The International Society of Precision Agriculture presents the
15th International Conference on Precision Agriculture
26–29 JUNE 2022
Minneapolis Marriott City Center | Minneapolis, Minnesota USA

COMPARISON OF DIFFERENT ASPATIAL AND SPATIAL INDICATORS TO ASSESS PERFORMANCE OF SPATIALIZED CROP MODELS AT DIFFERENT WITHIN-FIELD SCALES

Pasquel D.[1,*], Roux S.[2], Tisseyre B.[1], Taylor J.A.[1]
[1]ITAP, Univ. Montpellier, Institut Agro, INRAE, Montpellier, France
[2]MISTEA, Univ. Montpellier, Institut Agro, INRAE, Montpellier, France
*Corresponding author : daniel.pasquel@inrae.fr

A paper from the Proceedings of the
15th International Conference on Precision Agriculture
June 26-29, 2022
Minneapolis, Minnesota, United States

Abstract. Most current crop models are point-based models, i.e. they simulate agronomic variables at the spatial footprint on which they were initially designed (e.g. plant, field, region scale). Spatialization (i.e. using point-based crop models on a different scale than its native spatial footprint) represents a solution to use these crop models on a different scale. This is particularly interesting in a precision agriculture context where downscaling processes are involved to model agronomic variables on finer scale (e.g. within-field scale). To assess their performances, many indicators based on the comparison of estimated vs observed data, can be used. However, the use of classical, aspatial indicators may not be relevant to evaluate spatialized crop model performances. The objective of this work was to compare how different model performance indicators are able to evaluate the performance of a spatialized crop model at various within-field scales. The crop model spatialization processes were based on a spatial calibration of model parameters. This work focused on a case study using the crop model WaLIS (Water baLance for Intercropped Systems) to simulate vine water restriction (estimated through the predawn leaf water potential - Ψ_{PD}) for a vineyard in the South of France. The WaLIS model was employed at different spatial scales (field, site, within-field zone) to generate Ψ_{PD} maps. The management zones were generated from soil and vine ancillary data that are correlated with or directly influence vine water stress. Aspatial (RRMSE and D-index) and spatial (Cambardella index and Z-score) indicators were used to evaluate model performances at these different spatial scales. Results showed that these different indicators generated different 'best' simulation scales and there was no clear result of model performance from the spatial and aspatial indicators. This confirmed that current approaches to crop model evaluation were not well suited to evaluation the performance of spatialized crop models in a precision agricultural context. Evaluation in an operational context through decision-making evaluation and map comparison approaches provided a clearer understanding of model behavior and appeared to be a relevant method for evaluating downscaled spatialized crop model predictions for tactical, in-season and differential crop management.

Keywords. Spatialization, Spatial calibration, Downscaling, Evaluation indicators, Vine water restriction

Introduction

Most current crop models are point-based models, i.e. they simulate agronomic variables at the spatial footprint on which they were initially designed (e.g. plant, field, region scale) (Heuvelink et al., 2010). However, shifting model use from a strategic objective to tactical in-season management is becoming a significant issue for the agronomic community, especially in a precision agriculture context. Spatialization (i.e. using point-based crop models on a different scale than its native spatial footprint) represents a solution to address crop model use in a tactical and operational context. Calibration is an inevitable process to improve crop model performances (Seidel et al., 2018). This is especially true when using crop models over large areas (Jagtap & Jones, 2002), thus calibration is a critical step for models used in a spatialization context. In the case of precision agriculture applications that involve downscaling processes, it is important to consider how to calibrate crop models at the within-field scale. In this study, spatialization is based on a spatial calibration led at different spatial scales.

To assess crop model performance, many indicators based on the comparison of observed and modeled data, can be used, such as the relative root mean square error (RRMSE) or Willmott index of agreement (D-index) among others (Bennett et al., 2013; Wallach et al., 2019). However, the use of these indicators to evaluate spatialized crop models raises questions, and these indicators may not be relevant to evaluate their performances (Pasquel et al., 2022). Evaluation on the raw error of the crop model (error between observed and simulated data) using these indicators may not be sufficient to assess the spatial efficacy of calibration and/or prediction. Preliminary results using simulated data have indicated that the indicators currently used to evaluate spatialized crop models performance are not the most relevant (Pasquel et al., 2022). However, an evaluation and comparison of spatial and aspatial indicators on a real-world case study in a precision agricultural context has not yet been done, nor has there been any investigation of the effect of scale change (increasing levels of downscaling) on model performance. Note that if a spatialized crop model is used in an operational context associated with site-specific decision-making, it should be possible to mathematically define an error on the decision made using the simulations. Such indicators would clearly be the most adapted to assess model performances with respect to the targeted model use, but have not been strongly advocated to date.

Therefore, the objective of this work is to compare the evaluation of spatialized crop model performances using different indicators (aspatial: e.g. RRMSE or D-index; and spatial: e.g. Cambardella index or Moran index) for different simulation scales. The study is done on a vine water status crop model, WaLIS (Celette et al., 2010). As well as the different model statistics, the error on the decision taken is also used to evaluate the spatialized crop model performance. Model performance is defined by relationships between the observed and modeled data and the preservation of the spatial structure of the modeled variable in relation to the observed variable. It should be noted that the preservation of the spatial structure of the model output(s) is important in a precision agriculture objective.

Material and Methods

Field description and observed predawn leaf water potential ($\Psi_{PD.obs}$)

The predawn leaf water potential (Ψ_{PD}) of the vine was considered as the reference data in this case study. Ψ_{PD} measurements were carried out on a 1.2 ha Syrah vineyard block at INRAE's Pech Rouge estate (Gruissan, Aude, France) on 49 within-field sites (Fig. 1A) using a pressure chamber, these measurements were the observed data ($\Psi_{PD.obs}$). The $\Psi_{PD.obs}$ measurements were done for 7 dates in 2003 ($\Psi_{PD.obs.n-1}$) and 6 dates in 2004 ($\Psi_{PD.obs.n}$) (see Acevedo-Opazo et al. (2010) for full details on this data set).

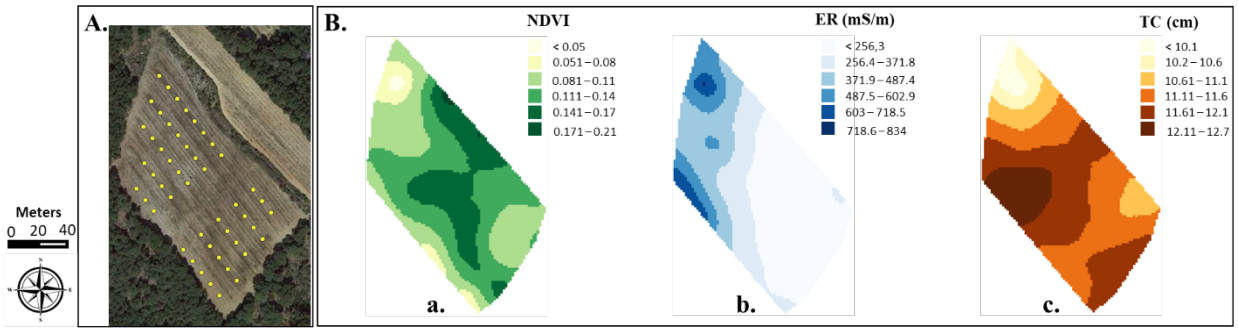

Fig 1. Experimental field with A. Locations of observed predawn leaf water potential ($\Psi_{PD.obs}$) within-field measurements in a 1.2 ha Syrah vineyard at INRAE Pech Rouge (Gruissan, Aude, France). B. Interpolated (kriged) maps of ancillary data used to define within-field zones: normalized difference vegetation index (NDVI) (a.), soil apparent electrical resistivity (ER, expressed as apparent electrical conductivity (EC_a)) (b.) and trunk circumference (TC) (c.) for the Syrah vineyard.

Three ancillary data were considered: soil apparent electrical resistivity (ER), trunk circumference (TC) and normalized difference vegetation index (NDVI). The ER and TC were both measured at the $\Psi_{PD.obs}$ measurement sites (Fig. 1A) in March 2006. The NDVI values were derived from an airborne multispectral image obtained in August 2006. Note that these ancillary data were used for their spatial pattern and not for their absolute values. The ancillary data were not measured in the same year as the $\Psi_{PD.obs}$; however, for a perennial crop like grapevine, it has been shown that NDVI and ER spatial patterns in this vineyard are temporally stable over short-time periods (3-5 years) (Kazmierski et al., 2011; Tisseyre et al., 2008). Thus, these ancillary data are assumed to present the same spatial pattern even in a different year. Ancillary data were interpolated by ordinary kriging using GeoFIS (Leroux et al., 2018) (Fig. 1B).

WaLIS and modeled predawn leaf water potential ($\Psi_{PD.mod}$)

The modeled predawn leaf water potential ($\Psi_{PD.mod}$) at multiple dates were simulated using a predictive model of vine water stress: Water baLance for Intercropped Systems – WaLIS (Celette et al., 2010). To run WaLIS, weather data for the years 2003 and 2004 were acquired through the weather station 11170004 (Gruissan) of the INRAE network via the Climatik application (Fig. 2). Measurements realized in 2003 were used to calibrate WaLIS and measurements realized in 2004 were used to evaluate modeling performances. Note that weather conditions for both years were relatively close. Daily mean temperature (T_{mean}) and daily precipitation (P) were recorded and daily evapo-transpiration (ET) was computed using the Penman-Monteith equation (Allen et al., 1989; Pereira et al., 1999). In reality, WaLIS simulates the fraction of transpirable soil water (FTSW) and, by using a conversion, FSTW was transformed into $\Psi_{PD.mod}$ (Eq. 1) (Lebon et al., 2003). However, this conversion contains a logarithmic expression and FTSW can be equal to 0, so a realistic $\Psi_{PD.mod}$ minimum had to be defined. The same $\Psi_{PD.obs}$ minimum for the field and year was used, -1.1 MPa.

$$\Psi_{PD.mod} = \frac{\log(FTSW) - \log(C_a)}{C_b} \quad (1)$$

with C_a is a constant equal to 1.0572 and C_b is a constant equal to 5.3452.

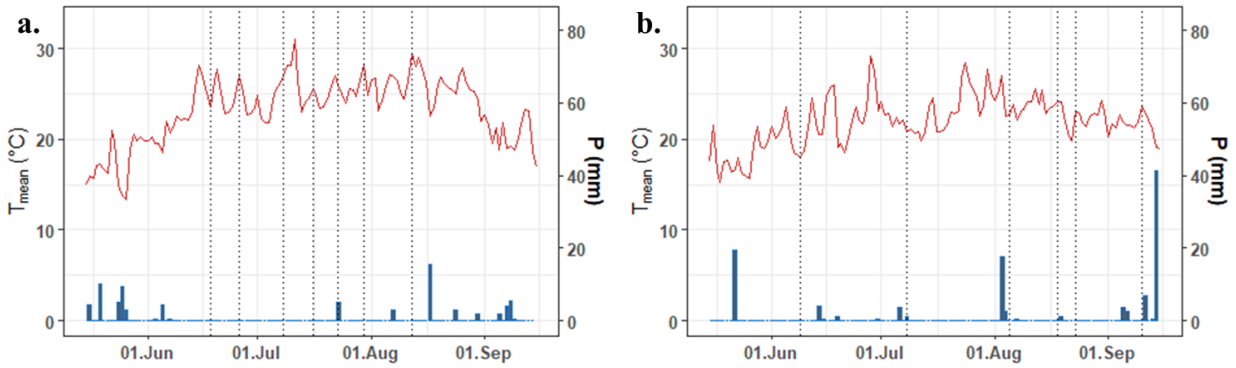

Fig 2. Weather conditions for 2003 (a.) and 2004 (b.) for the period of measurement dates used for respectively model calibration (2003) and performance evaluation (2004). The red line corresponds to daily mean temperature (T_{mean}) and the blue columns correspond to daily precipitation (P) events. Dates of measurement of predawn leaf water potential for calibration ($\Psi_{PD.obs.n-1}$) in 2003 were June 18th, June 26th, July 8th, July 16th, July 23rd, July 30th and August 12th. Dates of measurement of predawn leaf water potential for evaluation ($\Psi_{PD.obs.n}$) in 2004 were June 9th, July 8th, August 5th, August 18th, August 23rd and September 10th.

Spatial calibration of WaLIS

The intent of this study is to run the WaLIS model at different scales (at the measurement site-scale and at different within-field zone scales) than its native spatial footprint (i.e. field scale), thus WaLIS will be used in a spatialization process (Pasquel et al., 2022). The ER, TC and NDVI data were used for the realization of within-field zones via a segmentation algorithm (Pedroso et al., 2010) included in the GeoFIS software (Leroux et al., 2018) with the aim to define a base grid for the analyses. All three types of ancillary data were considered as potential surrogate to explain Ψ_{PD} spatial variability. The field was divided into 2 to 5 zones by the segmentation algorithm with these ancillary data (Fig. 3). Therefore, the spatial scales considered in this analysis are the individual measurement sites (n = 49) (Fig. 1A), the different zoning levels ($z \in [2;5]$) (Fig. 3) and the whole field (single value). The Ψ_{PD} is modeled by WaLIS ($\Psi_{PD.mod.n}$) on each of these considered spatial scales.

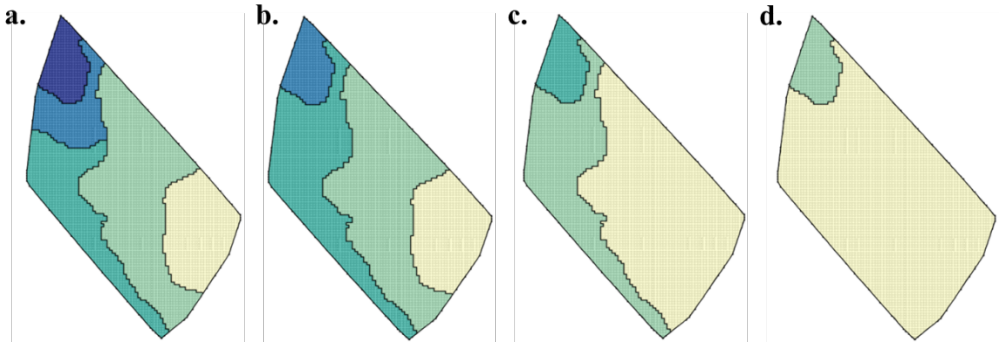

Fig. 3 Maps of different within-field zones defined with ancillary data using a segmentation algorithm; 5 zones (a.), 4 zones (b.), 3 zones (c.) and 2 zones (d.) for the Syrah vineyard. Note that the 1 zone solution is equivalent to the whole field scale.

Defining sub-units (zones) using ancillary data to apply crop models in order to spatialize the model is a process already seen in the literature (Basso et al., 2011; Cammarano et al., 2019, 2021; Guo et al., 2018). Spatial calibration for spatialization of the crop model is driven by the zoning process based on ancillary data, a process that is seldom seen in the literature. In an operational context here, a minimal size zone has been defined at 500 m², i.e. this is the minimum surface for which it is possible to set up a specific management action. With this constraint, 5 zones was the maximum number of zones that the segmentation algorithm was able to make (Fig. 3). Beyond 5 zones, the segmentation algorithm was unable to find a solution with the minimum surface constraint.

To enable spatialization, within each delineated zone at each spatial scale, an aggregation of the $\Psi_{PD.obs}$ from the measurement sites was performed to generate mean $\Psi_{PD.obs}$ values at larger spatial scales (zoning levels for $z \in [2;5]$ and the whole field scale). The local WaLIS calibration

was performed at either the original observation scale (for site-scale modeling) or after these aggregations (for modeling at scales larger than the site-scale), i.e. this is a kind of spatial calibration (Fig. 4). The calibrated parameters in the WaLIS model were the total transpirable soil water (TTSW) and the maximum crop coefficient of the vine (K_c). Other WaLIS parameters were kept at their default value for a Mediterranean context. Thus, only TTSW and K_c have been calibrated differently at the different scales because these parameters are known to be variable from one vineyard to another even at within-block scales (McClymont et al., 2019; Verdugo-Vásquez et al., 2022). The TTSW and K_c were calibrated with the data from the previous year ($\Psi_{PD.obs.n-1}$), and an optimization was performed with TTSW values ranging from 55 to 210 mm and K_c values ranging from 0.35 to 0.5 to identify optimal parameter values. The retained parameter values were those that minimized the mean absolute error (MAE) when compared to $\Psi_{PD.obs.n-1}$. For all the analyses, the output scale is disaggregated to the site-scale to assess the model performance (Fig. 4).

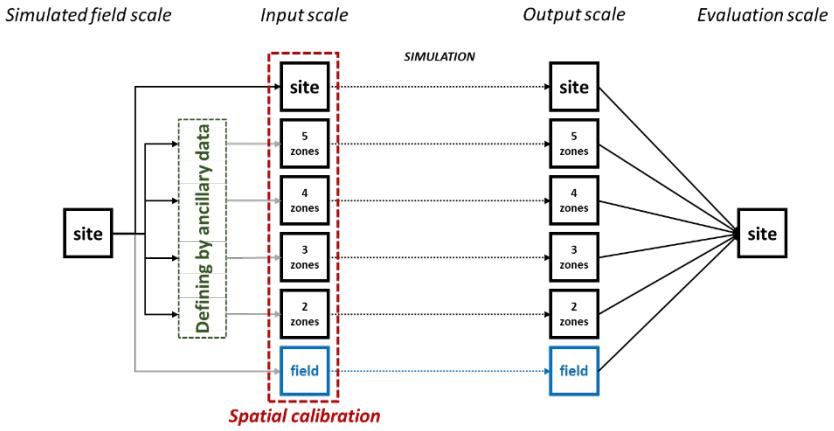

Fig. 4 Different modeling cases of predawn leaf water potential (Ψ_{PD}) by WaLIS according to different spatial scales defined at the observation scale, the whole field scale or intermediate zonal scales (2 to 5 zones) based on ancillary data. The native WaLIS spatial footprint is shown in blue and corresponds to the field scale. Site-scale corresponds to the original observation scale. The grey arrows correspond to the upscaling process associated with aggregations of the observed data to a higher spatial scale as model input. The spatial calibration is performed at this input scale.

Simulation performance evaluation using aspatial and spatial indicators

In order to estimate the WaLIS performance for the different simulation scales a set of aspatial indicators: relative root mean square error (RRMSE) (Eq. 2) and the Willmott index of agreement (D-index) (Eq. 3); and spatial indicators: Cambardella index (C_i) (Eq. 4) and Z-score (Eq. 6) were computed. RRMSE and D-index are assessment indicators of the fit between observed and simulated data. If the D-index is equal to 1, there is a perfect match between observed and simulated data, if D-index is equal to 0, there is no match at all. The C_i and Z-score are indicators of the spatial structure of the data. These indicators were calculated on both the observed and modeled data and on the residuals between the observed and modeled data. The C_i is derived from variographic analysis of $\Psi_{PD.obs}$ and $\Psi_{PD.mod}$, with values of $C_i < 0.25$ corresponding to a strong spatial structure, values between 0.25 and 0.75 corresponding to a moderate spatial structure and values > 0.75 corresponding to a weak spatial structure (Cambardella et al., 1994). Global Moran's index test (I) (Eq. 5) was also used to estimate spatial autocorrelation (Moran, 1948) and the values of I were then transformed into a Z-score. The Z-scores were interpreted to assess the magnitude and significance ($\alpha = 0.05$) of the spatial autocorrelation of Ψ_{PD} and residuals, knowing that -1.96 < Z-score < 1.96 corresponds to a non-significant spatial autocorrelation i.e. there is no clearly identifiable spatial structure.

$$\mathrm{RRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - S_i)^2}}{\bar{O}} \quad (2)$$

$$\text{D-index} = 1 - \frac{\sum_{i=1}^{n}(O_i - S_i)^2}{\sum_{i=1}^{n}(|S_i - \bar{O}| + |O_i - \bar{O}|)^2} \quad (3)$$

where O_i is the observed value, S_i is the corresponding simulated value, n is the number of observations (n = 49) and $\bar{O}$ is the average of observed values.

$$C_i = \frac{C_0}{C_0 + C_1}.100 \quad (4)$$

where C_0 is the variogram nugget and C_1 is the variogram partial sill.

$$I = \frac{n \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (5)$$

$$\text{Z-score} = \frac{I - E(I)}{\sqrt{V(I)}} \quad (6)$$

where y_i and y_j are the variable of interest at different spatial locations (i and j), $\bar{y}$ is the mean of the variable of interest, w_{ij} is a matrix of spatial weights quantifying the influence of j on i, n is the number of units indexed by i and j, $E(I)$ is the average of I and $V(I)$ is the variance of I.

Simulation performance evaluation from an operational context

The model outputs and the results of the indicators used to evaluate the performance of the modeling were also interpreted through the decision process to irrigate (or not) the vineyard at the within-field scale. The decision to irrigate was based on different Ψ_{PD} thresholds depending on the vine phenological cycle (Ojeda, 2007). The Ψ_{PD} thresholds were determined from the vine phenological stage assigned to the approximate corresponding dates (Fig. 5). Thus, the intent was not to compare the Ψ_{PD} values (modeled vs. observed) but to compare the decision made (to irrigate or not). The balanced accuracy statistic (BA) was used to summarize if the set of decisions taken at the site-scale to irrigate with $\Psi_{PD.mod}$ corresponded to the set of decisions taken based on $\Psi_{PD.obs}$ (Eq. 7). When BA is equal to 1 there is a perfect classification.

$$BA = \frac{\text{Sensitivity} + \text{Specificity}}{2} = \frac{1}{2}\left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right) \quad (7)$$

where TP is true positive, TN is true negative, FN is false negative and FP is false positive.

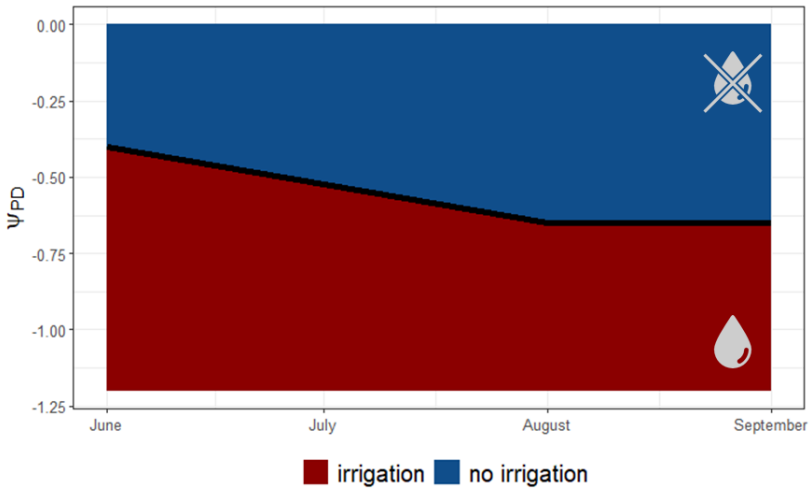

Fig. 5 Ψ_{PD} values below which irrigation is required over the season. The equation between June and August is Ψ_{PD} = -0.0033.Day number + 0.1 and is constant after August 1st at -0.65 MPa (adapted from Ojeda 2007).

Results and Discussion

Calibration values at varying spatial scales

Table 1 presents the results of the calibration optimization performed for the TTSW and K_c model parameters at different spatial scales. They represent the calibration of the model to the mean observation within each zone for the different levels of zoning ($z \in [2;5]$). The field-scale equated to a 1 zone scenario. The site-scale is not shown because it is too large to show here. However, for the site-scale calibration, the K_c values ranged from 0.368 to 0.5 and TTSW values ranged from 55 to 107.7 mm. The zone in the northern tip of the field (Fig. 3) was constant to all zonal modeling approaches and returned constant parameter values for all calibrations. However, it was interesting to notice that for each zone resulting from the merging of two other zones (Fig. 3), the calibrated parameter value was not necessarily an intermediate value of the parameter values of these two zones (Table 1). Notice that different zones within a spatial scale could have the same parameter values.

Table 1 Calibration values for each scale simulation (except site-scale) and for each WaLIS calibrated parameters: maximum crop coefficient of the vine (K_c) and total transpirable soil water (TTSW).

5 zones			4 zones			3 zones			2 zones			Field	
N°	K_c	TTSW	N°	K_c	TTSW	N°	K_c	TTSW	N°	K_c	TTSW	K_c	TTSW
1	0.494	67.4	1	0.494	67.4	1	0.494	67.4	1	0.494	67.4	0.467	70.5
2	0.458	67.4	2	0.5	73.6	2	0.5	73.6	2	0.485	73.6		
3	0.5	86											
4	0.491	89.1	3	0.491	89.1	3	0.5	89.1					
5	0.5	86	4	0.5	86								

Spatial structure preservation between $\Psi_{PD.obs}$ and $\Psi_{PD.mod}$

The $\Psi_{PD.obs}$ (in 2004) was moderately spatially structured over time in the Syrah vineyard (Fig. 6a), with the exception being the August 8th observation that exhibited no spatial structure. This measurement occurred shortly after a large (17.5 mm) and unusual precipitation event that will have had a short-term effect of homogenizing the vine water status within the vineyard. Two weeks after the event, the spatial structure in these data had returned. Therefore, the spatial structure was present when the vines were differentially stressed due to the different soil types in the vineyard that permit access to more or less soil water. The spatial structure of $\Psi_{PD.obs}$ and $\Psi_{PD.mod}$ realized at site-scale were comparable until early August, after which, the $\Psi_{PD.mod}$ no longer showed any spatial structure while the $\Psi_{PD.obs}$ continued to exhibit a moderate spatial structure. The WaLIS model tended to simulate more negative $\Psi_{PD.mod}$ than the $\Psi_{PD.obs}$, which tended to homogenize the $\Psi_{PD.mod}$ and, consequently, it did not preserve the spatial structure. The C_i and Z-score interpretations were complementary and presented the same trends, with the Z-score allowing for an assessment of the significance of the spatial autocorrelation. The spatial structure estimated by C_i should be discussed with the significance of the spatial autocorrelation estimated by I (through the Z-score), as the two principles cannot exist without each other (Tiefeldorf, 2000).

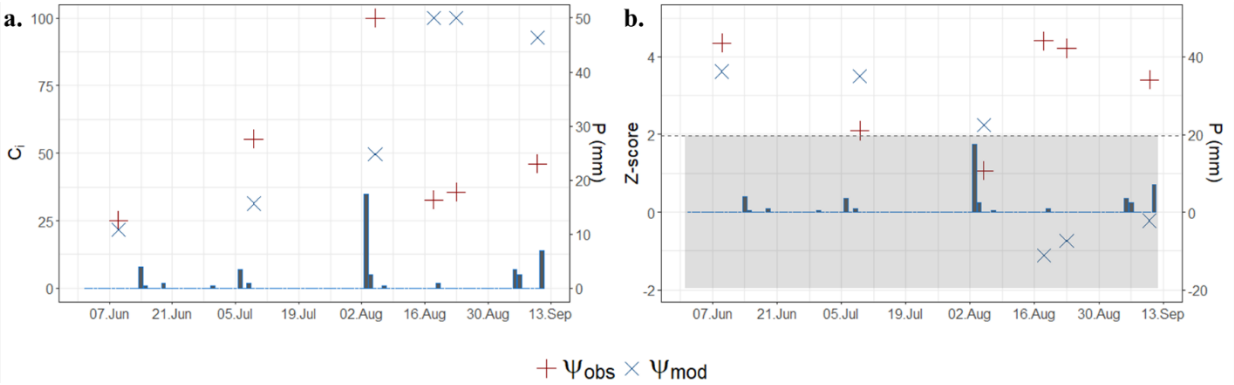

Fig. 6 Comparison between spatial structures of observed predawn leaf water potential ($\Psi_{PD.obs}$) and modeled predawn leaf water potential ($\Psi_{PD.mod}$) using the Cambardella index (C_i) (a.) and Z-score (b.). The grey area on the Z-scores plot refers to an area of non-significance of spatial autocorrelation. Blue columns correspond to daily precipitation (P) events.

Which modeling scale is the most relevant regarding aspatial and spatial indicators?

Depending on the modeling date (seasonal timing), the model performance was not constant. At the beginning of the season, the RRMSE values were relatively close regardless of the modeling scale considered (Fig. 7a). For late summer, when water stress was the highest, the RRMSE values increased because WaLIS underestimated the vine water restriction (overestimated Ψ_{PD}). In late summer, the RRMSE values were still the highest for the 2-zone and field scale simulations with 34% and 71% respectively. For the $\Psi_{PD.mod}$ generated at finer scales (3+ zones including individual sites), lower and similar RRMSE values were observed, showing a better agreement between $\Psi_{PD.obs}$ and $\Psi_{PD.mod}$. When using the D-index for assessing the agreement between $\Psi_{PD.obs}$ and $\Psi_{PD.mod}$, the finer the modeling scale at the beginning of summer, the better the agreement. However, this trend had disappeared by late summer (Fig. 7b) with the site-scale modeling dropping from the highest D-index to one of the lowest values in mid-August. The D-index for the field scale simulations were always equal to 0, i.e. there was no agreement between $\Psi_{PD.obs}$ and $\Psi_{PD.mod}$.

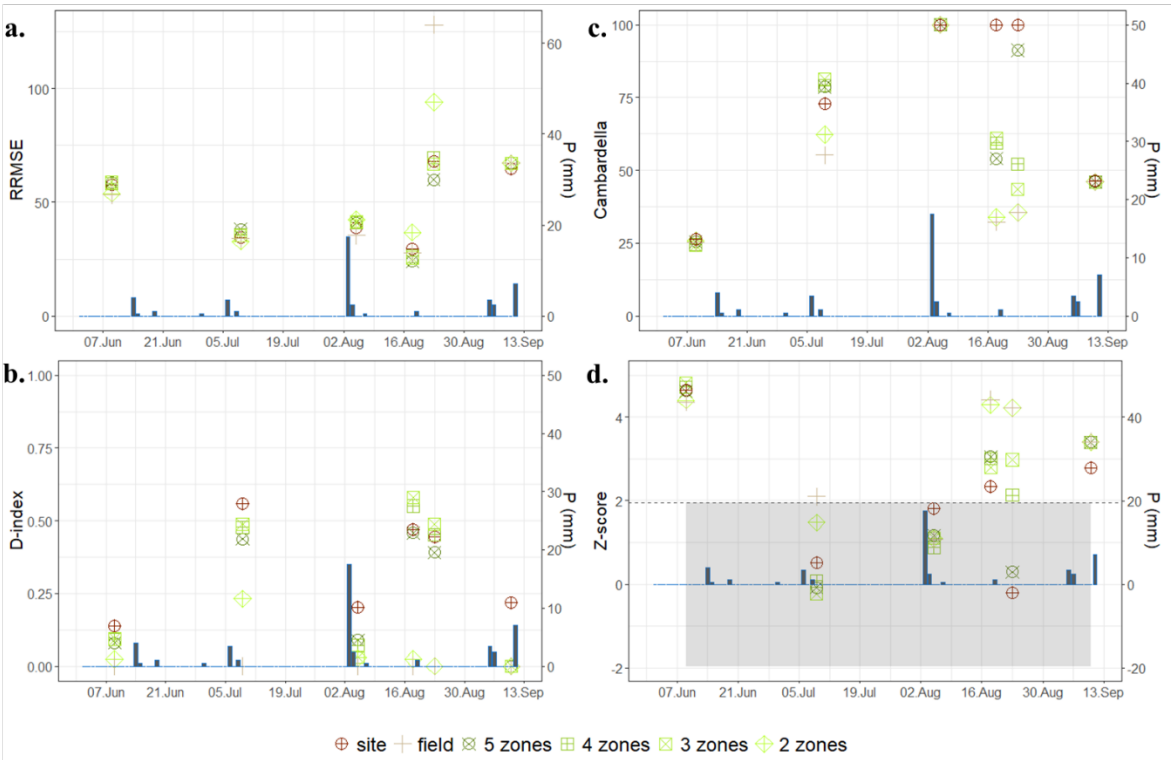

Fig. 7 Evaluation of the WaLIS simulations of the predawn leaf water potential (Ψ_{PD}) according to the different spatial scales along the measurement period. The RRMSE (a.) and D-index (b.) are calculated between observed and modeled data. The Cambardella index (C_i) (c.) and Z-scores (d.) are calculated on the residuals between observed and modeled data. The grey area on the Z-scores plot refers to an area of non-significance of spatial autocorrelation. Blue columns correspond to daily precipitation (P) events.

The spatial structure of the residuals is an indication of how well the model has been correctly spatialized. The target variable, $\Psi_{PD.obs}$, is spatially structured during some periods of the season (especially during dry periods) (Fig. 6a). If the modeling is able to replicate this, then the residuals should be spatially random (i.e. exhibit only a nugget effect). In the other extreme, as is the case for the whole field simulation, the removal of a constant (modeled) value from the observed data should retain the spatial structure in the $\Psi_{PD.obs}$ within the residuals. Therefore, a higher C_i value would indicate that the residuals were not spatially structured. A C_i value equal to 100 corresponds to a nugget effect of the variogram (random effects). Concerning the preservation of spatial structure with respect to $\Psi_{PD.obs}$, two trends were interesting to note.

The first trend concerned the period up until the beginning of August, with the spatial structure of the model residuals becoming less and less structured over time for all modeling scales, i.e. C_i increases (Fig. 7c). Thus, the residuals were randomly distributed by early August, as indicated

by an absence of spatial structure, and it appeared that the spatialized model was performing well. However, note the non-significance of the spatial autocorrelation with respect to the residuals (Fig. 7d). The same trend was observed for the spatial structure of $\Psi_{PD.obs}$ and $\Psi_{PD.mod}$ (Fig. 6), this loss of spatial structure was likely influenced by the precipitation events, which occurred just before the measurement dates, and had a homogenizing effect on the $\Psi_{PD.obs}$ values within the field.

The second result concerned the late summer period (mid-August to mid-September), when the residuals were randomly distributed for the site-scale modeling but spatially structured for modeling at higher spatial scales (zonal or whole field). The residual spatial structure at the higher spatial scales can be explained by the fact that the $\Psi_{PD.obs}$ were spatially structured and the $\Psi_{PD.mod}$ were not for this period (Fig. 6). The $\Psi_{PD.mod}$ were not structured because for all sites constituting the within-field zones and the whole field simulation scale, the $\Psi_{PD.mod}$ values were the same for all sites. Thus, modeling was not able to replicate the $\Psi_{PD.obs}$ spatial structure except for the case of the site-scale modeling, which seemed to be the best simulation scale regarding spatial indicators for this period.

The spatial structures of $\Psi_{PD.obs}$ and $\Psi_{PD.mod}$ were quite different, except at the beginning of the modeling period when the spatial structures were comparable, but a period that coincided with low vine water stress that is considered of little importance to producers. Additionally, even if the spatial structures reflected by the C_i were similar, it is important to note that the level of semivariance for $\Psi_{PD.obs}$ and $\Psi_{PD.mod}$ were not of the same magnitude. The $\Psi_{PD.mod}$ semivariance was nearly 75% lower than $\Psi_{PD.obs}$ semivariance, indicating a much smaller range of values in this dataset. The variance differences were not shown with the C_i value because it is a result of the ratio between the nugget and sill values. Another point of interest is the range of the variograms used to calculate the C_i. The $\Psi_{PD.obs}$ variogram range was twice that of the $\Psi_{PD.mod}$ variogram range. Thus, the $\Psi_{PD.mod}$ spatial patterns (using the site-scale modeling) were smaller than the $\Psi_{PD.obs}$ spatial patterns.

Looking holistically at the aspatial and spatial indicators, there is no clear pattern to identify the best modeling scale. The site-specific and 5-zone modeling (finer spatial resolution models) tended to indicate the best model spatialization, for example lower C_i and higher Z-score values in mid-August), but this was not supported by better aspatial metrics (lower RRMSE and higher D-Index) at this time and at these finer spatial resolutions (Fig. 7). Thus, the evaluation of model performance cannot be estimated by looking at only one or a combination of the indicators, i.e. different indicators point to a different best modeling scales, and none can be selected with certainty. For this reason, the metrics have also been interpreted relative to the decision process for within-field irrigation. This aimed to be able to decide which modeling scale was the most relevant, such that the $\Psi_{PD.mod}$ indicated the correct decision to be made (i.e. the decision that would have been taken with the $\Psi_{PD.obs}$ data).

Which modeling scale is the most relevant in regards to an operational context?

Figure 8 shows the translation of the model output, at each date and at each spatial scale, into an irrigation decision based on the date and the recommendation of Ojeda (2007) (Fig. 5). Note that the dates up until and including August 5th had a consistent non-irrigation decision for the real data and the modeling at all spatial scales regardless of the actual quality of the prediction (e.g. with the indices shown in Fig. 7). The operational decision-making to change the irrigation situation evolved in mid to late August, and the observed data indicated that the northern tip of the vineyard should be irrigated from August 18th 2004, and by September 10th the majority of the vineyard, except the southern third, should be receiving irrigation. For dates on or after August 18th, the scale at which the crop model was applied affected the quality of the irrigation decision. The whole field and 2-zone scale modeling flipped the whole field from non-irrigated to irrigated, although on same dates. The WaLIS model tended to simulate Ψ_{PD} more negative than $\Psi_{PD.obs}$, so all sites were predicted as requiring irrigation. By August 23rd, the overestimation of vine water stress by the WaLIS model led to an effective irrigation decision for the entire field for modeling at all spatial scales. Consequently, the remainder of this discussion will focus on the August 18th

results. The highest balanced accuracy statistic (BA) was obtained with modeling at the 5-zone scale, which outperformed the site-scale modeling according to the BA (Fig. 8). The BA was also higher for the 3 and 4-zone modeling than the site-scale modeling. These results showed that modeling at a spatial scale between 3 and 5 zones generated a better decision. It was not possible to identify this when using the proposed aspatial and spatial indicators. Evaluation in an operational context allowed an identification of which scale was the most relevant to model Ψ_{PD} as close as possible to the $\Psi_{PD.obs}$ decision.

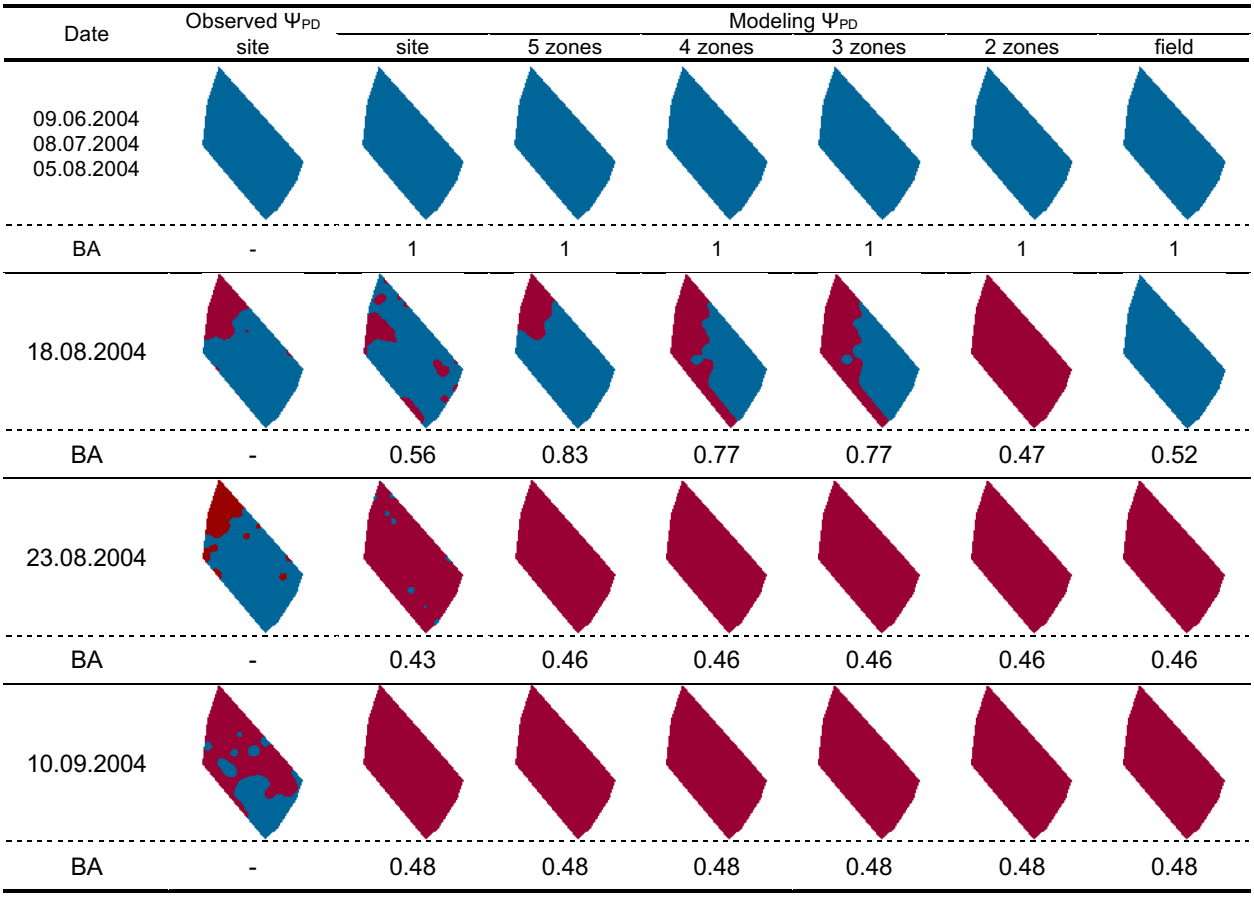

Fig. 8 Maps of within-field irrigation according to the date and the value of predawn leaf water potential (Ψ_{PD}) for the observed point data and the modeled data at each considered scale of simulation. BA = Balanced Accuracy. Observed and site-scale modelled data maps were interpolated using inverse distance weighting.

Conclusion

In a precision agriculture context, using crop models at a finer scale than the model's native spatial footprint is of principal interest. Spatialization using downscaling processes is one of the methods that could be used to achieve this goal. Resulting spatialized crop models are currently often evaluated using aspatial and spatial indicators. However, interpreted individually, these indicators indicated different best simulation scales, thus, this study showed that using these indicators was not the most relevant method for assessing this kind of model application. The evaluation of spatialized crop models for precision agriculture in an operational context seemed to be a better evaluation method. Based on a decision-making approach, identifying the best simulation scale that was closest to the observed data was much easier and more relevant for assessing model performance. Ideally, a spatial indicator able to indicate if the zoning level is more relevant than another level to simulate an agronomic variable could be a great improvement. The spatial indicators used in this study are blind to this goal, which is why evaluation in an operational context was more relevant in this case. Spatial calibration is the process key here in the spatialization process, it would also be interesting to see to what extent the spatial structure of the agronomic variable also influences the relevance of this spatialization method.

Acknowledgements

This work was supported by the French National Research Agency under the Investments for the Future Program, referred as ANR-16-CONV-0004.

References

Acevedo-Opazo, C., Tisseyre, B., Ojeda, H., & Guillaume, S. (2010). Spatial extrapolation of the vine (Vitis vinifera L.) water status: a first step towards a spatial prediction model. *Irrigation Science*, *28*(2), 143–155. https://doi.org/10.1007/s00271-009-0170-3

Allen, R. G., Jensen, M. E., Wright, J. L., & Burman, R. D. (1989). Operational Estimates of Reference Evapotranspiration. *Agronomy Journal*, *81*(4), 650–662. https://doi.org/10.2134/agronj1989.00021962008100040019x

Basso, B., Ritchie, J. T., Cammarano, D., & Sartori, L. (2011). A strategic and tactical management approach to select optimal N fertilizer rates for wheat in a spatially variable field. *European Journal of Agronomy*, *35*(4), 215–222. https://doi.org/10.1016/j.eja.2011.06.004

Bennett, N. D., Croke, B. F. W., Guariso, G., Guillaume, J. H. A., Hamilton, S. H., Jakeman, A. J., et al. (2013). Characterising performance of environmental models. *Environmental Modelling & Software*, *40*, 1–20. https://doi.org/10.1016/j.envsoft.2012.09.011

Cambardella, C. A., Moorman, T. B., Novak, J. M., Parkin, T. B., Karlen, D. L., Turco, R. F., & Konopka, A. E. (1994). Field-Scale Variability of Soil Properties in Central Iowa Soils. *Soil Science Society of America Journal*, *58*(5), 1501–1511. https://doi.org/10.2136/sssaj1994.03615995005800050033x

Cammarano, D., Basso, B., Holland, J., Gianinetti, A., Baronchelli, M., & Ronga, D. (2021). Modeling spatial and temporal optimal N fertilizer rates to reduce nitrate leaching while improving grain yield and quality in malting barley. *Computers and Electronics in Agriculture*, *182*, 105997. https://doi.org/10.1016/j.compag.2021.105997

Cammarano, D., Holland, J., Basso, B., Fontana, F., Murgia, T., Lange, C., et al. (2019). Integrating geospatial tools and a crop simulation model to understand spatial and temporal variability of cereals in Scotland. In *Precision agriculture '19* (pp. 29–35). Presented at the 12th European Conference on Precision Agriculture, Montpellier, France: Wageningen Academic Publishers. https://doi.org/10.3920/978-90-8686-888-9_2

Celette, F., Ripoche, A., & Gary, C. (2010). WaLIS—A simple model to simulate water partitioning in a crop association: The example of an intercropped vineyard. *Agricultural Water Management*, *97*(11), 1749–1759. https://doi.org/10.1016/j.agwat.2010.06.008

Guo, C., Zhang, L., Zhou, X., Zhu, Y., Cao, W., Qiu, X., et al. (2018). Integrating remote sensing information with crop model to monitor wheat growth and yield based on simulation zone partitioning. *Precision Agriculture*, *19*(1), 55–78. https://doi.org/10.1007/s11119-017-9498-5

Heuvelink, G. B. M., Brus, D. J., & Reinds, G. (2010). Accounting for spatial sampling effects in regional uncertainty propagation analysis (pp. 85–88). Presented at the 9th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Leicester, UK.

Jagtap, S. S., & Jones, J. W. (2002). Adaptation and evaluation of the CROPGRO-soybean model to predict regional yield and production. *Agriculture, Ecosystems & Environment*, *93*, 73–85.

Kazmierski, M., Glémas, P., Rousseau, J., & Tisseyre, B. (2011). Temporal stability of within-field patterns of NDVI in non irrigated Mediterranean vineyards. *OENO One*, *45*(2), 61. https://doi.org/10.20870/oeno-one.2011.45.2.1488

Lebon, E., Dumas, V., Pieri, P., & Schultz, H. R. (2003). Modelling the seasonal dynamics of the soil water balance of vineyards. *Functional Plant Biology*, *30*(6), 699. https://doi.org/10.1071/FP02222

Leroux, C., Jones, H., Pichon, L., Guillaume, S., Lamour, J., Taylor, J., et al. (2018). GeoFIS: An Open Source, Decision-Support Tool for Precision Agriculture Data. *Agriculture*, *8*(6), 73. https://doi.org/10.3390/agriculture8060073

McClymont, L., Goodwin, I., Whitfield, D. M., & O'Connell, M. G. (2019). Effects of within-block canopy cover variability on water use efficiency of grapevines in the Sunraysia irrigation region, Australia. *Agricultural Water Management*, *211*, 10–15. https://doi.org/10.1016/j.agwat.2018.09.028

Moran, P. A. P. (1948). The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society: Series B (Methodological)*, *10*(2), 243–251. https://doi.org/10.1111/j.2517-6161.1948.tb00012.x

Ojeda, H. (2007). Irrigation qualitative de précision de la vigne. *Le Progrès Agricole et Viticole*, *124*(7), 133–141.

Pasquel, D., Roux, S., Richetti, J., Cammarano, D., Tisseyre, B., & Taylor, J. A. (2022). A review of methods to evaluate crop model performance at multiple and changing spatial scales. *Precision Agriculture*. https://doi.org/10.1007/s11119-022-09885-4

Pedroso, M., Taylor, J., Tisseyre, B., Charnomordic, B., & Guillaume, S. (2010). A segmentation algorithm for the delineation of agricultural management zones. *Computers and Electronics in Agriculture*, *70*(1), 199–208. https://doi.org/10.1016/j.compag.2009.10.007

Pereira, L. S., Perrier, A., Allen, R. G., & Alves, I. (1999). Evapotranspiration: Concepts and Future Trends. *Journal of Irrigation and Drainage Engineering*, *125*(2), 45–51. https://doi.org/10.1061/(ASCE)0733-9437(1999)125:2(45)

Seidel, S. J., Palosuo, T., Thorburn, P., & Wallach, D. (2018). Towards improved calibration of crop models – Where are we now and where should we go? *European Journal of Agronomy*, *94*, 25–35. https://doi.org/10.1016/j.eja.2018.01.006

Tiefeldorf, M. (2000). *Modelling spatial processes - The identification and analysis of spatial relationships in regression residuals by means of Moran's I* (Springer.).

Tisseyre, B., Mazzoni, C., & Fonta, H. (2008). Within-field temporal stability of some parameters in Viticulture: potential toward a site specific management. *OENO One*, *42*(1), 27. https://doi.org/10.20870/oeno-one.2008.42.1.834

Verdugo-Vásquez, N., Acevedo-Opazo, C., Valdés-Gómez, H., Pañitrur-De la Fuente, C., Ingram, B., García de Cortázar-Atauri, I., & Tisseyre, B. (2022). Identification of main factors affecting the within-field spatial variability of grapevine phenology and total soluble solids accumulation: towards the vineyard zoning using auxiliary information. *Precision Agriculture*, *23*(1), 253–277. https://doi.org/10.1007/s11119-021-09836-5

Wallach, D., Makowski, D., Jones, J. W., & Brun, F. (2019). *Working with dynamic crop models - Methods, tools and examples for agriculture and environment* (3rd ed.). Elsevier.