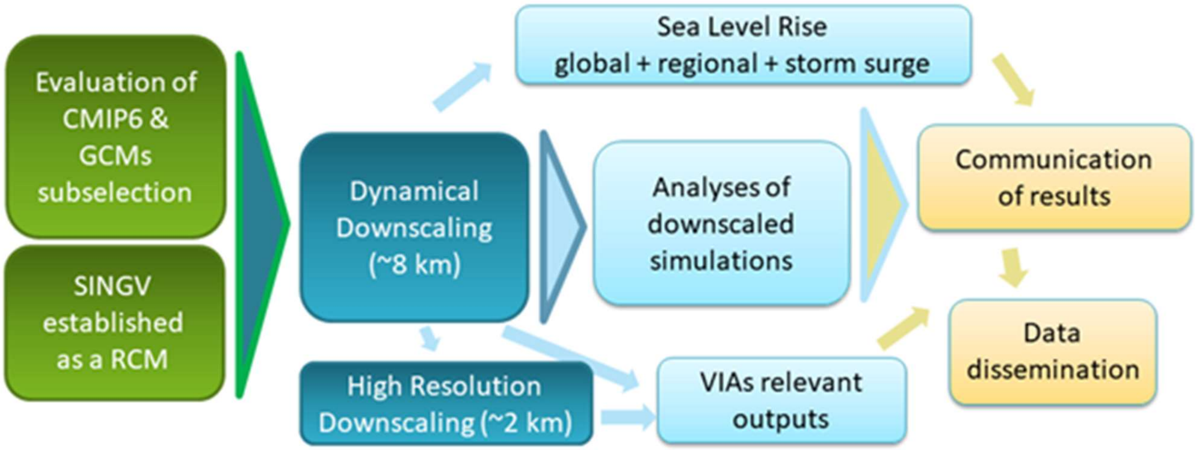

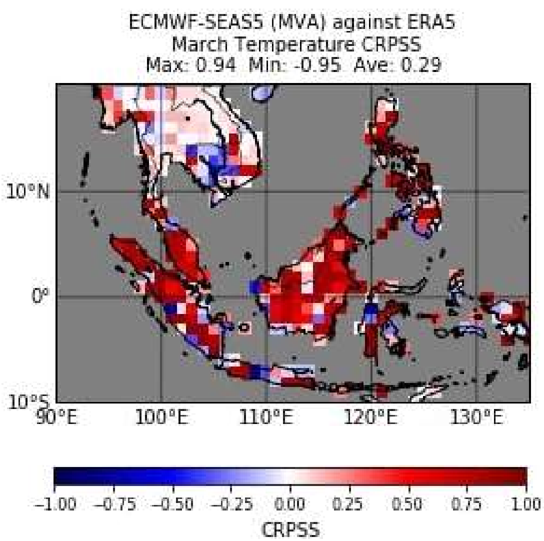

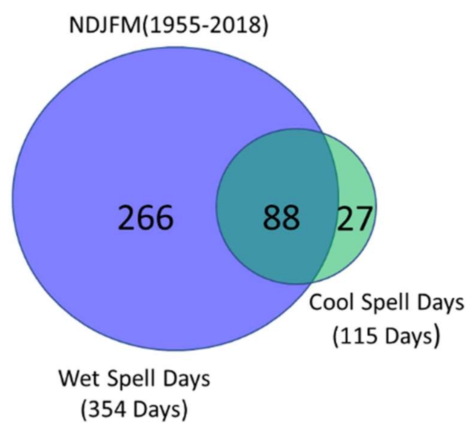

MSS Research Letters

Issue #4, December 2019

Meteorological Service Singapore

DISCLAIMER OF WARRANTIES AND LIABILITY

1. While National Environment Agency Meteorological Service Singapore (NEA MSS) has made every reasonable effort to ensure that the information contained in this publication has been obtained from reliable sources, NEA MSS shall not be responsible for any errors or omissions, or for the results obtained from the use of such information.

2. NEA MSS shall also not be liable for any damage or loss of any kind howsoever caused as a result of any reliance on the contents of this publication.

EDITOR'S NOTE.

After two years of publication, I am pleased to present the fourth issue of MSS Research Letters. In line with the end of the year, the theme of this issue is 'looking towards future climate research'. The first letter covers the preparation of Singapore's third national climate change study, an integral portion of the Centre for Climate Research Singapore's (CCRS's) research in the coming years. The other two letters are co-written by university interns to CCRS. These students spend two to five productive months at CCRS working on small research projects, contributing to the centre's research as well as building their own experience as researchers. We wish these future scientists all the best for the remainder of their studies and their future careers.

This issue also marks changes to the staff behind MSS Research Letters. In December, Dr Bertrand Timbal will be departing CCRS. Dr Timbal has been a key driving force in the launching of MSS Research Letters, and therefore we must thank him for his support for this excellent initiative. We also have Dr Hindumathi Palanisamy joining the editorial team as lead editor. Thank you for volunteering for this role, and we wish you well for future issues.

To all the authors and reviewers of this issue, I would like to thank you for working together to provide these interesting contributions. Thank you especially to our external reviewers: Chris Gordon, Gabriel Lau, and Andrew Robertson. Your contributions and guidance to ensuring the scientific integrity and improvement of the MSS Research Letters are much appreciated.

To all our readers, both inside and outside Meteorological Service Singapore, we hope you will enjoy this issue of MSS Research Letters, and that you will consider submitting suitable material for subsequent issues.

Warm regards,
Thea Turkington
Editor, MSS Research Letters

Cover figures: top – Schematic of the critical components to be undertaken in V3 (SINGV as a Regional Climate Model to deliver Singapore's 3rd National Climate Change Study, page 8); bottom left – Skill of monthly temperature forecasts using mean/variance adjustment calibration compared to the raw forecast (Calibration of ECMWF seasonal SEAS5 model's monthly temperature re-forecasts over the Southeast Asia region, page 25); bottom right – Overlap between days with wet spells and those with cool spells in Singapore (Exploration of the clustering of moisture flux fields and its relation to wet and cool spells in Singapore, page 16)

TABLE OF CONTENTS

SINGV AS A REGIONAL CLIMATE MODEL TO DELIVER SINGAPORE'S 3RD NATIONAL CLIMATE CHANGE STUDY

Bertrand Timbal, Venkatraman Prasanna, and Muhammad Eeqmal Hassim

Climate Modelling and Prediction Section, Centre for Climate Research Singapore

INTRODUCTION

In 2015, Singapore's Second National Climate Change Study (V2) was publicly released (Gordon et al. 2015; Marzin et al. 2015). V2 was a collaborative project between scientists from the Centre for Climate Research Singapore (CCRS) and the UK Met Office (UKMO) Hadley Centre. It provided a national authoritative perspective on future risks due to climate change, in sync with the international community and in particular the release of the 5th Assessment Report of the Intergovernmental Panel for Climate Change (IPCC, AR5, IPCC 2014). V2 relied on the international effort to generate global climate model (GCM) simulations of current and future climate, the so-called CMIP (Climate Model Intercomparison Project) dataset in its 5th phase (CMIP5), from which historical simulations and future climate projections for Singapore and the Western Maritime Continent (WMC) up to 2100 were produced.

As the international community has been progressing in the global science of climate change projections, it is important to keep up with the latest international developments. The IPCC is currently in its 6th cycle of global assessment, and the IPCC AR6 series of reports will be produced in 2021 and 2022 starting from the Working Group I (WGI), through the other two WGs and up to the Synthesis Report (SR) to be released in March 2022. It has thus been decided in Singapore to deliver an update in the form of a 3rd National Climate Change Study (or V3 for short) in sync with the IPCC calendar, aiming for the release of a new stakeholder report in sync with the IPCC cycle and SR release in early 2022.

Some of the important new international scientific developments need to be captured by V3. The first has to do with future Representative Concentration Pathways (RCPs), which are based on plausible global emission scenarios that take into account socio-economic and technological assumptions of how the future could evolve (van Vuuren et al. 2011). V2 used two plausible RCPs, RCP8.5 and RCP4.5, as global emission scenarios to illustrate the dependence of the projected changes on the trajectory of anthropogenic emissions in the foreseeable future. RCP8.5 is roughly a business-as-usual future scenario, while RCP4.5 involves some important global mitigation measures. However, these two RCPs not compatible with the newly signed Paris Agreement, and hence there is a need to explore a more stringent RCP, such as RCP2.6 which has a 50% chance of limiting the global warming below a 2°C target.

The second new development to be considered involves the underlying Global Climate Models (GCMs) database, which provides the scientific basis to explore global climate change projections, has continued to evolve. CMIP is now in its 6th instalment (CMIP6; Eyring et al. 2016). Since V2 was based on CMIP5 GCM simulations, it is therefore important that V3 makes use of the latest set of CMIP6 model simulations. In particular, early analysis of the CMIP6 database indicates that GCMs are now more sensitive and display a stronger warming in response to anthropogenic forcings (CMIP6; 2019). CMIP6 represents a 4-fold increase in the size of the database assembled compared to CMIP5, offering a large set of GCMs. More importantly, CMIP6 is populated with new types of experiments, which are directly in response to science questions. These new type of experiments offer the possibility of novel assessment of future risks related to climate change that V3 should endeavour to make use of. This is particularly relevant in the context of Singapore when it comes to sea level rise projections, as specific experiments will provide information about the various components relevant to sea level (e.g. change in the cryosphere).

Besides the progress in climate science across the international community, the research landscape in Singapore has also evolved. The most consequential change is the development over the last five years of a new limited area model, SINGV, developed in partnership between CCRS and the UK Met Office (Huang et al. 2019). SINGV is now run operationally by the Meteorological Service Singapore (MSS) and hence benefits from daily scrutiny by forecasters, and its performance is assessed using a range of objective evaluation metrics. This gives a very strong basis to establish SINGV as the Regional Climate Model (RCM) of choice to complete a new set of National Climate Change projections.

THE SINGV MODEL

As part of V2, GCM simulations were dynamically downscaled using the RCM developed by the UKMO Hadley Centre (HadGEM3-RA). In total a subset of 9 selected GCMs were downscaled to a spatial resolution of 12 km over the Southeast Asian region. At 12 km, important topographical features that affect the local climate such as coastlines and mountains across the WMC become much better resolved than in GCMs. For example, the island of Singapore is not represented in most current GCMs (Hassim et al. 2016).

In addition to the main V2 simulations at 12 km resolution, the first version of SINGV (SINGV v1) was used to run experimental 10-year convection-permitting simulations at 1.5 km resolution over a much smaller domain (390 x 390 grid points) centred on Singapore. This is turn was nested within a 4.5 km intermediate domain (800 x 800 grid points) inside the big 12 km RCM domain (see Figure 1). SINGV v1 was based on the most recent version of the UK Met Office 1.5 km numerical weather prediction (NWP) model at the time, UKV (Kendon et al. 2014), albeit with a number of enhancements. These are: i) using the ENDGAME dynamical core instead of New Dynamics; ii) a 'grey-zone' blended boundary layer scheme; iii) a revised warm rain microphysics scheme; iv) a shallow convection parameterisation; v) inclusion of graupel as a predicted variable; vi) increasing the number of model levels to 80 and vii) using the 100m SRTM (Shuttle Radar Topography Mission) dataset to derive the mean orography.

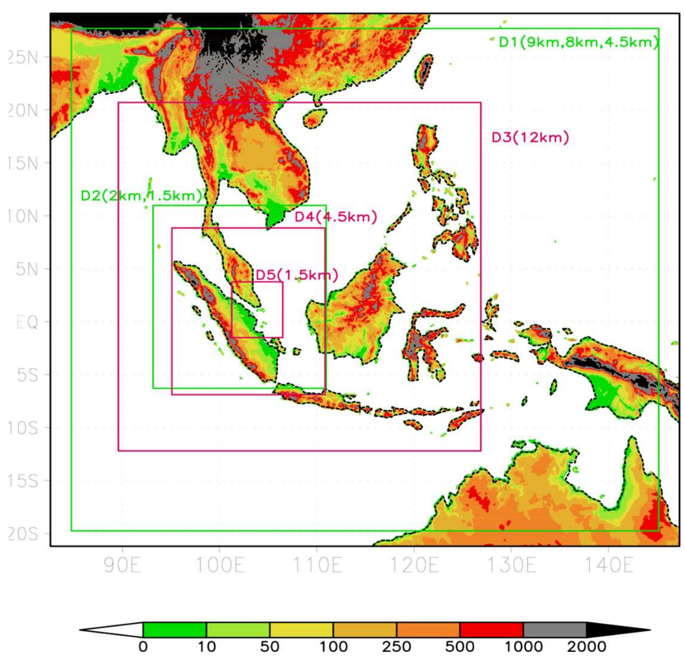

Figure 1 Map of the Western Maritime Continent showing land elevation (coloured background in m) with the various domain sizes considered: during V2 when the RCM was integrated at 12 (D3), 4.5 (D4) and 1.5km (D5), and tested for V3: at coarse resolution (D1: either 9, 8 or 4.5 km) and fine resolution (D2: either 2 or 1.5 km).

Since 2013, SINGV has continuously evolved as part of the development of a new NWP system for Singapore. From July 2019, SINGV became operational at version 5 (SINGV v5) and has since been issuing numerical weather forecasts for Singapore and the wider Malay Peninsula/Sumatra region. A time evolution of the SINGV model is presented in Figure 2.

In 2018, preliminary experiments were conducted to explore SINGV's potential as an RCM for the WMC; several versions were tested in-line with the ongoing development of the NWP system (Huang et al. 2019). Results presented here are primarily for versions v4.1 and v5.0, and these are compared to SINGV v1, used earlier as part of V2. At the time of writing, SINGV v5 is based on the Unified Model (UM) version 11.3, and is the basis for the tropical version of the UM known as RA1T (Regional Atmosphere 1 – Tropical). Performance evaluations among the different versions suggest that SINGV v5 is likely to be the definitive version of the RCM proposed for the delivery of V3. Some tests were also done with the most recent physics package that will form the next tropical version of the UM (RA2T), known here as SINGV v6.

To evolve SINGV from an NWP model to an RCM required further evaluation and tuning. These were needed in order to: (i) ensure the model is stable over long integrations for the likely domain(s) of interest; (ii) access suitable ancillary files such as the appropriate land-sea mask, orography, land cover type, time-evolving Sea Surface Temperatures (SSTs); and (iii) test the model with different boundary conditions to ensure it will be suitable as a downscaling tool embedded within a GCM.

Besides these technical aspects, some considerations have to be made to ensure the suitability of the tool for investigating scientific questions. These include the ideal domain configuration (balancing resolution with computing cost and taking into account the suitable ratio between outer domain boundary conditions, horizontal spacing and inner domain resolution) and suitable model configuration providing the most realistic regional climate patterns and characteristics.

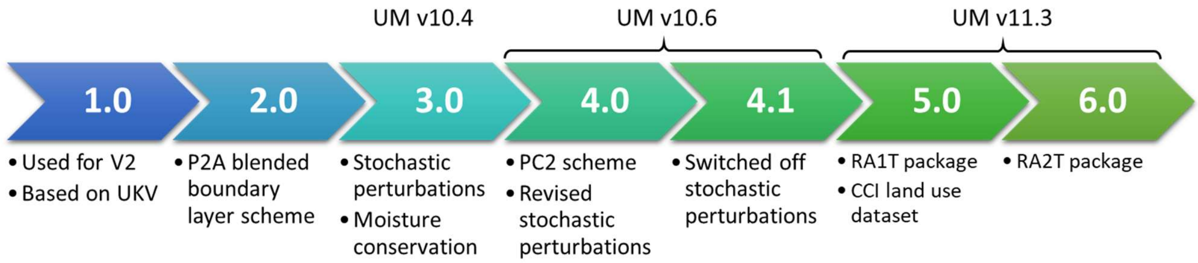

Figure 2 Schematic of the evolution with time of the SINGV model versions, indicating the major changes in the physics for each version and its relationship to the base Unified Model (UM) version used.

A series of integrations, up to a month long, of the model were completed to ensure that the technical aspects were successfully mastered and to decide on the optimal set-up to underpin the scientific research required as part of delivering V3. Both January and July were used in the tests in order to sample the annual cycle, and for some tests three years were used (2001, 2002 and 2003) to avoid results being unduly biased due to natural variability. A brief summary of the various tests is as follows.

The first climate mode experiments attempted with SINGV were with version 4.1, in which January 2001 was simulated with the latest ECMWF ReAnalysis, ERA5 (Hersbach et al. 2019; SST updated at 3 hour intervals). Three different grid resolutions were tested: 9, 4.5 and 1.5km. Both 9 and 4.5km were on a large domain (D1 close to the domain chosen by the international community (CORdinated Experiments for South-East Asia: CORDEX-SEA): 18S-26N; 90E-142E, see Figure 1) while the 1.5km simulation was on a small domain (D2 which was close to the domain used operationally for SINGV as an NWP model: 5S- 8N 95E-109E, see Figure 1).

To ensure the model's stability, gravity wave (Rayleigh) damping was implemented at model levels above 26km to minimise the reflection of upward-propagating gravity waves generated by convection. Over the D1 domain at 9km resolution, the model was tested both with and without parameterised convection. Early simulations were halted due to the model blowing up on a few occasions. The numerical instability was overcome by reducing the duration of the time step (the normal duration was 3 minutes, and on some occasions a time step as short as one and a half minutes was required).

After establishing the stability of the model over the chosen domains, integrations of two versions of SINGV, v4.1 and v5, were tested for 10 days in Jan 2001, driven by ERA5. Apart from changes to the physical package, an important difference between v4.1 and v5 is the vegetation coverage data used. SINGV v5 uses vegetation classification from the European Space Agency Climate Change Initiative (CCI), while v4.1 uses data from the International Geosphere-Biosphere Programme (IGBP). This was noted to be important for some regions where the classification of vegetation types was noticeably different (e.g. over Thailand). The CCI vegetation dataset has now become the default for the newer version of the UM model, and the local differences observed in the test run are likely to remain an improved feature in simulations with newer versions.

The final set of testing was performed using SINGV v5.0, the most recent version available from the UM model (version 11.3) and which included the most recent physical package (RA1T) that was accessible. SINGV v.5 had been tested in Singapore as part of the operational development of SINGV. A newer physical package (RA2T) was provided by the UM partnership and was also tested (SINGV v6). Only small changes in the physical package are of relevance to the region of the WMC, and so far no noticeable differences have emerged from the tests conducted. Hence results for SINGV v6 are not presented here, and it is likely that the version of SINGV chosen for V3 will remain SINGV v5.

The tests consisted of month-long simulations for both January and July, across the years 2001, 2002 and 2003. A 3-day spin-up time was used and simulations were completed for both D1 and D2 domains with a resolution of 8km and 2km, respectively. The Globe Digital Elevation Model (DEM) orography (30m resolution) was used for the simulation, the vegetation fraction was from the CCI dataset, although the land-sea was classified using the IGBP land-sea mask for both 8km and 2km simulations. The SST fields are updated at 3 hour intervals from the ERA5 skin temperature fields. Both 8km and 2km simulations were done with the explicit representation of convection.

Simulations at 2km in the D2 domain were forced by outputs from the 8km simulations on the D1 domain. For the simulations at 8km in the D1 domain, tests were conducted using ERA5 lateral boundary conditions, using either the 137 vertical model levels or the 37 standard pressure levels (from 1000hPa to 1hPa). The second setup was a downgrading of the quality of the information provided to the inner model, and was conducted in order to test a setup more aligned with what would be possible using GCM data from the CMIP database. Overall, only small differences were identified between the two setups (full 137 vertical model levels or reduction to the 37 standard pressure levels) of boundary conditions.

TESTING OF SINGV CONFIGURATION AS A SUITABLE RCM

One of the primary objectives of testing the new versions of SINGV was to establish if some of the known shortcomings of the RCM used as part of V2 were improved upon using more recent versions of SINGV. In particular, it has been well documented that extreme rainfall projections over land from V2 had a low level of confidence due to the inability of the model to reproduce the diurnal cycle of convection and reproduce the upper tail of the cumulative distribution function of daily rainfall (i.e. the wet extremes; Hassim et al. 2016).

The current evaluation was conducted with the first new version tested, SINGV v4.1 with and without convection parameterised at 9km resolution (Figure 3). Compared to SINGV v1, run at a higher resolution of 4.5km, SINGV v4.1 produced a more realistic peak of diurnal rainfall both in terms of timing and intensity. It was also evident that with the convection parameterisation switched on, the model initiated convection too early in the day and peaked too early compared to that shown in TRMM precipitation data. However at this coarse resolution of 9km, the magnitude was closer to what was observed. For all configurations with SINGV v4.1, an improvement in the diurnal cycle of convection over land was noted compared to the earlier SINGV v1.

The differences between SINGV v4.1 with and without convective parameterisation at 9km resolution were evaluated further by comparing the daily totals for the month-long integration in January for all the continental points across the larger domain D1 (Figure 4). While both configurations of SINGV v4.1 exhibited daily variability on par with observations (using TRMM estimates), SINGV v4.1 with parameterised convection appears to have a systematic overestimation of the daily rainfall. This result, combined with the previous finding that the daily timing of maximum precipitation is better captured when the convection is explicit, strongly suggests that even at a very coarse resolution of 9km the model is already “convection-enabling” and is performing better without the convection parameterisation. This result was confirmed with the SINGV v5 model run at 8km with explicit convection, which also matched observations for daily rainfall totals (Figure 4). Furthermore, this result is consistent with other studies for Western Africa using the UM model with a 4.5km horizontal resolution (e.g. Berthou et al. 2019) and for the WMC using the WRF model (Argueso et al. 2020).

Having assessed that recent versions of the SINGV model show a marked improvement compared to SINGV v1 and that the recent versions are performing as well and in fact better without the convection parameterisation even at coarse resolution (9km), two

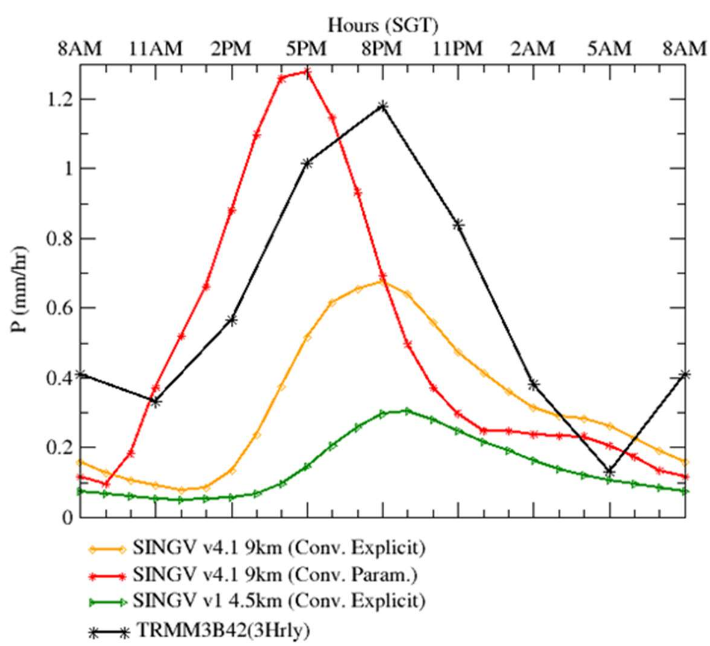

Figure 3 Diurnal cycle of land-only rainfall averaged across the D2 domain (see Figure 1 for details) for the earlier version of SINGV (v1) used as part of V2 and showing comparison with a more recent version of SINGV (4.1) with either explicit or parameterised convection and also TRMM observations. Diurnal cycles are computed for the month of January 2001 for both observations and SINGV simulations.

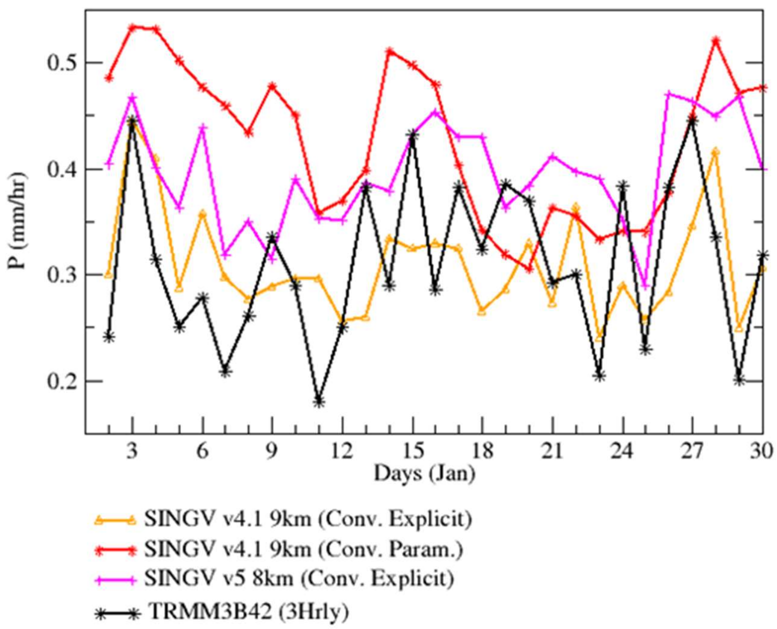

Figure 4 Day-to-day variability of daily total rainfall for land-only points averaged across the D1 domain (see Figure 1 for details) for different versions of SINGV (4.1 and 5.0) with explicit convection and comparison with parameterised convection for SINGV v4.1 and TRMM observations . Daily values are from a single month of simulation (January 2001).

questions remain to be evaluated: 1) the added value of running the model at much higher resolution (<= 2km) and 2) the selection of the most suitable version of the model for V3. The limited testing so far does not give a definitive answer to these two questions. This is particularly critical for the model outputs within the domain in the vicinity of Singapore (D5, Figure 1), as the ability of SINGV to reproduce the diurnal cycle of rainfall for land points across the very small domain varies amongst the model versions with explicit convection (v4.1, v5, as well as v6, which is not shown) giving no clear answer as to which might be the most suitable setup. These versions also seem to perform worse than those with parameterised convection. However, there are also large uncertainties in the observations (Figure 5). Nonetheless, there are some indications that for the same version of the model, a higher resolution tends to give a peak that is earlier and also has a higher daily maximum e.g. v4.1 between 1.5 km (green curve) and 9 km (yellow curve) and v5 between 2 km (pink curve) and 8 km (blue curve). However, the differences are small and in light of the very high computing cost, it may not be warranted to perform such high resolution simulations. Similarly, the differences between v5 and v4.1 are small but again point towards a small improvement with an earlier peak of daily maximum

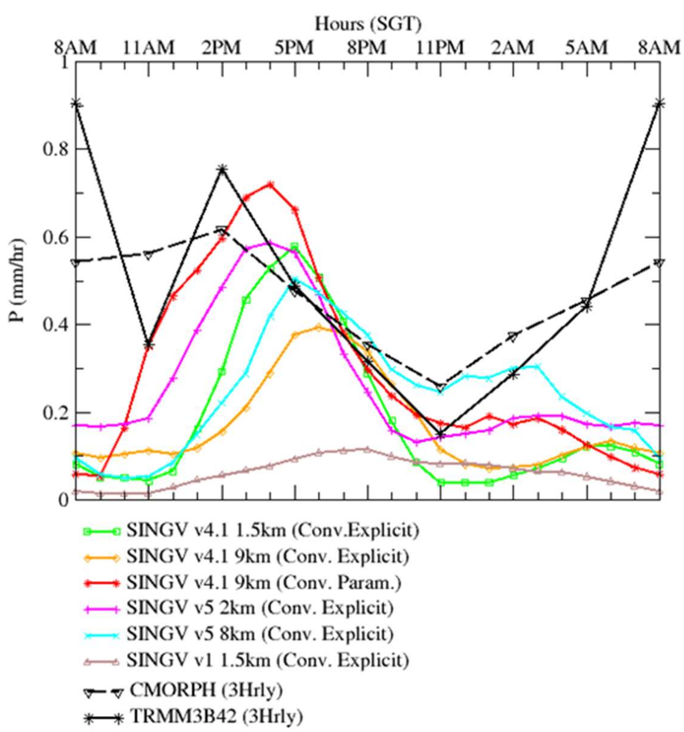

Figure 5 Diurnal cycle of land-only rainfall averaged across the D5 domain (see Figure 1 for details) for the various versions of SINGV (4.1, 5.0) for low resolution (8 or 9 km) and fine resolution (2 or 1.5 km) and compared with both TRMM (plain) and CMORPH (dashed) precipitation observations. Diurnal cycles are computed for the month of January 2001 for both the observations and SINGV simulations.

rainfall, thus suggesting that it makes sense to use SINGV v5, the most recent version and which has also been fully tested for NWP purposes.

The conclusion emerging from this series of tests to establish the most suitable version of SINGV as the RCM of choice for V3 indicates that SINGV v5 should be used for a large domain encompassing the WMC, for which the model is now demonstrably stable and benefits from up to date ancillary files. A resolution of 8 km, with the convection explicitly represented by the model instead of being parameterised, appears to be a suitable configuration. It performs better than what was used for V2 (HadGEM3-RA with a resolution of 12 km) as well as the first version of SINGV, which was integrated for short time-slices at 4.5 and 1.5 km. These findings will need to be confirmed with longer integrations using ERA5 as boundary conditions (so far only month-long integrations have been performed) and focusing on other aspects than the diurnal cycle of convection. These will be spelled out in the V3 plan described in the next section. However, at this stage of the testing, it is not obvious that the additional small gains in terms of producing a better diurnal cycle of rainfall when the resolution of SINGV v5 was increased to 2 km, warrants the integration of long and costly climate simulations on a smaller domain surrounding Singapore.

Planning of the 3rd National Climate Change Study

The aim of V3 is to conduct a national assessment of the effect of climate change on Singapore and the surrounding region, the WMC, based on the latest climate projections made available by the international community as part of CMIP6, in order to be consistent with the projections informing the 6th Assessment Report of the IPCC.

The remainder of the programme will run until 2022 when a high level report aimed at stakeholders and describing the new set of climate change projections will be released in conjunction with the IPCC synthesis report in March 2022. This stakeholders' report will be the main output of the entire programme and will serve to inform stakeholders about future projections and demonstrate the relevance of the latest IPCC report within the context of Singapore. The report will likely be supported by several publications in peer-reviewed literature to establish the scientific credentials of the underpinning work supporting the V3 projections.

Beside the documentation of the work, it is expected that a full set of climatological data tailored to stakeholder needs will be produced and made available. The variables of interest and their spatial and temporal resolutions will need to be decided in consultation with stakeholder agencies and will include, but not be limited to, temperature, rain and sea level information.

In order to successfully deliver new scientific insights and state-of-the-art national climate change projections, a number of mandatory components of the programme must be completed in the logical order depicted in Figure 6.

To start with, a limited number of GCMs will be selected and evaluated from the international datasets. These GCMS will come primarily from CMIP6, but it might be relevant to also include some of the models from CMIP5 which were used during V2 to support a continuity in the understanding of the uncertainties in the projections between V2 and V3. A suitable and robust methodology was established as part of the delivery of V2 where 9 GCMs were selected from CMIP5 that had an appropriate representation of the current climate and covered the range of sensitivities in global

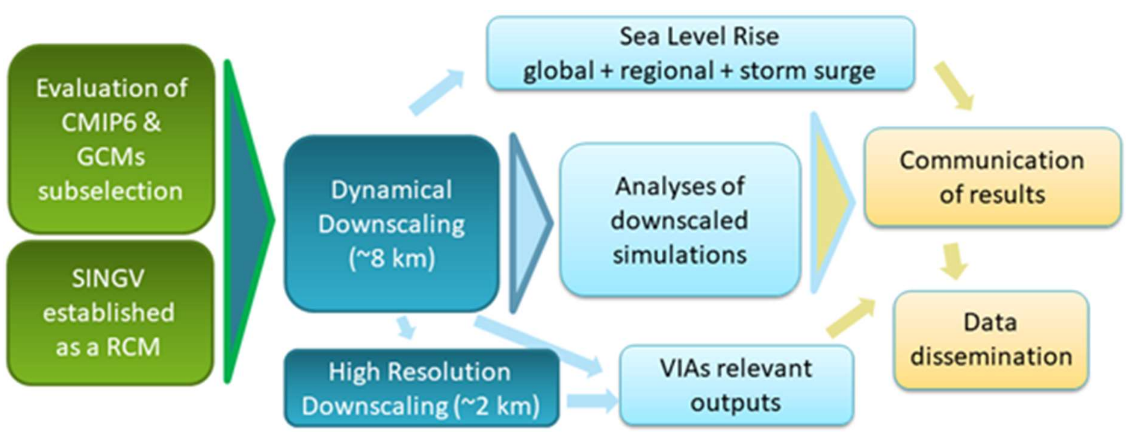

Figure 6 Schematic of the critical components to be undertaken to deliver V3 (Note: VIAs stands for Vulnerability and Impact Assessments).

response to external forcing (Marzin et al. 2015). By and large, the same methodology will be applied to CMIP6, ensuring that all models considered represent the key features of Southeast Asia's baseline climate realistically. In particular, the selected models must represent large scale climate circulation features relevant to Southeast Asia (e.g. the behaviour of the Monsoon, such as its annual cycle across the Maritime continent) and relevant features of interannual variability (e.g. the El Niño Southern Oscillation, ENSO, and the Madden-Julian Oscillation, MJO). In addition, an important aspect of exploring the CMIP6 dataset would be to establish if models' climate sensitivities are different, and ensure that the blend of selected models captures any change in sensitivity across the CMIP dataset and complements the choices made as part of V2.

In parallel, long climate simulations with the RCM will also proceed. As mentioned in the previous section, it has been established that SINGV is a suitable RCM for providing high resolution convection-enabling modelling capability across the WMC and, if required, at even higher resolution for a smaller domain around Singapore, although the benefits obtained with this higher resolution (up to 2km) come with a substantial additional computing cost. So far, the focus has been on the optimal setup in which to integrate SINGV and also on its performance for rainfall, in particular diurnal and annual cycles across Singapore and the region. But the extent of the model's ability has yet to be fully established by running a longer (30 to 50 years) integration forced by the latest generation of reanalyses, ERA5. This will provide a suitable benchmark to understand the model's behaviour and any possible shortcomings prior to forcing it with boundary conditions from the selected GCMs mentioned above. The second phase will then focus on using these GCMs' forcing data (instead of reanalyses) and observing the effect of this on model performance, starting with an evaluation of the impact on the model's skill and exploring methods to deal with the issues that are encountered with GCMs (i.e. lower resolution and systematic biases).

Once the model setup has been finalised, a major, computer-intensive component will be the integrations of the SINGV model to obtain several downscaled climate change projections across the WMC. For each selected GCM, a "current climate" simulation should be completed for a 50-year period (1965-2015). The outputs for Singapore are likely to require some further adjustments to ensure they are directly suitable for impact assessment, such as the bias correction which was introduced as part of V2. This is to be evaluated based on the earlier simulation of similar duration using ERA5 reanalyses as boundary conditions which will provide the upper boundary of the skill of the nested model when forced by unbiased external conditions. In that context, some limited high resolution simulations (e.g. 2km) might be justified to confirm the validity of the bias correction method. Then, for each GCM, several simulations will be performed for the remainder of the 21st century using selected emission scenarios. The likely list is RCP2.6 (consistent with Paris agreement targets), RCP4.5 (middle ground, likely outcome) and RCP8.5 (consistent with business-as-usual approach).

For the analyses of the RCM simulations, the first objective will be to validate the RCM simulations of the current climate, against both observations as well as the simulation of the current climate from the host model to evaluate issues such as possible drift from the host model climatology as well as quantifying the added value of the RCM compared to the host model. In that regard, this component will focus on the exploration on key mechanisms driving the climate of the WMC which are not well captured by GCMs due to insufficient spatial resolution in the shorter timescale (e.g. diurnal cycle of convection and contrast between land and sea), to large-scale phenomena (MJO propagation, monsoonal flow reversals, ENSO and Indian Ocean teleconnections). This analysis will build on the scientific knowledge accumulated in CCRS since the completion of V2. The evaluation of the RCM ability to add value in regard of these processes will be highly significant scientifically. From these simulations, time series will need to derive for Singapore for key variables with a proper bias correction. At the likely resolution used for these simulations (8 km, as used in the most recent test cases with SINGV v5) even if Singapore is covered by up to 10 grid boxes, the information for each grid boxes should not be treated as independent of each other due to numerical smoothing. Instead, an island-wide synthetic series should be created for all variables being considered, taking care not to reduce the reproduction of extremes. This latter point is likely to require having simulations at higher resolution (e.g. 2km or less).

In addition to running simulations across the WMC as a coarse convection-enabling resolution, there is a scientific interest in running simulations at a higher resolution for a smaller domain encompassing Singapore. A very fine resolution could enable the model to represent more accurately the convective nature of the rainfall affecting Singapore such as the diurnal cycle, and particularly relevant in the Singapore context, the possible change of properties of extremes rainfall events. Beside the scientific interest, the considerable computing cost of these very high resolution simulations would have to be balanced with the needs of the many Vulnerability and Impact Assessment (VIA) users who require outputs with very fine resolution. It is likely that only limited time slices can be afforded and will have to be selected to represent both the current climate and appropriate future timelines of interest to stakeholder agencies. The duration and number of time slices will need to be decided upon confirmation of a suitable resolution to obtain real gain in terms of the model ability to reproduce the key phenomena of relevance at that fine scale. It is expected that the model will have to be integrated at 2km or less in order to capture well the annual cycle of the diurnal cycle of convection around Singapore and its immediate vicinity. These simulations will be forced by boundary conditions from the SINGV simulations for the WMC domain. Such simulations will be able to provide output series with sufficient spatial resolution to depict plausible spatial variations around the island that will need to be validated against the high density network of observations across Singapore available for the most recent period. The focus is likely to be on rainfall but other variables may be of interest to users, such as temperature (in relation to thermal comfort studies) and wind (in relation to sea level rise and inundation studies).

A dedicated effort is also required to deliver a new set of projected sea levels for the 21st century, fully encompassing all uncertainties. Of particular focus will be evaluation of global sea level rise due to thermal expansion and ice mass loss from glaciers and ice sheets, including potential for high estimates due to uncertainties about continental ice sheets. Besides the global evolution of sea level, regional and local mean sea level variations driven by climate variability will need to be captured, as well as weather-related extreme sea levels (e.g. monsoonal flows, storm surges, waves) and corrected for the possible movement of the land (global patterns and local subsidence). Because of the existential threat to Singapore, particular emphasis needs to be placed on the low probability-high impact tail end of the distribution of projections. This was delivered in V2 under the H++ approach (Marzin et al. 2015) and will need to be improved on in V3 by delivering estimates of the probability of occurrences for the most extreme part of the distribution, especially the consideration of possible combined effects between regional signature of global sea level rise and changes in factors influencing extremes such as storm surges.

To ensure the effectiveness of V3 beside the scientific achievement, a dedicated effort will also be put into the documentation and communication of the outputs. Early in the programme, there will be a need to re-evaluate stakeholder needs. This was done comprehensively as part of V2, but will need to be updated. In particular the continuity between the projections across V2 and V3 will need to be communicated to stakeholders to help guide their uptake to the scientific findings. Besides the documentation (scientific papers, the stakeholder report and, as required, a more technical report), supporting science to help explore the projections data should be deployed. There is a need to develop a methodology to express the range of uncertainties in terms of climate risk or climate future (e.g. Whetton et al. 2012). Finally, some non-technical material needs to be delivered such as factsheets to explain in a very concise manner some key aspects of the project and its findings, and sets of slides to communicate essential V3 findings.

The final part of the programme will be the delivery of the output data on a suitable platform. The most effective mechanism would be a web-based data portal with several layers of access that provide: i) a basic level, freely available to all, providing access to high level information such as the synthesis report, some plots, simple data series, ii) a “registered” level, for registered users (registration will ensure there is some understanding of the users’ interests and the likely usage of the data) providing access to more in-depth information (e.g. technical report, most graphics, bias corrected model outputs) and iii) an “educated” level, accessible to users with whom some form of relationship has been established to ensure that their understanding of the science is compatible with getting access to the complete set of model outputs.

While all the components mentioned above are mandatory to ensure the successful delivery of a new set of national climate projections, some additional research directions could also be explored. While not essential to complete the main objectives, they would contribute toward a more comprehensive project (Figure 7).

The first additional component would be to complete downscaled climate change projections generated with a second RCM unrelated to SINGV. One of the main criticisms of V2 was that all the projections were based on a single RCM, and therefore it was impossible to evaluate the potential that the projections may be misleading due to some systematic deficiency of the model. While V3 will be based on SINGV that has

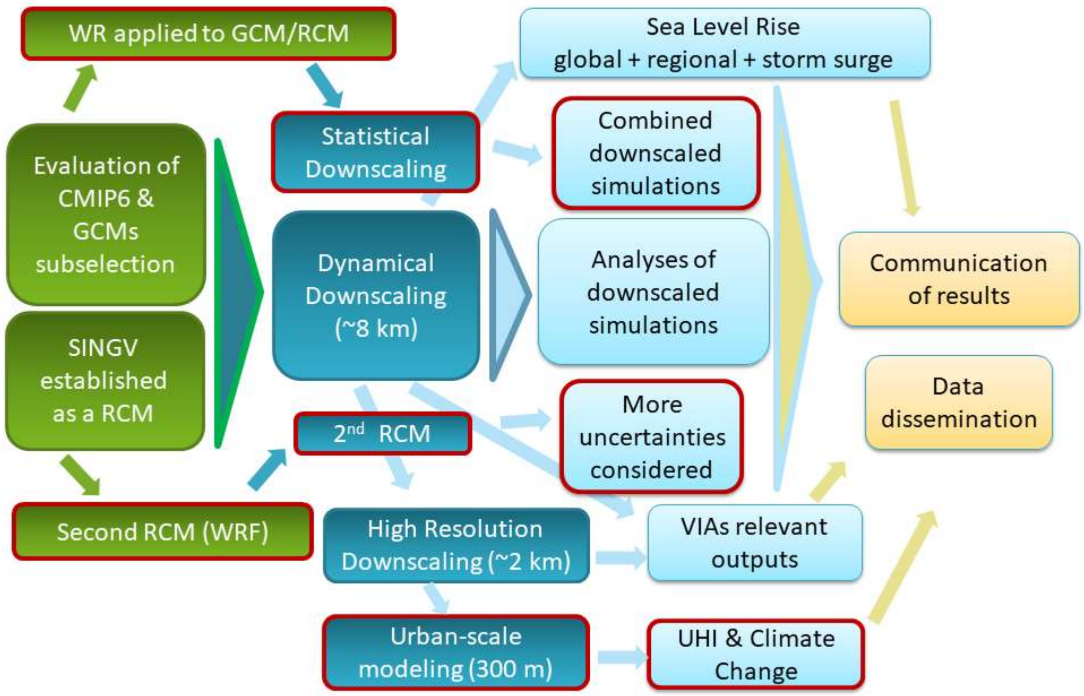

Figure 7 Schematic from Figure 6 completed with additional optional components (highlighted by red framed boxes) that could be undertaken for a more complete V3 project (Note: UHI stands for Urban Heat Island).

benefited from more in-depth evaluation than the previous model used during V2, it remains possible that some specific behaviour of SINGV could unduly influence the projected changes. It would therefore be useful to provide comparable projections with a different RCM. While several options exist for potential RCMs which have been used in the region, the most likely candidate is the WRF model, which has been extensively used in Singapore, in particular by a group in Tropical Marine Science Institute within National University of Singapore (Raghavan et al. 2016). Through a collaborative approach, parallel simulations using the same GCM boundary conditions will allow for a more comprehensive consideration of the uncertainties, including the uncertainties due to the downscaling method.

Dynamical downscaling allows higher resolution information to be obtained from a climate model. However, this information remains coarse from the perspective of a precise location (e.g. a measurement station on which an impact modelling tool might be based) and comes at a high computing cost. For this reason, statistical downscaling provides a useful alternative and is based on an empirical statistical relationship between the large-scale (predictors) and the local scale (predictands). CCRS has developed a methodology to classify daily meteorological situations into 8 clusters of weather regimes which are useful for describing day-to-day variability and relating it to local quantities such as rainfall in Singapore (Hassim and Timbal, 2019). Provided the global climate models are able to properly represent these weather regimes, this method could be used as a tool to interpret the model simulations in terms of their local signature in Singapore. It would then be possible to interpret changes in the weather regime frequency and/or local signatures between simulations of the current and future climate in terms of local changes in the predictands. This would thus provide a method to estimate local change due to climate change independently from the modelling approach which, while physically based, relies on estimation within the model parameterisations. Examples of studies done overseas (Climate Change In Australia, CSIRO and BoM, 2015) have shown that by comparing and contrasting statistically and dynamically downscaled climate change projections, it is possible to more appropriately determine the level of confidence to be placed on the projections, in particular because many more global simulations and scenarios can now be explored, due to the low computational cost.

Finally, understanding the combined effect of urbanisation and climate change for the city-state of Singapore would be highly beneficial. Very high resolution simulations (300m or less) could be carried out to explore the role of the local urban canopy scale and UHI dynamics in modifying the local climate. This would require SINGV to be further developed into an urban-scale model (uSINGV) and integrated at sufficiently high resolution (300m or less). Some early testing has already been completed (Simon-Moral et al. 2019), and if further long-term integration confirms that the model is suitable for climate integration, a series of simulations could be performed, using future climate boundary conditions and various urbanised landscapes to establish the combined effects of climate change and possible mitigation approaches. These simulations would come at a very high computing cost, and hence only very small domains would be feasible and only for short climatic integrations.

References

Argueso, D., Romero, R., Homar, V. (submitted) Precipitation features of the Maritime Continent in parameterized and explicit convection models. *Journal of Climate.*

Berthou S., Rowell, D.P., Kendon, E.J., Roberts, M.J., Stratton, R.A., Crook, J.A., Wilcox, C. (2019) Improved climatological precipitation characteristics over West Africa at convection-permitting scales. *Climate Dynamics*, 53:1991-2011.

CSIRO and Bureau of Meteorology (2015) Climate Change in Australia Information for Australia's Natural Resource Management Regions: Technical Report, CSIRO and Bureau of Meteorology, Australia. Available at: www.climatechangeinaustralia.gov.au/en/publications-library/technical-report/, last accessed: 12 December, 2019.

CMIP6 (2019) The CMIP6 landscape. *Nature Climate Change*, 9, 727, doi:10.1038/s41558-019-0599-1.

Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., Taylor, K.E. (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9:1937-1958.

Gordon C., Cheong, W.K., Marzin, C., Rahmat, R. (2015) Singapore's Second National Climate Change Study - Climate Projections to 2100 - Report to Stakeholders. Centre for Climate Research Singapore, Singapore. Available at: http://ccrs.weather.gov.sg/wp-content/uploads/2015/04/Publications-Second-National-Climate-Change-Study-Report-for-Stakeholders.pdf, last accessed: 12 December, 2019.

Hassim M.E., Rahmat, R., Lestari, R.K., Timbal, B. (2016) Generating Climate Change Rainfall Scenarios for Singapore: A Tale of Scale. *COSMOS Research highlights*, 12:8-13.

Hassim, M.E., Timbal, B. (2018) Observed rainfall trends over Singapore and the Maritime Continent from the perspective of regional-scale weather regimes. *Journal of Applied Meteorology and Climatology*, doi:10.1175/JAMC-D-18-0136.1.

Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Muñoz Sabater, J., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., Dee, D. (2019) Global reanalysis: goodbye ERA-Interim, hello ERA5. *ECMWF Newsletter No. 159 –* Spring 2019, pp. 17-24.

Huang, X.-Y., Barker, D., Webster, S., Dipankar, A., Mittermaier, M., Sun, X., North, R., Darvell, R., Boyd, D., Lo, J et al. (accepted) SINGV - the Convective-Scale Numerical Weather Prediction System for Singapore. *ASEAN Journal on Science & Technology for Development*.

IPCC (2014) Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, R. K. Pachauri and L. A. Meyer(eds.), IPCC, Geneva, Switzerland, 2014. Available at: www.ipcc.ch/report/ar5, last accessed: 12 December, 2019.

Kendon E.J., Roberts, N.M., Fowler, H.J., Roberts, M.J., Chan, S.C., Senior, C.A. (2014) Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nature Climate Change*, 4: 570-576.

Marzin C., Rahmat, R., Bernie, D., Bricheno, L., Buonomo, E., Calvert, D., Cannaby, H., Chan, S., Chattopadhyay, M., Cheong, W.K., Hassim M.E. et al. (2015) Singapore's Second National Climate Change Study - Phase 1. Centre for Climate Research Singapore, Singapore. Available at: http://ccrs.weather.gov.sg/Publications-Second-National-Climate-Change-Study-Science-Reports, last accessed: 12 December, 2019.

Simón-Moral A., Dipankar A., Roth M., Sánchez C., Velasco E., Huang X.Y. (under revision) Application of a single layer urban canopy model in a tropical city: preliminary results from Singapore. *Quarterly Journal of the Royal Meteorological Society*.

Raghavan S.V., Vu, M.T., Liong, S.Y. (2019) Regional climate simulations over Vietnam using the WRF model. *Theoretical and Applied Climatology,* 126:161-182.

Taylor K.E., Stouffer, R.J., Meehl, G.A. (2012) An Overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93:485-498.

van Vuuren D.P., Edmonds, J., Kainama M., e t al. (2011) The representative concentration pathways: an overview. *Climatic Change*, 109:5, doi:10.1007/s10584-011-0148-z.

Whetton, P., Hennessy, K., Clarke, J., McInnes, M., Kent, D. (2012) Use of Representative Climate Futures in impact and adaptation assessment. *Climatic Change*, 115: 433-442.

EXPLORATION OF THE CLUSTERING OF MOISTURE FLUX FIELDS AND ITS RELATION TO WET AND COOL SPELLS IN SINGAPORE

Han Yu Sit[1], Junhua Yang[2]

[1] Ecole Polytechnique, France, [2]Centre for Climate Research Singapore, Meteorological Service Singapore

INTRODUCTION

Wet spells are consecutive days of substantial amounts of rainfall and cool spells are consecutive days of low temperatures. These weather events are rare and often have significant impacts on Singapore. For instance, Singapore experienced a bout of cool and rainy weather from 10 - 14 January 2018 during which the daily maximum temperature ranged from 23.8°C to 27.3°C, belonging to 2% of the coldest days in Singapore over the past 40 years. Floods were triggered in parts of Singapore by the persistent rainfall (Channel News Asia, 2018) while people resorted to wearing winter clothing to resist the low temperatures (The Straits Times, 2018). With these significant impacts on the daily lives of the people in Singapore, the prediction of wet spells and cool spells are of meteorological and public interest.

The strong association between cold surges based on the definition by Lim et al. (2017) and wet spells has been investigated (Yang et al., 2019) by compositing the moisture flux patterns in the South China Sea over the wet spells. However, the characterization of wet spells based on cold surges in these studies relies on a cold surge index which applies several predefined thresholds to certain atmospheric fields in specified regions (Lim et. al, 2017). This study explores if an alternative characterization of cool spells and wet spells exists based on clustering of the moisture flux fields. Moisture flux plays an important role in weather forecasting (Banacos and Shultz, 2005) and is expected to also play a key role in the wet and cool spell process in Singapore. In this study, we wish to understand the variability of the regional moisture pattern in connection with the local temperature and rainfall observed in Singapore and identify the key patterns responsible for the observed wet and cool spells using clustering. The aim is to construct and understand the cool and wet spell climatology of Singapore. This will allow us to understand the variability of the regional moisture flux patterns during the northeast monsoon season and their connections to these spells, with the objective of improving the ability to predict these events.

Cool and wet spells in the northeast monsoon season from November to March (NDJFM) were first defined and identified based on historical data from Singapore's temperature and rain gauge records. A clustering algorithm was then used to separate each day in the monsoon season into different clusters based on its atmospheric fields. The aim is to isolate wet and cool spell days to a single cluster and hence characterize the spells using this cluster's atmospheric fields, with the ultimate objective of replacing the definition of spells with such a cluster.

DATA AND METHOD

A wet spell event is defined to be a period over which the daily rainfall duration is more than 6 hours and the average daily rainfall duration during the event is at least 8 hours for a minimum of two consecutive days. These durations represent the 90th and 95th percentile rainfall duration respectively during the NDJFM season from 1955-2018. The daily rainfall duration is given by the maximum rainfall duration recorded across all of Singapore's available rain gauges. Similarly, a cool spell is defined to be a period over which the daily maximum temperature is less than 27°C (lowest 4% of the NDJFM season) for at least two consecutive days. The daily maximum temperature is given by the average of the daily maximum temperature recorded across five weather stations in Singapore. Rainfall and temperature data from 1955 to 2018 were used to identify past wet and cool spell events.

It should be noted that there were fewer stations and rain gauges during the 1950s to the 1970s, and hence there was a chance of missing the detection of some of the wet and cool spell events during that period. However, due to the small size of Singapore and the large-scale impact of cold surges, it is reasonable to expect wet and cool spell events to be island-wide phenomena in most cases. As a result, the chance of missing them due to fewer observations in the past is low. It is therefore still possible to construct a reliable climatology of wet and cool spells despite limited observations in the earlier decades.

The identified dates of wet and cool spells were then compared against those dates obtained from the clustering of selected atmospheric fields. The fields used in this study were obtained from the European Centre for Medium Range Weather Forecasts (ECMWF) Interim Reanalysis dataset (ERA-I). The reanalysis dataset consists of daily mean values of various atmospheric fields obtained by averaging over 6-hourly data spanning from 1979 to 2018 at a spatial resolution of either 1x1 degree or 0.125 x 0.125 degree (Dee et al., 2011). The atmospheric fields used in this study include the two components of the vertically integrated moisture flux vector: the vertical integral of the eastward (VIQU) and northward (VIQV) moisture flux and the divergence of the moisture flux vector (div VIMF). These vertically integrated moisture flux datasets were of a resolution of 1x1 degree and were precomputed using 60 instantaneous analysed model levels from the surface to the top of the atmosphere (0.1 hPa), and were directly available as part of the ERA-I dataset.

Another atmospheric field variable used in this study is the 2-metre temperature (T2M) at a resolution of 0.125x0.125 degree. The domain for the clustering of the moisture flux variables (VIQU and VIQV) included all grid points bounded within the region -2°S−12°N and 98°E−115°E (270 grid points). The higher resolution T2M data were clustered over the grid points within a much smaller region around Singapore within 1.25°N−1.50°N and 103.625°E−104°E (12 grid points). The clustering domains are shown in figure 1. These domains were chosen so as to focus on the low temperature and wet conditions observed locally over Singapore during wet and cool spells, and the associated moisture flux pattern over a broader region in the South China Sea.

In this study, the K-Means clustering algorithm was used to cluster dates based on the similarity of their associated daily mean atmospheric fields. There were 270 grid point values for each of the 2 components of the moisture flux, but only 12 grid point values for the T2M variable. There was therefore a huge imbalance in the relative contribution of the moisture flux and the local T2M variables to the clustering algorithm. It was desirable to treat the local T2M variable and the moisture flux variables on equal footing during the clustering process, hence the 12 grid point values were duplicated 23 times (becoming 276 grid points) to match approximately the 270 grid points for the moisture flux variables. Since the Euclidean metric in K-Means clustering was used, the procedure of duplication is mathematically equivalent to the assignment of weights to the features used in the clustering, a technique commonly used in weighted clustering to assign importance to certain features. Without the duplication, the huge number of moisture flux variables will overwhelm the local temperature variable at the few grid points. It should be noted that the objective of the clustering is to separate the dates based on temperature, i.e. the clustering procedure is outcome-driven. Performing the duplication step helps to achieve a balance between the generality of unsupervised learning and the specificity of the desired outcome (a cluster with low temperatures and high rainfalls) to identify cool and wet spells. To verify the validity of duplication, tests were also performed where the number of duplications was varied, and the subsequent clustering results analysed.

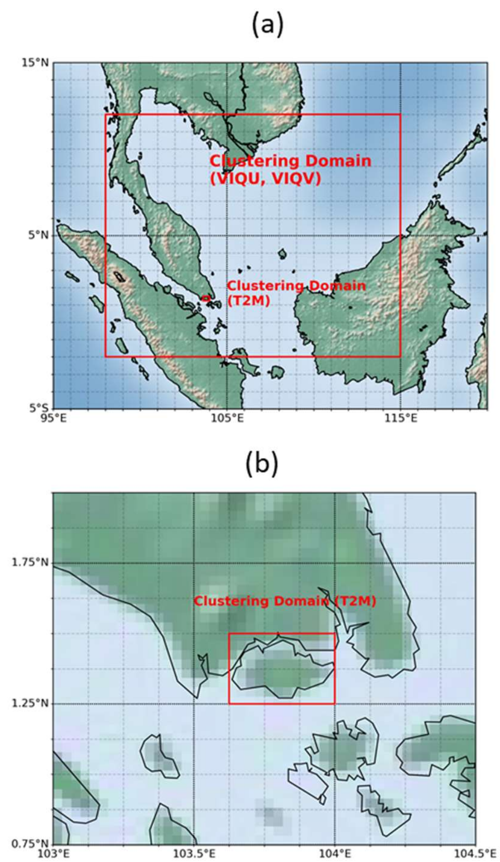

Figure 1 (a) The larger red rectangle represents the clustering domain (270 grid points) for the 2 components of the moisture flux variables (VIQU, VIQV at 1x1 deg). The smaller red rectangle (magnified in (b)) over Singapore represents the clustering domain for the 2-meter temperature (T2M at 0.125 x 0.125 degree, 12 grid points).

Prior to clustering, for each grid point in the atmospheric field, the field is standardized according to

$$X_{standard} = \frac{\mathrm{X} - \bar{X}}{\sigma_X}$$

where X is the original field, $\bar{X}$ is its mean and σ_X is its standard deviation. With two regional moisture flux variables (VIQU and VIQV) and a local (T2M) variable, there are altogether 816 (270x2+276=816) standardized grid point values used for the clustering of the 6050 data points or dates.

RESULTS AND DISCUSSION

By applying the definitions of wet and cool spells given in the previous section to the rain gauge and temperature records of Singapore, an annual average of 5.6 wet spell days and 1.8 cool spell days during the northeast monsoon season was obtained. The length of a wet spell and cool spell can range from 2-10 and 2-5 consecutive days respectively. 74% of all wet spells and 85% of all cool spells occur during the months of December and January. These two months can therefore be regarded as the wettest and coolest period for Singapore during the entire northeast monsoon season. Cool and wet spell events overlap strongly, as can be seen in figure 2c. 77% of all cool spell days are in fact wet spell days, but only 25% of all wet spell days are cool spell days.

The annual occurrence anomalies of wet spells and cool spells are shown respectively in figures 2a and 2b. The anomalies are plotted against ENSO years (from June to May). Each year is classified into one of the 3 categories: El Niño, La Niña or Neutral using a detrended (with background tropical warming signal removed) classical Nino3.4 Sea Surface Temperature (SST) index (Turkington et al., 2018). It was found that the occurrences of cool and wet spells vary from year to year. There seems to be no clear relationship between the main mode of natural variability affecting the Western Maritime Continent and Singapore (ENSO) and the occurrences of cool/wet spells. However, it is interesting to note that after the strong El Niño event in 1997, all subsequent El Niño years appear to be associated with negative anomalies.

In addition, there seems to be some form of variability occurring on the decadal time scale. Wet spells are seen to be more active during the late 1970s to the early 1990s and again from the late 1990s to the mid-2000s. From about 2007 onwards, the occurrences of both cool and wet spells are generally low, with the exception of a spike in cool spell days in the La Niña year 2017. There is also a relatively large number of cool spell

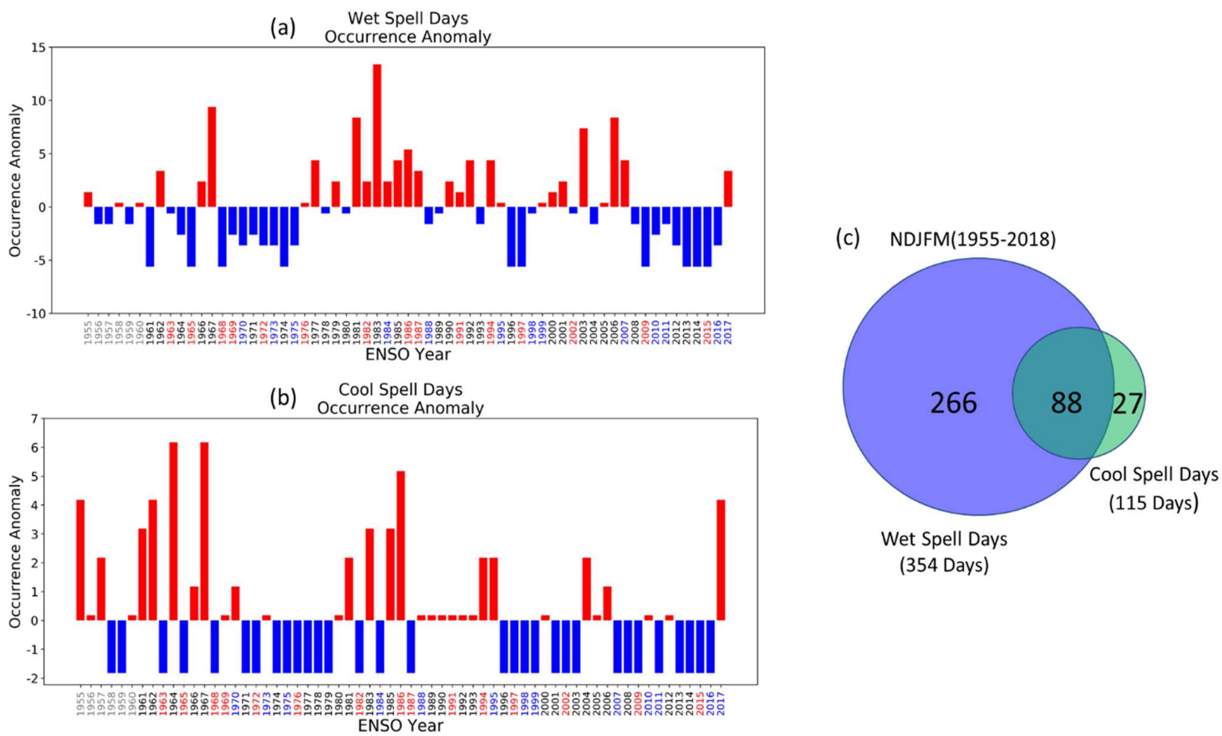

Figure 2: (a) and (b) shows the long-term trend in the anomaly (with mean from 1955-2018 subtracted) of the occurrences of wet spells and cool spells respectively in Singapore during the northeast monsoon season. Each year is labelled as an ENSO year (from June to May, red: El Niño, blue: La Niña, black: neutral, grey: no ENSO data). The ENSO classification is based on the detrended Nino3.4 SST index from Turkington et al. (2018). (c) is a Venn diagram showing the overlap between wet spell and cool spell days observed in Singapore.

days during the mid-1950s to 1970. This should however be interpreted with caution as this was a period when Singapore was largely rural and the recorded temperatures may have a low bias, resulting in an overestimate of cool spell events. Moreover, the lower number of wet spells observed during that period also does not seem to correspond with the high number of cool spells. Therefore, the positive anomaly in cool spell occurrence may not truly reflect the actual natural variability of the regional climate, although this cannot be ruled out completely due to the uncertainties involved.

These observed interdecadal variabilities might be linked to the interdecadal variability of the East Asian winter monsoon (Ding et al, 2014) and its impact in the deep tropics. However, a discussion of this would be beyond the scope of this work and further studies would be required to better understand the potential connection. It should also be noted that these analyses may be dependent on the choice of thresholds used in defining spells. However, we find that using different thresholds such as a minimum rainfall duration of 4 hours and an average of 6 hours resulted in similar interdecadal variations.

In the K-Means clustering algorithm, each date is represented by the 816 data points from the atmospheric variables VIQU, VIQV and T2M and can be considered as a point in an 816-dimensional vector space armed with the Euclidean metric. With this representation, the distance between points can be calculated, giving rise to the notions of the separation, the centres (centroids) and outliers of clusters. In this method, the number of clusters *n* was determined using the elbow method, where *n* is chosen such that a further increase in *n* does not significantly improve the separability of the clusters. The optimal number of clusters was estimated to be four.

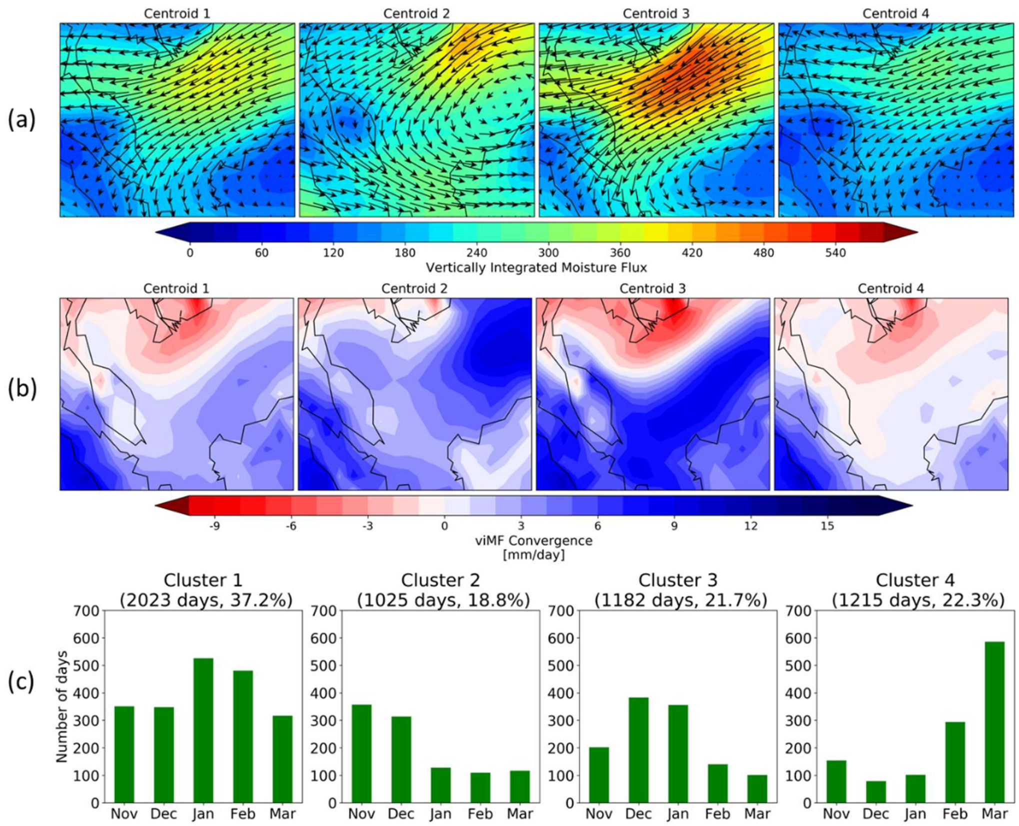

Figure 3: (a) and (b) show the vertically integrated moisture flux (VIMF) and its convergence respectively for the four cluster centroids computed using the K-Means algorithm. (c) shows the frequency of occurrences of each cluster during the northeast monsoon season. Cluster 3 is associated with the strongest moisture flux and convergence over the equatorial South China Sea. Majority of the wet spells (91%) and cool spells (100%) are associated with cluster 3.

Figure 3a, 3b and 3c show the centroid of each of the clusters for the moisture flux vector, moisture flux convergence and the number of days in each month across the clusters respectively. Cluster 1 shows a moderate amount of moisture flux and convergence over the South China Sea, implying a moderate amount of potential rainfall in the region. The dates in this cluster span almost evenly across the months with a slightly greater number of days in January and February. Among the four clusters, cluster 1 represents the largest proportion (37.2%) of all days during the northeast monsoon season.

Cluster 2 is characterized by the presence of a circulation of moisture flux in the South China Sea as shown in figure 3a. As a result of this circulation, there is a strong convergence of moisture in that region (figure 3b). The circulation might be associated with the Borneo vortex, a meso-α-scale cyclonic disturbance that forms to the northwest of Borneo Island from cold surges during the northeast monsoon season. This circulation redirects the prevailing northeasterly winds to become northwesterly winds around the vicinity of Singapore. Due to this shift in wind direction, there might also be a corresponding shift in local weather and diurnal pattern from the usual northeast monsoon setup.

Cluster 3 is characterized by the presence of a very strong moisture flux in the South China Sea. The strong flux transports a large amount of moisture towards Singapore and Peninsular Malaysia, and results in its strong convergence over the equatorial South China Sea as shown in figure 3b. This could result in heavy and persistent rainfall, and implies that most of the wet spells may actually lie in this cluster. This is further supported by the fact that the majority of the dates in this cluster fall in December and January during the wet phase of the northeast monsoon season.

In contrast to cluster 3, cluster 4 is characterized by generally weak moisture flux and moisture flux divergence instead of convergence in the region. Most of the days in cluster 4 occur during the late northeast monsoon, in the months of February and March during which rainfall over Singapore is low. Cluster 4 hence clearly represents the dry phase of the northeast monsoon. It should be emphasized that all four centroid patterns can be found across all months in the northeast monsoon season, but that certain patterns occur more frequently in certain months. Collectively, they represent the key variability of the moisture flux patterns in the season.

With these clustering results, wet and cool spell dates can now be matched with the dates in the clusters. 91% of all wet spell days and all cool spell days lie in cluster 3, consistent with the previous analysis of this cluster. Any dates falling outside of cluster 3 are unlikely to be part of a wet spell event. However, the inverse does not hold true because only 17% of the dates in cluster 3 are wet spell days. Nevertheless, the necessary large-scale moisture flux and moisture flux convergence patterns for the occurrence of wet spells has been established.

Also investigated were the changes in the clustering results due to the variation of the number of duplicates of the T2M dataset. In the case of having no duplicates, the moisture flux patterns of the centroid are extremely similar to the original case. However, there is a redistribution of the dates across the clusters. For instance, the percentage of dates in cluster 1 is reduced to 23.3% from the original 37.1%. The wet spell cluster (cluster 3) accounted for only 74% of all wet spell days compared to the original 91% and 70.9% of all cool spell days compared to the original 100%. With 11 duplicates, the percentage of all dates in cluster 1 was reduced to 24%, and the wet spell cluster accounted for 90.1% of all wet spell days and 96.4% of all cool spell days. Even though there is a redistribution of the dates, the overall trend in the distribution of the dates across the months remains similar (e.g. the month that a cluster peaks in). The physical interpretation of the moisture flux patterns remains unchanged. With this result, we have verified that increasing the number of duplicates up until equal weightage results in a better isolation of the spell days to a single cluster, justifying the necessity of duplication.

Figure 4a represents the boxplots of the rainfall duration, rainfall quantity, rainfall intensity and daily maximum temperature of the dates in each cluster. In the case of cluster 3, the interquartile range of the boxplot for the rainfall duration is significantly higher than the rest of the clusters, consistent with the analysis that cluster 3 represents the wet phase of the northeast monsoon season. It is also the coolest period, as shown by the significantly lower interquartile range of the daily maximum temperature. In contrast, the distribution of rainfall duration and quantity for cluster 4 is lower than the rest of the clusters, clearly showing cluster 4's association with the dry phase of the northeast monsoon season. Cluster 4 is also characterized by higher temperatures, with its interquartile range for the maximum daily temperature being the highest of all four clusters.

Given that wet spells are anomalous events, it was of interest to investigate if they correspond to the cluster outliers. The cluster outliers are outliers based on their distance from the centroids, and if they also correspond to the rainfall and temperature outliers, then wet spells can be identified directly. Figure 4b shows the boxplots for the dates that correspond to the top 5% of outliers in each cluster. A comparison with figure 4a shows a large overlap between the boxplots for the outliers and those for the entire set of dates. This implies that the distribution of temperature and rainfall is independent of whether the dates are outliers or not. For instance, if a date is a cluster outlier, it is not possible to tell if this date has a high rainfall duration. Hence, one

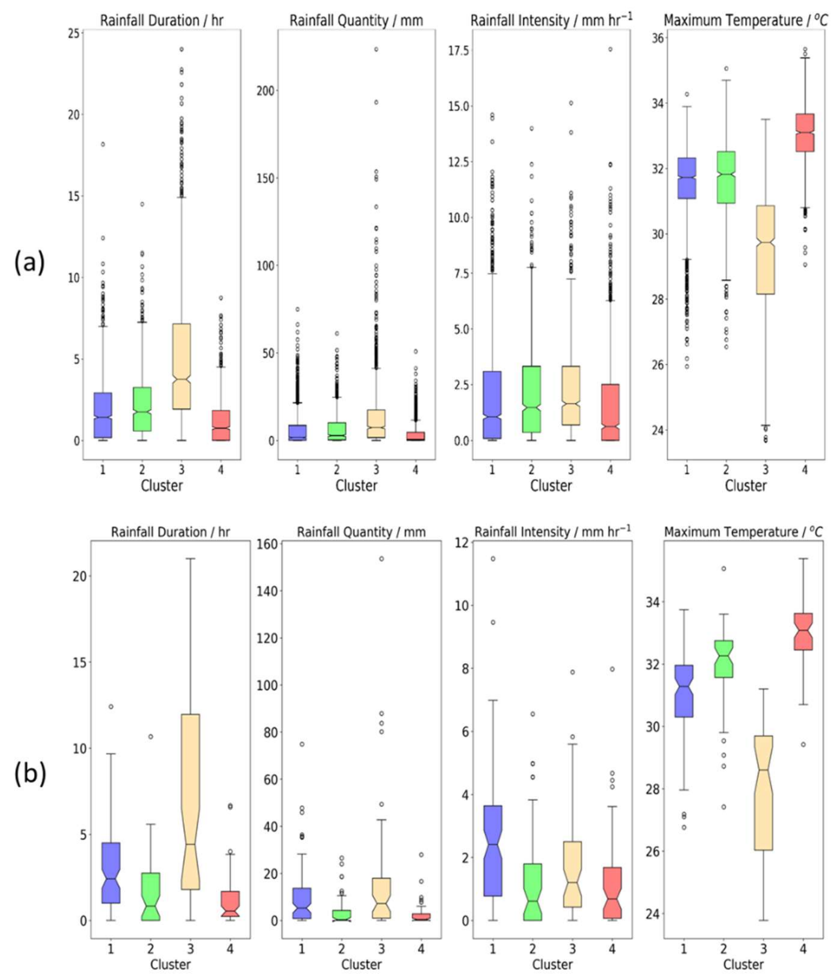

Figure 4: (a) Boxplots of daily rainfall and temperature observations made in Singapore and grouped into the 4 clusters identified using the K-Means algorithm. The whiskers represent the values at 1.5 times the interquartile range (Q3-Q1) below the first quantile (Q1) and above the third quantile (Q3). White circles beyond the whiskers are regarded as outliers. (b) shows the same box plots for the furthest 5% of the data points from the centroid of each cluster. The large overlap in the boxplots here and in (a) implies that the cluster outliers judged based on their distances from the centroids does not correspond directly to rainfall and temperature outliers.

cannot associate low temperatures and high rainfall duration with cluster outliers to identify spells.

The large overlap in the interquartile range across the clusters for each quantity in the boxplots also shows that the clustering procedure does not cleanly separate the local rainfall and temperature. Therefore, even after assigning a date to a cluster, predicting its rainfall and temperature remains a difficult task. At best, probabilistic predictions of rainfall and temperature can be given based on their distributions in each cluster. Nevertheless, we can expect significantly higher rainfall and lower daily maximum temperature predicted for dates classified under cluster 3 due to the skewness of their distributions. The converse would hold for those dates in cluster 4. Overall, the maximum daily temperature is the best-separated quantity as the interquartile ranges overlap the least. This is likely due to the fact that T2M is highly correlated with the daily maximum temperature in Singapore, thus the clustering based on T2M directly helps to improve the clustering of the daily maximum temperature.

CONCLUSION

A climatology of cool and wet spells in Singapore based on appropriate thresholds and criteria applied to the local rain gauge and temperature records has been established. Most cool and wet spells occur during December and January. It has also been observed that many cool spell days are simultaneously wet spell days, showing the strong connection between them. We have attempted to use the K-Means clustering approach to understand the key variability of the regional moisture flux pattern during the northeast monsoon season and the connections to wet and cool spells observed in Singapore. We managed to identify the key moisture flux pattern associated with wet spells. Cool and wet spell days were isolated to a single cluster (cluster 3) and it was thus reasoned that its associated moisture flux and the local 2-meter temperature fields can potentially serve as a good diagnostic for these spells.

However, as only 17% of the days in this particular cluster are wet spells and there is a large spread in the distribution of rainfall duration, further investigation is needed before this approach can be applied in forecasting. One possible avenue of further enquiry would be to study the difference between the wet spell days and non-wet spell days in this cluster, such as by trying to identify distinguishing features between the atmospheric fields in the two groups. Further optimization of the clustering algorithm, and better selection of clustering variables could also be carried out. Other clustering algorithms may also be explored to improve the clustering, as the K-Means method is sensitive to centroid initialisation and performs less optimally if the representation of the data in the 816-dimensional space is non-spherical.

As the high correlation between T2M and the daily maximum temperature led to a better separation of the daily maximum temperature among the clusters, another possible area of investigation could be the identification of new atmospheric variables that are highly correlated with the observed rainfall duration to achieve a better separation of it. With a clearer separation in the distribution of the rainfall and temperature indicators across the clusters, regression models that help to map these atmospheric field variables to rainfall and temperature for each cluster can then be developed to predict these quantities and thus wet and cool spells from their definitions. This would lead to a more direct identification of wet and cool spells, allowing for the possibility of predicting them by analysing the forecast atmospheric field patterns from numerical weather prediction models. Some successes and limitations of the method have also been discussed. The results of this study can potentially pave the way for the development of statistical tools for predicting wet and cool spells in Singapore.

References

Banacos, P., and Schulz, D. M. (2005) The use of moisture flux convergence in forecasting convective initiation: Historical and operational perspectives. *Weather and Forecasting*, 20: 351-366.

Channel News Asia (February 2, 2018) January an 'eventful' month for Singapore weather: Met Service, Available at: https://www.channelnewsasia.com/news/singapore/january-an-eventful-month-for-singapore-weather-met-service-9917370, last accessed: 23 September, 2019.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S. et al. (2011) The ERA-Interim Re-Analysis: Configuration and performance of the data assimilation system. *Quarterly Journal of Royal Meteorological Society*, 137: 553-597.

Ding, Y., Liu, Y., Liang, S. Ma, X., Zhang, Y., Si, D., Liang, P., Song, Y., Zhang, J. (2014) Interdecadal variability of the East Asian winter monsoon and its possible links to global climate change, *Journal of Meteorological Research*, 28: 693.

Yang, J.H., Qian, J.H., Timbal, B. (2019) A moisture flux perspective on cold surges and wet spells affecting Singapore, *MSS Research Letters*, Issue 3, pp 3-9. Available at http://ccrs.weather.gov.sg/wp- content/uploads/2019/07/MRL_Issue_3_Jul2019_FINAL.pdf, last accessed 23 September, 2019.

Lim, S. Y., C. Marzin, P. Xavier, C. P. Chang, and B. Timbal (2017) Impacts of Boreal Winter Monsoon Cold Surges and the interaction with MJO on Southeast Asia Rainfall, *Journal of Climate*, 30: 4267-4281.

The Straits Times (January 19, 2018) Last week's cool spell was Singapore's longest in a decade. Available at: https://www.straitstimes.com/singapore/environment/last-weeks-cool-spell-was-singapores-longest-in-a-decade, last accessed: 23 September, 2019.

Turkington, T., Timbal, B., Rahmat, R. (2019) The impact of global warming on sea surface temperature based El Niño–Southern Oscillation monitoring indices, *International Journal of Climatology*, 39: 1092-1103.

CALIBRATION OF ECMWF SEASONAL SEAS5 MODEL'S MONTHLY TEMPERATURE RE-FORECASTS OVER THE SOUTHEAST ASIA REGION

Regine Ho[1], Ryan Kang[2], and Raizan Rahmat[2]

[1]Nanyang Technological University, [2]Seasonal and Subseasonal Prediction Section, Centre for Climate Research Singapore

INTRODUCTION

State-of-the-art seasonal forecasting systems from Global Producing Centres (GPCs) are known to suffer from large model drifts and systematic model-dependent biases (Manzanas et al. 2019). These biases measure the correspondence between the average forecast and the average observed value of the predictand (Wilks, 2011). Re-forecasts of the past 20 to 30 years are produced using the same model by GPCs to bias-correct these drifts in real-time forecasts. However, model biases still exist in the re-forecasts and post-processing is needed to make the raw re-forecasts more usable.

Statistical post-processing has been a popular choice to reduce such biases, in order to improve the quality of probabilistic forecasts. They range from relatively simple methods such as bias adjustment (BA) and ensemble recalibration (RC) methods to more complex statistical downscaling techniques training on large-scale predictors. This letter will focus on BA and RC methods only. BA methods aim to generate the same mean and standard deviation of forecasts as the observations, whereas RC methods use the idea of inflation of the forecast ensemble mean and spread in order to resemble the observations as closely as possible.

Overall, relatively good seasonal skill has been documented for the Southeast Asia (SEA) region, especially for the 2-metre temperature (t2m) parameter. This letter investigates whether further skill improvements in the t2m parameter can be achieved by applying three calibration methods (MVA, CCR and LR) to the March to May (MAM) period. This period was chosen as records have shown that the region experienced more heatwaves during this season. Therefore, accurate and reliable monthly/seasonal temperature predictions during this period would be useful for alerting the population to any impending heatwaves, or for predicting the coming of relief conditions.

DATA AND METHODS

DATA

The ECMWF SEAS5 seasonal model (Johnson et al. 2019) was used in this study. This model is initialised on the first day of every month, producing 51 ensemble member forecasts up to 7 months into the future. Thirty-six years of re-forecasts from 1981 to 2016 are also produced on the same date as the forecasts, each consisting of 25 ensemble members. Re-forecasts are similar to forecasts, but are initialised for past dates using ERA-Interim reanalysis. These re-forecasts are used for assessing the model's skill and calibrating the model against its climatology.

ERA5 reanalysis (Hersbach and Dee, 2016) was used as the observational reference dataset for both the BA and RC calibration methods and also for the verification of raw/calibrated re-forecasts. ERA5 is the fifth generation reanalysis product from ECMWF, replacing ERA-Interim reanalysis which stopped production on 31 August 2019. It has a finer spatial resolution of 0.25° x 0.25°, is archived at hourly time steps and uses a more advanced data assimilation system than ERA-Interim (fourth generation).

In this study, a one-month lead forecast from the ECMWF SEAS5 model was analysed with predictions initialised in February 2019, for MAM. However, only the March results will be discussed in this letter. In addition, a more recent re-forecast period of 1993 to 2016 was used for calibration to avoid the long-term trend of climate change (Johnson et al. 2019) and also to be consistent with the calibration period used in the Copernicus Climate Change Service (C3S)'s multi-system seasonal forecast. Monthly-aggregated ECMWF SEAS5 re-forecasts and ERA5 2-m temperatures (t2m) that have been bi-linearly interpolated from their native horizontal resolution to the common 1° regular grid were downloaded using C3S-API Climate Date Store (Climate Data Store, 2019) over the common period of

24 years (1993-2016). The domains for both datasets spanned Southeast Asia (90°E – 140°E, 10°S – 20°N).

BIAS ADJUSTMENTS AND ENSEMBLE RECALIBRATION METHODS

BA methods adjust the raw model outputs towards a corresponding observational reference to make them compatible with the local climatology, regardless of the temporal pairing between the predictions and the observations. In the BA domain, methods range from simple adjustments in the mean/variance (MVA) to more complex quantile mapping (to adjust higher-order moments or non-normal probability density function). RC methods, on the other hand, transform the raw model outputs by temporal pairing of re-forecast ensemble means and the corresponding observations. Techniques for RC methods range from relatively simple algorithms such as climate conserving recalibration (CCR) to more general ensemble model output statistics (EMOS) using linear regression (LR) or non-homogeneous Gaussian regression.

This study uses a framework of ready-made statistical packages in R: ‘hyfo’ (Xu, 2015), and ‘climate4R’ and ‘calibratoR’ (Santander Meteorology Group, 2019). Jupyter Notebook (utilising Python) was used for assessing and extracting data, statistical computations and data analysis. The ‘calibratoR’ package consists of the BA and RC calibration methods for the seasonal re-forecast dataset, and the selected methods were mean/variance adjustment (MVA), climate conserving recalibration (CCR) and linear regression (LR). Detailed implementations of the three methods can be found in Appendix A.

All three methods of calibration (Manzanas et al, 2019) were applied under a leave-1-year-out cross validation (LOOCV) setting to avoid artificial skill due to small sample sizes (24 years of data in our study). Cross-validation simulates prediction for unknown data by repeating the fitting procedure on data subsets, and then testing the predictions on the data portions left out of each of these subsets. For example in LOOCV, if there are *n* data points in the original sample, then *n-1* samples will be used to train the model, with the remaining data point used as the validation (testing) set. This procedure is repeated *n* times.

The parameters in each calibration method were first determined based on the complete ensemble information and then applied to calibrate each individual ensemble member.

VERIFICATION SCORES

A good quality probabilistic forecast would illustrate close correspondence between the forecast conditions and the observed conditions (Murphy, 1993). Verification scores are used to determine different aspects of forecast quality, such as discrimination and accuracy. These two aspects of forecast quality were chosen for this study to verify the skill of the raw and calibrated re-forecasts against ERA5.

The first forecast quality, discrimination, measures the ability of the forecasts to distinguish between an event and the corresponding non-event, which can be described in terms of hit rate and false alarm rate respectively. Hit rate, also known as probability of detection, occurs when an observed “hit” event is correctly forecast. False alarm rate, also known as the probability of false detection, occurs when an observed “hit” event is incorrectly forecast as a “no hit”. Discrimination of the raw re-forecasts was assessed using the Relative Operating Characteristic (ROC) curve, by plotting the results of hit rate against false alarm rate between the forecasts and observations (Kharin and Zweirs, 2003). The verification score was measured by the area under the curve, computed as a scalar value known as the ROC score.

The ROC score ∈ [0,1], where 1 signifies a perfect score in discrimination, while a score of more than (less than) 0.5 represents better (poorer) skill of the forecast over the climatological reference. In addition, the skill of the calibrated re-forecasts was measured using the associated ROC Skill Score (ROCSS):

$$\text{ROCSS} = \frac{ROC_{cal}-ROC_{raw}}{1-RO\ _{raw}} \tag{1}$$

where ROCSS ∈ [-∞,1], ROC_{raw} is the ROC Score for the raw re-forecast, and ROC_{cal} the ROC Score for the calibrated re-forecast. ROCSS is based on tercile categories (above normal, near normal and below normal) that are independently computed for the re-forecasts and observations, effecting a bias correction in the forecasts, and making the ROCSS insensitive to biases as they are independent of calibration (Wilks, 2011).

The second forecast quality that will be analysed, accuracy, is a measure of the overall correspondence between forecasts and observations. The accuracy of the raw re-forecasts was assessed using the Continuous Ranked Probability Score (CRPS). CRPS is a quadratic integral measure of the difference between the forecast cumulative distribution function (CDF) and the empirical CDF of the observation (Hersbach, 2000), expressed by:

$$\text{CRPS} = \int_{-\infty}^{\infty} (\,\mathrm{F_{fcst}}(\mathrm{y}) - \mathrm{F_{obs}}(\mathrm{y}))^2 \,\mathrm{dy} \qquad (2)$$

where CRPS ∈ [0,1]. $F_{obs}(y)$ is a cumulative-probability step function going from 0 to 1 at the point where the forecast variable y equals the observation. Figure 1(b) illustrates how three different Gaussian probability distribution functions (PDF) in Figure 1(a) are expressed as CDFs, and compared against observations. A smaller area enclosed between the CDF and the step-function represents better correspondence, and thus, better CRPS. The value 0 denotes perfect score, thus, a lower (higher) score would mean better (poorer) forecast skill. CRPS is sensitive to biases unlike the ROC score, as it takes into account the continuous distribution function and does not divide the ensemble re-forecasts into probabilities of categorical events. The associated skill score, (CRPSS) is computed using the bias-corrected (calibrated) re-forecast's CRPS to verify against the raw re-forecast's CRPS (taken as the reference score):

$$\text{CRP Skill score (CRPSS)} = 1 - \frac{\mathrm{CRPS_{cal}}}{\mathrm{CRPS_{raw}}} \qquad (3)$$

where CRPSS ∈ [-∞,1]. ROCSS and CRPSS values above (below) 0 indicate the particular calibration improves (degrades) the raw re-forecast. These skill scores provide an objective and useful gauge of any skill improvements by the calibration.

Results and Discussion

Figure 2 shows the computed March mean temperature bias for the 1-month lead ensemble mean of the raw and calibrated (using the MVA method) SEAS5 re-forecast model against ERA5. In general, the raw re-forecast model displays negative bias (colder) over land, as expected based from similar results by Manzanas et al. 2019 for this region, and positive bias (warmer) over sea. Mean temperature bias values have been effectively reduced by the MVA method as values are close to 0 in Figure 1. The biases were reduced because all calibrations are applied under the LOOCV, which repeatedly fits the calibration re-forecasts using training sets that contained *n-1* samples, almost as many as there are in the entire dataset for testing. Plots of the other two methods (CCR and LR) are not shown here as they produced similar results to the MVA method. The maximum and minimum 'mean temperature bias' values are presented in Table 1, highlighting that further calibration is needed to correct the raw re-forecast model predictions.

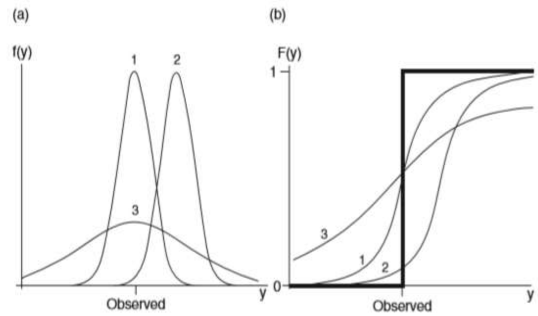

Figure 1 Schematic diagram of the continuous ranked probability skill score (Wilks, 2011): (a) shows three forecast PDFs and (b) illustrates their corresponding CDFs, with the step-function CDF for the observation.

Table 1 Calculated mean minimum and maximum temperature biases for raw and calibrated temperature re-forecasts based on the SEAS5 model.

Bias	Min. (°C)	Max. (°C)
Raw	-3.00	1.61
Calibrated MVA	-0.0269	0.0162
Calibrated CCR	-0.0300	0.0172
Calibrated LR	-0.0298	0.0172

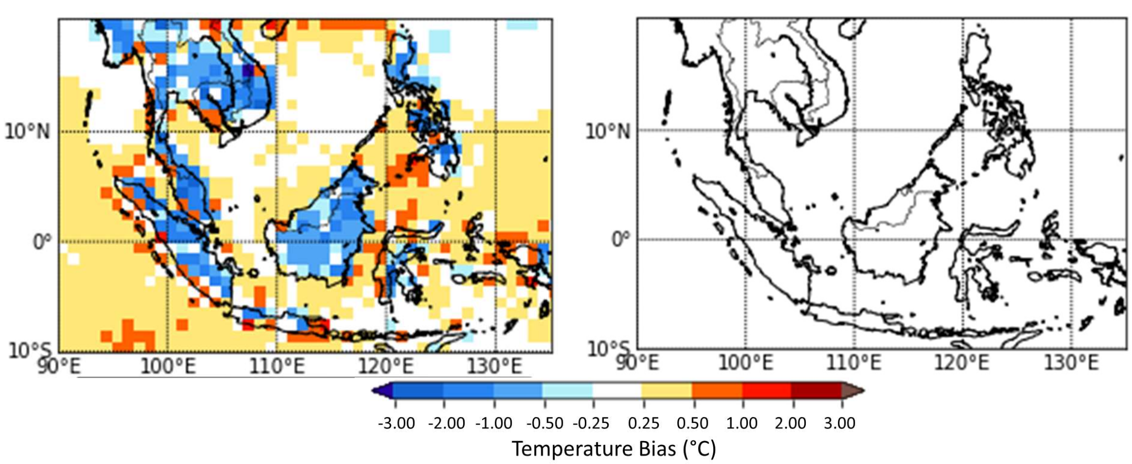

Figure 2 Raw (left) and MVA calibrated (right) mean temperature biases based on the ECMWF's SEAS5 model from ERA5 reanalysis reference. The biases are expressed in degrees Celsius (°C).

Above Normal **Below Normal**

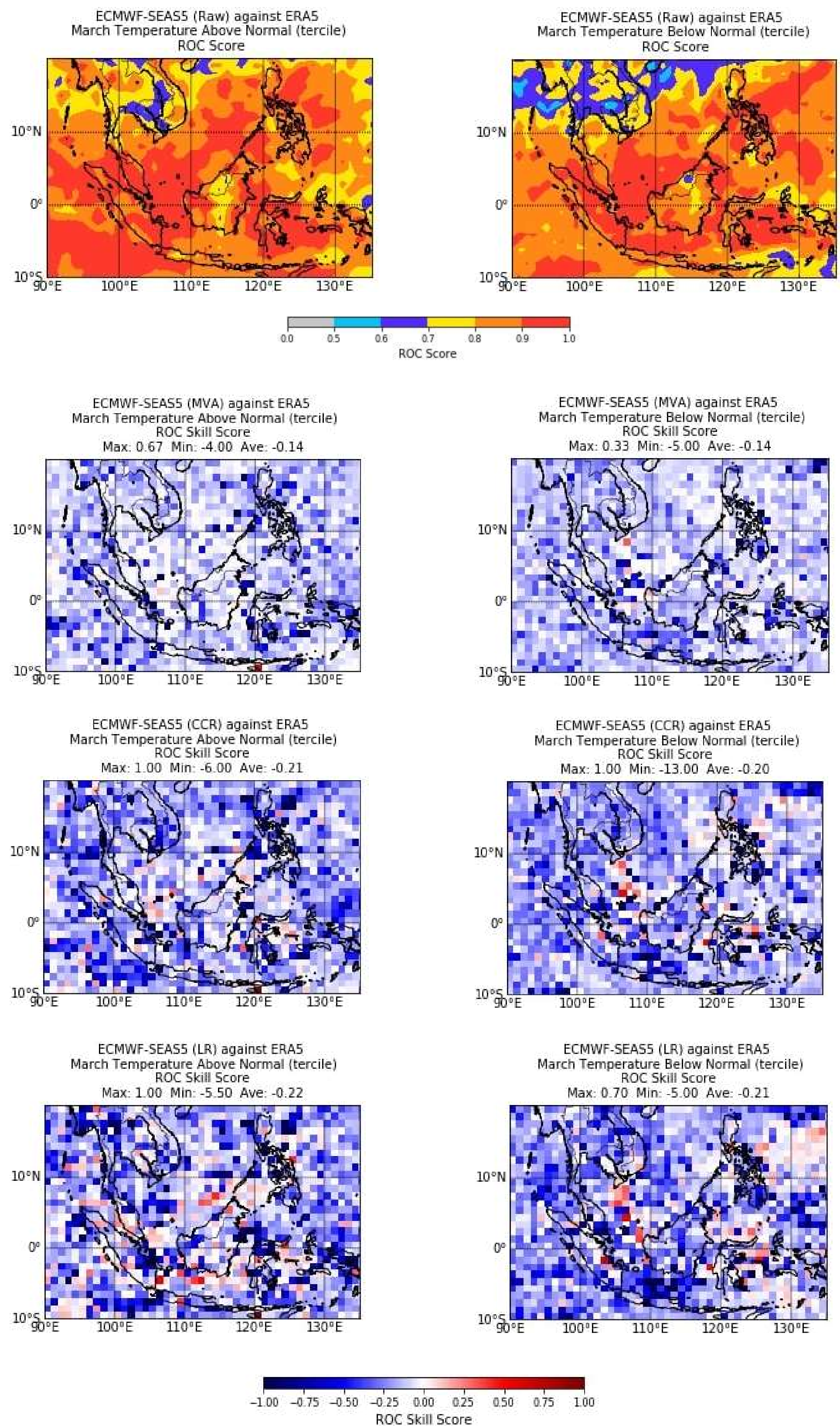

Figure 3 Above-normal Tercile (column 1) and Below-normal Tercile (column 2) for ROC Score of the raw re-forecast (row 1) and ROC Skill Score of calibrated re-forecasts using MVA, CCR and LR method respectively (rows 2-4).

Figure 3 illustrates ROC scores and skill scores for the raw and calibrated re-forecasts in the Southeast Asia region. The plots shown here are tercile-based probabilistic predictions, leaving out the near-normal tercile for brevity. Based on the performance of skill, most areas of the raw re-forecast model (row 1) achieved relatively good skill, as the majority of the ROC scores were > 0.5. This suggests that the raw model re-forecasts can be considered skilful.

On the other hand, the calibrated re-forecasts (rows 2-4) did not improve discrimination based on their ROC skill scores. The scores in most areas of Southeast Asia degraded after calibration, with ROCSS < 0. This was more evident for the RC (CCR, LR) methods than for the BA (MVA) method. The average ROC Skill Score achieved for MVA, CCA and LR-calibrated re-forecasts were all negative, -0.14, -0.21 and -0.22 for the above-normal tercile respectively, and -0.14, -0.20 and -0.21 for the below-normal tercile respectively.

There are some possible reasons for the poor skill reflected by ROCSS for the calibrated re-forecasts. Firstly, it could be due to the shorter re-forecast period (24 years) used for this study. A longer re-forecast period has been shown to result in a slight improvement in ROCSS (Manzanas et al. 2019). It would also mean a larger number of cross-validation points, resulting in a more robust estimate of skill. However, computing for a longer re-forecast period can be expensive and is hence not always feasible. Secondly, ROCSS is insensitive to biases. These biases are not reflected in the ROC curve because the actual numerical value of the forecast probabilities are not considered in the computations for the ROC curve, and are only used to sort the elements of the joint distribution (Wilks, 2011).

Overall, among the calibration methods, the MVA method displayed the least degraded ROC Skill Score.

Figure 4 shows the CRPS and its skill scores for the raw and calibrated re-forecasts respectively. The raw re-forecasts show almost close to perfect correspondence over sea areas as shown in Figure 4(a). However, there is a negative correspondence over land, as the CRPS for most land regions displays a score of approximately 1 or slightly more. This is due to the negative model temperature bias over the land (shown in Figure 2).

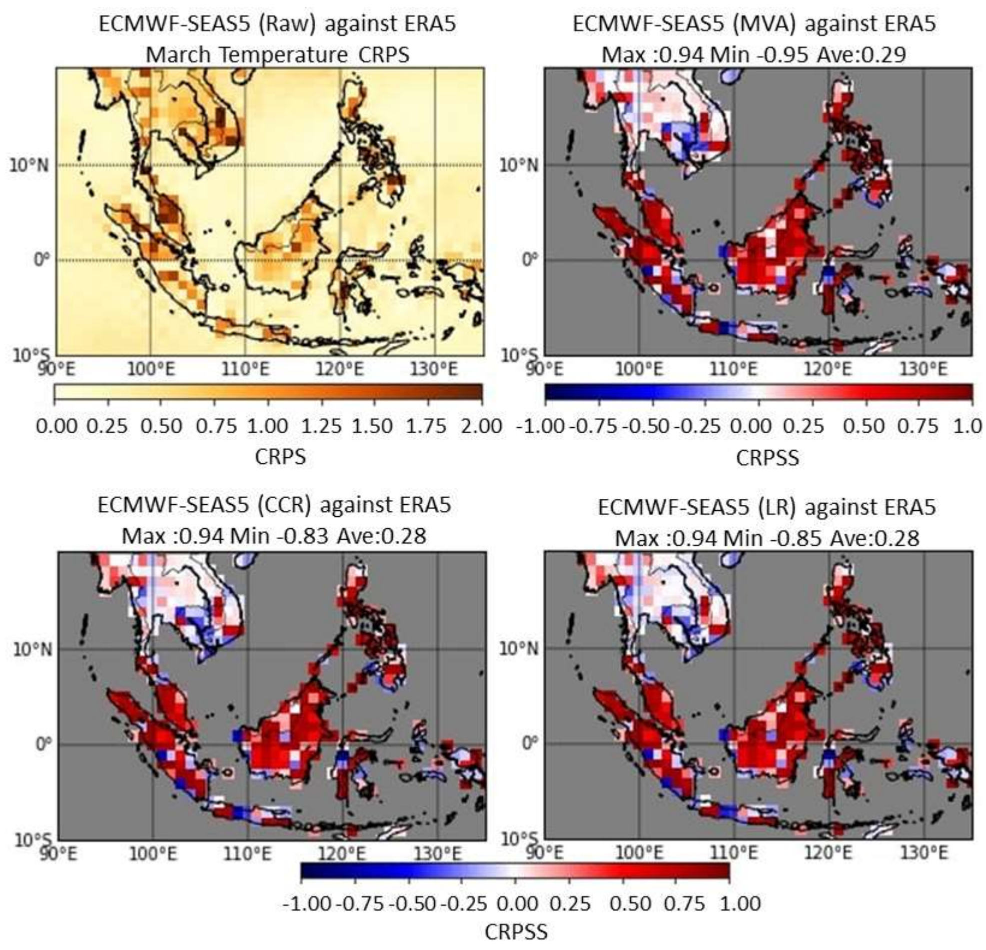

Figure 4 CRP Score plot of raw re-forecast (row 1, column 1) and CRP Skill Score plot of calibrated re-forecasts using MVA, CCR and LR method respectively (row 1 column 2, row 2 column 1 & 2). Land-sea mask (in grey) is applied for only the calibrated re-forecast plots to mask out the sea region.

After calibration, the re-forecasts displayed much better skill over land than over the sea. Hence, we decided to focus on studying the skill over land only, by applying a land-sea mask on the gridded calibrated re-forecasts as shown in Figure 4. Approximately 23% of the grids are considered land within our domain. The red (blue) areas on land imply that the calibration method degraded (improved) the raw model re-forecast. All three calibration methods showed similar CRPSS skill improvements for regions like Singapore, the Philippines, Malaysia, Sumatra and Borneo. However, some parts of the northern SEA region (Vietnam, Laos, Cambodia, and Thailand) displayed insignificant CRPSS skill improvements (scores were close to or less than 0). A similar observation was made for the northern SEA region where the ROC scores (above/below) were low to moderate. This could be due to the predictability limit in the current ECMWF-SEAS5 model at this time of the year, which cannot be compensated for by the calibration methods used in this study.

The percentage of total land area where the CRPSS scores were greater than 0 was also investigated. On average, 71% (66%, 64%) of the total land area indicated an improvement in CRPSS using the calibration method of MVA (CCR, LR). Overall, among the calibration methods surveyed, the MVA method displayed the highest percentage of CRPSS improvement over land regions.

CONCLUSION

This study explored three calibration methods (MVA, CCR and LR) for an ensemble re-forecast of 2-metre temperature over Southeast Asia using bias-adjustments and ensemble recalibration. All calibration methods effectively reduced the model's mean temperature biases and improved the corresponding bias-sensitive metrics using CRPS/CRPSS, which is crucial for end-users to compute specific climate indices based on absolute values/thresholds. As a result of the LOOCV setting, bias-insensitive verification metrics like the ROC Score and its corresponding skill score indicated more degradation than improvement in our calibrated re-forecast.

Overall, the MVA method displayed the best performance based on forecast discrimination (ROC) and accuracy (CRPS). One of the potential advantages of using the MVA method (which is a subset of BA methods) is its suitability for daily timescales. The daily data could be used to compute specific indices (e.g. heat waves, length of growing index, and thermal comfort index), or for various sectoral applications (e.g. health, agriculture, water resources, and energy).

Future work could extend the study to other times of the year (different seasons) and using other parameters like precipitation for more robust comparisons of different calibration methods. There is also potential scope to use bias-adjustment methods in sub-seasonal models (on a timescale of 2 weeks to 2 months) and for a combination of calibrated multi-model ensembles.

ACKNOWLEDGEMENTS

The authors would like to thank Copernicus Climate Change (C3S) Service for data access, and credits to the SantanderMetGroup and Yuanchao Xu for the R-library packages used in our study.

REFERENCES

Climate Data Store (2019) C3S API Climate Data Store. Available at: https://cds.climate.copernicus.eu, last accessed: 12 December, 2019.

Doblas-Reyes, F.J., Hagedorn, R., Palmer, T.N. (2005) The rationale behind the success of multi-model ensembles in seasonal forecasting — II. Calibration and combination. *Tellus A: Dynamic Meteorology and Oceanography*, 57:3, 234-252.

Hersbach, H. (2000) Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems. *Journal of Weather and Forecasting*, 15:559-570.

Hersbach, H., Dee, D. (2016) ERA5 reanalysis is in production. ECMWF Newsletter 147. Available at: https://www.ecmwf.int/en/newsletter/147/news/era5-reanalysis-production, last accessed: 12 December, 2019.

Johnson, S.J., Stockdale, T.N., Ferranti, L., Balmaseda, M.A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G. et al. (2019) SEAS5: the new ECMWF seasonal forecast system. *Geoscientific Model Development*, 12, 1087-1117.

Kharin, V.V., Zweirs, F.W. (2003) On the ROC Score of Probability Forecasts. *Journal of Climate*, 16, 4145-4150.

Manzanas, R., Gutiérrez, J.M., Bhend, J., Hemri, S., Doblas-Reyes, F.J., Torralba, V., Penabad, E., Brookshaw, A. (2019) Bias adjustment and ensemble recalibration methods for seasonal forecasting:

a comprehensive intercomparison using the C3S dataset. *Journal of Climate Dynamics*, 53: 1287.

Mason, I.B. (1999) Integrated verification procedures for forecasts and warnings. Available at: https://www.cawcr.gov.au/projects/verification/Mason/IntegratedVerificationProcedures.pdf, last accessed 23 September, 2019.

Murphy, A.H. (1993) What is a Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting. *Weather Forecasting*, 8:281-293.

Santander Meteorology Group (2019) Santander Meteorology Group Repositories. Available at: http://github.com/santandermetgroup, last accessed: 6 December, 2019.

Torralba, V., Doblas-Reyes, F.J., MacLeod, D., Christel, I., Davis, M. (2017) Seasonal Climate Prediction: A New Source of Information for the Management of Wind Energy Resources. *Journal of Applied Meteorology and Climatology*, 56: 1231-1247.

Weigel, A.P., Liniger, M.A., Appenzeller, C. (2009) Seasonal Ensemble Forecasts: Are Recalibrated Single Models Better than Multimodels? *Monthly Weather Review*, 137:1460-1479.

Wilks, D.S. (2011) Statistical methods in the atmospheric sciences, 3rd Edition (International geophysics series), Vol. 100. ISBN 978-0-12-385022-5.

Xu, Y (2015) hyfo: Hydrology and Climate Forecasting R Package for Data Analysis and Visualization. Available at: http://yuanchao-xu.github.io/hyfo, last accessed: 6 December, 2019.

APPENDIX

The following notations are used in the subsequent equations for each calibration method: $y_{m,t}$ and $y'_{m,t}$ are the original and calibrated values for the ensemble member *m* at the time *t (month)*, $\hat{y}$ is the average of the ensemble mean (y_t) over all times t, $\hat{o}$ is the mean of the observations over all times t, σ_f is the standard deviation of the complete ensemble forecast and σ_o is the standard deviation of the observations. ρ is the correlation between the ensemble mean and the observational reference.

MEAN/VARIANCE ADJUSTMENT (MVA)

The mean and variance of the ensemble forecast is adjusted to the observations via this formula (Torralba et al. 2017):

$$y'_{m,t} = (y_{m,t} - \hat{y})\frac{\sigma_f}{\sigma_o} + \hat{o} \qquad (4)$$

Monthly mean anomalies are calculated by subtracting the ensemble mean, $\hat{y}$, of the monthly averages, from the monthly average $y_{m,t}$ of each forecast for each year and member. The new monthly mean (calibrated value), $y'_{m,t}$, is calculated by multiplying the monthly mean anomaly by the ratio of the standard deviation of the reference observations to the standard deviation of the ensemble members, and adding the climatological mean $\hat{o}$ of the reference dataset (observations). In this equation, it is assumed that both the reference observations and re-forecasts are approximated well by a Gaussian (normal) distribution.

CLIMATE CONSERVING RECALIBRATION (CCR)

$$y'_{m,t} = \rho\frac{\sigma_o}{s.d.(\bar{y}_t)}\,\bar{y}_t + \sqrt{1-\rho^2}\,\frac{\sigma_o}{\sigma_f}\left(y_{m,t} - \bar{y}_t\right) + \hat{o} \quad (5)$$

Also known as variance inflation, this method modifies the predictions so that they have the same interannual variance as the observational reference at every grid point for all times t, by scaling the average of the ensemble mean of the forecast with a factor $\alpha = \rho\,\frac{\sigma_o}{s.d.(\bar{y}_t)}$ and correcting the underestimation or overestimation of the ensemble spread by scaling it with a factor of $\beta = \sqrt{1-\rho^2}\,\frac{\sigma_o}{\sigma_f}$, while preserving their interannual correlation (Doblas-Reyes et al. 2005; Weigel et al. 2009, Torralba et al. 2017) such that the re-forecast is identical to the observation climatology. The inflation in the ensemble spread ($y_{m,t} - \bar{y}_t$) by the factor β quantifies the uncertainty of the actual outcome and increases the reliability of the predicted probabilities. Factors α and β are found under constraints that 1) the standard deviation of the inflated prediction is the same as that for the observation, and 2) the forecast signal after the inflation is made equal to the correlation of the ensemble mean with the observation.

LINEAR REGRESSION RECALIBRATION (LR)

This method performs a linear regression between the ensemble mean and the corresponding observations: $o_t = \alpha + \beta\bar{y}_t + \epsilon$, where α is the intercept, β is the slope and ϵ is the residual error. It is similar to CCR, where the ensemble mean and ensemble spread are each being scaled by a particular factor, in order to achieve similar results for predictions and observations. To correct the forecast variance, the standardized anomalies are rescaled by the standard deviation of the predictive distribution (re-forecast) from the linear fitting, calculated using:

$$y'_{m,t} = \alpha + \beta\bar{y}_t + \gamma_t\left(y_{m,t} - \bar{y}_t\right) \qquad (6)$$

where $\gamma_t = std\left(\epsilon_{fit}\right)\sqrt{1 + \frac{1}{n} + \frac{\left(y_t - \bar{y}_t\right)^2}{(n-1)var(\epsilon_{obs})}}$,

ϵ_{fit} and ϵ_{obs} are the residuals from the regression and the observations respectively and *n* is the number of samples used.

GLOSSARY

Climate Conserving Recalibration (CCR): A statistical technique that corrects for biases in a dataset, where the correction varies in time (similar to LR).

Climate Model Intercomparison Project (CMIP): An international experimental framework set up by the World Climate Research Programme (WCRP) for the studying and assessment of the output from the many different global climate models (GCMs). It allows, among other things, the output of the many different climate models to be made publically available in a standardized format.

Cold surges: Strong northeasterly winds over the South China Sea that bring increased convection over the Maritime Continent during the Northeast Monsoon.

Continuous Ranked Probability Score (CRPS): A score that assesses the difference between what was observed and the probabilistic forecast. The skill score (CRPSS) compares the CRPS from the model with some reference, such as climatology.

CPC Morphing Technique (CMORPH): A gridded observational rainfall dataset based on satellite imagery.

European Centre for Medium-Range Weather Forecasts (ECMWF): A research institute and operational numerical weather prediction centre dedicated to improving forecasts in the 7 to 15-day window. It also provides additional forecasts, including those at the subseasonal to seasonal timescale.

Global Climate Models (GCMs): Models that use mathematical equations to describe the climate of the entire globe.

Global Producing Centres (GPCs): Centres that produce forecasts that cover the entire globe. Certain GPCs have been designated by the World Meteorological Organization as WMO Global Producing Centres of Long-Range Forecasts that adhere to required international standards.

Intergovernmental Panel for Climate Change (IPCC): An intergovernmental body of the United Nations for assessing the science related to climate change.

K-means clustering: A technique to partition a dataset into a certain number (k) of clusters.

Linear Regression Recalibration (LR): A statistical technique that corrects for biases in a dataset, where the correction varies in time and assumes there is a linear relationship between the observations and the dataset to be corrected.

Mean/Variance Adjustment (MVA): A statistical technique that corrects both the mean (average) bias, as well as error in the variance.

Moisture flux: The rate of flow of moisture (e.g. specific humidity).

Python: A general purpose programming language, in which code readability is important.

Re-forecasts: Similar to forecasts, but run for past dates and often uses reanalysis data as input. Re-forecasts are used in subseasonal and seasonal forecasting.

Regional Climate Models (RCMs): Similar to GCMs, but run for a region (or limited area) in order to achieve a finer resolution.

Relative Operating Characteristics (ROC) curve: Used to determine if a forecast is different (or able to discriminate) between instances when an event happens and when it does not happen (e.g. above normal temperature). The skill score (ROCS) compares the area under the ROC curve from the model with some reference, such as climatology.

Representative Concentration Pathways (RCPs): Greenhouse gas concentration trajectories (variations with time) that were adopted by the IPCC in 2014.

Singapore Variable Resolution (SINGV): MSS's integrated numerical weather prediction system, developed jointly in collaboration with the UK Met Office

Tropical Rainfall Measuring Mission (TRMM): A gridded rainfall observation dataset based on satellite imagery.

Unified Model (UM): A numerical model of the atmosphere developed by the UK Met Office that is used for both weather and climate applications.

Urban Heat Islands (UHIs): An urban or built up area that is considerably warmer than the surrounding areas.

V2, V3: Singapore's 2nd and 3rd National Climate Change studies, respectively.

Western Maritime Continent (WMC): The western part of the Maritime Continent, which includes Singapore, Peninsular Malaysia, Sumatra, as well as parts of Borneo and Java.

Meteorological Service Singapore

Website: ccrs.weather.gov.sg

Address: 36 Kim Chuan Road, Singapore, 537054