

Article

# Ensembles for Viticulture Climate Classifications of the Willamette Valley Wine Region

Brian Skahill <sup>1,\*</sup>, Bryan Berenguer <sup>1</sup> and Manfred Stoll <sup>2</sup>

<sup>1</sup> Northwest Wine Studies Center, Chemeketa Community College, Salem, OR 97304, USA; bryan.berenguer@chemeketa.edu

<sup>2</sup> Department of General and Organic Viticulture, Hochschule Geisenheim University, Von-Lade-Strasse 1, 65366 Geisenheim, Germany; Manfred.Stoll@hs-gm.de

\* Correspondence: bskahill@my.chemeketa.edu

**Abstract:** Future climate projections provide an opportunity to evaluate cultivar climate classification and preferred styles of wine production for a wine grape growing region. However, ensemble selection must account for downscaled archive model skills and interdependence rather than be arbitrary and subjective. Relatedly, methods for generalizing climate model choice remain uncertain, particularly for identifying optimal ensemble subsets. In this study we consider the complete archive of the thirty-two Coupled Model Intercomparison Project Phase 5 (CMIP5) daily Localized Constructed Analogs (LOCA) downscaled historic datasets and their observational data that were used for downscaling and bias corrections. We apply four model averaging methods to determine optimal ensembles for the computation of six common climate classification indices for the Willamette Valley (WV) American Viticultural Area (AVA). Among the four methods evaluated, elastic-net regularization consistently performed best with identifying optimal ensemble subsets. Variation exists among the optimal ensembles computed for each of the six bioclimatic indices. However, a subset of approximately seven to ten climate models were consistently excluded across all six indices' ensembles. While specific to the archive and wine region, optimal ensemble sizes were noticeably larger than ensemble sizes commonly employed in published studies. Results are reported such that they can be used by researchers to independently perform analyses involving any one of the six bioclimatic indices throughout the WV AVA while using historic and future LOCA CMIP5 climate projections. The data and methods employed herein are applicable for other wine regions.

**Keywords:** multi-model averaging; LOCA downscaled; CMIP5; climate projections; bioclimatic indices; Willamette Valley; American Viticultural Area



**Citation:** Skahill, B.; Berenguer, B.; Stoll, M. Ensembles for Viticulture Climate Classifications of the Willamette Valley Wine Region. *Climate* **2021**, *9*, 140. <https://doi.org/10.3390/cli9090140>

Academic Editor: Mário Gonzalez Pereira

Received: 26 July 2021

Accepted: 11 September 2021

Published: 15 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Climate is significant in determining cultivar suitability and the resulting wine profile typical for a wine grape growing region [1–4]. Temperature and water availability are recognized as two of the most important environmental factors influencing grapevine growth and the subsequent quality of berries and wine [5–9]. Various indices exist to classify viticulture climate [10], and several recent studies have evaluated the impacts of climate change to viticulture using bioclimatic indices and downscaled future climate projections for wine regions (see Table 1).

Quénol et al. [11] highlighted the importance of accounting for model uncertainty when performing studies that involve the application of future climate projections. Model simulations of the future from the Coupled Model Intercomparison Project Phase 5 (CMIP5) [12] archive are no longer necessarily considered as equally likely independent realizations of the future climate [13–18]. The models constituting the CMIP5 archive are now recognized as interdependent by virtue of shared codes and the use of similar parameterization schemes and calibration methods [17,19,20]. The assumption of archive

model independence results in a multi-model mean that is biased by the duplicative information that is contributed by the similar models. This can confound assessments of model agreement about changes in future climate and alter the statistics of identified correlations. Ignoring model dependence can result in poor accuracy and poor uncertainty quantification. The number of effective models in the archive is conjectured to be less than the number of simulations [17,21–24] and there is a need not only to better understand weighting strategies that constrain projection uncertainty [15–18], but also to efficiently and effectively generalize model choice in a manner that selects optimal ensemble subsets for specific applications [24,25].

A common approach to uncertainty quantification of future climate projections is to equally weight a multi-model ensemble, possibly formed by the subjective selection of a subset of the available models, wherein weights are assigned without any consideration of model skill [24]. While a given multi-model mean is generally considered a better choice than any single model [24,25], it is also potentially not optimal relative to alternative multi-model ensemble compositions and weights assignment strategies [13–18]. It has been noted that the equal weighted multi-model mean does not necessarily quantify projection uncertainty but rather is simply an uninformed ad hoc measure of ensemble spread [26].

It is now more commonplace to address the epistemic uncertainty associated with model choice by applying methods that account for model skill and interdependence wherein skill is measured by a comparison with observed data [15–17]. Weight assignments are further complicated by the learned understanding that individual model skill is dependent upon the modeling objective and study location [17]. For viticulture climate classification of a given area using future climate projections, this implies that weight assignments could vary by the bioclimatic index under consideration.

Sanderson et al. [17] introduced two weight assignment strategies to constrain projection uncertainty, one that accounted for model skill and another that combined the computed skill with an evaluation of model independence. These two approaches have been applied in recent studies [14,15,18]. A few recent studies have also applied Bayesian model averaging (BMA) to assign ensemble weights and constrain projection uncertainty [15,16,18]. BMA is attractive from the perspective that it robustly quantifies model uncertainty and its solution implicitly accounts for model interdependence. However, it is unattractive in that its application is highly compute intensive, and can also be technically challenging to diagnose, relative to most other model calibration strategies [27].

Herger et al. [24] distinguished between approaches that assign continuous weights to the entire ensemble archive with those that are focused on the identification of an ensemble subset. They applied mixed integer quadratic programming to efficiently and flexibly find optimal subsets while accounting for model interdependence. Yang et al. [25] used the Taylor diagram [28] to rank models and subsequently combined the higher performing models using simple averaging to find a nine member multi-model ensemble subset that best characterized the spatiotemporal variability of temperature and precipitation over China. In this study we apply a recent advance for regularizing linear models that is directed at automatic variable selection [29–31] to find prediction specific optimal ensemble subsets while simultaneously accounting for model skill and interdependence. Regularization of the ensemble linear model is needed given the thirty-two member multi-model archive we consider in this study and acknowledgment that there are  $2^n - 1$  possible models that involve subsets of  $n$  predictors [32]. If  $n = 32$ , then there are 4,294,967,295 possible models.

Specification of the selection criteria for the ensemble compositions or the ensemble weights is not common in published studies that report on viticultural climate impact assessments using downscaled CMIP5 ensemble subsets. The arbitrary selection and application of a single model or ensemble subset of downscaled CMIP5 models to quantify future climate impacts is most likely incomplete and suboptimal [24,25]. In this article we apply a directed equal weights strategy similar in design to that performed by Yang et al. [25], the two skill weighting strategies introduced by Sanderson et al. [17], and elastic-net regularization [33] to determine for each method optimal ensemble compositions and weights

relevant for the computation of the growing season average temperature (GST) [34], growing degree days (GDD) [1,3], Huglin's index (HI) [35], cool night index (CI) [36], dryness index (DI) [37], and the Géoviticulture multicriteria climate classification (GMCC) [38] viticultural climate indices for a study region containing the Willamette Valley (WV) American Viticultural Area (AVA).

Heat unit accumulation is important for suitability assessment in that each cultivar has a minimum summed temperature requirement to reach maturity. A late-ripening cultivar will not achieve full ripening in a cool climate and an early ripening cultivar will ripen too quickly in a warm climate. Optimal climate suitability occurs where cultivars ripen at the end of the growing season [8]. The most common applied heat unit accumulation index is the Winkler index (WI), more often referred to as growing degree days (GDD) [1,3,10]. It is the sum of the mean temperatures above a specified base for the growing season, defined as 1 April–31 October in the northern hemisphere [1]. The Huglin index (HI) [35] is similar in form to the GDD method but gives extra weight to the maximum temperature, includes an adjustment factor to account for day length as a function of latitude, and as originally defined only sums for the months April–September. Huglin and Schneider [39] classified cultivars as a function of HI. The GST climate index is the mean of the observed maximum and minimum daily surface air temperature values from the first of April through the end of October (for the Northern Hemisphere). Jones [34] demonstrated broad correlation between the GST and cultivar ripening potential across many wine regions.

While temperature-based heat unit accumulation indices are most used to address viticultural suitability [3], Tonietto and Carbonneau [38] introduced the Géoviticulture multicriteria climate classification system (GMCC) to describe and compare the climates of vineyards worldwide. The GMCC combines three complementary viticulture indices that not only capture climatology but also relate well to qualitative potential and characteristics relevant to wine production. The first index, HI, accounts for heat unit accumulation. The second index is the cool night index (CI) which is defined as the mean of the minimum night temperatures for the month before harvest. The third GMCC index is the dryness index (DI). It is the aggregate vertical accounting of potential soil water availability for April to September whose computation considers precipitation, soil evaporation, and vine transpiration while excluding surface runoff and drainage. It functions as an indicator of the level of presence-absence of dryness for a given area. The GMCC is applicable at multiple spatial scales and includes a comprehensive companion database to support worldwide climatic group comparisons, and one potential use of it is the study of the evolution of viticultural climate as a function of climate change [38].

We consider the complete archive of thirty-two CMIP5 daily Localized Constructed Analogs (LOCA) downscaled historic datasets [40,41] of minimum surface air temperature, maximum surface air temperature, and precipitation ([https://gdo-dcp.ucllnl.org/downscaled\\_cmip\\_projections/](https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/) accessed on 2 December 2020) and their observational data that were used for downscaling and bias correction [42]. The method comparison provides an opportunity to learn about their relative performance. It also reveals optimal ensemble subset sizes for the computation of the GST, GDD, HI, CI, DI, and GMCC indices in the WV AVA. Moreover, by considering multiple indices, we can explore whether LOCA CMIP5 ensemble compositions and weights vary by viticulture climate classification index.

The Willamette Valley is of interest not only given its recent and consistent observed warm to hot vintages [43], but also due to the very limited previous study of climate change impacts to viticulture in the AVA [44]. To our knowledge, this is the first study to comprehensively evaluate a complete downscaled CMIP5 model archive to determine optimal ensemble compositions and weights for the computation of future projections of viticulture climate classification for a wine region. Acknowledging that this type of study is technically challenging and infrequent [24], our results are provided in a manner such that they can be directly used by others to reliably analyze the impacts of daily LOCA CMIP5 future climate projections to viticulture for the WV AVA.

**Table 1.** A summary of recent studies that have evaluated the impacts of climate change to viticulture using bioclimatic indices and downscaled future climate projections.

| Authors                 | Wine Region/Study Area                                                                                                        | Bioclimatic Indices <sup>1</sup>                                                                               |
|-------------------------|-------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------|
| Blanco-Ward et al. [45] | The Portuguese Douro Demarcated Region                                                                                        | WI, HI, CI, DI                                                                                                 |
| Cabré and Nuñez [46]    | Argentinean provinces of Mendoza, San Juan, La Rioja, Salta and Catamarca                                                     | GST, CI, GSP, DSTmin                                                                                           |
| Cardell et al. [47]     | Europe                                                                                                                        | Tmax, AP, GST, WI, HI, RET, WB                                                                                 |
| Irimia et al. [48]      | Cotnari (Romania)                                                                                                             | GST, HI                                                                                                        |
| Koufos et al. [49]      | Aigialia, Chalkidiki, Crete, Drama, Kavala, Limnos, Marania, Naousa, Nemea, Pyrgos, Rodos, Samos, Santorini, Tripoli (Greece) | GDD                                                                                                            |
| Santos et al. [50]      | 50 protected denominations of origin and sub-regions throughout mainland Portugal                                             | HI, DI                                                                                                         |
| Sirnik [51]             | Valencia (Spain) and Goriška Brda (Slovenia)                                                                                  | WI, HI, DI                                                                                                     |
| Teslic et al. [52]      | Emilia-Romagna                                                                                                                | Tmean, HI, CI, GS <sub>l</sub> , GS <sub>ed</sub> , GS <sub>st</sub> , P <sub>tot</sub> , DI, DSI, FFP, FF, LF |
| Trbic et al. [53]       | 3 locations in Bosnia and Herzegovina                                                                                         | HI, CI, DI, GMCC                                                                                               |

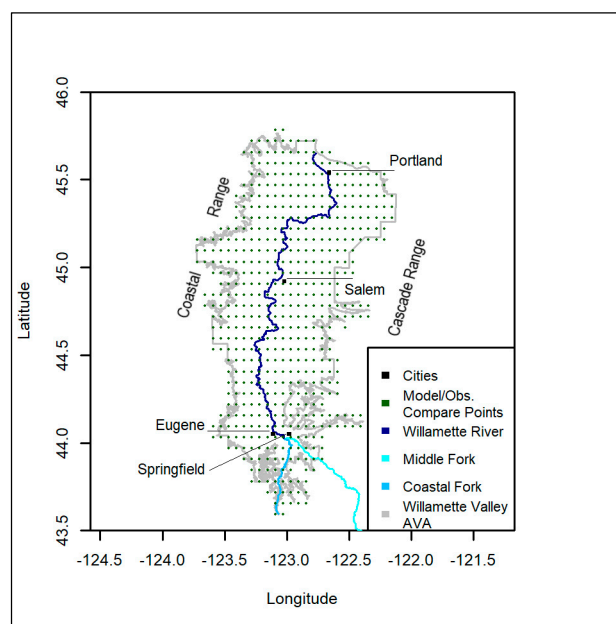
<sup>1</sup> WI = Winkler index, HI = Huglin index, CI = cool night index, DI = dryness index, GST = growing season average temperature, GSP = growing season average precipitation, DSTmin = monthly mean minimum temperature during the dormant season, Tmax = maximum temperature during summer, AP = annual precipitation, RET = real evapotranspiration, WB = water balance, GDD = growing degree days, Tmean = growing season mean temperature, GS<sub>l</sub> = growing season length, GS<sub>ed</sub> = climatologically defined end of growing season, GS<sub>st</sub> = climatologically defined start of growing season, P<sub>tot</sub> = growing season total rainfall, DSI = dry spell index, FFP = length of the frost-free period, FF = date of the first fall frost, LF = date of the last spring frost, and GMCC = the Géoviticulture multicriteria climate classification index.

## 2. Materials and Methods

### 2.1. Study Region

The study region consists of 553 model-observation comparison points which cover the 13,913 square kilometer Willamette Valley AVA which lies in the Willamette River Basin in the northwestern part of the State of Oregon (OR) in the U.S. (Figure 1). The WV AVA includes the entire main stem of the Willamette River, including parts of its Middle Fork and Coastal Fork which join near Springfield, OR just south of Eugene, OR. The Willamette River generally flows in a northly pattern to its confluence with the Columbia River in Portland, OR. The WV AVA is bounded to its west by the Oregon Coastal Range and to its east by the Cascade Range.

The Oregon Wine Board [54] reported for the 2019 vintage 25,197 harvested acres of the WV AVA's 25,452 planted acres. Total production for the WV AVA was estimated to be 73,560 tons, which accounted for approximately 69.7% of the total wine grape production for the State of Oregon. For the WV AVA, Pinot Noir, Pinot Gris, Chardonnay, Pinot Blanc, and Riesling accounted for 69.3, 16.9, 7.4, 1.2, and 1.2 percent of its total production, respectively.



**Figure 1.** The 553 model-observation data locations which cover the Willamette Valley American Viticultural Area (AVA) used to compute prediction specific model-to-measurement misfit evaluations and ensemble composition and weights assignments by comparing bioclimatic index relevant processed daily Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) model historic datasets [40,41] with their equivalently processed observed data counterparts [18,42].

## 2.2. Time Series Data

The complete archive of Coupled Model Intercomparison Project phase 5 (CMIP5) daily Localized Constructed Analogs (LOCA) multi-model historic (1950–2005) datasets [40,41] of minimum surface air temperature, maximum surface air temperature, and precipitation was collected from the Downscaled CMIP5 Climate and Hydrology Projections archive at [https://gdo-dcp.ucllnl.org/downscaled\\_cmip\\_projections/](https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/) (accessed on 2 December 2020). The 32 models and modeling groups that provided the data used in this study are listed in Table S1. The observation-based  $1/16^\circ$  longitude-latitude gridded Livneh data product [41] was used to train and downscale the daily LOCA CMIP5 datasets which span  $25.125^\circ$  N to  $52.875^\circ$  N and  $-124.625^\circ$  E to  $-67.000^\circ$  E [40,55]. In this study, version 1.2 of the Livneh dataset at the 553 grid locations indicated in Figure 1 functioned as the observation dataset. It was processed and compared with its corresponding LOCA CMIP5 processed model historic data counterparts [18] to compute prediction specific (i.e., growing season average temperature (GST), growing degree days (GDD), Huglin index (HI), cool night index (CI), dryness index (DI), and Géoviticulture multicriteria climate classification (GMCC)) model-to-measurement evaluations and resulting ensemble composition and weights.

## 2.3. Viticulture Climate Indices and Processed Datasets for Prediction Specific Modeling Analyses

Calculation of any one of the six viticulture climate classification indices involves one or more variables wherein each individual variable is a unique processed base model/ observation data product (i.e., precipitation, maximum surface air temperature, minimum surface air temperature). The model-observation datasets developed to support the prediction specific modeling analyses were not the specific viticultural climate indices but rather processed datasets that support their calculation as indicated in Equations (1)–(4), (5) and (7). The



growing season average temperature, GST [34], growing degree days, GDD [1,3], Huglin index, HI [35], and cool night index, CI [36], are defined in Equations (1)–(4), respectively.

$$GST = \frac{1}{n} \sum_{Apr\ 1}^{Oct\ 31} (T_{max} + T_{min})/2 = f\left(\sum_{Apr\ 1}^{Oct\ 31} T_{max}, \sum_{Apr\ 1}^{Oct\ 31} T_{min}\right), \quad (1)$$

$$GDD = \sum_{Apr\ 1}^{Oct\ 31} \max\left[\left(\frac{\overline{T_{max}} + \overline{T_{min}}}{2}\right) - 10, 0\right] = f(\overline{T_{max}}, \overline{T_{min}}), \quad (2)$$

$$HI = \sum_{Apr\ 1}^{Sep\ 30} \max\left[\left(\frac{(\overline{T_{mean}} - 10) + (\overline{T_{max}} - 10)}{2}\right), 0\right] \cdot K = f(\overline{T_{max}}, \overline{T_{min}}), \quad (3)$$

$$CI = \frac{1}{30} \sum_{Sep\ 1}^{Sep\ 30} T_{min} = f(T_{min}), \quad (4)$$

where  $n = 214$ ,  $T_{max}$  and  $T_{min}$ ,  $\overline{T_{max}}$ ,  $\overline{T_{min}}$ ,  $\overline{T_{mean}}$ , and  $K$  denote the number of days for the northern hemisphere growing season, the maximum and minimum daily surface air temperature data values in °C, the monthly means of the  $T_{max}$ , the monthly means of the  $T_{min}$ , the monthly means of the mean of the  $T_{max}$  and  $T_{min}$ , and an adjustment factor to account for day length as a function of latitude, respectively. In Equation (1) we see that the GST is a function of the sums of the  $T_{max}$  and  $T_{min}$  for the duration of the growing season. Equations (2) and (3) emphasize that the GDD and HI are each a function of the  $\overline{T_{max}}$  and  $\overline{T_{min}}$  for their respective periods. The adjustment factor,  $K$ , for the calculation of the HI was not included in the right-hand side of Equation (3) since it does not vary by climate model. The CI is a function of the  $T_{min}$  for the month of September (see Equation (4)). The dryness index [37] is a function of the April to September monthly means of the daily precipitation and maximum and minimum surface air temperature data,

$$DI = f(\overline{P}, \overline{T_{max}}, \overline{T_{min}}). \quad (5)$$

It is based on the following equation calculated monthly from April to September,

$$W = W_0 + P - T_v - E_s, \quad (6)$$

where  $W$ ,  $W_0$ ,  $P$ ,  $T_v$ , and  $E_s$  represent the soil water reserve in millimeters (mm) at the end of a given month, the initial soil water reserve in mm, monthly precipitation in mm, monthly potential transpiration in mm, and monthly direct evaporation from the soil in mm [38]. The value for DI each year is the value of  $W$  at the end of September (See [38] for additional details about calculation of the DI). The Géoviticulture multicriteria climate classification (GMCC) is defined as the cross product of the HI, CI, and DI [38]; hence,

$$GMCC = f(\overline{P}, \overline{T_{max}}, \overline{T_{min}}, CI). \quad (7)$$

Relative individual LOCA CMIP5 model performance may vary across each of a given prediction's respective functional variables [15,17,18,23]. For example, for computation of the GST, an individual LOCA CMIP5 model might possess better skill relative to the remaining members of the archive in summing daily maximum surface air temperatures from April through October but poor skill relative to the remaining members of the multi-model ensemble with summing the daily minimum surface air temperatures from April through October. Recently introduced methods for evaluating climate model skill require the consideration of each individual variable composing the prediction [15,17,18,23]. Table 2 summarizes the model/observation datasets that were developed to evaluate the performance of LOCA CMIP5 individual and multi-model ensembles for the calculation of the six viticulture climate classification indices throughout the WV AVA. Each processed model-observation dataset specified in Table 2 used data from the 553 model-observation grid locations indicated in Figure 1 for the entire 56 years (1950–2005) that define the LOCA CMIP5 historic period.

**Table 2.** Summary description of model-observation datasets that were developed for the Willamette Valley American Viticultural Area to evaluate the performance of Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) [40,41] ensembles to calculate six viticulture climate classification indices during the LOCA CMIP5 defined historic period (1950–2005).

| Prediction        | Dataset Description                                                | Model/Observation Dataset Sizes                |
|-------------------|--------------------------------------------------------------------|------------------------------------------------|
| GST               | $\sum_{Apr\ 1}^{Oct\ 31} T_{max}, \sum_{Apr\ 1}^{Oct\ 31} T_{min}$ | $553 \cdot 56 \cdot 2 = 61,936$                |
| GDD <sup>a</sup>  | $\overline{T_{max}}, \overline{T_{min}}$                           | $553 \cdot 56 \cdot 7 \cdot 2 = 433,552$       |
| HI <sup>b</sup>   | $\overline{T_{max}}, \overline{T_{min}}$                           | $553 \cdot 56 \cdot 2 = 371,616$               |
| CI                | $\frac{1}{30} \sum_{Sep\ 1}^{Sep\ 30} T_{min}$                     | $553 \cdot 56 \cdot 1 = 30,968$                |
| DI <sup>b</sup>   | $\overline{P}, \overline{T_{max}}, \overline{T_{min}}$             | $553 \cdot 56 \cdot 3 = 557,424$               |
| GMCC <sup>b</sup> | $\overline{P}, \overline{T_{max}}, \overline{T_{min}}, CI$         | $553 \cdot 56 \cdot (6 \cdot 3 + 1) = 588,392$ |

$T_{max}$  and  $T_{min}$ ,  $\overline{T_{max}}$ ,  $\overline{T_{min}}$ , and  $\overline{P}$  denote the maximum and minimum daily surface air temperature, monthly means of the  $T_{max}$ , monthly means of the  $T_{min}$ , and the mean monthly precipitation. <sup>a</sup> April to October. <sup>b</sup> April to September.

#### 2.4. Modeling Averaging Methods

Model averaging involves combining models to improve prediction accuracy and reduce forecast uncertainty, preferably while simultaneously accounting for model skill and interdependence [15,16,25,56]. Eyring et al. [13] expressed the need for further study and better understanding of advanced model averaging methods for climate projections.

##### 2.4.1. Model

The specified model is a general linear model without intercept,

$$M = \sum_{i=1}^{32} w_i M_i, \quad (8)$$

where  $M_i$  and  $w_i$  represent the  $i$ -th LOCA CMIP5 model and its assigned non-negative weight, respectively. The modeling objective is to minimize model-to-measurement misfit for each WV AVA viticulture climate classification index directed dataset defined in Table 2 by using one of four different methodologies as they are described in Sections 2.4.2–2.4.5.

##### 2.4.2. Skill Weighting

Sanderson et al. [23] and Sanderson et al. [17] introduced a weight assignment strategy for developing climate multi-model ensembles that addressed the assumption regarding CMIP5 model democracy; namely, that individual members of the CMIP5 multi-model archive can vary in their skill by location and prediction. Computation of the root mean squared error (RMSE) between the observations and each model, by variable, enables the ranking of the individual ensemble members. The RMSE is given by

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^N (S_j - O_j)^2}, \quad (9)$$

where  $S$ ,  $O$ , and  $N$  denote modeled data, their observed counterparts, and the number of model-observation comparisons. The skill weight assigned to the  $i$ -th LOCA CMIP5 model [17,23] is defined as

$$w_{skill,i} = A e^{-\left(\frac{\delta_{i(obs)}}{D_q}\right)^2}, \quad (10)$$

where  $A$ ,  $\delta_{i(obs)}$ , and  $D_q$  denote a normalizing constant such that the entire set of individual model skill weights sums to one, the median skill for each model [15], and a parameter called the radius of model equality [23] which determines the degree to which models

with poor skill are down-weighted. A model's median skill is defined as the median of its individual relative *RMSE* values,  $E^*$  [15,17,23], where

$$E^* = 1 + \frac{E - \bar{E}}{\bar{E}}, \quad (11)$$

and  $E$  and  $\bar{E}$  denote, for a given prediction specific functional variable, the model's *RMSE* and the mean *RMSE* across all ensemble members, respectively. A small value assigned to the radius of model equality will result in a weights assignment that favors an ensemble subset whose members possess the best median skill values; whereas, as  $D_q \rightarrow \infty$ , computed weights will approach the multi-model mean which does not account for model skill.

#### 2.4.3. Skill and Model Interdependence Weighting

Sanderson et al. [23] and Sanderson et al. [17] augmented their skill weight assignment strategy to also account for model similarity in attempts to develop ensemble compositions and weights that further reduce CMIP5 archive redundancy by virtue of shared parameterization schemes and training methods. As with their skill weight assignment strategy, *RMSE* values calculated between each model and its respective variable specific processed observations enables the calculation of independence weights for each LOCA CMIP5 ensemble member. A multi-variate inter-model distance matrix,  $\delta$ , is computed as the equally weighted linear combination of normalized variable specific inter-model distance matrices that are each composed of their *RMSE* distances between each pair of models. For any two models  $i$  and  $j$ , a similarity score matrix is defined as

$$S_{ij} = e^{-\left(\frac{\delta_{ij}}{D_u}\right)^2}, \quad (12)$$

where  $D_u$  is a parameter defined as the radius of similarity which specifies the distance scale over which models should be considered similar and down-weighted for co-dependence. We choose a value of 0.5 for  $D_u$  [15,23]. In theory, for two identical models,  $S_{ij} = 1$ . Also,  $S_{ij} \rightarrow 0$  as  $\delta_{ij} \rightarrow \infty$ . Using the similarity scores, the independence weight for the  $i$ -th LOCA CMIP5 model is defined as

$$w_u(i) = \left(1 + \sum_{i \neq j}^n S_{ij}\right)^{-1}, \quad (13)$$

where  $n$  is the total number of models [15,17,23]. A weight to assign to each individual model that simultaneously accounts for its skill and interdependence is the product of its skill weight and independence weight

$$w(i) = A \cdot w_{skill,i} \cdot w_u(i), \quad (14)$$

normalized by  $A$  such that the sum of the  $w(i)$  across all models equals one [15,17,23].

#### 2.4.4. Model Median Skill Directed Simple Model Averaging

For each prediction, the median skills computed for each model,  $\delta_{i(obs)}$ , were ranked and ensembles of size one to size equal to thirty-two were constructed using simple model averaging. This is like the approach presented by Yang et al. [25] who used skill scores obtained from application of the Taylor diagram [28] to build multi-model ensembles using simple model averaging.

#### 2.4.5. Elastic-Net Regularization

In their climate model subset study, Herger et al. [24] mentioned lasso regression from the field of machine learning as an alternative method to their approach. Zou and Hastie [33] introduced the elastic-net penalty as a compromise between ridge [57,58] and



lasso [59] regression. Given observations  $y_i$ ,  $i = 1, \dots, n$ , an  $n \times m$  matrix of covariates  $X$ , and an assumed linear model

$$y_i = \eta_0 + \eta_1 x_{i,1} + \dots + \eta_m x_{i,m}, \quad (15)$$

the elastic-net minimizes

$$\frac{1}{2n} \sum_{i=1}^n \tilde{w}_i (y_i - \eta_0 - \eta x_i^T)^2 + \lambda \sum_{j=1}^m \left[ \frac{1}{2} (1 - \alpha) \eta_j^2 + \alpha |\eta_j| \right], \quad (16)$$

where  $\lambda$  is a non-negative regularization parameter that is tuned to weight the overall strength of the penalty,  $\alpha \in [0, 1]$  is specified to control the penalty term to vary from ridge regression at  $\alpha = 0$  to lasso regression at  $\alpha = 1$ , and  $\tilde{w}_i$  is the weight assigned to the  $i$ th observation [29]. Ridge regression yields smooth solutions that include all the predictors, whereas application of lasso regression results in automatic variable selection (i.e., sparse, much more easily interpretable solutions) [32]. The elastic-net combines the two methods. As  $\alpha$  increases from 0 to 1 for a fixed  $\lambda$ , the number of zero-valued  $\eta_j$  increases from 0 to the sparsity of the lasso [29]. In this study, variable selection was desired, therefore  $\alpha$  was specified close to 1 for numerical stability [29]. With  $\alpha$  specified close to 1, the elastic-net performs much like lasso regression while retaining ridge regression's capacity to collectively shrink the coefficients for any highly correlated covariables [29,60]. Cross validation (CV) was employed to ensure that the minimizing value for  $\lambda$  was properly located for each elastic-net application. Google Scholar reports 14,411 citations for [33]. Implementations of the elastic net exist in R, Matlab, Apache Spark, and SAS.

### 3. Results and Discussion

For each viticulture climate classification directed prediction (i.e., GST, GDD, HI, CI, DI, GMCC), individual LOCA CMIP5 model RMSEs computed by variable are listed in Table 3. Together with the specification of  $D_q$  and  $D_u$ , these RMSE values can be used to independently calculate prediction specific ensemble skill weights and independence weights.

**Table 3.** Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) [40,41] model root mean squared error (RMSE) values by variable, as defined in Table 2, for calculation of the growing season average temperature (GST), growing degree days (GDD), Huglin index (HI), cool night index (CI), dryness index (DI), and the Géoviticulture multicriteria climate classification (GMCC) throughout the Willamette Valley American Viticultural Area during the LOCA CMIP5 defined historic period (1950–2005).

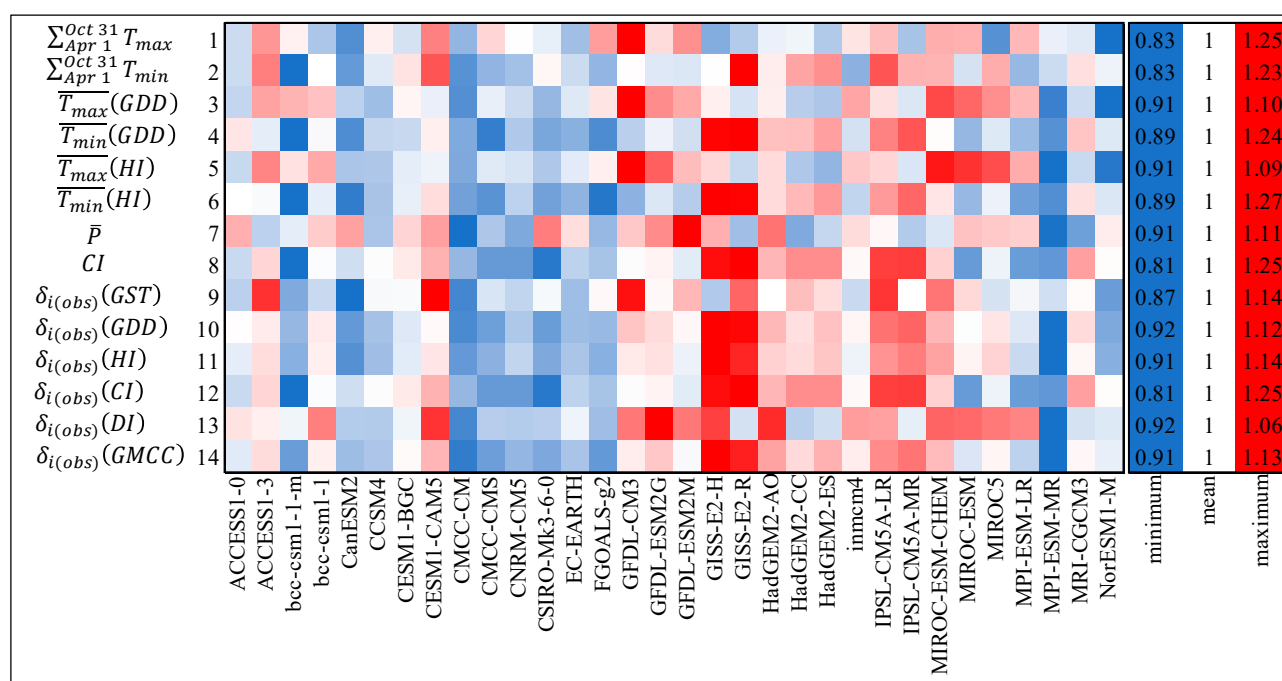
| Model         | GST                             |                                 | GDD                          |       | CI    | HI    | DI    |       |
|---------------|---------------------------------|---------------------------------|------------------------------|-------|-------|-------|-------|-------|
|               |                                 |                                 |                              |       |       |       |       |       |
|               | $\sum_{Apr 1}^{Oct 31} T_{max}$ | $\sum_{Apr 1}^{Oct 31} T_{min}$ | $\overline{T_{max} T_{min}}$ |       |       |       |       |       |
| ACCESS1-0     | 244.771                         | 175.139                         | 2.475                        | 1.570 | 1.520 | 2.517 | 1.509 | 1.677 |
| ACCESS1-3     | 282.841                         | 204.605                         | 2.631                        | 1.505 | 1.672 | 2.689 | 1.504 | 1.574 |
| bcc-csm1-1-m  | 259.278                         | 151.427                         | 2.614                        | 1.354 | 1.294 | 2.606 | 1.341 | 1.601 |
| bcc-csm1-1    | 238.086                         | 183.287                         | 2.602                        | 1.523 | 1.600 | 2.657 | 1.489 | 1.656 |
| CanESM2       | 221.319                         | 160.250                         | 2.469                        | 1.386 | 1.531 | 2.487 | 1.354 | 1.686 |
| CCSM4         | 260.058                         | 178.384                         | 2.439                        | 1.478 | 1.601 | 2.489 | 1.437 | 1.562 |
| CESM1-BGC     | 246.888                         | 188.197                         | 2.550                        | 1.481 | 1.640 | 2.549 | 1.490 | 1.650 |
| CESM1-CAM5    | 288.032                         | 211.412                         | 2.517                        | 1.551 | 1.730 | 2.561 | 1.566 | 1.687 |
| CMCC-CM       | 233.042                         | 158.486                         | 2.368                        | 1.421 | 1.445 | 2.440 | 1.396 | 1.479 |
| CMCC-CMS      | 266.691                         | 166.287                         | 2.516                        | 1.369 | 1.381 | 2.545 | 1.380 | 1.563 |
| CNRM-CM5      | 255.646                         | 169.015                         | 2.485                        | 1.461 | 1.380 | 2.535 | 1.454 | 1.532 |
| CSIRO-Mk3-6-0 | 250.895                         | 184.803                         | 2.429                        | 1.418 | 1.310 | 2.491 | 1.396 | 1.710 |
| EC-EARTH      | 236.074                         | 175.140                         | 2.505                        | 1.430 | 1.503 | 2.547 | 1.418 | 1.644 |

Table 3. Cont.

| Model          |                                   |                                   |                             |       | CI    | HI                   |                      |                |
|----------------|-----------------------------------|-----------------------------------|-----------------------------|-------|-------|----------------------|----------------------|----------------|
|                |                                   |                                   |                             |       |       |                      | DI                   |                |
|                | GST                               |                                   | GDD                         |       |       | GMCC                 |                      |                |
|                | $\sum_{Apr\ 1}^{Oct\ 31} T_{max}$ | $\sum_{Apr\ 1}^{Oct\ 31} T_{min}$ | $\overline{T_{max}T_{min}}$ |       | $CI$  | $\overline{T_{max}}$ | $\overline{T_{min}}$ | $\overline{P}$ |
| FGOALS-g2      | 281.068                           | 166.505                           | 2.567                       | 1.385 | 1.471 | 2.595                | 1.348                | 1.549          |
| GFDL-CM3       | 320.326                           | 183.522                           | 2.786                       | 1.469 | 1.602 | 2.805                | 1.416                | 1.659          |
| GFDL-ESM2G     | 265.095                           | 177.986                           | 2.652                       | 1.514 | 1.623 | 2.720                | 1.479                | 1.701          |
| GFDL-ESM2M     | 284.170                           | 177.658                           | 2.625                       | 1.489 | 1.561 | 2.640                | 1.448                | 1.796          |
| GISS-E2-H      | 230.396                           | 183.046                           | 2.556                       | 1.888 | 1.971 | 2.616                | 1.916                | 1.677          |
| GISS-E2-R      | 239.828                           | 225.547                           | 2.497                       | 1.903 | 2.006 | 2.518                | 1.902                | 1.556          |
| HadGEM2-AO     | 251.507                           | 186.377                           | 2.558                       | 1.623 | 1.721 | 2.612                | 1.568                | 1.717          |
| HadGEM2-CC     | 252.912                           | 198.520                           | 2.462                       | 1.625 | 1.786 | 2.491                | 1.623                | 1.533          |
| HadGEM2-ES     | 239.594                           | 201.824                           | 2.454                       | 1.672 | 1.785 | 2.461                | 1.675                | 1.579          |
| inmcm4         | 263.054                           | 165.371                           | 2.629                       | 1.491 | 1.615 | 2.629                | 1.458                | 1.645          |
| IPSL-CM5A-LR   | 272.671                           | 211.333                           | 2.572                       | 1.715 | 1.905 | 2.617                | 1.672                | 1.626          |
| IPSL-CM5A-MR   | 237.304                           | 196.434                           | 2.502                       | 1.776 | 1.908 | 2.538                | 1.752                | 1.567          |
| MIROC-ESM-CHEM | 277.020                           | 195.943                           | 2.715                       | 1.532 | 1.678 | 2.779                | 1.551                | 1.595          |
| MIROC-ESM      | 276.154                           | 176.654                           | 2.686                       | 1.439 | 1.383 | 2.757                | 1.424                | 1.663          |
| MIROC5         | 222.107                           | 197.340                           | 2.651                       | 1.500 | 1.580 | 2.736                | 1.493                | 1.658          |
| MPI-ESM-LR     | 274.320                           | 167.181                           | 2.610                       | 1.443 | 1.386 | 2.655                | 1.394                | 1.654          |
| MPI-ESM-MR     | 251.487                           | 175.201                           | 2.346                       | 1.392 | 1.375 | 2.355                | 1.375                | 1.480          |
| MRI-CGCM3      | 249.066                           | 188.684                           | 2.486                       | 1.616 | 1.757 | 2.519                | 1.560                | 1.523          |
| NorESM1-M      | 212.750                           | 180.521                           | 2.324                       | 1.501 | 1.608 | 2.365                | 1.480                | 1.634          |

Figure 2 presents viticulture climate classification directed LOCA CMIP5 model relative RMSE values,  $E^*$ , and median skill values,  $\delta_{i(obs)}$ , also normalized such that their means are one. Each row of Figure 2 has its own color scale wherein its minimum, mean, and maximum are blue, white, and red, respectively. Hence, for any given row, models color coded with shades of blue/red possess more/less skill than the ensemble mean. The first two/third and fourth/fifth and sixth/fifth through seventh/fifth through eighth rows of Figure 2 are the variables specific for the calculation of the GST/GDD/HI/DI/GMCC, respectively. The ninth through fourteenth rows of Figure 2 are the computed median skill values for calculation of the GST, GDD, HI, CI, DI, and GMCC, respectively.

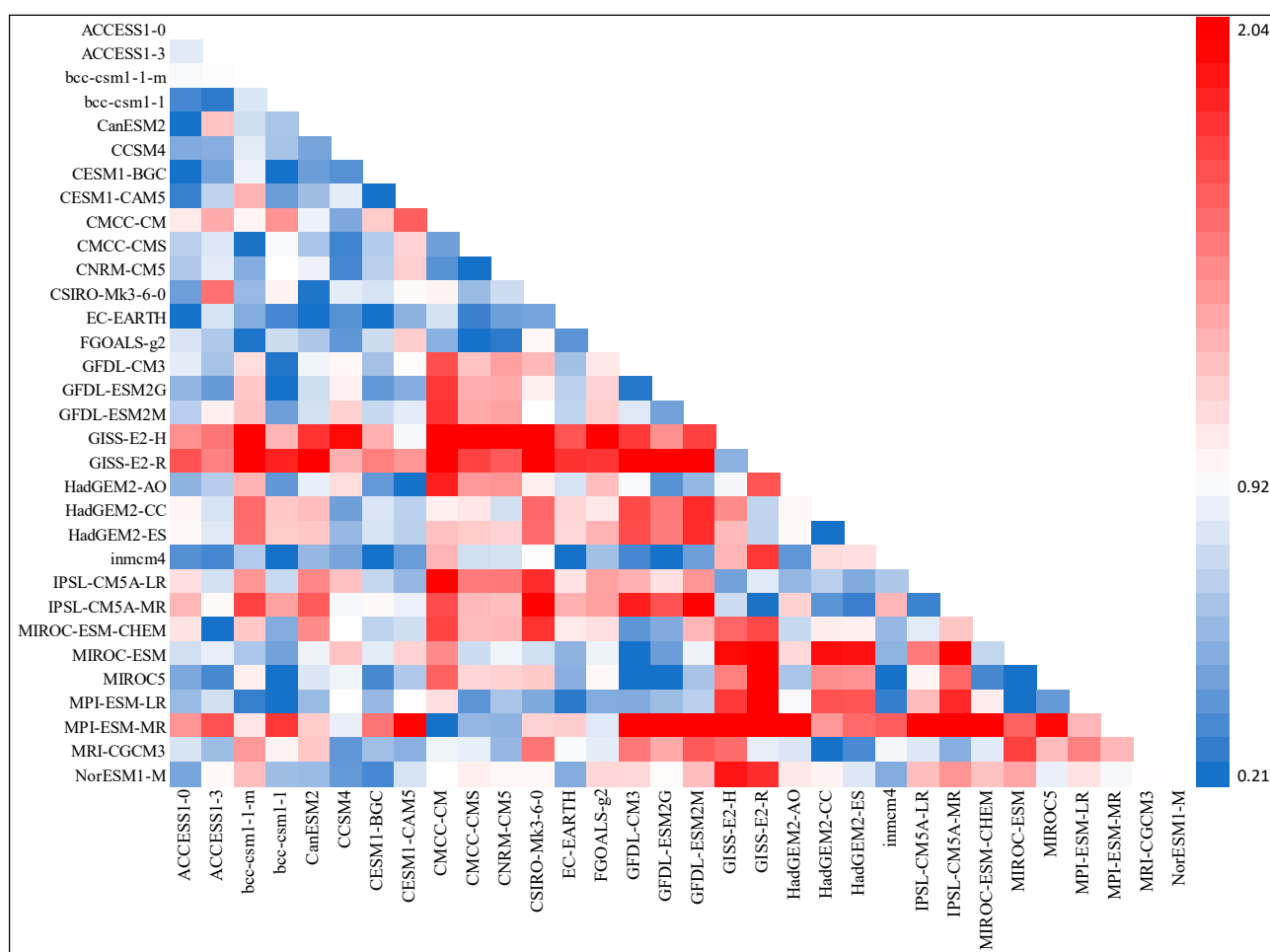
LOCA CMIP5 model median skill values vary by model and their individual rankings differ for each of the six viticulture climate classification indices that were considered for the Willamette Valley AVA. The results not only emphasize that it is necessary to account for model skill, but also, considering Equation (10), that skill weight assignments are prediction specific [17,23]. Model median skill values,  $\delta_{i(obs)}$ , are a function of their related relative RMSE values,  $E^*$ . The results presented in Figure 2 also emphasize that individual model skill varies for each of the functional variables relevant for the calculation of the GST, GDD, HI, CI, DI, and GMCC, respectively. For example, for calculation of the GST, its related relative RMSE model rankings emphasize that individual models substantively differ in their relative capacities to sum the daily (1) maximum and (2) minimum surface air temperatures from the beginning of April to the end of October.



**Figure 2.** A graphical representation of calculated viticulture climate classification specific (i.e., growing season average temperature (GST), growing degree days (GDD), Huglin index (HI), cool night index (CI), dryness index (DI), and the Géoviticulture multicriteria climate classification (GMCC)) Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) [40,41] model relative root mean squared error (RMSE) and normalized median skill values for the Willamette Valley American Viticultural Area during the LOCA CMIP5 defined historic period (1950–2005). The color scale is shown for each row, and blues/reds represent models with more/less skill than the ensemble mean (white).

Despite the previously mentioned observed differences with LOCA CMIP5 individual model skill for the calculation of six specific viticulture climate classification indices throughout the WV AVA, some observed similarities also exist. A strong consistency is observed for the rankings of model median skill values, and their related relative RMSE values,  $E^*$ , for the GDD and HI, which is understandable given their similar functional variable definitions (see Equations (2) and (3), and Table 2). In addition, across all six viticulture climate classification indices, there exists a group of nine LOCA CMIP5 models (bcc-csm1-1-m, CanESM2, CCSM4, CMCC-CM, CMCC-CMS, CNRM-CM5, CSIRO-Mk3-6-0, EC-EARTH, and MPI-ESM-MR) the median skill of which, in each case, is consistently greater than its ensemble mean.

Figure 3 is the multi-variate inter-model distance matrix that was developed using the RMSE values listed in Table 3 that are relevant for calculation of the GMCC viticulture climate classification index throughout the WV AVA. Together with specification of the radius of similarity, this matrix supports computation of model independence weights to account for, in addition to skill, archive model interdependence by down-weighting similar models [15,17,18,23]. Multi-variate inter-model distance matrices were also computed to calculate independence weights for the five other viticulture climate classification indices. In Figure 3, model pairs with small inter-model distances are denoted blue and considered dependent, whereas larger pairwise inter-model distances are red and signify independent models.



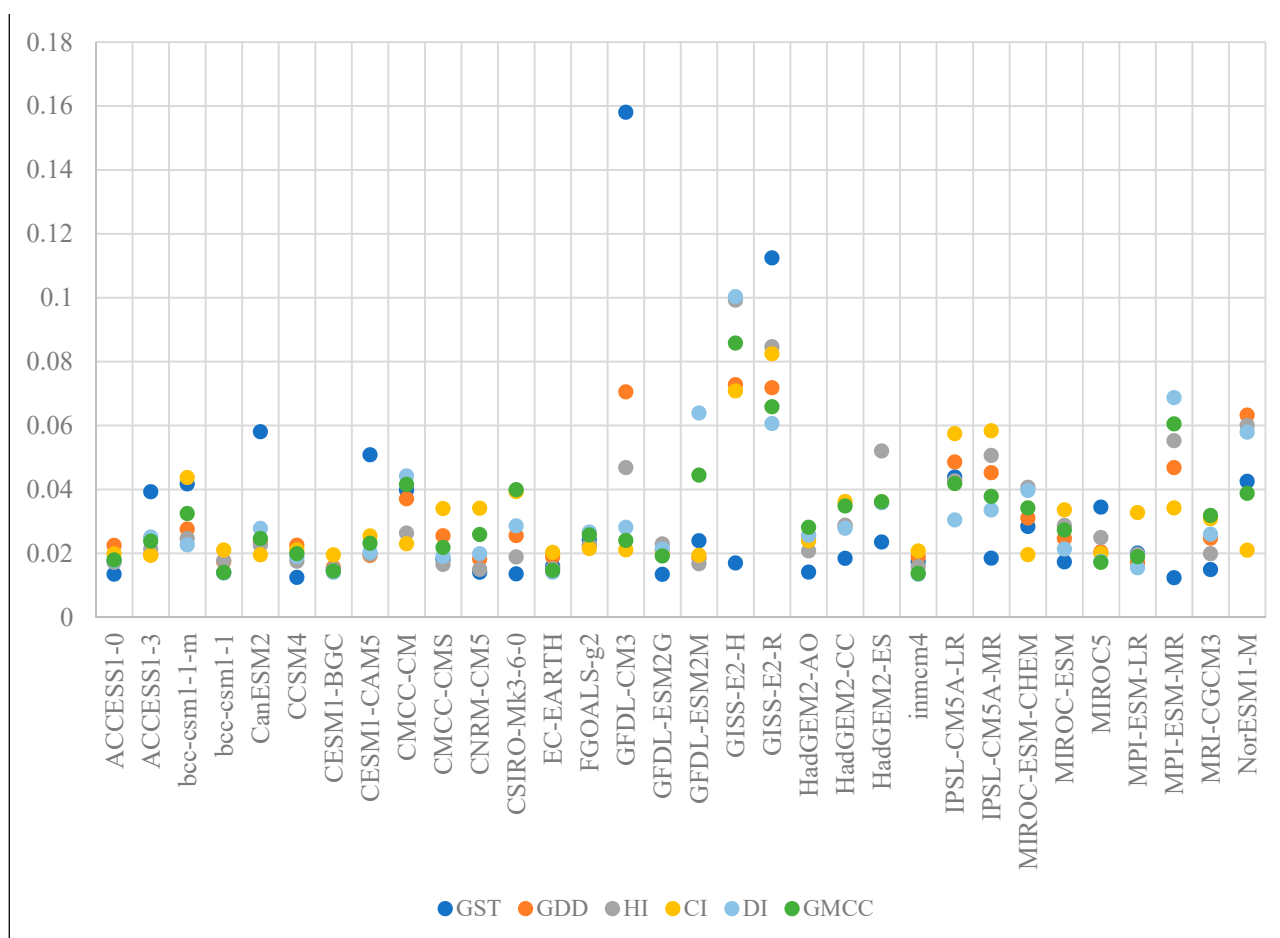
**Figure 3.** Multi-variate inter-model distance matrix [17,23] developed for calculation of the Géoviticulture multicriteria climate classification viticulture climate classification index for the Willamette Valley American Viticultural Area using the Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) model archive [40,41] for its defined historic period (1950–2005). Red matrix entries depict pairwise independent models, whereas entries colored blue denote pairwise models whose datasets are closer in agreement and demonstrate more co-dependence.

For all six viticulture climate classification indices considered in this study, calculated independence weights did not substantively alter results that were obtained which solely considered model skill weighting (see Table 4). This observation agrees with results presented by Wootten et al. [18]. Figure 4 is a plot of normalized independence weights that were calculated for each of the six viticulture climate classification indices. The normalized independence weights shown in Figure 4 emphasize that Equation (13) does not sufficiently down-weight model redundancy and only provides a slight nudge to weight the independent models more heavily. For this study, the slightly more heavily weighted independent models in general were also the models with poor relative skill. The skill as well as skill model interdependence weighting schemes generally produce weight assignments that cluster together [18].

**Table 4.** For the Willamette Valley American Viticultural Area (AVA) and growing season average temperature (GST), growing degree days (GDD), Huglin index (HI), cool night index (CI), dryness index (DI), and the Géoviticulture multicriteria climate classification (GMCC) viticulture climate classification indices, computed root mean squared error (RMSE) values obtained by comparing climate index specific processed daily Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) model historic datasets [40,41] with their equivalently processed observed data counterparts [24,41] using the skill weighting, and skill and model interdependence weighting methods [17,23]. N denotes the number of LOCA CMIP5 models defining the ensemble.

| GST                                       |        | GDD |      | HI |      | CI |      | DI |      | GMCC |      |
|-------------------------------------------|--------|-----|------|----|------|----|------|----|------|------|------|
| Skill Weighting                           |        |     |      |    |      |    |      |    |      |      |      |
| N                                         | RMSE   | N   | RMSE | N  | RMSE | N  | RMSE | N  | RMSE | N    | RMSE |
| 2                                         | 192.95 | 2   | 1.93 | 2  | 1.93 | 1  | 1.29 | 2  | 1.79 | 2    | 1.77 |
| 4                                         | 190.19 | 5   | 1.92 | 4  | 1.92 | 4  | 1.24 | 4  | 1.70 | 10   | 1.60 |
| 10                                        | 180.68 | 10  | 1.85 | 9  | 1.89 | 9  | 1.20 | 9  | 1.67 | 11   | 1.57 |
| 13                                        | 178.27 | 22  | 1.61 | 20 | 1.66 | 14 | 1.16 | 20 | 1.59 | 21   | 1.50 |
| 26                                        | 168.59 | 28  | 1.55 | 27 | 1.58 | 21 | 1.14 | 27 | 1.56 | 23   | 1.48 |
| 32                                        | 162.58 | 30  | 1.52 | 30 | 1.55 | 24 | 1.14 | 30 | 1.53 | 29   | 1.43 |
| 32                                        | 165.67 | 32  | 1.50 | 32 | 1.52 | 32 | 1.15 | 32 | 1.46 | 32   | 1.40 |
| 32                                        | 169.20 | 32  | 1.51 | 32 | 1.53 | 32 | 1.18 | 32 | 1.42 | 32   | 1.41 |
| 32                                        | 170.76 | 32  | 1.52 | 32 | 1.54 | 32 | 1.20 | 32 | 1.43 | 32   | 1.42 |
| Skill and Model Interdependence Weighting |        |     |      |    |      |    |      |    |      |      |      |
| N                                         | RMSE   | N   | RMSE | N  | RMSE | N  | RMSE | N  | RMSE | N    | RMSE |
| 2                                         | 193.03 | 2   | 1.93 | 2  | 1.93 | 1  | 1.29 | 1  | 1.79 | 2    | 1.77 |
| 3                                         | 191.09 | 4   | 1.92 | 3  | 1.92 | 3  | 1.25 | 6  | 1.73 | 10   | 1.64 |
| 6                                         | 183.59 | 9   | 1.87 | 8  | 1.91 | 8  | 1.20 | 12 | 1.71 | 11   | 1.62 |
| 10                                        | 181.50 | 18  | 1.65 | 18 | 1.75 | 13 | 1.19 | 19 | 1.65 | 23   | 1.55 |
| 24                                        | 172.79 | 28  | 1.59 | 27 | 1.67 | 20 | 1.15 | 21 | 1.62 | 25   | 1.53 |
| 32                                        | 164.74 | 30  | 1.55 | 30 | 1.61 | 23 | 1.14 | 28 | 1.60 | 30   | 1.46 |
| 32                                        | 167.06 | 32  | 1.50 | 32 | 1.53 | 32 | 1.14 | 32 | 1.53 | 32   | 1.41 |
| 32                                        | 174.76 | 32  | 1.51 | 32 | 1.53 | 32 | 1.18 | 32 | 1.42 | 32   | 1.41 |
| 32                                        | 178.34 | 32  | 1.53 | 32 | 1.56 | 32 | 1.21 | 32 | 1.43 | 32   | 1.42 |

Viticulture climate classification directed LOCA CMIP5 model median skill values,  $\delta_{i(obs)}$ , directly computed from the RMSE values listed in Table 3 and graphically represented in Figure 2, were ranked to develop ensemble compositions of size one to thirty-two. For each prediction of interest (i.e., GST, GDD, HI, CI, DI, and GMCC), the RMSE was subsequently computed for each of these thirty-two model compositions using simple model averaging (see Figure 5 for GST and Figures S1–S5 for GDD, HI, CI, DI, and GMCC, respectively). As previously mentioned, this simple method is like the approach presented by Yang et al. [25]; however, instead of the ranked skill scores from the Taylor diagram [28] we used the ranked model median skills.

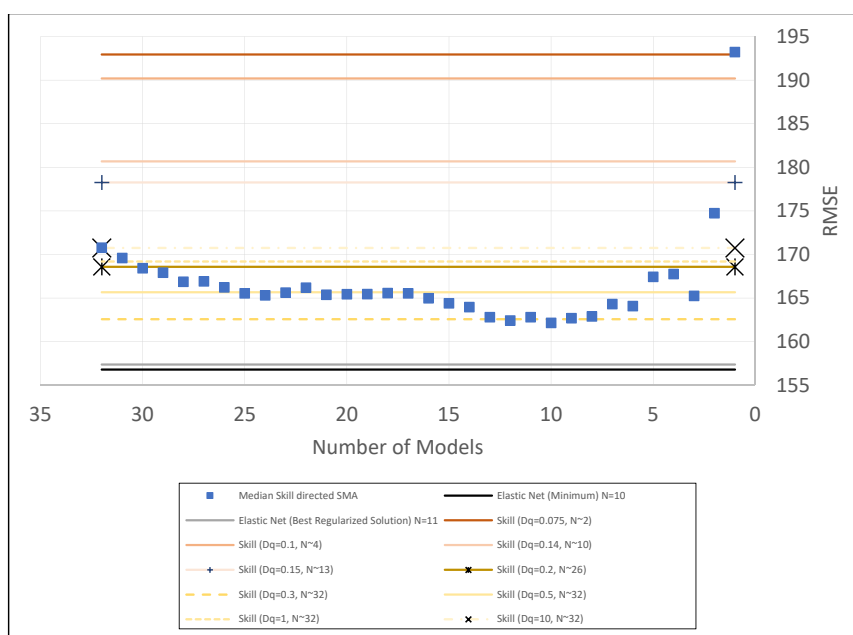


**Figure 4.** Normalized independence weights developed for the growing season average temperature (GST), growing degree days (GDD), Huglin index (HI), cool night index (CI), dryness index (DI), and the Géoviticulture multicriteria climate classification (GMCC) viticulture climate classification indices throughout the Willamette Valley American Viticultural Area using the Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) model archive [40,41] for its defined historic period (1950–2005).

Figure 5 includes a plot of the RMSEs computed using simple model averaging (SMA) for each of the thirty-two ensemble compositions formed from the ranked median skill values that are of relevance for the calculation of the growing season average temperature, GST, throughout the WV AVA. This median skill directed SMA plot demonstrates a quick drop from its single model RMSE to a minimum, associated with an ensemble subset of size 10, and subsequently rises to the ensemble mean composed of all thirty-two models. This pattern is generally also observed in Figures S1–S5; however, the ensemble sizes associated with the minimum RMSE for prediction of the GDD, HI, CI, DI, and GMCC are 11, 18, 12, 14, and 21, respectively.

Figure 5 and Figure S1–S5 also include results from application of the skill weighting method [17,23], with each figure including results associated with different parameterizations (i.e., values for the radius of model equality) that effectively result in ensemble compositions of varying sizes when zero is defined for a computed weight with a value less than  $1 \times 10^{-4}$  to estimate ensemble size for the given  $D_q$  value. In each figure, as the radius of model equality is increased from its initially assigned value which yields the most parsimonious ensemble composition considered for the skill weighting method, RMSE values for the skill weighting ensembles drop to a minimum and subsequently rise toward the no skill ensemble mean. For the skill weighting method, its minimum is always composed of all 32 models (see Figures 5 and S1–S5).





**Figure 5.** For the Willamette Valley American Viticultural Area (AVA) and growing season average temperature (GST), computed root mean squared error (RMSE) values obtained by comparing GST specific processed daily Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) model historic datasets [40,41] with their equivalently processed observed data counterparts [18,42] using skill weighting, with various values for its radius of model equality,  $D_q$  [17,23], model median skill directed simple model averaging (SMA), and elastic-net regularization [33].  $N$  denotes the number of LOCA CMIP5 models defining the ensemble.

Herger et al. [24] emphasized that the skill and model interdependence weighting methods are directed at assigning continuous weights as opposed to the identification of optimal ensemble subsets. Results presented herein and also by Wootten et al. [18] reinforce this observation (see Figure 5 and Figure S1–S5 and Table 4). However, it is notable to observe that the model median skill directed simple model averaging method generally yields lower RMSE values than those computed from the skill weighting method for ensembles of comparable sizes (see Figure 5 and Figure S1–S5). The method also consistently finds an optimal ensemble subset with an RMSE value less than or equal in value to the skill weighting method minimizing RMSE that includes the complete archive (see Figure 5 and Figure S1–S5).

While relatively straightforward methods such as skill weighting, skill weighting including interdependence, and model median skill SMA can be proposed and applied to assign weights for model averaging, one can instead apply optimization and inference methods that involve a systematic and iterative search of parameter space, including implementation of formal stopping/convergence criteria, to find prediction specific optimal ensemble compositions and weights to account for skill and interdependence. For example, Massoud et al. [15], Massoud et al. [16], and Wootten et al. [18] applied BMA [61], which involves Markov Chain Monte Carlo sampling, and Herger et al. [24] employed mixed integer quadratic programming for model averaging. We apply a calibration method that to our knowledge has not previously been considered for combining multi-model climate projections; viz., the elastic-net penalty [33]. It can be readily configured to either find smooth solutions or parsimonious models while preserving fit with the observational dataset.

Figure 5 and Figure S1–S5 further include results from applications of elastic-net regularization configured for automatic variable selection. Each elastic-net application was performed using the R software package ‘glmnet’ [29], employing  $k$ -fold CV with  $k = 32$  (i.e., leave-one-out CV),  $\alpha = 0.95$ , and  $\tilde{w}_i = 1$  for each observation. The best regularizing model is defined to be at the largest  $\lambda$  value within one standard error of the minimum [29,60,62]. Each figure includes the minimum, identified using leave-one-out

cross validation, and its associated best regularized solution. The best regularized solution in each case differed only slightly from the minimizing solution. Elastic-net regularization consistently found a minimizing optimal ensemble subset that outperformed the other three model averaging methods considered in this study (see Figures 5 and S1–S5). From a practical perspective, for a given prediction, application of elastic-net regularization is less resource and time intensive to apply than the skill weighting, skill and model interdependence weighting, and the model median skill SMA methods.

The calibrated models listed in Table 5 are prediction specific (i.e., GST, GDD, HI, CI, DI, GMCC) weight assignments computed from cross validation directed applications of elastic-net regularization configured for automatic variable selection. Equation (8) can be applied using the weights listed in Table 5 to optimally compute, within the framework presented herein, the GST, GDD, HI, CI, DI, or GMCC for the WV AVA while using LOCA CMIP5 historical and future climate projections. For example, the  $T_{max}$  and  $T_{min}$  associated with the ensemble model  $M$  defined using the weights listed in the second column of Table 5 can be used to optimally compute the GST throughout the WV AVA using Equation (1). The model weights listed in Table 5 are based on comparisons with the Livneh dataset [42] for the LOCA CMIP5 defined historical period [40,41]. While not shown, closely matching summary statistics were also obtained for each respective viticulture climate index, for calibration and verification, from cross validation directed elastic-net regularization fits which held back twenty-five percent of the LOCA CMIP5 historical data.

The models listed in Table 5 for optimized computation of the GST, GDD, HI, CI, DI, and GMCC throughout the WV AVA further reinforce statements regarding CMIP5 archive democracy and redundancy [15–17], and the need to develop ensemble compositions and weights assignments for a specific modeling objective and location [17]. For each viticulture climate classification index, unique ensemble subsets with variable sizes and weights were identified which outperform the complete archive ensemble mean in each case. Table 6 summarizes computed correlations among the models listed in Table 5. The optimal ensembles identified for the GST and CI viticulture climate classification indices both demonstrate poor correlation with the remaining four indices. Strong correlations exist among the optimal ensembles identified for computation of the GDD, HI, DI, and GMCC indices for the WV AVA when  $\tilde{w}_i = 1$  for each observation. The general functional definitions for the GST and the CI are each unique relative to the remaining four indices which share common variables (see Equations (1)–(4), (6) and (7)). However, alternate values could be assigned to the  $\tilde{w}_i$ , particularly to fit the DI and GMCC which include three and four distinct meteorological data requirements, respectively, to balance the contributions among the functional variables that define the indices [27,63]. For example, the final two columns of Table 5 include ensembles obtained for the DI and GMCC when the  $\tilde{w}_i$  are assigned such that model-to-measurement misfit for each individual functional variable is equal in value at the start of the optimization [27,63]. With the  $\tilde{w}_i$  assigned in this manner to fit the DI and GMCC, computed correlations among the ensembles decrease except for the CI and GMCC as expected (see Table 6).

Of the thirty-two models which constitute the entire LOCA CMIP5 archive, 10, 21, 21, 13, 22, and 21 of its members compose unique ensembles for optimal computation of the GST, GDD, HI, CI, DI, and GMCC throughout the WV AVA (see Table 5) when  $\tilde{w}_i = 1$  for each observation. Across these six viticulture climate classification index ensembles, seven LOCA CMIP5 models were consistently excluded from their compositions with a zero valued weights assignment (CESM1-BGC, CESM1-CAM5, GFDL-CM3, GFDL-ESM2G, GISS-E2-R, MIROC-ESM-CHEM, MPI-ESM-LR). If zero is defined for weights with a value less than 0.02, then the number of excluded models rises to ten (and includes EC-EARTH, IPSL-CM5A-LR, MIROC5). While unique to the WV AVA, the LOCA CMIP5 archive, and the six indices evaluated herein, these optimal ensemble composition sizes and weights differ relative to previously reported ensemble sizes and weights assignments which used downscaled climate models to evaluate the impacts of climate change to viticulture [45–53].

**Table 5.** Optimal weights obtained using elastic-net regularization [33] to assign when applying Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) historic and future climate projections [40,41] to compute the growing season average temperature (GST), growing degree days (GDD), Huglin index (HI), cool night index (CI), dryness index (DI), and the Géoviticulture multicriteria climate classification (GMCC) indices throughout the Willamette Valley American Viticultural Area (AVA). (\* indicates that each functional observation group defining the index was balanced to be seen of equal importance at the start of the optimization).

| Model          | Viticulture Climate Classification Index |       |       |       |                       |                       |                       | GMCC * |
|----------------|------------------------------------------|-------|-------|-------|-----------------------|-----------------------|-----------------------|--------|
|                | GST                                      | GDD   | HI    | CI    | DI                    | GMCC                  | DI *                  |        |
| ACCESS1-0      | 0                                        | 0.020 | 0.014 | 0.147 | 0.030                 | 0.030                 | 0.039                 | 0.047  |
| ACCESS1-3      | 0                                        | 0.015 | 0.011 | 0     | 0.023                 | 0.018                 | 0.053                 | 0.047  |
| bcc-csm1-1-m   | 0                                        | 0     | 0.013 | 0.264 | 0.023                 | 0.030                 | 0.028                 | 0.044  |
| bcc-csm1-1     | 0.084                                    | 0     | 0     | 0     | 0                     | 0                     | $7.59 \times 10^{-3}$ | 0.003  |
| CanESM2        | 0.173                                    | 0.090 | 0.095 | 0.090 | 0.080                 | 0.084                 | 0.040                 | 0.044  |
| CCSM4          | 0                                        | 0.034 | 0.013 | 0     | 0.026                 | 0.037                 | 0.065                 | 0.064  |
| CESM1-BGC      | 0                                        | 0     | 0     | 0     | 0                     | 0                     | 0                     | 0      |
| CESM1-CAM5     | 0                                        | 0     | 0     | 0     | 0                     | 0                     | 0                     | 0      |
| CMCC-CM        | 0.144                                    | 0.100 | 0.082 | 0.136 | 0.085                 | 0.098                 | 0.096                 | 0.101  |
| CMCC-CMS       | 0                                        | 0.038 | 0.047 | 0.025 | 0.051                 | 0.054                 | 0.049                 | 0.051  |
| CNRM-CM5       | 0.056                                    | 0.056 | 0.054 | 0.042 | 0.068                 | 0.075                 | 0.083                 | 0.091  |
| CSIRO-Mk3-6-0  | 0                                        | 0.078 | 0.068 | 0     | 0.059                 | 0.060                 | 0.037                 | 0.037  |
| EC-EARTH       | 0                                        | 0.005 | 0.016 | 0     | $5.39 \times 10^{-7}$ | $3.35 \times 10^{-6}$ | 0                     | 0      |
| FGOALS-g2      | 0                                        | 0.068 | 0.068 | 0.030 | 0.070                 | 0.074                 | 0.038                 | 0.046  |
| GFDL-CM3       | 0                                        | 0     | 0     | 0     | 0                     | 0                     | 0.010                 | 0.004  |
| GFDL-ESM2G     | 0                                        | 0     | 0     | 0     | 0                     | 0                     | 0                     | 0      |
| GFDL-ESM2M     | 0                                        | 0.019 | 0.022 | 0     | 0.015                 | 0.015                 | 0                     | 0      |
| GISS-E2-H      | 0.172                                    | 0.015 | 0     | 0     | 0                     | 0                     | 0                     | 0      |
| GISS-E2-R      | 0                                        | 0     | 0     | 0     | 0                     | 0                     | 0.011                 | 0.001  |
| HadGEM2-AO     | 0                                        | 0.011 | 0.011 | 0.031 | 0.009                 | 0.006                 | 0                     | 0      |
| HadGEM2-CC     | 0                                        | 0.030 | 0.023 | 0     | 0.055                 | 0.050                 | 0.129                 | 0.115  |
| HadGEM2-ES     | 0.089                                    | 0.055 | 0.067 | 0.021 | 0.071                 | 0.070                 | 0.072                 | 0.068  |
| inmcm4         | 0.037                                    | 0.055 | 0.067 | 0     | 0.043                 | 0.041                 | 0                     | 0      |
| IPSL-CM5A-LR   | 0                                        | 0     | 0     | 0     | 0.002                 | 0                     | 0.052                 | 0.036  |
| IPSL-CM5A-MR   | 0                                        | 0.038 | 0.034 | 0.075 | 0.036                 | 0.035                 | 0.041                 | 0.043  |
| MIROC-ESM-CHEM | 0                                        | 0     | 0     | 0     | 0                     | 0                     | 0.014                 | 0.009  |
| MIROC-ESM      | 0.091                                    | 0.053 | 0.062 | 0.001 | 0.045                 | 0.040                 | 0                     | 0.002  |
| MIROC5         | 0.008                                    | 0     | 0     | 0     | 0                     | 0                     | 0                     | 0      |
| MPI-ESM-LR     | 0                                        | 0     | 0     | 0     | 0                     | 0                     | 0.012                 | 0.011  |
| MPI-ESM-MR     | 0                                        | 0.086 | 0.099 | 0.060 | 0.092                 | 0.082                 | 0.065                 | 0.072  |
| MRI-CGCM3      | 0                                        | 0.040 | 0.049 | 0.072 | 0.043                 | 0.036                 | 0.009                 | 0.019  |
| NorESM1-M      | 0.146                                    | 0.092 | 0.085 | 0     | 0.073                 | 0.064                 | 0.037                 | 0.034  |
| Non-Zero Count | 10                                       | 21    | 21    | 13    | 22                    | 21                    | 22                    | 23     |

**Table 6.** Correlations among the Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) [40,41] ensemble compositions and weights obtained using elastic-net regularization [33] to compute historic and future climate projections for the growing season average temperature (GST), growing degree days (GDD), Huglin index (HI), cool night index (CI), dryness index (DI), and the Géoviticulture multicriteria climate classification (GMCC) indices throughout the Willamette Valley American Viticultural Area (AVA). Values above the main diagonal treated each individual processed observation equally. Values below the main diagonal were obtained by balancing each functional observation group defining the index to be seen of equal importance at the start of the optimization).

|      | GST   | GDD   | HI    | CI    | DI    | GMCC  |
|------|-------|-------|-------|-------|-------|-------|
| GST  | 1     | 0.518 | 0.456 | 0.083 | 0.394 | 0.408 |
| GDD  |       | 1     | 0.966 | 0.192 | 0.946 | 0.936 |
| HI   |       |       | 1     | 0.219 | 0.944 | 0.919 |
| CI   |       |       |       | 1     | 0.311 | 0.361 |
| DI   | 0.121 | 0.501 | 0.415 | 0.234 | 1     | 0.987 |
| GMCC | 0.134 | 0.565 | 0.490 | 0.373 | 0.980 | 1     |

The data and methods employed in our study are applicable for other wine regions. The approach outlined in this study for developing model-observation comparisons is of value from a practical perspective in that it is potentially far easier to compute and apply a viticulture climate classification directed optimal ensemble based on its underlying data types than the index itself, particularly for compute intensive indices such as the dryness index [37]. A potentially promising related research direction would be to explore updating the cultivar specific phenological understanding associated with one or more of the easy to compute and widely used temperature only based indices considered in this study by leveraging a recent contribution about the modeling of grapevine phenology [64]. Further application and evaluation of the two methods that were applied for the first time in this study to generalize climate model choice for the WV AVA using the LOCA CMIP5 archive (i.e., the simple model averaging method of [25] that uses the ranked model median skills and elastic-net regularization [33]) is encouraged and would not necessarily be limited to our study area or predictions of viticulture climate classification.

The optimal ensemble subsets computed in this study for each of the six indices are unique prediction and location specific model combinations that adhere to the principle of parsimony by simultaneously accounting for ensemble model skill and similarity for the LOCA CMIP5 historical period. Our approach to ensemble model selection is like other similar studies in that it depends on historical calibration training data [18,24,25]. It is preferable to use an optimal ensemble subset rather than its comparable no skill ensemble mean from the perspective that the parsimonious model combination does not contain duplicative information that can bias results and correlation statistics [23] and also in that the ensemble mean is prone to poor model accuracy and poor model uncertainty quantification [24]. The skill of each prediction and location specific parsimonious optimal ensemble subset for predicting future climates is dependent upon whether the calibrated model was overfit [24,65]. Further related study is encouraged to better understand wine region specific model archive weighting strategies that not only select optimal ensemble subsets for viticulture climate impact assessments but that also constrain projection uncertainty. Although application of the elastic net identifies the best regularized solution and yields robust estimates of model uncertainty, one potentially promising direction to explore, for example, is to combine and compare the elastic net and Reliability Ensemble Averaging (REA) methods [33,65]. Combining results from application of the elastic net with the REA method would result in an REA model performance criterion that simultaneously accounts for model skill and interdependence and an REA model convergence criterion that is less likely to be confounded due to model replication [23].

#### 4. Conclusions

This study was motivated by recent contributions which underscored the need for prediction and location specific Coupled Model Intercomparison Project Phase 5 ensemble compositions and weights that account for model skill and similarity and that address ensemble subset selection. It was also motivated by the expressed need to further study methods that generalize climate model choice. This study considered the complete archive of the thirty-two Coupled Model Intercomparison Project Phase 5 Localized Constructed Analogs downscaled daily historic datasets. It explored, for the first time, ensemble subset selection for the optimal computation of the growing season average temperature, growing degree days, Huglin's index, cool night index, dryness index, and the multicriteria climate classification indices throughout the Willamette Valley American Viticultural Area. Four model averaging methods were evaluated including two recently introduced from the climate science community, viz., skill weighting, and skill and model interdependence weighting. Two additional methods, neither previously considered, that were also applied to generalize climate model selection included model median skill directed simple model averaging, a modification of a recently introduced approach to combine climate models, and elastic-net regularization, a formal optimization technique that is well designed for application to automatic variable selection problems. The scope of the method comparison

performed in this study is limited to these four methods, the 553 model-observation data locations which cover the WV AVA as shown in Figure 1, and the model-observation datasets listed in Table 2. Further research of methods directed at ensemble subset selection is encouraged.

For a given viticulture climate classification index, application of the skill as well as skill model interdependence weighting methods involves evaluating the performance of each indices' individual functional variables. Application of the two climate model weighting strategies to compute the GST, GDD, HI, CI, DI, and GMCC throughout the WV AVA revealed that LOCA CMIP5 individual model skill varies by each indices' defined processed meteorological variables. Moreover, also considering independence in addition to skill did not substantively alter ensemble compositions or weights relative to solely considering model skill. While results obtained from their application did vary as a function of their parameterization, optimal ensembles from these two methods consistently, across all six indices, involved the complete archive and were less optimal than those obtained from the other two model averaging methods. Interestingly, model median skill directed SMA, which can be directly applied following application of the skill weighting strategy, was able to consistently compute LOCA CMIP5 ensemble subsets for each of the six viticulture climate classification indices. The RMSE values listed in Table 3 permit researchers to independently apply the skill weighting, skill weighting including interdependence, and model median skill directed SMA strategies to develop ensembles to compute the GST, GDD, HI, CI, DI, and GMCC indices throughout the WV AVA using historic and future LOCA CMIP5 climate projections.

Of the four model averaging methods considered in this study, results obtained from application of elastic-net regularization consistently outperformed the other three methods with finding optimal ensemble subsets from among the complete LOCA CMIP5 model archive. Its application, much like the other three methods, is straightforward. The LOCA CMIP5 ensemble compositions and weights did vary by viticulture climate classification index. While specific to the archive and location, it was noteworthy to observe that the optimal ensemble composition sizes and weights across the six viticulture climate classification indices differed relative to ensemble sizes and weights assignments reported in published studies that focused on climate change impacts to viticulture. The weights listed in Table 5, derived from applications of elastic-net regularization, permit researchers to independently develop ensembles to compute the GST, GDD, HI, CI, DI, and GMCC throughout the WV AVA using historic and future LOCA CMIP5 climate projections.

Future projections of the bioclimatic indices considered in this study, computed using optimal compositions and weights, enable robust, adaptive evaluations of existing and new sites for viticulture throughout the Willamette Valley American Viticultural Area. The climate model ensemble weights assignment strategies presented herein are transferable to other wine regions. Further related study is encouraged, not only for ensemble subset selection, but also to constrain forecast uncertainty. For example, while compute intensive for some indices such as the DI and the GMCC, further related study that seeks to find optimal ensemble subsets using the computed indices instead is also encouraged. This would enable comparison with the results contained herein as well as provide opportunity to find optimal ensembles across predefined sets of viticulture climate classification indices for a given wine region.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/cli9090140/s1>, Figure S1: For the Willamette Valley American Viticultural Area (AVA) and growing degree days (GDD), computed root mean squared error (RMSE) values obtained by comparing GDD specific processed daily Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) model historic datasets [40,41] with their equivalently processed observed data counterparts [18,42] using skill weighting, with various values for its radius of model equality,  $D_q$  [17,23], model median skill directed simple model averaging (SMA), and elastic-net regularization [33].  $N$  denotes the number of LOCA CMIP5 models defining the ensemble, Figure S2: For the Willamette Valley American Viticultural Area (AVA) and Huglin index (HI), com-



puted root mean squared error (RMSE) values obtained by comparing HI specific processed daily Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) model historic datasets [40,41] with their equivalently processed observed data counterparts [18,42] using skill weighting, with various values for its radius of model equality,  $D_q$  [17,23], model median skill directed simple model averaging (SMA), and elastic-net regularization [33].  $N$  denotes the number of LOCA CMIP5 models defining the ensemble, Figure S3: For the Willamette Valley American Viticultural Area (AVA) and cool night index (CI), computed root mean squared error (RMSE) values obtained by comparing CI specific processed daily Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) model historic datasets [40,41] with their equivalently processed observed data counterparts [18,42] using skill weighting, with various values for its radius of model equality,  $D_q$  [17,23], model median skill directed simple model averaging (SMA), and elastic-net regularization [33].  $N$  denotes the number of LOCA CMIP5 models defining the ensemble, Figure S4: For the Willamette Valley American Viticultural Area (AVA) and dryness index (DI), computed root mean squared error (RMSE) values obtained by comparing DI specific processed daily Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) model historic datasets [40,41] with their equivalently processed observed data counterparts [18,42] using skill weighting, with various values for its radius of model equality,  $D_q$  [17,23], model median skill directed simple model averaging (SMA), and elastic-net regularization [33].  $N$  denotes the number of LOCA CMIP5 models defining the ensemble, Figure S5: For the Willamette Valley American Viticultural Area (AVA) and Géoviticulture multicriteria climate classification (GMCC), computed root mean squared error (RMSE) values obtained by comparing GMCC specific processed daily Localized Constructed Analogs (LOCA) Coupled Model Intercomparison Project phase 5 (CMIP5) model historic datasets [40,41] with their equivalently processed observed data counterparts [18,42] using skill weighting, with various values for its radius of model equality,  $D_q$  [17,23], model median skill directed simple model averaging (SMA), and elastic-net regularization [33].  $N$  denotes the number of LOCA CMIP5 models defining the ensemble, Table S1: The 32 models and modeling groups that provided Coupled Model Intercomparison Project phase 5 (CMIP5) data used in this study.

**Author Contributions:** Conceptualization, B.S., B.B. and M.S.; methodology, B.S., B.B. and M.S.; formal analysis, B.S.; investigation, B.S., B.B. and M.S.; writing—original draft preparation, B.S.; writing—review and editing, B.S., B.B. and M.S.; visualization, B.S., B.B. and M.S.; supervision, B.B. and M.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** [https://gdo-dcp.ucllnl.org/downscaled\\_cmip\\_projections/](https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/); <https://psl.noaa.gov/> (accessed on 2 December 2020).

**Acknowledgments:** The first author would like to gratefully thank the Northwest Wine Studies Center Wine Studies Program located at Chemeketa Eola in the Eola-Amity Hills sub-AVA of the WV AVA for their support of this capstone project research work activity. The authors acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups listed in Table S1 for producing and making available their model output. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provided coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. Livneh data were provided by the NOAA/OAR/ESRL PSL, Boulder, Colorado, USA, from their website at <https://psl.noaa.gov/> (accessed on 2 December 2020). We also thank the three anonymous reviewers for their comments, which improved this article.

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

## References

1. Winkler, J.; Cook, J.; Kliewer, W.M.; Lider, L.A. *General Viticulture*, 4th ed.; University of California: Berkeley, CA, USA, 1974.
2. Carbonneau, A. Ecophysiologie de la vigne et terroir. In *Terroir, Zonazione Viticoltura*; Fregoni, M., Schuster, D., Paoletti, A., Eds.; Phytoline: Piacenza, Italy, 2003; pp. 61–102.



3. Jones, G.V.; Duff, A.; Hall, A.; Myers, J. Spatial Analysis of Climate in Winegrape Growing Regions in the Western United States. *Am. J. Enol. Vitic.* **2010**, *61*, 313–326.
4. Shaw, T.B. Climate change and the evolution of the Ontario cool climate wine regions in Canada. *J. Wine Res.* **2016**, *28*, 13–45. [[CrossRef](#)]
5. Coombe, B. Influence of temperature on composition and quality of grapes. *Acta Hort.* **1987**, *206*, 23–36. [[CrossRef](#)]
6. Hunter, J.J.; Bonnardot, V. Climatic requirements for optimal physiological processes: A factor in viticultural zoning. In Proceedings of the 4th International Symposium on Viticultural Zoning, Avignon, France, 17–20 June 2002; pp. 553–565.
7. Hunter, J.J.; Deloire, A. Terroir and vine water relation effects on grape ripening and wine quality of Syrah/R99. In Proceedings of the VIth International Terroir Congress, Bordeaux and Montpellier, France, 3–7 July 2006.
8. Deloire, A.; Vaudour, E.; Carey, V.A.; Bonnardot, V.; van Leeuwen, C. Grapevine responses to terroir: A global approach. *OENO One* **2005**, *39*, 149–162. [[CrossRef](#)]
9. Van Leeuwen, C.; Garnier, C.; Agut, C.; Baculat, B.; Barbeau, G.; Besnard, E.; Bois, B.; Boursiquot, J.-M.; Chuine, I.; Dessup, T.; et al. Heat requirements for grapevine varieties is essential information to adapt plant material in a changing climate. In Proceedings of the 7th International Terroir Congress, Nyon, Switzerland, 19–23 May 2008.
10. Badr, G.; Hoogenboom, G.; Abouali, M.; Moyer, M.; Keller, M. Analysis of several bioclimatic indices for viticultural zoning in the Pacific Northwest. *Clim. Res.* **2018**, *76*, 203–223. [[CrossRef](#)]
11. Quenol, H.; Atauri, I.G.D.C.; Bois, B.; Sturman, A.; Bonnardot, V.; Le Roux, R. Which climatic modeling to assess climate change impacts on vineyards? *OENO One* **2017**, *51*, 91–97. [[CrossRef](#)]
12. 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](#)]
13. Eyring, V.; Cox, P.M.; Flato, G.M.; Gleckler, P.J.; Abramowitz, G.; Caldwell, P.; Collins, W.D.; Gier, B.K.; Hall, A.D.; Hoffman, F.M.; et al. Taking climate model evaluation to the next level. *Nat. Clim. Chang.* **2019**, *9*, 102–110. [[CrossRef](#)]
14. Knutti, R.; Sedláček, J.; Sanderson, B.M.; Lorenz, R.; Fischer, E.M.; Eyring, V. A climate model projection weighting scheme accounting for performance and interdependence. *Geophys. Res. Lett.* **2017**, *44*, 1909–1918. [[CrossRef](#)]
15. Massoud, E.; Espinoza, V.; Guan, B.; Waliser, D. Global Climate Model Ensemble Approaches for Future Projections of Atmospheric Rivers. *Earth's Future* **2019**, *7*, 1136–1151. [[CrossRef](#)]
16. Massoud, E.C.; Lee, H.; Gibson, P.B.; Loikith, P.; Waliser, D.E. Bayesian Model Averaging of Climate Model Projections Constrained by Precipitation Observations over the Contiguous United States. *J. Hydrometeorol.* **2020**, *21*, 2401–2418. [[CrossRef](#)]
17. Sanderson, B.M.; Wehner, M.; Knutti, R. Skill and independence weighting for multi-model assessments. *Geosci. Model Dev.* **2017**, *10*, 2379–2395. [[CrossRef](#)]
18. Wootten, A.; Massoud, E.; Sengupta, A.; Waliser, D.; Lee, H. The Effect of Statistical Downscaling on the Weighting of Multi-Model Ensembles of Precipitation. *Climate* **2020**, *8*, 138. [[CrossRef](#)]
19. Knutti, R.; Masson, D.; Gettelman, A. Climate model genealogy: Generation CMIP5 and how we got there. *Geophys. Res. Lett.* **2013**, *40*, 1194–1199. [[CrossRef](#)]
20. Masson, D.; Knutti, R. Climate model genealogy. *Geophys. Res. Lett.* **2011**, *38*, 08703. [[CrossRef](#)]
21. Annan, J.D.; Hargreaves, J. Understanding the CMIP3 Multimodel Ensemble. *J. Clim.* **2011**, *24*, 4529–4538. [[CrossRef](#)]
22. Sanderson, B.; Knutti, R. On the interpretation of constrained climate model ensembles. *Geophys. Res. Lett.* **2012**, *39*, 16708. [[CrossRef](#)]
23. Sanderson, B.; Knutti, R.; Caldwell, P. A Representative Democracy to Reduce Interdependency in a Multimodel Ensemble. *J. Clim.* **2015**, *28*, 5171–5194. [[CrossRef](#)]
24. Herger, N.; Abramowitz, G.; Knutti, R.; Angéilil, O.; Lehmann, K.; Sanderson, B.M. Selecting a climate model subset to optimise key ensemble properties. *Earth Syst. Dyn.* **2018**, *9*, 135–151. [[CrossRef](#)]
25. Yang, X.; Yu, X.; Wang, Y.; He, X.; Pan, M.; Zhang, M.; Liu, Y.; Ren, L.; Sheffield, J. The Optimal Multimodel Ensemble of Bias-Corrected CMIP5 Climate Models over China. *J. Hydrometeorol.* **2020**, *21*, 845–863. [[CrossRef](#)]
26. Lorenz, R.; Herger, N.; Sedláček, J.; Eyring, V.; Fischer, E.M.; Knutti, R. Prospects and Caveats of Weighting Climate Models for Summer Maximum Temperature Projections Over North America. *J. Geophys. Res. Atmos.* **2018**, *123*, 4509–4526. [[CrossRef](#)]
27. Doherty, J. *Calibration and Uncertainty Analysis for Complex Environmental Models*; Watermark Numerical Computing: Brisbane, Australia, 2015; ISBN 978-0-9943786-0-6.
28. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **2001**, *106*, 7183–7192. [[CrossRef](#)]
29. Friedman, J.; Hastie, T.; Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **2010**, *33*, 1–22. [[CrossRef](#)]
30. Simon, N.; Friedman, J.H.; Hastie, T.; Tibshirani, R. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *J. Stat. Softw.* **2011**, *39*, 1–13. [[CrossRef](#)]
31. Tibshirani, R.; Bien, J.; Friedman, J.H.; Hastie, T.; Simon, N.; Taylor, J.; Tibshirani, R.J. Strong rules for discarding predictors in lasso-type problems. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2012**, *74*, 245–266. [[CrossRef](#)]
32. Gareth, J.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning: With Applications in R*; Springer: New York, NY, USA, 2013.

33. Zou, H.; Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2005**, *67*, 301–320. [CrossRef]
34. Jones, G.V. Climate and terroir: Impacts of climate variability and change on wine. In *Fine Wine and Terroir: The Geoscience Perspective*; Macqueen, R.W., Meinert, L.D., Eds.; Geoscience Canada Reprint Series No.; Geo-logical Association of Canada: St. John's, NL, Canada, 2006; p. 203.
35. Huglin, M. Nouveau Mode d'Évaluation des Possibilités Héliothermiques d'un Milieu Viticole. *C.R. Acad. Agr. Fr.* **1978**, *64*, 1117–1126.
36. Tonietto, J. Les macroclimats viticoles mondiaux et l'influence du mésoclimat sur la typicité de la Syrah et du Muscat de Hambourg dans le sud de la France: Méthodologie de caractérisation. Ph.D. Thesis, Ecole Nationale Supérieure Agronomique, Montpellier, France, 1999; p. 233.
37. Riou, C.; Becker, N.; Sotes Ruiz, V.; Gomez-Miguel, V.; Carbonneau, A.; Panagiotou, M.; Calo, A.; Costacurta, A.; de Castro, R.; Pinto, A.; et al. *Le Déterminisme Climatique de la Maturation du Raisin: Application au Zonage de la Teneur en Sucre Dans la Communauté Européenne*; Office des Publications Officielles des Communautés Européennes: Luxembourg, 1994; p. 322.
38. Tonietto, J.; Carbonneau, A. A multicriteria climatic classification system for grape-growing regions worldwide. *Agric. For. Meteorol.* **2004**, *124*, 81–97. [CrossRef]
39. Huglin, P.; Schneider, C. *Biologie et Écologie de la Vigne*; Lavoisier: Paris, France, 1998.
40. Pierce, D.W.; Cayan, D.R.; Thrasher, B. Statistical Downscaling Using Localized Constructed Analogs (LOCA)\*. *J. Hydrometeorol.* **2014**, *15*, 2558–2585. [CrossRef]
41. Pierce, D.W.; Cayan, D.R.; Maurer, E.P.; Abatzoglou, J.; Hegewisch, K.C. Improved Bias Correction Techniques for Hydrological Simulations of Climate Change\*. *J. Hydrometeorol.* **2015**, *16*, 2421–2442. [CrossRef]
42. Livneh, B.; Rosenberg, E.A.; Lin, C.; Nijssen, B.; Mishra, V.; Andreadis, K.M.; Maurer, E.; Lettenmaier, D.P. A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions. *J. Clim.* **2013**, *26*, 9384–9392. [CrossRef]
43. Wine Spectator. Vintage Charts. Oregon: Pinot Noir. Available online: <https://www.winespectator.com/vintage-charts/region/oregon> (accessed on 1 December 2020).
44. Schultze, S.R.; Sabbatini, P. Implications of a Climate-Changed Atmosphere on Cool-Climate Viticulture. *J. Appl. Meteorol. Clim.* **2019**, *58*, 1141–1153. [CrossRef]
45. Blanco-Ward, D.; Ribeiro, A.; Barreales, D.; Castro, J.; Verdial, J.; Feliciano, M.; Viceto, C.; Rocha, A.; Carlos, C.; Silveira, C.; et al. Climate change potential effects on grapevine bioclimatic indices: A case study for the Portuguese demarcated Douro Region (Portugal). *BIO Web Conf.* **2019**, *12*, 01013. [CrossRef]
46. Cabré, F.; Nuñez, M. Impacts of climate change on viticulture in Argentina. *Reg. Environ. Chang.* **2020**, *20*, 1–12. [CrossRef]
47. Cardell, M.F.; Amengual, A.; Romero, R. Future effects of climate change on the suitability of wine grape production across Europe. *Reg. Environ. Chang.* **2019**, *19*, 2299–2310. [CrossRef]
48. Mihai, I.L.; Valeriu, P.C.; Renan, L.; Herve, Q.; Cyril, T.; Lucian, S. Projections of Climate Suitability for Wine Production for the Cotnari Wine Region (Romania). *Present. Environ. Sustain. Dev.* **2019**, *13*, 5–18. [CrossRef]
49. Mavromatis, T.; Koufos, G.C.; Koundouras, S.; Jones, G.V. Adaptive capacity of winegrape varieties cultivated in Greece to climate change: Current trends and future projections. *OENO One* **2020**, *54*, 1201–1219. [CrossRef]
50. Santos, M.; Fonseca, A.; Fraga, H.; Jones, G.V.; Santos, J.A. Bioclimatic conditions of the Portuguese wine denominations of origin under changing climates. *Int. J. Clim.* **2020**, *40*, 927–941. [CrossRef]
51. Sirnik, I. Spatial-temporal analysis of climate change impact on viticultural regions Valencia DO and Goriška Brda. Ph.D. Thesis, Universitat Politècnica València, Valencia, Spain, 2019.
52. Teslić, N.; Vujadinović, M.; Ruml, M.; Ricci, A.; Vuković, A.; Parpinello, G.P.; Versari, A. Future climatic suitability of the Emilia-Romagna (Italy) region for grape production. *Reg. Environ. Chang.* **2019**, *19*, 599–614. [CrossRef]
53. Trbic, G.; Djurdjevic, V.I.; Mandic, M.V.; Ivanisevic, M.; Cupac, R.; Bajic, D.; Zahirovic, E.; Filipovic, D.; Dekic, R.; Popov, T.; et al. The impact of climate change on grapevines in Bosnia and Herzegovina. *Euro-Mediterr. J. Environ. Integr.* **2021**, *6*, 1–9. [CrossRef]
54. Oregon Wine Board. 2019 Oregon Vineyard and Winery Report. Available online: <https://industry.oregonwine.org/resources/reports-studies/2019-oregon-vineyard-and-winery-report/> (accessed on 2 September 2020).
55. Vano, J.; Hamman, J.; Gutmann, E.; Wood, A.; Mizukami, N.; Clark, M.; Pierce, D.W.; Cayan, D.R.; Wobus, C.; Nowak, K.; et al. Comparing Downscaled LOCA and BCSD CMIP5 Climate and Hydrology Projections-Release of Downscaled LOCA CMIP5 Hydrology. Available online: [https://gdo-dcp.ucllnl.org/downscaled\\_cmip\\_projections/techmemo/LOCA\\_BCSD\\_hydrology\\_tech\\_memo.pdf](https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/techmemo/LOCA_BCSD_hydrology_tech_memo.pdf) (accessed on 2 December 2020).
56. Shin, Y.; Lee, Y.; Park, J.-S. A Weighting Scheme in A Multi-Model Ensemble for Bias-Corrected Climate Simulation. *Atmosphere* **2020**, *11*, 775. [CrossRef]
57. Hoerl, A.E.; Kennard, R.W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970**, *12*, 55–67. [CrossRef]
58. Tikhonov, A.N. On the stability of inverse problems. *Dokl. Akad. Nauk. SSSR* **1943**, *39*, 195–198.
59. Tibshirani, R. Regression Shrinkage and Selection Via the Lasso. *J. R. Stat. Soc. Ser. B* **1996**, *58*, 267–288. [CrossRef]
60. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Prediction, Inference and Data Mining*, 2nd ed.; Springer: New York, NY, USA, 2009.

- 
61. Raftery, A.E.; Gneiting, T.; Balabdaoui, F.; Polakowski, M. Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *Mon. Weather. Rev.* **2005**, *133*, 1155–1174. [[CrossRef](#)]
  62. Hastie, T.; Qian, J. An Introduction to Glmnet. Available online: <https://glmnet.stanford.edu/articles/glmnet.html> (accessed on 23 January 2021).
  63. Doherty, J.; Welter, D. A short exploration of structural noise. *Water Resour. Res.* **2010**, *46*. [[CrossRef](#)]
  64. Parker, A.K.; de Cortázar-Atauri, I.G.; Géný, L.; Spring, J.-L.; Destrac, A.; Schultz, H.; Molitor, D.; Lacombe, T.; Graça, A.; Monamy, C.; et al. Temperature-based grapevine sugar ripeness modelling for a wide range of *Vitis vinifera* L. cultivars. *Agric. For. Meteorol.* **2020**, *285–286*, 107902. [[CrossRef](#)]
  65. Giorgi, F.; Mearns, L.O. Calculation of Average, Uncertainty Range, and Reliability of Regional Climate Changes from AOGCM Simulations via the “Reliability Ensemble Averaging” (REA) Method. *J. Clim.* **2002**, *15*, 1141–1158. [[CrossRef](#)]