water

MDPI

Article

Impact of Geology on Seasonal Hydrological Predictability in Alpine Regions by a Sensitivity Analysis Framework

Maria Stergiadi [1,*], Nicola Di Marco [1], Diego Avesani [1,*], Maurizio Righetti [1] and Marco Borga [2]

[1] Faculty of Science and Technology, Free University of Bozen-Bolzano, Piazza Università 5, 39100 Bozen-Bolzano, Italy; nicola.dimarco@unibz.it (N.D.M.); maurizio.righetti@unibz.it (M.R.)

[2] Department of Land, Environment, Agriculture and Forestry, University of Padua, Viale dell' Università 16, 35020 Legnaro, Italy; marco.borga@unipd.it

* Correspondence: maria.stergiadi@natec.unibz.it (M.S.); diego.avesani@unibz.it (D.A.); Tel.: +39-329-4483-052 (M.S.)

Received: 24 June 2020; Accepted: 6 August 2020; Published: 11 August 2020

Abstract: Catchment geology has a major influence on the relative impact of the main seasonal hydrological predictability sources (initial conditions (IC), climate forcing (CF)) on the forecast skill as it defines the system's persistence. A quantification of its effect, though, on the contribution of the predictability sources to the forecast skill has not been previously investigated. In this work we apply the End Point Blending (EPB) framework to assess the contribution of IC and CF to the seasonal streamflow forecast skill over two catchments that represent the end members of a set of catchments of contrasting geology, hence contrasting hydrological response: a highly-permeable, hence slow-responding catchment and a fast-responding catchment of low permeability. Our results show that the contribution of IC in the slow-responding catchment is higher by up to 44% for forecasts initialized in winter and spring and by up to 21% for forecasts initialized in summer. IC are important for up to 4 months of lead in the slow-responding catchment and 2 months of lead in the flashier catchment. Our analysis highlights the added value of the EPB in comparison to the traditional ESP/revESP approach for identifying the sources of seasonal hydrological predictability, on the basis of catchment geology.

Keywords: seasonal streamflow forecast; initial conditions; sensitivity analysis

1. Introduction

Seasonal streamflow predictability is a key element to the efficient design of water-related management strategies as it has a major impact on the resilience of several sectors like hydropower production, water supply and agriculture. Seasonal hydrological forecasts can provide knowledge of the future land surface hydrological conditions several months in advance [1,2], contributing to early preparedness against disasters and to optimal water usage. Therefore, identifying the sources of skill of such forecasts can be crucial for human safety as well as for sustainable water management [3–7].

Seasonal hydrological forecasting relies on both statistical and dynamic methods. The former link land surface conditions or climate precursors to hydrological predictands [8–11], whereas the latter employ hydrological models to predict seasonal hydrology [12,13]. The predictive skill of such models depends principally on the initial hydrological conditions (IC) at the start date of the simulation and the climate forcing (CF) during the forecast period [6,7,14,15]. Predictability related to IC arises from the relatively slow evolution of land surface moisture states, mainly subsurface water content (SWC) and snow water equivalent (SWE), that defines the hydrological system's persistence

(i.e., its memory) [1,16]. CF is another source of skill influencing seasonal hydrological predictions and it can either be resampled from the historical climatology (Ensemble Streamflow Prediction approach, ESP) [17] or generated by climate forecast models [18–20].

Establishment of the dominant factor determining the seasonal forecast skill is crucial to funnel the efforts into improving the appropriate predictive tools, either by developing data assimilation techniques to improve the IC estimates or by enhancing the accuracy of the CF. The relative contribution of the two major sources of seasonal hydrological predictability (i.e., IC, CF) has been investigated in several catchments worldwide [2,14,21–24] by applying the ESP/reverse ESP (revESP) framework [25] that compares the predictability stemming from perfect IC and climatological CF (ESP) with the predictability related to known CF and climatological IC (revESP). Recent advances in the ESP/revESP methodology involve the End Point Blending (EPB) [1], a computationally-light alternative to the Variational Ensemble Streamflow Prediction Assessment (VESPA) [15] that accounts for intermediate levels of knowledge between perfect knowledge and climatological (total lack of knowledge), allowing for the calculation of the effect of changes in IC and CF skill on the streamflow forecast skill—a metric termed "skill elasticity" [15].

Many studies have focused on the impact of catchment location [2,22,25], size [21], elevation [23,24], season and lead time (LT) [2,22,25] on the ESP/revESP skill. A number of studies have also evaluated the effect of geology on the ESP seasonal predictive skill [14,23,24,26]. Geological characteristics have a substantial impact on catchment permeability, subsurface storage and evolution of moisture states, thereby determine flow variability and drive IC-related seasonal hydrological predictability [26–28]. Water stored in the soil and the snowpack establishes a memory effect controlled by the timescale required to dismiss the impact of varying soil wetness [29]. In permeable catchments with large groundwater storage that dampens the high-frequency rainfall fluctuations to a slower and smoother hydrological response, this memory effect persists longer whereas catchments of low permeability and storage have shorter memory of hydrological states, hence smaller persistence [24]. Therefore, seasonal hydrological predictability stemming from knowledge of IC is higher for regions with longer IC persistence, as defined by their geological characteristics. For example, Paiva et al. [14] suggested that the geological characteristics in the southeastern part of the Amazon could constitute groundwater storage important for seasonal discharge predictability during the low-flow period. The presence of high groundwater storages was related to increased streamflow forecast predictability related to known IC in the Seine catchment in northern France [23] and the UK [26]. Staudinger and Seibert [24] also confirmed longer persistence of IC in mountainous catchments in Switzerland with large groundwater storage.

The aforementioned studies, however, focused on identifying regions (and seasons) where (and when) predictability was related to perfect knowledge of IC, hence associated with skillful or unskillful ESP. On the other hand, the recently developed VESPA [15] and EPB [1] methodologies that evaluate the effect of varying levels of IC and CF uncertainty on the seasonal forecast skill, have been applied to regions that differ in terms of hydroclimatic conditions, without considering their geological properties. Therefore, to the best of our knowledge, no previous work has quantified the contribution of the hydrological predictability sources (IC, CF) to the seasonal streamflow forecast skill in regions of contrasting geology, hence differing persistence.

In this study we build upon the work of Norbiato et al. [27] who analyzed the streamflow characteristics of a set of alpine catchments of differing geology, hence of diverse subsurface storage capacity and hydrological response (with a focus on event runoff coefficients, though). From the set of catchments investigated in the above-mentioned study, we selected two catchments that represent the end members as far as geology-induced subsurface water storage capacity is concerned, for which we applied the EPB methodology to identify how geology affects the dominant factors (IC, CF) controlling the seasonal streamflow forecast skill. These two case studies are representative of two extreme hydrological behaviors, with one being characterized by a fast-responding hydrological regime due to the presence of low-permeability rocks and the other being dominated by highly-permeable lithological

units resulting in a slower response. Moreover, the case studies were chosen to be uninfluenced by human interventions (e.g., dams, irrigation, water diversion) and have similar characteristics (similar size, mean elevation and climatic conditions), differing only in terms of geology. This way we isolate the effect of the contrasting hydrological regime induced by the differing geological properties on the diverse contribution of the predictability sources to the seasonal streamflow forecast skill. Specifically, the experimental work addresses the following scientific questions: (1) what is the effect of geology on the contribution of each predictability source to the forecast skill; (2) how does the contribution of each predictability source to the forecast skill evolve with season and LT over catchments of contrasting geology; and (3) what is the added value of the EPB methodology in comparison to the traditional ESP/revESP approach for identifying the sources of seasonal hydrological predictability in regions of contrasting geological characteristics? The insights provided by answering these questions can set the guidelines for effectively selecting the components of a forecasting system for a given catchment on the basis of its geological characteristics, including investments in more sophisticated models, assimilation of higher-quality observational or remote-sensing data and better seasonal forecasts.

2. Materials and Methods

2.1. Case Studies

The case studies selected for this work are located in the upper Adige River Basin (Eastern Italian Alps; Figure 1a) and constitute the end members of a set of catchments that are characterized by diverse geology, hence by diverse water storage capacity and hydrological response. As aforementioned, this set has been previously analyzed by Norbiato et al. [27] who used the ratio of daily flow exceeded 90% of the time to the median daily flow as an indicator of subsurface water storage capacity. The reader is referred to the above-cited study for detailed information on the characteristics of the set of catchments. The case studies selected for the purpose of this paper involve two unregulated catchments of this set of similar size and mean elevation and fairly similar climate pattern but differing lithology, hence contrasting water storage capacity and hydrological response: the Passirio catchment closed at Merano (Figure 1b) that is fast-responding and the Gadera catchment closed at Mantana (Figure 1c) that is slow-responding. An analysis of these case studies allows for the embracement of relevant aspects of hydrological behavior within the river basin in relation to its geological characteristics.

The case studies feature the typical characteristics of alpine environments such as complex topography, high altitude and a snow-dominated hydrological regime but are diverse in terms of geology, implying diverse permeability and storage capacity. The Passirio catchment, located at the northwestern part of the upper Adige River Basin, is part of the Austroalpine geological group containing medium- to high-grade alpine and variscan metamorphic rocks like gneiss (orthogneiss and paragneiss) that are characterized by low permeability, resulting in low groundwater storage capacity. In the Gadera catchment at the eastern part of the basin the lithology changes to formations belonging to the Southern Alps geological group that comprise of metamorphic and carbonate rocks of low-grade metamorphism like phyllites (including limestones and dolostones), while karst processes prevail indicating high permeability and groundwater storage. The diverse properties of the case studies result in a variable hydrological response that renders them representative of the two extreme hydrological behaviors that characterize the basin: a fast-responding hydrological regime governed by quick subsurface flow and surface runoff generation dominates the Passirio catchment, whereas a slow-responding hydrological regime caused by the damping effect of the large groundwater storage prevails over the Gadera [27,30,31]. Land use consists mainly of grassland, sparse vegetation and coniferous forests in the Passirio catchment, while the Gadera catchment is mainly dominated by coniferous forests, meadows and sparsely vegetated areas including outcrop rocks of high permeability [27].

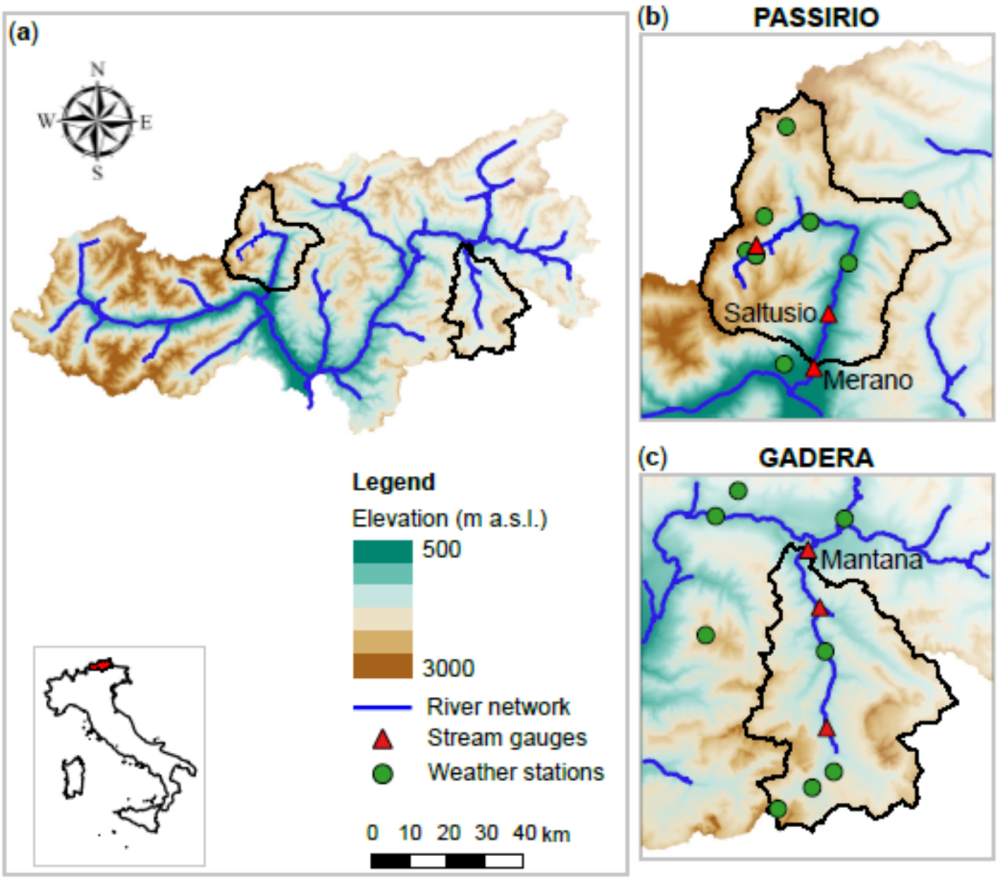

Figure 1. (**a**) Location of the two study catchments within the upper Adige River Basin. The inlet at the bottom left shows the location of the upper Adige River Basin in Italy; (**b**) Passirio catchment; and (**c**) Gadera catchment.

An overview of the main characteristics of the two case studies is presented in Table 1. The climate in the study areas is continental with dry winters and wet summers characterized by intense convective thunderstorms. The monthly distribution of precipitation exhibits two maxima during August and November, with a lower value of mean annual precipitation over the Gadera catchment due to the sheltering effect caused by the mountainous ranges located in the north and the south and a higher value over the Passirio attributed to the stau effect. For the period from October until March snow is accumulated in both catchments, with the end of March marking the onset of the melting period during which the largest part of runoff is generated. Seasonal streamflow generally follows the same pattern as precipitation, with low flows during winter due to snow accumulation in the high-altitude areas and peak flows due to snow melting during summer and cyclonic storms during autumn [27,32].

Table 1. Main characteristics of the two case studies.

Catchment	Passirio	Gadera
Gauging station	Merano	Mantana
Contributing area (km^2)	402	390
Mean elevation (m a.s.l.)	1874	1858
Elevation range (m a.s.l.)	360–3500	810–3050
Mean annual precipitation (mm)	1124	932
Mean annual temperature (°C)	3.1	2.6

Three streamflow gauging stations are present in the Passirio catchment (Figure 1b), with the station of Merano denoting the outlet section with a contributing area of approximately 402 km^2. A set

of eight precipitation and temperature stations located in an elevation range of 300–3000 m provide weather data for this catchment. In the Gadera catchment (Figure 1c) there are five precipitation stations in an elevation range of 700–1600 m and eight temperature stations covering a range of 700–3000 m, while three streamflow gauging stations exist also in this catchment closing at Mantana, with a contributing area of approximately 390 km^2.

For catchments located in similar climatic regions that imply comparable boundary forcing, catchment geology can exert a major influence on seasonal hydrological predictability as it defines the persistence of the catchment's land surface IC, i.e., the temporal dependence of sequential values of moisture states like SWC and SWE [33]. IC persistence can be quantified by calculating the autocorrelation of the above-mentioned moisture states for different time lags [29], as shown in Figure 2, where the autocorrelation was computed for time lags ranging from 1 to 6 months in accordance with the LT of the seasonal streamflow forecasting exercise presented in this study. For the calculation of the autocorrelation, data from all months in the historical record were considered for the SWC as well as for the SWE, that even in summer had extremely low monthly cumulative values which, however, never equaled zero.

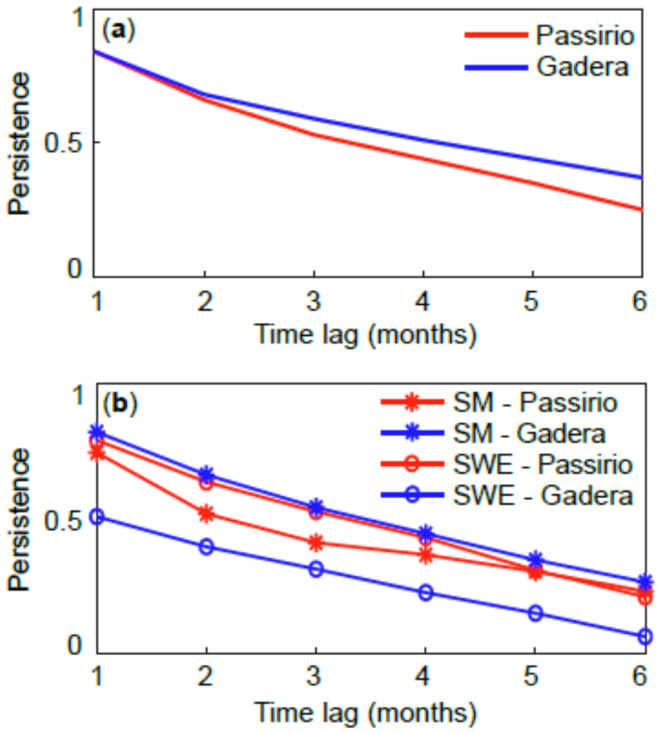

Figure 2. (**a**) Persistence of the land surface moisture states (sum of snow water equivalent (SWE) and subsurface water content (SWC)); and (**b**) persistence of the individual moisture states (SWE, SWC) in the two case studies.

The calculated autocorrelation (i.e., persistence) of the land surface hydrological states (sum of SWC and SWE) after time lag 2 was found to be higher in the Gadera catchment due to the longer memory imposed by large storages that render the statistical interdependence between successive time periods more important (Figure 2a). A further analysis of the individual components (i.e., SWC, SWE) affecting IC persistence indicated a higher SWE persistence for all time lags and a lower SWC persistence mainly for times lags 2–4 in the Passirio catchment, compared to the respective values calculated for the Gadera (Figure 2b). These findings imply that the hydrological regime in the Passirio catchment is strongly snow-dominated, in contrast to the hydrological regime in the Gadera where both snow and precipitation drive the hydrological response.

2.2. Hydrological Model

The Integrated Catchment-Scale Hydrological Model (ICHYMOD) [34–36] was employed in this study to perform the ESP and revESP experiments as prerequisites for the application of the EPB. The reader is referred to the above-cited literature for a detailed description of the model. Briefly, ICHYMOD is a continuous, conceptual, semi-distributed rainfall-runoff model operating on an hourly or daily time step and it comprises of modules describing snow accumulation and melt, glacier melt, soil moisture, groundwater and flow generation. Snow and glacier melt are calculated applying a distribution function approach according to a combined radiation index degree-day concept [37], as adjusted by Zaramella et al. [36]. Potential evapotranspiration is calculated based on the Hargreaves method [38]. The water storage capacity across the basin is described using a probability distribution function [39].

Saturation at any point in the basin leads to the generation of surface runoff, which integrated over the basin results in the total direct runoff entering the fast response pathways to the basin outlet. The direct runoff routing is based on the Muskingum–Cunge method [40]. Water drainage from the soil due to percolation enters the slow response pathways and is represented by a function of basin moisture storage [41]. The slow components (i.e., base flow) of the total runoff are routed through an exponential store. The total flow at the outlet of the basin is calculated as the summation of the storage representations of the fast and slow response pathways. For the purpose of this study the version of the model operating on a daily time step was used. Following the methodological approach adopted by previous studies that employed hydrological models running on a daily time step [2,21,25,26], the results were aggregated to monthly values given that aggregate (cumulative) values of hydrological variables are important for seasonal forecasting.

2.3. Model Calibration and Validation

Model parameter calibration was performed using the Particle Swarm Optimization method [42]. The Kling–Gupta efficiency (KGE) [43] was utilized for model calibration and validation. A simple split-sample test [44] was implemented by splitting the available dataset in two parts, one for calibration and one for validation. Observed streamflow data was available throughout the period 2000–2018 for the gauging station at Mantana (outlet of the Gadera catchment), whereas for the gauging station at Merano (outlet of the Passirio catchment) observational data was available only for the period 2013–2018. Hence, for the Gadera catchment the calibration period was set from 1 October 2010 until 1 October 2018 (due to the availability of observational data of better quality with less gaps) and the validation period from 1 October 2000 until 1 October 2010, whereas for the Passirio the calibration was performed from 1 October 2013 until 1 October 2015 and the validation from 1 October 2015 until 1 October 2018. The KGE values for the calibration were 0.76 and 0.79 for the Passirio and the Gadera, respectively, whereas in validation mode the KGE dropped to 0.75 for the Passirio and 0.78 for the Gadera. The efficiency over the whole period that data was available was 0.77 for the Passirio and 0.81 for the Gadera.

Given the short time series of available streamflow data for the Merano gauging station and to verify the model's performance over the Passirio catchment, simulated streamflow values (i.e., after model calibration) were compared to observed values at the gauging station of Saltusio (Figure 1b) (with a contributing area of approximately 342 km^2), located approximately 6 km upstream of Merano (contributing area of 402 km^2). The KGE over the whole simulation period for the Saltusio station was equal to 0.73, hence we concluded that the ICHYMOD model is able to efficiently reproduce the temporal changes in streamflow in both case studies.

The differing geology of the two case studies translated into differential parameterization of the subsurface processes that define the contribution of the main flow components (surface flow, subsurface flow, base flow) to the total runoff simulated by the hydrological model. Geologic characteristics in the flashier Passirio catchment resulted in a low fraction of the simulated total flow originating from base flow (nearly 30%), while in the more permeable Gadera catchment approximately 70% of the

simulated total flow stemmed from the large groundwater aquifers. These values are in line with the base flow index (BFI) [45] calculated from the observed streamflow data, after applying a simple separation method.

2.4. Experimental Design

This study is focused on the implementation of the EPB elasticity framework for the generation of hindcasts (i.e., retrospective forecasts) with a LT of 1 to 6 months, initialized on the first day of each calendar month of the period 2002–2018. To this end, we utilized ESP and revESP hindcasts as well as a simulation of the climatology (CLIM). The period 2000–2001 was used for model spin-up. Moreover, following the common approach adopted by previous studies [2,15,25,26,46–48], a retrospective simulation spanning the entire period was performed to create surrogate observations of streamflow data in order to exclude model error. This simulation (hereon referred to as perfect) was employed both for the application of the EPB and for the evaluation of the performance of the EPB-generated hindcasts. The four above-stated experiments (ESP, revESP, CLIM and perfect) represent the four end points of uncertainty in the predictability sources and constitute the prerequisites for the application of the EPB analysis.

As aforementioned, the daily forecasted streamflow time series resulting from the ESP, revESP and CLIM experiments as well as from the perfect simulation were aggregated to monthly values as, in view of the relevant uncertainties within seasonal hydrological applications (e.g., for reservoir management), the interest lies mostly in monthly uncertainties. Model simulations, however, were performed on a daily time step to provide accurate estimates of IC on the first day of each calendar month, when the hindcasts were initialized. In this work, month one of LT refers to the first month of the forecast. For example, for a forecast initialized on 1 January, month one of LT refers to January, month two to February and so on.

2.4.1. EPB Implementation

Both the ESP and the revESP rely on the assumption of perfect knowledge of either IC or CF, respectively. In reality, however, even the best estimates or measurements of hydrological and meteorological variables encompass a certain degree of uncertainty. Hence, an analysis of intermediate levels of uncertainty between perfect knowledge and climatological (i.e., deducted by the distribution of all data in the historical record) can provide a more realistic insight on the seasonal hydrological predictability under different levels of accuracy in the predictability source (IC and/or CF). The strategy involves blending linearly the perfect and climatological IC and CF, using weights [15]. The weights utilized for the blending were $w = (0,\ 0.05,\ 0.10,\ 0.25,\ 0.50,\ 0.75,\ 0.90,\ 0.95,\ 1.0)$. A weight of zero represents zero uncertainty (perfect knowledge), while a weight equal to one represents climatological uncertainty (total lack of knowledge). Weights denoted as w_{IC} refer to the weights used to blend the IC, while weights denoted as w_{CF} refer to the weights used to blend the CFs. Thus, using $w_{IC} = 0$ and $w_{CF} = 1$ yields the ESP forecast, $w_{IC} = 1$ and $w_{CF} = 0$ yields the revESP forecast, $w_{IC} = 1$ and $w_{CF} = 1$ yields the CLIM forecast and $w_{IC} = 0$ and $w_{CF} = 0$ yields the perfect forecast. For every combination of w_{IC} and w_{CF} the forecast skill can be calculated, resulting in skill plots displaying the forecast skill as a function of the IC and CF skill (Figure S1 in the Supplementary Materials). The horizontal and vertical axes represent the IC and CF skill, respectively, corresponding to weights w_{IC} and w_{CF}, and result from the following equations:

$$\text{IC skill} = 100 \times \left(1 - w_{IC}^2\right) \tag{1}$$

$$\text{CF skill} = 100 \times \left(1 - w_{CF}^2\right) \tag{2}$$

Hence, for a weight equal to zero (i.e., zero uncertainty) the skill of the predictability source (IC, CF) will have its maximum value (100%), corresponding to a forecast with perfect knowledge of

the specific predictability source. On the contrary, for a weight equaling unity the skill will be null, indicating complete lack of knowledge in the predictability source.

As shown in Figure S1, the grey rhomboid shapes represent a specific forecast for each combination of IC and CF skill, whereas the red circles are the four end points of uncertainty: the ESP, the revESP, the perfect simulation and the CLIM. The EPB approach [1] involves combining the four end points of the variable weight assessment [15] using for the blending the same weights as the VESPA methodology. Thus, new hindcasts are generated for each w_{IC}–w_{CF} combination, without the necessity to perform the additional computationally-heavy simulations required by the VESPA. For each combination point, the four end points are blended given the specific weights to create a new ensemble of 100 members. The percentage of each point used to generate the new ensemble (i.e., the number of members randomly resampled from each end point) is provided by the following equation:

$$EP(\%) = (1 - |x_{EP} - w_{IC}|) \times \left(1 - \left|y_{EP} - w_{CF}\right|\right) \quad (3)$$

where w_{IC} and w_{CF} refer to the weights of the specific combination point for which the new ensemble is created, while x_{EP} and y_{EP} refer to the w_{IC} and w_{CF} values of the end point for which the percentage is calculated, respectively (e.g., if the combination point reproduced refers to the ESP, then $w_{IC} = x_{EP} = 0$ and $w_{CF} = y_{EP} = 1$; hence 100% of the EPB hindcasts are resampled from the ESP point).

The skill plots obtained with the EPB method allow for the quantification of the change in forecast skill related to a change in IC and/or CF skill. To this end, the skill elasticities for the IC and the CF can be calculated as the gradient in streamflow prediction skill relative to changes in IC and CF skill. For the purposes of this study, the skill elasticities were calculated across three transects in the center of the response surface (highlighted with the magenta-colored rectangle in Figure S1 in the Supplementary Materials) and averaged according to the equations:

$$E_{IC} = \left\{ \frac{S[F(75,19)] - S[F(19,19)]}{75\% - 19\%} + \frac{S[F(75,44)] - S[F(19,44)]}{75\% - 19\%} + \frac{S[F(75,75)] - S[F(19,75)]}{75\% - 19\%} \right\} / 3 \quad (4)$$

$$E_{CF} = \left\{ \frac{S[F(19,75)] - S[F(19,19)]}{75\% - 19\%} + \frac{S[F(44,75)] - S[F(44,19)]}{75\% - 19\%} + \frac{S[F(75,75)] - S[F(75,19)]}{75\% - 19\%} \right\} / 3 \quad (5)$$

where the numerators, expressed as $S[F(\text{IC skill}, \text{CF skill})] - S[F(\text{IC skill}, \text{CF skill})]$, comprise of the gradients in streamflow forecast skill between IC skill (or CF skill) values of 75% and 19% (the denominator). A positive skill elasticity value implies that an improvement in the predictability source (IC or CF) results in an improvement in the streamflow forecast skill. Negative skill elasticity values imply that improving the skill of the predictability source leads to a deterioration of the forecast skill, whereas a skill elasticity of zero indicates that improving the predictability source has no influence on the forecast skill. The calculated skill elasticities were transformed to percentages of contribution of each predictability source to the streamflow forecast skill according to the equations:

$$\text{IC contribution} = (E_{IC}/(E_{IC} + E_{CF})) \times 100 \quad (6)$$

$$\text{CF contribution} = (E_{CF}/(E_{IC} + E_{CF})) \times 100 \quad (7)$$

2.5. Forecast Evaluation

In this study the skill of the EPB streamflow hindcasts is calculated as a skill score that indicates the improvement of a specific forecast over a reference forecast. The generic equation for the calculation of a skill score is given by:

$$\text{Skill score}_{HC} = 1 - \frac{\text{score}_{HC}}{\text{score}_{Ref}} \quad (8)$$

where *HC* refers to the EPB hindcast and *Ref* to the reference forecast against which the EPB hindcasts were benchmarked, here the CLIM. The skill score equals unity for a perfect forecast. Declining values of the skill score indicate declining forecast skill until the value of zero where the forecast skill equals

that of the reference forecast. Negative skill score values imply that the forecast is less skillful than the reference (CLIM).

For the evaluation of the EPB hindcasts, deterministic and probabilistic metrics were employed to assess different forecast attributes like accuracy, sharpness and reliability. The probabilistic metric used is the Continuous Ranked Probability Score (CRPS, and its corresponding skill score, CRPSS), that is the most common score exploited for the evaluation of the general performance of hydrological ensemble forecasts [48]. The deterministic metrics involve the mean squared error skill score (MSEss) and the mean absolute error skill score (MAEss, the deterministic equivalent of the CRPSS) and were calculated using the ensemble mean. The pattern of the results regarding the contribution of the predictability sources to the seasonal streamflow forecast skill over the two case studies of contrasting geology was found to be practically irrelevant of the forecast metric applied, hence, to the interest of parsimony, the rest of the paper is based upon the MAEss only.

3. Results

3.1. Hydroclimatic Regime

The observed temperature and precipitation as well as the simulated hydrological variables for the two case studies averaged over the simulation period (2002–2018) are presented in Figure 3. The climate pattern in the two case studies is fairly similar, with increased precipitation during summer and autumn and decreased precipitation during winter and spring. The average annual precipitation is higher over the Passirio catchment, with monthly precipitation values being equal to that over the Gadera only for the period from June until September. Snow accumulation, as indicated by the increasing values of SWE in both catchments, starts in October and reaches its maximum values in February and March, after which the increased temperature leads to melting and to the generation of the largest part of runoff volume, which at monthly time scales in small catchments equals approximately streamflow [25]. Increased spring and summer temperatures also cause increased losses due to evapotranspiration. Snow accumulation is higher over the Passirio catchment resulting in a mostly snow-dominated hydrological regime, with declining SWC and streamflow trends after the major part of the snowpack has melted. The Gadera catchment, on the other hand, is dominated by both snow and precipitation leading to increased SWC and streamflow values even after the melting season is over. Simulated SWC exhibits its maximum value during June in the Passirio catchment and during September in the Gadera, while the highest SWC variability (in terms of interannual standard deviation) in both case studies is manifested in May.

The low groundwater storage capacity over the Passirio catchment results in a quick response to rainfall, hence in a highly variable hydrological regime as illustrated in Figure 3a. In contrast, the variance of monthly flow over the Gadera catchment (Figure 3b) is lower due to the attenuating effect of large groundwater storages sustained by the karstified and limestone/dolostone aquifers that prevail in the subsurface. The contrasting hydrological behavior of the two case studies is also reflected in the value of the coefficient of variation (CV) of monthly flow that equals 0.74 for the flashier Passirio catchment, while it drops to 0.44 for the slower-responding Gadera catchment.

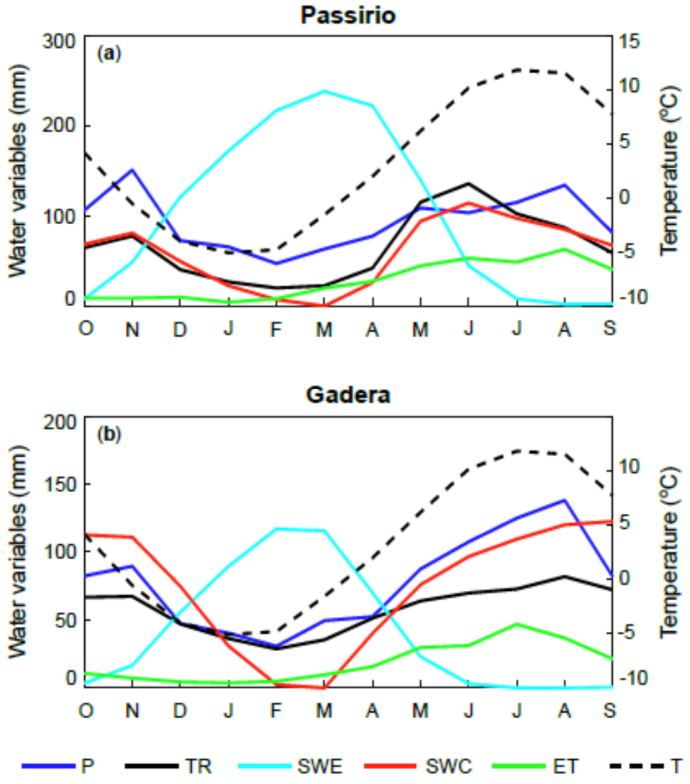

Figure 3. Mean monthly hydrological variables (left vertical axis) and temperature (right vertical axis) for (**a**) the Passirio catchment; and (**b**) the Gadera catchment, over the period 2002–2018. Mean precipitation (P) and temperature (T) refer to observed values; snow water equivalent (SWE), evapotranspiration (ET), subsurface water content (SWC) and total runoff (TR) refer to simulated values. The simulated SWC drawn has been reduced by the lowest monthly value to show only the active range. Note the difference in the left vertical axes between the two case studies, as well as the difference in the left and right vertical axes for each case study.

3.2. Predictability Variations

As shown in Figure 3, the land surface hydrological conditions and the climate conditions vary considerably with the season under consideration. The variation of streamflow, however, is affected not only by the climatic settings but also by the subsurface processes that highly depend on the catchment's geological characteristics, ultimately affecting the water storage capacity and the catchment's memory. These variations are shown in Figures 4 and 5 and Figure S2 in the Supplementary Materials for chained time series of ensemble streamflow hindcasts with varying levels of skill in the predictability sources (IC, CF), as generated by the EPB approach for the two case studies of differing geology. Figures 4 and 5 refer to LT equal to 1 and 6 months, respectively, while Figure S2 refers to LT 3 months. For better visual inspection, only the results for the period 2002–2007 are presented. The black line represents the time series of surrogate observed streamflow data (perfect simulation) while the red line illustrates the forecast ensemble mean. The uncertainty band between the 5th and 95th percentiles is also shown. The 90% interquartile range (i.e., the difference between the 95th and the 5th percentiles of the forecast distribution) is used as a measure of the forecast spread, with lower spreads indicating less uncertain, hence sharper forecasts [49]. The left panel of the graph for each of the two case studies depicts forecasts with decreasing level of IC skill (from top to bottom) considering climatological uncertainty for the CF, while the right panel shows forecasts with decreasing level of CF skill, considering unskillful IC. Thus, the top left graph for each case study represents the ESP forecast, the top right the revESP and the bottom graph the CLIM.

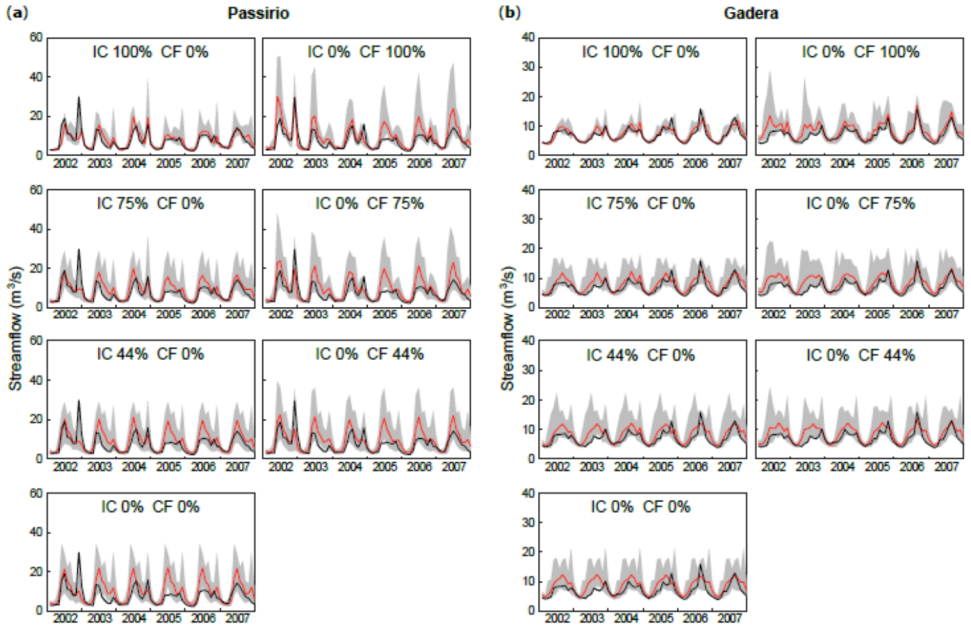

Figure 4. Chained time series of ensemble streamflow forecasts initialized on the first day of each month for varying levels of initial conditions (IC) and climate forcing (CF) skill (assuming null skill in the alternate predictability source), for lead time (LT) 1 month (**a**) over the Passirio catchment; and (**b**) over the Gadera catchment. The black line represents the surrogate streamflow observations created by the perfect simulation and the red line represents the forecast ensemble mean. The range of the forecast streamflow values (forecast spread) is shown by the grey 5th and 95th percentile band. Note the difference in the vertical axis between the two case studies.

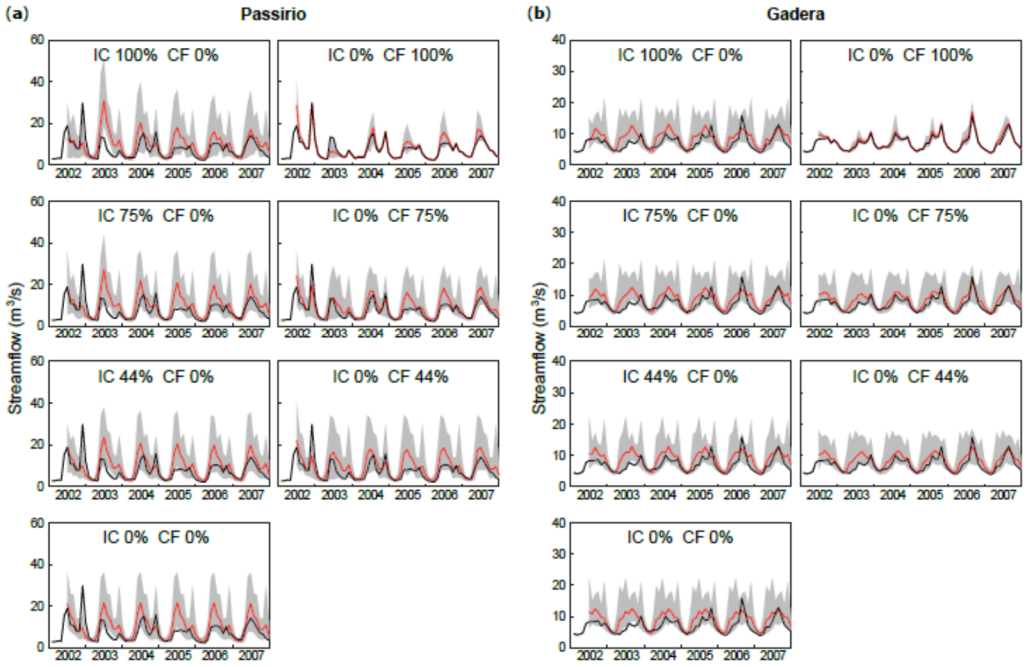

Figure 5. Chained time series of ensemble streamflow forecasts initialized on the first day of each month for varying levels of initial conditions (IC) and climate forcing (CF) skill (assuming null skill in the alternate predictability source), for LT 6 months (**a**) over the Passirio catchment; and (**b**) over the Gadera catchment. Line colors and forecast spreads are explained in Figure 4. Note the difference in the vertical axis between the two case studies.

In the Passirio catchment, a decrease in the IC skill from 100% (ESP) to 75% for the first month of LT (Figure 4a) magnified substantially the spread and reduced the forecast skill, whereas an additional reduction in the IC skill to 44% and 0% (the latter representing the CLIM) resulted in a small further amplification of the forecast spread and in additional deterioration of the forecast skill. The same experiments for the Gadera catchment (Figure 4b) revealed an even more dramatical increase in the spread and a sharper decline in the forecast skill for a decrease in the ESP IC skill from 100% (ESP) to 75%, whereas also for this catchment further reductions in the IC skill led to a slight spread magnification but higher forecast error.

A gradual decrease in the CF skill, ranging from 100% (revESP) to 0% (CLIM), led to a rather small increase in the forecast spread and error for all cases in both case studies (Figure 4a,b), implying that, for such a short lead, CF does not exert a major effect on the forecast's predictive capability. IC clearly dominate streamflow forecast predictability in both case studies for the first month of LT as indicated by the narrower spreads of the graphs on the left panel compared to the graphs on the right panel. A reduction in the IC skill below 75% led to spreads comparable to that of the CLIM forecast but resulted in a considerable increase of the forecast error, suggesting the importance of IC persistence in providing forecast skill in both the case studies of contrasting geological characteristics. In the Gadera catchment, however, IC predictability plays a more critical role for the forecast skill and sharpness in comparison to the Passirio, as even a small decrease in the IC skill resulted in a much wider amplification of the spread (hence in reduced sharpness) and in a vast decrease in the forecast skill.

Forecast predictions for the third month of LT in the Passirio catchment (Figure S2a in the Supplementary Materials) induced no notable differences in the forecast spread when IC skill diminished from perfect (ESP) to climatological, whereas the forecast skill decreased slightly when IC skill dropped below 44%. On the contrary, a marginal decrease in the CF skill from perfect (revESP) to 75% resulted in a considerable increase in the spread and the forecast error. Further reduction in CF skill led to an additional sharp decline in the forecast skill but had little effect on the spread. The stronger memory of IC in the Gadera catchment (Figure S2b in the Supplementary Materials) caused an increase in the spread and a considerable gradual decrease in the forecast skill when reduced knowledge of IC ranging from 75% to climatological was introduced. However, a reduction in the CF skill also yielded an increase in the spread, albeit lower than the one noticed in the Passirio and a major deterioration of the forecast skill. From these results it is evident that by the third month of LT the memory of antecedent moisture conditions in the Passirio catchment has waned while predictability related to knowledge of CF starts dominating the hydrological regime. In the Gadera, however, both IC and CF contribute to the forecast skill, with IC-related predictability favoring the forecast skill during the low-flow period and CF skill benefitting the forecast skill during the high-flow period.

By the sixth month of LT and as deducted by the graphs shown in Figure 5, all predictability in both case studies arises from knowledge of CF. Forecasts of any level of knowledge of IC (combined with unskillful CF), as depicted in the left panel of the graphs, are characterized by an ample ensemble spread (hence large uncertainty) and a low forecast skill, as a result of the land surface memory in the two case studies not extending to such long leads. In the Passirio catchment (Figure 5a), a decrease in IC skill generally resulted in a slightly narrower forecast spread but had no impact on the forecast skill, whereas a decline in the CF skill from perfect (revESP) to climatological resulted in a vast increase in the spread even for a small reduction (by 25%) and in gradual diminution of the forecast skill. In the Gadera catchment (Figure 5b), declining IC skill led to a negligible increase in the forecast spread and error, while even a small degradation in the CF skill caused the forecast spread to reach the spread of the CLIM forecast and the forecast skill to decay dramatically.

3.3. Plots of Streamflow Forecast Skill

Better understanding of the predictive capability of the hindcasts of varying level of IC and CF skill can be acquired by drawing skill plots displaying the streamflow forecast skill as a function of the

IC and CF skill. Such plots can provide a more comprehensive qualitative assessment of the primary predictability source affecting the streamflow forecast skill in the two case studies of contrasting geology. Figures 6–9 show plots of the streamflow forecast skill (MAEss) for the two case studies and for forecast initialization on the first day of December, March, June and September, respectively. These forecasts were selected to represent the changes in streamflow forecast skill over each season. Skill plots for other indicative initialization months are displayed in Figures S3–S6 in the Supplementary Materials. Along the horizontal axis, IC skill increases from left to right with the end point representing the ESP (perfect IC with climatological CF). Likewise, along the vertical axis CF skill increases from bottom to top, with the end point corresponding to the revESP (perfect CF with climatological IC). The perfect forecast is situated in the top right corner of each panel (perfect IC and CF), whereas the CLIM is located in the bottom left corner (climatological uncertainty in both IC and CF). Generally, (more) vertical gradients indicate predictability stemming from knowledge of the IC, while (more) horizontal gradients suggest predictability arising from knowledge of CF.

For both catchments and for forecasts initialized during early to mid-winter (December, Figure 6; and January (figure not included for the sake of brevity), IC are generally important for streamflow predictability for the first two months of LT that include the low-flow period when streamflow originates principally from SWC-driven base flow, with some IC predictability also extending until the third month of lead over the Gadera catchment. The forecast skill related to IC is generally higher over the highly-permeable Gadera catchment (indicated by the steepest gradient of the contours) as a result of the higher SWC persistence in this catchment (as shown in Figure 2b) during the low-flow period when the major part of streamflow consists of base flow. Forecasts initialized in February (Figure S3 in the Supplementary Materials) show the longest predictability arising from IC as melt-dominated months with a major contribution to streamflow enter the forecast. The importance of IC is notable up to the fifth month of lead, with increased IC predictability suggested for the Gadera catchment. During March as the second month of LT (Figure S3) and for both case studies, predictability originating from IC results somehow lower compared to the subsequent month of April (as the third month of lead) due to the lack of snowmelt contributing to streamflow.

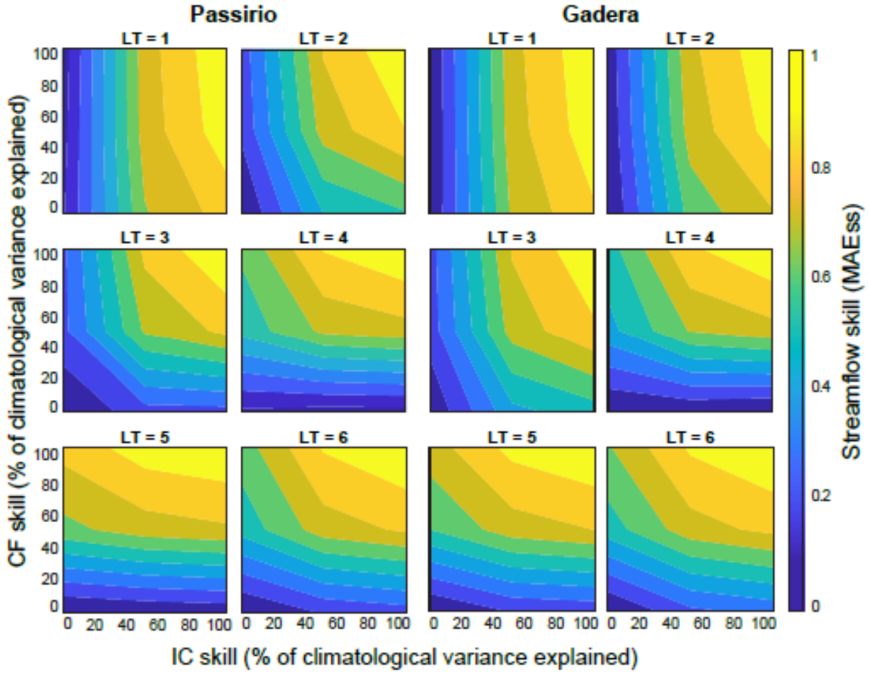

Figure 6. Plots of the mean absolute error skill score (MAEss) for the two case studies, for hindcasts generated over the period 2002–2018 and forecast initialization on 1 December, versus the skill in the two predictability sources (IC, CF). The results are presented for a LT of 1 to 6 months.

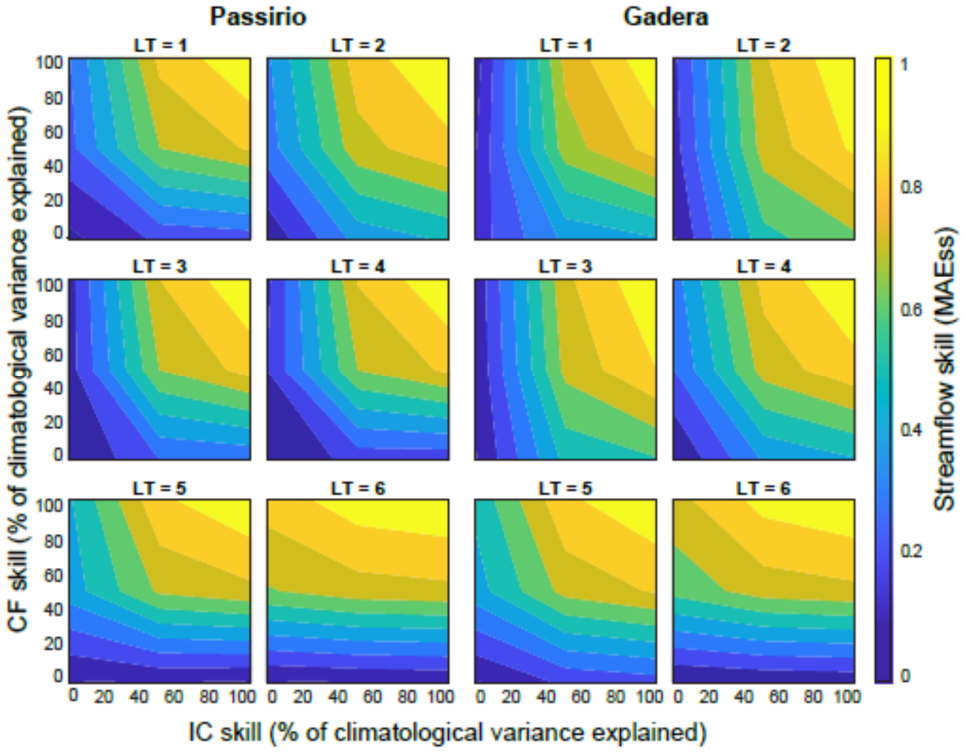

Figure 7. As in Figure 6 but for forecast initialization on 1 March.

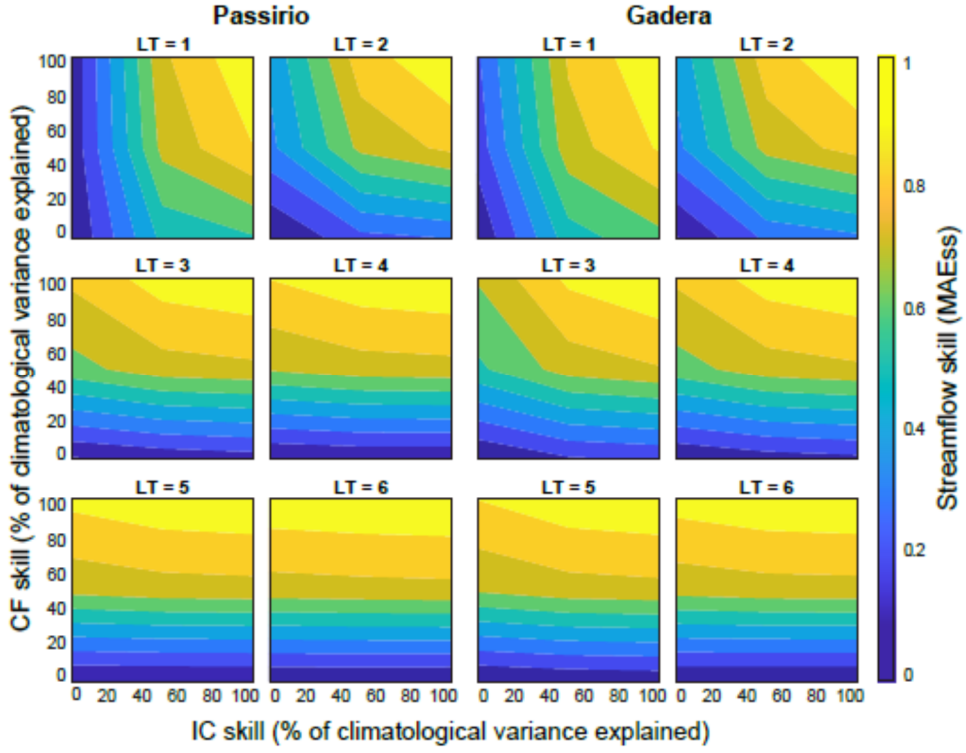

Figure 8. As in Figure 6 but for forecast initialization on 1 June.

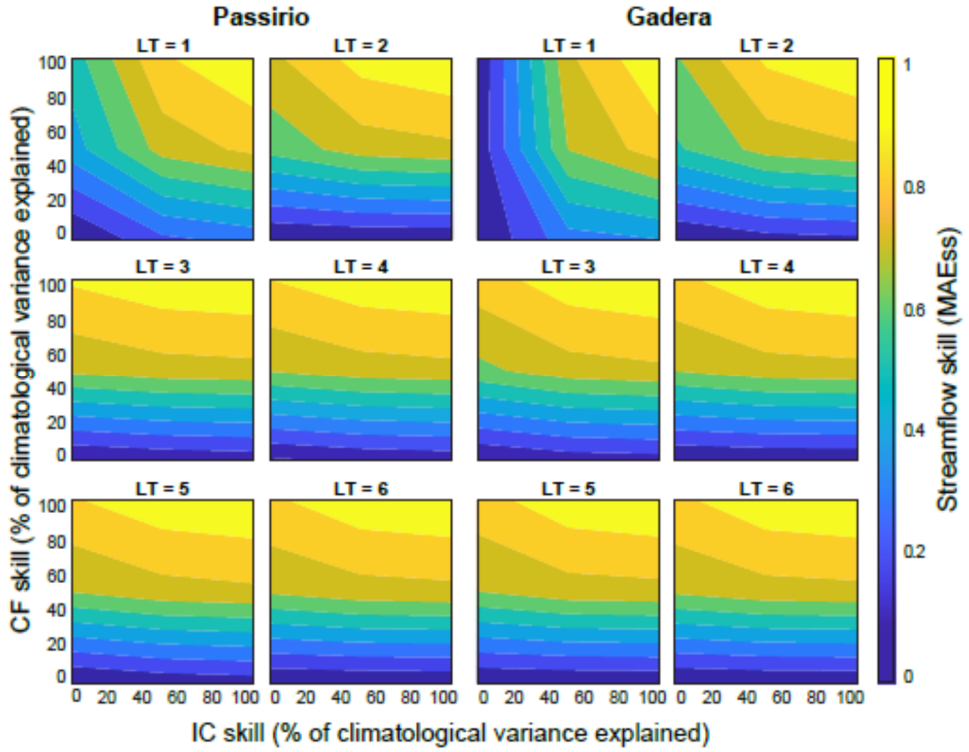

Figure 9. As in Figure 6 but for forecast initialization on 1 September.

For forecasts initialized in spring, IC are critical for one to four months of LT due to the onset of melt-based streamflow generation (with diminishing duration of IC predictability moving from forecast initialization in March through May). However, the plots reveal a considerably higher IC predictability in the Gadera catchment compared to the Passirio for forecasts initialized in March (Figure 7) and April (Figure S4 in the Supplementary Materials) as a result of the higher IC persistence exhibited in the Gadera, whereas for initialization in May (figure not included for the sake of brevity) the forecast skill arising from knowledge of IC in the two case studies results quite similar. Generally, the high SWC variability manifested in May in both case studies leads to increased forecast predictability arising from knowledge of IC for all forecasts that include this month in the forecasting period. As expected, this effect wanes as May shifts away from the initialization date.

During early to mid- summer (June, Figure 8; and July, Figure S5 in the Supplementary Materials), IC are important only for the first month of LT, after which CF becomes critical as increased precipitation controls the hydrological regime during the high-flow period. Knowledge of IC in the Passirio catchment in June seems to be more important than in the Gadera (as implied by the slightly steeper gradient of the contours for LT 1 month) as a result of both increased SWC states and the ongoing contribution of snowmelt to streamflow. For forecasts initialized in August (figure not shown in the interest of brevity), September (Figure 9), October (Figure S6 in the Supplementary Materials) and November (figure not presented for the same of brevity), CF is generally fundamental as a consequence of the rainfall effects on streamflow. Two exceptions are noted: for September as the first month of LT (Figure 9) over the Gadera catchment where IC predictability results to be highly important due to the increased SWC levels and SWC persistence, combined with a decline in precipitation, as well as for October as the first month of lead (Figure S6 in the Supplementary Materials) over the Passirio, increased and faster snow accumulation and higher SWE persistence compared to the Gadera lead to increased streamflow predictability related to IC.

The above qualitative analysis based on the forecast skill plots solely gives a general outlook of the sensitivity of the streamflow forecast skill to changes in the IC and CF skill and how this differs between the two case studies of contrasting geology, given the season and LT under consideration.

Overall, predictability originating from knowledge of IC is more important for the slower-responding Gadera catchment for forecasts initialized in winter and spring, whereas summer and autumn forecasts show nearly the same sensitivity to changes in the IC and CF skill (with the exceptions of September and October as LT 1 month, as stated in the previous paragraph). Moreover, IC play a major role to forecast predictability for longer leads over the Gadera catchment as a result of its longer memory of moisture states that transfers predictability to longer leads.

3.4. Contribution of Predictability Sources to Forecast Skill

A deeper insight into the benefits of improving the IC or CF skill as depicted in the potential improvements in streamflow forecast skill can be achieved by calculating the forecast skill elasticities as described in Section 2.4.1. Equations (4)–(7) allow for a quantitative assessment of the contribution of each predictability source to the seasonal streamflow forecast skill, ultimately indicating how geology affects the hydrological predictability in the two case studies of contrasting geological characteristics. Figure 10 shows the contribution of IC (Figure 10a) and CF (Figure 10b) to the seasonal streamflow forecast skill (expressed in percentage of contribution to the total forecast skill) over the Passirio catchment (red line) and the Gadera catchment (blue line), for each initialization month and for LT equal to 1, 3 and 6 months.

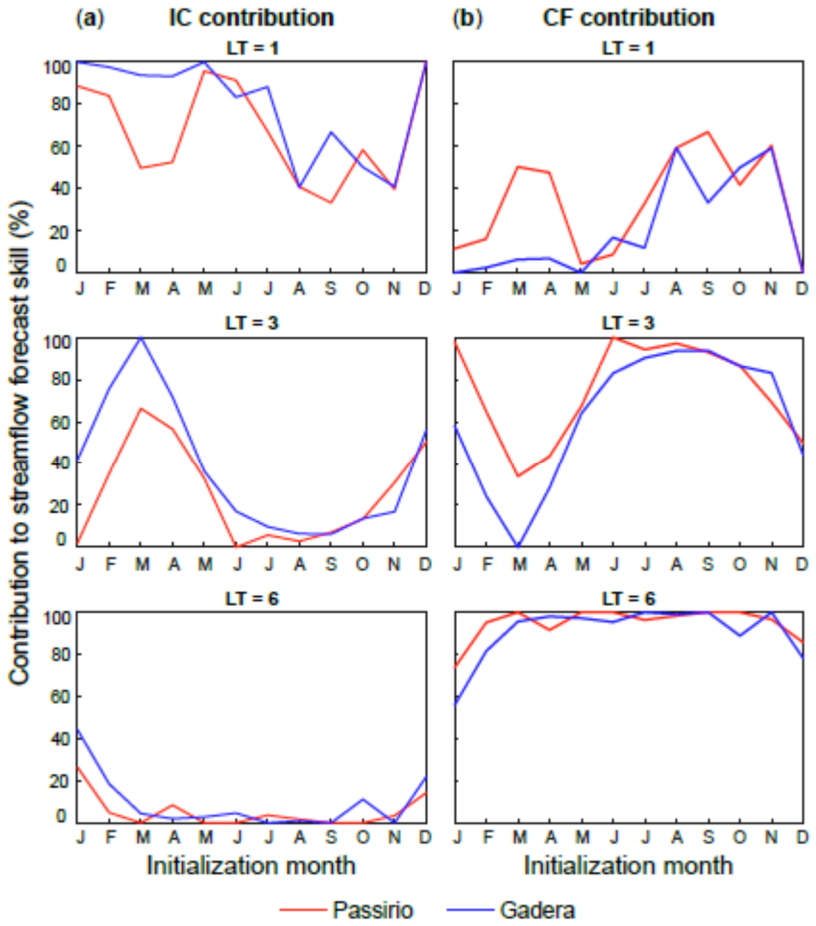

Figure 10. Contribution of (**a**) IC; and (**b**) CF to the streamflow forecast skill (expressed in percentage of the total forecast skill) over the two case studies, for each initialization month and LT of 1, 3 and 6 months.

It is evident that differing geology results in different persistence of the IC, which in turn leads to a different contribution of IC and CF to the seasonal streamflow forecast skill. The contribution of

IC over the more-permeable Gadera catchment is generally higher than the contribution of IC over the less-permeable Passirio for all seasons and LTs, with the difference being more pronounced for shorter leads. The contribution of CF is generally higher over the Passirio catchment due to the shorter persistence of moisture states that transfers predictability to the meteorological forcing. A comparison of the relative importance of the two predictability sources over the case studies suggests that IC dominate the forecast skill for forecasts initialized during the low-flow season and the beginning of the melting season for up to four months of lead in the slow-responding Gadera catchment and for up to two months of lead in the faster-responding Passirio, with their contribution to the forecast skill exceeding 75%. For longer leads as well as for forecasts initialized during summer and autumn, mainly as of LT 2, the contribution of CF to the forecast skill is generally higher than that of IC. However, for the first month of lead of late summer and autumn forecasts and over both case studies, the contribution of IC and CF to the forecast skill is comparable.

Given that the focus of this work is the analysis of the impact of geology on the contribution of the predictability sources to the seasonal streamflow forecast skill, it would be useful to investigate the difference in the contribution of the IC to the forecast skill between the two case studies of contrasting geological characteristics. This difference is presented in Figure 11a for each initialization month and LT. Positive values denote that the contribution of IC over the Gadera catchment exceeds that over the Passirio whereas negative values (highlighted with the red-colored boxes) demonstrate the opposite. The results have also been averaged over each initialization month (Figure 11b) and LT (Figure 11c), to gain an overall perspective of the evolution of the contribution of IC to the forecast skill over the different seasons and LTs.

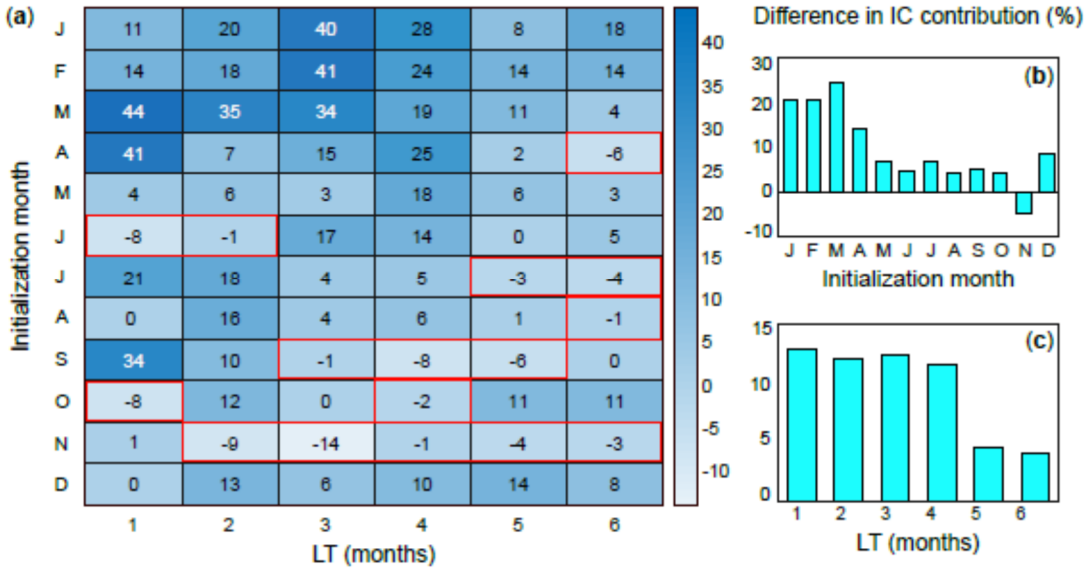

Figure 11. (**a**) Difference in IC contribution (in %) to the streamflow forecast skill for each initialization month (along each line) and LT (along each column) between the two case studies. Positive values indicate that the contribution of IC over the Gadera exceeds that over the Passirio, whereas negative values (included within the red-colored boxes) indicate the opposite. The averaged values over each initialization month (**b**) and LT (**c**) are also shown.

The calculated contribution of IC in the Gadera catchment is higher than in the Passirio for forecasts initialized in winter, spring and summer (Figure 11a), with more notable differences for winter and spring (Figure 11a,b). The higher SWC persistence in the slow-responding Gadera catchment during the low-flow period (January till March) and at the beginning of the melting period (April) renders memory of hydrological states more important than in the fast-responding Passirio catchment, resulting in up to 44% more streamflow forecast predictability originating from knowledge of IC. For forecasts initialized in summer, IC contribute by up to 21% more to the streamflow forecast skill in the Gadera

catchment as a result of elevated SWC levels. For forecasts initialized in June, however, and for the first two months of LT, the contribution of IC to the forecast skill in the Passirio is slightly higher than in the Gadera (by 1% to 8%). On average, the higher importance of IC in the Gadera catchment starts diminishing as of the fifth month of LT (Figure 11c), implying that the different memory of land surface hydrological conditions in the two case studies induced by the contrasting geology tends to become unimportant for such long leads. The (positive) difference in IC contribution between the two case studies swaps sign in autumn (Figure 11a,b), when the higher accumulation of snow in the Passirio catchment leads to the contribution of IC in the Passirio becoming more important, with up to 14% more forecast skill stemming from IC.

3.5. Classification of ESP Efficiency

Given that the common practice in seasonal hydrological forecasting involves the application of the ESP, the EPB can allow for the classification of the efficiency of the ESP over catchments of differing geology, resulting in a guide map that could be utilized by water practitioners relying on the ESP as a tool for seasonal hydrological forecasting, informing them when (in terms of season and LT) and where (in terms of regions of specific geological characteristics) it could be skillful. Figure 12 illustrates such a classification of the ESP efficiency over each initialization month and LT for the two case studies. The ESP efficiency is classified into four classes according to the contribution of IC to the streamflow forecast skill as, very good: more than 75% of the forecast skill originates from knowledge of IC; good: 50–75% of the forecast skill originates from knowledge of IC; moderate: 25–50% of the forecast skill originates from knowledge of IC; and poor: less than 25% of the forecast skill originates from knowledge of IC.

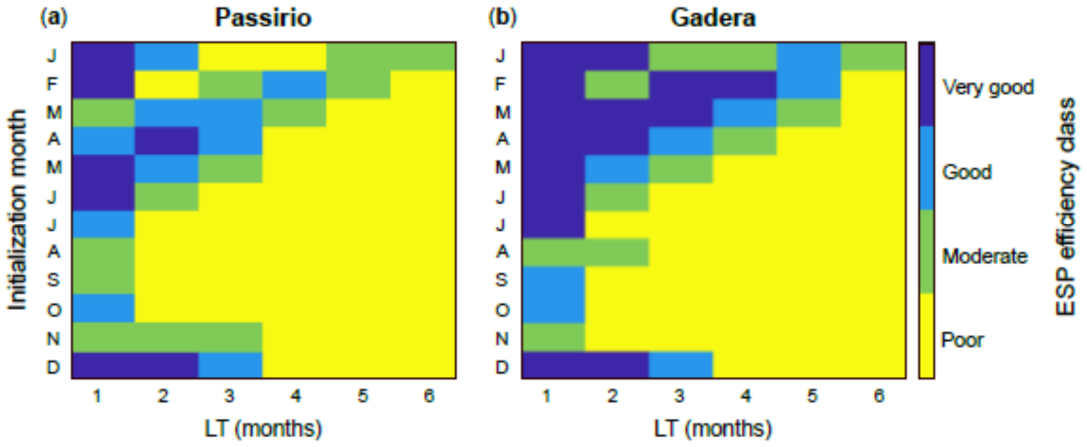

Figure 12. Classification of the ESP efficiency (**a**) for the Passirio catchment; and (**b**) for the Gadera catchment, for each initialization month (along each line) and LT (along each column). Four ESP efficiency classes can be distinguished; very good: more than 75% of the forecast skill originates from knowledge of IC; good: 50% to 75% of the forecast skill originates from knowledge of IC; moderate: 25% to 50% of the forecast skill originates from knowledge of IC; poor: less than 25% of the forecast skill originates from knowledge of IC.

Considering the ESP generally skillful when the IC contribution to the forecast skill is higher than 50%, Figure 12 indicates that, overall, the ESP is expected to be more skillful and for longer LTs over the slow-responding Gadera catchment. For forecasts initialized in winter, the skill of the ESP extends until up to five months of LT in the slow-responding Gadera catchment and until up to three months in the flashier Passirio (note that the high IC-related forecast predictability in May as the fourth month of LT is discussed in Section 4 below). Spring forecasts show predictability for up to four months of LT and three months of LT in the Gadera and the Passirio catchment, respectively,

while for forecasts initialized in summer and autumn and in both case studies, the ESP results skillful only for the first month of lead. A stricter evaluation of IC-related predictability (involving more than 75% of the forecast skill stemming from IC) reveals that the ESP would be highly skillful for forecasts initialized during the low-flow season (December–March) and the melting season (April–June) for up to four months of LT in the Gadera catchment and for up to two months of LT in the Passirio. For initialization in late summer and autumn, when precipitation controls the hydrological regime in both case studies, the contribution of CF to the forecast skill is paramount.

The above analysis illustrates the added value of the EPB sensitivity analysis framework in comparison to the traditional ESP/revESP approach as the EPB does not only show when and where the main predictability sources (IC, CF) dominate the forecast skill but also quantifies the contribution of each of them on the streamflow forecast skill. Therefore, it can effectively identify the locations (and periods) where (and when) the application of the ESP is expected to be successful, while also indicating the suitability of climate-model-based seasonal hydrological forecasting when predictability originates from skillful CF. The former possibility (i.e., the ESP) would be beneficial for regions where the contribution of IC is high, like regions where geological characteristics cause high and long persistence of moisture states. On the contrary, the latter approach (i.e., climate-model-based forecasting) is recommended for regions where the persistence of IC is low as a result of geology-induced low permeability and groundwater storage capacity. However, in cases when predictability originates from both IC and CF that the ESP/revESP framework would be able to identify whether solely perfect IC (ESP) or perfect CF (revESP) could provide skill to the forecast, the EPB allows for a more in-depth analysis of the contribution of the predictability sources, revealing when and where practices could benefit equally from both proper model initialization and skillful boundary forcing.

4. Discussion

The present study applied the EPB framework to identify how geology affects seasonal streamflow predictability related to IC and CF over two case studies of contrasting lithology, representing the end members of a set of catchments of diverse hydrological behavior [27]. Despite the fact that the selection criteria of the two case studies relied on the similarity of the catchments in terms of climatic conditions, size and mean elevation so that they would only differentiate in terms of geological characteristics, some climate-induced effects were inevitable within the present analysis. Generally, IC predictability in March in both case studies is low compared to the rest of the spring months as a result of the depleted SWC levels encountered during this month as well as because of the negative temperatures that prohibit snow from melting. In April, the increase in temperature above zero signifies the onset of the melting period that leads to an increase in SWC and thus in streamflow, transferring predictability to the IC. Moreover, the high SWC variability manifested in May in both case studies results in increased streamflow forecast skill stemming from knowledge of IC for all forecasts that include this specific month, regardless of the geology-induced hydrological behavior of the case study under consideration. Thus, the forecast initialized in February in the Passirio catchment, which would have led to skillful ESP only for the first month of LT shows skill also for the fourth month of lead that involves the month of May.

Increased SWC levels in the Passirio in June and the ongoing contribution of snowmelt to streamflow leads to a slightly higher contribution of the IC to the forecast skill compared to the Gadera, where all snow has melted by June. Better IC-related predictability for autumn forecasts in the Passirio is related to a rapid increase in snow depth, in combination with higher snow dominance over the hydrological regime in this catchment that transfer predictability to the IC. Water stored within the soil and the snowpack establishes a memory effect determined by the timescale necessary to dismiss the impact of varying wet or dry hydrological conditions like precipitation or drought, respectively [29]. In catchments with large groundwater storage that dampens the high-frequency rainfall fluctuations to a slower and smoother hydrological response, this memory effect persists longer whereas catchments with low storage have shorter memory of hydrological states, hence smaller

persistence [24]. Therefore, seasonal hydrological predictability stemming from knowledge of IC is higher for regions with longer IC persistence, as defined by their geological characteristics.

The EPB method provides a further insight into where and when forecasting investments have the greatest ability to improve streamflow predictions made to assist water managers. Although for some months or seasons the seasonal hydrological predictability source is evident like, for example, during winter when IC play a key role due to low flow, there are cases when the EPB results contradict the common expectancy. For instance, operational forecasts for May in the Passirio catchment as early as three months in advance would typically rely on IC predictability given snowmelt-dominance over the hydrological regime during spring. However, the calculated contribution of the two predictability sources to the streamflow forecast skill within the EPB analysis is similar, implying that equal improvements in IC and in CF would result in equal improvements to the forecast skill. Hence, despite the melt-based generation of streamflow in May, the EPB shows that skillful IC estimates as well as the incorporation of climate model-based streamflow predictions would equally enhance the forecast skill. The same conclusion can be drawn for the one month-lead late summer and autumn forecasts over both case studies, implying that, for certain seasons and LTs, even in highly persistent catchments with large storages CF can be equally important as IC. Therefore, for forecasts made to provide information regarding, e.g., reservoir management, the calculated contribution of IC and CF based on the EPB can clearly indicate both towards which direction attempts to improve the forecast skill should be made as well as how much improvement could be achieved. Such an improvement could be accomplished according to the current reservoir management rules and rationally attainable improvements, on the basis of their costs.

As already discussed earlier, a "perfect-model" assumption was adopted by generating surrogate observations of streamflow time series, utilized both as an EPB end point (perfect simulation) as well as a reference to estimate the performance of the EPB hindcasts. The rationale behind this choice is the exclusion of errors related to model structure or parameter estimation that would introduce a further source of uncertainty, complicating the identification of uncertainty originating from the IC and the CF. Errors related to model parameterization can be reduced by properly calibrating the hydrological model so that the simulated outputs are as close as possible to the observations [50–53]. In practice, the calibration procedure cannot completely remove the model errors as some level of uncertainty is still included in the model predictability. Nonetheless, many studies focusing on seasonal hydrological forecasting have performed "perfect-model" experiments [14,21,22,25] since the impact of the residual (i.e., after calibration) model error on the model's idealized forecast capability is hard to evaluate [15]. Future assessments should move towards the incorporation of observed data in an effort to determine the impact of model uncertainty in seasonal dynamic forecasting applications. Moreover, further work involves the extension of the current analysis to include catchments of geological characteristics that result in intermediate levels of hydrological response, to allow a thorough evaluation of the impact of geology on the contribution of the two main predictability sources to the forecast skill.

5. Conclusions

In this study, we applied the EPB methodology to analyze the effect of geology on the contribution of the seasonal hydrological predictability sources (IC, CF) to the seasonal streamflow forecast skill over two case studies of contrasting hydrological behavior. The case studies were extracted from a set of catchments of differing geology previously investigated by Norbiato et al. [27] and represent the end members of the set in terms of geology: a slow-responding catchment dominated by low-permeability rocks, hence low water storage and a fast-responding catchment characterized by high permeability and high storage capacity. Our work resulted in the following main conclusions:

1. Geological characteristics have a substantial impact on the persistence of moisture states, hence on the seasonal hydrological predictability arising from knowledge of IC. An analysis of the difference in IC contribution to the seasonal streamflow forecast skill over the two case studies of contrasting geology suggested that, generally, when compared to the contribution of IC over

the fast-responding Passirio catchment, the contribution of IC in the slow-responding Gadera catchment is higher by up to 44% for forecasts initialized in winter and spring and by up to 21% for forecasts initialized in summer. In autumn, however, the higher and rapid accumulation of snow in the Passirio catchment, in combination with the higher SWE persistence exhibited compared to respective value in the Gadera, leads to the contribution of IC in the Passirio becoming more important by up to 14%.

2. The sensitivity analysis based on the EPB indicated that knowledge of IC is more important than CF for forecasts initialized during the low-flow season (in winter) and the beginning of the melting season (in spring) for up to four months of LT in the slow-responding Gadera catchment and for up to two months of LT in the flashier Passirio catchment. During these seasons and LTs, more than 75% of the forecast skill stems from IC. For longer LTs as well as for forecasts initialized during summer and autumn (as of LT 2), the contribution of CF to the forecast skill is higher than that of IC, with generally similar contribution for the one month-lead forecasts. Our results corroborate the key role of geology in the IC-dominance over the seasonal streamflow forecast skill. Catchments with geological characteristics that induce large water storages have a slow response in precipitation events, leading to temporal dependence of streamflow for longer leads [28] during the low-flow period and the melting period. In summer and autumn, however, CF controls the forecast skill in both case studies indicating the importance of climate even in regions characterized by large subsurface storage.
3. The sensitivity analysis framework (EPB) provides an added value in comparison to the traditional ESP/revESP approach. Whereas the latter methodology can only individuate predictability stemming from perfect knowledge of either IC or CF, the EPB allows for a quantification of the contribution of each predictability source to the forecast skill. Hence, the EPB can identify regions (and periods) where (and when) predictability can be attributed to both predictability sources and clarify the predictability source for periods when the ESP/revESP approach shows unclear or ambiguous results. In this work, the EPB allowed for a classification of the efficiency of the ESP over the two cases studies representing the end members of a set of catchments of differing geology. The presented framework could be utilized as a guide map for a successful application of the ESP or of a climate-model-based hydrological forecasting tool, according to the geological characteristics of the area under consideration. For regions where the ESP results highly skillful (i.e., regions where the major part of the forecast skill stems from knowledge of IC), water managers could invest in data assimilation techniques that merge information from model simulations and observational data (either in-situ [51,54] or satellite [55–58]) of hydrological variables to reduce IC error. This could be the case for e.g., snow-dominated catchments where accurate estimates of initial SWE levels could lead to a potential improvement in forecasts initialized in spring, when snowmelt is a major component of streamflow [25,59]. On the other hand, in regions of low-skilled ESP, the focus should be on climate forecast products as precipitation and/or temperature are the main drivers of forecast predictability. In the cases, however, when the forecast skill is related to both IC and CF, attempts for forecast improvements should carefully consider the necessity for both advanced data acquisition tools and meteorological forecasts.

This study complements previous findings on the importance of IC for seasonal hydrological predictability over catchments with large subsurface storage [23,26]. Moreover, our results underpin the importance of CF for certain seasons and LTs also over catchments with large subsurface storage, where IC typically dominate. This finding is of particular importance for operational practice that traditionally relies on IC predictability, with seasonal climate forecasts being only recently incorporated to operational hydro-meteorological forecasting chains [15]. Our work provides an insight into the contribution of the dominant predictability sources over catchments of diverse geology, highlighting the importance of the EPB as a powerful tool that could advance efforts to improve seasonal forecast accuracy over specific regions, on the basis of their geological characteristics.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4441/12/8/2255/s1, Figure S1: Schematic illustration of a skill plot according to the Variational Ensemble Streamflow Prediction Assessment (VESPA) and the End Point Blending (EPB) methodologies. The x and y axes represent the initial conditions (IC) and climate forcing (CF) skill, respectively (%). The skill values are included within the dark grey rectangles, while the respective blending weights w_{IC} and w_{CF} from which the skill values are derived are shown in the light grey rectangles. The grey rhomboid shapes represent a forecast for a specific w_{IC}–w_{CF} combination point, while the red circles are the four end points of forecast uncertainty: the Ensemble Streamflow Prediction (ESP), the reverse Ensemble Streamflow Prediction (revESP), the perfect forecast and the climatology (CLIM). The magenta-colored rectangle defines the center of the response surface within which skill elasticities were calculated in this study. Figure S2: Chained time series of ensemble streamflow forecasts initialized on the first day of each month for varying levels of IC and CF skill (assuming null skill in the alternate predictability source), for lead time (LT) 3 months (a) over the Passirio catchment; and (b) over the Gadera catchment. The black line represents the surrogate streamflow observations created by the perfect simulation, while the red line represents the forecast ensemble mean. The range of the forecast streamflow values (forecast spread) is shown by the grey 5th and 95th percentile band. Note the difference in the vertical axis between the two case studies. Figure S3: Plots of the mean absolute error skill score (MAEss) for the two case studies, for hindcasts generated over the period 2002–2018 and forecast initialization on 1 February versus the skill in the two predictability sources (IC, CF). The results are presented for a LT of 1 to 6 months. Figure S4: As in Figure S3 but for forecast initialization on 1 April. Figure S5: As in Figure S3 but for forecast initialization on 1 July. Figure S6: As in Figure S3 but for forecast initialization on 1 October.

Author Contributions: Conceptualization, M.S. and M.B.; methodology, M.S. and N.D.M.; validation, M.S.; investigation, M.S. and D.A.; data curation, M.S.; writing—original draft preparation, M.S.; writing—review and editing, all authors; supervision, D.A., M.R. and M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Open Access Publishing Fund of the Free University of Bozen-Bolzano.

Acknowledgments: The authors would like to acknowledge the Hydrographic Office of the Autonomous Province of Bozen-Bolzano for providing meteorological and streamflow data, as well as Maria–Helena Ramos (IRSTEA, France) and Ilias Pechlivanidis (SMHI, Sweden) for early discussions and suggestions on the present work. M.R. and D.A. would also like to acknowledge support from the project "Applied Thermo-Fluid Dynamics Laboratories, Applied Research Infrastructures for Companies and Industry in South Tyrol" (FESR1029), financed by the European Regional Development Fund (ERDF) Investment for Growth and Jobs Program 2014–2020 and the Autonomous Province of Bolzano.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Arnal, L.; Wood, A.W.; Stephens, E.; Cloke, H.L.; Pappenberger, F. An efficient approach for estimating streamflow forecast skill elasticity. *J. Hydrometeorol.* **2017**, *18*, 1715–1729. [CrossRef]
2. Zhang, X.; Tang, Q.; Leng, G.; Liu, X.; Li, Z.; Huang, Z. On the dominant factor controlling seasonal hydrological forecast skill in China. *Water* **2017**, *9*, 902. [CrossRef]
3. Anghileri, D.; Voisin, N.; Castelletti, A.; Pianosi, F.; Nijssen, B.; Lettenmaier, D.P. Value of long-term streamflow forecasts to reservoir operations for water supply in snow-dominated river catchments. *Water Resour. Res.* **2016**, *52*, 4209–4225. [CrossRef]
4. Hamlet, A.F.; Huppert, D.; Lettenmaier, D.P. Economic value of long-lead streamflow forecasts for Columbia River hydropower. *J. Water Resour. Plan. Manag.* **2002**, *128*, 91–101. [CrossRef]
5. Viel, C.; Beaulant, A.-L.; Soubeyroux, J.-M.; Céron, J.-P. How seasonal forecast could help a decision maker: An example of climate service for water resource management. *Adv. Sci. Res.* **2016**, *13*, 51–55. [CrossRef]
6. Yossef, N.C.; Winsemius, H.; Weerts, A.; Beek, R.; Bierkens, M.F.P. Skill of a global seasonal streamflow forecasting system, relative roles of initial conditions and meteorological forcing. *Water Resour. Res.* **2013**, *49*, 4687–4699. [CrossRef]
7. Yuan, X.; Wood, E.F.; Ma, Z. A review on climate-model-based seasonal hydrologic forecasting: Physical understanding and system development. *Wiley Interdiscip. Rev. Water* **2015**, *2*, 523–536. [CrossRef]
8. Iliopoulou, T.; Aguilar, C.; Arheimer, B.; Bermúdez, M.; Bezak, N.; Ficchì, A.; Koutsoyiannis, D.; Parajka, J.; Polo, M.J.; Thirel, G.; et al. A large sample analysis of European rivers on seasonal river flow correlation and its physical drivers. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 73–91. [CrossRef]
9. Maurer, E.P.; Lettenmaier, D.P. Predictability of seasonal runoff in the Mississippi River basin. *J. Geophys. Res. Atmos.* **2003**, *108*. [CrossRef]

10. Svensson, C. Seasonal river flow forecasts for the United Kingdom using persistence and historical analogues. *Hydrol. Sci. J.* **2016**, *61*, 19–35. [CrossRef]
11. Wang, Q.J.; Robertson, D.E.; Chiew, F.H.S. A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites. *Water Resour. Res.* **2009**, *45*. [CrossRef]
12. Bazile, R.; Boucher, M.-A.; Perreault, L.; Leconte, R. Verification of ECMWF System 4 for seasonal hydrological forecasting in a northern climate. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 5747–5762. [CrossRef]
13. Cloke, H.L.; Pappenberger, F. Ensemble flood forecasting: A review. *J. Hydrol.* **2009**, *375*, 613–626. [CrossRef]
14. Paiva, R.C.D.; Collischonn, W.; Bonnet, M.P.; de Gonçalves, L.G.G. On the sources of hydrological prediction uncertainty in the Amazon. *Hydrol. Earth Syst. Sci.* **2012**, *16*, 3127–3137. [CrossRef]
15. Wood, A.W.; Hopson, T.; Newman, A.; Brekke, L.; Arnold, J.; Clark, M. Quantifying streamflow forecast skill elasticity to initial condition and climate prediction skill. *J. Hydrometeorol.* **2016**, *17*, 651–668. [CrossRef]
16. Orth, R.; Koster, R.D.; Seneviratne, S.I. Inferring soil moisture memory from streamflow observations using a simple water balance model. *J. Hydrometeorol.* **2013**, *14*, 1773–1790. [CrossRef]
17. Day, G.N. Extended streamflow forecasting using NWSRFS. *J. Water Resour. Plan. Manag.* **1985**, *111*, 157–170. [CrossRef]
18. Förster, K.; Oesterle, F.; Hanzer, F.; Schöber, J.; Huttenlau, M.; Strasser, U. A snow and ice melt seasonal prediction modelling system for Alpine reservoirs. *Proc. Int. Assoc. Hydrol. Sci.* **2016**, *374*, 143–150. [CrossRef]
19. Maurer, E.P.; Wood, A.W.; Adam, J.C.; Lettenmaier, D.P.; Nijssen, B. A long-term hydrologically based dataset of land surface fluxes and states for the Conterminous United States. *J. Clim.* **2002**, *15*, 3237–3251. [CrossRef]
20. Wanders, N.; Thober, S.; Kumar, R.; Pan, M.; Sheffield, J.; Samaniego, L.; Wood, E.F. Development and evaluation of a pan-European multimodel seasonal hydrological forecasting system. *J. Hydrometeorol.* **2019**, *20*, 99–115. [CrossRef]
21. Li, H.; Luo, L.; Wood, E.F.; Schaake, J.C. The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. *J. Geophys. Res. Atmos.* **2009**, *114*. [CrossRef]
22. Shukla, S.; Lettenmaier, D.P. Seasonal hydrologic prediction in the United States: Understanding the role of initial hydrologic conditions and seasonal climate forecast skill. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 3529–3538. [CrossRef]
23. Singla, S.; Céron, J.-P.; Martin, E.; Regimbeau, F.; Déqué, M.; Habets, F.; Vidal, J.-P. Predictability of soil moisture and river flows over France for the spring season. *Hydrol. Earth Syst. Sci.* **2012**, *16*, 201–216. [CrossRef]
24. Staudinger, M.; Seibert, J. Predictability of low flow—An assessment with simulation experiments. *J. Hydrol.* **2014**, *519*, 1383–1393. [CrossRef]
25. Wood, A.W.; Lettenmaier, D.P. An ensemble approach for attribution of hydrologic prediction uncertainty. *Geophys. Res. Lett.* **2008**, *35*. [CrossRef]
26. Harrigan, S.; Prudhomme, C.; Parry, S.; Smith, K.; Tanguy, M. Benchmarking ensemble streamflow prediction skill in the UK. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 2023–2039. [CrossRef]
27. Norbiato, D.; Borga, M.; Merz, R.; Blöschl, G.; Carton, A. Controls on event runoff coefficients in the eastern Italian Alps. *J. Hydrol.* **2009**, *375*, 312–325. [CrossRef]
28. Chiverton, A.; Hannaford, J.; Holman, I.; Corstanje, R.; Prudhomme, C.; Bloomfield, J.; Hess, T.M. Which catchment characteristics control the temporal dependence structure of daily river flows? *Hydrol. Process.* **2015**, *29*, 1353–1369. [CrossRef]
29. Ghannam, K.; Nakai, T.; Paschalis, A.; Oishi, C.A.; Kotani, A.; Igarashi, Y.; Kumagai, T.; Katul, G.G. Persistence and memory timescales in root-zone soil moisture dynamics. *Water Resour. Res.* **2016**, *52*, 1427–1445. [CrossRef]
30. Chiogna, G.; Majone, B.; Cano Paoli, K.; Diamantini, E.; Stella, E.; Mallucci, S.; Lencioni, V.; Zandonai, F.; Bellin, A. A review of hydrological and chemical stressors in the Adige catchment and its ecological status. *Sci. Total Environ.* **2016**, *540*, 429–443. [CrossRef]
31. Van de Griend, A.A.; Seyhan, E.; Engelen, G.B.; Geirnaert, W. Hydrological characteristics of an Alpine glacial valley in the North Italian Dolomites. *J. Hydrol.* **1986**, *88*, 275–299. [CrossRef]
32. Mallucci, S.; Majone, B.; Bellin, A. Detection and attribution of hydrological changes in a large Alpine river basin. *J. Hydrol.* **2019**, *575*, 1214–1229. [CrossRef]
33. Koutsoyiannis, D. Hydrologic persistence and the Hurst phenomenon. *Water Encycl.* **2005**, 210–221. [CrossRef]

34. Borga, M. Accuracy of radar rainfall estimates for streamflow simulation. *J. Hydrol.* **2002**, *267*, 26–39. [CrossRef]
35. Norbiato, D.; Borga, M.; Degli Esposti, S.; Gaume, E.; Anquetin, S. Flash flood warning based on rainfall thresholds and soil moisture conditions: An assessment for gauged and ungauged basins. *J. Hydrol.* **2008**, *362*, 274–290. [CrossRef]
36. Zaramella, M.; Borga, M.; Zoccatelli, D.; Carturan, L. TOPMELT 1.0: A topography-based distribution function approach to snowmelt simulation for hydrological modelling at basin scale. *Geosci. Model Dev.* **2019**, *12*, 5251–5265. [CrossRef]
37. Cazorzi, F.; Dalla Fontana, G. Snowmelt modelling by combining air temperature and a distributed radiation index. *J. Hydrol.* **1996**, *181*, 169–187. [CrossRef]
38. Hargreaves, G.H.; Samani, Z.A. Estimation of potential evapotranspiration. *J. Irrig. Drain. Div. Proc. Am. Soc. Civ. Eng.* **1982**, *108*, 223–230.
39. Moore, R.J. The probability-distributed principle and runoff production at point and basin scales. *Hydrol. Sci. J.* **1985**, *30*, 273–297. [CrossRef]
40. Cunge, J.A. On the subject of a flood propagation computation method (Musklngum Method). *J. Hydraul. Res.* **1969**, *7*, 205–230. [CrossRef]
41. Moore, R.J. The PDM rainfall-runoff model. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 483–499. [CrossRef]
42. Eberhart, R.; Kennedy, J. A new optimizer using particle swarm theory. In Proceedings of the 6th International Symposium on Micro Machine and Human Science, Nagoya, Japan, 4–6 October 1995; IEEE Service Center: Piscataway, NJ, USA, 1995; pp. 39–43. [CrossRef]
43. 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]
44. Klemes, V. Operational testing of hydrological simulation models. *Hydrol. Sci. J.* **1986**, *31*, 13–24. [CrossRef]
45. Kelly, L.; Kalin, M.R.; Bertram, D.; Kanjaye, M.; Nkhata, M.; Sibande, H. Quantification of temporal variations in Base Flow Index using sporadic river data: Application to the Bua catchment, Malawi. *Water* **2019**, *11*, 901. [CrossRef]
46. Alfieri, L.; Pappenberger, F.; Wetterhall, F.; Haiden, T.; Richardson, D.; Salamon, P. Evaluation of ensemble streamflow predictions in Europe. *J. Hydrol.* **2014**, *517*, 913–922. [CrossRef]
47. Monhart, S.; Zappa, M.; Spirig, C.; Schär, C.; Bogner, K. Subseasonal hydrometeorological ensemble predictions in small- and medium-sized mountainous catchments: Benefits of the NWP approach. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 493–513. [CrossRef]
48. Pappenberger, F.; Ramos, M.H.; Cloke, H.L.; Wetterhall, F.; Alfieri, L.; Bogner, K.; Mueller, A.; Salamon, P. How do I know if my forecasts are better? Using benchmarks in hydrological ensemble prediction. *J. Hydrol.* **2015**, *522*, 697–713. [CrossRef]
49. Arnal, L.; Cloke, H.L.; Stephens, E.; Wetterhall, F.; Prudhomme, C.; Neumann, J.; Krzeminski, B.; Pappenberger, F. Skilful seasonal forecasts of streamflow over Europe? *Hydrol. Earth Syst. Sci.* **2018**, *22*, 2057–2072. [CrossRef]
50. Moradkhani, H.; Sorooshian, S. General review of rainfall-runoff modeling: Model calibration, data assimilation, and uncertainty analysis. In *Hydrological Modelling and the Water Cycle: Coupling the Atmospheric and Hydrological Models*; Sorooshian, S., Hsu, K.-L., Coppola, E., Tomassetti, B., Verdecchia, M., Visconti, G., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 1–24. ISBN 978-3-540-77843-1. [CrossRef]
51. Seo, D.-J.; Koren, V.; Cajina, N. Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting. *J. Hydrometeorol.* **2003**, *4*, 627–641. [CrossRef]
52. Shi, X.; Wood, A.W.; Lettenmaier, D.P. How essential is hydrologic model calibration to seasonal streamflow forecasting? *J. Hydrometeorol.* **2008**, *9*, 1350–1363. [CrossRef]
53. Wu, Q.; Liu, S.; Cai, Y.; Li, X.; Jiang, Y. Improvement of hydrological model calibration by selecting multiple parameter ranges. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 393–407. [CrossRef]
54. Clark, M.P.; Rupp, D.E.; Woods, R.A.; Zheng, X.; Ibbitt, R.P.; Slater, A.G.; Schmidt, J.; Uddstrom, M.J. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. *Adv. Water Resour.* **2008**, *31*, 1309–1324. [CrossRef]
55. Andreadis, K.M.; Lettenmaier, D.P. Assimilating remotely sensed snow observations into a macroscale hydrology model. *Adv. Water Resour.* **2006**, *29*, 872–886. [CrossRef]

56. Brocca, L.; Melone, F.; Moramarco, T.; Wagner, W.; Naeimi, V.; Bartalis, Z.; Hasenauer, S. Improving runoff prediction through the assimilation of the ASCAT soil moisture product. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 1881–1893. [CrossRef]
57. Liu, Y.Y.; Dorigo, W.A.; Parinussa, R.M.; de Jeu, R.A.M.; Wagner, W.; McCabe, M.F.; Evans, J.P.; van Dijk, A.I.J.M. Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sens. Environ.* **2012**, *123*, 280–297. [CrossRef]
58. Rodell, M.; Houser, P.R. Updating a land surface model with MODIS-derived snow cover. *J. Hydrometeorol.* **2004**, *5*, 1064–1075. [CrossRef]
59. Mahanama, S.; Livneh, B.; Koster, R.; Lettenmaier, D.; Reichle, R. Soil moisture, snow, and seasonal streamflow forecasts in the United States. *J. Hydrometeorol.* **2012**, *13*, 189–203. [CrossRef]