Article

Application of High-Resolution Regional Climate Model Simulations for Crop Yield Estimation in Southern Brazil

Santiago Vianna Cuadra [1,*], Monique Pires Gravina de Oliveira [1], Daniel de Castro Victoria [1], Fabiani Denise Bender [1], Maria L. Bettolli [2], Silvina Solman [3], Rosmeri Porfírio da Rocha [4], Jesús Fernández [5], Josipa Milovac [5], Erika Coppola [6] and Moira Doyle [7]

1 EMBRAPA Digital Agriculture, Campinas 13083-886, SP, Brazil; moniquepgoliveira@gmail.com (M.P.G.d.O.); daniel.victoria@embrapa.br (D.d.C.V.); fabianidenise@gmail.com (F.D.B.)

2 Departamento de Ciencias de la Atmósfera y los Océanos, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Instituto Franco-Argentino para el Estudio del Clima y sus Impactos (IRL 3351 IFAECI), CNRS-IRD-CONICET-UBA, Ciudad Universitaria Pabellón II Piso 2, Ciudad Autónoma de Buenos Aires C1428EGA, Argentina; bettolli@at.fcen.uba.ar

3 Departamento de Ciencias de la Atmósfera y los Océanos, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Centro de Investigaciones del Mar y la Atmósfera (CIMA), CONICET-Universidad de Buenos Aires, Instituto Franco-Argentino para el Estudio del Clima y sus Impactos (IRL 3351 IFAECI), CNRS-IRD-CONICET-UBA, Ciudad Universitaria Pabellón II Piso 2, Ciudad Autónoma de Buenos Aires C1428EGA, Argentina; solman@cima.fcen.uba.ar

4 Departamento de Ciências Atmosféricas, Universidade de São Paulo, São Paulo 03178-200, SP, Brazil; rosmerir.rocha@iag.usp.br

5 Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, E-39005 Santander, Spain; jesus.fernandez@unican.es (J.F.); milovacj@unican.es (J.M.)

6 International Centre for Theoretical Physics, 34151 Trieste, Italy; coppolae@ictp.it

7 Centro de Investigaciones del Mar y la Atmósfera (CIMA), CONICET—Departamento de Ciencias de la Atmósfera y los Océanos, National University of La Plata, La Plata C1428, Argentina; doyle@cima.fcen.uba.ar

* Correspondence: santiago.cuadra@embrapa.br

Abstract: This study is focused on assessing the impacts of different regional climate model targeted simulations performed at convection-permitting resolution (CPRCM) in the AgS crop model yield simulations, evaluating to what extent climate model uncertainty impacts the modeled yield—considering the spatial and temporal variability of crop yield simulations over central-south Brazil. The ensemble of CPRCMs has been produced as part of a Flagship Pilot Study (FPS-SESA) framework, endorsed by the Coordinated Regional Climate Downscaling Experiment (CORDEX). The AgS simulated crop yield exhibited significant differences, in both space and time, among the simulations driven by the different CPRCMs as well as when compared with the simulations driven by observations. Rainfall showed the highest uncertainty in CPRCM simulations, particularly in its spatial variability, whereas modeled temperature and solar radiation were generally more accurate and exhibited smaller spatial and temporal differences. The results evidenced the need for multi-model simulations to account for different uncertainty, from different climate models and climate models parameterizations, in crop yield estimations. Inter-institutional collaboration and coordinated science are key aspects to address these end-to-end studies in South America, since there is no single institution able to produce such CPRCM-CropModels ensembles.

Keywords: regional climate model; convection permitting; crop growth model; soybean; maize

1. Introduction

Agriculture production is central to the Brazilian economy, with estimates indicating that up to 20% of the Gross Domestic Product (GDP) is related to the agricultural sector. In 2022, Brazil was the world leader in Soybean production (120 M tons) and the third largest Maize producer (109 M tons) [1]. Although production has consistently increased in recent decades, production variability is frequent and is mostly related to climate variability, especially dry spells that affect major soy-producing states. In 2012, a major drought reduced soy yield in Paraná to 2307 kg/ha in contrast to 3973 kg/ha in the previous year. Similar reductions were observed in Santa Catarina (from 3259 kg/ha in 2011 to 2393 kg/ha in 2012) and Rio Grande do Sul (from 2876 kg/ha in 2011 to 1430 kg/ha in 2012) [2]. At the end of 2015, another drought hit central Brazil, reducing second season maize yields in the states of Mato Grosso do Sul (3593 kg/ha in 2016 against 5785 kg/ha in the previous year), Mato Grosso (4106 kg/ha in 2016 vs. 5980 kg/ha in 2015), Goiás (4339 kg/ha in 2016 vs. 6786 kg/ha in 2015), and the Federal District (4168 kg/ha in 2016 vs. 8098 kg/ha in 2015) [2]. Thus, a good understanding of climatic variations and weather characteristics are critical for understanding crop production variation and risk assessment evaluation. In turn, those are central to food security, farmers' economic stability, and the overall country economy, being of great interest to both private and governmental sectors.

The major instrument to mitigate economic losses is insurance. In Brazil, there are two major insurance programs for the agricultural sector, Proagro, a direct insurance provided by the Brazilian Central Bank for small farmers, and PSR, a subvention program that provides financial support for the acquisition of private insurance, focused on medium to large producers. During the last decade, agricultural losses due to adverse weather have raised the insurance loss ratio (insurance paid/earned premium), reaching 90% in 2015, 126% in 2021, and 153% in 2022. The average loss ratio from 2014 to 2023 is above 90% [3], posing a challenge for the crop insurance sector.

In this context, crop models are valuable tools for understanding the impacts of climate variability on agricultural failures. Being weather-dependent, daily weather inputs (e.g., rainfall, temperature, and solar radiation) are key drivers of crop growth and significant sources of uncertainty [4,5]. Therefore, combining seasonal climate predictions with crop yield simulations offers a powerful tool to support both private and governmental sectors in planning, prior to and during the crop season. Moreover, high-resolution climate data produced with regional climate models (RCMs) may be particularly preferable, given their added value in the representation of climate variables compared with coarser global climate models [6]. Hence, high-resolution RCMs are expected to capture the spatial and temporal evolution of key climate variables and may therefore be used for providing reliable inputs to crop yield simulations and, therefore, building reliable yield products [7,8]. Additionally, using an ensemble of RCMs helps address the downscaling uncertainties associated with climate data inputs for crop models [9,10]. Projections of climate change impacts are also essential to guide governmental planning for climate change mitigation and adaptation [11,12].

Several studies have shown that climate simulations produced by RCMs operating at resolutions of tens of kilometers (e.g., 50 to 20 km) improve the representation of several regional climatic features compared with the global climate models that provide the boundary driving data. This has been shown for any region in the world and particularly for South America [13,14]. In the last decade, it has been demonstrated that RCMs operating at resolutions of a few kilometers that explicitly resolve convective processes can capture several features of daily and sub daily precipitation statistics compared with coarser RCMs [15–17]. This so-called convection permitting regional climate models (CPRCMs) was first implemented over several regions in the world (UK, Europe, Australia, among

others) and it has been demonstrated that coordinated ensembles of CPRCMs are able to reduce the uncertainty of future climate projections [15,18], among others. In the framework of the CORDEX-FPS-SESA [19], an unprecedented effort allowed the production of an ensemble of coordinated CPRCM simulations for subtropical South America. A preliminary analysis of this ensemble has shown that it has the ability of capturing the sub-daily statistics of extreme precipitation that were not captured by coarser RCMs [19]. However, these simulations were not evaluated from the perspective of applications for the agriculture sector, probably the major economic activity over this region.

This study aims at evaluating crop yield simulations driven by the coordinated set of CPRCM simulations performed in the framework of the CORDEX FPS-SESA for assessing its impact in capturing the spatio-temporal variability of crop yields over south Brazil. We present the results and discuss the simulation aggregation process and the impacts of CPRCM biases in different variables on the crop yield simulations.

2. Material and Methods

2.1. The CORDEX FPS-Southeastern South America (SESA) CPRCM Simulations

The CORDEX Flagship Pilot Study—Southeastern South America (FPS-SESA) initiative aims at improving the capacity of simulating extreme precipitation events in SESA with a coordinated approach involving convection permitting RCMs and empirical statistical downscaling (ESD) approaches with the goal of producing actionable climate information for assessing the impact of adverse climate conditions on the agricultural production and hydrology on the region [19]. The initiative also pursues promoting inter-institutional collaborations among interdisciplinary research groups from a suite of institutions from South America and Europe.

Under the CORDEX FPS-SESA, a series of coordinated simulations with RCMs operating at km-scale resolution were performed, driven by the ERA5 reanalysis. The simulations spanned a 3-year-long period that was chosen based on the most impactful conditions for crops in the region. Hence, the 2018–2021 period, was selected since it was characterized by unprecedented dry conditions and anomalous high temperatures over the area, two conditions that have a strong impact on crops, but at the same time several extreme precipitation events occurred [6]. The CPRCM simulations performed in the framework of the FPS-SESA were configured at a resolution of 4 km, covering a common domain displayed in Figure 1. The simulations available for driving the crop model are summarized in Table 1. Two versions of the regional climate model RegCM4 [20] and RegCM5 [21] were used. Additionally, the RegCM5 simulations were performed using two different planetary boundary layer (PBL) schemes. The RegCM5-ICTP-pbl1 uses the Holtslag scheme and RegCM5-ICTP-pbl2 uses the UW PBL scheme. The major difference among these two PBL schemes is that the Holstag scheme is characterized by a local closure while the UW scheme has a non-local closure. Local and non-local closures impact on the vertical mixing within the PBL mostly for unstable conditions and, therefore, impact on the simulated near surface temperature. Additional CPRCM simulations were performed by two versions of the WRF model [version 4.3.3 and 4.1.5] [22], one of them covering the whole South American domain and executed by the National Center for Atmospheric Research (WRF-NCAR) and another covering the FPS-SESA domain (Figure 1) and executed by the University of Cantabria (WRF-UCAN); a similar physical parameterization setting was used by UCAN and NCAR, except for the microphysics and shallow cumulus scheme.
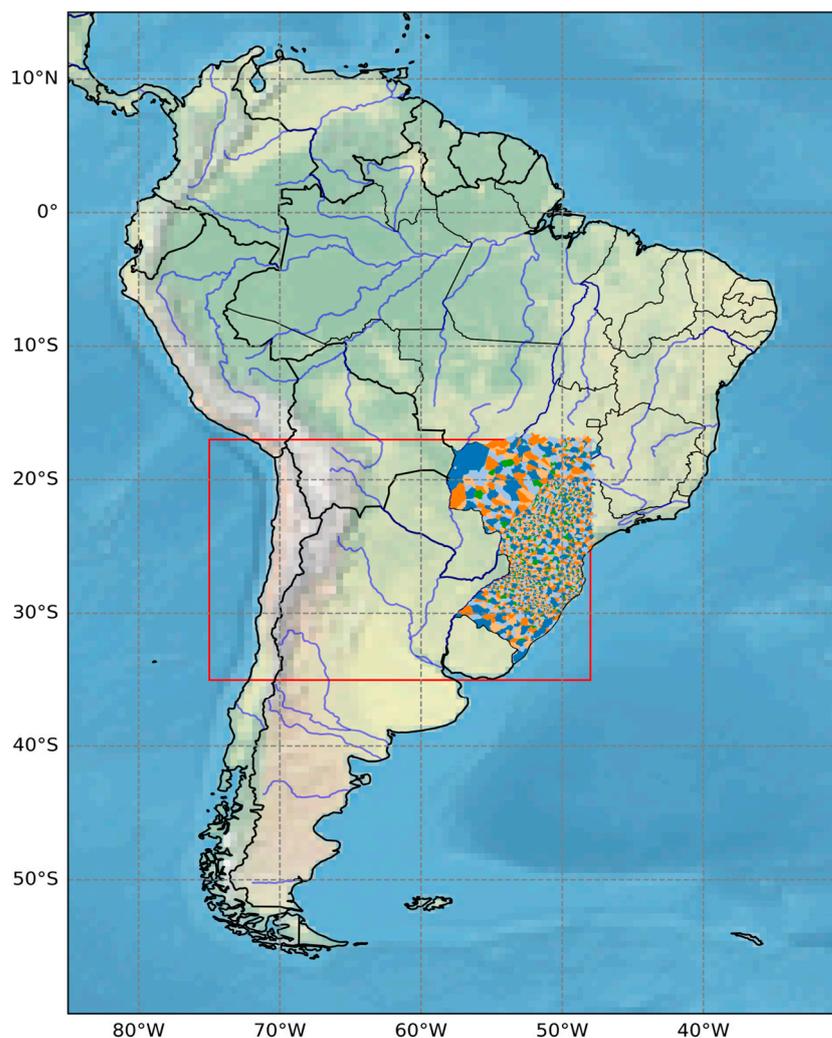
Figure 1. Region of study: minimum domain covered by the CPRCM simulations (red rectangle) conducted in the FPS-SESA initiative and counties with yield data over Brazil (colored).

Table 1. List of CPRCM simulations used for driving the crop models.

Climate Data/Model	Institution	Label	Reference
NASA POWER/GPM	NASA	BASE	https://power.larc.nasa.gov * https://gpm.nasa.gov/ *
RegCM5	ICTP (Italy)	RegCM5-ICTP-pbl1	Giorgi et al. [21]
RegCM5	ICTP (Italy)	RegCM5-ICTP-pbl2	Giorgi et al. [21]
RegCM4	USP (Brazil)	RegCM4-USP	Giorgi et al. [21]
WRF433	UCAN (Spain)	WRF-UCAN	Skamarock et al. [22]
WRF415	NCAR (USA)	WRF-NCAR	Skamarock et al. [22]

* URL (accessed on 2 March 2025).

2.2. Crop Model

2.2.1. AgS

The AgS (Agricultural Crop Simulator) is a biophysical crop model simulation platform that computes the biophysical and physiological processes of development, growth, and yield of various agricultural crops, associated with a set of equations to represent water flows in the soil. In general, the AgS simulates daily biomass accumulation based on the balance between photosynthesis and the maintenance and growth respiration of each plant organ. This balance is affected by each crop's morphological characteristics and by the de-

velopment and growth responses to carbon dioxide concentration in the air, air temperature, solar radiation interception, evaporative demand, and water availability. Photoassimilates' translocation depends on the plant development stage, vegetative or reproductive, and on growth conditions, such as air temperature and water stress. For example, sugarcane increases the allocation of carbon in the form of carbohydrates under conditions of low temperatures or water deficiency. The capacity to produce photoassimilates is a function of canopy growth, which in turn is a function of phenological development, leaf area, photosynthetic efficiency of the species and water status of the crop. Final yield is the result of the accumulation of biomass in grains and associated structures, starting when a predefined thermal time to reproductive stage is reached.

The state variables defined in the AgS represent the morphological characteristics of plants, so that their evolution characterizes plant growth and development. Leaf area index, fraction of solar radiation intercepted, maintenance respiration, root system depth, and biomass are represented. Biomass in the model is divided into green and dead leaves, thin and thick roots, reserve carbohydrates, stem, and grains and their associated structures, such as pods and ears. The variables that represent biomass are connected to water availability in the soil through evapotranspiration which, in turn, interacts with the physical properties of the soil profile, water balance, and root depth, these two also simulated by the model.

2.2.2. Soybean and Maize Parametrizations

Three datasets were used for the parameterization/evaluation, and calibration/evaluation of the maize and soybean models. The model was calibrated with flux tower experiment datasets for species parameters associated with photosynthesis, evapotranspiration, and stress effects on canopy and growth. Cultivar-specific parameters were determined in the calibration of generic cultivars for each crop. For soybean, these generic cultivars were grouped by growth habit—one group with determinate growth and one with an indeterminate or semi-determinate growth habit—and three maturity groups—early, medium, and late, leading to six generic cultivars. For maize, only maturity groups were relevant, leading to three generic cultivars. Data used for calibration were obtained in the rainfed breeding experiments across Brazil and included as variables cycle length, flowering date, and yield. For maize, cycle length was estimated based on maturity date. These observations were used to determine phenology, canopy-specific, and yield-related parameters. Calibration was performed by brute force optimization and the parameters obtained were evaluated in a simple holdout approach to ensure their generalization capabilities, Figure A1 presents the results for the AgS simulations used for the model parameterization. Soil properties were locally determined and weather data were obtained from local meteorological stations, except from breeding data, which used a generic high-water holding capacity soil profile and NASA POWER/GPM data.

2.3. AgS Simulation Settings and Data

AgS uses soil properties, planting date, species/cultivar planted, rainfall, solar radiation, and minimum and maximum temperature to simulate crop growth and development on a daily time step.

The soil physical and hydraulic properties used by the model are the soil total porosity (volume fraction), field capacity (volume fraction), wilting point (volume fraction), saturated hydraulic conductivity (m s^{-1}), and granulometric fractions, for each soil layer. When the saturated upper limit, drained upper limit and lower limit of plant extractable water were not available, a pedotransfer function was used [23]. For the saturated hydraulic conductivity estimation, the model uses the functions proposed by Tomasella and Hodnett [24].

For soybean, six simulations were run for each grid point, with three sowing dates (1 October, 1 November, and 1 December) and three growing crop seasons (sowing in 2018, 2019, and 2020). For maize, six simulations were also run for each grid point, three sowing dates (1 February, 1 March, and 1 April) and three crop seasons (sowing in 2019, 2020, and 2021). To account for soil water content initialization, all simulations started 180 days before the planting date.

Daily meteorological data were derived from National Aeronautics and Space Administration Prediction Worldwide Energy Resources (NASA POWER, https://power.larc.nasa.gov/, accessed on 2 March 2025), for maximum and minimum temperature, and solar radiation, and global precipitation measurement (GPM, https://gpm.nasa.gov/, accessed on 2 March 2025) for precipitation. These satellite-based products provide reliable estimates that are consistent with in situ data for Brazil [25–27], being considered as the base simulations—that are used as the references for evaluating AgS simulations that utilize RCMs (described in Section 2.1). We present the AgS simulations derived from daily climate variables provided by (see Table 1 and Figure A2) (i) BASE, NASAPOWER and GPM observations; (ii) WRF-NCAR, from WRF-NCAR simulation; (iii) WRF-UCAN, from WRF-UCAN simulations; and (iv) RegCM5-ICTP-pbl1, from RegCM5-ICTP-pbl1 simulations; (v) RegCM5-ICTP-pbl2, from RegCM5-ICTP-pbl2 simulations; and (vi) RegCM4-USP, from RegCM4-USP simulations.

2.4. Crop Yield Simulation Evaluations

Root mean square error (*RMSE*; Equation (1)) and Mean Error (*ME*, Equation (2)) were used to evaluate the simulations (three sowing dates and three growing season) of crop yield (yield), solar radiation (SRAD), daily average temperature (TEMP), and precipitation (PREC).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(s_i - o_i)^2}{n}} \tag{1}$$

$$ME = \frac{\sum_{i}^{n}(s_i - o_i)}{n} \tag{2}$$

where o_i are the observed values, i.e., values obtained with BASE data; s_i are the simulated values using the evaluated data sources (it is worth noting that the meteorological variables are directly simulated by the CPRCMs and yield is simulated by the AgS model); n is the number of observations. Additionally, following Doi et al. [28], we calculate the anomaly correlation coefficients (*ACCs*), over space and time, as a deterministic accuracy score as follows.

$$ACC = \frac{\sum_{i}^{n}(D_{s,i}D_{o,i})}{\sqrt{\frac{1}{n^2}D_{s,i}{}^2 D_{o,i}{}^2}} \tag{3}$$

where D_s and D_o are defined as follows

$$D_{s,i} = \frac{\sum_{i}^{n}(s_i - \hat{s}_i)}{\hat{s}_i} \tag{4}$$

$$D_{o,i} = \frac{\sum_{i}^{n}(o_i - \hat{o}_i)}{\hat{o}_i} \tag{5}$$

All analyses were performed using the R Core Team (2020).

2.5. Effect of the Number of Grid Points

We consider daily climate data time series for the BASE simulations using a $0.1° \times 0.1°$ horizontal grid, i.e., the original grids of the GPM precipitation; NASAPOWER, the source for air temperature and solar radiation, has a half degree resolution. In the assessment

described in Section 2.4, for the BASE simulations, we had a total of 11,217 grid points, covering the counties shown in Figure 1. The RCM simulations were interpolated to a common grid of 0.04° × 0.04°, totaling 71,351 grid points. Considering that we run nine simulations for soybean (three sowing dates and three crop seasons) and nine for maize, the total number of simulations were 201,906 for the BASE simulations and 6,421,590 for the RCMs simulations. Many of these simulations were averaged, i.e., all grid points where the grid center was inside the county, to determine the yield associated with each county.

We therefore also evaluate the impacts of reducing the number of grid points to determine county yield estimates. In each county, we randomly subsampled the simulations associated with its grid points from 1 to 40 or its maximum number of grid points. We repeated the procedure 150 times and each repetition (j) resulted in a yield estimate for the county ($Y_{i,j}$) averaged from all subsampled points (i). We calculated the average absolute difference (Z_i) between the yield estimated with the subsampled number of points, for the respective climate data source, and using all points (Y_{max}) over all repetitions (Equation (6)). This assumes that the simulation with the maximum resolution represents the minimum error achievable by the model and aims at evaluating the loss in accuracy. We also calculated the standard deviation of these absolute differences across repetitions to determine the uncertainty associated with the choice of grid points. Additionally, for each number of points, we determined the difference between the average yield from the different data sources and the average yield of BASE simulations ($X_{i,j}$) (Equation (7)).

$$\underline{Z}_i = \frac{\sum_{j=1}^{150}\left|Y_{i,j} - Y_{max}\right|}{150} \tag{6}$$

$$\underline{E}_i = \frac{\sum_{j=1}^{150}\left|Y_{i,j} - X_{i,j}\right|}{150} \tag{7}$$

3. Results

3.1. Impacts of Regional Climate Data on Crop Yield Simulations

Our results suggested that the impact associated with regional climate data differs by the regional model, meteorological variable, and crop (Table 2 and Figures 2–7). AgS simulations for soybean yield using daily data from CPRCMs (Figure 2b–f) show a similar pattern to yields achieved by running AgS with BASE climate (Figure 2a), except for the simulations with RegCM5-ICTP_pbl2 (Figure 2e). Comparatively, WRF-NCAR reaches the highest *ACC* (0.82, Table 2) calculated over space—i.e., *ACC* calculated with the fields plotted in Figure 2. All models present average deviations lower than 10% (Table 2, MB), except for the simulations with RegCM5-ICTP-pbl2 that presented *MB* of −20% (−547 kg ha⁻¹).

Table 2. Simulation's statistics based on AgS soybean (SB) and maize simulations (MZ): Mean (*M*); Mean bias (*MB*); Root mean square error (*RMSE*); Anomaly correlation coefficients over average field (*ACCs*) and time (*ACCt*) for averaged growing season rainfall (RAIN, mm m⁻²), daily air temperature (TEMP, °C), daily incident radiation (SRAD, MJ m⁻² day⁻¹), and crop yield (YIELD, kg ha⁻¹).

Metric	Experiment	RAIN (SB)	TEMP (SB)	SRAD (SB)	YIELD (SB)	RAIN (MZ)	TEMP(MZ)	SRAD (MZ)	YIELD (MZ)
M	BASE	609	24.52	21.85	2734	423	21.83	16.99	2411
M	WRF-NCAR	650	24.47	24.66	2868	379	21.51	18.76	1992
M	WRF-UCAN	846	25.37	23.00	2970	536	21.52	17.16	2695
M	RegCM5-ICTP-pbl1	602	25.21	22.27	2627	395	22.43	17.63	2532
M	RegCM5-ICTP-pbl2	451	25.10	24.45	2187	292	21.97	18.82	1663
M	RegCM4-USP	642	25.50	23.21	2751	406	22.94	17.87	2609

Table 2. *Cont.*

| Metric	Experiment	RAIN (SB)	TEMP (SB)	SRAD (SB)	YIELD (SB)	RAIN (MZ)	TEMP(MZ)	SRAD (MZ)	YIELD (MZ)
MB	NCAR_WRF	41	−0.05	2.81	134	−45	−0.32	1.77	−419
MB	UCAN_WRF	236	0.85	1.15	236	112	−0.31	0.17	284
MB	ICTP_RegCM5_pbl1	−8	0.69	0.42	−107	−28	0.60	0.64	120
MB	ICTP_RegCM5_pbl2	−158	0.58	2.59	−547	−131	0.14	1.83	−748
MB	USP_RegCM4	33	0.98	1.35	17	−18	1.11	0.89	198
RMSE	NCAR_WRF	176	0.77	2.93	374	115	0.80	1.82	622
RMSE	UCAN_WRF	284	1.17	1.32	408	145	0.75	0.43	476
RMSE	ICTP_RegCM5_pbl1	193	1.00	0.85	431	149	0.93	0.83	642
RMSE	ICTP_RegCM5_pbl2	202	0.98	2.67	679	163	0.82	1.87	945
RMSE	USP_RegCM4	200	1.22	1.61	380	142	1.33	1.02	794
ACCs	NCAR_WRF	0.26	0.95	0.80	0.82	0.40	0.97	0.98	0.66
ACCs	UCAN_WRF	0.31	0.94	0.87	0.58	0.61	0.96	0.96	0.67
ACCs	ICTP_RegCM5_pbl1	0.07	0.94	0.84	0.41	0.35	0.96	0.94	0.30
ACCs	ICTP_RegCM5_pbl2	0.19	0.95	0.81	0.43	0.51	0.96	0.97	0.40
ACCs	USP_RegCM4	0.18	0.94	0.81	0.56	0.41	0.96	0.94	0.11
ACCt	NCAR_WRF	0.57	0.64	0.85	0.66	0.75	0.93	0.96	0.77
ACCt	UCAN_WRF	0.61	0.50	0.73	0.46	0.74	0.75	0.85	0.71
ACCt	ICTP_RegCM5_pbl1	0.46	0.60	0.69	0.37	0.74	0.93	0.98	0.76
ACCt	ICTP_RegCM5_pbl2	0.61	0.53	0.77	0.47	0.82	0.93	0.98	0.77
ACCt	USP_RegCM4	0.50	0.58	0.44	0.46	0.77	0.92	0.98	0.75
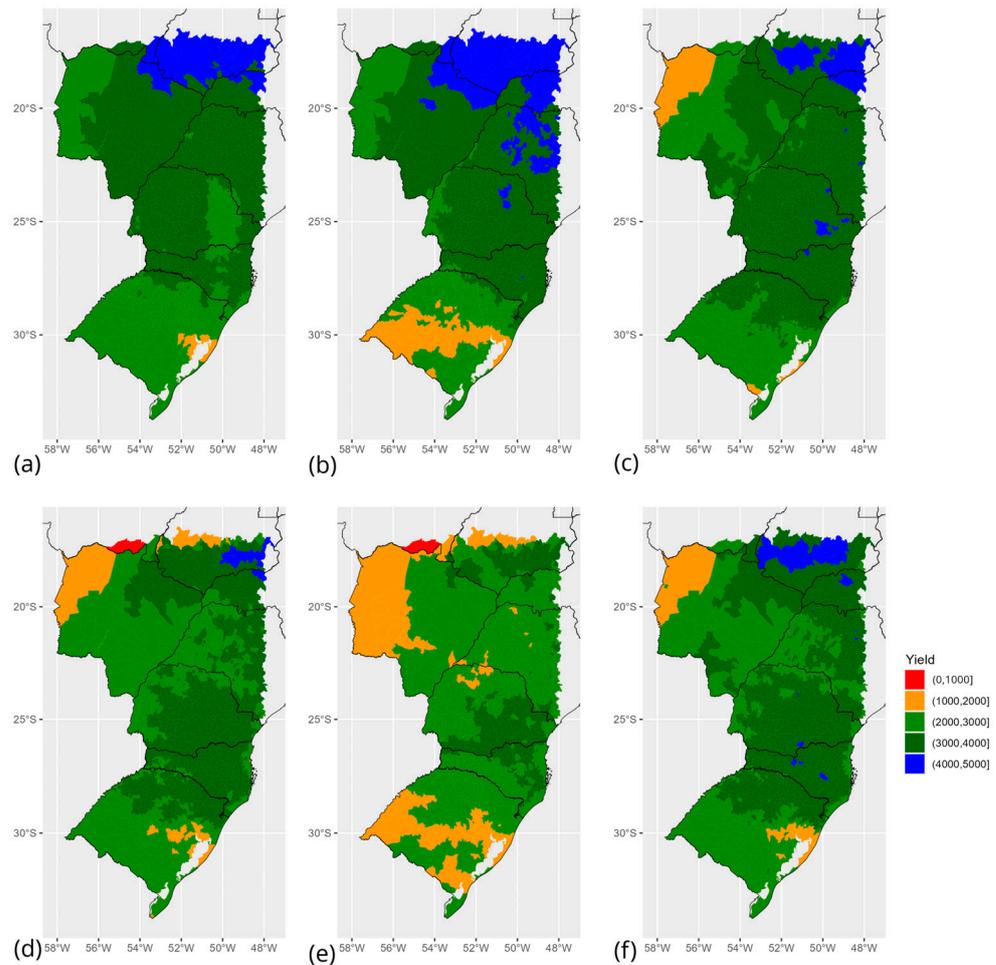
Figure 2. Average soybean yield (kg ha^{-1}) across three sowing dates (1 October, 1 November, and 1 December) and three growing seasons (sowing in 2018, 2019, and 2020) as simulated by AgS crop model with daily climate from (**a**) BASE, (**b**) WRF-NCAR, (**c**) WRF-UCAN, (**d**) RegCM5-ICTP-pbl1, (**e**) RegCM5-ICTP-pbl2, and (**f**) RegCM4-USP.
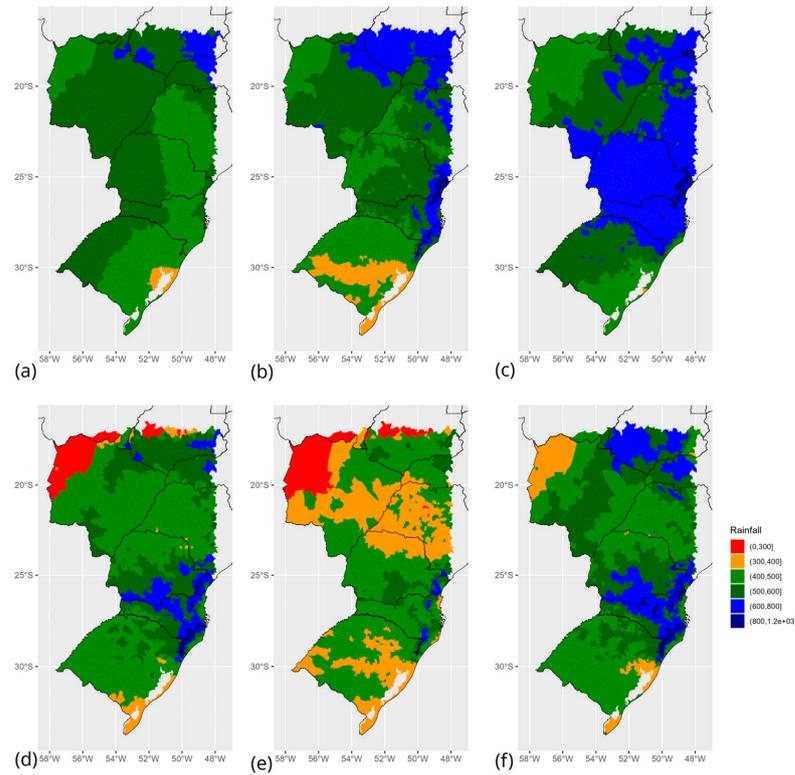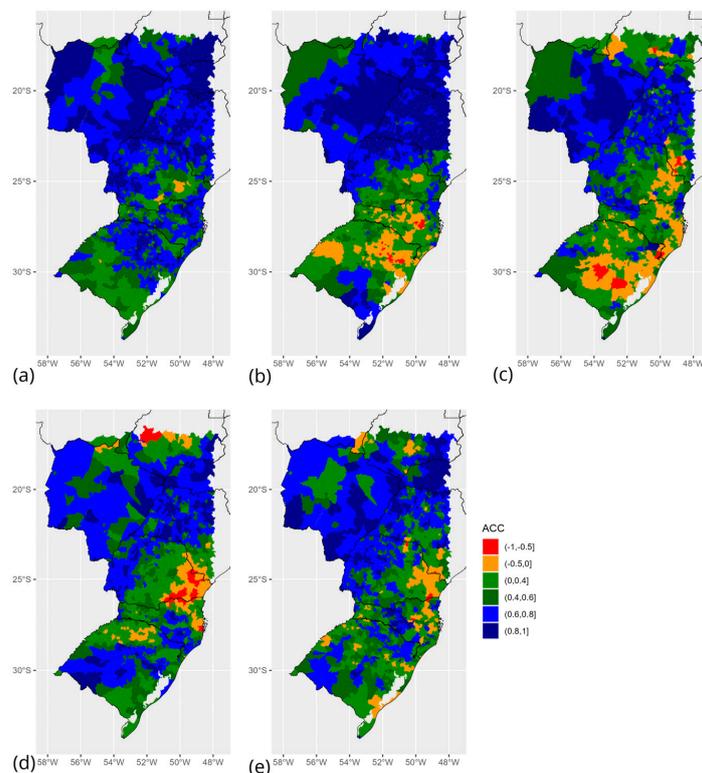
Figure 3. Average total rainfall (mm m^{-2}) during the soybean growth season, from sowing to harvest dates, across three sowing dates (1 October, 1 November, and 1 December) and three growing seasons (sowing in 2018, 2019, and 2020), as simulated by the CPRCMs models: (**a**) BASE, (**b**) WRF-NCAR, (**c**) WRF-UCAN, (**d**) RegCM5-ICTP-pbl1, (**e**) RegCM5-ICTP-pbl2, and (**f**) RegCM4-USP.
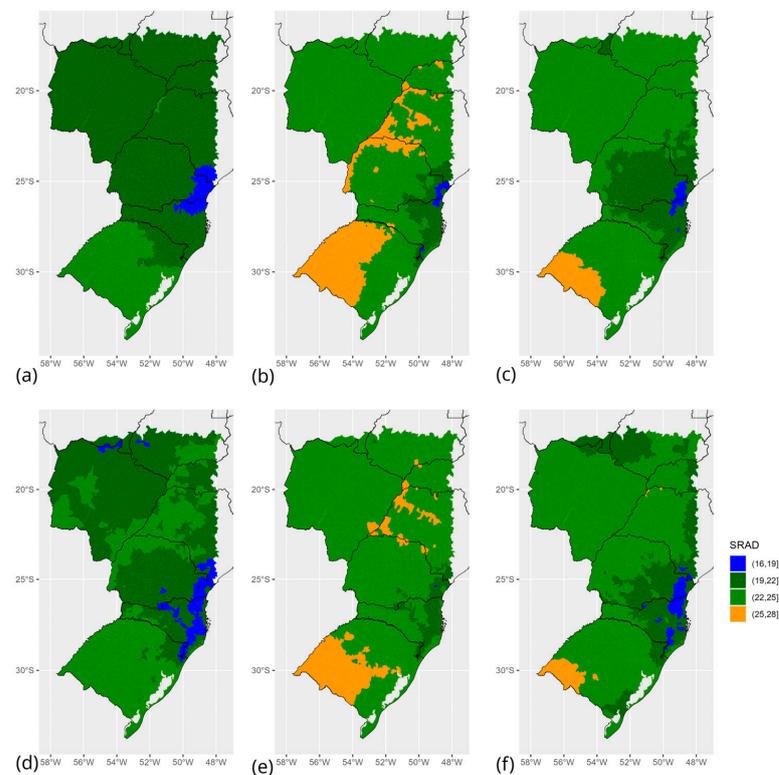

Figure 4. Average daily incident solar radiation (MJ m^{-2} day^{-1}) during the soybean growth season, from sowing to harvest dates, across three sowing dates (1 October, 1 November, and 1 December) and three crop seasons (sowing in 2018, 2019, and 2020) as simulated by the CPRCMs: (**a**) BASE, (**b**) WRF-NCAR, (**c**) WRF-UCAN, (**d**) RegCM5-ICTP-pbl1, (**e**) RegCM5-ICTP-pbl2, and (**f**) RegCM4-USP.
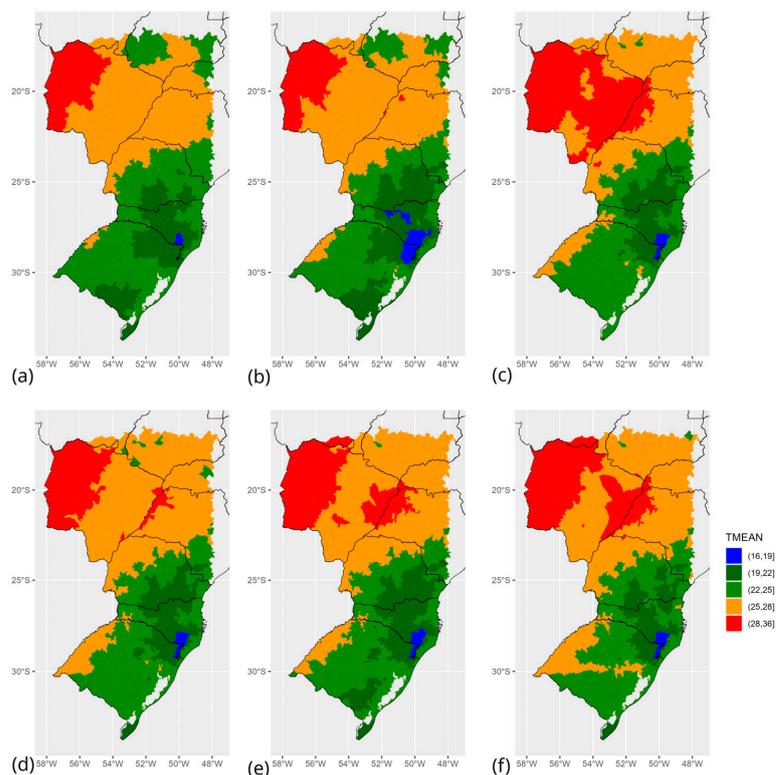
Figure 5. Average daily air temperature (°C) during the soybean growth season, from sowing to harvest dates, across three sowing dates (1 October, 1 November, and 1 December) and three crop seasons (sowing in 2018, 2019, and 2020) as simulated the CPRCMs: (**a**) BASE, (**b**) WRF-NCAR, (**c**) WRF-UCAN, (**d**) RegCM5-ICTP-pbl1, (**e**) RegCM5-ICTP-pbl2, and (**f**) RegCM5-USP.

Figure 6. Average soybean yield anomaly correlation coefficients (*ACCt*) over time across three sowing dates (1 October, 1 November, and 1 December) and three growing seasons (sowing in 2018, 2019, and 2020): (**a**) WRF-NCAR, (**b**) WRF-UCAN, (**c**) RegCM5-ICTP-pbl1, (**d**) RegCM5-ICTP-pbl2, and (**e**) RegCM5-USP.
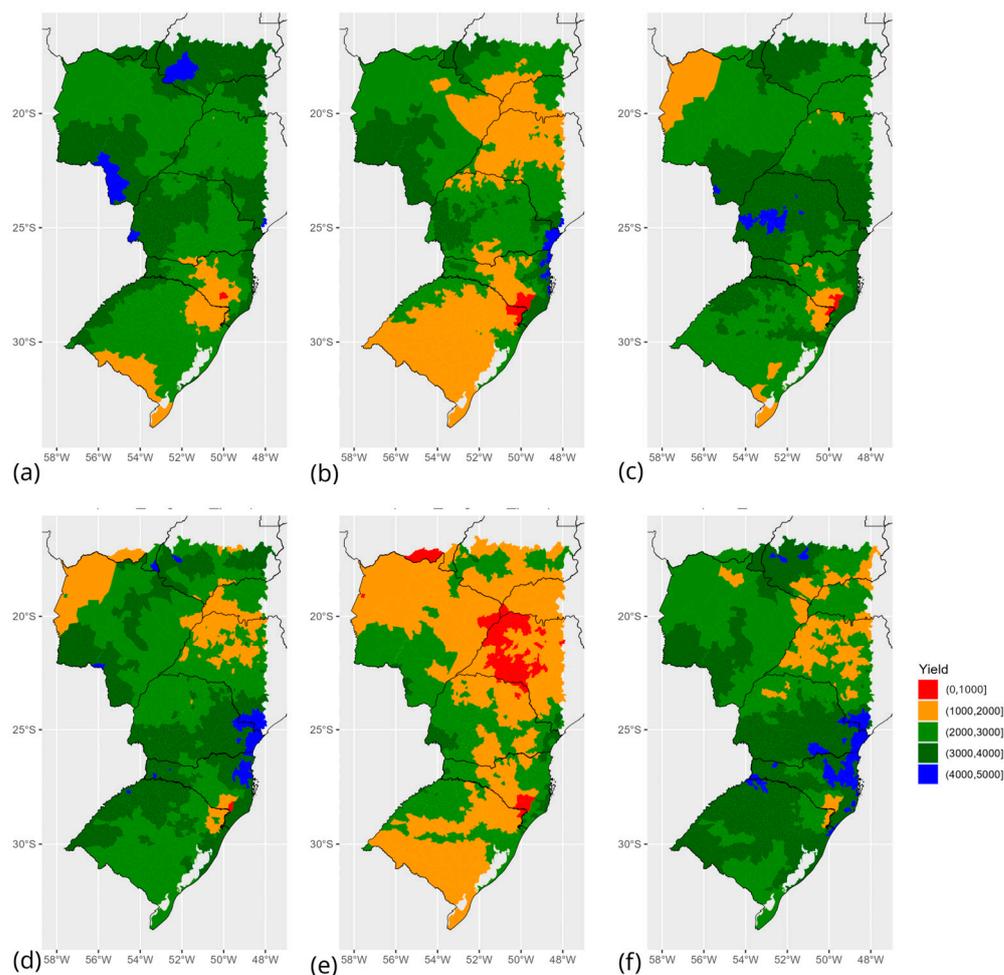
Figure 7. Average maize yield across three sowing dates (1 February, 1 March, and 1 April) and three growing seasons (sowing in 2019, 2020, and 2021), as simulated by the AgS crop model with daily climate from (**a**) BASE, (**b**) WRF-NCAR, (**c**) WRF-UCAN, (**d**) RegCM5-ICTP-pbl1, (**e**) RegCM5-ICTP-pbl2, and (**f**) RegCM4-USP.

Average total rainfall (Figure 3) during crop growth cycle shows substantial differences between CPRCM simulations and the reference. Closest average rainfall is simulated by WRF-NCAR, presenting the lowest *RMSE* (Table 2). The negative highlight was the WRF-UCAN, that systematically overestimates the rainfall—highest *MB* and *RMSE* (Table 2). RegCM5-ICTP simulations are the two models that underestimate the rainfall on average (*MB*, Table 2) and particularly the RegCM5-ICTP-pbl2. Although RegCM5-ICTP-pbl2 systematically underestimates the rainfall, the intraseasonal and interannual variabilities were well captured, as reflected by the highest average (for all counties) *ACC* calculated over time.

The best model in simulating incident solar radiation was the RegCM5-ICTP-pbl1 (Figure 4) which has the lowest *MB* and *RMSE* (Table 2). The WRF-NCAR and RegCM5-ICTP-pbl2 systematically overestimate the SRAD by ~12% (Table 2). Although presenting high average deviations, WRF-NCAR has the best temporal correlation (*ACCt*, Table 2).

Average air temperature was very well simulated by CPRCMs (Figure 5). All models presented a very similar average field. Average bias ranges from −0.05 °C (for WRF-NCAR) to +0.98 °C (for RegCM4-USP). Spatial and temporal *ACC* are also relatively high for all models, except for the *ACCt* of RegCM4-USP (Table 2).

Yield anomaly correlation calculated over time in Figure 6 (intraseasonal and inter-annual variability) shows that, in general, yield simulations with CPRCMs inputs are

more consistent over the more tropical region of the domain, north of latitude ~25° S. Low temporal correlation is particularly noted over the regions that have more complex terrain, around 25° W and 50° W.

Similarly to soybean simulations, the WRF-UCAN, RegCM5-ICTP-pbl1, and RegCM4-USP models presented comparable average maize yield results when compared to the BASE simulations (Figure 7). As with the soybean results, the RegCM5-ICTP-pbl2 simulations exhibited the largest deviation (highest *MB* and *RMSE*, Table 2). However, unlike the soybean simulations, WRF-NCAR showed a high error in the maize simulations, with both *MB* and *RMSE* being notably higher compared to the BASE simulations (Table 2).

3.2. Effect of the Number of Grid Points on the Average of Crop Yield at County Level

We assessed the effect of sampling grid points on yield estimates over all counties (Figure 8). Almost 80% of the counties had more than 10 grid points and 23% had at least 40 points (Table S1, Supplementary Material). Choosing more or fewer points leads to an average difference from the finer grid, for most counties, ranging from about 20 kg ha^{-1} for soybean and 40 kg ha^{-1} for maize to almost 200 kg ha^{-1} for both crops. The magnitude of the variability caused by point choice, represented by the standard deviation, is, for both crops, in most counties, over 60% of the average loss in accuracy, which suggests that, in some cases, the loss in accuracy may be dominated by a poor point choice. Nevertheless, from the yields presented for BASE estimates in Figures 2 and 7, for neither crop this value customarily would result in more than 15% of loss of accuracy.
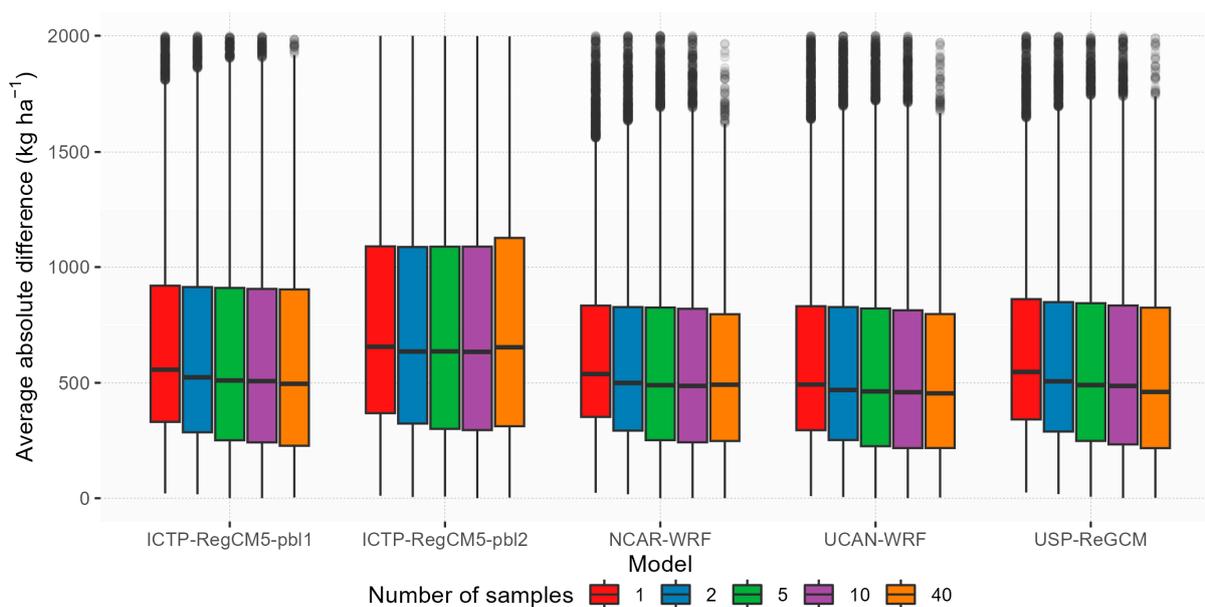
Figure 8. Effect of the number of grid points on the absolute difference between the soybean yield estimates obtained by the AgS model using daily climate from each model—WRF-NCAR, WRF-UCAN, RegCM5-ICTP-pbl1, RegCM5-ICTP-pbl2, and RegCM4-USP—and using the BASE reference (kg ha^{-1}) at the county level considering three sowing dates (1 October, 1 November, and 1 December) and three growing seasons (sowing in 2018, 2019, and 2020) Boxplots represent the median and first and third quartiles. Outliers were characterized as those points that exceeded or were inferior to quartiles by one and half times over the interquartile range.

With regard to the difference from the BASE estimates when using more or fewer points, we note that the distributions over all counties are similar within models for both crops, regardless of the number of points. Up to 10 points, for most counties and for both crops, the differences are mostly lower than 10%. While there are differences, likely caused by skews in the distribution, which distance the mean from the median, the best

performances were noted for NCAR-WRF and UCAN-WRF, and the highest error, for RegCM5-ICTP-pbl2, similarly to the result of Section 3.1.

4. Discussion

4.1. Impacts of Climate Input on Crop Models

Climate directly affects the crop yield, especially in tropical regions, where most of the crops are cultivated without the use of irrigation; most of the yield variability is related to climate oscillations. Therefore, the source of the climate data for crop yield estimation directly impacts the results. The impact of using the daily atmospheric variables derived from CPRCMs indicates that water availability is the primary factor influencing crop yield in simulations, particularly when demand surpasses supply, with variation in accuracy across regions. This finding aligns with previous studies across different crops and regions, such as Watsn and Challinor [29] in India, Ramarohetra et al. [9] in West Africa, Ahmad et al. [30] in South Asia, and Doi et al. [28] on a global scale. All these studies emphasize the need for precise rainfall estimations, while acknowledging the regional variability introduced by RCMs. Together, these works highlight the intersection among climate dynamics, water availability, and crop modeling, underscoring the importance of addressing uncertainties in water-related climate inputs to improve agricultural forecasting.

Although a common approach to addressing bias in climate models is to apply bias correction techniques to their outputs [7,12,31], the present study aimed to compare crop model responses to distinct CPRCMs. This approach allows the identification of strengths and weaknesses in CPRCM models and parametrizations, providing valuable insights to improve these models, reduce uncertainty, and enhance climate predictions for agricultural applications.

Despite lower water demand for maize (~400–700 mm) compared to soybean (~500–800 mm), maize is typically cultivated with a higher risk of water scarcity season, leading to more frequent soil moisture deficits which are the main driver of crop failure. Scarcer rains, with lower associated rainfall amounts of RGMs (especially below 25° S latitude), is linked to the reduced accuracy of spatial grain yields.

4.2. Impacts of Subsampling Grid Points

Usually, the applications of crop growth models are performed at the local scale, i.e., for few specific sites or counties [32,33]. When applied at a regional scale, many users take the central point of the county or exploit homogeneous climatic zones (HCZs), using, for example, the van Wart et al. [34] protocol. Major reasons for that are that either advanced crop growth model users are not necessarily experts in computer programming or that there are computational limitations for running large scale simulations. For example, in our study, as previously mentioned, more than six million simulations had to be performed.

From our results, we noted that, when choosing only one point, random sampling led to the magnitudes of the loss in accuracy corresponding to an average of 32% of the difference from the BASE simulations for all models in soybean estimates and 21% in maize estimates, for most counties. This suggests that, while not negligible, overall the differences are likely to be dominated by the differences in model estimates themselves, in particular in the higher end of the distribution.

Indeed, when accounting for the variability associated with the point choice, randomly choosing one or two points per county could increase the uncertainty in the estimates to the point of generating incorrect results. However, if climatic zones are used instead of random choice, it is possible that even with a few points the results could be satisfactory. Zhao et al. [35] assessed the precision gains of different stratification approaches over random sampling. They noted that, for their approach with the more consistent performance, i.e., the one based on coordinates using 4, 8, or 16 regions, often showed an improvement

of less than 10% in precision, regardless of the number of points. However, they also noted gains of over 20% in some cases, and as they assessed the different numbers of groups for each approach, using multiple crop models and evaluating the results for two crops, and they also noted that the best approach is model-dependent. An additional issue that can be neglected by poor sampling is related to capturing extremes. Van Bussel et al. [36] assessed stratified and random sampling weather observations for multiple wheat models in Germany. They observed that, for stratified samples, the temporal variability across the evaluated years was only lightly affected by the sample size, but for some models, the frequency of occurrence of lower yields decreased with fewer sampling points. These studies suggest that, to improve the choice of representative grid points in the region, for each of the assessed climate models, for both soybean and maize, using a different crop model, a dedicated study would be required.

5. Conclusions

This study is unique in the use of a state-of-the-art coordinated ensemble of convection-permitting regional climate model (CPRCM) simulations to drive an impact model focused on simulating crop yield in southern Brazil. The CPRCM ensemble has been built with grid spacings of only a few kilometers produced over southeastern South America (SESA) under the umbrella of the FPS-SESA initiative. The relevance of using a number of coordinated CPRCM simulations allowed exploring to what extent the uncertainty in the simulated climate is further affecting the uncertainty in the simulated crop yield. The evaluation of CPRCM uncertainty in representing major climate variables and the quantification of their impacts on crop model simulations show that, although crop growth simulations using CPRCMs generally reproduce the primary spatial and temporal variability of crop yields, some model configurations provide low accuracy. For example, the AgS crop model simulations driven by RegCM5-ICTP-pbl2 climate data perform poorly, which may be attributed to the fact that this simulation employs a planetary boundary layer scheme based on a local closure assumption. Moreover, the simulated rainfall uncertainty during the growing season emerged as a dominant factor affecting crop growth simulations.

The effect of the number of grid points on average crop yield at the county level reveals that selecting a single random point can lead to deviations of up to 32% in soybean estimates and 21% in maize estimates compared to the calculations using all available CPRCM grid points. These results indicate that common approaches—such as using the county's central point or exploiting homogeneous climatic zones (HCZs)—to estimate climate impacts at the county and regional scales can introduce random uncertainties, especially when using high-resolution climate simulations.

In the future, it will be important to assess not only the uncertainties associated with climate model simulations but also those introduced by crop growth models—specifically, by running an ensemble of crop models. Inter-institutional collaboration and coordinated scientific efforts are essential to address these comprehensive end-to-end studies.

Author Contributions: All author's works in the conceptualization, methodology, investigation, writing and reviewing the manuscript. All authors have read and agreed to the published version of the manuscript.
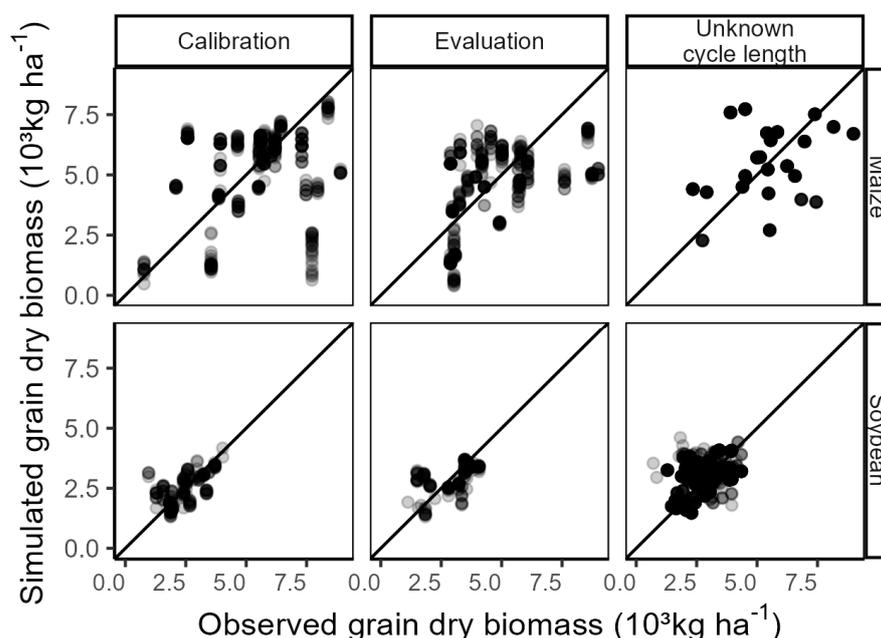
Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

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

Appendix A

Appendix A.1 Calibration and Evaluation Results for the AgS Crop Growth Model

Figure A1. Simulated and measured soybean and maize yield (kg ha^{-1}) using the breeding program dataset (sites used for calibration and evaluation are all located in the domain region of the CPRCM simulations).

Appendix A.2 Methodological Flowchart
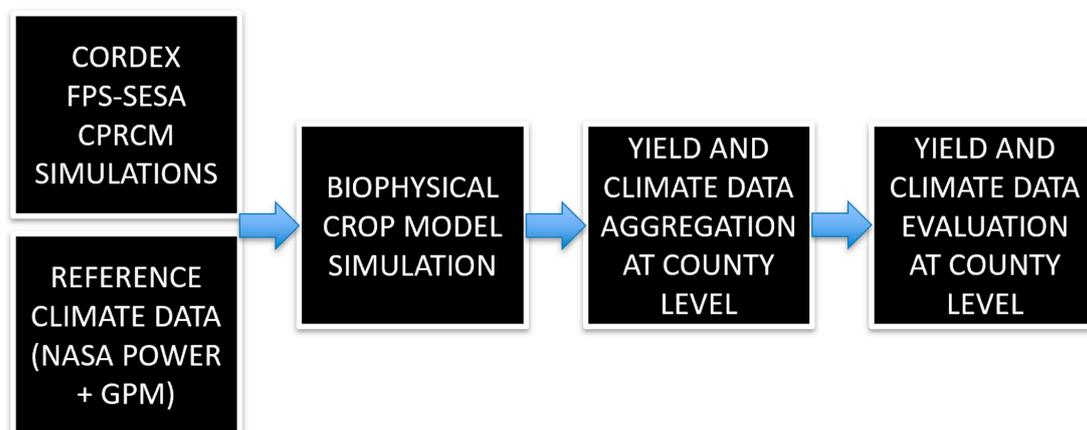
Figure A2. Visual representation of methodological flowchart.

References

1. FAOSTAT. Food and Agriculture Organization of the United Nations. 2024. Available online: https://www.fao.org/faostat/en/#data/QCL/visualize (accessed on 28 October 2024).

2. IBGE. Instituto Brasileiro de Geografia e Estatística. Sistema de Recuperação Automática (SIDRA). 2024. Available online: https://sidra.ibge.gov.br (accessed on 28 October 2024).

3. MAPA. Ministério da Agricultura Pecuária e Abastecimento—Programa de Subvenção ao Prêmio do Seguro Rural. Relatório 2023. 2024. Available online: https://www.gov.br/agricultura/pt-br/assuntos/riscos-seguro/seguro-rural/dados/relatorios/RelatorioGeralPSR2023.pdf (accessed on 28 October 2024).

4. Kim, K.S.; Yoo, B. Comparison of Regional Climate Scenario Data by a Spatial Resolution for the Impact Assessment of the Uncertainty Associated with Meteorological Inputs Data on Crop Yield Simulations in Korea. *J. Crop Sci. Biotechnol.* **2015**, *18*, 249–255. [CrossRef]

5. Rezaei, E.E.; Webber, H.; Asseng, S.; Boote, K.; Durand, J.L.; Ewert, F.; Martre, P.; MacCarthy, D.S. Climate Change Impacts on Crop Yields. *Nat. Rev. Earth Environ.* **2023**, *4*, 831–846. [CrossRef]

6. da Rocha, R.P.; Llopart, M.; Reboita, M.S.; Bettolli, M.L.; Solman, S.; Fernández, J.; Milovac, J.; Feijoó, M.; Coppola, E. Precipitation Diurnal Cycle Assessment in Convection-Permitting Simulations in Southeastern South America. *Earth Syst. Environ.* **2024**, *8*, 1–19. [CrossRef]

7. Chapagain, R.; Remenyi, T.A.; Harris, R.M.B.; Mohammed, C.L.; Huth, N.; Wallach, D.; Rezaei, E.E.; Ojeda, J.J. Decomposing Crop Model Uncertainty: A Systematic Review. *Field Crops Res.* **2022**, *279*, 108448. [CrossRef]

8. Pan, H.; Chen, Z. *Crop Growth Modeling and Yield Forecasting BT-Agro-Geoinformatics: Theory and Practice*; Di, L., Üstündağ, B., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 205–220. [CrossRef]

9. Ramarohetra, J.; Pohl, B.; Sultan, B. Errors and Uncertainties Introduced by a Regional Climate Model in Climate Impact Assessments: Example of Crop Yield Simulations in West Africa. *Environ. Res. Lett.* **2015**, *10*, 124014. [CrossRef]

10. Vesely, F.M.; Paleari, L.; Movedi, E.; Bellocchi, G.; Confalonieri, R. Quantifying Uncertainty Due to Stochastic Weather Generators in Climate Change Impact Studies. *Sci. Rep.* **2019**, *9*, 9258. [CrossRef] [PubMed]

11. Hasegawa, T.; Wakatsuki, H.; Ju, H.; Vyas, S.; Nelson, G.C.; Farrell, A.; Deryng, D.; Meza, F.; Makowski, D. A Global Dataset for the Projected Impacts of Climate Change on Four Major Crops. *Sci. Data* **2022**, *9*, 58. [CrossRef]

12. Lafferty, D.C.; Sriver, R.L. Downscaling and Bias-Correction Contribute Considerable Uncertainty to Local Climate Projections in CMIP6. *npj Clim. Atmos. Sci.* **2023**, *6*, 158. [CrossRef]

13. Falco, M.; Carril, A.F.; Menéndez, C.G.; Zaninelli, P.G.; Li, L.Z.X. Assessment of CORDEX Simulations over South America: Added Value on Seasonal Climatology and Resolution Considerations. *Clim. Dyn.* **2019**, *52*, 4771–4786. [CrossRef]

14. Reboita, M.S.; Dutra, L.M.M.; Dias, C.G. Diurnal Cycle of Precipitation Simulated by RegCM4 over South America: Present and Future Scenarios. *Clim. Res.* **2016**, *70*, 39–55. [CrossRef]

15. Coppola, E.; Sobolowski, S.; Pichelli, E.; Raffaele, F.; Ahrens, B.; Anders, I.; Ban, N.; Bastin, S.; Belda, M.; Belusic, D.; et al. A First-of-Its-Kind Multi-Model Convection Permitting Ensemble for Investigating Convective Phenomena over Europe and the Mediterranean. *Clim. Dyn.* **2020**, *55*, 3–34. [CrossRef]

16. Prein, A.F.; Langhans, W.; Fosser, G.; Ferrone, A.; Ban, N.; Goergen, K.; Keller, M.; Tölle, M.; Gutjahr, O.; Feser, F.; et al. A Review on Regional Convection-Permitting Climate Modeling: Demonstrations, Prospects, and Challenges. *Rev. Geophys.* **2015**, *53*, 323–361. [CrossRef] [PubMed]

17. Scaff, L.; Prein, A.F.; Li, Y.; Liu, C.; Rasmussen, R.; Ikeda, K. Simulating the Convective Precipitation Diurnal Cycle in North America's Current and Future Climate. *Clim. Dyn.* **2019**, *55*, 369–382. [CrossRef]

18. Kendon, E.J.; Ban, N.; Roberts, N.M.; Fowler, H.J.; Roberts, M.J.; Chan, S.C.; Evans, J.P.; Fosser, G.; Wilkinson, J.M. Do Convection-Permitting Regional Climate Models Improve Projections of Future Precipitation Change? *Bull. Am. Meteorol. Soc.* **2017**, *98*, 79–93. [CrossRef]

19. Bettolli, M.L.; Solman, S.A.; da Rocha, R.P.; Llopart, M.; Gutierrez, J.M.; Fernández, J.; Olmo, M.E.; Lavin-Gullon, A.; Chou, S.C.; Carneiro Rodrigues, D.; et al. The CORDEX Flagship Pilot Study in Southeastern South America: A Comparative Study of Statistical and Dynamical Downscaling Models in Simulating Daily Extreme Precipitation Events. *Clim. Dyn.* **2021**, *56*, 1589–1608. [CrossRef]

20. Giorgi, F.; Coppola, E.; Solmon, F.; Mariotti, L.; Sylla, M.B.; Bi, X.; Elguindi, N.; Diro, G.T.; Nair, V.; Giuliani, G.; et al. RegCM4: Model Description and Preliminary Tests over Multiple CORDEX Domains. *Clim. Res.* **2012**, *52*, 7–29. [CrossRef]

21. Giorgi, F.; Coppola, E.; Giuliani, G.; Ciarlo, J.M.; Pichelli, E.; Nogherotto, R.; Raffaele, F.; Malguzzi, P.; Davolio, S.; Stocchi, P.; et al. The Fifth Generation Regional Climate Modeling System, RegCM5: Description and Illustrative Examples at Parameterized Convection and Convection-Permitting Resolutions. *J. Geophys. Res. Atmos.* **2023**, *128*, e2022JD038199. [CrossRef]

22. Skamarock, W.C.; Klemp, J.B.; Dudhia, J.B.; Gill, D.O.; Barker, D.M.; Duda, M.G.; Huang, X.-Y.; Wang, W.; Powers, J.G. *A Description of the Advanced Research WRF Model Version 4.3*; (No. NCAR/TN556+STR); NSF: Alexandria, VA, USA, 2021. [CrossRef]

23. Tomasella, J.; Hodnett, M.G.; Rossato, L. Pedotransfer Functions for the Estimation of Soil Water Retention in Brazilian Soils. *Soil Sci. Soc. Am. J.* **2000**, *64*, 327–338. [CrossRef]

24. Tomasella, J.; Hodnett, M.G. Estimating Unsaturated Hydraulic Conductivity of Brazilian Soils Using Soil-Water Retention Data. *Soil Sci.* **1997**, *162*, 703–712. [CrossRef]

25. Carrara, E.R.; Lopes, P.S.; Reis, A.C.Z.; Silva, J.X.; Dias, L.C.d.C.M.; Schultz, É.B.; Marques, D.B.D.; da Silva, D.A.; Veroneze, R.; Andrade, R.G.; et al. NASA POWER Satellite Meteorological System Is a Good Tool for Obtaining Estimates of the Temperature-Humidity Index under Brazilian Conditions Compared to INMET Weather Stations Data. *Int. J. Biometeorol.* **2023**, *67*, 1273–1277. [CrossRef]

26. Rozante, J.R.; Vila, D.A.; Barboza Chiquetto, J.; Fernandes, A.D.A.; Souza Alvim, D. Evaluation of TRMM/GPM Blended Daily Products over Brazil. *Remote Sens.* **2018**, *10*, 882. [CrossRef]

27. Duarte, Y.C.N.; Sentelhas, P.C. NASA/POWER and DailyGridded Weather Datasets—How Good They Are for Estimating Maize Yields in Brazil? *Int. J. Biometeorol.* **2020**, *64*, 319–329. [CrossRef] [PubMed]

28. Doi, T.; Sakurai, G.; Iizumi, T. Seasonal Predictability of Four Major Crop Yields Worldwide by a Hybrid System of Dynamical Climate Prediction and Eco-Physiological Crop-Growth Simulation. *Front. Sustain. Food Syst.* **2020**, *4*, 84. [CrossRef]

29. Watson, J.; Challinor, A. The Relative Importance of Rainfall, Temperature and Yield Data for a Regional-Scale Crop Model. *Agric. For. Meteorol.* **2013**, *170*, 47–57. [CrossRef]

30. Ahmad, I.; Ahmad, B.; Boote, K.; Hoogenboom, G. Adaptation Strategies for Maize Production under Climate Change for Semi-Arid Environments. *Eur. J. Agron.* **2020**, *115*, 126040. [CrossRef]

31. Kim, Y.; Evans, J.P.; Sharma, A. Multivariate Bias Correction of Regional Climate Model Boundary Conditions. *Clim. Dyn.* **2023**, *61*, 3253–3269. [CrossRef]

32. Battisti, R.; Sentelhas, P.C.; Pascoalino, J.A.L.; Sako, H.; de Sá Dantas, J.P.; Moraes, M.F. Soybean Yield Gap in the Areas of Yield Contest in Brazil. *Int. J. Plant Prod.* **2018**, *12*, 159–168. [CrossRef]

33. Dias, H.B.; Sentelhas, P.C. Sugarcane Yield Gap Analysis in Brazil—A Multi-Model Approach for Determining Magnitudes and Causes. *Sci. Total Environ.* **2018**, *637–638*, 1127–1136. [CrossRef]

34. Van Wart, J.; van Bussel, L.G.J.; Wolf, J.; Licker, R.; Grassini, P.; Nelson, A.; Boogaard, H.; Gerber, J.; Mueller, N.D.; Claessens, L.; et al. Use of Agro-Climatic Zones to Upscale Simulated Crop Yield Potential. *Field Crops Res.* **2013**, *143*, 44–55. [CrossRef]

35. Zhao, G.; Hoffmann, H.; Yeluripati, J.; Xenia, S.; Nendel, C.; Coucheney, E.; Kuhnert, M.; Tao, F.; Constantin, J.; Raynal, H.; et al. Evaluating the Precision of Eight Spatial Sampling Schemes in Estimating Regional Means of Simulated Yield for Two Crops. *Environ. Model. Softw.* **2016**, *80*, 100–112. [CrossRef]

36. Van Bussel, L.G.J.; Ewert, F.; Zhao, G.; Hoffmann, H.; Enders, A.; Wallach, D.; Asseng, S.; Baigorria, G.A.; Basso, B.; Biernath, C.; et al. Spatial Sampling of Weather Data for Regional Crop Yield Simulations. *Agric. For. Meteorol.* **2016**, *220*, 101–115. [CrossRef]