UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ÉVALUATION DU CYCLE JOURNALIER DANS LES SIMULATIONS DU GRAND ENSEMBLE CLIMEX

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN SCIENCES DE L'ATMOSPHÈRE

PAR
ANNE-MARIE BÉGIN

MARS 2021

REMERCIEMENTS

Je tiens à remercier mes deux directeurs de mémoire, Martin Leduc et Annie Poulin, pour m'avoir encouragée et orientée tout au long de mes travaux. Un merci à tous mes professeurs qui m'ont transmis leur savoir et leur passion pour la science. Finalement, merci aux membres de ma famille pour leur support inconditionnel, ainsi qu'à mon conjoint pour ses encouragements et son amour.

TABLE DES MATIÈRES

LISTE DES FIGURES

RÉSUMÉ

Pour estimer l'impact des changements climatiques sur notre société, les scientifiques se basent habituellement sur des projections climatiques produites à l'aide de modèles numériques. Ces modèles permettent entre autres d'évaluer les effets de l'augmentation des concentrations atmosphériques de gaz à effet de serre (GES) sur le climat ainsi que la variabilité naturelle du système climatique. Il est connu que la température moyenne mondiale va augmenter et que l'occurrence, l'intensité et la distribution spatio-temporelle des précipitations extrêmes vont changer. Ces évènements météorologiques extrêmes causent des sécheresses, des inondations et d'autres catastrophes naturelles qui engendrent des conséquences importantes sur notre vie et notre environnement. En particulier, la précipitation est une variable clé dans l'adaptation aux changements climatiques.

Cette étude se base sur le grand ensemble ClimEx, composé de 50 simulations indépendantes permettant d'étudier l'effet des changements climatiques et de la variabilité naturelle sur l'hydrologie des bassins versants du Québec. Cet ensemble a été produit en utilisant le Modèle Régional Canadien du Climat (MRCC5) sur une grille à 12 km de résolution en utilisant comme conditions aux frontières latérales des simulations produites par le modèle Canandien du système terrestre de deuxième génération (CanESM2) à 310 km de résolution spatiale horizontale.

L'objectif de ce projet est d'évaluer les performances de l'ensemble ClimEx à simuler le cycle journalier et à représenter les valeurs extrêmes. Pour y arriver, 30 ans de données horaires pour la précipitation et aux trois heures pour la température sont analysées. Les données des 50 membres de ClimEx sont comparées à celles d'une simulation du MRCC5 pilotée par la réanalyse ERA-Interim, à la réanalyse ERA5 et à des stations météorologiques d'Environnement et Changement climatique Canada (ECCC). Une évaluation de la sensibilité de différentes statistiques climatiques au nombre de membres dans l'ensemble est aussi effectuée.

Le cycle journalier de la précipitation montre une corrélation généralement non significative avec une amplitude qui est similaire à celles des données d'observation des stations d'ECCC. Pour la température de l'air près de la surface, la corrélation est forte et l'amplitude du cycle est semblable à celles des observations. ClimEx représente généralement bien les 95, 97 ,99$^{\text{ième}}$ quantiles pour la précipitation. Pour la température, il représente souvent bien les quantiles, mais avec un biais chaud sur la partie sud du Québec. Toutefois, des écarts importants ont été observés pour les valeurs maximales de ClimEx, qui sont jusqu'à 10 fois supérieures à celles d'ERA5 pour la précipitation mais l'écart est moins grand avec les données d'ECCC. Pour la température, les valeurs maximales et minimales montrent jusqu'à 20°C d'écart au-dessus des valeurs d'ERA5 et l'écart est inférieur par rapport à ECCC. Le seuil de sensibilité au nombre de membres sur les 95, $99^{\text{ième}}$ quantiles et sur le cycle moyen pour la précipitation se situe entre 15 et 50 membres de l'ensemble, pour obtenir une erreur inférieure à 5%. Pour la température, une erreur inférieure à 0.5°C implique l'utilisation de 1 et 17 membres pour les 95, $99^{\text{ième}}$ quantiles et de 1 à 2 membres pour le cycle moyen.

INTRODUCTION

Dans le contexte actuel des changements climatiques, il est devenu essentiel d'améliorer notre capacité à évaluer le climat futur avec une plus grande précision (Dessai *et al.*, 2009). Les gouvernements et les différentes autorités ont besoin d'informations claires sur le climat pour le futur proche et lointain, afin de planifier l'adaptation de notre société face aux changements climatiques (Porter *et al.*, 2015). Le climat est le résultat d'un système thermodynamique complexe, ce qui rend sa prévision difficile et bien sûr incertaine. Pour comprendre pourquoi et comment le climat va changer en termes de conditions météorologiques, les scientifiques utilisent des projections climatiques basées sur des modèles numériques. Les modèles climatiques permettent d'évaluer l'impact de l'augmentation des concentrations de gaz à effet de serre (GES) ainsi que la variabilité naturelle sur le climat (Deser *et al.*, 2012).

De nos jours, les phénomènes météorologiques extrêmes sont de plus en plus visibles (Trenberth *et al.*, 2015). Actuellement, il est bien connu que l'augmentation des concentrations atmosphériques de GES entraîne une augmentation de la température moyenne mondiale (Pachauri *et al.*, 2014). En termes de précipitation, l'occurrence, l'intensité et la distribution spatio-temporelle des événements extrêmes vont également changer (Innocenti *et al.*, 2019). Afin de nous adapter à ces changements, nous devons estimer la probabilité d'occurrence de ces événements extrêmes. La précipitation est une variable clé pour l'adaptation aux changements climatiques (Hawkins et Sutton, 2009), car elle a d'importantes conséquences sur notre vie et notre environnement. Il est aussi connu que les précipitations sont largement affectées par les modes naturels de variabilité du climat (Hoerling *et al.*,

2010) comme l'oscillation australe El Niño (ENSO) (Ropelewski et Halpert, 1986) et l'oscillation nord-atlantique (Zhang *et al.*, 2010). De ce fait, il est primordial d'étudier la variabilité naturelle du climat à travers les modèles climatiques.

Dans cette étude, les simulations du grand ensemble ClimEx sont utilisées (Leduc *et al.*, 2019). Ce grand ensemble de données a été créé afin d'étudier l'effet de la variabilité naturelle et des changements climatiques sur l'hydrologie des bassins versants au Québec et en Bavière, avec un accent particulier sur les applications liées aux événements extrêmes. Cet ensemble de simulations a été produit à l'aide du MRCC5 sur une grille à 12 km de résolution, piloté à ses frontières latérales par 50 membres de conditions initiales différentes du modèle global CanESM2. Des simulations régionales ont été produites sur le domaine du nord-est de l'Amérique du Nord (NNA) avec une période horaire d'archivage pour le taux de précipitation et une période de trois heures pour la température de l'air près de la surface.

Les projections provenant des modèles climatiques sont soumises à plusieurs sources d'incertitude. Les trois principales sources d'incertitude sont : la formulation du modèle, le scénario d'émission (par exemple RCP 8.5) et la variabilité naturelle du climat (Hawkins et Sutton, 2009). À l'échelle régionale, la variabilité naturelle est plus dominante (Hawkins et Sutton, 2009). Les membres de l'ensemble ClimEx sont le produit d'un seul modèle et d'un seul scénario d'émissions, ce qui ne permet pas d'évaluer l'incertitude due à la formulation du modèle et au scénario, mais plutôt d'étudier en détail la variabilité naturelle du climat.

Comme le démontre Lorenz (1963), le fait de modifier légèrement les conditions initiales d'un système dynamique non linéaire comme l'atmosphère peut conduire à des états très différents après un certain temps. Dans le cadre de ClimEx, le même principe a été appliqué à l'initialisation d'un modèle climatique global pour générer un grand ensemble de réalisations climatiques indépendantes.

Cette approche est directement liée aux événements extrêmes qui émergent de cette variabilité naturelle et qui provoquent des inondations, des sécheresses et d'autres catastrophes naturelles. Afin d'analyser plus en détail les conséquences directes de ces phénomènes, des modèles hydrologiques sont souvent utilisés. Ils permettent de simuler le cycle hydrologique à l'échelle du bassin versant afin d'étudier l'effet des conditions météorologiques sur le débit des cours d'eau. La plupart de ces modèles utilisent comme intrants des données quotidiennes (Azarnia, 2017), plus particulièrement les températures maximales et minimales quotidiennes ainsi que la précipitation totale. Les modèles hydrologiques sont habituellement calibrés en comparant les débits historiques observés avec les débits simulés obtenus, en utilisant comme données d'entrée les précipitations et la température observées sur la même période historique. Pour les petits bassins versants (< 500 km^2), les données quotidiennes peuvent avoir une résolution temporelle insuffisante pour capturer correctement la réponse hydrologique des bassins versants (p.ex. crues et étiages). Afin d'améliorer la prévision du débit des rivières au niveau quotidien et donc de faire progresser la science en hydrologie, plusieurs études ont souligné l'importance d'utiliser des informations sous-journalières telles que l'ensemble de données ClimEx pour saisir de manière plus réaliste le débordement des rivières dans les petits bassins versants (Cortés-Hernández *et al.*, 2016; Bevelhimer *et al.*, 2015; Olsson *et al.*, 2015).

L'analyse du cycle journalier passe par différentes méthodes, comme l'utilisation de l'heure à laquelle se produisent les précipitations maximales et minimales (Pfeifroth *et al.*, 2016) et l'étude de l'amplitude et de la phase du cycle journalier (Liang *et al.*, 2004). Pour analyser le cycle journalier, il est nécessaire de disposer de données horaires comme référence avec laquelle comparer. Afin de fournir des données horaires aux modèles hydrologiques, il existe plusieurs modèles régionaux du climat comme par exemple ceux présentés par NA-CORDEX (voir `https://na-cordex.org/sub-daily-precipitation.html`), notamment le Canadian Regional Climate Model version 4 (CanRCM4), le Regional Climate Model system version 4 (RegCM4) ou le Weather Research and Forecasting Model (WRF).

L'énorme coût de calcul lié à la production de grands ensembles comme ClimEx et le grand volume de données résultant de tels projets mettent en évidence l'importance d'estimer la taille de l'ensemble requise pour différentes applications. Dans l'article de Milinski *et al.* (2020), on aborde la question de la taille nécessaire d'un ensemble pour estimer différentes statistiques climatiques à partir de différents critères de précision. Ce type d'analyse permet en particulier d'évaluer l'impact de la variabilité naturelle du climat sur le calcul des statistiques. Il est clair qu'il n'y a pas encore de consensus quant au nombre de membres nécessaires dans les ensembles afin d'estimer l'effet de la variabilité naturelle du climat sur le calcul des statistiques puisque ce nombre de membres dépend fortement du type d'application et de la variable impliquée. Par exemple, Milinski *et al.* (2020) ont montré que plus de 100 membres sont nécessaires pour estimer la réponse au signal des changements climatiques des précipitations sur les régions océaniques avec une erreur acceptable de 0,2 mm/jour. Alors que 25 membres peuvent fournir une bonne estimation des tendances au forçage des GES du réchauffement avec un erreur acceptable de 5%. Les effets de la variabilité naturelle sur différentes statistiques ont aussi été étudiés (Deser *et al.*, 2012; Hoerling *et al.*, 2010). On y mentionne que la part de la variabilité naturelle varie selon les régions et que les précipitations sont plus sensibles que la température (Deser *et al.*, 2012). Hoerling *et al.* (2010) indique que ENSO est une source importante de la variabilité naturelle pour les précipitations à l'échelle rérionale. En particulier en modélisation hydrologique, l'incertitude liée à la variabilité naturelle est importante (Seiller et Anctil, 2014). Selon les travaux de Giuntoli *et al.* (2018), les incertitudes pour la précipitation sont dominées par la variabilité du modèle climatique et la variabilité interne, à quelques exceptions près. D'autres part, la température est plus sensible au scénario RCP que les précipitations. Comme chaque membre représente une trajectoire différente du système climatique, l'effet du nombre de membres ainsi que le seuil minimal quant à la taille de l'ensemble doivent être évalués.

Ce mémoire s'inscrit à l'intérieur d'un projet PSR-SIIRI avec l'École de technologie supérieur (ÉTS) dans le cadre d'une collaboration entre le Québec et la Bavière. Ces deux régions sont influencées par des inondations, généralement causées par la fonte des neiges au printemps ainsi que les précipitations extrêmes en été et en automne. Dans un contexte de changments climatiques, on peut s'attendre à une diminution de la neige en hiver et à une augmentation des précipitations extrêmes. C'est pourquoi dans le cadre de ce mémoire nous nous concentrons sur les saisons d'été et d'automne. Dans le but d'étudier l'influence de ces phénomènes sur l'hydrologie, les données de l'ensemble ClimEx sont utilisées afin de piloter des modèles hydrologiques. Ce mémoire permettra d'évaluer la qualité des données de cet ensemble en vue d'effectuer des simulations hydrologiques à l'échelle sous-journalière.

L'objectif de ce mémoire est d'évaluer la simulation du cycle journalier par le modèle MRCC5 dans l'ensemble ClimEx, en vue d'évaluer la possibilité d'utiliser ces données pour piloter des modèles hydrologiques à l'échelle sous-journalière, dans le cadre d'études d'impacts des changments climatiques sur l'hydrologie. L'évaluation du cycle journalier de ClimEx implique une comparaison avec des données de stations météorologiques d'environnement et changement climatique Canada, la réanalyse ERA5 et une simulation MRCC5 pilotée par la réanalyse ERA-Interim. Une analyse du cycle journalier, de différents quantiles ainsi que des valeurs maximales et minimales de l'ensemble y est présentée, suivie d'une étude de la sensibilité de différentes statistiques au nombre de membres.

Le corps de ce mémoire est écrit sous la forme d'un article qui sera soumis à une revue spécialisée. L'article est structuré comme suit : introduction (section 1.2), méthodologie (section 1.3), résultats (section 1.4), discussion (section 1.5) et conclusion (section 1.6).

CHAPITRE I

EVALUATION OF THE DAILY CYCLE IN THE SIMULATIONS OF THE CLIMEX LARGE-ENSEMBLE

By

Anne-Marie Bégin, Martin Leduc, Annie Poulin

To be submitted to Atmosphere-Ocean

Corresponding author address :
Anne-Marie Bégin
UQAM/Ouranos 550 Sherbrooke St West,19thfloor, West Tower
Montréal, QC, Canada, H3A 1B9
E-mail :begin.anne_marie.2@courrier.uqam.ca

1.1 Abstract

To estimate the impact of climate change on our society we need to use climate projections based on numerical models. These models make it possible to assess the effects on climate of the increase in greenhouse gases (GHG) as well as natural variability. We know that the global average temperature will increase and that the occurrence, intensity and spatio-temporal distribution of extreme precipitations will change. These extreme weather events cause droughts, floods and other natural disasters that have significant consequences on our life and environment. Precipitation is a key variable in adapting to climate change.

This study focuses on the ClimEx large ensemble, a set of 50 independent simulations created to study the effect of climate change and natural variability on the water network in Quebec. This dataset consists of simulations produced using the Canadian Regional Climate Model version 5 (CRCM5) at 12 km of horizontal grid spacing driven by simulations from the second generation Canadian Earth System Model (CanESM2) global model at 310 km of horizontal grid spacing.

The aim of the project is to evaluate the performance of the ClimEx ensemble in simulating the daily cycle and representing extreme values. To get there, 30 years of hourly time series for precipitation and 3-hourly for temperature are analyzed. The simulations are compared with the values from the simulation of CRCM5 driven by ERA-Interim reanalysis, the ERA5 reanalysis and Environment and Climate Change Canada (ECCC) stations. An evaluation of the sensitivity of different statistics to the number of members is also performed.

The daily cycle of precipitation from ClimEx shows mainly non-significant correlations with the other datasets and its amplitude is similar to the observation data from ECCC stations. For temperature, the correlation is strong and the amplitude

of the cycle is similar to observations. ClimEx provides a fairly good representation of the 95, 97, 99th quantiles for precipitation. For temperature it represents often a good distribution of quantiles but with a warm bias in southern Quebec. For precipitation hourly maximum, ClimEx shows values 10 times higher than ERA5 but compared to ECCC the difference is less. For temperature, minimum and maximum values may exceed the ERA5 limit by up to 20°C but compared with ECCC the difference is less. For precipitation, the minimum number of members for the estimation of the 95 and 99th quantiles and the mean cycle is between 15 and 50 for an estimation error of less than 5%. For the 95, 99th quantiles of temperature, the minimum number of members is between 1 and 17 and for the mean cycle 1 to 2 members are necessary to obtain an estimation error of less than 0.5°C.

1.2 Introduction

In the current context of climate change, it has become essential to improve our ability to predict the future climate with greater accuracy (Dessai *et al.*, 2009). Governments and policy makers need clear information about the near and far future climate in order to plan adaptation for our society (Porter *et al.*, 2015). Climate is the result of a complex thermodynamic system which makes its prediction difficult and, of course, uncertain. To understand why and how the future climate will change in terms of meteorological conditions, scientists use climate projections based on numerical models. Climate models allow to evaluate the impact of the increase in greenhouse gas (GHG) concentration as well as natural climate variability (Deser *et al.*, 2012).

Extreme weather events are more and more visible nowadays (Trenberth *et al.*, 2015). We know that the increase in atmospheric concentration of GHG cause the global mean temperature increase (Pachauri *et al.*, 2014). In terms of precipitation, the occurrence, intensity and spatio-temporal distribution of extremes will change as well due to the climate change (Innocenti *et al.*, 2019). To adapt to these changes, we need to estimate the probability of occurrence of these extreme events. Precipitation is a key variable for adaptation to climate change because it has a lot of consequences in our life and environment (Hawkins et Sutton, 2009). We know precipitation is largely affected by natural modes of climate variability (Hoerling *et al.*, 2010) like El Niño southern oscillations (ENSO) (Ropelewski et Halpert, 1986) and the North Atlantic oscillation (Zhang *et al.*, 2010). Therefore, it is essential to study the natural variability of the climate through climate models.

In this study, the set of simulations from the ClimEx large ensemble is used (Leduc *et al.*, 2019). This large dataset was created to study the effect of natural variability and climate change on water networks in Quebec and Bavaria, with a specific focus on applications related to extreme events. This dataset was produced using the CRCM5 at 12 km of resolution driven by 50 different initial condition members from the CanESM2 global climate model. Regional simulations were produced over northeastern North America (NNA) domain with an hourly archival time period for precipitation and a three-hourly period for surface-air temperature.

The accuracy of future projections is linked to the fact that climate model simulations are subject to several sources of uncertainty. The three main sources of uncertainty are : the model formulation, the emission scenario (e.g. RCP 8.5) and natural climate variability (Hawkins et Sutton, 2009). At the regional scale, natural variability is the most dominant source of uncertainty (Hawkins et Sutton, 2009; Deser *et al.*, 2012). Despite that ClimEx ensemble is the product of a single pair of global and regional models and one emission scenario, thus limiting any evaluation of model and scenario uncertainties, the large number of members in the ClimEx ensemble allows a detailed study of the internal climate variability associated with this model chain.

As demonstrated by Lorenz (1963), slightly differing initial conditions of a dynamical non-linear system like the atmosphere can lead to very different states after a certain period of time. In the ClimEx framework, the same principle was applied to the initialization of a global climate model to generate a large ensemble of independent climate realizations.

This approach is directly linked to the extreme events that emerge from this natural variability and that cause floods, droughts and other natural disasters. To visualize in more detail the direct consequences of these phenomena, hydrological models are widely used. They allow to simulate the water cycle at the watershed scale to study the effect of meteorological conditions on stream flow. Currently, most of these models use daily data as input (Azarnia, 2017), which generally consists in daily maximum and minimum temperature and total precipitation. Hydrological models are typically calibrated by comparing the observed historical stream flows with the simulated ones obtained when using observed precipitation and temperature as an input over the same historical time period. In small watersheds (< 500 km^2), daily data may have insufficient temporal resolution to correctly capture the watershed reactivity and stream flow variations and peaks. To improve the prediction of the river flows at a sub-daily level and thus to advance the science of hydrology, several studies highlighted the importance of using sub-daily information such as the ClimEx dataset to more realistically capture the runoff behaviour in small watershed (Cortés-Hernández *et al.*, 2016; Bevelhimer *et al.*, 2015; Olsson *et al.*, 2015).

The analysis of the sub-daily cycle goes through different methods, such as using the time at which the maximum and minimum precipitation occur (Pfeifroth *et al.*, 2016).The study of the amplitude and phase of the daily cycle is a method used by Liang *et al.* (2004). To evaluate the daily cycle it is necessary to have hourly data as a reference to compare with. To provide hourly data to hydrological models, there are several models presented in NA-CORDEX (`https://na-cordex.org/sub-daily-precipitation.html`) like the Canadian Regional Climate Model version 4 (CanRCM4), the Regional Climate Model system version 4 (RegCM4) or the Weather Research and Forecasting (WRF) Model.

The huge computational cost related to the production of a large ensemble like ClimEx and the big data volume resulting from such projects emphasize the importance of estimating the required ensemble size for different applications. In Milinski *et al.* (2020), the question of the minimal ensemble size required to estimate different statistics based on different precision criteria is investigated. In particular, this kind of analysis allows an interpretation of the impact of natural variability in the calculation of climate statistics. It is clear that there is no consensus yet on the number of members that are necessary to adequately capture the effect of natural variability as it strongly depends on the type of application and variables involved. For example Milinski *et al.* (2020) have shown that more than 100 members are necessary to estimated the forced rainfall response over the ocean regions with an acceptable error of 0.2 mm/day while 25 members can provide a good estimation of the forced warming trends with an acceptable error of 5%.The effects of natural variability on different statistics were also studied in Deser *et al.* (2012) and Hoerling *et al.* (2010). It is shown that the magnitude of natural variability varies by region and that precipitation is more sensitive than temperature (Deser *et al.*, 2012). Hoerling *et al.* (2010) highlight that ENSO is an important source of the natural variability of precipitation at regional scales. In particular for hydrological modelling, the uncertainty related to natural variability can be important according to Seiller et Anctil (2014). According to Giuntoli *et al.* (2018), precipitation uncertainty is dominated by model and internal variability with a few exceptions and temperature is more sensitive to the RCP scenario than precipitation. Since the effect of internal variability depends on many factors such as the variable, season and region of interest, the minimal ensemble size required for different applications must be evaluated.

This paper is organised as follows : introduction (section 1.2), methods (section 1.3), results (section 1.4), discussion (section 1.5) and conclusion (section 1.6).

1.3 Methods

1.3.1 Ensemble description

The Climate Change and Hydrological Extremes (ClimEx) large ensemble was designed to study the effects of natural variability and climate change on hydrographic networks in Quebec and Bavaria (Leduc *et al.*, 2019). It was produced by dynamically downscaling the Canadian Earth System Model large ensemble (CanESM2-LE) at 310 km of resolution to 12 km of resolution using the CRCM5. The CanESM2-LE is formed by an ensemble of simulations differing by small differences in their initial conditions. This large ensemble was produced from a 1000 years equilibrium simulation forced by preindustrial conditions. To this simulation were applied five random perturbations (on cloud overlap) in 1850, thus leading to five simulations that were run for 100 years until 1950. Then, on each of these five runs were applied ten new random perturbations in 1950. This framework results in fifty members (labelled as ClimEx (50) in the following) aiming to span the range of natural variability from 1950 to 2100. Between 1950 and 2005, the forcing applied corresponds to the observed concentration of GHG, aerosols and to solar forcing. After 2005, the RCP 8.5 emission scenario is used. Another CRCM5 simulation was also produced between 1979 and 2013 using ERA-Interim reanalysis as lateral boundary conditions (labelled as ClimEx ERA-I). The latter simulation allows to assess the model performance when driven by lateral boundary conditions that closely represent observed climate. The complete information about run labelling and outputted variables can be found at `www.climex-project.org/en/ensemble-documentation`.

1.3.2 Study area and reference dataset

The domain studied in this research based on the ClimEx project is northeastern North America (NNA). Over a period of 30 years between 1981 and 2010, for summer (JJA) and fall (SON), the daily cycle has been evaluated for surface air temperature and precipitation, which are two variables of high interest in hydroclimatology. The two seasons chosen are in relation with the interests of a PSI RIISI project in collaboration with the École de technologie supérieure (ÉTS) as part of a Quebec and Bavaria collaboration. The project focuses is on extreme precipitation events during the summer and fall seasons. The CRCM5 simulations are evaluated for various statistics by comparing with the ERA5 reanalysis at 31 km of resolution (for information see `www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5`), which is increasingly used in the literature as a reference dataset (Tetzner *et al.*, 2019; Olauson, 2018; Graham *et al.*, 2019). In a second step, some regions are compared with data from the ECCC meteorological stations.

The stations were chosen north and south of the 50 th parallel to represent two types of climate in Quebec (see figure 1.2). In each region, only the stations with more than 80% of available data over the evaluation period are selected. By applying this selection criterion for precipitation, 7 stations are found to be usable in the South region and 2 for the North region. For temperature, 5 stations are selected for the south and 2 for the north. To compare ClimEx with the ECCC and ERA5, the cycle is evaluated using the nearest grid point to each station. For the two regions of interest, we have averaged the different stations and corresponding nearest grid points for model and reanalysis are calculated and then compared. A quality control procedure was applied to hourly precipitation data, where some outlier values (>100 mm/hr) were removed from the time series.

We evaluate the ClimEx dataset by comparing with three reference datasets, that is ClimEx ERA-I, ERA5 and ECCC stations to consider the uncertainties. ERA5 is a reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) using the 4d-var data assimilation method. This method uses different atmospheric observation data and processes it numerically to produce the most similar conditions to observed climate. For stations from ECCC, we have selected those with more than 80% available data although these may still contain some outliers, for instance due to measurement errors.

1.3.3 Dataset comparison and ClimEx evaluation

This section describes all the methods that were used to compare the different datasets and evaluate the ClimEx precipitation and temperature daily cycle.

Regional differences

The first step of the methodology is to represent the seasonal climatic averages for temperature and monthly total precipitation over the entire simulation domain. The ClimEx ERA-I simulation and the average of the 50 members are then compared with ERA5. ERA5 was produced at 31 km resolution and a cubic interpolation method is applied to compare with the CRCM5 model on a common grid at 12 km resolution.

Daily cycle

In connection with the availability of the ClimEx dataset at sub-daily time scales, the daily cycle is determined for temperature using 3-hourly archives and hourly archives for precipitation. The seasonal mean daily cycle is obtained by taking the hourly or 3-hourly averages over thirty years for each season. The daily cycle is then calculated separately for each station and nearest grid points (for ClimEx (50), ClimEx ERA-I and ERA5). All stations (nearest grid points) are then averaged over each region. This approach results in one daily cycle statistics (e.g. seasonal mean daily cycle) per data source, region and season. To represent climatic extremes, the minimum and maximum values are determined for each dataset and then the distribution of data for the 95, 97 and 99 th quantiles are analyzed.

Second-order statistics

Statistics are calculated on the datasets to evaluate the seasonal mean values, the cycle amplitude and its phase. To represent the different daily cycles coming from the 50 members and to compare them with the reference data, Taylor diagrams (Taylor, 2001) are used. This diagram allows us to see the following three characteristics on the same diagram. As shown below, the centred root mean square difference (RMSD) can be expressed as a function of cycle correlation and standard deviation. First, let x_i and r_i be the hourly mean values calculated over the evaluation period for the simulation and reference data respectively with the i index representing the hour of the day, $\bar{x}$ and $\bar{r}$ the mean daily cycle values and N the number of points in the daily cycle. The centred RMSD can thus be calculated as :

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[(x_i - \bar{x}) - (r_i - \bar{r})]^2}. \qquad (1.1)$$

The standard deviation (σ) indicates the amplitude of the daily cycle and can be written as :

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}. \qquad (1.2)$$

Finally, the correlation coefficient (R) assesses the co-variability between the daily cycles of the two series and corresponds to :

$$R = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(r_i - \bar{r})}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(r_i - \bar{r})^2}}. \tag{1.3}$$

It can be shown that (1.1) can be written as a function of both (1.2) and (1.3), leading to the following geometric relationship between the three parameters (Taylor, 2001)

$$RMSD^2 = \sigma_x^2 + \sigma_r^2 - 2\sigma_x\sigma_r R, \tag{1.4}$$

$$c^2 = a^2 + b^2 - 2abcos\phi. \tag{1.5}$$

As (1.5) represents the cosines law, (1.4) can be schematized as in figure 1.1, where a, b, c represent, respectively σ_x, σ_r and RMSD and cosϕ represent R in the triangle.

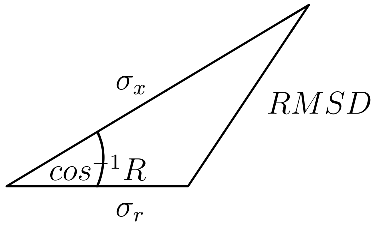


FIGURE 1.1 – Geometrical relationship between the correlation coefficient R, the centred RMSD, and the standard deviations σ_x and σ_r of the test field and reference data, respectively.

Sensitivity of the statistics to the number of members

To estimate the sensitivity of the previous statistics to the number of ensemble members, the bootstrap method is used. Using this method it is possible to estimate the error made in the estimation of a given statistic (e.g. 99th quantile) as a function of the numbers of members sampled from the ClimEx large ensemble. The method consists in randomly sampling with replacement of a certain number of members and determining the statistics based on this ensemble sample. In the present study, 200 samples are drawn for different ensemble sizes from 1 to 50 members, resulting in 200 estimations of the statistics for each ensemble size. By taking the maximum and minimum among these estimated values, the maximum possible error from the best estimate of statistics (using the full 50-member ensemble) is then obtained. The relative error from the best estimate is then analyzed for precipitation while the absolute error is used for temperature.

1.4 Results

1.4.1 Regional difference

The first step of this study is to evaluate the monthly mean precipitation and temperature over the entire ClimEx NNA domain. ERA5 will be compared with the 50 members of ClimEx (ClimEx (50)) and with the run driven by ERA-Interim (ClimEx ERA-I).

In figures 1.3a and 1.4a, we see the mean values for a 30 years period of the interpolated ERA5 data for the seasonal mean precipitation during summer (JJA) and fall (SON) seasons, respectively. In figures 1.3b and 1.4b are shown the differences between ClimEx ERA-I and ERA5 and in 1.3c and 1.4c the difference between ClimEx (50) and ERA5. Over the land in summer, ClimEx ERA-I is generally wetter with some drier spots while ClimEx (50) is much drier over the western part of the domain by more than 50 mm/month. Over the ocean, both datasets are generally wetter with a more important difference for ClimEx (50) than ClimEx ERA-I. In fall, ClimEx ERA-I is generally wetter over land and the storm track while ClimEx (50) is drier over northeastern USA and generally wetter everywhere else.

In figures 1.5a and 1.6a, we see the mean temperature for 30 years period of the interpolated ERA5 data during summer and fall, respectively. In figures 1.5b and 1.6b are shown the differences between ClimEx ERA-I and ERA5 and in 1.5c and 1.6c the difference between ClimEx (50) and ERA5. In summer, ClimEx ERA-I shows relatively small differences ranging between 2°C over the US to -2°C over Quebec. ClimEx (50) shows a similar pattern of differences but with much higher values over the western part of the domain (by around 5°C). In fall, ClimEx ERA-I is again relatively similar to ERA5 while ClimEx (50) is systematically warmer over most land areas.

1.4.2 Daily cycle

The hourly mean values for precipitation rate and temperature for 30 years are shown in figures 1.7 to 1.10 for summer and fall over South and North stations. The curves for ERA5, ECCC and ClimEx ERA-I are plotted in red, blue and green, respectively. For ClimEx (50), the first member (kda) is shown in purple while the 49 remaining members are in grey. In this section we will analyze the diurnal cycle according to its general characteristics, while a more detailed analysis based on the Taylor diagram will follow in the next section.

Figures 1.7a and 1.7b show the seasonal mean daily cycle of precipitation rate in South for summer and fall, respectively. In summer, the shape of the curves is similar, a general increase in the precipitation rate appears at the end of the day, although for ERA5 the maximum is delayed toward the early evening while it appears during the afternoon for ECCC. ERA5 reach a maximum around 0.34 mm/hr, which is very large when compared with the other datasets. In fall, the cycle is generally flatter and similar among datasets. Figures 1.8a and 1.8b show the daily cycle of precipitation rate in North for summer and fall, respectively. In summer, the shape of the curves is similar except for ERA5. In fall, the shape is flatter and similar again but ECCC is drier than the others. It is worth noting from figures 1.8a and 1.8b that ClimEx ERA-I simulations generally lie between the ERA5 and ECCC results.

Figures 1.9a and 1.9b show the daily cycle of temperature in South for summer and fall, respectively. In summer, the shape of the curves is similar but the ClimEx members are warmer than the others. In fall, the shape is similar too with ECCC being warmer while ERA5 and ClimEx ERA-I are colder. Figures 1.10a and 1.10b show the daily cycle of temperature in North for summer and fall, respectively. In summer, the shape and values of the cycles are similar, although the ClimEx members tend to be slightly colder during the afternoon. In fall, the shape of the cycles are similar but there are large offsets between the curves, which is of similar magnitude than the amplitude of the cycles. It is worth noting that the amplitude of the cycles for fall in North is relatively small as compared with the other cycles investigated previously.

1.4.3 Taylor diagram

The Taylor diagram is used to summarize multiple aspects of analysis on a single diagram Taylor (2001). In this study we use it to have a more complete interpretation of the daily cycle, more specifically according to its amplitude and phase. Results presented in this section refer to the figures of the daily cycle presented in the previous section.

In every figure from 1.11 to 1.12, the blue line refers to the ECCC stations reference daily cycle and the grey dots to each member of ClimEx and the ERA5 reanalysis. The green lines refer to the standard deviation (e.g. the amplitude of the daily cycle), the black lines to the correlation coefficient (e.g. the phase of the cycle) and the red lines to the centred RMSD.

The level of correlation is determined by the following absolute criteria from `www.psystat.at.ua/Articles/Table_Pearson.PDF`. For precipitation (df = 24-2) and a significance level of 95%, a correlation greater than R = 0.34 is considered significant. For the temperature (df = 8-2) this value is 0.62 because of the smaller sample. A one-tail test is used here because a negative correlation does not make sense in this context.

Figures 1.11a to 1.11d refer to the daily cycle of precipitation in South and North for summer and fall, respectively. The correlation is mainly non-significant for precipitation. For the amplitude of the cycle, it is generally similar. It should be noted that the correlation for ERA5 is also non-significant while its amplitude is generally larger than ClimEx.

Figures 1.12a to 1.12d refer to the daily cycle of temperature in South and North for summer and fall, respectively. The correlation is strong for temperature and slightly more in summer. For the amplitude of the cycle, it is generally similar to observation. It should be noted that the correlation for ERA5 is also strong while its amplitude is generally smaller than ClimEx.

1.4.4 95, 97, 99th quantiles

In this section, we evaluated how the extreme values are simulated by the CRCM5 in the ClimEx large ensemble. More specifically, the distributions of data are analyzed in terms of quantiles. In this section a comparison is performed between ERA5, ClimEx(50) and ECCC for the 95, 97, 99th quantiles. In figures 1.13 to 1.16, the red line represents ERA5, the black line ClimEx (50) and the blue line ECCC and the pink envelope represents the limit of the ClimEx quantiles.

Figures 1.13a and 1.13b show quantile values of precipitation in South for summer and fall, respectively. In summer, the shape of the daily cycles are generally similar, although ERA5 shows a bit noisy values between 10 am and 8 pm in summer. It can also be noted that the ClimEx quantiles show a slightly shifted daily cycle with a maximum around 8 pm while the observed maximum is rather around 3 pm. For fall, the cycles are generally flat while the quantile values are relatively similar among datasets. Figures 1.14a and 1.14b refer to quantiles of precipitation in North for summer and fall, respectively. In summer, the shape of the lines for ClimEx (50) and ECCC are similar, while ERA5 shows a different behaviour where the 95 to 99th quantiles are much closer from one another (between 0 am and 3 pm) as compared with ClimEx (50) and ECCC. For fall, the cycles are also similar, although noisy.

It is important to note that ClimEx shows smoother curves because the quantiles are evaluated using all 50 members rather than one single realization as for ERA5 and ECCC. Estimation errors related to the sampling will be analyzed in more details in the section 1.4.6.

Figures 1.15a and 1.15b refer to the 95, 97, 99th quantiles of temperature in South for summer and fall, respectively. In summer, the shape of lines is similar, although ClimEx is warmer. For fall, the shape is similar too with the difference between ClimEx and ECCC data being less important but ERA5 is still generally colder. Figures 1.16a and 1.16b refer to quantiles of temperature in North for summer and fall respectively. In summer, the shape and amplitude of cycles are similar as the three quantiles generally overlap across datasets. In fall, the shape is also similar but ClimEx is colder.

1.4.5 Maximum and minimum values

Beyond the 99th quantile, there are larger extremes that inform us about the maximum possible values within the ClimEx set. In figures 1.17 to 1.20, the red line represents ERA5, the black line ClimEx (50), the green line ClimEx ERA-I, the blue line ECCC and the pink envelope represents the limit of the ClimEx (50) maximum to minimum.

Figures 1.17a to 1.18b refer to the maximum and minimum values of precipitation in South and North for summer and fall, respectively. Each of the figures shows that the ClimEx (50) dataset produces extreme maximum values higher than ERA5, up to ten times higher. We note and exception for fall in the South region where ECCC stations show much higher maximum values, which might be related to the quality of station data as will be discussed in section 1.5.

Figures 1.19a to 1.20b refer to the maximum and minimum values of temperature in South and North for summer and fall, respectively. Each of the figures shows that the ClimEx (50) set produces extreme maximum and minimum values higher and lower than ERA5, respectively, leading to a range of values higher by 20°C in ClimEx (50).

It is worth noting that the ClimEx ERA-I and ECCC generally lies between ClimEx (50) and ERA5.

1.4.6 Sensitivity of the statistics to the number of members

In this section, we are interested in estimating the relative error of precipitation and absolute error for temperature that can be obtained as a function of the number of members on 95, 99^{th} quantiles and the mean cycle. To get there, we use the bootstrap technique.

As an example for precipitation, the results are obtained as follows. First, for each statistic we determine the maximum and minimum values among the 200 random samples, leading to a graph as in figure 1.21. Then the absolute error is obtained by subtracting the maximum value minus the minimum value and divided by two. Then, we retain one value of absolute error per ensemble size and statistics by selecting the maximum error found in each set, which are presented in figure 1.22. As expected, the absolute error is larger for higher quantiles. The relative error is then calculated by dividing the absolute error by the real value from the full

50-member ensemble. From the relative error, we define a 5% error threshold to determine the minimal number of members that are required to reach this level of statistics precision. Here, we analyze only the relative error for precipitation while the results of the absolute error are presented in appendix A. For temperature, the absolute error was selected instead in order to determine the minimal number of members necessary to obtain an error of less than 0.5°C.

Figures 1.23a and 1.23b refer to the relative error of precipitation associated with South for summer and fall, respectively. In summer, it takes 23 members to get an error of less than 5% for the 99^{th} quantile and the mean, while the 95^{th} quantile gets below the 5% error threshold only around 50 members. In fall, it takes 15 to 22 members to get an error of less than 5%. Figures 1.24a and 1.24b refer to the relative error of precipitation associated with North for summer and fall, respectively. In summer, it takes 22 to 34 members to get an error of less than 5% based on the different statistics. In fall, it takes 18 to 32 members to get an error of less than 5%.

Figures 1.25a and 1.25b refer to the absolute error of temperature associated with South for summer and fall, respectively. In summer, it takes 1 to 4 members to get an error of less than 0.5°C based on the different statistics. In fall, it takes 1 to 8 members to get an error of less than 0.5°C. Figures 1.26a and 1.26b refer to the absolute error of temperature associated with North for summer and fall respectively. In summer, it takes 2 to 9 members to get an error of less than 0.5°C based on the different statistics. In fall, it takes 2 to 17 members to get an error of less than 0.5°C.

1.5 Discussion

The analysis of the daily cycle in ClimEx involves several aspects, especially the amplitude and phase of the cycle, and its related spatio-temporal distribution. Sometimes it is difficult to evaluate hourly data, particularly for extremes, since only one realization of reality is available, and therefore sampling errors are generally large. A general observation is that temperature shows a higher level of correlation than precipitation in terms of reproducing the observed daily cycle.

For precipitation, it can be seen that the simulation driven by ERA-Interim has a similar behaviour to the ClimEx simulations driven by CanESM2. The correlation between ClimEx members and ECCC is mainly non-significant, probably due to the influence of the high natural climate variability of the precipitation field. However, as the ECCC curve leaves the ClimEx envelope in particular for the maximum values of precipitation during fall in South, it is likely that it is the consequence of other sources of uncertainty such as errors in the observation data or the difficulty of the model to simulate the cycle correctly. It is worth noting that while some outlier values of hourly precipitation were removed from the ECCC station data, this quality control procedure might be insufficient and therefore some unrealistic values could still affect these results. It was observed that the daily cycle of ECCC precipitation was mostly drier than the others probably due to measurement problems at the stations. The type of equipment used to measure rainfall tend to under catch the precipitation (Sieck *et al.*, 2007). By comparing the different quantiles it is interesting to see that ClimEx, ECCC and ERA5 show similar behaviours with one another despite some exceptions even if ClimEx shows smoother curves. As far as the maximum values are concerned, it is clear that ClimEx produces greater precipitation extremes than ERA5 and ECCC. However, this important difference might not be only due to model structural differences, but could also be partly explained by the large number of members available in ClimEx that provide more potential to approach higher maximum values as compared with the unique realization of ERA5 or ECCC. There is an exception in fall for South stations, the ECCC data shows higher maximums than ClimEx (50) probably due to the erroneous values included in the data at the stations.

To estimate the minimum number of members that must be used to obtain an error lower than a given threshold for different statistics, the bootstrap method is used. As expected for the absolute error, the higher the quantile is, the larger the error of estimation is (see figure 1.22). However, the relative error is much less sensitive to the quantile because larger absolute errors are partly compensated by their higher associated quantile (figures 1.23a to 1.24b). For precipitation, it was found that 15 to 50 members are required to obtain a relative error lower than 5% when considering the 95 and 99th quantiles and the mean of the cycle. The maximum relative error reached is 32% for all statistics. These results show that precipitation statistics are quite sensitive to the number of members.

For temperature, the daily cycle of ClimEx (50) over 30 years is very similar according to observations data from ECCC and ERA5 reanalysis. In our study, we find strong correlations between cycles from different datasets. The main difference lies more in the bias between the cycles, rather than in its amplitude and phase. This result is observed for both mean cycle and different quantiles. Again, we see that ClimEx produces maximum and minimum values beyond the ERA5 and ECCC limits. We also note that the differences between the cycles are lower in summer. To evaluate the 95 and 99th quantiles, very few members are necessary for temperature. We estimate that 1 to 17 members are required for a maximum error of 0.5°C. To estimate the mean cycle, 1 to 2 members is needed for keeping the error below 0.5°C. The maximum absolute error reached is 1.4°C for all statistics.

The temperature is not very sensitive to the number of members.

1.6 Conclusion

Improving our ability to simulate and predict with accuracy the multiple aspects of the daily cycle is essential to better prepare for climate change. The objective of this study was to evaluate the daily cycle of simulations from the ClimEx large ensemble, such as its mean and extreme values. Our adaptation involves, among other things, knowing the impacts of climate change on the water network. Especially, the current study is important to evaluate and improve hydrological models on a sub-daily scale.

The daily cycle was calculated over 30 years for the summer and fall seasons in two regions, located in the North and South parts of Quebec province of Canada. The simulated mean cycle and its extreme values were compared with those obtained from three reference datasets, namely ECCC stations, ClimEx ERA-I and the ERA5 reanalysis. As the objective of this study was to compare ClimEx with station data and since the data from ECCC stations are sparsely and not uniformly distributed across Quebec, the study cannot be representative of the whole ClimEx simulation domain. It is also sometimes difficult to make strong conclusions given the uncertainty of station data and the difficulty of judging aberrant values. It is for this reason that we applied a partial quality control and applied a threshold of 100 mm/hr for the data at the stations. There is also the issue related to the spatial scales involved, as ClimEx and ERA5 use grid meshes with different spatial resolutions, but also since the stations only represent point estimates.

The daily precipitation cycle shows generally non-significant correlation with ECCC data. The amplitude of the cycle is similar than observation from ECCC stations. For temperature, the daily cycle shows a strong correlation with the reference datasets and the amplitude of the cycles is very similar to the ECCC data.

The representativeness of the extreme values as simulated in the ClimEx dataset was assessed. This analysis involves the comparison of the 95, 97, and 99$^{\text{th}}$ quantiles between ClimEx (50), ERA5 and ECCC. For precipitation, the distribution of quantiles is often similar to the two reference datasets but with few exceptions. For the temperature, the distribution of the quantiles is generally similar despite a warm bias in southern Quebec. As for the maximum and minimum values produced by ClimEx, they are higher than those of ERA5 by more than 10 times for precipitation and by more than 20°C for temperature, but the difference is smaller with the values of ECCC.

This study also relates to the evaluation of the minimum number of members required to assess climate statistics below a given error threshold. This can be seen as another way to evaluate the effect of natural variability on different statistics (mean cycle, 95$^{\text{th}}$ quantile, 99$^{\text{th}}$ quantile). The results demonstrated that precipitation requires from 15 to 50 members according to the region and season in order to obtain an error lower than 5% relative to the real estimate obtained from the entire 50-member ensemble. For temperature, using 1 to 17 members is sufficient to estimate the 95 and 99$^{\text{th}}$ quantile with an error of less than 0.5°C. For the temperature mean cycle 1 to 2 members are required. As expected, precipitation is more sensitive to the number of members than temperature in relation to the effect of natural climatic variability on the estimate of climate statistics. It should be noted that the results of random peak could be slightly different if the bootstrap procedure is repeated.

As future work, further studies could compare with other observation datasets like the Canadian Precipitation Analyze (CaPA) that shows better performance in many regions of Quebec for most watersheds according to the study of Bajamgnigni Gbambie *et al.* (2017). The evaluation of other variables of interest in hydrology (e.g. wind and radiation variables) could be conducted. It would be of interest for the scientific community to pursue the development of methods for analyzing large ensembles. Especially, because large ensembles provide estimates that are much more precise then they are for a single realization of reality. This is does not mean that the simulations are of better quality than reality since they could be very precise but far from reality. One advantage of continuing work on these large datasets would be to better define the limits of their use and evaluation. While high-resolution large ensembles like ClimEx imply an enormous amount of data which may rise some technical limitations for a complete analysis, they offer quite a large range of possibilities to the climate science community in terms of new analysis methods for model evaluation and the characterization of climate variability and extreme events.

1.7 Figures

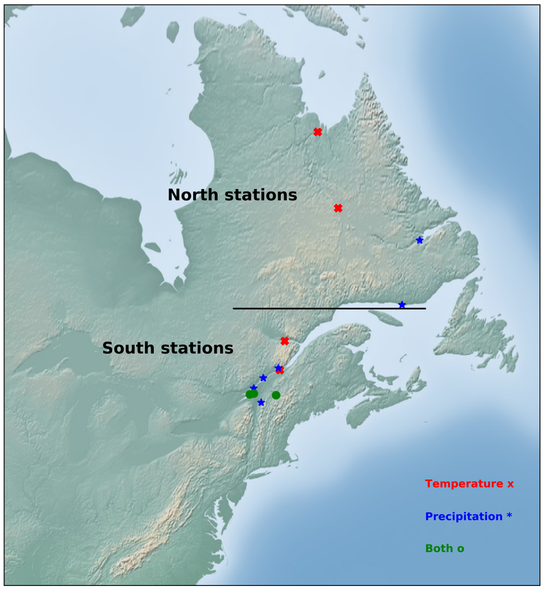


Figure 1.2 – Map of ECCC stations used for temperature and precipitation, respectively. The analysis of the results is separated in two regions, labeled as North and South.

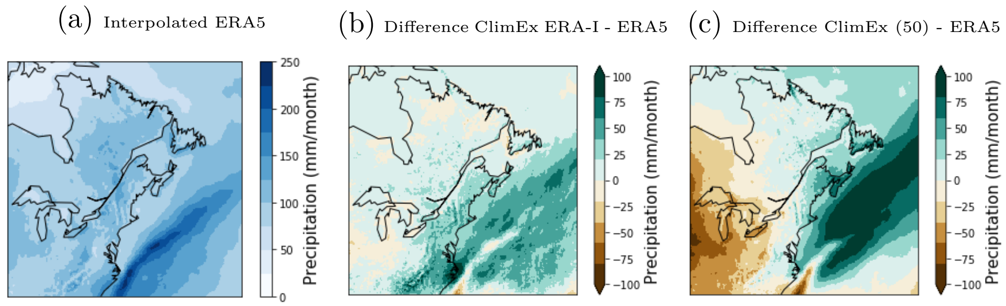


FIGURE 1.3 – Monthly mean precipitation (mm/month) in summer (JJA) over the ClimEx domain for (a) ERA5 reanalysis (b) ClimEx ERA-I (c) ClimEx (50).

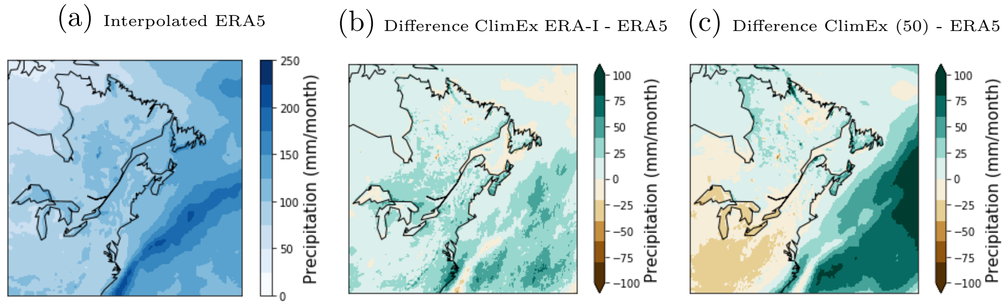
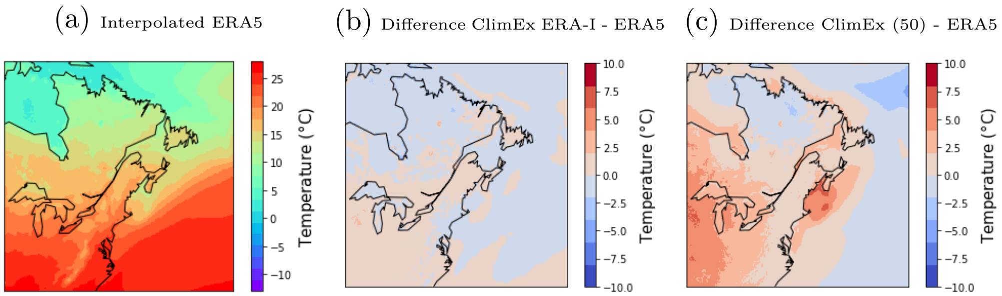
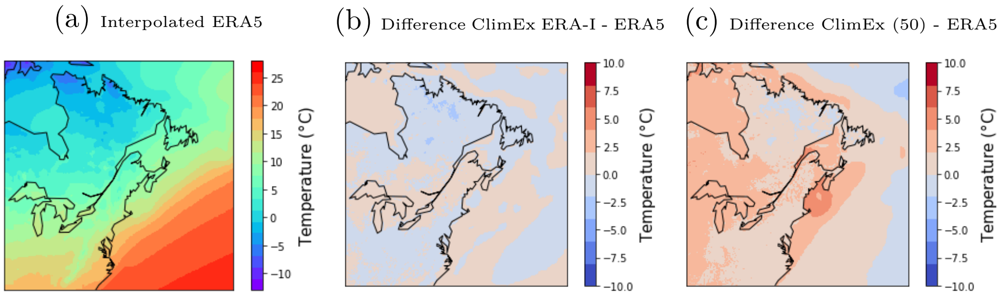


FIGURE 1.4 – Same as figure 1.3 but for fall (SON).



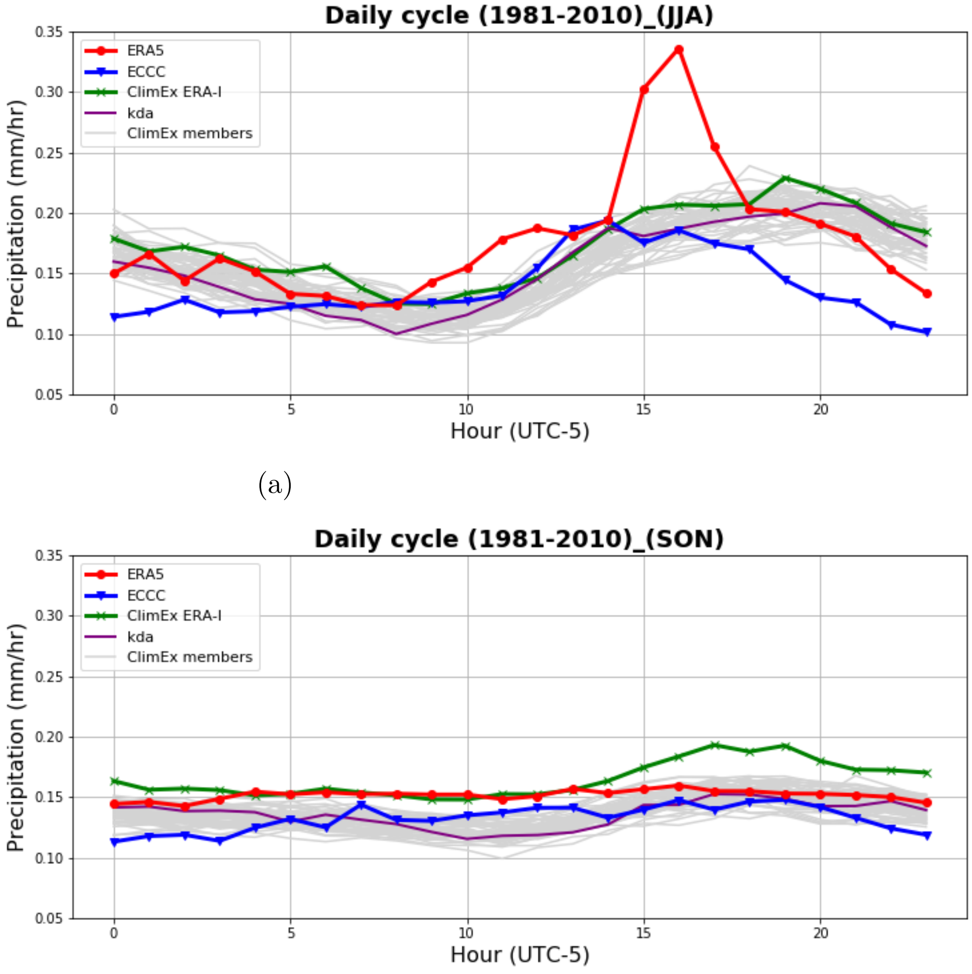
(a) Interpolated ERA5 (b) Difference ClimEx ERA-I - ERA5 (c) Difference ClimEx (50) - ERA5

Figure 1.5 – Average temperature (°C) in summer (JJA) over the ClimEx domain for (a) ERA5 reanalysis (b) ClimEx ERA-I (c) ClimEx (50).



(a) Interpolated ERA5 (b) Difference ClimEx ERA-I - ERA5 (c) Difference ClimEx (50) - ERA5

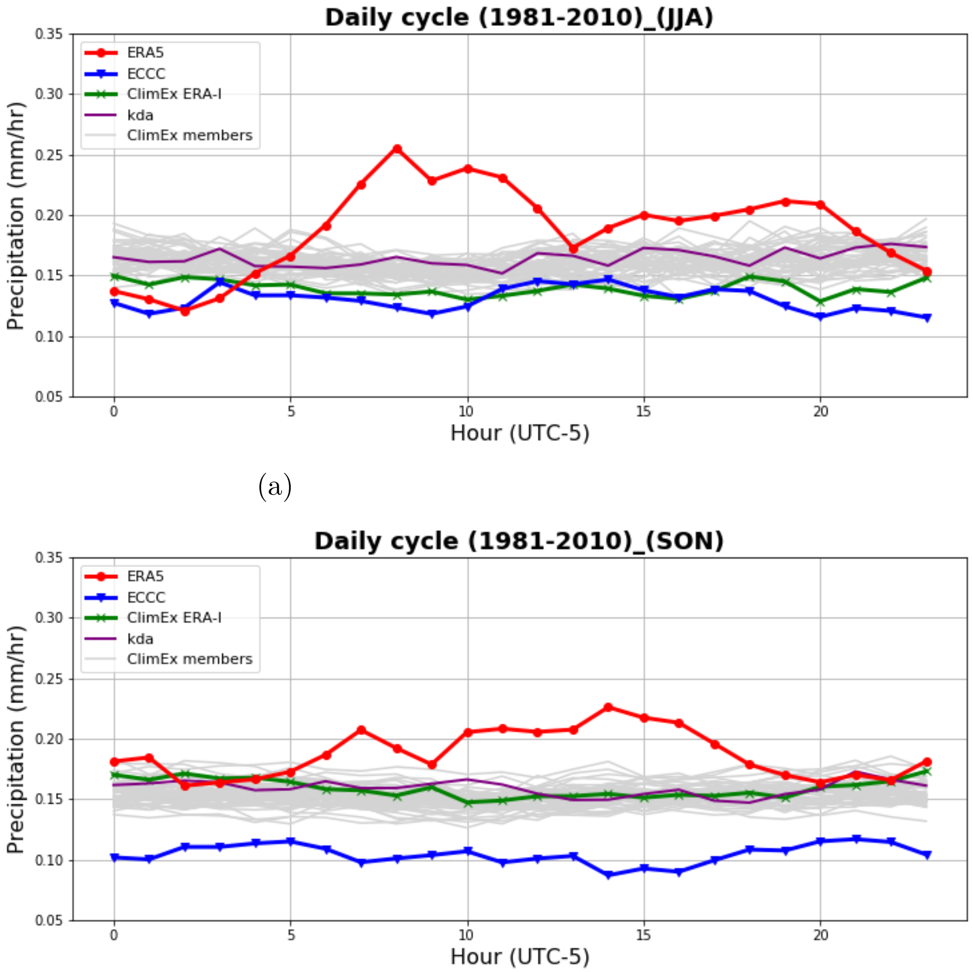
Figure 1.6 – Same as figure 1.5 but for fall (SON).



(a)

(b)

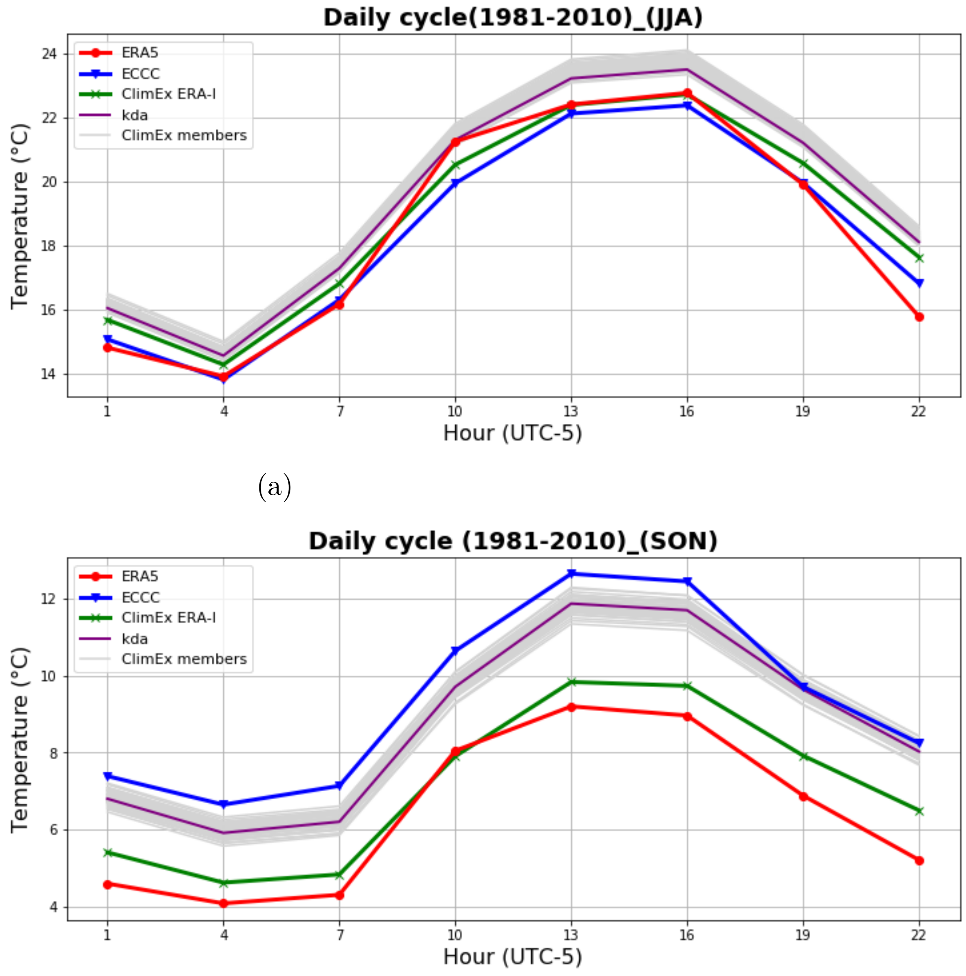
FIGURE 1.7 – Daily cycle climatological mean calculated for the hourly precipitation rate in the South region for (a) summer (JJA) and (b) fall (SON). The ECCC stations data located in the South region are averaged and the same approach is applied to the corresponding nearest grid points for model and reanalysis.



(a)

(b)

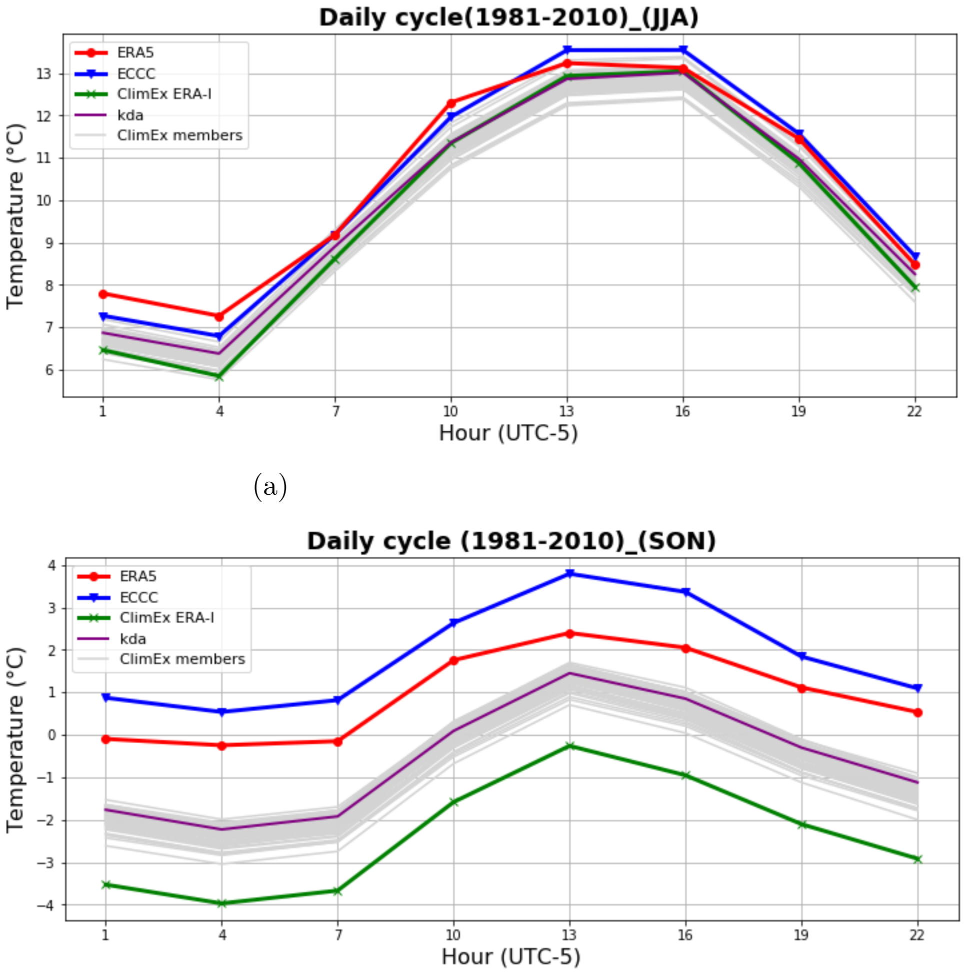
FIGURE 1.8 – Same as figure 1.7 but for North stations.



(a)

(b)

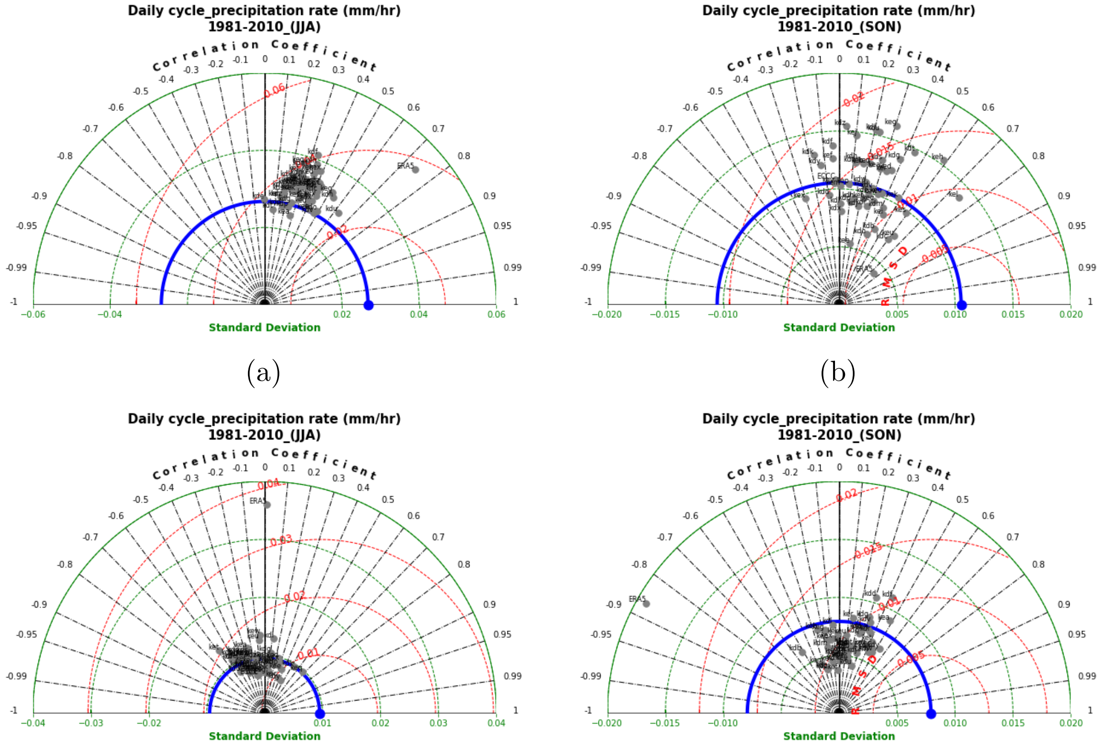
FIGURE 1.9 – Daily cycle climatological mean calculated for temperature in the South region for (a) summer (JJA) and (b) fall (SON). The ECCC stations data located in the South region are averaged and the same approach is applied to the corresponding nearest grid points for model and reanalysis.



(a)

(b)

Figure 1.10 – Same as figure 1.9 but for North stations.



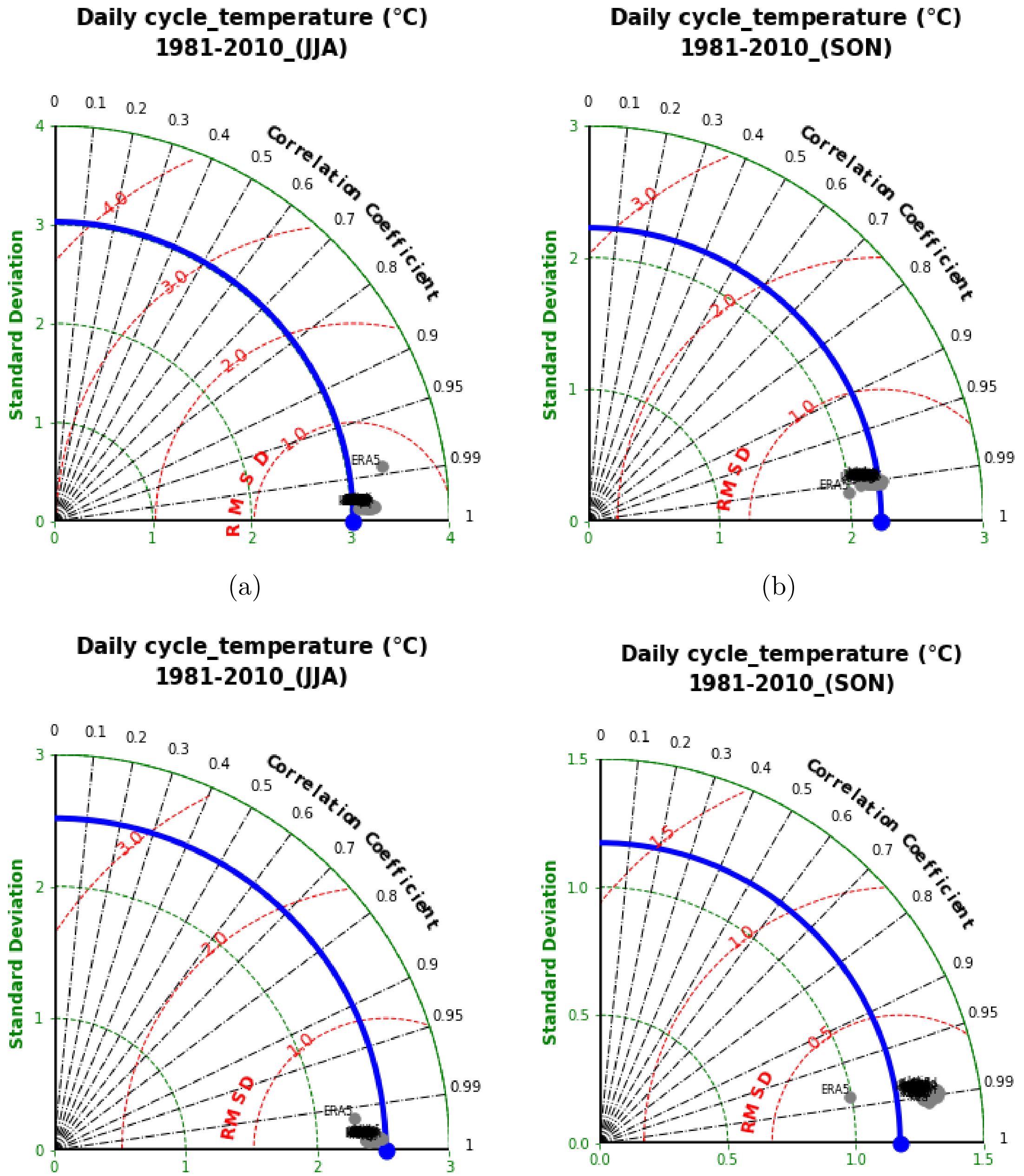
(a)

(b)

(c)

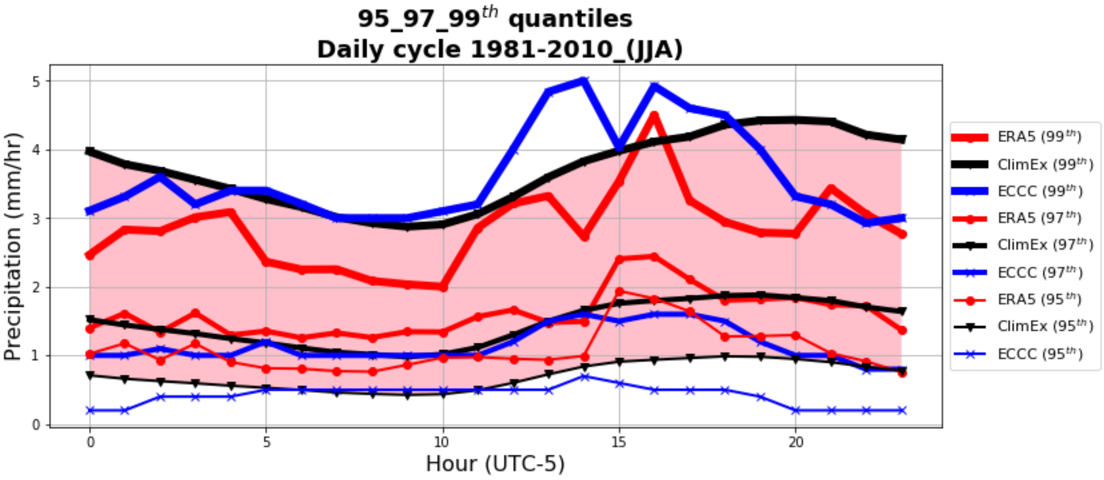
(d)

FIGURE 1.11 – Taylor diagram - Precipitation rate (mm/h). The blue line is the observation, green lines (standard deviation), black lines (correlation), red lines (RMSD). (a) Summer for South stations (b) Fall for South stations (c) Summer for North stations (d) Fall for North stations.

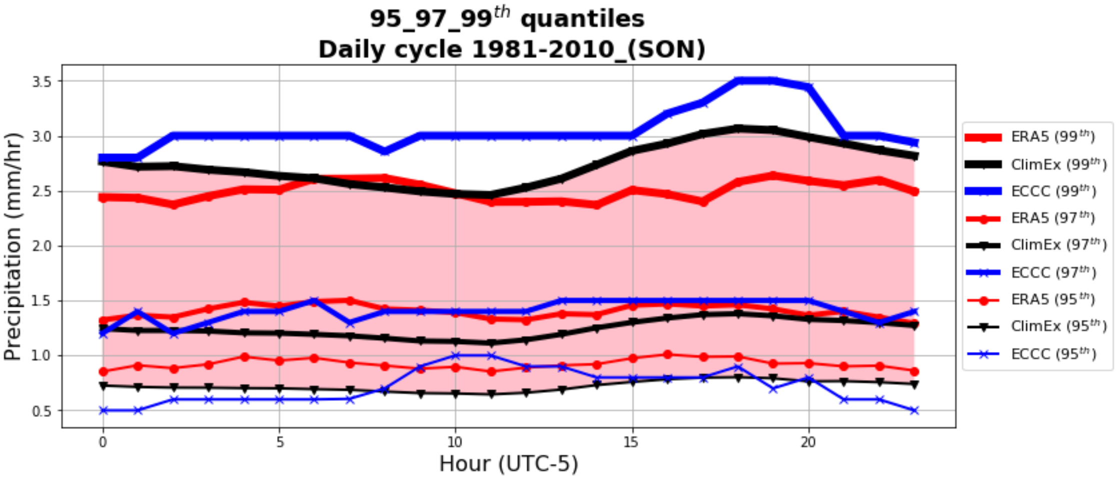


(a) (b) (c) (d)

FIGURE 1.12 – Taylor diagram - Temperature (°C). The blue line is the observation, green lines (standard deviation), black lines (correlation), red lines (RMSD). (a) Summer for South stations (b) Fall for South stations (c) Summer for North stations (d) Fall for North stations.

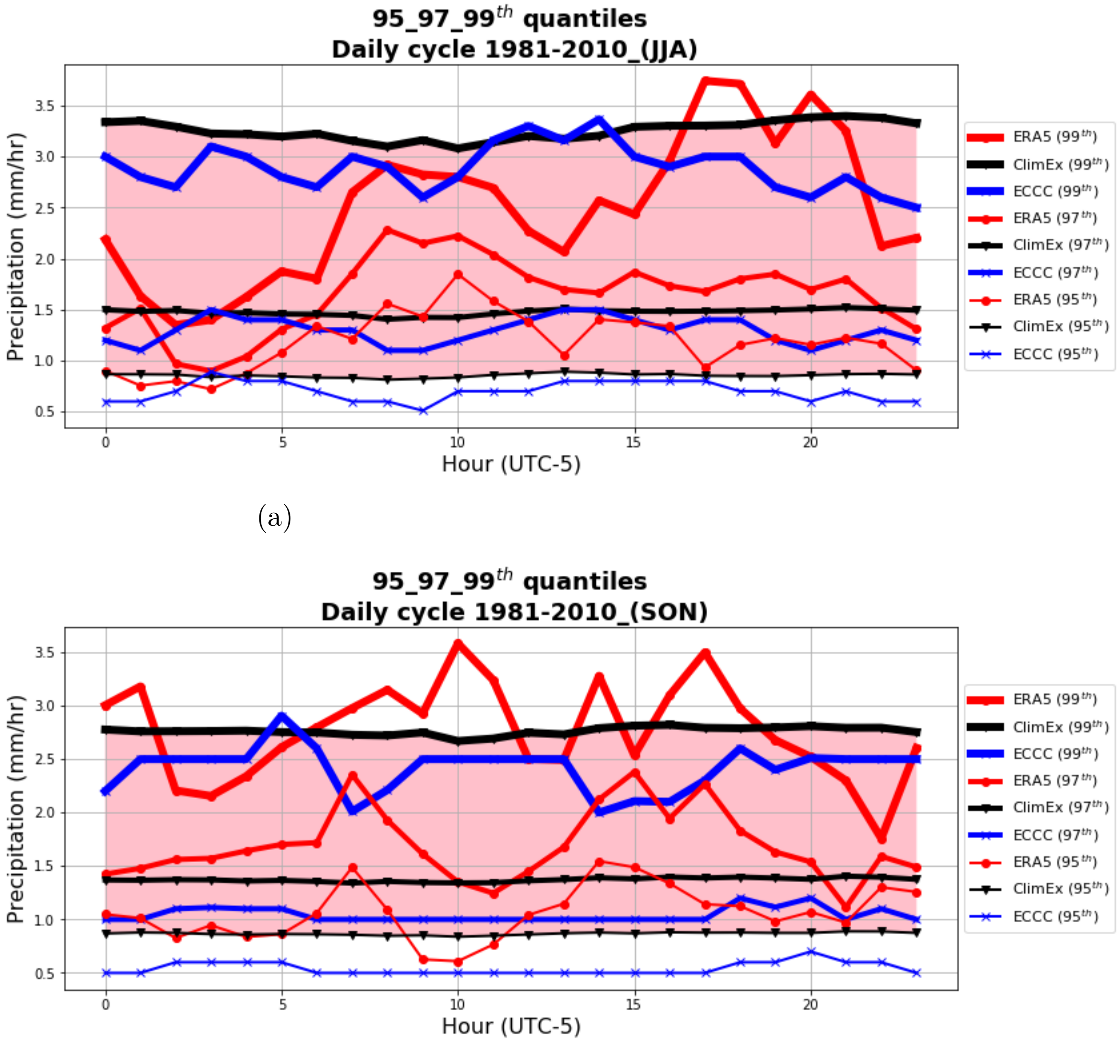


(a)



(b)

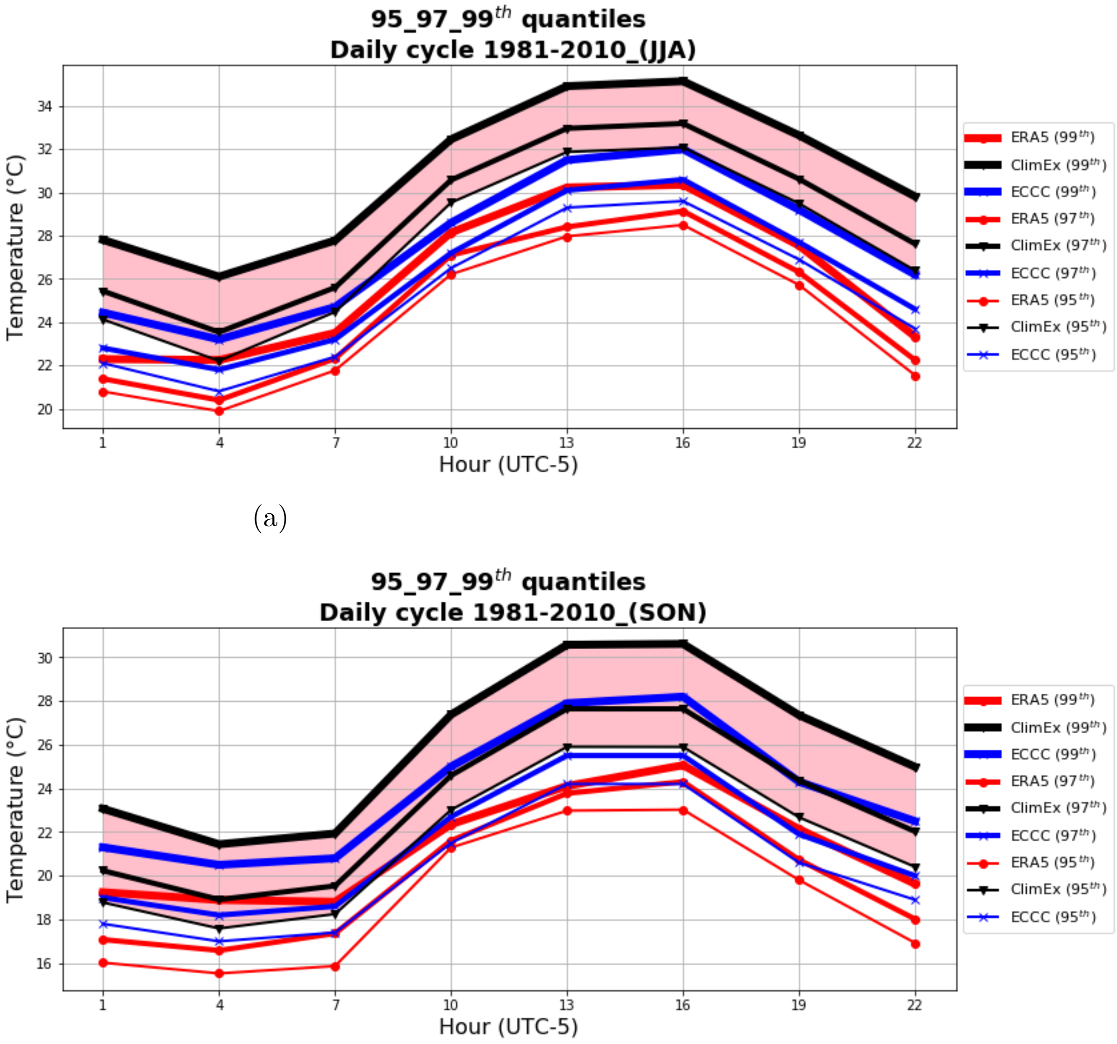
FIGURE 1.13 – Quantiles 95,97,99th - Precipitation rate (mm/hr) - South stations. The red lines are ERA5, the blue lines are ECCC, the black lines are ClimEx (50). (a) Summer (b) Fall.



(a)

(b)

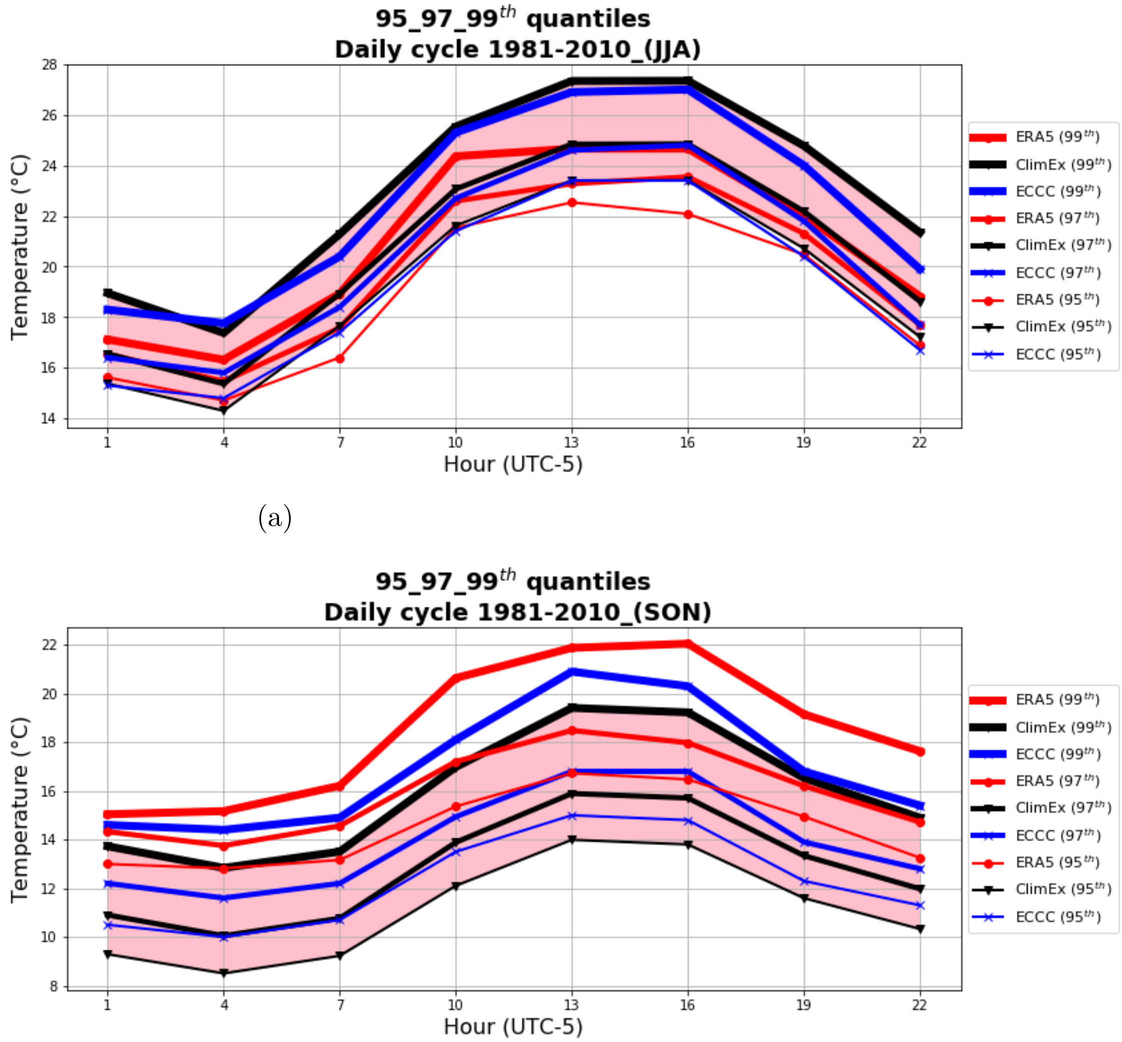
FIGURE 1.14 – Same as figure 1.13 but for North stations.



(a)

(b)

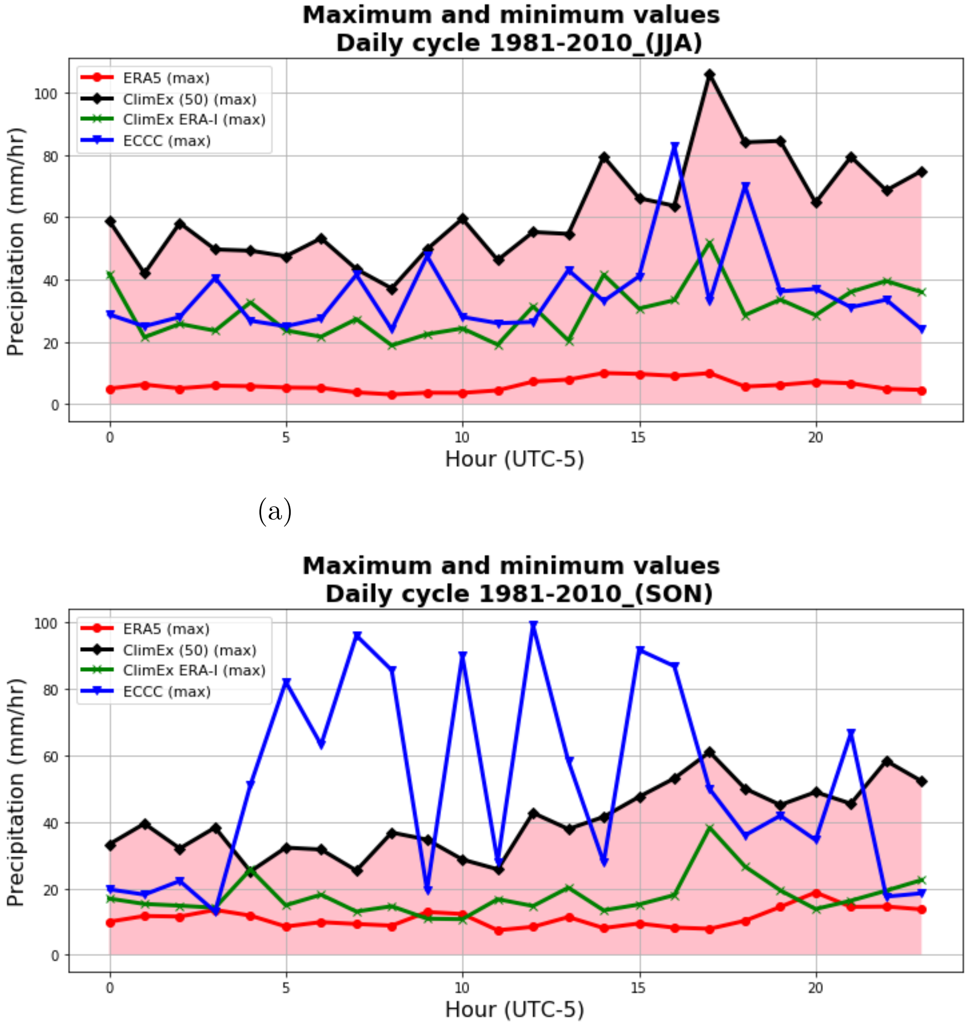
FIGURE 1.15 – Quantiles 95,97,99th - Temperature (°C) - South stations. The red lines are ERA5, the blue lines are ECCC, the black lines are ClimEx (50). (a) Summer (b) Fall.



(a)

(b)

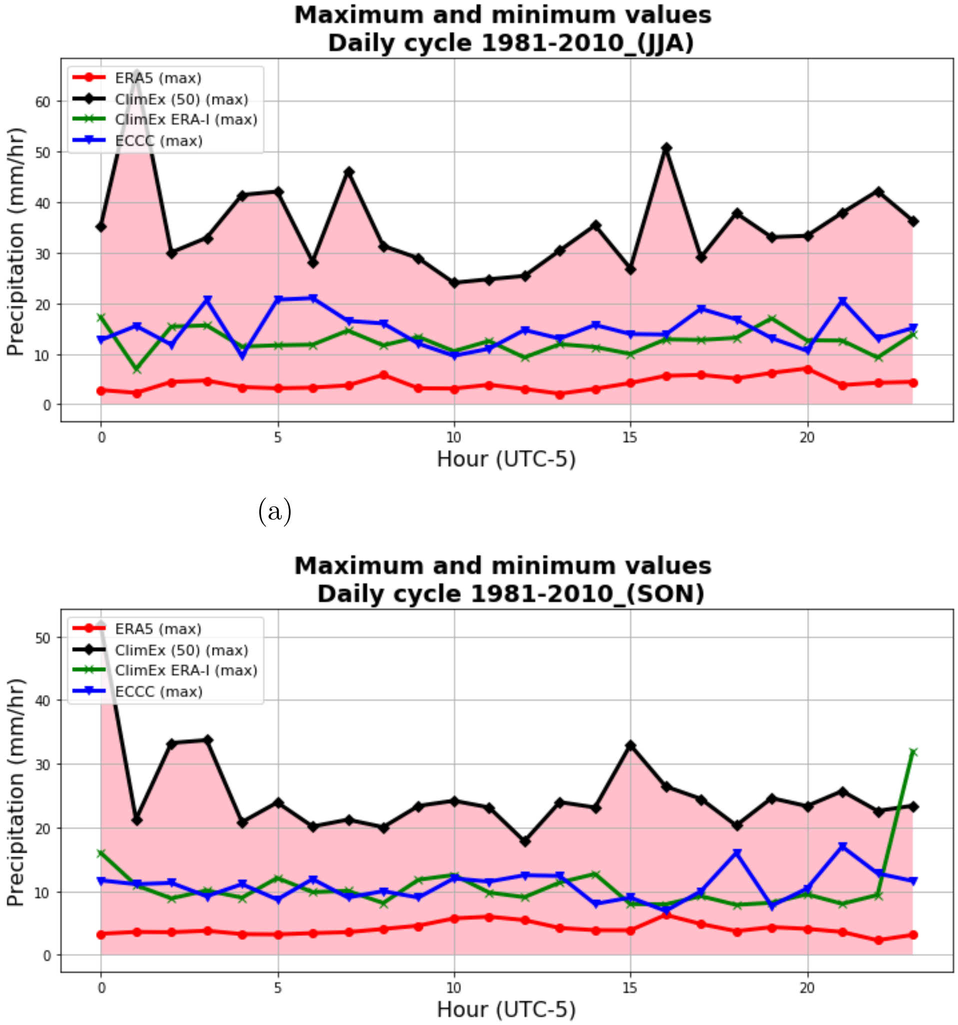
FIGURE 1.16 – Same as figure 1.15 but for North stations.



(a)

(b)

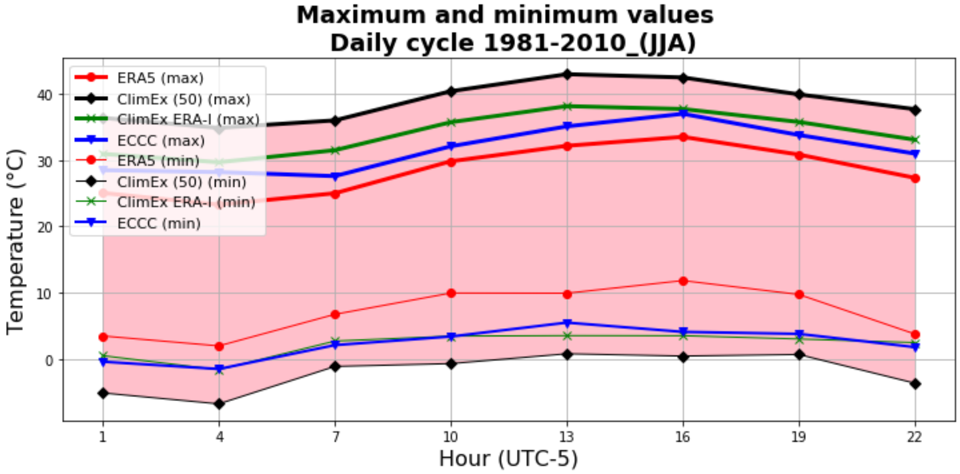
Figure 1.17 – Maximum values - Precipitation rate (mm/hr) - South stations. The red line is ERA5, the blue line is ECCC, the green line is ClimEx ERA-I, the black line is ClimEx (50). (a) Summer (b) Fall.



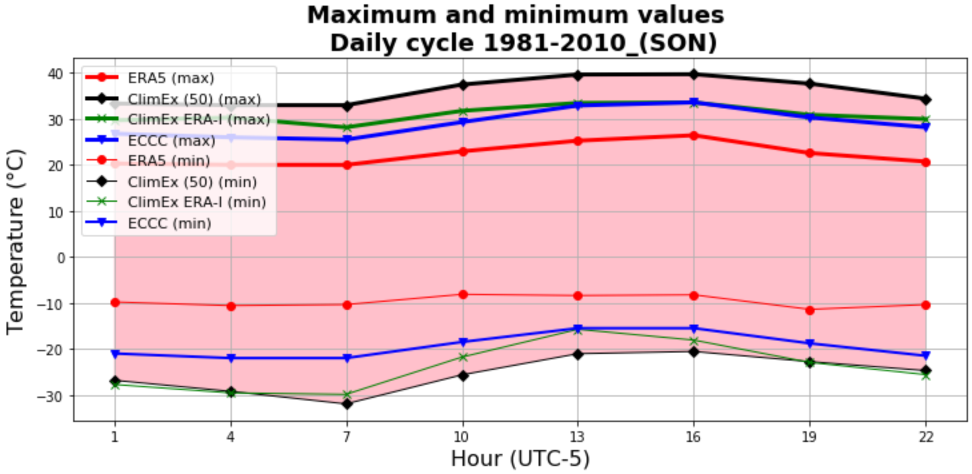
(a)

(b)

FIGURE 1.18 – Same as figure 1.17 but for North stations.

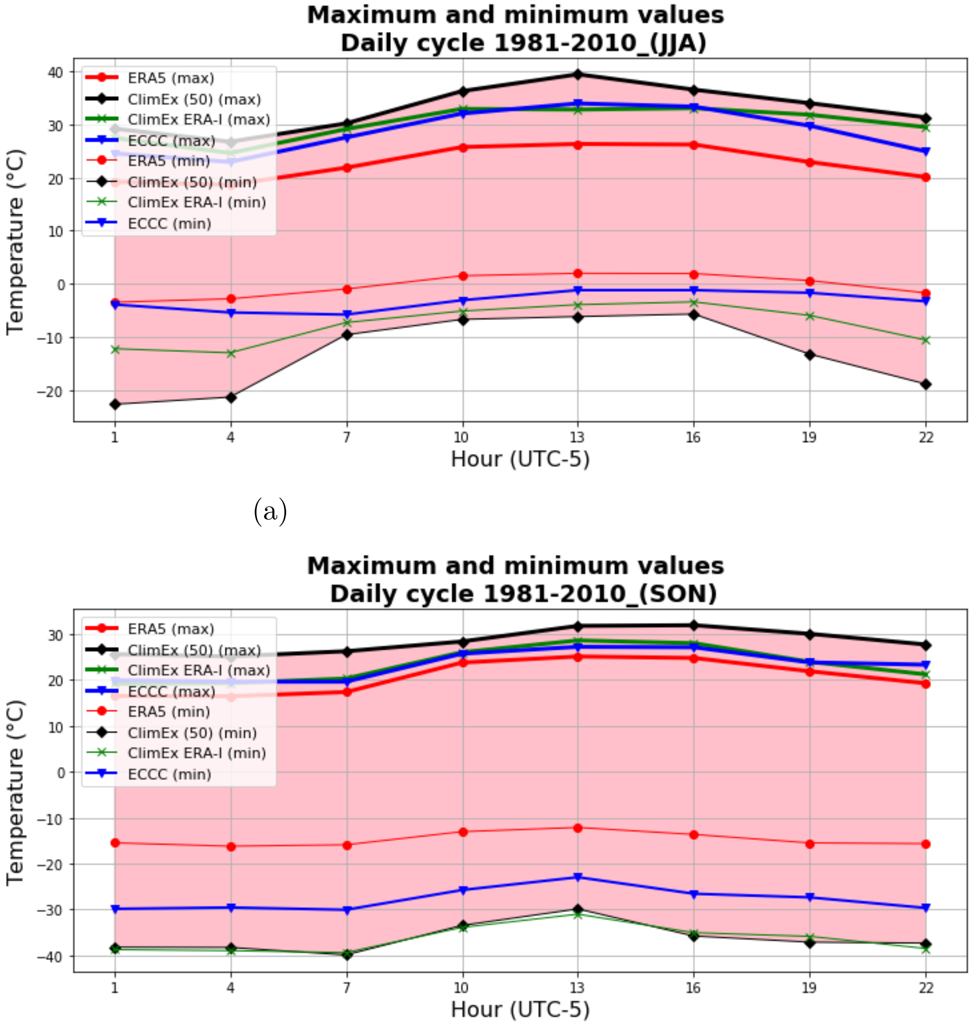


(a)



(b)

FIGURE 1.19 – Maximum and minimum values - Temperature (°C) - South stations. The red lines are ERA5, the blue lines are ECCC, the green lines are ClimEx ERA-I, the black lines are ClimEx (50). (a) Summer (b) Fall.



(a)

(b)

FIGURE 1.20 – Same as figure 1.19 but for North stations.

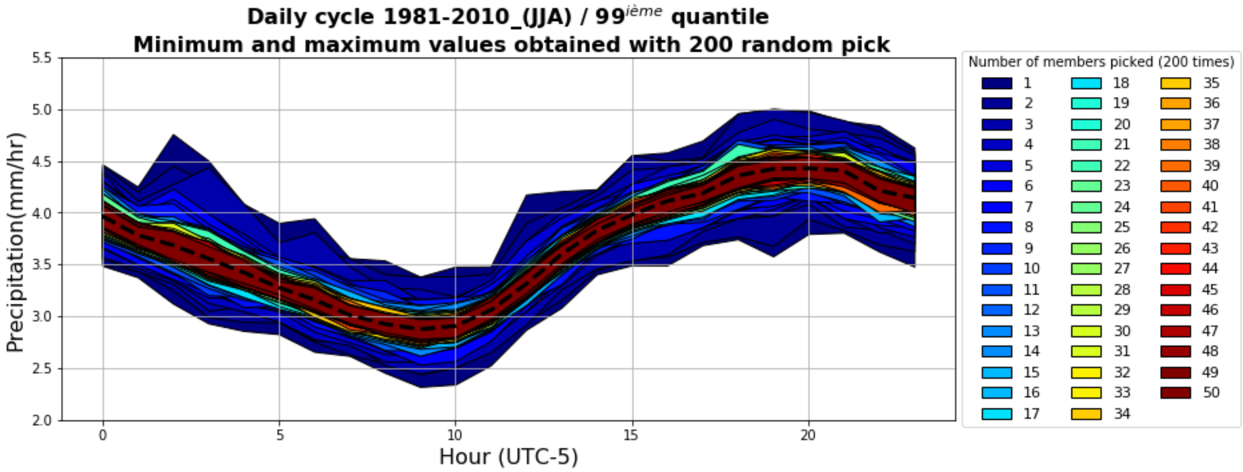


FIGURE 1.21 – Maximum and minimum values from random pick - 99 th quantile, precipitation rate (mm/hr)(JJA) - South stations. The dash line represent the value of the ClimEx ensemble and the color the number of members picked 200 times.

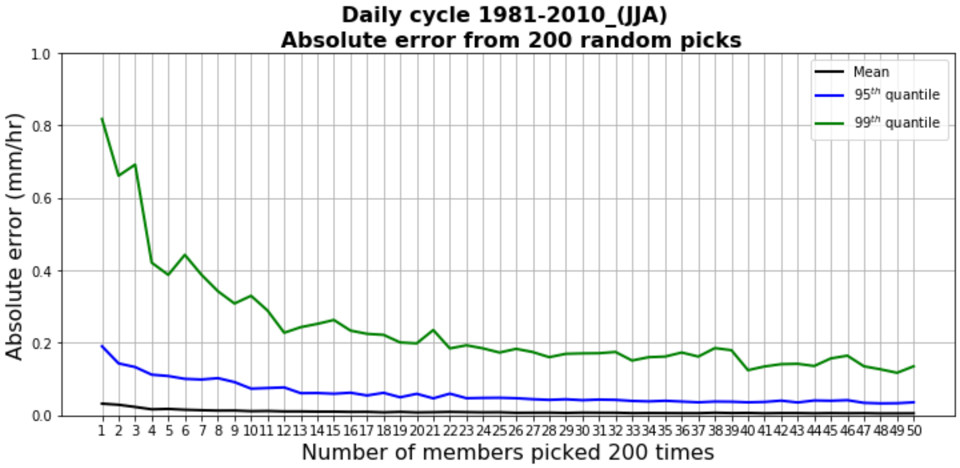
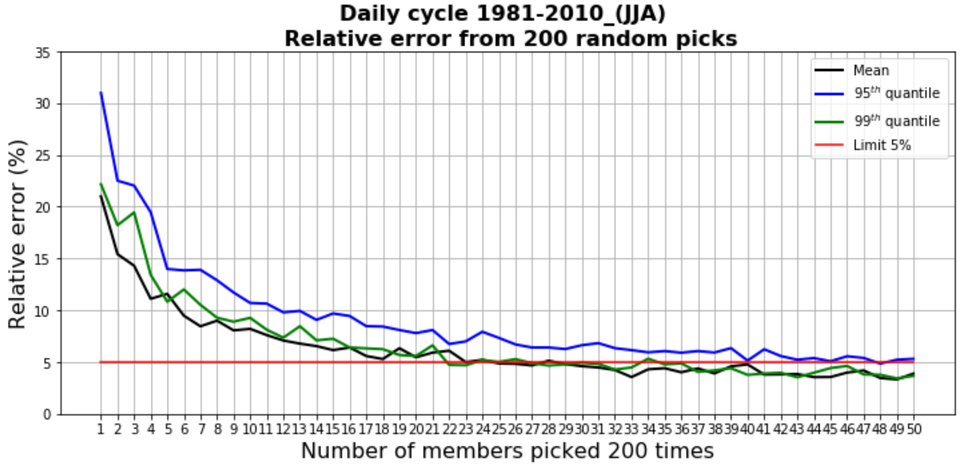
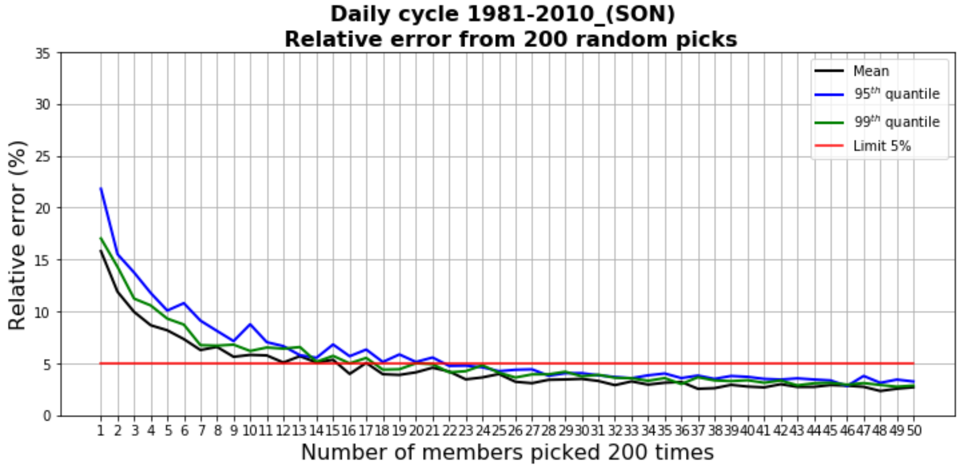


FIGURE 1.22 – Absolute error from random pick - Daily cycle, precipitation rate (mm/hr) (JJA) - South stations.

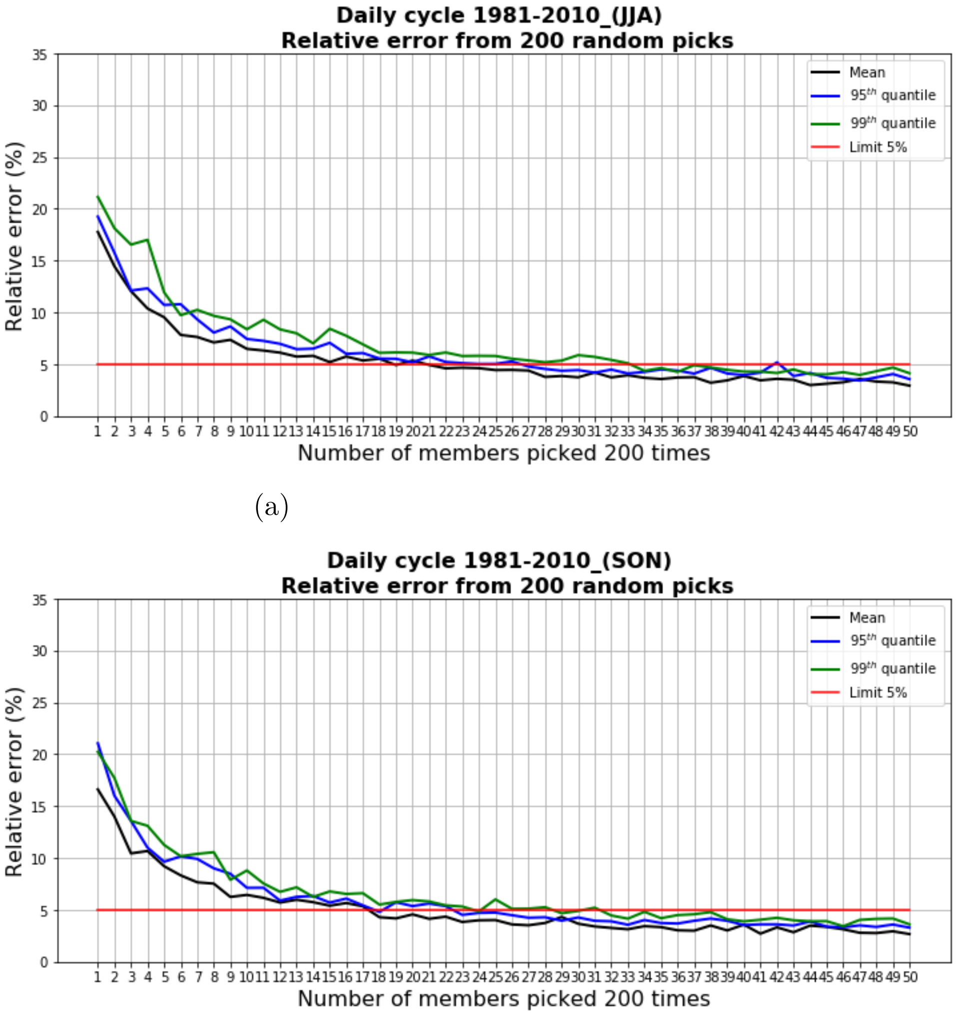


(a)



(b)

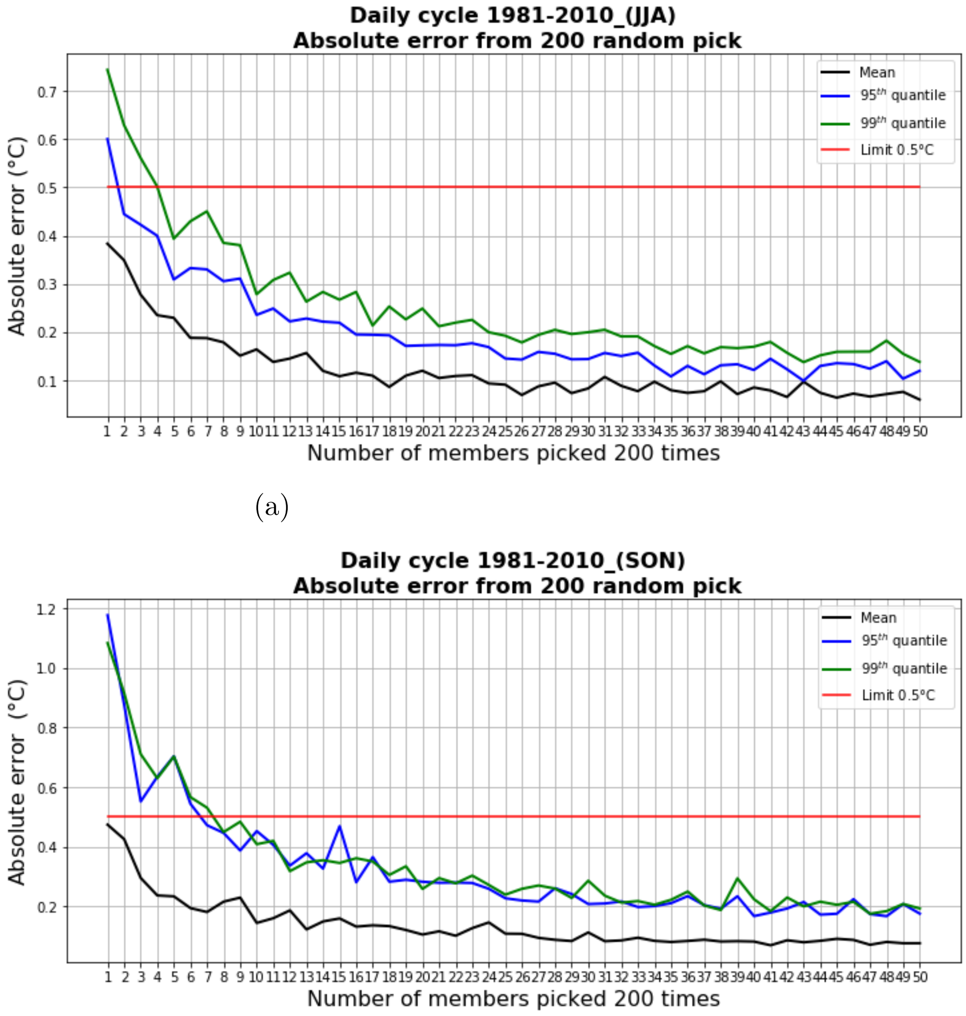
FIGURE 1.23 – Relative error from random pick - Daily cycle, precipitation (%) - South stations. (a) Summer (b) Fall.



(a)

(b)

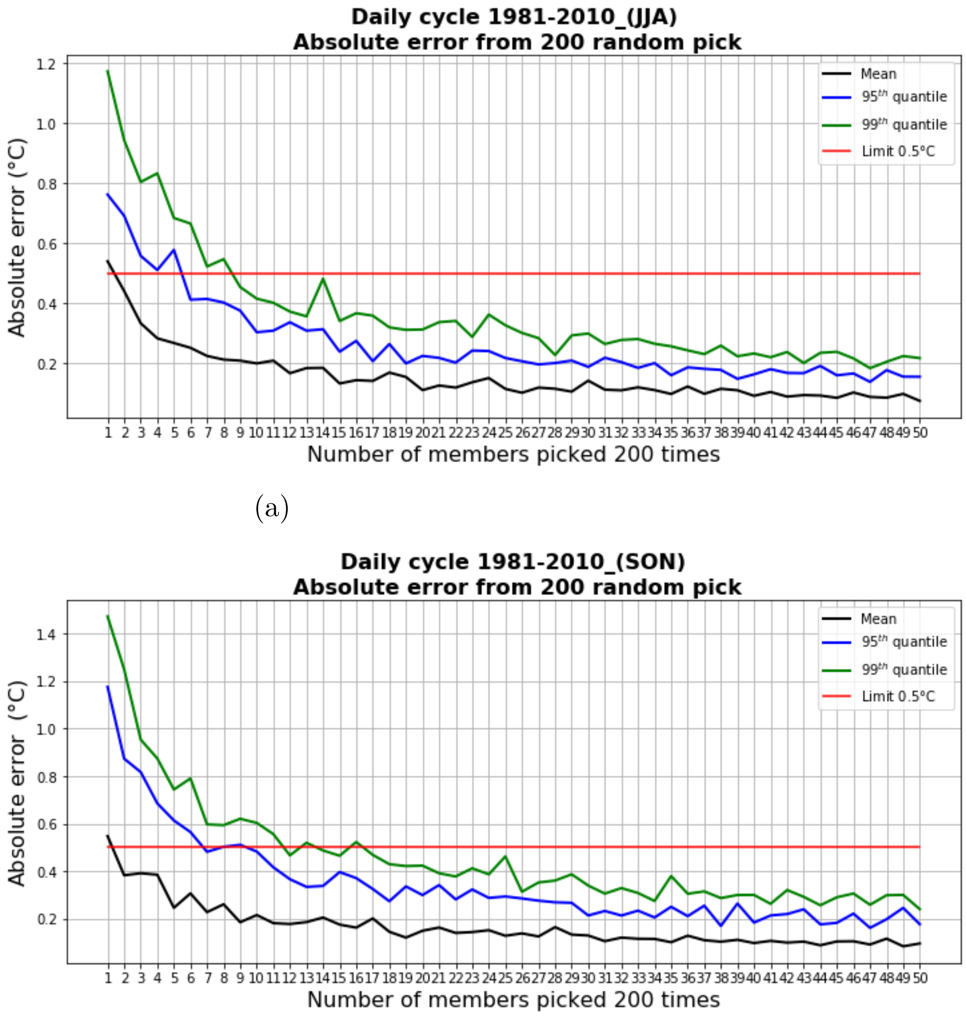
FIGURE 1.24 – Relative error from random pick - Daily cycle, precipitation (%) - North stations. (a) Summer (b) Fall.



(a)

(b)

FIGURE 1.25 – Absolute error from random pick - Daily cycle, temperature (°C) - South stations. (a) Summer (b) Fall.



Daily cycle 1981-2010_(JJA)
Absolute error from 200 random pick

(a)

Daily cycle 1981-2010_(SON)
Absolute error from 200 random pick

(b)

FIGURE 1.26 – Absolute error from random pick - Daily cycle, temperature (°C) - North stations. (a) Summer (b) Fall.

CONCLUSION

Afin de mieux nous préparer à faire face aux changements climatiques, il est essentiel d'améliorer notre capacité à simuler et à prévoir avec précision les multiples aspects du cycle journalier. L'objectif de cette étude est d'évaluer le cycle journalier des simulations du grand ensemble ClimEx au niveau de ses valeurs moyennes et extrêmes. Notre adaptation implique aussi de connaître les impacts des changements climatiques sur le réseau hydrique. L'étude actuelle se veut une contribution au développement de simulations de modèles hydrologiques à l'échelle sous-journalière en contexte de changment climatique.

Dans le cadre de cette étude, le cycle journalier a été calculé sur 30 ans pour les saisons d'été et d'automne sur deux zones géographiques, situées dans les parties nord et sud de la province canadienne du Québec. Le cycle journalier moyen et ses valeurs extrêmes ont été comparés à trois jeux de données de référence, les stations d'ECCC, la réanalyse ERA5 et une simulation du MRCC5 pilotée par la réanalyse ERA-Interim (ClimEx ERA-I). Comme l'objectif de cette étude était de comparer ClimEx avec les données des stations et comme les données des stations d'ECCC sont peu nombreuses en plus d'être réparties de manière non uniforme sur le territoire québécois, l'étude ne peut être représentative de l'ensemble du domaine de simulation de ClimEx. De plus, il est parfois difficile de tirer des conclusions solides étant donné l'incertitude des données des stations et la difficulté de juger des valeurs aberrantes. Il y a aussi le problème lié aux échelles spatiales impliquées, non seulement puisque ClimEx et ERA5 utilisent des grilles de résolutions différentes, mais aussi parce que les stations représentent des estimations ponctuelles (c.-à-d. sans dimension).

Les résultats pour le cycle journalier des précipitations montrent une corrélation essentiellement non significative avec les données d'ECCC. L'amplitude du cycle est similaire par rapport aux observations des stations d'ECCC. Pour la température, le cycle journalier montre une forte corrélation avec les données de référence et l'amplitude des cycles est très similaire à celles données d'ECCC.

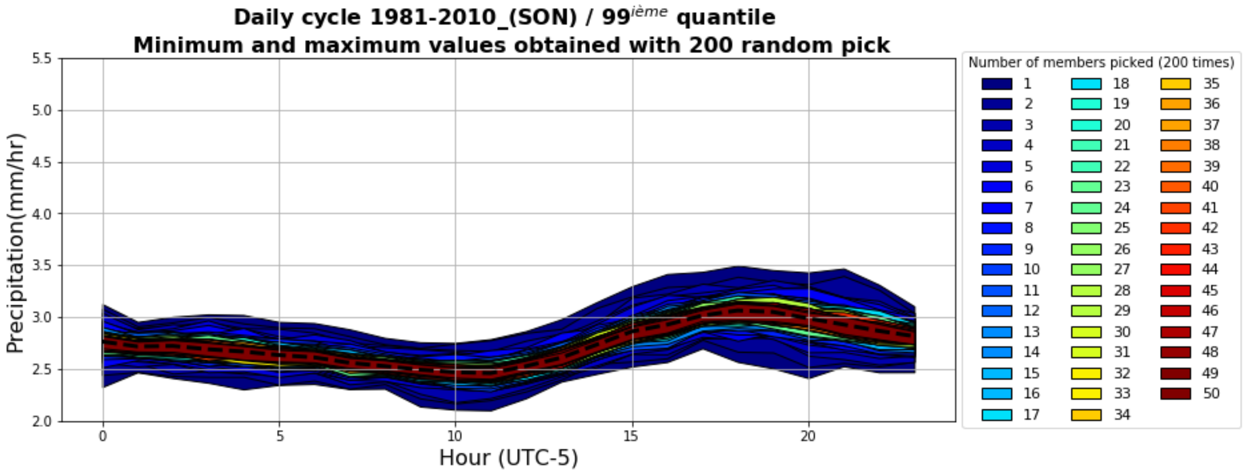
Un des objectifs était d'évaluer la représentativité des valeurs extrêmes, telles que simulées dans l'ensemble ClimEx. Cette analyse implique la comparaison des quantiles 95, 97 et 99$^{\text{ième}}$ entre ClimEx (50), ERA5 et ECCC. Pour la précipitation, la distribution des quantiles est très similaire à celles des produits de référence, mais avec quelques exceptions. Pour la température, la distribution des quantiles est également très similaire malgré un biais chaud sur la partie sud du Québec. Quant aux valeurs maximales et minimales produites par ClimEx, elles sont jusqu'à 10 fois supérieures à celles d'ERA5 pour les précipitations et montrent jusqu'à environ 20°C d'écart pour la température, mais l'écart est moins grand avec les données d'ECCC.

Cette étude porte également sur l'évaluation du seuil minimal quant au nombre de membres requis pour évaluer différentes statistiques pour une marge d'erreur statistique donnée. Cela peut être considéré comme une autre façon d'évaluer l'effet de la variabilité naturelle sur différentes statistiques (cycle moyen, 95$^{\text{ième}}$ quantile, 99$^{\text{ième}}$ quantile). Les résultats ont démontré que les précipitations nécessitent de 15 à 50 membres selon la région et la saison pour obtenir une erreur relative inférieure à 5% par rapport à l'estimation de référence obtenue de l'ensemble des 50 membres. Pour la température, l'utilisation de 1 à 17 membres est suffisante pour estimer les quantiles 95 et 99$^{\text{ième}}$ avec une erreur inférieure à 0,5°C. Pour le cycle moyen de la température, de 1 à 2 membres seulement sont nécessaires. Comme prévu, les précipitations sont plus sensibles au nombre de membres que la température, en raison de l'effet de la variabilité climatique naturelle qui diffère selon les variables. Il convient de noter que les résultats de la procédure de pige aléatoire seraient légèrement différents si l'expérience était répétée.

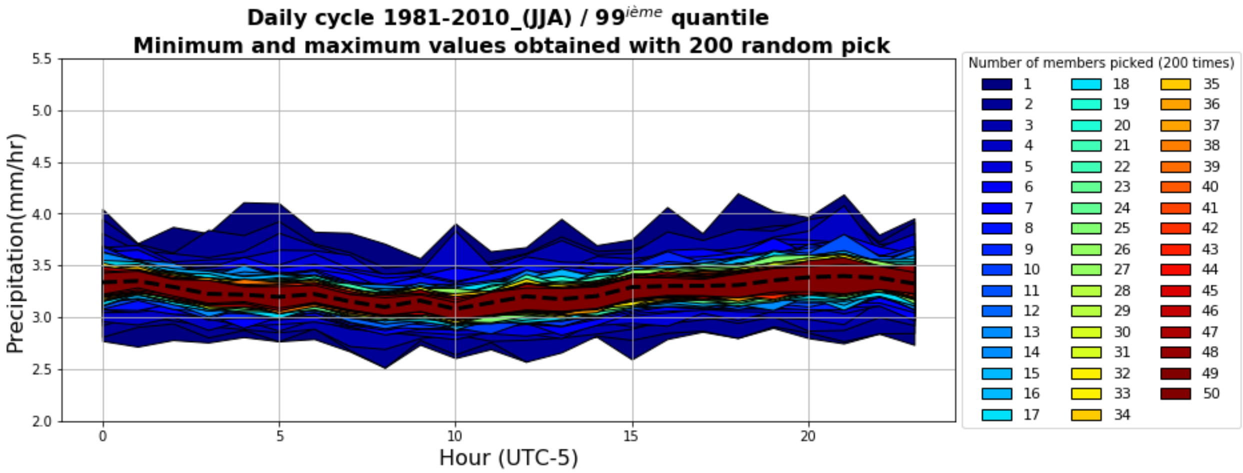
Comme travaux futurs, il serait intéressant de poursuivre cette étude en comparant ClimEx avec d'autres jeux de données comme le Canadian Precipitation Analysis (CaPA) qui montre de bonnes performances dans de nombreuses régions du Québec pour la plupart des bassins versants selon l'étude de Bajamgnigni Gbambie *et al.* (2017). Une autre option intéressante serait d'évaluer le cycle journalier pour d'autres variables d'intérêt en hydrologie (par exemple le vent et le rayonnement solaire). Il serait important pour la communauté scientifique de poursuivre le développement des méthodes d'analyse des grands ensembles. Tel que démontré dans cette étude, les grands ensembles fournissent des estimés avec une plus grande précision statistique en comparaison avec une seule réalisation ou trajectoire du sytème climatique. Cela ne signifie pas pour autant que les simulations sont de meilleure qualité que la réalité elle-même, puisqu'il est également possible d'être très précis mais loin de la réalité. L'avantage de poursuivre les travaux sur ces grands ensembles serait de permettre de mieux définir les limites quant à leur utilisation et leur évaluation. Alors que les grands ensembles à haute résolution comme ClimEx impliquent une énorme quantité de données, le coût informatique pour le traitement de ces données peut limiter l'application de certaines techniques d'analyse approfondies des simulations. Par contre, les grands ensembles offrent aussi à la communauté scientifique sur le climat un large éventail de possibilités et d'innovations en termes de méthodes d'analyse pour l'évaluation des modèles de climat, la caractérisation de la variabilité naturelle du climat et de l'étude des phénomènes extrêmes.

APPENDICE A

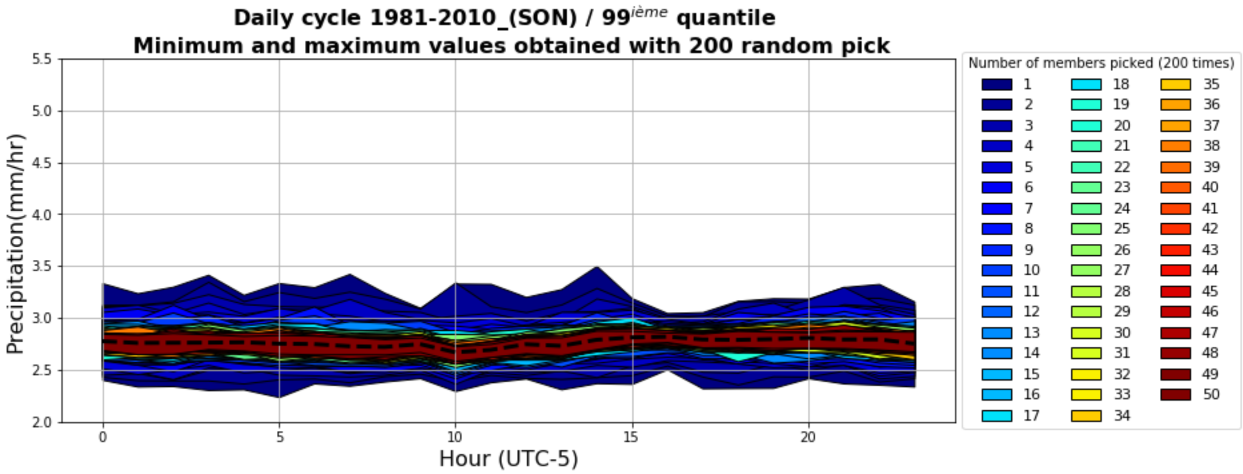
FIGURES COMPLÉMENTAIRES



(a) Summer, South stations

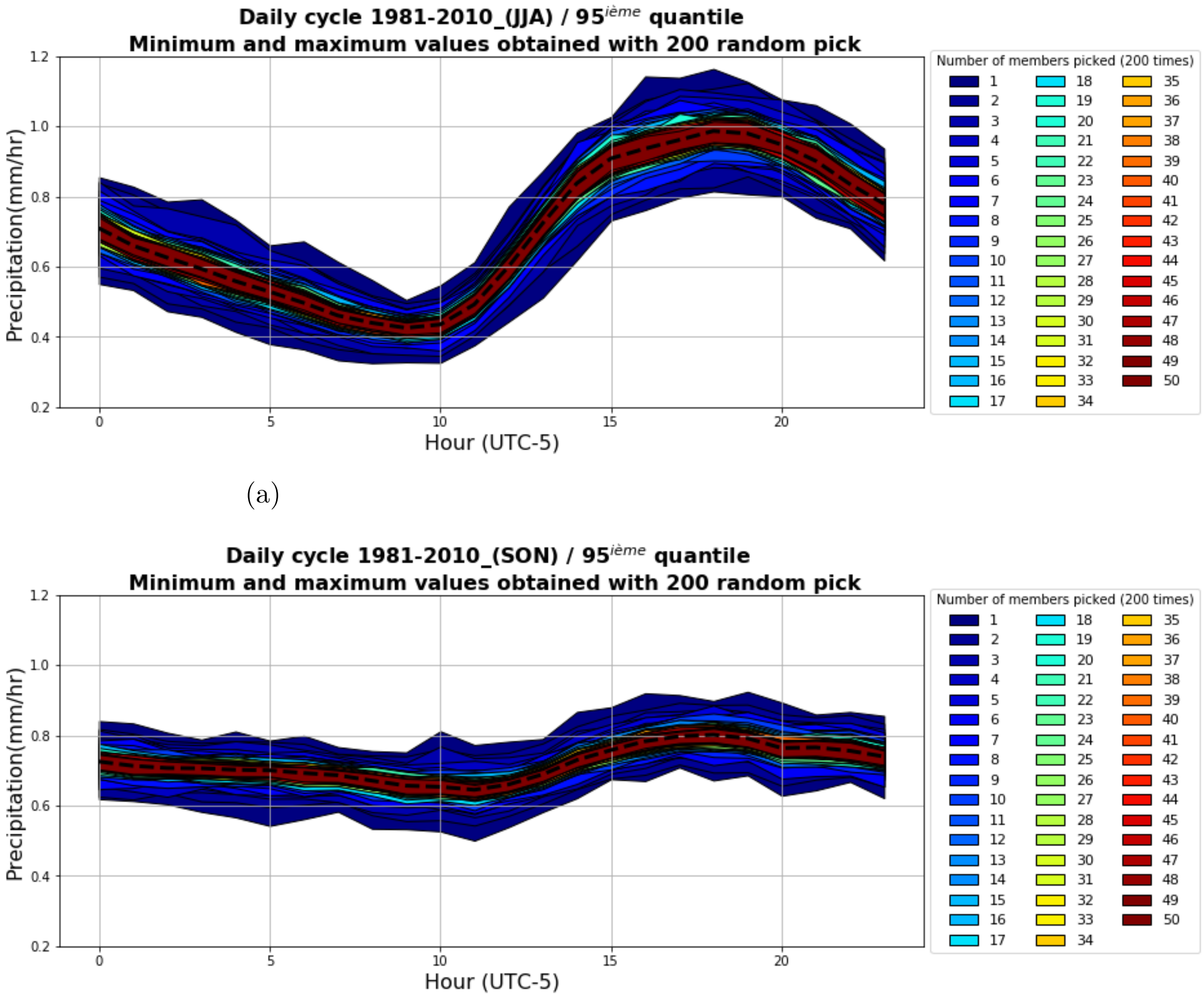


(b) Summer, North stations.



(c) Fall, North stations.

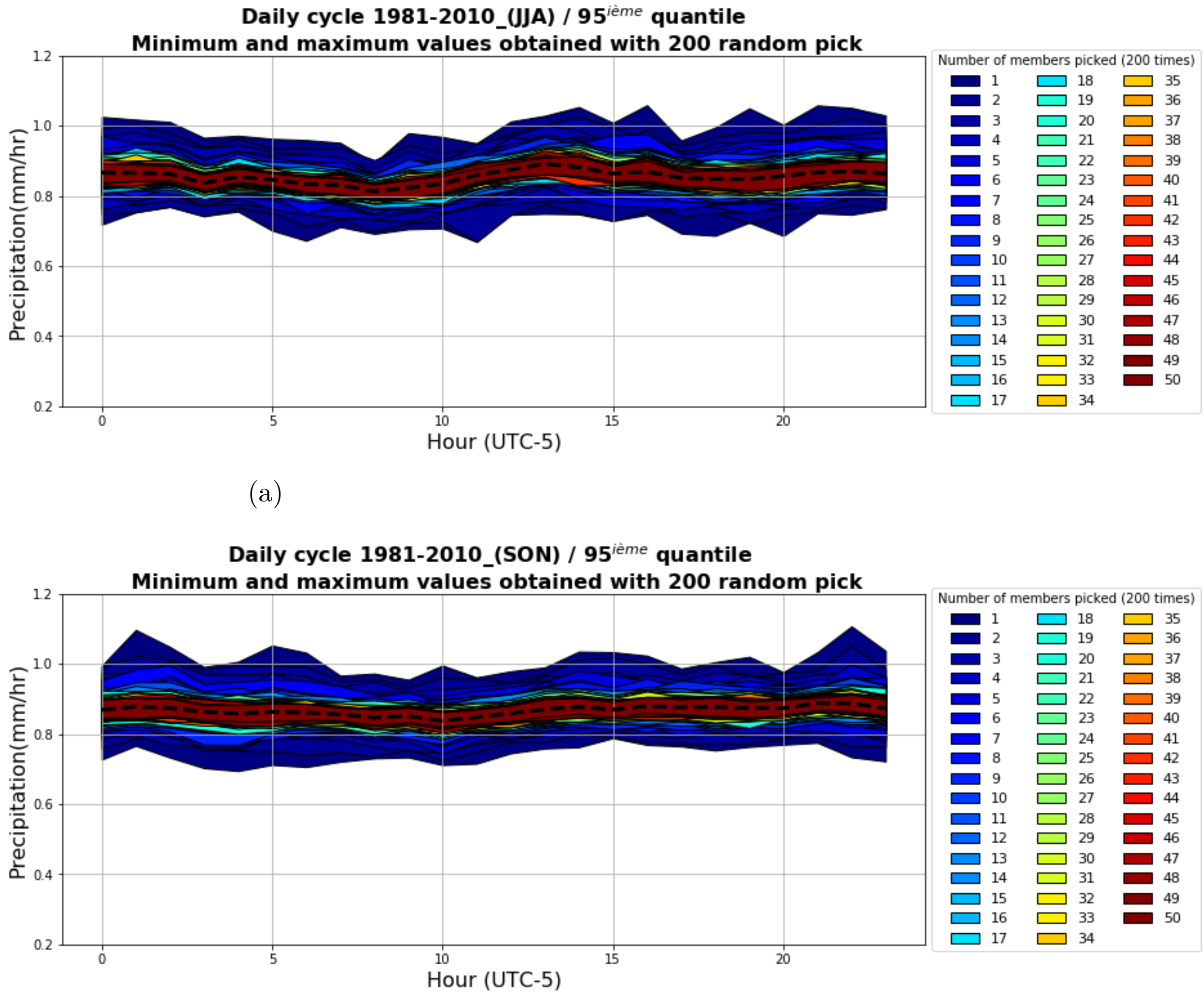
FIGURE A.1 – Maximum and minimum values from random pick - 99 $^{\text{th}}$ quantile, precipitation rate (mm/hr). The dash line represent the value of the ClimEx ensemble and the color the number of members picked 200 times.



(a)

(b)

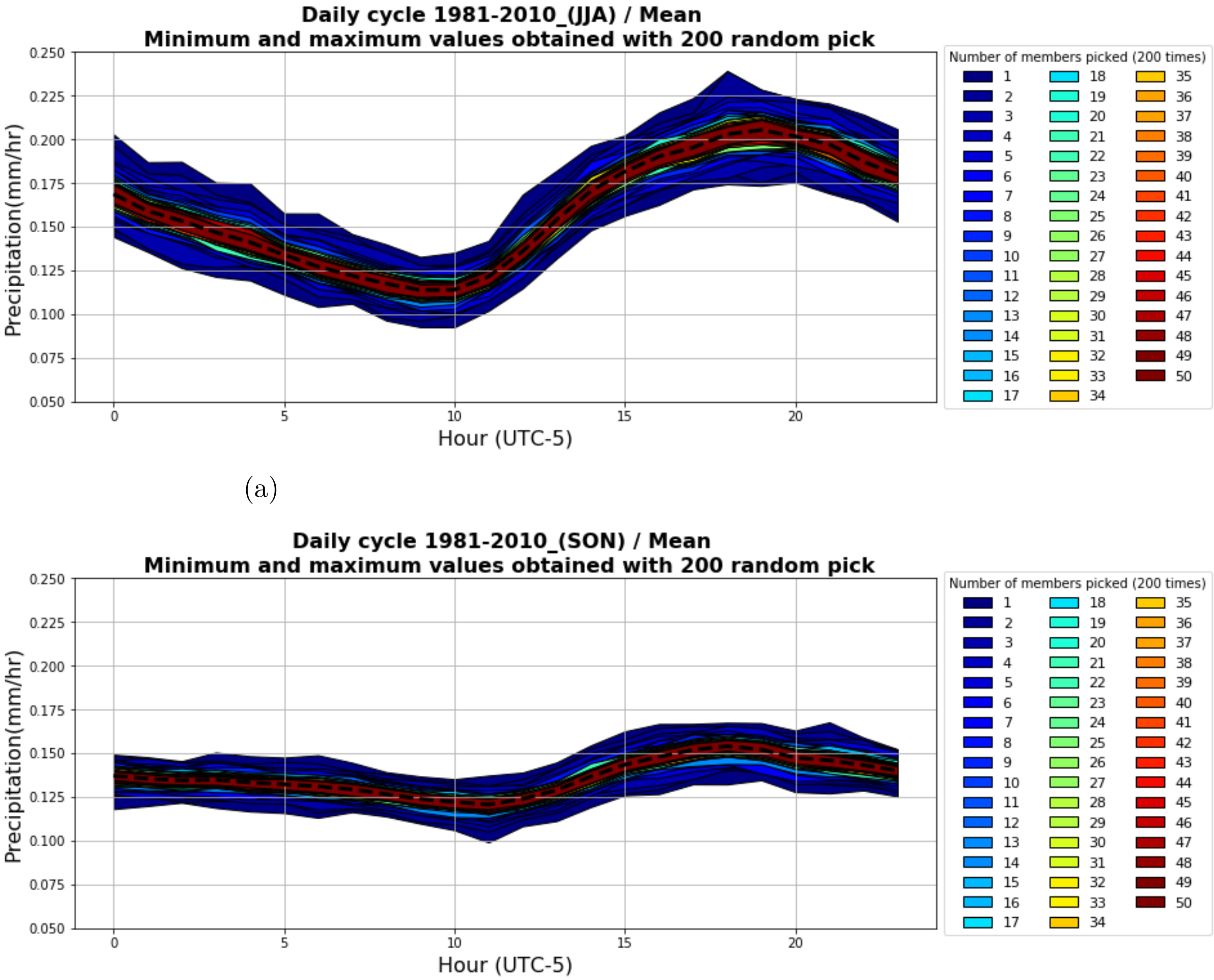
Figure A.2 – Maximum and minimum values from random pick - 95 th quantile, precipitation rate (mm/hr), South stations. The dash line represent the value of the ClimEx ensemble and the color the number of members picked 200 times. (a) Summer (b) Fall.



(a)

(b)

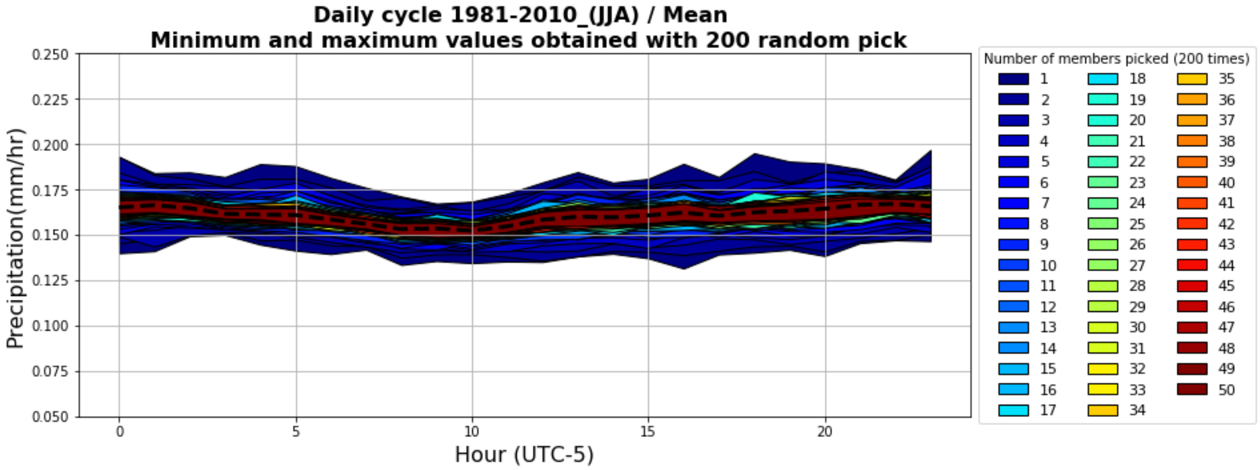
Figure A.3 – Same as A.2 but for North stations.



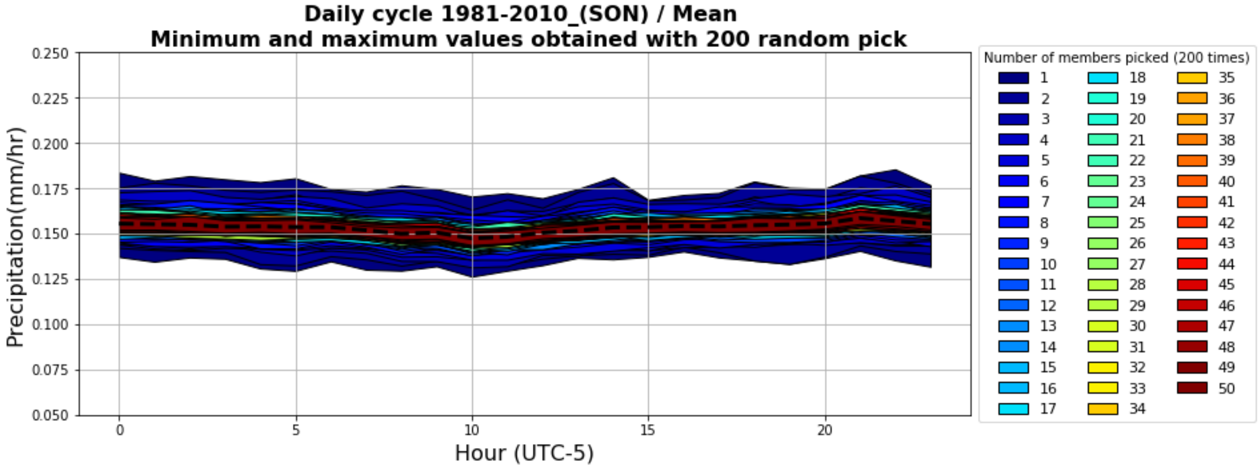
(a)

(b)

FIGURE A.4 – Maximum and minimum values from random pick - Daily cycle, precipitation rate (mm/hr), South stations. The dash line represent the value of the ClimEx ensemble and the color the number of members picked 200 times. (a) Summer (b) Fall.

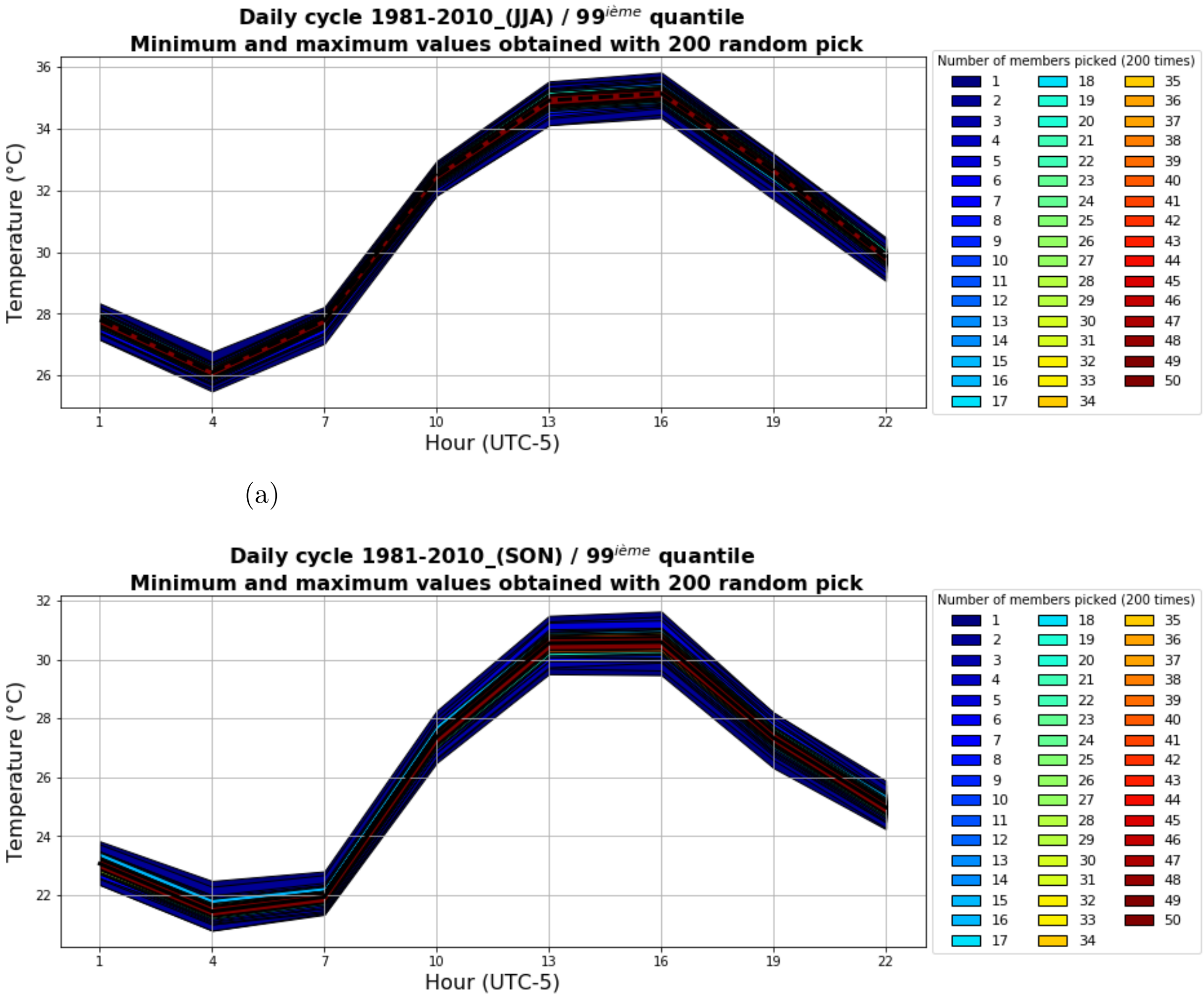


(a)



(b)

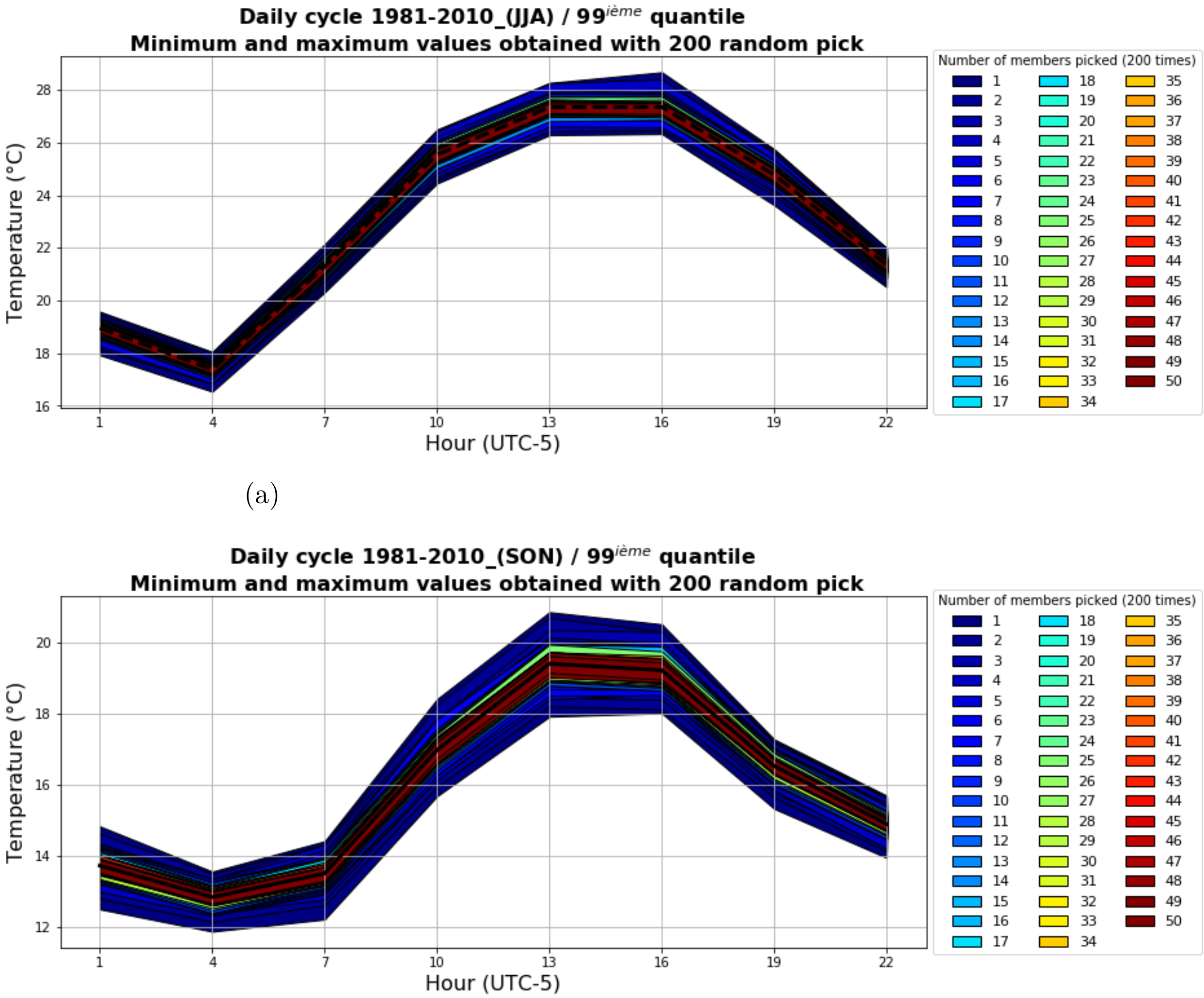
Figure A.5 – Same as A.4 but for North stations.



(a)

(b)

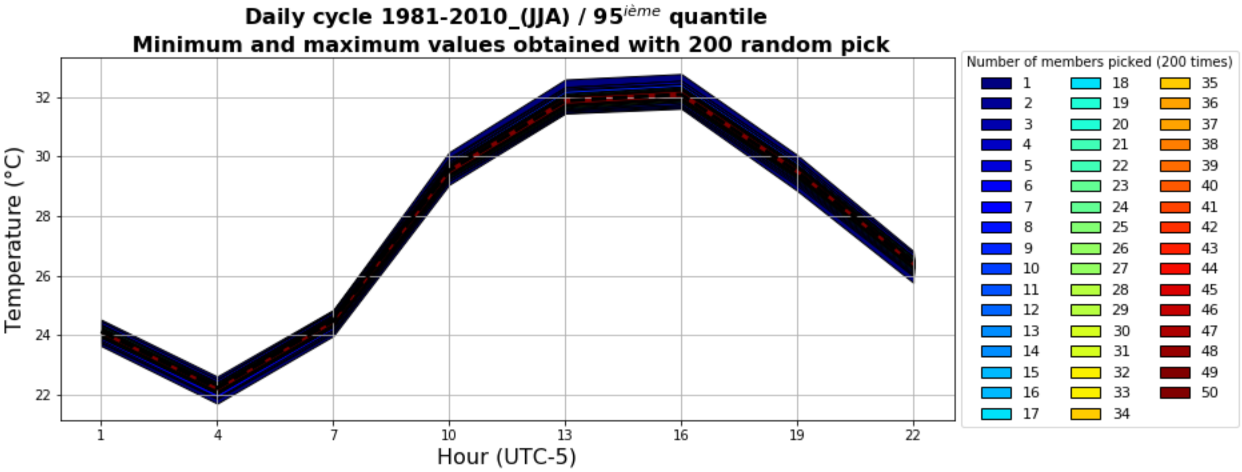
FIGURE A.6 – Maximum and minimum values from random pick - 99th quantile, temperature (°C), South stations. The dash line represent the value of the ClimEx ensemble and the color the number of members picked 200 times. (a) Summer (b) Fall.



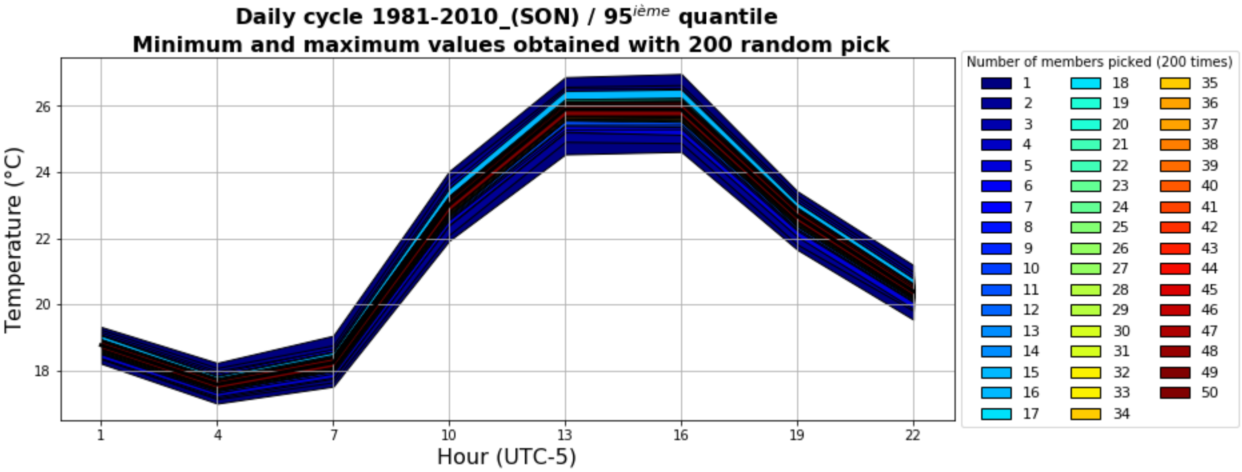
(a)

(b)

FIGURE A.7 – Same as A.6 but for North stations.

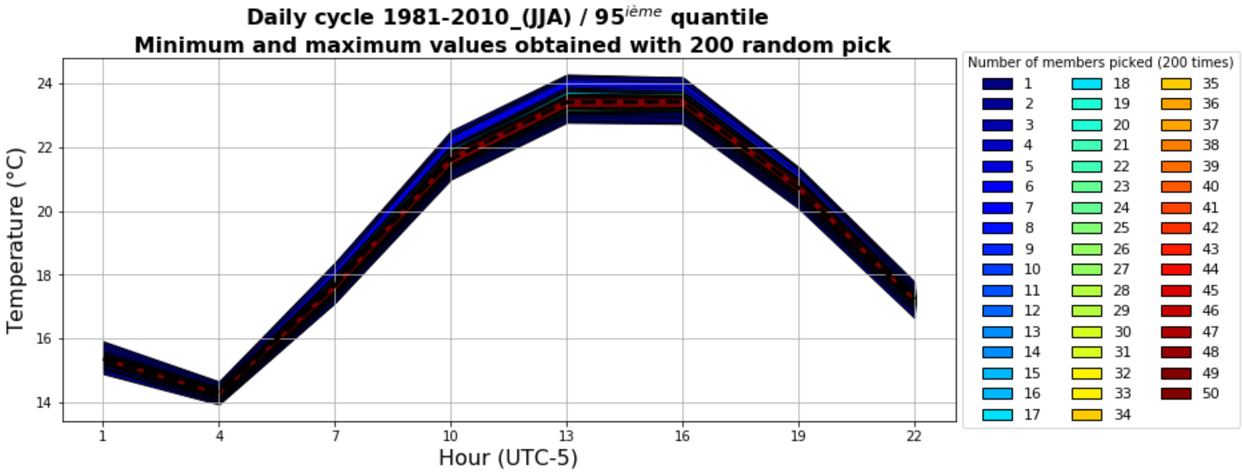


(a)

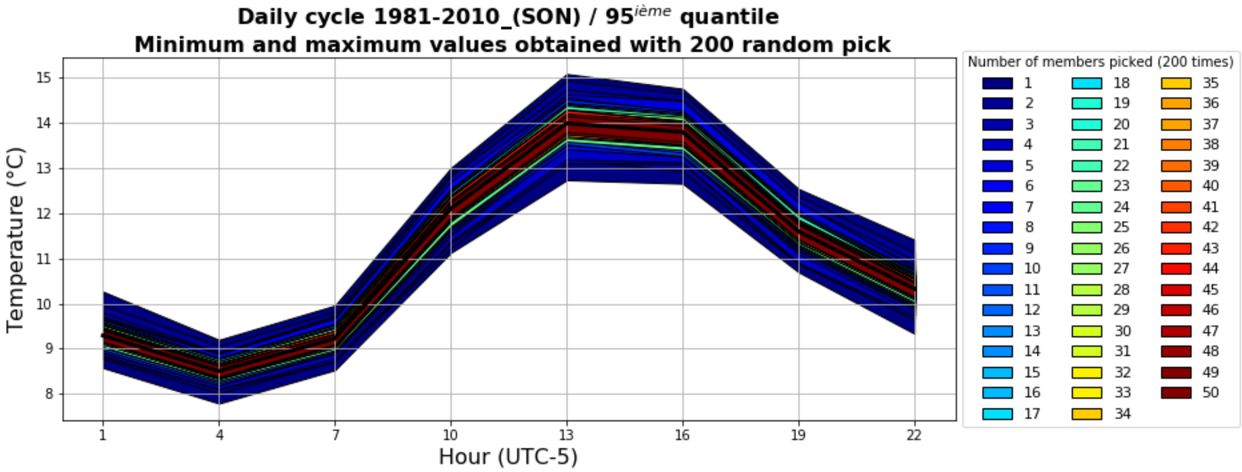


(b)

Figure A.8 – Maximum and minimum values from random pick - 95th quantile, temperature (°C), South stations. The dash line represent the value of the ClimEx ensemble and the color the number of members picked 200 times. (a) Summer (b) Fall.

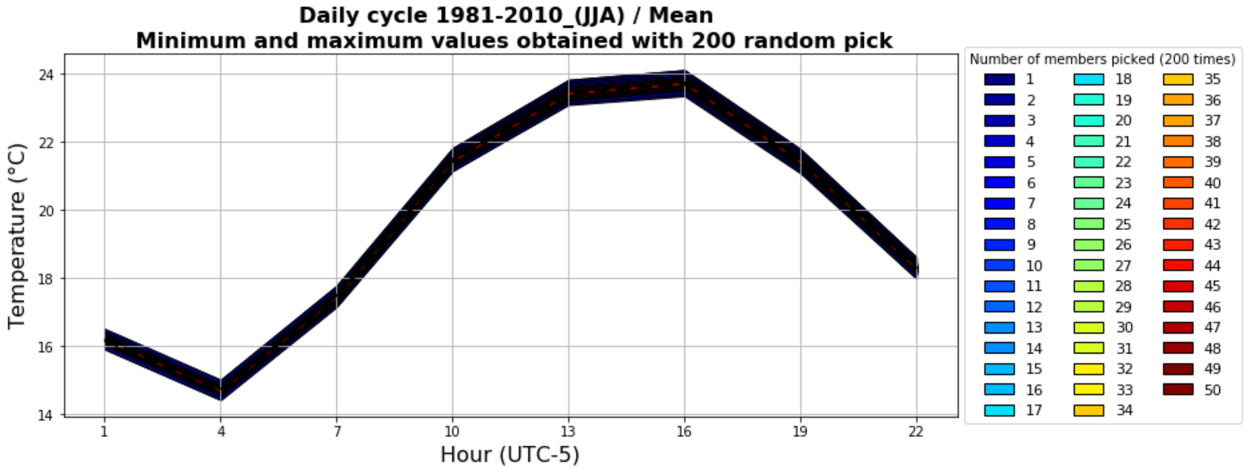


(a)

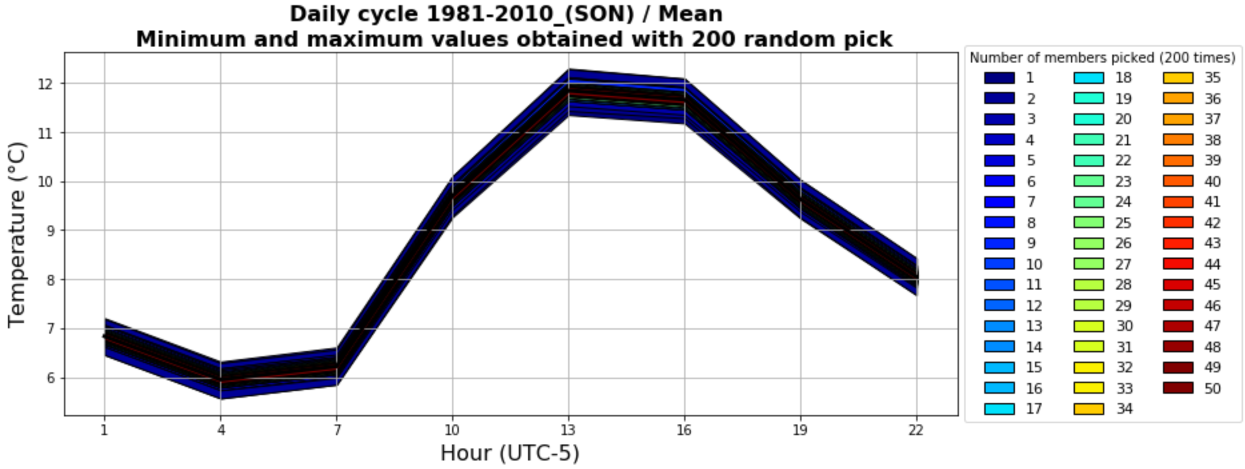


(b)

Figure A.9 – Same as A.8 but for North stations.

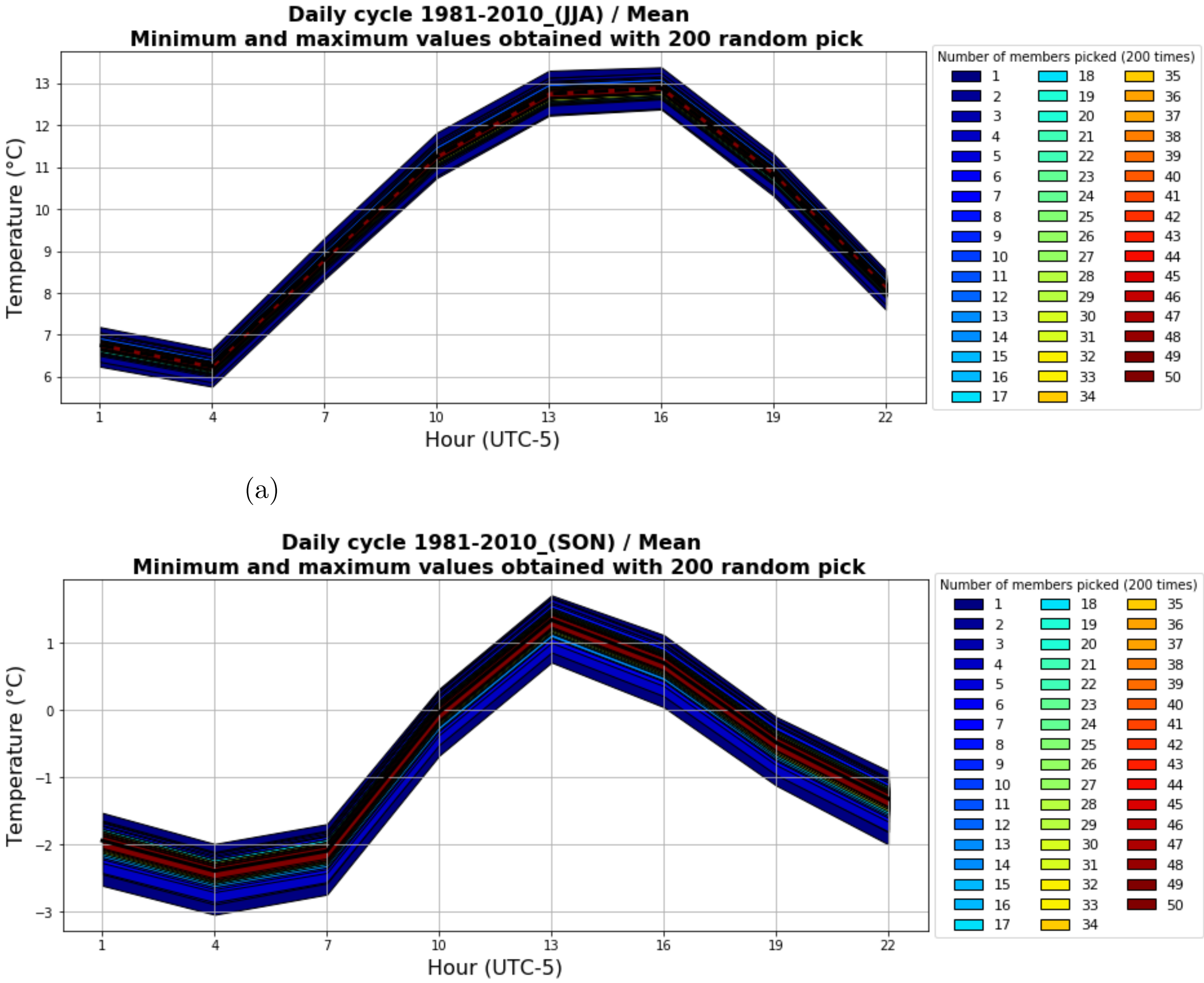


(a)



(b)

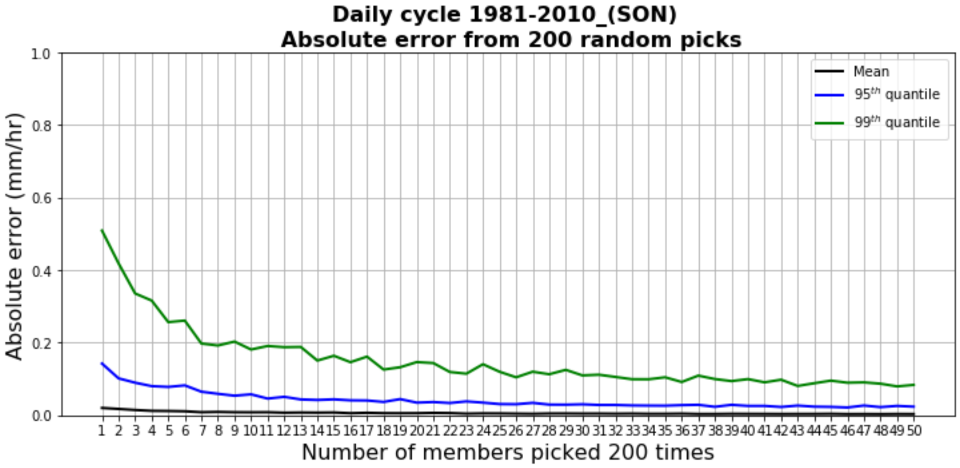
FIGURE A.10 – Maximum and minimum values from random pick - Daily cycle, temperature (°C), South stations. The dash line represent the value of the ClimEx ensemble and the color the number of members picked 200 times. (a) Summer (b) Fall.



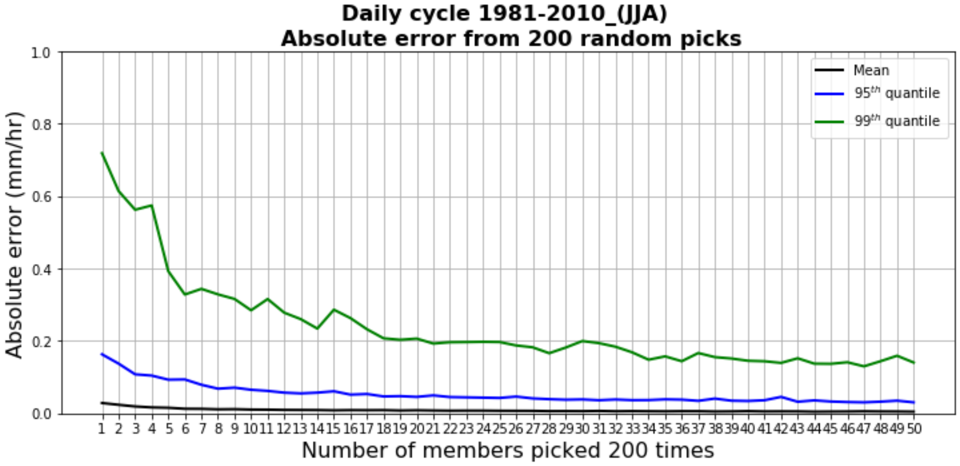
(a)

(b)

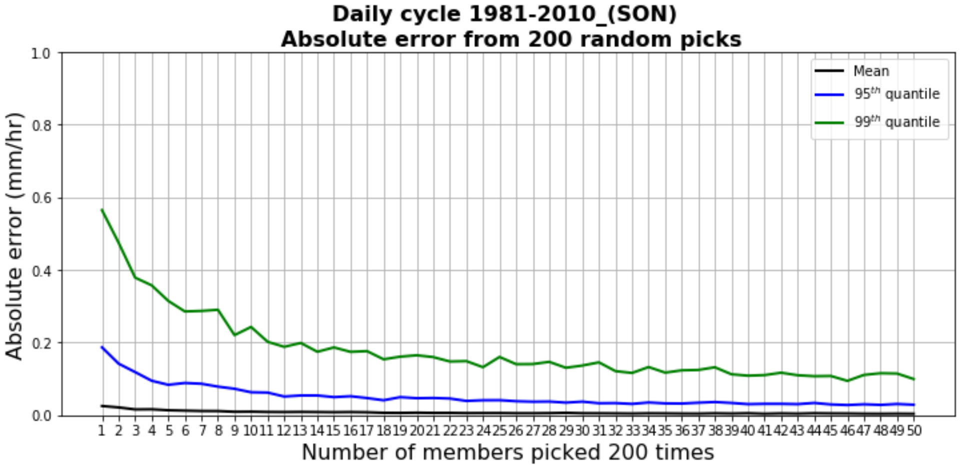
Figure A.11 – Same as A.10 but for North stations.



(a) Fall, South stations.



(b) Summer, North stations.



(c) Fall, North stations.

FIGURE A.12 – Absolute error from random pick - Daily cycle, precipitation rate (mm/hr).

RÉFÉRENCES

Azarnia, S. (2017). *Évaluation hydrologique de jeu de données de précipitations et températures journalières issues du modèle régional de climat" CROQ".* (Thesis).

Bajamgnigni Gbambie, A. S., Poulin, A., Boucher, M.-A. et Arsenault, R. (2017). Added value of alternative information in interpolated precipitation datasets for hydrology. *Journal of Hydrometeorology*, *18*(1), 247–264.

Bevelhimer, M. S., McManamay, R. A. et O'connor, B. (2015). Characterizing sub-daily flow regimes : implications of hydrologic resolution on ecohydrology studies. *River research and applications*, *31*(7), 867–879.

Cortés-Hernández, V. E., Zheng, F., Evans, J., Lambert, M., Sharma, A. et Westra, S. (2016). Evaluating regional climate models for simulating sub-daily rainfall extremes. *Climate Dynamics*, *47*(5-6), 1613–1628.

Deser, C., Knutti, R., Solomon, S. et Phillips, A. S. (2012). Communication of the role of natural variability in future north american climate. *Nature Climate Change*, *2*(11), 775–779.

Dessai, S., Hulme, M., Lempert, R. et Pielke Jr, R. (2009). Climate prediction : a limit to adaptation. *Adapting to climate change : thresholds, values, governance*, *64*, 78.

Giuntoli, I., Villarini, G., Prudhomme, C. et Hannah, D. M. (2018). Uncertainties in projected runoff over the conterminous united states. *Climatic Change*, *150*(3-4), 149–162.

Graham, R. M., Hudson, S. R. et Maturilli, M. (2019). Improved performance of era5 in arctic gateway relative to four global atmospheric reanalyses. *Geophysical Research Letters*, *46*(11), 6138–6147.

Hawkins, E. et Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, *90*(8), 1095–1108.

Hoerling, M., Eischeid, J. et Perlwitz, J. (2010). Regional precipitation trends : Distinguishing natural variability from anthropogenic forcing. *Journal of Climate*, *23*(8), 2131–2145. `http://dx.doi.org/10.1175/2009jcli3420.1`. Récupéré de `https://journals.ametsoc.org/doi/abs/10.1175/2009JCLI3420.1`

Innocenti, S., Mailhot, A., Frigon, A., Cannon, A. J. et Leduc, M. (2019). Observed and simulated precipitation over northeastern north america : how do daily and sub-daily extremes scale in space and time ? *Journal of Climate*, (2019).

Leduc, M., Mailhot, A., Frigon, A., Martel, J.-L., Ludwig, R., Brietzke, G. B., Giguère, M., Brissette, F., Turcotte, R., Braun, M. et Scinocca, J. (2019). The climex project : A 50-member ensemble of climate change projections at 12-km resolution over europe and northeastern north america with the canadian regional climate model (crcm5). *Journal of Applied Meteorology and Climatology*, *58*(4), 663–693. `http://dx.doi.org/10.1175/jamc-d-18-0021.1`. Récupéré de `https://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-18-0021.1`

Liang, X., Li, L., Dai, A. et Kunkel, K. E. (2004). Regional climate model simulation of summer precipitation diurnal cycle over the united states. *Geophysical Research Letters*, *31*(24).

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, *20*(2), 130–141.

Milinski, S., Maher, N. et Olonscheck, D. (2020). How large does a large ensemble need to be ? *Earth System Dynamics*, *11*(4), 885–901.

Olauson, J. (2018). Era5 : The new champion of wind power modelling ? *Renewable energy*, *126*, 322–331.

Olsson, J., Berg, P. et Kawamura, A. (2015). Impact of rcm spatial resolution on the reproduction of local, subdaily precipitation. *Journal of hydrometeorology*, *16*(2), 534–547.

Pachauri, R. K., Allen, M. R., Barros, V. R., Broome, J., Cramer, W., Christ, R., Church, J. A., Clarke, L., Dahe, Q. et Dasgupta, P. (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.* Ipcc.

Pfeifroth, U., Trentmann, J., Fink, A. H. et Ahrens, B. (2016). Evaluating satellite-based diurnal cycles of precipitation in the african tropics. *Journal of Applied Meteorology and Climatology*, *55*(1), 23–39.

Porter, J. J., Demeritt, D. et Dessai, S. (2015). The right stuff ? informing adaptation to climate change in british local government. *Global Environmental Change*, *35*, 411–422.

Ropelewski, C. F. et Halpert, M. S. (1986). North american precipitation and temperature patterns associated with the el niño/southern oscillation (enso). *Monthly Weather Review*, *114*(12), 2352–2362.

Seiller, G. et Anctil, F. (2014). Climate change impacts on the hydrologic regime of a canadian river : comparing uncertainties arising from climate natural variability and lumped hydrological model structures. *Hydrology and Earth System Sciences*, *18*(6), 2033.

Sieck, L. C., Burges, S. J. et Steiner, M. (2007). Challenges in obtaining reliable measurements of point rainfall. *Water Resources Research*, *43*(1).

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research : Atmospheres*, *106*(D7), 7183–7192.

Tetzner, D., Thomas, E. et Allen, C. (2019). A validation of era5 reanalysis data in the southern antarctic peninsula—ellsworth land region, and its implications for ice core studies. *Geosciences*, *9*(7), 289.

Trenberth, K. E., Fasullo, J. T. et Shepherd, T. G. (2015). Attribution of climate extreme events. *Nature Climate Change*, *5*(8), 725–730.

Zhang, X., Wang, J., Zwiers, F. W. et Groisman, P. Y. (2010). The influence of large-scale climate variability on winter maximum daily precipitation over north america. *Journal of Climate*, *23*(11), 2902–2915.