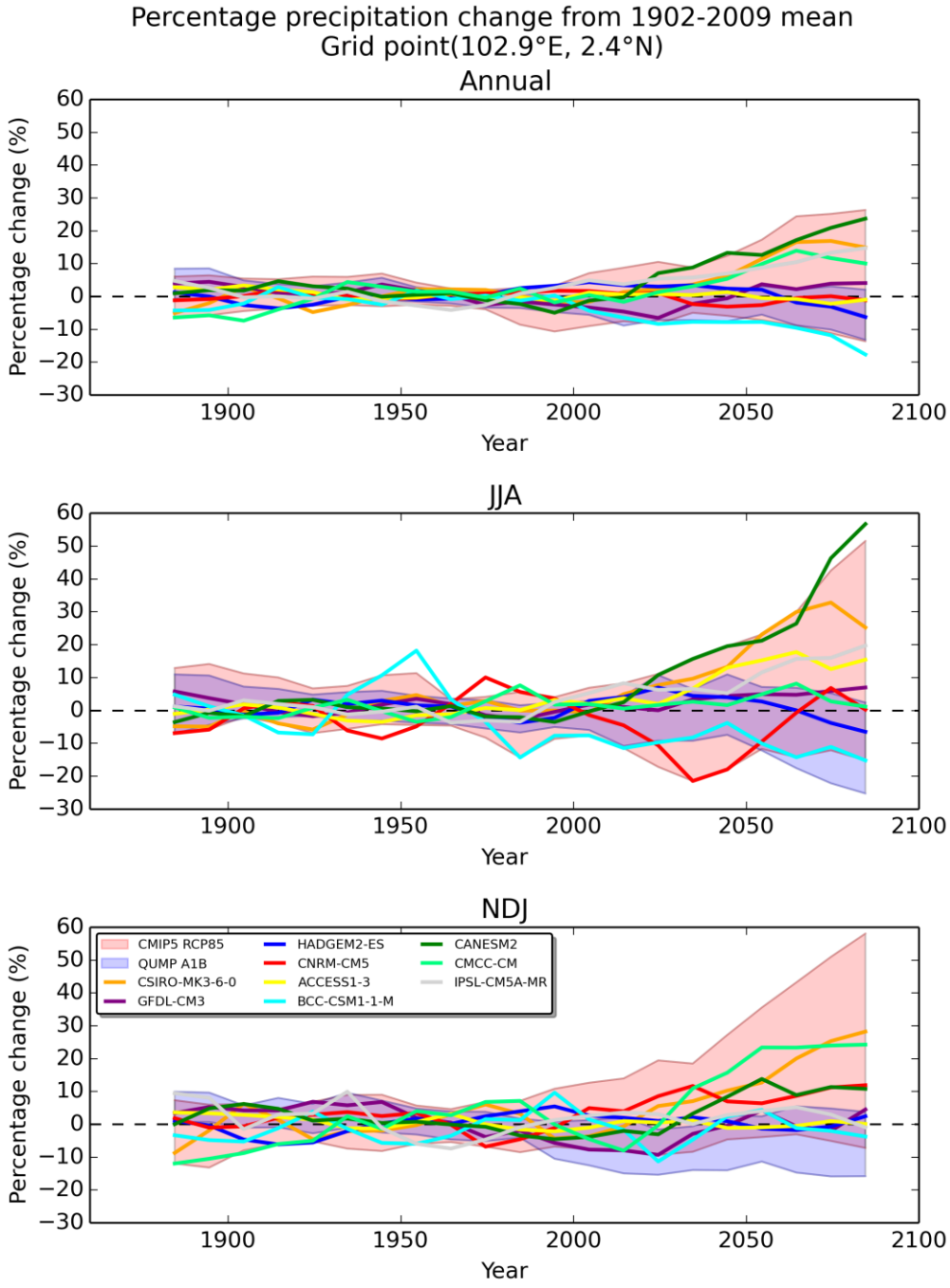

Figure 7.4.3 Percentage change in 30-year mean precipitation (stepping every 10 years) with respect to the 1902-2009 mean for the 9 CMIP5 GCMs selected in Chapter 3 under RCP8.5 emissions scenario, compared against the time-varying envelope of 30-year mean percentage precipitation changes from CMIP5 RCP8.5 (pink shading) and HadCM3Q SRES A1B(blue shading). At each time point, the envelope is the range of the ensemble after both extreme values have been removed to reduce the effect of outliers.

For NDJ annual mean percentage precipitation changes, the driest observed anomalies from the GPCP dataset are only slightly less dry than the driest years in either the CMIP5 or HadCM3Q ensembles. Therefore stakeholders who base their decisions on the worst observed cases in the 1970s for NDJ mean precipitation should find their decisions reasonably resilient to future changes, according to the current state of the art climate modelling. This also applies to 5-year running means. The subset of CMIP5 models used in Chapter 5 are successful at sampling the driest years explored by the full set of acceptable CMIP5 models. Therefore, stakeholders should be able to use the 9 regional climate simulations from Chapter 5 to sample the uncertainty in the driest years.

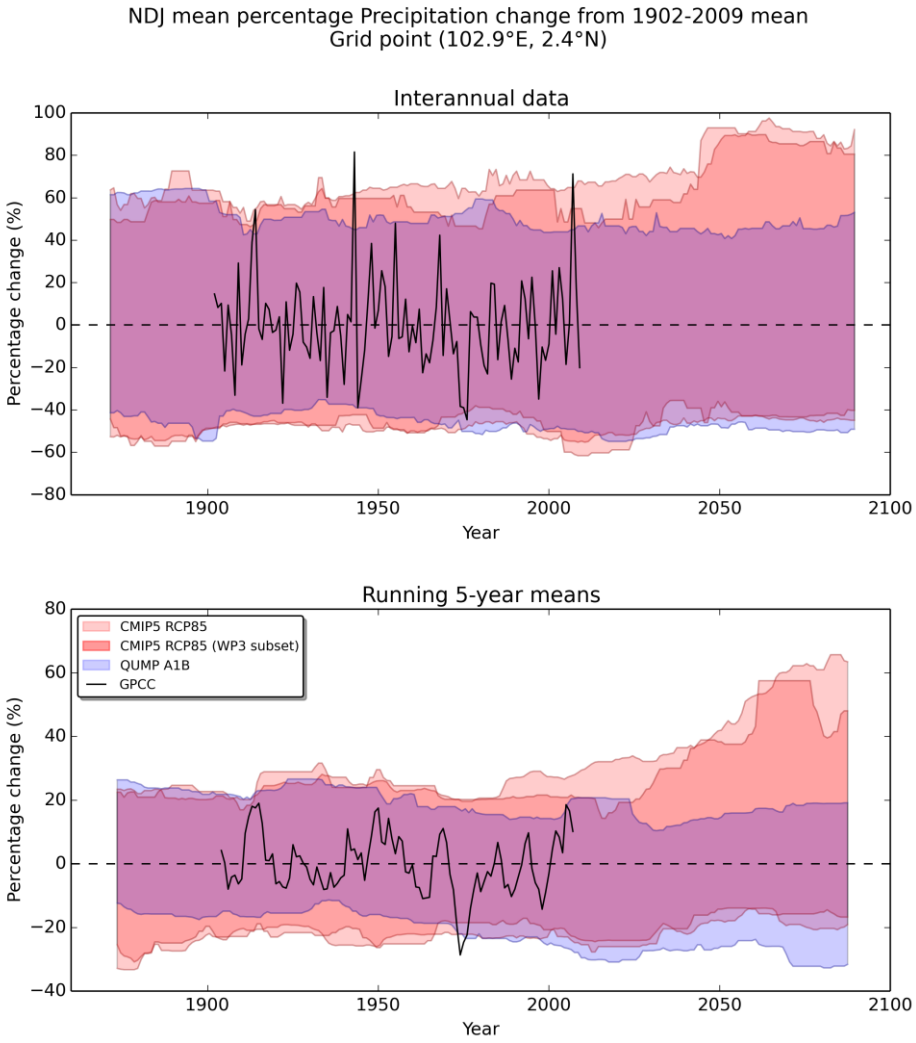

Figure 7.4.4 The top panel panel shows envelopes of year-to-year variability for percentage changes in NDJ mean precipitation relative to 1902-2009 for three ensembles at South Malaysia GCM point, found to be the best overall predictor of Singapore rainfall in Section 7.3. The lighter red shading is for the full set of CMIP5 models found to be acceptable in Chapter 3. The darker red shading is for the 9 CMIP5 models used in Chapters 4 and 5. The blue shading is for the 17 HadCM3Q PPE simulations, driven by the SRES A1B scenario. The envelopes at a given time point are the 5-95% confidence interval based on the data from a 20-year window centred on that time point; therefore the last time point at which an envelope is plotted is at 2090. Bottom panel: as for top panel, but for 5-year running means.

Finally, we repeat our caution (Section 7.2.1) that there may be structural errors common to all these climate models and therefore it is important to research the key processes involved in such dry years. Also, it is important for stakeholders to consider if a critical decision threshold, at which some adaptation plan may cease to be resilient, may exist just below the worst observed or projected case.

For JJA mean precipitation changes, for both 1-year and running 5-year averages, the HadCM3Q and CMIP5 simulations predict there could be JJA seasons in second half of 21st century drier than the observed record. In the historical period all three envelopes are consistently wider than the observed spread, so the climate models might have too much year-to-year variability in JJA. Despite this, HadCM3Q simulations do show that there could be years in the future when JJA is drier than its own historical record.

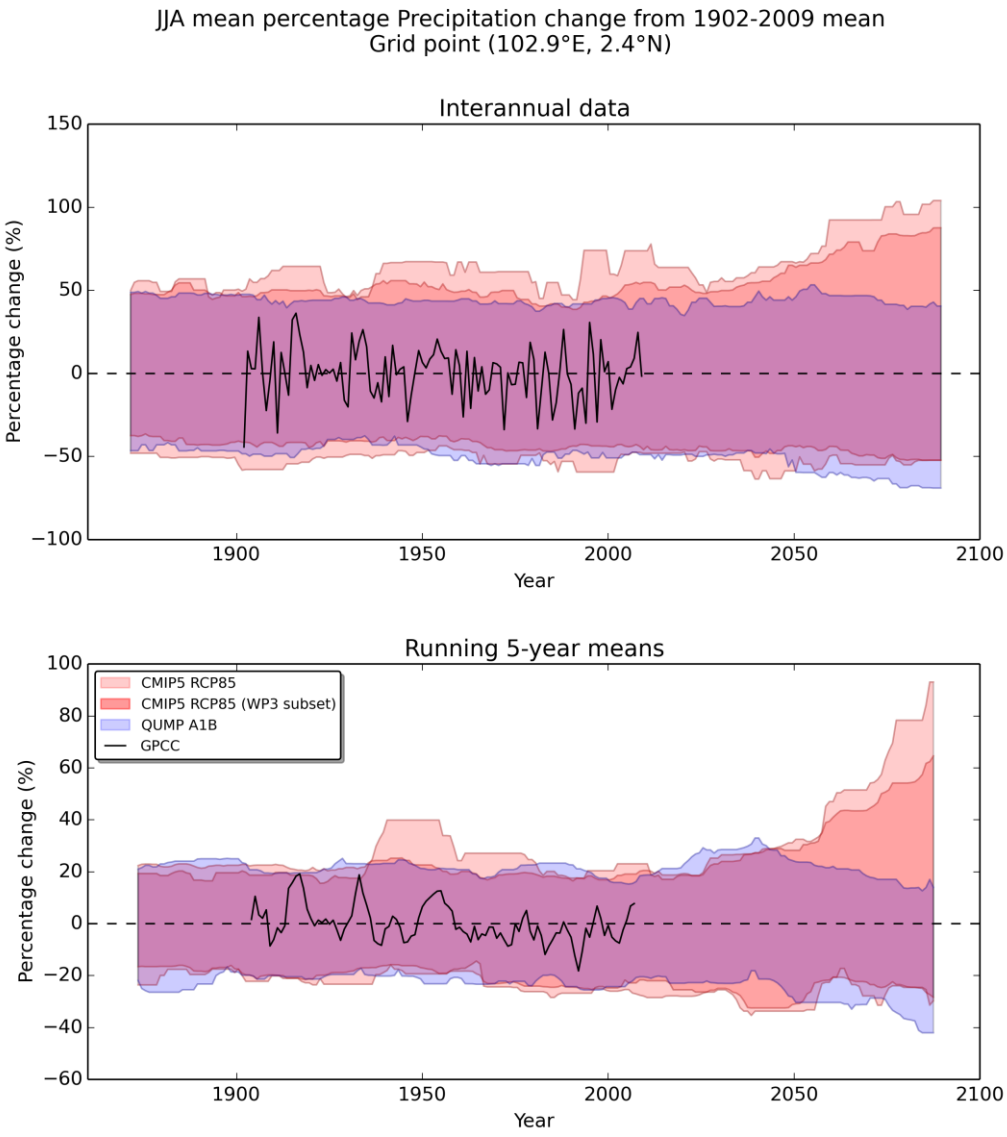

Figure 7.4.5 As Figure 7.4.4 but for JJA.

The mechanism for this in HadCM3Q simulations is linked to the enhanced surface ocean warming towards the end of the 21st century in the tropical eastern Pacific relative to the rest of the tropical oceans (see Section 7.3). There is no reason to think this

response is implausible, and so stakeholders might want to plan for an increased risk for even drier JJA seasons later in the century.

7.5. Summary

The summary is divided into two sections, on temperature and precipitation changes respectively.

7.5.1 Temperature

Projections are provided that consider the response to carbon *emissions* in the SRES A1B scenario. An important source of uncertainty arising from the role of carbon cycle feedbacks in converting emissions to atmospheric CO_2 concentrations is therefore included, in contrast to the results from Chapter 5. Singapore temperatures are projected to increase throughout the 21st century with a typical median change of 4^{o}C by 2100. The uncertainty increases throughout the 21st century. Table 7.5.1 shows that the 10-90% credible interval (the Bayesian counterpart to the confidence interval) of annual mean temperature change is 1.6^{o}C in the 2050s but grows to 2.4^{o}C by the 2080s.

Table 7.5.1: A summary of the 10th, 50th and 90th percentiles of temperature change relative to 1961-90 estimated from the 1-year average sampled data in ^{o}C that are representative of two 30-year periods centred on 2055 and 2085 and averaged across all 9 Singapore locations.

Percentile	2050s	2080s
10th	1.7	2.3
50th	2.5	3.5
90th	3.3	4.7

Section 7.6 provides some guidance on what these SRES A1B results mean for stakeholders who are using the regional climate simulations for RCP4.5 and RCP8.5 generated in Chapter 5.

Table 7.5.2: As Table 7.5.1 but for temperature changes relative to 1980-2005.

Percentile	2050s	2080s
10th	1.4	2.0
50th	2.1	3.1
90th	2.8	4.2

For Chapter 5, a different baseline period of 1980-2005 has been used. The sampled time series of temperature changes (see Sampled Data in Section 7.6) used to estimate Table 7.5.1 can be re-centred to any baseline of choice. Table 7.5.2 shows that the percentiles of temperature change relative to 1980-2005 are 0.3^{o}C to 0.5^{o}C less than the temperature changes relative to the cooler baseline period of 1961-90.

7.5.2 Precipitation

The last of the three assessment criteria for production of PDFs (see Section 7.3) failed for precipitation, as HadCM3Q model variants differed in their precipitation response to the majority of CMIP5 models. It was not possible therefore to provide reliable probabilistic projections. Instead the CMIP5 and HadCM3Q ensembles have been pooled, and statistically processed to reduce the influence of outliers to provide estimates of the wider uncertainty. This provides a wider estimate of uncertainty, which accounts for the possibility that different ensembles of alternative climate models, or model variants, may not simulate the same envelopes of plausible future outcomes. The 9 CMIP5 simulations used to the drive the RCMs in Chapter 5 can then be assessed in the context of this wider uncertainty. The precipitation results are presented at the GCM grid-box scale, and results from the RCP8.5 and SRES A1B scenario simulations (both driven by prescribed CO2 concentrations rather than from emissions) are pooled. This is acceptable to first order, because the regional precipitation changes show limited dependence on globally averaged temperature rise. In summary:

The GCM simulations used to drive the regional climate simulations in Chapter 5 have been well chosen to represent the range of annual mean precipitation changes simulated by both CMIP5 and HadCM3Q. The same applies to the July-August wet season, while noting that the HadCM3Q results suggest a slightly drier tail of plausible outcomes beyond 2050 (see below). For example, there is evidence from HadCM3Q simulations of an increasing risk after 2050 that the current observed record JJA dry record will be beaten.

For the wet season NDJ, the Chapter 5 GCM simulations become somewhat less representative of the wider spread of possible outcomes.

For JJA, HadCM3Q samples a scenario where tropical eastern central Pacific SSTs warm more rapidly than elsewhere in the tropics. This causes a more extreme reduction in precipitation over the 21st century than any of the CMIP5 models. Stakeholders are advised to consider that there is a possible scenario where precipitation changes could be drier than in the driest of the 9 RCM runs. However CNRM-CM5 in 2030 does seem to cover the HadCM3Q range at the low end.

For NDJ, there are a number of CMIP5 models not used in Chapter 5 that suggest rainfall could increase more rapidly than the two models used in Chapter 5 that have the wettest response, CSIRO-MK3-6-0 and CMCC-CM. Again, stakeholders should consider the effect of this.

7.6. Recommendations and Limitations

7.6.1 The Product

The probabilistic climate projections for 1.5m temperature change relative to 1961-90 (see cumulative distribution function (CDF) data below) are provided as a netcdf file for each seasonal (Annual, NDJ, FM, AM, JJA, SO) and monthly meaning periods, for each of the nine 12km x 12km RCM grid boxes around Singapore specified by MSS, and for a number of time points. The probabilistic projections are provided for 1-year averages for years 1861 to 2100. Also, each CDF data file is paired with another netcdf file storing an associated set of sampled data (see below). See Appendix to Chapter 7 for netcdf data formats.

This information has been used to make Figures 7.4.2 and 7.6.1 and the code to make these plots is available. A complete set of plume plots similar to Figure 7.4.2 for all 18 meaning periods has been collated as a set of pre-prepared graphics in PDF format.

7.6.2 Uses

There are a number of ways in which the probabilistic information can be used by stakeholders.

1. Compared with a limited set of plausible climate change simulations, such as those of Chapter 5, the probabilistic projections are suitable for showing a wider range of possible outcomes consistent with a more comprehensive assessment of uncertainties in drivers of future climate change consistent with current knowledge. The PDFs can therefore be used as a context for simulations like those generated in Chapter 5, that will be very useful but will under-represent the full implications of current uncertainties.
2. The probabilistic projections can provide a risk assessment for realising some specified threshold of climate change that has been identified by some earlier vulnerability assessment by the stakeholder as being a critical point when climate change has a significant impact.
3. The probabilistic projections can be sampled to provide inputs to impacts models to provide a consistent set of projections in the impact variable, although the regional climate output in Chapter 5 may be more suitable in some cases e.g. when coherency across time and space is needed.
4. The CDFs are based on individual seasons and so include year-to-year climate variability, and are comparable to historical interannual mean time series of seasonal mean observations or new seasonal mean observations as they arise. This allows the probabilistic projections to be more easily related to people's experiences (Sexton and Harris, submitted).

The product has been divided into CDF data and Sampled Data. CDF data is useful for the first and second uses detailed above. The Sampled Data is suitable for the third use, and when more than one variable is being considered (a single variable here is temperature change at a single location, a single time point, and a single meaning period).

A limitation of this product is that the probabilistic projections are provided for the A1B SRES emissions scenario used in the 4th IPCC assessment report. This was because the HadCM3Q runs are not available for RCP scenarios. We provide some advice below on how to deal with this when stakeholders are using RCP4.5 and RCP8.5.

7.6.3 CDF data

The main diagnostics that the probabilistic data can support include:

1. Generation of graphics in Section 7.4.1 and pre-prepared maps (see Figure 7.4.2) that show how the uncertainty in temperature change evolves over time
2. Plotting the probability density function (see right panel of Figure 7.6.1) that shows the relative likelihood of any given level of climate change. This is a useful visual tool.
3. Estimation of the credible interval between two percentiles e.g. the credible interval between the 5th and 95th percentiles
4. Estimation of the risk of temperature change going above or below a user-defined threshold. This can be done graphically using left panel of Figure 7.6.1, by reading off the value of the y-axis that corresponds to the threshold on the x-axis. The plot shows the probability of being below any value on the x-axis so the y-value would have to be subtracted from 1.0 to obtain the probability of going above the threshold.

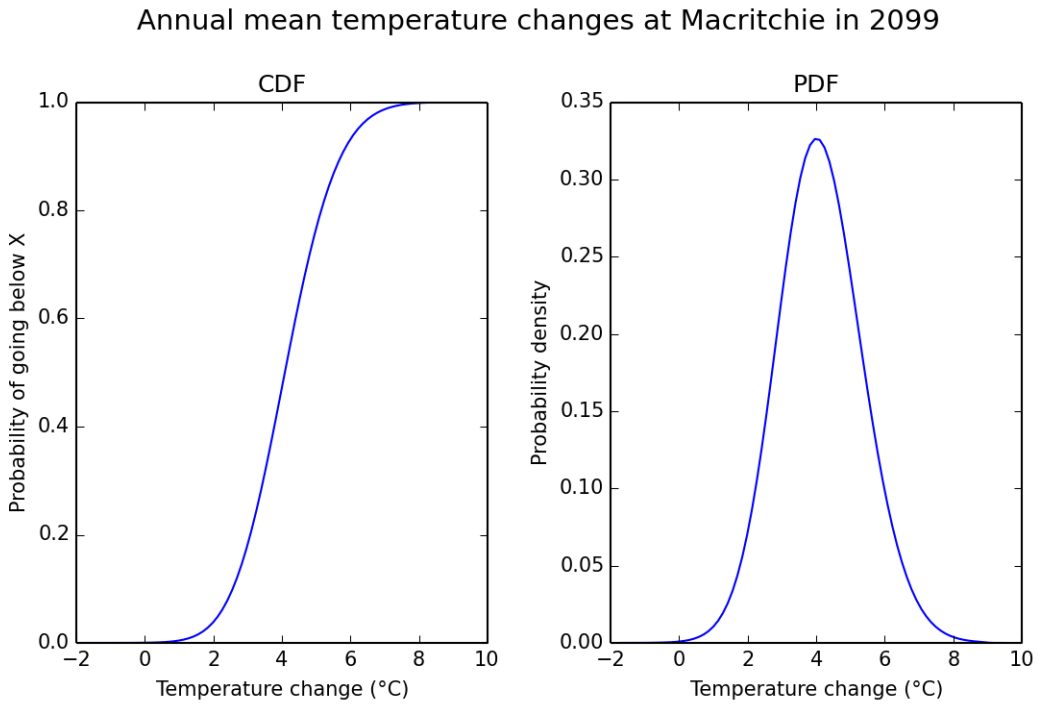

Figure 7.6.1 CDF (left panel) and PDF (right panel) for temperature changes relative to 1961-90 for Macritchie in 2099.

7.6.4 Sampled data

Sampled data has been produced to provide 2500 plausible realisations that are consistent with the probabilistic projections for 1-year averages. Compared with the CDF data, sampled data has three main uses:

- It gives users the flexibility to calculate a wider range of indicators of interest, akin to an ensemble of time series of raw climate model output. One example would be the changing risk of runs of consecutive hot or cool years at a given location. Sampled data also supports applications requiring analysis of uncertainty in more than one variable at once, in a way that requires the coherency between the variables to be accounted for. For example, coherency may be required at different locations, at different seasonal meaning periods, or at different periods in the 21st century. As the same GCM grid point i.e. “South Malaysia” is used for all meaning periods to downscale to all 9 Singapore RCM locations, the sampled data should be reasonably coherent. It should be noted that coherency across space and time is not perfectly represented in the statistical method, as the emulator required in stage one of the method (see Section 7.2 and Sexton et al 2012) is only approximately multivariate.

- Sampled data can be averaged into multi-annual means. For example, Tables 7.5.1 and 7.5.2 were based on the sampled data averaged into two 30-year mean periods, 2040-2069 and 2070-2099.

- The probabilistic projections are for temperature changes relative to 1961-90 baseline. Often this choice of baseline might be inconsistent with the one that is most useful to the stakeholder, maybe because 1961-90 is not recent enough and therefore not relevant, or because some relevant empirical data is available for a different period. The sampled data has the advantage over the CDF data that it can easily be re-centred, realisation by realisation, to any user-specified baseline period, allowing the probability distributions to be re-estimated from the adjusted sampled data. Table 7.5.2 was generated using this technique.

7.6.5 Limitations and caveats

There are several key limitations associated with the probabilistic projections, some of which are common to non-probabilistic projections such as those in Chapter 5. These are summarised in Box 7.6.1. Consequently there are caveats on how the probabilistic information can be used (see Box 7.6.2).

Box 7.6.1 Interpretation and Limitations

- The PDFs are based on model and observational evidence available around the time of CMIP3 for the SRES A1B emissions scenario. They also depend on some expert judgements, including the choice of methodology used to make them. Therefore the PDFs are liable to change in the future, given better models, revised forcing estimates, updates and additions to the observational constraints, or improvements to the methodology.
- Not all uncertainties can be quantified. For example, some earth system processes are not yet well enough understood to be included in contemporary climate models.There are no statements in the 5th IPCC assessment to suggest with any confidence that these missing sources of uncertainty are as or more important than the ones we have quantified here. The following uncertainties are not sampled:
    - external natural forcing from the solar cycle and volcanic eruptions,
    - land-use emissions and historical fossil fuel emissions,
    - ocean carbon, methane and nitrogen cycles,
    - permafrost response, although there is confidence this will increase global temperature somewhat.
- The method does not quantify systematic errors that are common to all HadCM3Q and CMIP3 climate models. Therefore, the probabilistic projections are potentially sensitive to these common systematic errors. However, in the 5th IPCC assessment, there have been modest improvements in simulation of some of the systematic errors associated with drivers of climate variability, although there is little evidence to suggest how these may affect climate change over Singapore.
- Errors in global climate model projections cannot be compensated by statistical procedures no matter how complex, and will be reflected in uncertainties at all scales.
- There is a cascade of confidence in climate projections, with moderate confidence in those at continental scale; those at 12 km resolution are indicative of climate change to the extent that they reflect large-scale changes modified by local conditions such as mountains and coasts. The level of confidence is different for different variables.
- The PDFs for temperature change are made available for the SRES A1B emission scenario. We suggest (see text below) that PDFs for RPC8.5 can be approximated by scaling up the temperature changes by a factor of 1.5. This works when the baseline period is either 1961-90 or 1980-2005 (the Chapter 5 baseline).

Box 7.6.2 Caveats on use

The PDFs are based on the statistical synthesis of multiple lines of evidence. The statistical modelling requires assumptions to achieve this. Whilst the bulk of the range of outcomes defined by the PDFs are not much affected by these assumptions, the tails of the PDFs may well be affected. Sampled data has been restricted to values within the 1st and 99th percentiles. However, we recommend that PDFs and sampled data outside the 5th and 95th percentile are not considered to be reliable and used with extreme caution, especially where the tails of the PDF are long.

- Individual sampled realisations should not be considered in isolation.
- When using 1-year PDFs, it needs to be recognised that the year-to-year climate variability comes from the HadCM3Q simulations, and will not be perfectly simulated. It may be possible to bias-correct the sampled data using historical observations.
- The multivariate statistical approach used in the method is sophisticated but approximate. Therefore, if coherency is required across locations, time points, or meaning periods, the sampled data should first of all be checked to see if it is similar to the level of coherency found in the regional climate simulations from WP3 or observations.
- We do not recommend using the probabilistic projections in an optimal cost-benefit decision-making framework as PDFs will evolve over time with new models, observations and improvements to methodology. The probabilistic projections are suitable for estimating the risk of a given level of climate change, and for providing a estimate of the range of climate change over the 21st century. These can be used in resilient decision-making frameworks.

Can we say anything about response to the RCP8.5 emissions scenario based on the probabilistic projections for the SRES A1B emissions?

As the HadCM3Q coupled model ensemble (Ensemble 2 in Table 7.2.1) has not been run for RCP8.5, it is only possible to make probabilistic projections for the A1B SRES scenario used in the fourth IPCC assessment. However, RCM simulations have been generated in Chapter 5 with the RPC4.5 and RPC8.5 emissions scenarios. So, given the PDFs are for the SRES A1B scenario, is it possible to say anything about RCP4.5 and RCP8.5?

The Met Office has generated additional 57-member perturbed parameter ensembles with HadCM3C (HadCM3Q ocean-atmosphere model, plus a terrestrial carbon cycle component as used in Ensemble 6 in Table 7.2.1) forced by A1B (Lambert et al 2013), and RCP2.6 and RCP8.5 emissions scenarios (Booth et al 2013). These emissions-driven ensembles account for the combined uncertainties in both climate sensitivity, carbon cycle feedbacks, ocean heat uptake and aerosol chemistry, the exact same set of Earth system uncertainties accounted for in the PDFs. As well as showing that

simulations forced with the RCP8.5 scenario warm more rapidly during the 21st century than when forced by SRES A1B due to higher CO_2 levels, Booth et al (2013) show that the emissions-driven simulations warm faster and have a greater spread than their concentration-driven counterparts.

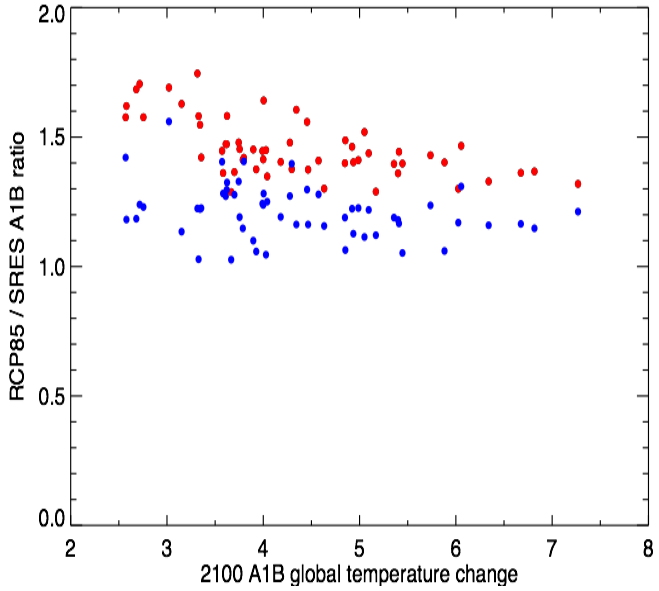

Figure 7.6.2 The ratio of global annual mean temperature changes at 2050 (blue dots) and 2100 (red dots) from RCP85 and SRES A1B scenarios for 57 members of HadCM3C ensembles from Booth et al (2013).

Given that the 57-member ensembles sample the same uncertainties quantified in the PDFs, we can look for an emergent relationship between temperature changes under SRES A1B and RCP8.5 emissions scenarios across the 57-member ensemble. If a clear relationship exists, there is potential to estimate what might happen to the temperature change under RCP8.5 based on the probabilistic projections for SRES A1B. Taking the global temperature changes from Fig 1 of Booth et al (2013) at 2050 and 2100, Figure 7.6.2 shows that the ratio between the RCP85 and SRES A1B values are approximately constant for A1B temperature changes at 2100 of 3.5^{o}C or more, for both 2050 (blue dots) and 2100 (red dots). For lower temperature changes by 2100 in SRES A1B, the ratios are somewhat higher. Therefore these ensembles allow us to suggest a 'rule of thumb' for inferring the RCP8.5 (not RCP4.5) PDFs from the SRES A1B ones.

We suggest that, to first order, the normalised response in the second half of 21st century will be roughly the same under both SRES A1B and RCP8.5, and that the main difference would be due to changes in global temperature. If so, then it is possible to suggest that a rule of thumb for approximating PDFs under RCP85 from those under SRES A1B is to scale the latter up by a factor of 1.5 for 2100 and 1.2 for 2050. Note that this could be refined if a decision maker is just interested in one particular tail; in this case for 2100, the scaling factors would be 1.7 for the lower tail and 1.4 for the upper tail. The advantage of suggesting one number is that it is simple, and that the lower tail will not scale up so much, while the upper tail will scale up slightly too much. This conservative approach widens the range somewhat , reducing the risk of this simplified adjustment underestimating the spread that one would find in properly estimated PDFs for RCP8.5. Note, that the values in Figure 7.6.2 are slightly dependent on the choice of baseline period, but that a change from 1961-90 to 1980-2005 (the Chapter 5 baseline) does not affect the values sufficiently to change the scaling factors above.

Acknowledgements

Thanks to Robin Chadwick and Peter Good for their valuable insights and good advice during the analysis of the difference between the precipitation response in HadCM3Q and CMIP5. Thanks to James Murphy and Glen Harris for reviewing the report.

References

Ackerley, D and coauthors. Sensitivity of Twentieth-Century Sahel Rainfall to Sulfate Aerosol and CO2 Forcing. (2011) J. Climate, 24, 4999–5014 DOI: http://dx.doi.org/10.1175/JCLI-D-11-00019.1

Allan, R. J., and Ansell, T. J. (2006) A new globally complete monthly historical gridded mean sea level pressure data set (HadSLP2): 1850-2003: J. Clim 19 p.5816-5842

Ashfaq, M., Skinner, C. B., and Diffenbaugh, N. S. Influence of SST biases on future climate change projections. (2011) Clim. Dyn., 36, 1303–1319. doi:10.1007/s00382-010-0875-2.

Booth, B. B. B and coauthors. (2012) High sensitivity of future global warming to land carbon uncertainties. Environ. Res. Lett. 7 8pp doi:10.1088/1748-9326/7/2/024002

Booth, B. B. B., Bernie, D., McNeall D., Hawkins, E., Caesar, J., Boulton, C., Friedlingstein, P., and Sexton, D. M. H. (2013) Scenario and modelling uncertainty in global mean temperature change derived from emission-driven global climate models. Earth Sys. Dyn., 4, 95–108. doi:10.5194/esd-4-95-2013.

Brierley C. M., Collins M., and Thorpe, A. J. (2010) The impact of perturbations to ocean-model parameters on climate and climate change in a coupled model. Clim. Dyn., 34, 325-343.

Chadwick, R., Boutle, I., and Martin, G. (2013) Spatial patterns of precipitation change in CMIP5: Why the rich do not get richer in the Tropics. J. Clim 26 11 p. 3803-3822.

Collins, M., Brierley, C. M., MacVean, M., Booth, B. B. B. and Harris, G. R. (2007) The sensitivity of the rate of transient climate change to ocean physics perturbations J. Climate 20 10 2315-2320.

Collins M., Booth B. B. B., Bhaskaran B., Harris G. R., Murphy J. M., Sexton D. M. H., and Webb M. J. (2011) Climate mode, errors, feedbacks, and forcings: a comparison of perturbed physics and multi-model ensembles. Clim. Dyn., 36 (9-10), p. 1737-1766, DOI:10.1007/s00382-010-0808-0.

Dee, D. P. and coauthors (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 137 (656) p. 553-597 doi:10.1002/qj.828

Flato, G., J. and co-authors. (2013) Evalutation of climate models. IPCC WGI Fifth Assessment Report, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, chapter 9

Harris, G.R., Sexton, D. M. H., Booth, B. B. B., Collins, M., Murphy, J. M., and Webb, M. J. (2006) Frequency distributions of transient regional climate change from perturbed physics ensembles of General Circulation Model Simulations. Clim Dyn 27 p. 357-375

Harris, G. R., Sexton, D. M. H., Booth, B. B. B., Collins, M., & Murphy, J. M. Probabilistic projections of transient climate change. (2012) Clim. Dyn. doi:10.1007/s00382-012-1647-y.

Hawkins, E., Joshi, M. and Frame, D. (2014) Nature Climate Change, 4. pp. 646-647. ISSN 1758-678X

Jones, P. D., D. H. Lister, T. J. Osborn, C. Harpham, M. Salmon, and C. P. Morice (2012) Hemispheric and large-scale land surface air temperature variations: An extensive revision and an update to 2010, J. Geophys. Res., 117, D05127, doi:10.1029/2011JD017139.

Lambert F. H., Harris G. R., Collins M., Murphy J. M., Sexton D. M. H., and Booth B. B. B. (2013) Interactions between perturbations to different Earth system components simulated by a fully-coupled climate model, Clim. Dyn. 41 (11-12), p.3055-3072.

Carol F. McSweeney, Richard G. Jones, and Ben B. B. Booth (2012): Selecting Ensemble Members to Provide Regional Climate Change Information. J. Climate, 25, 7100–7121. doi:10.1175/JCLI-D-11-00526.1.

Meehl, G. A. et al. (2007) The WCRP CMIP3 multi-model dataset: A new era in climate change research. Bull. Am. Meteorol. Soc. 88, 1383-1394.

Murphy, J. M. and Sexton, D. M. H. and Barnett, D. N. and Jones, G. S. and Webb, M. J. and Collins, M. and Stainforth, D. A. (2004) Quantification of modelling uncertainties in a large ensemble of climate change simulations Nature 430 768-772.

Murphy, J. M. et al . UK climate projections science report: Climate change projections. (2009) (Met Office Hadley Centre, Exeter, UK).

Murphy, J. M. and co-authors. Transient climate changes in a perturbed parameter ensemble of emissions-driven earth system model simulations. (2014) *Clim. Dyn.* doi:10.1007/s00382-014-2097-5.

Nakicenovic, N. et al. (2000) Emission scenarios. A special report of Working Group III of the Intergovernmental Panel on Climate Change. (Cambridge University Press). 599pp.

Reichler, T. and Kim, J. (2008) How well do coupled models simulate today's climate? Bull. Am. Meteorol. Soc., 89, 303–311. doi:10.1175/BAMS-89-3-303.

Rougier, J. (2007) Probabilistic inference for future climate using an ensemble of climate model evaluations. Clim. Change 81, 247-264.

Rougier, J., Sexton, D. M. H, Murphy, J. M., and Stainforth, D. A. (2009) Analyzing the climate sensitivity of the HadSM3 climate model using ensembles from different but related experiments J. Climate 22 13 1327-1353.

Sexton, D. M. H., Murphy, J., Collins, M., & Webb, M.. Multivariate prediction using imperfect climate models part I: outline of methodology. (2011) Clim. Dyn.. DOI: 10.1007/s00382-011-1208-9

Sexton, D. M. H., and Harris, G. R., The importance of including variability in climate change projections used for adaptation (submitted to Nature Climate Change).

Smith, D. Scaife, A. and Kirtman, B. (2012) What is the current state of scientific knowledge with regard to seasonal and decadal forecasting? ERL 7 DOI:10.1088/1748-9326/7/1/015602

Spencer, H., R. Sutton and J.M. Slingo (2007) El Nino in a coupled climate model: Sensitivity to changes in mean state induced by heat flux and wind stress corrections. J. Clim. 20, 2273-2298.

Webb, M. J and coauthors (2006) On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. Clim. Dyn. 27:17–38. DOI:10.1007/s00382- 006-0111-2