Article

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

Paolina Bongioannini Cerlini [1], **Miriam Saraceni [2,*]**, **Lorenzo Silvestri [1,2]**, **Silvia Meniconi [2]** and **Bruno Brunone [2]**

[1] Centro Interuniversitario di Ricerca sull'Inquinamento e sull'Ambiente Mauro Felli (CIRIAF)-Centro di Ricerca sul Clima e Cambiamenti Climatici (CRC), University of Perugia, 06123 Perugia, Italy; paolina.cerlini@unipg.it (P.B.C.); lorenzo.silvestri@unipg.it (L.S.)

[2] Department of Civil and Environmental Engineering (DICA), Centro di Ricerca sul Clima e Cambiamenti Climatici (CRC), University of Perugia, 06123 Perugia, Italy; silvia.meniconi@unipg.it (S.M.); bruno.brunone@unipg.it (B.B.)

* Correspondence: miriam.saraceni@studenti.unipg.it

Abstract: Climate change has a strong impact on inland water bodies such as lakes. This means that the increase in lake temperature recorded in recent decades-in Europe as well-can change the evaporation regime of the lakes. This, together with the variation of the water cycle, in particular precipitation, implies that the water mass balance of lakes may vary due to climate change. Water mass balance modeling is therefore of paramount importance to monitor lakes in the context of global warming. Although many studies have focused on such a modeling, there is no shared approach that can be used for any lake across the globe, irrespective of the size. This becomes even more problematic for shallow and small lakes, for which few studies exist. For this reason, in this paper the use of reanalysis data, in particular ERA5-Land provided by the European Centre for Medium-Range Weather Forecasts (ECMWF), is proposed for the mass balance modeling. In fact, ERA5-Land has a global coverage and it is the only data source comprising a specific model for lakes, the Fresh-water Lake model (FLake). The chosen case study is the Trasimeno lake, a small and shallow lake located in Central Italy. The use of the reanalysis was preceded by data validation by considering both ground-based and satellite observations. The results show that there is a good agreement between the observed monthly variation of the lake level, ΔH, and the corresponding values of the water storage, δ, computed by means of the ERA5-Land data (Pearson coefficient larger than 70%). Discrepancies between observations and the ERA5-Land data happen in periods characterized in Europe by an extreme climate anomaly. This promising result encourages the use of ERA5-Land for other lakes.

Keywords: ERA5-land; lake water mass balance; climate extremes; satellite data; FLake

1. Introduction

Climate change has a significant impact on lake water bodies worldwide [1]. At the same time, according to their size, lakes may behave as sentinels or regulators of climate change [2,3]. The impact of climate change has been recognized globally for many aspects concerning lakes. The temperature has been increasing [4–6] and ice phenology has been changing, with ice cover loss [7] due to earlier ice breakup, later ice freeze-up, and shorter ice duration, at least in the northern hemisphere [8]. Evaporation of lakes has been changing pattern, due to surface warming [9], shortening of the ice cover periods, and decreasing the ratio of sensible to latent heat flux [10]. Specifically in Europe, between 1966–2015, the average lake surface temperature has been increasing by about 0.58 °C/decade [11]. Such a feature can change the lake mixing regime-both globally [12] and regionally [13] 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 [10,14–16]. However, the sensitivity of the water level to climate change and the extent of future variations remains quite uncertain in quantitative terms [1].

To improve the understanding of lakes' response to climate change-in particular, the behaviour in time of their level-in situ measurements should be coupled to numerical modeling of the physical processes. In the literature [17–19], the shared approach for evaluating the water level change, ΔH, in a given time interval, Δt, is based on the water mass balance equation:

$$H(t) - H(t - \Delta t) = \Delta H = P(t) - E(t) + R(t) \tag{1}$$

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 neighbouring 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 [18,20], in this paper, the water mass balance in Equation (1) has been evaluated on a monthly time scale (Δt = one month), that is appropriate from the lake monitoring point of view.

Different approaches can be followed for evaluating quantities in Equation (1). Precipitation over the lake is usually measured at the ground base [17] even if recently satellite observations are used as gridded data instead of local scattered measurements [19,21–23]. Unlike precipitation, evaporation and runoff are not usually measured at the ground base, but parametric formulas are often utilised 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 [17,24–26] and Thornthwaite [27] 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 [19,28], as well as the remotely sensed observations [22,23].

For the runoff, the simplest method is to calibrate the runoff coefficient, that depends on precipitation [17,19,25,27], or to use satellite observations [29]. More complex parameterizations are those characterizing "ad hoc" models as the Soil Water Assessment Tool (SWAT) [20], and the Variable Infiltration Capacity (VIC) [30,31] 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 utilised in [18,32], 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 [18] and Victoria lake [19] 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 [19] or open, as in the case of the harmonic regression model [32]. 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-hereafter referred to as ERA5L [33] the land component of the third generation ERA5 reanalysis [34], provided by the European Centre for Medium-Range Weather Forecast

(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 Freshwater Lake model-hereafter referred to as FLake [35]. 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 the lakes' surface more predictable [36].

In this paper, ERA5L data have been used for evaluating the lake level evolution by means of Equation (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 paper aims at the performance evaluation of the reanalyses as a needed action prior to their extensive use.

This paper is structured as follows. A brief description of the Trasimeno lake characteristics and previous studies is given in Section 2. The three different sources of data used in this work (reanalysis, ground-based and satellite) are described in Section 3. The methodology adopted for data validation and lake level modeling is described in Section 4. Results about data validation, model calibration and verification are presented in Section 5. The obtained results are discussed in Section 6. Finally, the Conclusions are offered with regard also to the use of the proposed model for other lakes in the world.

2. Site Description and Previous Studies

The Trasimeno lake is located in the Umbria region, Central Italy (Figure 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 1), respectively; accordingly, it is characterized by a large surface to volume ratio [37]. Thus, Trasimeno lake is smaller than those addressed in the literature, with an extent of the order of 10^3 km^2 [17,23,31] or 10^4 km^2 [18]. 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 [38,39]. 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 [27] 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 [40].
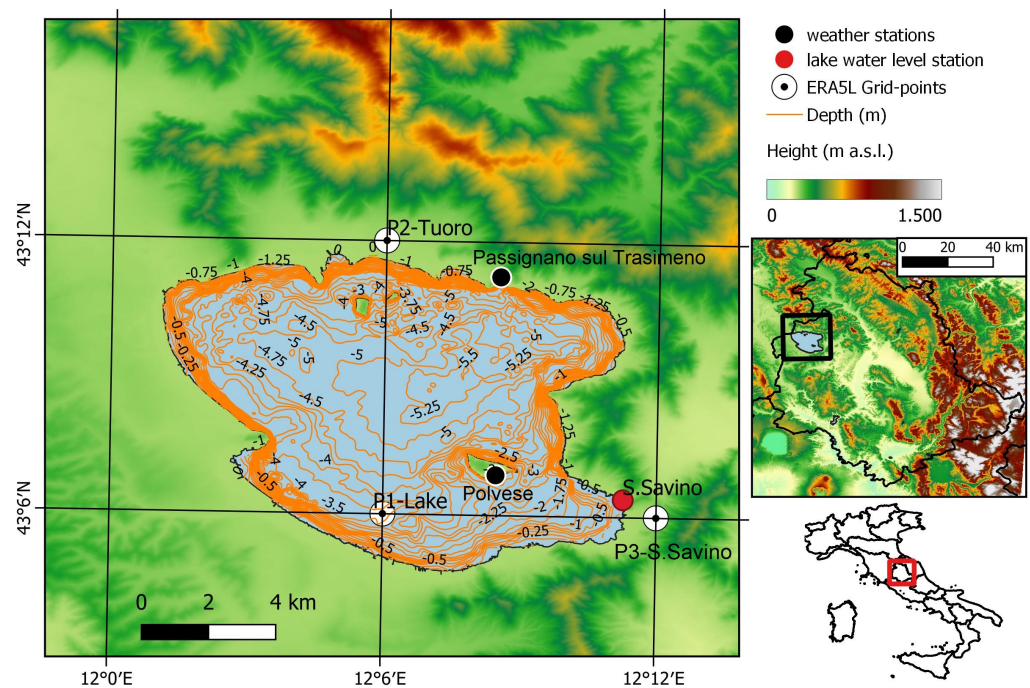
Figure 1. Map of the Trasimeno lake area and bathymetry with: ERA5L grid and points (P1-Lake, P2-Tuoro, and P3-S. Savino, indicated by white circles); the Umbria region weather stations measuring temperature and precipitation (Polvese, and Passignano sul Trasimeno, indicated by black circles), and the lake level measurement station (S. Savino, indicated by a red circle).

Throughout the years, some parametric models have been proposed for evaluating the Trasimeno lake water mass balance: the LAGO [27], the SimBaT [41], and the HLM [42]. 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 [43], 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 [44] (LAGO)-is shown to be negligible compared to evaporation and precipitation on a yearly average [38]. 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 backwards 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 [41].

The approach proposed in this paper, based on data from reanalysis (ERA5L), aims at providing an alternative to the use of the parametric models.

3. Data Sources: Reanalysis and Observations

In this paper, 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.

3.1. Era5l Reanalysis and the Flake Model

3.1.1. Ifs and Era5l

The Integrated Forecast System (IFS) is a General Circulation Model (GCM) provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). 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. Such a complex procedure, called 4D Var (4 Dimensional Variational assimilation), is run continuously at a horizontal resolution of 0.25° ($\simeq$31 km) to provide data consistency from 1950 to nowadays. In IFS, the water cycle is consistent with observations and physical fluxes, given the consistency and interaction between all the above-cited model components. In particular, the atmospheric model interacts with the land-surface model H-TESSEL through surface fluxes and all the meteoric species [45]. This complex system produces the ERA5 reanalysis [34].

The land-surface component of the ERA5 reanalysis is ERA5L [33]. 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 [35].

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.

3.1.2. Flake Model and Era5l Fluxes

FLake is a one-dimensional model which represents fresh-water lakes [46–48], 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 behaviour 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 short wave 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 [36,49,50].

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 in the underlying surface. For a more detailed explanation of FLake, the reader can refer to the IFS documentation [51] and the theory in [35,52].

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 [53] 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 vapour, E. Specifically, for each water tile, i, and grid point, j, it is:

$$E_{i,j} = \rho_a |U_{L,j}| C_{H,i,j} [q_{L,j} - q_{sat,j}(T_{sk,i,j})] \tag{2}$$

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 [51] for a detailed description). For each water tile, q_{sat} is calculated using the mean skin temperature of the tile, $T_{sk,i}$ (°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 the convective and the 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 by large-scale weather patterns (i.e., fronts). Large-scale and convective precipitation are calculated in the cloud and convective schemes, parts of IFS, parameterizing rain and snow water content, cloud liquid, cloud ice, and fractional cloud cover with prognostic equations. The convective and cloud schemes represent the sources and sinks of cloud 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 [54,55], conversion from cloud water into rain, detrainment from mass updraft, and evaporation from precipitation [56,57]. All these processes are integrated over the atmosphere column at each grid point, j, to get the precipitation over the lake, P, as:

$$P_j = \int_{p_0}^{p_{top}} (G_j - D_j) \frac{dp}{g} \tag{3}$$

where p = column pressure (hPa) with the subscripts *0* and *top* indicating the surface and top of the 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.

3.1.3. 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 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).

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 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 ± 1 cm (over the whole measurement range), respectively.

The time series of the daily mean level are shown in Figure 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.
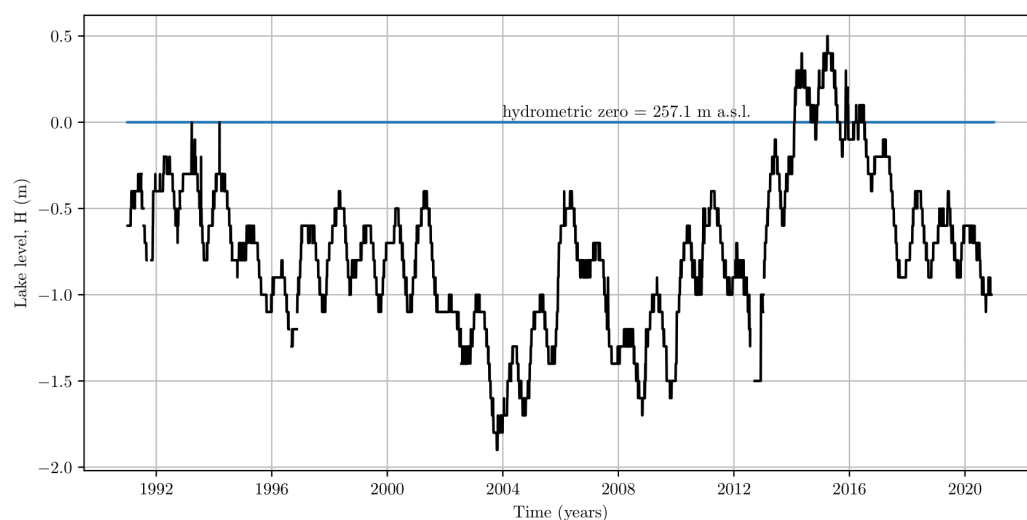

Figure 2. Values of the daily mean level, *H*, of the Trasimeno lake, measured at the S. Savino weather station, with in blue the hydrometric zero.

Before using lake level data, they have been checked for internal consistency on the basis of the precipitation measured at the same weather station [58,59]. 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 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. 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) [60], 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 the 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 3). The wind set-up, evaluated by considering such average wind speed and direction in the formula proposed in [61], 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.

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 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 4b for the period January 2002–January 2003, as an example.
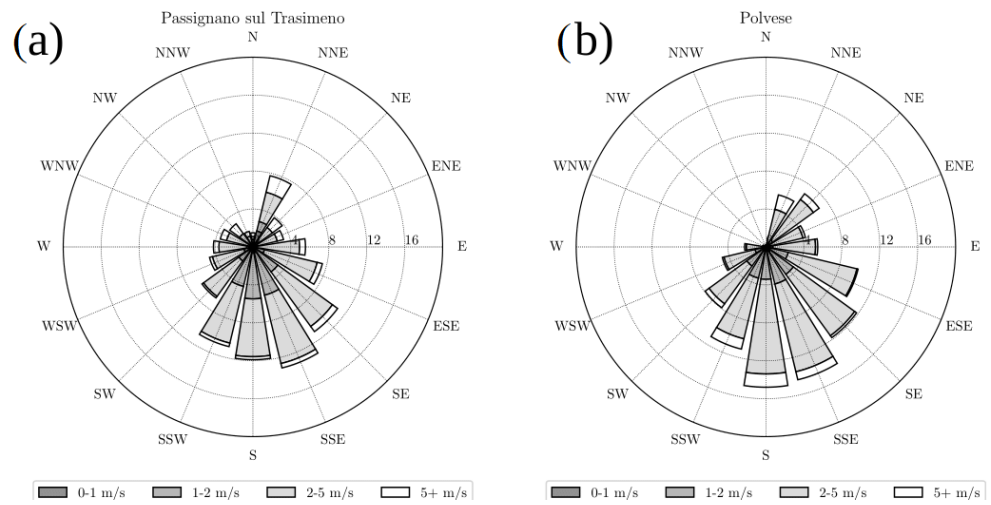

Figure 3. Ten meters wind rose at the two different weather stations in the lake area (black circles in Figure 1): (**a**) Passignano sul Trasimeno; (**b**) Polvese. Radial distance and shading indicate the percentage of days with wind direction in the sector and wind speed intervals, respectively.
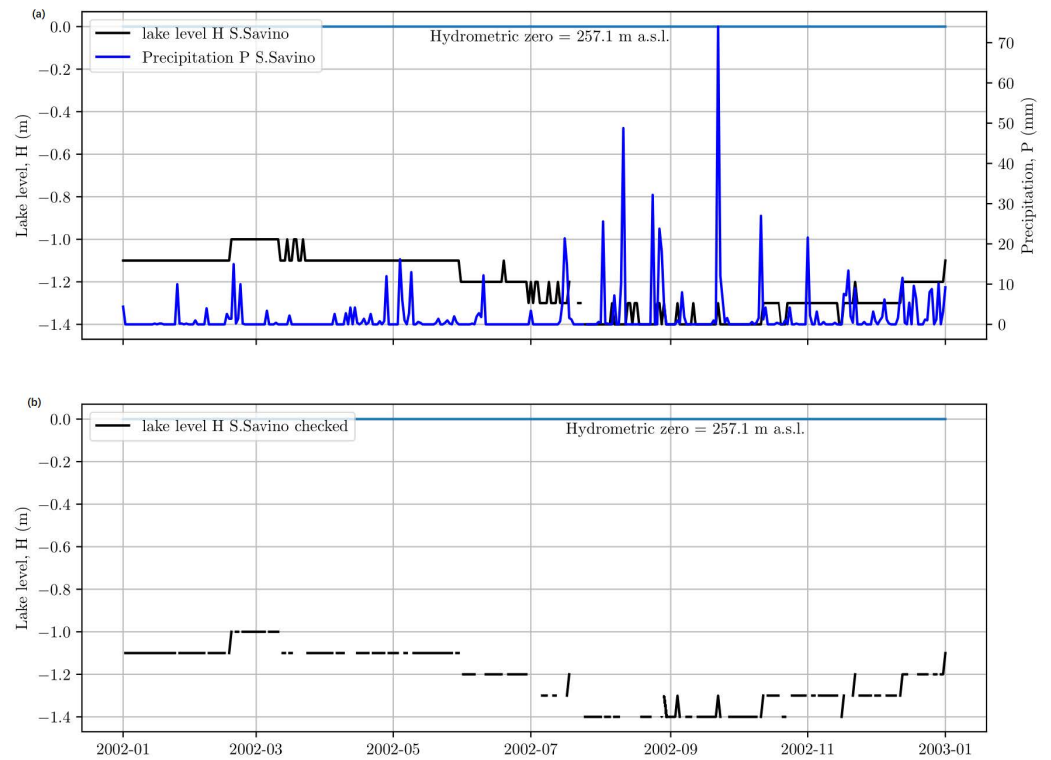
Figure 4. Cross-check between precipitation and level of the Trasimeno lake at S. Savino for the period January 2002–January 2003. (**a,b**) In Figure 3a, the unchecked lake level is reported together with precipitation, whereas in Figure 3b the validated level is shown.

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) [62]. This 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 multipurposes 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 paper, only LSWT values with high quality level (i.e., 4 or 5) have been considered.

4. Methods

4.1. Era5l Data Validation

Before using for modeling the water mass balance, ERA5L data have been validated against observations for what it 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 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:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i}^{N}(m_i - o_i)^2} \qquad (4)$$

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

4.2. Modeling the Lake Level: The Conceptual Model

As mentioned, for the Trasimeno lake it has been shown [38,39] 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 Equation (1), it is assumed $R(t) = 0$ and then it can be rewritten as:

$$\Delta H \simeq P(t) - E(t) = \delta(t) \qquad (5)$$

where ΔH = monthly variation of the lake level, and $\delta = P - E$ 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 [63], but it may be important in wet years, during extreme precipitation events, when it recharges the lake [63]. On the other hand, some shallow lakes are highly influenced by surface runoff from rivers, which significantly influences the lake's water balance [20,63], becoming more important under conditions of extreme precipitation [64]. 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 [17].

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 Equation (4), and the Kling–Gupta efficiency, KGE_S, defined as [65]:

$$KGE_S = 1 - \sqrt{(r_S - 1)^2 + (\frac{\sigma_m}{\sigma_o} - 1)^2 + (\frac{\sigma_m}{\sigma_o} - 1)^2} \tag{6}$$

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.

5. Results

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 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 temperature given in ERA5L.
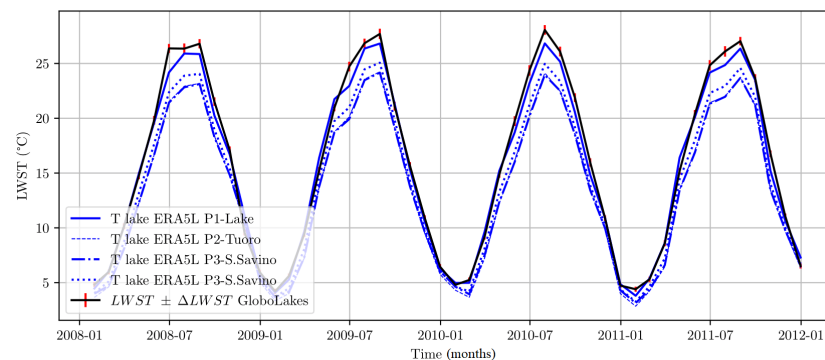

Figure 5. Comparison for the period 2008-2011 between the lake surface water temperature, LSWT (°C) from satellite (reported in black), and the mixed layer temperature, T_{ML}(°C), as modeled by ERA5L (in blue), for each ERA5L point and their weighted average.

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 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 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 raingauge observations is shown in Figure 6, where only the period from 2008 to 2011 is reported as an example.
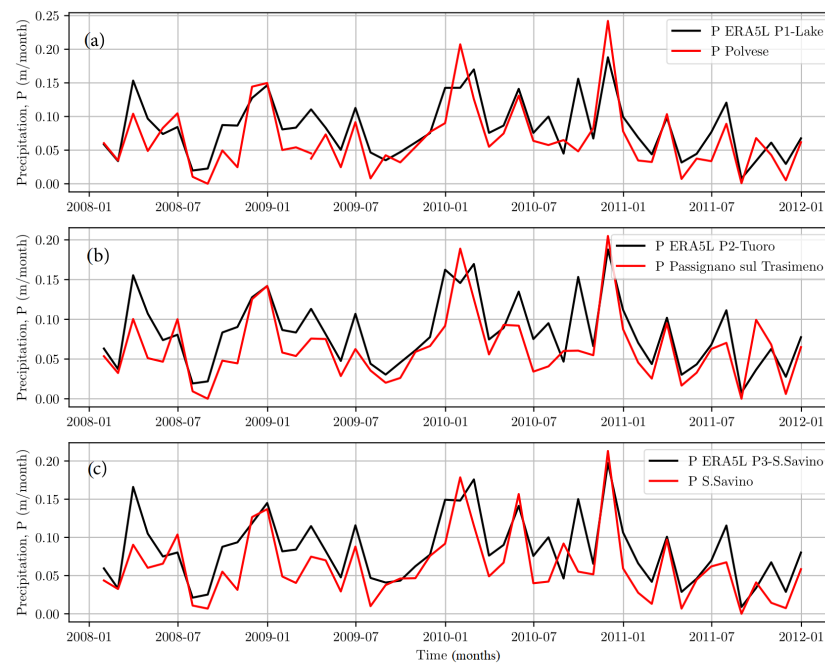

Figure 6. Comparison for the period 2008–2011 between the precipitation over the lake, *P* (m/month), observed at ground base by weather stations (reported in red), and the one, provided by ERA5L (in black), at: (**a**) P1-Lake vs. Polvese, (**b**) P2-Tuoro vs. Passignano sul Trasimeno, and (**c**) P3-S. 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), with the largest value reached at P2-Tuoro ($r_P^2 = 0.80$, where the subscript *P* indicates precipitation).

Table 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 [66,67]), this lays the basis for a reliable use of evaporation and precipitation fluxes from ERA5L.

5.2. Modeling the Lake Level: Correlation Analysis

The validity of the model proposed in Equation (5) has been checked by comparing the time series of the modeled δ with the observed ΔH, as described in Section 4.2. From the comparison shown in Figure 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 7 and Table 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 7a).
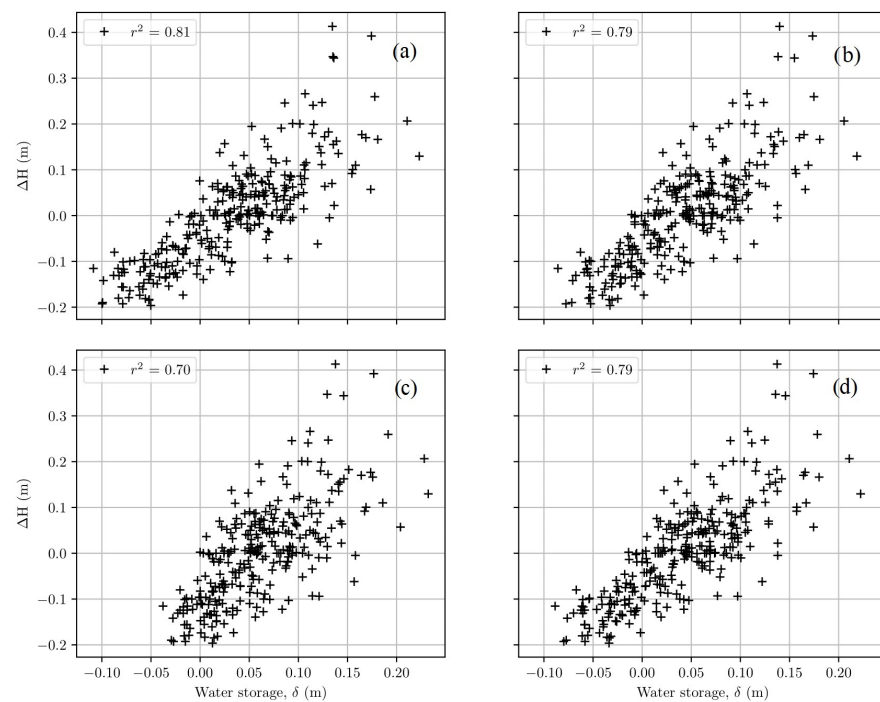

Figure 7. Correlation analysis between the monthly variations of the lake level, ΔH, and water storage, δ, for the three ERA5L points closest to the lake: (**a**) P1-Lake, (**b**) P2-Tuoro, (**c**) P3-S. Savino, and (**d**) their weighted average.

Table 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 1). The slightly better performance of the point P1-Lake can be ascribed to the larger values of evaporation from the lake accounted by ERA5L (not shown). In fact, in the case of P1-Lake, the evaporation from the lake accounts for the 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 on the correlation between the P1-Lake time-serie, δ, and the observed, ΔH, especially in the summer season when the evaporation is prominent 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 Equation (5) has been deemed valid for simulating the variations of the lake level, ΔH.

5.3. Model Calibration and Verification

According to Figure 7, a linear relationship between δ, given by ERA5L P1-Lake, and the observed ΔH has been assumed:

$$\Delta H(t) = \alpha \delta(t) + \beta \tag{7}$$

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 8, the values of the monthly lake level, given by Equation (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.
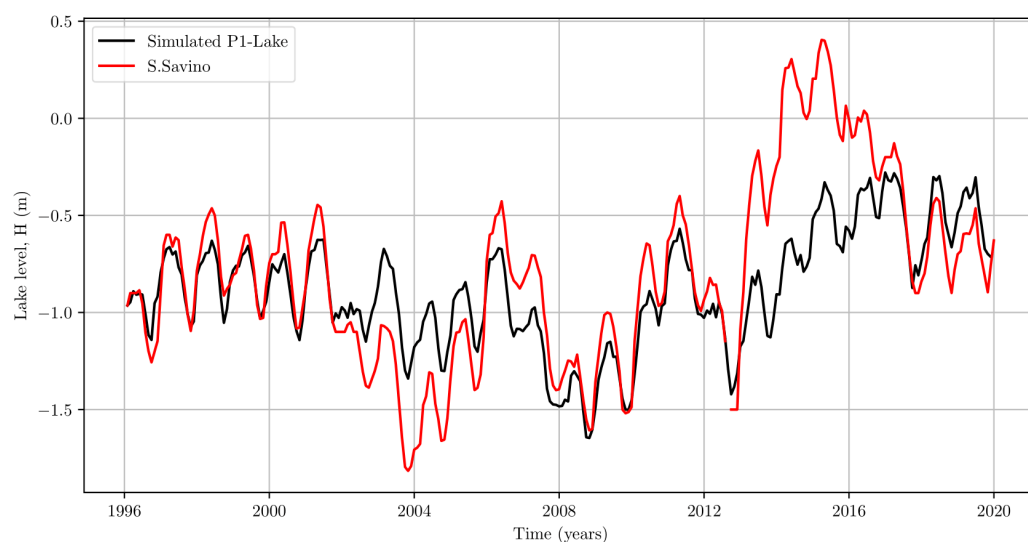

Figure 8. Monthly lake level, *H*, simulated (in black) by Equation (7) vs. observed (in red) values at S. Savino. In the model, the validation period is from 1996 to 2000, whereas the simulation one is from 2000 to 2019.

The values reported in Table 4 indicate that, apart from the validation period, where the simulation overlaps the observation, the periods that are better simulated 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 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 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 are 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 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 [68,69]. 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).

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.

6. Discussion

Figure 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 the July temperatures anomaly reaching 6° [70,71]. In this period, the model returns values of δ that are much higher than the observations ($RMSE_S$ = 0.30 m in Table 4). The second period, late 2012-early 2013, with the poorest agreement ($RMSE_S$ = 0.57 m in Table 4), is characterized by intense precipitation over central Italy [72] and, then, also over the Trasimeno lake area.

To understand the reasons for these discrepancies, the monthly precipitation of ERA5L has been compared with the ground-based observations. This 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 1). The results of such an analysis are shown in Figures 9 and 10 for the period 2002–2005 and 2012–2013, respectively.
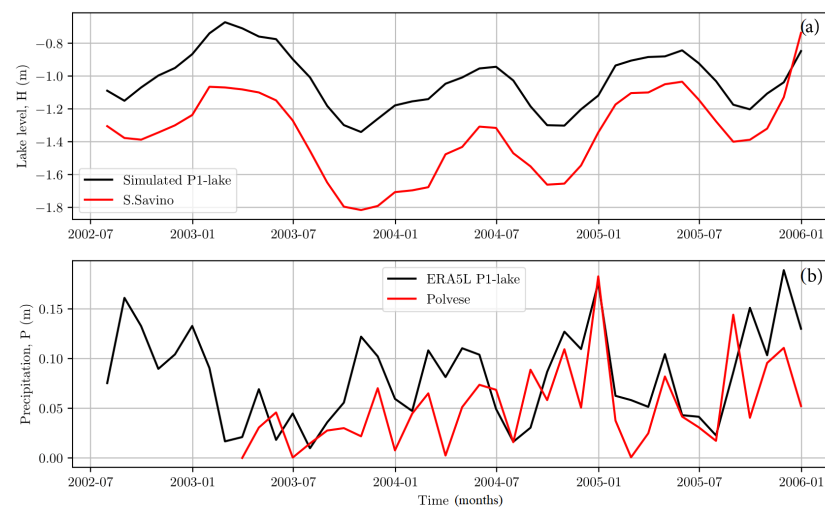
Figure 9. Comparison between ERA5L and ground-based observed data for the drought period 2002–2006: (**a**) monthly lake level, *H*, from ERA5L P1-Lake (in black) and the observed one at S. Savino station (in red); (**b**) precipitation over the lake (m/month), *P*, from ERA5L P1-Lake (in black) and the observed one at Polvese station (in red).

Figure 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).
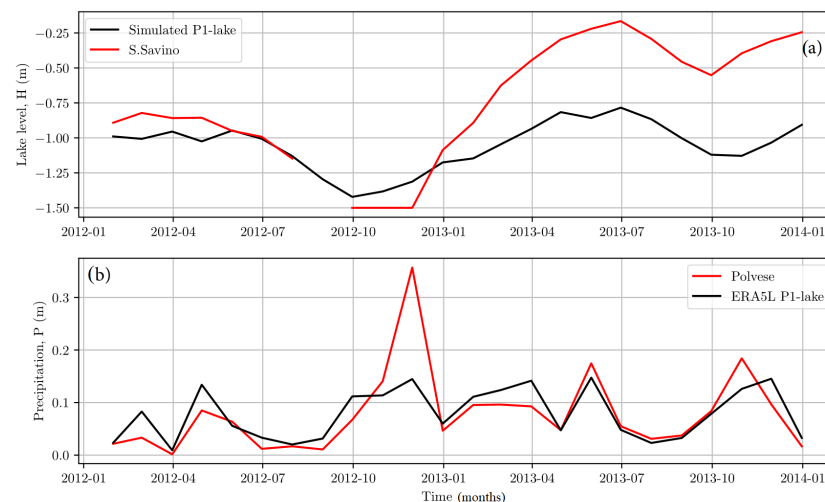
The drought period is not properly captured by reanalysis precipitation data (Figure 9). Indeed, since April 2003, the comparison between the precipitation given by ERA5L at P1-Lake and the one observed at Polvese, points out that the former is much larger (0.10 m/month vs. 0.05 m/month). This feature 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 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 a $RMSE_T$ 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 [63].

For the intense precipitation period, the focus has been put on late 2012 and early 2013. In fact, in December 2012 (Figure 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 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, in May 2019 [73], 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 [20]). 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 has been 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 [74]. Moreover, in the literature, precipitation biases of ERA5 have been noted in other parts of the world. In particular, in [75] 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 [76] 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 General Circulation Models (GCM) worldwide like IFS [77,78].

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) [79], which have shown a general underestimation of the persistence of drought periods [78]. 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 [80]. Indeed, the causes for the 2003 drought period lie in a positive feedback between soil moisture and precipitation [81], inaccurately represented by GCMs, that show contrasting behaviour with observations [82].

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 [77]. 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 [72]. 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 [83,84].

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 [85] and in other regions of the world [86].

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 on average an evaporation from the lake that is the 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.

7. Conclusions

In this paper, 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 the European Centre of Medium-Range Weather Forecast (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 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 paper is among the first (see also [87]) 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 [20] and Tana lake [28], with a catchment much larger than the basin. In these cases, groundwater flow cannot be considered negligible, nor runoff [20]. Moreover, some shallow lakes have tributaries, e.g., rivers, and therefore their influence must be taken into account [28,88]. Furthermore, many shallow and small lakes are in subpolar regions where the presence of ice cover is not negligible [89] and this could change the evaporation regime [10] (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 [17,69].

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 the literature [90]. This would allow 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 paper, 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 [91,92], reanalyses are an alternative to other methods for modeling the water mass balance of inland water bodies.

Author Contributions: All authors contributed to the conceptualization of the research. P.B.C., M.S. and B.B. worked on the methodology and the model implementation and validation. L.S. contributed to the reanalysis data validation. P.B.C. and S.M. supervised all the research group work. M.S., B.B. and P.B.C. wrote the original draft. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

References

1. Woolway, R.I.; Kraemer, B.M.; Lenters, J.D.; Merchant, C.J.; O'Reilly, C.M.; Sharma, S. Global lake responses to climate change. *Nat. Rev. Earth Environ.* **2020**, *1*, 388–403. [CrossRef]
2. Williamson, C.E.; Saros, J.E.; Vincent, W.F.; Smol, J.P. Lakes and reservoirs as sentinels, integrators, and regulators of climate change. *Limnol. Oceanogr.* **2009**, *54*, 2273–2282. [CrossRef]
3. Kirillin, G.B.; Shatwell, T.; Wen, L. Ice-Covered Lakes of Tibetan Plateau as Solar Heat Collectors. *Geophys. Res. Lett.* **2021**, *48*, e2021GL093429. [CrossRef]
4. 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. *Geophys. Res. Lett.* **2015**, *42*, 10–773. [CrossRef]
5. 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. *Nat. Commun.* **2020**, *11*, 1–9. [CrossRef]
6. Rai, M.R.; Chidthaisong, A.; Ekkawatpanit, C.; Varnakovida, P. Assessing Climate Change Trends and Their Relationships with Alpine Vegetation and Surface Water Dynamics in the Everest Region, Nepal. *Atmosphere* **2021**, *12*, 987. [CrossRef]
7. 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. *Nat. Clim. Chang.* **2019**, *9*, 227–231. [CrossRef]
8. 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.; et al. Historical Trends in Lake and River Ice Cover in the Northern Hemisphere. *Science* **2000**, *289*, 1743–1746. [CrossRef]
9. Schneider, P.; Hook, S.J. Space observations of inland water bodies show rapid surface warming since 1985. *Geophys. Res. Lett.* **2010**, *37* . [CrossRef]
10. Wang, W.; Lee, X.; Xiao, W.; Liu, S.; Schultz, N.; Wang, Y.; Zhang, M.; Zhao, L. Global lake evaporation accelerated by changes in surface energy allocation in a warmer climate. *Nat. Geosci.* **2018**, *11*, 410–414. [CrossRef]
11. Dokulil, M.T.; de Eyto, E.; Maberly, S.C.; May, L.; Weyhenmeyer, G.A.; Woolway, R.I. Increasing maximum lake surface temperature under climate change. *Clim. Chang.* **2021**, *165*, 1–17. [CrossRef]
12. Woolway, R.I.; Merchant, C.J. Worldwide alteration of lake mixing regimes in response to climate change. *Nat. Geosci.* **2019**, *12*, 271–276. [CrossRef]
13. Kirillin, G. Modeling the impact of global warming on water temperature and seasonal mixing regimes in small temperate lakes. *Boreal Environ. Res. Publ. Board* **2010**, *15*, 279–293.
14. Nowak, B.; 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* **2019**, *11*, 2651. [CrossRef]
15. Lei, Y.; Zhu, Y.; Wang, B.; Yao, T.; Yang, K.; Zhang, X.; Zhai, J.; Ma, N. Extreme lake level changes on the Tibetan Plateau associated with the 2015/2016 El Niño. *Geophys. Res. Lett.* **2019**, *46*, 5889–5898. [CrossRef]
16. Schulz, S.; Darehshouri, S.; Hassanzadeh, E.; Tajrishy, M.; Schüth, C. Climate change or irrigated agriculture–what drives the water level decline of Lake Urmia. *Sci. Rep.* **2020**, *10*, 1–10. [CrossRef]
17. Kebede, S.; Travi, Y.; Alemayehu, T.; Marc, V. Water balance of lake Tana and its sensitivity to fluctuations in rainfall, Blue Nile basin, Ethiopia. *J. Hydrol.* **2006**, *316*, 233–247. [CrossRef]
18. Gronewold, A.D.; Smith, J.P.; Read, L.K.; Crooks, J.L. Reconciling the water balance of large lake systems. *Adv. Water Resour.* **2020**, *137*, 103505. [CrossRef]
19. Vanderkelen, I.; Van Lipzig, N.P.; Thiery, W. Modelling the water balance of Lake Victoria (East Africa)–part 1: Observational analysis. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 5509–5525. [CrossRef]
20. Cao, Y.; Fu, C.; Wang, X.; Dong, L.; Yao, S.; Xue, B.; Wu, H.; Wu, H. Decoding the dramatic hundred-year water level variations of a typical great lake in semi-arid region of northeastern Asia. *Sci. Total. Environ.* **2021**, *770*, 145353. [CrossRef]
21. Lei, Y.; Yao, T.; Yang, K.; Sheng, Y.; Kleinherenbrink, M.; Yi, S.; Bird, B.W.; Zhang, X.; Zhu, L.; Zhang, G. Lake seasonality across the Tibetan Plateau and their varying relationship with regional mass changes and local hydrology. *Geophys. Res. Lett.* **2017**, *44*, 892–900. [CrossRef]
22. Smith, J.P.; Gronewold, A.D. Development and analysis of a Bayesian water balance model for large lake systems. *arXiv* **2017**, arXiv:1710.10161.
23. Lu, J.; Chen, X.; Zhang, L.; Sauvage, S.; Sánchez-Pérez, J.M. Water balance assessment of an ungauged area in Poyang lake watershed using a spatially distributed runoff coefficient model. *J. Hydroinform.* **2018**, *20*, 1009–1024. [CrossRef]

24. Chebud, Y.A.; Melesse, A.M. Modelling lake stage and water balance of lake Tana, Ethiopia. *Hydrol. Process. Int. J.* **2009**, *23*, 3534–3544. [CrossRef]

25. Dessie, M.; Verhoest, N.E.; Pauwels, V.R.; Adgo, E.; Deckers, J.; Poesen, J.; Nyssen, J. Water balance of a lake with floodplain buffering: Lake Tana, Blue Nile basin, Ethiopia. *J. Hydrol.* **2015**, *522*, 174–186. [CrossRef]

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

27. Dragoni, W.; Piscopo, V.; Di Matteo, L.; Gnucci, L.; Leone, A.; Lotti, F.; Melillo, M.; Petitta, M. Risultati del progetto di ricerca PRIN laghi 2003–2005. *G. Geol. Appl.* **2006**, *3*, 39–46. (In Italian)

28. Setegn, S.G.; Rayner, D.; Melesse, A.M.; Dargahi, B.; Srinivasan, R. Impact of climate change on the hydroclimatology of Lake Tana Basin, Ethiopia. *Water Resour. Res.* **2011**, *47*, W04511. [CrossRef]

29. Syed, T.; Famiglietti, J.; Chen, J.; Rodell, M.; Seneviratne, S.I.; Viterbo, P.; Wilson, C. Total basin discharge for the Amazon and Mississippi River basins from GRACE and a land-atmosphere water balance. *Geophys. Res. Lett.* **2005**, *32*. [CrossRef]

30. 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. *Geophys. Res. Lett.* **2017**, *44*, 5550–5560. [CrossRef]

31. Tong, K.; Su, F.; Li, C. Modeling of water fluxes and budget in Nam Co Basin during 1979–2013. *J. Hydrometeor.* **2020**, *21*, 829–844. [CrossRef]

32. Mahmood, R.; Jia, S.; Mahmood, T.; Mehmood, A. Predicted and projected water resources changes in the chari catchment, the lake Chad basin, Africa. *J. Hydrometeor.* **2020**, *21*, 73–91. [CrossRef]

33. 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 Syst. Sci. Data Discuss.* **2021**, *13*, 4349–4383. [CrossRef]

34. 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. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [CrossRef]

35. Mironov, D.; Heise, E.; Kourzeneva, E.; Ritter, B.; Schneider, N.; Terzhevik, A. Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. *Boreal Environ. Res. Publ. Board* **2010**, *15*, 218–230.

36. Balsamo, G. Interactive lakes in the integrated forecasting system. *ECMWF Newsl.* **2013**, *137*, 30–34.

37. Vincent, W. Effects of climate change on lakes. *Encyclopedia of Inland Waters*; Elsevier: Amsterdam, The Netherlands, 2009.

38. Dragoni, W.; Melillo, M.; Giontella, C. Bilancio idrico del lago Trasimeno. *Tutela Ambient. Lago Trasimeno* **2012**, *1*. (In Italian)

39. Ludovisi, A.; Gaino, E. Meteorological and water quality changes in Lake Trasimeno (Umbria, Italy) during the last fifty years. *J. Limnol.* **2010**, *69*, 174. [CrossRef]

40. Downing, J.A. Emerging global role of small lakes and ponds: Little things mean a lot. *Limnetica* **2010**, *29*, 9–24. [CrossRef]

41. Casadei, S.; Pierleoni, A.; Bellezza, M. Integrated water resources management in a lake system: A case study in central Italy. *Water* **2016**, *8*, 570. [CrossRef]

42. Ubertini, L.; Casadei, S.; Pierleoni, A. Lakes Management in Umbria Region, Central Italy. In Proceedings of the Taal 2007: The 12th World Lake Conference, Jaipur, India, 29 October–2 November 2007; pp. 1356–1364.

43. Thornthwaite, C.; Mather, J. Instructions and tables for computing potential evapotranspiration and the water balance. *Publ. Climatol.* **1957**, *1*, 206–208, 230–240 .

44. Vandewiele, G.; Xu, C.Y. Methodology and comparative study of monthly water balance models in Belgium, China and Burma. *J. Hydrol.* **1992**, *134*, 315–347. [CrossRef]

45. Balsamo, G.; Beljaars, A.; Scipal, K.; Viterbo, P.; van den Hurk, B.; Hirschi, M.; 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. *J. Hydrometeor.* **2009**, *10*, 623–643. [CrossRef]

46. Thiery, W.; Martynov, A.; Darchambeau, F.; Descy, J.P.; Plisnier, P.D.; Sushama, L.; van Lipzig, N.P. Understanding the performance of the FLake model over two African Great Lakes. *Geosci. Model Dev.* **2014**, *7*, 317–337. [CrossRef]

47. Shevnina, E.; Kourzeneva, E. Thermal regime and components of water balance of lakes in Antarctica at the Fildes peninsula and the Larsemann Hills. *Tellus Dyn. Meteorol. Oceanogr.* **2017**, *69*, 1317202. [CrossRef]

48. Betts, A.K.; Reid, D.; Crossett, C. Evaluation of the FLake model in ERA5 for Lake Champlain. *Front. Environ. Sci.* **2020**, *8*, 250. [CrossRef]

49. Manrique Suñén, A.; Nordbo, A.; Balsamo, G.; Beljaars, A.; Mammarella, I. Land surface model over forest and lake surfaces in a boreal site-evaluation of the tiling method. In Proceedings of the EGU General Assembly Conference Abstracts, Vienna, Austria, 22–27 April 2012; p. 2392.

50. Samuelsson, P.; Kourzeneva, E.; Mironov, D. The impact of lakes on the European climate as simulated by a regional climate model. *Boreal Environ. Res. Publ. Board* **2010**, *15*, 113–219.

51. ECMWF. *IFS Documentation CY45R1-Part IV: Physical Processes*; IFS Documentation; ECMWF: Reading, UK, 2018.

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

53. Balsamo, G.; Salgado, R.; Dutra, E.; Boussetta, S.; Stockdale, T.; Potes, M. On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model. *Tellus Dyn. Meteorol. Oceanogr.* **2012**, *64*, 15829. [CrossRef]

54. Tiedtke, M. Representation of clouds in large-scale models. *Mon. Weather. Rev.* **1993**, *121*, 3040–3061. [CrossRef]

55. Forbes, R.; Tompkins, A.; Untch, A. *A New Prognostic Bulk Microphysics Scheme for the IFS, ECMWF Technical Memoranda*; ECMWF: Reading, UK, 2011.

56. Tiedtke, M. A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Weather. Rev.* **1989**, *117*, 1779–1800. [CrossRef]

57. Bechtold, P.; Semane, N.; Lopez, P.; Chaboureau, J.P.; Beljaars, A.; Bormann, N. Representing equilibrium and nonequilibrium convection in large-scale models. *J. Atmos. Sci.* **2014**, *71*, 734–753. [CrossRef]

58. Bongioannini Cerlini, P.; Silvestri, L.; Saraceni, M. Quality control and gap-filling methods applied to hourly temperature observations over central Italy. *Meteorol. Appl.* **2020**, *27*, e1913.

59. Silvestri, L.; Saraceni, M.; Bongioannini Cerlini, P. Quality management system and design of an integrated mesoscale meteorological network in Central Italy. *Meteorol. Appl.* **2022**, *29*, e2060. [CrossRef]

60. Braca, G.; Bussettini, M.; Lastoria, B.; Mariani, S.; Piva, F. Il Bilancio Idrologico Gis BAsed a scala Nazionale su Griglia regolare—BIGBANG: Metodologia e stime. *Rapp. Sulla Dispon. Nat. Della Risorsa Idrica. Rapp. ISPRA* **2021**, *339*, 1–181.

61. Colvin, J.; Lazarus, S.; Splitt, M.; Weaver, R.; Taeb, P. Wind driven setup in east central Florida's Indian River Lagoon: Forcings and parameterizations. *Estuar. Coast. Shelf Sci.* **2018**, *213*, 40–48. [CrossRef]

62. Carrea, L.; Merchant, C. GloboLakes: Lake surface water temperature (LSWT) v4. 0 (1995–2016). *Cent. Environ. Data Anal.* **2019**, *10*, 29.

63. Klimaszyk, P.; Rzymski, P.; Piotrowicz, R.; Joniak, T. Contribution of surface runoff from forested areas to the chemistry of a through-flow lake. *Environ. Earth Sci.* **2015**, *73*, 3963–3973. [CrossRef]

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

65. Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* **2009**, *377*, 80–91. [CrossRef]

66. Sun, H.; Su, F.; Yao, T.; He, Z.; Tang, G.; Huang, J.; Zheng, B.; Meng, F.; Ou, T.; Chen, D. General overestimation of ERA5 precipitation in flow simulations for High Mountain Asia basins. *Environ. Res. Commun.* **2021**, *3*, 121003. [CrossRef]

67. Jiao, D.; Xu, N.; Yang, F.; Xu, K. Evaluation of spatial-temporal variation performance of ERA5 precipitation data in China. *Sci. Rep.* **2021**, *11*, 1–13. [CrossRef] [PubMed]

68. Muala, E.; Mohamed, Y.A.; Duan, Z.; 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 Sens.* **2014**, *6*, 7522–7545. [CrossRef]

69. Li, Y.; Zhang, Q.; Yao, J.; Tan, Z.; Liu, X. Assessment of water storage response to surface hydrological connectivity in a large floodplain system (Poyang Lake, China) using hydrodynamic and geostatistical analysis. *Stoch. Environ. Res. Risk Assess.* **2019**, *33*, 2071–2088. [CrossRef]

70. Beniston, M. The 2003 heat wave in Europe: A shape of things to come? An analysis based on Swiss climatological data and model simulations. *Geophys. Res. Lett.* **2004**, *31*. [CrossRef]

71. Black, E.; Blackburn, M.; Harrison, G.; Hoskins, B.; Methven, J. Factors contributing to the summer 2003 European heatwave. *Weather* **2004**, *59*, 217–223. [CrossRef]

72. 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. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 1953–1977. [CrossRef]

73. ISPRA. *Gli indicatori del Clima in Italia nel 2019. 20° Rapporto Annuale ISPRA*; ISPRA: Roma, Italy, 2020. (In Italian)

74. Silvestri, L.; Saraceni, M.; Bongioannini Cerlini, P. Links between precipitation, circulation weather types and orography in central Italy. *Int. J. Climatol.* **2022**, 1–19. [CrossRef]

75. Jiang, Q.; Li, W.; Fan, Z.; He, X.; Sun, W.; Chen, S.; Wen, J.; Gao, J.; Wang, J. Evaluation of the ERA5 reanalysis precipitation dataset over Chinese Mainland. *J. Hydrol.* **2021**, *595*, 125660. [CrossRef]

76. Xu, J.; Ma, Z.; Yan, S.; 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. *J. Hydrol.* **2021**, *605*, 127353. [CrossRef]

77. Sillmann, J.; Kharin, V.; Zhang, X.; Zwiers, F.; Bronaugh, D. Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *J. Geophys. Res. Atmos.* **2013**, *118*, 1716–1733. [CrossRef]

78. Moon, H.; Gudmundsson, L.; Seneviratne, S.I. Drought persistence errors in global climate models. *J. Geophys. Res. Atmos.* **2018**, *123*, 3483–3496. [CrossRef] [PubMed]

79. Taylor, K.E.; Stouffer, R.J.; Meehl, G.A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 485–498. [CrossRef]

80. 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. *Nat. Commun.* **2019**, *10*, 1–14. [CrossRef] [PubMed]

81. Schär, C.; Vidale, P.L.; Lüthi, D.; Frei, C.; Häberli, C.; Liniger, M.A.; Appenzeller, C. The role of increasing temperature variability in European summer heatwaves. *Nature* **2004**, *427*, 332–336. [CrossRef] [PubMed]

82. Taylor, C.M.; de Jeu, R.A.; Guichard, F.; Harris, P.P.; Dorigo, W.A. Afternoon rain more likely over drier soils. *Nature* **2012**, *489*, 423–426. [CrossRef]

83. Lee, K.O.; Flamant, C.; Ducrocq, V.; Duffourg, F.; Fourrié, N.; 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. *Q. J. R. Meteorol. Soc.* **2016**, *142*, 2623–2635. [CrossRef]

84. Barthlott, C.; Davolio, S. Mechanisms initiating heavy precipitation over Italy during HyMeX Special Observation Period 1: A numerical case study using two mesoscale models. *Q. J. R. Meteorol. Soc.* **2016**, *142*, 238–258. [CrossRef]

85. Hassler, B.; Lauer, A. Comparison of Reanalysis and Observational Precipitation Datasets Including ERA5 and WFDE5. *Atmosphere* **2021**, *12*, 1462. [CrossRef]

86. Tang, G.; Clark, M.P.; Papalexiou, S.M.; Ma, Z.; Hong, Y. Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. *Remote. Sens. Environ.* **2020**, *240*, 111697. [CrossRef]

87. Ayzel, G.; Izhitskiy, A. Climate change impact assessment on freshwater inflow into the Small Aral Sea. *Water* **2019**, *11*, 2377. [CrossRef]

88. Soja, G.; Züger, J.; Knoflacher, M.; Kinner, P.; Soja, A.M. Climate impacts on water balance of a shallow steppe lake in Eastern Austria (Lake Neusiedl). *J. Hydrol.* **2013**, *480*, 115–124. [CrossRef]

89. Gibson, J. Short-term evaporation and water budget comparisons in shallow Arctic lakes using non-steady isotope mass balance. *J. Hydrol.* **2002**, *264*, 242–261. [CrossRef]

90. Li, X.Y.; Xu, H.Y.; Sun, Y.L.; Zhang, D.S.; Yang, Z.P. Lake-level change and water balance analysis at Lake Qinghai, west China during recent decades. *Water Resour. Manag.* **2007**, *21*, 1505–1516. [CrossRef]

91. Bongioannini Cerlini, P.; Meniconi, S.; Brunone, B. Groundwater supply and climate change management by means of global atmospheric datasets. Preliminary results. *Procedia Eng.* **2017**, *186*, 420–427. [CrossRef]

92. Bongioannini Cerlini, P.; Silvestri, L.; Meniconi, S.; 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 Interact.* **2021**, *25*, 15–32. [CrossRef]

93. C3S. ERA5-Land hourly data from 1950 to present. *Copernic. Clim. Chang. Serv.* **2020**.