University of Perugia
Department of Civil and Eviromental Engineering

INTERNATIONAL DOCTORAL PROGRAM

IN **CIVIL** AND **ENVIRONMENTAL ENGINEERING**

PhD thesis
XXXV cycle

Multi-scale modeling of precipitation processes: analysis of the water cycle with the use of both cloud-resolving models and ensemble forecasting

PhD Candidate
Miriam Saraceni

Tutor
Prof. Ass. Silvia Meniconi

International Tutor
Dr. Peter Bechtold

Co - Tutor
Dr. Paolina Bongioannini Cerlini

International Faculty Board

PhD Coordinator

Prof. Filippo Ubertini

Italian members

Belardi Paolo
Bianconi Fabio
Brunone Bruno
Breccolotti Marco
Cavalagli Nicola
Cerni Gianluca
Cluni Federico
Falcinelli Stefano
Ferrante Marco
Flammini Alessia
Formisano Antonio
Fornaciari Da Passano
Gigliotti Giovanni
Gioffrè Massimiliano
Giorgi Giacomo
Gusella Vittorio
Kenny Jose Maria
Manciola Piergiorgio
Materazzi A. Luigi
Meniconi Silvia
Menchetelli Valeria
Mochi Giovanni
Morbidelli Renato
Orlandi Fabio
Pezzolla Daniela
Puglia Debora
Salciarini Diana
Saltalippi Carla
Terenzi Benedetta
Torre Luigi
Valentini Luca
Venanzi Ilaria
Ventura Flaminia

International members

Arwade Sanjay
Caracoglia Luca
Covas Didia
Denoël Vincent
Glisic Branko
Grigoriu Mircea
Halevi Yoram
Laflamme Simon
Rodríguez-Tembleque Luis
Saez Andres
Sritharan Sri

Contents

List of Figures

List of Tables

Abbreviations

ARPS	Advanced Regional Prediction System
CRMs	Cloud Resolving Models
ECMWF	European Centre for Medium-Range Weather Forecasting
ERA5L	ERA5 - Land Reanalysis
FLake	Fresh-water Lake model
GCM	General Circulation Models
IFS	Integrated Forecasting System
RCE	Radiative Convective Equilibrium
RCEMIP	Radiative Convective Equilibrium Modeling Intercomparison Project
SAM	System of Atmospheric Modeling
SCM	Single Column Models
SPP	Stochastically Perturbed Parameterizations

1. Introduction

Human-caused climate change has driven changes that are now detectable in the global water cycle since mid- 20th century (Chapter 8 of the last report of the Intergovernmental Panel on Climate Change, IPCC (Masson-Delmotte et al., 2021)). It is further projected with high confidence that there will be an increased variability of the water cycle in most regions of the world and under all emission scenarios (Masson-Delmotte et al., 2021). The IPCC also concluded that there is evidence of observed changes and attribution to the human influence for several types of extremes since the last report (IPCC, 2014), particularly regarding extreme precipitation, drought, tropical cyclones, and compound extremes.

The robustness of the response of each term of the water budget (evaporation, precipitation, and the inflow and outflow of water vapor) to a warming climate has been assessed in many studies (Held and Soden, 2006) matching observation with General Circulation Models (GCM) simulations. The predicted changes in the water cycle by the direct influence of global warming are linked to surface heating and the increased water holding capacity. The former induces increased evaporation and surface drying, increasing the risk of drought. The latter leads to increased water vapor in the atmosphere, hence more intense precipitation events, increasing the flooding risk (Trenberth, 2011). At global scales, global climate modeling studies suggest that extreme precipitation events will intensify in proportion to their intensity (O'Gorman and Schneider, 2009; Pendergrass, 2018), with the most extreme events exhibiting the highest rate of intensification.

More specifically, extreme precipitation events are the most common natural hazards that affect the Mediterranean region and are considered to have the greatest negative impact due to their frequency and the high vulnerability of densely populated areas (Llasat et al., 2010; Tramblay and Somot, 2018). The Mediterranean region is also considered a 'hot spot' for the impacts of ongoing climate change (Giorgi, 2006), making it particularly vulnerable to the effects of extreme precipitation events. Extreme events will become more frequent, greatly impacting water resources. This is extremely relevant in terms of water scarcity (IPCC, 2014; Lionello, 2012). Specifically for the Mediterranean region, the trend confirmed by climate scenarios is that of drought and decreasing soil moisture (Grillakis, 2019; Wilhite and Pulwarty, 2017). One of the most important and most impacted water resources by climate change in Europe are lakes (George, 2010).

Thus, the first aspect addressed by the thesis is if there is a possibility of monitoring the water budget of lakes in the Mediterranean region under climate change and if it is possible to infer the climate extremes' impact on the lakes water mass balance variations with the use of modeling, specifically GCMs. Indeed, GCMs and modeling in general, by describing all the components of the atmosphere from fundamental physical-mathematical energy-momentum and mass conservation laws, provide physical quantification that enables us to examine the driving forces of atmospheric phenomena.

Moreover, the analysis is carried out to control how GCMs are able to reproduce such events in the region. The use of GCMs is also justified by the fact that the latter allows for the use of consistent spatial and temporal data and are available for long periods of time. Furthermore, GCMs provide relevant climate datasets, known as reanalyses (for example ERA5 and ERA5 Land, the reanalysis of the European Centre for Medium-Range Weather Forecasting (ECMWF) (Hersbach et al., 2020)), which can be used for this purpose and have proven useful in many contexts of hydrogeological modeling of inland water basins (Mahto and Mishra, 2019; Tarek et al., 2020) that are greatly affected by climate change (Schneider and Hook, 2010; Wang et al., 2018).

However, by using GCMs and reanalysis it is found that, even if they actually provide a reliable source for hydrological modeling and are capable of reproducing well the water cycle with its variability, there are some drawbacks. Indeed, all the assessed (and projected) changes in the water cycle budget of the atmosphere, specifically regarding precipitation, are uncertain, due to model uncertainty (Masson-Delmotte et al., 2021). Indeed, although GCMs have made significant progress in simulating the Earth's climate, there are still many challenges in accurately representing the regional patterns, temporal variations, and combination of frequency and intensity of precipitation (Trenberth et al., 2003;

Dai, 2006; Meehl et al., 2005). Especially regarding precipitation extremes, are not simulated reliably, since there is a discrepancy between the observed and modeled rate of increase of precipitation with warming (O'Gorman and Schneider, 2009).

The increase in precipitation is correlated with the increase in water vapor amount. The latter is constrained by the Clausius-Clapeyron (CC) relationship, which predicts an increase in the saturation water vapor pressure (the water holding capacity of air) of approximately 7% per °C rise in temperature (Held and Soden, 2006). For surface temperatures near those of the present-day tropics, the precipitation extremes are thought to increase at a rate close to this thermodynamic rate (Muller et al., 2011) In contrast, model-predicted changes to global precipitation increase happen at a rate of 2% per °C rise (CMIP3), significantly less than the CC predicted rate (Naegele, 2016). This discrepancy firstly highlights the fact that there are other processes, mainly energetic ones, that constrain precipitation besides the moisture availability (Naegele, 2016) and secondly, underlines the fact that GCMs have produced unrealistically small precipitation increases with warming. In summary, models are not able to accurately reproduce extreme events, particularly those related to deep convection such as precipitation. This is because convective precipitation remains a major source of systematic errors in weather prediction and climate models (Reynolds et al., 2020). Moreover, GCMs do not resolve convection explicitly, but their convective processes, thus cloud formation and precipitation, are parametrized (expressed as a function of the model prognostic variables). Thus, errors in simulated precipitation fields often indicate weaknesses in the representation of these physical processes in the model.

Given this, it makes more pressing the issue to make better predictions of precipitation events and carefully analyze simulated precipitation fields for model evaluation and development. It is of primary importance to investigate in detail physical processes and more specifically convective ones and how they are modeled, to unravel sources of uncertainties and understand how convection is reproduced in models. Thus, the second aspect under focus in the thesis is improving the understanding of the uncertainties behind convective processes and precipitation. This is done, on one hand, by analyzing in detail convective processes in an idealized setup, the Radiative Convective Equilibrium (RCE), with the use of Cloud Resolving Models (CRMs) and, on the other hand, by investigating the reliability and performance of GCMs for extreme convective mesoscale systems in our region like the Mediterranean tropical-like cyclones also called "medicanes" (Miglietta et al., 2013). Indeed, the use of models of different levels of complexity, as suggested in the literature (Maher et al., 2019), provides an understanding of the processes involved in modeling convection and helps in making useful and valuable predictions.

The study of convection at resolved scale is widely recognized as important, specifically because CRMs, whose grid-spacing is fine enough, allow explicit convection simulation (Guichard and Couvreux, 2017; Randall et al., 1996), hence providing detailed process information about convection (Petch, 2006). CRMs have been extensively used for simulating mesoscale convective systems at mid-latitudes (Barthlott and Davolio, 2016; Ferretti et al., 2014; Miglietta and Rotunno, 2019) and to study climate equilibrium and its instabilities (Muller and Held, 2012; Emanuel et al., 2014; Yano and Plant, 2012). In this context, the chosen approach is to run CRMs at RCE. Indeed, RCE, which is a statistical equilibrium between the atmosphere and surface, in the absence of lateral energy transport, thus representing the tropical climate (Held et al., 1993), has been widely applied for studying climate change (Wing et al., 2020) and convection (Bretherton et al., 2005).

This framework is commonly used because the instability of the Earth's climate and the ability of climate models to simulate it may be connected to the instability of RCE states (Raymond and Zeng, 2000), which are characterized by deep convection and a balance between radiative cooling and convective heating. Understanding the physics underlying RCE instability could be critical for accurately simulating a climate in which deep convection is prevalent, as is the case in the current climate. Indeed, the observed convective organization (including squall lines, mesoscale convective systems (Emanuel et al., 1994; Holton, 1973), the Madden–Julian oscillation (Grabowski and Moncrieff, 2004)) can also spontaneously appear in cloud-resolving models that are forced homogeneously to RCE. This type of phenomenon in models is relevant to real-world convection and climate (Holloway et al., 2017) and more specifically is thought to be related to the formation of tropical cyclones (Muller and Romps, 2018).

Thus, the study of convection in idealized CRM simulation can give insight into what feedback are more relevant to the formation of the convective organization, using it as a proxy for the above-mentioned mesoscale convective systems. This indeed represents the analysis carried out in the thesis. In particular, the study of the organization of convection, by running CRMs to RCE, has underlined its role in cyclogenesis (Muller and Romps, 2018). Muller and Romps (2018) identified the important role of radiative feedback in cyclogenesis, similar to what happens in idealized RCE simulations. This is particularly true for tropical cyclones (Landsea and Cangialosi, 2018; Zhang et al., 2019b) and their counterpart in the Mediterranean area, the medicanes. Indeed, medicanes are a particular type of Mediterranean cyclone that, for a certain period of their life, can exhibit tropical-like conditions. As a result, similar feedback mechanisms to those of tropical cyclones have been recognized to maintain them

(Flaounas et al., 2022). It is precisely these feedback mechanisms that are generally present in RCE studies and are worth studying in idealized conditions as they would be reproduced in CRMs.

However, in convective organization studies using RCE regarding cyclogenesis (Muller and Romps, 2018), besides the similarities between the idealized conditions and the real atmosphere, it is recognized that, in the latter, cyclogenesis can be quite different and is influenced by large-scale environmental conditions. Thus, if on one hand, the improvement in modeling convection and investigating convective organization can be a tool to understand the processes and feedback behind mesoscale convective systems such as cyclones, on the other hand, the improvement of the formation and intensification predictions of these latter is vital, especially regarding the effects of the large scale environment.

For this reason, the focus is also put on how the convective phenomena are reproduced by models and what aspects of mesoscale convective systems are more sensitive to the model simulation. This type of event is indeed representative of deep moist convection within a strong vertical circulation. In this context, tropical cyclones have been identified among the most impactful extreme events, same as for the corresponding phenomenon in the Mediterranean region, the medicanes. These phenomena have been chosen as they have an important hydrological impact in our region, which is intensifying with global warming (Trenberth et al., 2003). Thus, their prediction needs to be improved.

The use of GCMs ensemble forecasting is proposed. Indeed, in the literature, the advantages of using ensemble forecasts instead of high-resolution deterministic forecasts for studying this type of phenomenon have been recognized (Magnusson et al., 2015; Buizza and Hollingsworth, 2002). In the case of the Mediterranean cyclone, and in particular medicanes, there has been a long history of international coordination to improve the predictability (Flaounas et al., 2022). There are inherent limits to the predictability of these cyclones, due to the predictability of mid-latitude weather (Zhang et al., 2019a). More specifically, studies have shown that large-scale conditions often pose a challenge to the predictability of Mediterranean cyclogenesis (Chaboureau et al., 2012; Portmann et al., 2020; Pantillon et al., 2013; Di Muzio et al., 2019). The limits of predictability are also due to the multiscale nature of these cyclones, the contrast in the density of observations between land and sea, and the diverse geography of the region (Flaounas et al., 2022). Furthermore, the challenges that are faced when dealing with cyclones intensification, regard the influence by (and the interaction between) various processes that happen at different scales, including small, convective-scale processes (Judt et al., 2016) and larger, synoptic-scale environmental conditions. Improving the accuracy, reliability, and timeliness of forecasts for these cyclones is a complex task that requires advances in both

large and convective-scale initial condition uncertainty representations. This implies once again the use of ensemble forecasting, as the chosen approach.

Furthermore, the interest behind the study of both the RCE at the tropics and the tropical-like cyclones in the Mediterranean lies in the fact that both are verified equilibrium states. RCE at the tropics is equilibrium over large space and time scales between radiative cooling and radiative heating, and tropical cyclones represent local equilibrium between the heating (coming mainly from the latent heat of vaporization from the ocean's surface, absorbed by the inward airflow of the cyclone) and the turbulent dissipation in the storm's atmospheric boundary layer.

Improving understanding of climate feedback and convection and how they are modeled on one hand, and improving forecasting of extreme convective events, on the other, can eventually lead to better precipitation prediction. This, in turn, will benefit the assessment of the changes in the hydrological cycle and their effects, helping to reduce the uncertainty connected to the water cycle variability and to hydrological modeling and thus can inform better climate change mitigation measures.

1.1.Purpose of the study

The purpose of the study has been to investigate the hydrological cycle of the atmosphere and its changes, specifically regarding mesoscale convective systems at different spatial and temporal scales with a multi-modeling approach. More specifically the goals of the work have been:

- Utilizing data from climate models in order to support the identification of mitigation measures against the impact of climate change on water resources, in particular lakes. This is done together with the combination of observational data and GCMs output, reanalysis data, of the hydrological cycle and its changes in the context of modeling the water mass balance of inland water bodies;

- Examining convective organization, by using RCE simulations with different convective permitting models, specifically investigating the different roles of climate feedback (moisture feedback and radiative feedback) behind the occurrence of the above-mentioned RCE instability, which is crucial for convection in mesoscale convective systems, in order to find common feature in models used in climate sensitivities studies;

- Investigating the simulation of Mediterranean cyclones, specifically "medicanes" and the effect of including the uncertainty on the physical parameters in addition to that of initial conditions, by means of ensemble

forecasting. This is done in order to analyze the improvements or drawbacks in precipitation, intensity, and cyclone tracking predictions for these phenomena;

Indeed, the outcome of the global circulation at different space-time scales is precipitation. This makes it important to look at different scales of precipitation with a multi-modeling approach. Thus, in this research, the focus has been put not on a single type of model, but both on CRMs, and GCMs in combination with observational data and GCMs output. Indeed, one of the thesis assumptions is that the main way to study the hydrological cycle and its variability is to combine observations with modeling.

1.2.Overview of the chapters

This research has been focused firstly on the combination of observational data and GCMs output, reanalysis data, to be used in modeling the water mass balance of one of the most important water resources of the central Italy region of Umbria, Trasimeno lake, to be used as a case study. Thus, the second chapter of the research is concerned with water cycle budget modeling that has been based on the combination of observational data of different kinds, meteorological observations (in our case from the Umbrian ground-based network), observations from satellites, and climate data set (reanalysis) as observations. Indeed, observations and reanalysis have been used as forcing for regional hydrological simulations and water mass balance comparison. Specifically, the ERA5 Land reanalysis, the ECMWF reanalysis, has been used (Muñoz-Sabater et al., 2021). Its importance not only lies in the fact that it represents the output of a GCM and gives a global description of the climate in time, which is reflected in local atmospheric conditions, but also because it is the first reanalysis that includes the representation of lakes at the surface, with the model Fresh-water Lake model (FLake) (Mironov, 2008). The combination of modeling and the vast use of data turned out to be a suitable tool to study climate change in the region, reflected in the variability of the water cycle.

It has to be noted here, that the research conducted in the second chapter concerns the use of observations belonging to the Umbria region meteorological ground-based network. The use of this data involved previous work that has been carried out for their validation. Indeed the ground-based network provides hourly measurements which are affected by different types of errors and present missing data, and some work has been done to validate and reconstruct those data (Bongioannini Cerlini et al., 2020) and to integrate all the observations present in the territory in an Integrated Meteorological Network (Silvestri et al., 2022b).

After becoming familiar with reanalyses and the underlying climate processes of

climate change and the variation of the hydrologic cycle in the context of global warming, the study proceeded to investigate the hydrological cycle with CRMs, by modeling explicitly deep convection, making it possible to resolve sub-grid deep convective processes. This is the content of the third chapter of the work. This high-resolution modeling is very important since it is becoming possible also for climate modeling to increase its resolution (Stevens et al., 2019).

It is noted here that, before using CRMs, the global water cycle and its equilibrium state, have been studied through the use of Single Column Models (SCM), mimicking RCE equilibrium. SCM means simulating a single vertical column of the GCM. It represents a grid point of the parent GCM which is evolving in time, considered in isolation from the rest of the model. Indeed this part of the study was thought of as a fundamental step in understanding idealized simulations and initialization, models output managing as well as getting an understanding of how the tropical circulation works and what type of equilibrium it can reach. This work paved the path toward convective simulation with CRMs.

The study then proceeded with the simulation and comparison of two different CRMs in their simulation of RCE, investigating the aggregation phenomenon. The simulations were carried out with two different models for numerical integration and physics conception, in an attempt to try to isolate the processes leading to aggregation that was common to the two models, considering, instead of sensitivity experiments with the same model, the differences between the two models. Indeed, in the light of the last Radiative Convective Equilibrium Modeling Intercomparison Project (RCEMIP) project on the RCE simulations of several different models under various initialization, it was seen the difference in the models' numerics and physics plays the most important role (Wing et al., 2020). The used models are the Advanced Regional Prediction System (ARPS) model (Xue et al., 2000) built for the simulation of convective cells and supercells and the System of Atmospheric Modeling (SAM) climate model (Khairoutdinov and Kogan, 1999). The two models are both non-hydrostatic and in both of them, convection is solved explicitly. They both present the conservation equations of mass, momentum, and energy, and both are constructed with humid variables (water vapor content, cloud ice, cloud water, precipitating water content, hail, and graupel) as prognostics. The simulations all start from a tropical equilibrium profile condition (in temperature, humidity, and pressure) and have been run until equilibrium was reached. The phenomenon of aggregation has then been assessed and its sensitivity to many models properties have been studied.

Lastly, from the study of convective organization in three-dimensional models run in RCE and the associated climate feedback, we moved to the study of mesoscale convective systems, such as tropical cyclones, through ensembles,

moving beyond RCE and considering the effect of rotation. Thus the fourth chapter is concerned with going from modeling convection with high-resolution 3D models to global mesoscale models such as GCMs. In particular, attention has been focused on the Mediterranean region and medicanes, which are tropical-like cyclone phenomena. The study investigated how ensembles, which contain a measure of uncertainty in initial conditions and physical parameters of convection, turbulence, radiation, and cloud process parameterizations, can accurately predict the position and intensity of these phenomena. To do this, a set of three experiments was used with the high-resolution (around 9 km) Integrated Forecasting System (IFS) model of the ECMWF. The three experiments compared an ensemble forecast with perturbation in initial conditions, an ensemble forecast with perturbation in the physical parameters of the convection parameterization (such as entrainment), and an ensemble forecast with perturbation in the parameters of all relevant model parameterizations. The technique used for physical perturbation is a relatively new technique called Stochastically Perturbed Parameterizations (SPP), and it is not yet used operationally by ECMWF. For this reason, the study also looked at the differences in performance between the use of SPP and the use of operational ensembles.

2. Monitoring the water mass balance of small shallow lakes by an ERA5-Land reanalysis and water level measurement-based model

This chapter[1] regards the simulation of the lake water levels of small shallow lakes where Precipitation and Evaporation are considered the main fluxes affecting the lake water mass balance. The goal of the work is to provide a model that can inform lake management and help understand the future changes in the lake water balance, following its past. The approach used for the modeling is based on the use of widely available sets of data: time series of precipitation and evaporation and lake water levels measurements. The former is comprised of global atmospheric data sets (reanalysis) and the latter is provided by regional authorities on public websites. The tool developed using these data is an alternative to parametric modeling of fluxes and to the use of various sources of data (satellites, local weather stations, output of general circulation models) to compose the lake water mass balance. The model is based on the use of ERA5

[1] This chapter is a slightly modified version of the paper: Cerlini PB, Saraceni M, Silvestri L, Meniconi S, Brunone B., Monitoring the Water Mass Balance Variability of Small Shallow Lakes by an ERA5-Land Reanalysis and Water Level Measurement-Based Model. An Application to the Trasimeno Lake, Italy. *Atmosphere*, 2022. DOI: https://doi.org/10.3390/atmos13060949.

Land, the land component of the third-generation ERA5 reanalysis, provided by the ECMWF. With respect to other available reanalyses, ERA5 Land has been adopted because it is the only reanalysis including lake representation, by means of the Fresh-water Lake model, FLake. In the proposed method, the lake water levels are correlated with the water mass balance obtained by the reanalysis. This is preceded by data validation by considering both ground-based and satellite observations. The outcome of the modeling is analyzed by looking at the models strength and flaws, reflecting climate extreme period effects on the lake water resources. The chosen case study is the Trasimeno lake, a small and shallow lake located in Central Italy. The results show that there is a good agreement between the observed monthly variation of the lake level and the corresponding values of the water storage. This promising result encourages the use of ERA5-Land for other lakes.

2.1.Introduction

Climate change has a significant impact on lake water bodies worldwide (Woolway et al., 2020). At the same time, according to their size, lakes may behave as sentinels or regulators of climate change (Williamson et al., 2009; Kirillin et al., 2021). The impact of climate change has been recognized globally for many aspects concerning lakes. The temperature has been increasing (O'Reilly et al., 2015; Maberly et al., 2020; Rai et al., 2021) and ice phenology has been changing, with ice cover loss (Sharma et al., 2019) due to earlier ice breakup, later ice freeze-up, and shorter ice duration, at least in the northern hemisphere (Magnuson et al., 2000). Evaporation of lakes has been changing pattern, due to surface warming (Schneider and Hook, 2010), shortening of the ice cover periods, and decreasing the ratio of sensible to latent heat flux (Wang et al., 2018). Specifically in Europe, between 1966-2015, the average lake surface temperature has been increasing by about 0.58 ^{o}C/decade (Dokulil et al., 2021). Such a feature can change the lake mixing regime – both globally (Woolway and Merchant, 2019) and regionally (Kirillin, 2010) – and accelerate the evaporation process. Together with intense precipitation, evaporation affects the lake's water storage and then its level. Shifts in precipitation due to evaporation cause changes in the water budget and residence time of lakes, as well as in their depth and areal extent (Nowak and Lawniczak-Malińska, 2019; Lei et al., 2019; Schulz et al., 2020; Wang et al., 2018). However, the sensitivity of the water level to climate change and the extent of future variations remains quite uncertain in quantitative terms (Woolway et al., 2020).

To improve the understanding of lakes' response to climate change – in particular, the behavior in time of their level – in situ measurements should be coupled to numerical modeling of the physical processes. In the literature (Kebede et al., 2006; Gronewold et al., 2020; Vanderkelen et al., 2018), the shared approach for evaluating the water level change, ΔH, in a given time interval, Δt, is based

on the water mass balance equation:

$$H(t) - H(t - \Delta t) = \Delta H = P(t) - E(t) + R(t) \tag{2.1}$$

where $P =$ precipitation over the lake, $E=$ evaporation from the lake, and the term R comprises the surface water inflow that is not absorbed by the soil in the neighboring area and then it passes into the lake; the groundwater flow, and – where there is any – water from tributaries of the lake, such as rivers or other connected lakes. According to the literature (Gronewold et al., 2020; Cao et al., 2021), in this study, the water mass balance in Eq. (2.1) has been evaluated on a monthly time scale ($\Delta t =$ one month), that is appropriate from the lake monitoring point of view.

Different approaches can be followed for evaluating quantities in Eq. (2.1). Precipitation over the lake is usually measured at the ground base (Kebede et al., 2006) even if, recently, satellite observations are used as gridded data instead of local scattered measurements (Lei et al., 2017; Smith and Gronewold, 2017; Vanderkelen et al., 2018; Lu et al., 2018). Unlike precipitation, evaporation and runoff are not usually measured at the ground base, but parametric formulas are often utilized instead. These are based on ground-based measurements of temperature, relative humidity, and wind.

For the evaporation, the most popular relationships are the Penman-Monteith (Kebede et al., 2006; Chebud and Melesse, 2009; Dessie et al., 2015; Lükő et al., 2022) and Thornthwaite (Dragoni et al., 2006) ones. Possible alternatives for evaluating the evaporation are the use of the output of the numerical weather prediction (NWP) models, run over the lake basin area (Vanderkelen et al., 2018; Setegn et al., 2011), as well as the remotely sensed observations (Lu et al., 2018; Smith and Gronewold, 2017).

For the runoff, the simplest method is to calibrate the runoff coefficient, which depends on precipitation (Kebede et al., 2006; Dragoni et al., 2006; Vanderkelen et al., 2018; Dessie et al., 2015), or to use satellite observations (Syed et al., 2005). More complex parameterizations are those characterizing "ad hoc" models as the Soil Water Assessment Tool (SWAT) (Cao et al., 2021), and the Variable Infiltration Capacity (VIC) (Zhang et al., 2017; Tong et al., 2020) ones.

In a different approach, the water mass balance is given by a general model, instead of combining parametric formulas for the evaporation and runoff. As examples, a Bayesian framework model and a harmonic regression model are utilized in Gronewold et al. (2020) and Mahmood et al. (2020), respectively; in both these models, satellite data are considered for estimating the quantities.

The mentioned approaches have proven to be good candidates for modeling the lake water mass balance. However, for two main reasons, it is not certain that a methodology, refined for a given lake, will provide good results when applied to other lakes around the globe. The first reason is that most of these recent studies are carried out on large lakes, for which a large amount of data is available, with good spatial coverage – as, for example, Great Lakes (Gronewold et al., 2020) and Victoria lake (Vanderkelen et al., 2018) – which allows representing adequately all the fluxes involved in the water mass balance. The second reason is that they use data and models that might not be readily available – as the model ICON used for evaluating evaporation in (Vanderkelen et al., 2018) – or open, as in the case of the harmonic regression model (Mahmood et al., 2020). Moreover, with regard to observations, the mentioned techniques usually consider satellite data that are not available for all the globe, or local ground-based observations that are not provided by any regional meteorological or hydrographic services. In conclusion, there is no comprehensive approach that can be straightforwardly applicable to basins of different geographic/physical characteristics.

For the above reasons, a different approach is proposed here, based on the use of ERA5 - Land Reanalysis (ERA5L)– hereafter referred to as ERA5L (Muñoz-Sabater et al., 2021) – the land component of the third generation ERA5 reanalysis (Hersbach et al., 2020), provided by ECMWF, which is one of the most well-established global atmospheric datasets. With respect to other available reanalyses, ERA5L has been adopted not only because it covers the entire globe and provides climate data that can reproduce physical processes, but also because it is the only reanalysis including lake representation, by means of the Fresh-water Lake model - hereafter referred to as FLake (Mironov et al., 2010). FLake predicts vertical temperature structure and mixing in lakes for time scales from a few hours to years. In ERA5, lake temperatures are then used for evaluating the interaction between the land and atmosphere, thus making the fluxes over and from lakes' surface more predictable (Balsamo, 2013).

In this study, ERA5L data have been used for evaluating the lake level evolution by means of Eq. (2.1). The analysis is focused on the Trasimeno lake, a relatively small and shallow lake, in the Umbria region, Central Italy, for which the water level measurements are available since 1991. Moreover, this study aims at the performance evaluation of the reanalyses as a needed action prior to their extensive use.

This chapter is structured as follows. A brief description of the Trasimeno lake characteristics and previous studies is given in Section 2.2. The three different sources of data used in this work (reanalysis, ground-based, and satellite) are

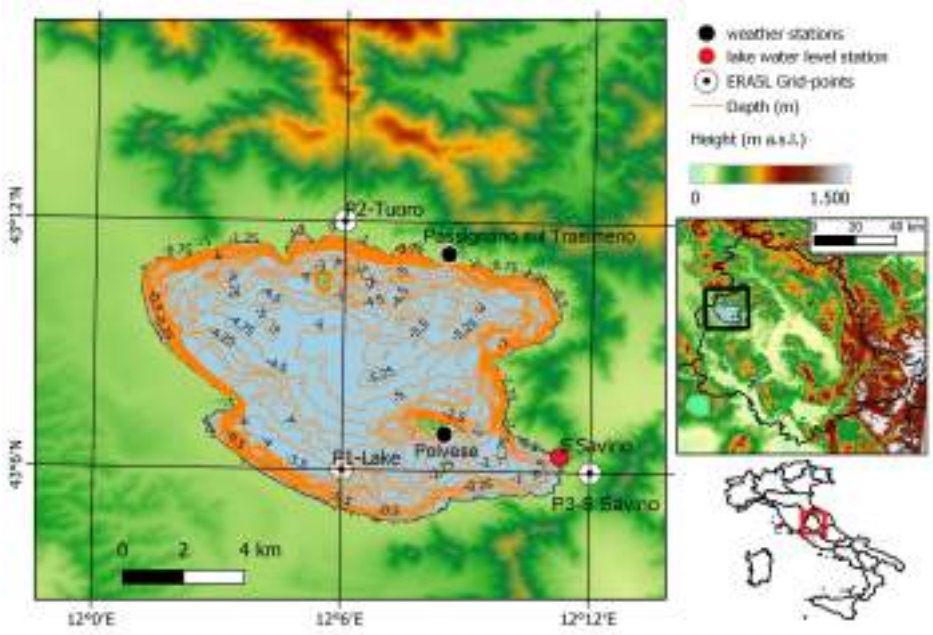

Fig. 2.1: Map of the Trasimeno lake area and bathymetry with: ERA5L grid and points (P1-Lake, P2-Tuoro, and P3-S.Savino, indicated by white circles); the Umbria region weather stations measuring temperature and precipitation (Polvese, and Passignano sul Trasimeno, indicated by black circles), and the lake level measurement station (S.Savino, indicated by a red circle).

described in Section 2.3. The methodology adopted for data validation and lake level modeling is described in Section 2.4. Results about data validation, model calibration, and verification are presented in Section 3.4. The obtained results are discussed in Section 3.5. Finally, the Conclusions are offered with regard also to the use of the proposed model for other lakes in the world.

2.2.Site description and previous studies

The Trasimeno lake is located in the Umbria region, Central Italy (Figure 2.1). The origin of the lake basin is linked to the tectonic phenomena concomitant to the final phases of the Apennine orogeny. It is a relatively small and shallow lake, with an average surface and depth of 122 km^2 and 4.2 m (Figure 2.1), respectively; accordingly, it is characterized by a large surface to volume ratio (Vincent, 2009). Thus, Trasimeno lake is smaller than those addressed in the literature, with an extent of the order of 10^3 km^2 (Kebede et al., 2006; Lu et al., 2018; Tong et al., 2020) or 10^4 km^2 (Gronewold et al., 2020). For the reasons discussed above, about the possible use of models calibrated for larger lakes, this makes the Trasimeno lake case study of particular interest for evaluating the performance of the proposed approach.

The Trasimeno lake plays a very important role from the economic point of view, i.e. for tourism and agriculture, since it is the major water resource in the surrounding region. Due to its role in terms of water supply, its level is influenced also by human management, such as the irrigation withdrawals by means of the artificial outlet, built in the XIX century. However, the major influence on the lake level is represented by climatic factors, such as basin precipitation and evaporation (Dragoni et al., 2012; Ludovisi and Gaino, 2010). On the contrary, due to the region's climate, the ice plays no role.

The catchment area of the Trasimeno lake is approximately equal to the hydrogeological basin (Dragoni et al., 2006) and it has no natural emissaries. By looking at the CORINE (Coordination of Information on the Environment) Land Cover database – provided by the European Copernicus project land monitoring services – the area around the lake consists mainly of some forests and arable areas with a few small regions covered with vineyards and olive groves. It is known that water from the lake is extracted both for on-site cultivation as well as cultivation in the areas surrounding the lake outside the catchment area. This is assumed to be particularly true in the summer months, even though there are no data available to confirm this hypothesis.

For all the above reasons, particularly the small depth, the level of the Trasimeno lake is highly variable and then particularly sensitive to climate change. Precisely, it is especially vulnerable to changes in the precipitation-to-evaporation ratio. Accordingly, it represents a very challenging water body for modeling water mass balance and could be a representative case study for many other lakes of similar size around the globe. These are many more in number than the large ones and much less studied, but they may have a significant impact on the global climate and human activities (Downing, 2010).

Throughout the years, some parametric models have been proposed for evaluating the Trasimeno lake water mass balance: the LAGO (Dragoni et al., 2006), the SimbaT (Casadei et al., 2016), and the HLM (Ubertini et al., 2007). These models are very similar since they all consider the same physical processes for the lake water mass balance and similar parameterizations. LAGO simulates the mass balance by considering the inflow-outflow processes in the lake and its catchment area. The lake basin is divided into two reservoirs representing the lake and the catchment area and aquifer, put in communication, respectively. On the contrary, SimBaT and HLM consider a unique reservoir and are based on a single water mass balance (flows-outflows) equation. In these models, the precipitation over the lake area is divided between the one falling over the lake and the one over the catchment (LAGO) or by multiplying the total precipitation to their respective areas (SimBaT and HLM). They also take into account the values observed by the local ground-based weather stations.

Evaporation instead is estimated by the Thornthwaite formula (Thornthwaite and Mather, 1957), based on the measured ground-based temperature. The runoff – evaluated simply by multiplying the precipitation over the catchment area by a calibrated runoff coefficient (SimBaT and HLM) or by the parametric formula of Vandewiele (Vandewiele et al., 1992) (LAGO) – is shown to be negligible compared to evaporation and precipitation on a yearly average (Dragoni et al., 2012). However, this may not hold in the case of extreme precipitation events, in which surface runoff and higher-than-normal groundwater levels due to floods may contribute significantly to the lake water mass balance.

The above-mentioned models, based on parametric formulas and with no specific data validation procedure, require long periods of time for calibrating the coefficients. For example, LAGO was calibrated from 2006 until 1984 (i.e., for more than twenty years) and then used to simulate the lake level backward to 1966 whereas the SimBaT model was calibrated from 1963 to 1999 (i.e., for more than thirty years) and validated from 1999 to 2010. However, these simulations fail to represent the water balance in climatically anomalous periods; for SimBaT, as an example, the drought period around 2003 (Casadei et al., 2016).
The approach proposed in this study, based on data from reanalysis (ERA5L), aims at providing an alternative to the use of the parametric models.

2.3.Data sources: Reanalysis and Observations

In this study, three different sources of data have been used for modeling the lake water mass balance: the ERA5 Land (ERA5L) reanalysis data of lake temperature, precipitation, and evaporation; the ground-based observations of lake level, precipitation, and wind; the satellite-based observations of lake surface water temperature. Each source is detailed in the following subsections. For the analysis, all data have been grouped monthly, by averaging for temperature, wind, and lake level and accumulating for precipitation and evaporation.

2.3.1.ERA5L reanalysis and the FLake model

IFS and ERA5L

The Integrated Forecast System (IFS) is a GCM provided by ECMWF, used for numerical weather predictions. The latest version of IFS (CY45R1) includes several components which interact together: the atmospheric model, the land-surface model, the ocean model, the ocean wave model, and the data assimilation of the initial conditions for the IFS system. The latter, which is a complex procedure, called 4D Var (4 Dimensional Variational assimilation), is run to provide consistency with observations. Indeed, the 4D Var is a data assimilation method of estimating the set of unknown parameters contained in the model by optimizing the fit between the solution of the model and a set of observations that the model is meant to predict, reducing the difference between the

initial conditions and the observations. The procedure of adjusting the parameters until the model 'best predicts' the observables, is known as optimization and it is run over a twelve-hour assimilation window. The "four-dimensional" nature of 4D-Var reflects the fact that the observation set spans not only three-dimensional space but also a time domain (Rabier et al., 2000). Given that, in the IFS, the water cycle is consistent with observations and physical fluxes, given the consistency and interaction between all the above-cited model components and the 4D Var process. In particular, the atmospheric model interacts with the land-surface model H-TESSEL through surface fluxes and all the meteoric species (Balsamo et al., 2009). The same complex system produces the ERA5 reanalysis (Hersbach et al., 2020), at a horizontal resolution of 0.25° ($\simeq$ 31 km) from 1950 to nowadays, where the data assimilation is run continuously with past observations instead of only using a short assimilation-window.

The land-surface component of the ERA5 reanalysis is ERA5L (Muñoz-Sabater et al., 2021). It is produced hourly using H-TESSEL, run at a higher horizontal resolution ($= 0.1° \simeq$ 11 km), driven by the downscaled meteorological forcing, via a linear interpolation from ERA5, including an elevation correction for the thermodynamic near-surface state.

Within each grid point of the IFS model, H-TESSEL consists of several tiles/fractions, each representing a different ground interface: one bare soil fraction, two vegetated fractions (high and low vegetation without snow), three snow/ice fractions (snow on bare ground/low vegetation, high vegetation with snow beneath, and lake-ice), and two water fractions (including interception reservoir and ocean). Since 2015, the specific Freshwater Lake model (FLake) – described in the next subsection – has been added to H-TESSEL to represent lakes and their interface with the atmosphere (Mironov et al., 2010).

The effectiveness of using ERA5L data derives from taking under consideration time series of points which, unlike observations, are not local data, but are spatially consistent, since they are the result of a large-scale climate model. Specifically for the Trasimeno lake area, the considered ERA5L points are homogeneous in the type of terrain.

FLake model and ERA5L fluxes

FLake is a one-dimensional model which represents fresh-water lakes (Thiery et al., 2014; Shevnina and Kourzeneva, 2017; Betts et al., 2020), by parameterizing them as basins of fixed depth divided into a surface mixed layer, which is directly in contact with the atmosphere, and a deep layer, at the bottom. By means of the heat balance, it computes the behavior over time of the temperature of the two layers: the surface mixed layer temperature, T_{ML}, and the bottom layer temperature, T_{BL}. In such a context, the shortwave radiation

reaching the surface – including both the absorbed and reflected one –, the latent heat, the sensible heat, and the long-wave radiation flux from the lake surface are taken into account. FLake provides the seasonal evolution of the lake temperature to IFS to improve the calculation of the fluxes to and out of the surface, including evaporation and precipitation (Balsamo, 2013; Manrique Suñén et al., 2012; Samuelsson et al., 2010).

The decision to use ERA5L is based on the fact that FLake, integrated with H-TESSEL, provides evaporation data from the lake since it models the lake layer temperatures. Such data are not available for the other reanalyses. Furthermore, ERA5L provides precipitation data which are calculated in interaction with the presence of the lake on the underlying surface. For a more detailed explanation of FLake, given the brief description in Appendix A, the reader can refer to the IFS documentation (ECMWF, 2018) and the theory in Mironov (2008) and Mironov et al. (2010).

FLake takes into account all inland water bodies occupying at least 1% surface area of each model grid box. For the ERA5L grid points in the Trasimeno lake area, the lake cover fraction is around 0.1. The lake depth – specified according to the bathymetry (Balsamo et al., 2012) – is assumed as equal to 4 m.

As mentioned above, FLake provides the lake temperature – specifically T_{ML} – to the land surface model for evaluating the mean skin temperature of the lake tile, T_{sk}, and then the evaporation from the lake. Surface heat, momentum, and water fluxes are calculated by H-TESSEL for each tile based on their resistance, humidity, and temperature properties. For what concerns the evaporation, in IFS a bulk aerodynamic formula is applied to calculate the turbulent flux of water vapor, E. Specifically, for each water tile, *i*, and grid point, *j*, it is:

$$E_{i,j} = \rho_a |U_{L,j}| C_{H,i,j} [q_{L,j} - q_{sat,j}(T_{sk,i,j})] \tag{2.2}$$

with ρ_a (kgm^{-3}) = air density; U_L (ms^{-1}) and q_L (g/kg) being the wind speed and humidity of the nearest to the surface atmospheric model level, respectively; q_{sat} (g/kg) = saturation specific humidity; and $C_{H,i}$ = turbulent exchange coefficient. Such a coefficient, which represents the surface resistance, is computed considering different roughness lengths for heat and momentum by the Monin–Obukhov formulation (see technical support from ECMWF (2018) for a detailed description). For each water tile, q_{sat} is calculated using the mean skin temperature of the tile, $T_{sk,i}(^\circ C)$.

The atmospheric component of IFS provides the precipitation, that interacts with the land surface model, as the result of multiple processes in the atmosphere. Accordingly, it comprises both convective and large-scale rainfall. The

former is generated by convection (when atmospheric air at low levels is warmer and less dense than the air above, so it rises), whereas the latter is by large-scale weather patterns (i.e., fronts). Large-scale and convective precipitation are calculated in two different schemes within the IFS model, the cloud and convective schemes. The moist convection scheme (Chapter 6 of (IFS Documentation CY47R3, 2021a)) is based on the mass-flux approach and represents deep, shallow, and mid-level convection. The convection scheme represents convection at spatial scales smaller than the grid box and generates precipitation based on the amount of downdraught and updraught mass fluxes and the change that they make to the moisture, heat, and momentum budgets.

The cloud scheme (Chapter 7 of (IFS Documentation CY47R3, 2021a)) parameterizes clouds and large-scale precipitation with prognostic equations for cloud liquid, cloud ice, rain, and snow water contents and a sub-grid fractional cloud cover. The cloud schemes represent the sources and sinks of clouds and precipitation due to the major generation and destruction processes. These include condensation, deposition, evaporation, collection, melting, freezing, cloud formation by detrainment from cumulus convection (Tiedtke, 1993; Forbes et al., 2011), conversion from cloud water into the rain, detrainment from mass updraft, and evaporation from precipitation (Tiedtke, 1989; Bechtold et al., 2014). All these processes are integrated over the atmosphere column at each grid point, j, to get the precipitation over the lake, P, as:

$$P_j = \int_{p_0}^{p_{top}} (G_j - D_j) \frac{dp}{g} \tag{2.3}$$

where p = column pressure (hPa) with the subscripts *0* and *top* indicating the surface and top of atmosphere, respectively, g = acceleration of gravity (= 9.8 m/s^2), and G_j and D_j represent all the precipitation sources and sinks, respectively.

Selected ERA5L data

Accordingly to ERA5L horizontal resolution, in the area surrounding the lake, three grid points have been considered (indicated by white circles in Figure 2.1): the two closest to the Trasimeno lake (i.e., P2-Tuoro and P3-S.Savino), and the one located in the lake (P1-Lake). P1-Lake has the largest lake cover fraction (= 0.1), followed by P2-Tuoro (= 0.08), and P3-S.Savino (= 0.05).

2.3.2.Ground-based data

The region around the lake is monitored by three weather stations managed by the Umbria regional hydrographic service (indicated by black circles in Figure 2.1): Polvese, Passignano sul Trasimeno, and S. Savino. Polvese and Passignano sul Trasimeno weather stations have been measuring temperature, precipitation, and wind since 2002 and 1992, respectively. The S.Savino station

measured precipitation from 1996 to 2019 and the lake level since 1991, with only two data missing periods (i.e., between 2012 and 2013) due to maintenance failure of the weather station hydrometer. The lake level is measured hourly by the ultrasonic hydrometer CAE ULM20. However, only the daily mean values are provided by the Umbria regional hydrographic service. The instrument resolution, sensitivity, and accuracy are 1 cm, 0.1 cm, and $\pm$ 1 cm (over the whole measurement range), respectively.

The time series of the daily mean level are shown in Figure 2.2 where the values are referred to the hydrometric zero value (= 257.1 m a.s.l.) that has been reached only in the early nineties and exceeded only from 2014 to 2017.

Before using lake level data, they have been checked for internal consistency on the basis of the precipitation measured at the same weather station (Bongioannini Cerlini et al., 2020; Silvestri et al., 2022b). Precisely, the lake level data have been considered invalid (thus creating a masked value) if on the same day, it rained and the lake level remained unchanged and vice-versa. This level

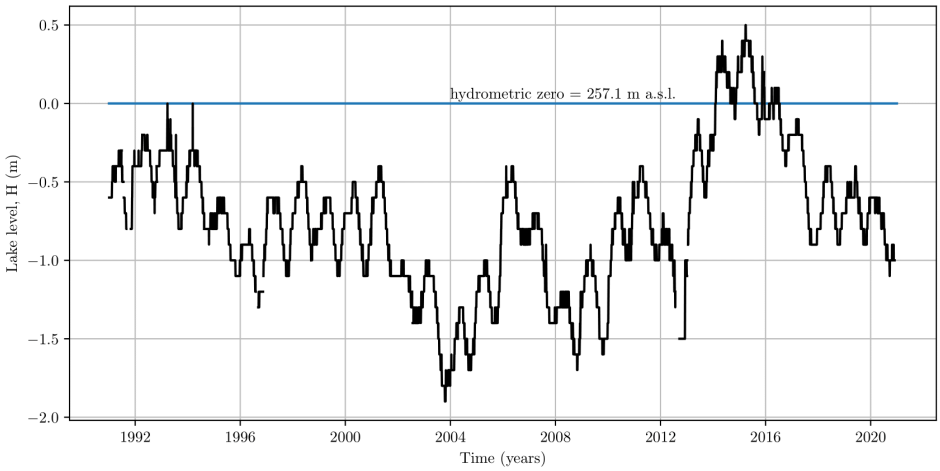

Fig. 2.2: Values of the daily mean level, H, of the Trasimeno lake, measured at the S.Savino weather station, within blue the hydrometric zero.

check was dictated by the fact that precipitation is the flux that contributes the most to the variation of the lake level. The threshold of minimum precipitation to change the lake level was set to 1.5 mm, taken as the difference between mean precipitation and evaporation over the period. Choosing the difference between daily precipitation and evaporation as the threshold is dictated by the assumption that the minimum change to the lake level is the amount of water that precipitates on the lake net of that which evaporates on a daily basis. Since no ground-based evaporation measurements are available, to choose the

threshold value, the climatological values from 1950 to 2019, given by the Italian Institute for Environmental Protection and Research (ISPRA) (Braca et al., 2021), have been considered. ISPRA reported that in Umbria, the daily mean rainfall value during the period 1996-2019 is 2.5 mm/day, and the daily mean value for evaporation is 1 mm/day, thus giving a threshold value of 1.5 mm/day.

On the basis of data from the closest weather stations to the S.Savino hydrometer (i.e., Polvese and Passignano sul Trasimeno), the daily mean wind speed on the Trasimeno lake is 2.9 m/s and the predominant wind direction is from the south (Figure 2.3). The wind set-up, evaluated by considering such average wind speed and direction in the formula proposed in (Colvin et al., 2018), amounts to approximately 1.9 cm. Such a value is comparable to the instrument resolution and accuracy and therefore its effect has been neglected. In the future, for a further and more detailed analysis, a second hydrometer should be located on the opposite side of the lake.

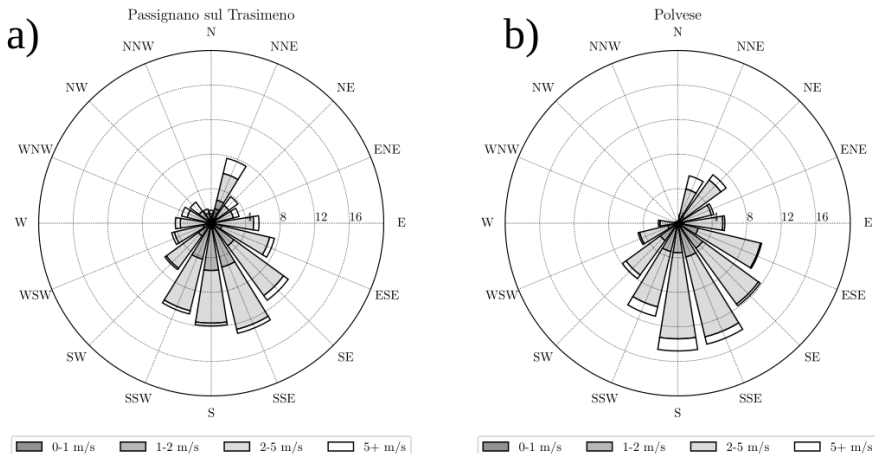

Fig. 2.3: Ten meters wind rose at the two different weather stations in the lake area (black circles in Figure 1): a) Passignano sul Trasimeno; b) Polvese. Radial distance and shading indicate the percentage of days with wind direction in the sector and wind speed intervals, respectively.

The validation check has been specifically applied to periods where the lake level is erroneously fluctuating between two values but it was not raining, as for example, around March 2002 (Figure 2.4). However, quite long periods of persistence of the lake level are probably due to the fact that indeed some weak rainfalls do not change the lake level significantly. This control, although elementary, provides a way of eliminating invalid data, which have been then disregarded in the following analysis. The final result, i.e. the validated data series of the lake level, is shown in Figure 2.4b for the period January 2002 -

January 2003, as an example.

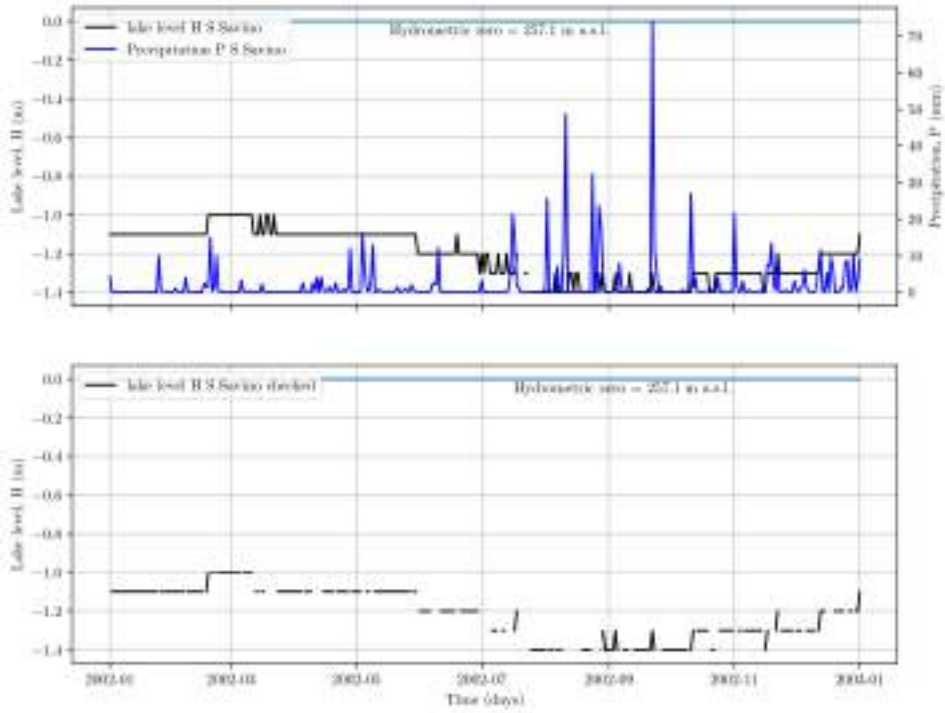

Fig. 2.4: Cross-check between precipitation and level of the Trasimeno lake at S.Savino for the period January 2002 - January 2003. In Figure 3a, the unchecked lake level is reported together with precipitation, whereas in Figure 3b the validated level is shown.

2.3.3.Satellite data

Due to the lack of ground-based evaporation observations, the ERA5L lake surface mixed layer temperature, T_{ML}, that is modeled prognostically by FLake, has been compared to the satellite observed Lake Surface Water Temperature (LSWT, °C) provided by the Natural Environment Research Council (NERC) project Global Observatory of Lake Responses to Environmental Change (GloboLakes) (Carrea and Merchant, 2019). This is because T_{ML} is the FLake parameter related to the atmosphere in the evaluation of the evaporation from the lake and it is assumed to correspond to the lake surface temperature observations.

The LSWT values are obtained by combining orbit data from different sources of multipurpose imaging radiometers. These are the AVHRR (Advanced Very High-Resolution Radiometer) in MetOpA, AATSR (Advanced Along Track Scanning Radiometer) in Envisat, and ATSR-2 (Along Track Scanning Radiometer)

in ERS-2 (European Remote Sensing Satellite). From 1995 to 2016, for all the lakes of the globe, these data are distributed on a grid with a horizontal resolution of 0.05° ($\simeq$ 5 km).

The GloboLakes dataset (v4.0) consists of the LSWT daily observations, with their uncertainty and quality levels. The uncertainty comprises an evaluation of the total uncertainty in the LSWT retrieval process. The quality level, an index ranging from 2 (suspect/marginal quality) to 5 (best quality), is a measure of the quality with which the retrieval process is executed. In this study, only LSWT values with high-quality levels (i.e., 4 or 5) have been considered.

2.4.Methods

2.4.1.ERA5L data validation

Before using for modeling the water mass balance, ERA5L data have been validated against observations for what concerns precipitation and evaporation, the most influential fluxes in the water mass balance.
The comparison between reanalysis T_{ML} and satellite LSWT has been carried out for the period 1995-2016 for each of the three ERA5L grid points and their weighted average (by lake cover fraction).

The monthly ERA5L precipitation data for each of the selected ERA5L grid points have been compared to the precipitation provided by the ground-based weather network stations (Figure 2.1). The nearest to the ERA5L point weather stations have been chosen. Precisely, the ERA5L P1-Lake precipitation has been compared to the Polvese one since 2002; the ERA5L P2-Tuoro to the Passignano sul Trasimeno one since 1996; and the ERA5L P3-S.Savino to the S.Savino one since 1996.

The correspondence between reanalysis and observations has been measured by the Pearson correlation coefficient, r^2, and the Root Mean Square Error, RMSE, defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i}^{N}(m_i - o_i)^2} \tag{2.4}$$

where m_i are the modeled values from reanalysis, o_i are the observations, and N is the number of the values sample.

2.4.2.Modeling the lake level: the conceptual model

As mentioned, for the Trasimeno lake it has been shown (Ludovisi and Gaino, 2010; Dragoni et al., 2012) that, monthly, the groundwater and surface runoff can be neglected given that, on average, the catchment area is approximately

equal to the hydro-geological basin. Accordingly in Eq. (2.1), it is assumed $R(t) = 0$ and then it can be rewritten as:

$$\Delta H \simeq P(t) - E(t) = \delta(t) \tag{2.5}$$

where ΔH = monthly variation of the lake level, and δ = *P-E* is the water storage. In other words, in this model, the basic assumption is that the precipitation and evaporation are the fluxes that mostly influence the water balance of a closed lake, like the Trasimeno one. This assumption has been verified by checking that the monthly mean runoff flux (surface and subsurface runoff), provided by the reanalysis, was of an order of magnitude smaller than the precipitation and evaporation fluxes (not shown). However, as already pointed out, this assumption might not hold either for other shallow lakes or under extreme hydrologic conditions during a period of heavy rainfall, when also inflow from groundwater does occur.

Indeed, on one hand, the groundwater inflow from the aquifer is often negligible on an annual and monthly basis, as shown for other shallow lakes in the literature (Klimaszyk et al., 2015), but it may be important in wet years, during extreme precipitation events, when it recharges the lake (Klimaszyk et al., 2015). On the other hand, some shallow lakes are highly influenced by surface runoff from rivers, which significantly influences the lake's water balance (Cao et al., 2021; Klimaszyk et al., 2015), becoming more important under conditions of extreme precipitation (Yihdego et al., 2017). However, surface runoff input in shallow lakes, as in the case of Trasimeno lake, is an order of magnitude smaller than the actual runoff from rivers or tributaries (Kebede et al., 2006).

The monthly time series of δ have been computed for each of the three ERA5L grid points and for their weighted average, by considering their respective lake cover fraction. Successively, such time-series of δ have been compared to the observed ΔH one to check for the model validity.
The performance of the proposed model has been evaluated on the basis of the RMSE, as defined by Eq. (2.4), and the Kling–Gupta efficiency, KGE_S, defined as in Gupta et al. (2009):

$$KGE_S = 1 - \sqrt{(r_S - 1)^2 + (\frac{\sigma_m}{\sigma_o} - 1)^2 + (\frac{\sigma_m}{\sigma_o} - 1)^2} \tag{2.6}$$

where r_S is the linear correlation between observations and simulations; $\frac{\sigma_m}{\sigma_o}$ and $\frac{\sigma_m}{\sigma_o}$ are a measure of the error and bias, respectively, with σ_m and σ_o being the standard deviation in simulation and observations, respectively, and μ_m and μ_o are the simulation and observation mean, respectively. Values of KGE_S equal to 1 indicate a perfect agreement between simulations and observations, whereas negative KGE_S values imply a disagreement between them.

2.5.Results

2.5.1.Lake temperature and precipitation validated data

An example of the comparison made between LSWT data from satellite and T_{ML} from reanalysis is reported in Figure 2.5 for the period from 2008 to 2011. From such a comparison it emerges that on the whole, the LSWT trend is in good agreement with the one of the temperatures given in ERA5L.

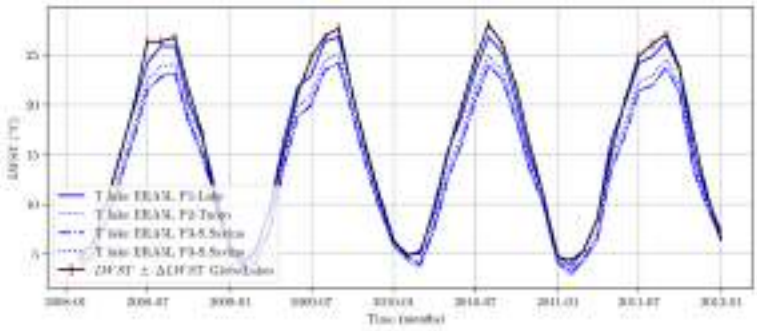

Fig. 2.5: Comparison for the period 2008-2011 between the lake surface water temperature, LSWT ($°C$) from satellite (reported in black), and the mixed layer temperature, $T_{ML}(°C)$, as modeled by ERA5L (in blue), for each ERA5L point and their weighted average.

The best correspondence concerns the ERA5L P1-Lake data, i.e. at the grid point located in the lake, which falls within the LSWT associated uncertainty. This is underlined in Table 2.1 by the values of the Pearson correlation, r_T^2, and Root Mean Square Error, RMSE_T, where the subscript T indicates the temperature.

Table 2.1: Pearson correlation coefficient, r_T^2, and RMSE_T for the lake temperature in the period 1995-2016.

	P1-Lake	**P2-Tuoro**	**P3-S.Savino**	**3 Points Mean**
r_T^2	0.99	0.99	0.99	0.99
RMSE_T (°C)	1.1	2.6	2.6	2.0

While the correlation is close to unity, the RMSE_T values indicate that the ERA5L temperatures are on average 2.6 °C different from those observed in the P2-Tuoro and P3-S.Savino grid points. The closest series of T_{ML} from ERA5L to the observed one is that from P1-Lake, which deviates by 1.1 °C. This is possibly due to the fact that point P1-Lake is located at the lake and then it has a lake fractional cover larger than the other two points.

The comparison between monthly ERA5L precipitation data and rain gauge

observations is shown in Figure 2.6, where only the period from 2008 to 2011 is reported as an example.

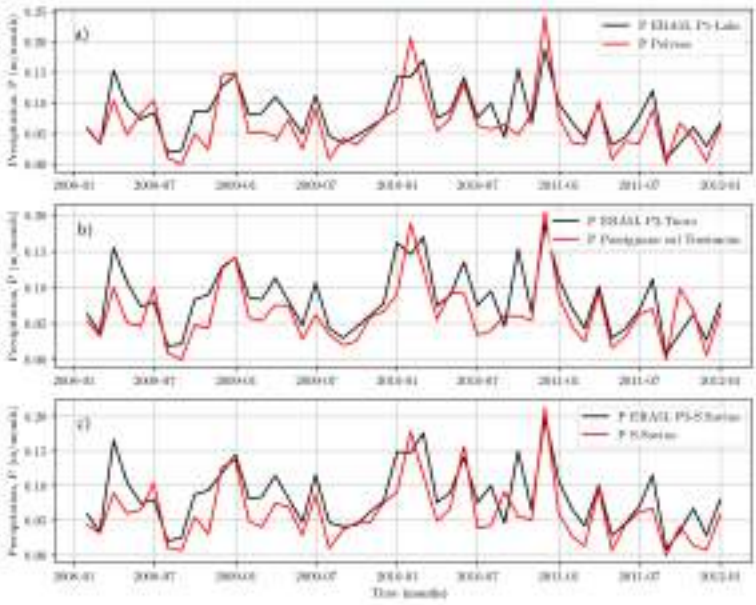

Fig. 2.6: Comparison for the period 2008-2011 between the precipitation over the lake, P (m/month), observed at ground base by weather stations (reported in red), and the one, provided by ERA5L (in black), at: a) P1-Lake vs. Polvese, b) P2-Tuoro vs. Passignano sul Trasimeno, and c) P3-S.Savino vs. S.Savino.

From such a comparison it can be noted that on the whole the precipitation trends of ERA5L and the observed one are in good agreement. In fact, even if discrepancies occur in some months, the monthly precipitation pattern is captured reasonably well by ERA5L. Indeed, for all grid points the correlation is above 0.70 (Table 2.2), with the largest value reached at P2-Tuoro ($r_P^2 =$ 0.80, where the subscript P indicates precipitation).

Table 2.2: Pearson correlation coefficient,r_P^2, and RMSE$_P$ for the precipitation over the lake in the period 1995-2016.

	P1-Lake	**P2-Tuoro**	**P3-S.Savino**
r_P^2	0.70	0.80	0.75
RMSE$_P$ (m/month)	0.04	0.03	0.04

When the RMSE$_P$ values are considered, the difference in precipitation is quite small in the three grid points. In fact, it is about 0.04 m/month for P1-Lake and P3-S.Savino, while it decreases slightly for P2-Tuoro (0.03 m/month). This analysis suggests that the precipitation range of ERA5L is in good agreement

with the values measured on the ground.
Since the ERA5L validation process has given satisfactory results, given the value of the Pearson correlation ($r_P^2 \geq 70\%$ and $r_T^2 \geq 70\%$ as in (Sun et al., 2021a; Jiao et al., 2021)), this lays the basis for reliable use of evaporation and precipitation fluxes from ERA5L.

2.5.2.Modeling the lake level: correlation analysis

The validity of the model proposed in Eq. (2.5) has been checked by comparing the time series of the modeled δ with the observed ΔH, as described in Section 2.4.2. From the comparison shown in Figure 2.7, it can be seen that for the considered period there is a good correspondence in the trend. However, at the higher values, the scattering increases, thus implying a worse correspondence between δ and ΔH. This is probably because the runoff has been neglected. Indeed, these higher levels correspond to periods in which the lake level has risen due to intense rainfall. This feature probably contributed to the increase of surface runoff and groundwater level and, consequently, of the groundwater inflow into the lake.

To assess whether the linear relationship between ΔH and δ is valid, the Pearson correlation coefficient, r_δ^2, has been evaluated, by assuming both δ and ΔH normally distributed (Figure 2.7 and Table 2.3). The correlation analysis reveals that r_δ^2 is always larger than 70 $\%$ and it reaches its maximum value ($=$ 81%) at the grid point P1-Lake (Figure 2.7a).

Table 2.3: Pearson correlation coefficient, r_δ^2, between ΔH and δ in the period 1996-2019.

	P1-Lake	**P2-Tuoro**	**P3-S.Savino**	**3 Points Mean**
r_δ^2	0.81	0.79	0.70	0.79

Although the δ series are very similar to each other – as expected since the considered grid points are located in a limited area – as for temperature and precipitation, the strongest correlation is definitely for the grid point P1-Lake, located in the Trasimeno lake (Figure 2.1).

The slightly better performance of point P1-Lake can be ascribed to the larger values of evaporation from the lake accounted for by ERA5L (not shown). In fact, in the case of P1-Lake, the evaporation from the lake accounts for 60% of the total evaporation, while in the other points, this percentage is smaller (45% for P2-Tuoro and 20% for P3-S.Savino). Thus, this might have positively impacted the correlation between the P1-Lake time-series, δ, and the observed, ΔH, especially in the summer season when evaporation is prominent

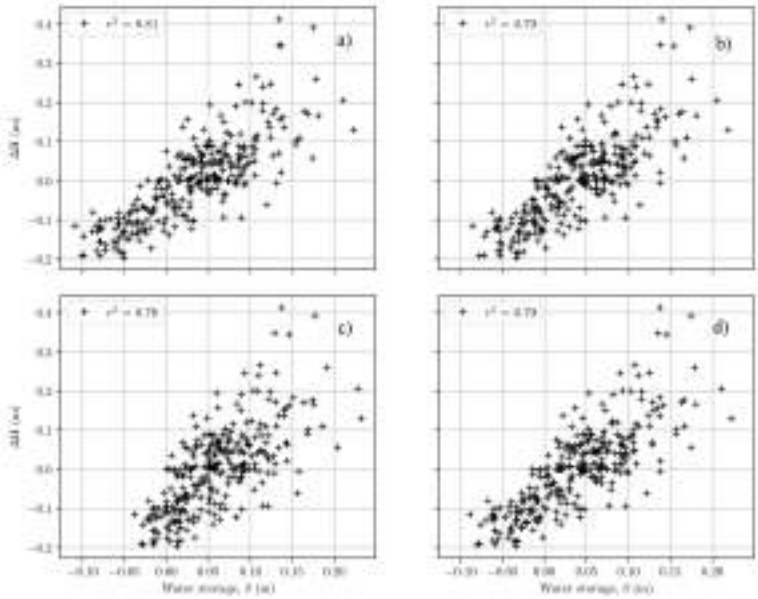

Fig. 2.7: Correlation analysis between the monthly variations of the lake level, ΔH, and water storage, δ, for the three ERA5L points closest to the lake: a) P1-Lake, b) P2-Tuoro, c) P3-S.Savino, and d) their weighted average.

in modifying the lake level. Accordingly, in the successive phase of the model calibration, it has been decided to focus the attention only on the ERA5L P1-Lake grid point. As a result of the above correlation analysis, the simplified model of Eq. (2.5) has been deemed valid for simulating the variations of the lake level, ΔH.

2.5.3.Model calibration and verification

According to Figure 2.7, a linear relationship between δ, given by ERA5L P1-Lake, and the observed ΔH has been assumed:

$$\Delta H(t) = \alpha\delta(t) + \beta \tag{2.7}$$

where the value of the coefficients α (= 1.33) and β (=-0.04) has been calibrated (with $r_\delta^2 = 0.85$) by considering only the values of ΔH in the initial five years (i.e., from 1996 to 2000, included) at the S.Savino station. The successive period 2001-2019 has then been used for model validation.
In Figure 2.8, the values of the monthly lake level, given by Eq. (2.7), are compared to the observed ones. The simulated values are in good agreement with the observed ones and allow capturing the annual trend of the observed lake level.

The values reported in Table 2.4 indicate that, apart from the validation pe-

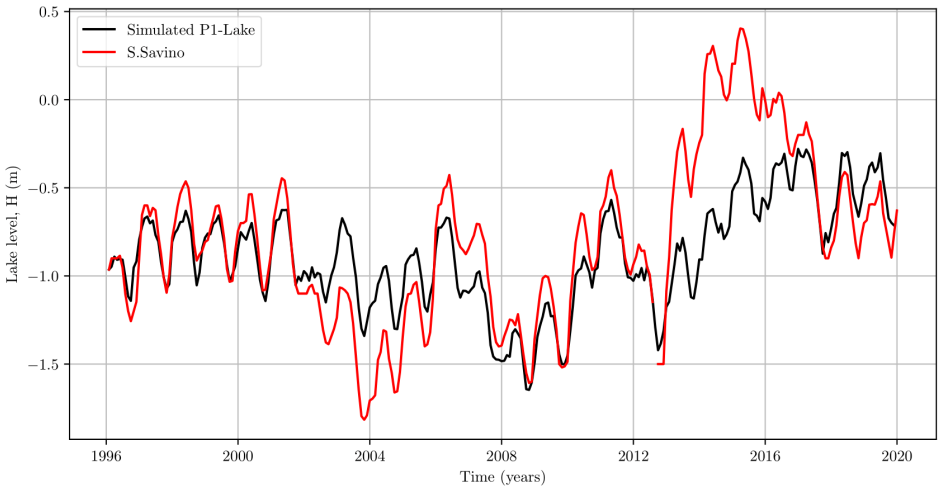

Fig. 2.8: Monthly lake level, H, simulated (in black) by Eq. (2.7) vs. observed (in red) values at S.Savino. In the model, the validation period is from 1996 to 2000, whereas the simulation one is from 2000 to 2019.

riod, where the simulation overlaps the observation, the periods that are better stimulated by the linear model are those from 2006 to 2012 and 2017 to 2019 when small RMSE_S and the closest to one KGE_S values occur (the subscript S indicates simulations).

Table 2.4: Performance of the model simulating the monthly lake level, H: Root Mean Square Error, RMSE_S, and Kling–Gupta efficiency, KGE_S, in different periods.

Period	RMSE_S (m)	KGE_S
1996-2019	0.31	0.56
2002-2005	0.30	0.55
2006-2012	0.17	0.79
2013-2016	0.57	-2.1
2017-2019	0.18	0.65

Indeed, in these two periods, RMSE_S is actually smaller (Table 2.4) than the value computed for the total simulation period 1996-2019, and the KGE_S is larger, reaching 0.79 for 2006-2012. On the contrary, in the periods 2002-2005 and 2013-2016, there is a disagreement between the simulated values and observations. Precisely, the former is larger (smaller) than the observations in 2002-2005 (2013-2016). In the first case, RMSE_S is 0.30 m while in the second case, it reaches 0.57 m, values that reflect the discrepancy between simulation and observation. Such a feature is also confirmed by the values of KGE_S which, in the period 2002-2005, is smaller than in the other ones. The negative value

in 2013-2016, makes this period the worst simulated one by the model.

An aspect that might have had an impact on the discrepancy at higher water level variation, for example in 2013-2016, is related to the assumed relationship between the water balance and the lake level, which might not be linear. Indeed, the lake surface area increases disproportionally at higher water levels (Figure 2.7), so the increase in water level per increase in water storage diminishes, implying a parabolic relationship between the lake level and the surface area, as reported in the literature (Muala et al., 2014; Li et al., 2019). However, this aspect has been neglected since the quadratic interpolation of the data points gives almost equal results in terms of the Pearson correlation (not shown). Moreover, the contribution of groundwater inflow was probably prevailing in this period.

As pointed out above, this misrepresentation, especially for the period 2013-2016, could also be related to the fact that the runoff has been neglected. In fact, runoff, in the form of groundwater inflow, could be quite influential in an extremely wet period. Unlikely, the lack of a piezometric network in the lake area does not allow getting at a certain determination about the role of the groundwater inflow.

2.6. Discussion

Figure 2.8 plots indicate that the proposed model is able to reproduce the trend of the observed lake level, managing to simulate adequately the dynamics of the water balance of the Trasimeno lake. However, as mentioned, the model (almost) consistently underestimates the water level and there are few specific periods where the simulation fails to capture the observed values. The first one is from 2002 to 2005, and the second one is from 2013 to 2016.

The first period was a drought period: at the end of 2003, the lake level reached -1.9 m, i.e., the minimum value in 1996-2019. It is worth noting that 2003 is the year when Europe experienced a particularly extreme climate anomaly with July temperatures anomaly reaching 6° (Beniston, 2004; Black et al., 2004). In this period, the model returns values of δ that are much higher than the observations ($\text{RMSE}_S = 0.30$ m in Table 2.4). The second period, late 2012-early 2013, with the poorest agreement ($\text{RMSE}_S = 0.57$ m in Table 2.4), is characterized by intense precipitation over central Italy (Ferretti et al., 2014) and, then, also over the Trasimeno lake area.

To understand the reasons of these discrepancies, the monthly precipitation of ERA5L has been compared with the ground-based observations. This is because, although the ERA5L precipitation data have been validated on the whole against observations, there may be errors during periods of climate extremes.

With this aim, the weather station of Polvese (in operation since 2003) has been considered, as it is the closest to the ERA5L point P1-Lake (Figure 2.1). The results of such an analysis are shown in Figures 2.9 and 2.10 for the period 2002-2005 and 2012-2013, respectively.

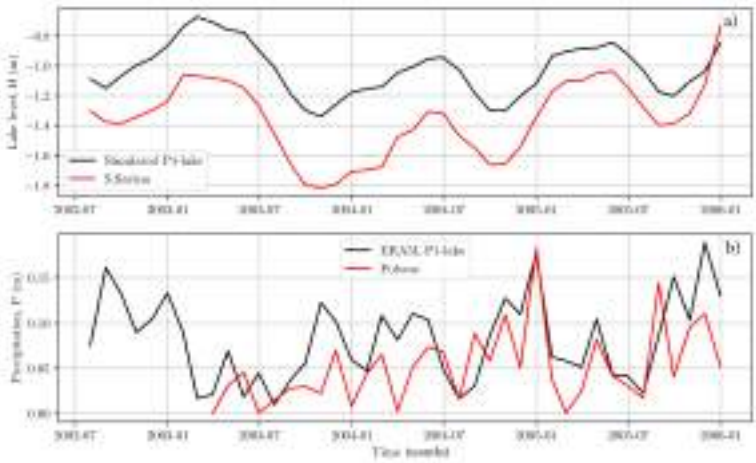

Fig. 2.9: Comparison between ERA5L and ground-based observed data for the drought period 2002-2006: a) monthly lake level, H, from ERA5L P1-Lake (in black) and the observed one at S. Savino station (in red); b) precipitation over the lake (m/month), P, from ERA5L P1-Lake (in black) and the observed one at Polvese station (in red).

The drought period is not properly captured by reanalysis precipitation data (Figure 2.9). Indeed, since April 2003, the comparison between the precipitation given by ERA5L at P1-Lake and the one observed at Polvese, points out that the former is much larger (0.10 m/month vs. 0.05 m/month). This feature persists throughout most of 2004 and 2005. At the beginning of 2003, precipitation data were not available for the Polvese station. However, for the other stations of S.Savino and Passignano sul Trasimeno, both in the lake area (Figure 2.1), an overestimation of the precipitation by ERA5L has been noticed (not shown).

Therefore, a similar overestimation can be assumed also for the Polvese station. Such discrepancies between simulated and observed data in drought periods could also be due to possible errors in the evaporation data, which cannot be directly checked as there are no evaporation measurements. However, the check of LSWT vs. the temperature from ERA5L indicates an RMSE_T of about 1 degree, which could have influenced a difference in the evaporation. Moreover, another element of misrepresentation in this drought period might come from a non-negligible groundwater outflow from the lake, which supplies the aquifer in drought periods, as documented for other shallow lakes (Klimaszyk et al., 2015).

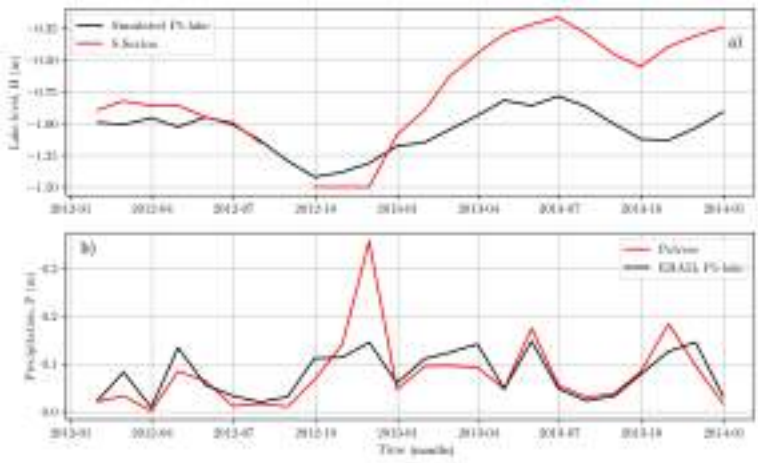

Fig. 2.10: Comparison between ERA5L and ground-based observed data for the intense precipitation period 2012-2013: a) monthly lake level, H, from ERA5L P1-Lake (in black) and the observed one at S. Savino station (in red); b) precipitation over the lake (m/month), P, from ERA5L P1-Lake (in black) and the observed one at Polvese station (in red).

For the intense precipitation period, the focus has been put on late 2012 and early 2013. In fact, in December 2012 (Figure 2.10), a monthly peak of precipitation equal to 0.30 m has been reached at the Polvese station, which exceeds also the climatological average of precipitation in 1981-2010 ($\simeq$ 0.1 m), evaluated by means of data collected by the hydrographic service. This peak in the ground-based precipitation observation is not captured by the ERA5L precipitation series, with a bias of 0.20 m/month between the observation and reanalysis. Failing to be captured by the reanalysis, this discrepancy has affected the reproduction of the subsequent monthly lake level maintaining it lower also in the following period, until 2016, when there is again a good match between reanalysis and observations.

As a result of the above analysis, it can be affirmed that the precipitation biases are responsible for the incorrect simulation of the lake level. However, as anticipated in the results section 2.5.2, also the omission of the groundwater inflow could have contributed to the incorrect simulation of the peak between 2012 and 2013. Indeed, the runoff (surface and subsurface) has been neglected, by comparing it with the evaporation and precipitation fluxes on a monthly time scale. This inaccuracy might have affected the lake level modeling in this period. Indeed, in the period between 2017 and 2019, there were some intense precipitation events (for example, on May 2019 (ISPRA, 2020), with a monthly average of 0.18 m), that, although not comparable to those of 2012, are quite

well simulated by the proposed model. Thus, the modeling has failed in an extreme context, which, differently than in other cases, has probably caused an increase in the groundwater level and, as a consequence, a non-negligible groundwater flux. Thus, the proposed model could be expanded with the inclusion of a finer resolution groundwater fluxes modeling (e.g., by the SWAT model (Cao et al., 2021)). This could be important, especially for future model development in its application to other lakes around the globe, where the runoff might not be negligible.

For what it concerns the biases in precipitation, the results obtained here are in line with what was found in Central Italy where discrepancies between observations and ERA5 reanalysis precipitation have been noticed in general, and specifically in the Trasimeno lake area, where yearly mean precipitation is overestimated by ERA5 reanalyses (Silvestri et al., 2022a). Moreover, in the literature, precipitation biases of ERA5 have been noted in other parts of the world. In particular, in (Jiang et al., 2021) it is shown that even if the spatial and temporal patterns of ERA5 are consistent with rain gauges, ERA5 tends to overestimate the light precipitation events and underestimate moderate and heavy precipitation events in China. Similarly, in (Xu et al., 2021) it is pointed out that in certain seasons the precipitation data provided by ERA5 and ERA5L can be outperformed by satellite data. However, it must be noted that this type of limitations, i.e. the simulation of extreme events (drought and intense precipitation) are shared by other reanalyses, thus by the other GCMs worldwide like IFS (Sillmann et al., 2013; Moon et al., 2018).

Regarding drought periods, the underestimation of the extent and persistence of dryness is confirmed by the results of the Coupled Model Intercomparison Project 5 (CMIP5) (Taylor et al., 2012b), which have shown a general underestimation of the persistence of drought periods (Moon et al., 2018). Furthermore, if the specific drought event that occurred in 2003 is considered, it has been found that many hydrological models based on reanalysis or GCM input underestimate the intensity of the event (Schewe et al., 2019). Indeed, the causes for the 2003 drought period lie in a positive feedback between soil moisture and precipitation (Schär et al., 2004), inaccurately represented by GCMs, that show contrasting behavior with observations (Taylor et al., 2012a).

With regard to the intense precipitation events, it has been assessed that, in relation to their low resolution, GCMs tend to under-simulate the magnitude of precipitation extremes (Sillmann et al., 2013). More specifically, for the December 2012 event, it must be noted that, while there is no specific literature about it, there are some studies on events of similar nature, occurring in the same period (October 2012), in some cases even passing over the lake area (Ferretti et al., 2014). In these studies, although the diurnal cycle of precipitation has

been adequately represented, an underestimation of extreme precipitation has been noted, by using the numerical weather prediction version of two different GCMs (Lee et al., 2016; Barthlott and Davolio, 2016).

Given the existence of these errors shared by the GCMs that produce the reanalyses, it should however be noted that ERA5 and ERA5L reanalyses perform better than other reanalyses when referring to precipitation in Europe (Hassler and Lauer, 2021) and in other regions of the world (Tang et al., 2020).

Another simulation bias could be attributed to the ERA5L evaporation field. As already mentioned, it coincides with the evaporation from inland water bodies and it strongly depends on the lake cover fraction at the grid point. Such a quantity, even in the case of P1-Lake, is never equal to 1, giving an average evaporation from the lake that is 60% of the total evaporation from the grid point. Thus, this might have resulted in an evaporation deficiency, which could have had an impact on the simulation of the water balance, especially in the drought periods, when the evaporation could have been underestimated.
Finally, the water mass balance does not take into account withdrawals. These are human forcings that, in a closed and shallow basin such as the Trasimeno lake, can be influential. Indeed, the water extractions from the lake are not evenly distributed throughout the year but are supposed to occur mainly in the summer season. In this dry period, there could have been withdrawals from the lake to guarantee water for irrigating the crops in the surrounding area. Even if the lake withdrawals are probably an order of magnitude smaller than the contribution of precipitation and evaporation, their inclusion could have a positive impact on the lake level simulation. Unfortunately, withdrawals, as the groundwater level, are not monitored by local authorities.

2.7.Conclusions

In this study, a model based on data from global atmospheric datasets (reanalysis) and water level measurements is proposed for evaluating the water budget for small, shallow lakes. Precisely, data from ERA5 Land (ERA5L), provided by ECMWF, are used for the case study of the Trasimeno lake in the Umbria region, Central Italy, where level measurements are available since 1991. Reanalysis data have been validated by considering both ground-based and satellite observations. Compared to other possible approaches, what makes the proposed one relevant is the fact that it uses a single source of climate data (i.e., ERA5L), which has the advantage of being spatially consistent. In addition, ERA5L is the only global reanalysis including a lake model (i.e., Fresh-water Lake model, FLake), that provides the lake heat balance and, then it guarantees the reliability of both the evaporation and precipitation fluxes.

The main results of the executed analysis are synthesized below:

- lake surface temperature provided by FLake is in good agreement with the satellite product GloboLakes (i.e., the lake surface water temperature, LSWT) with a Root Mean Square Error RMSE_T of about 1 °C;
- precipitation provided by ERA5L is in good agreement with the ground-based observations, with a Root Mean Square Error RMSE_P of about 0.04 m;
- there is a strong link between the observed variation of the lake level, ΔH, and the values of the water storage, δ, evaluated as the difference between precipitation and evaporation, with a Pearson coefficient larger than 70% and a maximum value of 81% when the grid point located over the lake (P1-lake) is considered;
- the series of the observed ΔH vs. δ, from ERA5L, indicate that a linear link can be assumed with an RMSE_S equal to 0.31 m;
- discrepancies in the linear behavior happen in two periods characterized in Europe by an extreme climate anomaly: the drought 2003-2005 and the intense precipitation period 2013-2016, when reanalysis overestimates and underestimates precipitation, respectively, and the groundwater inflow (not considered in the model).

This last aspect is related to the bias of the GCMs, like IFS, that produce the reanalysis. Although ERA5L represents one of the best-performing reanalyses, the water mass balance modeling needs to be improved to take into account both the higher resolution fluxes and the reproduction of groundwater and surface runoff, which may have influenced the poor simulation of the above-mentioned extreme events.

Even with the mentioned shortcomings, the proposed model allows capturing the monthly variability of the level of the Trasimeno lake, a lake where groundwater in and outflow are in general negligible. As far as we know, this study is among the first (see also Ayzel and Izhitskiy (2019)) to use reanalysis for modeling the water mass balance of a lake. The relevance of this approach lies in the fact that it can be used for other shallow lakes over the globe, where similar conditions to the Trasimeno lake are met. Indeed, before applying this method to other shallow lakes, the potential impact of factors other than precipitation and evaporation needs to be evaluated: inflow and outflow, specifically considering groundwater, lake ice cover, the relation between the level and the surface area, and the catchment use. In fact, Trasimeno lake has a very small basin, compared to its watershed, and therefore the groundwater flow is negligible. This is not true for other shallow lakes, such as Hulun lake (Cao et al., 2021) and Tana lake (Setegn et al., 2011), with a catchment much larger than the basin. In these cases, groundwater flow cannot be considered negligible, nor

runoff (Cao et al., 2021). Moreover, some shallow lakes have tributaries, e.g. rivers, and therefore their influence must be taken into account (Soja et al., 2013; Setegn et al., 2011). Furthermore, many shallow and small lakes are in subpolar regions where the presence of ice cover is not negligible (Gibson, 2002) and this could change the evaporation regime (Wang et al., 2018) (although in FLake such an effect is taken into account). The influence of the morphometry of the lake, which is neglected in FLake, could be relevant for other lakes (Kebede et al., 2006; Li et al., 2019).

The overall performance of the model could improve by including lake withdrawals, recharges, and specifically groundwater inflow. Both the proposed model and the regional services would benefit from constant monitoring of lake withdrawals, groundwater inflow (by a piezometric network), and level observations. Regarding this last aspect, at least one more station should be added on the opposite side of the lake, as in literature (Li et al., 2007). This would allow for increasing the reliability of the lake level observations and considering the effect of the wind on the level measurement. Moreover, the consideration in FLake of the water mass balance and a finer resolution of ERA5 – which both could be achieved in the next few years by ECMWF – will certainly contribute to a better prediction of evaporation, precipitation, and other feedback at the surface, needed for the type of modeling reported in this work. Indeed, the work presented here will be extended in the future by considering – where possible – all the aspects mentioned above.

Beyond the above-discussed shortcomings, the model proposed in this study, based on ERA5L data and lake level measurements, can simulate the behavior over time of the level of the Trasimeno lake. Such a result confirms that as for the simulation of the water table of shallow unconfined aquifers in the Umbria region (Bongioannini Cerlini et al., 2017, 2021), reanalyses are an alternative to other methods for modeling the water mass balance of inland water bodies.

3. Convective Organization in Cloud Resolving Models: The competing effect of radiative and moisture feedback

This chapter[1] is concerned with the comparison of models run to a Radiative Convective Equilibrium (RCE). The goal of the work is to provide insight into convection, as reproduced by different Cloud Resolving Models. The goal of the work, following the suggestion from the Radiative-Convective Equilibrium Model Intercomparison Project (RCEMIP), is to identify key parameters controlling self-aggregation in RCE for both models, to discuss the processes controlled by these parameters and to underline the model's similarities and differences. The two CRMs studied, the SAM (System for Atmospheric Modeling) and the ARPS (Advanced Regional Prediction System), present similar statistics concerning precipitation, but different warming, and drying of the atmosphere, within the spread of the RCEMIP values. The focus has been put particularly on the different sensitivity of the two models on the moisture feedback, due to the different saturation of the sub-cloud layer. The effect of cold pools is what determines the difference between these two models, especially concerning the localization of convection and the weaker effect of the organized state on the

[1] This chapter is a slightly different version of the paper that has been accepted for publication at the *Journal of Advances in Modeling Earth Systems, JAMES*: Cerlini PB, Saraceni M, Silvestri L, Competing effect of radiative and moisture feedback in convective-aggregation states in two CRMs, https://doi.org/10.1029/2022MS003323

average domain statistics in one model compared to the other. The findings of this chapter might be relevant to other RCEMIP CRMs.

3.1.Introduction

Radiative-convective equilibrium (RCE) of an ensemble of clouds has been used as an example of a statistical equilibrium state of the atmosphere able to mimic the tropical part of the climate system (Emanuel et al., 2014). Indeed, given the crucial importance of the balance between radiative cooling and latent heating due to condensation and surface-sensitive heat flux in the Tropics in the Earth's global balance, the RCE state of the atmosphere has become a proxy to study the link between global circulation and convection (Held et al., 1993; Randall et al., 1994; Pauluis and Held, 2002a,b; Yano and Plant, 2012).

The implicit assumption in this is that the tropical circulation may be larger than the small-scale convective processes of an ensemble of clouds and that the statistical equilibrium implied by the RCE state is based on processes in near-equilibrium on different space and time scales with the large-scale flow. This point was brought to general attention when the increased computing capability available made it possible to run three-dimensional high-resolution simulations (Tompkins and Craig, 1998; Bretherton et al., 2005) and to study the sensitivity of RCE states using models with enhanced dimensions of the grid reaching the dimensions of mesoscale processes, with explicit moist variables and different physics parameterizations. Such models can simulate the space-time statistics of an ensemble of clouds (Khairoutdinov and Randall, 2003) over domain sizes with spatial extension up to hundreds of kilometers and for a length of time much longer than that of a single cloud. This happens over homogeneous surface conditions allowing the study of cloud interaction with the spontaneous development of the convective organization (self-aggregation). Despite the differences in parametrizations packages (e.g. microphysics, radiation, turbulence) between models, they showed in some cases spontaneous self-aggregation of clouds (Tompkins and Semie, 2017; Khairoutdinov and Emanuel, 2010; Jeevanjee and Romps, 2013; Ruppert Jr and Hohenegger, 2018; Holloway and Woolnough, 2016; Hohenegger and Stevens, 2016).

The different model representation of cloud physics and convective processes has been recognized as having a relevant impact both on the closeness between modeled self-aggregation and observed atmospheric convective organization and on the modeled climate sensitivity due to self-aggregation feedback (Wing et al., 2020).In fact, within RCEMIP, the analyzed models, one-dimensional SCM, three-dimensional CRM, and GCM were driven to RCE equilibrium, using a predefined setup and initialization process to start from a similar RCE configuration to provide a baseline for their comparison. Nevertheless, given the statistical nature of the RCE equilibrium, which implies the dynamical equilib-

rium of convection (Yano et al., 2012), and given the spread in the RCEMIP models in their reproduction of convection, different results were produced by the models due to their distinct sensitivity to the various climate parameters. Furthermore, RCEMIP underlined that different model responses are linked to differences in models physics and numerics (Wing et al., 2020). Thus, the question remains as to which factors in the models are prevalent in aggregation.

Bretherton et al. (2005) and Muller and Held (2012) found that a low-level circulation from the dry to moist regions, forced by longwave radiative cooling in the lower troposphere, is responsible for self-aggregation, by transporting moist static energy (MSE) up-gradient. Wing and Emanuel (2014), by using an MSE variance budget confirmed that such a mechanism sustains self-aggregation in its intermediate stages. Wing et al. (2017) refer to it as the "indirect longwave feedback", in contrast to the direct diabatic feedback, which is caused by enhanced clear-sky cooling in the dry region relative to the moist regions. Besides radiative feedback, another important role in convective aggregation is played by moisture feedback, which can be related to three correlated factors (Wing et al., 2017): anomalously moist boundary layer, cold pool activity, and entrainment. In particular, Muller and Bony (2015) found that aggregation could be obtained by suppressing rain evaporation, even in the absence of radiative feedback. This mechanism was called "moisture-memory aggregation", where moist regions remain moist, thus more favorable to convection (Tompkins, 2001b; Craig and Mack, 2013). In the following, as usually adopted in literature, we will refer to the moisture-convection feedback as positive "moisture-memory feedback". However, a strong cold pool activity, which in general acts to disperse convection, can be regarded as negative moisture feedback. Such feedback can be considered as a factor changing the organization of convection when, given a moist atmosphere, the efficiency of precipitation is high and less precipitation evaporates (Emanuel, 2021; Held et al., 1993).

The results of RCEMIP and the literature regarding convective organization experiments (Wing and Emanuel, 2014; Yang and Tan, 2020), brought us to hypothesize that there can be different sensitivities to the moisture feedback. In particular, the content of humidity of the mid-troposphere, related to the properties of the cloud and the sub-cloud layers, can change the prevailing factors that bring to self-aggregation. While this has already been proven by suppressing the rain evaporation artificially within the same model (Jeevanjee and Romps, 2013; Muller and Bony, 2015), it might occur within the RCE experiments due to the wide spread of the relative humidity and temperature profiles, observed in the RCEMIP experiment.

For these reasons, this study sets out to compare two models in their reproduction of convection statistics: the Advanced Regional Prediction System (ARPS,

Oklahoma University (OU), (Xue et al., 2000, 2001)) and the System of Atmospheric Modeling (SAM) (Khairoutdinov and Randall, 2003). ARPS is a state-of-the-art reference model from its use in three-dimensional simulations based on a non-hydrostatic formulation of conservation equation for momentum, energy, and water variables used for Numerical Weather Predictions (NWP) (Xue et al., 2014; Sun et al., 2021b). It is recalled here that this model, although very similar to the most used WRF model (Skamarock et al., 2005), was not included in the RCEMIP (Wing et al., 2020). Therefore, this is the first study that investigates self-aggregation with such a model. The SAM model, based on the anelastic approximation, is formulated to conserve the liquid/ice static energy, which is a standard variable to study an ensemble of clouds that is continuously forced in a RCE simulation. Thus, SAM has been extensively used to study convective self-aggregation (Bretherton et al., 2005; Wing and Emanuel, 2014; Muller and Held, 2012; Emanuel et al., 2014).

The aggregation of convection is in fact an indication of the internal oscillation of the model in an RCE configuration, where the boundary conditions are periodic and the lateral energy transport is absent, causing the model to reproduce a statistical oscillation of convection within the system. By reaching the statistical equilibrium of precipitation, one can study the statistical oscillation of convection within two families of models, and their intrinsic process of convective organization. In the SAM model, this oscillation is obtained by conserving the liquid/ice static energy while in the ARPS model, there is a separation of dynamic and thermodynamic variables. The comparison starts from an initial model neutral state, crucial to start the dominant processes on the model domain since how convection is organized depends on the intrinsic characteristics of the models (Wing et al., 2020; Yang and Tan, 2020; Pope et al., 2021).

Furthermore, the results of RCEMIP show that SAM-CRM and the WRF models (WRF-CRM and WRF-COL-CRM) are positioned differently in the RCEMIP spread regarding the degree of aggregation metrics, with WRF models presenting lower values for all aggregation metrics. Indeed ARPS, as mentioned above, is very similar to WRF and specifically to WRF-CRM, because, besides the numerical similarities they share the same radiation scheme, the Goddard Scheme (Chou and Suarez, 1999). Different values of aggregation metrics are a sign of the underlying differences in the feedback roles on self-aggregation shown by the two models.

Thus, the questions posed here, stemming from the main research question, are:

- What are the statistical properties of convection when each of the models reaches a stable state?

- Are there similar aggregation processes in the two models?
- Are there different strengths of the moisture feedback in the two models?

In Section 3.2 the two models, the numerical simulation setup and the initialization are described. In Sections 3.4 and 3.5 the results of the convective organization statistics, the cloud properties, and the convective organization feedback are described and discussed. In Section 3.6 a summary of the work is given.

3.2.Numerical Simulations

3.2.1.The SAM model

The first simulation is performed by using the System of Atmospheric Modeling (SAM), version 6.10.6, (Khairoutdinov and Randall, 2003). SAM solves the anelastic continuity, momentum, and scalar conservation equations. The prognostic thermodynamic variables are the total non-precipitating water ($q_T = q_v + q_c + q_i$ = water vapor + cloud water + cloud ice), the total precipitating water ($q_p = q_r + q_s + q_g$ = rain + snow + graupel) and the liquid/ice static energy, $h_L = c_p T + gz - L_v(q_c + q_r) - L_s(q_i + q_s + q_g)$, with L_v and L_s being the latent heat of vaporization and sublimation, respectively. By definition, h_L is conserved during the moist adiabatic processes (including freezing/melting of precipitation).

Given h_L, q_T and q_p, the mixing ratio of the various hydrometeors (q_c, q_i, q_r, q_s, q_g) is diagnosed by partitioning relationships that depend only on temperature. The diagnosed mixing ratios are used to compute the water sedimentation and hydrometeor conversion rates through a bulk microphysics scheme, where the autoconversion of cloud water into rain is evaluated through the Kessler scheme, while ice aggregation is parametrized similarly to (Lin et al., 1983). Cloud ice is considered non-precipitating water but it is allowed to fall with its own terminal velocity $V_{TI} = 0.4$ m/s (Khairoutdinov and Randall, 2003). Longwave and shortwave radiative fluxes are computed using the radiation code from the National Center for Atmospheric Research (NCAR) Community Atmosphere Model (CAM) version 3.0, (Collins et al., 2006).

We choose a first-order Smagorinsky closure scheme for subgrid-scale (SGS) turbulence. The same SGS parametrization was used in previous studies by (Bretherton et al., 2005; Muller and Held, 2012; Wing and Emanuel, 2014). Surface fluxes are interactively computed according to the Monin-Obukhov similarity theory.

3.2.2.The ARPS model

The second simulation is performed by using the Advanced Regional Prediction System (ARPS), version 5.3.4, (Xue et al., 2000, 2001). ARPS solves the

fully compressible conservation equations for mass, momentum, heat, and water substances (water vapor, liquid, and ice). The thermodynamic prognostic variables are the potential temperature, pressure, and the mixing ratio for six water species (water vapor, q_v, cloud water, q_c, cloud ice, q_i, rain, q_r, snow, q_s and hail,q_h).

Precipitation is computed through a bulk microphysics scheme where autoconversion of cloud water into rain is evaluated through the Kessler scheme (Kessler, 1969) and ice aggregation is treated with the three ice categories (cloud ice, snow, and hail or graupel) scheme of Lin et al. (1983).

The radiation code is adopted from the NASA/Goddard Space Flight Center, with shortwave radiative fluxes based on the model of Chou (1990) and longwave radiative fluxes based on the model of Chou and Suarez (1994). Surface fluxes are computed according to the Monin-Obukhov similarity theory and a first-order Smagorinsky scheme has been chosen for turbulence closure.

Table 3.1: Main properties of the two numerical models and simulations: the model version; the horizontal resolution, Δx; the size of the squared domain; the total running time; the radiation, microphysics, sub-grid scale mixing and surface fluxes parametrizations.

	SAM	ARPS
Model version	6.10.6	5.3.4
Δx (km)	3	3
Domain size (km)	1152	1152
Run time (days)	160	158
Radiation (Fully interactive)	CAM version 3.0 (Collins et al., 2006)	NASA/ Goddard (Chou, 1990; Chou and Suarez, 1994)
Microphysics	Original SAM single-moment (Khairoutdinov and Randall, 2003)	Warm-rain Kessler scheme (Kessler, 1969), Ice Lin scheme (Lin et al., 1983)
Subgrid-scale mixing	First-order Smagorinsky	First-order Smagorinsky
Surface fluxes (Fully interactive)	Monin Obukhov similarity	Monin Obukhov similarity

3.2.3.Numerical setup and initialization

The SAM simulation is performed over a doubly periodic domain with a size of 1152 $\times$ 1152 km^2 and a uniform horizontal resolution of 3 km. We use 64 vertical grid levels with a rigid lid at the top at about 27 km. The first level is at 25 m and grid spacing gradually increases from 50 m near the surface to 500 m above 5 km. Then, it increases again from 500 m to 1 km above 20 km. Newtonian damping is applied to all prognostic variables in the upper third of the model domain (above 18 km). At the bottom, there is an oceanic surface with a constant sea surface temperature of 302 K, which was considered as the lower limit for self-aggregation to happen in Wing and Emanuel (2014). The simulation is run with fully interactive radiation and diurnal cycle as done in Stephens et al. (2008); Ruppert Jr and Hohenegger (2018); Tompkins and Semie (2017). The domain central latitude is set to zero. There is no mean wind and no rotation.

The ARPS simulation has the same horizontal resolution and domain size as the SAM simulation. We use 62 vertical levels with a rigid lid at the top at about 25 km. The first level is at 35 m and grid spacing is 35 m up to 140 m. Then the vertical grid is gradually stretched from about 70 m to about 700 meters up to 20 km. Above 20 km the grid spacing is about 800 m. Rayleigh damping is applied above 19 km. As done for the SAM simulation, ARPS is run with a diurnal cycle and fully interactive radiation. The domain central latitude is set to zero and there are no mean wind and no rotation. The oceanic surface has a constant SST of about 302 K.

The main properties of numerical models and simulations are summarized in Table 3.1. Both simulations run for about 160 days. SAM runs with a time-step of 10 s, while ARPS runs with a time-step of 6 s. Output fields are generated every 6 hours. ARPS and SAM large domain simulations are initialized from the equilibrium sounding of a corresponding smaller domain RCE simulation. Then, convection in the large domain SAM simulations is initiated by adding white noise to the h_L in the lowest five levels, with an amplitude of 0.1 K in the lowest level linearly decreasing to 0.02 K in the fifth level. On the other hand, in large-domain ARPS simulations, convective motions are initiated by applying a random perturbation of magnitude 0.2 K to the potential temperature field over the whole domain.

3.3.Small domain simulations

ARPS and SAM large domain simulations are initialized from the equilibrium sounding of a corresponding smaller domain RCE simulation (see ARPS and SAM profiles in Figures 3.1c and 3.1d). The small domain of ARPS has a size of 100 km x 100 km, almost equal to that of SAM (96 km x 96 km). Both small domain simulations are run for 80 days and they are initialized by using

a sounding obtained from a previous run of the SAM model in RCE equilibrium without self-aggregation. The final equilibrium sounding is obtained by averaging over the last 10 days of the small domain simulations. Averaging over longer time periods, such as 20 or 30 days, made a small difference in the final soundings (not shown) since the equilibrium state was very stable for both models.

The initial conditions of ARPS and SAM large simulations, corresponding to the final equilibrium states of small domain RCE simulations, turned out to be very different, even if the latter started from the same initial sounding. Figure 3.1 shows how the equilibrium state of ARPS is colder (Figure 3.1c, vertical profile of temperature) and drier (Figure 3.1d, vertical profile of water vapor mixing ratio) than that of SAM.

One could ask whether this equilibrium state in ARPS is sensitive to the initial conditions or to the model settings. Therefore, we perform three sensitivity experiments on the small domain simulations: first, we initialize the simulations by using the analytic sounding and the model levels as prescribed in the RCEMIP protocol by Wing et al. (2018) (ARPS-rcemip, SAM-rcemip in Figure 3.1); second, we increase the CO_2 concentration of ARPS from 300 ppmv to 355 ppmv (ARPS-co2) in order to be equal to the standard values of SAM (and closer to the value of 348 ppmv as prescribed by Wing et al. (2018)); third, we add together the changes described in the first and second experiment (ARPS-rcemip-co2). Figure 3.1 shows that the final equilibrium state is similar among all the experiments, both for ARPS and SAM models. Therefore, we can state that the drier and colder equilibrium state of ARPS would have occurred even if we had slightly different initial soundings or higher CO_2 concentrations.

The colder and drier profile of ARPS turns out to be crucial for later stages of convective aggregation. Therefore we briefly introduce here some elements leading to the decrease in temperature and humidity of the ARPS domain. Figure 3.2a shows the non-aggregated state of the small domain simulation after 75 days. Precipitable water (Figure 3.2b) drops very quickly from about 60 mm to 42 mm, while the daily precipitation rate exhibits an opposite behavior by increasing abruptly to about 6.5 mm/day (Figure 3.2b). After a few days of simulation, the small domain is entirely covered by a very thin anvil cloud which remains there until the end of the simulation (Figure 3.2c). Figure 3.2 confirms the presence of a large anvil cloud fraction also on the ARPS-rcemip experiments, thereby confirming the role of anvil clouds in cooling the ARPS domain, as described above.

The average cooling and drying of the ARPS domain are due to the presence of such an anvil which blocks the incoming solar radiation. Such high cloud

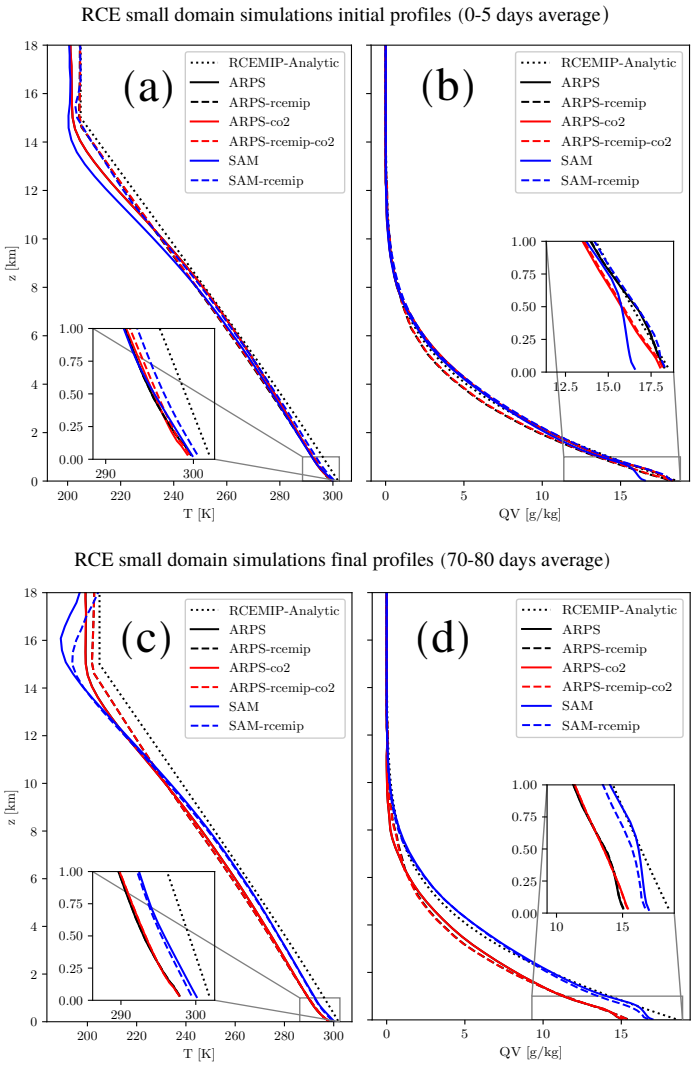

Fig. 3.1: Initial (a,b) and final (c,d) vertical profiles of absolute temperature (T) and water vapor mixing ratio (QV) for all small domain simulations used to initialize the large domain simulations. Each Figure contains also the comparison with the analytic profiles dictated by the RCEMIP protocol, together with results from ARPS and SAM simulations initialized by the RCEMIP Protocol (ARPS-RCE and SAM-RCE, respectively).

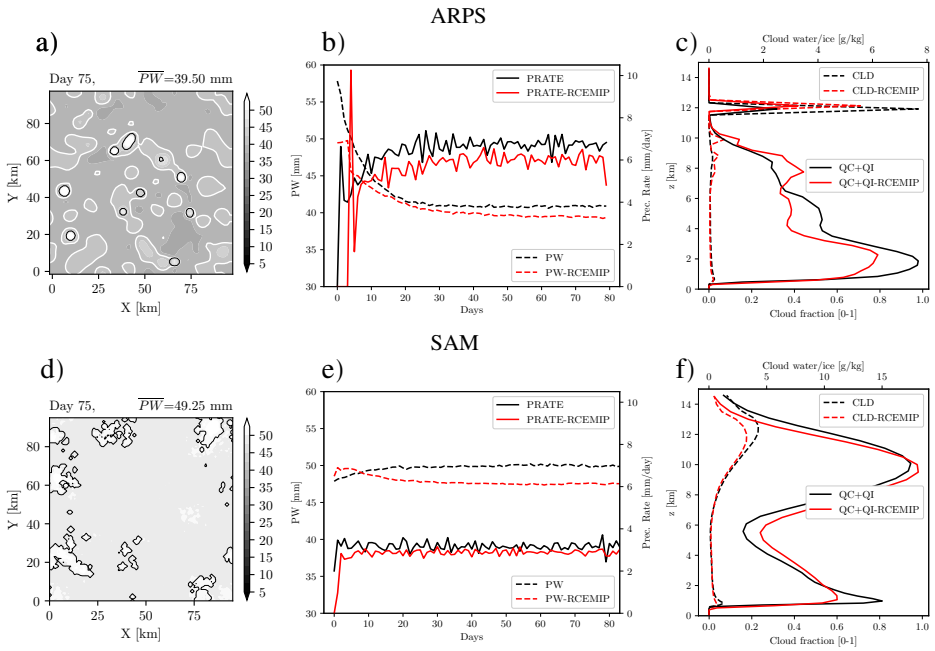

Fig. 3.2: Snapshots of PW at day 75 (midnight) on the small domain used to initialize ARPS simulation (same contours and colors used in Figure 3.3) (a). Time evolution of daily averaged precipitation rate and precipitable water over the ARPS small domain (b). Cloud fraction and total cloud condensate averaged over the last 20 days of the small domain simulations (c). As adopted in RCEMIP (Wing et al., 2020), a cloud is defined according to a threshold value of cloud condensate (10^{-5} kg kg^{-1} or 1% of the saturation mixing ratio over water, whichever is smaller), and a cloud fraction of 1.0 at a particular model level indicates that the entire domain is covered in clouds at that level.

fraction over small domain simulations of RCE has been found also during the RCEMIP project by Wing et al. (2020) (see Figure 9 in the article) and therefore it is not related only to the specific model configuration.

When initializing the large domain ARPS simulation, the cloud water and ice at 12 km, produced by the smaller domain, are removed, removing the large anvil, while leaving its effect on the vertical profile of temperature and water vapor. Therefore convective motions of ARPS start in a drier and colder domain than those in SAM.
The main mechanisms behind the anvil formation in the ARPS small domain rely on the properties of the microphysics scheme adopted by the model, as mentioned in the previous section. Further details are provided in Section 3.5.

3.4.Results

3.4.1.Statistics of convective organization

In SAM and ARPS large domain simulations the precipitation rate reaches a statistical equilibrium, with similar values of 4.2 $mm\ day^{-1}$ in SAM and of 4.1 $mm\ day^{-1}$ in ARPS. The domain average statistics of the simulations final stages are reported in Table 3.2. In SAM, an RCE state is reached, where the total heat flux (THF, sum of the latent heat flux LHF and sensible heat flux SHF) is in balance with net column radiative cooling (R_{NET}), and the LHF, which dominates the THF, is in balance with precipitation (Precip) (see Table 3.2). In ARPS, instead, the net atmospheric energy imbalance, F_{NET}, is greater than in SAM ($F_{NET} = 26\ Wm^{-2}$ for ARPS and $F_{NET} = 7.32\ Wm^{-2}$ for SAM), reaching a value similar to that obtained for the model WRF-CRM in RCEMIP (see Table 3.2).

Both model simulations present the convective organization as it is shown by

Table 3.2: RCE average statistics over the aggregated state (days 115-120) of simulations, following Table A2 of (Wing et al., 2020). The values for the RCEMIP SAM-CRM model, RCEMIP WRF-CRM model, and the average ($\pm$ the standard deviation) of all RCEMIP models are reported in the last three columns for a direct comparison. Such values are directly taken from Table A2 or the text of Wing et al. (2020). $\mathbf{F}_{NET}$ is the atmospheric energy imbalance, that is the magnitude of the difference between R_{NET} and the total surface thermal fluxes; R_{NET} is the column integrated atmospheric radiative forcing (negative values indicates net atmospheric radiative cooling) which is obtained directly by column integration of the radiative forcing (qrad, prognostic variable); LHF and SHF are surface latent and sensible heat (positive values indicates fluxes into the atmosphere); PW is the precipitable water; Precip. is the daily precipitation rate; LWP and IWP are the cloud liquid water path and cloud ice water path respectively. LR is the tropospheric (15 km) Lapse Rate; T_s, RH_s are respectively the absolute temperature and the relative humidity at the lowest model level; Iorg, σ^2_{CRH} and f_{sub} are the three metrics of aggregation used in the RCEMIP experiment.

Var	Unit	SAM	ARPS	RCEMIP-SAM	RCEMIP-WRF	RCEMIP-AVG (STD)
$\mathbf{F}_{NET}$	W m^{-2}	7.32	26	3.87	21.73	4.12 ($\pm$5.66)
R_{NET}	W m^{-2}	-123.03	-102.7	-118.05	-106.54	-110.17 ($\pm$16.08)
LHF	W m^{-2}	122.24	65.7	113.15	90.37	101.93 ($\pm$15.29)
SHF	W m^{-2}	8.1	11	8.77	37.90	11.16 ($\pm$5.74)
PW	mm	27	38.1	31.2	41.2	32.8 ($\pm$4.1)
Precip.	mm day^{-1}	4.3	4.2	3.9	3.1	3.5 ($\pm$0.5)
LWP	mm	0.051	0.015	0.048	0.065	0.041 ($\pm$0.028)
IWP	mm	0.021	0.007	0.025	0.097	0.037 ($\pm$0.038)
LR	K km^{-1}	-6.68	-7.08	-7.2	-6.91	-6.83 ($\pm$0.65)
T_s	K	299.9	298.3	297.9	299.5	n/a
RH_s	%	66	75	69	58	73 (n/a)
Iorg		0.9	0.6	0.91	0.47	0.60 (0.163)
σ^2_{CRH}		0.04	0.02	0.03	0.001	0.027 (0.015)
f_{sub}		0.74	0.66	0.75	0.52	0.69 (0.108)

the Precipitable Water (PW) pattern evolution in Figure 3.3a-3.3h. In both

simulations, moister regions are clustered and present the highest value of total integrated water condensate (see black contours in Figure 3.3a-3.3h) and extreme values of precipitation rate (see Figure 3.3i-3.3p).

As underlined in Yano et al. (2012), such features can be interpreted as a signature of self-organized criticality. The time evolution of the precipitation rate of ARPS and SAM (Figure 3.3i-3.3p), clearly shows a different spatial distribution that corresponds to a more limited precipitation range in intensity in ARPS than in SAM. This is further analyzed in Figure 3.4, which shows the distribution of number occurrence as a function of column water vapor and the precipitation rate.
In fact, Figure 3.4 underlines the simulations similarity and differences. Indeed, on one hand, it points out that in both simulations water vapor presents a homeostasis behavior, confirming the analysis of Yano et al. (2012). The frequency distribution of column water vapor is asymmetric as a function of the corresponding amount of precipitation, and in the critical self-organized state, precipitation is forced into a strongly limited range even with a rapid increase in column water vapor. On the other hand, a dramatic difference between ARPS and SAM stands out from Figure 3.4: a much narrower range of column water vapor in ARPS than in SAM, for a given precipitation.

Convective organization in both SAM and ARPS has an impact, as a whole, on the simulated atmosphere, leading to its warming and drying, as it is reported in Figure 3.5a, 3.5b and 3.5c.
The warming produced by the organization process can be inferred from the mean state profiles of MSE, Temperature (T), and Relative humidity (RH) averaged at equilibrium (between 115-120 days) over the whole domain with respect to the initial ones (averaged between the first 5-10 days). These results are in line with the main results of the RCEMIP (Wing et al., 2020) who found that there is a robustness of the results on heating and drying of the mean state with convective organization among models. The temperature and relative humidity profiles of both SAM and ARPS are within the ensemble spread of RCEMIP mean state profiles (see Figures 7 and 8 of Wing et al. (2020)). However, SAM final state is warmer and drier than the ARPS one (Figure 3.5). This is also evident from the values at the surface shown in Table 3.2, which are near the RCEMIP range values. The surface relative humidity (RH_s) is 66% for SAM and 75% for ARPS and the surface temperature (T_s) is 299.9 K for SAM and 298.4 K for ARPS. As already stated in the previous section, given the different initialization, ARPS starts already with an initial colder profile compared to SAM.

The warmer and drier final state of SAM is reflected also in the values of the surface fluxes (see Table 3.2). Given the smaller RH_s in the SAM model, the

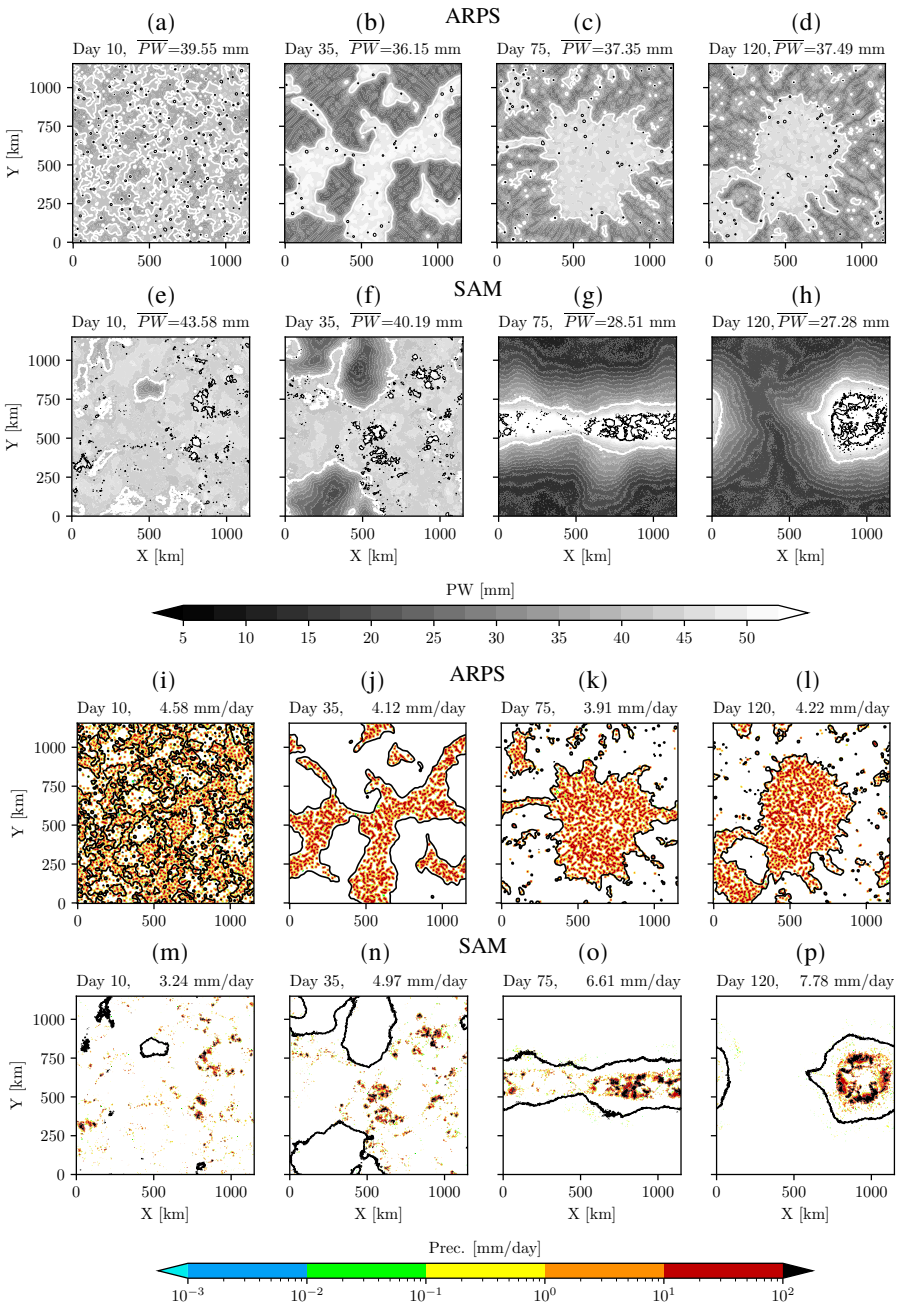

Fig. 3.3: Time evolution of Precipitable Water (PW, filled contours) for ARPS (a,b,c,d) and SAM (e,f,g,h) simulations. The region where aggregation occurs is moister and presents a higher PW (lighter colors). For both models, snapshots are taken at midnight after 10 (a,e), 35 (b,f), 75 (c,g), and 120 (d,h) days. The thick white line represents the boundary between moist and dry patches, taken as PW=40 mm. Black lines are contours of column-integrated water condensate of 1 mm. Time evolution of Precipitation rate (Prec, filled contours) for ARPS (i,j,k,l) and SAM (m,n,o,p) simulations. For both models, snapshots are taken at midnight after 10 (i,m), 35 (j,n), 75 (k,o), and 120 (l,p) days. The thick black line represents the boundary between moist and dry patches, taken as PW=40 mm.

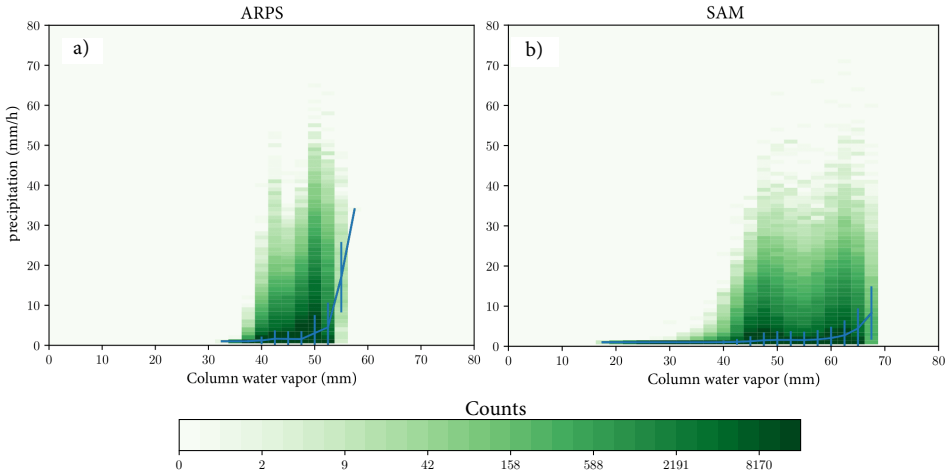

Fig. 3.4: Distribution of number occurrence as a function of column water vapor (horizontal axis, mm) and the precipitation rate (vertical axis, mm/h) for ARPS (a) and for SAM (b). The bin sizes are 2.5 mm for the water species and 1 mm/h for precipitation. The solid curve in each panel shows the mean precipitation rate as defined for each column-integrated water bin.

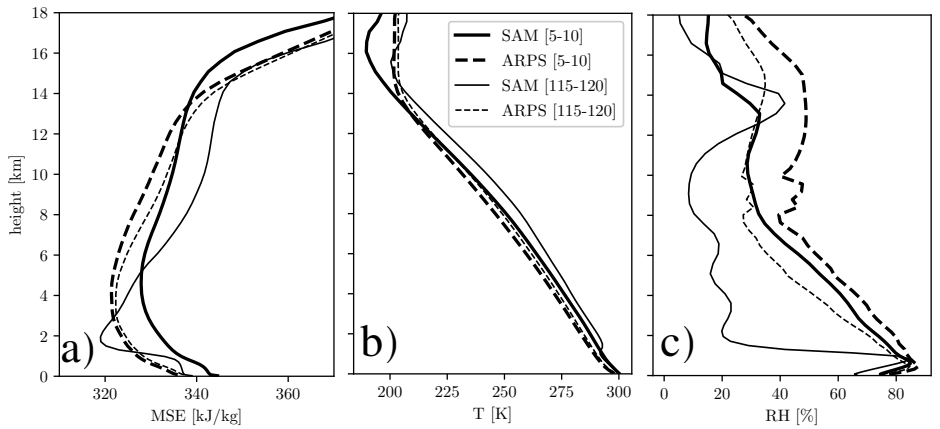

Fig. 3.5: Horizontally averaged profiles of Moist Static Energy (MSE) (a), absolute Temperature (T) (b), and Relative Humidity (RH) (c) for the two simulations, averaged at the initial stage (5-10 days) and at the final aggregated state 115-120 days).

LHFs are larger (LHF $= 122.24\ W\ m^{-2}$) than those of ARPS (LHF $= 65.7\ W\ m^{-2}$). On the other hand, given the smaller T_s of ARPS, the SHFs are larger (SHF $= 11\ W\ m^{-2}$) than those of SAM (SHF $= 9.1\ W\ m^{-2}$).

By looking at the time evolution of PW in the two simulations (Figure 3.6), it

can be noted that in SAM the convective organization leads to a decrease of PW of 10 mm, while the ARPS PW decreases by only 2 mm.

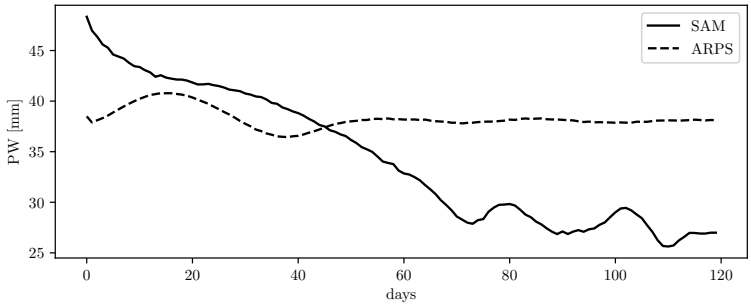

Fig. 3.6: Time evolution of Precipitable Water (PW) for both simulations. Precipitable water is evaluated by considering all condensates.

In SAM such a high decrease is driven mainly by the expansion of dry areas, as shown by the moisture-sorted time series of the Water Vapor Path (WVP) (Figure 3.7). These are computed by dividing the two simulations domain into blocks of equal area (96 km^2), and then sorting them into four quartiles from driest to moistest, based on their daily WVP.

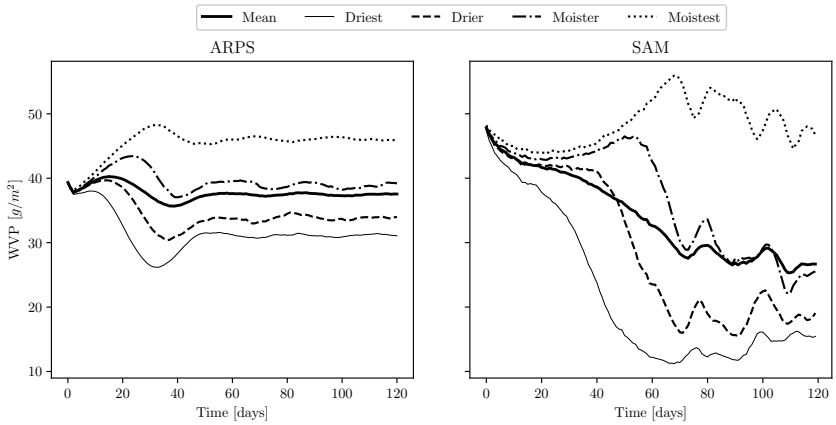

Fig. 3.7: Moisture-sorted time series of the daily averaged Water Vapor Path (WVP) g/m^2 for ARPS (a) and for SAM (b). The thick lines are the domain mean and the other curves are the means over the 96 km^2 blocks sorted into four quartiles based on their daily WVP.

Figure 3.7b shows that in SAM there is a very large inter-quartile difference, driven by a rapid decrease of WVP in the driest quartile which starts suddenly after the beginning of the simulation. On the other hand, moist regions get moister only after 40 days. In ARPS (Figure 3.7a), the interquartile difference is

smaller, since WVP in dry regions does not exhibit a drastic decrease as in SAM. Moreover, moist regions in ARPS get moister before dry regions start to get drier. This striking difference between the two models indicates that different organization feedback are at play. The same concepts can be drawn by looking at the column relative humidity (CRH) evolution which is included in Figure 3.9, for the sake of completeness. While Figure 3.7 shows that SAM simulation starts from a higher value of WVP with respect to ARPS, the average CRH of SAM (Figure 3.9) is close to the ARPS one, indicating that ARPS columns are drier and colder. This may play a significant role in the initial stages of convective aggregation, where the initial convective oscillation in ARPS (Figure 3.3a) shows the presence of many dry regions with similar extent to the moist ones, while in SAM, the strong decrease of PW is concentrated in a few regions (Figure 3.3e).

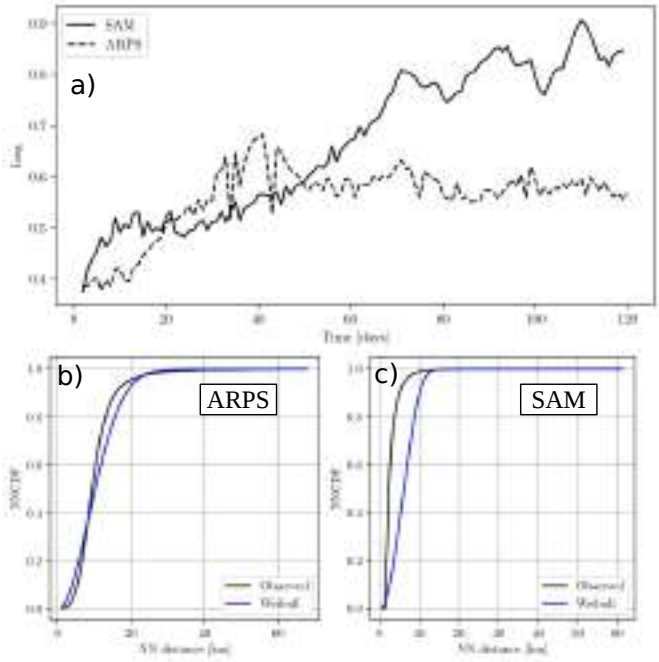

Fig. 3.8: The daily Organization index (Iorg) for SAM and ARPS, as computed by (Tompkins and Semie, 2017) is reported in (a), while the corresponding cumulative density function of the calculated Nearest Neighbor distances (NNCDF) versus nearest neighbor distance of observed and Weibull distribution is displayed for ARPS in (b) and for SAM in (c).

In order to quantify the degree of convective organization we followed the same

approach of (Wing et al., 2020), by computing three different metrics: the Organisation index (Iorg) (Tompkins and Semie, 2017), the subsidence fraction (f_{sub}) (Coppin and Bony, 2015) and the spatial variance of column relative humidity (σ^2_{CRH}) (Wing and Cronin, 2016). Iorg compares the cumulative density

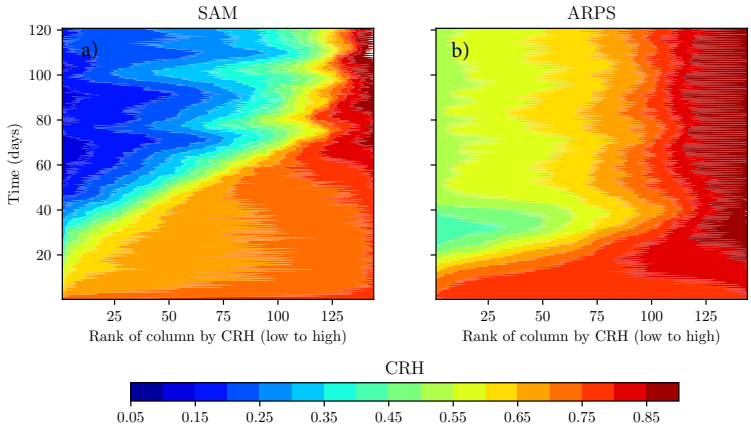

Fig. 3.9: Block averaged daily mean column relative humidity (CRH) for SAM in (a) and for ARPS in (b).

function of the nearest-neighbor distances between convective cores (NNCDF) of the simulated convection to the one expected for a random distribution of the same number of convective cores. It is based on calculating the NNCDF of the simulation updraft grid cells, considering the vertical velocity field, w, at 500 hPa and the threshold between updraft and downdraft at w = 0.5 m/s. Then the NNCDF of the model simulation is compared to the theoretical random distribution of the same cores, the Weibull distribution (NNCDF = $1 - \exp(-\lambda\pi r^2)$, (Weger et al., 1992)), where λ is the number of points per unit area and r is the nearest neighbor distance. Iorg is obtained by calculating the area below the line of the graph of simulated versus theoretical NNCDF. A value of 0.5 or less represents random convection over the domain. The higher the value, the more organized the convection is. It must be noted that the Iorg values of RCEMIP reported in Table 3.2 were computed by using Outgoing Longwave Radiation as opposed to the vertical velocity used here. f_{sub} is the area fraction of the domain where the daily average and block average (96 x 96 km^2) large-scale vertical velocity at 500 hPa is directed downward. Finally, σ^2_{CRH} is self-explanatory.

As can be seen in Table 3.2, SAM has larger values for all the metrics, indicating a more aggregated convection. While the statistics for our SAM simulation are nearly the same as RCEMIP-SAM, the ARPS ones are different from RCEMIP-WRF since the latter model did not attain a state of convective organization.

The lower value of σ^2_{CRH} at the end of the simulation in ARPS confirms the lower interquartile range of WVP between dry and moist regions, while the lower value of f_{sub} at the end of the simulation in ARPS is a clear indication of fewer subsidence regions in the domain compared to SAM. Lastly, the Iorg indicates that SAM undergoes a strongly organized convection, while in ARPS the convection is less aggregated (Figure 3.8a).

This result holds for different thresholds of updraft velocity, ranging from 0.5 to 0.7 m/s (not shown). Indeed, for SAM after 70 days, when the threshold exceeds 0.7 m/s Iorg presents very large oscillations, due to the small amount of selected updraft cores. However, up to 70 days, SAM always presents a greater Iorg than ARPS for the same thresholds. The NNCDF in ARPS (Figure 3.8b) indicates the presence of regular convection at distances less than 10 km, while the clustering occurs at larger spatial scales, up to 60 km. This regular convection is noticeable also in Figure 3.3c and Figure 3.3d, where total column water condensate is also present outside the main aggregated region. A similar distribution was obtained from WRF-RCE simulations (Tompkins and Semie, 2017) and also from satellite observations of tropical convection (Semie and Bony, 2020). This regular convection is absent in SAM (Figure 3.8c), since any new convective core develops only in the already aggregated zone. Indeed, in Figure 3.3h only one cluster is present, made by very small and packed convective structures.

3.4.2. Cloud properties

The difference in the cloud properties in the two models is underlined in Figures 3.10a, 3.10b, 3.10c and 3.10d, which show respectively the radiative heating rate, the cloud fraction, the cloud water and the cloud ice at the initial stage (averaged between 5-10 days) of the considered simulations. As adopted in RCEMIP (Wing et al., 2020), a cloud is defined according to a threshold value of cloud condensate (10^{-5} kg kg^{-1} or 1% of the saturation mixing ratio over water, whichever is smaller), and a cloud fraction of 1.0 at a particular model level indicates that the entire domain is covered in clouds at that level.

The cloud fraction profiles at the initial state are very different among the two simulations, especially regarding the high-level clouds ($>$ 8 km). This is also visible in Figure 3.11, where the anvil clouds evolution is shown. The peak high cloud fraction ("anvil") is very large for ARPS: the ARPS anvil is located between 10 and 15 km and occupies almost the whole domain, reaching an average value of cloud fraction of 0.9. On the other hand, SAM presents fewer anvil clouds (peak cloud fraction around 0.13 at 12 km height), containing a higher amount of ice than those of ARPS (see Figure 3.10d). These features depend on the different ice distributions over the domain, on the cloud definition, and on the different microphysics of the models.

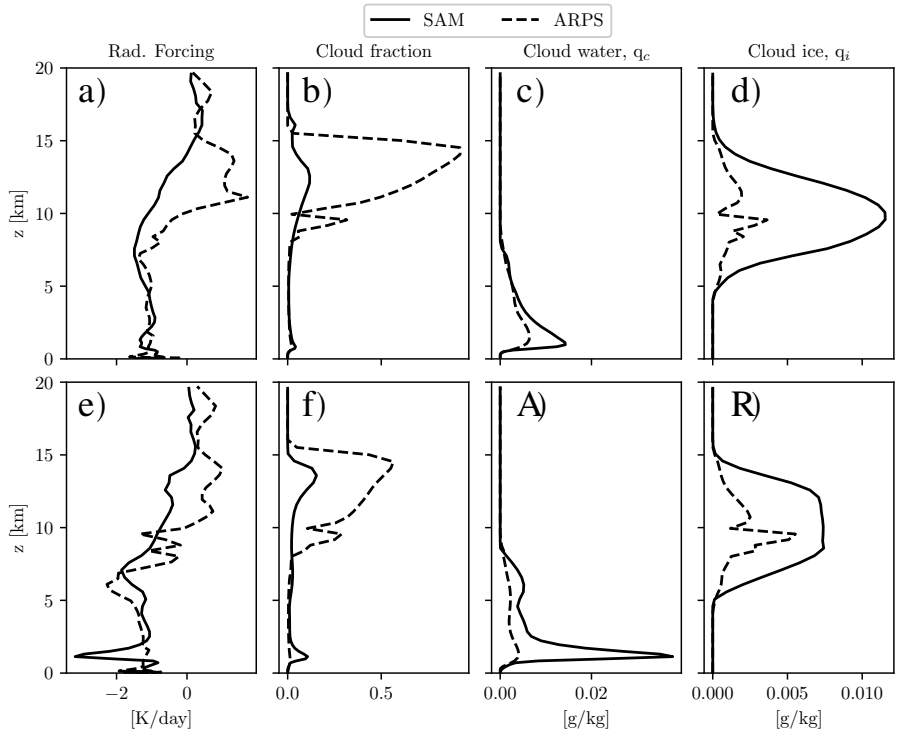

Fig. 3.10: Radiative Heating rate, cloud fraction, cloud water, and cloud ice for the initial state (a,b,c,d) and the final state (e,f,g,h) of the two simulations. The initial (final) state is averaged over the days 5-10 (115-120) days for both simulations.

Firstly, in ARPS, anvil clouds can be classified into two types: lower anvils at the top of deep convective towers (secondary peak cloud fraction at 10 km) which contain a higher amount of ice (more than 10^{-5} kg/kg); high cirrus clouds with a lower amount of ice (less than 10^{-5} kg/kg) that remain after the dissipation of convective systems and occupy most of the domain. Both cloud types are homogeneously distributed over the moister and the drier zones of the domain. In SAM, on the other hand, a high amount of cloud ice is concentrated in a few moister portions of the domain. Therefore, when computing the cloud fraction, the threshold in the cloud definition produces a higher anvil in ARPS due to the small values attained by the cirrus remnants clouds, which are comparable to the threshold of the 1% of saturation mixing ratio and are spread all over the domain. The threshold used in the cloud definition is responsible for producing too much cloud fraction as underlined by Stauffer and Wing (2022). By using a fixed threshold of 10^{-5} kg/kg throughout the atmosphere, the anvil cloud fraction of ARPS reduces, but it remains larger than that of SAM (not shown).

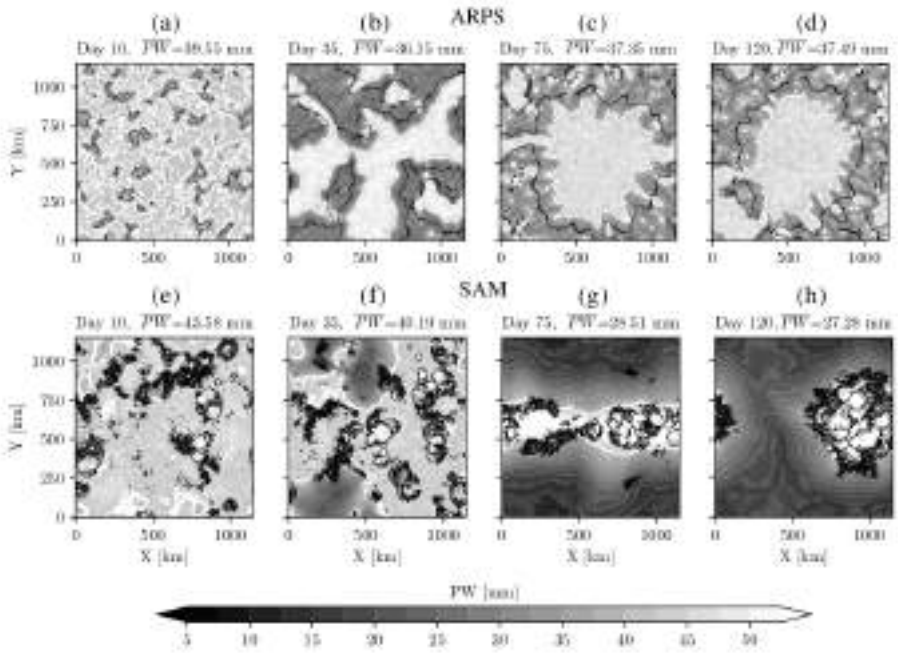

Fig. 3.11: Time evolution of Precipitable Water (PW, filled contours) for ARPS (a,b,c,d) and SAM (e,f,g,h) simulations and total water condensate at height of 12 km. The region where aggregation occurs is moister and presents a higher PW (lighter colors). For both models, snapshots are taken at midnight after 10 (a,e), 35 (b,f), 75 (c,g), and 120 (d,h) days. The thick white line represents the boundary between moist and dry patches, taken as PW=40 mm. Black lines are contours of total water condensate of 0.4 g/kg at a height of 12 km, representing high-level anvil clouds.

Another important difference is in the time evolution of cloud ice, which in SAM it is already present at the first time-steps of the simulation, while in ARPS it is almost null. The reasons behind the huge difference in cloud ice at the initial stages reside in the way ice is created, advected, and transformed into the two models. Firstly, in ARPS, ice is a prognostic quantity and it is directly advected and mixed. The only source of ice is through microphysical processes. In SAM the ice amount is conserved together with cloud water amount and water vapor and is diagnostically obtained from temperature and saturation mixing ratio at each level. Thus, initially, the cloud ice amount in SAM is higher than in ARPS because in ARPS it takes some time for the microphysical processes to produce it, while in SAM it is already present at the second time-step.

The difference in microphysics also plays a role. Khairoutdinov et al. (2022) showed that the single moment microphysics of SAM, as used in this article, underestimates the amount of high cloud. This is also visible from our results, where the high cloud fraction is much less than in the ARPS model, where

different microphysics is used. An important difference in the microphysics parameterization between the two models is the presence of ice sedimentation, which in SAM is permitted, with an ice terminal velocity of 0.4 m/s. In (Khairoutdinov and Randall, 2003) it was verified that the presence of ice sedimentation in SAM leads to a reduced anvil above 9km. Thus, in the absence of ice sedimentation, the anvil in SAM would have been more extensive.

The different radiative heating profiles between ARPS and SAM, as shown in Figure 3.10a, demonstrate that the above-discussed differences in high-level cloud properties are relative to radiatively active clouds and not only to clouds artificially generated by the chosen cloud definition. The presence of a thick anvil in ARPS heats the upper troposphere, while it cools the middle and lower troposphere (Figure 3.10a). In SAM there is cooling in all troposphere except for the top of the anvil at 15 km. This is because ice cirrus clouds act to reflect incoming shortwave radiation and entrap long-wave radiation from the clouds below (Liou, 1986; Schlimme et al., 2005). At the beginning (0-5 days), both models present similar radiative heating rate profiles between moist and dry regions (not shown).

However, after day 10 the situation has changed in ARPS, especially in the mid-level troposphere, where strong differential heating is developed between the dry and moist regions (see Figure 3.12a). Deep clouds are mostly concentrated in moist regions causing a strong reduction of longwave cooling. This direct diabatic effect of high-level clouds is known to be a strong longwave positive feedback for the maintenance of convective aggregation (Wing et al., 2017). However, in our simulation, this effect could be important also in the triggering phase of aggregation, since it is visible from the beginning of the ARPS simulation. The same cannot be said for the SAM model, where the main differential radiative heating is driven by low-level clouds after 20-25 days (see Figure 3.12b). Indeed at early time-steps, at the top of the boundary layer, SAM presents a simulated cloud fraction slightly higher than ARPS (see also Figure 3.13), with a low-level cloud cover fraction > 0.1 and a higher content of q_c ($q_c = 0.015$ g/kg for SAM and $q_c = 0.009$ g/kg for ARPS). The presence of low clouds in SAM is known to be a necessary factor for convective organization (Muller and Held, 2012; Wing and Emanuel, 2014; Muller and Bony, 2015).

With the convective organization, the anvil cloud fraction is greatly reduced both in ARPS and SAM as is shown in Figure 3.10f. This may be a signal of the Iris effect presence in both models. However, to properly assess this phenomenon, additional analyses would have been needed, going beyond the scope of the article. At the same time, the cloud ice decreases in SAM, almost by half, while it increases in ARPS (Figure 3.10h). Such increase in ARPS is concentrated in the moist regions, where cloud ice reaches values of 0.020

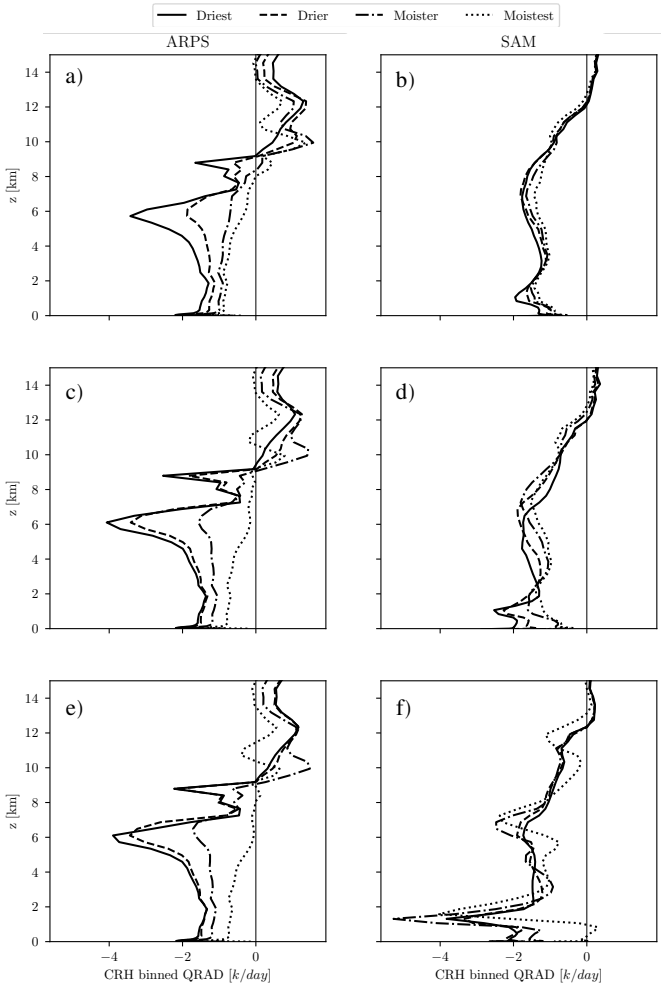

Fig. 3.12: Average values of the CRH blocks-quartiles binned of the radiative heating (QRAD) between the 20th and 25th day for ARPS (a) and SAM (b), between the 50th and 55th day for ARPS (c) and SAM (d) and between the 115th and 120th day for ARPS (e) and SAM (f).

g/kg (not shown), comparable to those reached by SAM at the initial stages. Such an increase in anvil cloud ice amplifies the differential radiative heating between moist and dry regions at later stages of aggregation (see Figure 3.12c and 3.12e).

Thicker anvils are more efficient in blocking the removal of heat in the convective region, with respect to dry regions. Thus, a larger amount of heat

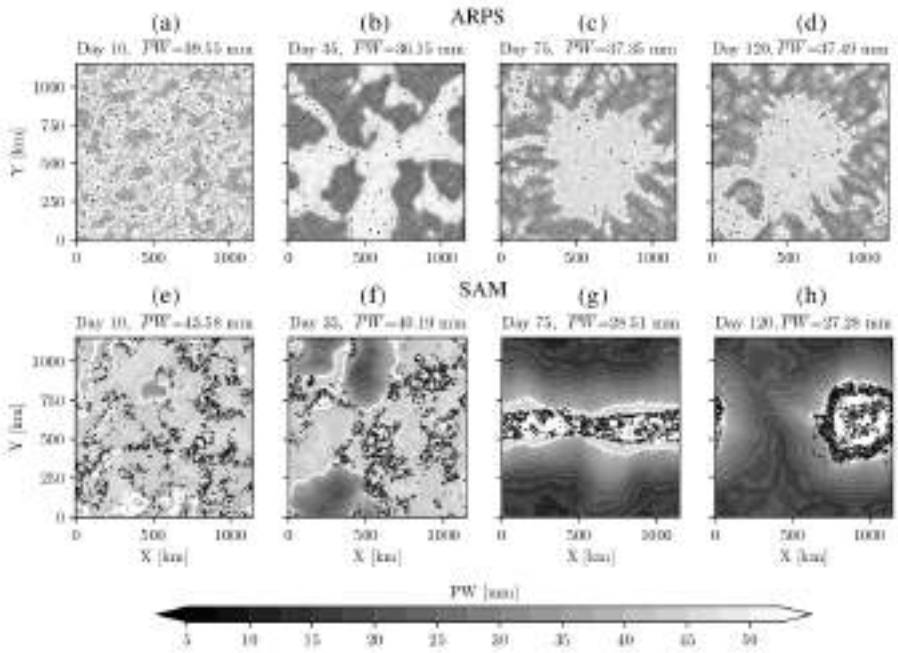

Fig. 3.13: Time evolution of Precipitable Water (PW, filled contours) for ARPS (a,b,c,d) and SAM (e,f,g,h) simulations and total water condensate at height of 1.5 km. The region where aggregation occurs is moister and presents a higher PW (lighter colors). For both models, snapshots are taken on midnight after 10 (a,e), 35 (b,f), 75 (c,g) and 120 (d,h) days. The thick white line represents the boundary between moist and dry patches, taken as PW=40 mm. Black lines are contours of total water condensate of 0.4 g/kg at a height of 1.5 km, representing low-level clouds.

must be transported to the dry regions and radiated out to space (Wing et al., 2017; Yang and Tan, 2020). On the other hand, in SAM cloud ice decreases throughout the whole domain. This is probably due to the large warming of the whole troposphere, which decreases the amount of cloud ice while increasing the amount of cloud water. Figure 3.10g shows a drastic increase of low-level cloud fraction and q_c in SAM, reaching an equilibrium value of 0.04 g/kg after aggregation.

The correspondent radiative forcing in SAM is a pronounced cooling at 2 km of almost -4 K/day. This marked cooling is absent in ARPS simulation because low clouds are too few and the cloud water is low. Therefore at intermediate and final stages, in SAM, the main differential radiative heating rate is observed at low-level and it is generated by a strong cooling at the top of low-level clouds in the driest regions (see Figure 3.12d and Figure 3.12f).

3.4.3. Convective organization feedback

Another signature that characterizes convective organization in both models is the increase of the spatial variance of the vertically integrated Frozen Moist Static Energy (FMSE) (see Figure 3.14a). This increase is exponential in the early stages of the simulations, however, it presents different time scales of aggregation and different magnitudes. ARPS reaches a steady state after 40 days with an increase that is almost one order of magnitude lower than SAM. In SAM the organization process is slower, reaching its steady state after 60 days. In order to quantify the feedback mechanisms leading to such an increase, we computed the budget of the vertically integrated FMSE, following Wing and Emanuel (2014). The FMSE, h, is conserved during moist adiabatic processes and is given by:

$$h = c_p T + gz + L_v q_v - L_f q_i \tag{3.1}$$

Where c_P is the specific heat of dry air, T is the air temperature, g is gravitational acceleration, z is the height above the surface, L_v is the latent heat of vaporization, q_v is the specific humidity with respect to water vapor, L_f is the latent heat of fusion, and q_i is the specific humidity with respect to ice condensates. The budget equation for the mass-weighted vertical integral of *h*, $\widehat{h}$, is:

$$\frac{\partial \widehat{h}}{\partial t} = THF + RAD - \nabla_h \cdot \widehat{\vec{\mathbf{u}} h} \tag{3.2}$$

where THF is the total surface heat flux (the sum of the latent and the sensible heat flux), RAD is the column-integrated net radiative heating, and $\nabla_h \cdot \widehat{\vec{\mathbf{u}} h}$ is the horizontal divergence of the vertically integrated flux of h. For a quantity x, we denote "$\widehat{x}$" as its density-weighted vertical integral $\int_0^{z_{top}} x\rho dz$.

The horizontal domain mean of x is instead denoted as $\{x\}$ and the anomaly as x'. Following (Wing and Emanuel, 2014) to get to their Eq. (9), the horizontal mean of Eq. (3.2) is subtracted from the total budget and multiplied by $\widehat{h}'$ to get the budget for the horizontal spatial variance of $\widehat{h}$:

$$\frac{1}{2}\frac{\partial \widehat{h}'^2}{\partial t} = THF'\widehat{h}' + RAD'\widehat{h}' - \widehat{h}'\nabla_h \cdot \widehat{\vec{\mathbf{u}} h} \tag{3.3}$$

The terms on the right-hand side of Eq. (3.3) are the spatial covariances between $\widehat{h}$ and the budget terms of Eq. (3.2). The last term, due to advective processes or convergence of h, is computed as a residual from the remaining budget, as in Wing and Emanuel (2014). The analysis of these covariance terms determines how each process of the FMSE budget equation contributes to the observed changes of the FMSE variance, $\widehat{h}'^2$, hence to changes in the convective organization inside the models. All the budget terms are computed as six-hourly mean fields, with the convergence term computed as a residual,

then the daily and domain mean is computed and it is normalized by the factor $\{\widehat{h}'^2\}$. The outcome of this process for both models is shown in Figure 3.14b for ARPS and Figure 3.14c for SAM.

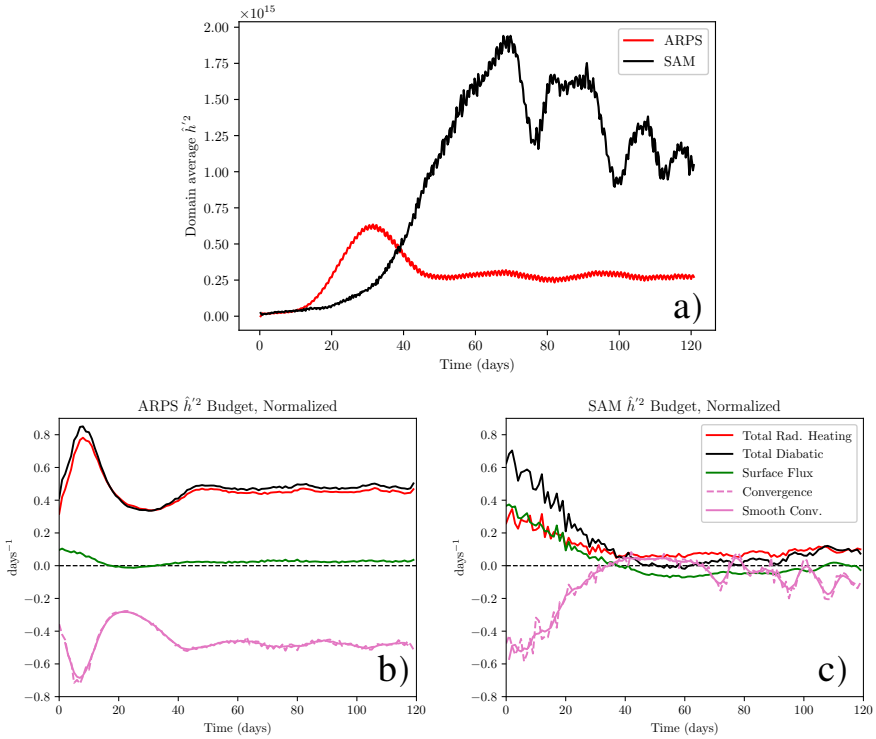

Fig. 3.14: a) Time evolution of the 6-hourly domain mean spatial variance of the vertical-integrated Frozen Moist Static Energy (FMSE) (J^2/m^4) for SAM in black and for ARPS in red. b) and c) Time evolution of the domain mean terms in the FMSE budget, each normalized by $\{h'^2\}$, with units of day^{-1}, for ARPS and SAM, respectively. The terms of the budget are the sum of all diabatic correlation terms (black), the sum of the longwave and shortwave fluxes correlation (red), the surface enthalpy flux correlation (green), and the horizontal convergence of flux of FMSE (pink dashed). A 5-day running average is applied to the horizontal convergence term to provide a smoothed version (solid pink). The black dashed line is the zero line, plotted as a reference.

In SAM for the first 40 days, the total diabatic term is positive, with both surface fluxes (latent and sensible) and radiative fluxes (shortwave and longwave) contributing to the increase of $\widehat{h}'^2$, thus to the organization. The convergence term, instead, is negative, representing negative feedback on the organization. After the 40th day, the convergence term becomes positive, acting as the main

source of organization maintenance, when the total diabatic term tends toward zero. The convergence term is related to what is found in Muller and Held (2012), a mesoscale circulation that transports moist static energy up-gradient. At later stages of the simulation, the convective organization is maintained by the radiative heating feedback. These results resemble the one obtained for SAM by Wing and Emanuel (2014).

However, the same can not be said for ARPS. Indeed, in ARPS the total diabatic term is always highly positive, with the total radiative heating being dominant on the surface fluxes. The convergence term is always negative. This means that while radiative feedback help increases $\widehat{h}'^2$, the convergence term is always acting as damping on the convective organization. Interestingly, these results point out that ARPS behaves differently than SAM. Similar results were obtained in Wing and Cronin (2016) in their colder long channel simulations, where the bulk circulation never transports moist static energy up-gradient and the radiative feedback are helping maintain organization.

The radiative feedback which initially helps convective aggregation in SAM involves especially low clouds, while in ARPS is due to the presence of high clouds, as pointed out in the previous Section. In SAM the radiative feedback, predominant at the beginning of the simulation, is firstly due to higher longwave cooling in the dry regions compared to the moist regions, and only secondly by the shortwave heating in the moist regions compared to the dry regions. This is similar to what is found in Wing and Emanuel (2014). After the initial stages of simulation, the longwave radiation exerts an indirect feedback, where the longwave cooling in dry regions generates a shallow circulation between dry and moist regions (positive convergence term, pink line in Figure 3.14c). In ARPS instead, the initial presence of a deck of high clouds exerts positive longwave feedback, primarily because the column longwave cooling is strongly reduced by the longwave opacity and low temperature of high clouds (Wing et al., 2017), as demonstrated by Wing and Cronin (2016). This is inferred, by Figure 3.12a for ARPS differential cooling in the inter-quartile range. This feedback is maintained throughout the whole simulation (see Figure 3.12 at later stages in ARPS).

Regarding surface fluxes, in SAM the feedback is stronger than in ARPS (Figure 3.14), and it is comparable in magnitude to the effect of radiative feedback. For both models the surface fluxes are positively contributing to the organization at early stages, moistening the moist regions. At later stages in SAM, the feedback becomes negative, due to air-sea enthalpy disequilibrium (Wing and Emanuel, 2014), while in ARPS it remains weakly positive.

The mesoscale circulation that develops with the organization in SAM can be

visualized using the stream function Ψ (Bretherton et al., 2005), derived as the horizontal integral over vertical velocity starting from the driest column to the moistest, after having sorted them from lowest to highest CRH. The same sorting described in the previous section 3.1 is applied here, but in this case, it is done based on the CRH. By looking at the advective tendencies of MSE, implied by the stream function, one can capture the general mechanism of energy exchange between the columns.

In SAM, the MSE "circulation" occurs between the moist and dry columns only after the 40th day as underlined by the $\widehat{h}'^2$ budget in Figure 3.14. By day 50, SAM is attaining a state of convective organization. An up-gradient transport of MSE is implied by Figure 3.15b, with the low MSE being accumulated in the dry columns. There is a low-level circulation in SAM, which transports low-level high MSE air into the moist regions (at around 1 km), and mid-level low MSE air into the dry region (at around 2 km). This results in a net reduction of MSE in the dry regions, and an increase of MSE in the moist regions, thereby enhancing the gradient of $\widehat{h}$. This happens in correspondence with the presence of a deck of low clouds (Figure 3.15b) in dry regions, as already pointed out in the previous Section. This is also visible in Figure 3.15b where the low clouds in SAM reach more than half of the sorted block throughout the whole simulation (Figure 3.16).

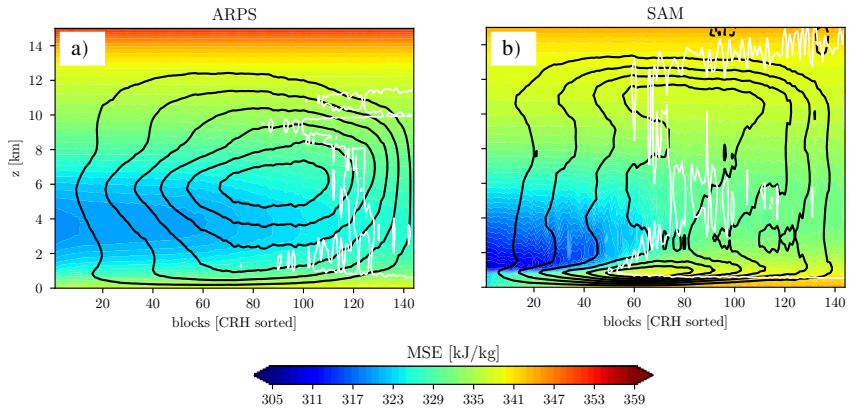

Fig. 3.15: The average value of the MSE "circulation" between 50-55 days for ARPS (a) and SAM (b), with columns ranked by Column Relative Humidity (CRH), from driest to moistest. Black contours show the stream function Ψ (contour interval 0.05 $kg\ m^{-2}s^{-1}$, starting at 0.01 $kg\ m^{-2}s^{-1}$, solid for positive values and dashed for negative values) as a function of CRH and height. White contours show cloud condensate (cloud ice and cloud water) $q_N = q_i + q_c$ (contour interval 0.005 g/kg, starting at 0.001 mg/kg). Shaded contours represent MSE.

In ARPS, on the other hand, the "circulation" of MSE is not as strong and bottom-heavy as the one in SAM, both in the days before the organization (as shown in Figure 3.16b) or once the simulation has reached the organized equilibrium (as shown in Figure 3.16d). Therefore, Figure 3.15a suggests that the low-level circulation of MSE, which leads to the organization in the SAM model, is absent in the ARPS model. As mentioned above, the $\widehat{h}'^2$ budget demonstrates that the convergence term is always damping the organization. This is also visible from the CRH sorted mass flux (taken as $M = \rho w$) of ARPS (Figure 3.18d and Figure 3.18f), where the updrafts in the moist regions, reach their maximum below 8 km, and not at low level, as for the SAM model.

The absence of low clouds in ARPS correlates with the absence of low-level cooling in dry regions, which can be seen in the radiative heating quartiles (Figure 3.12). Muller and Bony (2015) pointed out that the vertical structure of the radiative cooling/heating is relevant to the convective organization since it determines the strength and height of the low-level circulation and hence the associated MSE transport. Since the main differential radiative heating in ARPS is found at a high level, this points out that the organization is primarily due to the high cloud radiative effects, as also demonstrated by Coppin and Bony (2015). The fact that the low-level circulation never appears in ARPS is also demonstrated by looking at the bottom layer wind speed at the boundary between the moist and the dry regions (see Figure 3.17). While in SAM such velocity increases (which is a clear signal of a radiatively driven aggregation as shown by Windmiller and Craig (2019), see their Figure 8), in ARPS it remains almost constant.

While the thicker layer of anvil clouds and the absence of low-level cloud can explain the different radiative feedback triggering organization, they can not explain why ARPS presents a faster aggregation process compared to SAM, as clearly seen from the increase of $\widehat{h}'^2$ (Figure 3.14a). In the literature, (Jeevanjee and Romps, 2013; Muller and Bony, 2015; Holloway and Woolnough, 2016), a faster aggregation is usually found by suppressing rain evaporation and therefore by weakening evaporation-driven downdraft and cold pools. This feature, which is not realistic in standard conditions, may occur when the sub-cloud layer is nearly saturated and rain cannot evaporate (Wing et al., 2017). Cold pools slow down the organization process by increasing low-level mixing between moist and dry regions (Jeevanjee and Romps, 2013), killing convective cells within a few hours (Muller and Bony, 2015). As a consequence, when cold pools are weaker, even in the absence of radiative feedback, convection remains active in the same moist regions throughout the simulation, giving rise to a "moisture-memory aggregation".

The low-level mixing by cold pools is evident in SAM since the beginning of the

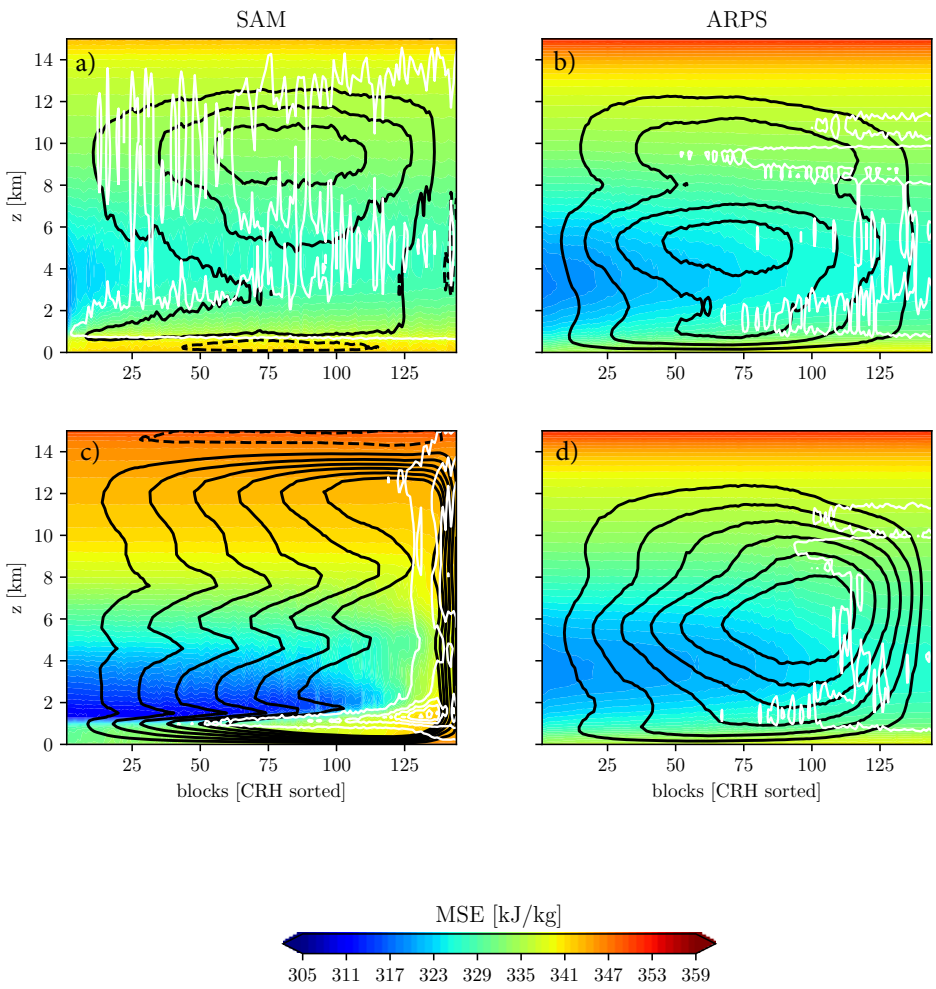

Fig. 3.16: The average value of the MSE circulation between 20-25 days for SAM (a) and ARPS (b) and between 115-120 days for SAM (c) and ARPS (d), with columns ranked by Column Relative Humidity (CRH), from driest to moistest. Black contours show the stream function Ψ (contour interval 0.05 $kg\ m^{-2}s^{-1}$, starting at 0.01 $kg\ m^{-2}s^{-1}$, solid for positive values and dashed for negative values) as a function of CRH and height. White contours show cloud condensate (cloud ice and cloud water) $q_N = q_i + q_c$ (contour interval 0.005 g/kg, starting at 0.001 mg/kg). Shaded contours represent the FMSE.

simulation and up to 30 days: Figure 3.16a shows the presence of a circulation in the sub-cloud layer (below 1 km) which acts to transport MSE from moist regions to dry regions, therefore slowing down the aggregation process. Such boundary layer moistening in the dry regions is also confirmed by the negative domain-average mass flux in the moist regions below 1 km (see Figure 3.18a). The same cannot be said for ARPS, where these signs of boundary layer moist-

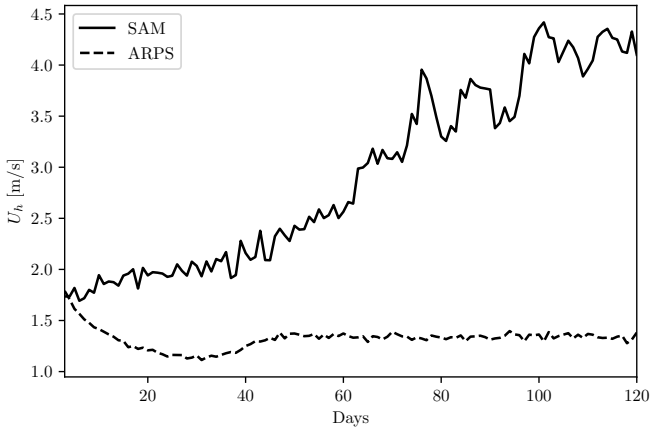

Fig. 3.17: Time evolution of bottom layer horizontal wind speed (U_h) averaged over the boundary between moist and dry patch (identified with the same criterion as in Figure 3.3).

ening in the dry regions never appear. The absence of these features must be found in the different saturation of the sub-cloud layer. Figures 3.19a and 3.19b show that ARPS has a saturated sub-cloud layer both at the intermediate stages (not shown) and at the end (Figure 3.19a) of the simulation. Instead, SAM never reaches such a condition (Figure 3.19b). The saturation of the sub-cloud layer in ARPS directly influences the downdrafts properties: ARPS downdrafts are weaker than those in SAM and they cover a smaller fraction of the domain (see Figure 3.20).

It is underlined here that, the sign of the localized convective activity in ARPS, mentioned in previous sections, corresponds to the lack of the creation at the surface of the radiatively-driven cold pool as in SAM and to the presence of many dry patches, reaching a maximum extension when the sub-cloud layer is fully saturated (see the behavior of the driest quartile of ARPS in Figure 3.7a in comparison with Figure 3.14b at day 35).

Another signature of the role of moisture-convection feedback in ARPS is visible in the strong positive humidity anomalies which develop in the moist regions (Figure 3.21a). Such anomaly is always present, from the beginning to the final stages of the simulation (see Figure 3.22b, 3.22d and 3.22f). In SAM, instead, at the beginning of the simulation, only negative anomalies are present outside the boundary layer, demonstrating that the main feedback occurs in the driest regions of the domain. Positive anomalies are present only inside the boundary layer of dry regions at the initial stages (Figure 3.22a), while they develop in moist regions, only at later stages (see Figure 3.22c and 3.22e).

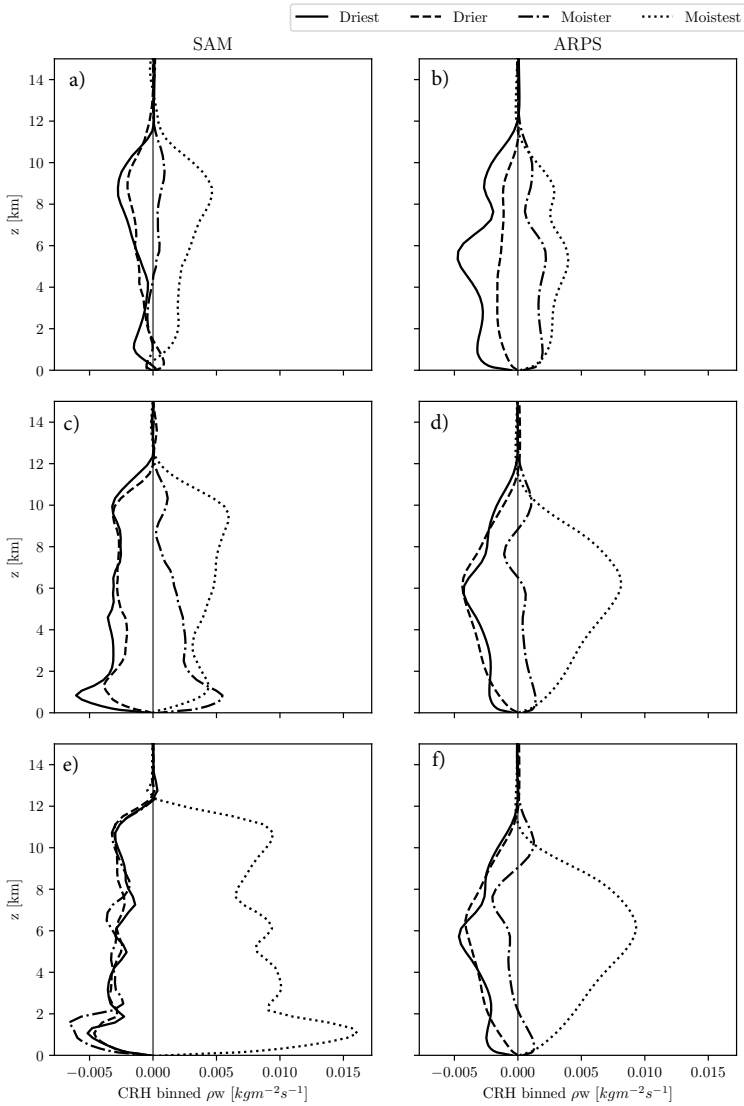

Fig. 3.18: Average values of the CRH blocks-quartiles binned of the mass-weighted vertical velocity between the 20th and 25th day for SAM (a) and ARPS (b), between the 50th and 55th day for SAM (c) and ARPS (d) and between the 115th and 120th day for SAM (e) and ARPS (f).

The absence of a low-level circulation in ARPS, which exports high-MSE from dry regions to moist regions, prevents the formation of extremely dry regions as in SAM.

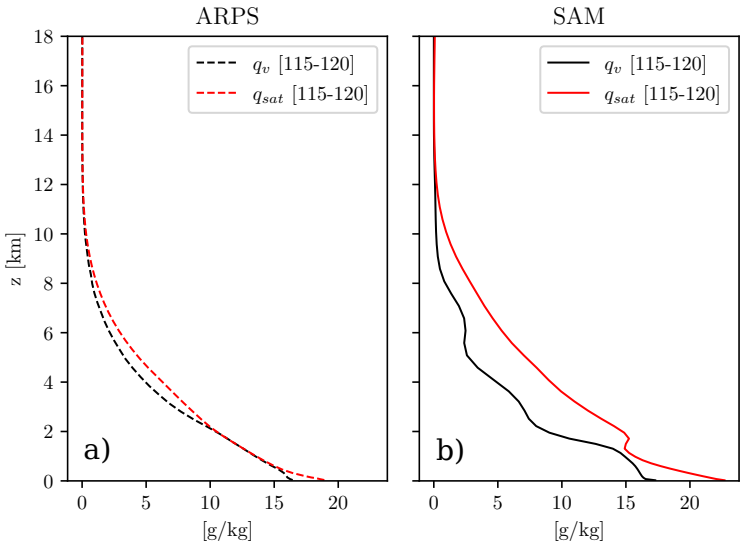

Fig. 3.19: Vertical profiles of ARPS saturation water vapor mixing ratio (q_{sat}) and water vapor mixing ratio (q_v) over cloudy grid points (defined as grid points at 1.5 km height where the total cloud water condensate exceeded 10^{-3} g/kg) between days 115 and 120 a); b) same as a) but for SAM simulation.

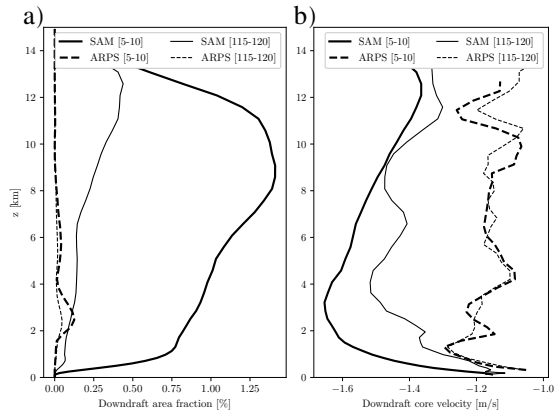

Fig. 3.20: Vertical profiles of domain area fraction covered by downdrafts (a); downdrafts core velocity (b). Thick lines denote the initial state (5-10 days average), while thin lines denote the aggregated state (115-120) days average). Continuous lines are SAM simulations, while dashed lines are ARPS simulations. Downdrafts cores are taken as grid points where the vertical velocity is less than -1 m/s.

3.5.Discussion

In general, cold pool acts to disperse convection (Zuidema et al., 2017). A saturated sub-cloud layer causes less rain evaporation, weaker downdrafts, and weaker cold pools (Wing et al., 2017). Weaker cold pools decrease the low-

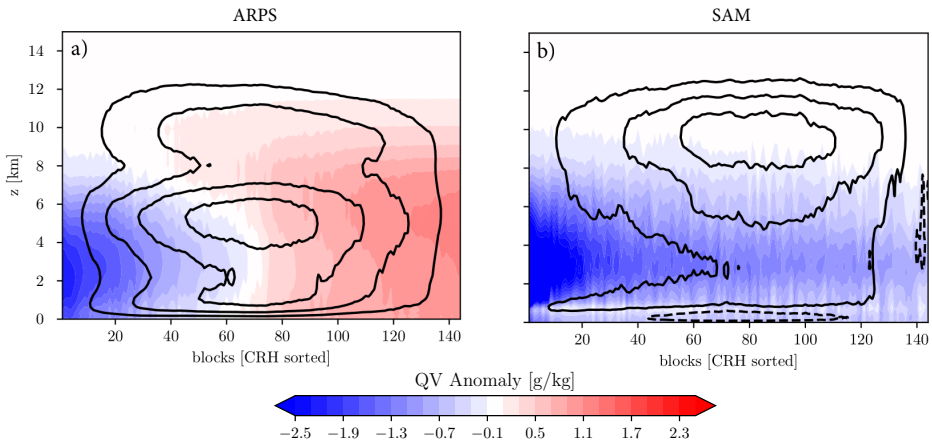

Fig. 3.21: Specific humidity (qv) anomaly from the initial day of simulation (qv_0) in ARPS (a) and in SAM (b) between the 20th and the 25th day of the simulation, with columns ranked by Column Relative Humidity (CRH), from driest to moistest. Black contours show the stream function Ψ (contour interval 0.05 $kg\ m^{-2}s^{-1}$, starting at 0.01 $kg\ m^{-2}s^{-1}$, solid for positive values and dashed for negative values) as a function of CRH and height.

level mixing between moist convective regions and dry regions (Jeevanjee and Romps, 2013). This low-level mixing is important for the moistening of the boundary layer and for the triggering of new convection in dry regions, where radiative subsidence is predominant. This first cold pool effect can be regarded as a negative moisture feedback to convective organization, since it produces a down-gradient transport of moisture in dry regions, therefore slowing down convective aggregation. However, if cold pools (in moist regions) and radiative subsidence (in dry regions) are strong enough to generate low clouds at the boundary between moist and dry regions (Tompkins, 2001a), then an up-gradient low-level transport of MSE from dry regions to moist regions develop by an indirect positive longwave feedback (Muller and Held, 2012). This circulation enforces the already existent radiative subsidence. Therefore, as a secondary effect, evaporatively-driven cold pools can help the expansion of radiatively-driven cold pools ("dry pools") by triggering low clouds in dry regions.

Such competing effects of moisture and radiative feedback are clearly visible in the SAM model in the early stages of aggregation (Figure 3.23b). The up-gradient transport, then amplifies the already existent direct diabatic longwave feedback (Wing et al., 2017), helping to maintain convective organization. Instead, in ARPS, the higher saturation of the sub-cloud layer weakens the cold pool effect and, as a consequence, this low-level mixing is weaker. Therefore, this produces a circulation of dry air from dry regions to moist regions, driven

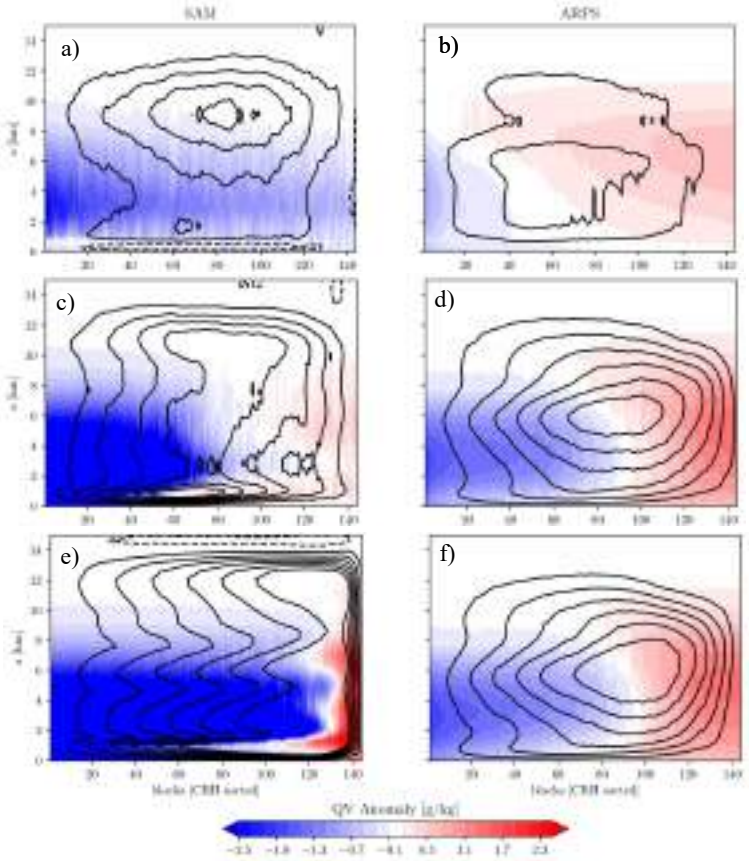

Fig. 3.22: Specific humidity (qv) anomaly from the initial day of simulation (qv_0) in SAM and in in ARPS between the 10th and the 15th day (SAM in (a) and ARPS in (b)), between the 50th and the 55th day of simulation (SAM in (c) and ARPS in (d)) and between the 115th and the 120th day (SAM in (e) and ARPS in (f)), with columns ranked by Column Relative Humidity (CRH), from driest to moistest. Black contours show the stream function Ψ (contour interval 0.05 $kg\ m^{-2}s^{-1}$, starting at 0.01 $kg\ m^{-2}s^{-1}$, solid for positive values and dashed for negative values) as a function of CRH and height.

by the presence of a direct diabatic longwave feedback, which is stronger compared with that of SAM also due to the presence of a high anvil cloud fraction. This is shown in the early stages of ARPS simulation (Figure 3.23a).

In SAM, while cold pools in the moist regions tend to disperse convection, dry regions start to expand due to a positive indirect longwave feedback (longwave cooling in dry regions drives a shallow circulation between the dry and moist regions and this circulation transports MSE up-gradient (Wing et al., 2017)) which is sustained by low clouds. The MSE export from the moist boundary

layer in the dry region causes strong drying and further increase of outgoing longwave radiation (Figure 3.23d for SAM at intermediate stages). Therefore, dry regions expand until the point in which they lose all their water vapor (also visible in the qv anomaly plot at later stages of the simulation in Figures 3.22c and 3.22e). Dry regions expansion seclude all moist convection in a single region since evaporatively-driven cold pools are not strong enough to counteract the radiatively-driven cold pool once it has strengthened.

However, when ARPS reaches saturation and evaporatively driven downdraft and cold pools are switched off, there are no mechanisms counteracting the high-cloud diabatic longwave direct feedback, which is driving convective aggregation (Figure 3.14a). Indeed, the ARPS model needs several time-steps to produce ice and deep convection throughout the whole domain. Therefore, the strongest longwave direct feedback is reached when the cloud ice has reached its maximum. The radiative feedback and the entire $\widehat{h}'^2$ budget for ARPS found here is very similar to the findings of Yang and Tan (2020) for the WRF model. They found a larger positive contribution from the longwave radiation term than that found for SAM in Wing and Emanuel (2014). There, similarly to our result, they also found an initial increase (before 10 days) in the longwave radiative feedback. Instead, SAM already starts with the presence of cloud ice and deep convection throughout the whole domain due to its diagnostic microphysics, but the increase in the longwave feedback is counteracted and reduced (as shown in 3.14a) by the beginning of cold pool activity and low-level mixing which acts to moisten dry regions. This explains the behavior of the $\widehat{h}'^2$ budget in SAM, where the initial radiative feedback and surface fluxes feedback start decreasing at the expense of the convergence feedback term. In ARPS, instead, cold pools are weaker and they can not generate low-level clouds inside the dry region. Therefore, the enhancement of the low-level longwave feedback and the consequent expansion of the dry region to form a larger radiatively driven cold pool, can not happen. After 30 days of simulation cold pool effect weakens, and the high-cloud radiative feedback dominates the maintenance of the convective organization.

These findings are in line with what was found in the literature. Yanase et al. (2020) found a competing effect between a positive feedback induced by radiative cooling in the subsidence region and a negative feedback involving cold pool spreading over the preceding subsidence area and homogenizing the boundary layer moisture (see their Figure 4). This is induced by the evaporative cooling of falling rain in the convective region. Indeed, Khairoutdinov et al. (2009) showed that removing the evaporation of rain in their simulation (thus weakening cold pools), also results in a lower shallow cloud fraction covering the domain. In this case, new deep clouds were found to develop at the sites of previous deep clouds, resembling the "moisture-memory aggregation". When

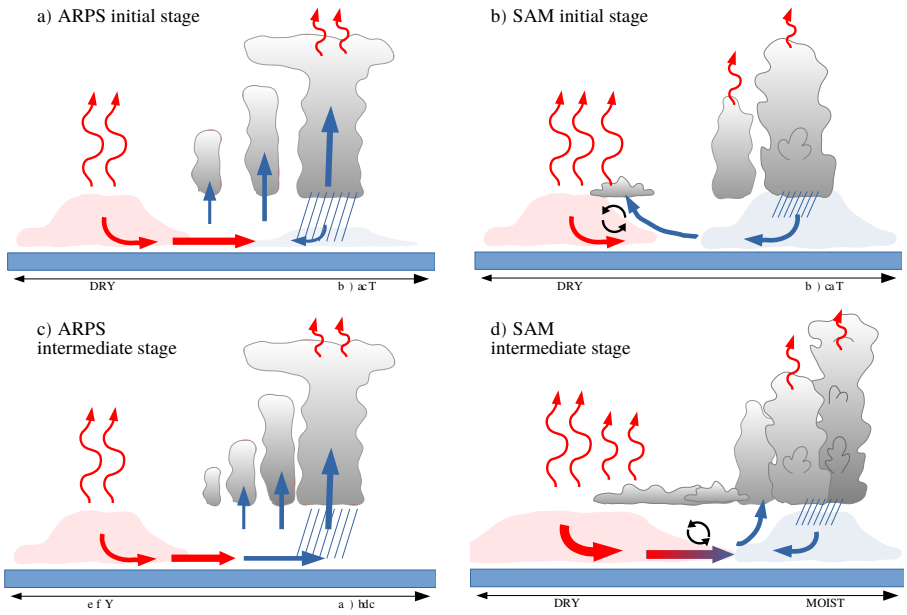

Fig. 3.23: Schematic illustration of the mechanisms proposed for self-aggregation in the two models. Red wavy arrows indicate outgoing longwave radiation; red shading indicates radiatively-driven cold pools generated in the drier zones; blue shading indicates evaporatively-driven cold pools generated in the moister zones; black arrows indicate low-level mixing; red and blue arrows indicate the predominant sense of MSE circulation. In ARPS between 5 and 30 days, at the initial stage (a), the creation of a larger anvil and the presence of many dry regions causes a strong differential heating between moist and dry zones (direct longwave feedback). The weaker cold pools are not able to transport moisture inside the dry regions, to moisten the boundary layer by low-level mixing, and to create shallow clouds in the dry regions. Instead, this happens in SAM at the initial stage, until day 30 (b), where dry patches start to expand due to the enhancement of radiative cooling by shallow clouds. At intermediate stages in ARPS, between 30 and 40 days (c), due to the weaker low-level mixing of the boundary layer, moist zones get saturated and evaporatively-driven cold pools are switched off. At this point, the direct longwave feedback maintains the moist region moister and the dry regions drier. In SAM, at the intermediate stage, between 40 and 60 days (d), when radiatively-driven cold pools have become strong enough to invert the low-level circulation created by evaporatively-driven cold pools, the high MSE which were stored at the boundary of the dry region by low-level mixing, starts to be transported into the moist regions, causing them to get moister (after day 60). Such up-gradient MSE circulation ends when the driest region has lost all its water vapor, therefore preventing any further expansion of dry regions. From this point forward (after day 70), the expansion (contraction) of moist regions alternates with the contraction (expansion) of dry regions due to competition between radiative and evaporatively-driven cold pools.

rain evaporation was present, deep clouds tended to appear along the edges of spreading cold pools, favoring also the formation of shallow clouds. Therefore, the convective organization can occur even with a low amount of shallow clouds and weak MSE circulation, once it is ensured that the sub-cloud layer is enough saturated to weaken downdrafts (Wing et al., 2017).

Different sub-cloud layer properties can arise spontaneously from different models even when starting from a similar setting as in the RCEMIP project (Wing et al., 2018). Indeed, Wing et al. (2020) found a substantial spread in the domain average temperature and humidity profiles after reaching equilibrium. Another signal of a weaker cold pool effect is the weaker surface fluxes in ARPS, compared to those in SAM (see Table 3.2). As shown by Tompkins (2001a); Schlemmer and Hohenegger (2014); Drager and van den Heever (2017), the gusty wind brought by cold pools generally enhances surface fluxes.

Differences in the way the convective organization is achieved in CRM, by using other models than SAM, have been noticed in previous studies (Jeevanjee and Romps, 2013; Yang and Tan, 2020; Tompkins and Semie, 2017; Holloway and Woolnough, 2016). For example, Holloway and Woolnough (2016) found that a low-level circulation was present in the Met Office Unified Model, but was driven mainly by anomalies in low-level diabatic heating from convection and other microphysical processes, and not by radiative cooling gradients between the moist and dry regions. Furthermore, they found that this was not crucial in organizing convection. Similarly, this has been found by Yang and Tan (2020) with WRF. For them, the expansion of dry areas was due to the dry-subsidence feedback. Tompkins and Semie (2017), using WRF, found that water vapor feedback with convection is a necessary but not sufficient condition for convective aggregation, similar to what is found here.

One could ask why in the ARPS model the sub-cloud layer is saturated. First of all, the ARPS and SAM models reach their equilibrium in very different ways. In particular, the small domain simulation of ARPS is entirely covered by a large anvil (Figure 3.2) when reaching its equilibrium. Such an anvil blocks incoming radiation and the simulation domain starts to get colder and drier with a high precipitation rate. When initializing the new large simulation, the cloud water and ice at 12 km are removed, removing the large anvil, while leaving its effect on the vertical profile of temperature and water vapor. Therefore, an aggregated state is obtained, but this occurs in a drier and colder domain, with a nearly saturated sub-cloud layer.

The usually adopted procedure of initialization by a small domain (see RCEMIP protocol (Wing et al., 2018)), is thought to eliminate a long spin-up period to reach the model's RCE state without large adjustments (Wing et al., 2018).

However, such a procedure could be affected by the presence of a large optically thin clouds anvil, which will dry and cool the whole domain. Such presence is evident also in other models of the RCEMIP project, as shown by large cloud fractions in Figure 9 of Wing et al. (2020). The reason behind the large anvil cloud fraction and cloud ice in the ARPS small domain simulation was found in cloud microphysics properties and the water species conservation process, similar to the large domain simulation. In SAM the ice is diagnostic, while in ARPS is prognostic. This difference is particularly important at the beginning of the simulation, while the microphysical features become important during the simulation, with the lower aggregation of ice to form snow and the absence of sedimentation helping to maintain and increase anvil cloud fraction in ARPS, compared to SAM.

The updrafts number and velocity are lower in ARPS than in SAM. They could be diluted by the larger lateral mixing of ARPS (not shown). Following Tompkins and Semie (2017) greater lateral mixing would help the convective organization. The effect of mixing will be investigated in more detail in the following. However, we note here that an organized state in SAM is reached with a very small lateral mixing, in contrast to what was predicted by Tompkins and Semie (2017). SAM is likely compensating the mixing effect with numerical diffusion due to the second-order accurate advection scheme (Smolarkiewicz and Grabowski, 1990), or the radiative feedback is so strong that aggregation can occur also in an environment where deep convection is not sensitive to entrainment (as occurring in the SAM model).

Since the feedback mechanisms studied in the two models report differences in both microphysics and moisture-convection feedback in the boundary layer, it can be thought that the statistical difference in aggregation indexes is related to a different dominant mode in achieving self-aggregation. Since the statistics of the RCEMIP experiment on CRMs include these feedback, they must be taken into account in the discussion of the existence of a relationship between the idealized RCE experiments and the observations of tropical convection.

Comparison of the frequency distribution between a parameter such as precipitation (see Bretherton et al. (2005)) and column water content (see Yano et al. (2012)) can give insight into the observed statistics of precipitation (Peters and Neelin, 2006; Bechtold, 2008) with respect to observations of RCE in the tropics.

3.6. Conclusions

In this study, we performed two RCE simulations with two different CRM (SAM and ARPS) and we compared their properties while reaching a statistical equilibrium of precipitation. The two models reach an organized state presenting

different warming and drying of the whole domain. In particular, a saturated sub-cloud layer in ARPS allows the localization of convection in moist regions, by weakening the negative influence of cold pools. Stronger cold pools in SAM, instead, help the creation of shallow clouds in dry regions, increasing the long-wave feedback responsible for their expansion; while delocalizing convection in moist regions and therefore opposing high-cloud radiative feedback. We found that negative moisture feedback leads to lower values of aggregation metrics in ARPS compared to SAM, mainly due to the strong drying of dry regions in the SAM model which corresponds to stronger warming and drying of the domain.

There is a chance that this is a common feature of CRMs models, where the closer the sub-cloud layer to saturation, the lower the degree of aggregation and the weaker the effect of the organized state on the average domain statistics. Moreover, in ARPS weaker cold pools lead to fewer surface fluxes, hence a greater imbalance between them and radiative cooling, F_{NET}. This makes ARPS closer to the RCE equilibrium observed in the tropics on a daily basis (Jakob et al., 2019).

Given the criticality of precipitation and its evaporation at the resolved scales, further experiments are needed with different microphysics and turbulence schemes and increased resolution, in order to resolve small scales turbulence fluxes and convective motions. Furthermore, from our results, it seems that the higher saturation in ARPS leads to a different distribution of precipitation against total water condensate, compared to that of SAM. This underlines different ways to reach a self-organized critical state in the two models, which are both contained in the statistical oscillation of atmospheric tropical convection and could be included in climate sensitivity studies.

4. Mediterranean tropical-like cyclones simulated by ensemble forecasting

This chapter[1] is concerned with the study of medicanes by means of ensemble forecasting. The objectives of the work are to investigate the prediction of medicanes intensity and tracks using a general circulation model and, compare ensembles with the inclusion of the physics uncertainties and ensembles with the inclusion of initial conditions uncertainty. Indeed, there are inherent limits to the predictions of Mediterranean cyclones, and medicanes in particular, due to the predictability of mid-latitude weather, to the multiscale nature of these cyclones, to the contrast in the density of observations between land and sea, and to the diverse complex orography of the region. Improving the accuracy, and reliability of forecasts for these cyclones is a complex task that requires advances in both large and convective-scale initial condition uncertainty representations. Thus, a comparison between three ensemble forecast experiments is set up with the use of the ECMWF Integrated Forecasting System, IFS, ensemble forecasting system. One ensemble is run with only initial condition perturbation (Ensemble Data Assimilation, EDA), one is run with the perturbation of physical parameterizations and one is run with only the convective parameterization

[1] This chapter covers the research carried out in collaboration with ECMWF. It focuses on the analysis of ensemble forecasting applied to Mediterranean tropical-like cyclones. It is concerned with assessing the differences between the types of uncertainty included in the ensemble, the initial conditions one or the physical parameterization one, and their impact on the characteristics of these cyclones.

perturbations. The physical parameterization perturbation has been carried out by means of Stochastically Perturbed Parameterizations (SPP), a novel but promising scheme developed at ECMWF. The study is carried out for three medicanes, among the strongest in recent years: Ianos, Zorbas, and Trixie. It was generally found that the forecasts accurately reproduced both the thermal structure and symmetry of the cyclones, compared to the analysis value. It was also found that the use of SPP, and specifically, the perturbation of convective parameters, improves the forecast, in terms of tracking spread and precipitation intensity (based on statistical scores). Finally, it was verified that among the three medicanes, the best predicted is Ianos. Given, that the ECMWF ensemble forecasts model with high resolution can adequately reproduce medicanes with their tropical-like features, these results pose the basis for investigating the physical mechanism behind these results, especially regarding the deepening of the cyclone, its tracking, and its transition to the tropical phase.

4.1.Introduction

The Mediterranean region is a small but geographically complex area characterized by sharp land and sea transitions and surrounded by high mountain ranges. It is known for its frequent cyclogenesis. A small number of the intense cyclones that originate in the region present tropical-like features (Flaounas et al., 2022). They are a very significant phenomenon, due to their similarity with tropical cyclones, and while they are typically shorter-lived than North Atlantic hurricanes, they may exhibit several tropical-like characteristics in their mature phase, such as a high degree of axial symmetry, a warm core, a tendency to weaken after making landfall, and a cloud-free "eye" at the center of the storm of mostly calm weather, as inferred from satellite images. Such vortices are better known as Tropical-Like Cyclones or Mediterranean hurricanes (medicanes). Medicanes have been documented in the Mediterranean region since the beginning of the satellite era (Ernst and Matson, 1983) and have been associated with polar lows (Rasmussen and Zick, 1987). These storms pose a significant threat due to their intense winds, heavy rainfall, and associated flooding.

Medicanes features have been commonly observed in the literature (Cavicchia et al., 2014; Romero and Emanuel, 2013; Emanuel, 2005; Zhang et al., 2019c; Miglietta and Rotunno, 2019). They occur very infrequently, with an average of about $1/2$ events per year over the entire Mediterranean region. They are most commonly formed in the western Mediterranean and in the area between the Ionian Sea and the North African coast. Medicanes have a distinct seasonal pattern, with a peak at the start of winter, a significant number of events during fall, a few during spring, and very little activity in summer. As pointed out by Miglietta and Rotunno (2019) they only have a lifespan of a few days due to the limited size of the Mediterranean Sea, which is their main source of energy. Furthermore, they only exhibit fully tropical characteristics for a short period,

with extratropical features predominating for most of their lifetime (Miglietta and Rotunno, 2019).

Medicanes differ from other Mediterranean cyclones in the complexity of their formation and maintenance. Indeed, unlike hurricanes, which develop in regions with near-zero baroclinicity and draw their energy from warm tropical oceans, medicanes form from pressure lows under moderate to strong baroclinicity, which is a typical condition of midlatitudes. The latter aspect leads to a low rate of occurrence, given that the environmental conditions of weak vertical wind shear, which are necessary for their development, are unusual (Nastos et al., 2018). The interaction between warm sea and cold air associated with a deep upper-level trough creates the necessary thermodynamic disequilibrium for these storms to develop a warm core (Emanuel, 2005; Miglietta and Rotunno, 2019). The type of mechanics that can be deemed responsible for the creation of a medicane in the Mediterranean is similar to the one responsible for the formation of tropical cyclones farther from the equator near the tropics (McTaggart-Cowan et al., 2015).

The development and maintenance of medicanes are the results of a synergy between synoptic-scale processes, which provide the necessary environment, and mesoscale processes such as deep convection and latent heat fluxes from the sea (Emanuel, 2005). Because of their small size, limited data availability, low frequency of occurrence (Cavicchia et al., 2014), and the complex geography of the Mediterranean region, studying and predicting medicanes is a challenge for numerical weather forecasting. For this reason, besides some climatological studies on medicanes, using synthetic production of tracks and 3D numerical simulation (Romero and Emanuel, 2013; Cavicchia et al., 2014) there is a higher rate of studies in the literature that has been focused on modeling medicanes using convective permitting models, due to the fact that they can be more effective at reproducing the small-scale processes that contribute to medicane formation and maintenance (Davolio et al., 2009; Miglietta et al., 2011, 2013; Mazza et al., 2017; Cioni et al., 2016; Ricchi et al., 2019) rather than observational aspects (Pytharoulis et al., 2000; Moscatello et al., 2008).

There have been relatively few studies that have analyzed medicanes using ensemble forecasts (Chaboureau et al., 2012; Mazza et al., 2017) and more specifically, by using the ECMWF ensemble forecasting system (Pantillon et al., 2013; Di Muzio et al., 2019; Portmann et al., 2020). Indeed, the ensemble forecast of ECMWF has proven to be a useful tool for predicting extreme weather events (Buizza and Hollingsworth, 2002; Buizza, 2008; Magnusson et al., 2015), for analyzing tropical cyclones (Torn and Cook, 2013) and their predictability (Munsell et al., 2013). Moreover, the model has demonstrated high predictive skill also for medicanes (Di Muzio et al., 2019). Pantillon et al. (2013)

used ECMWF operational ensemble forecasts to study the predictability of a medicane in 2006 and found that they were more successful at consistently capturing early signals of its occurrence compared to ECMWF deterministic forecasts. Di Muzio et al. (2019), who used ECMWF ensemble forecasts to systematically analyze the predictability of medicanes, found that the ensemble members noted a marked drop in predictive skill beyond 5-7 lead days, indicating the existence of predictability barriers. Portmann et al. (2020) used ensemble forecasting to assess upstream uncertainties in the prediction of medicanes, finding also that the uncertainties were reduced with later initializations.

Research conducted on medicanes with ensemble forecasting has generally been carried out through the use of the perturbations to initial conditions only (Di Muzio et al., 2019; Portmann et al., 2020). However, an important part of the uncertainty associated with forecasting comes from uncertainty related to the physics of the model. For these reasons, this present study is concerned with ensemble forecasting that takes into account not only the uncertainty of initial conditions but also the uncertainty of model parameters. We present an assessment of the prediction of medicanes, with the use, not only of the Integrated Forecasting System (IFS) operational ensemble forecasting system at ECMWF, with initial conditions perturbation, but using also the physical parameterization perturbations, the Stochastically Perturbed Parameterizations (SPP) ensemble forecast. Indeed, this is a novel scheme of stochastic representation of model uncertainties which is still under development at ECMWF in order to replace the Stochastically Perturbed Parameterization Tendency scheme (SPPT) (Palmer et al., 2009). SPP has been found relevant in improving forecasts, specifically with ECMWF ensembles (Ollinaho et al., 2017a) and tropical cyclones. This new representation of uncertainties associated with model physics consists of a set of physical parameters in the model being perturbed (Ollinaho et al., 2017b). The use of this method is rising in the literature (Frogner et al., 2022) precisely because it allows the reproduction of uncertainty close to the actual source of error and maintains physical consistency, particularly with local conservation of energy and humidity (Lang et al., 2021).

Thus a comparison between three ensemble forecast experiments is set up. One ensemble is run with only initial condition perturbations, through the Ensemble Data Assimilation (EDA), one is run with the entire physical parameterizations perturbed and one is run with only the convective parameterization perturbed. The goal of this study is to determine whether these forecasts can accurately predict medicanes and if there are possible biases presented by the ensemble forecasts. The goal of the research is also to assess which of the perturbation experiments can capture the medicane more accurately. Differently from (Di Muzio et al., 2019), only three medicanes, among the strongest in recent years, were chosen: Ianos (September 15-20, 2020), Zorbas (September 27-31,

2018), and Trixie (October 28-November 1, 2016). The three medicanes have been selected because they are very different from each other, with Trixie being the weakest but also the longest lasting of the three (Di Muzio et al., 2019) and generally among the longest lasting medicanes, Zorbas one of the shortest-lived and presenting high variability in predictability, as documented by (Portmann et al., 2020) and Ianos one of the most intense medicanes ever observed, reaching category 2 hurricane status (Lagouvardos et al., 2022).

In section 4.2 the data and methods used are described, with an in-depth description of the ensemble forecast experiments carried out, a description of the SPP, and of the tracking method. In section 4.3 a brief overview of the three storms studied is provided, highlighting their important characteristics. In section 4.4 the results are presented for the track position and intensity, precipitation, and thermal structure. In the final section 4.5, the results are discussed and the concluding remarks are given.

4.2.Data and Methods

In this section, a description of the methods and techniques used to analyze the medicanes with the ensemble forecast and the operational analysis with the ECMWF model Integrated Forecasting System (IFS), is given. Firstly, a description of the Stochastically Perturbed Parameterizations ensemble forecast of IFS is provided, together with a brief description of the carried-out perturbation experiments. Secondly, the data, used to validate the simulation besides the operational analysis, are described. Finally, the tracking algorithm is described.

4.2.1.Ensemble Forecast Simulation

For this work, the ensemble forecast experiments with the ECMWF IFS (Cycle 47r3: ECMWF,(IFS Documentation CY47R3, 2021b)) and the ECMWF operational analysis have been used. Both the ensemble forecast and the operational analysis have a $\simeq$ 9 km horizontal grid spacing (TCo1279, for a more in-depth description of the horizontal grid, see Malardel et al. (2016)) and are run with 137 vertical levels. The forecasted period used in this work is 9 days. Three different sets of experiments have been conducted, all of them consisting of an 8-members ensemble. The ensemble forecasts are initialized, amounting to 3 initial dates, each day at 0000 UTC. For Ianos the three dates are the 15th, the 16th, and the 17th of September 2020, for Zorbas the three dates are the 25th, the 26th, and the 27th of September 2018 and for Trixie the three dates are the 25th, the 26th, and the 27th of October 2016. Regarding the physical parameterization, a detailed description can be found in the IFS documentation (IFS Documentation CY47R3, 2021a). The ensemble forecasts are coupled to the ECMWF Wave Model (ecWAM:(IFS Documentation CY47R3, 2021c)), to the Nucleus for European Modelling of the Ocean (NEMO) ocean model (Mogensen et al., 2012) and the LIM2 sea-ice model (Goosse and Fichefet, 1999).

The Different types of experiments that have been carried out are reported in Table 4.1.

Table 4.1: Description of the different ensemble forecast experiments

Experiment ID	Experiment Setup
INI	Initial perturbations only - no model uncertainty representation
$SPP-Conv$	No initial perturbations - convective parameterization uncertainty representation
SPP	No initial perturbations - physical parameterizations uncertainty representation

The first experiment is the ensemble forecast with initial condition perturbation only (INI experiment). This is done by adding perturbation to a 4D-Var (Rabier et al., 2000) analysis. The perturbations are constructed from an ensemble of 4D-Var data assimilations (Ensemble Data Assimilation, EDA, (Buizza et al., 2008)). The second and third sets of experiments are conducted by running the ensemble with SSP applied. In the former (SPP-Conv) only the convective parameterization parameters are perturbed. In the latter (SPP), the whole physics parameterization, convection, radiation, clouds and large-scale precipitation and turbulence diffusion and sub-grid orography parameters are perturbed, as discussed below.

4.2.2.Stochastically Perturbed Parameterizations Ensemble

The Stochastically Perturbed Parameterizations scheme represents model uncertainty in numerical weather prediction by introducing stochastic perturbations into the physical parameterization schemes (Lang et al., 2021) as mentioned above. SPP is a new scheme, aimed at replacing the currently used Stochastically Perturbed Parameterization Tendency scheme (SPPT) (Palmer et al., 2009) in the ECMWF ensemble forecast in June 2023.

One could ask why the need for a new scheme to represent model uncertainty. The SPPT scheme randomly perturbs the net of the physical tendencies at each time-step. The physical tendencies are defined as the amount of change of the prognostic model variables (winds, temperature, and humidity) due to physical parameterizations. The perturbation is produced with multiplicative noise. This leads to greater uncertainty attributed to larger tendencies while preserving the relative balances between the tendencies of different physical processes. However, SPPT is not able to trigger new states (for a single time-step), but it is only able to enhance or diminish the effect of the represented physical processes,

and it introduces inconsistency between the perturbed physics tendencies and fluxes that are computed from the unperturbed tendencies. Since there is no correction for the surface fluxes, after perturbing the atmospheric tendencies, an energy imbalance is introduced into the system.

For these reasons, Ollinaho et al. (2017b) started developing a new scheme of stochastic representation of model uncertainties, following the work of Baker et al. (2014); Christensen et al. (2015). This new scheme of perturbation is based on applying perturbations directly to well-known poorly constrained parameters and variables within the parameterization schemes. This way the model uncertainty representation can be directly related to known sources of model uncertainties associated with specific processes. Thus, with the SPP, some individual IFS parameters and variables are perturbed at each time-step with in-space varying noise derived from in-time evolving 2D random number fields. The SPP scheme is designed such that it converges to the unperturbed (i.e. deterministic) forecast model in the limit of vanishing variance of the noise.

Perturbed parameters

The implementation of SPP allows simultaneous perturbations of up to 27 parameters and variables in the deterministic IFS parametrizations of (a) turbulent diffusion (Köhler et al., 2011), subgrid orography (Beljaars et al., 2004), convection (Tiedtke, 1989; Bechtold et al., 2008), cloud processes and large-scale precipitation (Tiedtke, 1993; Forbes et al., 2011), and radiation (ecRad, (Hogan and Bozzo, 2018)). These 27 parameters and variables are reported in Tables 4.2 - 4.5.

Regarding the boundary layer processes, one of the most important aspects is the turbulent transfer, with which the atmospheric model knows about the surface boundary condition. Land surfaces tend to be heterogeneous and it is very difficult to characterize real terrain in terms of the surface drag that it exerts on the flow, especially in areas with orography. Usually, the Monin Obukhov Similarity is used, which, however, applies to stationary homogeneous conditions. Furthermore, there is much uncertainty regarding the interaction between turbulence and ocean waves. Given that and the fact that the atmospheric circulation is sensitive to surface drag (Sandu et al., 2016), the chosen parameters of the scheme to be perturbed are related to the surface drag and the sub-grid orography scheme. Furthermore, the von Kármán constant is perturbed, because several uncertain quantities in the turbulent diffusion scheme are proportional to the von Kármán constant. These are the standard deviation of vertical velocity, which is estimated via an empirical formula, the Obukhov length used to compute the stability functions in the outer layer for surface-driven diffusion, the surface-layer (the level below the lowest model level) mixing length used in the outer layer (the level above the lowest model level) and the

cloud-top driven diffusion. It has to be underlined here that the perturbation to the transfer coefficient for momentum over land and over the ocean and the von Kármán constant is changed to be dependent on the type of planetary boundary layer (PBL) type.

Table 4.2: Perturbed parameter settings for the turbulent diffusion and subgrid orography parameterization. The first column is the parameter identifier, the second is a brief explanation of the role of the parameter in the model. The third column (dist.) indicates the sampled distribution type: LN and N refer to the log-normal and normal distribution, respectively. The fourth column presents the standard deviations of the two underlying Gaussian distributions.

Parameter ID	Role of the parameter	Dist	σ	Scheme
Turbulent diffusion and subgrid orography				
CFM_{OC}	Transfer coefficient for momentum over ocean	LN	0.26, for dry convective PBL 0.33	Mean/Median
CFM_{LA}	Transfer coefficient for momentum over land	LN	0.65, for dry convective PBL 0.78	Mean/Median
$RKAP$	von Kármán constant	LN	0.65, for dry convective PBL 0.26	Mean/Median
$TOFDC$	Coefficient in turbulent orographic form drag scheme	LN	0.78	Mean
$HSDT$	St. dev. of subgrid orography	LN	0.52	Mean
$VDEXC_{LEN}$	Length-scale for vertical mixing in stable boundary layer	LN	1.04	Mean

Regarding the convective parameterization, the most uncertain processes are the mixing of the cloud with the environment (detrainment and entrainment rates) and the mixing in the sub-cloud layer, the autoconversion in the microphysics of clouds, and the convective momentum transport. The uncertainty of the latter is of particular interest because it depends on the organization of convection and on the fact that it usually acts to reduce the shear, while in the mesoscale convective system, it can also have an opposite effect (Ollinaho

et al., 2017b). Indeed, the perturbation regards the zonal and meridional convective momentum transport due to shallow and deep convection.

Table 4.3: Perturbed parameter settings for the convection parameterization. The first column is the parameter identifier, the second is a brief explanation of the role of the parameter in the model. The third column (dist.) indicates the sampled distribution type: LN and N refer to the log-normal and normal distribution, respectively. The fourth column presents the standard deviations of the two underlying Gaussian distributions.

Parameter ID	Role of the parameter	Dist	σ	Scheme
	Convection			
$ENTRORG$	Entrainment rate	LN	0.39	Mean
$ENTSHALP$	Shallow entrainment rate	LN	0.39	Mean
$DETRPEN$	Detrainment rate for penetrative convection	LN	0.39	Mean
$RPRCON$	Conversion coefficient cloud to rain	LN	0.52	Mean
$CUDU$	Zonal convective momentum transport, deep convection	N	1.22	Mean
$CUDV$	Meridional convective momentum transport, deep convection	N	1.22	Mean
$CUDUS$	Zonal convective momentum transport, shallow convection	N	1.33	Mean
$CUDVS$	Meridional convective momentum transport, shallow convection	N	1.33	Mean
$RTAU$	Adjustment time-scale in CAPE closure	LN	0.78	Mean
$ENTSTPC1$	Shallow entrainment test parcel entrainment	LN	0.39	Mean

Concerning the parameterization of clouds and precipitation microphysics, perturbation is added to the different processes of cloud formation, cloud dissipation, rain formation, and snow aggregation. The rain evaporation rate, the snow sublimation rate, and the vertical velocity that is used to calculate the adiabatic temperature change for saturation adjustment of condensation and evaporation are perturbed. These perturbations are chosen as they directly change latent heating, which affects atmospheric stability and tropospheric dynamics, particularly in the storm tracks (Attinger et al., 2019, 2021). Indeed,

these processes' uncertainty comes from the low understanding of the rain and snow particle size distributions, terminal fall speeds, and particle properties (Boutle and Abel, 2012). Other uncertainties come from the variability of adiabatic heating and cooling, which drives condensation and evaporation. Other perturbations regard the critical relative humidity parameter characterizing the sub-grid heterogeneity of humidity in a grid box and define the relative humidity at which condensation begins to form in the respective grid box.

Table 4.4: Perturbed parameter settings for the cloud and large-scale precipitation parameterization. The first column is the parameter identifier, the second is a brief explanation of the role of the parameter in the model. The third column (dist.) indicates the sampled distribution type: LN and N refer to the log-normal distribution and the normal distribution, respectively. The fourth column presents the standard deviations of the two underlying Gaussian distributions.

Parameter ID	Role of the parameter	Dist	σ	Scheme
Cloud and large-scale precipitation				
$RAMID$	Relative humidity threshold for the onset of stratiform condensation	LN	0.13	Mean
$RCLDIFF$	Diffusion coefficient for the evaporation of cloud at subgrid cloud edges	LN	1.04	Mean
$RLCRITSNOW$	Cloud ice threshold for autoconversion to snow	LN	0.78	Mean
$RAINEVAP$	Rain evaporation rate	LN	0.65	Mean
$SNOWSUBLIM$	Snow sublimation rate	LN	0.65	Mean
$QSATVERVE$	Vertical velocity used to calculate the adiabatic temperature change for saturation adjustment	LN	0.39	Mean

Finally, regarding radiation parameterization, the greatest source of uncertainty in the IFS radiation scheme is the treatment of clouds and aerosols (Morcrette et al., 2008). Indeed, with the use of the Monte Carlo Independent Column Approximation (McICA) to represent sub-grid cloud distribution from the cloud fraction and grid box mean cloud water content, uncertainty may arise due to the vertical decorrelation length scale governing the degree to which clouds in adjacent layers overlap, and due to the fractional standard deviation of water content in a given model layer. regarding aerosols, their optical depth for 5

different species is prescribed from a climatology derived from (Tegen et al., 1997). The climatology varies as a function of longitude, latitude, and month. To compute the 3D aerosol distribution, concentrations are assumed to decrease exponentially with height according to a specified scale height for each species. SPP perturbs aerosol optical depth and this scale height independently. Moreover, the other important aspect of parameterization is converting water content to the optical depth of a layer. This is done by parameterizing it as a function of effective radius, which for liquid clouds is calculated as a function of liquid water content (Martin et al., 1994), while ice is a function of temperature. Both parameters are perturbed.

Table 4.5: Perturbed parameter settings for the radiation parameterization. The first column is the parameter identifier, the second is a brief explanation of the role of the parameter in the model. The third column (dist.) indicates the sampled distribution type: LN and N refer to the log-normal and normal distribution, respectively. The fourth column presents the standard deviations of the two underlying Gaussian distributions.

Parameter ID	Role of the parameter	Dist	σ	Scheme
	Radiation			
$ZDECORR$	Cloud vertical decorrelation height	LN	0.78	Mean
$ZSIGQCW$	Fractional st. dev. of horiz. distribution of water content	LN	0.52	Mean
$ZRADEFF$	Effective radius of cloud water and ice	LN	0.78	Mean
ZHS_VDAERO	Scale height of aerosol normal vertical distribution	LN	1.04	Mean
$DELTA_AERO$	Optical thickness of aerosol	LN	0.78	Mean

Perturbation distribution

As mentioned above, the SPP perturbation is applied through a 2D random field generator. In its implementation, SPP uses a single scale with a decorrelation length scale of 1,000 km and a decorrelation time of 3 days (Figure 1 of Lang et al. (2021)).

If the unperturbed parameters are defined as $\hat{\xi_j}$ (the value of the parameter used in the deterministic forecasts.) for each of the jth parameters, the perturbed parameter is defined as ξ_j, where there are two options of drawing it from either a Gaussian distribution or log-normal distribution, where either the mean or the median is set equal to the unperturbed quantity. For most parameters, SPP

samples from a log-normal distribution, thus:

$$\xi_j = exp(\Psi_j)\hat{\xi}_j, \quad \Psi_j \sim \mathcal{N}(\mu_j, \sigma_j^2) \tag{4.1}$$

Here, the perturbations Ψ_j sample a Gaussian distribution with a mean μ_j and a standard deviation σ_j, both determined individually for each perturbed parameter j. The above-mentioned two options considered for determining the mean μ_j are:

1. μ_j = -1/2 σ_j, where the mean of the log-normal distribution of the perturbed parameter ξ_j is equal to the value of the unperturbed parameter $\hat{\xi}_j$;
2. μ_j = -σ_j, where the median of the log-normal distribution of the perturbed parameter ξ_j is equal to the value of the unperturbed parameter $\hat{\xi}_j$;

Each perturbed parameter is assigned an individual random field and different random fields are statistically independent. The log-normal distribution has been chosen for practical reasons since it is ensured that the perturbed parameter values ξ_j retain their original sign. The only exception to the use of independent log-normal distribution are the perturbations applied to the zonal and meridional components of the convective momentum transport, which are perturbed with multiplicative noise drawn from a bi-variate normal distribution and they share the same 2D random field.

The perturbations $exp(\Psi_j)$ vary in space and time using independent patterns for each parameter j, and for each ensemble member (Figures 1 and 2 of Ollinaho et al. (2017b)). The same horizontal pattern is applied to all model levels, in order to maintain physical consistency within a grid column.

4.2.3. Validation Data

In order to analyze the predictive skill of the ensemble, besides the operational analysis, the ensemble forecast perturbation experiments have been compared with the satellite-based, globally-gridded Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG) (Huffman et al., 2020). In GPM-IMERG, the retrievals from geostationary satellites are blended seamlessly with information from the passive microwave (PMW) sensors from low-orbit satellites. this is done in order to provide for both a high accuracy and a high temporal (30 min) and spatial ($\simeq$ 10 km) resolution since the precipitation estimation based on only the PMW suffers from a low sampling rate. Furthermore, the data are also calibrated by using rain gauges at the ground base. In this research, the 24-hr accumulated precipitation values are used.

4.2.4.Cyclone tracking

The method described here has been used to evaluate the tracks for both operational analyses, used as verification and ensemble forecasts. The tracking method is based on Picornell et al. (2014) and Ragone et al. (2018). The algorithm is firstly aimed at finding the local minima of the sea level pressure field at each time-step. Then, for each minimum, the gradient of the sea level pressure along eight principal directions (E, NE, N, NW, W, SW, S, SE) inside a circle of radius 200 km is computed. The computed gradient is then chosen to be lower than $5hPa/200km$ in at least 6 directions. After the minimum detection and filtering via selection through seal level pressure gradient, a proximity condition is applied to construct the complete trajectory. Starting from the first time-step, each minimum is connected to the following one, at the following timestep, that respect the condition of being closer than $\Delta x = V\Delta t$, with $V = 50km/h$ and $\Delta t = 3h$. If this condition is met, the two consecutive minima are considered to belong to the same trajectory. This condition has been considered suitable and chosen according to the results of Ragone et al. (2018). Once the trajectories have been found, only the trajectories that last longer than 24 hours and those that spend more than 12 hours over the sea are selected. Trajectories that spend less than half their duration over land or within 100 kilometers of the coast will be discarded.

4.3.Overview of the Storms

In this section, an overview of the three medicanes chosen is given. A brief analysis of the synoptic environment is provided, following the literature. A summary of the main features as retrieved by the analysis data: the storm duration, the period, the region of occurrence and the asymmetry *B*, and upper-level thermal wind $-V_T^U$ (the last two parameters will be explained in the following sections) is provided in Table 4.6. The intensity (central pressure) and trajectory of each storm are shown in Figure 4.1.

Table 4.6: Duration, region of occurrence, central pressure, CP, asymmetry parameter, B, and upper-level thermal wind from the operational analysis for each storm. The upper-level thermal wind, $-V_T^U$ and the thermal asymmetry parameter, *B* are taken from the Hart cyclones phase-space (Hart, 2003).

Storm	Region	Period	Duration (d)	CP (hPa)	B	$-V_T^U$
Ianos	SM	September 2020	7	994	0	73
Zorbas	SM	September/October 2018	5	993	4	39
Trixie	SM	October 2016	4	1009	1	38

*SM = Southern Mediterranean

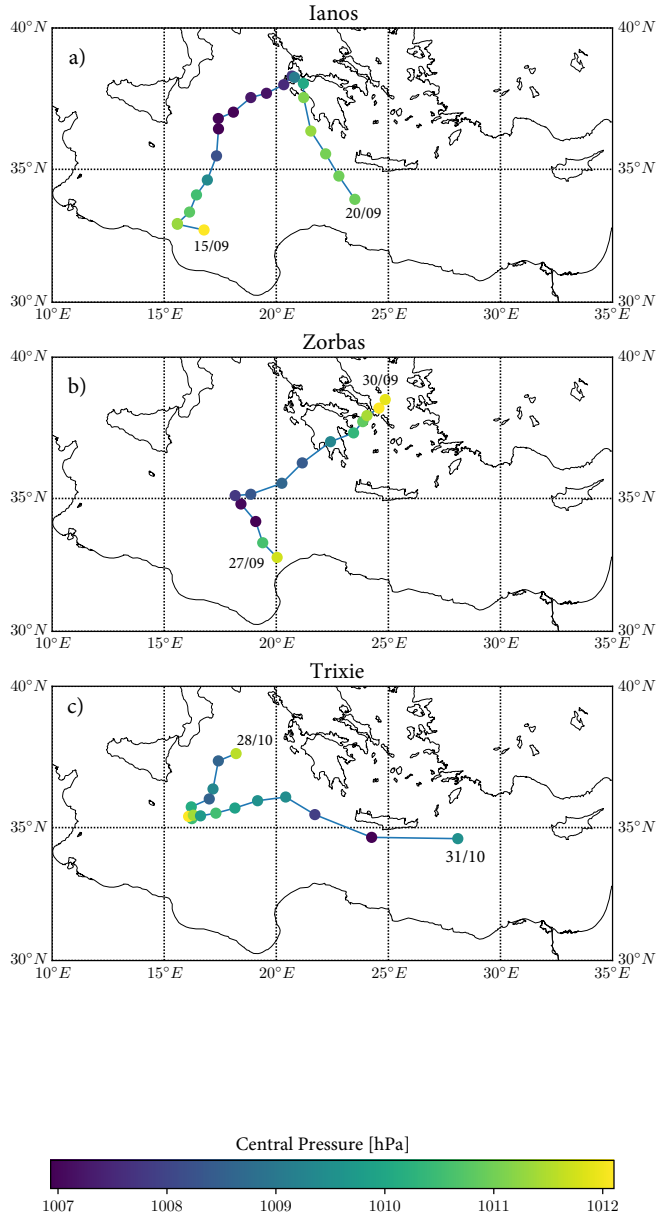

Fig. 4.1: Track of the three storms for the operational analysis. The colour of each circle represents the intensity, meaning the central pressure in hPa.

4.3.1. Ianos

As the track suggests (Figure 4.1a) Ianos originated in the Gulf of Sidra after an upper-level cutoff low that formed on the 9th of September moved eastward to the western Mediterranean and northern Africa. Then between the 14th-15th

of September, it emerged in the Gulf of Sidra, due to stratospheric dry intrusion with high vorticity values (Comellas Prat et al., 2021). This process was also accompanied by an anomalous value of SST which was 1.5 °C higher than the normal (27-28 °C).

On the 16th of September, it intensified, and it moved to the Ionian sea, and then on the 17th of September, the system showed a clear mid-level warm core, lining up with the cut-off low. Ianos obtained tropical characteristics at this point (CP = 994, B = 0, and $-V_T^U = 73$). The analysis is capable to reproduce a value similar to the observed pressure minimum of 995 hPa (Comellas Prat et al., 2021).

4.3.2. Zorbas

Zorbas formed at 12:00UTC on 27 September 2018 close to North Africa and then moved into the central Mediterranean, turning eastward and moving over Greece into the Aegean Sea, where it finally decayed 4 d after its formation (Figure 4.1b). Medicane Zorbas originated from a PV streamer created by the elongation of a large-scale trough in eastern Europe, which reached the central Mediterranean Sea at 00:00UTC on 27 September 2018 (Portmann et al., 2020). This PV streamer broke up at the time of cyclogenesis, resulting in the formation of a PV cutoff.

At the same time, as for Ianos, a large SST anomaly was taking place in the Gulf of Sidra, with similar intensity as in the Ianos case. Zorbas cyclogenesis was also accompanied by a strong low-level advection of air with low virtual potential temperature, Θ_e, across the Aegean sea. Then this air was substantially moistened by sea-surface fluxes as it traveled across the Aegean Sea. One day after formation (28th of September), Zorbas presented a deep warm core (Table 4.6). After reaching the state of the warm core when Zorbas reached Greece on 29 September 2018, this was sustained for the next few days. Zorbas reached its maximum intensity (observed 992 hPa), which is well captured by the analysis.

4.3.3. Trixie

Medicane Trixie formed on the 28th of October, as the consequence of a deep cut-off low which emerged on 26–27 October and moved from northern to southern Italy in the following days, triggering deep convective storms along the Italian west coast. This PV anomaly crossed the Adriatic Sea on 28 October at 04:00 UTC (EUMETSAT analysis by Scott Bachmeier, Jochen Kerkmann, and Djordje Gencic) and then quickly moved to Sicily and approached Tunisia and Algeria. Then between the 26 and the 28 of October, the medicane started to develop in the area of the old PV anomaly. On the 29th of October, it deep-

ened and moved to the east of Malta, then on the 30th of October, it moved eastward towards Greece (Figure 4.1c).

Trixie reached the warm core structure on the 30th of October. This is in line with observation (EUMETSTAT report). Eventually, it passed near Crete on the 31st of October. The most intense convective activity started after 12 h of 30 October 2016 and lasted until 31 October 2016 (Dafis et al., 2020). In the analysis, there was only a short intensification period evident and the minimum pressure was fluctuating between 1010 and 1014 hPa during the period from the 29th and the 30th of October, which might have been highly underestimated.

The three storms formed and developed in the same area, the Southern Mediterranean, in the Ionian and Aegean Seas. This region has one of the highest medicanes occurrence, as recognized in the literature (Cavicchia et al., 2014; Zhang et al., 2021). They occurred in the same period of the year, between September and November, the most frequent period for medicane occurrence (Romero and Emanuel, 2013). From Figure 4.1 it can be gathered that there are some differences in duration and intensity, with Trixie being the longest-lasting of the three medicanes (in terms of the deepening phase) and Zorbas and Ianos being deeper than Trixie.

4.4.Results

This section examines the ensemble forecast experiments for certain aspects that are crucial for cyclones. The primary focus is on the cyclone track, as it is largely influenced by large-scale processes. The second feature under investigation is the cyclone's intensity, as measured by central pressure. This value is considered the most stable and robust metric for assessing the intensity of a cyclone on a global scale (Davis, 2018). Moreover, the precipitation is analyzed. By looking not only at the tracking but also at the intensity and precipitation it is possible to assess the model's capability to reproduce multi-scale processes. Then the precipitation is analyzed.

4.4.1.Tracking

The ensemble tracking results are reported in Figure 4.2. As examples, the tracks starting from the 16th of September are shown for Ianos, the ones starting from the 27th of September are shown for Zorbas, and the ones starting from the 27th of October are shown for Trixie.

By looking at Figure 4.2a to c and Figure4.2d to f the ensemble tracks follow sensibly the references for Ianos and for Zorbas. On the contrary for Trixie, the tracking, which starts one day prior to the starting date of the reference track (starting on the 28th) follows the track until the 29th early hours (Figure 4.2g

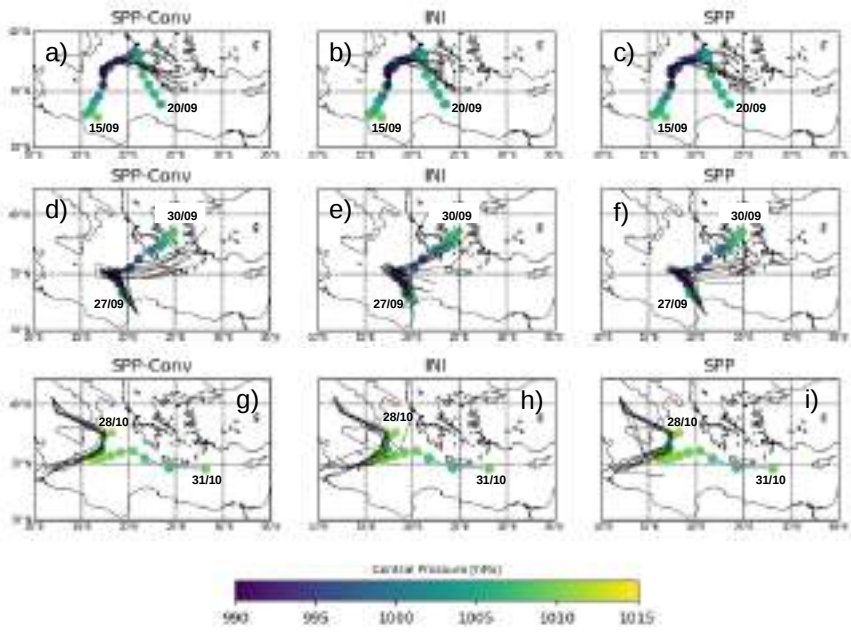

Fig. 4.2: Track of the three storms for the operational analysis as reference track and for the ensemble members belonging to each experiment (SPP-Conv on the first column, INI on the second column and SPP on the third column) for the three storms, Ianos in a, b and c, Zorbas in d, e and f and Trixie in g, h and i. As background the operation analysis is reported with the colours representing the intensity, meaning the central pressure in hPa. For Ianos the experiments starting on the 16/09 have been chosen, for Zorbas the ones starting on the 27/09 and for Trixie the ones starting from the 27/10.

to i), then it diverges and ends up in North Africa, underlying a missed forecast. For earlier starting dates (the 25th) only some members are able to follow the reference track further in the Aegean Sea, but with a great error in terms of position (not shown). Figure 4.2 shows that, as expected, the simulations with initial conditions perturbations present usually more spread at the starting point. However, the spread at later stages in the simulation seems to be similar for both the INI experiment and the two experiments with the perturbation on the physics, SPP, and SPP-Conv.

For Ianos the 16th of September was chosen as an example, but the behavior for the three starting dates is quite similar, with them being able to reproduce the trajectory quite well for the first days and the error increasing with time. It never exceeds 400 km (not shown) and up to 48 hours is always below 100 km (not shown). The error values are similar in all three experiments, but slightly lower for the INI experiment. In the case of Zorbas and Trixie, since the starting dates are earlier than when the reference track starts, the earlier the simulations

start, the more uncertainty there is with respect to the starting position, thus the more spread. In Figure 4.2 the latest starting dates simulations are shown for both cyclones. For Zorbas, starting from the 27th, the obtained tracks are following the reference with a small error at least for the first days, which is anyway always under 300 km, at later stages. As in the case of Ianos the error committed by the three experiments is similar. For Trixie, the error, regarding the 27th, the last starting date shown in Figure 4.2g, h and i, goes up to 700/800 km for the three experiments.
The tracking results shown in Figure 4.2 are mirrored by the ensemble spread and the relationship between the spread and the error presented in Figure 4.3 and Figure 4.4. Following the approach of Hamill et al. (2011), the spread of the ensemble for a single cyclone ensemble forecast at a given forecast time is defined as:

$$S(t) = \frac{1}{n} \sum_{i=1}^{n} D_i \tag{4.2}$$

where D_i denotes the great-circle distance of the ith ensemble member position of the cyclone from the ensemble mean cyclone position. The total number of ensemble members used for the calculation is 8 and the spread has been computed for each starting date (3 dates for each medicane). The results, reported in Figure 4.3, regard the same starting date as in Figure 4.2, thus for Ianos is the 16th of September, for Zorbas the 27th of September, and for Trixie the 27th of October, as also representative of the spread of the other starting dates. It is shown that for Ianos (Figure 4.3a) the SPP and the INI experiment show the largest mean track spread within the first 60-80 hours when the SPP-Conv track spread increases. For Zorbas (Figure 4.3b) the spread in the track is higher in the INI and the SPP experiments, compared to the SPP-Conv for the whole forecast duration, while for Trixie, the INI experiment shows a higher spread, especially in the first hours, becoming lower only than the other experiments only after 100 h. Zorbas and Ianos present less spread than Trixie upon initialization for the INI experiment in particular, compared to Trixie. This might be due to the fact that for Trixie, which analysis track starts on the 28th of October (Figure 4.2c) a great uncertainty in the initial position of the track exists (not shown). Indeed the spread decreases after 48 hr, coinciding with the 28th of October (Figure 4.2c).

At later forecast steps, Trixie is also the cyclone with the highest spread for all three perturbations compared to Ianos and Zorbas. By looking at the general spread trends, the INI experiment spread is highest at initialization, as also seen in Lang et al. (2012), since the EDA perturbations (INI experiment) are associated with a shift and intensification/weakening of the cyclone. Even considering the case-to-case variability, in terms of the spread of the track, the SPP-Conv experiment is the one with consistently less spread for all the cyclones.

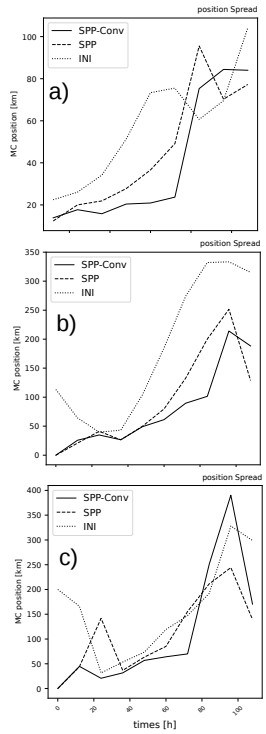

Fig. 4.3: Mean ensemble spread of the medicanes track for each ensemble perturbation experiment for Ianos (a), Zorbas (b) and Trixie (c). The track spread is computed as described in Eq. 4.2 and is reported in km. For Ianos the experiments starting on the 16/09 have been chosen, for Zorbas the ones starting on the 27/09 and for Trixie the ones starting from the 27/10, in order to be consistent with the ensemble tracks shown in Figure 4.2

The low error values obtained by Ianos and Zorbas, compared to the values reached in the case of Trixie ($\geq$ 800 km in Figure 4.2), make it not only the medicane with the largest spread but also, the most significant error. Indeed, it has been verified that this particular simulated cyclone diverges from the analysis track. As mentioned above, the Trixie simulation track is probably a result of the ensemble forecast not capturing properly the processes connected to the cyclogenesis, but it can also be related to the simulations starting too early before the medicane appearance. This is an aspect also recognized by (Di Muzio et al., 2019) in the simulation of these events.

It has to be noted here, that both experiments with SPP, and specifically the SPP-Conv one, manage to produce a spread comparable to the one produced by the initial condition perturbation. Similar results have been found in Lang et al.

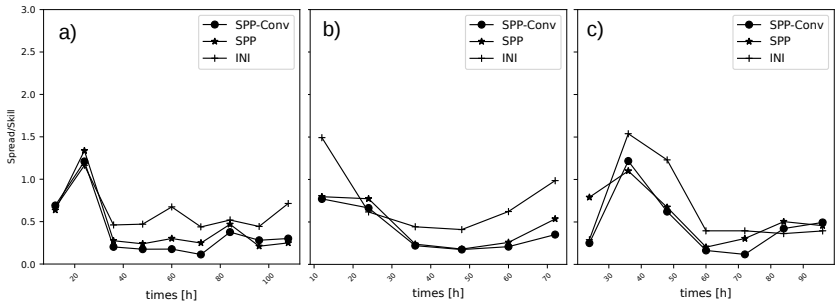

Fig. 4.4: Ensemble Spread/skill relationship for each ensemble perturbation experiment for Ianos (a), Zorbas (b) and Trixie (c). The spread is computed as in Eq. 4.2 and the skill is the root mean squared error between the ensemble mean and the analysis tracks. For Ianos the experiments starting on the 16/09 have been chosen, for Zorbas the ones starting on the 27/09 and for Trixie the ones starting from the 27/10, in order to be consistent with the ensemble tracks shown in Figure 4.2.

(2012). These results underline the importance of the role of convection and convective heating in explaining the source of uncertainty in the development of these cyclones.

However, when looking at the spread-skill relationship (shown in Figure 4.4), measured by the ratio of the ensemble spread and the root mean squared error between the ensemble mean and the operational analysis, it can be said that as mentioned before the error is always greater than the ensemble spread and they are only comparable in the first hours of simulations when the spread/skill values are nearly 1 (Figure 4.4). Both the SPP and the SPP-Conv are usually more under-dispersed than the INI experiment in all three cases, but the three experiments perform generally in the same way, with the INI experiment presenting lower error after the second day for Ianos and Zorbas, but behaving similarly in Trixie.

4.4.2.Intensity

The other important aspect investigated is the intensity of the cyclones analyzed. The intensity is assessed by studying the minimum core pressure at the cyclone position.

The development of the core pressure with the simulations is reported in Figure 4.5 where the mean sea level pressure at the centre of the cyclone (Minimum Core Pressure in Figure 4.5) is shown, with the ensemble mean, the 25%-75% percentile and the 5%-95% percentile compared to the analysis for the three ensemble experiment for chosen starting dates, as an example. For Ianos there is an overestimation of the deepening of the cyclone (Figure 4.5a, d and g), with a temporal shift of one day, compared to the analysis, which is consistent in all three experiments. For Zorbas the minimum pressure is underestimated

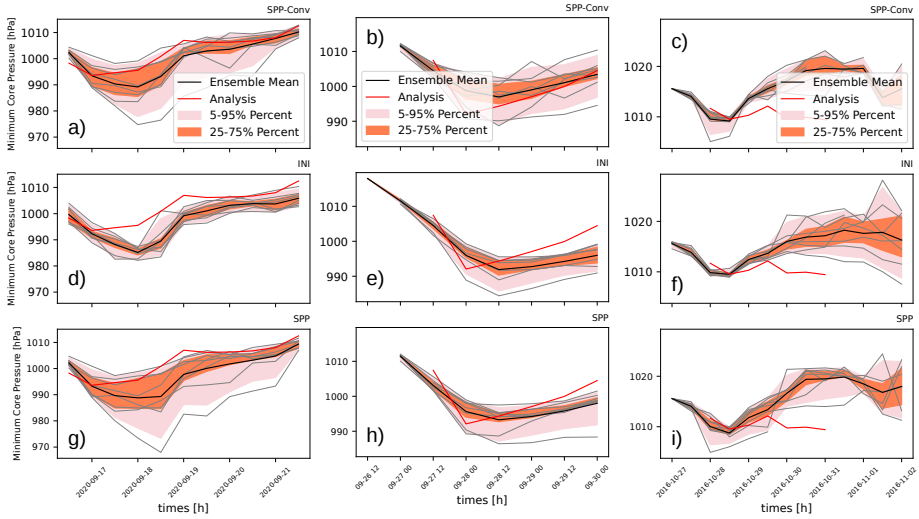

Fig. 4.5: Analysis of the mean seal level central pressure, for Ianos in the first column, for Zorbas in the second column, and for Trixie in the third column. The plots show the ensemble members' development throughout the simulation. In each Figure the ensemble mean is reported in black, the operational analysis is reported in red and the two shaded areas represent the 25-75 % percentile and the 9-95 % percentile. The SPP-Conv experiment is reported for each medicane in Figures (a), (b), and (c). The INI experiment is reported in Figures (d), (e) and (f) and the SPP experiment in Figures (g), (h) and (i). in (b), and the SPP experiment in (c).

only in the case of the SPP-Conv experiment, however, there is a shift of 12 hours compared to the analysis in the three experiments (Figure 4.5b, e and h). In the case of Trixie, there is an underestimation (Figure 4.5c, f and i) and the ensemble experiments are not able to capture the entire development of the cyclone pressure.

In all three cyclones, the spread is larger for the SPP and the SPP-Conv experiment compared to the INI one. This allows the reference analysis core pressure to be included in the ensemble spread of these experiments compared to what happens with the INI experiment. This is particularly true for Ianos and Zorbas. Indeed, the development of these cyclones intensity is best captured by the SPP-Conv experiment (Figure 4.5a and b). As pointed out before, there is a shift of the minimum pressure of 12 to 24 hours compared to the analysis. For Ianos this is also present in the simulation starting on the 15th but is absent when starting the simulation on the 17th. This is due to the improved initial conditions with the forecast start date closer to occurrence. The same can be said for Zorbas.

As for Trixie, besides the first deepening happening on the 28th of October,

which is well captured by all experiments, the second minimum, which is the deepest, is not captured by any of the members of every experiment, even if starting from the 27th of October (the last starting date). In general, there is an underestimation of the pressure minimum and the trend up to the 29th is captured better by the SPP experiment (Figure 4.5i).

4.4.3.Precipitation

The precipitation field has also been analyzed and verified against observation. The verification field is chosen to be the precipitation, as in Vich et al. (2011); Montani et al. (2011). Matching the forecast and the verifying data is made difficult by the irregularly spaced ground network of observations and by the fact that precipitation is not a continuous field, thus a point-to-point verification presents several problems and is guaranteed by using satellite data. Thus, in this work observations from satellites, specifically the above-discussed Integrated Multi-satellite Retrievals for GPM (GPM-IMERG), are used.

The precipitation structure of the three medicanes, in terms of intensity and position, is similar to what is observed in tropical cyclones (Zhang et al., 2021, 2019c). Indeed, there are similarities in their rainfall structures and those for Mediterranean cyclones (Flaounas et al., 2018) where most precipitation associated with medicanes is concentrated to the northeast side of the cyclone centre, as is shown in Figure 4.6 for the three cyclones. For the three cyclone, the ensemble mean of the last starting date is shown. For each cyclone, the daily accumulation on the day of the "tropical-like" phase is shown in Figure 4.6. For Ianos is the 17th of September, for Zorbas is the 28th of September, and for Trixie is the 28th of October. The daily accumulation values from the ensemble forecast experiments are comparable, within the error to the observed values only regarding Ianos. In general, the maximum is better captured by the INI experiment (Figure 4.6b, f, l). The standard deviation of each ensemble experiment reveals that there is higher uncertainty related to the higher values of the precipitation distribution (Supplementary Figure S1). This is consistent for all three cyclones. There is little difference between the SPP-Conv and the SPP experiment in terms of precipitation distribution. This mirrors what was shown in Figure 4.5.

The positioning of the maximum precipitation in the presented distributions (Figure 4.6), in general, is in accordance with the observed GPM-IMERG distribution. However, there are some other secondary maximums in the distribution that are not well captured, for all three cyclones. The daily accumulated precipitation shown in Figure 4.6 belongs to the simulations with the latest starting date, where the maximum precipitation is better captured. Indeed, the error in the simulated ensemble mean precipitation maximum compared to the observation is decreasing with the forecast start date closer to occurrence. This is

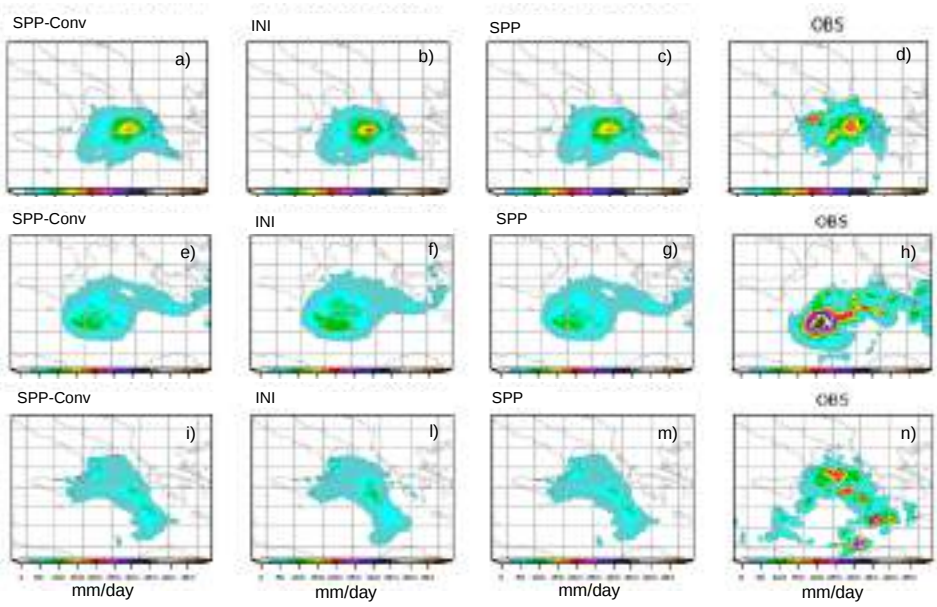

Fig. 4.6: Daily accumulated precipitation (mm/day) for the three ensemble experiments ensemble means compared to the satellite observation GPM-IMERG. For Ianos the 17th is shown in Figures (a), (b) (c) and (d). For Zorbas the 28th is shown in Figures (e), (f) (g) and (h). For Trixie the 28th is shown in Figures (i), (l) (m) and (n). the SPP-Conv ensemble forecast accumulated precipitation is reported in the first column, the SPP ensemble in the second column, the INI ensemble in the third column and the observations in the fourth column. For Ianos the experiments starting on the 17th is shown, for Zorbas the ones starting on the 27th and for Trixie the ones starting from the 27th.

specifically true for Ianos and Zorbas (as it is shown in Supplementary Figure S2), while for Trixie, the maximum is equally underestimated. For Trixie, the simulation is not able to capture the intensity of the precipitation. This is linked to the absence of the deepening of the cyclone and the precipitation starts to decrease sensibly after the 28th. This is seen within all the starting dates.
In order to evaluate the whole performance of the ensemble several probabilistic precipitation thresholds have been chosen from 1 to 150 mm, and all the ensemble members have been considered. Following the cyclone position at each timestep, an area of 500 km^2 around the cyclone centre has been considered for evaluation. The analysis has been carried out using many verification measures, following the literature (Montani et al., 2011; Marsigli et al., 2008; Vich et al., 2011; Buizza et al., 1999). A brief description of these measures is given in the Supplementary Information. All the verification results have been calculated for the entire simulation period, according to their starting date.

Starting with the Brier Score (BS) (Brier et al., 1950) it can be said that the

three ensemble experiments for the three medicanes are generally better than the control forecast at simulating the observed precipitation at all thresholds, as it is shown in Figure 4.7. Indeed, BS indicates the root mean squared error between the observed and the forecast probabilities, thus the lower the better, and in our study, the BS is higher in the control run than in the three perturbations runs.

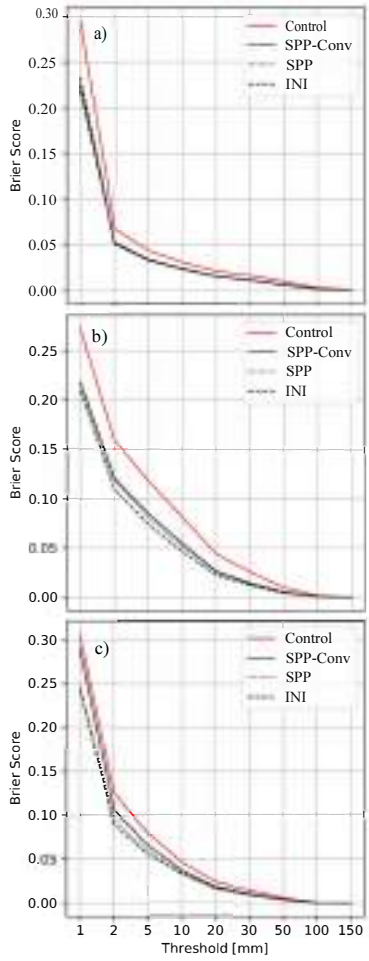

Fig. 4.7: Brier score for the SPP-Conv ensemble forecast experiment, for the SPP ensemble forecast experiment, and for the INI ensemble experiment for Ianos (a), Zorbas (b) and Trixie (c).

For Ianos, which presents the lowest values, the three perturbation ensembles are almost indistinguishable (Figure 4.7a), while for both Zorbas and Trixie the perturbation experiment SPP-Conv has a slightly higher value than the SPP and the INI experiment (Figure 4.7b and 4.7c). The difference between the three experiments decreases with increasing thresholds, and in general, the im-

provement of the ensemble experiments compared to the control run decreases for extreme events when the Brier score tends to zero as the events become increasingly rare.

To further analyze the respective skill of the three ensemble experiments, the Relative Operating Curve (ROC) has been computed and it is shown in Figure 4.8. It measures the ability of the forecast to discriminate between two alternatives outcome and thus is a measure of resolution. In general, the ROC Area shows that the SPP-Conv is better at discriminating between events and non-events than the SPP and INI perturbations in Zorbas and Ianos (but only between 10 and 50 mm), while for Trixie all three perturbations perform similarly (Figure 4.8c).

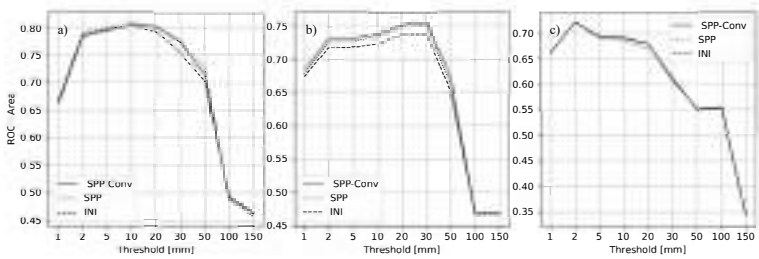

Fig. 4.8: ROC Area for the SPP-Conv ensemble forecast experiment, for the SPP ensemble forecast experiment, and for the INI ensemble experiment for Ianos (a), Zorbas (b) and Trixie (c).

The ROC area presents values above 0.70 for Ianos from 2 mm to 50 mm (Figure 4.8a), which is a good value according to Buizza et al. (1999) for a useful forecast. This is true also for Zorbas from 2 mm and 30, and for Trixie, but only for 2 mm. The ROC area value remains however greater than 0.50 until 50 mm (100 mm for Trixie). The difficulties in predicting extreme precipitation, values over 50 mm may be probably due to the grid length of 9 km. This is also possibly linked to the lack of the deepening of the cyclone in the ensemble forecast experiments, compared to the observations, as it will be shown in the next section. In general, Trixie is the simulated medicane with a lower skill for all the perturbations.

4.4.4. Thermal structure and asymmetry

To complete the analysis of these three storms, the thermal structure, and the thermal asymmetry have been investigated. The chosen parameter to quantify the latter has been recognized in the upper-level thermal wind, $-V_T^U$, which is considered to be a relevant parameter in distinguishing tropical-like cyclones from fully baroclinic cyclones (Mazza et al., 2017), and secondly on the thermal asymmetry, B. These parameters belong to the theory of the three-dimensional

cyclone phase space introduced by Hart (2003). There, the thermal asymmetry is defined as the storm-motion-relative 900–600 hPa thickness asymmetry across the cyclone within its radius:

$$B = (\overline{Z_{600\,hPa} - Z_{900\,hPa}}|_R - \overline{Z_{600\,hPa} - Z_{900\,hPa}}|_L) \tag{4.3}$$

where Z is the geopotential height, R indicates the right of current storm motion, L indicates the left of storm motion, and the overbar indicates the areal mean over a semicircle around the cyclone centre.

Instead, the cyclone's upper-level thermal structure (i.e., its cold or warm core) is indicated by $-V_T^U$. If it attains a positive sign, the cyclone attains an upper-level warm core. Indeed the $-V_T^U$ is defined as:

$$-|V_T^U| = \frac{\partial(\Delta Z)}{\partial ln(p)}^{300\,hPa}_{600\,hPa} \tag{4.4}$$

The two pressure levels have been changed from 600 hPa to 700 hPa and from 300 hPa to 400 hPa due to the lower height of the tropopause in the mid-latitudes with respect to the tropics (Picornell et al., 2014). These values are computed within a 200 km radius around the detected cyclone center. In the Hart (2003) formulation, this radius was chosen to be 500 km, but given the smaller size of medicanes compared to tropical cyclones (Miglietta et al., 2013), the radius used in this study is smaller, 200 km. The choice to consider important only the upper-level structure, compared to the lower-level (thus using also the lower-level thermal wind $-V_T^L$) is because its positive value can characterize not only medicanes but also extratropical cyclones with warm seclusion (Hart, 2003). In the case of the thermal asymmetry, the threshold value of $B = 10\,m$ has been determined by analyzing ECMWF reanalyses ERA40 at 1.125° of the resolution, from which no major hurricane (winds of greater than 210 km/h) had associated with it a value of B that exceeded 10 m (Hart and Evans, 2001). Even if the threshold value of B has been originally determined for larger tropical/extratropical cyclones, previous studies have shown that such value is also useful in the case of medicanes (Miglietta et al., 2011).

The Hart parameters, $-V_T^U$ and B, are analyzed in Figure 4.9 for Ianos, in Figure 4.10 for Zorbas and in Figure 4.11 for Trixie. The focus has been put on the values where the deepening of the cyclone occurred in the observation, thus the values in Table 4.6 are reported as references in Figures 4.9-4.11. There, the violin plots of the forecast distribution for each experiment at different starting dates are shown, with the median (white dot) and the interquartile range (gray bar). The sections at the sides of each violin plot represent the kernel density estimation to show the distribution shape of the data (wider sections of the

violin plot represent a higher probability that members of the population will take on the given value; the skinnier sections represent a lower probability). The choice for the violin plot has been made because, unlike a box plot that can only show summary statistics, violin plots depict summary statistics and the density of each variable.

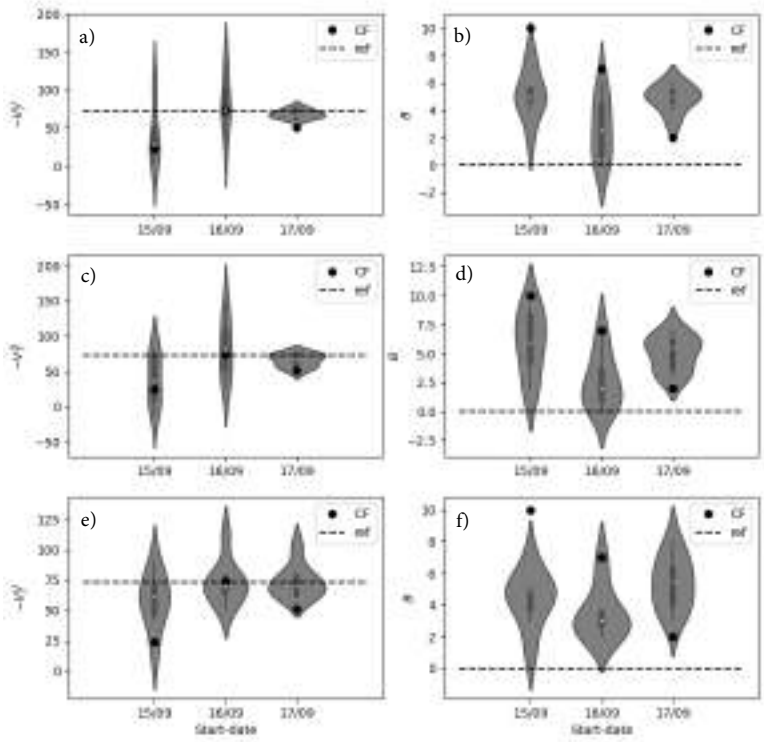

Fig. 4.9: Ensemble forecasts violin plots of thermal wind, $-V_T^U$, and thermal symmetry, B, for each start date for the Ianos storm. The violin plot is a hybrid of a box plot and a kernel density plot, which shows peaks in the data. The white dot represents the median, the thick gray bar in the center represents the interquartile range (25th-75th) and the thin gray line represents the rest of the distribution. On each side of the gray line is a kernel density estimation to show the distribution shape of the data. At the top, in Figure (a) and Figure (b) the SPP-Conv ensemble experiment violin plots are shown, in the middle, in Figure (c) and Figure (d) the SPP ensemble experiment violin plots are shown, and at the bottom, and in Figure (e) and Figure (f) the INI ensemble experiment violin plots are shown.

The spread of both the thermal wind and asymmetry does not present a monotonic behavior, a decrease or increase with different start-dates in all three storms. In the case of Ianos, it decreases with the start date for the thermal wind (Figures 4.9a, 4.9c and 4.9e.), but it is comparable for the thermal asymmetry, as shown in Figures 4.9b, 4.9d, and 4.9f. Zorbas exhibits a usually

smaller spread at the last start date compared to the previous ones regarding the thermal wind (Figure 4.10a and 4.10e). Similarly to what happens for Ianos, there is a comparable spread for the thermal asymmetry, as can be seen in Figure 4.10b, 4.10d and 4.10f. In Trixie, the spread is once again similar for the thermal asymmetry parameter (Figures4.11b-d-f) and it oscillates between the three start dates regarding the thermal wind, as it is shown in Figures 4.11a-c-e.

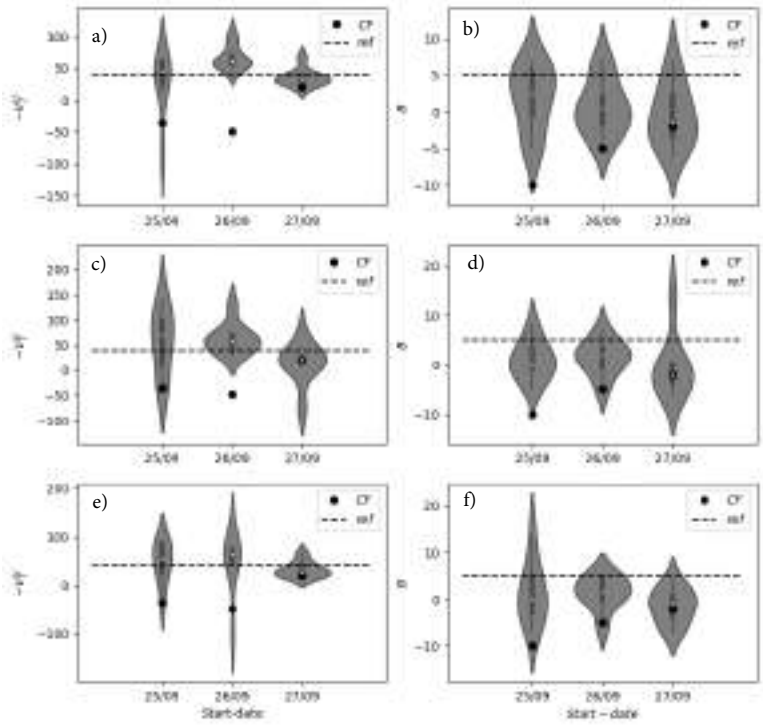

Fig. 4.10: Ensemble forecasts violin plots of thermal wind, $-V_T^U$, and thermal symmetry, B, for each start date for the Zorbas storm. The violin plot is a hybrid of a box plot and a kernel density plot, which shows peaks in the data. The white dot represents the median, the thick gray bar in the center represents the interquartile range (25th-75th) and the thin gray line represents the rest of the distribution. On each side of the gray line is a kernel density estimation to show the distribution shape of the data. At the top, in Figure (a) and Figure (b) the SPP-Conv ensemble experiment violin plots are shown, in the middle, in Figure (c) and Figure (d) the SPP ensemble experiment violin plots are shown, and at the bottom, and in Figure (e) and Figure (f) the INI ensemble experiment violin plots are shown.

The forecast distribution in the three ensemble experiments presents different behaviors in comparison with the analysis reference value (ref in Figures 4.9-4.11) and with the start date in the three experiments.
Ianos shows a reduced spread with later start dates accompanied by the dis-

tribution value approaching the analysis value. This is true in particular for the SPP-Conv (Figure 4.9a) experiment and for the SPP one (Figure 4.9c). Nonetheless, all three experiments distribution $-V_T^U$ values manage to be close to the analysis one. Regarding the thermal asymmetry parameter the distribution spread, which is only slightly lower in the third start date, the latest, is accompanied by a consistent slight overestimation (B around 4-6) of the analysis values ($B = 0$). This is verified for all three sets of ensemble experiments (Figures 4.9b-d-f). However, even with this underestimated by the ensemble forecast experiments, the median values of the distributions are always below the threshold of 10 m which would define the system as a frontal system rather than a non-frontal one (Miglietta et al., 2013).

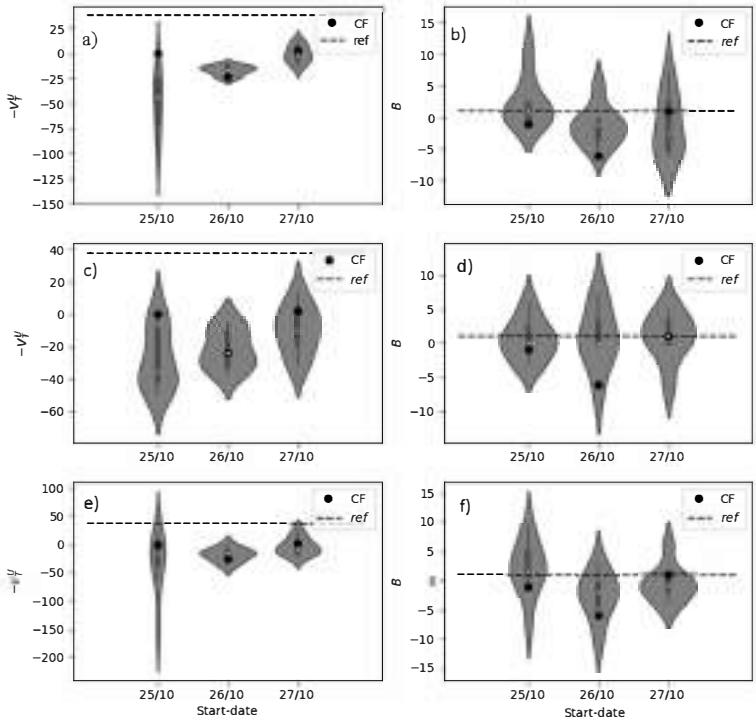

Fig. 4.11: Ensemble forecasts violin plots of thermal wind, $-V_T^U$, and thermal symmetry, B, for each start date for the Trixie storm. The violin plot is a hybrid of a box plot and a kernel density plot, which shows peaks in the data. The white dot represents the median, the thick gray bar in the center represents the interquartile range (25th-75th) and the thin gray line represents the rest of the distribution. On each side of the gray line is a kernel density estimation to show the distribution shape of the data. At the top, in Figure (a) and Figure (b) the SPP-Conv ensemble experiment violin plots are shown, in the middle, in Figure (c) and Figure (d) the SPP ensemble experiment violin plots are shown, and at the bottom, and in Figure (e) and Figure (f) the INI ensemble experiment violin plots are shown.

In the case of Zorbas, the forecast distribution of three experiments contains the reference analysis values for both the thermal wind and the thermal asymmetry (Figure 4.10). however, while for $-V_T^U$ the spread of the distribution is lower, especially for later start dates, thus the median values compare well with the analysis value, for B, the spread is greater and thus the median values of the distribution underestimate the analysis value in most cases (see Figures 4.10b-d-f), with outliers associated to a value of B greater than 20. For Trixie, instead, there is a consistent underestimation (with $-V_T^U$ around -25 to -50) of the thermal wind, in all three ensemble experiments. The forecast distribution is closer to the analysis values ($-V_T^U = 39$) for the third start date, especially for the INI experiment, as it is shown in Figures 4.11a-c-e. For asymmetry, the forecast distribution contains the analysis value for all three ensemble experiments, similar to Zorbas. However, differently from Zorbas the value here is only slightly underestimated, with the median value of the distribution closer to the analysis value (see Figures 4.11b-d-f).

Generally, the three ensemble experiments are able to reproduce the thermal structure of the storms, showing in some cases decreased spread with increasing start date, probably linked to the fact that these start dates coincide with the period in which the cyclone had already developed (see Figures 4.9a-c-e for Ianos). This behavior is consistent for all three perturbation experiments. However, while for Ianos the spread is lower in the SPP and SPP-Conv experiments than in the INI experiment, for Zorbas and Trixie, the spread is lower in the SPP-Conv and INI experiment, than in the SPP one. Furthermore, for Trixie, the thermal structure is underestimated (Figure 4.11), meaning that a warm core was never reached in the ensemble forecast experiments. Looking at the three start dates used in this study they are probably too early to produce the deepening of the cyclone, as also underlined previously when looking at the central pressure MAE and ME and the low precipitation scores. Furthermore, the limited intensification of Trixie, compared to Ianos and Zorbas, could be limiting the proper simulation of the cyclone itself.

Regarding thermal asymmetry, it is usually predicted quite well by the three ensemble forecasts experiments, given the reference analysis values lie in most cases in the interquartile range of the forecast distributions. In the case of Ianos, there is an underestimation of the asymmetry, with the storm attaining lower thermal asymmetry than in the analysis, whereas for Trixie and Zorbas the forecast distribution deviates less from the reference values. There is a slight reduction of the spread in the last start date, as in Di Muzio et al. (2019), especially for Ianos, given that, the appearance of a symmetrical storm can benefit from improved initial conditions. Differently from the thermal structure, in this case, the three ensemble forecasts experiments behave similarly in terms of spread and simulation of the reference values. Lastly, the presence in some

cases of a bias between the distribution of the ensemble forecast and the control run (as shown for example in Figure 4.10a and c), could be related mainly to the small ensemble size.

4.5.Conclusion

Predicting and simulating medicanes is a difficult task due to them being extreme events found near the tail of the forecast distribution (Majumdar and Torn, 2014) and due to the complexity of the processes involved. The specific barriers to predictability and the atmospheric conditions that lead to medicanes formation and evolution are not fully understood. Research utilizing multi-physics approaches has found that the track, intensity, and duration of medicanes are heavily sensitive to factors such as convection, microphysics, and boundary-layer parameterizations (Ragone et al., 2018; Miglietta et al., 2015). This study indeed is a first step towards the understanding of this sensitivity by using ensemble forecast simulation.

The analysis has been focused on three medicanes with the use of the ECMWF model IFS ensemble forecast system. Three experiments were conducted, the operational ensemble forecast with the perturbation at initial conditions, INI, and two experiments with SPP, in one case perturbing only the parameters of the convection parameterization, SPP-Conv, and in the other perturbing the parameters of all relevant physical parameterizations, SPP. The used approach has been aimed at the analysis first and foremost of tracking and intensity, meaning central pressure. Secondly, the precipitation field has been analyzed, and finally, the focus has been put on the parameters characterizing the thermal structure of the cyclone and its asymmetry.

From the study, it was found that the use of SPP, and specifically the perturbation of convective parameters (SPP-Conv), in terms of tracking spread and precipitation intensity compare well to the INI experiment. Similar results have been obtained also in Ollinaho et al. (2017b), especially regarding precipitation. The three experiments are usually under dispersive when it comes to tracking position, even if the INI experiment is giving slightly better results. It has to be noted that the ensemble spread and mean are usually lower in the SPP experiments. Nonetheless, the latter can produce the same spread as the initial condition perturbation experiment, underlying the benefit of the introduction of perturbations of physical parameters, especially regarding convection, in comparison with only using the perturbations of the initial conditions (Lang et al., 2012; Ollinaho et al., 2017b).

Regarding the reproduction of the minimum central pressure by the ensemble forecasts (Figure 4.5) it is found that in general there is a time shift in the reproduction of the minim intensity, where both Ianos and Zorbas ensemble mean

reach the maximum intensity with a delay. Since the delay is decreasing with the forecast start date closer to the occurrence, this is attributable to improved initial conditions. Regarding Ianos and Zorbas, the ensemble spread for the SPP experiments is always able to include the analysis one, compared to the INI experiment (Figure 4.5).

By looking at the simulated precipitation distribution, compared to the GPM-IMERG satellite observations, show that only for Ianos, the precipitation field is well reproduced by the ensembles, with the INI experiment being slightly more intense than the SPP and SPP-Conv experiments. Even if the GPM-IMERG dataset, tends to overestimate precipitation over the Mediterranean, the precipitation simulated for Zorbas and Trixie is too weak. However, it is discussed that in general, the precipitation maximum compares better with the forecast start date closer to the occurrence, thus for Zorbas, starting the ensemble forecasts on the 28th would have shown better-simulated precipitation. However, the precipitation maximum values are slightly better simulated by the INI experiment also for Zorbas. The analysis of the precipitation field deepened by computing some statistical scores and the results show that the SPP-Conv is slightly better than the SPP and INI experiment. The BS score results are quite similar between the three experiments and generally show how the three ensemble forecasts perform better than the control forecast. The ROC-Area results (Figure 4.8) instead show that the SPP-Conv and SPP ensemble experiments present higher values than the INI one. Furthermore, all three forecasts present useful values in terms of precipitation prediction (Buizza et al., 2008), being higher than 0.5 and in most cases higher than 0.7. The most significant values, occur in Ianos for a wide range of precipitation thresholds (from 2 to 50 mm), making it the most well-captured storm of the three.

Regarding the tracking, the precipitation, and the intensity of the cyclone there is a common gradual decrease of the error with the forecast start date closer to the occurrence of the medicane. This is consistent for all three ensemble experiments and it is specific for Ianos and Zorbas. Usually, the later starting forecasts tend to be more accurate (lower error and spread). For Trixie, this decrease in the error is weaker and in some cases nonexistent. Indeed, as pointed out in the results the simulation of this particular cyclone can be considered a missed forecast.

Through tracking, it was verified that with regard to the Trixie simulation, there is a southeastward shift of the trajectory with respect to the analysis and, that, among the three storms, the best predicted is Ianos. The southeastward shift was identified after the analysis of the ensemble spread and error. A lower spread in the tracked position was found for Zorbas and Ianos compared to Trixie, but the error of Trixie exceeds 800 km. This is attributable to the fact

that the ensemble forecast starts too early with respect to the cyclone intensification phase for Trixie. This result, in particular, aligns with previous studies that had pointed to Trixie's low occurrence probability up to two days earlier (Di Muzio et al., 2019). Indeed, There is an inherently low probability of medicanes occurrence (as seen in Di Muzio et al. (2019)) and the development of a warm core cyclone depends on many factors, large-scale factors and small-scale factors like surface fluxes, that can be improved by the initial conditions of a preexisting cyclone. The decrease of the spread at later starting dates in Trixie can be correlated with the higher probability of occurrence and the better reproduction of the physical processes leading to cyclone intensification.

By the analysis of the cyclone thermal structure and asymmetry, it is shown that if for Ianos and Zorbas the tropical-like phase is reached, with a gradual evolution of the error and spread, where there is a lowering of the spread accompanied by the distribution approaching the reference values (upper-level thermal wind in Figure 4.9 and 4.10). In those cases, the forecasts are accurate at reproducing both the thermal structure and symmetry of the cyclones, compared to the analysis value. Both Zorbas and Ianos deep warm core phase is well captured, especially with the forecast start date closer to the occurrence. The reduction of the spread with the forecast start date closer to occurrence is in line with what was found in the literature, specifically for Zorbas. Indeed, in Portmann et al. (2020) it was found that later forecasts can capture better the thermal structure of Zorbas, due to the lower uncertainty of the positioning of potential vorticity streamer connected to the generation of the medicane. This aspect will be explored in the following work, especially in relation to the SPP perturbations.

The only exception to the reasonable reproduction of the storm thermal structure is the upper-level thermal wind in Trixie (Figures 4.11a-c-e). This means that for Trixie, where the analysis value reported the presence of a warm core ($-V_T^U = 25$ in Table 4.6), the three ensemble forecast presents a cold core cyclone (negative $-V_T^U$ values). This, together with the underestimation of the deepening of the cyclone, can explain the low scores for the precipitation analysis since the simulated cyclone is not able to reach the warm core. It has to be noted here that also the limited intensification (core minimum pressure comparison in Table 4.6) of Trixie may have represented a limit in its simulation by the ensemble.

Moreover, the analysis of the thermal structure and symmetry indicates that for the upper-level thermal wind, the smaller spread is in most cases obtained in the SPP-Conv experiment (in particular at later forecasts), while the spread is comparable for the three ensemble experiments regarding the thermal symmetry. When looking at the distance between the ensemble mean and the

reference values in Figures 4.9-4.11, there seems to be a similarity between The INI ensembles and the two SPP ones. This, similarly to the spread/skill relationship presented for the tracking, points out the possibility of using SPP in ensemble forecasting to generate a similar spread/error in the ensemble distributions compared to only using initial conditions perturbations. Indeed the ensemble experiments behave similarly regarding the simulation of the tropical-like characteristics of the three cyclones. Furthermore, the SPP and SPP-Conv experiments are found to be extremely similar, possibly meaning that the uncertainties linked to the convection parameterization are predominant in the simulation of these types of phenomena. It would also be interesting, as future development, to look into the combined effect on spread and error of initial conditions and model perturbation, to get the complete ensemble set-up.

Considering all the parameters used in this analysis, it can be assessed that the ECMWF ensemble forecasts model can adequately reproduce medicanes with their tropical-like features. The limitations of this work are related to the fact that only 8 members and three start dates are used for each forecast. Nonetheless, the result of this study, compared to previous ones (Di Muzio et al., 2019), which used only the operation ensemble forecast of ECMWF, points out the benefit of using the Stochastically Perturbed Parameterization, SPP ensemble forecast compared to only perturbing the initial conditions. The SPP and especially the SPP-Conv experiments show lower but comparable spread and better precipitation scores than the INI experiment. It is also pointed out that there is an improvement in the simulation of medicanes tropical traits by initializing the simulation when the cyclone already exists, due to improved initial conditions. The expansion of this work will be aimed at investigating the physical mechanism behind these results, especially regarding the deepening of the cyclone, its tracking, and its transition to the tropical phase. This will be done by examining the interplay between the physics perturbations and the affected processes at different spatial and temporal scales leading to the medicanes formation and evolution. Furthermore, since it is known that there are different categories of medicanes Miglietta and Rotunno (2019), an investigation will be done to assess to which category these medicanes belong and to expand the ensemble experiments also considering a greater number of medicanes, to cover all categories.

5. Conclusion

5.1.Summary of the chapters

This work tackles different types of modeling of the water cycle budget of the atmosphere. In the second chapter, the modeling is applied to connect the atmosphere water cycle, by means of precipitation and evaporation, to inland water bodies' mass balance. In the third chapter, the modeling is done with Cloud Resolving Models, to study the climate feedback connected to the organization of convection (the instability of the Radiative Convective Equilibrium) in order to understand possible common feedback of this type of model when simulating convective motions. In the fourth chapter, the modeling is carried out with a General Circulation Model, through ensemble forecasting, to look at the simulation of mesoscale convective systems, Mediterranean tropical-like cyclones focusing on how the cyclones feature are reproduced, especially when the uncertainty of the convective parameters is taken into account.

Chapter 2 is concerned with the use of ERA5 reanalysis for the modeling of water level proposed for evaluating the water budget for small, shallow lakes. The modeling has been applied specifically to the case study of the Trasimeno lake in the Umbria region, Central Italy, for the period 1991-2019. The model has been developed to answer the first research question. Indeed, the model can be used not only to reproduce the water mass balance of a shallow lake in the past, but can also be used to look at the impact of climate extremes on the water balance, thus informing also future possible climate change effects on the lake. This is particularly true if one considers the link found in the chapter regarding the dramatic change (increase and decrease) in lake levels and two periods of climate extremes, the intense heat wave during 2003-2005 which decreased the

lake levels to their ever-recorded minimum (-1.9 m), and the period of extreme precipitation (2012-2013) which made the lake levels rise around 0.30 m in one month.

The modeling is based on the basic assumption that precipitation and evaporation are the fluxes that mostly influence the water balance of a closed lake and it is evaluated on a monthly time scale. The model can adequately reproduce the lake water level, and there is a strong linear link between the observed variation of the lake level, ΔH, and the values of the water storage, δ, evaluated as the difference between precipitation and evaporation. However, some discrepancies arise precisely in two periods characterized by an extreme climate anomaly: the drought 2003-2005 and the intense precipitation period 2013-2016, when reanalysis overestimates and underestimates precipitation, respectively, and the groundwater inflow (not considered in the model). This last aspect has been attributed to the low resolution of ERA5 Land, which even if representing one of the best-performing reanalyses, does not resolve small-scale processes but parameterized them, affecting the water mass balance modeling. More specifically, the results here presented underline that in Central Italy the high resolution of ERA5 and ERA5 Land is not able to capture the complex orography of the region and the processes connected to the latter. However, the proposed model allows capturing the monthly variability of the level of the Trasimeno lake.

Indeed, compared to other possible approaches, what makes the proposed one relevant is the fact that it uses a single source of climate data (i.e., ERA5 Land), which has the advantage of being spatially consistent and being the only global reanalysis including a lake model (i.e., Fresh-water Lake model, FLake), that provides the lake heat balance and, then it guarantees the reliability of both the evaporation and precipitation fluxes.

In Chapter 3 the second research question has been tackled by means of performing two Radiative Convective Equilibrium simulations with two different Cloud Resolving Models (SAM and ARPS) and comparing their properties while reaching a statistical equilibrium of precipitation. The study has been specifically aimed at studying the convective organization occurrence in the two models, the instability of RCE, and the mechanism leading up to it. The study of the differences between the two simulations has been carried out by looking at many aspects of the simulations, the total condensed water, and precipitation distributions, the standard deviation of the column relative humidity, the organization indexes, the cloud fraction, cloud condensate and radiative heating profiles. In particular, it has been evinced that the two models reached an organized state presenting different warming and drying of the whole domain.

This has been linked to the different properties of the sub-cloud layer in the two models and the effect that this has on the cold pools and the feedback that helps maintain convection organization. The latter has been assessed by computing the frozen moist static energy budget and its variance, looking at the competing effects of moisture and radiative feedback affecting the convective organization. In the study, it was found that in SAM, longwave radiative feedback is important in the initial stages of the organization, while an up-gradient transport of moist static energy is necessary for the maintenance of the organized state. on the other hand, in ARPS the radiative feedback induced by high clouds is the main feedback sustaining convective organization, together with the moisture convection feedback localizing convection.

These results are found to be connected to the cold pool effect. The latter is regarded as a negative moisture feedback to convective organization, since it produces a down-gradient transport of moisture in dry regions, therefore slowing down convective aggregation. On the other hand, if cold pools, in moist regions, and radiative subsidence, in dry regions, are strong enough to generate low clouds at the boundary between moist and dry regions, then an up-gradient low-level transport of moist static energy from dry regions to moist regions develops by an indirect positive longwave feedback. This mechanism is then responsible for maintaining convective organization.

Therefore, evaporatively-driven cold pools can help the expansion of radiatively-driven cold pools, as a secondary effect, by triggering low clouds in dry regions. In ARPS, the sub-cloud layer becomes saturated, causing less rain evaporation, weaker downdrafts, and weaker cold pools, thus allowing the localization of the convection in moist regions by weakening the negative influence of cold pools. Stronger cold pools in SAM, instead, help the creation of shallow clouds in dry regions, increasing the longwave feedback responsible for their expansion; while delocalizing convection in moist regions and therefore opposing high-cloud radiative feedback. The main finding is that negative moisture feedback in ARPS can lead to lower values of aggregation metrics compared to SAM, mainly due to the strong drying of dry regions in the SAM model which corresponds to stronger warming and drying of the domain.

Lastly, Chapter 4 is concerned with answering the third research question by performing ensemble forecast simulations of medicanes. A comparison between three ensemble forecast experiments has been set up, by using the ECMWF IFS ensemble forecasts. Three different ensemble forecasts have been performed by using perturbation of the initial condition only, by Ensemble Data Assimilation and by perturbing parameters belonging to the convective parameterization only, and by perturbing parameters belonging to all physical parameterizations, by means of the Stochastically Perturbed Parameterizations. The chosen med-

icanes under study are among the strongest in recent years: Ianos, Zorbas, and Trixie. The study is carried out to investigate the simulation of the key measurements for a cyclone. First of all the central pressure error and track position spread is analyzed. It is found that the use of SPP and specifically the perturbation of convective parameters shows comparable results to the use of initial conditions perturbations, especially regarding the spread/skill relationships for tracking and by looking at the analysis of the central minimum pressure. Thus, the change in the convective parameterization alone (changes in convective heating in the simulations) is able to produce the same spread as introducing errors in the initial conditions.

The analysis then passed on to precipitation with the use of probabilistic statistical scores and once again showed that the precipitation characteristics have been reproduced well and the ensemble with convective parameterization perturbation performed better that the other two experiments. Generally, the three ensemble forecasts performed better than the control run and the IFS ensemble forecasts produced useful values in terms of precipitation prediction. It is also found that there is a link between the error present for central pressure, the low scores for precipitation, and the absence of a warm core in the ensemble forecasts for Trixie. Finally, the cyclone thermal structure and symmetry have been analyzed in the simulations. The three ensemble forecasts experiments exhibit a non-gradual evolution with increasing start-date. In some cases, there is a lowering of the spread accompanied by the distribution approaching the reference values, the operational analysis. However, the forecasts are pretty accurate at reproducing both the thermal structure and symmetry of the cyclones, compared to the analysis value, with the smaller spread and error in most cases obtained in the perturbed convective parameterization experiment for the upper-level thermal wind.

5.2.Concluding remarks

The model in the second chapter, even with the mentioned shortcomings, allows for capturing the monthly variability of the level of the Trasimeno lake and could be applied to other shallow lakes with similar features, and therefore makes reanalyses an alternative to other methods for modeling the water mass balance of inland water bodies. Before applying this method to other shallow lakes, the potential impact of factors other than precipitation and evaporation needs to be evaluated: inflow and outflow, specifically considering groundwater, lake ice cover, and the presence of tributaries, e.g. rivers. The performance of the model could improve by including lake withdrawals, recharges, and specifically groundwater inflow. On one hand, the outcome of this work urged the Umbria regional authorities to increase the data and their reliability, around the lake area, by constant monitoring of lake withdrawals, groundwater inflow (by a piezometric network), and level observations. Regarding this last aspect, the

addition of at least one more weather station (including a hydrometer) on the opposite side of the lake, as the literature suggests (especially for accounting for the wind effects on lake levels), has been requested. On the other hand, the work is now proceeding to the application of the modeling to other lakes in different regions of the world, specifically trying to account for groundwater and surface runoff.

The third chapter points out that the saturation of the sub-cloud layer in CRMs allows the localization of the convection in moist regions by weakening the negative influence of cold pools. This negative moisture feedback in a cloud-resolving model can lead to lower values of aggregation metrics and to a weaker effect of the organized state on the average domain statistics (less drying). There is a chance that this effect, found in these two models, could be found also in other CRMs. Thus, the research will be expanded with the analysis of the whole CRMs used in the RCEMIP project, by correlating the degree of organization in the models and the degree of saturation of the sub-cloud layer. The other possible extension of the work will be relating the cold pool effect and presence to the mixing in the boundary layer, given their known tight relationship in the literature. Moreover, it has to be noted here that the two cloud-resolving models represent two different states of the tropical climate. While the SAM model, is more similar to what is observed in the tropics, regarding the radiative heating profile, it also presents an unrealistic amount of low clouds. In the same way, the ARPS model presents an unrealistic amount of high clouds, but its greater imbalance between the radiative cooling and the surface fluxes makes it closer to the RCE equilibrium observed in the tropics on a daily basis. This underlines the fact that both models are contained in the statistical oscillation of atmospheric tropical convection and therefore could be included in climate sensitivity studies.

The fourth chapter assessed that the ECMWF ensemble forecasts model can adequately reproduce medicanes with their tropical-like features, even with the use of only 8 members and three starting dates. It also points out that the inclusion of the uncertainty in the physical, and more specifically, in the convective parameterization, leads to an improvement in the simulation of these events. Indeed, the ensemble forecasts with the SPP, generally show similar results to that obtained with the initial condition perturbation ensemble forecast. There is an improvement in the simulation of medicanes tropical traits by initializing the simulation when the cyclone already exists, due to improved initial conditions. Indeed, specifically in the case of Trixie, the early start produced a southeastern shift of the cyclone compared to the analysis track and a lack of reproduction of the warm core transition. The expansion of the work will be aimed at understanding the physical processes behind these results. In specific, the research objective will be to understand which processes in the model physics are behind the intensification of these three storms and their transition

to the tropical-like phase or lack thereof (specifically in the Trixie case). This will be carried out with a specific investigation of the effect of the physical parameterizations perturbations, focusing on convection and convective heating.

The work conducted in this thesis used a different approach from the historically used approach of simulating phenomena at the synoptic scale (cyclones) and then descending to the convective scale with more observationally and numerically resolved modeling. In contrast, the use of global-scale modeling (reanalysis) by including local data and the use of non-hydrostatic modeling with prognostic equations of moist variables, made it possible to study convective phenomena and the resulting statistical variability by including processes occurring at the microphysics and turbulence scales in the energy balance in the scales solved by the energy flux conservation equations. The approach used follows the strand of global modeling that seeks to build a digital twin of the Earth to study its climate and climate variations by including in the modeling every relevant aspect of the processes occurring in it, and in particular all the scales that contribute to the spectrum of climate variability that results from convective processes. Some of the climate feedback that makes convection-related events more extreme can be derived from the modeling that is the result of the research conducted and will be the subject of further research.

Funding: "This research has been funded by the Italian Ministry of University and Research (MIUR) and University of Perugia within the program *Dipartimenti di Eccellenza 2018-2022*."

Data Availability Statement:
Regarding the modeling of the lake water mass balance modeling, the lake level data that support the findings of the study are freely available at `https://annali.regione.umbria.it/`, the LSWT dataset GloboLakes is freely available on the CEDA (Centre for Environmental Data Analysis) Archive (Carrea and Merchant, 2019) and the ERA5 reanalysis data are freely available on the Copernicus Climate Data Store (C3S, 2020).

Regarding the Cloud Resolving Modeling, the SAM and the ARPS models output used in the study are publicly available via the Zenodo online repository. The SAM output is available at `https://doi.org/10.5281/zenodo.7310136` and the ARPS output is available at `https://doi.org/10.5281/zenodo.6953873`. The SAM model source code is freely available and can be found at `http://rossby.msrc.sunysb.edu/~marat/SAM.html`. The ARPS model source code is freely available and can be found at `https://arps.caps.ou.edu/arpsdown.html`. The Jupyter Notebooks used to produce the Figures and Tables are available at `https://doi.org/10.5281/zenodo.7357619`.

Regarding the ensemble forecast simulations, the output of the IFS ensemble forecasts and the operational analysis data are available upon request on the MARS Archive at `https://www.ecmwf.int/en/forecasts/dataset/ecmwf-research-experiments` and the precipitation dataset GPM-IMERG used for validation is freely available on the NASA data archives at `https://gpm.nasa.gov/data/imerg`.

Acknowledgments

I would like to express my sincere gratitude to both my supervisors, Prof. Silvia Meniconi and Dr. Paolina Bongioannini Cerlini, for the great support and guidance they provided throughout these years and for the patients with which they have followed me. I am especially grateful to Dr. Paolina Bongioannini Cerlini, who took me under her wing since the day we met after my master's graduation, and thought me everything I know about atmospheric convection. By giving me the chance to undertake a Ph.D. in atmospheric physics, she opened a door that I never thought was possible to be opened, allowing me to pursue my passion at the same University that raised me as a physicist. I am also grateful for their unwavering dedication and insightful discussions over the past three years that helped develop my research. I must also thank Prof. Bruno Brunone for his invaluable scientific input in my research. I would like to thank my international supervisor, Dr. Peter Bechtold, profusely, because he was able to support me throughout my thesis even from afar and during a delicate period such as the pandemic. I also want to thank him for his dedicated attention and the vast knowledge he has tried to pass on to me during the crucial final months of my work. Lastly, I want to thank Dr. Lorenzo Silvestri, who was already enrolled in the Ph.D. program when I arrived and has been constant support not only by providing technical help but also by always keeping the door open for discussion and doubts throughout these three years of Ph.D.
During the last three years, many things in my life have changed and I have grown tremendously, both professionally and in character and strength. For this, I must thank many people in my life, my Ph.D. fellow colleagues, my friends, my family, and my partner. I want to thank the group of Ph.D. colleagues, that became my work family, because without you, I would never have had so much fun during these years, and being together in our small kitchen or going out in the evenings, made every bleak day pass. To my friends, you know who you are, both those nearby and those far away, my thanks go to you because I have always felt everybody close and supportive, especially in the periods when I was busier and overwhelmed by work. A special thanks go to my dear friend Sara who has never avoided endless weekly calls to alternatively cry or laugh together. I also want to thank my family, my mother Silvia, my father Alfonso, and my brother Michele, for all the constant support I have felt during these years of ups and downs, a little part of my success is also yours. Finally, I want to thank my partner Mattia because, without him, I would not be who I am today, not only because you taught me that there is another way to see and experience life, but because you are the best adventurer companion I could ask for, and I am fortunate to call you as such.

A. The FLake model equations

The Fresh-water Lake model (FLake) (Mironov et al., 2010) is a one-dimensional model representing fresh-water lakes, which predicts their vertical temperature structure on timescales from hours to years, for lakes of various depths (up to 60 m). FLake has been implemented on IFS in subsequent steps, with some degree of simplification given that the IFS is a global model which can't describe global lakes in detail. FLake has been implemented and tested at a global scale in the IFS, (Dutra et al., 2010) and its forecast sensitivity has been assessed (Balsamo et al., 2012). The introduction of FLake in the IFS has permitted the treatment of lakes, not as ocean points, where only grid-scale lakes (lake fraction larger than 50%) were considered, but as individual entities. With FLake, inland water bodies that are fully or partially covering the grid box are considered and modeled prognostically, evolving with the same grid resolution and timestep of the atmospheric and land surface processes in the IFS. A new tile has been added to the land surface model H-TESSEL to represent the lake interface with the atmosphere (Figure A.1).

The vertical temperature profile specified in FLake consists of a mixed layer at top of a bottom layer as it is shown in Figure A.1. In the mixed layer, there is a uniform distribution of temperature, and at the bottom layer, there is a thermocline, with its upper boundary located at the mixed layer bottom, and the lower boundary at the lake bottom.The prognostic variables are the mixed layer depth $d_{ML}(m)$ and temperature $T_{ML}(^{\circ}C)$, the mean layer water temperature $T_{MN}(^{\circ}C)$ and the bottom layer water temperature $T_{BL}(^{\circ}C)$, the shape factor ξ_{ML} adimensional.

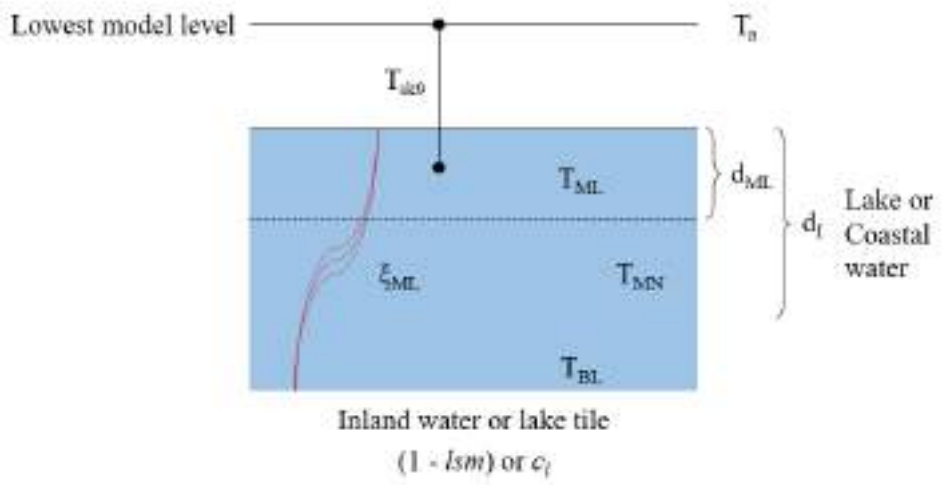

Fig. A.1: Schematic representation of the new lake tile modeled by Flake in the IFS. The lake is parameterized in two layers, a mixed layer at the top with uniform temperature and a bottom layer with a thermocline. The skin temperature$T_{s,k}$ of the lake is then passed to the atmospheric model to compute surface fluxes.

The parameterization for the mixed layer depth is done by taking into account the mixing that is generated by wind and by volumetric heating by solar radiation (Mironov, 2008). The parameterization for the thermocline is done according to the self-similarity concept (assumed shape of a temperature-depth curve) (Mironov, 2008). This assumption is based on observational studies.

FLake requires as input the lake location and the lake depth and the lake initial conditions. It runs over every grid box or the tile fraction with no bottoms sediments and no turbulent flux, with fractional ice on the surface and snow parameters used mainly for albedo calculation. The two-layer parameterization of the temperature profile is:

$$T_l(z) = \begin{cases} T_{ML} \quad if \;\; 0 < z \leq d_{ML} \\ T_{ML} - (T_{ML} - T_{BL})\Phi_{T_l}(\zeta) \quad if \;\; d_{ML} < z < d_l \end{cases} \tag{A.1}$$

as a function of depth z. $T_{BL}(^{\circ}C)$ is the temperature at the bottom of water body ($z = d_l(m)$), and $\Phi_T = [T_{ML} - T_l(z,t)]/[T_{ML} - T_{BL}]$ is a dimensionless function of dimensionless depth $\zeta = [z/d_{ML}(t)]/[d_l/d_{ML}(t)]$ which satisfies the boundary conditions of $\Phi_T(0) = 0$ and $\Phi_T(1) = 1$.
Then $T_{MN}(^{\circ}C)$ the mean layer temperature, is obtained integrating A.1:

$$T_{MN} = T_{ML} - \xi_{ML}\left(1 - \frac{d_{ML}}{dl}\right)(T_{ML} - TBL) \tag{A.2}$$

where $\xi_{ML}(m)$ is the shape factor of the thermocline temperature profile, linked to the dimensionless function of dimensionless depth ζ (Mironov, 2008). This equation links together the temperature of the two layers and the depth of the two layers.

By integrating the heat transfer equation over the vertical z, from 0 to d_l (depth of the lake), the equation for the mean temperature of the water column is obtained as:

$$d_l \frac{dT_{MN}}{dt} = \frac{1}{\rho_w c_w}\left[G_s + I_s - G_{BL} - I(d_l)\right] \tag{A.3}$$

where $c_w = 4.210^3 Jkg^{-1}K^{-1}$ is the specific heat of water, $G_s(Wm^{-2})$ is the sum of the sensible $SH(Wm^{-2})$, latent heat fluxes $LH(Wm^{-2})$ and the net heat flux due to long-wave radiation $I_{LW}(Wm^{-2})$ at the air-water interface: $G_s = LH + SH + I_{LW}$, I_s (Wm^{-2}) is the value of the heat flux due to solar radiation, and $G_{BL}(Wm^{-2})$ is the heat flux through the lake bottom. The radiation heat flux $I(d_l)$ that penetrates the water is the surface value of the incident solar radiation flux from the atmosphere multiplied by $1-\alpha$, where α is the albedo of the water surface with respect to solar radiation. The equation of heat budget of the mixed layer instead can be written as:

$$d_{ML} \frac{dT_{ML}}{dt} = \frac{1}{\rho_w c_w}\left[G_s + I_s - G_{ML} - I(d_{ML})\right] \tag{A.4}$$

where $G_{ML}(Wm^{-2})$ is the heat flux at the bottom of the mixed layer. This equation assumes two variants for stationary and retreating mixed layer (where $dd_{ML}/dt \leq 0$) and for deepening mixed layers (where $dd_{ML}/dt > 0$). A detailed description of the equations is behind the scope of this article, more information is reported in (ECMWF, 2018) and (Mironov, 2008).

B. Verification scores

Wilks (2011) provide extensive details on many verification scores, as well as the ECMWF web page (https://confluence.ecmwf.int/x/4fbUBw) on statistical concepts and probabilistic data. The verification procedure involves a comparison between matched pairs of forecasts and the observations to which they pertain. In the following, the verification scores used in the manuscript are described.

B.1.Relative Operating Curve

The Receiver Operating Characteristic (ROC) is a measurement of the forecast's ability to distinguish between two alternative outcomes and thus, measures resolution. The ROC was first introduced to meteorology by Mason in 1982 and is obtained by plotting the probability of detection (POD) against the probability of false detection (POFD) using a set of increasing probability thresholds. The ROC area, which is the area under the ROC curve, is often used as a score. A score of 0.5 indicates no skill and a score of 1 indicates perfect skill. It's important to note that ROC can also be computed for deterministic forecasts using different criteria such as median, certain percentiles, or chosen thresholds of interest instead of the probability thresholds used in probabilistic forecasts.

B.2.Frequency Bias score

The Bias is a measure of how the frequency of forecasted events (F) compares to the frequency of observed events (O):

$$FB = \frac{F}{O} \tag{B.1}$$

A forecast is considered unbiased if the value is 1. A value greater than 1 indicates an overpredicting forecast, and a value less than 1 indicates an underpredicting forecast.

B.3.Brier score

The Brier score (Brier et al., 1950), introduced by Brier in 1950, is a measure of the accuracy of probability forecasts in terms of predicting whether or not an event occurred. It is defined as the mean squared error of probabilistic forecasts, with events assigned a value of 1 and non-events assigned a value of zero. It has a mathematical structure similar to the Mean Square Error (MSE).

$$BS = \overline{(p-o)^2} \tag{B.2}$$

Thus, BS measures the difference between the forecast probability of an event (p), and its occurrence (o). The Brier score ranges from 0 to 1 with a perfect score of 0. Indeed it is a measure of the mean squared probability error of a forecast compared to a reference dataset such as observations, analyses, or climatology. The Brier Score for the forecast is calculated as the mean squared probability error of the forecast against observations or analyses over a given period. A lower Brier Score indicates a more accurate forecast, with a score of 0 indicating a wholly accurate forecast and a score of 1 indicating a wholly inaccurate forecast.

B.4.Continuous Ranked Probability Score

The Continuous Ranked Probability Score (CRPS) is a measure of how good forecasts are in matching observed outcomes and it generalizes the concept of the Mean Absolute Error (MAE) to the case of probabilistic forecasts. The CRPS is commonly employed to assess the accuracy of probabilistic forecasting models. Unlike metrics such as MAE, it takes into account the fact that forecasts are probabilistic, but observations are definite. The CRPS evaluates the overall forecast distribution, rather than focusing on a particular point. CRPS is calculated by comparing the Cumulative Distribution Functions (CDF) for the forecast against a reference dataset (observations, or analyses, or climatology) over a given period. It is defined as the integral of the difference between the cumulative distribution function (CDF) of the ensemble forecast and the CDF of the observed value, over all possible forecasted values. The mathematical formulation of the CRPS for an ensemble forecast is given by:

$$CRPS = \int_{-\infty}^{\infty} (F(x) - H(x-o))^2 dx \tag{B.3}$$

Where *F(x)* is the predicted cumulative distribution function of the ensemble forecast *H(x - o)* is the Heaviside step function, with o being the observed value

and x is the forecasted value. Alternatively, it can be written as:

$$CRPS = \frac{1}{N}\sum_{i=1}^{N}(F(x_i) - H(x_i - o))^2 \tag{B.4}$$

Where N is the number of ensemble members and xi is the ith ensemble member.

References

Attinger, R., Spreitzer, E., Boettcher, M., Forbes, R., Wernli, H., and Joos, H.: Quantifying the role of individual diabatic processes for the formation of PV anomalies in a North Pacific cyclone, Quarterly Journal of the Royal Meteorological Society, 145, 2454–2476, 2019.

Attinger, R., Spreitzer, E., Boettcher, M., Wernli, H., and Joos, H.: Systematic assessment of the diabatic processes that modify low-level potential vorticity in extratropical cyclones, Weather and Climate Dynamics, 2, 1073–1091, 2021.

Ayzel, G. and Izhitskiy, A.: Climate change impact assessment on freshwater inflow into the Small Aral Sea, Water, 11, 2377, 2019.

Baker, L., Rudd, A., Migliorini, S., and Bannister, R.: Representation of model error in a convective-scale ensemble prediction system, Nonlinear Processes in Geophysics, 21, 19–39, 2014.

Balsamo, G.: Interactive lakes in the integrated forecasting system, ECMWF Newsletter, 137, 30–34, 2013.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., and Betts, A. K.: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System, Journal of Hydrometeorology, 10, 623–643, 2009.

Balsamo, G., Salgado, R., Dutra, E., Boussetta, S., Stockdale, T., and Potes, M.: On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model, Tellus A: Dynamic Meteorology and Oceanography, 64, 15 829, 2012.

Barthlott, C. and Davolio, S.: Mechanisms initiating heavy precipitation over Italy during HyMeX Special Observation Period 1: a numerical case study using two mesoscale models, Quarterly Journal of the Royal Meteorological Society, 142, 238–258, 2016.

Bechtold, P.: Convection parameterization, in: Proc. Seminar on Parameterization of Subgrid Physical Processes, pp. 63–86, 2008.

Bechtold, P., Köhler, M., Jung, T., Doblas-Reyes, F., Leutbecher, M., Rodwell, M. J., Vitart, F., and Balsamo, G.: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 134, 1337–1351, 2008.

Bechtold, P., Semane, N., Lopez, P., Chaboureau, J.-P., Beljaars, A., and Bormann, N.: Representing equilibrium and nonequilibrium convection in large-scale models, Journal of the Atmospheric Sciences, 71, 734–753, 2014.

Beljaars, A. C., Brown, A. R., and Wood, N.: A new parametrization of turbulent orographic form drag, Quarterly Journal of the Royal Meteorological Society, 130, 1327–1347, 2004.

Beniston, M.: The 2003 heat wave in Europe: A shape of things to come? An analysis based on Swiss climatological data and model simulations, Geophysical Research Letters, 31, 2004.

Betts, A. K., Reid, D., and Crossett, C.: Evaluation of the FLake model in ERA5 for Lake Champlain, Frontiers in Environmental Science, 8, 250, 2020.

Black, E., Blackburn, M., Harrison, G., Hoskins, B., and Methven, J.: Factors contributing to the summer 2003 European heatwave, Weather, 59, 217–223, 2004.

Bongioannini Cerlini, P., Meniconi, S., and Brunone, B.: Groundwater supply and climate change management by means of global atmospheric datasets. Preliminary results, Procedia Engineering, 186, 420–427, 2017.

Bongioannini Cerlini, P., Silvestri, L., and Saraceni, M.: Quality control and gap-filling methods applied to hourly temperature observations over central Italy, Meteorological Applications, 27, e1913, 2020.

Bongioannini Cerlini, P., Silvestri, L., Meniconi, S., and Brunone, B.: Simulation of the water table elevation in shallow unconfined aquifers by means of the ERA5 soil moisture dataset: The Umbria Region case study, Earth Interactions, 25, 15–32, 2021.

Boutle, I. and Abel, S.: Microphysical controls on the stratocumulus topped boundary-layer structure during VOCALS-REx, Atmospheric Chemistry and Physics, 12, 2849–2863, 2012.

Braca, G., Bussettini, M., Lastoria, B., Mariani, S., and Piva, F.: Il Bilancio Idrologico Gis BAsed a scala Nazionale su Griglia regolare—BIGBANG: Metodologia e stime, Rapporto sulla disponibilità naturale della risorsa idrica. Rapp. ISPRA, 339, 1–181, 2021.

Bretherton, C. S., Blossey, P. N., and Khairoutdinov, M.: An energy-balance analysis of deep convective self-aggregation above uniform SST, Journal of the atmospheric sciences, 62, 4273–4292, 2005.

Brier, G. W. et al.: Verification of forecasts expressed in terms of probability, Monthly weather review, 78, 1–3, 1950.

Buizza, R.: The value of probabilistic prediction, Atmospheric Science Letters, 9, 36–42, 2008.

Buizza, R. and Hollingsworth, A.: Storm prediction over Europe using the ECMWF ensemble prediction system, Meteorological Applications, 9, 289–305, 2002.

Buizza, R., Milleer, M., and Palmer, T. N.: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system, Quarterly Journal of the Royal Meteorological Society, 125, 2887–2908, 1999.

Buizza, R., Leutbecher, M., and Isaksen, L.: Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 134, 2051–2066, 2008.

C3S: ERA5-Land hourly data from 1950 to present, Copernicus Climate Change Service, URL `https://cds.climate.copernicus.eu/#!/home`, 2020.

Cao, Y., Fu, C., Wang, X., Dong, L., Yao, S., Xue, B., Wu, H., and Wu, H.: Decoding the dramatic hundred-year water level variations of a typical great lake in semi-arid region of northeastern Asia, Science of The Total Environment, 770, 145 353, 2021.

Carrea, L. and Merchant, C.: GloboLakes: lake surface water temperature (LSWT) v4. 0 (1995–2016), Centre for Environmental Data Analysis, 29, 2019.

Casadei, S., Pierleoni, A., and Bellezza, M.: Integrated water resources management in a lake system: A case study in central Italy, Water, 8, 570, 2016.

Cavicchia, L., von Storch, H., and Gualdi, S.: A long-term climatology of medicanes, Climate dynamics, 43, 1183–1195, 2014.

Chaboureau, J. P., Pantillon, F., Lambert, D., Richard, E., and Claud, C.: Tropical transition of a Mediterranean storm by jet crossing, Quarterly Journal of the Royal Meteorological Society, 138, 596–611, https://doi.org/10.1002/qj.960, 2012.

Chebud, Y. A. and Melesse, A. M.: Modelling lake stage and water balance of lake Tana, Ethiopia, Hydrological Processes: An International Journal, 23, 3534–3544, 2009.

Chou, M.-D.: Parameterizations for the absorption of solar radiation by O 2 and CO 2 with application to climate studies, Journal of Climate, 3, 209–217, 1990.

Chou, M.-D. and Suarez, M. J.: An efficient thermal infrared radiation parameterization for use in general circulation models, NASA Technical Memorandum, 3, Article ID 104 606., 1994.

Chou, M.-D. and Suarez, M. J.: A solar radiation parameterization for atmospheric studies, Tech. rep., 1999.

Christensen, H. M., Moroz, I., and Palmer, T.: Stochastic and perturbed parameter representations of model uncertainty in convection parameterization, Journal of the Atmospheric Sciences, 72, 2525–2544, 2015.

Cioni, G., Malguzzi, P., and Buzzi, A.: Thermal structure and dynamical precursor of a Mediterranean tropical-like cyclone, Quarterly Journal of the Royal Meteorological Society, 142, 1757–1766, 2016.

Collins, W. D., Rasch, P. J., Boville, B. A., Hack, J. J., McCaa, J. R., Williamson, D. L., Briegleb, B. P., Bitz, C. M., Lin, S.-J., and Zhang, M.: The formulation and atmospheric simulation of the Community Atmosphere Model version 3 (CAM3), Journal of Climate, 19, 2144–2161, 2006.

Colvin, J., Lazarus, S., Splitt, M., Weaver, R., and Taeb, P.: Wind driven setup in east central Florida's Indian River Lagoon: Forcings and parameterizations, Estuarine, Coastal and Shelf Science, 213, 40–48, 2018.

Comellas Prat, A., Federico, S., Torcasio, R. C., D'Adderio, L. P., Dietrich, S., and Panegrossi, G.: Evaluation of the Sensitivity of Medicane Ianos to Model Microphysics and Initial Conditions Using Satellite Measurements, Remote Sensing, 13, 4984, 2021.

Coppin, D. and Bony, S.: Physical mechanisms controlling the initiation of convective self-aggregation in a general circulation model, Journal of Advances in Modeling Earth Systems, 7, 2060–2078, 2015.

Craig, G. C. and Mack, J. M.: A coarsening model for self-organization of tropical convection, Journal of Geophysical Research: Atmospheres, 118, 8761–8769, 2013.

Dafis, S., Claud, C., Kotroni, V., Lagouvardos, K., and Rysman, J.-F.: Insights into the convective evolution of Mediterranean tropical-like cyclones, Quarterly Journal of the Royal Meteorological Society, 146, 4147–4169, 2020.

Dai, A.: Precipitation characteristics in eighteen coupled climate models, Journal of climate, 19, 4605–4630, 2006.

Davis, C.: Resolving tropical cyclone intensity in models, Geophysical Research Letters, 45, 2082–2087, 2018.

Davolio, S., Miglietta, M., Moscatello, A., Pacifico, F., Buzzi, A., and Rotunno, R.: Numerical forecast and analysis of a tropical-like cyclone in the Ionian Sea, Natural Hazards and Earth System Sciences, 9, 551–562, 2009.

Dessie, M., Verhoest, N. E., Pauwels, V. R., Adgo, E., Deckers, J., Poesen, J., and Nyssen, J.: Water balance of a lake with floodplain buffering: lake Tana, Blue Nile basin, Ethiopia, Journal of Hydrology, 522, 174–186, 2015.

Di Muzio, E., Riemer, M., Fink, A. H., and Maier-Gerber, M.: Assessing the predictability of Medicanes in ECMWF ensemble forecasts using an object-based approach, Quarterly Journal of the Royal Meteorological Society, 145, 1202–1217, 2019.

Dokulil, M. T., de Eyto, E., Maberly, S. C., May, L., Weyhenmeyer, G. A., and Woolway, R. I.: Increasing maximum lake surface temperature under climate change, Climatic Change, 165, 1–17, 2021.

Downing, J. A.: Emerging global role of small lakes and ponds: little things mean a lot, Limnetica, 29, 9–24, 2010.

Drager, A. J. and van den Heever, S. C.: Characterizing convective cold pools, Journal of Advances in Modeling Earth Systems, 9, 1091–1115, 2017.

Dragoni, W., Piscopo, V., Di Matteo, L., Gnucci, L., Leone, A., Lotti, F., Melillo, M., and Petitta, M.: Risultati del progetto di ricerca PRIN laghi 2003–2005, Giornale di Geologia Applicata, 3, 39–46 *(In Italian)*, 2006.

Dragoni, W., Melillo, M., and Giontella, C.: Bilancio idrico del lago Trasimeno, Tutela ambientale del lago Trasimeno, 1 (In Italian), 2012.

Dutra, E., Stepanenko, V. M., Balsamo, G., Viterbo, P., Miranda, P., Mironov, D., and Schär, C.: An offline study of the impact of lakes on the performance of the ECMWF surface scheme, Boreal Environment Research Publishing Board, 15, 100–112, 2010.

ECMWF: IFS Documentation CY45R1 - Part IV: Physical processes, 4, URL https://www.ecmwf.int/node/18714, 2018.

Emanuel, K.: Genesis and maintenance of" Mediterranean hurricanes", Advances in Geosciences, 2, 217–220, 2005.

Emanuel, K., Wing, A. A., and Vincent, E. M.: Radiative-convective instability, Journal of Advances in Modeling Earth Systems, 6, 75–90, 2014.

Emanuel, K. A.: Quasi-Equilibrium Dynamics of the Tropical Atmosphere, The Global Circulation of the Atmosphere, 2021.

Emanuel, K. A. et al.: Atmospheric convection, Oxford University Press on Demand, 1994.

Ernst, J. and Matson, M.: A Mediterranean tropical storm?, Weather, 38, 332–337, 1983.

Ferretti, R., Pichelli, E., Gentile, S., Maiello, I., Cimini, D., Davolio, S., Miglietta, M., Panegrossi, G., Baldini, L., Pasi, F., et al.: Overview of the first HyMeX Special Observation Period over Italy: observations and model results, Hydrology and Earth System Sciences, 18, 1953–1977, 2014.

Flaounas, E., Kotroni, V., Lagouvardos, K., Gray, S. L., Rysman, J.-F., and Claud, C.: Heavy rainfall in Mediterranean cyclones. Part I: contribution of deep convection and warm conveyor belt, Climate dynamics, 50, 2935–2949, 2018.

Flaounas, E., Davolio, S., Raveh-Rubin, S., Pantillon, F., Miglietta, M. M., Gaertner, M. A., Hatzaki, M., Homar, V., Khodayar, S., Korres, G., et al.: Mediterranean cyclones: Current knowledge and open questions on dynamics, prediction, climatology and impacts, Weather and Climate Dynamics, 3, 173–208, 2022.

Forbes, R., Tompkins, A., and Untch, A.: A new prognostic bulk microphysics scheme for the IFS, ECMWF Technical Memoranda No. 649, 2011.

Frogner, I.-L., Andrae, U., Ollinaho, P., Hally, A., Hämäläinen, K., Kauhanen, J., Ivarsson, K.-I., and Yazgi, D.: Model uncertainty representation in a convection-permitting ensemble—SPP and SPPT in HarmonEPS, Monthly Weather Review, 150, 775–795, 2022.

George, G.: The impact of climate change on European lakes, Springer, 2010.

Gibson, J.: Short-term evaporation and water budget comparisons in shallow Arctic lakes using non-steady isotope mass balance, Journal of Hydrology, 264, 242–261, 2002.

Giorgi, F.: Climate change hot-spots, Geophysical research letters, 33, 2006.

Goosse, H. and Fichefet, T.: Importance of ice-ocean interactions for the global ocean circulation: A model study, Journal of Geophysical Research: Oceans, 104, 23 337–23 355, 1999.

Grabowski, W. W. and Moncrieff, M.: Moisture–convection feedback in the tropics, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 130, 3081–3104, 2004.

Grillakis, M. G.: Increase in severe and extreme soil moisture droughts for Europe under climate change, Science of The Total Environment, 660, 1245–1255, 2019.

Gronewold, A. D., Smith, J. P., Read, L. K., and Crooks, J. L.: Reconciling the water balance of large lake systems, Advances in Water Resources, 137, 103 505, 2020.

Guichard, F. and Couvreux, F.: A short review of numerical cloud-resolving models, Tellus A: Dynamic Meteorology and Oceanography, 69, 1373 578, 2017.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80–91, 2009.

Hamill, T. M., Whitaker, J. S., Fiorino, M., and Benjamin, S. G.: Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble Kalman filter, Monthly Weather Review, 139, 668–688, 2011.

Hart, R. E.: A cyclone phase space derived from thermal wind and thermal asymmetry, Monthly weather review, 131, 585–616, 2003.

Hart, R. E. and Evans, J. L.: A climatology of the extratropical transition of Atlantic tropical cyclones, Journal of Climate, 14, 546–564, 2001.

Hassler, B. and Lauer, A.: Comparison of Reanalysis and Observational Precipitation Datasets Including ERA5 and WFDE5, Atmosphere, 12, 1462, 2021.

Held, I. M. and Soden, B. J.: Robust responses of the hydrological cycle to global warming, Journal of climate, 19, 5686–5699, 2006.

Held, I. M., Hemler, R. S., and Ramaswamy, V.: Radiative-convective equilibrium with explicit two-dimensional moist convection, Journal of Atmospheric Sciences, 50, 3909–3927, 1993.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, 2020.

Hogan, R. J. and Bozzo, A.: A flexible and efficient radiation scheme for the ECMWF model, Journal of Advances in Modeling Earth Systems, 10, 1990–2008, 2018.

Hohenegger, C. and Stevens, B.: Coupled radiative convective equilibrium simulations with explicit and parameterized convection, Journal of Advances in Modeling Earth Systems, 8, 1468–1482, 2016.

Holloway, C. E. and Woolnough, S. J.: The sensitivity of convective aggregation to diabatic processes in idealized radiative-convective equilibrium simulations, Journal of Advances in Modeling Earth Systems, 8, 166–195, 2016.

Holloway, C. E., Wing, A. A., Bony, S., Muller, C., Masunaga, H., L'Ecuyer, T. S., Turner, D. D., and Zuidema, P.: Observing convective aggregation, Surveys in Geophysics, 38, 1199–1236, 2017.

Holton, J. R.: An introduction to dynamic meteorology, American Journal of Physics, 41, 752–754, 1973.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K.-L., Joyce, R. J., Kidd, C., Nelkin, E. J., Sorooshian, S., Stocker, E. F., Tan, J., et al.: Integrated multi-satellite retrievals for the global precipitation measurement (GPM) mission (IMERG), in: Satellite precipitation measurement, pp. 343–353, Springer, 2020.

IFS Documentation CY47R3, E.: IFS Documentation CY47R3 Part IV Physical processes, 2021a.

IFS Documentation CY47R3, E.: IFS Documentation CY47R3 Part V Ensemble prediction system, 2021b.

IFS Documentation CY47R3, E.: IFS Documentation CY47R3 Part VII ECMWF Wave model, 2021c.

IPCC, A.: IPCC Fifth Assessment Report—Synthesis Report, 2014.

ISPRA: Gli indicatori del clima in Italia nel 2019., 2020.

Jakob, C., Singh, M., and Jungandreas, L.: Radiative convective equilibrium and organized convection: An observational perspective, Journal of Geophysical Research: Atmospheres, 124, 5418–5430, 2019.

Jeevanjee, N. and Romps, D. M.: Convective self-aggregation, cold pools, and domain size, Geophysical Research Letters, 40, 994–998, 2013.

Jiang, Q., Li, W., Fan, Z., He, X., Sun, W., Chen, S., Wen, J., Gao, J., and Wang, J.: Evaluation of the ERA5 reanalysis precipitation dataset over Chinese Mainland, Journal of Hydrology, 595, 125 660, 2021.

Jiao, D., Xu, N., Yang, F., and Xu, K.: Evaluation of spatial-temporal variation performance of ERA5 precipitation data in China, Scientific Reports, 11, 1–13, 2021.

Judt, F., Chen, S. S., and Berner, J.: Predictability of tropical cyclone intensity: Scale-dependent forecast error growth in high-resolution stochastic kinetic-energy backscatter ensembles, Quarterly Journal of the Royal Meteorological Society, 142, 43–57, 2016.

Kebede, S., Travi, Y., Alemayehu, T., and Marc, V.: Water balance of lake Tana and its sensitivity to fluctuations in rainfall, Blue Nile basin, Ethiopia, Journal of Hydrology, 316, 233–247, 2006.

Kessler, E.: On the distribution and continuity of water substance in atmospheric circulations, in: On the distribution and continuity of water substance in atmospheric circulations, pp. 1–84, Springer, 1969.

Khairoutdinov, M. F. and Emanuel, K.: Aggregated convection and the regulation of tropical climate, in: 29th Conf. on Hurricanes and Tropical Meteorology, pp. P2–69, Amer. Meteor. Soc., 2010.

Khairoutdinov, M. F. and Kogan, Y. L.: A large eddy simulation model with explicit microphysics: Validation against aircraft observations of a stratocumulus-topped boundary layer, Journal of the atmospheric sciences, 56, 2115–2131, 1999.

Khairoutdinov, M. F. and Randall, D. A.: Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities, Journal of Atmospheric Sciences, 60, 607–625, 2003.

Khairoutdinov, M. F., Krueger, S. K., Moeng, C.-H., Bogenschutz, P. A., and Randall, D. A.: Large-eddy simulation of maritime deep tropical convection, Journal of Advances in Modeling Earth Systems, 1, 2009.

Khairoutdinov, M. F., Blossey, P. N., and Bretherton, C. S.: Global System for Atmospheric Modeling: Model Description and Preliminary Results, Journal of Advances in Modeling Earth Systems, https://doi.org/https://doi.org/10.1029/2021MS002968, e2021MS002968 2021MS002968, 2022.

Kirillin, G.: Modeling the impact of global warming on water temperature and seasonal mixing regimes in small temperate lakes, Boreal Environment Research Publishing Board, 2010.

Kirillin, G. B., Shatwell, T., and Wen, L.: Ice-Covered Lakes of Tibetan Plateau as Solar Heat Collectors, Geophysical Research Letters, 48, e2021GL093 429, 2021.

Klimaszyk, P., Rzymski, P., Piotrowicz, R., and Joniak, T.: Contribution of surface runoff from forested areas to the chemistry of a through-flow lake, Environmental Earth Sciences, 73, 3963–3973, 2015.

Köhler, M., Ahlgrimm, M., and Beljaars, A.: Unified treatment of dry convective and stratocumulus-topped boundary layers in the ECMWF model, Quarterly Journal of the Royal Meteorological Society, 137, 43–57, 2011.

Lagouvardos, K., Karagiannidis, A., Dafis, S., Kalimeris, A., and Kotroni, V.: Ianos—A hurricane in the Mediterranean, Bulletin of the American Meteorological Society, 103, E1621–E1636, 2022.

Landsea, C. W. and Cangialosi, J. P.: Have we reached the limits of predictability for tropical cyclone track forecasting?, Bulletin of the American Meteorological Society, 99, 2237–2243, 2018.

Lang, S., Leutbecher, M., and Jones, S.: Impact of perturbation methods in the ECMWF ensemble prediction system on tropical cyclone forecasts, Quarterly Journal of the Royal Meteorological Society, 138, 2030–2046, 2012.

Lang, S. T., Lock, S.-J., Leutbecher, M., Bechtold, P., and Forbes, R. M.: Revision of the stochastically perturbed parametrisations model uncertainty scheme in the integrated forecasting system, Quarterly Journal of the Royal Meteorological Society, 147, 1364–1381, 2021.

Lee, K.-O., Flamant, C., Ducrocq, V., Duffourg, F., Fourrié, N., and Davolio, S.: Convective initiation and maintenance processes of two back-building mesoscale convective systems leading to heavy precipitation events in Southern Italy during HyMeX IOP 13, Quarterly Journal of the Royal Meteorological Society, 142, 2623–2635, 2016.

Lei, Y., Yao, T., Yang, K., Sheng, Y., Kleinherenbrink, M., Yi, S., Bird, B. W., Zhang, X., Zhu, L., and Zhang, G.: Lake seasonality across the Tibetan Plateau and their varying relationship with regional mass changes and local hydrology, Geophysical Research Letters, 44, 892–900, 2017.

Lei, Y., Zhu, Y., Wang, B., Yao, T., Yang, K., Zhang, X., Zhai, J., and Ma, N.: Extreme lake level changes on the Tibetan Plateau associated with the 2015/2016 El Niño, Geophysical Research Letters, 46, 5889–5898, 2019.

Li, X.-Y., Xu, H.-Y., Sun, Y.-L., Zhang, D.-S., and Yang, Z.-P.: Lake-level change and water balance analysis at Lake Qinghai, west China during recent decades, Water Resources Management, 21, 1505–1516, 2007.

Li, Y., Zhang, Q., Yao, J., Tan, Z., and Liu, X.: Assessment of water storage response to surface hydrological connectivity in a large floodplain system (Poyang Lake, China) using hydrodynamic and geostatistical analysis, Stochastic Environmental Research and Risk Assessment, 33, 2071–2088, 2019.

Lin, Y.-L., Farley, R. D., and Orville, H. D.: Bulk parameterization of the snow field in a cloud model, Journal of Applied Meteorology and climatology, 22, 1065–1092, 1983.

Lionello, P.: The climate of the Mediterranean region: From the past to the future, Elsevier, 2012.

Liou, K.-N.: Influence of cirrus clouds on weather and climate processes: A global perspective, Monthly Weather Review, 114, 1167–1199, 1986.

Llasat, M. C., Llasat-Botija, M., Prat, M., Porcu, F., Price, C., Mugnai, A., Lagouvardos, K., Kotroni, V., Katsanos, D., Michaelides, S., et al.: High-impact floods and flash floods in Mediterranean countries: the FLASH preliminary database, Advances in Geosciences, 23, 47–55, 2010.

Lu, J., Chen, X., Zhang, L., Sauvage, S., and Sánchez-Pérez, J.-M.: Water balance assessment of an ungauged area in Poyang lake watershed using a spatially distributed runoff coefficient model, Journal of Hydroinformatics, 20, 1009–1024, 2018.

Ludovisi, A. and Gaino, E.: Meteorological and water quality changes in Lake Trasimeno (Umbria, Italy) during the last fifty years, Journal of Limnology, 69, 174, 2010.

Lükő, G., Torma, P., and Weidinger, T.: Intra-Seasonal and Intra-Annual Variation of the Latent Heat Flux Transfer Coefficient for a Freshwater Lake, Atmosphere, 13, 352, 2022.

Maberly, S. C., O'Donnell, R. A., Woolway, R. I., Cutler, M. E., Gong, M., Jones, I. D., Merchant, C. J., Miller, C. A., Politi, E., Scott, E. M., et al.: Global lake thermal regions shift under climate change, Nature Communications, 11, 1–9, 2020.

Magnuson, J. J., Robertson, D. M., Benson, B. J., Wynne, R. H., Livingstone, D. M., Arai, T., Assel, R. A., Barry, R. G., Card, V., Kuusisto, E., Granin, N. G., Prowse, T. D., Stewart, K. M., and Vuglinski, V. S.: Historical Trends in Lake and River Ice Cover in the Northern Hemisphere, Science, 289, 1743–1746, 2000.

Magnusson, L., Thorpe, A., Buizza, R., Rabier, F., and Nicolau, J.: Predicting this year's European heat wave, ECMWF Newsletter, 145, 4–5, 2015.

Maher, P., Gerber, E. P., Medeiros, B., Merlis, T. M., Sherwood, S., Sheshadri, A., Sobel, A. H., Vallis, G. K., Voigt, A., and Zurita-Gotor, P.: Model hierarchies for understanding atmospheric circulation, Reviews of Geophysics, 57, 250–280, 2019.

Mahmood, R., Jia, S., Mahmood, T., and Mehmood, A.: Predicted and projected water resources changes in the chari catchment, the lake Chad basin, Africa, Journal of Hydrometeorology, 21, 73–91, 2020.

Mahto, S. S. and Mishra, V.: Does ERA-5 outperform other reanalysis products for hydrologic applications in India?, Journal of Geophysical Research: Atmospheres, 124, 9423–9441, 2019.

Majumdar, S. J. and Torn, R. D.: Probabilistic verification of global and mesoscale ensemble forecasts of tropical cyclogenesis, Weather and Forecasting, 29, 1181–1198, 2014.

Malardel, S., Wedi, N., Deconinck, W., Diamantakis, M., Kühnlein, C., Mozdzynski, G., Hamrud, M., and Smolarkiewicz, P.: A new grid for the IFS, ECMWF newsletter, 146, 321, 2016.

Manrique Suñén, A., Nordbo, A., Balsamo, G., Beljaars, A., and Mammarella, I.: Land surface model over forest and lake surfaces in a boreal site-evaluation of the tiling method, in: EGU General Assembly Conference Abstracts, p. 2392, 2012.

Marsigli, C., Montani, A., and Paccangnella, T.: A spatial verification method applied to the evaluation of high-resolution ensemble forecasts, Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling, 15, 125–143, 2008.

Martin, G., Johnson, D., and Spice, A.: The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds, Journal of Atmospheric Sciences, 51, 1823–1842, 1994.

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M., et al.: Climate change 2021: the physical science basis, Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change, p. 2, 2021.

Mazza, E., Ulbrich, U., and Klein, R.: The tropical transition of the October 1996 medicane in the western Mediterranean Sea: A warm seclusion event, Monthly Weather Review, 145, 2575–2595, 2017.

McTaggart-Cowan, R., Davies, E. L., Fairman, J. G., Galarneau, T. J., and Schultz, D. M.: Revisiting the 26.5° C sea surface temperature threshold for tropical cyclone development, Bulletin of the American Meteorological Society, 96, 1929–1943, 2015.

Meehl, G. A., Covey, C., McAvaney, B., Latif, M., and Stouffer, R. J.: Overview of the coupled model intercomparison project, Bulletin of the American Meteorological Society, 86, 89–93, 2005.

Miglietta, M., Laviola, S., Malvaldi, A., Conte, D., Levizzani, V., and Price, C.: Analysis of tropical-like cyclones over the Mediterranean Sea through a combined modeling and satellite approach, Geophysical Research Letters, 40, 2400–2405, 2013.

Miglietta, M. M. and Rotunno, R.: Development mechanisms for Mediterranean tropical-like cyclones (medicanes), Quarterly Journal of the Royal Meteorological Society, 145, 1444–1460, 2019.

Miglietta, M. M., Moscatello, A., Conte, D., Mannarini, G., Lacorata, G., and Rotunno, R.: Numerical analysis of a Mediterranean 'hurricane'over southeastern Italy: Sensitivity experiments to sea surface temperature, Atmospheric research, 101, 412–426, 2011.

Miglietta, M. M., Mastrangelo, D., and Conte, D.: Influence of physics parameterization schemes on the simulation of a tropical-like cyclone in the Mediterranean Sea, Atmospheric Research, 153, 360–375, 2015.

Mironov, D.: Parameterization of Lakes in Numerical Weather Prediction. Description of a Lake Model (No. 11), Tech. rep., COSMO Technical Report. German Weather Service, Offenbach am Main, Germany, 2008.

Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., Schneider, N., and Terzhevik, A.: Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO, Boreal Environment Research Publishing Board, 2010.

Mogensen, K., Keeley, S., and Towers, P.: Coupling of the NEMO and IFS models in a single executable, ECMWF Reading, United Kingdom, 2012.

Montani, A., Cesari, D., Marsigli, C., and Paccagnella, T.: Seven years of activity in the field of mesoscale ensemble forecasting by the COSMO-LEPS system: main achievements and open challenges, Tellus A: Dynamic Meteorology and Oceanography, 63, 605–624, 2011.

Moon, H., Gudmundsson, L., and Seneviratne, S. I.: Drought persistence errors in global climate models, Journal of Geophysical Research: Atmospheres, 123, 3483–3496, 2018.

Morcrette, J., Barker, H. W., Cole, J., Iacono, M. J., and Pincus, R.: Impact of a new radiation package, McRad, in the ECMWF Integrated Forecasting System, Monthly weather review, 136, 4773–4798, 2008.

Moscatello, A., Marcello Miglietta, M., and Rotunno, R.: Observational analysis of a Mediterranean'hurricane'over south-eastern Italy, Weather, 63, 306, 2008.

Muala, E., Mohamed, Y. A., Duan, Z., and Van der Zaag, P.: Estimation of reservoir discharges from Lake Nasser and Roseires Reservoir in the Nile Basin using satellite altimetry and imagery data, Remote Sensing, 6, 7522–7545, 2014.

Muller, C. and Bony, S.: What favors convective aggregation and why?, Geophysical Research Letters, 42, 5626–5634, 2015.

Muller, C. J. and Held, I. M.: Detailed investigation of the self-aggregation of convection in cloud-resolving simulations, Journal of the Atmospheric Sciences, 69, 2551–2565, 2012.

Muller, C. J. and Romps, D. M.: Acceleration of tropical cyclogenesis by self-aggregation feedbacks, Proceedings of the National Academy of Sciences, 115, 2930–2935, 2018.

Muller, C. J., O'Gorman, P. A., and Back, L. E.: Intensification of precipitation extremes with warming in a cloud-resolving model, Journal of Climate, 24, 2784–2800, 2011.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., et al.: ERA5-Land: A state-of-the-art global reanalysis dataset for land applications, Earth System Science Data Discussions, 1–50, 2021.

Munsell, E. B., Zhang, F., and Stern, D. P.: Predictability and dynamics of a nonintensifying tropical storm: Erika (2009), Journal of the atmospheric sciences, 70, 2505–2524, 2013.

Naegele, A. C.: Observations and simulations of the interactions between clouds, radiation, and precipitation, Ph.D. thesis, Colorado State University, 2016.

Nastos, P., Papadimou, K. K., and Matsangouras, I.: Mediterranean tropical-like cyclones: Impacts and composite daily means and anomalies of synoptic patterns, Atmospheric Research, 208, 156–166, 2018.

Nowak, B. and Lawniczak-Malińska, A. E.: The influence of hydrometeorological conditions on changes in littoral and riparian vegetation of a meromictic lake in the last half-century, Water, 11, 2651, 2019.

O'Gorman, P. A. and Schneider, T.: The physical basis for increases in precipitation extremes in simulations of 21st-century climate change, Proceedings of the National Academy of Sciences, 106, 14 773–14 777, 2009.

Ollinaho, P., Lock, S.-J., Leutbecher, M., Bechtold, P., Beljaars, A., Bozzo, A., Forbes, R. M., Haiden, T., Hogan, R. J., and Sandu, I.: Stochastically Perturbed Parametrizations (SPP)-representing model uncertainties on the process-level, in: EGU General Assembly Conference Abstracts, p. 17706, 2017a.

Ollinaho, P., Lock, S.-J., Leutbecher, M., Bechtold, P., Beljaars, A., Bozzo, A., Forbes, R. M., Haiden, T., Hogan, R. J., and Sandu, I.: Towards process-level representation of model uncertainties: stochastically perturbed parametrizations in the ECMWF ensemble, Quarterly Journal of the Royal Meteorological Society, 143, 408–422, 2017b.

O'Reilly, C. M., Sharma, S., Gray, D. K., Hampton, S. E., Read, J. S., Rowley, R. J., Schneider, P., Lenters, J. D., McIntyre, P. B., Kraemer, B. M., et al.: Rapid and highly variable warming of lake surface waters around the globe, Geophysical Research Letters, 42, 10–773, 2015.

Palmer, T. N., Buizza, R., Doblas-Reyes, F., Jung, T., Leutbecher, M., Shutts, G. J., Steinheimer, M., and Weisheimer, A.: Stochastic parametrization and model uncertainty, 2009.

Pantillon, F., Chaboureau, J.-P., Mascart, P., and Lac, C.: Predictability of a Mediterranean tropical-like storm downstream of the extratropical transition of Hurricane Helene (2006), Monthly Weather Review, 141, 1943–1962, 2013.

Pauluis, O. and Held, I. M.: Entropy budget of an atmosphere in radiative–convective equilibrium. Part I: Maximum work and frictional dissipation, Journal of the Atmospheric Sciences, 59, 125–139, 2002a.

Pauluis, O. and Held, I. M.: Entropy budget of an atmosphere in radiative–convective equilibrium. Part II: Latent heat transport and moist processes, Journal of Atmospheric Sciences, 59, 140–149, 2002b.

Pendergrass, A. G.: What precipitation is extreme?, Science, 360, 1072–1073, 2018.

Petch, J.: Sensitivity studies of developing convection in a cloud-resolving model, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 132, 345–358, 2006.

Peters, O. and Neelin, J. D.: Critical phenomena in atmospheric precipitation, Nature physics, 2, 393–396, 2006.

Picornell, M., Campins, J., and Jansà, A.: Detection and thermal description of medicanes from numerical simulation, Natural Hazards and Earth System Sciences, 14, 1059–1070, 2014.

Pope, K. N., Holloway, C. E., Jones, T. R., and Stein, T. H.: Cloud-Radiation Interactions and Their Contributions to Convective Self-Aggregation, Journal of Advances in Modeling Earth Systems, 13, e2021MS002 535, 2021.

Portmann, R., González-Alemán, J. J., Sprenger, M., and Wernli, H.: How an uncertain short-wave perturbation on the North Atlantic wave guide affects the forecast of an intense Mediterranean cyclone (Medicane Zorbas), Weather and Climate Dynamics, 1, 597–615, 2020.

Pytharoulis, I., Craig, G. C., and Ballard, S. P.: The hurricane-like Mediterranean cyclone of January 1995, Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling, 7, 261–279, 2000.

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F., and Simmons, A.: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics, Quarterly Journal of the Royal Meteorological Society, 126, 1143–1170, 2000.

Ragone, F., Mariotti, M., Parodi, A., Von Hardenberg, J., and Pasquero, C.: A climatological study of western mediterranean medicanes in numerical simulations with explicit and parameterized convection, Atmosphere, 9, 397, 2018.

Rai, M. R., Chidthaisong, A., Ekkawatpanit, C., and Varnakovida, P.: Assessing Climate Change Trends and Their Relationships with Alpine Vegetation and Surface Water Dynamics in the Everest Region, Nepal, Atmosphere, 12, 987, 2021.

Randall, D. A., Hu, Q., Xu, K.-M., and Krueger, S. K.: Radiative-convective disequilibrium, Atmospheric Research, 31, 315–327, 1994.

Randall, D. A., Xu, K.-M., Somerville, R. J., and Iacobellis, S.: Single-column models and cloud ensemble models as links between observations and climate models, Journal of Climate, 9, 1683–1697, 1996.

Rasmussen, E. and Zick, C.: A subsynoptic vortex over the Mediterranean with some resemblance to polar lows, Tellus A, 39, 408–425, 1987.

Raymond, D. J. and Zeng, X.: Instability and large-scale circulations in a two-column model of the tropical troposphere, Quarterly Journal of the Royal Meteorological Society, 126, 3117–3135, 2000.

Reynolds, C., Williams, K. D., and Zadra, A.: WGNE systematic error survey results summary, in: 100th American Meteorological Society Annual Meeting, AMS, 2020.

Ricchi, A., Miglietta, M. M., Bonaldo, D., Cioni, G., Rizza, U., and Carniel, S.: Multi-physics ensemble versus atmosphere–ocean coupled model simulations for a tropical-like cyclone in the Mediterranean Sea, Atmosphere, 10, 202, 2019.

Romero, R. and Emanuel, K.: Medicane risk in a changing climate, Journal of Geophysical Research: Atmospheres, 118, 5992–6001, 2013.

Ruppert Jr, J. H. and Hohenegger, C.: Diurnal circulation adjustment and organized deep convection, Journal of Climate, 31, 4899–4916, 2018.

Samuelsson, P., Kourzeneva, E., and Mironov, D.: The impact of lakes on the European climate as simulated by a regional climate model, Boreal Environment Research Publishing Board, 2010.

Sandu, I., Bechtold, P., Beljaars, A., Bozzo, A., Pithan, F., Shepherd, T. G., and Zadra, A.: Impacts of parameterized orographic drag on the N orthern H emisphere winter circulation, Journal of Advances in Modeling Earth Systems, 8, 196–211, 2016.

Schär, C., Vidale, P. L., Lüthi, D., Frei, C., Häberli, C., Liniger, M. A., and Appenzeller, C.: The role of increasing temperature variability in European summer heatwaves, Nature, 427, 332–336, 2004.

Schewe, J., Gosling, S. N., Reyer, C., Zhao, F., Ciais, P., Elliott, J., Francois, L., Huber, V., Lotze, H. K., Seneviratne, S. I., et al.: State-of-the-art global models underestimate impacts from climate extremes, Nature Communications, 10, 1–14, 2019.

Schlemmer, L. and Hohenegger, C.: The formation of wider and deeper clouds as a result of cold-pool dynamics, Journal of the Atmospheric Sciences, 71, 2842–2858, 2014.

Schlimme, I., Macke, A., and Reichardt, J.: The impact of ice crystal shapes, size distributions, and spatial structures of cirrus clouds on solar radiative fluxes, Journal of the Atmospheric Sciences, 62, 2274–2283, 2005.

Schneider, P. and Hook, S. J.: Space observations of inland water bodies show rapid surface warming since 1985, Geophysical Research Letters, 37, 2010.

Schulz, S., Darehshouri, S., Hassanzadeh, E., Tajrishy, M., and Schüth, C.: Climate change or irrigated agriculture–what drives the water level decline of Lake Urmia, Scientific Reports, 10, 1–10, 2020.

Semie, A. G. and Bony, S.: Relationship between precipitation extremes and convective organization inferred from satellite observations, Geophysical Research Letters, 47, e2019GL086 927, 2020.

Setegn, S. G., Rayner, D., Melesse, A. M., Dargahi, B., and Srinivasan, R.: Impact of climate change on the hydroclimatology of Lake Tana Basin, Ethiopia, Water Resources Research, 47, 2011.

Sharma, S., Blagrave, K., Magnuson, J. J., O'Reilly, C. M., Oliver, S., Batt, R. D., Magee, M. R., Straile, D., Weyhenmeyer, G. A., Winslow, L., et al.: Widespread loss of lake ice around the Northern Hemisphere in a warming world, Nature Climate Change, 9, 227–231, 2019.

Shevnina, E. and Kourzeneva, E.: Thermal regime and components of water balance of lakes in Antarctica at the Fildes peninsula and the Larsemann Hills, Tellus A: Dynamic Meteorology and Oceanography, 69, 1317 202, 2017.

Sillmann, J., Kharin, V., Zhang, X., Zwiers, F., and Bronaugh, D.: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate, Journal of Geophysical Research: Atmospheres, 118, 1716–1733, 2013.

Silvestri, L., Saraceni, M., and Bongioannini Cerlini, P.: Links between precipitation, circulation weather types and orography in central Italy, International Journal of Climatology, 1–19, 2022a.

Silvestri, L., Saraceni, M., and Bongioannini Cerlini, P.: Quality management system and design of an integrated mesoscale meteorological network in Central Italy, Meteorological Applications, 29, e2060, 2022b.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G.: A description of the advanced research WRF version 2, Tech. rep., National Center For Atmospheric Research Boulder Co Mesoscale and Microscale , 2005.

Smith, J. P. and Gronewold, A. D.: Development and analysis of a Bayesian water balance model for large lake systems, arXiv: Applications, 2017.

Smolarkiewicz, P. K. and Grabowski, W. W.: The multidimensional positive definite advection transport algorithm: Nonoscillatory option, Journal of Computational Physics, 86, 355–375, 1990.

Soja, G., Züger, J., Knoflacher, M., Kinner, P., and Soja, A.-M.: Climate impacts on water balance of a shallow steppe lake in Eastern Austria (Lake Neusiedl), Journal of Hydrology, 480, 115–124, 2013.

Stauffer, C. L. and Wing, A. A.: Properties, Changes, and Controls of Deep-Convecting Clouds in Radiative-Convective Equilibrium, Journal of Advances in Modeling Earth Systems, 14, e2021MS002 917, 2022.

Stephens, G. L., Van Den Heever, S., and Pakula, L.: Radiative–convective feedbacks in idealized states of radiative–convective equilibrium, Journal of the Atmospheric Sciences, 65, 3899–3916, 2008.

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., Düben, P., Judt, F., Khairoutdinov, M., Klocke, D., et al.: DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains, Progress in Earth and Planetary Science, 6, 1–17, 2019.

Sun, H., Su, F., Yao, T., He, Z., Tang, G., Huang, J., Zheng, B., Meng, F., Ou, T., and Chen, D.: General overestimation of ERA5 precipitation in flow simulations for High Mountain Asia basins, Environmental Research Communications, 3, 121 003, 2021a.

Sun, S., Zhou, B., Xue, M., and Zhu, K.: Scale-Similarity Subgrid-Scale Turbulence Closure for Supercell Simulations at Kilometer-Scale Resolutions: Comparison against a Large-Eddy Simulation, Journal of the Atmospheric Sciences, 78, 417–437, 2021b.

Syed, T., Famiglietti, J., Chen, J., Rodell, M., Seneviratne, S. I., Viterbo, P., and Wilson, C.: Total basin discharge for the Amazon and Mississippi River basins from GRACE and a land-atmosphere water balance, Geophysical Research Letters, 32, 2005.

Tang, G., Clark, M. P., Papalexiou, S. M., Ma, Z., and Hong, Y.: Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets, Remote sensing of environment, 240, 111 697, 2020.

Tarek, M., Brissette, F. P., and Arsenault, R.: Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America, Hydrology and Earth System Sciences, 24, 2527–2544, 2020.

Taylor, C. M., de Jeu, R. A., Guichard, F., Harris, P. P., and Dorigo, W. A.: Afternoon rain more likely over drier soils, Nature, 489, 423–426, 2012a.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American meteorological Society, 93, 485–498, 2012b.

Tegen, I., Hollrig, P., Chin, M., Fung, I., Jacob, D., and Penner, J.: Contribution of different aerosol species to the global aerosol extinction optical thickness:

Estimates from model results, Journal of Geophysical Research: Atmospheres, 102, 23 895–23 915, 1997.

Thiery, W., Martynov, A., Darchambeau, F., Descy, J.-P., Plisnier, P.-D., Sushama, L., and van Lipzig, N. P.: Understanding the performance of the FLake model over two African Great Lakes, Geoscientific Model Development, 7, 317–337, 2014.

Thornthwaite, C. and Mather, J.: Instructions and tables for computing potential evapotranspiration and the water balance, Publications in Climatology, Technical report no. 5, 1957.

Tiedtke, M.: A comprehensive mass flux scheme for cumulus parameterization in large-scale models, Monthly weather review, 117, 1779–1800, 1989.

Tiedtke, M.: Representation of clouds in large-scale models, Monthly Weather Review, 121, 3040–3061, 1993.

Tompkins, A. M.: Organization of tropical convection in low vertical wind shears: The role of cold pools, Journal of the atmospheric sciences, 58, 1650–1672, 2001a.

Tompkins, A. M.: Organization of tropical convection in low vertical wind shears: The role of water vapor, Journal of the atmospheric sciences, 58, 529–545, 2001b.

Tompkins, A. M. and Craig, G. C.: Radiative–convective equilibrium in a three-dimensional cloud-ensemble model, Quarterly Journal of the Royal Meteorological Society, 124, 2073–2097, 1998.

Tompkins, A. M. and Semie, A. G.: Organization of tropical convection in low vertical wind shears: Role of updraft entrainment, Journal of Advances in Modeling Earth Systems, 9, 1046–1068, 2017.

Tong, K., Su, F., and Li, C.: Modeling of water fluxes and budget in Nam Co Basin during 1979–2013, Journal of Hydrometeorology, 21, 829–844, 2020.

Torn, R. D. and Cook, D.: The role of vortex and environment errors in genesis forecasts of Hurricanes Danielle and Karl (2010), Monthly weather review, 141, 232–251, 2013.

Tramblay, Y. and Somot, S.: Future evolution of extreme precipitation in the Mediterranean, Climatic Change, 151, 289–302, 2018.

Trenberth, K. E.: Changes in precipitation with climate change, Climate Research, 47, 123–138, 2011.

Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B.: The changing character of precipitation, Bulletin of the American Meteorological Society, 84, 1205–1218, 2003.

Ubertini, L., Casadei, S., and Pierleoni, A.: Lakes Management in Umbria Region, Central Italy, in: Proceedings of Taal 2007: The 12th World Lake Conference, pp. 1356–1364, 2007.

Vanderkelen, I., Van Lipzig, N. P., and Thiery, W.: Modelling the water balance of Lake Victoria (East Africa)–part 1: observational analysis, Hydrology and Earth System Sciences, 22, 5509–5525, 2018.

Vandewiele, G., Xu, C.-Y., et al.: Methodology and comparative study of monthly water balance models in Belgium, China and Burma, Journal of Hydrology, 134, 315–347, 1992.

Vich, M., Romero, R., and Brooks, H.: Ensemble prediction of Mediterranean high-impact events using potential vorticity perturbations. Part I: Comparison against the multiphysics approach, Atmospheric research, 102, 227–241, 2011.

Vincent, W.: Effects of climate change on lakes, Encyclopedia of Inland Waters. Elsevier, 2009.

Wang, W., Lee, X., Xiao, W., Liu, S., Schultz, N., Wang, Y., Zhang, M., and Zhao, L.: Global lake evaporation accelerated by changes in surface energy allocation in a warmer climate, Nature Geoscience, 11, 410–414, 2018.

Weger, R., Lee, J., Zhu, T., and Welch, R.: Clustering, randomness and regularity in cloud fields: 1. Theoretical considerations, Journal of Geophysical Research: Atmospheres, 97, 20 519–20 536, 1992.

Wilhite, D. and Pulwarty, R. S.: Drought and water crises: integrating science, management, and policy, CRC Press, 2017.

Wilks, D. S.: Statistical methods in the atmospheric sciences, vol. 100, Academic press, 2011.

Williamson, C. E., Saros, J. E., Vincent, W. F., and Smol, J. P.: Lakes and reservoirs as sentinels, integrators, and regulators of climate change, Limnology and Oceanography, 54, 2273–2282, 2009.

Windmiller, J. M. and Craig, G. C.: Universality in the spatial evolution of self-aggregation of tropical convection, Journal of the Atmospheric Sciences, 76, 1677–1696, 2019.

Wing, A. A. and Cronin, T. W.: Self-aggregation of convection in long channel geometry, Quarterly Journal of the Royal Meteorological Society, 142, 1–15, 2016.

Wing, A. A. and Emanuel, K. A.: Physical mechanisms controlling self-aggregation of convection in idealized numerical modeling simulations, Journal of Advances in Modeling Earth Systems, 6, 59–74, 2014.

Wing, A. A., Emanuel, K., Holloway, C. E., and Muller, C.: Convective self-aggregation in numerical simulations: A review, Shallow clouds, water vapor, circulation, and climate sensitivity, pp. 1–25, 2017.

Wing, A. A., Reed, K. A., Satoh, M., Stevens, B., Bony, S., and Ohno, T.: Radiative–convective equilibrium model intercomparison project, Geoscientific Model Development, 11, 793–813, 2018.

Wing, A. A., Stauffer, C. L., Becker, T., Reed, K. A., Ahn, M.-S., Arnold, N. P., Bony, S., Branson, M., Bryan, G. H., Chaboureau, J.-P., De Roode, S. R., Gayatri, K., Hohenegger, C., Hu, I.-K., Jansson, F., Jones, T. R., Khairoutdinov, M., Kim, D., Martin, Z. K., Matsugishi, S., Medeiros, B., Miura, H., Moon, Y., Müller, S. K., Ohno, T., Popp, M., Prabhakaran, T., Randall, D., Rios-Berrios, R., Rochetin, N., Roehrig, R., Romps, D. M., Ruppert Jr., J. H., Satoh, M., Silvers, L. G., Singh, M. S., Stevens, B., Tomassini, L., van Heerwaarden, C. C., Wang, S., and Zhao, M.: Clouds and Convective Self-Aggregation in a Multimodel Ensemble of Radiative-Convective Equilibrium Simulations, Journal of Advances in Modeling Earth Systems, 12, e2020MS002 138, 2020.

Woolway, R. I. and Merchant, C. J.: Worldwide alteration of lake mixing regimes in response to climate change, Nature Geoscience, 12, 271–276, 2019.

Woolway, R. I., Kraemer, B. M., Lenters, J. D., Merchant, C. J., O'Reilly, C. M., and Sharma, S.: Global lake responses to climate change, Nature Reviews Earth & Environment, 1, 388–403, 2020.

Xu, J., Ma, Z., Yan, S., and Peng, J.: Do ERA5 and ERA5-Land precipitation estimates outperform satellite-based precipitation products? A comprehensive comparison between state-of-the-art Model-based and satellite-based precipitation products over mainland China, Journal of Hydrology, 605, 127 353, 2021.

Xue, M., Droegemeier, K. K., and Wong, V.: The Advanced Regional Prediction System (ARPS)–A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification, Meteorology and Atmospheric Physics, 75, 161–193, 2000.

Xue, M., Droegemeier, K. K., Wong, V., Shapiro, A., Brewster, K., Carr, F., Weber, D., Liu, Y., and Wang, D.: The Advanced Regional Prediction System (ARPS)–A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications, Meteorology and atmospheric physics, 76, 143–165, 2001.

Xue, M., Hu, M., and Schenkman, A. D.: Numerical prediction of the 8 May 2003 Oklahoma City tornadic supercell and embedded tornado using ARPS with the assimilation of WSR-88D data, Weather and Forecasting, 29, 39–62, 2014.

Yanase, T., Nishizawa, S., Miura, H., Takemi, T., and Tomita, H.: New critical length for the onset of self-aggregation of moist convection, Geophysical Research Letters, 47, e2020GL088 763, 2020.

Yang, B. and Tan, Z.-M.: The initiation of dry patches in cloud-resolving convective self-aggregation simulations: Boundary layer dry-subsidence feedback, Journal of the Atmospheric Sciences, 77, 4129–4141, 2020.

Yano, J.-I. and Plant, R.: Convective quasi-equilibrium, Reviews of Geophysics, 50, 2012.

Yano, J.-I., Liu, C., and Moncrieff, M. W.: Self-organized criticality and homeostasis in atmospheric convective organization, Journal of the atmospheric sciences, 69, 3449–3462, 2012.

Yihdego, Y., Webb, J. A., and Vaheddoost, B.: Highlighting the role of groundwater in lake–aquifer interaction to reduce vulnerability and enhance resilience to climate change, Hydrology, 4, 10, 2017.

Zhang, F., Sun, Y. Q., Magnusson, L., Buizza, R., Lin, S.-J., Chen, J.-H., and Emanuel, K.: What is the predictability limit of midlatitude weather?, Journal of the Atmospheric Sciences, 76, 1077–1091, 2019a.

Zhang, G., Yao, T., Shum, C., Yi, S., Yang, K., Xie, H., Feng, W., Bolch, T., Wang, L., Behrangi, A., et al.: Lake volume and groundwater storage variations in Tibetan Plateau's endorheic basin, Geophysical Research Letters, 44, 5550–5560, 2017.

Zhang, G., Murakami, H., Gudgel, R., and Yang, X.: Dynamical seasonal prediction of tropical cyclone activity: Robust assessment of prediction skill and predictability, Geophysical Research Letters, 46, 5506–5515, 2019b.

Zhang, W., Villarini, G., Vecchi, G. A., and Murakami, H.: Rainfall from tropical cyclones: high-resolution simulations and seasonal forecasts, Climate dynamics, 52, 5269–5289, 2019c.

Zhang, W., Villarini, G., Scoccimarro, E., and Napolitano, F.: Examining the precipitation associated with medicanes in the high-resolution ERA-5 reanalysis data, International Journal of Climatology, 41, E126–E132, 2021.

Zuidema, P., Torri, G., Muller, C., and Chandra, A.: A survey of precipitation-induced atmospheric cold pools over oceans and their interactions with the larger-scale environment, Surveys in Geophysics, 38, 1283–1305, 2017.