

# **Spatial crop-water variations in rainfed wheat systems**

**From simulation modelling to site-specific management**

**Tomás Roquette Tenreiro**

## Thesis committee

### Dissertation supervisors

**Prof. dr. ir. Elías Fereres Castiel, Department of Agronomy – University of Córdoba and Institute for Sustainable Agriculture – Spanish National Research Council (CSIC), Spain.**

**Dr. ir. José Alfonso Gómez Calero, Institute for Sustainable Agriculture – Spanish National Research Council (CSIC), Spain.**

### Other collaborators

Dr. ir. Margarita García-Vila, Department of Agronomy – University of Córdoba, Spain || Dr. ir. José Jimenez-Berni, Institute for Sustainable Agriculture – Spanish National Research Council (CSIC), Spain || Ir. Jakub Jerabek, Department of Landscape Water Conservation, Czech Technical University in Prague, Czech Republic || Dr. ir. David Zumr, Department of Landscape Water Conservation, Czech Technical University in Prague, Czech Republic || Dr. ir. Gonzalo Martínez, Department of Applied Physics, University of Córdoba, Spain || Prof. dr. ir. Francisco Avillez, Higher Institute of Agronomy, University of Lisbon, Portugal || Prof. dr. ir. José C. Coelho , Higher Institute of Agronomy, University of Lisbon, Portugal || Ir. Manuel Penteado, Agro-analitica – Precision Agriculture and Systems Optimization, Portugal.

# Spatial crop-water variations in rainfed wheat systems

From simulation modelling to site-specific management

**Tomás Roquette Tenreiro**

This Doctoral Thesis was conducted at the Institute for Sustainable Agriculture (CSIC) and received funding from the European Commission under project SHui – Grant agreement ID 773903.

Córdoba, July 2022



*"El camino hacia el saber no es acumular ciencia como algunos acumulan riquezas; es reconocer la propia realidad de uno en el mundo y juzgarla; es renovar en sí mismo el misterio de la creación"* - Maimónides (Córdoba, siglo XII).





TÍTULO DE LA TESIS

**Spatial crop-water variations in rainfed wheat systems**  
From simulation modelling to site-specific management

DOCTORANDO

**Tomás Ferreira Roquette da Cruz Tenreiro**

**INFORME RAZONADO DE LOS DIRECTORES DE LA TESIS**

(evolución y desarrollo de la tesis, así como a trabajos y publicaciones derivados de la misma)

Esta tesis doctoral representa una contribución original a la agronomía de los sistemas agrícolas de secano, con énfasis en el papel que juegan los flujos de agua en áreas de topografía ondulada en la determinación de las variaciones espaciales del rendimiento del trigo, dentro de una parcela que se maneja de forma homogénea. La tesis se ha realizado en capítulos que están asociados en un enfoque integrador y parte de un análisis crítico del conocimiento disponible en la modelación de los flujos horizontales y de la respuesta de la producción de trigo al suministro de agua. El segundo capítulo explora el uso de un indicador simple de la cobertura vegetal para determinar las variaciones espaciales de este determinante de rendimiento, aplicado a diversos cultivos. Como las variaciones espaciales no sólo se deben a las diferencias en crecimiento sino a otras diferencias en el factor suelo, como en el suministro hídrico debido a movimientos laterales de agua, el tercer capítulo aborda la determinación experimental de dichos flujos en una parcela de trigo durante dos años y su influencia en las variaciones espaciales de producción dentro de la parcela. Finalmente, el cuarto capítulo valora la viabilidad económica de usar agricultura de precisión en el sistema agrícola estudiado bajo distintos escenarios.

El doctorando ha realizado su labor con el asesoramiento de sus directores, pero, muy mayoritariamente, ha trabajado de forma independiente, como debería ser el caso en la realización de una tesis doctoral, además los directores informan que todo el trabajo es original y ha sido supervisado adecuadamente a lo largo de los años de trabajo, por lo que se informa favorablemente esta tesis.

## **Publicaciones derivadas de la tesis:**

- Tenreiro, T. R., García-Vila, M., Gómez, J. A., Jimenez-Berni, J. A., & Fereres, E. (2020). Water modelling approaches and opportunities to simulate spatial water variations at crop field level. *Agricultural Water Management*, 240, 106254.
- Tenreiro, T. R., García-Vila, M., Gómez, J. A., Jiménez-Berni, J. A., & Fereres, E. (2021). Using NDVI for the assessment of canopy cover in agricultural crops within modelling research. *Computers and Electronics in Agriculture*, 182, 106038.
- Tenreiro, T. R., Jeřábek, J., Gómez, J. A., Zumr, D., Martínez, G., García-Vila, M., & Fereres, E. (2022). Simulating water lateral inflow and its contribution to spatial variations of rainfed wheat yields. *European Journal of Agronomy*, 137, 126515.

## **Otras contribuciones derivadas de la tesis:**

- Tenreiro, T. R. (2019). Modelling Seminar on AquaCrop at Fuzhou (Fujian Agriculture University – EU Shui H2020 Project General Assembly).
- Tenreiro, T. R., García-Vila, M., Gómez, J. A., & Fereres, E. (2020). Uncertainties associated with the delineation of management zones in precision agriculture. In EGU General Assembly Conference Abstracts (p. 5709).
- Tenreiro, T. R., García-Vila, M., Gómez, J. A., & Fereres, E. (2020). From point to field scale – uncertainties associated with the upscaling of modelling for field heterogeneity assessment. In European Society of Agronomy - XVI conference in Seville.
- Jeřábek, J., Zumr, D., Dostál, T., Tenreiro, T. R., Strauss, P., & Vaverková, M. D. (2021, November). The effects of management practices and fires on soil water dynamics at three locations across Europe. In *2021 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor)* (pp. 197-202). IEEE.
- Tenreiro, T. R. (2022). Digitalização na Agricultura - Novas Perspectivas: oportunidades para la tasa de aplicación variable de nitrógeno en los sistemas de trigo de secano cordobeses – un análisis económico. Instituto Superior de Agronomía, Universidad de Lisboa.
- Tenreiro, T. R., Avillez, F., Gómez, J. A., Penteado, M., Coelho, J. C., Fereres, E. (2022). Opportunities for variable application rate of nitrogen under spatial water variations in rainfed wheat systems – an economic analysis. [Paper under review in Precision Agriculture].

Por todo ello, se autoriza la presentación de la tesis doctoral.

**FERERES**  
**CASTIEL**  
**ELIAS -**  
**05393383K**

Firmado  
digitalmente por  
FERERES CASTIEL  
ELIAS - 05393383K  
Fecha: 2022.05.11  
10:48:55 +02'00'

Fdo.: Elias Fereres Castiel

**GOMEZ**  
**CALERO JOSE**  
**ALFONSO -**  
**30533940Z**

Firmado digitalmente  
por GOMEZ CALERO  
JOSE ALFONSO -  
30533940Z  
Fecha: 2022.05.20  
12:59:10 +02'00'

Fdo.: José Alfonso Gómez Calero





# Contents

<b>1 Introduction</b>	<b>2</b>
1.1 The spatial dilemma within rainfed crop systems modelling . . . . .	2
1.2 New data assimilation methods . . . . .	3
1.3 Crop-water modelling approaches and opportunities to simulate spatial water variations at crop field level . . . . .	4
1.4 Opportunities for site-specific crop management . . . . .	5
1.5 Thesis research framework . . . . .	6
FIGURES - Chapter 1 . . . . .	8
<b>2 Water modelling approaches and opportunities to simulate spatial water variations at crop field level</b>	<b>10</b>
2.1 Introduction . . . . .	11
2.2 Methodology . . . . .	14
2.2.1 Review approach . . . . .	14
2.2.2 Model selection and soil-plant water processes description . . . . .	14
2.3 Modelling soil-plant water processes . . . . .	15
2.3.1 The pre-infiltration phase . . . . .	15
2.3.2 Infiltration . . . . .	18
2.3.3 Surface-water flow . . . . .	21
2.3.4 Evaporation . . . . .	22
2.3.5 Transpiration and root water uptake . . . . .	27
2.3.6 Redistribution and drainage . . . . .	31
2.3.7 Capillary rise . . . . .	32
2.3.8 Subsurface lateral flow . . . . .	33
2.3.9 Solute transport . . . . .	34
2.4 Crop and hydrologic models: what sets them apart? . . . . .	36
2.5 Opportunities to simulate spatial water variation . . . . .	38
FIGURES - Chapter 2 . . . . .	42
APPENDIX - Chapter 2 . . . . .	45
SUPPLEMENTARY MATERIAL - Chapter 2 . . . . .	46
<b>3 Using NDVI for the assessment of canopy cover in agricultural crops within modelling research</b>	<b>60</b>
3.1 Introduction . . . . .	61
3.2 Materials & Methods . . . . .	63
3.2.1 Meta-analysis . . . . .	63
3.2.2 A case study – testing NDVI-CC correlations in simulations with the AquaCrop model . . . . .	66
3.3 Results & Discussion . . . . .	68
3.3.1 General results of the meta-analysis . . . . .	68

3.3.2	Applications of NDVI-CC correlations . . . . .	71
3.3.3	Application to the simulation of wheat yields . . . . .	73
3.3.4	Closing remarks . . . . .	74
3.4	Conclusion . . . . .	76
FIGURES - Chapter 3	. . . . .	77
TABLES - Chapter 3	. . . . .	83
APPENDIX - Chapter 3	. . . . .	85
<b>4</b>	<b>Simulating water lateral inflow and its contribution to spatial variations of rainfed wheat yields</b>	<b>88</b>
4.1	Introduction . . . . .	89
4.2	Materials & Methods . . . . .	90
4.2.1	Experimental sites . . . . .	90
4.2.2	Sampling scheme and experimental design . . . . .	91
4.2.3	Modelling approach . . . . .	96
4.2.4	Statistical analysis . . . . .	102
4.3	Results . . . . .	102
4.3.1	Experimental data . . . . .	102
4.3.2	HYDRUS run-off simulation and ANN-LIF forecasting . . . . .	103
4.3.3	Yield simulations with AquaCrop . . . . .	104
4.4	Discussion . . . . .	105
4.4.1	SWC measurements and lateral inflow calculations . . . . .	105
4.4.2	HYDRUS simulations and lateral inflow predictions . . . . .	107
4.4.3	Contribution of lateral flows to crop yield . . . . .	109
4.4.4	Agronomic implications . . . . .	112
4.5	Conclusion . . . . .	114
FIGURES - Chapter 4	. . . . .	115
TABLES - Chapter 4	. . . . .	124
APPENDIX - Chapter 4	. . . . .	126
<b>5</b>	<b>Opportunities for variable application rate of nitrogen under spatial water variations in rainfed wheat systems – an economic analysis</b>	<b>132</b>
5.1	Introduction . . . . .	133
5.2	Materials & Methods . . . . .	135
5.2.1	Experimental conditions and on-farm data collection . . . . .	135
5.2.2	Lateral inflow (LIF) zones mapping . . . . .	136
5.2.3	Intra-plot spatial assessment of Yield Gaps . . . . .	137
5.2.4	Quantifying the relative advantage of VAR adoption . . . . .	138
5.2.5	Economic modelling and analysis of future scenarios . . . . .	140
5.3	Results . . . . .	143
5.3.1	YG analysis . . . . .	143
5.3.2	The economics of VAR adoption . . . . .	144
5.3.3	Economic trade-offs for viability of VAR . . . . .	145
5.4	Discussion . . . . .	146
5.4.1	Crop yields and yield gaps . . . . .	146
5.4.2	The economics of VAR adoption . . . . .	147
5.4.3	Methodological considerations and practical issues . . . . .	150
5.5	Conclusion . . . . .	152
FIGURES - Chapter 5	. . . . .	154
TABLES - Chapter 5	. . . . .	159

APPENDIX - Chapter 5 . . . . .	164
<b>6 Conclusions</b>	<b>170</b>
<b>References</b>	<b>176</b>
<b>Summary</b>	<b>216</b>
<b>Acknowledgements</b>	<b>222</b>
<b>Curriculum Vitae</b>	<b>226</b>

# List of Figures

1.1	Schematic representation of grain yield of wheat, biomass at harvest, and harvest index, in relation to the proportion of available water used by flowering time. The scale of the y-axis is arbitrary because it varies according to each crop parameter. Figure adapted from Passioura (2002). . . . .	8
2.1	Chronological evolution of modelling approaches and theoretical fundamentals; in blue: methodological approaches related to infiltration processes and soil-water movement; in orange: related to evaporation; in green: to transpiration; in dark-purple: plant-root water uptake; in dark-red: crop growth simulation models; in dark-pink: hydrology based models. PE means Penman equation, 2-stage-M corresponds to the 2 stage method proposed by Ritchie. (1972), PT means Priestley and Taylor, HG represents the Hargreaves equation and PM is Penman-Monteith. FSPM means 'functional structural plant modelling. The Beer-Lambert Law, which embraces a wider scope than crop-hydrological issues but it has influenced many modelling approaches of evaporation and transpiration is represented in black color. All horizontal arrows refer to the time-scale below, except in the case of 'preferential flow' and '3D root modelling'. . . . .	42
2.2	Number of models simulating a specific process (N=12), the most common modelling approaches used. The horizontal bars show the number of models that simulate (use) the corresponding process (approach). 'NLD' means that non-linear differential equations are used in the estimation of the extraction sink term (as an alternative to linear or exponential approaches). . . . .	43
2.3	'Spatial scores' (left); spatial water processes considered in each model (right). Spatial scores represent the relative amount of spatial processes found in each model (expressed as the amount of spatial processes considered by a model, divided by the total amount of spatial processes that we identified). . . . .	44
3.1	Our case study methodological scheme. Simulations at point based scale were conducted for 28 different sites at three different seasons. . . . .	77
3.2	NDVI plotted against CC for all datasets. The dots are colored according to the data source. Grey vertical lines represent the residuals (error) for each observation and points are sized according to the corresponding error level, the bigger the point the larger the error. . . . .	78
3.3	NDVI data plotted against CC, dots colored according to NDVI source (A) or crop type (D). Cumulative probability curves (B and E) and best fitted regression lines (C and F) for each group of data. . . . .	79
3.4	NDVI data plotted against CC (A, D, G and J), cumulative probability curves (B, E, H and K) and smooth regression lines (C, F, I, L) for each group of data. Dots and lines are colored according to crop species (A-I) or growing season (J-L). Crop species with N<10 were excluded from regressions. . . . .	80

3.5	Statistical indicators of model performance: $R^2$ and RMSE plotted against sampling size (A and B, respectively) and against the number of data sources used for each model (C and D, respectively). Both indicators followed a logarithmic response curve, negative for the case of $R^2$ and positive for RMSE. All regression coefficients were significant and $R^2$ values ranged around 0.6 (results not shown). . . . .	81
3.6	A) Evaluation of simulated CC in comparison with the ground measurements taken in Trial A. The bars represent the RMSE, the Pearson correlation coefficient (r) and the Willmott's index of agreement (d) of estimated CC at each observation point; B) Simulated vs. observed yield (Mg/ha). Green dots correspond to the estimated yield values obtained from yield mapping (i.e. according to the equation of Reitz & Kutzbach. (1996)) plotted against manual samples taken in Trial A, orange dots represent (AquaCrop) simulated vs. observed yields, units are expressed in Mg/ha; The satellite NDVI estimated CC time-series, plotted in calendar days, and the CC curve assimilated into the AquaCrop model: C) Point B2 in Trial B and D) Point C1 in Trial D. . . . .	82
4.1	Visual symptoms of early crop senescence apparently caused by spatial water variations in Córdoba, Spain. The higher elevation zones show yellowing patterns due to lower water availability, which limits crop yield (Sadras et al., 2016). Photo credits: T. R. Tenreiro. . . . .	115
4.2	Maps of the experimental catchments: the elevation maps with contour lines of catchment one and two, respectively (A and B); the Normalized Flow Accumulation Index (NFAI) rasterized with 5m spatial resolution for catchment one and two, respectively (C and D). Sampling zones (A, B, C) and sampling points (A1, A2, A3, B1, B2, B3, C1, C2, C3) are represented by solid black lines and purple circles, respectively. . . . .	116
4.3	Daily mean values for (A) rainfall (dashed lines represent cumulative rainfall computed from the 1st of October, values shown by right vertical axis), (B) temperature and (C) reference evapotranspiration (ET <sub>0</sub> ), respectively expressed in mm, degree Celsius and mm. Lines represent season time-series expressed in terms of days after sowing (DAS). Grey areas represent the monitoring time window of each observation year (i.e., from DAS 20 to 157 in 2019/20 and from DAS 5 to 158 in 2020/21). Daily mean temperature values are represented by the heavy solid lines. Maximum and minimum temperature values are respectively represented by a solid and a dashed skinny line. Each season values are shown in different colors (2019/20 in green and 2020/21 in brown). . . . .	117

4.4 Sketch of the methodological design: Two field experiments were conducted, each sampling point was defined according to a spatial analysis (step 0). A daily calendar of lateral water inflow (LIF) was calculated based on field observations (step 1) and simulations of LIF were conducted for 30 years through a hydrologic modelling approach (step 2). LIF predictions were assimilated into the crop modelling stage (step 3) and the net yield responses were simulated (step 4). Dashed lines delineate different methodological stages, rounded parallelograms indicate experimental sites and data, solid line rectangles indicate different sub-steps, solid line circles represent simulation tools, losenge and arrows indicate conditional steps. More information on the simulation settings of HYDRUS-1D and AquaCrop is respectively provided by Šimůnek et al. (2018) and Steduto et al. (2009). Additional details related to the use of artificial neural networks (ANN) for hydrological modelling are provided by Maier & Dandy (2000). . . . .	118
4.5 Corrected soil water content ( $SWC_c$ ) for each sampling zone (A-C), error bars are shown for daily $SWC_c$ values. Daily lateral inflows (LIF) per sampling zone (inferred from probes data). Values are expressed in mm. . . . .	119
4.6 HYDRUS-1D simulations performance according to the means of measured soil water content and the best fitted soil parameters. "n.a." indicates non-applicable situations, which correspond to the sampling point A3 and C3 that were not considered in catchment one (Figure 4.2). Both the Nash-Sutcliffe and the $R^2$ coefficients are indicated by the left axis, while the RMSE is indicated by the right axis. Mean $R^2$ values ranged from 0.65 ( $\pm 0.13$ ) in 2019/20 to 0.93 ( $\pm 0.03$ ) in 2020/21 and RMSE ( $cm^3 cm^{-3}$ ) from 0.04 ( $\pm 0.01$ ) to 0.03 ( $\pm 0.01$ ), by the same order. . . . .	120
4.7 The feedforward Artificial Neural Network (ANN) used for forecasting lateral inflow (LIF) over 30 years: A) schematic representation of the $R^2$ assessment based on trial-and-error procedure that was taken for model architecture delineation; B) three-dimensional representation of the $R^2$ values obtained for each combination of model structure (the best performing ANN corresponded to an ANN composed by six different features and four hidden layers); C) ANN statistical evaluation, observed LIF plotted against predicted LIF (values are expressed in mm and represent cumulative LIF), blue dots correspond to the training set (2019/20) and red triangles to the testing set (2020/21). Both the training and the testing set fitted into a linear regression, respectively expressed as $y = 1.432x + 0.528$ and $y = 1.122x + 2.919$ , and subsequently with a $R^2$ of 0.97 and 0.78, and a RMSE of 10.1 and 14.8 mm; D) the schematic representation of the best performing ANN, solved in 9962 steps: green circles represent input features, grey squares indicate the hidden layers' coefficients, each node is represented by a grey circle, the blue small circles represent the transfer functions and the blue squares indicate the linked weight factors. Selected features (i.e., x1, x3, x5, x6, x10, x11) are shown in Table 4.4. . . . .	121

4.8	Cumulative lateral inflow ( <i>CUM.LIF</i> ) and cumulative precipitation ( <i>CUM.P</i> ), respectively represented by solid and dashed lines (expressed in mm). Facet-plots show the two forms of water supply to zone C in catchment two (shown in Figure 4.2). Values were computed in daily time-steps over a period of 30 crop seasons, from sowing to harvesting date. <i>CUM.P</i> values were computed from daily weather records for Cordoba [1990–2020], obtained from the same weather station introduced in Section 4.2.2. <i>CUM.LIF</i> values were forecasted by the ANN model shown in Figure 4.7. Vertical black bars represent daily LIF events (daily LIF magnitudes are multiplied by a factor of 4.5 for better visualization). . . . .	122
4.9	Main AquaCrop simulation outcomes: A) cumulative probability distribution curves for simulated $Y_2$ yields ( $\text{Mg ha}^{-1}$ ) in the two significantly distinct zones (i.e., A and C in catchment two); B) cumulative probability of $NYR_{LIF}$ , expressed in $\text{Mg ha}^{-1}$ ; C) time-series of $NYR_{LIF}$ expressed in $\text{Mg ha}^{-1}$ over 30 years; D) cumulative probability of $NYR_{LIF-rel}$ , expressed in %. Values correspond to zone C in catchment two for the period of 1990–2021. Dashed lines represent the median values, in plots B and D (i.e., $0.265 \text{ Mg ha}^{-1}$ and 6.3%, respectively), and the mean $NYR_{LIF}$ in plot C (i.e., $0.383 \text{ Mg ha}^{-1}$ ). . . . .	123
5.1	Experimental fields. A) Total farm area (320 ha); B) Digital Elevation Model (DEM), values expressed in m amsl.; C) Soil types according to USDA classification system; D) TOPMODEL Topographic Index (TMI) as described in detail by Beven et al. (2021), values are unitless; E) Yw zones map (i.e., 'LIF' and 'no-LIF' zones). According to Tenreiro et al. (2022), 'LIF zones' are characterized by significant water supplied through lateral inflow, while 'no-LIF zones' are characterized by null or insignificant lateral inflow. . . . .	154
5.2	Water supply: seasonal precipitation (P) and lateral inflow (LIF) for the experimental dataset. Values are expressed in mm. More information is provided in Tenreiro et al. (2022). . . . .	155
5.3	Yield variations (values expressed in $\text{Mg DM ha}^{-1}$ ): A) Cumulative probability distribution of actual yields ( $Y_a$ ) for each season (2015/16 to 2020/21) and within each zone (i.e., solid lines correspond to 'no-LIF zones' while dashed lines relate to 'LIF-zones'); B) Cumulative probability distribution of yield gaps ( $YG$ ) for each season and within each zone. Vertical lines indicate Yw levels obtained from Tenreiro et al. (2022) simulations. . . . .	156
5.4	The negative relation between return time on investment (ROlt, expressed in years) and the annual wheat area (expressed in ha). The annual wheat area is subjected to the farm conditions supporting our analysis (i.e., an average LIF area share of 17.4%). Scenario S-2 represents the effect of a 35% increase in LIF area share within the cropping system. More information regarding the remaining scenarios considered is provided in Table 5.2. . . . .	157
5.5	The 10-year overall total gain (OTG) under multiple scenarios (units expressed in thousand €). The OTG was estimated by solving the NPV series over a period of 10 years, which was a function of both the discount rate used and the price/payment scenarios adopted (Table 5.2). . . . .	158

# List of Tables

3.1 Sampling groups characterization: N [data sources], N [data points] and N [crop species] refer respectively to the number of data sources, data points and crop species included in each sampling group. The canopy structure was expressed by the leaf angle distribution parameter ( $\chi$ ), obtained from Campbell & Norman. (1998). . . . .	83
3.2 Model coefficient estimates of each group for linear and quadratic regressions. The best fitted regressions are highlighted in bold, which were selected as those maximizing $R^2$ while keeping a minimal amount of significant coefficients. Root mean square error (RMSE) was estimated for the best fitted regression of each group. Significance codes: '***' 0.1% '**' 1% '*' 5%. . . . .	84
3.3 Simulation error assessment. Simulated yield RMSE and Coefficient of Variation (CV). '% Total yield variation' corresponds to the fraction of total observed variation that was captured by the assimilation of CC into modelling simulations. . . . .	84
4.1 Main geomorphological attributes within each sampling zone (mean values and standard deviations). The standard deviations are presented in parentheses. Codes: EC <sub>a</sub> = soil apparent electrical conductivity, %Clay = percentage of clay content, %Sand = percentage of sand content, NFAI = "Normalized flow accumulation index" (Tarboton et al., 1991; Jenson & Domingue., 1988). The amount of chosen sampling points was supported by Chanzy et al. (1998). %Clay, %Sand and bulk density were averaged for the three zones in catchment-2 according to farm records. . . . .	124
4.2 Crop data used for the parameterization of the HYDRUS and the AquaCrop models. CC <sub>MAX</sub> is maximum green canopy cover (Steduto et al., 2009), used for parameterization of the soil cover partitioning method, as described in Tenreiro et al. (2020). Mean sowing rate was 180 kg ha <sup>-1</sup> for all trials. The standard deviations are presented in parentheses. . . . .	124
4.3 The parameter intervals used for model optimization. Values were adjusted during the optimization process for each individual point (Figure 4.2). . . . .	125
4.4 Artificial neural network (ANN) potential predictors assessed through principal component analysis (PCA) and trial-and-error procedure (Roadknight et al., 1997). Main PCA outcomes are shown in Appendix-A3. More information regarding the PCA is provided in Supplementary material. . . . .	125
4.5 Mean soil water content (SWC), mean daily lateral inflow ( $LIF_n$ ) and mean cumulative LIF ( $CUM.LIF_n$ ) over season, for the two years of field data. All values are expressed in mm. The standard deviations are presented in parentheses. Mean values followed by a common letter are not significantly different according to the HSD-test conducted at the 5% level of significance ( $p - value < 0.05$ ). . . . .	126

5.1	Crop management information: Fertilizers DAP and CAN correspond to diammonium phosphate (18-46-0) and calcium ammonium nitrate (24-0-0 + 8% Ca), respectively. Urea (46% N) was top-dressing applied. Yw values were obtained from Tenreiro et al. (2022). Sowing and harvest dates indicate day plus month and mean N applied is expressed in kg N ha <sup>-1</sup> . Water-limited yield (Yw) values are expressed in Mg DM (grain) ha <sup>-1</sup> . Mean plant density was 230 plants ha <sup>-1</sup> . 'LIF': downslope zones with significant amount of water supplied through lateral flow coming from upslope areas of the same field; 'No-LIF': upslope zones where no significant amount of water is supplied through lateral flow. More information on LIF and No-LIF zones is provided in Tenreiro et al. (2022), see also section 5.2.2. . . . .	159
5.2	The ten different scenarios considered to analyse the impact of price support policies, extra direct payments, and different market prices on VAR economic relevance. For each scenario, three different sub-scenarios were considered by setting the discount rate ( $\gamma$ ) at 2.5%, 5% and 7.5%. LPP and DPA mean linked to production payment and direct payment on crop area, respectively. $AC_{VAR}$ means annual cost of VAR use and $II_{VAR}$ is the initial acquisition cost of VAR technology (Appendix – Table A1). LPP was based on a total available budget equal to 46€ ha <sup>-1</sup> (as considered in S-7 and S-8 scenarios) divided by the average yield. . . . .	160
5.3	Mean yield values, yield gaps, and coefficients of variation. The standard deviations are presented in parentheses. Means followed by a common letter are not significantly different, according to the HSD-test conducted at the 5% level of significance (p-value < 0.05). . . . .	161
5.4	The economics of VAR adoption, under spatial water variations in rainfed wheat systems according to our experimental conditions. $ADGM_t$ is the annual differential gross margin in year $t$ , $ADR_t$ is the annual differential revenue in year $t$ , and $ADC_t$ is the annual differential cost in year $t$ . $DGM$ , $ADR$ , $ADC$ are, respectively, the average differential gross margin, the average differential revenue and the average differential costs. Economic scenarios are described in Table 5.2. Values are expressed in € ha <sup>-1</sup> of LIF area. . . . .	162
5.5	This table shows the best fitted models' coefficient-a values (i.e., a power model adjusted to the data), the minimum area for adoption of VAR (expressed in ha per year), the mean return on investment time ( $ROI_t$ , expressed in years) and the mean overall total gain ( $OTG$ , expressed in thousand € per 10-year period). The minimum area for adoption corresponds to the value below which the return on investment takes longer than the amortization of equipment (i.e., 10 years). $OTG$ means followed by a common letter are not significantly different, according to the HSD-test conducted at the 5% level of significance. The internal rate of return ( $IRR$ ) values assume an average annual sown area equal to 92 ha. . . . .	163

## List of Abbreviations

- $ADC$ : annual differential cost ( $\text{€ ha}^{-1} \text{ year}^{-1}$ );
- $ADGM$ : annual differential gross margin ( $\text{€ ha}^{-1} \text{ year}^{-1}$ );
- $ADR$ : annual differential revenue ( $\text{€ ha}^{-1} \text{ year}^{-1}$ );
- $ANN$ : artificial neural network;
- $CC$ : canopy cover, expressed in %;
- $CP$ : capacitance probe;
- $CUM.ET_0$ : season cumulative evapotranspiration of reference ( $ET_0$ ), expressed in mm;
- $CUM.LIF$ : season cumulative lateral inflow (LIF), expressed in mm;
- $CUM.P$ : season cumulative precipitation, expressed in mm;
- $DEM$ : digital elevation model (raster);
- $DGM$ : differential gross margin ( $\text{€ ha}^{-1} \text{ year}^{-1}$ );
- $FAI$ : flow accumulation index (the absolute number of upslope cells flowing to each assigned cell of the DEM raster);
- $GY$ : grain yield, expressed in Mg (dry mass)  $\text{ha}^{-1}$
- $LIF$ : lateral inflow, expressed in mm;
- $LIF.MWP = LIF$  marginal water productivity (expressed in  $\text{kg GY ha}^{-1} \text{ mm}^{-1}$ );
- $K_{SAT}$ : saturated hydraulic conductivity, expressed in  $\text{mm day}^{-1}$ ;
- $NDVI$ : normalized difference vegetation index;
- $NFAI$ : normalized flow accumulation index;
- $NP$ : neutron probe;
- $NYR_{LIF}$ : net yield response to LIF, expressed in  $\text{Mg GY ha}^{-1}$ );
- $P$ : daily precipitation, expressed in  $\text{mm day}^{-1}$ ;
- $PAW_0$ : plant available water at sowing date, expressed in mm;
- $PF.LIF$ : post-flowering LIF (the fraction of  $CUM.LIF$  taking place at post-flowering stages), expressed in %;
- $Relative.T$ : mean relative crop transpiration (estimated as the season average of daily crop actual transpiration divided by potential transpiration), expressed in %;
- $SWC$ : soil water content, expressed in mm;
- $VAR$ : variable application rate;
- $Ya$ : actual yield ( $\text{Mg ha}^{-1}$ );
- $YG$ : yield gap ( $\text{Mg ha}^{-1}$ ), i.e., the difference between the water-limited yield and the actual yield;
- $Yw$ : water-limited yield ( $\text{Mg ha}^{-1}$ );

# Chapter 1

## Introduction

### 1.1 The spatial dilemma within rainfed crop systems modelling

Rainfed agriculture plays a decisive role in world food production, accounting for more than 75% of global cropped area, and being responsible for more than 60% of total cereals' production (Cassman et al., 2003; Connor & Mínguez, 2012). Sustaining food production by rainfed crops in the years ahead will require productivity gains in resource use (Fischer & Connor, 2018). Specific challenges consist of estimating the magnitude and thus the value of yield gaps, identifying limiting factors, and implementing profitable and sustainable strategies (Fischer, 2015; Rattalino-Edreira et al., 2018; Silva, 2017).

Recently, yield gaps in rainfed farming have been assessed with the use of crop simulation models considering the most representative biophysical conditions and, to a lesser extent, crop management practices observed in farmers' fields (Spiertz, 2014). However, the considerable spatial variability that exists in soil hydraulic properties within a field, and the accurate modelling of crop heterogeneity requires assessing the spatial variations of water as they affect crop behaviour (Nielsen et al., 1973; Miller et al., 1988).

Extraordinary advances in computer engineering and programming languages have been accomplished, particularly over the last three decades. These have facilitated crop modelling, which have led to the adoption of such tools for many applications in agronomy (Passioura, 1996; Jones et al., 2017). However, recent advances have not yet succeeded in scaling up mechanisms

from point to field level in crop models. While significant progress has been achieved in the engineering aspects of spatial variation, such as increasing spatial resolution of data systems, variable rate technologies, and automation, much less effort has been dedicated to the simulation of within field crop responses to spatial variations.

In this thesis, ‘field level’ (or ‘field scale’) is defined as the entire crop plot, and thus the terms ‘intra-plot’ and ‘within-field’ are used as synonyms. While ‘point scale’ focuses on the dominant characteristics to represent an area of interest, ‘field scale’ considers the entire field/plot area and the inherent spatial variation of growing conditions.

Most studies on rainfed yield gaps ignore intra-plot variability (Fischer et al., 2014; Guilpart et al., 2017; Lobell et al., 2009; Schils et al., 2018). This is partially due to data availability constraints (Beza et al., 2017) and to the limitations of crop models to simulate processes such as spatial water distribution (Tenreiro et al., 2020). Essentially, if crop models are to be used in assisting crop management, they may greatly benefit from spatial water modelling approaches capable of accurately representing and simulating within-field variation of water-related processes.

## 1.2 New data assimilation methods

New data assimilation methods can improve the assessment of crop attributes, which reveal new opportunities for modelling spatial variations. Examples are the opportunities of very-high-resolution remote sensing, thanks to the advent of unmanned aerial vehicles and other airborne platforms, novel computer vision techniques, machine learning methods, spectral analysis, and object-based classification algorithms (Bendig et al., 2015; Berni et al., 2009; Chianucci et al., 2016; Gao et al., 2020; Tenreiro et al., 2021; Viña et al., 2011; Waldner et al., 2019).

For the simulation of water processes at crop field level, several methodologies for data assimilation, geospatial simulation, visualization and validation of models have been proposed (e.g., geospatial interpolation of

point based simulations, zonal statistics applied to mapped simulation results, integration of modelling with remote sensing, and other new data assimilation methods). However, most of these cases have been focused at regional scales and do not address within-field spatial variation (Droogers & Bastiaanssen, 2000; Grassini et al., 2015; Han et al., 2019; Jia et al., 2011; Lobell & Azzari, 2017; Lobell et al., 2015; Lorite et al., 2013; Sadler & Russell, 1997; Shu et al., 2018; Zwart & Bastiaanssen, 2007). Other promising cases reveal some advances in the spatial simulation of water and vegetation (Moiling et al., 2005; Booker et al., 2015; Wallor et al., 2018), but still neglecting spatial behaviour of yield determining factors, such as variations of the harvest index caused by differences in available soil water (Figure 1.1).

In terms of spatial variations in water distribution, and their effects on crop growth and development, most of the geostatistical methods applied to point-based (or partially distributed) models smooth considerably the actual spatial heterogeneity. This occurs because lateral water movement and cause-effect relations between neighbouring areas are ignored in the water balance calculation schemes. In addition, relying solely on geostatistics to deal with spatial heterogeneity does not resolve the existing knowledge gaps regarding the driving mechanisms of spatial variations. This issue was raised by Nielsen & Wendoroth (2003), who suggested that statistical methods should not replace research inventiveness in the assessment of spatial and temporal variations.

### **1.3 Crop-water modelling approaches and opportunities to simulate spatial water variations at crop field level**

The present doctoral thesis reviewed some of the most widely adopted crop (e.g., WOFOST, DSSAT, APSIM, DAISY, STICS, AquaCrop and MONICA) and hydrologic models (HYDRUS-1D, HYDRUS-2D, SWAP, MIKE-SHE and SWIM), from the standpoint of identifying opportunities for simulating spatial water variations at crop field level through the incorporation of both surface and subsurface lateral flows as they affect crop performance (Tenreiro et al., 2020;

2022).

The incorporation of lateral flows within the water balance allows a better simulation of the mechanisms that determine yield variations over space. However, this requires innovative experimental approaches which should account for spatially distributed processes and which should be related to the geomorphological properties with implications for plant available water.

Data collection and field experimentation must be conducted at “real scales”, which is relatively costly and difficult to replicate over long periods of time (Sadras et al., 2020). Therefore, it is accepted that the combination of both experimentation and simulation modelling is a valid strategy for making progress (Kirkegaard & Hunt, 2010). By combining field experimentation, simulation models, and the use of new data assimilation methods (i.e., remote sensing or artificial neural networks), it should be possible to investigate the relevance of lateral flows.

## 1.4 Opportunities for site-specific crop management

Progress in the simulation of crop response to spatial variations can deliver new opportunities for better crop production through site-specific management (Basso & Antle, 2020). Site-specific management, linked to the concept of precision agriculture, is defined as the agricultural crop management that is conducted at a lower spatial scale than the whole field (Mulla & Schepers, 1997). This implies that intra-plot variations are considered, and that crops are managed accordingly, considering the site-specific conditions that vary within a single field.

In fields of undulating topography, where rainfed crops experience different degrees of water stress due to spatial water variations caused by lateral flows, yields vary spatially within the same field (Halvorson & Doll, 1991; Tenreiro et al., 2022).

The pattern of crop water use can be greatly affected by management (Passioura, 2002), as over fertilization may result in exceedingly vigorous cereal

crops that are prone to consume an excessive amount of water in the vegetative stage before yield-defining critical stages (Figure 1.1). This may increase water-stress during grain formation reducing the harvest index, especially under rainfed conditions when crop yields are water-limited (van Herwaarden et al., 1998; Figure 1.1).

Water shortages during grain formation results in yield losses (Figure 1.1) because the crop initiates senescence prematurely in zones with the least available water. Therefore, the spatial variations in soil water supply can justify a variable fertilization rate over space. From both an economic and an environmental perspective (Lowenberg-DeBoer & Erickson, 2019; Robertson et al., 2008), different nutrient requirements and application rates represent an opportunity for better farm management with productivity gains in resource use and net margins (Fischer & Connor, 2018; Nielsen & Halvorson, 1991; Sadras, 2002; Whelan & McBratney, 2000).

## 1.5 Thesis research framework

The present thesis focused on the “spatial dilemma” regarding crop simulation under spatial heterogeneous conditions. This thesis builds upon the main hypotheses that point-based crop modelling is a limited representation of the field, that the spatial interpolation of point simulations ignores the existing hydrological interactions and feedback between adjacent areas within the field, and that up-scaling such processes is a promising step to integrate crop simulation and decision-making in site-specific management.

**Main objective** To study the spatial crop-water variations in rainfed wheat systems in Mediterranean conditions, and to evaluate the potential of hydrologic and crop simulation modelling to predict within-field yield variation and its implications for site-specific management.

Within this context, the following subsequent research objectives were addressed:

## **First objective**

1. To investigate opportunities to simulate spatial water variations at crop field level by reviewing the main crop-water modelling approaches and identifying opportunities to simulate spatial water variations at crop field level.

This specific objective will be covered in chapter two.

## **Second objective**

2. To explore new data assimilation methods within crop modelling research. Combining new data assimilation methods with crop simulation modelling for determining crop yield variability at field level.

This specific objective will be covered in chapter three.

## **Third objective**

3. To evaluate the agronomic relevance of spatial water variations by:
  - 3.1) assessing experimentally the magnitude and frequency of water lateral flows occurring in crop fields of undulating topography;
  - 3.2) simulating water lateral inflow over multiple years and its contribution to spatial variations of rainfed wheat yields.
  - 3.3) determining the net contribution of lateral water inflows to spatial variations of rainfed wheat yields in fields of undulating topography.

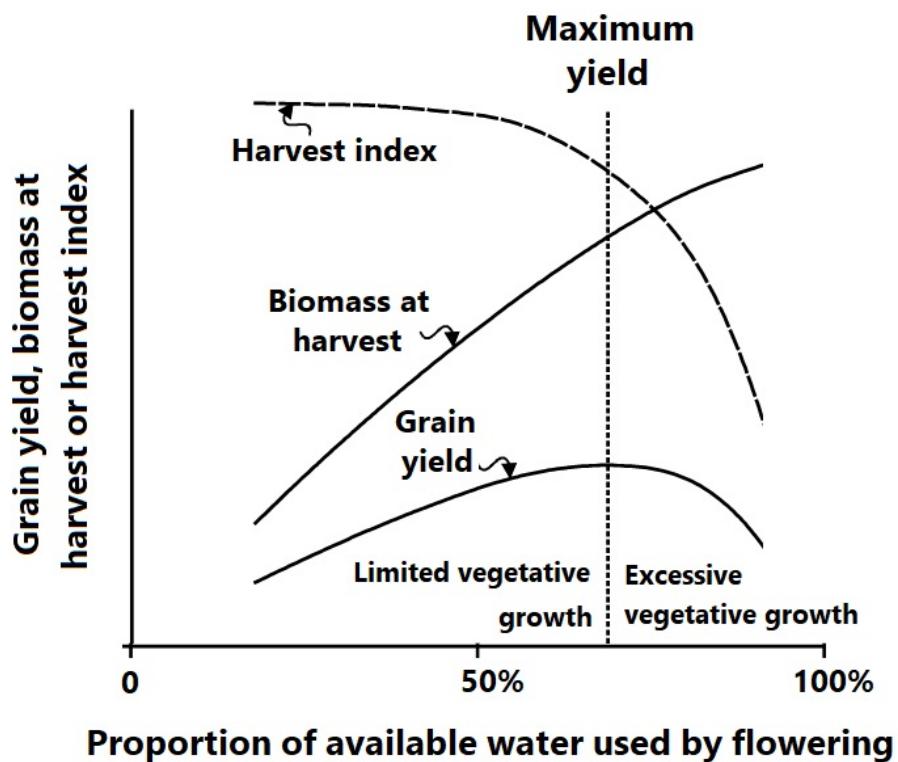
This specific objective will be covered in chapter four.

## **Fourth objective**

4. To analyse the economic relevance of site-specific management as a strategy to deal with spatial water variations by investigating the main opportunities for variable application rate of nitrogen under spatial water variations in rainfed wheat systems from an economic perspective.

This specific objective will be covered in chapter five.

## FIGURES - Chapter 1



**Figure 1.1:** Schematic representation of grain yield of wheat, biomass at harvest, and harvest index, in relation to the proportion of available water used by flowering time. The scale of the y-axis is arbitrary because it varies according to each crop parameter. Figure adapted from Passioura (2002).



# Chapter 2

## Water modelling approaches and opportunities to simulate spatial water variations at crop field level

This chapter has been published as:

Tenreiro, T. R., García-Vila, M., Gómez, J. A., Jimenez-Berni, J. A., & Fereres, E. (2020). Water modelling approaches and opportunities to simulate spatial water variations at crop field level. Agricultural Water Management, 240, 106254.

### Abstract

Considerable spatial variability in soil hydraulic properties exists within a field, even in those considered homogeneous. Spatial variability of water as a major driver of crop heterogeneity gains particular relevance within the context of precision agriculture, but modelling has devoted insufficient efforts to scale up from point to field the associated 'cause-effect' relations of water spatial variations. Seven crop simulation models (WOFOST, DSSAT, AP-SIM, DAISY, STICS, AquaCrop and MONICA) and five hydrologic models (HYDRUS-1D, HYDRUS-2D, SWAP, MIKE-SHE and SWIM) were selected and their water modelling approaches were systematically reviewed for comparison. Crop models rely mainly on 'discrete' and empirical approaches for modelling soil water movement while hydrologic models emphasize more 'continuous' and mechanistic ones. Combining both types of models may not be the best way forward as none of the models consider all of the processes which are relevant for the simulation of spatial variations. Hydrologic models pay more attention to spatially variable water processes than crop simulation models, although their focus is on scales higher than the field which is the relevant scale for assessing the influence of such variations on crop behaviour. Opportunities for progress in the spatial simulation of water processes at field level will probably come from two different directions. One implying a stronger synergism between both model families by using continuous-type approaches to simulate some mechanisms in existing crop models, and the other through the integration of lateral flows in the simulation of discrete water movement approaches.

## 2.1 Introduction

An inherent action of biological sciences is the conceptual representation of systems through hierarchical levels (De Wit, 1982; Loomis et al., 1979). In agronomy related studies, it implies different levels of complexity determined by the nature of the issue addressed (Ahuja et al., 2019). Over the last 60 years, crop scientists have dedicated particular attention to modelling in an attempt to mathematically represent the functioning of agricultural systems at different levels of complexity, and to simulate their response to multiple factors in an 'easy-fast' and 'low-cost' way (Carberry, 2003; Fischer & Connor, 2018; Jin et al., 2018; Jones et al., 2017a; Lobell et al., 2009).

Conceptually, models can be divided into 'functional-empirical' or 'mechanistic' and distinguished according to their spatial scale, being classified as 'point-based' or 'distributed' (ASCE, 1982; Passioura, 1996; Thomas and Smith, 2003). While 'engineering-oriented' models tend to be classified as functional-empirical, 'science-oriented' models are mostly considered mechanistic (or process-based). While point-based scales ignore spatial variability by averaging or using 'dominant' characteristics to model an area of interest, distributed scales consider the spatial distribution of resources and the consequent crop response. Functional-empirical models have shown potential to support benchmarking, decision and policy making at different temporal-spatial scales (Boote et al., 1996; García-Vila et al., 2009; Mateos et al., 2002; Passioura, 1973). Mechanistic models have been mostly used to assist plant breeding for specific environments (Fischer and Connor, 2018; Struik, 2016; Yin & Struik, 2007), the identification of global yield- gaps (Byerlee et al., 2014; Boogaard et al., 2013; Grassini et al., 2015); (<http://www.yieldgap.org/>), and for agro-ecological resource management (Boote et al., 1996; Booker et al., 2015; Fischer et al., 2002; Thorp et al., 2008).

The extraordinary advances in computer engineering and programming languages, particularly over the last three decades, have intensified the modelling processes contributing to an increased adoption of such tools for many applications (Jones et al., 2017a; Passioura, 1996; Seidel et al., 2018;

Thorp et al., 2012). The use of models in large cooperative efforts (such as The Agricultural Model Intercomparison and Improvement Project; [www.agmip.org](http://www.agmip.org)) has intensified recently, too, partially in response to new needs (Asseng et al., 2014). However, recent advances have not yet succeeded to scale up mechanisms from point to field level in crop models (Ahuja et al., 2019; Fischer and Connor, 2018). Such limitation may be leading to an impasse in modelling, compromising the adoption of these tools, mostly in the context of precision agriculture (Jones et al., 2017b). In fact, while significant advances have been made in the engineering aspects of precision agriculture, such as increasing spatial resolution, variable rate technologies and automation, much less effort has been devoted to understand the crop mechanisms in response to spatial variations (Cassman, 1999; McBratney et al., 2005; Monzon et al., 2018).

As considerable spatial variability in soil hydraulic properties exists within a field, even in those considered homogeneous (Nielsen et al., 1973), the accurate modelling of crop heterogeneity requires assessing the spatial variability of water as it affects crop behaviour (Ritchie, 1981; Sadras et al., 2016; Verhagen and Bouma, 1997). This aspect is considered a serious limitation in current crop models but it has received limited attention (Ahuja et al., 2014; Jones et al., 2017b).

In the modelling of water, two ‘families’ of models can be distinguished: crop models and hydrology based models. Both families have been widely used worldwide and different arguments are employed to promote the adoption of each one depending on the specifications of each case-study (Ahuja et al., 2014; Devia et al., 2015; Jones et al., 2017b). While crop models are centered on the growth and development as affected by the environment, hydrological modelling emphasizes mostly systems’ water dynamics at different scales. In this sense, and in regard to water-related processes, crop models tend to be more empirically-based while hydrologic models are more mechanistic. In relation to spatial scales, while crop models are limited to point-based scales, some hydrological models distribute partially water processes. However, practically all distributed models have in fact ‘discrete characteristics’ (e.g. input parameters, boundary conditions) and follow non-linear relations for

multi-dimensional representations that lead to important trade-offs between accuracy and data requirements that must be considered (Passioura, 1996). Just as distributed models rely partially on discrete characteristics, the distinction between mechanistic and empirical models is not clear-cut in this case, as there is a continuum between the two approaches as well, with mechanistic models having always some empirical components.

Essentially, if crop models are to be used to improve water management in precision agriculture, they may greatly benefit from spatial water modelling approaches capable of accurately represent and simulate within-field variation of water-related processes. It is important to reflect if this will be more likely achieved by distributing water processes in crop models (i.e. identifying conceptual gaps in the water balance structure) or by coupling both families.

Contemporary reviews on crop modelling (Ahuja et al., 2019; Boote et al., 2013; Bouman et al., 1996; de Wit et al., 2018; Holzworth et al., 2015; Jin et al., 2018; Jones et al., 2017a,b; Whisler et al., 1986) have tended to cover most main variables governing crop growth and development (at point-based scales), models structure and software details, but do not dwell on the modelling approaches with sufficient detail to be able to identify the main conceptual gaps that constrain the use of models for spatial variable applications. Also, most reviews focus on one or only a few models without reaching out to other different types of models (Boote et al., 1996; de Wit et al., 2018; Holzworth et al., 2015; van Ittersum et al., 2003). Rarely, a single variable, such as water, has been the subject of specific analyses in crop modelling reviews.

In our case, we have focused solely on water because it is a major determinant of spatial heterogeneity in the field (Nielsen et al., 1973, 1987; Ritchie, 1981; Ahuja et al., 1984; Sadler and Russell, 1997; Wallor et al., 2018). In an effort to assist in the scaling up of crop simulation models, we have carried out a systematic review of the approaches taken to simulate water in a number of selected crop and hydrologic models.

## 2.2 Methodology

### 2.2.1 Review approach

Since most of the selected models are based on similar fundamentals, this review followed a 'process-based' structure to avoid repetitive comparisons among models. The analysis was carried out in three consecutive steps:

- 1) selection of models based on a literature search of recent reviews;
- 2) identification and description of plant-soil-water processes addressed by the selected models;
- 3) compilation of results and a comparative analysis.

### 2.2.2 Model selection and soil-plant water processes description

The first step consisted of conducting a web search of the last ten years of published reviews on modelling that were either crop-based or hydrologic-based. Priority was given to 'multiple-species' and 'comprehensive' models since we focused on models complying with the following two criteria: (1) related literature is accessible and clear regarding fundamentals, equations and assumptions; and (2) the calibration and parameterization for multiple field crops is possible.

A total of 42 articles were found, out of which 34 were rejected because they were focused on topics different than a crop model/s review. Eight documents were selected: one technical report (Kirby et al., 2013) and seven scientific papers (Donatelli et al., 2017; Holzworth et al., 2015; Jin et al., 2018; Jones et al., 2017a,b; Rauff and Bello, 2015; Shaw et al., 2013). The following seven 'crop-based' models were selected: WOFOST, DSSAT, APSIM, DAISY, STICS, AquaCrop and MONICA. An equivalent approach was followed to select the hydrology-based models. The initial search yielded a total of 99 results out of which 90 were also rejected for a similar reason as to the crop models. The final nine documents selected (Devia et al., 2015; Dwarakish and Ganasri, 2015; Gao and Li, 2014; Golden et al., 2014; Hallouin et al., 2018; Kauffeldt et al., 2016; Salvadore et al., 2015; Song et al., 2015; Sood and Smakhtin, 2015) referred to

12 models from which a final sample of five models was chosen based on the intended scale of analysis (i.e. small catchment plot or crop field level) and the potential to be coupled with ‘crop-based’ models in regard to water processes. The models are: HYDRUS-1D, HYDRUS-2D, SWAP, MIKE-SHE and SWIM.

Once models were selected, we proceeded with the identification of all processes where water moves along the soil–plant–atmosphere system, called here soil–plant water processes. Subsequently, a description of the modelling approaches followed by the selected models was obtained from the literature going back to the initial publication of each model. Soil–plant water processes were structured following the fate of water in a hypothetical hydrological unit: (1) pre-infiltration, (2) infiltration, (3) surface–water flow, (4) evaporation, (5) root water uptake and transpiration, (6) internal drainage, (7) capillary rise, (8) subsurface lateral flows, and (9) solute transport.

Following a basic description of all models, every water process was described in detail as simulated in each of the models. The fundamentals of the modelling approaches were described based on a literature search that was not limited by any time-frame. All results were synthesized in a table for a comparison among models (Appendix – Table 1). The tabled results are fully integrated with the text, following the same nomenclature and acronyms. Considering our sample ( $N = 12$  models), a descriptive statistical analysis was conducted to explore differences among models in regard to the degree of spatial components and an association plot was produced to justify our discussion.

## 2.3 Modelling soil–plant water processes

### 2.3.1 The pre-infiltration phase

The pre-infiltration phase involves all water processes taking place above the soil surface (i.e. precipitation, irrigation, surface run-on, canopy or mulch interception, and gravitational flow through plant surfaces, commonly called stem-flow). This phase determines the amount of water supplied from

precipitation ( $P$ ), including outflows from the snowpack (DAISY, MONICA, HYDRUS-1D/2D, SWAP, MIKE-SHE) and irrigation ( $I$ ) after subtracting the evaporated fractions of intercepted water by canopy and other surfaces (e.g. mulches). Some models operating at catchment scale (i.e. MIKE-SHE), consider superficial inflow ( $SIF$ ) from run-on as a supply form, too (DHI, 2017b).

Precipitation is taken as an input (Beaudoin et al., 2009; Hansen et al., 1990; Jones et al., 2003; Keating et al., 2003; Nendel et al., 2011; Raes et al., 2009a; Simunek et al., 2008; van Dam et al., 1997; van Van Diepen et al., 1989; Verburg et al., 1996), which can be represented at point- or field-scale, depending on the spatial variability of rainfall and the availability of spatially distributed data (Basso et al., 2001; Thorp et al., 2008; Zhou & Zhao, 2019).

Irrigation, when considered (DSSAT, APSIM, DAISY, STICS, AquaCrop, MONICA, HYDRUS, SWAP, SWIM), must be previously set up within a management module that can be activated when necessary (Boote et al., 1996; García-Vila and Fereres, 2012; Hussein et al., 2011). Irrigation water supply is taken as a net inflow, either assuming no losses (DSSAT, APSIM, DAISY, AquaCrop, MONICA, HYDRUS, SWAP, SWIM) or by subtracting the corresponding application losses (STICS). Four irrigation methods may be considered: (1) Surface; (2) Sprinkler; (3) Drip; and (4) Subsurface drip. Irrigation applications may be simulated in three different ways: (1) a calendar defined by the user; (2) a planned irrigation schedule applying constant or variable rates once a threshold of soil water content is reached; (3) a planned schedule based on multiple-criteria (e.g. crop phenological stage, soil water content, water availability constraints).

The method influences whether irrigation is applied above canopy (i.e. sprinkler pivot) or below (i.e. furrow or drip irrigation). Irrigation applied above canopy implies the simulation of pre-infiltration processes such as canopy or mulch interception and evaporation from plant surfaces (STICS, HYDRUS, SWAP). Some models allow users to define the fraction of soil surface wetted by irrigation (AquaCrop).

The fraction of water intercepted by the canopy may be simulated by an

analogy of the 'Beer-Lambert law' (B-L) (Murphy & Knoerr, 1975). Some models (APSIM) use the B-L approach to simulate 'rainfall attenuation', i.e. interception of rainfall or irrigation water applied over the canopy. The fraction of intercepted water by the canopy (or mulch) may be assumed as part of a direct evaporation 'pool' (Murphy & Knoerr, 1975) or recovered into the 'infiltration pool' (i.e. the amount reaching the soil surface) in form of stem flow down to the soil surface. For the recovered fraction, Brisson et al. (2003) proposed the simulation of stemflow ( $SF$ ) as a function of  $LAI$ , light extinction coefficient ( $k$ ) and an empirical crop coefficient ( $SF_{MAX}$ ) that depends on the architecture and wettability of plant surfaces (Wang et al., 2015) and the total water supply, i.e. irrigation ( $I$ ) and/or precipitation ( $P$ ) according to:

$$SF = SF_{MAX}[1 - e^{(-kLAI)}](P + I) \quad (2.1)$$

Alternatively, the method proposed by Braden. (1985) and von Hoyningen-Huene. (1981) can also be used to estimate the fraction of intercepted water (e.g. HYDRUS, SWAP):

$$INT = aLAI[1 - (1 + \frac{bAbC_{pool}}{aLAI})^{-1}] \quad (2.2)$$

where  $LAI$  is leaf area index,  $a$  is an empirical coefficient (assumed as 0.25 by default),  $b$  is the soil cover fraction (assumed as 0.33 of  $LAI$ ), and  $AbC_{pool}$  is the 'above canopy pool'.

Other models (DAISY, STICS) represent the effects of mulch residues on the modelling of water interception dynamics (Brisson et al., 2003; Hansen et al., 2012). Such advancements have focused on the development of empirical equations that estimate the quantity of soil cover with time (e.g. STICS). According to the calibration of Scopel et al. (1998), the effect can be represented by a negative logarithmic relation that determines the decomposition rate of the mulch type with time.

Some models (DAISY, MONICA, HYDRUS-1D/2D, SWAP, MIKE-SHE) integrate snow accumulation and melting processes within the pre-infiltration phase. As described by Abrahamsen and Hansen (2000), these processes can

be determined as a function of precipitation, air temperature, global radiation, ground heat flux, albedo and depth of the snowpack. Water losses from the snowpack, occur in the form of evaporation, sublimation and percolation (when the retention capacity is exceeded), and evaporation tends to have priority over sublimation.

In this sense, the *INFpool* (expressed in units of length per time, as mm day<sup>-1</sup> or cm day<sup>-1</sup> in case of daily time-steps) can be calculated as the sum of non-intercepted water from rainfall (*P*), irrigation (*I*), and when considered, also superficial inflow from run-on (*SIF*) and stemflow (*SF*). All contribute eventually as an input to a snow pack module (*SPM*), if considered, from which the melted fractions recover, counting as an input on the estimation of infiltration in the subsequent time-step. Non-recovered fractions through *SF* are likely to be lost through direct evaporation. The *INFpool* is therefore the amount of available water at soil surface to be infiltrated.

### 2.3.2 Infiltration

Effective infiltration (*INFeff*) is the fraction of *INFpool* that infiltrates into the soil at a given time step. The remaining fraction is considered a surface water surplus, that may originate water ponding (leading to accumulation, evaporation or infiltration in following time-steps) and surface runoff depending on the surface conditions and topographic characteristics of the area (Allen, 1991).

The simplest approach to estimate *INFeff* applies a simple capacity model (CAP) in which maximum infiltration capacity is defined as the difference between the soil saturation water content ( $\theta_{SAT}$ ) and actual water content ( $\theta$ ), expressed as a fraction of a volume. In this case, the infiltration capacity is defined as the maximum amount of *INFpool* that infiltrates in a given soil under specific conditions (WOFOST and MONICA; Appendix).

An alternative and widely adopted approach (DSSAT, APSIM, STICS, AquaCrop; Appendix) is the USDA curve number method (CN-method). As discussed in detail by Allen (1991), this 'infiltration-loss based method', calculates *INFeff* as a function of the potential maximum retention (*S*). *S*

(expressed in mm) is defined according to the curve number (CN), which is an empirical parameter, determined from tabled empirical values, according to land cover and soil hydrological group (Cronshey, 1986; Rallison, 1980; Woodward et al., 2003):

$$S = 254\left(\frac{100}{CN} - 1\right) \quad (2.3)$$

The soil hydrologic group classification is based on probability distribution curves of measured infiltration rates related to soil antecedent conditions (Allen, 1991; Cronshey, 1986). Therefore, CN is not a constant but varies from event to event. The robustness of this method is related to the vast quantity of field measurements that support it. However, it reveals some weakness when surface runoff ( $SRn$ ) is a small fraction of the  $INFpool$  (i.e. arid or semi-arid conditions), a situation where a wider range of CN is observed (Allen, 1991). According to the CN-method,  $INF_{eff}$  can be calculated as:

$$INF_{eff} = INFpool - \left[ \frac{(INFpool - \chi S)^2}{INFpool + (1 - \chi)S} \right] \quad (2.4)$$

where the subtracted fraction corresponds to  $SRn$  and corresponds to the 'initial abstraction' which is the initial fraction of  $S$  that can infiltrate before starting surface runoff (Allen, 1991). More mechanistic approaches compute infiltration according to the formulations of Richards (1931) and Richardson (1922) for transient flow conditions (DAISY, HYDRUS-1D/2D, SWAP, MIKE-SHE, SWIM; Appendix).  $INF_{eff}$  can be defined by unsaturated soil water movement, which is estimated through numerical solutions of the Richards equation (Buchan, 2003):

$$INF_{eff} = \frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} [K(\theta) \frac{\partial \psi}{\partial z} - K(\theta)] \quad (2.5)$$

where  $\theta$  is the volumetric water content of the soil top layer (expressed as a fraction of a volume),  $t$  is time,  $z$  is the vertical coordinate (expressed in units of length), positive when water flows downwards,  $K$  is the unsaturated hydraulic conductivity as a function of  $\theta$ , and  $\psi$  is capillary pressure head relative to atmospheric pressure (unit of length). While the space  $z$  and time  $t$

are independent variables,  $\psi$  and  $\theta$  are dependent variables.

The methods used to solve Richards' equation are an important issue in hydrological research (van Dam, 2000; van Dam et al., 1997). Many alternative mathematical solutions have been proposed (e.g. geometric means, arithmetic means, iterative methods), but the parabolic form of this equation in combination with the strong non-linearity of the soil hydraulic functions (i.e. functions relating water content, soil pressure head and hydraulic conductivity) makes it a non-consensus and difficult task. According to van Genuchten (1980) and Mualem (1976), the soil hydraulic functions can be represented as:

$$\theta(h) = \theta_{PWP} + \frac{\theta_{SAT} - \theta_{PWP}}{(1 + |\alpha h|^n)^{\frac{n-1}{n}}} \quad (2.6)$$

$$K(\theta) = K_{SAT} \left( \frac{\theta - \theta_{PWP}}{\theta_{SAT} - \theta_{PWP}} \right)^{\lambda} [1 - [1 - (\frac{\theta - \theta_{PWP}}{\theta_{SAT} - \theta_{PWP}})^{\frac{n}{n-1}}]^{\frac{n-1}{n}}]^2 \quad (2.7)$$

where  $\theta$  is the volumetric soil water content as a function of the soil pressure head ( $h$ ),  $\theta_{PWP}$  is the volumetric soil water content at permanent wilting point,  $\theta_{SAT}$  is the volumetric soil water content at saturation point,  $\alpha$  and  $n$  are empirical shape factors (respectively, expressed in units of length and unitless),  $K_{SAT}$  is the saturated hydraulic conductivity (expressed in units of length per time), and  $h$  is expressed in units of length. When  $INFpool$  increases at a higher rate than the maximum infiltration capacity (i.e.  $K(\theta)=K_{SAT}$ ), water accumulates at soil surface (DAISY, HYDRUS-1D/2D, SWAP, MIKE-SHE). The ponded infiltration can be modeled through a solution of Darcy's equation (DAISY) or according to the Green-Ampt approach (HYDRUS-1D/2D, SWAP, MIKE-SHE, Appendix), which is based on Darcy's equation for continuous saturated conditions and considers the wetting front as the reference elevation, where gravitational head is zero (Green & Ampt, 1911). The decreasing hydraulic gradient caused by the wetting front drives the drop in infiltration rate over time, which can be mathematically represented as:

$$INF_{eff} = K \left( \frac{L_f + H_0 + H_F}{L_f} \right) \quad (2.8)$$

where  $K$  represents the hydraulic conductivity,  $L_f$  is the thickness of the soil

being considered (i.e. the depth of the wetting front),  $H_0$  is the pressure head at soil surface and  $H_F$  is the pressure head at the wetting front ( $H_0$  and  $H_F$ , both assumed to be constant). Despite being originally developed for flat land conditions, the Green-Ampt approach has also been applied to sloping surfaces (Chen & Young, 2006).

The  $INF_{eff}$  estimation approaches that are based on the hydraulic conductivity of the top soil layer assume that after infiltrating into the soil, water is stored in successive layers downward according to a physical constraint that is imposed by the drainage ability of the soil (APSIM, STICS, AquaCrop).

### 2.3.3 Surface-water flow

Surface-water flow can be classified as a loss flow (i.e. when represented as  $SR_n$ ), a re-distribution flow (i.e. when represented as a lateral flow affecting the water balance of neighbouring hydrological units) or as a channel flow (i.e., for furrow irrigation simulation applications, as described by van Dam (2000a) and van Dam et al. (1997). However, in most of the selected models (i.e., all except MIKE-SHE), surface water flow is only represented as  $SR_n$  and therefore considered as an outflow of the system. It is therefore relevant to discuss  $SR_n$  in some detail below.

According to Ponce & Hawkins (1996),  $SR_n$  can be distinguished into different forms (Appendix): Hortonian overland flow ( $HRT_f$ ), saturation overland flow ( $SAT_f$ ), throughflow ( $THR_f$ ), the direct channel interception flow ( $DCI_f$ ) and surface phenomena flow ( $SUR_f$ ). Hortonian overland flow is the water flow occurring when rainfall or irrigation (or the combination of both by analogy with  $INF_{pool}$ ) exceeds soil infiltration capacity (i.e. typically the case of a rainfall storm). Saturation flow occurs when the profile gets saturated. While Hortonian flow is a 'pre-infiltration' process, saturation flow is a 'post-infiltration' one. Throughflow is the horizontal water flow beneath the land surface, usually when the soil is saturated. The direct channel interception flow is a type of runoff that refers to the spatial redistribution of rainfall directly intercepted by channels. This is an important type of flow in high dense and humid channel areas where channel interception may be the main source of

surface-water flow, as reported by Hawkins (1973), who assessed watersheds characterized by frequent storm precipitation events, hilly landscapes and large areas of stream channels, where direct interception occurs in great extents. Under flat conditions, typical rice landscapes in the Philippines or channelled plots in the Netherlands can be taken as an example of these areas, too. Surface phenomena flow is all the flow driven by crust development, hydrophobic layers and frozen ground that do not allow vertical flow to occur. While some models simulate  $SR_n$  (mostly in forms of  $HRT_f$ ,  $SAT_f$ ,  $THR_f$ ) through empirical approaches based on the CN-method (DSSAT, APSIM, STICS, AquaCrop), others derive  $SR_n$  from Richards' based approaches (DAISY, HYDRUS-1D/2D, SWAP, SWIM). 'Overland flow models' such as MIKE-SHE dedicate particular attention to the simulation of  $SR_n$  by dividing it into  $HRT_f$ ,  $SAT_f$ ,  $SUR_f$ ,  $THR_f$  estimated through a 'diffusive wave' approach which considers a Manning's type roughness coefficient, and through the St. Venant equations (Saint-Venant, 1871), as explained in detail by DHI (2017a,b).

### 2.3.4 Evaporation

Evaporation ( $E$ ) modelling has three different components: direct evaporation of water intercepted by the crop canopy ( $E_c$ ), from the soil surface ( $E_s$ ), and from mulches ( $E_m$ ).

$$E = E_c + E_s + E_m \quad (2.9)$$

However not all models calculate the three  $E$  components separately (DSSAT, MONICA), because, depending on the calculation procedure, evaporative demand may include all components together. The estimation of evaporation is conceptually divided into two steps for most of the selected models (WOFOST, APSIM, AquaCrop, DAISY, HYDRUS, SWAP): 1) the calculation of evaporative demand ( $ED_e$ ), i.e. mass transfer based on latent heat; and 2) a 'partitioning' according to the corresponding evaporative surface area (i.e. fraction of crop canopy, fraction of uncovered soil, fraction of mulch).

The main formulations used to calculate the evaporative demand ( $ED_e$ ) are

based on the energy balance balance (Penman, 1948, 1956), which evolved as the PM equation (Allen et al., 1998; Monteith, 1976; Monteith and Unsworth, 1990; Eq. (2.10); WOFOST, DSSAT, APSIM, DAISY, STICS, AquaCrop, MONICA, HYDRUS-1D/2D, SWAP). Other approaches used in the selected models are the Priestley-Taylor (PT) (Priestley & Taylor, 1972; Eq. (2.11); DSSAT, APSIM, STICS, MONICA), and Hargreaves (HG) (Hargreaves and Samani, 1982; Eq. (2.12); DAISY, HYDRUS-1D/2D) equations. The formulations of the equations are:

$$ED_e = \frac{0.408\Delta(Rn - G) + (\frac{\varphi \times 900}{T+273})U_2(es - ea)}{\Delta + \varphi(1 + 0.34U_2)} \quad (2.10)$$

$$ED_e = \frac{\Delta(Rn - G)}{\Delta + \varphi} \alpha_{PT} \quad (2.11)$$

$$ED_e = C_H Ra \sqrt{(T_{MAX} - T_{MIN})(T_{MEAN} + 17.8)} \quad (2.12)$$

where  $\Delta$  is the slope of the saturation vapor pressure function versus temperature ( $\text{kPa } ^\circ\text{C}^{-1}$ ),  $Rn$  is the daily net radiation at the soil surface (i.e. incoming minus reflected radiation expressed in  $\text{MJ m}^{-2} \text{ day}^{-1}$ ),  $G$  is the soil heat flux ( $\text{MJ m}^{-2} \text{ day}^{-1}$ ),  $\varphi$  is the psychrometric constant that is calculated according to the altitude (set by default as  $66 \text{ Pa K}^{-1}$ ),  $T$  represents the mean temperature of the air (measured at 2 m height and expressed in  $^\circ\text{C}$ ),  $U_2$  represents the wind speed (also measured at 2 m height and expressed in  $\text{m s}^{-1}$ ),  $es$  represents the saturation vapor pressure,  $ea$  the air vapor pressure (the difference of both equals vapour pressure deficit -  $VPD$ ), both expressed in kPa,  $\alpha_{PT}$  is an empirically derived factor that depends on the season and location in relation to large water bodies (Castellvi et al., 2001) and varies from 1.26 (Priestley and Taylor, 1972) in minimal advection conditions, to maximum reported values that vary from 1.74 to 3.12 (Eaton et al., 2001; Jensen et al., 1990; Viswanadham et al., 1991),  $C_H$  is a constant parameter, assumed as 0.0023 according to Hargreaves and Samani (1982),  $Ra$  is the extraterrestrial solar radiation ( $\text{MJ m}^{-2} d^{-1}$ ), and  $T_{MAX}$  and  $T_{MIN}$  are respectively maximum and minimum air temperatures (both expressed in  $^\circ\text{C}$ ).

Once  $EDe$  is estimated (expressed in units of length per time, as mm

day<sup>-1</sup>), some models follow a Beer-Lambert type approach (INT-BL) using  $LAI$  and the extinction coefficient ( $k$ ) to calculate an evaporation coefficient ( $K_e$ ) to estimate  $Es$  (WOFOST, DSSAT, APSIM, DAISY, STICS, MONICA, HYDRUS-1D/2D, SWAP, SWIM), while others follow a soil-cover based method (SC-M), whether using soil cover ( $SC$ ) instead (AquaCrop, HYDRUS-1D/2D, SWAP, MIKE-SHE, SWIM). Some models deliver the option to follow both approaches. The two approaches are expressed as follows:

$$Es = \sum (ED_{ei} K_e \times e^{-kLAI_i}) \quad (2.13)$$

$$Es = \sum [ED_{ei} K_e (1 - SC_i)] \quad (2.14)$$

where the subscript  $i$  corresponds to the model time-step (e.g. hourly, daily), and  $K_e$  is the evaporation coefficient for wet surfaces set by default as 1.10 according to Allen et al. (1998). Alternative approaches derive both daily crop ( $Ec_i$ ) and soil water evaporation ( $Es_i$ ) directly from an adaptation of the original Penman's equation (PE) (Penman, 1963) as function of the fraction of uncovered soil (DSSAT):

$$E = Ec_i + Es_i = R_{si}[(4.88 \times 10^{-3}) - (GS_i 4.37 \times 10^{-3})](T_i + 29) \quad (2.15)$$

where  $T_i$  is the daily air temperature (°C),  $R_{si}$  is daily solar radiation (MJ m<sup>-2</sup>) and  $GS_i$  is a ground surface correction factor defined according to the light extinction coefficient ( $k$ ) and daily values of leaf area index ( $LAI_i$ ) as following:

$$GS_i = (0.1 e^{-kLAI_i}) + [0.2(1 - e^{-kLAI_i})] \quad (2.16)$$

Similar approaches can be followed to estimate  $Em$  (STICS). However, in the case of inert mulching, the extinction coefficient decreases in time due to the decomposition of those residues as well as soil cover fraction (Scopel et al., 1998).

In addition to the previous methods, some models (MIKE-SHE, SWIM, DAISY, STICS) estimate  $E$  through soil vegetation atmosphere transfer schemes

(SVAT, Appendix). The SVAT approach (Brisson et al., 2003; Hansen et al., 2012; van der Keur et al., 2001; Verburg et al., 1996; Verburg, 1996) is based on the aerodynamic transfer of water vapor and integrates the effect of the aerodynamic resistance to water vapour transport ( $r_a$ ), the soil–vegetation–atmosphere pathways components resistance ( $r_c$ ), the saturated vapour pressure at canopy temperature ( $e_{S(T_c)}$ ), the vapour pressure in the overlying air boundary ( $e_{ref}$ ), the air density, the ratio of the molecular mass of water vapour to that of dry air ( $\rho_\epsilon$ ) and the canopy pressure ( $\rho_s$ ) as follows:

$$E = \frac{\rho_\epsilon}{\rho_s} \left( \frac{e_{S(T_c)} - e_{ref}}{r_a + r_c} \right) \quad (2.17)$$

where the calculation of  $E$  accounts for the different evaporation pathways. These approaches are also called networks of resistances (Campbell, 1985; Koster and Suarez, 1994). Some models divide soil evaporation in two consecutive stages (APSIM, STICS, AquaCrop, SWAP, MIKE-SHE, SWIM), others integrate it into a single formulation (e.g. WOFOST, DSSAT, DAISY, MONICA, HYDRUS, SWAP). Models integrating one single formulation apply directly to  $EDe$  a 'Beer-Lambert' type integrated approach (INT-BL) whether using  $LAI$  (see equation 2.13) or a soil-cover based method (SC-M) if using SC (see equation 2.14).

The 'two-stage method' (2-stage-M) proposed by Ritchie (1972) and based on Philip and De Vries (1957), considers that evaporation occurs in two consecutive stages: the first limited by the energy available, and the second, limited by water availability. While in the first stage, the evaporative rate is a function of the potential evaporative demand ( $EDe$ ), in the second stage, a falling response takes place since the surface soil water content decreases with time and  $Es$  depends on the flow of water to the soil surface which decreases exponentially with time.

The '2-stage-M' approach may be integrated into a single formulation considering a reduction factor that equals 1 for the first stage, and decreases to 0 during the second stage (AquaCrop):

$$E_s = (1 - SC)K_r K_e E D_e \quad (2.18)$$

When the second stage starts,  $K_r$  is calculated through an exponential relation that depends on a decline factor ( $f_k$ ) related to the relative soil water content ( $W_{rel}$ ) (AquaCrop):

$$K_r = \frac{e^{f_k \times W_{rel}} - 1}{e^{f_k} - 1} \quad (2.19)$$

where  $W_{rel}$  is a relative weighting factor, estimated according to Raes et al. (2017). For a given soil type, the second stage evaporation can also be empirically related to the square root of an independent variable, such as time (APSIM):

$$E_{s_{stage-II}} = \eta \sqrt{t} \quad (2.20)$$

where  $\eta$  represents a parameter related to the soil type and  $t$  is time; Others use an empirical parameter ( $A$ ) as following (STICS):

$$E_{s_{stage-II}} = \sqrt{(2A \sum E D_i) + A^2 + A} \quad (2.21)$$

where parameter ( $A$ ) depends on the aerodynamic resistance, the latent heat of vaporization, the water vapour pressure, air temperature, and a diffusion coefficient that is related to the bulk density of the evaporative soil layer and the surface temperature (Brisson and Perrier, 1991); The second stage  $E_s$  rate can still be modeled as a function of soil water flow ( $q$ ). This tends to be the case of approaches operating at smaller time-steps (SWAP, MIKE-SHE). In this sense,  $E_s$  gets boundary limited by the maximum upward flow, modelled through a numerical solution of Richards' equation that is simplified in the following form (van Dam and Feddes, 2000):

$$E_{s_{stage-II}} < -K_{1/2} \left[ \frac{(h_{atm} - h_1) - z_1}{z_1} \right] \quad (2.22)$$

where  $K_{1/2}$  represents the average hydraulic conductivity in the top soil evaporative layer (expressed in units of length per time),  $z_1$  corresponds to the thickness of the top soil evaporative layer (expressed in units of length),

$h_{atm} - h_1$  is the pressure gradient between atmospheric and top layer pressure (expressed in units of length).

### 2.3.5 Transpiration and root water uptake

Crop transpiration ( $T$ ) is determined by the atmospheric-evaporative demand ( $EDt$ ) and limited by root water uptake. Atmospheric conditions govern  $EDt$  while root water uptake is a function of both soil water availability and resource capture dynamics (Passioura, 1983). In general, models estimate first potential transpiration demand and then actual  $T$  rates according to canopy and root water uptake related factors, as well as to soil water status. Despite the atmospheric demand for transpiration ( $EDt$ ) being conceptually the same as the evaporative demand ( $EDe$ ), some models have the possibility to treat them separately as two different calculation procedures depending on different approaches and methods (DSSAT, APSIM, DAISY, STICS, AquaCrop, MONICA, HYDRUS-1D/2D). For this reason, we represent atmospheric demand for transpiration as  $EDt$  and evaporative demand for evaporation as  $EDe$  (Appendix).

Similarly to  $EDe$ ,  $EDt$  is estimated by the selected models according to one of the following approaches: Penman-Monteith (WOFOST, DSSAT, APSIM, DAISY, STICS, AquaCrop, MONICA, HYDRUS-1D/2D, SWAP), Priestley-Taylor (DSSAT, APSIM, STICS and MONICA as well), Hargreaves (DAISY, HYDRUS-1D/2D) or a SVAT scheme (MIKE-SHE, SWIM, DAISY and STICS as well). In order to estimate potential crop transpiration ( $Tc$ ), some models multiply  $EDt$  by a crop specific coefficient ( $Kc$ ) (MONICA), while others use a transpiration coefficient that is equivalent to the crop basal coefficient ( $Kc_b$ ) (AquaCrop). While  $Kc$  includes both soil evaporation and crop transpiration, the  $Kc_b$  is a specific parameter representing the transpiration fraction and a residual diffusive evaporation component supplied beneath vegetation (Allen et al., 1998, 2005; Raes et al., 2017).  $Tc$  is then adjusted to the transpiration surface through a 'Beer-Lambert type' integrated approach (INT-BL) using  $LAI$  (MONICA) or through a green canopy soil cover based method (SC-M) (AquaCrop), both equivalent to the evaporation approaches and respectively

expressed as:

$$T_c = ED_t Kc (1 - e^{-kLA}) \quad (2.23)$$

$$T_c = ED_t Kc_b SC \quad (2.24)$$

where  $SC$  is the green canopy soil cover adjusted for micro-advection effects. According to van Dam et al. (1997), an alternative  $T_c$  approach depending on the ratio between the daily amount of intercepted precipitation and the evaporation rate of water intercepted by the canopy ( $\beta$ ) is used (SWAP). This approach assumes that  $T_c$  is reduced by the water evaporation from the wet canopy ( $Ec$ ), since part of the latent heat flux is 'consumed' on leaf evaporation processes. The canopy transpiration through the leaf stomata gets maximum when  $\beta$  gets equal to zero (i.e. when  $Ec$  equals zero).

Actual transpiration ( $Ta$ ) can be reduced in multiple stress situations: soil saturation, low soil moisture, salinity, and excessive temperatures inducing stomata closure (Hsiao, 1973). This can be modeled through the use of stress coefficients ( $Ks$ ), which are calculated as (AquaCrop):

$$Ks = 1 - S_{rel} \quad (2.25)$$

$$Ks = 1 - \left[ \frac{e^{S_{rel}fshape} - 1}{e^{fshape} - 1} \right] \quad (2.26)$$

$$Ks = \frac{S_n S_x}{S_n + (S_x - S_n) e^{-r(1-S_{rel})}} \quad (2.27)$$

where  $S_{rel}$  is the relative stress level and  $f_{shape}$  is the curve shape factor,  $S_n$  and  $S_x$  are respectively the relative stress level at the lower and upper threshold and  $r$  is a rate factor (Raes et al., 2009b). AquaCrop is the only of the selected models which appears to use both transpiration coefficients equivalent to crop basal coefficients ( $Kcb$ ) and multiple stress coefficients ( $Ks$ ). For this specific case,  $Ta$  is estimated as:

$$T_a = ED_t K c_b SCK s \leq S_i \quad (2.28)$$

where  $K_s$  varies from 0 to 1 according to three main different approaches.

The actual crop transpiration ( $T_a$ ) is limited by an extraction sink term ( $T_a < S_i$ ), which, in the case of multi-layer models, is computed separately for each individual soil layer. According to Ritchie (1972, 1981) and Feddes et al. (1978),  $S_i$  can be calculated as a linear function (LIN) of soil moisture content ( $\theta$ ) or pressure head ( $h$ ) and the maximum extraction rate ( $S_{MAX}$ ):

$$S_i = \alpha(h) S_{MAX}(z) \quad (2.29)$$

where  $S_{MAX}$  depends on the vertical rooting depth and  $\alpha(h)$  is a coefficient that depends linearly on  $h$  in three different phases: 1)  $h$  is considered to increase linearly from 0 to 1, between a h-minimum threshold (i.e. saturation conditions) and an intermediate h-threshold; 2)  $h$  equals 1 for an intermediate interval of  $h$  (i.e. optimal soil moisture content pressure head for plant uptake); and 3)  $h$  decreases linearly from 1 to 0 (i.e. at permanent wilting point). The coefficient  $\alpha$  can also be represented as a function of soil water content (Raes et al., 2017; van Genuchten, 1980). This approach can be employed in both 'discrete' or 'continuous' representation schemes of the vadose zone. 'Discrete' schemes determine  $S_{MAX}(z)$  as the product of  $T_a$  and a root density term ( $D_{root}$ ), which can be computed separately for each individual layer (APSIM, STICS, AquaCrop, MONICA) or for the whole root zone (DSSAT, WOFOST), abbreviated as following:

$$S_{MAX} = T_a D_{root} \quad (2.30)$$

where  $D_{root}$  represents the fraction of total root density in each layer, when computed individually for multiple layers, or the rooted fraction of a single layer depth, when computed for the whole root zone. On the other hand, some 'continuous' schemes (DAISY, SWAP, SWIM) define  $S_{MAX}$  through an integral equation (from root depth to soil surface) that can be simplified as:

$$S_{MAX} = \left[ \frac{T_a \pi_{root}}{\int_{-D_{root}}^0 \pi_{root} dz} \right] \quad (2.31)$$

where  $\pi_{root}$  is the root length density (expressed in mm mm<sup>-3</sup>), defined as function of both space and time (van Dam et al., 1997). Space can be represented in one ( $\pi_{root}(z, t)$ ) or two dimensions ( $\pi_{root}(x, z, t)$ ) as respectively described by van Dam et al. (1997) and Simunek & Hopmans (2009).

An alternative approach (APSIM) describes  $S_i$  through an exponential relation (EXP):

$$S_i = [1 - e^{kl(t-tc)}] S_{MAX} \quad (2.32)$$

where  $k$  is a diffusivity constant (expressed in cm<sup>2</sup> day<sup>-1</sup>),  $l$  is the root length density (equivalent to  $\pi_{root}$  but here expressed in cm of root per cm<sup>3</sup> of soil),  $t$  is time and  $tc$  is the beginning time of water extraction (Passioura, 1983; Tinker, 1976). According to DHI (2017b), an alternative to this relation is to simplify the root depth as a linear function of time while assuming root length density as constant (MIKE-SHE).

According to van Genuchten (1987), an osmotic pressure term ( $h_\phi$ ) can be included in the calculation of  $\alpha$  (Equation 2.29) that becomes a nonlinear function (HYDRUS-1D/2D, SWIM), also time dependent  $\alpha(h, h_\phi, z, t)$  (Simunek et al., 2018a; van Genuchten, 1987). For the specific case of HYDRUS-2D, a horizontal coordinate is also incorporated into the  $\alpha(h, h_\phi, x, z, t)$  extraction function (Simunek & Hopmans, 2009), and  $S_i$  is calculated as:

$$S_i = \int_{\Omega R} [\alpha(h, h_\phi, x, z, t) \frac{L_t}{L_x L_z}] T_a \quad (2.33)$$

where  $L_x$  is the width of the root zone ( $\Omega R$ ),  $L_z$  is the depth of  $\Omega R$ , and  $L_t$  is the soil surface associated with the transpiration process, all expressed in units of length (Simunek & Hopmans, 2009).

Apart from the three general modelling approaches described for  $S_i$  calculation, some models (HYDRUS-1D/2D) also include a module for compensatory mechanisms (Appendix) regulating root water extraction

(Simunek & Hopmans, 2009). This enables the simulation of physiological responses at the root level under spatially distributed stress conditions (Bouten, 1995; Hsiao, 1973; Li et al., 2001). In these cases, a root adaptability factor, defined as the threshold value above which reduced root water (or nutrient) uptake in water (or nutrient) stressed parts of the root zone, is fully compensated by increased uptake in other root zones that are less stressed.

### 2.3.6 Redistribution and drainage

Modelling drainage processes has been a central issue in hydrology for centuries (Skaggs & Chescheir, 1999). For point-based models, drainage is represented as a vertical flow (Chescheir, 2003), generally simulated in two main ways (Appendix): with a 'tipping-bucket' approach (TBA), or based on Darcy's or Richards' equations.

The *TBA*, as described by Emerman (1995), implicitly considers that macropore water flow is the only mechanism of water transport between each 'tipping bucket' (i.e. soil layer). Each 'bucket' is boundary defined by a lower and an upper limit; the  $\theta_{PWP}$  (when empty) and the  $\theta_{FC}$  (when full). If water content exceeds  $\theta_{FC}$ , water excess flows vertically downwards to the next layer for a given time-step. The TBA models (WOFOST, DSSAT, APSIM, STICS, AquaCrop, MONICA) are simple and fully discrete in time (constant conditions are assumed for a certain time interval). A notable limitation of the TBA approach is the fact that the chosen time step is critical for an accurate prediction of the observation (Emerman, 1995). The minimum effective time step is the minimum period over which an appropriate fraction of the soil water excess (when  $\theta > \theta_{FC}$ ) drains down to the next unit. The fundamental equation describing internal drainage ( $D$ ) under *TBA* is:

$$D = \alpha(\theta - \theta_{FC}) \quad (2.34)$$

where  $\alpha$  is the drainage coefficient and  $\theta > \theta_{FC}$  (Emerman, 1995; Ritchie, 1984).

Modelling approaches based on Darcy's and Richards' equations allow

a continuous representation of soil water movement, for saturated and unsaturated conditions. While Darcy's is used for steady-state flow modelling (Buchan & Cameron, 2003), Richards' is used for transient flows (Buchan, 2003; Richards, 1931; Simunek and van Genuchten, 2008). The model formulations are dependent on the spatial-scale, as the flow term ( $q$ ) can be defined as a one, two or three dimensional vector (Buchan, 2003), leading to different calibration requirements and computation times.

Soil water movement also depends on the wetting/drying history of the soil, a phenomenon called hysteresis (Hillel, 1980). In general, hysteresis retards water movement, while preferential flow enhances water movement. In all crop models described here, hysteresis is ignored since only one curve is used to describe the  $h(\theta)$  relationship (WOFOST, DSSAT, APSIM, DAISY, STICS, AquaCrop, MONICA). This is mostly due to the time and cost associated with the inclusion of hysteresis in the calibration of this relationship. However, this might lead to considerable uncertainties regarding the simulation of infiltration and lateral flow rates, mostly at larger time-steps (van Dam, 2000b). However, soil water hysteresis effects can be simulated using the Scott's scaling method (SCOTS), which requires only the calibration of the main drying and wetting water retention curves to calculate the scanning curves (Scott, 1983). The scanning curves are derived by linear scaling of the main curves as follows (HYDRUS-1D/2D, SWAP):

$$\frac{\theta_{SAT}^* - \theta_{res}}{\theta_{SAT} - \theta_{res}} = \frac{\theta_{act} - \theta_{res}}{\theta_{md} - \theta_{res}} \quad (2.35)$$

where  $\theta_{SAT}^*$  is the adapted  $\theta_{SAT}$ ,  $\theta_{act}$  is the actual water content,  $\theta_{md}$  is the water content of the main drying curve at the actual soil water pressure head, and  $\theta_{res}$  is the residual water content of the wetting scanning curve (Kroes et al., 2017b).

### 2.3.7 Capillary rise

Quantification of capillary rise ( $CR$ ) is of great importance for the accurate simulation of the water balance, particularly in areas with shallow groundwater tables (Kroes et al., 2017a). However, not all selected models conceptually

consider  $CR$  on the calculations of the water balance. Those following a 'tipping-bucket approach' TBA and considering  $CR$ , either take it as an input (STICS, MONICA), or simulate it (DSSAT, AquaCrop) according to relations (soil texture specific) between water table depth, soil hydraulic properties and actual soil water content (or pressure head) of the unsaturated receiving layers (Raes et al., 2017). For the first case (STICS, MONICA),  $CR$  is defined by a Neumann type lower boundary condition (i.e., flux is a function of time, as  $CR$  depends on a defined calendar). For the second case (DSSAT, AquaCrop), both a Dirichlet type (i.e.,  $CR$  as a function of soil water content) and a Cauchy type condition (i.e.,  $CR$  flux as a function of groundwater level) are considered (Raes et al., 2017; Ritchie, 1998). However, none of these cases (DSSAT, STICS, AquaCrop, MONICA), simulate the feedback between the vadose zone and the water table (i.e., the water table depth is not updated).

Models using Richards' equation (DAISY, HYDRUS-1D/2D, SWAP, MIKE-SHE, SWIM), have different approaches to simulate  $CR$  flow while updating water table depth (Hansen et al., 2012; Simunek et al., 1999; van Dam, 2000a; van Dam and Feddes, 2000; Verburg, 1996). In the selected models there are four different types (Appendix): 1) models that do not-consider  $CR$  ( $X$ ) (WOFOST, APSIM); 2) models with predefined  $CR$  ( $D - CR$ ) (STICS, MONICA); 3) models simulating  $CR$  but without updating water table depth ( $SnU - CR$ ) (DSSAT, AquaCrop), and 4) models simulating  $CR$  and updating water table depth ( $SU - CR$ ) (DAISY, HYDRUS-1D/2D, SWAP, MIKE-SHE, SWIM).

### 2.3.8 Subsurface lateral flow

Among the models, subsurface (water) lateral flow ( $SSLF$ ) is simulated only by Richards' equation based models. However, most of these models (DAISY, HYDRUS-1D, SWAP, MIKE-SHE, SWIM) limit  $SSLF$  simulations to lateral drainage processes, such as lateral out flows between the simulated plot and neighbouring drainage canals (Simunek et al., 2018b; Simunek et al., 1998; van Dam, 2000a; van Dam et al., 1997; Verburg, 1996). For these cases (Appendix), lateral flow to drains ( $q_{drain}$ ) is represented by the Hooghoudt equation (Ritzema, 1994) which can be simplified as follows:

$$q_{drain} = \frac{\phi_{WL} - \phi_{drain}}{\gamma_{drain}} = \frac{8K_{SAT}^h d \Delta h_{tot} + 4K_{SAT}^h d \Delta h_{tot}^2}{L^2} \quad (2.36)$$

where  $\phi_{WL}$  (cm) represents mean groundwater level,  $\phi_{drain}$  is the drain level (cm), and  $\gamma_{drain}$  is the resistance to drainage (cm day<sup>-1</sup>);  $q_{drain}$  (cm d<sup>-1</sup>) is the drain discharge rate,  $K_{SAT}^h$  (cm d<sup>-1</sup>) is the horizontal saturated hydraulic conductivity,  $d$  (cm) is the equivalent depth, which is a reduced value of the impermeable layer depth below the drain level,  $\Delta h_{tot}$  (cm) is the total hydraulic head difference between the drain level and the phreatic level at midpoint, and  $L$  (cm) is the drain spacing. This approach is not considered fully distributed since  $q_{drain}$  is assumed as a system water loss but not as a re-distributive process. An alternative approach is used in HYDRUS-2D, which considers SSLF within the water balance calculation by adding a horizontal term to the Richards' equation as follows:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}[K(\theta)\frac{\partial \psi}{\partial z} - K(\theta)] + \frac{\partial}{\partial x}[K(h)\frac{\partial h}{\partial x}] \quad (2.37)$$

where the hypothetical horizontal gradient of both hydraulic conductivity and pressure head is added to the 1-D vertical formulation. For this case, a 'Galerkin finite-element method' is used to convert the differential equation into a discrete type problem (Mohsen, 1982; Simunek et al., 1999). A similar approach is proposed by van Dam et al. (1997) for introducing 'Neumann-type' conditions in the SWAP model to define the lower boundary in the calculations of capillary rise (*CR*) through Richards' equation.

### 2.3.9 Solute transport

Within the scope of this review, and considering the existing relations between solute concentration and root water uptake (e.g. salinity, co-limitation as discussed by Cossani & Sadras (2003)), we included a general description of solute transport processes. We limited our analysis to identify whether salts and nutrient transport processes are addressed by the selected models.

In AquaCrop, particular attention is given to the salt balance and consequently, to crop yield response to salinity (Raes et al., 2009a,b). According

to Raes et al. (2012), incoming and outgoing salt fluxes can be simulated with downward (i.e. vertical leaching) and upward water movement (i.e. flow of saline water through capillary rise from a shallow water table). Conceptually, the salinity concentration for a given layer ( $\sigma_k$ ) can be updated at each time step, every time water moves in ( $\Delta\theta_{in}$ ) or moves out ( $\Delta\theta_{out}$ ). For these cases, the salt balance of a given layer is determined for a particular time step as:

$$\frac{\partial\sigma_k}{\partial t} = \frac{(\sigma_k\theta_k) + (\sigma_{in}\Delta\theta_{in}) - (\sigma_{out}\Delta\theta_{out})}{\theta_k + \Delta\theta_{in} - \Delta\theta_{out}} \quad (2.38)$$

where  $\sigma_k$  is the specific layer salt content (expressed in g) and  $\theta_k$  is the actual water content (expressed in mm) of layer  $k$ . Other models, based on Richards' equation (e.g. HYDRUS, SWAP, SWIM), use differential equations based on the convective-dispersive transport (C-D) theory, which can be simplified as:

$$\frac{\partial c_\theta}{\partial t} = \left[ \frac{\partial}{\partial z} \theta D \frac{\partial c}{\partial z} \right] - \frac{\partial q_c}{\partial z} \quad (2.39)$$

where  $c_\theta$  (mg L<sup>-1</sup>) is the solute concentration (i.e. salts and nutrients in inorganic form) in soil solution,  $\theta$  is the soil volumetric water content (cm<sup>3</sup> cm<sup>-3</sup>),  $q$  is the water flux (cm day<sup>-1</sup>) and  $D$  is the dispersion coefficient which, according to Kersebaum (1989), can be estimated as:

$$D = D_0\tau^{-1} + D_v\left(\frac{q}{\theta}\right) \quad (2.40)$$

where  $D_0$  is the solute diffusion coefficient (which can be assumed as 2.14 cm day<sup>-1</sup> for the case of nitrate),  $\tau$  represents the tortuosity and  $D_v$  is the standard dispersion factor (assumed as 25 cm for the case of nitrate). According to van Genuchten (1985), solute adsorption effects can be incorporated by considering the adsorbed concentrations as a linear function of solute concentrations. This has great importance for nutrient transport modelling as following:

$$\frac{\partial}{\partial t}(\theta_{ci} + \rho S_i) = \frac{\partial}{\partial z}\left(D \frac{\partial c_i}{\partial z} - q C_i\right) - \mu_{wi}\theta_{ci} - \mu_{si}\rho S_i \quad (2.41)$$

where  $c_i$  is the solute concentration (g cm<sup>-3</sup>),  $S_i$  is the adsorbed concentration (mg mg<sup>-1</sup> or %),  $\theta$  is the volumetric water content (%),  $q$  the

volumetric flux ( $\text{cm day}^{-1}$ ),  $D$  is the dispersion coefficient ( $\text{cm}^2 \text{ day}^{-1}$ ),  $\rho$  the porous medium bulk density ( $\text{g cm}^{-3}$ ),  $z$  is distance (cm), and  $t$  is time (day); the subscript  $i$  delineates the  $i^{th}$  chain member. The coefficients  $w_i$  and  $s_i$  correspond to rate constants for the first-order decay in the liquid and solid phases of the soil respectively. Units can be adjusted to multiple temporal scales.

## 2.4 Crop and hydrologic models: what sets them apart?

The diversity found in the employed methods to simulate the role of water varied among models and among the different processes, which is partly related to the historical development of the models, as shown in the chronological map of modelling approaches presented in Figure 2.1. Note that while hydrologic models have their foundations mostly on research that started in the XIX Century, crop models are sustained by more recent approaches, whose fundamentals evolved from the 1950-60's (Jin et al., 2018; Jones et al., 2017a). After the publication of Darcy's equation and Beer-Lambert law (Figure 2.1), we note that hydrologists devoted most of their subsequent efforts to the development of modelling approaches of soil-water movement (e.g. infiltration, capillary forces, drainage processes). However, crop plants were still excluded from the hydrologic system at the time, with modelling prioritizing the representation of soil-water processes without focusing on related plant processes such as transpiration and root water uptake.

This paradigm changed substantially in the 1960's, when the pioneering works of some agronomists and physicists working on photosynthesis (de Wit, 1965; Duncan et al., 1967), broadened perspectives and brought biological variables into the water modelling context (Bouman et al., 1996; Jones et al., 2017a). In fact, the study of photosynthesis was at the root of the development of the first crop models, leading to an uniform approach where most crop models today are radiation driven. This common approach has oriented modelling towards the simulation of optimal conditions, not paying sufficient attention to the responses to environmental stress, thus limiting their use in crop management research (Loomis et al., 1979). The uncertainty regarding

future climate scenarios (Hansen and Jones, 2000; Rosenzweig et al., 2014) and the growing demand for decision support tools within the context of precision agriculture (McBratney et al., 2005; Cassman, 1999) will very likely require additional efforts to improve crop modelling under water-limited conditions, i.e. to integrate more water-driven mechanisms in crop models or, as discussed by Passioura (1996), to transform source-limited approaches into sink-limited.

Regarding the diversity of employed methods among models, there is a higher diversity in the simulation approaches of pre-infiltration processes (Figure 2.2). The number of components considered in the calculation of the infiltration pool varies substantially among models (Figure 2.2). The main issues are related to the incorporation (or exclusion) of snow pack modules and calculation methods related to canopy interception and surface inflow (or outflow) from run-on (or runoff) processes (Figure 2.2 and Appendix). Such discrepancies might be the result of a longer scientific heritage, since it appears that more time has been dedicated to the study of infiltration-related processes than to other water balance processes (Figure 2.1), promoting the observed diversification of methodologies and modelling approaches. Another case of methodological discrepancy among models is related to evaporation from soil, where several different methods are used: one or two stages, with the second stage limited either by time or by soil water flow, which can be modeled in several different forms (Section 2.3.4 and Appendix).

The highest degree of concordance among models is related to the calculation of evaporative demand (Figure 2.2). From the five methods identified (i.e. PE, PM, PT, HG and SVAT schemes; Appendix), the large majority of models have adopted PM equation (Section 2.3.4), with the exception of MIKE-SHE and SWIM (Appendix). Similarly, in most cases reviewed (Figure 2.2), the root water uptake is based on a linear model relating relative uptake to soil water content between the upper and the lower limit (Section 2.3.5).

While evaporative demand is more or less uniformly treated, this is not the case for the partitioning of  $ET$  into evaporation and transpiration. Most models follow a Beer-Lambert type equation depending on the canopy extinction coefficient and leaf area index (Section 2.3.4 and 2.3.5), but for some of the

hydrologic models (HYDRUS-1D, HYDRUS-2D, SWAP, SWIM) an option is offered to use the soil cover method instead. In general, in the case of most crop models, there is a clear agreement on the use of Beer-Lambert formulation, but for hydrologic models both options seem to be equally adopted (Appendix).

While there is substantial agreement in the fundamental approaches of the reviewed models, there are also major differences among the two model families. The main differences found between crop and hydrologic models are related to temporal and spatial resolution of processes in the soil-plant-atmosphere continuum, and to the degree of mechanistic or empirical-based approaches used (Appendix), implying considerable differences in terms of complexity as well. While the crop models (with the exception of DAISY) follow a TBA, hydrologic models are based on numerical solutions of Richards' equation (Section 2.3.6). Such a divergence implies structural differences between both families not only in terms of spatial resolution but also in terms of temporal scales.

The TBA limits models to a point-based scale where drainage is assumed to be a steady flow (Section 2.3.6), only vertical and discrete in time (resulting in longer time-steps, e.g. daily). The degree of empiricism involved in TBA based models (i.e. most crop models) is also higher (e.g. CN-method, drainage coefficients, capillary rise defined by Neumann type conditions). On the other hand, hydrologic models, based on numerical solutions of Richards, are capable of simulating the water balance at shorter time-steps (e.g. hourly) and of integrating some multi-dimensionality in the simulation of water flows by distributing partially water over the horizontal space.

## 2.5 Opportunities to simulate spatial water variation

The ability to simulate continuity and multi-dimensionality does not necessarily imply the simulation of a full distribution of water over space, as none of the hydrologic models (and none of the crop models) simulates all spatial processes that we have identified (Figure 2.3). All spatial processes are covered by at least one model (Figure 2.2), but none of these models covers all of them

simultaneously (Figure 2.3). Some models consider subsurface lateral flows but still ignore surface inflow from run-on (Figure 2.3). Additionally, as more generally discussed (Passioura, 1996; Nielsen & Alemi, 1989), the apparent continuity associated with hydrologic models can also be a point of discussion as these models follow 'discrete characteristics' too (e.g. input parameters, boundary conditions, reduced dimensions through the Galerkin finite-element method, Scott's scaling method for simulating hysteresis), becoming eventually more stochastic rather than deterministic.

Regarding the simulation of water processes spatially at crop field level, some methodologies for geospatial simulation, visualization and validation of models (e.g. geospatial interpolation of point based simulations, zonal statistics applied to mapped simulation results, integration of modelling with remote sensing) have been proposed (Basso et al., 2001; Booker et al., 2015; Campos et al., 2019; Casa et al., 2015; Droogers and Bastiaanssen, 2000; Er-Raki et al., 2007; Grassini et al., 2015; Han et al., 2019; Jégo et al., 2012; Jia et al., 2011; Lobell et al., 2015; Lorite et al., 2013; McBratney et al., 2005; Moiling et al., 2005; Silvestro et al., 2017; Shu et al., 2018; Ward et al., 2018). However, while most of these cases have been developed at regional scales, not addressing within-field spatial variation (Droogers and Bastiaanssen, 2000; Grassini et al., 2015; Han et al., 2019; Jia et al., 2011; Lobell and Azzari, 2017; Lobell et al., 2015; Lorite et al., 2013; Sadler and Russell, 1997; Shu et al., 2018; Zwart and Bastiaanssen, 2007), others, that reveal some promising advances in respect to the spatial simulation of water and vegetation, do still neglect spatial compensations of yield determining factors such as variations of the harvest index or root growth (Moiling et al., 2005; Booker et al., 2015; Ward et al., 2018). In addition, we acknowledge that the assimilation of remote sensing to quantify spatial variations is also problematic as ample variation can be observed when using reflectance signals to derive canopy structure parameters with implications on crop transpiration and photosynthetic activity (e.g. *LAI*, *SC*), as done by as done by Campos et al. (2019), Casa et al. (2015), Er-Raki et al. (2007), Silvestro et al. (2017), requiring in-situ and crop specific calibration that is not straight forward (Gao et al., 2020).

We emphasize that in terms of water spatial distribution and its effects on crop growth and development, geostatistical methods applied to point-based (or partially distributed) water balances might smooth considerably the actual spatial heterogeneity, because lateral water movement and 'cause-effect' relations between neighbouring cells are still partly ignored (Fig. 2.3). This awareness is in line with the observations reported by Wallor et al. (2018). In addition to this, relying solely on geostatistics to deal with spatial heterogeneity does not resolve the existing knowledge gaps in regard to the driving mechanisms (McBratney et al., 2005). This was also raised by Nielsen and Wendoroth (2003), who suggested that statistical methods should not replace research inventiveness in the assessment of spatial and temporal variations.

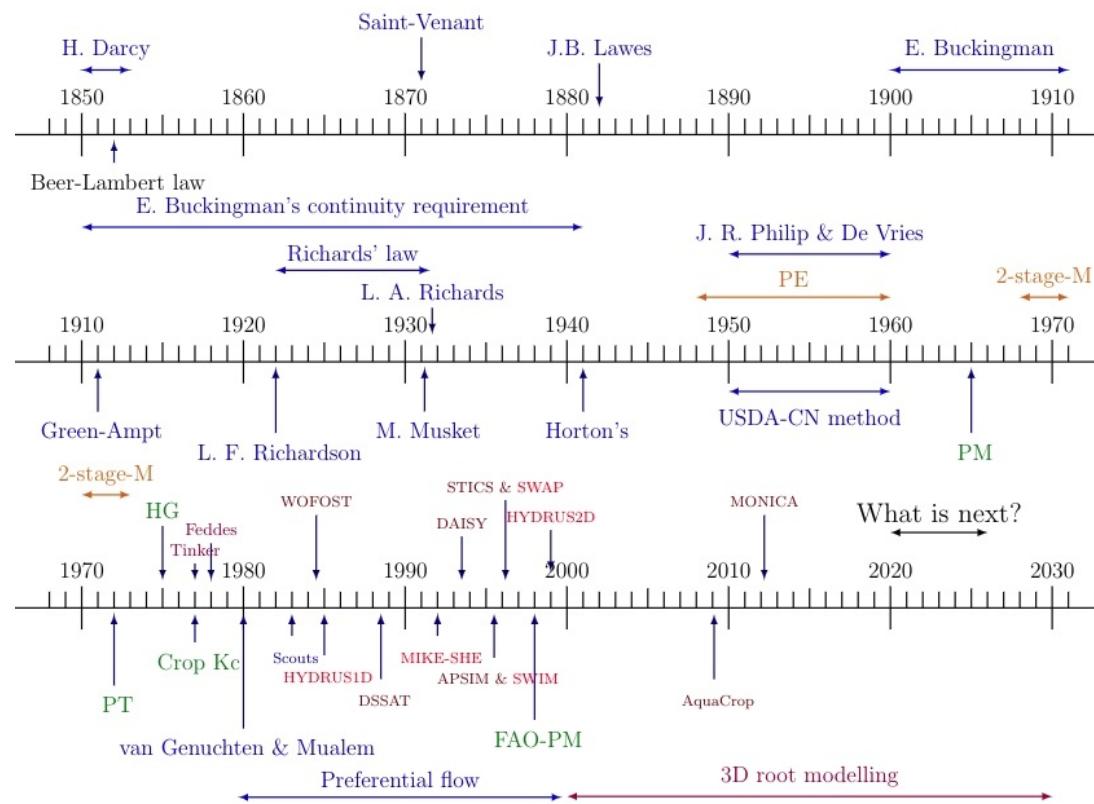
In order to distribute spatially water processes in crop models, further steps might be foreseen in two different directions. One implying a stronger synergism between both model families, that might result in the addition of spatial and continuous mechanisms to crop models, other through the integration of lateral flows in current TBA based discrete approaches. The specific processes and approaches that hold the most promise for advances are related to the incorporation of surface inflow and subsurface lateral flows (Figure 2.3), by using differential equations (Section 2.3.3 and 2.3.8) or through novel water spatial partitioning relations that must be developed for TBA based discrete approaches.

The future will surely be determined by the existing trade-offs between models complexity and adoption. The excessive simulation time and the calculation complexity associated with mechanistic structures that was sometimes seen in the past as a constraint to adoption (Loomis et al., 1979; Nielsen and Alemi, 1989; Passioura, 1996), is very likely to be overcome by today's enormous computational capacity of alternative operational systems (Thorp et al., 2012). However, larger calibration and parameterization requirements associated to mechanistic approaches that depend on complex numerical and analytical solutions of nonlinear equations may not meet the small 'appetite for data' that we aim for in an attractive tool. Therefore, both ways imply important trade-offs between accuracy and data requirements that

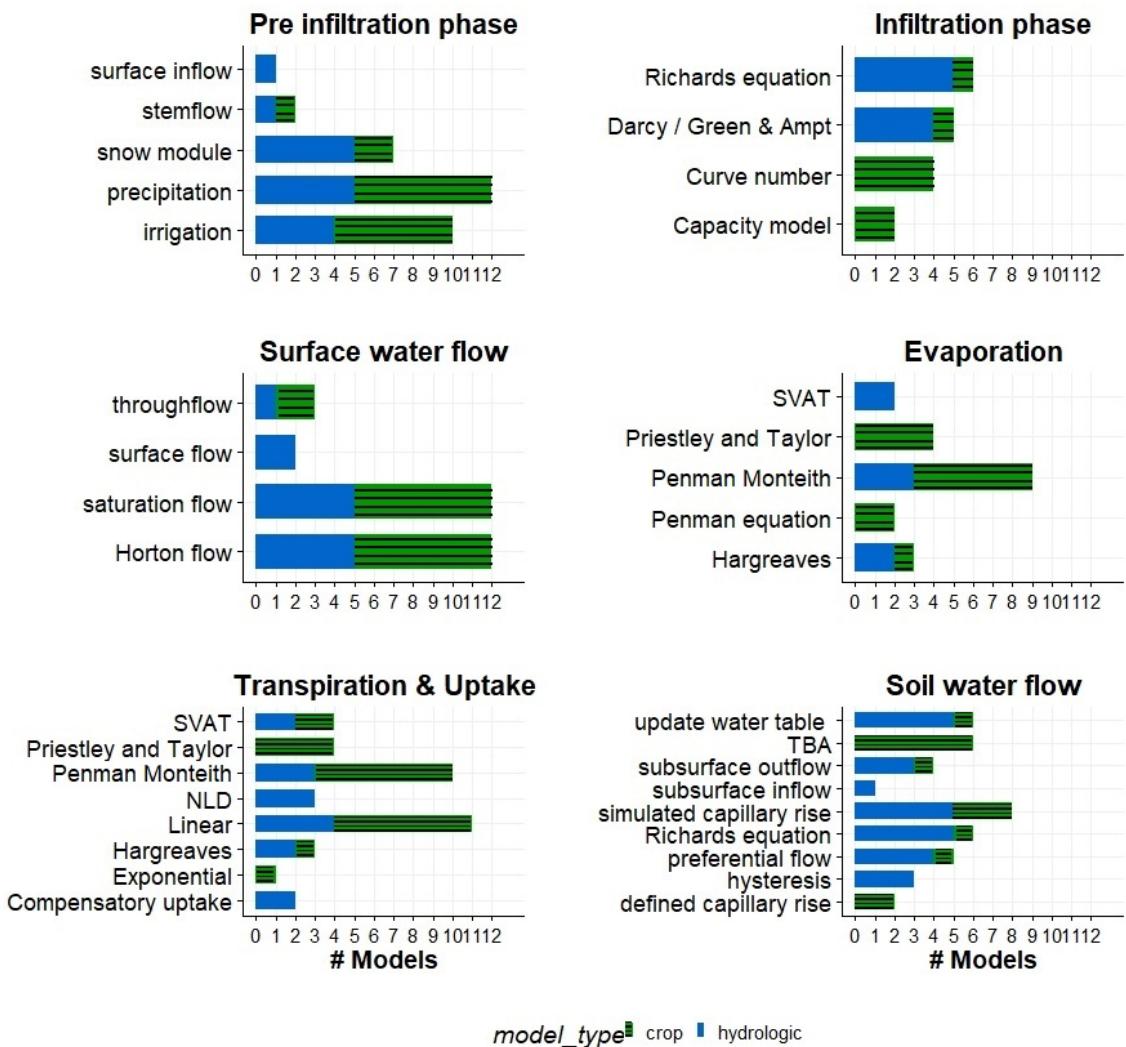
must be considered. In any case, we conclude that further steps are in need of experimental datasets for the calibration and validation of new upscaling efforts (as also raised by Sadler and Russell (1997)). Spatially distributed data related to subsurface soil texture and plant available water will be essential to achieve a better performance of modelling (Wallor et al., 2018). In this sense, crop modelers are strongly encouraged to come up with innovative databases, suitable for upscaling and spatially calibrating modelling tools at field level.

The success of precision agriculture and spatial management will surely benefit from new advances in the spatial modelling of water as we identify scope for conceptual improvements. Further (coordinated) research efforts are definitely needed, empowering linkages between researchers, farmers, sensing manufacturers and consultants is highly recommended in order to promote field experiments at 'real scales' capable of capturing satisfactory levels of spatial variation.

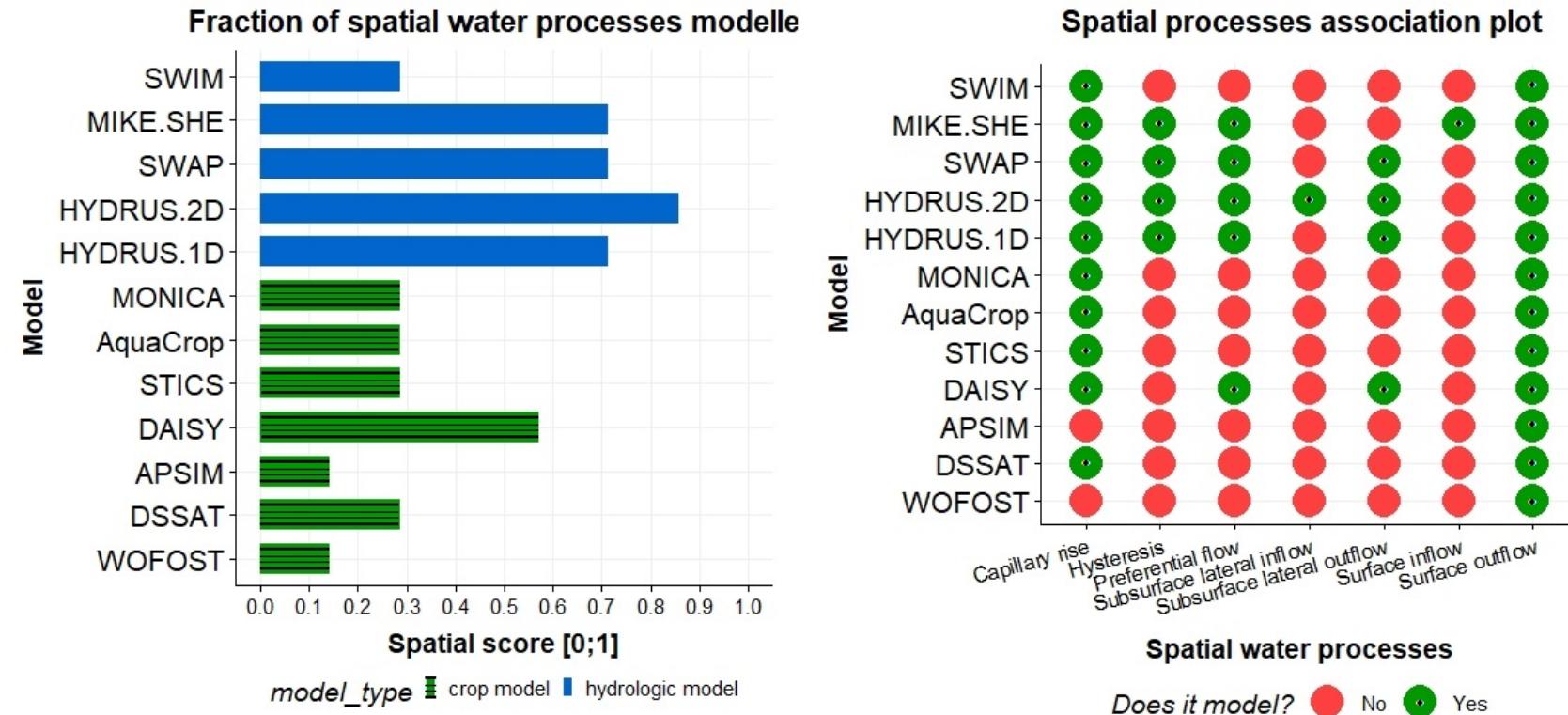
## FIGURES - Chapter 2



**Figure 2.1:** Chronological evolution of modelling approaches and theoretical fundamentals; in blue: methodological approaches related to infiltration processes and soil-water movement; in orange: related to evaporation; in green: to transpiration; in dark-purple: plant-root water uptake; in dark-red: crop growth simulation models; in dark-pink: hydrology based models. PE means Penman equation, 2-stage-M corresponds to the 2 stage method proposed by Ritchie. (1972), PT means Priestley and Taylor, HG represents the Hargreaves equation and PM is Penman-Monteith. FSPM means 'functional structural plant modelling'. The Beer-Lambert Law, which embraces a wider scope than crop-hydrological issues but it has influenced many modelling approaches of evaporation and transpiration is represented in black color. All horizontal arrows refer to the time-scale below, except in the case of 'preferential flow' and '3D root modelling'.



**Figure 2.2:** Number of models simulating a specific process (N=12), the most common modelling approaches used. The horizontal bars show the number of models that simulate (use) the corresponding process (approach). 'NLD' means that non-linear differential equations are used in the estimation of the extraction sink term (as an alternative to linear or exponential approaches).



**Figure 2.3:** 'Spatial scores' (left); spatial water processes considered in each model (right). Spatial scores represent the relative amount of spatial processes found in each model (expressed as the amount of spatial processes considered by a model, divided by the total amount of spatial processes that we identified).

## APPENDIX - Chapter 2

Table 1: Synthesis of models. *Codes:* INFpool = Infiltration pool, INFeff = effective infiltration, INFponding = ponding infiltration, SRn = Forms of surface runoff (2: Horton and saturation flow; 3a: previous plus throughflow; 3b: previous plus surface flow; All: all escribed forms of SRn, P = precipitation, I = irrigation, SPM = snowpack module, SF = stemflow, INT = intercepted pool, SIF = surface inflow (from run on), CAP = capacity model, CN = curve number, G&A = Green and Ampt, PE = Penman equation, PM = Penman-Monteith, PT = Priestly and Taylor, HG = Hargreaves, SVAT = Soil vegetation atmosphere transfer scheme, CU = compensatory uptake, SSLF = subsurface lateral flow, PF = preferential flow, INT-BL = Integrated through Beer-Lambert equation, SC-M = Soil cover method, LIN = Linear, EXP = Exponential, NLD = Non-linear differential equations,  $ED_e$  (or  $ED_t$ ) = Evaporation demand to estimate evaporation (or transpiration), TBA = Tipping bucket approach, D-CR = predefined CR, SnU-CR = simulated CR without updating water table, SU-CR = simulated CR and updated water table.

Model	WOFOST	DSSAT	APSIM	DAISY	STICS	AquaCrop	MONICA	HYDRUS-1D	HYDRUS-2D	SWAP	MIKE-SHE	SWIM
INFpool	P	P,I	P,I	P,I,SPM	P,I,SF	P,I	P,I,SPM	P,I,INT,SPM	P,I,INT,SPM	P,I,INT,SPM	P,I,INT,SPM,SIF	P,I,SPM
INFeff	CAP	CN	CN	Richards	CN	CN	CAP	Richards	Richards	Richards	Richards	Richards
INFponding	×	×	×	Darcy	×	×	×	G&A	G&A	G&A	G&A	×
45	SRn	2	2	2	3a	3a	2	2	2	2	All	3b
	ED <sub>e</sub>	PE,PM	PE,PT	PM,PT	PM,HG	PM,PT	PM	PM,PT	PM,HG	PM	SVAT	SVAT
	ED <sub>t</sub>	PM	PM,PT	PM,PT	PM,HG, SVAT	PM,PT,SVAT	PM	PM,PT	PM,HG	PM	SVAT	SVAT
	2-stage	×	×	✓	×	✓	✓	×	×	×	✓	✓
	Partitioning	INT-BL	INT-BL	INT-BL	INT-BL	INT-BL	SC-M	INT-BL	INT-BL,SC-M	INT-BL,SC-M	SC-M	INT-BL,SC-M
	K <sub>s</sub>	×	✓	✓	×	✓	✓	×	×	×	×	×
	S <sub>i</sub>	LIN	LIN	LIN,EXP	LIN	LIN	LIN	LIN,NLD	LIN,NLD	LIN	LIN	NLD
	CU	×	×	×	×	×	×	✓	✓	×	×	×
	Drainage	TBA	TBA	TBA	Richards	TBA	TBA	Richards	Richards	Richards	Richards	Richards
	CR	×	SnU-CR	×	SU-CR	D-CR	SnU-CR	D-CR	SU-CR	SU-CR	SU-CR	SU-CR
	Hysteresis	×	×	×	×	×	×	✓	✓	✓	×	×
	SSLF	×	×	×	OUT	×	×	OUT	OUT,IN	OUT	×	×
	PF	×	×	×	✓	×	×	✓	✓	✓	✓	×

Main sources: (1) Boogaard et al. (2014); (2) Hoogenboom et al. (2017); (3) Verburg (1996), Keating et al. (2003); (4) Hansen et al. (1990), Abrahamsen and Hansen (2000), Hansen et al. (2012); (5) Brisson et al. (2003); (6) Steduto et al. (2009), Raes et al. (2009a, 2017); (7) Nendel et al. (2011); (8) Simunek et al. (1998), Simunek et al. (2018a); (9) Simunek et al. (1999); (10) van Dam et al. (1997), van Dam (2000c), Kroes et al. (2017c); (11) Abbott et al. (1986), DHI (2017b); (12) Verburg et al. (1996).

## SUPPLEMENTARY MATERIAL - Chapter 2

**General description of selected models:** The following sections are dedicated to introducing each model regarding historical context, basic structure, and functioning, including some relevant calibration and simulation case-studies.

### 1. The WOrld FOod STudies (WOFOST)

The 'WOrld FOod STudies' (WOFOST) is a crop simulation model initially developed by the Centre for World Studies (Wageningen University & Research) in cooperation with the Centre for Agrobiological Research (CABO) both located in Wageningen, The Netherlands (Boogaard et al., 2014; de Wit et al., 2018; van Ittersum et al., 2003; van Van Diepen et al., 1989). WOFOST is a radiation-driven model which shares many of the fundamentals and algorithms with the earlier SUCROS model (Laar et al., 1997). The model estimates crop yield at three main production levels (Boogaard et al., 2014; van Van Diepen et al., 1989): 1) Potential crop yield; 2) Water-limited crop yield, assuming optimal nutrient supply; and 3) Nutrient-limited crop yield. In this review, we focused on the water-limited production level ( $Y_W$ ) since WOFOST only simulates and delivers the water balance components when crops grow under water-limited conditions (Boogaard et al., 2014; van Van Diepen et al., 1989). WOFOST computes the water balance through different approaches. Starting with the original single layer 'tipping bucket' approach (van Van Diepen et al., 1989), a more advanced soil water balance is also currently available coupled with the Soil Water Atmosphere Plant (SWAP) model, by applying a numerical solution to the Richard's equation at a variable time basis in a multi-layer representation (Boogaard et al., 2014; de Wit et al., 2018; Eitzinger et al., 2004). However, here we will consider only the original description of the water processes in WOFOST for comparative purposes. Many calibration and simulation case-studies are found in literature (Boogaard et al., 2013; de Koning et al., 1995; Dobermann et al., 2000; Eitzinger et al., 2004; Haberle and van Diepen, 1999; Kassie et al., 2014; Roetter, 1993; Supit, 1997; Todorovic et al., 2009; Wang et al., 2011; Wolf and van Diepen, 1994).

## 2. The decision support system for agrotechnology transfer (DSSAT)

The decision support system for agrotechnology transfer (DSSAT) is an integrated system of computer software coupling independent models for simulating cropping systems. DSSAT has evolved from previously available models such as CERES and CROPGRO (Jones et al., 2017a). This modelling effort started in the 1980's, when different groups produced models in isolation, leading to the divergence of the modelling community and approaches (Jones et al., 2003). Since that time, efforts were applied to create the present DSSAT (Jones et al., 2017a). Currently, the DSSAT is continuously tested and improved, partly inspired by the modelling structure of APSIM which resulted in the revision of the CROPGRO models (Jones et al., 2001). The present version of DSSAT (i.e. CERES-based) is radiation-driven and simulates the effects of management practices on cropping system production, both short and long term, delivering predictions for more than 20 crop species (Ahmad et al., 2012; Amiri et al., 2013; Bastos et al., 2002; Boote et al., 2018, 2002; Boote and Scholberg, 2006; Cammarano et al., 2012; Eitzinger et al., 2004; Giraldo et al., 1988; Griffin, 1993; Hartkamp et al., 2002; Hoogenboom et al., 1994; Yiwen Jiang et al., 2016; Jones et al., 2003; Keating et al., 1992; Liu et al., 2011; Malik et al., 2018; Matthews and Hunt, 1994; Modala et al., 2015; Pedreira et al., 2011; Robertson et al., 2002; Saseendran et al., 2010; Singh et al., 2017; Singh and Virmani, 1996; Šťastná et al., 2010; Travasso and Magrin, 1998; White et al., 1995; Woli et al., 2017).

### 3. The Agricultural Production Systems Simulator (APSIM)

The Agricultural Production Systems Simulator (APSIM) has been developed by the Agricultural Production Systems Research Unit (APSRU), a collaborative group from CSIRO and Queensland State Government agencies in Australia (Keating et al., 2003). APSIM was designed for a wide range of applications, from on-farm decision support for Australian farmers (<http://www.yieldprophet.com.au/>), to support risk assessment for government policy, also making it a guide to scientific research and education (Keating et al., 2003). The model is radiation-driven and follows two major modelling conceptual approaches for simulating the soil water balance. It follows a 'tipping bucket' approach (the SOILWAT module) or it applies a solution of the Richards' equation for one-dimension, when coupled with the soil and water integrated model (SWIM) (Connolly et al., 2002; Stewart et al., 2006). As for WOFOST, here, we focused only on the SOILWAT module of APSIM. The model has been calibrated for several crop species (Ahmed et al., 2016; Delve et al., 2009; Hearn, 1994; Keating et al., 1999; Manschadi et al., 2006; Mohanty et al., 2012; Robertson et al., 2002; Robertson and Lilley, 2016; Snow et al., 1999; Zheng et al., 2014). Additional documentation found at: <https://www.apsim.info/documentation/model-documentation/soil-modules-documentation/soilwat/>

### 4. The open soil-crop-atmosphere system model (DAISY)

The open soil-crop-atmosphere system model (DAISY) has been developed by the University of Copenhagen back in the 1990's and is focused on the simulation of soil organic matter dynamics, water-nitrogen balance and crop productivity (Hansen et al., 1991, 1990). The model is centered on the functioning of a crop-carbon module that simulates plant growth as a function of canopy photosynthesis and plant respiration, as affected by weather variables (i.e. global radiation, air temperature), plant state variables (i.e. LAI), and water and N stress coefficients (Abrahamsen and Hansen, 2000). The photosynthesis follows the structure of SUCROS model (Van Keulen et al., 1982). DAISY is particularly interesting for carbon and nutrient-balance studies with users such as policy makers interested in nitrogen leaching, carbon fixation and soil quality, as well as advisers focused on yield prediction and fertilization planning. The model has been used in multiple case-studies (Hansen et al., 1991; Jensen et al., 1994; Manevski et al., 2016; Rötter et al., 2012).

## 5. Simulateur multidisciplinaire pour les cultures standard (STICS)

'Simulateur multidisciplinaire pour les cultures standard' (STICS) is a model developed by a consortium back in 1996 (Brisson et al., 2003). STICS simulates daily crop growth and soil, water and nitrogen balances relying on radiation-driven relationships (Beaudoin et al., 2009; Brisson et al., 2003; Whisler et al., 1986). The model reveals high adaptability to several crop species by using generic parameters relevant for most crops, while keeping options concerning both physiological and management parameters that can be adjusted for each crop (Brisson et al., 2003). STICS has promoted a rich communication environment between users and developers leading to a constant improvement of the model (Hunink et al., 2011). Despite being initially developed for two main crops (wheat and maize), many simulation studies for multiple species may be found in the literature (Brisson et al., 2002; Garcia de Cortazar Atauri, 2006; Jégo et al., 2012; Rötter et al., 2012; Salo et al., 2016). STICS is able to simulate dormancy periods, bud breaking mechanisms, symbiotic N fixation, and detailed effects of mulching, which makes it possible to be used in simulations of perennial crops (Garcia de Cortazar Atauri, 2006).

## 6. The FAO crop model to simulate yield response to water (AquaCrop)

AquaCrop is a water-driven model for simulating productivity of major annual crops (Steduto et al., 2009). The model simulates biomass production by assuming a linear relation between biomass and transpiration through a water productivity (WP) conservative parameter (Steduto et al., 2012, 2009; Tanner and Sinclair, 1983). The WP value is normalized for reference evapotranspiration and CO<sub>2</sub> concentration (Steduto et al., 2007). Contrary to other models, AquaCrop uses canopy cover (CC) instead of LAI to characterize crop growth. Despite its relatively simplicity, AquaCrop is rooted in the fundamental processes involved in crop productivity with special emphasis on the responses to water deficits both from a physiological and an agronomic perspective (Dirk Raes et al., 2009). The model simulates yield responses to salinity and has been applied to the simulations of irrigated cropping systems (García-Vila et al., 2009; García-Vila and Fereres, 2012) and for decision support at the irrigation district level (Lorite et al., 2007). AquaCrop has been widely calibrated and validated during its lifetime (Araya et al., 2016, 2010a b, 2010b a; Ćosić et al., 2017; Espadafor et al., 2017; García-Vila et al., 2009; Geerts et al., 2009, 2008; Heng et al., 2009; Hsiao et al., 2009; Linker et al., 2016; Maniruzzaman et al., 2015; Montoya et al., 2016; Paredes et al., 2015; Rinaldi et al., 2011; Stricevic et al., 2011; Todorovic et al., 2009; Wellens et al., 2013; Zeleke et al., 2011).

## 7. The MOdel for Nltrogen and CArbon dynamics (MONICA)

MONICA is the acronym for "MOdel of Nltrogen and CArbon dynamics in Agro-ecosystems" and has been developed by the Institute of Landscape Systems Analysis from the Leibniz Centre for Agricultural Landscape Research (ZALF). It was first released in 2011 and since then it has been updated to the V2.0 version published in 2016 (Berg, 2018; Nendel, 2014; Nendel et al., 2011). MONICA describes mechanistically the transport and bio-chemical cycling of carbon, nitrogen and water in agro-ecosystems at daily time steps (Berg, 2018). The model simulates crop growth with a radiation-driven approach as in the SUCROS model (Van Keulen et al., 1982). The model emerged from improvements on the carbon cycle of the HERMES model (Kersebaum, 2007, 1995). HERMES simulation of nitrogen mineralization from the soil organic pool, was replaced by an approach similar to that of DAISY, which also takes into account the dynamics of soil microbial biomass (Ahuja et al., 2014; ISMC, 2017; Wallor et al., 2018). MONICA has also been used to develop decision support systems (<http://www.landcare2020.de>), used on both tropical and temperate conditions (Wenkel et al., 2013). The model has been calibrated for several crop species used in several case-studies (Nendel et al., 2013, 2011; Rötter et al., 2012; Salo et al., 2016; Speck et al., 2015).

## 8. The HYDRUS-1D package

The HYDRUS-1D is a software package developed by the International Groundwater Modelling Center and the University of California at Riverside. The model aims primarily to generate a finite element solution capable to simulate one-dimensional movement of water (i.e. lumped scale), heat, and multiple solutes in variably saturated media (Simunek et al., 1998). The HYDRUS packages use a Microsoft Windows based graphical user interface for input data management and discretization of nodes, parameterization, execution and output visualization. Recent applications of the HYDRUS-1D include evaluations of irrigation schemes, studies of root water uptake, groundwater recharge and nutrient transport (Li et al., 2017; Šimůnek et al., 2016; Slama et al., 2019; Tafteh and Sepaskhah, 2012; Wang et al., 2014; Zhou and Zhao, 2019). The model estimates crop water uptake based on potential evapotranspiration as an input or through radiation-driven crop growth, such as the one in SUCROS or WOFOST used in parallel (Li et al., 2017). Crop-simulation with HYDRUS-1D has been done for multiple crop species (Hou et al., 2017; Li et al., 2017; Tafteh and Sepaskhah, 2012; Zhou and Zhao, 2019).

## 9. The HYDRUS-2D package

HYDRUS-2D emerged from HYDRUS-1D. It is also a Microsoft Windows based Modelling environment for analyzing water flow and solute transport in variably saturated porous media, but in this case for two-dimensions. The software package includes a 2D finite element approach for simulating the movement of water, heat, and multiple solutes. The model includes a parameter optimization algorithm for inverse estimation of soil hydraulic and/or solute transport parameters. The model is supported by an interactive graphics-based interface for data-preprocessing, generation of a multidimensional structured grid, and graphic presentation of the results. Regarding crop growth and water uptake, a similar approach as in HYDRUS-1D is followed. Several applications in regard to irrigation management, simulation of flow to drainage systems, movement of salts and agro-chemicals, and seasonal simulation of water flow can (Deb Sanjit K. et al., 2016; Karandish and Šimůnek, 2016; Siyal et al., 2012; Skaggs et al., 2004). The software package has been upgraded and it is now fully integrated with the HYDRUS-2D/3D version which is the current extension developed for both 2D and 3D applications (Šimůnek et al., 2018). The new version has a new graphical environment and it is commercially distributed (i.e. not for free) by the International Groundwater Modelling Center (IGWMC).

## 10. The Soil Water Atmosphere and Plant model (SWAP)

SWAP was designed in Wageningen to simulate flow and transport processes at field scale level (Van Dam et al., 1997). The model simulates transport of water, solutes and heat in unsaturated/saturated soils, and it addresses both research and practical questions in the field of agriculture, water management and environmental analyses (Kroes and Sutip, 2011; van Dam, 2000). The crop growth module is based on that of WOFOST (Kroes and Sutip, 2011; van Dam, 2000), but the model can also be coupled with other crop models for estimation of crop related parameters (Dokoochaki et al., 2016). The soil moisture, heat and solute modules exchange status information each time step. Crop growth is affected by actual soil moisture and salinity on a daily basis but hydrological processes (i.e. infiltration, runoff, etc) are simulated at time-steps less than one minute. The model accounts for macro-porous flow and water repellency situations. Main applications are related to surface water management and drainage studies (Ben-Asher et al., 2006; Yao Jiang et al., 2016; Kroes and Sutip, 2011; Sarwar et al., 2000; Smets et al., 1997; van Schaik et al., 2010).

## 11. The integrated catchment model - 'Système Hydrologique Européen' (MIKE-SHE)

MIKE-SHE emerged from the 'Système Hydrologique Européen (SHE)' (Abbott et al., 1986) as the result of an European consortium formed by the Institute of Hydrology (UK), the engineering company SOGREAH (FR) and the Danish Hydraulic Institute (DHI). Its tailor-based approach allows the user to adjust the complexity of the modelling approaches, according to the conditions of the intended case-study (DHI, 2017a a, 2017b b). Within the current review and for simplification purposes, MIKE-SHE is described according to the approaches set by default (i.e. demo model). Information on other options may be found in the latest version of the manual (DHI, 2017a a, 2017b b). MIKE SHE has been developed for integrated catchment hydrology applications, such as the conjunctive use and management of surface and groundwater, irrigation and drought management, wetland management and restoration, environmental river flows modelling, floodplain management, groundwater-induced flooding simulation, nutrient transportation and management and integrated mine water management (Butts and Graham, 2005). The model has been used at the crop field scale, too (Hashemi et al., 2018; Singh et al., 1999).

## 12. Soil Water Infiltration and Movement model (SWIM)

SWIM is a software package developed in Australia within the CSIRO Division of Soils (Verburg et al., 1996b a). The model has been used to simulate runoff, infiltration, redistribution flow and transport of nutrients, plant uptake and transpiration, soil evaporation, and vertical drainage (Ross, 1990; Verburg et al., 1996a b, 1996b a). The model does not compute evapotranspiration which must be taken as an input parameter, or from a radiation-driven crop model, such as APSIM that must be run in parallel. This has been done for some crop species under different growing conditions (Stewart et al., 2006; Verburg et al., 1996a).

## 13. Synthesis of models

Table: Synthesis of models. *Codes:* Synthesis of models. Codes: (NL) = The Netherlands, (USA) = United States of America, (AUS) = Australia, (DK) = Denmark, (FR) = France, (DE) = Germany, WUR = Wageningen University and Research, CABO = Centre for Agrobiological Research, CSIRO = Commonwealth Scientific and Industrial Research Organisation, DHI – Danish Hydraulic Institute for Water and Environment, USAID = US Agency for International Development, USDA = US Department of Agriculture, DA-DBSE = Department of Agriculture and Department of Basic Sciences and Environment from the University of Copenhagen, INRA = French National Institute for Agricultural Research, FAO (UN) = Food and Agriculture Organization of United Nations, ZALF = Institute of Leibniz Centre for Agricultural Landscape Research, IGWMC = International Groundwater Modelling Center, UCR = University of California Riverside. The 'time-scale' indicates the extent of each simulation time-step (i.e. hourly or less than one-hour when models simulate water processes in a scale of minutes). The 'spatial-scale' is characterized as 'point-based' for cases where all water flows occur vertically or as 'semi-distributed' for models that partially distribute water over the horizontal space.

ID	Model	Type	Released	Research institution	Country	Spatial-scale	Time-scale
1	WOFOST	Crop	1980's	WUR/CABO	NL	Point-based	daily
2	DSSAT	Crop	1980's	USAID/USA	USA	Point-based	daily
3	APSIM	Crop	1990's	CSIRO	AUS	Point-based	daily
4	DAISY	Crop	1990's	DA-DBSE	DK	Semi-Distributed	daily
5	STICS	Crop	1990's	INRA	FR	Point-based	daily
6	AquaCrop	Crop	2000's	FAO	UN	Point-based	daily
7	MONICA	Crop	2010's	ZALF	DE	Point-based	daily
8	HYDRUS-1D	Hydrology	1980's	IGWMC/UCR	USA	Semi-Distributed	<hour
9	HYDRUS-2D	Hydrology	1990's	IGWMC/UCR	USA	Semi-Distributed	<hour
10	SWAP	Hydrology	1990's	WUR	NL	Semi-Distributed	<hour
11	MIKE-SHE	Hydrology	1990's	DHI	DK	Semi-Distributed	<hour
12	SWIM	Hydrology	1990's	CSIRO	AUS	Point-based	<hour

*Original sources:* 1) van Van Diepen et al. (1989); 2) Jones et al. (2003); 3) Keating et al. (2003); 4) Hansen et al. (1990); 5) Brisson et al. (2003); 6) Steduto et al. (2009); 7) Nendel et al. (2011); 8) Simunek et al. (1998); 9) Simunek et al. (1999); 10) van Dam et al. (1997); 11) Abbott et al. (1986); 12) Verburg et al. (1996).

## 14. Supplementary references [general description of models]

Ahmad, S., Ahmad, A., Soler, C.M.T., Ali, H., Zia-Ul-Haq, M., Anothai, J., Hussain, A., Hoogenboom, G., Hasanuzzaman, M., 2012. Application of the CSM-CERES-Rice model for evaluation of plant density and nitrogen management of fine transplanted rice for an irrigated semiarid environment. Precis. Agric. 13, 200–218.

Ahmed, M., Akram, M.N., Asim, M., Aslam, M., Hassan, F.-U., Higgins, S., Stöckle, C.O., Hoogenboom, G., 2016. Calibration and validation of APSIM-Wheat and CERES-Wheat for spring wheat under rainfed conditions: Models evaluation and application. Comput. Electron. Agric. 123, 384–401.

Ahuja, L.R., Ma, L., Lascano, R.J., Saseendran, S.A., Fang, Q.X., Nielsen, D.C., Wang, E., Colaizzi, P.D., 2014. Syntheses of the Current Model Applications for Managing Water and Needs for Experimental Data and Model Improvements to Enhance these Applications, in: Practical Applications of Agricultural System Models to Optimize the Use of Limited Water, Advances in Agricultural Systems Modeling. American Society of Agronomy, Inc., Crop Science Society of America, Inc., and Soil Science Society of America, Inc., Madison, WI, pp. 399–438.

Amiri, E., Rezaei, M., Bannayan, M., Soufizadeh, S., 2013. Calibration and Evaluation of CERES Rice Model under Different Nitrogen- and Water-Management Options in Semi-Mediterranean Climate Condition. Commun. Soil Sci. Plant Anal. 44, 1814–1830.

- Araya, A., Habtu, S., Hadgu, K.M., Kebede, A., Dejene, T., 2010a. Test of AquaCrop model in simulating biomass and yield of water deficient and irrigated barley (*Hordeum vulgare*). Agric. Water Manage. 97, 1838–1846.
- Araya, A., Keesstra, S.D., Stroosnijder, L., 2010b. Simulating yield response to water of Teff (*Eragrostis tef*) with FAO's AquaCrop model. Field Crops Res. 116, 196–204.
- Araya, A., Kisekka, I., Holman, J., 2016. Evaluating deficit irrigation management strategies for grain sorghum using AquaCrop. Irrig. Sci. 34, 465–481.
- Ben-Asher, J., van Dam, J., Feddes, R.A., Jhorar, R.K., 2006. Irrigation of grapevines with saline water: II. Mathematical simulation of vine growth and yield. Agric. Water Manage. 83, 22–29.
- Berg, M., 2018. MONICA Agrosystems-model [WWW Document]. URL <https://github.com/zalf-rpm/monica> (accessed 12.12.18).
- Boote, K.J., Mínguez, M.I., Sau, F., 2002. Adapting the CROPGRO legume model to simulate growth of faba bean. Agron. J. 94, 743–756.
- Boote, K.J., Prasad, V., Allen, L.H., Singh, P., Jones, J.W., 2018. Modeling sensitivity of grain yield to elevated temperature in the DSSAT crop models for peanut, soybean, dry bean, chickpea, sorghum, and millet. Eur. J. Agron. 100, 99–109.
- Boote, K.J., Scholberg, J.M.S., 2006. Developing, parameterizing, and testing of dynamic crop growth models for horticultural crops, in: III International Symposium on Models for Plant Growth, Environmental Control and Farm Management in Protected Cultivation 718. actahort.org, pp. 23–34.
- Butts, M., Graham, D., 2005. Flexible Integrated Watershed Modeling with MIKE SHE, in: Frevert, D., Singh, V. (Eds.), Watershed Models. CRC Press, pp. 245–271.
- Cammarano, D., Payero, J., Basso, B., Wilkens, P., Grace, P., 2012. Agronomic and economic evaluation of irrigation strategies on cotton lint yield in Australia. Crop Pasture Sci. 63, 647–655.
- Connolly, R.D., Bell, M., Huth, N., Freebairn, D.M., Thomas, G., 2002. Simulating infiltration and the water balance in cropping systems with APSIM-SWIM. Soil Res. 40, 221–242.
- Ćosić, M., Stričević, R., Djurović, N., Moravčević, D., Pavlović, M., Todorović, M., 2017. Predicting biomass and yield of sweet pepper grown with and without plastic film mulching under different water supply and weather conditions. Agric. Water Manage. 188, 91–100.
- Deb Sanjit K., Sharma Parmodh, Shukla Manoj K., Ashigh Jamshid, Šimůnek Jiří, 2016. Numerical Evaluation of Nitrate Distributions in the Onion Root Zone under Conventional Furrow Fertigation. J. Hydrol. Eng. 21, 05015026.
- Delve, R.J., Probert, M.E., Cobo, J.G., Ricaurte, J., Rivera, M., Barrios, E., Rao, I.M., 2009. Simulating phosphorus responses in annual crops using APSIM: model evaluation on contrasting soil types. Nutr. Cycling Agroecosyst. 84, 293–306.
- Dokoochaki, H., Gheysari, M., Mousavi, S.-F., Zand-Parsa, S., Miguez, F.E., Archontoulis, S.V., Hoogenboom, G., 2016. Coupling and testing a new soil water module in DSSAT CERES-Maize model for maize production under semi-arid condition. Agric. Water Manage. 163, 90–99.
- Eitzinger, J., Trnka, M., Hösch, J., Žalud, Z., Dubrovský, M., 2004. Comparison of CERES, WOFOST and SWAP models in simulating soil water content during growing season under different soil conditions. Ecol. Modell. 171, 223–246.
- Espadafor, M., Couto, L., Resende, M., Henderson, D.W., García-Vila, M., Fereres, E., 2017. Simulation of the Responses of Dry Beans (*Phaseolus vulgaris* L.) to Irrigation. Transactions of the ASABE 60, 1983–1994.
- Garcia de Cortazar Atauri, I., 2006. Adaptation du modèle STICS à la vigne (*Vitis vinifera* L.): utilisation dans le cadre d'une étude d'impact du changement climatique à l'échelle de la France. École nationale supérieure agronomique (Montpellier).
- Geerts, S., Raes, D., Garcia, M., Condori, O., Mamani, J., Miranda, R., Cusicanqui, J., Taboada, C., Yucra, E., Vacher, J., 2008. Could deficit irrigation be a sustainable practice for quinoa (*Chenopodium quinoa* Willd.) in the Southern Bolivian Altiplano? Agric. Water Manage. 95, 909–917.
- Geerts, S., Raes, D., Garcia, M., Miranda, R., Cusicanqui, J.A., Taboada, C., Mendoza, J., Huanca, R., Mamani, A., Condori, O., Mamani, J., Morales, B., Osco, V., Steduto, P., 2009. Simulating Yield Response of Quinoa to Water

Availability with AquaCrop. *Agron. J.* 101, 499–508.

Giraldo, L.M., Lizcano, L.J., Gijsman, A.J., Rivera, B., Franco, L.H., 1988. Adaptación del modelo DSSAT para simular la producción de Brachiaria decumbens.

Griffin, T.S., 1993. simulation model for potato growth and development.

Haberle, J., van Diepen, C.A., oct1999. Variability of potential and water-limited sugar beet and spring wheat yields simulated with WOFOST model. Rostlinna Vyroba – UZPI (Czech Republic) 45.

Hartkamp, A.D., Hoogenboom, G., White, J.W., 2002. Adaptation of the CROPGRO growth model to velvet bean (*Mucuna pruriens*): I. Model development. *Field Crops Res.* 78, 9–25.

Hashemi, F., Olesen, J.E., Hansen, A.L., Børgesen, C.D., Dalgaard, T., 2018. Spatially differentiated strategies for reducing nitrate loads from agriculture in two Danish catchments. *J. Environ. Manage.* 208, 77–91.

Hearn, A.B., 1994. OZCOT: A simulation model for cotton crop management. *Agric. Syst.* 44, 257–299.

Heng, L.K., Hsiao, T., Evett, S., Howell, T., Steduto, P., 2009. Validating the FAO AquaCrop Model for Irrigated and Water Deficient Field Maize. *Agron. J.* 101, 488–498.

Hoogenboom, G., White, J.W., Jones, J.W., Boote, K.J., 1994. BEANGRO: A Process-Oriented Dry Bean Model with a Versatile User Interface. *Agron. J.* 86, 182–190.

Hou, L., Zhou, Y., Bao, H., Wenninger, J., 2017. Simulation of maize (*Zea mays* L.) water use with the HYDRUS-1D model in the semi-arid Hailiutu River catchment, Northwest China. *Hydrol. Sci. J.* 62, 93–103.

Hsiao, T.C., Heng, L., Steduto, P., Rojas-Lara, B., Raes, D., Fereres, E., 2009. AquaCrop—The FAO Crop Model to Simulate Yield Response to Water: III. Parameterization and Testing for Maize. *Agron. J.* 101, 448–459.

ISMС, 2017. MONICA — International Soil Modeling Consortium [WWW Document]. URL <https://soil-modeling.org/resources-links/model-portal/monica> (accessed 12.12.18).

Jégo, G., Pattey, E., Liu, J., 2012. Using Leaf Area Index, retrieved from optical imagery, in the STICS crop model for predicting yield and biomass of field crops. *Field Crops Res.* 131, 63–74.

Jensen, C., Stougaard, B., Olsen, P., 1994. Simulation of Water and Nitrogen Dynamics at Three Danish Locations by Use of the DAISY Model. *Acta Agric. Scand. Sect. B. Soil Plant Sci.* 44, 75–83.

Jiang, Y., Xu, X., Huang, Q., Huo, Z., Huang, G., 2016. Optimizing regional irrigation water use by integrating a two-level optimization model and an agro-hydrological model. *Agric. Water Manage.* 178, 76–88.

Jiang, Y., Zhang, L., Zhang, B., He, C., Jin, X., Bai, X., 2016. Modeling irrigation management for water conservation by DSSAT-maize model in arid northwestern China. *Agric. Water Manage.* 177, 37–45.

Karandish, F., Šimůnek, J., 2016. A field-modeling study for assessing temporal variations of soil-water-crop interactions under water-saving irrigation strategies. *Agric. Water Manage.* 178, 291–303.

Kassie, B.T., Van Ittersum, M.K., Hengsdijk, H., Asseng, S., Wolf, J., Rötter, R.P., 2014. Climate-induced yield variability and yield gaps of maize (*Zea mays* L.) in the Central Rift Valley of Ethiopia. *Field Crops Res.* 160, 41–53.

Keating, B.A., Robertson, M.J., Muchow, R.C., Huth, N.I., 1999. Modelling sugarcane production systems I. Development and performance of the sugarcane module. *Field Crops Res.* 61, 253–271.

Keating, B.A., Wafula, B.M., Watiki, J.M., 1992. Development of a modelling capability for maize in semi-arid eastern Kenya, in: A Search for Strategies for Sustainable Dryland Cropping in Semi-Arid Eastern Kenya. ACIAR Proceedings. bob-mccown.com, pp. 26–33.

Kersebaum, K.C., 2007. Modelling nitrogen dynamics in soil–crop systems with HERMES, in: Modelling Water and Nutrient Dynamics in Soil–crop Systems. Springer Netherlands, pp. 147–160.

Kersebaum, K.C., 1995. Application of a simple management model to simulate water and nitrogen dynamics. *Ecol. Modell.* 81, 145–156.

Kroes, J., Supit, I., Van Dam, J., Van Walsum, P., Mulder, M., 2017. Impact of capillary rise and recirculation on crop yields. <https://doi.org/10.5194/hess-2017-223>

Kroes, J.G., Supit, I., 2011. Impact analysis of drought, water excess and salinity on grass production in The Netherlands using historical and future climate data. *Agric. Ecosyst. Environ.* 144, 370–381.

Laar, H.H., van Goudriaan, J., van Keulen, H., 1997. SUCROS97: Simulation of crop growth for potential and water-limited production situations. As applied to spring wheat.

- Linker, R., Ioslovich, I., Sylaios, G., Plauborg, F., Battilani, A., 2016. Optimal model-based deficit irrigation scheduling using AquaCrop: A simulation study with cotton, potato and tomato. *Agric. Water Manage.* 163, 236–243.
- Liu, H.L., Yang, J.Y., Drury, C.F., Reynolds, W.D., Tan, C.S., Bai, Y.L., He, P., Jin, J., Hoogenboom, G., 2011. Using the DSSAT-CERES-Maize model to simulate crop yield and nitrogen cycling in fields under long-term continuous maize production. *Nutr. Cycling Agroecosyst.* 89, 313–328.
- Malik, W., Boote, K.J., Hoogenboom, G., Cavero, J., Dechmi, F., 2018. Adapting the CROPGRO model to simulate alfalfa growth and yield. *Agron. J.*
- Manevski, K., Børgesen, C.D., Li, X., Andersen, M.N., Abrahamsen, P., Hu, C., Hansen, S., 2016. Integrated modelling of crop production and nitrate leaching with the Daisy model. *MethodsX* 3, 350–363.
- Maniruzzaman, M., Talukder, M.S.U., Khan, M.H., Biswas, J.C., Nemes, A., 2015. Validation of the AquaCrop model for irrigated rice production under varied water regimes in Bangladesh. *Agric. Water Manage.* 159, 331–340.
- Manschadi, A.M., Hochman, Z., Mclean, G., DeVool, P., Holzworth, D., Meinke, H., 2006. APSIM-Barley model—adaptation of a wheat model to simulate barley growth and development.
- Matthews, R.B., Hunt, L.A., 1994. GUMCAS: a model describing the growth of cassava (*Manihot esculenta* L. Crantz). *Field Crops Res.* 36, 69–84.
- Modala, N.R., Ale, S., Rajan, N., Munster, C.L., DeLaune, P.B., Thorp, K.R., Nair, S.S., Barnes, E.M., 2015. Evaluation of the CSM-CROPGRO-Cotton model for the Texas rolling plains region and simulation of deficit irrigation strategies for increasing water use efficiency. *Transactions of the ASABE* 58, 685–696.
- Mohanty, M., Probert, M.E., Reddy, K.S., Dalal, R.C., Mishra, A.K., Subba Rao, A., Singh, M., Menzies, N.W., 2012. Simulating soybean–wheat cropping system: APSIM model parameterization and validation. *Agric. Ecosyst. Environ.* 152, 68–78.
- Montoya, F., Camargo, D., Ortega, J.F., Córcoles, J.I., Domínguez, A., 2016. Evaluation of Aquacrop model for a potato crop under different irrigation conditions. *Agric. Water Manage.* 164, 267–280.
- Nendel, C., 2014. MONICA: A Simulation Model for Nitrogen and Carbon Dynamics in Agro-Ecosystems, in: Mueller, L., Saparov, A., Lischeid, G. (Eds.), *Novel Measurement and Assessment Tools for Monitoring and Management of Land and Water Resources in Agricultural Landscapes of Central Asia*. Springer International Publishing, Cham, pp. 389–405.
- Nendel, C., Berg, M., Kersebaum, K.C., Mirschel, W., Specka, X., Wegehenkel, M., Wenkel, K.O., Wieland, R., 2011. The MONICA model: Testing predictability for crop growth, soil moisture and nitrogen dynamics. *Ecol. Modell.* 222, 1614–1625.
- Nendel, C., Wieland, R., Mirschel, W., Specka, X., Guddat, C., Kersebaum, K.C., 2013. Simulating regional winter wheat yields using input data of different spatial resolution. *Field Crops Res.* 145, 67–77.
- Paredes, P., Wei, Z., Liu, Y., Xu, D., Xin, Y., Zhang, B., Pereira, L.S., 2015. Performance assessment of the FAO AquaCrop model for soil water, soil evaporation, biomass and yield of soybeans in North China Plain. *Agric. Water Manage.* 152, 57–71.
- Pedreira, B.C., Pedreira, C.G.S., Boote, K.J., Lara, M.A.S., Alderman, P.D., 2011. Adapting the CROPGRO perennial forage model to predict growth of Brachiaria brizantha. *Field Crops Res.* 120, 370–379.
- Rinaldi, M., Garofalo, P., Rubino, P., Steduto, P., 2011. Processing tomatoes under different irrigation regimes in Southern Italy: agronomic and economic assessments in a simulation case study. *J Agrometeorol* 3, 39–56.
- Robertson, M.J., Carberry, P.S., Huth, N.I., Turpin, J.E., Probert, M.E., Poulton, P.L., Bell, M., Wright, G.C., Yeates, S.J., Brinsmead, R.B., 2002. Simulation of growth and development of diverse legume species in APSIM. *Aust. J. Agric. Res.* 53, 429–446.
- Robertson, M.J., Lilley, J.M., 2016. Simulation of growth, development and yield of canola (*Brassica napus*) in APSIM. *Crop Pasture Sci.* 67, 332–344.
- Roetter, R., 1993. Simulation of the biophysical limitations to maize production under rainfed conditions in Kenya: evaluation and application of the model WOFOST.
- Ross, P.J., 1990. SWIM: a simulation model for soil water infiltration and movement: reference manual.

Rötter, R.P., Palosuo, T., Kersebaum, K.C., Angulo, C., Bindi, M., Ewert, F., Ferrise, R., Hlavinka, P., Moriondo, M., Nendel, C., Olesen, J.E., Patil, R.H., Ruget, F., Takáč, J., Trnka, M., 2012. Simulation of spring barley yield in different climatic zones of Northern and Central Europe: A comparison of nine crop models. *Field Crops Res.* 133, 23–36.

Salo, T.J., Palosuo, T., Kersebaum, K.C., Nendel, C., Angulo, C., Ewert, F., Bindi, M., Calanca, P., Klein, T., Moriondo, M., Ferrise, R., Olesen, J.E., Patil, R.H., Ruget, F., Takáč, J., Hlavinka, P., Trnka, M., Rötter, R.P., 2016. Comparing the performance of 11 crop simulation models in predicting yield response to nitrogen fertilization. *J. Agric. Sci.* 154, 1218–1240.

Sarwar, A., Bastiaanssen, W.G.M., Boers, T.M., Van Dam, J.C., 2000. Devaluating drainage design parameters for the fourth drainage project, Pakistan by using SWAP model: part I—calibration. *Irrig. Drain. Syst.* 14, 257–280.

Saseendran, S.A., Nielsen, D.C., Ma, L., Ahuja, L.R., 2010. Adapting CROPGRO for Simulating Spring Canola Growth with Both RZWQM2 and DSSAT 4.0. *Agron. J.* 102, 1606–1621.

Singh, P., Boote, K.J., Kadiyala, M.D.M., Nedumaran, S., Gupta, S.K., Srinivas, K., Bantilan, M.C.S., 2017. An assessment of yield gains under climate change due to genetic modification of pearl millet. *Sci. Total Environ.* 601–602, 1226–1237.

Singh, P., Virmani, S.M., 1996. Modeling growth and yield of chickpea (*Cicer arietinum* L.). *Field Crops Res.* 46, 41–59.

Singh, R., Subramanian, K., Refsgaard, J.C., 1999. Hydrological modelling of a small watershed using MIKE SHE for irrigation planning. *Agric. Water Manage.* 41, 149–166.

Siyal, A.A., Bristow, K.L., Šimůnek, J., 2012. Minimizing nitrogen leaching from furrow irrigation through novel fertilizer placement and soil surface management strategies. *Agric. Water Manage.* 115, 242–251.

Skaggs, T.H., Trout, T.J., Simunek, J., Shouse, P.J., 2004. Comparison of HYDRUS-2D Simulations of Drip Irrigation with Experimental Observations. *J. Irrig. Drain. Eng.* 130, 304–310.

Slama, F., Zemni, N., Bouksila, F., De Mascellis, R., Bouhlila, R., 2019. Modelling the Impact on Root Water Uptake and Solute Return Flow of Different Drip Irrigation Regimes with Brackish Water. *Water* 11, 425.

Smets, S.M.P., Kuper, M., Van Dam, J.C., Feddes, R.A., 1997. Salinization and crop transpiration of irrigated fields in Pakistan's Punjab. *Agric. Water Manage.* 35, 43–60.

Snow, V.O., Bond, W.J., Myers, B.J., Theiveyanathan, S., Smith, C.J., Benyon, R.G., 1999. Modelling the water balance of effluent-irrigated trees. *Agric. Water Manage.* 39, 47–67.

Specka, X., Nendel, C., Wieland, R., 2015. Analysing the parameter sensitivity of the agro-ecosystem model MONICA for different crops. *Eur. J. Agron.* 71, 73–87.

Stasna, M., Toman, F., Dufkova, J., 2010. Usage of SUBSTOR model in potato yield prediction. *Agric. Water Manage.* 97, 286–290.

Stewart, L.K., Charlesworth, P.B., Bristow, K.L., Thorburn, P.J., 2006. Estimating deep drainage and nitrate leaching from the root zone under sugarcane using APSIM-SWIM. *Agric. Water Manage.* 81, 315–334.

Stricevic, R., Cosic, M., Djurovic, N., Pejic, B., Maksimovic, L., 2011. Assessment of the FAO AquaCrop model in the simulation of rainfed and supplementally irrigated maize, sugar beet and sunflower. *Agric. Water Manage.* 98, 1615–1621.

Supit, I., 1997. Predicting national wheat yields using a crop simulation and trend models. *Agric. For. Meteorol.* 88, 199–214.

Tafteh, A., Sepaskhah, A.R., 2012. Application of HYDRUS-1D model for simulating water and nitrate leaching from continuous and alternate furrow irrigated rapeseed and maize fields. *Agric. Water Manage.* 113, 19–29.

Todorovic, M., Albrizio, R., Zivotic, L., Saab, M.-T.A., Stöckle, C., Steduto, P., 2009. Assessment of AquaCrop, CropSyst, and WOFOST Models in the Simulation of Sunflower Growth under Different Water Regimes. *Agron. J.* 101, 509–521.

Travasso, M.I., Caldiz, D.O., Saluzzo, J.A., 1996. Yield prediction using the SUBSTOR-potato model under Argentinian conditions. *Potato Res.* 39, 305–312.

Travasso, M.I., Magrin, G.O., 1998. Utility of CERES-Barley under Argentine conditions. *Field Crops Res.* 57, 329–333.

- van Keulen, H., Seligman, H.G., 1987. Simulation of water use, nitrogen nutrition and growth of a spring wheat crop. Pudoc: Wageningen, The Netherlands).
- van Schaik, N.L.M.B., Hendriks, R.F.A., van Dam, J.C., 2010. Parameterization of Macropore Flow Using Dye-Tracer Infiltration Patterns in the SWAP Model. *Vadose Zone J.* 9, 95–106.
- Wallor, E., Kersebaum, K.-C., Ventrella, D., Bindi, M., Cammarano, D., Coucheney, E., Gaiser, T., Garofalo, P., Giglio, L., Giola, P., Hoffmann, M.P., Iocola, I., Lana, M., Lewan, E., Maharjan, G.R., Moriondo, M., Mula, L., Nendel, C., Pohankova, E., Roggero, P.P., Trnka, M., Trombi, G., 2018. The response of process-based agro-ecosystem models to within-field variability in site conditions. *Field Crops Res.* 228, 1–19.
- Wang, C., Mao, X., Hatano, R., 2014. Modeling Ponded Infiltration in Fine Textured Soils with Coarse Interlayer. *Soil Sci. Soc. Am. J.* 78, 745–753.
- Wang, T., Lu, C., Yu, B., 2011. Production potential and yield gaps of summer maize in the Beijing-Tianjin-Hebei Region. *J. Geogr. Sci.* 21, 677–688.
- Wellens, J., Raes, D., Traore, F., Denis, A., Djaby, B., Tychon, B., 2013. Performance assessment of the FAO AquaCrop model for irrigated cabbage on farmer plots in a semi-arid environment. *Agric. Water Manage.* 127, 40–47.
- Wenkel, K.-O., Berg, M., Mirschel, W., Wieland, R., Nendel, C., Köstner, B., 2013. LandCaRe DSS—An interactive decision support system for climate change impact assessment and the analysis of potential agricultural land use adaptation strategies. *J. Environ. Manage.* 127, S168–S183.
- Whisler, F.D., Acock, B., Baker, D.N., Fye, R.E., Hodges, H.F., Lambert, J.R., Lemmon, H.E., McKinnon, J.M., Reddy, V.R., 1986. Crop simulation models in agronomic systems. *Adv. Agron.* 40, 41–208.
- White, J.W., Hoogenboom, G., Jones, J.W., Boote, K.J., 1995. Evaluation of the Dry Bean Model BEANGRO V1.01 for Crop Production Research in a Tropical Environment. *Exp. Agric.* 31, 241–254.
- Wolf, J., van Diepen, C.A., 1994. Effects of climate change on silage maize production potential in the European community. *Agric. For. Meteorol.* 71, 33–60.
- Woli, P., Rouquette, F.M., Long, C.R., Gowda, P., Pequeno, D.N.L., 2017. Simulated bermudagrass production and nitrate leaching affected by El Niño-Southern oscillation, soil, and clipping frequency. *Agron. J.* 109, 2649–2661.
- Zeleke, K.T., Luckett, D., Cowley, R., 2011. Calibration and Testing of the FAO AquaCrop Model for Canola. *Agron. J.* 103, 1610–1618.
- Zheng, B., Chenu, K., Doherty, A., Chapman, S., 2014. The APSIM-Wheat Module (7.5 R3008). Agricultural Production Systems Simulator (APSIM) Initiative: Toowoomba, Australian.
- Zhou, H., Zhao, W.Z., 2019. Modeling soil water balance and irrigation strategies in a flood-irrigated wheat-maize rotation system. A case in dry climate, China. *Agric. Water Manage.* 221, 286–302.



# Chapter 3

## Using NDVI for the assessment of canopy cover in agricultural crops within modelling research

This chapter has been published as:

Tenreiro, T. R., García-Vila, M., Gómez, J. A., Jiménez-Berni, J. A., & Fereres, E. (2021). Using NDVI for the assessment of canopy cover in agricultural crops within modelling research. Computers and Electronics in Agriculture, 182, 106038.

### Abstract

The fraction of green canopy cover (CC) is an important feature commonly used to characterize crop growth and for calibration of crop and hydrological models. It is well accepted that there is a relation between CC and NDVI through linear or quadratic models, however a straight-forward empirical approach, to derive CC from NDVI observations, is still lacking. In this study, we conducted a meta-analysis of the NDVI-CC relationships with data collected from 19 different studies ( $N=1397$ ). Generic models are proposed here for 13 different agricultural crops, and the associated degree of uncertainty, together with the magnitude of error were quantified for each model (RMSE around 6–18% of CC). We observed that correlations are adequate for the majority of crops as  $R^2$  values were above 75% for most cases, and coefficient estimates were significant for most of the linear and quadratic models. Extrapolation to conditions different than those found in the studies may require local validation, as obtained regressions are affected by non-sampling errors or sources of systematic error that need further investigation. In a case study with wheat, we tested the use of NDVI as a proxy to estimate CC and to calibrate the AquaCrop model. Simulation outcomes were validated with field data collected from three growing seasons and confirmed that the NDVI-CC relationship was useful for modelling research. We highlight that the overall applicability of these relationships to modelling is promising as the RMSE are in line with acceptable levels published in several sensitivity analyses.

### 3.1 Introduction

The fraction of green canopy cover (CC) is defined as the fraction of projected ground area covered by photosynthetically active vegetation (Wittich and Hansing, 1995). This key vegetation feature is used to quantify crop canopy growth, radiation interception, and evapotranspiration partitioning in both crop and hydrological modelling applications (Allen and Pereira, 2009; Bouman, 1995; Dorigo et al., 2007; Gómez et al., 2009; Steduto et al., 2009; Steven et al., 1986; Tenreiro et al., 2020).

Proximal and remote sensing have both been used to characterize CC through the use of the 'Normalized Difference Vegetation Index' (NDVI) (Jasinski, 1990; Hatfield et al., 2008; Plant, 2001; Trout et al., 2008). In addition to the NDVI, other indices have been developed to characterize the vegetation through remote sensing observations. Examples are the Soil Adjusted Vegetation Index (SAVI), the Atmospherically Resistant Vegetation Index (ARVI), the Global Environment Monitoring Index (GEMI), the Enhanced Vegetation Index (EVI), the Green Chlorophyll Index ( $CI_{Green}$ ), the Red-edge Chlorophyll Index ( $CI_{Red-edge}$ ), the Weighted Difference Vegetation Index (WDVI) (Basso et al., 2004; Baret & Guyot, 1991; Citelson, 2013; Huete, 1988; Jiang et al., 2008; Rondeaux et al., 1996; Viña et al., 2011; Wiegand et al., 1991). Theoretically, many advantages can be attributed to some of these indices in comparison with the NDVI, which is commonly known for its limitations in dealing with soil background and atmospheric effects (Purevdorj et al., 1998; Rondeaux et al., 1996; Viña et al., 2011). However, NDVI is still one of the most widely adopted vegetation indices due to its simplicity of use and interpretation, thus its popularity in the literature, and the fact that it is readily available from most satellite and other remote sensing providers (Pettorelli et al., 2005; Van Leeuwen et al., 2006; Scheftic et al., 2014; Maestrini & Basso, 2018; Meng et al., 2013; Weiss et al., 2001).

Gao et al. (2020) have found in their review generally good agreement on the relation between CC and NDVI through linear or quadratic models, although they reported considerable variability among and within different vegetation

types. It has been well accepted that, aside from other factors, correlations between NDVI and CC vary among crop species (Gitelson, 2016), but it is unclear which are the standard correlations to be considered for different groups of crop types, how tight is the association between the two variables, and thus the degree of uncertainty in predicting CC from NDVI.

The identification of standard models, i.e. correlations with similar regression coefficients and NDVI saturation thresholds (Gutman and Ignatov, 1997), associated with similar light extinction coefficients (Campbell and Norman, 1998), would enable us to propose generic algorithms that would not require re-parameterization for some crop types, allowing the extension of existing NDVI-CC correlations into modelling applications for many other crop species (Gitelson, 2013).

One increasingly important application for the spatial analysis of cropping systems, is the estimation of CC via remote sensing to calibrate simulation models (Campos et al., 2019; Casa et al., 2015; Er-Raki et al., 2007; Jin et al., 2018a; Mohamed Sallah et al., 2019; Silvestro et al., 2017). In order to explore how correlations between NDVI and CC relate to specific crop species and groups of crop types, we conducted a meta-analysis in which we compiled information that correlated remote and proximal sensing NDVI with field observations of CC for different crop species and types. With this analysis, we also aimed to contribute to the improvement and standardization of CC model calibration with NDVI, by suggesting generic and robust correlations to be used for different agricultural crops, as an interesting alternative to in situ measurements of CC that can be costly and time consuming.

A meta-analysis has generally multiple applications in applied research. Within this particular study we highlight its value for exploring heterogeneity among crop species and types regarding NDVI-CC correlations, to identify general patterns in existing data and opportunities for future research (Krupnik et al., 2019; Stewart, 2010).

## 3.2 Materials & Methods

### 3.2.1 Meta-analysis

The meta-analysis was developed from a systematic review of all calibration studies published with the following keyword combinations on the title: "vegetation index + canopy cover", "vegetation indices + canopy cover", "NDVI + canopy cover", "NDVI + ground cover", "NDVI + coverage", "vegetation indices + cover", "NDVI + cover", and "remote sensing + groundcover". A total of 22 published articles (Calera et al., 2001; Carlson et al., 1994; Carlson and Ripley, 1997; de la Casa et al., 2018, 2014; Derrien et al., 1992; Er-Raki et al., 2007; Citelson, 2013; Citelson et al., 2002; Goodwin et al., 2018; Gutman and Ignatov, 1998; Imukova et al., 2015; Jasinski, 1990; Jiang et al., 2006; Jiménez-Muñoz et al., 2009; Johnson and Trout, 2012; Lukina et al., 1999; Prabhakara et al., 2015; Purevdorj et al., 1998; Todd and Hoffer, 1998; Trout et al., 2008; Verger et al., 2009) were selected according to the abstract, from which five were rejected due to lack of data on measured CC (Carlson and Ripley, 1997; Derrien et al., 1992; Gutman and Ignatov, 1998; Jasinski, 1990; Todd and Hoffer, 1998).

Additionally, we included in the analysis two unpublished databases collected by us as follows: 1) Data collected in 2012-15 ( $N= 33$ ) for multiple annual crops in the region of Córdoba, Spain; 2) Data collected in 2019-20 ( $N=16$ ) in a commercial wheat plot, located in the region of Córdoba, Spain. Following is a brief description of both databases.

The first unpublished database contains 33 observation units: winter wheat ( $N=6$ ), sunflower ( $N=5$ ), cotton ( $N=5$ ), maize ( $N=5$ ), basil ( $N=3$ ), common bean ( $N=2$ ), garlic ( $N=1$ ), watermelon ( $N=1$ ), asparagus ( $N=1$ ), sorghum ( $N=1$ ), onion ( $N=1$ ), chickpea ( $N=1$ ), and rosemary ( $N=1$ ). CC data were collected in spring-summer cloud free days (i.e., April-July) at plot level (2-5 ha), in visually homogeneous zones and excluding field border stripes with the same width of the satellite spatial resolution (30 m). The zonal average of satellite NDVI (Landsat-7) was estimated from 25-55 pixels, varying from case to case according to the area of observation. CC was measured with the GreenCrop Tracker software at different growth stages, at a height of 1 m above the canopy,

perpendicularly to the ground to minimize the significance of angular effects, with a digital camera (Nikon Coolpix S7000 16 MP) between 10:00–15:00 local time (Cihlar et al., 1987).

The second unpublished database has 16 observation units and was collected in a commercial wheat field at four different dates, from crop emergence to anthesis. The field was divided in four observation zones, according to a clustering analysis of historical NDVI patterns and field geomorphological properties. Each zone was approximately 2 ha, excluding border stripes with the same width of each satellite pixel (10 m). Border stripes were excluded to guarantee that all satellite pixels considered within each zone were entirely located within the same zone. At each zone, the average of approximately 200 pixels of NDVI was plotted against the average of 10 observation points of CC, following a random sampling scheme. CC ground measurements were taken with a digital camera (Canon EOS 550D + EFS 18-135 mm CMOS APS-C 18.7 MP) at 1.8 m height (at 10:00–12:00 local time) using an image processing package (Patrignani & Ochsner, 2015). All observations were conducted under clear sky conditions (0–2% cloud cover) and satellite data were atmospherically corrected using the Sen2Cor processor and Planet-DEM (i.e., Sentinel-2 Level-2A products).

Fiji Image-J software (<https://imagej.net/>) was used to extract the data from each document. The final database (N=1397) combined information on NDVI, measured CC, crop species, location of the study, NDVI source (i.e., satellite or in-situ measured with a digital camera or a spectroradiometer) and coefficient of determination ( $R^2$ ). Both published and unpublished sources of data were considered together because it is accepted that the inclusion of unpublished data minimizes the effects of publication bias in meta-analysis, while it maximizes the total sample size, thus enabling the extrapolation of results to a larger extent of case-studies (Krupnik et al., 2019).

All data were grouped according to crop species, crop type (i.e., cereals, grain legumes, grasslands, horticultural, industrial crops, and forage legumes), growing season (i.e., winter-spring, spring-summer, perennial crops) and NDVI source. Crop species with less than 10 observations entered the general

correlations but were excluded from the regression analyses of specific crop species due to excessive data skewness (i.e., asymmetry of the cumulative probability distribution). Crop types were defined according to taxonomic criteria (e.g. cereals, legumes), agronomic use (e.g. grain, forage) and canopy structure. Regarding canopy structure, crop species were clustered according to the mean values of the leaf angle distribution parameter (extracted from Table 15.1 in Campbell & Norman. (1998)). Regression analysis was used to estimate model coefficients, expressed linearly as following:

$$CC(\%) = a \cdot NDVI + b \quad (3.1)$$

where  $a$  represents the inverse of the difference between the NDVI value of an area of bare soil and the NDVI value of a pure vegetation pixel, and  $b$  corresponds to the negative fraction of the same bare soil NDVI divided by the difference of the previous two (Qi et al., 2000). Alternative models were also tested, including quadratic models through polynomial regression analysis (Gao et al., 2020), as well as logarithmic and exponential models, respectively expressed as:

$$CC(\%) = a \cdot NDVI^2 + b \cdot NDVI + c \quad (3.2)$$

$$CC(\%) = a \cdot \log(NDVI) + b \quad (3.3)$$

$$CC(\%) = a \cdot e^{b \cdot NDVI} \quad (3.4)$$

Least Squares Fitting (LSF) and statistical hypothesis testing were respectively used to estimate the regression coefficients and their significance level (stats package in R; Team, 2002). For each group correlation, the null hypothesis was tested with the non-parametric Mann–Whitney U test since samples were not normally distributed. Non-normality was checked with the Shapiro-Wilk test (Acutis et al., 2012). The NDVI saturation threshold was estimated by solving the best fitted regression with CC equal to 100% (Gutman & Ignatov, 1997). The performance of each model was evaluated by comparing

simulated sets of CC against measured CC. Root mean square error (RMSE) and  $R^2$  were used as statistical indicators of performance for model evaluation. The RMSE quantifies the weighted variations in error (residual) between the predicted and observed values of CC, while the  $R^2$  indicates the percentage of variance that is explained by each model.

### 3.2.2 A case study – testing NDVI-CC correlations in simulations with the AquaCrop model

The applicability of our results was tested using four independent sets of experimental data. The NDVI-CC model obtained for wheat was used to assess site-specific time-series of NDVI, used to estimate CC values, and to calibrate CC curves at multiple locations in simulations of the AquaCrop model (Steduto et al., 2009). The site-specific time-series of NDVI were defined through interpolation of discrete values, obtained from Sentinel-2A imagery, at dates with clear sky conditions and for each selected observation point ( $N=28$ ). Three different seasons of satellite data (i.e. 2015/16, 2017/18 and 2019/20) were used to assess four different '*field × year*' dataset combinations in two wheat fields grown in 2019/20 (i.e. trial A and B), one in 2017/18 (trial C), and one in 2015/16 (trial D). While trials B, C and D correspond to the same field at different seasons, trial A was located in a different field nearby (Appendix – Table A.3.1).

The two experimental fields are located in Cordoba, southern Spain ( $37.8^\circ$  N,  $4.8^\circ$  W, mean altitude 170 m). Field one (trial A) and field two (trials B, C and D) are 42 and 36 ha, respectively. The soils are of clay texture with high bulk density (1.65–1.88 g cm<sup>3</sup>) and of 1.2–1.6 m depth. The spatial variation of soil properties was characterized in field one by using an electromagnetic induction sensor (DUALEM-21S) to measure soil EC<sub>a</sub> (dS/m) at 35 and 85 cm depth. Soil samples (%Clay, %Sand, pH) were collected at 35 cm depth following a multistage sampling scheme that was based on two different clusters of superficial EC<sub>a</sub>. In field two, soil properties were averaged for the entire field according to on-farm records.

The wheat cultivars KIKO-NICK-R1, Anthoris-R1, KIKO-NICK-R1 and Amilcar were sown in trial A, B, C and D, respectively. Seeding rates were 200 ( $\pm 20$ ) kg/ha. Trial A was fertilized with two applications of Calcium Ammonium Nitrate plus Sulphur (230 + 180 kg/ha). Trials B, C and D were respectively fertilized with 165, 180 and 190 kg N/ha (Ubesol + Urea), and 60 kg P/ha (Diammonium phosphate).

Ground measurements of CC were conducted in trial A, every 15-20 days, from sowing to harvest. CC was measured using a digital camera (Canon EOS 550D + EFS 18-135 mm CMOS APS-C 18.7 MP) at 1.8 m height and an image processing package (Patrignani and Ochsner, 2015). Yield was spatially assessed and it was determined with the 'New Holland' Precision Land Manager (PLM) software, taking as an input the shapefiles generated by the combine harvester monitor (Fendt PLI C 5275). Yield values were calculated with a spatial resolution of 100 m<sup>2</sup>, following the equation of Reitz & Kutzbach (1996). An evaluation of the estimated yield data from the combine monitor was performed by comparison with manual samples taken at each point in trial A (sampled areas of 0.9 m<sup>2</sup>). Spatial yield data was computed with R-studio (Lovelace et al., 2019).

The AquaCrop model was parameterized with data collected at 28 different observation points, 10 corresponding to trial A and six to each of the remaining trials (Appendix - Table A.3.1). Soil parameterization was divided into two different soil types for trial A, while only one soil type was considered for the other trials (Appendix - Table A.3.1). Soil hydraulic parameters were estimated with the 'USDA-rosetta program' (Schaap et al., 2001). Daily weather data were obtained from a weather station nearby. Crop parameterization included sowing and harvesting dates, mean emergence date, seeding rate, site-specific plant density, root growth rate and crop stages duration (Appendix - Table A.3.1). Crop stages duration were obtained from field observations of phenological development and adjusted according to the CC curves obtained from satellite NDVI, i.e. the vegetative stage duration and the beginning date of crop senescence, both expressed in calendar terms (Figure 3.1). Both the fitted maximum canopy cover (CC<sub>MAX</sub>) and the corresponding date, when crop

reaches maximum CC, were estimated by solving the first derivative of the fitted CC curve. More information regarding the parametrization procedures of the AquaCrop model may be found in Steduto et al. (2012). Our methodological approach is synthesized in Figure 3.1.

Simulated yield values were plotted against point-based observations, estimated from the corresponding yield maps. RMSE of simulated yield was calculated for each field and the coefficient of variation (CV) of simulated yields of different sites was assessed and compared with the CV of observed yields. The CC values, estimated from NDVI, were compared with the fitted curve values used in the model parameterization. Ground measures of CC, taken in field A, were used for validation of CC values estimated from NDVI. The Willmott index of agreement ( $d$ ), the Pearson correlation coefficient ( $r$ ) and the RMSE (expressed in % of CC) were estimated and used for this evaluation.

### 3.3 Results & Discussion

#### 3.3.1 General results of the meta-analysis

The meta-analysis yielded a total of 1397 data points (Table 3.1 and Figure 3.2). Within a total of 26 different crop species recorded, 13 had more than 10 observations ( $N>10$ ) and followed a log-concave cumulative probability function of NDVI (Figure 3.3). For these 13 crop species, the NDVI values followed a bimodal distribution and the obtained correlations had a normal distribution of residuals according to the corresponding histogram for each regression.

Within the total universe of data points collected ( $N=1397$ ), 34 were classified as 'unknown' crop species due to lack of information regarding the species in the publications. These points (2.5% of total sample size) were used for the general correlations but discarded for the specific groups regressions. A general standard model ( $N=1397$ ) was established. The best fitted model follows a linear structure (Table 3.2), in which the effects of soil background and crop traits are highly simplified (Gao et al., 2020; Gutman and Ignatov, 1997).

In this meta-analysis we found that the satellite NDVI ( $N=524$ ) tends

to overestimate CC for low levels of NDVI (Figure 3.3-C and Table 3.2) in comparison to in-situ proximal sensing ( $N=873$ ). This is justified by the noise effect that disturbs the reflected signals of low vegetated surfaces (Todd & Hoffer, 1998). This effect is more relevant when using satellite data because the spatial resolution of the sensors is commonly larger than the scale of individual vegetation objects, which enhances the noise effect of soil background. The lower  $R^2$  of the satellite generic model may also be explained by a larger number of crop species being considered (Table 3.1). However, when considering only species with more than 10 observations, no differences regarding the species composition of these two groups (i.e. satellite and in-situ) were observed (results not shown).

The cumulative probabilities of NDVI for each crop type followed logarithmic shapes. All best fitted models were either linear or quadratic (Table 3.2). By contrast, the  $R^2$  values of logarithmic and exponential models were considerably lower than those for linear and quadratic regressions, and RMSE (of logarithmic and exponential models) were on average 10–25% higher (results not shown).

For most crop types, within a NDVI range of 0.25–0.75, the regressions did not differ considerably from each other (Figure 3.3). Below and above those limits, soil background noise and NDVI saturation effects make it difficult to estimate CC accurately (Carlson and Ripley, 1997; Prabhakara et al., 2015; Xue & Su, 2017). However, despite the general linear trend, an exceptional behaviour is observed for the case of grain legumes (i.e., soybean) because the best fitted correlation was quadratic (Table 3.2). For this specific case the quadratic model was a better alternative to the linear one (de la Casa et al., 2018; Gitelson et al., 2013). By contrast, in cereals and other field crops (i.e., horticultural and industrial crops), the linear model seems to be adequate (Table 3.2). However, we must highlight that, independently on the best fitted model, most models have considerable RMSE (Table 3.2), suggesting that for applications that require high accuracy, local calibration must be conducted.

According to the proportion of total data variation, the obtained  $R^2$  values indicate a fair fitness of observed data in the models proposed (Figure 3.4 and Table 3.2). The correlations were satisfactory as  $R^2$  values were above

70–75% for most cases and coefficient estimates were significant for most models (Table 3.2), but sources of systematic error need to be identified to increase the accuracy of NDVI-CC correlations. The observed trend of RMSE increasing (and  $R^2$  decreasing) with sampling size reveals the limitations to extrapolate these models to different environments than those in calibration studies (Figure 3.5). A similar effect is observed when increasing the number of databases considered in each model, which might be associated to non-sampling errors or sources of systematic error among studies (Poate & Daplyn, 1993). The uncertainty of each individual observation was not captured in our analysis because it was not possible to identify a common and objective indicator of uncertainty, which would be independent on sampling size and directly comparable among all observations considered. Nevertheless, we acknowledge that the variability of the random disturbance of observations did not differ greatly across groups of input data as the residuals plotted against fitted values were randomly distributed within comparable ranges, and the 'Cook's distance', which was plotted for each group of input data, did not show influential outliers (results not shown).

Saturation effects of NDVI were only observed for grasslands, broccoli, lettuce and sunflower (Table 3.2), which is likely due to a stronger asymptotic behavior of CC, caused by more horizontal leaf angle distribution (Table 3.1) and higher canopy expansion rate, both typical of these crops (Campbell & Norman, 1998; Johnson & Trout, 2012). The slight saturation effect that is observed in these crops might also be associated with lower variability of the red reflectance. At advanced phenological stages, the NDVI can become insensitive to the variation in red reflectance, when NIR reflectance surpasses largely the reflectance of red wavelength (Gitelson, 2016). For these crops, the NDVI saturated above 0.9, which is equivalent to a Leaf Area Index (LAI) between three and four, depending on the crop species (Bouman et al., 1992). Under optimal conditions, the LAI of upright canopy crops (e.g. cereals) can easily reach a value of six, with substantial mutual shading. Therefore, saturation effects are likely to be observed in the NDVI-LAI relations, but the same may not apply to CC where these relations appear to be mostly linear (Table 3.2).

and Figure 3.4). There are slight saturation effects in the NDVI-CC of some cases due to a higher leaf angle distribution parameter (Table 3.1 and 3.2). Our results suggest that it may be more interesting to correlate NDVI with CC than with LAI, because, in relative terms, the exploitable range between both variables is larger in the case of CC.

Despite the overall satisfactory goodness-of-fit, the coefficients of determination varied among crop types and species, as well as the RMSE (Figure 3.5). Different models can be proposed for several crops species and groups of crop types, but under levels of uncertainty that range from 6% to 18% (Table 3.2 and Figure 3.4). However, we recognize that the highest accuracy does not necessarily imply the best option. Existing trade-offs between temporal and spatial resolution of input data must also be considered (Lobell, 2013). While remote sensing NDVI overestimates CC, mostly in situations of lower CC (Figure 3.3-C), the higher temporal resolution of NDVI satellite data and the lower cost associated to its access and use are offsetting reasons to support its use.

### 3.3.2 Applications of NDVI-CC correlations

The capacity to model agronomic mechanisms causing spatial variation in yield, with practical implications for site-specific management remains a challenge that is far from being resolved (Lamb et al., 1997; Leroux et al., 2017, 2018). However, with the widespread advances in yield monitoring, the suitable equilibrium between temporal and spatial resolution of freely available remote sensing NDVI data, and using Table 3.2 regressions, it may be possible to estimate CC with a fine-resolution for various crop types, and use it for management applications in precision agriculture.

We believe that the use of NDVI-CC generic correlations will contribute to the standardization of the main approaches taken in the assimilation of CC into crop modelling applications (e.g. Jin et al., 2020). For these specific cases, the use of present regressions will have different implications in the various applications of crop simulation models. CC values used to update crop

simulation models will introduce inevitably uncertainty, however, even with the observed RMSE levels, the use of input canopy data is likely to result in better yield predictions than without any data assimilation of this kind (Doraiswamy et al., 2003; Huang et al., 2015). The average error observed will propagate in different ways towards the estimates, not only depending on what is being simulated but also on the range of variation for both CC and final estimates (Guo et al., 2019; Jin et al., 2018b). However, we believe that for many practical applications at field level, the observed RMSE are acceptable, as the temporal integration of high frequency canopy data, interpolated under similar levels of input error, have resulted in tolerable RMSE yield estimates for decision making (Dente et al., 2008; Waldner et al., 2019).

In crop simulation models that integrate multiple processes (e.g. Steduto et al., 2009), CC dependent parameters have a low first order effect on yield because simulated yield is mostly regulated by second order effects and interactions among multiple parameters through different simulated processes (Silvestro et al., 2017). Therefore, the assimilation of Table 3.2 regressions does not imply an error propagation towards the final estimates in the same order of magnitude due to model plasticity (i.e. the aptitude of a model to vary the sensitivity to input parameters under variable application conditions). Silvestro et al. (2017) conducted a global sensitivity analysis of wheat yield simulated with the model AquaCrop, and for a variation of CC-dependent parameters within a range of  $\pm 33\%$ , the final estimates showed an acceptable sensitivity index (i.e. 0.1–0.7 ton/ha). Considering the  $R^2$  and RMSE values of Table 3.2 models, we hypothesize that the final yield estimates will have less uncertainty than that reported by Silvestro et al. (2017). This is mostly valid for the simulation of potential yield because, under crop stress conditions, larger uncertainty in the final estimates is expected for the same level of input error (Guo et al., 2019; Vanuytrecht et al., 2014).

Additional potential is seen in the use of the NDVI-CC correlations herein for assessment of spatial variations, where absolute values are less critical than spatial-temporal relative variations, such as site-specific emergence dates derived from zonal CC, spatial variation of vegetative growth rates, the

estimation of different dates of CC peak in different management zones, the relative variation of starting dates for crop senescence within the same field. These are also examples of CC dependent parameters that have minimal first order effects on the final yield estimates (Vanuytrecht et al., 2014).

The proposed models may be considered adequate for many applications, but there is an effect of irregular sampling sizes that must be considered (Figure 3.5), as well as sources of systematic error that affect extrapolation of Table 3.2 models. Our analysis shows that CC may be estimated from NDVI albeit at different levels of accuracy, depending on the crop. The NDVI-CC relations for some crop species are supported by enough data availability, while in other crop types more data is needed (e.g., industrial and horticultural crops). Even though more attention has been devoted to cereals than other crop types (Table 3.1), there is insufficient data for important cereals such as rice, oat or sorghum. The same applies to several other horticultural and field crops such as potato, onion and sugar beet.

### 3.3.3 Application to the simulation of wheat yields

The spatial-temporal assimilation of CC data into the AquaCrop model was performed in a case study with wheat, eight to 10 satellite images per season were used to obtain NDVI at each observation site (Supplementary material). Daily values of CC were estimated with the wheat NDVI-CC model developed here. The ground measurements of CC ( $N=12$ ), collected at each observation site in trial A ( $N=10$ ), indicated a suitable goodness-of-fit as the estimated CC curves showed a mean RMSE of 12.03 % ( $\pm 2.9$ ) which is line with the RMSE values of the meta-analysis results (Table 3.2). Both the Pearson correlation coefficient ( $r$ ) and the Willmott's index of agreement ( $d$ ) were bar-plotted, separately for each observation site (Figure 3.6-A). These two correlation indices indicated a good level of model performance, with mean values of 0.92 and 0.93 ( $\pm 0.03$ ).

The AquaCrop model, calibrated with CC curves that were adjusted with the results of our meta-analysis (Figure 3.6-C and -D), was capable to simulate

accurately crop yield (Figure 3.6-B). Our modelling captured a fair fraction of the overall variation of observations (Table 3.3), both in terms of space and time (i.e. two different fields at the same year vs. the same field at multiple years).

The RMSE of our yield estimates (Table 3.3) was within the acceptable sensitivity index range reported by Silvestro et al. (2017). In relative terms, we observed that our modelling approach was capable to explain most of the total observed variation. Our simulated yields showed, for each '*field × year*' combination, CV values ranging from 52 to 87% of the observed CV (Table 3.3). The combine harvest data was also well correlated with the manual sampling records (RMSE = 0.327 Mg/ha, Appendix – Table A.3.2). This indicates that the error magnitude of simulated yields using the NDVI-CC relationship may be acceptable in many practical applications. As measured, both the simulated yield and the combine harvest data RMSE's were within comparable ranges (0.327–0.504 Mg/ha), which indicates that assimilating NDVI-CC into simulation modelling results in a similar level of uncertainty with yield mapping from combine harvest data, used in precision agriculture.

It must be highlighted that AquaCrop simulated yields correspond to water-limited yields. It is known that these can deviate from reality due to the model point-based structure (Tenreiro et al., 2020). We believe that the inclusion of spatial variations in water availability is likely to close the gap between simulated and observed yield CV. In this case-study, the use of the NDVI-CC empirical model, to calibrate a crop simulation model, provided simulation results that were very much in line with measured observations, suggesting that our meta-analysis results will contribute to new advances in modelling research applications.

### 3.3.4 Closing remarks

We evaluated the robustness of several NDVI-CC relations as proposed by Dorigo et al. (2007), who recommended to test the predictive capacity of statistical relationships, within this context, over independent data sets, including other crops and observations at both different locations and

phenological stages. Our analysis provides substantial empirical evidence to support the use of linear and quadratic models to estimate CC from NDVI (Carlson and Ripley, 1997; Qi et al., 2000). It also contributes new findings to recently published ones (e.g., Gao et al., 2020). Gao et al. (2020) addressed the NDVI-CC correlation from a mechanistic perspective, focusing on theoretical considerations of linear and quadratic models, and on reviewing relative vegetation abundance algorithms and correction methods of NDVI estimations. Our study approaches these models empirically by compiling multiple experimental data and proposes generic and ‘easy-to-use’ correlations, which are computationally undemanding (Dorigo et al., 2007), and enable canopy data assimilation into modelling applications for many crop species.

While our study has only considered the relationships between NDVI and CC, we recognize that the use of other vegetation indices may be a viable alternative, more precise in some cases. However, the NDVI was the only index that is common across all the studies used in this analysis and the use of alternative indices would require access to the raw reflectance data, which would not be possible for many studies. The recent advances in very-high-resolution (VHR) remote sensing thanks to the advent of unmanned aerial vehicles (UAVs) and other airborne platforms, combined with novel computer vision techniques (e.g. machine learning image recognition, spectral analysis, object-based classification), can improve the study of canopy attributes such as CC (Chianucci et al., 2016). However, the broader adoption of these technologies may be limited in countries where the use of UAVs is banned or strictly regulated. Moreover, the area throughput of airborne platforms is still very limited, particularly when compared with satellite imagery. In addition, estimating CC from historical satellite data through computationally undemanding models, such as the ones presented here, extends the time window of many of the applications discussed here.

### 3.4 Conclusion

Several practical advantages of using NDVI for the assessment of CC were identified and discussed. Our results were also experimentally tested, providing a quantitative evidence that our models can be used in multiple applications within modelling research. We concluded that, despite the overall uncertainty of the models presented, our results can be adopted with fair confidence in modelling applications, mostly in cases where the relative variations of predictions are prioritized over the absolute accuracy level. Examples of these are the spatial assessment of the vegetative state of a crop as well as the use of simulation models to deal with relative variations within the context of precision agriculture. We believe that the empirical models presented here will contribute to the use of NDVI for determining CC, thus improving crop growth estimates in experimental and modelling approaches that will assist in decision-making in agricultural systems.

## FIGURES - Chapter 3

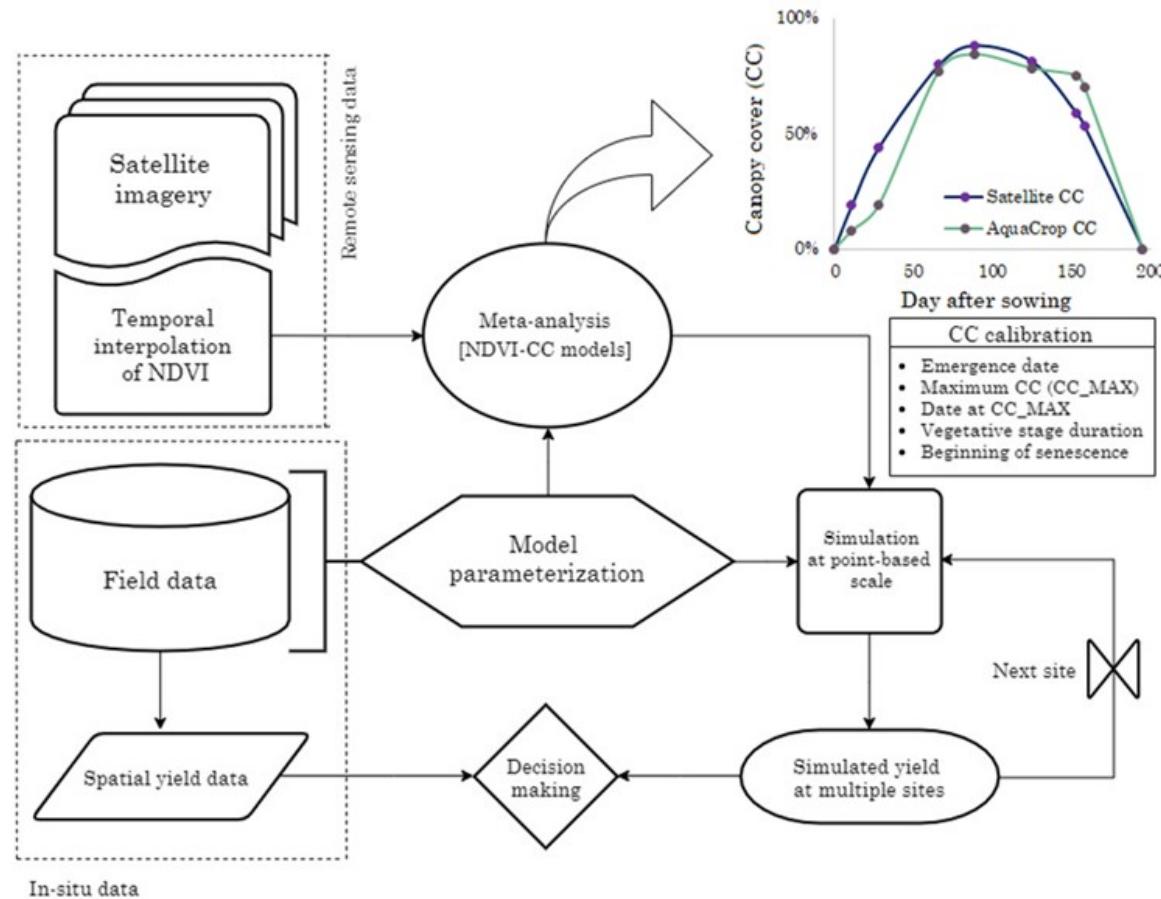
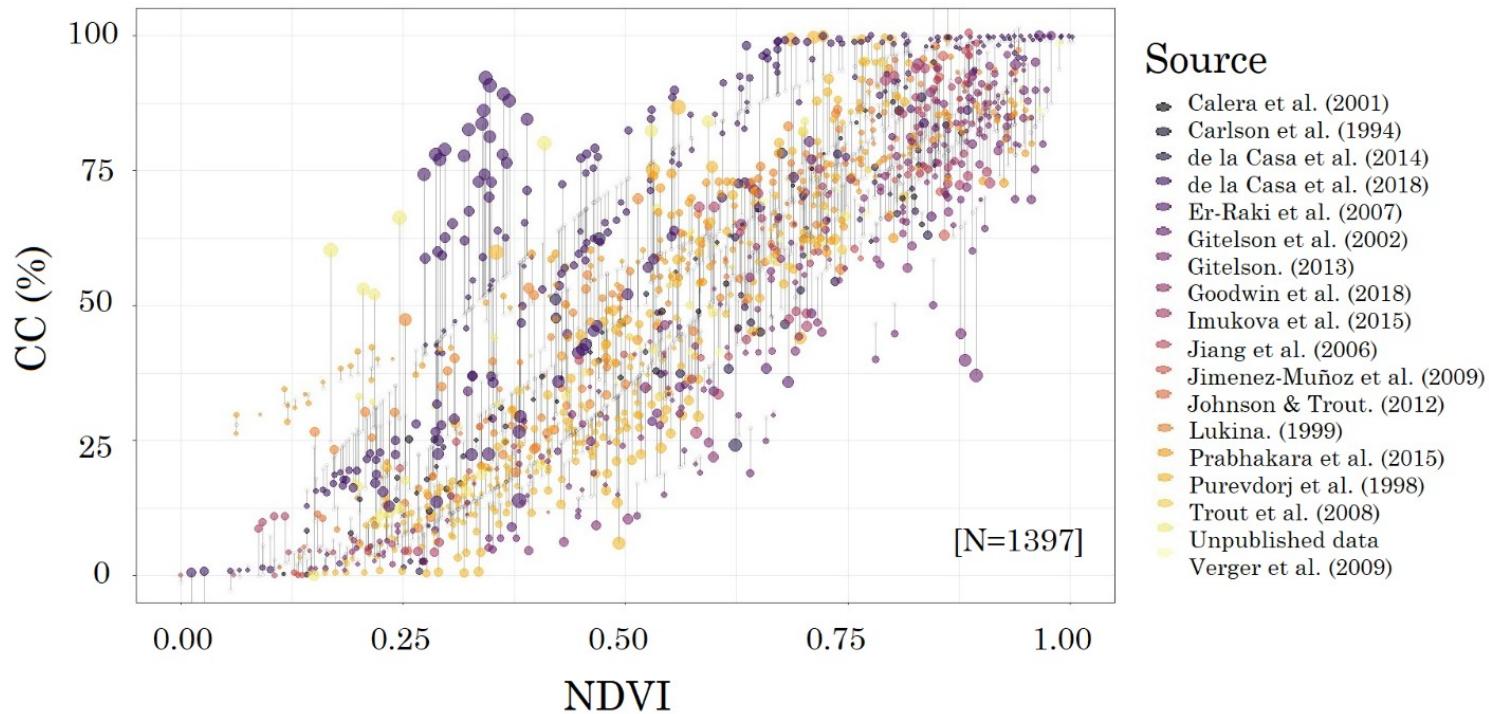
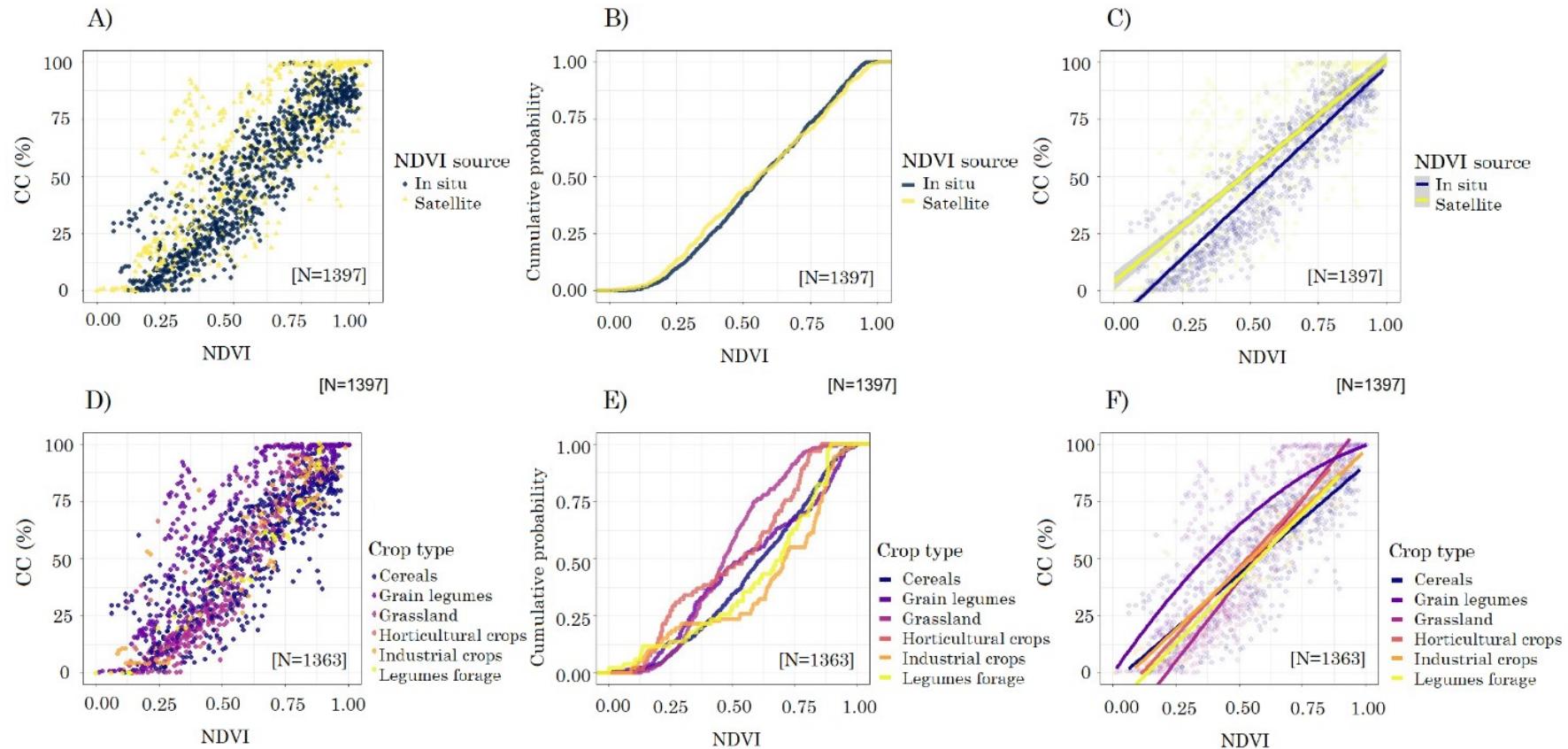


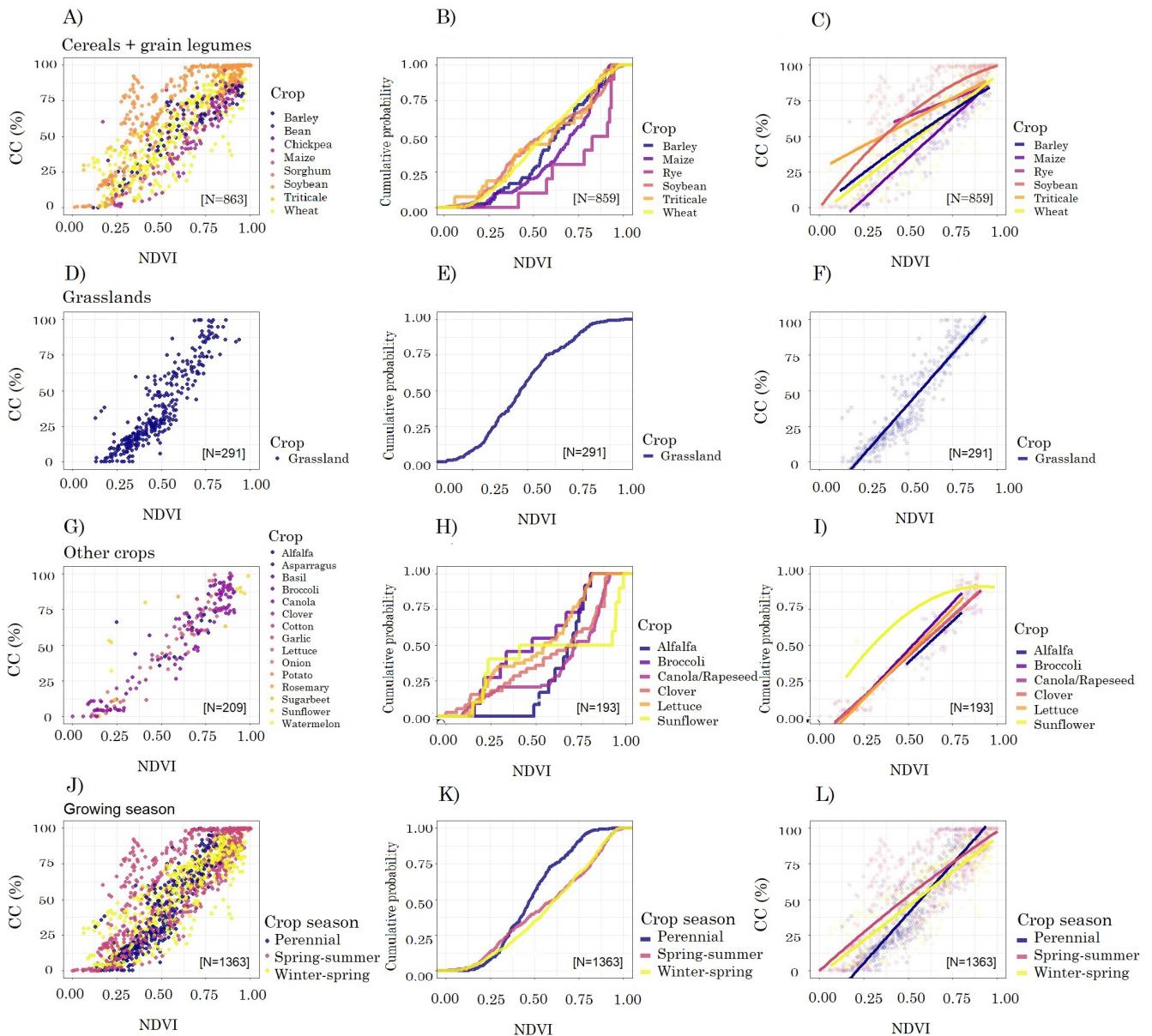
Figure 3.1: Our case study methodological scheme. Simulations at point based scale were conducted for 28 different sites at three different seasons.



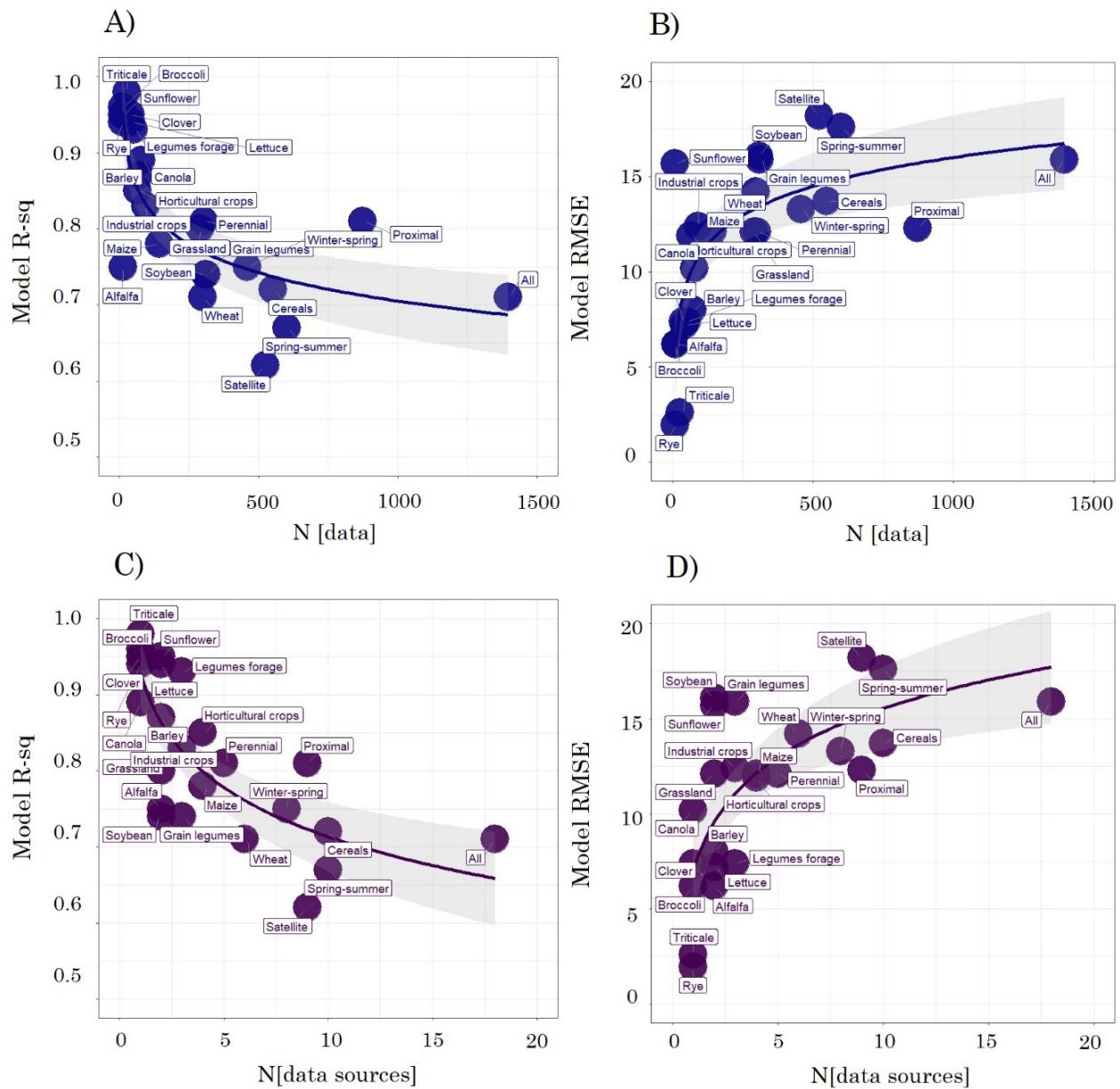
**Figure 3.2:** NDVI plotted against CC for all datasets. The dots are colored according to the data source. Grey vertical lines represent the residuals (error) for each observation and points are sized according to the corresponding error level, the bigger the point the larger the error.



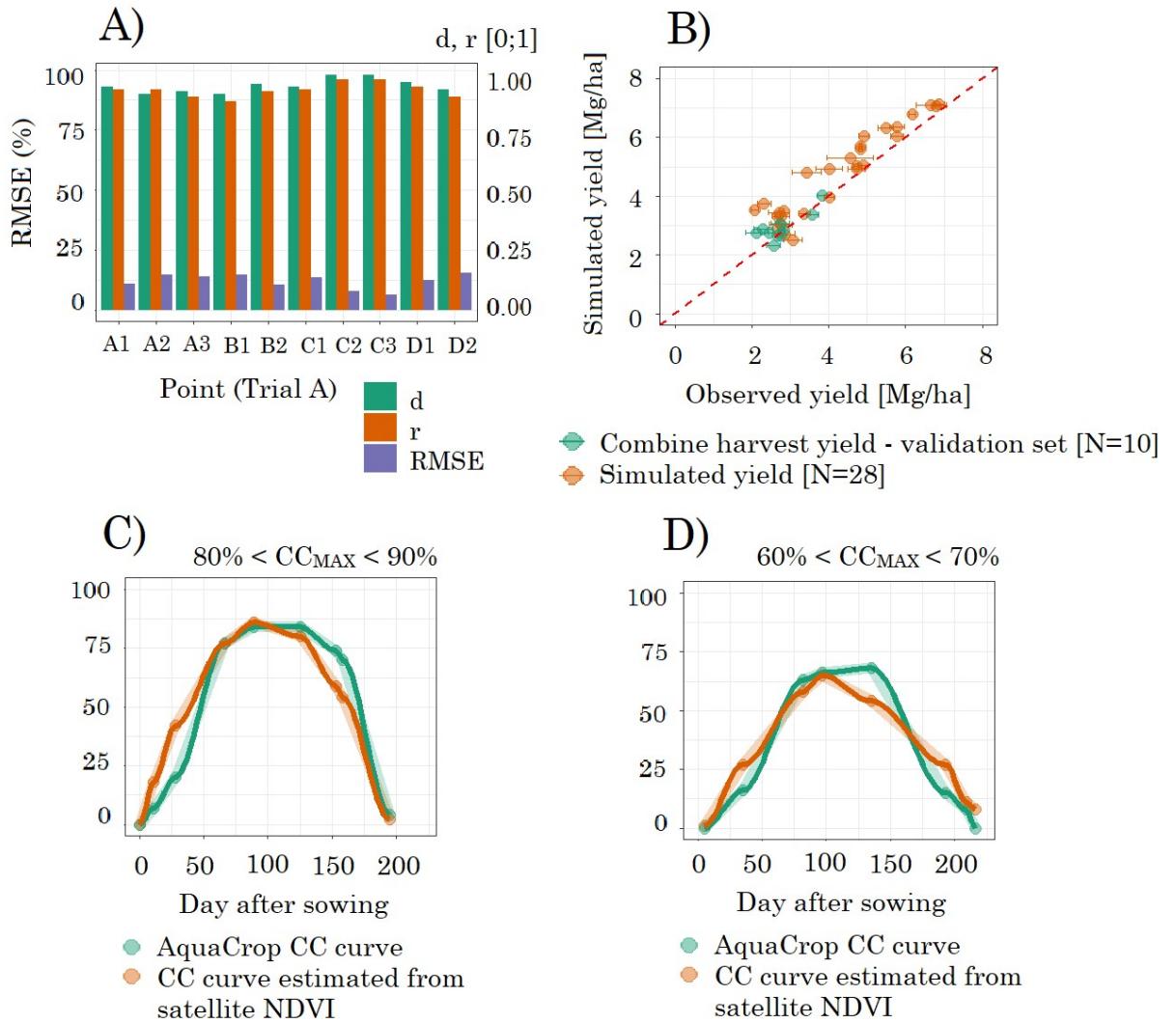
**Figure 3.3:** NDVI data plotted against CC, dots colored according to NDVI source (A) or crop type (D). Cumulative probability curves (B and E) and best fitted regression lines (C and F) for each group of data.



**Figure 3.4:** NDVI data plotted against CC (%) (A, D, G and J), cumulative probability curves (B, E, H and K) and smooth regression lines (C, F, I, L) for each group of data. Dots and lines are colored according to crop species (A-I) or growing season (J-L). Crop species with  $N < 10$  were excluded from regressions.



**Figure 3.5:** Statistical indicators of model performance:  $R^2$  and RMSE plotted against sampling size (A and B, respectively) and against the number of data sources used for each model (C and D, respectively). Both indicators followed a logarithmic response curve, negative for the case of  $R^2$  and positive for RMSE. All regression coefficients were significant and  $R^2$  values ranged around 0.6 (results not shown).



**Figure 3.6:** A) Evaluation of simulated CC in comparison with the ground measurements taken in Trial A. The bars represent the RMSE, the Pearson correlation coefficient ( $r$ ) and the Willmott's index of agreement ( $d$ ) of estimated CC at each observation point; B) Simulated vs. observed yield (Mg/ha). Green dots correspond to the estimated yield values obtained from yield mapping (i.e. according to the equation of Reitz & Kutzbach. (1996)) plotted against manual samples taken in Trial A, orange dots represent (AquaCrop) simulated vs. observed yields, units are expressed in Mg/ha; The satellite NDVI estimated CC time-series, plotted in calendar days, and the CC curve assimilated into the AquaCrop model: C) Point B2 in Trial B and D) Point C1 in Trial D..

## TABLES - Chapter 3

**Table 3.1:** Sampling groups characterization: N [data sources], N [data points] and N [crop species] refer respectively to the number of data sources, data points and crop species included in each sampling group. The canopy structure was expressed by the leaf angle distribution parameter ( $\chi$ ), obtained from Campbell & Norman. (1998).

Model	N [data sources]	N [data points]	N [Crop species]	Canopy structure ( $\chi$ )
General	18	1397	26	-
Satellite	9	524	22	-
In-situ	9	873	10	-
Cereals	10	551	6	0.9-1.65
Grain legumes	3	312	3	<0.85
Grassland	2	291	a.	0.7-2.5
Horticultural crops	4	65	10	1.5-1.9
Industrial crops	3	93	3	2-3
Legumes forage	3	51	2	$\simeq$ 2.5
Winter-spring crops	8	459	6	-
Spring-summer crops	10	601	18	-
Perennial crops	5	303	2	-

*Crop types species (N>10):* 1) Cereals (Barley, Maize, Triticale, Rye, Wheat), 2) Grain legumes (Soybean), 3) Grassland (a. non-specified), 4) Horticultural crops (Broccoli, Lettuce), 5) Industrial crops (Canola, Sunflower), 6) Legumes forage (Alfalfa, Clover). The table values also consider crop species with less than 10 observations.

**Table 3.2:** Model coefficient estimates of each group for linear and quadratic regressions. The best fitted regressions are highlighted in bold, which were selected as those maximizing  $R^2$  while keeping a minimal amount of significant coefficients. Root mean square error (RMSE) was estimated for the best fitted regression of each group. Significance codes: '\*\*\*' 0.1% '\*\*' 1% '\*' 5%.

Model	Linear				Quadratic				RMSE	NDVI saturation threshold
	a	b	$R^2$	a	b	c	$R^2$			
General [N=1397]	105.427***	-6.501***	0.71	4.257	100.719***	-5.439*	0.71	15.9	1.01	
Satellite [N=524]	97.088***	4.106*	0.62	-44.549**	146.135***	-6.606	0.63	18.2	0.97	
In-situ [N=873]	112.023***	-13.813***	0.81	27.786***	81.174***	-6.705**	0.81	12.3	1.01	
Cereals [N=551]	95.241***	-4.118*	0.72	27.916*	63.478***	3.445	0.72	13.7	1.09	
Grain legumes [N=312]	97.268***	11.079***	0.69	-113.005***	225.360***	-17.147***	0.74	15.9	0.99	
Grassland [N=291]	141.287***	-30.004***	0.80	107.530***	32.505	-5.786	0.82	12.1	0.92	
Horticultural crops [N=65]	109.663***	-13.552***	0.85	29.041	91.741*	-8.496	0.84	11.9	1.04	
Industrial crops [N=93]	104.355***	-6.822*	0.83	0.659	103.659***	-6.694	0.83	12.4	1.02	
Legumes forage [N=51]	113.290***	-15.240***	0.93	77.263***	35.651*	-1.626	0.95	7.4	1.02	
Barley [N=72]	93.243***	-1.58	0.82	<b>-84.847***</b>	<b>193.324***</b>	<b>-27.172***</b>	<b>0.87</b>	8.0	1.13	
Maize [N=143]	113.956***	-22.268***	0.78	55.121*	49.001	-5.697	0.78	12.1	1.07	
Soybean [N=309]	97.003***	11.290***	0.69	<b>-113.388***</b>	<b>225.74***</b>	<b>-17.189***</b>	<b>0.74</b>	16.1	0.99	
Triticale [N=27]	66.410***	26.722***	0.98	-20.198*	87.413***	22.736***	0.98	2.6	1.10	
Rye [N=10]	48.520***	39.733***	0.94	-7.791	59.550	36.128*	0.93	1.95	1.24	
Wheat [N=298]	97.368***	-4.492*	0.71	36.124*	56.9**	4.94	0.72	14.2	1.07	
Alfalfa [N=12]	115.79***	-20.07	0.75	-145.82	304.17	-79.42	0.73	6.2	1.04	
Broccoli [N=11]	132.291***	-19.550**	0.96	124.463*	12.387	2.216	0.96	6.2	0.90	
Canola [N=78]	112.335***	-13.869***	0.89	41.193	70.062**	-6.601	0.89	10.2	1.01	
Clover [N=39]	114.154***	-14.786***	0.95	81.896***	31.767*	-1.270	0.96	7.4	1.01	
Lettuce [N=43]	126.706***	-19.464***	0.94	81.638*	50.287	-6.304	0.94	7.2	0.94	
Sunflower [N=10]	77.37**	1.58	0.69	<b>-108.61*</b>	<b>198.96**</b>	<b>0.011*</b>	<b>0.95</b>	15.7	0.91	
Winter-spring crops [N=459]	96.337***	-2.734	0.75	14.657	79.952***	1.002	0.75	13.3	1.07	
Spring-summer crops [N=601]	98.762***	1.56	0.67	<b>-36.585**</b>	<b>139.268***</b>	<b>-7.2*</b>	<b>0.67</b>	17.6	1.07	
Perennial crops [N=303]	139.314***	-29.320***	0.81	100.603***	37.054	-6.44***	0.81	12.1	0.92	

**Table 3.3:** Simulation error assessment. Simulated yield RMSE and Coefficient of Variation (CV). '% Total yield variation' corresponds to the fraction of total observed variation that was captured by the assimilation of CC into modelling simulations.

Trial	Year	Yield.RMSE (Mg/ha)	CV (Simulated yield)	CV (Observed yield)	% Total yield variation
A	2019/20	0.013	14.0%	16.0%	87.4%
B	2019/20	0.837	9.7%	18.1%	53.9%
C	2017/18	0.662	6.7%	13.0%	52.0%
D	2015/16	0.782	17.2%	31.0%	55.3%
Mean	-	0.504	10.2%	15.7%	64.4%

## APPENDIX - Chapter 3

**Table A.3.1.** Experimental data sets used for parameterization of the AquaCrop model. Bulk density (BD) mean values were used, in combination with clay and sand contents, to estimate hydraulic parameters for each soil type (Schaap et al., 2001). The standard deviations are indicated between brackets. The hydraulic conductivity values used for the soil paramterization were the mean values reported in this table ( $K_{SAT}$  mean, expressed in  $\text{mm day}^{-1}$ ). One single soil horizon was considered (140 cm depth). The initial curve number was set at a value of 78 (i.e. hydrologic group D). More information regarding the water balance approach that is followed by AquaCrop found in Tenreiro et al. (2020). Mean sowing rate was set at 200 kg/ha for all trials.

Data		Field 1 - Trial A (2019/20)										Field 2 - Trial B (2019/20)						Field 2 - Trial C (2017/18)						Field 2 - Trial D (2015/16)																				
Parameter	Units	A1	A2	A3	B1	B2	C1	C2	C3	D1	D2	A1	A2	B1	B2	C1	C2	A1	A2	B1	B2	C1	C2	A1	A2	B1	B2	C1	C2															
ECa	$\text{dS m}^{-1}$	0.45	0.55	0.45	0.20	0.20	0.20	0.20	0.20	0.35	0.40	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-															
Clay	%	50					38			50					44				44							44																		
Sand	%	15					22			15					22				22							22																		
Texture	USDA class	Clay					Clay-loam			Clay					Clay				Clay							Clay																		
BD	$\text{g cm}^{-3}$	1.76	1.68	1.78	1.88	1.78	1.80	1.81	1.78	1.81	1.82					-				-																								
$BD_{mean}$	$\text{g cm}^{-3}$	1.77 (0.05)			1.81 (0.04)			1.77 (0.05)			1.66 (0.05)			1.66 (0.05)			1.66 (0.05)			1.66 (0.05)			1.66 (0.05)			1.66 (0.05)																		
$K_{SAT}$ range	$\text{mm day}^{-1}$	4.8-50			6.2-65			4.8-50			5.5-60			5.5-60			5.5-60			5.5-60			5.5-60			5.5-60																		
$K_{SAT}$ mean	$\text{mm day}^{-1}$																																											
$\theta_{PWP}$	%	26			22			26			18			18			18			18			18			18																		
$\theta_{FC}$	%	39			35			39			35			35			35			35			35			35																		
$\theta_{SAT}$	%	41			40			41			40			40			40			40			40			40																		
Sowing date	date	13-Dec										18-Nov						24-Nov						10-Nov																				
Crop emergence	DAS	9										8						10						10																				
Plant density	plants $\text{m}^{-2}$	150	150	250	250	300	300	150	300	300	300	360	360	300	300	300	300	360	360	360	300	300	300	300	300	300	300	300	300	300														
$CC_{MAX}$	%	75	77	80	75	66	84	82	70	84	85	87	86	86	86	83	84	86	84	85	81	78	81	75	74	72	72	68	71															
$Root_{growth}$	$\text{cm day}^{-1}$	0.7			0.8			0.7			0.8			0.8			0.8			0.8			0.8			0.8																		
Vegetative stage	days	120										105						135						105																				
Anthesis duration	days	10										14						18						16																				
Reproductive stage	days	58										84						80						85																				
Senescence duration	days	20										35						38						40																				
Harvest date	date	9-Jun										13-Jun						27-Jun						19-Jun																				

**Table A.3.2.** Yield simulation outcomes at each observation point [N=28]. Manual sampling yield data are provided for Trial A [N=10]. Standard deviations are indicated between brackets.

Field	Season	Trial	Point	Observed yield ( $\text{Mg ha}^{-1}$ )	Simulated yield ( $\text{Mg ha}^{-1}$ )	Manual sampled yield ( $\text{Mg ha}^{-1}$ )
1	2019/20	A	A1	2.75 (0.30)	3.02	2.12
1	2019/20	A	A2	3.06 (0.23)	2.50	2.73
1	2019/20	A	A3	2.86 (0.26)	2.70	2.30
1	2019/20	A	B1	3.36 (0.15)	3.40	3.58
1	2019/20	A	B2	2.66 (0.01)	2.90	2.67
1	2019/20	A	C1	4.01 (0.13)	3.95	3.83
1	2019/20	A	C2	2.75 (0.22)	3.30	2.44
1	2019/20	A	C3	2.66 (0.06)	3.30	2.75
1	2019/20	A	D1	2.83 (0.01)	3.49	2.84
1	2019/20	A	D2	2.31 (0.18)	3.73	2.56
2	2019/20	B	A1	5.78 (0.20)	6.35	-
2	2019/20	B	A2	5.77 (0.16)	6.03	-
2	2019/20	B	B1	4.83 (0.05)	5.59	-
2	2019/20	B	B2	4.82 (0.03)	5.71	-
2	2019/20	B	C1	3.41 (0.38)	4.79	-
2	2019/20	B	C2	4.56 (0.61)	5.30	-
2	2017/18	C	A1	6.88 (0.13)	7.12	-
2	2017/18	C	A2	6.79 (0.09)	7.05	-
2	2017/18	C	B1	6.66 (0.40)	7.08	-
2	2017/18	C	B2	6.17 (0.08)	6.77	-
2	2017/18	C	C1	4.91 (0.13)	6.04	-
2	2017/18	C	C2	5.48 (0.20)	6.31	-
2	2015/16	D	A1	4.89 (0.13)	5.05	-
2	2015/16	D	A2	4.74 (0.11)	5.01	-
2	2015/16	D	B1	4.01 (0.35)	4.92	-
2	2015/16	D	B2	4.72 (0.23)	4.91	-
2	2015/16	D	C1	2.70 (0.28)	3.44	-
2	2015/16	D	C2	2.07 (0.09)	3.54	-



# Chapter 4

## Simulating water lateral inflow and its contribution to spatial variations of rainfed wheat yields

This chapter has been published as:

**Tenreiro, T. R., Jeřábek, J., Gómez, J. A., Zumr, D., Martínez, G., García-Vila, M., Fereres, E.** (2022). Simulating water lateral inflow and its contribution to spatial variations of rainfed wheat yields, European Journal of Agronomy, 137, 126515.

### Abstract

Spatial variations of crop yields are commonly observed in typical rainfed systems worldwide. It is accepted that such variations are likely to be associated, among other factors, with water spatial variations due to lateral water flows occurring in fields with undulating topography. However, some of the main processes governing water spatial distribution such as lateral flow are not entirely considered by the most commonly adopted crop simulation models. This brings uncertainty to the process of yield simulation at field-scale, especially under water-limited conditions. Although it is expected that lateral water movement determines spatial variations of crop yields, it is still unclear what is the net contribution of lateral water inflows (LIF) to spatial variations of rainfed yields in fields of undulating topography. In this sense, by combining field experimentation, simulation models (HYDRUS-1D and AquaCrop), and the use of artificial neural networks, we assessed the occurrence and magnitude of LIF, and their impact on wheat yields in Córdoba, Spain, over a 30-year period. Seasonal precipitation varied over 30 years from 212.8 to 759.5 mm, and cumulative LIF ranged from 30 to 125 mm. The ratio of seasonal cumulative LIF divided by seasonal precipitation varied from 10.7 to 38.9% over the 30 years. The net contribution of LIF to spatial variations of rainfed potential yields showed to be relevant but highly irregular among years. Despite the inter-annual variability, typical of Mediterranean conditions, the occurrence of LIF caused simulated wheat yields to vary +16% from up to downslope areas of the field. The net yield responses to LIF, in downslope areas were on average 383 kg grain yield (GY)  $\text{ha}^{-1}$ , and the LIF marginal water productivity reached  $24.6 (\pm 13.2)$  kg GY  $\text{ha}^{-1} \text{ mm}^{-1}$  in years of maximum responsiveness. Decision makers are encouraged to take water spatial variations into account when adjusting management to different potential yielding zones within the same field. However, this process is expected to benefit from further advances in in-season weather forecasting that should be coupled with a methodological approach such as the one presented here.

## 4.1 Introduction

Modern agriculture aims to optimize the efficiency and profitability of farming systems while sustaining the increase in food production needed for a growing population (Connor & Mínguez, 2012; Fischer & Connor, 2018; Kirkegaard & Hunt, 2010). In most typical rainfed systems worldwide, where spatial variations of crop yields are commonly observed (Bramley, 2009; Maestrini & Basso, 2018; Sadler & Russell, 1997; Sadras & Bongiovanni, 2004; Sida et al., 2021), there is an opportunity for increasing productivity of resource use by determining the management options to exploit the site-specific conditions within fields (Cassman, 1999; McBratney et al., 2005). Site-specific variations of crop yields caused by differences in water availability due to lateral inflow from up to downslope areas have been identified (Ciha, 1984; Batchelor et al., 2002; Halvorson & Doll, 1991; Rockström & Valentin, 1997; Schmitter et al., 2015). However, the contribution of lateral inflow to spatial variations of yields has not been systematically explored.

The intra-plot heterogeneity associated with lateral water flows has implications in input allocation, allowing for spatial variations in crop management in the context of precision agriculture (Ahuja et al., 2019; Nielsen et al., 1973; Sadler & Russell, 1997; Verhagen et al., 1995; Wallor et al., 2018; Ward et al., 2018). Precision agriculture would surely benefit from advances in the spatial simulation of water variations over fields. However, to model accurately rainfed yields in fields of undulating topography, we need simulation tools capable of forecasting spatial variations in water availability within a field for assessing their impact on crop performance.

Over the last decades, there has been great expansion in modelling agricultural processes at the point scale (Jones et al., 2017; Spiertz, H., 2014), but insufficient efforts have been devoted to scale up water-related processes, which vary spatially, from point to field level (Ahuja et al., 2019; Wallor et al., 2018). Tenreiro et al. (2020) have recently reviewed some of the most widely adopted crop and hydrologic models and the main opportunities to simulate spatial water variations at crop field level, and concluded that the

most promising strategies for scaling up are related to the incorporation of both surface and subsurface lateral flows when simulating crop performance. The incorporation of surface and subsurface lateral flows within simulation modelling requires innovative approaches and datasets, which should be spatially distributed and related to the geomorphological properties with implications for plant available water (Nielsen & Wendoroth, 2003; Wallor et al., 2018). However, data collection requires field experimentation conducted at "real scales", which are relatively expensive and difficult to replicate over long periods of time. Therefore, the combination of both experimentation and modelling is a valid strategy for making progress (Jones et al., 2017; Kamilaris et al., 2017; Toreti et al., 2018; Wolfert et al., 2017).

Considering the typical inter-annual variation of the processes governing crop-water spatial relations (de Wit & van Keulen, 1987), the present study investigated the following question: "*what is the net contribution of lateral water inflows to spatial variations of rainfed wheat yields in fields of undulating topography?*" Figure 4.1 illustrates graphically the relevance of our research question. To address our research question, we developed a novel methodology to explore the linkage between lateral inflows (including both surface and subsurface flows) and yield variations in specific zones within a field. By combining field experimentation, simulation models and the use of artificial intelligence, we assessed the occurrence and magnitude of lateral inflows, and their impact on wheat yields in Córdoba, Spain, over a 30-year period.

## 4.2 Materials & Methods

### 4.2.1 Experimental sites

The experimental sites consisted of two nearby hydrologically independent fields, located in Córdoba, southern Spain ( $37.8^{\circ}$  N,  $4.8^{\circ}$  W, mean altitude 170 m amsl.) of 42 and 36 ha, respectively. Two catchment areas within the selected fields, 9.5 and 6.2 ha respectively, were delineated from a flow direction raster obtained with the SAGA - Wang Liu algorithm (Wang and Liu, 2006), from a Digital Elevation Model (DEM) with 5 m spatial resolution collected with

LiDAR (CNIG, 2019). The soils are formed on Miocene marls and have been classified as Typic Haploxerert (Soil Survey Staff, 1999), or Vertisol according to the FAO classification, characterized by clay texture with a shrinking-swelling nature, high bulk density ( $1.65\text{--}1.88 \text{ g cm}^{-3}$ ) and of 1.2–1.6 m depth. The catchments have mean slopes of 2–6%, respectively N-S and W-E oriented, and with elevation varying from 140 to 195 m. Due to crop rotations, wheat was monitored in catchment one in 2019/2020 (Figure 4.2-A and 4.2-C), and in catchment two in 2020/21 (Figure 4.2-B and 4.2-D).

#### 4.2.2 Sampling scheme and experimental design

##### Geomorphological properties and sampling points

The spatial variation of soil geomorphological properties was characterized in catchment one by using an electromagnetic induction sensor (DUALEM-21S) to measure soil apparent electrical conductivity (EC<sub>a</sub>, dS m<sup>-1</sup>) within the top 0–50 and 0–90 cm soil layers, four days after a rainfall event of approximately 10 mm (McCutcheon et al., 2006) and before sowing. Soil samples (%Clay, %Sand, pH) were collected at 35 cm depth following a multistage sampling scheme that was based on two different EC<sub>a</sub>-based clusters. In catchment two, soil properties were averaged for the entire field using farm records (Table 4.1).

Topographic attributes were computed with SAGA GIS (version 2.3.2) from the DEM raster. Soil moisture sampling zones were delineated according to both elevation and the flow accumulation index (FAI). The flow accumulation index (FAI) is expressed as the absolute number of upslope cells flowing to each assigned cell of the DEM raster (Tarboton et al., 1991; Jenson & Domingue., 1988). Since it is dependent on field scale and input data spatial resolution, a normalization of the index (NFAI) was computed as follows:

$$NFAI = 1 - \frac{FAI_{MAX} - FAI}{FAI_{MAX}} \quad (4.1)$$

The NFAI was estimated with SAGA-GIS (Conrad et al., 2015) and it was represented as a raster with the same spatial resolution of the input DEM.

Three sampling zones ( $N=3$ ) were distributed along two converging flow pathways in 2019/20 (with 2-3 replicate samples of soil moisture per zone, Figure 4.2-A and -C), and along the same flow channel in 2020/21 (with three replicates per zone, Figure 4.2-B and -D). In catchment one (2019/20), the sampling points (2-3 per zone) were positioned according to the peak of the mean EC<sub>a</sub> histogram, which was done to select sites of maximum representativeness of soil properties within each zone. In catchment two (2020/21), where soil properties were averaged, the sampling points were simply positioned within each zone according to the flow direction and spaced in 1 m intervals.

## Rainfall and meteorological data

Rainfall was monitored upslope with an autonomous rain gauge system (ECRN-100, ZENTRA Cloud ZL6, 16 cm collector diameter), with 10 min time resolution. The rain gauge was installed in an intermediate site, located 600–900 m from each catchment, at a spot above the highest point of each catchment (Appendix-A1). In 2019/20, manual pluviometers (TFA 47.1008, 12 cm collector diameter) were positioned at each observation point ( $N=7$ ) to capture rainfall coefficient of variation. The coefficient of variation was computed in relation to the autonomous rain gauge measurements for two separate rainfall events before crop emergence in 2019/20.

Weather data (Figure 4.3) were obtained from a meteorological station nearby, located less than 5 km away from each field (Appendix-A1). These included global radiation, wind speed, air temperature and relative humidity, which were daily averaged over half hourly measurements, and used to compute daily mean E<sub>To</sub> values according to FAO Penman-Monteith (Allen et al., 1998).

## Soil water content and lateral inflow calculations

Soil water content (SWC), expressed in mm, was monitored with multisensor capacitance probes (SENTEK-D&D 90 cm, Sentek Technologies Ltd., Australia), installed at each observation point ( $N=7$  in 2019/20 and  $N=9$  in 2020/21). Probes were respectively installed at day after sowing (DAS) 20 and 5 in

2019/20 and 2020/21. Each probe integrates nine sensors, distributed at every 10 cm, from 5 to 85 cm depth. The sensors provide near real-time data on soil temperature and SWC (30 min time resolution), which were accessed through the software IrriMAX Live ([www.irrimaxlive.com/](http://www.irrimaxlive.com/)).

An error assessment was conducted for each probe, based on discrete SWC measurements ( $N=20 \ date \times depth(z)$  per observation point) made with a neutron probe in 2019/20 (NP, Campbell Pacific Nuclear Scientific, Model 503). A relative error assessment of the capacitance probes (CP) was conducted by taking the NP measurements as control. The NP access tubes were installed 40 cm apart from the CP. One tube per observation point was considered sufficient to meet the requirements for precision and statistical power (Evett et al., 2009). NP measurements were taken at five different dates, and at 15, 30, 60 and 90 cm depth, and those values were correlated with the capacitance probe sensors located at 15 cm depth, the average of 25–35 cm, the average of 55–65 cm, and the last sensor located at 85 cm depth, respectively. The NP was calibrated with gravimetric measurements of SWC, obtained in the same soil type in a farm nearby (Appendix-A1). The NP calibration linear functions varied from depth to depth (Appendix-Table A) and were characterised by  $R^2$  values of 0.96–0.98 for deep soil layers (30–90 cm) and 0.81–0.85 for the surface depth (0–30 cm). Multiple calibration functions were tested in the IrriMAX Live software in order to minimize error fluctuations per probe and depth. The ‘Sentek D&D cracking clays’ calibration function, available in IrriMAX Live software, was chosen as the most suitable option for our soil type (Paltineanu & Starr, 1997; RoTimi Ojo et al., 2015). IrriMAX Live values (i.e., from CP measurements) were corrected with the NP measurements as follows:

$$SWC_c = \sum_{z=0cm}^{z=90cm} (SWCr_z \cdot \Omega_z) \quad (4.2)$$

where  $SWC_c$  represents the corrected SWC (expressed in mm), for the entire profile and  $SWCr_z$  represents the SWC value (in mm) provided by the IrriMAX calibration function from raw input data, measured with the CP at depth  $z$ .  $\Omega_z$  is the correction factor (unitless) estimated for each  $probe \times depth(z)$  combination

and it was computed as the mean ratio between the SWC measurements with the NP and the capacitance probes for depth  $z$ . The same calibration was used in 2020/21. Daily values of SWC were computed for each observation point as the simple average of all  $SWC_c$  values registered within the same day.

Lateral inflows (LIF) were calculated in daily time-steps for each observation point, from daily values of  $SWC_c$ . Daily variations of  $SWC_c$  were computed as the difference between  $SWC_c$  in day  $n$  and day  $n-1$ . Observed lateral inflows were assumed to be the absolute difference between the daily  $SWC_c$  variation and the daily rainfall registered by the rain gauge, which was computed as follows:

$$SWC_{c(n)} - SWC_{c(n-1)} > P_{(n-1)} \Rightarrow LIF_n = SWC_{c(n)} - SWC_{c(n-1)} - P_{(n-1)} \quad (4.3)$$

Every day  $SWC_c$  varied by an amount greater than the registered precipitation ( $P$ ), a lateral inflow was assumed to take place of a magnitude equivalent to the difference between the increase in SWC and the rainfall amount. For the calculation of LIF, deep percolation and evapotranspiration (ET) were not considered for the following reasons. In our soil type, deep percolation approaches zero (Giraldez & Sposito, 1985). In the case of ET, the crop water extraction following a rainfall event is quite limited, as intercepted water evaporates first from wet canopies as shown by Tolk et al. (1995). Therefore, considering their very small magnitudes in our case, we did not attempt to measure or estimate either deep percolation or ET, as it would have added uncertainty to our LIF calculations. Our approach determines the minimum LIF quantity that actually takes place under field conditions. Therefore, if under our approach LIF is relevant for determining crop yield, then it must play an even more important role in actual yield variations within a field.

## Crop data and field observations

The wheat cultivars *KIKO-NICK-R1* and *Avispa-R1* were sown in catchment one (2019/20 season) and two (2020/21 season), respectively. Seeding rates were 180 ( $\pm 20$ ) kg ha $^{-1}$ . Catchment one was fertilized with two applications of calcium ammonium nitrate plus sulphur (62 + 50 kg N ha $^{-1}$ ) and catchment two was fertilized with ammonium sulfate plus urea (160 kg N ha $^{-1}$ ), and diammonium phosphate (60 kg P ha $^{-1}$ ). Crop nutrient status was controlled through foliar analysis conducted at flowering, with no critical deficiencies observed.

Ground measurements of canopy cover (CC) were conducted in both trials, every 10-20 days, during the monitoring period. CC was measured at each observation point using a digital camera (Canon EOS 550D + EFS 18-135 mm CMOS APS-C 18.7 MP) at 1.5 m height and an image processing package (Patrignani & Ochsner, 2015). CC curves were used to parameterize crop growth related factors in the simulations with both the HYDRUS and the AquaCrop models (Table 4.2). Crop stages duration were obtained from field observations of phenological development and adjusted according to the CC curves obtained from satellite NDVI, as described in detail by Tenreiro et al. (2021). Both seeding rates and site-specific plant density values were registered and used for modelling parameterization as well. Rooting depth trend was inferred in each observation point using the SWC information (Table 4.2). Maximum rooting depth was considered to be equal to average soil depth (1.4 m).

Grain yield was harvested by combine, using the '*New Holland*' *Precision Land Manager* (PLM) software which took as an input the shapefiles generated by the combine harvester monitor (Fendt PLI C 5275). Yield values were computed with R-studio (Lovelace et al., 2019), under a spatial resolution of 100 m $^2$  with the equation of Reitz & Kutzbach (1996). The accuracy of the yield data from the combine monitor was assessed by comparing manual samples taken at each observation point in catchment one (sampled areas of 0.9 m $^2$ ) against the combined harvest data. More information regarding the yield spatial assessment is provided in Tenreiro et al. (2021). Yield observations

were obtained from yield maps of catchment two (2015-2020 harvests). Zonal means and standard deviations were estimated and values were plotted against simulated yield values.

#### 4.2.3 Modelling approach

To simulate water lateral flow and its contribution to spatial variations of crop yield, a multi-stage modelling approach was designed (Figure 4.4). The modelling approach was based on field experimental data (step-0, Figure 4.4) and it was divided in four subsequent steps (steps 1-4, Figure 4.4). Field experimental data were collected according to a spatial analysis based on GIS data aimed at identifying observation points according to standard hydrological connectivity rules (i.e., field channel networks and flow accumulation index), topographic attributes (i.e., elevation, slope orientation) and other geomorphological properties (step-0, Figure 4.4). The first step consisted on lateral inflow calculations from field observation data, in daily time-steps and for each observation point (step-1, Figure 4.4). Then, a hydrologic routine simulated the lateral outflows (with HYDRUS-1D, Šimůnek et al., 2018), generated upslope in form of surface run-off, and then the occurrence of lateral water inflows (LIF) was predicted at each sampling point (step-2, Figure 4.4). For each sampling point, daily calendars of LIF were determined according to a hydrological analysis that combined field measured data with both HYDRUS simulations and a machine learning (empirically based) approach (step-2, Figure 4.4). An Artificial Neural Network (ANN) was used to simulate LIF over a period of 30 years (1990-2020). Maier & Dandy (2000) reviewed in detail the main applications of ANN models for the prediction and forecasting of hydrological variables. The ANN model architecture was delineated according to a trial-and-error procedure (step-2, Figure 4.4). Then, the outputs of the hydrology-based modelling routine (step-2, Figure 4.4) were used as inputs to the crop-based modelling stage (step-3, Figure 4.4). The crop routine incorporated the calendars of daily LIF values as additional water supply (step-3 and -4, Figure 4.4). The forecasted LIF over a 30-year period, were then introduced as additional water supply through the irrigation module

in wheat yield simulations with the AquaCrop model (Steduto et al., 2009). Both the HYDRUS and the AquaCrop modelling routines functioned at one dimensional space. Every simulation run was performed at one-dimension, but we followed a ‘feedforward scheme’, which allowed us to simulate and incorporate LIF throughout consecutive modelling steps.

## Simulating the hydrology with HYDRUS

The HYDRUS-1D model (Šimůnek et al., 2018) was used to simulate water infiltration and estimate surface run-off and the spatial variation of soil hydraulic properties. The HYDRUS-1D is a physically-based model that solves Richards’ equation for transient water transport in variably saturated porous media, and incorporates processes such as soil evaporation, crop transpiration, root growth, and plant water uptake (Šimůnek et al., 2018, Tenreiro et al., 2020). The standard ‘van Genuchten-Mualem model’ was used to represent the soil hydraulic characteristics (Mualem, 1976; van Genuchten 1980).

The soil profile was modelled in one dimension at each of the measured points in both experimental years. Two soil layers were considered: surface and sub-surface. The surface layer was set at the first 10 cm of the modelled profile, while the sub-surface layer was set down to 140 cm depth. The meteorological conditions governing evaporative demand were set as the upper boundary condition. A ‘free drainage’ condition was considered at the bottom boundary of the soil profile because well drainage conditions with lack of soil reduction symptoms were observed at the BC horizon below 140 cm depth, during a pit excavation conducted prior to this study, and no indications of a water table were found. A steady state of SWC based on the measured data was used to set initial conditions (i.e.,  $0.20 \text{ cm}^3 \text{ cm}^{-3}$  in the season 2019/20 and  $0.27 \text{ cm}^3 \text{ cm}^{-3}$  in the season 2020/21, following an average 60 day long warm-up period to minimize the interference of soil-water altered conditions due to probes installation).

The potential transpiration was estimated through the soil cover partitioning method as described in Tenreiro et al. (2020). The same canopy cover curves

used in simulations with the AquaCrop model were used for this purpose. Root growth rates were assumed to be linearly constant (Table 4.2). The rooting depth distribution function was based on the model of Hoffman & van Genuchten (1983) and the water stress function of Feddes (1978), adapted for wheat and described in Wesseling et al. (1991), was considered. Since the inverse optimization procedure is sensitive to the rooting depth distribution, this was adjusted by means of SWC (Zumr et al., 2006).

HYDRUS-1D parameterization was optimized using the measured soil water content and the best fitted saturated hydraulic conductivity ( $K_{SAT}$ ) values estimated for each soil layer at each observation point. The parameters were optimized by the *Marquardt-Levenberg Optimization Algorithm* (Šimunek et al., 2012). The initial estimation of parameters was based on field measurements of soil texture and bulk density (Table 4.1) using the *Rosetta* neural network predictor (Schaap et al., 2001). Initial  $K_{SAT}$  ranged according to van Genuchten's  $\alpha$  and  $n$  shape parameters, which were estimated through the *Rosetta* predictions (more information is provided in Table A.1 in Tenreiro et al. (2021)). The minimum and maximum values of saturated and residual water content were estimated from the measured SWC data. Parameter intervals used for the optimization are given in Table 4.3. Due to spatial variability in measured SWC, these ranges were further adjusted for each sampling point separately. Several initial estimations were used to better explore the parameter spatial dimension and to avoid falling in local minimum values (Šimůnek et al. 2018). The objective function minimum was found typically after 7–20 iterations of the *Marquardt-Levenberg Optimization Algorithm*. However, due to the numerical instabilities during several optimization runs, multiple non-converging model runs needed to be performed for each sampling point (i.e., typically 5 to 50 runs). A daily calendar of surface run-off was estimated for each sampling point and experimental year.

## Forecasting lateral inflows

Daily calendars of lateral inflows ( $LIF_n$ , expressed in  $\text{mm day}^{-1}$ ) at each sampling point were computed as a function of the run-off generated at that

same point, as well as the run-off flowing from neighbouring and upslope points within the same field. The daily values of simulated run-off with HYDRUS were used as input features into a static multi-layer (feedforward) ANN model (de Vos & Rientjes, 2005). Different performance criteria were used to determine network inputs and model architecture (Maier & Dandy, 2000). Input features ( $N=6$ ) were selected according to a principal component analysis (PCA) associating different potential predictors ( $N=12$ ) with LIF observations at each sampling point (Table 4). Model architecture was delineated according to a trial-and-error procedure (Roadknight et al., 1997; Senthil-Kumar et al., 2005; Shukla et al., 1996). The ANN prediction accuracy and its computation-training speed were assessed with R-studio (Günther & Fritsch, 2010).

The ANN processes multiple algebraic operations over several input features which are defined by a single column vector ( $\vec{X}$ ). Each of the inputs is attenuated by a weight factor ( $w$ ) that is linked to a transfer function (i.e., a logistic transformation of the data). The calculation scheme of an ANN using  $i$  input features (computed in daily time steps  $n$ ) and  $k$  hidden layers can be simplified as following:

$$CUM.LIF_n = \sum^k \begin{pmatrix} x_{n1} \\ \vdots \\ x_{ni} \end{pmatrix} \cdot \begin{pmatrix} (\alpha_{11} \cdot w_{11}) \dots (\alpha_{1i} \cdot w_{1i}) \\ \vdots \\ (\alpha_{j1} \cdot w_{j1}) \dots (\alpha_{ji} \cdot w_{ji}) \end{pmatrix} + \epsilon \quad (4.4)$$

where  $\alpha$  corresponds to the transfer function operator and the subscript  $j$  delineates the number of nodes in each layer (de Vos & Rientjes, 2005). The bias coefficient ( $\epsilon$ ) corresponds to the overall error of the network, i.e., the sum of each node residuals, and the dependent variable is expressed in daily cumulative terms ( $CUM.LIF_n$ ) to reduce the effects caused by temporal deviations between observations and predictions of LIF. The relation between daily ( $n$ ) cumulative LIF ( $CUM.LIF_n$ ) and daily LIF ( $LIF_n$ ), at each point-site, computed from the beginning of simulation until day  $n = i$  is expressed as following:

$$CUM.LIF_n = \sum_0^{n=i} LIF_n \quad (4.5)$$

Daily  $LIF_n$  was therefore estimated as the absolute difference between  $CUM.LIF_n$  at day  $n$  and  $CUM.LIF_{n-1}$  at day  $n - 1$ :

$$LIF_n = CUM.LIF_n - CUM.LIF_{n-1} \quad (4.6)$$

Catchment one data (2019/20) were used to train the ANN model while catchment two data (2020/21) were used for testing. For each trial, ANN-LIF predictions were plotted against the observations of LIF and the Willmott index of agreement ( $d$ ), the  $R^2$  and the RMSE were used as statistical indicators of ANN performance. The best-performing ANN was used to forecast  $CUM.LIF_n$  time-series over a period of 30 years, from which the  $LIF_n$  values were derived for each zone.  $CUM.LIF_n$  values were smoothed by a 'monotonically increasing function' to preserve a strictly increasing pattern over season.

The best performing ANN was applied to forecast daily LIF in catchment two over the same period of 30 years. The model HYDRUS-1D was used to simulate run-off at each sampling point and for each growing season [1990-2020]. Catchment two properties (Table 4.1) and mean values for crop data (Table 4.2) were used as input features to the best performing ANN. Weather records (i.e., global radiation, wind speed, air temperature, relative humidity and rainfall) for Cordoba [1990-2020] were obtained from the same weather station (Appendix-A1).

## **The crop model stage with AquaCrop**

The AquaCrop v6.1 model (Steduto et al., 2009; 2012) was used to simulate net yield responses to lateral inflows ( $NYR_{LIF}$ ).  $NYR_{LIF}$  was assumed to be the absolute difference in terms of simulated (water-limited) yield, between these two different scenarios:

- 1) yield simulation without lateral water inflow, following the standard

assumption in AquaCrop where water inflow takes only place due to vertical infiltration ( $Y_1$ );

2) yield simulation including lateral inflow as an additional water supply, which was set in the model through the irrigation module and computed as the LIF forecasted by the ANN ( $Y_2$ ).

$NYR_{LIF}$  was calculated as following:

$$NYR_{LIF} = Y_2 - Y_1 \quad (4.7)$$

where all terms are expressed in  $Mg\ ha^{-1}$ . Yield response to lateral inflow was also estimated in relative terms ( $NYR_{LIF-rel}$ ) as following:

$$NYR_{LIF-rel} = \frac{NYR_{LIF}}{Y_{Mean}} \quad (4.8)$$

where  $Y_{Mean}$  represents the mean yield estimated in scenario 1.

A sequence of 30 years LIF [1990-2020] and its impact on grain yield was simulated to derive probability distribution functions of both  $NYR_{LIF}$  and  $NYR_{LIF-rel}$ . From the obtained series of  $NYR_{LIF}$ , the LIF marginal water productivity ( $LIF.MWP$ ) was computed as following:

$$LIF.MWP = \frac{NYR_{LIF} \times 1000}{CUM.LIF} \quad (4.9)$$

where for each marginal unit of water supplied as LIF (expressed in mm), the crop was assumed to respond with additional grain yield (Passioura & Angus, 2010).  $LIF.MWP$  was expressed in  $kg\ grain\ ha^{-1}\ mm^{-1}$ .

The AquaCrop model was parameterized with field data (catchment two values, Table 4.1) and crop data (mean values, Table 4.2). Long-term simulations were conducted for catchment two because it showed lower standard deviations for zonal means (Table 4.1 and 4.2). The hydraulic conductivity values used in AquaCrop simulations ( $K_{SAT}$  mean, expressed in  $mm\ day^{-1}$ ) were the mean values reported in Table A.1 in Tenreiro et al. (2021). Two soil horizons were also considered (surface above 30 cm depth and sub-surface from 30 to 140 cm depth). The initial curve number was set at a value of 84 (i.e., hydrologic group

D). More information regarding the water balance approach that is followed by AquaCrop found in Tenreiro et al. (2020). Simulated canopy growth and grain yield were validated through field observations (Section 4.2.2).

#### 4.2.4 Statistical analysis

Differences among sampling zones were tested for significance, the null hypothesis was checked for the mean differences of both observed SWC and LIF with the non-parametric Tukey's HSD (honestly significant difference) test because these variables were not normally distributed. Non-normality was checked with the Shapiro-Wilk test (Acutis et al., 2012). Within field spatial variations (among the three sampling zones) were assessed with standard coefficients of variation. The residuals of the ANN features were checked to be randomly distributed (Supplementary material). HYDRUS simulations of SWC were tested with the Nash–Sutcliffe model efficiency coefficient, the  $R^2$  and the RMSE (Moriasi et al., 2007; Nash & Sutcliffe, 1970; Yang et al., 2014). AquaCrop simulations of grain yield were tested with the Willmott  $d$  index, the  $R^2$  and the RMSE (Willmott, 1981). Spearman's correlation analysis was used to explore relationships among several crop variables and both  $NYR_{LIF}$  and  $LIF.MWP$ .

### 4.3 Results

#### 4.3.1 Experimental data

CPs overestimated SWC in comparison with the NP measurements. The mean correction factor ( $\Omega_z$ ) varied with both probe location and sensor depth (Appendix-A2). The amplitude of error variation was greatest at the surface layers (0–30 cm) but it tended to stabilize for deeper layers (Appendix-A2).

The soil water content ( $SWC_c$ ) varied spatially in both experimental catchments/years (Table 4.5 and Figure 4.5). According to the HSD-test conducted at a 5% level of significance, the annual mean SWC values varied significantly among sampling zones (Table 4.5). Our sampling scheme did also capture significant differences among zones in LIF values (Table 4.5), both in

daily ( $LIF_n$ ) and cumulative terms ( $CUM.LIF_n$ ). However, as observed for SWC, zone B showed a ‘transient profile’, where both the SWC and LIF values were not ‘systematically and significantly’ different from the adjacent sampling zones. For an accurate distinction among zones within our experimental sets, only zones A and C were consistently different in terms of SWC (and LIF) variations (Figure 4.2 and Table 4.5).

The magnitude of both SWC and LIF values also varied among catchments and/or years (Table 4.5 and Figure 4.5). Catchment one showed lower mean values than catchment two (Table 4.5), partly explained by the differences in initial SWC (Figure 4.5). There might be also differences in soil properties among sampling zones influencing SWC measurements in catchment two, because clay and sand content, and bulk density were averaged for all zones (Table 4.1). These differences can also be attributed to the performance of the CP readings according to the calibration functions used. The mean coefficient of variation of daily rainfall, measured for two events in 2019/20, was 9.5%, which was lower than the LIF coefficient of variation (13%) estimated from probes data among zones in the same season.

#### 4.3.2 HYDRUS run-off simulation and ANN-LIF forecasting

The HYDRUS-1D simulations of SWC showed a mean Nash–Sutcliffe model efficiency of 0.61 ( $\pm 0.14$ ) and 0.92 ( $\pm 0.04$ ) for each of the experimental years, respectively (Figure 4.6). In addition, the model has also increased the  $R^2$  (while reducing the RMSE), from the first to the second year (Figure 4.6). A higher variation of performance indicators was also observed in catchment one than in two. The run-off coefficients (cumulative simulated run-off divided by the cumulative precipitation over the same period) varied from 4 to 13% in 2019/20 year and from 7 to 24% in 2020/21.

The best-performing ANN had six input features, four hidden layers with three to four nodes per layer (Figure 4.7-A and -B). The best-performing features were  $x_1, x_3, x_5, x_6, x_{10}, x_{11}$  (Table 4.4). According to the PCA euclidean scores, these features were the most correlated with LIF, significantly contributing to

the first and second components, which explained together 52.2% of LIF variance (Appendix-A3). The network was solved in 9962 steps (Figure 4.7-D). The best performing ANN had a  $R^2$  of 0.97 and 0.78, respectively for the training and the testing set, and the RMSE of predicted *CUM.LIF* values were 10.1 and 14.8 mm, respectively (Figure 4.7-C). Observed *CUM.LIF* were scatter-plotted against simulated values and no biased error trend was observed for any of the subsets, as the residuals did not vary with the level of predicted LIF. Despite showing a solid forecasting capacity, the ANN exhibits a general trend to overpredict LIF in comparison to measured values (Figure 4.7-C).

Daily LIF calendars were predicted with the ANN (Figure 4.7) over a 30-year period (Figure 4.8). While seasonal precipitation varied over 30 years from 212.8 to 759.5 mm, cumulative LIF ranged from 30 to 125 mm (Figure 4.8). The ratio of seasonal cumulative LIF divided by seasonal precipitation varied from 10.7 to 38.9% over the 30 years.

### 4.3.3 Yield simulations with AquaCrop

The AquaCrop simulation outcomes in terms of yield response to LIF were highly variable from year to year. Simulated plotted against observed yields [2015-20] showed RMSE,  $R^2$  and Willmott  $d$  index respectively equal to 0.374 ( $\text{Mg GY ha}^{-1}$ ), 0.35, 0.76 (Appendix-A4). Differences among zones were only experimentally tested for the two systematically distinct zones (i.e., A and C according to Table 4.5). In this sense, the cumulative probability of  $NYR_{LIF}$  for zone C can be also interpreted as the absolute difference between the two curves shown in Figure 4.9-A. Since cumulative LIF in zone A was negligible (Appendix-A6), the  $NYR_{LIF}$  was only estimated for the water-receiving zone C (Figure 4.9-B). In this case,  $NYR_{LIF}$  corresponds to  $Y_2$  minus  $Y_1$ , in zone C, or  $Y_2$  in zone C minus  $Y_2$  in zone A (Figure 4.9-A). Figure 4.9 shows the cumulative probability distribution curves for simulated yield responses in the two significantly distinct zones (A and C). Results are shown both in absolute and relative terms (Figure 4.9-B and -D). Mean values of simulated  $NYR_{LIF}$  and  $NYR_{LIF-rel}$  were 383  $\text{kg ha}^{-1}$  (Figure 4.9-C) and 16.2%, respectively. Absolute

values are expressed in terms of grain yield (GY) in dry mass (DM).

Following crop yield simulations over a period of 30 years, the simulated seasonal  $NYR_{LIF}$  varied from -0.14 to 2.08 Mg ha<sup>-1</sup>, corresponding to a relative net contribution of LIF to grain yield ( $NYR_{LIF-rel}$ ) that ranged from -3% up to +168%.  $NYR_{LIF}$  was larger than 800 kg GY ha<sup>-1</sup> in five out of 30 years (i.e., 1992/93, 1999/00, 2003/04, 2004/05, 2008/09). By contrast, in other five years (i.e., 1990/91, 1996/97, 2001/02, 2007/08 and 2019/20), LIF caused yield losses due to water excess (Figure 4.9-C and Appendix-Table B). For the remaining years, the  $NYR_{LIF}$  ranged from none to 670 kg GY ha<sup>-1</sup> (e.g., 2014/15 as shown in Figure 4.9-C), being below 265 kg GY ha<sup>-1</sup> for at least 50% of the years (Figure 4.9-B).

According to the Spearman's correlation matrix, shown in Appendix (Figure A7),  $NYR_{LIF}$  was negatively correlated with  $CUM.P$  and simulated yields. The lower the yields simulated in higher zones (i.e., zone A), the higher was  $NYR_{LIF}$  at lower zones (i.e., zone C), which is caused by marginal water productivity gains (Appendix-A7 and -Table B). According to our simulations, the occurrence of LIF over fields in our conditions has resulted in a mean  $LIF.MWP$  of 24.6 ( $\pm 13.2$ ) kg ha<sup>-1</sup> mm<sup>-1</sup> in years of maximum responsiveness (Appendix-Table B).  $LIF.MWP$  were also significantly and negatively affected by  $CUM.LIF$ , which indicates that lower  $CUM.LIF$  also increases  $LIF.MWP$  (Appendix-A7).

## 4.4 Discussion

### 4.4.1 SWC measurements and lateral inflow calculations

The accuracy of the SWC measured with the CPs indicate that there is a clear 'trade-off' between the precision of SWC measurements and the time frequency/autonomy of the data collection system (Appendix-A2). The amplitude of error variation was the greatest at surface layers (0-30 cm) but it appeared to stabilize for deeper layers, which is likely to be explained by a better relation between the moisture in the sensing volume and the average zonal moisture (Chanzy et al., 1998) . Lighter hysteresis effects at deeper soil layers may

also contribute to the error stabilization with depth (Evett et al., 2009). The SWC correction method that we followed assumes that a ‘linearity condition’ is fulfilled, as described in Campbell (1990) and Chanzy et al. (1998). Similar behaviour was observed by Mwale et al. (2005), who analysed this issue more in detail. However, although Mwale et al. (2005) reported a 20% average error between capacitance and NPs (with a mean  $R^2$  value of 0.7), their measurements were conducted in lighter soils and under semi-controlled conditions. Our results show a systematic overestimation of CP measured SWC (Appendix-A2), which highlights that CP should not be used to monitor SWC in absolute terms without previous calibration.

Although there is uncertainty on  $SWC_c$  measurements, the observed pattern of inferred LIF among zones is in line with both mean zonal elevation and  $NFAI$  values (Table 4.1 and 4.5). The higher is the  $NFAI$  (and the lower the elevation), the larger is  $CUM.LIF$  inferred from field observations of SWC (Table 4.1 and Figure 4.5). It is accepted that a fraction of these differences could be associated with rainfall spatial variation within field as one single pluviometer was used for monitoring rainfall input. However, the fact that the rainfall coefficient of variation (9.5%) was lower than the relative variation of LIF among zones (13%) indicates that our modelling scheme is appropriate to assess our research question.

Neither ET nor deep percolation were directly considered for LIF calculations as done by others (Klaij & Vachaud, 1992; Rockström & Valentin, 1997). We recognize that our approach may lead to an underestimation of LIF values because these flows play an additive role when computing LIF from a standard water balance approach. However, it was assumed that the inclusion of such flows would add uncertainty to our analysis because they were not measured. We consider that plant water uptake tends to be very low in rainy days due to transpiration suppression by evaporation of canopy intercepted water (Tolk et al., 1995). In addition, deep percolation approaches zero in our soil conditions (Giraldez & Sposito, 1985), which contrasts with the case of the sandy soils assessed by Klaij & Vachaud. (1992). We decided to calculate LIF by relating measured SWC daily variations with daily measurements of

rainfall according to a conditional term (i.e., LIF was only considered when daily SWC varied by an amount greater than the registered rainfall) because this method selects only days when LIF took certainly place. In this sense, we focused only on measured values avoiding uncertainty on calendarization of LIF due to the inclusion of non-measured flows. Our results must be interpreted as the minimum LIF quantity that actually took place under our experimental conditions, suggesting that if LIF is relevant for determining crop yield under our approach, then it must play an important role in yield variations within a field.

#### 4.4.2 HYDRUS simulations and lateral inflow predictions

The HYDRUS-1D simulations showed a solid ‘goodness of variance fit’ according to the statistical indicators presented in Figure 4.6. The observed ‘Nash–Sutcliffe’ model efficiency coefficients indicate a robust match between the simulated and the observed SWC (Moriasi et al., 2007; Yang et al., 2014). The deviation statistics RMSE indicate also that the simulated results were sufficiently accurate, despite the lower  $R^2$  values in the training set (i.e., 2019/20 as shown in Figure 4.6). These are likely to be explained by the  $R^2$  insensitiveness to additive and proportional differences between the simulated and measured data SWC (Yang et al., 2014). The observed mean run-off coefficients (i.e., 8.5 and 15.5%, respectively for each experimental catchment/year) were in line with other published results obtained for cereals, under conventional systems in our conditions and with similar year precipitation patterns (Lasanta et al., 2000; Torralba, 2013). Mediterranean environments are particularly prone to run-off due to a combination of multiple factors. García-Ruiz. (2010) identified some of these factors from which we highlight high rainfall intensity, the presence of steep slopes, frequent land-use changes and cultivation systems.

According to our experimental data, the best-performing ANN was capable of reproducing LIF patterns with enough accuracy among independent years and sites of observation (Figure 4.7). The forecasted cumulative LIF was increased in years of relatively high cumulative precipitation distributed through intensive

events (Figure 4.8). Clear examples were observed in the seasons of 1995/96, 1996/97 and 1997/98. Despite the relatively low run-off coefficients and distributed among few events (Appendix-A5), the cumulative LIF forecasted by the ANN was comprehended within considerable levels (Figure 4.8 and Appendix-A6). Although no significant LIF was captured in high elevation zones of our fields (Table 4.5 and Appendix-A6), the amount of LIF that was forecasted at receiving areas (i.e., zone C) varied from 10.7 to 38.9% of seasonal cumulative precipitation, which is explained by the flow accumulation effect (in line with the *NFAI* values, as shown in Table 4.1 and Figure 4.2).

Our best-performing ANN was fed by six different features, which could show partial collinearity among them when used in classical regression models. The existing relation between the saturated hydraulic conductivity ( $x10$ ) and run-off features ( $x1$  and  $x5$ ) is a good example of this issue. However, ANN's are known for dealing with the problem of collinearity in different forms than classical regression models (Kempenaar et al., 2016). The outputs of each layer are functions of multiple combinations of logistic transformations that involve higher orders of interactions than in the original predictors. The neural networks prevent the problems of multicollinearity and over-parameterization at the cost of interpretability (De Veux & Ungar, 1994). In addition, no clear symptoms of overfitting were observed as the ANN "goodness of variance fit" decreases from the training to the testing set and increasing the network size did not minimize the training error without also reducing the testing one (Figure 4.7-A and -B). Maier & Dandy (2000) defined the 'optimal network geometry' as the smallest network that adequately captures the relationship in the training data, and our best-performing network was chosen based on the same criteria.

The best-performing ANN was trained on 959 observations of SWC, while 1377 observations were used for testing. Although it is normally assumed that there is a tradeoff between the sizes of the training and testing set of data, sustained by the idea that "more testing data" imply "less training data" (Maier & Dandy, 2000), it is naturally arguable that this rule of thumb does not apply to all cases. In our case, no "random initialization" or bootstrapping methods were used for data division as both data sets were split by years and

according to quality components. By quality components, we refer to differences in the standard deviations associated with the means for geomorphological attributes (Table 4.1), crop data (Table 4.2) and in respect to the performance of run-off simulations (Figure 4.6). Therefore, the fact that our model was tested with a 43% larger data set than the one used for training, should not be seen as downside of our approach. The objective of a supervised model is not simply learning on the largest dataset possible but doing it in a way that maximizes its performance on unknown data. We tried to capture a larger scope of observations for validation by selecting 2020/21 data for testing, which was done to explore the ANN potential to be extrapolated over 30 years. Considering the small increase of RMSE and the acceptable decrease of  $R^2$  (Figure 4.7-C), that were observed from the training to the testing stage, we rejected the hypothesis of undertraining (de Vos & Rientjes, 2005) and decided to use the selected ANN for forecasting multiple year scenarios. In addition, the predicted cumulative LIF values were smoothed by a 'monotonically increasing function', which resulted in a mean error for individual LIF events lower than the uncertainty level associated with rainfall input (i.e., here quantified as a spatial coefficient of variation equal to 9.5%). The two experimental catchment/years also presented different initial conditions, indicating that the model performed well in both situations. While catchment one (2021) showed homogeneous initial SWC among zones, catchment two (2020) monitoring was initiated under heterogeneous conditions (Figure 4.5). Therefore, no evidence was found to support the idea that an alternative modelling approach would deal with a lower magnitude of both input and outcome uncertainty.

#### 4.4.3 Contribution of lateral flows to crop yield

The AquaCrop simulations of rainfed wheat yield showed typical fluctuations under Mediterranean conditions (Figure 4.9-C and Appendix-Table B).  $NYR_{LIF}$  was larger than 800 kg GY ha<sup>-1</sup> in five out of 30 years (Figure 4.9-B and -C). These were characterized by medium-low rainfall during the vegetative stage in combination with considerable LIF events occurring at post-anthesis (Figure 4.8), when the impact on yield of additional water is highest (Abbate

et al., 2004). A significant negative correlation was found between  $NYR_{LIF}$  and  $CUM.P$  (Appendix-A7), indicating that the yield response to LIF increases with water-deficit in low rainfall years. These results are in line with the sensitivity analysis conducted by Maina & Siirila-Woodburn (2020), who concluded that crop transpiration responds the most to lateral flows in periods of lower water supply and higher demand. By contrast, in other five years, LIF caused yield losses due to water excess (Figure 4.9-C and Appendix-Table B). These were typically characterized by intense rainfall events at early vegetative stages in combination with an absence of LIF events taking place after flowering (Figure 4.8 and Appendix-Table B). In these cases, the early intense rainfall events have generated run-off upslope (Appendix-A5), and consequently high LIF at the receiving zones, promoting soil water saturation which have impacted crop growth negatively. According to our simulations for these five years showing yield losses (i.e., 1990/91, 1996/97, 2001/02, 2007/08 and 2019/20), soil water saturation periods causing stomatal closure, and associated with high LIF, were on average 48% ( $\pm 27$ ) longer in zone C than in zone A (Supplementary material). For the remaining years, a relatively moderate and highly variable contribution of LIF to crop yield variations was observed (Figure 4.9-A and -C). We highlight that  $NYR_{LIF}$  was highly positive (or negative) for about one third of the years, being relatively moderate for the remaining two thirds of the simulation period.

In general, the simulated yields were well correlated with field observations, showing a RMSE of 374 kg GY  $ha^{-1}$  (Appendix-A4). Our yield correlations are acceptable because the RMSE of simulations is included within the error range of yield observations associated with the process of yield mapping (i.e., 172–809 kg GY  $ha^{-1}$ ). In addition, it must be considered that not all the error in our deviation statistics is contained within the simulated variable because observations are not error free (Appendix-Figure A4). The observed yields were also subjected to errors through the process of combine yield monitoring/mapping. The relatively low  $R^2$  and the modest Willmott index of agreement ( $d$ ) that characterize our yield correlations (Appendix-A4) should not be seen as a drawback of our simulation results because these are due to the short range of yield observations used, which affects the proportional

differences between simulated and measured yields (Yang et al., 2014) .

In terms of marginal water productivity (Passioura & Angus, 2010) and according to our simulations, it is considered that for each marginal unit of water supplied as LIF (expressed in mm), the crop responded with 24.6 kg ha<sup>-1</sup> of additional grain yield in years of maximum responsiveness (i.e., 1992/93, 1999/00, 2003/04, 2004/05, 2008/09). Our results are slightly above the benchmark of 20 kg ha<sup>-1</sup> mm<sup>-1</sup> that was originally proposed by French & Schultz. (1984). However, several other studies on water productivity for winter wheat grown at specific geographic regions have presented similar values (Passioura & Angus, 2010; Rattalino Edreira et al., 2018; Sadras & Angus, 2006; Silva et al., 2020). It was observed that  $NYR_{LIF}$  is more relevant in years of moderate-strong water stress due to lower rainfall (Figure 4.9-C and Appendix-A7).  $LIF.MWP$  is also negatively correlated with the simulated yield (Appendix-A7), indicating that  $LIF.MWP$  increases in years characterized by lower potential yields. In these cases, we may expect relatively higher marginal water productivity rates because the net yield response is computed over a 'water stressed baseline situation' (Abbate et al., 2004; Kirkegaard et al., 2007).

In mean terms, there was no significant correlation between  $CUM.LIF$  and  $NYR_{LIF}$  (Appendix-A7). It must be highlighted that a lack of correlation between  $NYR_{LIF}$  and  $CUM.LIF$  (over 30 years) does not mean an absence of yield response to lateral inflow.  $CUM.LIF$  represents the total amount of additional water supplied through LIF over the entire season, but such water input may impact positively or negatively crop yield, depending on several other factors which vary seasonally (e.g., rainfall amount and distribution, LIF timing, soil water status, crop developmental stage). One example is the effect of  $CUM.P$  over the  $CUM.LIF \times NYR_{LIF}$  relation.  $CUM.LIF$  and  $NYR_{LIF}$  are significantly correlated when controlling for specific  $CUM.P$  levels. For years of low  $CUM.P$  (within the 25<sup>th</sup> lowest percentile),  $NYR_{LIF}$  increased significantly with  $CUM.LIF$  (data not shown) although no significant correlation is found when  $CUM.P$  levels are not controlled (Appendix-A7). The same applies to plant available water at sowing date ( $PAW_0$ ), which is highly correlated with yield for low  $CUM.P$  years (data not shown) but this relationship has no

significance over a 30-year period (Appendix-A7). The cumulative precipitation was the most significant variable governing  $NYR_{LIF}$ , not only because it is strongly correlated with  $CUM.LIF$  and  $LIF.MWP$  but also because it affects the yield response to  $PAW_0$  and  $CUM.LIF$ .

#### 4.4.4 Agronomic implications

The use of artificial intelligence methods for agronomic research requires assumptions on key aspects and processes that vary largely across fields and among different years (Smith, 2020; Wolfert et al., 2017). Our ANN-forecasting showed a general trend for over predicting LIF (Figure 4.7-C). This could bias the yield simulations under scenario number two. The observed over-predicting trend can be associated with confounding effects of our algorithm or by the presence of unobserved causes of some input variables, which are seen as major constraints to draw 'causal inference' from data when using artificial intelligence methods such as ANN (Pearl, 2019). One limitation of our approach is that it strongly depends on the fitness of the training set. In our case, the training set included some input variables that are highly year dependent (e.g., cumulative rainfall -  $x11$ ), limiting model extrapolation due to inter-annual variability of rainfall, highly typical of Mediterranean conditions. However, we consider that the data-generating distribution of predicted LIF patterns is acceptable for the purpose of the present study (Appendix-A6) and the simulated yields were in line with field observations (Appendix-A4). To evaluate our network performance, we took advantage of the experimental training/testing datasets where both the predicted and the observed LIF outcomes were available. Both the RMSE and the  $R^2$  of our ANN model are seen as adequate given the scale of analysis and the fact that each set was obtained from independent years and sites. Alternatively, the data requirements for an entirely causal approach are vast and generally difficult or costly to satisfy in practice, particularly under the scale of our analysis.

This is the first modelling study to our knowledge, conducted at commercial fields' scale and built over experimental data, delivering LIF patterns and their simulated impact on wheat yields over 30 years. Despite of the overall

uncertainty associated with our analysis, which is associated with both scale and methodological constraints, the net contribution of lateral inflows to spatial variations of rainfed potential yields showed to be relevant but highly variable among years. Despite the inter-annual variability, typical of Mediterranean conditions, the mean  $NYR_{LIF-rel}$  over 30 years was 16%. This value is in line with the coefficients of variation obtained from yield maps in the same sites (Tenreiro et al., 2021). Other studies reported similar levels of yield variance (Batchelor et al., 2002; Florin et al., 2009). Batchelor et al. (2002) assessed (intra-plot) spatial yield variations in soybean, caused by differences in water availability, and reported mean values of 15.6% variance. Florin et al. (2009) observed spatial yield variations in wheat of 10–25%, depending on the field scale. We recognize that spatial variations of yields are naturally determined by many other factors apart from water spatial variations due to run-off and/or LIF. Examples are spatial differences in soil fertility, in soil organic matter content and/or in soil depth (Franz et al., 2020; Kravchenko & Bullock, 2000; Monzon et al., 2018). However, the present study supports with evidence that water spatial variations may play a central role within this context because the main yield limiting factors were controlled in both the experiments and the simulations taken.

The simulated  $NYR_{LIF}$  varied from year to year, depending on the combination of both year specific-meteorological conditions and crop related factors (Figure 4.8 and 4.9-C). These were mostly explained by *CUM.P* and simulated yields (Appendix-A7). We highlight that under present conditions, the estimated mean  $NYR_{LIF}$  (i.e., 383 kg GY ha<sup>-1</sup>), forecasted over the last 30 years, should not be taken as a stable benchmark for decision making. The estimation of variable application rates as a direct function of our mean results would likely imply worse management decisions than those supported by real time observations. Rainfed winter-cereals, such as winter wheat, are known for being highly plastic crops in the expression of yields (Sadras et al., 2009; Fischer et al., 2019). This aspect partially explains the relatively moderate and highly variable contribution of LIF to crop yield variations over multiple years (Figure 4.9). Therefore, we suggest that mechanistic agronomy

and continuous field visiting are necessary to deal with temporal asymmetries of spatial variations in rainfed precision agriculture.

Although farmers and agronomists are not encouraged to adjust management to (within field) spatial variation of potential yields, by directly using the present results as a reference, the integration of our methodology with in-seasonal weather data will definitely contribute to solve the trade-offs determining the adoption of precision agriculture in local conditions. Further research is needed, focusing on future changes, such as atmospheric CO<sub>2</sub> trends with impact on plants water productivity rates (Allen et al., 2011; Hsiao, 1993; Passioura & Angus, 2010; Steduto et al., 2007) and technical or commercial modifications (Balafoutis et al., 2017; Hochman et al., 2009; Kirkegaard & Hunt, 2010; McBratney et al., 2005) that might impact the existing trade-offs involved in the adoption of precision agriculture at multiple scales (i.e., field, cropping system, farm level). Agronomic implications and opportunities for precision management will naturally depend on the proportional relations between water supplying and receiving zones, which need to be assessed at both field and farm level.

## 4.5 Conclusion

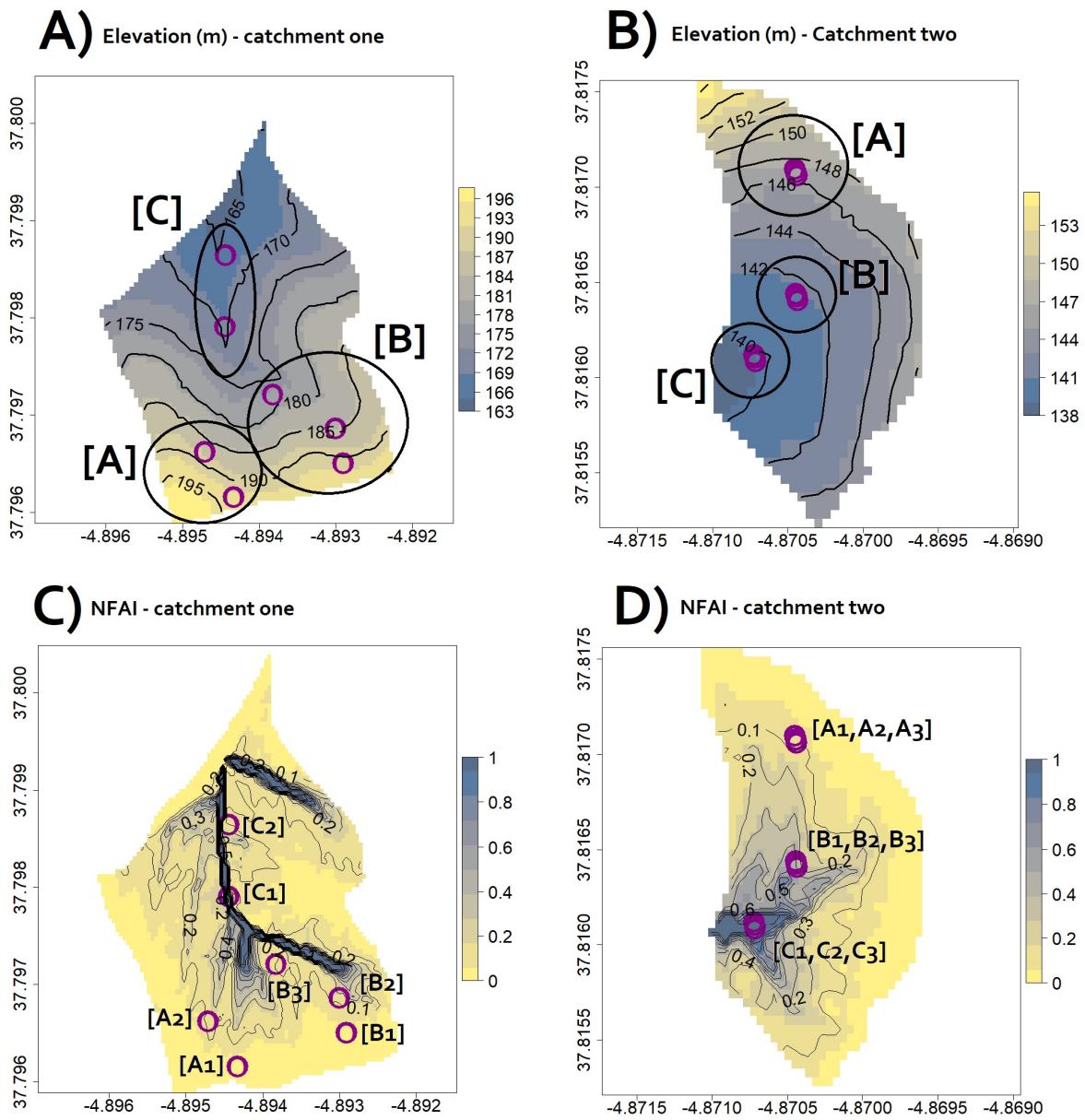
Water lateral inflows (LIF) contribute to yield variations in rainfed wheat production systems such as the one studied here. Both the forecasted LIF patterns and the crop modelling approach indicated a considerable inter-annual variation of the principal mechanisms involved in this relation. The ratio of seasonal cumulative LIF divided by seasonal precipitation varied from 10.7 to 38.9% over the 30 years. The net yield responses to LIF were on average 383 kg GY ha<sup>-1</sup> and, in years of maximum responsiveness, the LIF marginal water productivity reached 24.6 ( $\pm 13.2$ ) kg GY ha<sup>-1</sup> mm<sup>-1</sup>. Such years of maximum responsiveness were associated with low rainfall during the vegetative stages of the crop in combination with LIF occurring at post-flowering stages. Decision makers are encouraged to take water spatial variations into account when adjusting management to different potential yielding zones within the same

field. However, while some limitations for extrapolating results were identified, this process is also expected to benefit from further advances in in-season weather forecasting that should be coupled with a methodological approach such as the one presented here.

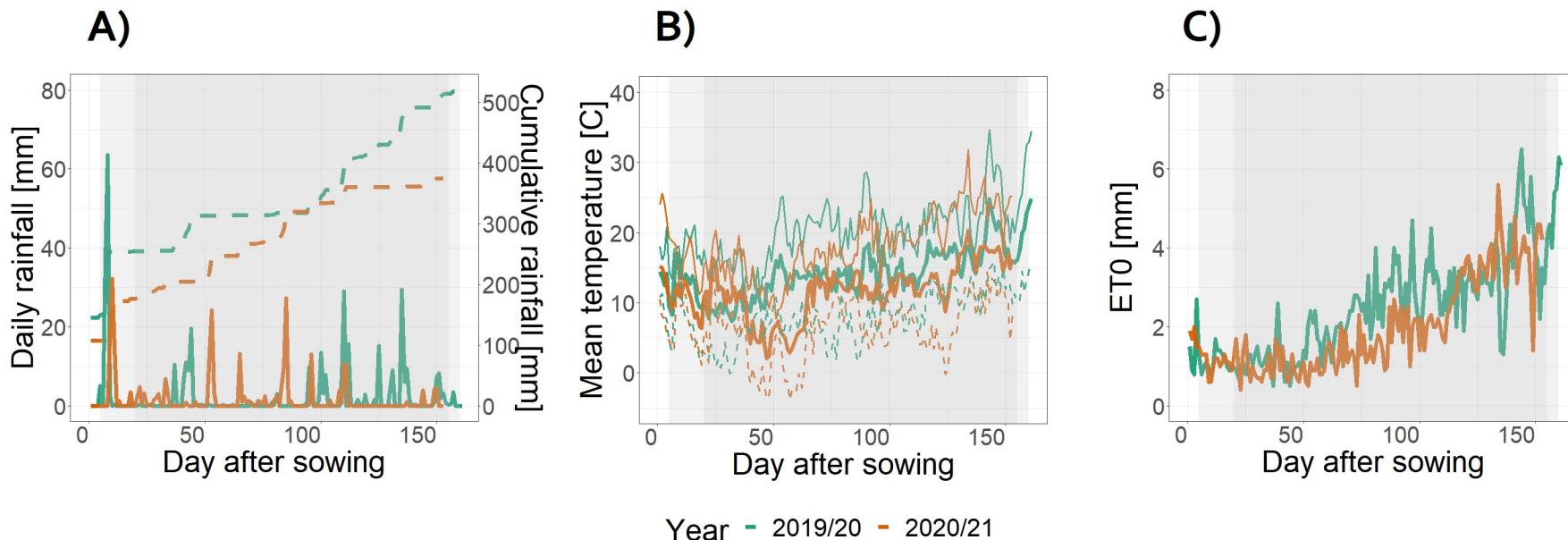
## FIGURES - Chapter 4



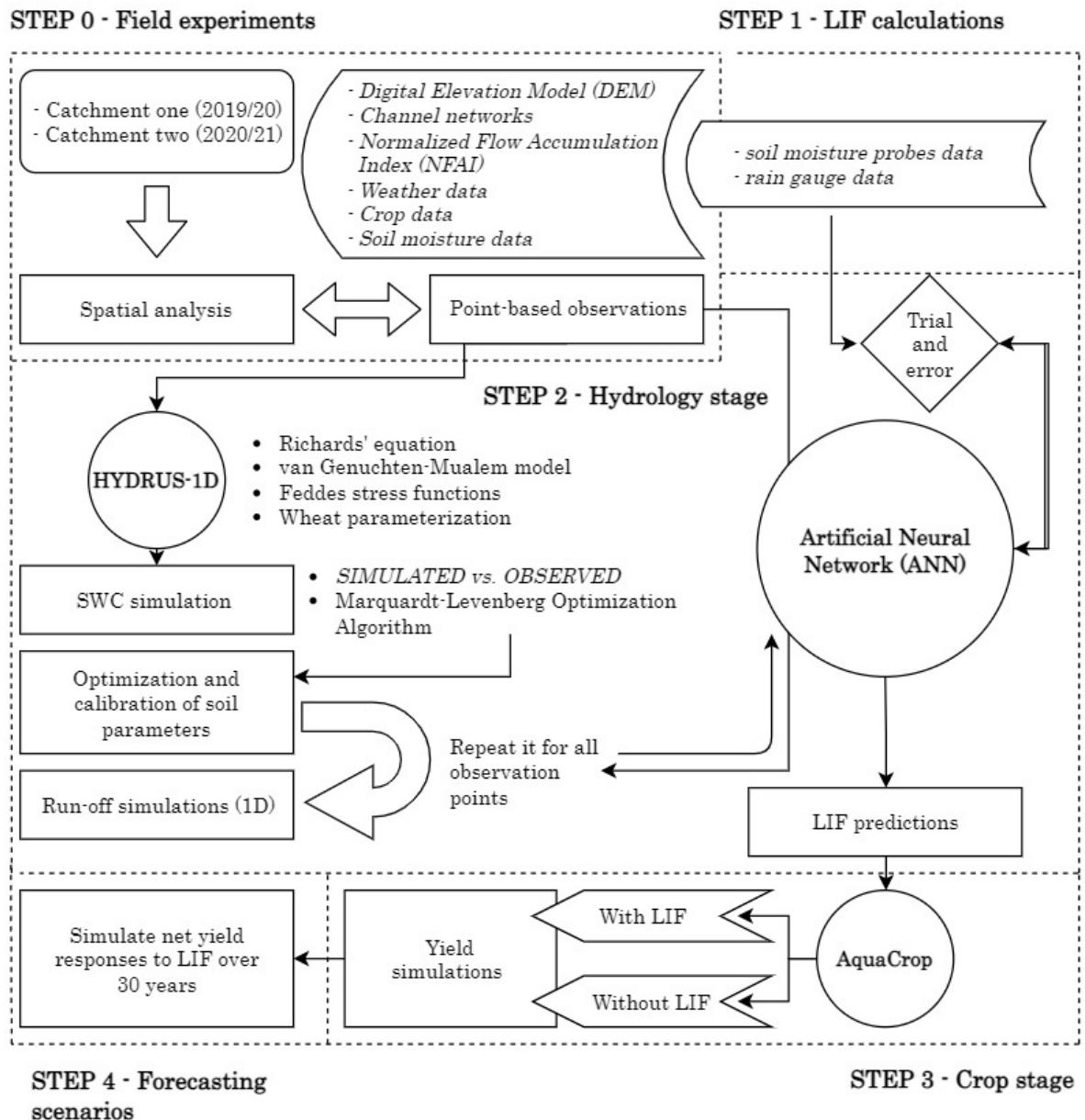
**Figure 4.1:** Visual symptoms of early crop senescence apparently caused by spatial water variations in Córdoba, Spain. The higher elevation zones show yellowing patterns due to lower water availability, which limits crop yield (Sadras et al., 2016). Photo credits: T. R. Tenreiro.



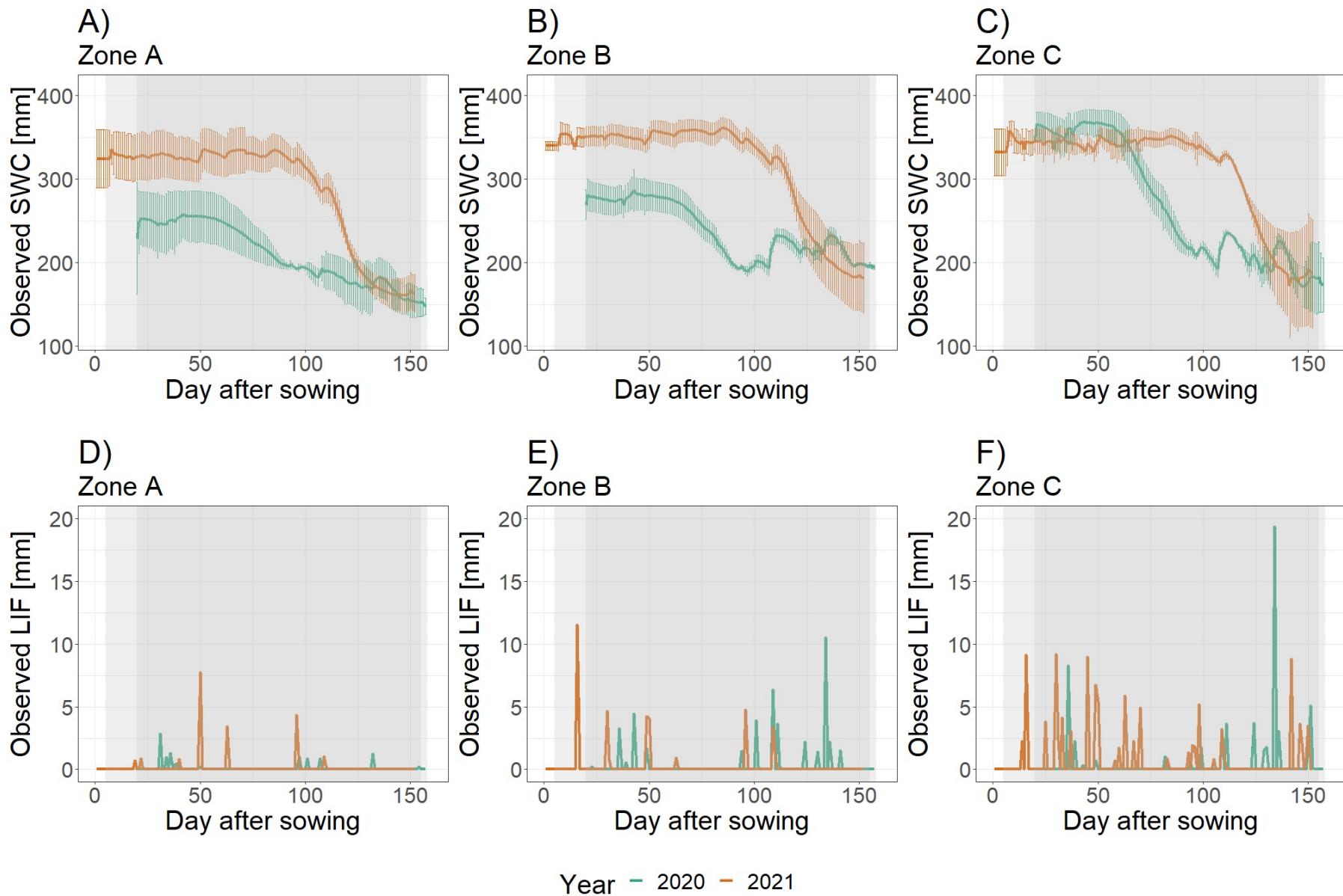
**Figure 4.2:** Maps of the experimental catchments: the elevation maps with contour lines of catchment one and two, respectively (A and B); the Normalized Flow Accumulation Index (NFAI) rasterized with 5m spatial resolution for catchment one and two, respectively (C and D). Sampling zones (A, B, C) and sampling points ( $A_1, A_2, A_3, B_1, B_2, B_3, C_1, C_2, C_3$ ) are represented by solid black lines and purple circles, respectively.



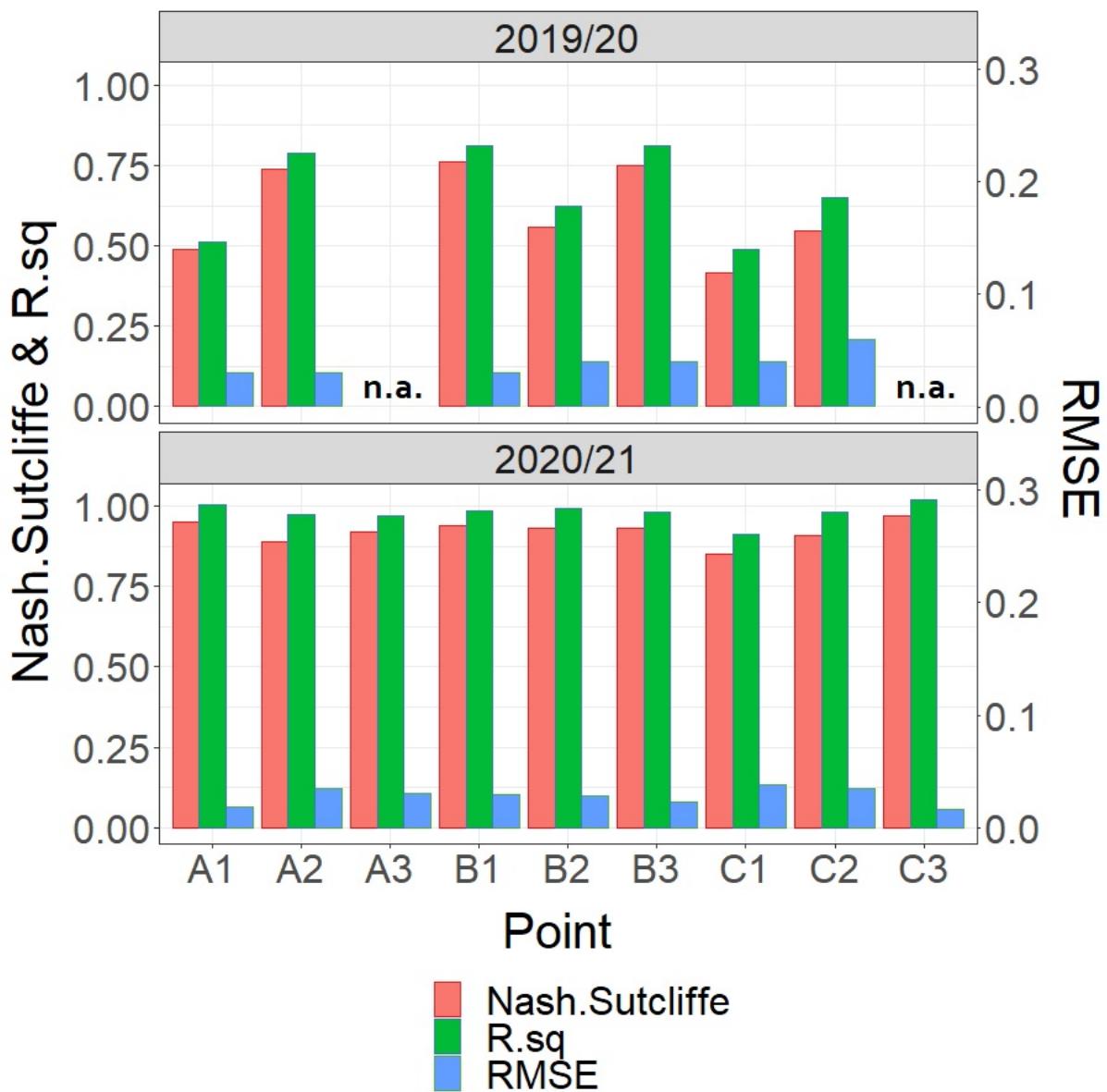
**Figure 4.3:** Daily mean values for (A) rainfall (dashed lines represent cumulative rainfall computed from the 1st of October, values shown by right vertical axis), (B) temperature and (C) reference evapotranspiration (ETo), respectively expressed in mm, degree Celsius and mm. Lines represent season time-series expressed in terms of days after sowing (DAS). Grey areas represent the monitoring time window of each observation year (i.e., from DAS 20 to 157 in 2019/20 and from DAS 5 to 158 in 2020/21). Daily mean temperature values are represented by the heavy solid lines. Maximum and minimum temperature values are respectively represented by a solid and a dashed skinny line. Each season values are shown in different colors (2019/20 in green and 2020/21 in brown).



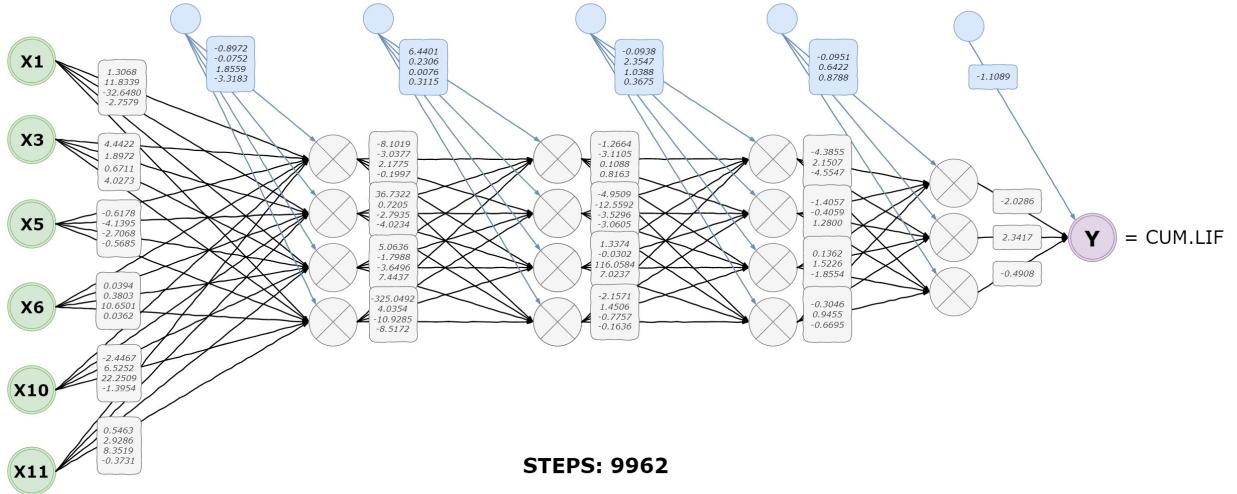
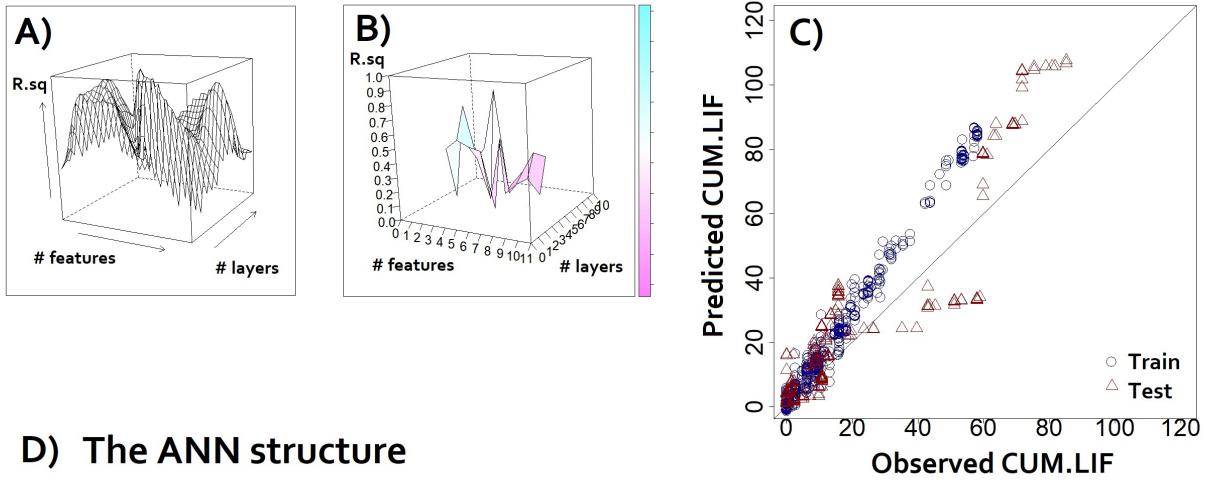
**Figure 4.4:** Sketch of the methodological design: Two field experiments were conducted, each sampling point was defined according to a spatial analysis (step 0). A daily calendar of lateral water inflow (LIF) was calculated based on field observations (step 1) and simulations of LIF were conducted for 30 years through a hydrologic modelling approach (step 2). LIF predictions were assimilated into the crop modelling stage (step 3) and the net yield responses were simulated (step 4). Dashed lines delineate different methodological stages, rounded parallelograms indicate experimental sites and data, solid line rectangles indicate different sub-steps, solid line circles represent simulation tools, losengete and arrows indicate conditional steps. More information on the simulation settings of HYDRUS-1D and AquaCrop is respectively provided by Šimůnek et al. (2018) and Steduto et al. (2009). Additional details related to the use of artificial neural networks (ANN) for hydrological modelling are provided by Maier & Dandy (2000).



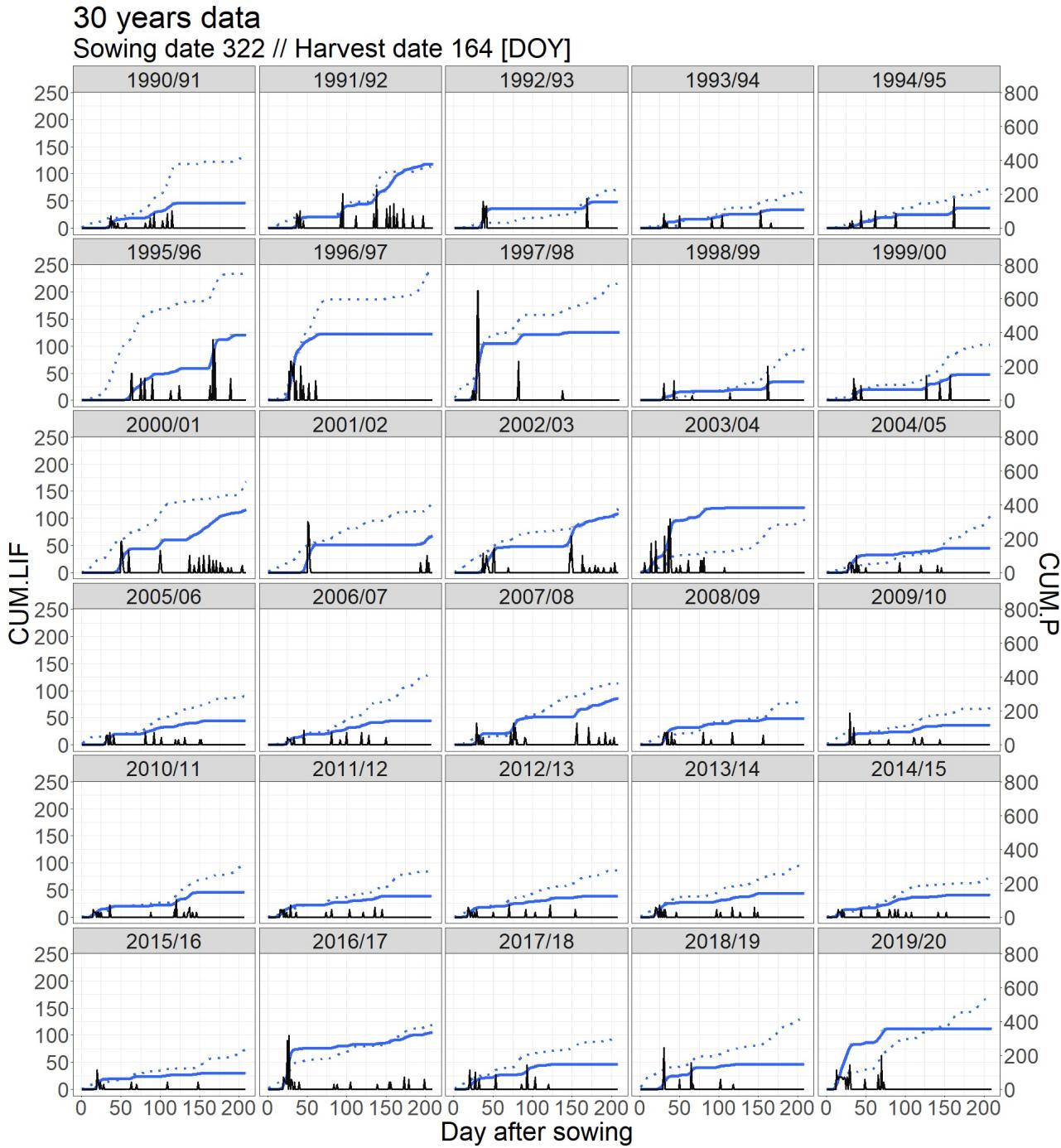
**Figure 4.5:** Corrected soil water content ( $SWC_c$ ) for each sampling zone (A-C), error bars are shown for daily  $SWC_c$  values. Daily lateral inflows (LIF) per sampling zone (inferred from probes data). Values are expressed in mm.



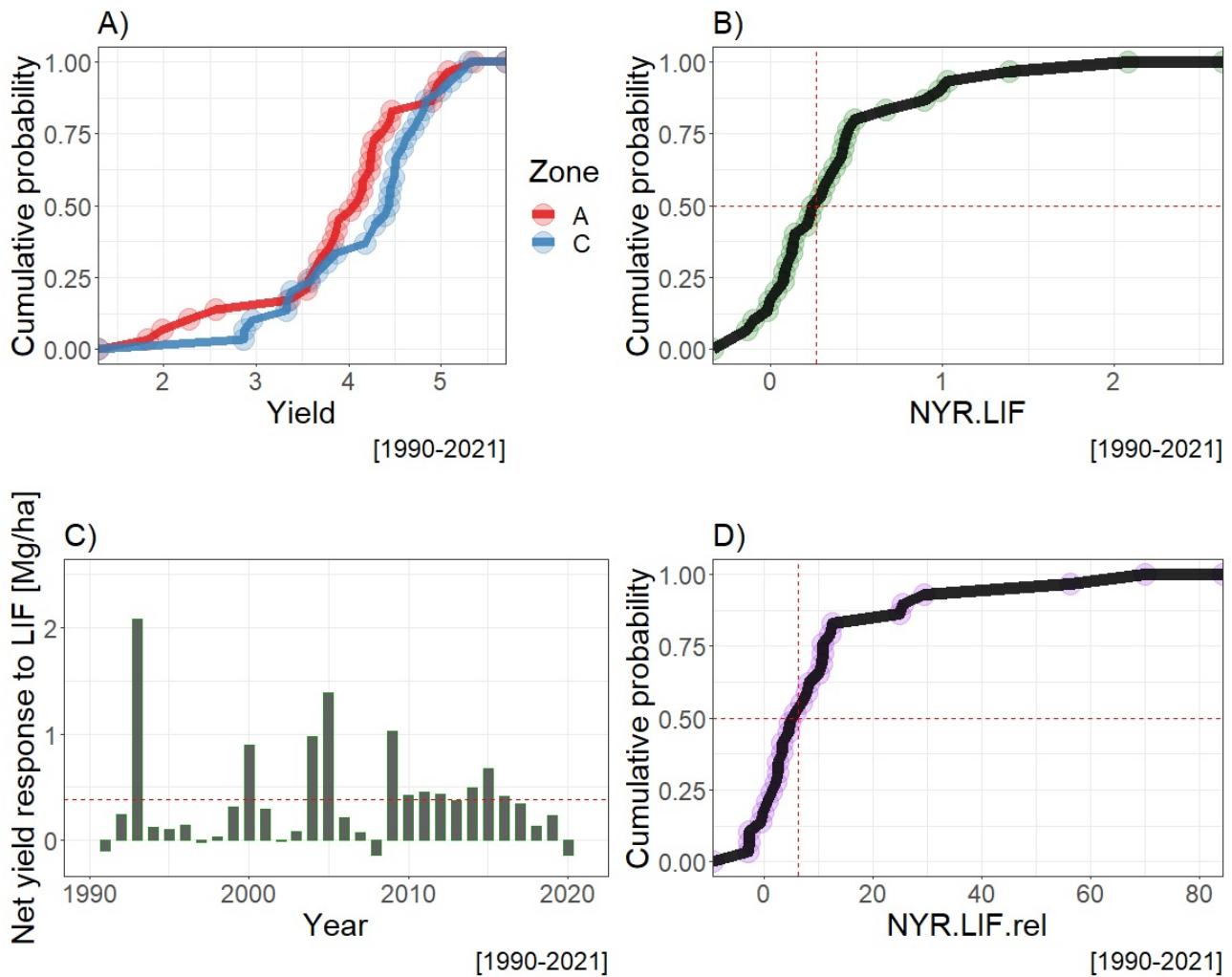
**Figure 4.6:** HYDRUS-1D simulations performance according to the means of measured soil water content and the best fitted soil parameters. “n.a.” indicates non-applicable situations, which correspond to the sampling point A3 and C3 that were not considered in catchment one (Figure 4.2). Both the Nash-Sutcliffe and the  $R^2$  coefficients are indicated by the left axis, while the RMSE is indicated by the right axis. Mean  $R^2$  values ranged from 0.65 ( $\pm 0.13$ ) in 2019/20 to 0.93 ( $\pm 0.03$ ) in 2020/21 and RMSE ( $\text{cm}^3 \text{cm}^{-3}$ ) from 0.04 ( $\pm 0.01$ ) to 0.03 ( $\pm 0.01$ ), by the same order.



**Figure 4.7:** The feedforward Artificial Neural Network (ANN) used for forecasting lateral inflow (LIF) over 30 years: A) schematic representation of the  $R^2$  assessment based on trial-and-error procedure that was taken for model architecture delineation; B) three-dimensional representation of the  $R^2$  values obtained for each combination of model structure (the best performing ANN corresponded to an ANN composed by six different features and four hidden layers); C) ANN statistical evaluation, observed LIF plotted against predicted LIF (values are expressed in mm and represent cumulative LIF), blue dots correspond to the training set (2019/20) and red triangles to the testing set (2020/21). Both the training and the testing set fitted into a linear regression, respectively expressed as  $y = 1.432x + 0.528$  and  $y = 1.122x + 2.919$ , and subsequently with a  $R^2$  of 0.97 and 0.78, and a RMSE of 10.1 and 14.8 mm; D) the schematic representation of the best performing ANN, solved in 9962 steps: green circles represent input features, grey squares indicate the hidden layers' coefficients, each node is represented by a grey circle, the blue small circles represent the transfer functions and the blue squares indicate the linked weight factors. Selected features (i.e.,  $x_1$ ,  $x_3$ ,  $x_5$ ,  $x_6$ ,  $x_{10}$ ,  $x_{11}$ ) are shown in Table 4.4.



**Figure 4.8:** Cumulative lateral inflow ( $CUM.LIF$ ) and cumulative precipitation ( $CUM.P$ ), respectively represented by solid and dashed lines (expressed in mm). Facet-plots show the two forms of water supply to zone C in catchment two (shown in Figure 4.2). Values were computed in daily time-steps over a period of 30 crop seasons, from sowing to harvesting date.  $CUM.P$  values were computed from daily weather records for Cordoba [1990–2020], obtained from the same weather station introduced in Section 4.2.2.  $CUM.LIF$  values were forecasted by the ANN model shown in Figure 4.7. Vertical black bars represent daily LIF events (daily LIF magnitudes are multiplied by a factor of 4.5 for better visualization).



**Figure 4.9:** Main AquaCrop simulation outcomes: A) cumulative probability distribution curves for simulated  $Y_2$  yields ( $\text{Mg ha}^{-1}$ ) in the two significantly distinct zones (i.e., A and C in catchment two); B) cumulative probability of  $NYR_{LIF}$ , expressed in  $\text{Mg ha}^{-1}$ ; C) time-series of  $NYR_{LIF}$  expressed in  $\text{Mg ha}^{-1}$  over 30 years; D) cumulative probability of  $NYR_{LIF-rel}$ , expressed in %. Values correspond to zone C in catchment two for the period of 1990-2021. Dashed lines represent the median values, in plots B and D (i.e.,  $0.265 \text{ Mg ha}^{-1}$  and  $6.3\%$ , respectively), and the mean  $NYR_{LIF}$  in plot C (i.e.,  $0.383 \text{ Mg ha}^{-1}$ ).

## TABLES - Chapter 4

**Table 4.1:** Main geomorphological attributes within each sampling zone (mean values and standard deviations). The standard deviations are presented in parentheses. Codes: ECa = soil apparent electrical conductivity, %Clay = percentage of clay content, %Sand = percentage of sand content, NFAI = “Normalized flow accumulation index” (Tarboton et al., 1991; Jenson & Domingue., 1988). The amount of chosen sampling points was supported by Chanzy et al. (1998). %Clay, %Sand and bulk density were averaged for the three zones in catchment-2 according to farm records.

Parameter	Units	Catchment-1			Catchment-2		
		A	B	C	A	B	C
Zone							
Sampling points [N]		2	3	2	3	3	3
Elevation	m (amsl.)	187 (2.4)	184 (5.1)	168 (2.3)	147 (0.5)	141 (0.3)	139 (0.3)
ECa	dS m <sup>-1</sup>	0.31 (0.08)	0.28 (0.09)	0.48 (0.06)	-	-	-
%Clay	%	45 (3.4)	42 (3.3)	50 (3.4)		44 (2.8)	
%Sand	%	18 (2.6)	22 (2.9)	15 (2.8)		22 (3.2)	
Bulk density	g cm <sup>-3</sup>	1.78 (0.06)	1.81 (0.04)	1.74 (0.05)		1.66 (0.05)	
NFAI	[0;1]	0.09 (0.021)	0.12 (0.082)	0.72 (0.218)	0.07 (0.005)	0.42 (0.002)	0.98 (0.011)

**Table 4.2:** Crop data used for the parameterization of the HYDRUS and the AquaCrop models.  $CC_{MAX}$  is maximum green canopy cover (Steduto et al., 2009), used for parameterization of the soil cover partitioning method, as described in Tenreiro et al. (2020). Mean sowing rate was 180 kg ha<sup>-1</sup> for all trials. The standard deviations are presented in parentheses.

Data	Parameter	Zone [units]	Catchment-1			Catchment-2			Mean
			A	B	C	A	B	C	
Sowing date		date		13-Dec			18-Nov		1-Dec
Crop emergence		DAS		10			8		9
Plant density		plants m <sup>-2</sup>	200 (21.8)	280 (19.2)	225 (16.3)	250 (18.2)	225 (16.44)	225 (10.4)	236 (27.4)
$CC_{MAX}$		%	80 (3.5)	82 (4.6)	84 (2.1)	90 (1.7)	95 (0.6)	92 (0.5)	87 (5.9)
Root growth		cm day <sup>-1</sup>		0.7 (0.1)			0.9 (0.1)		0.8 (0.1)
Vegetative stage		days		120			120		120
Anthesis stage		days		10			14		12
Reproductive stage		days		58			76		67
Senescence duration		days		20			40		30
Crop maturity		date		8-Jun			1-Jun		4-Jun
Harvest date		date		10-Jun			8-Jun		9-Jun

**Table 4.3:** The parameter intervals used for model optimization. Values were adjusted during the optimization process for each individual point (Figure 4.2).

Parameter	units	Min	Max
residual vol. water content ( $\theta_r$ )	$\text{cm}^3 \text{ cm}^{-3}$	0	0.2
saturated vol. water content ( $\theta_s$ )	$\text{cm}^3 \text{ cm}^{-3}$	0.3	0.5
van Genuchten's parameter ( $\alpha$ )	$\text{cm}^{-1}$	0.0001	0.1
van Genuchten's parameter ( $n$ )	-	1.2	1.5
saturated hydraulic conductivity ( $K_{SAT}$ )	$\text{mm day}^{-1}$	2	100

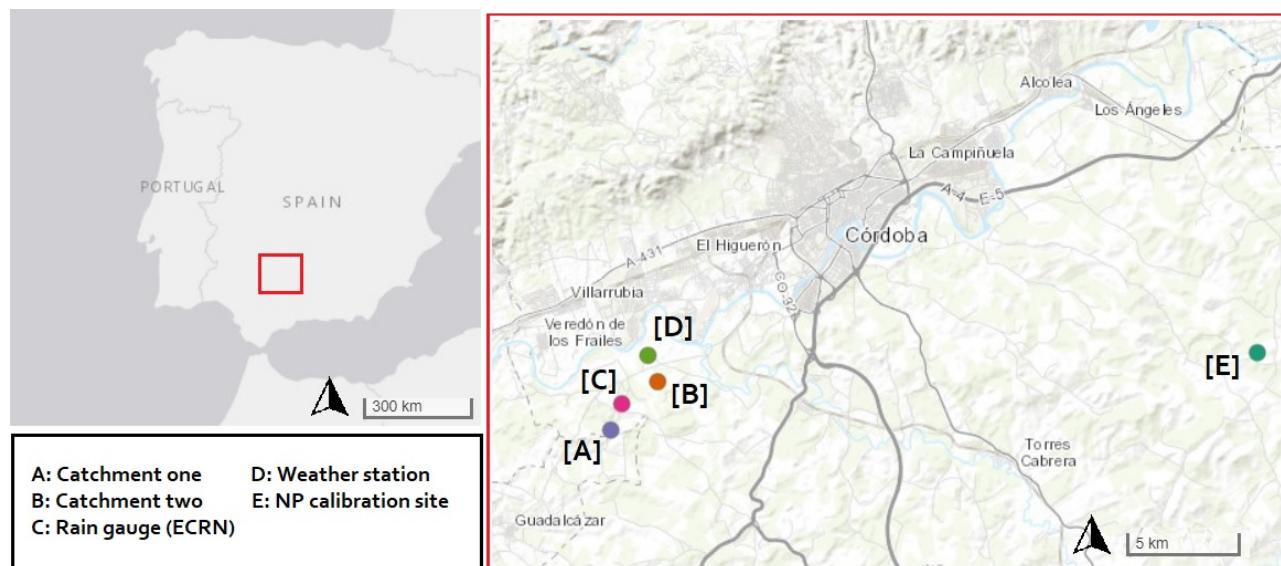
**Table 4.4:** Artificial neural network (ANN) potential predictors assessed through principal component analysis (PCA) and trial-and-error procedure (Roadknight et al., 1997). Main PCA outcomes are shown in Appendix-A3. More information regarding the PCA is provided in Supplementary material.

id	Input feature	Estimation procedure/description	Contribution to LIF
x1	Cumulative surface run-off [mm] at location $(x, y)$	Simulated with HYDRUS-1D through multiple iterations optimized by means of the measured soil water content	High
x2	Day after run-off at location $(x, y)$	Computed with R-studio as a function of x1 calendarization	Low
x3	NFAI [0;1] at location $(x, y)$	Estimated with SAGA-GIS (2.3.2) and computed with R-studio (Lovelace et al., 2019)	High
x4	Slope [%] at site $i$	Estimated with SAGA-GIS (2.3.2) and computed with R-studio (Lovelace et al., 2019)	Medium
x5	Cumulative surface run-off at upslope contributing points [mm]	The daily median of x1 values, simulated with HYDRUS-1D for upslope hydrologically contributing points	High
x6	Canopy cover [%]	According to Tenreiro et al. (2021)	High
x7	Surface saturated hydraulic conductivity [ $\text{mm day}^{-1}$ ]	Optimized $K_{SAT}$ at location $(x, y)$ with HYDRUS-1D for surface layer (0-10 cm depth)	Medium
x8	Day after run-off	Computed with R-studio as a function of x5 calendarization	Medium
x9	Day after precipitation event	Computed with R-studio as a function of x11 calendarization	Low
x10	Saturated hydraulic conductivity (mean) [ $\text{mm day}^{-1}$ ]	The average between surface and sub-surface $K_{SAT}$ (optimized with HYDRUS-1D)	High
x11	Cumulative daily precipitation [mm]	Field measured with a rain gauge (section 2.2.2)	High
x12	Boolean 'satunsat' parameter	Boolean parameter set to define days of saturated vs. unsaturated conditions in the vadose zone	Medium

**Table 4.5:** Mean soil water content (SWC), mean daily lateral inflow ( $LIF_n$ ) and mean cumulative LIF ( $CUM.LIF_n$ ) over season, for the two years of field data. All values are expressed in mm. The standard deviations are presented in parentheses. Mean values followed by a common letter are not significantly different according to the HSD-test conducted at the 5% level of significance ( $p - value < 0.05$ ).

Zone	Catchment-1 (2019/20)	Catchment-2 (2020/21)	
SWC [mm]	A B C	209.36 (39.6)c 237.36 (34.4)b 269.37 (63.1)a	289.19 (65.1)b 316.52 (61.2)a 310.87 (60.4)a
$LIF_n$ [mm]	A B C	0.08 (0.46)b 0.31 (1.52)ab 0.42 (2.09)a	0.13 (0.8)b 0.23 (1.6)b 0.71 (2.7)a
$CUM.LIF_n$ [mm]	A B C	6.98 (3.9)b 17.88 (15.7)a 19.76 (16.9)a	11.24 (11.2)c 23.94 (17.4)b 60.28 (42.1)a

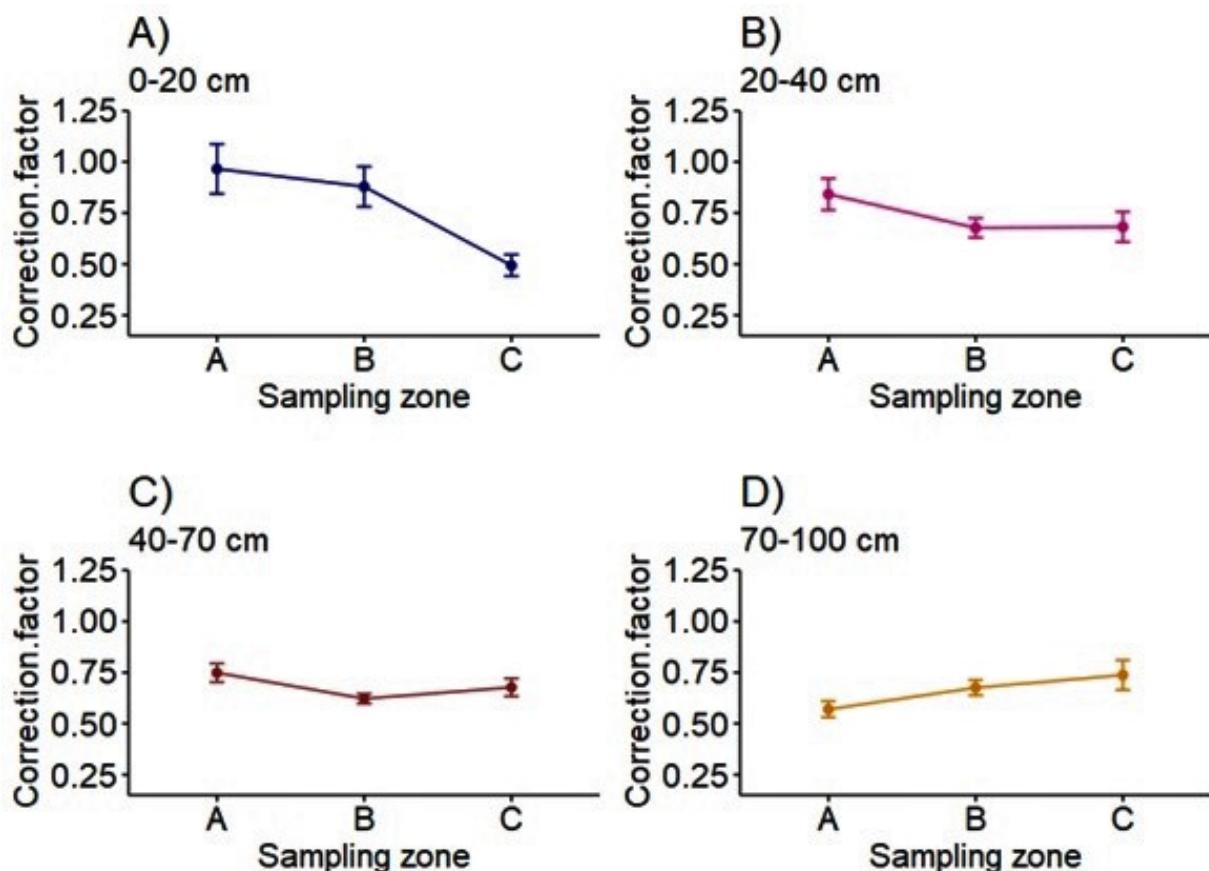
## APPENDIX - Chapter 4



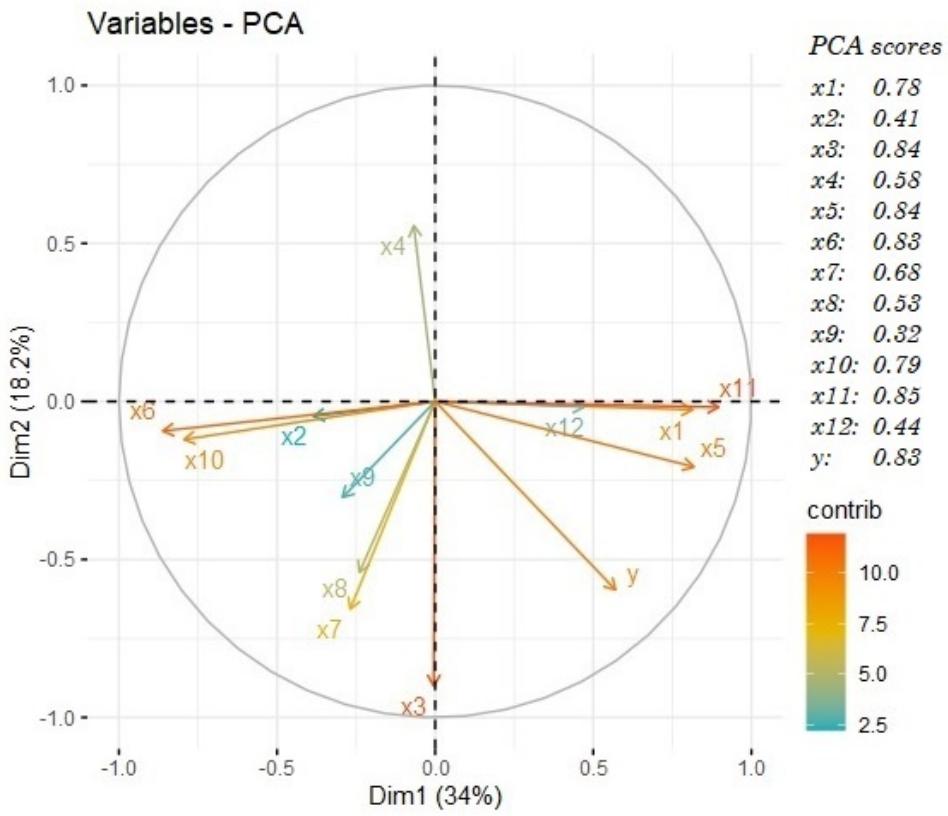
**Figure A1.** Map of the study locations: catchment one (A), catchment two (B), rain gauge ECRN-100 ZENTRA system (C), weather station (D), neutron probe calibration site (E). Further information on the neutron probe calibration site is provided by Soriano et al. (2018).

**Table A.** The parameters of the neutron probe (NP) calibration functions ( $y = a[\frac{x}{std}] - b$ ).  $y$  is volumetric SWC, expressed in %,  $x$  is probe measured value (unitless) and 'std' is the standard correction value. More information on the neutron probe calibration site is provided by Soriano et al. (2018).

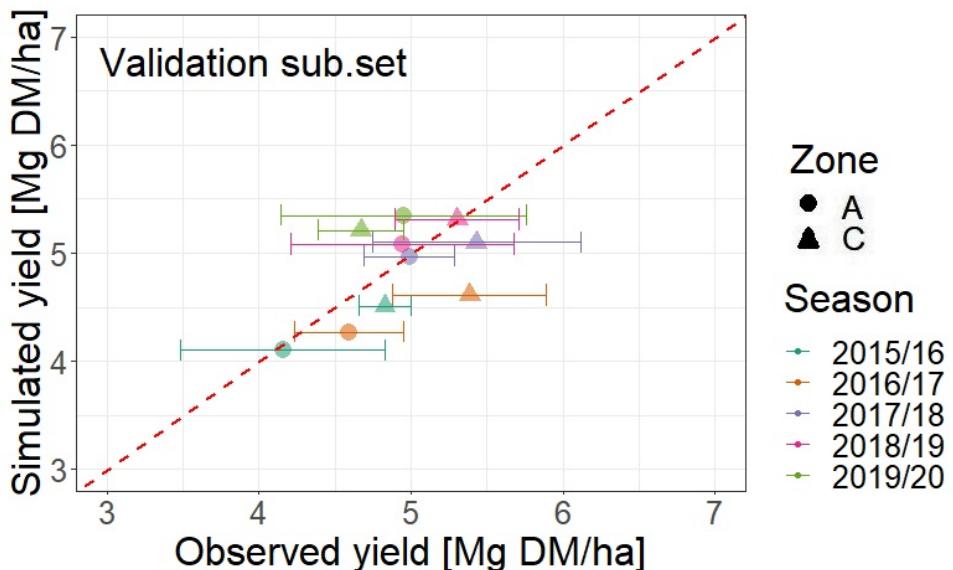
Depth	a	b	std	$R^2$
0-15 cm	25.416	0.264	7830	0.82
15-30 cm	21.974	7.381	7830	0.98
30-90 cm	27.210	16.897	7434	0.96



**Figure A2.** The mean correction factor ( $\Omega_z$ ) plotted for each probe location and sensor depth in 2019/20, at the 95% confidence level.  $\Omega_z$  was computed as the mean ratio between the SWC measurements with the NP and the capacitance probes for depth  $z$  (Equation 4.2).

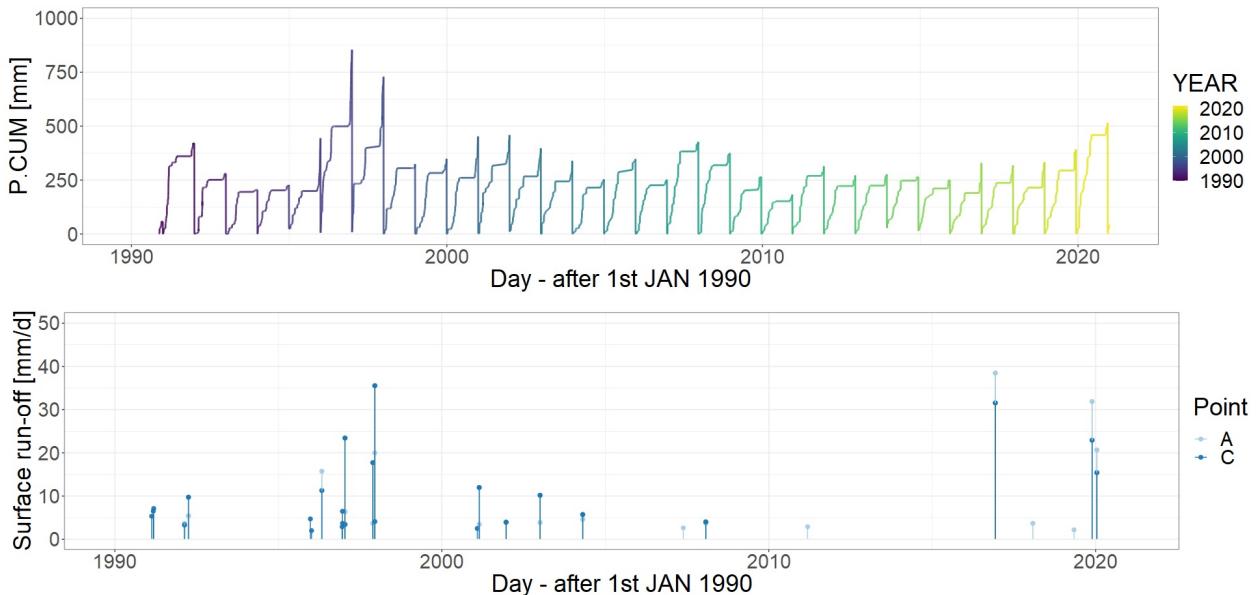


**Figure A3.** The principal component analysis (PCA) plot. Vectors are colored according to the contribution degree to LIF. PCA scores are equal to the module of each vector, indicating the weight associated with the combination of the two principal components (i.e., Dim 1 and Dim2, respectively explaining 34 and 18.2% of LIF variance). Variables are defined in Table 4.4.

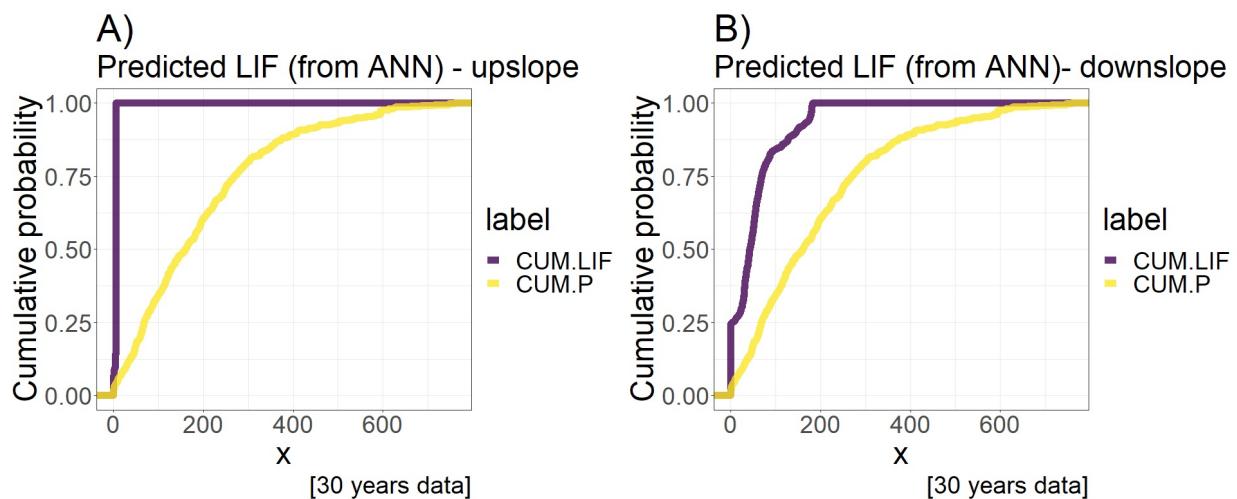


**Figure A4.** Simulated vs. observed yields (obtained from historical yield maps). Units are expressed in  $\text{Mg GY ha}^{-1}$ . Circles and triangles represent yields in zone A and C, respectively. Simulated yields correspond to the  $Y_2$  scenario. RMSE,  $R^2$  and Willmott  $d$  index are respectively equal to  $0.374$  ( $\text{Mg GY ha}^{-1}$ ),  $0.35$ ,  $0.76$ . The horizontal bars indicate the error associated with the process of yield mapping (i.e.,  $172\text{--}809 \text{ kg GY ha}^{-1}$ ).

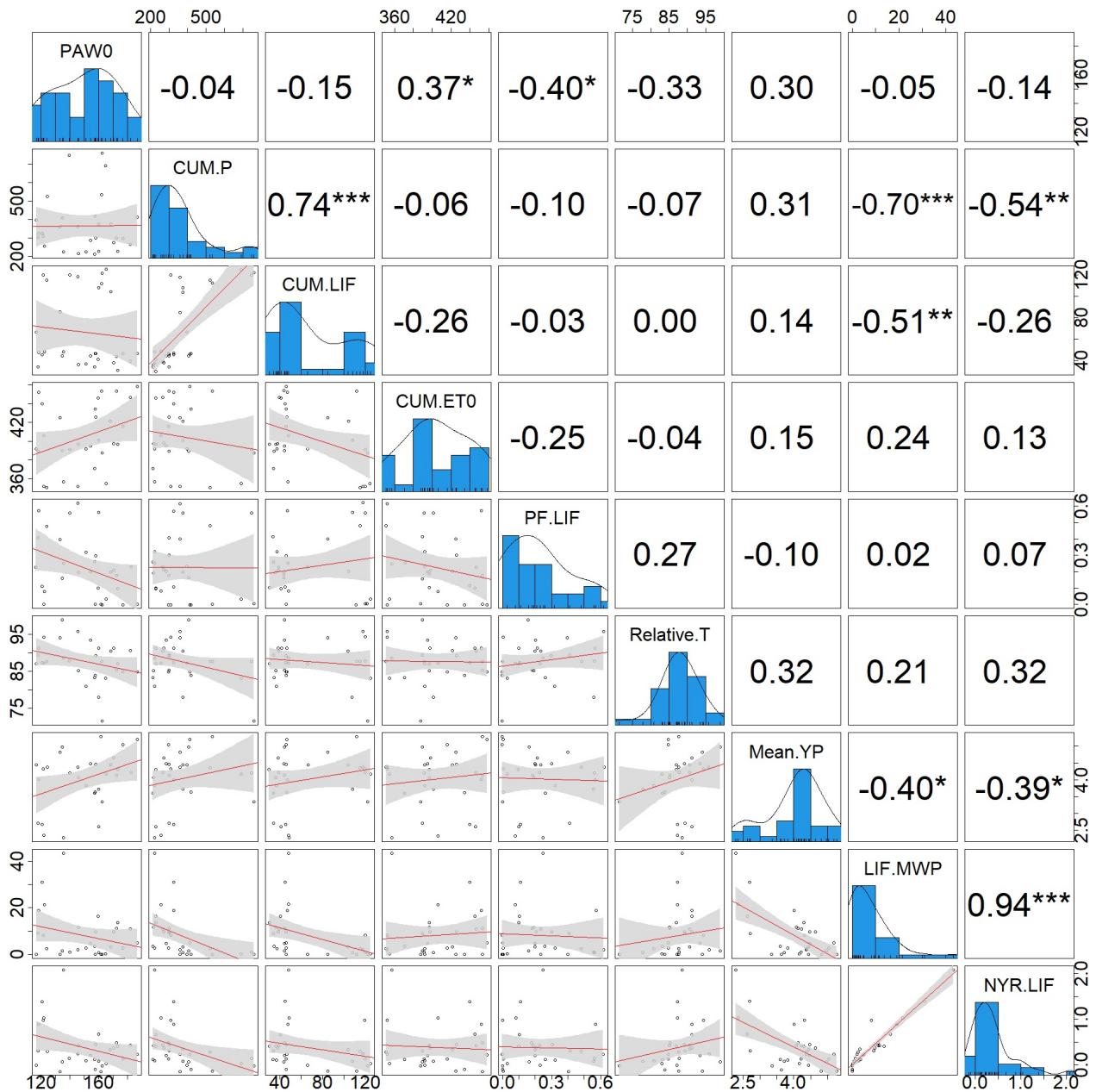
30 years data



**Figure A5.** Cumulative precipitation ( $CUM.P$ ) and HYDRUS run-off simulations over 30 years [1990–2020].  $CUM.P$  is expressed in mm. Blue bars represent daily surface run-off values, expressed in  $\text{mm day}^{-1}$ .



**Figure A6.** Cumulative probability distribution curves of  $CUM.LIF$  and  $CUM.P$  over 30 years. Curves shown for zone A (A) and zone C (B), representing both cumulative precipitation and cumulative LIF, values expressed in  $\text{mm day}^{-1}$ .



**Figure A7.** Spearman's correlation pairs-plot. Codes:  $PAW_0$  = initial plant available water (at sowing date);  $CUM.P$  = season cumulative precipitation;  $CUM.LIF$  = season cumulative lateral inflow (LIF);  $CUM.ETO$  = season cumulative evapotranspiration ( $ET_0$ );  $PF.LIF$  = post-flowering LIF, corresponding to the fraction of  $CUM.LIF$  taking place at post-flowering stages; 'Relative.T' = mean relative crop transpiration (the simple average of the relative crop transpiration between zone A and C, expressed in %); Mean YP = mean yield potential (the simple average of simulated yield between zone A and C, expressed in  $Mg\text{ GY}^{-1}\text{ ha}^{-1}$ );  $LIF.MWP$  = LIF marginal water productivity (expressed in  $kg\text{ GY}^{-1}\text{ mm}^{-1}$ );  $NYR.LIF$  = Net yield response to LIF (expressed in  $Mg\text{ GY}^{-1}\text{ ha}^{-1}$ );  $PAW_0$ ,  $CUM.P$ ,  $CUM.LIF$  and  $CUM.ETO$  are expressed in mm. 'Relative.T' is estimated at the season average of daily crop actual transpiration divided by daily potential transpiration. Simulation files are provided in Supplementary materials. Season is defined from sowing to harvesting date. Input values are synthesized in Appendix-Table B. Significant correlations at the 5% level of significance are highlighted with the symbol '\*'. Significance codes: '\*\*\*' 0.1%, '\*\*' 1%, '\*' 5%.

**Table B.** AquaCrop simulation outcomes. Codes:  $PAW_0$  = Initial plant available water (at sowing date);  $CUM.ETo$  = Season cumulative  $ETo$  (from sowing to harvesting date);  $CUM.P$  = Season cumulative precipitation;  $CUM.LIF$  = Season cumulative lateral inflow (LIF); Post-flowering LIF = The fraction of season LIF taking place at post-flowering stages (expressed in %); Relative T [Zone A] = Mean relative crop transpiration in zone A (or in the absence of LIF); Relative T [Zone C] = Mean relative crop transpiration in zone C (including LIF); Yield.1 [Zone A] = Yield in scenario 1 (in the absence of LIF); Yield.2 [Zone C] = Yield in scenario 2 (including LIF);  $NYR.LIF$  = Net yield response to LIF;  $NYR.LIF_{rel}$  = Net yield response to LIF in relative terms;  $LIF.MWP$  = LIF marginal water productivity (expressed in  $\text{kg GY ha}^{-1} \text{mm}^{-1}$ ). Relative T is estimated as the season average of daily crop actual transpiration divided by potential transpiration. Simulation files are provided in Supplementary materials.

Season	$PAW_0$	$CUM.ETo$	$CUM.P$	$CUM.LIF$	Post-flowering LIF	Relative T [Zone A]	Relative T [Zone C]	Yield.1 [Zone A]	Yield.2 [Zone C]	$NYR.LIF$	$NYR.LIF_{rel}$	$LIF.MWP$
	[ $\text{mm}$ ]	[ $\text{mm}$ ]	[ $\text{mm}$ ]	[ $\text{mm}$ ]	[%]	[%]	[%]	[ $\text{Mg GY ha}^{-1}$ ]	[ $\text{Mg GY ha}^{-1}$ ]	[ $\text{Mg GY ha}^{-1}$ ]	[%]	[ $\text{kg GY ha}^{-1} \text{mm}^{-1}$ ]
1990/91	157.5	434.4	419.1	46	0	84.2	83.4	3.68	3.58	-0.10	-3	-
1991/92	145.5	351.8	361.4	118	62	93.4	98.2	4.25	4.49	0.24	6	2.03
1992/93	135.4	356.8	224.3	48	25	87.0	90.0	1.24	3.32	2.08	168	43.33
1993/94	157.3	370.8	212.8	34	26	82.6	83.8	3.54	3.66	0.12	3	3.53
1994/95	151.2	391.8	223.8	37	32	78.3	83.4	3.77	3.87	0.10	3	2.70
1995/96	139.7	391.3	747.0	120	56	87.0	88.4	4.14	4.28	0.14	3	1.17
1996/97	162.4	387.5	759.5	122	0	72.5	70.4	3.36	3.34	-0.02	-1	-
1997/98	165.0	354.8	689.6	125	3	83.0	83.1	4.15	4.18	0.03	1	0.24
1998/99	117.9	389.6	301.6	34	41	88.6	90.1	2.56	2.87	0.31	12	9.12
1999/00	118.2	452.1	326.2	47	57	89.5	92.9	3.56	4.45	0.89	25	18.94
2000/01	124.0	350.7	523.9	115	48	85.7	89.6	4.22	4.51	0.29	7	2.52
2001/02	116.4	392.0	398.2	66	23	87.1	87.0	4.44	4.43	-0.01	0	-
2002/03	160.1	401.3	376.2	108	56	77.7	78.1	3.68	3.76	0.08	2	0.74
2003/04	121.6	352.1	310.0	120	0	91.1	97.4	3.83	4.81	0.98	26	8.17
2004/05	120.6	439.8	325.5	45	11	85.5	89.0	1.99	3.38	1.39	70	30.89
2005/06	154.3	394.9	286.4	44	20	91.6	95.2	4.37	4.58	0.21	5	4.77
2006/07	134.6	389.8	407.8	44	16	98.4	99.4	4.93	5.00	0.07	1	1.59
2007/08	133.1	424.9	359.8	85	40	94.9	93.2	4.89	4.75	-0.14	-3	-
2008/09	122.0	406.1	257.7	48	8	86.1	88.3	1.83	2.86	1.03	56	21.46
2009/10	145.9	397.1	215.8	36	19	82.1	88.0	3.86	4.28	0.42	11	11.67
2010/11	157.9	425.6	300.0	46	30	87.4	90.9	4.24	4.69	0.45	11	9.78
2011/12	182.3	446.6	266.9	39	23	88.9	92.1	4.01	4.44	0.43	11	11.03
2012/13	170.4	397.8	273.8	38	21	89.6	92.9	4.46	4.83	0.37	8	9.74
2013/14	172.5	453.6	299.8	44	18	83.5	87.0	3.89	4.38	0.49	13	11.14
2014/15	159.8	445.7	228.6	41	10	90.2	92.3	2.28	2.95	0.67	29	16.34
2015/16	173.2	396.7	228.9	30	10	85.8	88.1	4.10	4.51	0.41	10	13.67
2016/17	168.3	421.4	375.4	104	20	89.8	91.2	4.27	4.61	0.34	8	3.27
2017/18	176.9	416.0	297.1	46	0	84.0	85.6	4.97	5.10	0.13	3	2.83
2018/19	186.9	458.4	412.3	46	0	83.7	85.8	5.08	5.31	0.23	5	5.00
2019/20	161.8	453.2	535.2	112	0	88.6	86.5	5.35	5.21	-0.14	-3	-

# Chapter 5

## Opportunities for variable application rate of nitrogen under spatial water variations in rainfed wheat systems – an economic analysis

This chapter has been under review as:

**Tenreiro, T. R.**, Avillez, F., Gómez, J. A., Penteado, M., Coelho, J. C., Fereres, E. (2022). Opportunities for variable application rate of nitrogen under spatial water variations in rainfed wheat systems – an economic analysis. *Precision Agriculture*.

### Abstract

In fields of undulating topography, where rainfed crops experience different degrees of water stress caused by spatial water variations, yields vary spatially within the same field, thus offering opportunities for variable application rates (VAR) of nitrogen fertilizer. This study assessed the spatial variations of yield gaps in rainfed wheat caused by lateral flows from high to low points, grown in Córdoba, Spain, over six consecutive seasons (2016–2021). The economic implications associated with multiple scenarios of VAR adoption were explored through a case study and recommendations were proposed. Both farm size (i.e., annual sown area) and topographic structure impacted the dynamics of investment returns. Under current policy-prices conditions, VAR adoption would have an economic advantage in farms similar to that of the case study with an annual sown area greater than  $567 \text{ ha year}^{-1}$ . Nevertheless, current trends on energy prices, transportation costs and impacts on both cereal prices and fertilizers costs enhance the viability of VAR adoption for a wider population of farm types. The profitability of adopting VAR improves under such scenarios and, in the absence of additional policy support, the minimum area for adoption of VAR decreases to a range of  $68\text{--}177 \text{ ha year}^{-1}$ . The combination of price increases with the introduction of an additional subsidy on crop area could substantially lower the adoption threshold down to  $46 \text{ ha year}^{-1}$ , turning VAR technology economically viable for a much wider population of farmers.

## 5.1 Introduction

Rainfed agriculture plays a determinant role in food production worldwide as it accounts for more than 75% of global cropped area, being responsible for more than 60% of the global cereals' production (Cassman et al., 2003; Connor & Mínguez, 2012). Sustaining food production by rainfed crops in the years ahead will require productivity gains in resource use (Fischer & Connor, 2018). Specific challenges include estimating the magnitude and thus the value of yield gaps, identifying limiting factors, and implementing profitable and sustainable strategies. Closing yield gaps in rainfed farming, while improving resource use efficiency, is also expected to minimize the expansion of arable land and emissions while cutting on other undesirable externalities (Cassman et al., 2003; Snyder et al., 2009). Over the last decades, new technologies have been evolving within the context of rainfed farming (e.g., precision agriculture), aiming to increase the productivity of resources while offering substantial environmental benefits (Griffin & Shockley, 2018).

Spatial variations in rainfed crop yields are commonly observed worldwide (Florin et al., 2009; Griffin et al., 2020), which reveal, theoretically, opportunities for optimizing resource use (e.g., fertilizer) through precision agriculture (PA). A central concept of PA is the spatially variable application rate (VAR) of fertilizer according to the intra-plot variations of yield levels (Basso et al., 2013; Bullock & Lowenberg-DeBoer, 2007; Robertson et al., 2008; Pedersen et al., 2021). In fields of undulating topography, where rainfed crops experience different degrees of water stress due to spatial water variations caused by lateral flows, yields vary spatially within the same field (Halvorson & Doll, 1991; Tenreiro et al., 2022). These variations may imply different nutrient requirements and application rates over fields (Nielsen & Halvorson, 1991; Sadras, 2002). However, although the theoretical reasons for adopting VAR technology in rainfed systems are well accepted, its adoption has not been widespread (Basso et al., 2013; Lowenberg-DeBoer & Erickson, 2019; Robertson et al., 2012).

Considerable attention has been devoted to VAR adoption constraints

(Robertson et al., 2008; Schimmelpfennig et al., 2016; Welsh et al., 2003). Sceptics argue that there is a lack of evidence of relative advantage on VAR adoption. The principal issues are related to technical obstacles, complexity on the use of equipment and software, problems with data availability and access, operational incompatibility and ambiguity regarding financial benefits (Pathak et al., 2019; Robertson et al., 2007). Nevertheless, current trends on energy prices, transportation costs and impacts on both cereal prices and fertilizers costs (Chowdhury et al., 2021; Deloitte, 2021; EUC, 2021; Glauber & Laborde, 2022; Khalfaoui et al., 2021; FAO, 2021; USDA, 2021) may alter the trade-offs involved in the adoption of VAR in rainfed systems. Generally, there are other advantages attributed to the increase of fertilizer use efficiency that go beyond economic reasons which should not be ignored. Adequate use of fertilizer is not only needed to sustain yields and to increase the efficiency of water and energy use, but it also contributes to minimize the overall emissions over the supply chain (Snyder et al., 2009).

Understanding and capitalizing on yield variability is one of the main objectives of PA and the yield gap (YG) is an important concept to be used in this context as it provides a benchmark to explore yield variations (Cassman, 1999). YG is defined as the difference between the potential ( $Y_p$ ), or water-limited yield ( $Y_w$ ) in the case of rainfed cropping systems, and the actual yield ( $Y_a$ ) achieved by farmers (Loomis & Connor, 1992; Fischer, 2015). However, most studies on YG analysis ignore intra-plot variability (Fischer et al., 2014; Lobell et al., 2009; Schils et al., 2018), which is partially due to data availability constraints (Beza et al., 2017) and to the limitations of crop models to simulate processes such as spatial water distribution (Tenreiro et al., 2020).

In fields of undulating topography where water flows from higher to lower elevation zones, the assessment of spatial variations of  $Y_w$  may be combined with yield mapping from combine harvesting to understand how YG varies within a field. Assuming that  $Y_a$  is achieved under current management practices (i.e., uniform nutrient application), the magnitude of YG and its variation over space (i.e., how site-specific  $Y_a$  relates to  $Y_w$  and how it varies from upslope to downslope zones) could be taken as an indication of the relative

advantage of VAR adoption. The larger the differences in YG among different zones, the greater would be the opportunity for adopting VAR (Robertson et al., 2008).

This study builds upon a methodological framework that assessed both Yw and Ya spatial variations in rainfed wheat fields (Tenreiro et al., 2022) which explored how YG varied within fields and from year to year. Intra-plot yield variation was modelled, and management units were determined according to the YG variation. Some economic implications associated with multiple scenarios of VAR adoption were explored through a case study and recommendations were proposed (Tenreiro et al., 2022). The present study performs a novel economic analysis on VAR adoption to benefit from the spatial water variations in rainfed wheat systems. The hypothesis that spatial variations in soil water supply may justify a variable fertilization rate over space is tested, focusing on the following objectives:

1. The identification of economic trade-offs between scale of cropping and opportunities for VAR adoption;
2. The assessment of return on investments at farm scale in different scenarios, where the spatial water variations over fields determine yield variability and scope for adopting VAR.

## 5.2 Materials & Methods

### 5.2.1 Experimental conditions and on-farm data collection

The experimental farm is located in Córdoba province, southern Spain ( $37.8^{\circ}$  N,  $4.8^{\circ}$  W, mean altitude 165 m amsl., mean precipitation 605 mm year $^{-1}$ , range of slope steepness 2–6%), and it has a total arable area of 320 ha (Figure 5.1-A). At farm level, crop rotations include autumn and spring sown crops (e.g., winter wheat, rapeseed, sunflower, chickpea), with winter wheat being grown 1–2 times every four seasons. This study is limited to wheat (*Triticum durum* L.), which is a major crop within the farming system with a relative area share of 0.25–0.3 year $^{-1}$ . The soils are predominantly of clay texture (40–50% clay

and 15–22% sand) with high bulk density (1.65–1.88 g cm<sup>3</sup>) and moderate depth (1.2–1.6 m). More details on soil properties and other geophysical conditions are provided in Tenreiro et al. (2022).

Depending on the year conditions and operational aspects at farm level, sowing dates vary from middle November to late December (Table 5.1). Wheat is sown under direct seeding at an average rate of 200 ( $\pm 20$ ) kg/ha. The current management system is spatially uniform, following a fertilization plan that consists of two applications per season (one pre-sowing and the other near the end of the vegetative stage). Mean N-P-K application rates to wheat are 165-60-0 units per ha year<sup>-1</sup>. Crop nutrient status is controlled every season through foliar analysis conducted at flowering. More information on crop management is provided in Table 5.1.

### 5.2.2 Lateral inflow (LIF) zones mapping

Spatial water variations caused by lateral flows imply yield variations. Different potential yielding zones, within the same field, were delineated according to lateral inflow (LIF) levels. For zone delineation, we used the TOPMODEL Topographic Index (TMTI) as described in detail by Beven et al. (2021) and expressed as following:

$$TMTI_{(x,y)} = \ln\left(\frac{a}{\tan\beta}\right) \quad (5.1)$$

where  $a$  represents the flow accumulation, defined as the upslope contributing area to that point  $(x, y)$ , and  $\beta$  represents the field slope angle in the same point. While the hydraulic gradient ' $\tan\beta$ ' is defined with respect to the plan distance, the flow accumulation rate is defined with respect to the plan unit area.  $a$  is expressed in m<sup>2</sup> and  $\beta$  in degrees.

We scaled up to the farm level the results of Tenreiro et al. (2022), who followed an experimental design that consisted of choosing three different sampling zones per field, where the TMTI values ranged from 0 to 6 with median values varying from 1–2 in ‘no-LIF zones’ up to 4–6 in ‘LIF zones’ (Appendix A1). The Yw simulation results in Tenreiro et al. (2022) are assumed to be dependent

on the scale of the experimental fields. To avoid an extrapolation of results to different scales from those in experimental conditions, all points with a TMTI falling out of the distribution range in experimental conditions (Appendix A1) were rejected (i.e., points with  $TMTI > 6$  were discarded). For the same reason, only those fields with a predominance of clay texture were selected (Figure 5.1). According to Tenreiro et al. (2022), only two zones were significantly different in terms of soil water content variations due to LIF (i.e., zones A and C, Appendix A1). Therefore, in the present study, only two different zones were considered:

1. 'LIF zones' (i.e., downslope zones with significant amount of water supplied through lateral flow taking place from upslope areas of the same field);
2. 'No-LIF zones' (i.e., upslope zones where no significant amount of water is supplied through lateral flow);

According to the standard deviation of each significantly different zone (Appendix A1), a TMTI cutting threshold value equal to 5 was considered and each location was classified as following:

1. IF  $TMTI_{(x,y)} < 5$  [point classified as 'no-LIF zone'];
2. ELSE [point classified as 'LIF zone'];

Using the previous criteria, a subset plot of 92 ha was selected (Figure 5.1-D and -E), from which 76 ha were classified as 'no-LIF zones' while the remaining 16 ha were considered to be 'LIF zones'. Therefore, the selected farm area has an overall share of 'LIF zones' equal to 17.4% (Figure 1-E).

### 5.2.3 Intra-plot spatial assessment of Yield Gaps

YG were spatially assessed at intra-plot level with a resolution of  $100 \text{ m}^2$  according to the following equation:

$$YG_{(x,y)} = Yw_{(x,y)} - Ya_{(x,y)} \quad (5.2)$$

where  $Y_G$ ,  $Y_w$  and  $Y_a$  represent the yield gap, the water-limited yield and the actual yield, respectively. The subscripts ( $x, y$ ) indicate the corresponding cell coordinates. Yields are expressed in terms of Mg DM (grain)  $\text{ha}^{-1}$ .

Two zones were considered according to lateral inflow (LIF) water supply (i.e., LIF and no-LIF zones) and  $Y_w$  was assumed to be constant within each zone.  $Y_w$  values (2016–2021 period) were obtained from Tenreiro et al. (2022) and  $Y_a$  was determined with the ‘New Holland’ Precision Land Manager (PLM) software, taking as an input the shapefiles generated by the combine harvester monitor (Fendt PLI C 5275). To obtain LIF zones’  $Y_w$ , lateral inflow was simulated as an additional water supply to the effective rainfall using the AquaCrop simulation model (Steduto et al., 2009) as described in Tenreiro et al. (2022). No-LIF zones’  $Y_w$  were simulated, assuming that water inflow takes place only due to vertical infiltration (Tenreiro et al. 2021).

#### 5.2.4 Quantifying the relative advantage of VAR adoption

Our assessment focused on the relation between intra- and inter-season economic benefits of VAR adoption, and on the total investment costs in relation with the amortization of equipment. We did not explore the concept of relative advantage (Robertson et al., 2012) in terms of income risk reduction or environmental benefits, as done by Swinton & Ahmad (1996).

In our perspective, exploring the relative advantage of VAR adoption must dedicate attention to the scope for improvement, not only regarding the target scenario but also the current management system (i.e., the baseline scenario). Our baseline scenario is characterized by uniform nutrient application rate thus its agronomic performance is captured by the determination of  $Y_a$ . Our target scenario implies VAR adoption and it is benchmarked by the simulated  $Y_w$  level. Therefore, the relative advantage is estimated as the difference between the target and the baseline scenario. This was expressed as the differential gross margin (DCM), which is the difference between the gross margin obtained with VAR and without it (Pedersen et al., 2021). Since the gross margin equals revenue minus expenditure, it was computed as a function of grain yield. Grain

yield determines both the output revenue and the input costs (through fertilizer requirements). The yield difference between the target ( $Y_w$ ) and the baseline scenario ( $Y_a$ ) is defined as the YG.

The relative advantage was measured as the average DGM. The average DGM ( $\text{€ ha}^{-1}$ ) was estimated by averaging the annual ADGM values ( $\text{€ ha}^{-1} \text{year}^{-1}$ ), expressed as:

$$DGM = \sum_{t=1}^T \frac{ADGM_t}{T} \quad (5.3)$$

where  $T$  is the number of seasons ( $N=6$  years) of our dataset (i.e., 2016–2021), and the subscript  $t$  indicates the specific season.  $ADGM_t$  is the difference between the annual differential revenue ( $ADR_t$ ) and the annual differential costs ( $ADC_t$ ), expressed as:

$$ADGM_t = ADR_t - ADC_t \quad (5.4)$$

where both  $ADR_t$  and  $ADC_t$  are expressed in  $\text{€ ha}^{-1} \text{year}^{-1}$ . Since the DGM represents the differential gross margin, it applies to the differential area, which is defined by the LIF area share. Therefore, both  $ADR_t$  and  $ADC_t$  are defined per ha of LIF zones. The  $ADR_t$  was computed as a function of YG, expressed as following:

$$ADR_t = YGz_t \cdot 1000 \cdot (Price_{WHEAT} + LPP) + DPA \quad (5.5)$$

where  $YGz_t$  represents the YG difference among zones (i.e., between LIF and no-LIF zones) for a specific year ( $t$ ), and the  $Price_{WHEAT}$  corresponds to the grain price ( $\text{€ kg}^{-1}$  DM grain). LPP represents a linked production payment in some scenarios (i.e., a subsidy linked to yield and expressed in  $\text{€ kg}^{-1}$  DM grain), and DPA represents a direct payment on annual wheat grown area ( $\text{€ year}^{-1}$ ). DPA does not consider current direct payments, but it explores the introduction of an additional payment on area over current subsidy levels in one scenario. The larger is  $YGz_t$ , the higher is the expected  $ADR_t$ . The  $YGz_t$  was computed as follows:

$$YG_{Zt} = \sqrt{\left( \left[ \int_{cum.d=0}^{cum.d=1} ECDF(YG_{LIF}) \right] - \left[ \int_{cum.d=0}^{cum.d=1} ECDF(YG_{No-LIF}) \right] \right)^2} \quad (5.6)$$

where  $ECDF(YG_{LIF})$  is the ‘empirical cumulative distribution function’ of YG in ‘LIF zones’ and  $ECDF(YG_{No-LIF})$  is the ‘empirical cumulative distribution function’ of YG in ‘No-LIF zones’.  $Cum.d$  delineates the cumulative density, varying from 0 to 1. The ‘empirical cumulative distribution function’ is the distribution function of a sample measure. In our case, this applies to the yield gap distribution over field. It expresses the fraction of yield gap observations that are less than or equal to a specified value. For  $ECDF$  computation, we used the function ‘stat\_ecdf’ from the ‘ggplot2’ library in R-studio (Wickham, 2007).

The  $ADC_t$  was estimated by the following equation:

$$ADC_t = (YG_{Zt} \cdot 1000 \cdot N_{GRAIN} \cdot Price_N) + AC_{VAR} \quad (5.7)$$

where  $N_{GRAIN}$  is the nitrogen content in grain (expressed in terms of % DM), the  $Price_N$  is the price of N (expressed in € per kg N) and the  $AC_{VAR}$  is the annual cost (Table 5.2) associated with VAR use (e.g., technology renting plus external consultancy costs).  $N_{GRAIN}$  was set equal to 2.8% of DM (Quemada et al., 2016).  $Price_N$  was estimated according to the 2020/21 fertilizer price index for Diammonium Phosphate (DAP), Calcium Ammonium Nitrate (CAN) and Urea (EUC, 2021), and considering the mean N content of fertilizer which was estimated as the simple average of N content in DAP, Urea and CAN. The baseline  $Price_N$  was set equal to 1.10 € kg<sup>-1</sup> N.

## 5.2.5 Economic modelling and analysis of future scenarios

The capital recovery of VAR adoption was modelled according to 10 future scenarios (Table 5.2) which considered different prices and payments (i.e., application of extra subsidies). The Net Present Value (NPV) was estimated over a period of 10 years which was assumed to be the total amortization period of VAR equipment (Drabik & Peerlings, 2016; Tozer, 2009). NPV was estimated

as following:

$$NPV = -II_{VAR} + \sum_{i=1}^N \frac{DGM_i \cdot LIF_{AreaSHARE}}{(1 + \gamma)^i} \quad (5.8)$$

where  $II_{VAR}$  is the initial investment (i.e., the acquisition cost),  $\gamma$  is the discount rate,  $i$  is the year of the cash-flow and  $N$  represents the total amortization period (10 years). The acquisition costs (Appendix – Table A1) were based on information compiled from different sources (AAEA, 2000; Batte & Ehsani, 2006; Finco et al., 2021; Griffin, 2006; Tozer, 2009), including the acquisition costs of the GPS guiding system, the precision application system RTK, the GPS receiver, the base station, the replicators and the N application controller (Appendix – Table A1). The overall total gain (OTG) was estimated by solving the NPV series over a period of 10 years ( $N=10$ ), which was a function of both the discount rate used and the price/payment scenario adopted (Table 5.2). The Internal Rate of Return (IRR) and the return on investment payback time ( $ROI_t$ ) were respectively estimated by solving the following equations:

$$II_{VAR} = \sum_{i=1}^N \frac{DGM_i \cdot LIF_{AreaSHARE}}{(1 + IRR)^i} \quad (5.9)$$

$$ROI_t = \frac{II_{VAR}}{(DGM \cdot LIF_{AreaSHARE} \cdot Wheat_{area})} \quad (5.10)$$

where  $Wheat_{area}$  is the annual wheat sown area (ha). The relation between  $ROI_t$  and  $Wheat_{area}$  was obtained through regression analysis, considering the price/payment scenarios (Table 5.2).

The ten different scenarios analysed the impact of price support policies, extra direct payments on crop area and different market prices (Table 5.2). The ratios considered for increased prices scenarios were defined according to the observed connectedness between fertilizer and grain market prices. By averaging the values reported by Chowdhury et al. (2021) and Khalfaoui et al. (2021), we assumed a ratio among product prices of 0.46, meaning that a 100% increase in fertilizer prices would increase wheat prices by 46%. The comparison of capital recovery among different scenarios was made by adopting

the current situation as the baseline scenario. For each scenario, three different sub-scenarios were considered by setting the discount rate at 2.5%, 5% and 7.5% (Table 5.2). The following 10 scenarios were considered:

1. Baseline scenario (S-1): overall 'LIF-zones' share within the cropping system equal to 17.4%, according to our farm conditions, current CAP direct payments and product prices.
2. Enhanced LIF area share scenario (S-2): overall 'LIF-zones' share within the cropping system increased by 35% in relation to the baseline scenario, current CAP direct payments and product prices.
3. Moderate increased prices scenario (S-3): baseline scenario plus product prices increased by 30% and 66%, respectively for wheat grain and N fertilizer price.
4. Drastically increased prices scenario (S-4): baseline scenario plus product prices increased by 100% and 219%, respectively for wheat grain and N fertilizer price.
5. Price support scenario (S-5): baseline scenario plus introduction of a direct payment linked to production (LPP) equal to 0.02 € kg<sup>-1</sup>.
6. Price support plus drastically increased prices scenario (S-6): scenario LIF0.DPP.P0 plus product prices increased by 100-219% (i.e., wheat grain and N fertilizer price).
7. Additional direct payment on cropped area scenario (S-7): baseline scenario plus introduction of an additional payment on cropped area equal to 46 € ha<sup>-1</sup>.
8. Additional direct payment on crop area plus drastically increased prices scenario (S-8): baseline scenario plus introduction of an additional payment on crop area equal to 46 € ha<sup>-1</sup> plus product prices increased by 100-219% (i.e., wheat grain and N fertilizer price).
9. Support VAR investment (S-9): baseline scenario plus the introduction of a support on investment corresponding to 50% of initial investment covered by a subsidy.
10. Support VAR investment plus drastically increased prices scenario

(S-10): baseline scenario plus the introduction of a 50% subsidy of initial investment covered by a subsidy, plus product prices increased by 100-219% (wheat grain and N fertilizer prices).

Currently, there is no specific support for cereals in Spain, where the present study was conducted, but we followed the example of what was allocated as crop-specific aid per ha in the cases of oilseeds and legume-crops during the year 2020 (FECA, 2021). According to existing crop-specific aids, a mean aid of 40-55 € ha<sup>-1</sup> was assumed. This interval supposes that for our mean yield levels an additional aid of 0.02 € kg<sup>-1</sup> could be applied as LPP, equivalent to a support of 46 € ha<sup>-1</sup> as DPA.

The relation between return time on investment ( $ROI_t$ , expressed in years) and the annual wheat area (expressed in ha) was modelled. Regression analysis was used to estimate model coefficients, several alternative models were tested, including linear, quadratic, logarithmic, exponential and power models. Least Squares Fitting (LSF) and statistical hypothesis testing were respectively used to estimate the regression coefficients and their significance level (stats package in R; Team, 2000). Root mean square error (RMSE) and  $R^2$  were used as statistical indicators of performance for the best-model selection. For each scenario, the minimum area for adoption, corresponding to the threshold below which the return on investment takes longer than the amortization of VAR equipment, was also estimated. The null hypothesis was tested with the non-parametric Tukey's range test (HSD-test).

## 5.3 Results

### 5.3.1 YG analysis

Figure 5.2 shows seasonal water supply over the six seasons considered. Seasonal precipitation (P) from sowing to harvest varied from 228 to 535 mm and seasonal LIF (only considered in 'LIF zones') ranged from 30 to 112 mm.

Figure 5.3 shows simulated Yw values (2016-2021), also reported in Table 5.1, which were obtained from Tenreiro et al. (2022). Measured Ya values ranged

from 0.3 to 6.4 Mg DM ha<sup>-1</sup> over space and time. The intervals of intra-plot variation for each season are shown in Figure 5.3. Ya showed a mean coefficient of variation of 14.7% ( $\pm 3.1\%$ ) among the six seasons considered. The probability distribution curves of YG among zones and within each season are also shown in Figure 5.3. Median YG was 1.7 Mg DM ha<sup>-1</sup> with a standard deviation of 0.135 Mg DM ha<sup>-1</sup>. YG values ranged from 0 to 4.8 Mg DM ha<sup>-1</sup> in different parts of the 92 ha field. In relative terms, the average YG was 30.5% of Yw. The YG's were larger in LIF zones in four out of six seasons (Figure 5.3). The yield maps are shown in Appendix (Figure A2). The YG results are synthesized in Table 5.3.

### 5.3.2 The economics of VAR adoption

The capital recovery of VAR adoption was expressed as a function of both the scale of adoption (i.e., annual wheat sown area) and the financial discount rate. This relation was directly affected by the differential gross margin (DGM), which was computed as the average of Table 5.4 values.

On average, DGM ranged from 12.1 to 147.5 € ha<sup>-1</sup>, depending on the scenario considered. Under current conditions (see S-1 and S-2 in Table 5.4), three out of six seasons (i.e., 2017/18, 2018/19, 2019/20) presented a negative  $ADGM_t$ , indicating risk of economic inefficiencies associated with VAR adoption for 50% of the years investigated. DGM was negative when  $ADR_t$  was lower than  $ADC_t$ . For each year specific conditions, there was a great variation of  $ADGM_t$ , which ranged from -42.3 (baseline scenario) to 405.2 € ha<sup>-1</sup> (scenario S-8, i.e., a policy support through an additional direct payment of +46€ ha<sup>-1</sup> on crop area plus high wheat grain and N fertilizer prices; Table 5.2).

The economics of VAR adoption varied considerably, not only among different years but also for the different scenarios (Table 5.4). The  $ADGM_t$  and  $ADR_t$  showed stronger sensitiveness to the differences among years and economic scenarios than  $ADC_t$ , which presented a lower range of variation (Table 5.4). LIF coefficients (i.e., season LIF divided by season precipitation) varied from 11 to 27% according to Figure 5.2 values. Those years with higher LIF coefficients

(e.g., 2016/17 and 2020/21) showed higher  $ADR_t$  and  $ADGM_t$ . The opportunity cost showed a trend to decrease with the increase of LIF coefficient but regressions were not capable to explain more than 30% of overall variation (results not shown).

### 5.3.3 Economic trade-offs for viability of VAR

Model coefficients were obtained through regression analysis (Table 5.5). From the set of models tested, power models were selected as the best fitted ones. Power models maximized  $R^2$  values and minimized the RMSE. While the  $R^2$  values of linear, quadratic, logarithmic and exponential models were respectively <0.1, 0.1-0.2, 0.2-0.4 and 0.7-0.8, the RMSE (expressed in years) were larger than 10 for the first three model types and approximated 0.5 for the exponential models. Power models ( $a \cdot x^{-1}$ ) were the best fitted models with  $R^2$  above 0.98 and RMSE lower than 0.1. Model coefficients varied from scenario to scenario but the statistics of model performance did not change among scenarios. Model coefficient  $a$  and the minimum area for adoption of VAR are shown in Table 5.5 for each of the 10 scenarios. The minimum area for adoption corresponds to the cutting threshold value below which the return on investment takes longer than the amortization of VAR equipment.

There was a negative relation between the return time on investment ( $ROI_t$ , expressed in years) and the annual wheat sown area (expressed in ha) which was best fitted by a power model. Figure 5.4 shows the power models that were fitted for each economic scenario. The 10-year overall total gain (OTG) changed significantly from case to case, and according to the different discount rates (Figure 5.5). The lower the discount rate, the larger the profitability of VAR investment for addressing the spatial variations of yield gaps due to water redistribution by lateral inflow. Over a 10-year period and under the baseline scenario, at a discount rate of 2.5%, the OTG is expected to vary from -100.9 thousand € (for a total of 50 ha of wheat sown every year) to 310.3 thousand € (in the case of 2000 ha of wheat sown annually). Under the most profitable scenario (S-8), for a median farmer with 100 ha wheat sown per year, the 10-year OTG is forecasted as 125.3, 134.2 and 144.6 thousand € respectively

with a discount rate of 7.5%, 5% and 2.5% (Figure 5.5 and Supplementary material).

In our case (annual sown area equal to 92 ha, with an overall share of 'LIF zones' equal to 17.4%), the internal rate of return (IRR) was negative for almost all cases considered (Table 5.5 and Figure 5.5), indicating a lack of relative advantage (i.e., additional profitability) for all the scenarios and discount rates considered. Over the ten different scenarios explored, and assuming a median case of 92 ha sown, we observed that the capital return could increase up to 5% in the case of scenario S-8 (Table 5.5 and Appendix - Table A2), but this would depend on a drastic price increase or on changes in agricultural policies.

## 5.4 Discussion

### 5.4.1 Crop yields and yield gaps

Actual wheat yields ( $Y_a$ ) varied within the range of other studies (Padilla et al., 2012; Schils et al., 2018) and exhibited significant differences among zones in two out of six seasons (Table 5.3). This is associated with yield maps fitting into variograms with irregular spatial structure as both CV and differences among means did not show consistency from year to year. This highlights that temporal instability is an important issue for site-specific management because the agronomic implications of asymmetric spatial variations differ greatly with the crop  $\times$  year setting, as also discussed by Tenreiro et al. (2020b).

Coefficients of variation were consistently higher in 'no-LIF zones', which is in line with the results of Tenreiro et al. (2022), who highlighted that the contribution of LIF to yield spatial variations tends to be stronger in years of relatively low water supply. This indicates that the degree of variation tends to increase with the level of water stress. Observed CVs of  $Y_a$  were similar to those reported by others (Batchelor et al., 2002; Florin et al., 2009; Whelan & McBratney, 2000).

YG's CV were notably larger than  $Y_a$ 's CV, which is attributed to a larger variation sources affecting the process of YG mapping. According to Table

5.3, YG's were statistically (and systematically) different among zones. This indicates that the YG appears to be a more precise benchmark (instead of using solely Ya or Yw) for precision management decision-making, because it captures a higher degree of spatial variation and it delivers the magnitude of the expected response under a VAR for each year-specific conditions.

#### 5.4.2 The economics of VAR adoption

The results presented here are conditioned by the specific topographic conditions of the chosen farm. Under different geomorphological conditions (e.g., increased LIF area share), the returns on investment are expected to change. The larger the relative share of LIF zones within the farming system, the lower the minimum area required for VAR adoption (see S-2 in Table 5.5).

Under current conditions (S1), we computed a relative advantage associated with VAR adoption but only for an annual wheat sown area larger than 567 ha year<sup>-1</sup> (Table 5.5). This is considerably larger than typical European (arable) farm sizes, which range from 4 to 62 ha (Andersen, 2017). For cases with less sown area, the investment costs could still be recovered but it would not be due to the relative additional gain. This means that a farm with a lower annual sown area and currently profitable with uniform N applications, could pay for the VAR investment, but the overall profitability of the farm would decrease if VAR is adopted.

All the alternative scenarios considered accelerated capital recovery (Table 5.4 and Figure 5.4). However, wheat and N prices were the most determinant factors for VAR viability (Table 5.5 and Figure 5.4). The slope of the 10-year OTG increased notably with both wheat and N prices increasing (Figure 5.5). However, this is determined by the N-grain prices' relation adopted. The most promising scenario, which showed the largest gain on capital recovery (see S-8 in Table 5.4 and Figure 5.5), was the inclusion of an additional direct payment on crop area (+46 € ha<sup>-1</sup>) plus a drastic evolution of both input and output prices.

An additional direct payment on area could also deliver advantages for

a wider range of farmers as the minimum annual wheat area for adoption decreases more than 75% under scenarios S-7 and S-8 (Table 5.5). For the same available budget considered, the introduction of an additional direct payment on area would imply a more robust advantage for VAR adoption than payments linked to production (Table 5.4 and 5.5). Payments linked to grain yield would be mostly diluted over the Ya level, which relates to the current management system, and not over the differential YG closing effort that is attributed to VAR adoption. Since current mean YG's are approximately 30% of mean Yw's (Table 5.3), the major part of a subsidy support linked to production would not apply to the differential gross revenue, caused by the technological shift, but rather to the actual yield level that is already achieved under the current management (i.e., uniform nutrient application). In this sense, a much wider fraction of financial inflow would be directly attributed to the technological shift in the case of an additional payment on area (Table 5.4). This indicates that the direct payment of 46 € ha<sup>-1</sup> year<sup>-1</sup> scenario could be a better option in terms of policy support on investment.

We did not investigate the opportunity of lower initial investment costs associated with lower VAR acquisition costs of equipment, but this was partly explored through the scenarios S-9 and S-10. These scenarios explored the impacts of a policy support on investment through a support-payment equal to 50% of the initial acquisition costs (Table 5.2). This strategy impacts the return on investment by decreasing  $ROI_t$  (Figure 5.4-E), but achieving significant impacts on OTG requires further prices' changes (Figure 5.5-E). In the absence of further price changes, this option does not guarantee incentives for farmers' adoption and this could be a constraint from a strategic point of view.

Our results showed that the profitability of VAR investment would respond more to changes in market prices than to policy supports (Table 5.4 and 5.5). In the absence of additional policy supports, the minimum area for adoption of VAR decreases substantially under both price evolution scenarios (i.e., respectively 69% and 88% according to Table 5.5). The evolution of prices that we assumed is likely to turn VAR into a viable technology for a much wider population of farmers, as the minimum area for adoption, according to the considered range

of prices' increase, is expected to decrease from 567 ha to an interval of 67 to 177 ha year<sup>-1</sup>. This clearly indicates that, given current trends on both wheat and N prices observed in early 2022 (Glauber & Laborde, 2022; Vos et al., 2022), more farmers would be inclined to adopt VAR in local rainfed systems.

The mean OTG values shown in Table 5.5 were estimated as the simple average of OTG obtained for the following series of sown areas: 50, 100, 250, 500, 1000 and 2000 ha. According to the HSD-test results shown in Table 5.5, we observed that scenarios S-1, S-2, S-5 and S-9 did not significantly differ from each other, which indicates that different market prices are the most determinant condition to improve VAR economic viability.

The obtained  $ADGM_t$  fluctuated in a wider range than the values reported by Robertson et al. (2007, 2012). Our results ranged from -42.3 to 405.2 € ha<sup>-1</sup> year<sup>-1</sup> (Table 5.4). Nevertheless, our results are expressed in € ha<sup>-1</sup> of LIF area. Since our conditions are characterized by a mean LIF area share of 17.4%, the  $ADGM_t$  values must be extrapolated to the total crop area for a direct comparison with the results of Robertson et al. (2007, 2009). In terms of total crop area, the obtained  $ADGM_t$  ranged from 7.2 to 70.5 € ha<sup>-1</sup> which is more in agreement with the literature. The average DGM was equal to 12.2 € ha<sup>-1</sup> year<sup>-1</sup> for the baseline-scenario (S-1) and, considering all the scenarios explored, the average DGM was 61.1 € ha<sup>-1</sup> year<sup>-1</sup>, which very much in line with the range of values reported by Robertson et al. (2007, 2009).

Seasons characterized by a negative  $ADGM_t$  showed lower annual revenues than costs. Annual costs are often characterized by larger fixed costs than variable costs, because the variable costs are insensitive to farm size and structure (Pedersen et al., 2021). When the annual revenue did not overcome the 90 € ha<sup>-1</sup> year<sup>-1</sup> threshold (Appendix – Table A1), the net margin was negative. This was observed in three out of six years, indicating a risk of economic losses caused by VAR adoption in half of the years investigated. However, since both Ya and YG were significantly different among zones for those same years, this could still justify the adoption of VAR from both an agronomic and environmental perspective (Mulla & Schepers, 1997; Pathak et al., 2019; Plant, 2001).

We highlight that from a financial point of view, the viability of VAR is strongly dependent on the annual sown area, which depends on how farmers value the return on capital. Under current price conditions (S-1 and S-2), we highlight that a convincing IRR was only obtained for annual sown areas above 500 ha year<sup>-1</sup> (Appendix - Table A2). The IRR can be profitable for sown areas between 125 and 500 ha year<sup>-1</sup>, but it would depend on changes over product prices (see S-4, S-7 and S.10 in Appendix - Table A2).

#### 5.4.3 Methodological considerations and practical issues

VAR may increase farmers' profit by reducing costs or increasing the value of production, because fertilizer rates can be both increased or decreased among differentiated zones. Under rainfed conditions, inter-annual climatic variation leads to considerable asymmetries in crop yield patterns and financial returns on VAR. Some years show advantages on increasing N rates downslope, some benefit from decreasing applications, and others from applying N uniformly. We assume that, independently on the year type, our analysis succeeded well in modelling the expected marginal returns on VAR because the  $YGZ_t$  magnitude was computed as a module. This is valid for both years of reduced or increased N rates in LIF zones. In both situations, the differential application rates are captured by the present methodology.

Our analysis assumes that N application rates approach the crop net N requirements, considering that most N inputs are recovered in the harvested grain. Considering that mean Ya range from 3.1 to 4.5 Mg DM ha<sup>-1</sup> (Table 5.3), crop net N-requirements would be 88-126 kg N ha<sup>-1</sup>. In our study-case conditions, we took N use efficiency into consideration as 25-40% more N is applied on average (Table 5.2). However, we consider that this is in line with the 'characteristic operating space' for N use efficiency that is found in literature addressing European commercial (cereal) farms (Panel, 2015; Quemada et al., 2020; Silva et al., 2021).

This is an 'ex-post analysis' which may be a disadvantage for guiding decisions under new seasons' conditions (Bullock & Lowenberg-DeBoer., 2007).

Our study dealt with historical data (i.e., pre-collected information) and obtained from six consecutive seasons which may not be sufficient to capture the entire variability of rainfed systems. In reality, farmers face much greater uncertainty because decisions need to be taken according to specific year conditions which are highly variable. In addition, it is difficult to follow our approach as 'on-season' YG assessment because it requires post-harvest yield information.

Therefore, the following practical question arises: how could farmers manage N applications under VAR, when their decisions must be taken without full access to the same kind of information as here presented?

It has been shown by Tenreiro et al. (2022) that the net yield response to LIF varies substantially from year to year. This turns decisions on VAR adoption challenging because farmers face possible asymmetries in crop responses patterns to differential N rates. However, some important guidelines can be proposed from our results. Under our study conditions, the maximum differential YG was  $0.83 \text{ Mg DM ha}^{-1}$  (Figure 5.3-B), corresponding to a maximum differential N rate of  $\pm 23.2 \text{ kg N ha}^{-1}$  (i.e., considering crop N requirements equal to grain uptake and assuming a N concentration equal to 2.8%  $\text{DM}^{-1}$  grain, Quemada et al., 2016). This represents a variation of approximately  $\pm 16\% \text{ N ha}^{-1}$  among zones. In addition, we suggested that the cost of opportunity for VAR adoption tends to decrease with the LIF coefficient (i.e., LIF divided by season precipitation). The larger is the fraction of LIF contribution to total season water supply, the greater could be the economic advantages of VAR adoption. Our results have practical implications for nutrient management in areas of undulating topography. In this sense, the following recommendations are offered:

1. Use VAR for basal fertilization, applying up to +8% more N in 'LIF zones' in comparison to 'No-LIF zones'.
2. Adjust that pattern on top-dressing applications, conducted at dates prior to flowering and according to the following criteria:
  1. If the LIF coefficient falls within the top 25% percentile near flowering (i.e.,

if LIF coefficient is larger than 25%, according to Tenreiro et al. (2022) findings), up to +8% more N could be applied in ‘LIF zones’ in comparison to ‘No-LIF zones’.

2. If the LIF coefficient falls within the bottom 25% percentile at flowering date (i.e., if LIF coefficient is smaller than 15%, according to Tenreiro et al. (2022) findings), the application pattern could be inverted by lowering N rates to a minimum of -8% less N in ‘LIF zones’, in comparison to ‘No-LIF zones’.
3. For seasons with a LIF coefficient ranging from 15% to 25% at flowering, a plausible recommendation would be to not vary the N rate for top-dressing applications.

Our results are largely conditioned by the yield simulations of Tenreiro et al. (2022) and must not be directly extrapolated to other cases without further cautious considerations. For many farms, either small sized or presenting low yield variations among zones, none to minor economic advantages associated with VAR are expected for N management. It is therefore essential to farmers and advisers to take into consideration the scale effects here addressed before promoting a technological shift of this kind.

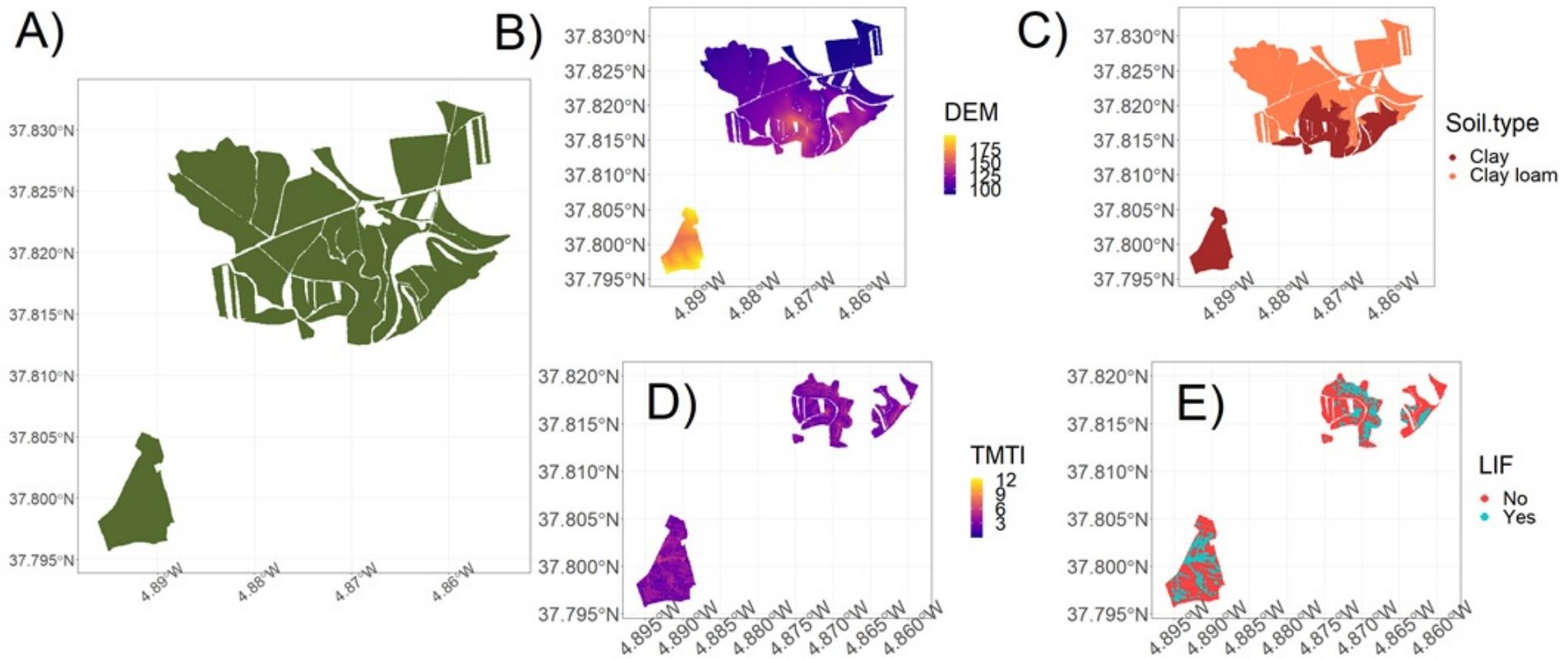
## 5.5 Conclusion

This study demonstrated how the relative (economic) advantages of VAR adoption in rainfed wheat systems of undulating topography would change greatly from year to year and from farm to farm. Both farm size (i.e., annual sown area) and topographic structure (influencing the redistribution of water from high to low parts of the fields) impacted the dynamics of investment returns. Considerable effects of scale were observed and the minimum area for adoption varied widely among different economic scenarios. Our study suggests that there are economic opportunities for N management through VAR as a strategy for bridging yield gaps at intra-plot level, which are caused

by lateral inflows from high to low parts of a field. In the case considered here, VAR adoption shows, for the current policy-prices scenario, that VAR adoption would have an economic advantage in farms with an annual sown area greater than 567 ha year<sup>-1</sup>, which is considerably larger than typical European cereal farm sizes. The profitability of adopting VAR is expected to respond largely to future market prices, and, in the absence of additional policy supports, the minimum area for adoption of VAR could decrease to a range of 68-177 ha year<sup>-1</sup>, depending on the price increases scenario. The effects of policy support on VAR adoption were also investigated with additional payments on crop area being the most promising from both public and private interest perspectives. The combination of further price increases and an additional payment on crop area could lower the adoption threshold down to 46 ha year<sup>-1</sup>, turning VAR technology economically viable for a much wider population of farmers. Over the total amortization period, the (mean) differential gross margin of this case study that is attributed to VAR adoption was 12.2 € ha<sup>-1</sup> year<sup>-1</sup>. Nevertheless, considerable inter-annual variation is expected and farmers might experience net financial losses in some specific years.

## FIGURES - Chapter 5

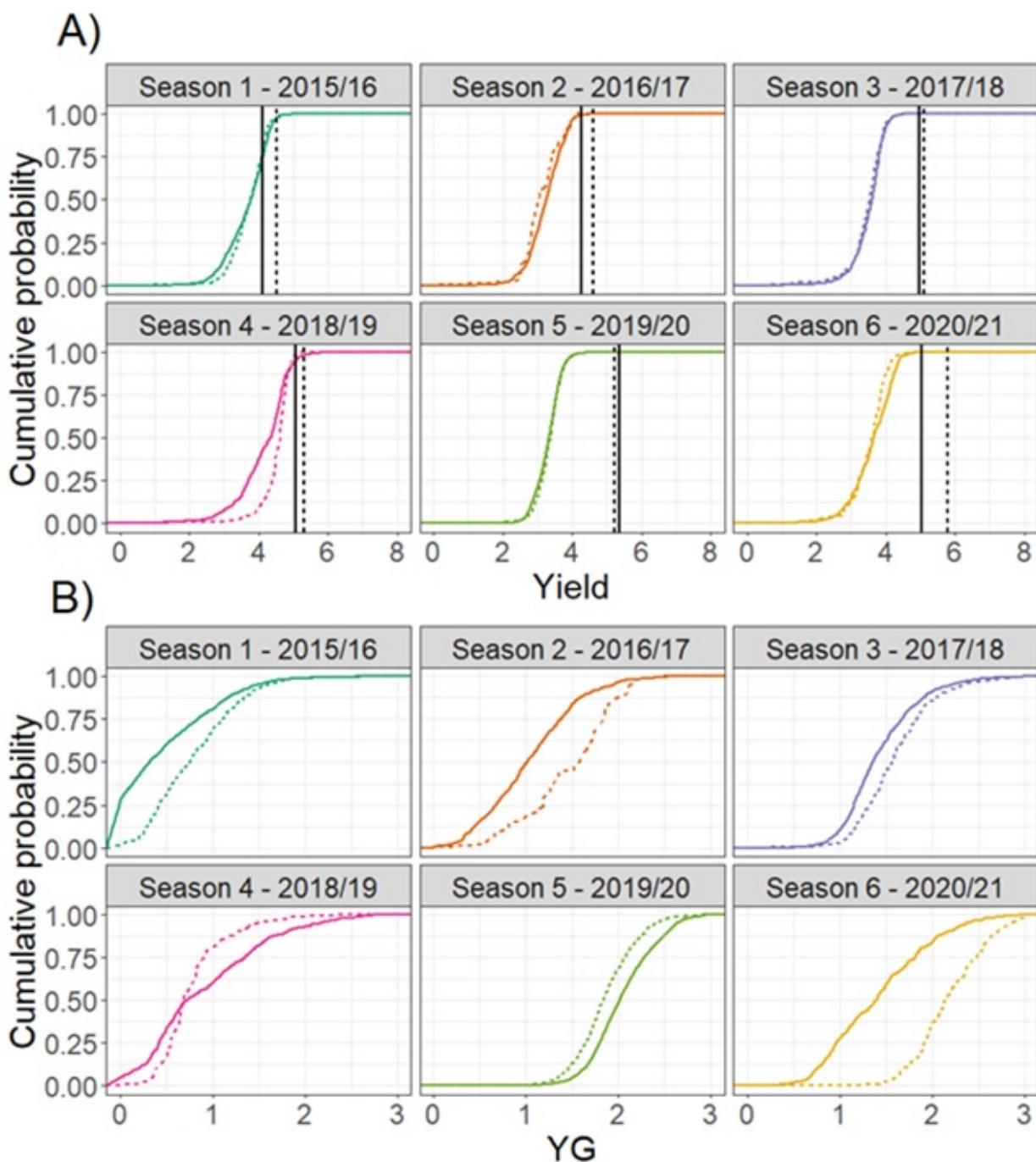
154



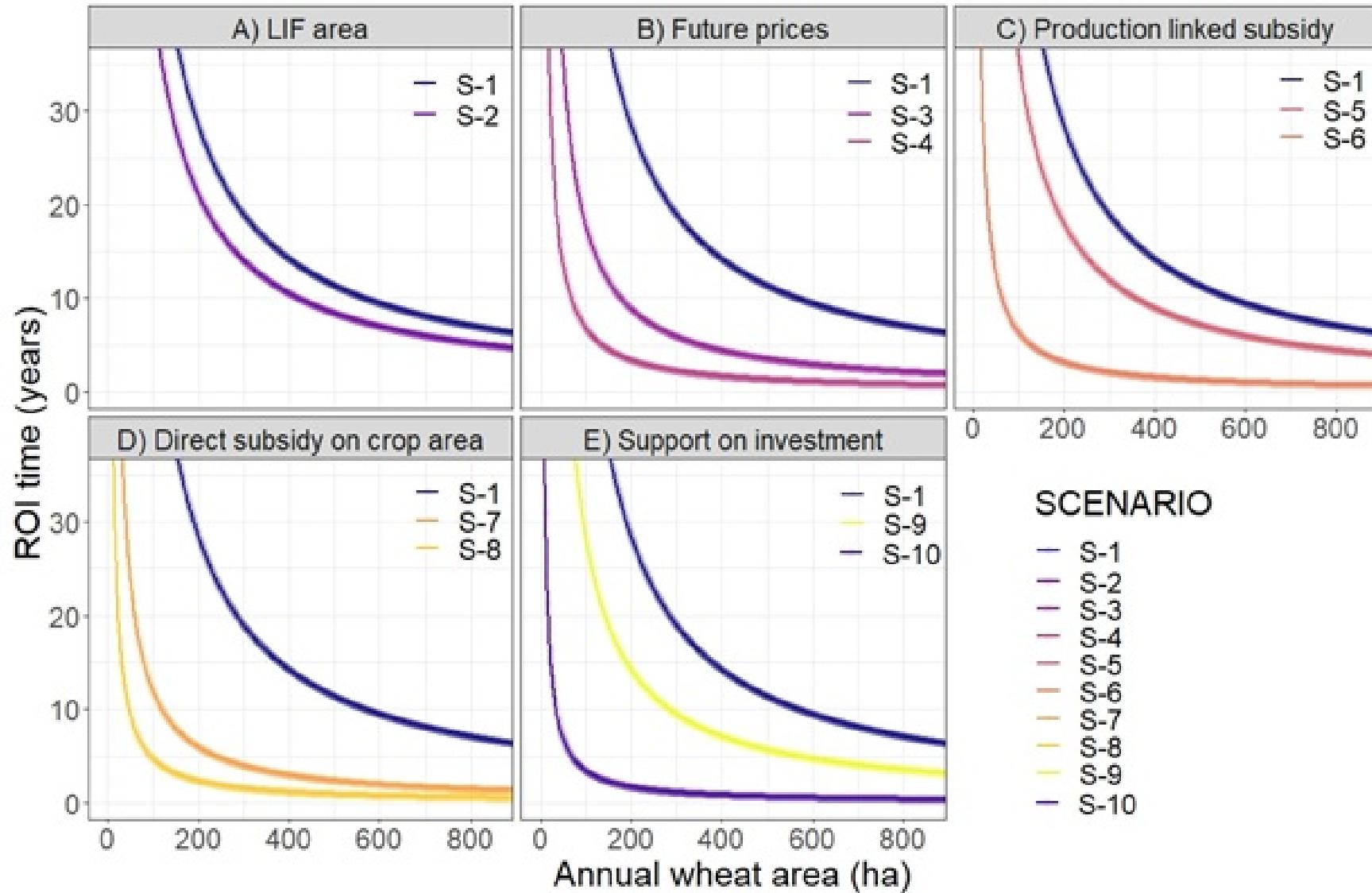
**Figure 5.1:** Experimental fields. A) Total farm area (320 ha); B) Digital Elevation Model (DEM), values expressed in m amsl; C) Soil types according to USDA classification system; D) TOPMODEL Topographic Index (TMTI) as described in detail by Beven et al. (2021), values are unitless; E) Yw zones map (i.e., 'LIF' and 'no-LIF' zones). According to Tenreiro et al. (2022), 'LIF zones' are characterized by significant water supplied through lateral inflow, while 'no-LIF zones' are characterized by null or insignificant lateral inflow.



Figure 5.2: Water supply: seasonal precipitation (P) and lateral inflow (LIF) for the experimental dataset. Values are expressed in mm. More information is provided in Tenreiro et al. (2022).



**Figure 5.3:** Yield variations (values expressed in  $\text{Mg DM ha}^{-1}$ ): A) Cumulative probability distribution of actual yields ( $Y_a$ ) for each season (2015/16 to 2020/21) and within each zone (i.e., solid lines correspond to 'no-LIF zones' while dashed lines relate to 'LIF-zones'); B) Cumulative probability distribution of yield gaps ( $Y_G$ ) for each season and within each zone. Vertical lines indicate  $Y_w$  levels obtained from Tenreiro et al. (2022) simulations.



**Figure 5.4:** The negative relation between return time on investment (ROIt, expressed in years) and the annual wheat area (expressed in ha). The annual wheat area is subjected to the farm conditions supporting our analysis (i.e., an average LIF area share of 17.4%). Scenario S-2 represents the effect of a 35% increase in LIF area share within the cropping system. More information regarding the remaining scenarios considered is provided in Table 5.2.

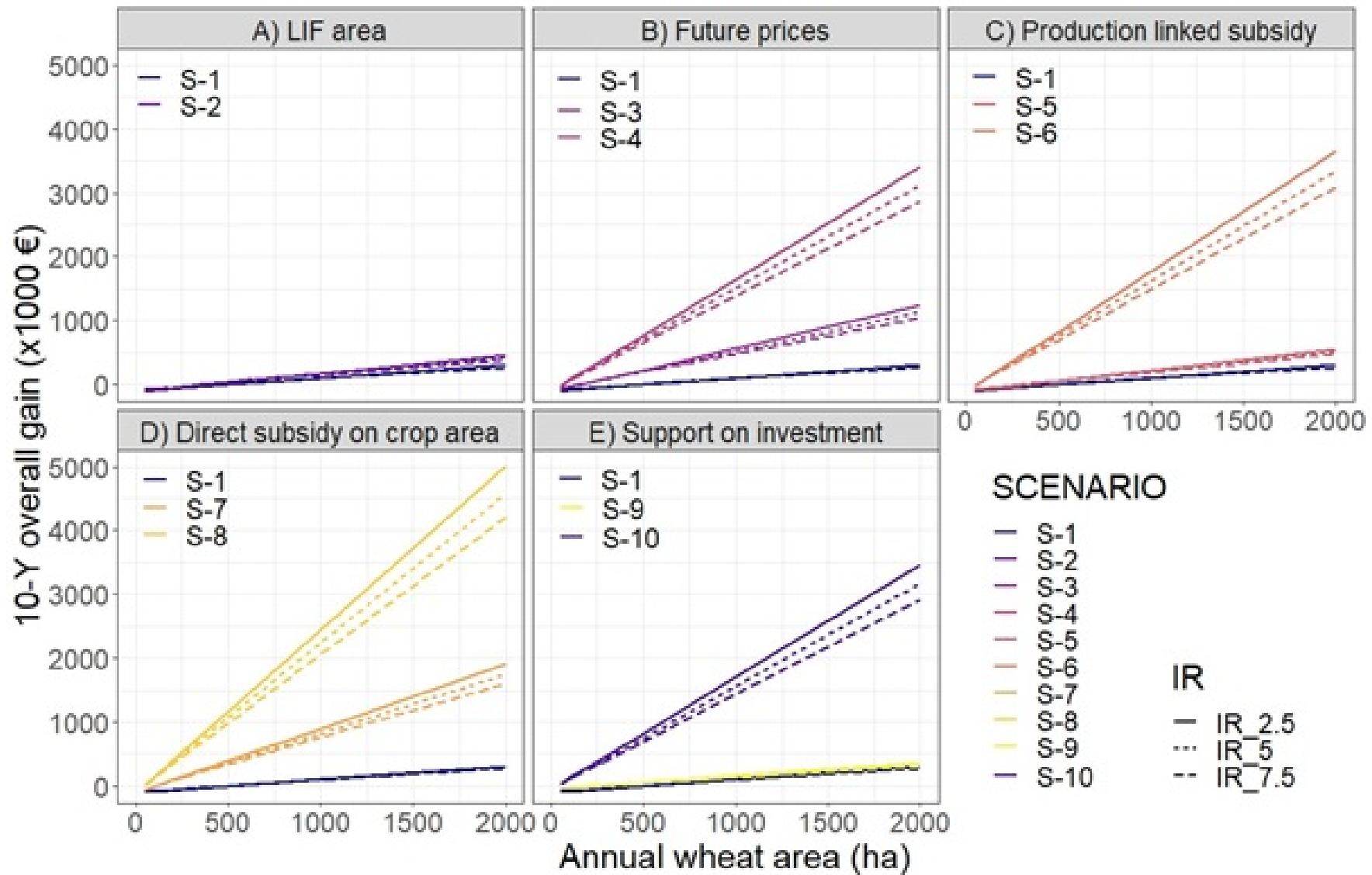


Figure 5.5: The 10-year overall total gain (OTG) under multiple scenarios (units expressed in thousand €). The OTG was estimated by solving the NPV series over a period of 10 years, which was a function of both the discount rate used and the price/payment scenarios adopted (Table 5.2).

## TABLES - Chapter 5

**Table 5.1:** Crop management information: Fertilizers DAP and CAN correspond to diammonium phosphate (18-46-0) and calcium ammonium nitrate (24-0-0 + 8% Ca), respectively. Urea (46% N) was top-dressing applied.  $Y_w$  values were obtained from Tenreiro et al. (2022). Sowing and harvest dates indicate day plus month and mean N applied is expressed in  $\text{kg N ha}^{-1}$ . Water-limited yield ( $Y_w$ ) values are expressed in  $\text{Mg DM (grain) ha}^{-1}$ . Mean plant density was 230 plants  $\text{ha}^{-1}$ . 'LIF': downslope zones with significant amount of water supplied through lateral flow coming from upslope areas of the same field; 'No-LIF': upslope zones where no significant amount of water is supplied through lateral flow. More information on LIF and No-LIF zones is provided in Tenreiro et al. (2022), see also section 5.2.2.

Season	Cultivar	Area [ha]	Sowing date	Harvest date	Previous crop	Fertilization	Mean N applied [ $\text{kg N ha}^{-1}$ ]	$Y_{w_{LIF}}$ [ $\text{Mg DM ha}^{-1}$ ]	$Y_{w_{No-LIF}}$ [ $\text{Mg DM ha}^{-1}$ ]
2015/16	Amilcar	39.5	10.11	23.06	Chickpea	DAP + Urea	172	4.5	4.1
2015/16	Don Ricardo	43.0	9.11	20.06	Sunflower	DAP + Urea	188	4.5	4.1
2016/17	Athoris	20.0	16.11	21.06	Onions	DAP + Urea	170	4.6	4.3
2017/18	Kiko-Nick	50.3	24.11	21.06	Sunflower	DAP + Urea	182	5.1	4.9
2018/19	Antalis	39.9	17.11	16.06	Sunflower	DAP + Urea	152	5.3	5.1
2019/20	Kiko-Nick	50.2	13.12	10.06	Rapeseed	CAN + DAP	110	5.2	5.3
2020/21	Avispa-R1	40.3	18.11	8.06	Rapeseed	DAP + Urea	160	5.8	5.1

**Table 5.2:** The ten different scenarios considered to analyse the impact of price support policies, extra direct payments, and different market prices on VAR economic relevance. For each scenario, three different sub-scenarios were considered by setting the discount rate ( $\gamma$ ) at 2.5%, 5% and 7.5%. LPP and DPA mean linked to production payment and direct payment on crop area, respectively.  $AC_{VAR}$  means annual cost of VAR use and  $II_{VAR}$  is the initial acquisition cost of VAR technology (Appendix – Table A1). LPP was based on a total available budget equal to 46€ ha $^{-1}$  (as considered in S-7 and S-8 scenarios) divided by the average yield.

Scenario	[S-1]	[S-2]	[S-3]	[S-4]	[S-5]	[S-6]	[S-7]	[S-8]	[S-9]	[S-10]
$LIF_{AreaSHARE}$ (%)	17.4	17.4 (+35%)	17.4	17.4	17.4	17.4	17.4	17.4	17.4	17.4
$Price_{WHEAT}$ (€ kg $^{-1}$ DM grain)	0.322	0.322	0.416	0.640	0.322	0.640	0.322	0.640	0.322	0.640
LPP (€ kg $^{-1}$ )	0	0	0	0	+0.02	+0.02	0	0	0	0
DPA (€ ha $^{-1}$ )	0	0	0	0	0	0	46	46	0	0
PriceN (€ kg $^{-1}$ N)	1.093	1.093	1.812	3.490	1.093	3.490	1.093	3.490	1.093	3.490
$AC_{VAR}$ (€/ha)	90	90	90	90	90	90	90	90	90	90
$II_{VAR}$ (€)	11,900	11,900	11,900	11,900	11,900	11,900	11,900	11,900	5,950	5,950
$\gamma$ (%)	[2.5 - 7]	[2.5 - 7]	[2.5 - 7]	[2.5 - 7]	[2.5 - 7]	[2.5 - 7]	[2.5 - 7]	[2.5 - 7]	[2.5 - 7]	[2.5 - 7]

**Table 5.3:** Mean yield values, yield gaps, and coefficients of variation. The standard deviations are presented in parentheses. Means followed by a common letter are not significantly different, according to the HSD-test conducted at the 5% level of significance (p-value < 0.05).

Season	Ya (Mg DM ha <sup>-1</sup> )		CV% (Ya)		YG (Mg DM ha <sup>-1</sup> )		CV% (YG)	
	LIF	No-LIF	LIF	No-LIF	LIF	No-LIF	LIF	No-LIF
2015/16	3.72 (0.4) a	3.66 (0.6) a	12.4	15.8	0.79 (0.4) a	0.49 (0.5) b	57.3	102.9
2016/17	3.13 (0.5) a	3.23 (0.5) a	15.0	15.5	1.48 (0.4) a	1.04 (0.5) b	31.7	47.6
2017/18	3.45 (0.5) a	3.51 (0.4) a	14.3	12.5	1.65 (0.5) a	1.45 (0.4) b	29.9	30.2
2018/19	4.52 (0.4) a	4.14 (0.6) b	8.3	17.1	0.78 (0.3) b	0.95 (0.7) a	47.6	75.6
2019/20	3.34 (0.3) a	3.29 (0.3) b	10.8	11.1	1.87 (0.3) b	2.05 (0.3) a	19.3	17.8
2020/21	3.52 (0.5) a	3.60 (0.6) a	14.6	16.9	2.27 (0.5) a	1.44 (0.6) b	22.7	41.8

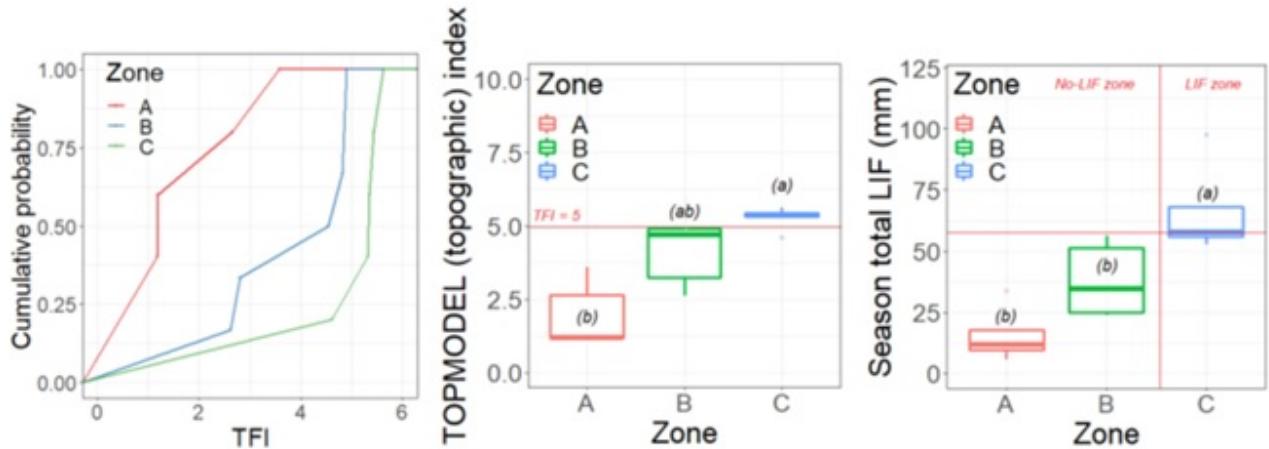
**Table 5.4:** The economics of VAR adoption, under spatial water variations in rainfed wheat systems according to our experimental conditions.  $ADGM_t$  is the annual differential gross margin in year  $t$ ,  $ADR_t$  is the annual differential revenue in year  $t$ , and  $ADC_t$  is the annual differential cost in year  $t$ .  $DGM$ ,  $ADR$ ,  $ADC$  are, respectively, the average differential gross margin, the average differential revenue and the average differential costs. Economic scenarios are described in Table 5.2. Values are expressed in € ha<sup>-1</sup> of LIF area.

Term	Season	Economic scenario									
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10
$ADGM_t$	2015/16	-0.4	-0.4	23.1	77.9	5.8	84.2	45.6	123.9	-0.4	77.9
	2016/17	38.7	38.7	72.4	151.2	47.6	160.1	84.7	197.2	38.7	151.1
	2017/18	-33.8	-33.8	-19.1	15.3	-29.9	19.2	12.2	61.3	-33.8	15.2
	2018/19	-42.3	-42.3	-29.8	-0.6	-38.9	2.8	3.7	45.5	-42.3	-0.6
	2019/20	-39.1	-39.1	-25.7	5.6	-35.5	9.1	6.9	51.6	-39.0	5.5
	2020/21	149.7	149.7	212.4	359.2	166.2	375.7	195.7	405.2	149.7	358.9
$ADR_t$	2015/16	99.1	99.1	128.8	198.2	105.3	204.4	145.1	244.2	99.1	198.2
	2016/17	142.3	142.3	185.0	284.6	151.2	293.5	188.3	330.6	142.3	284.6
	2017/18	62.1	62.1	80.8	124.2	66.0	128.1	108.1	170.2	62.1	124.2
	2018/19	52.8	52.8	68.6	105.5	56.1	108.8	98.8	151.5	52.8	105.5
	2019/20	56.4	56.4	73.3	112.8	59.9	116.3	102.4	158.8	56.4	112.8
	2020/21	265.0	265.0	344.5	530.0	281.6	546.6	311.0	576.0	265.0	530.0
$ADC_t$	2015/16	99.5	99.5	105.7	120.2	99.5	120.2	99.5	120.2	99.5	120.3
	2016/17	103.6	103.6	112.6	133.4	103.6	133.4	103.6	133.4	103.6	133.6
	2017/18	95.9	95.9	99.9	109.0	95.9	109.0	95.9	109.0	95.9	109.0
	2018/19	95.0	95.0	98.4	106.1	95.0	106.1	95.0	106.1	95.0	106.2
	2019/20	95.4	95.4	99.0	107.2	95.4	107.2	95.4	107.2	95.4	107.3
	2020/21	115.3	115.3	132.1	170.8	115.3	170.8	115.3	170.8	115.3	171.1
$DGM$	[average]	12.1	12.1	38.9	101.4	19.2	108.5	58.1	147.4	12.1	101.3
$ADR$	[average]	113.0	113.0	146.8	225.9	120.0	233.0	159.0	271.9	113.0	225.9
$ADC$	[average]	100.8	100.8	108.0	124.5	100.8	124.5	100.8	124.5	100.8	124.6

**Table 5.5:** This table shows the best fitted models' coefficient-a values (i.e., a power model adjusted to the data), the minimum area for adoption of VAR (expressed in ha per year), the mean return on investment time ( $ROI_t$ , expressed in years) and the mean overall total gain ( $OTG$ , expressed in thousand € per 10-year period). The minimum area for adoption corresponds to the value below which the return on investment takes longer than the amortization of equipment (i.e., 10 years).  $OTG$  means followed by a common letter are not significantly different, according to the HSD-test conducted at the 5% level of significance. The internal rate of return ( $IRR$ ) values assume an average annual sown area equal to 92 ha.

Scenario -	Best fitted model (a/x) (coefficient a)	Minimum area for adoption [ha]	Mean $ROI_t$ [years]	Mean $OTG$ [thousand €/10-years]	$IRR$ [%]
[S-1]	5677.4	567.74	35.484	25.592 e	-14%
[S-2]	4205.5	420.55	26.284	69.579 e	-11%
[S-3]	1772.8	177.28	11.079	302.406 d	-2%
[S-4]	679.87	67.987	4.249	949.423 b	-13%
[S-5]	3590.7	359.07	22.441	98.629 e	-1%
[S-6]	635.63	63.563	3.972	1022.459 b	0%
[S-7]	1186	118.6	7.412	501.513 c	-9%
[S-8]	467.75	46.775	2.923	1425.343 a	5%
[S-9]	2838.7	283.87	17.742	75.634 e	-14%
[S-10]	340.3	34.03	2.126	998.347 b	-2%

## APPENDIX - Chapter 5



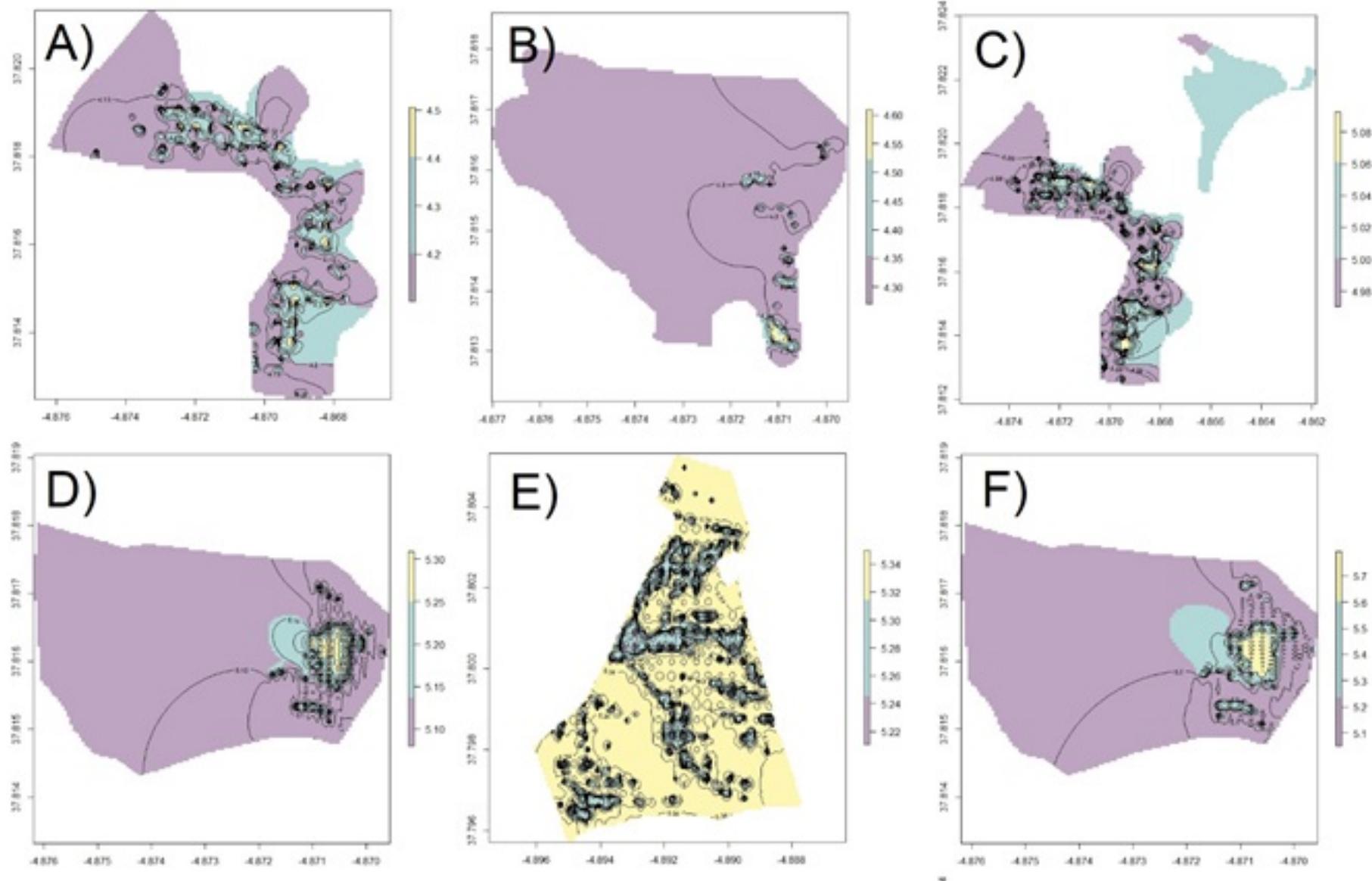
**Figure A1.** Left-plot: probability distribution curves of TOPMODEL topographic index values (Beven et al., 2021) for each sampling zone. Centre-plot: boxplots of mean TOPMODEL topographic index for each sampling zone. Right-plot: boxplots of mean lateral inflow (LIF) within each sampling zone. Values correspond to the experimental measurements taken by Tenreiro et al. (2022). Boxes indicate the lower and upper quartiles. The solid line within the box is the median. Whiskers indicate the most extreme data point which is no more than 1.5 times the interquartile range from the box, and the outlier dots are those observations that are beyond that range.

**Table A1.** Acquisition costs: information compiled from different sources (AAEA, 2000; Batte & Ehsani, 2006; Finco et al., 2021; Griffin, 2006; Tozer, 2009). (\*)VAR annual costs were estimated by considering the (fixed + variable) costs of a tractor (with an amortization through 18 years) plus the combine harvesting and yield monitoring renting cost, expressed in € ha<sup>-1</sup>. VAR technology includes a precision application system RTK, GPS receiver, base station, replicators and application controller.

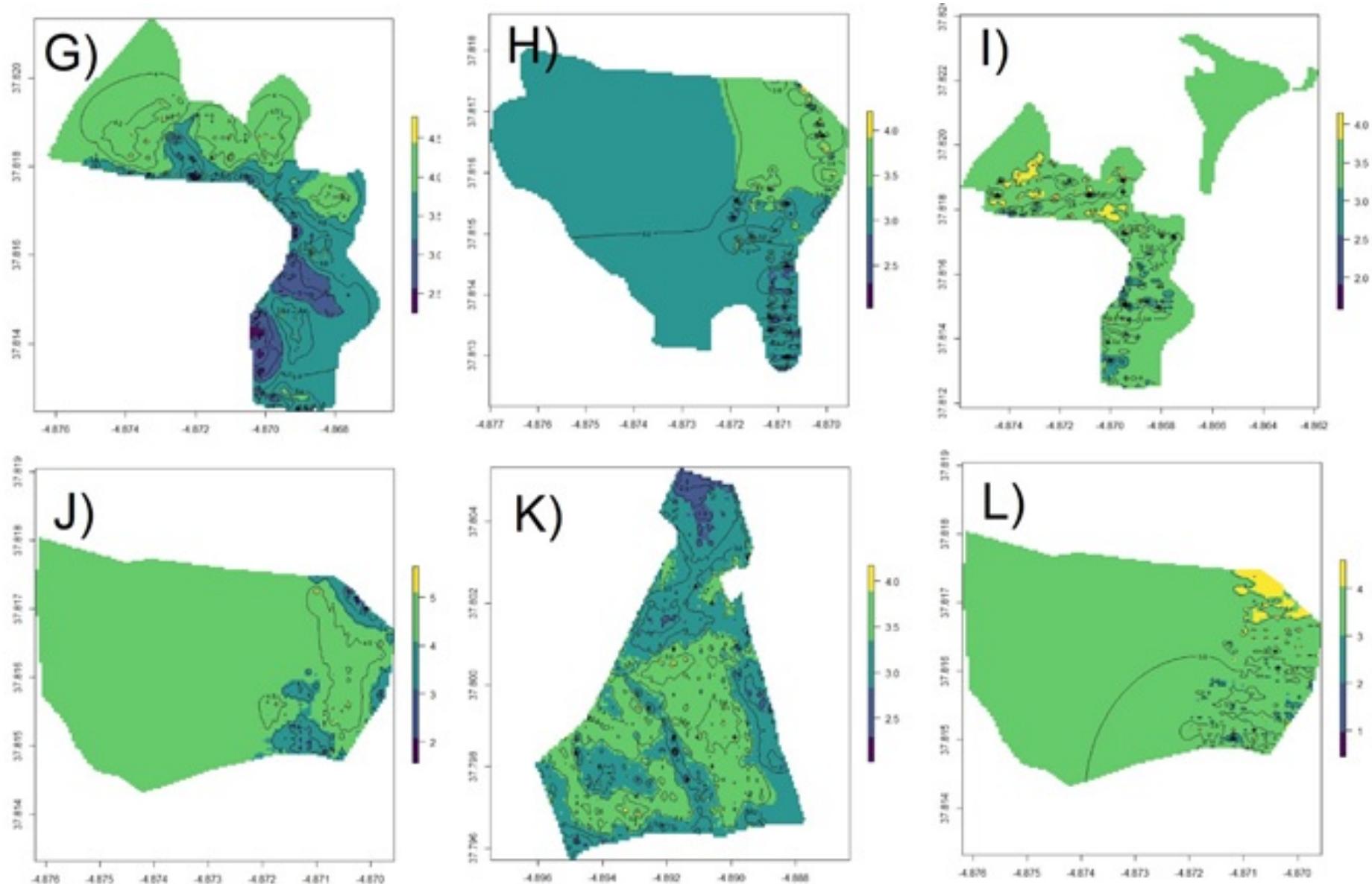
Equipment/operation	Mode	Cost	Lifetime	Period	Units
Tractor 270HP	Acquisition	170,000	16,000 h	> 10 years	€
GPS guiding system	Acquisition	4,000	2,000 h	> 10 years	€
Combine harvester + yield monitor	Renting	60	2,000 h	Season	€ ha <sup>-1</sup>
VAR technology	Acquisition	5,200	1,500 h	≈ 10 years	€
N application controller (18 m bar)	Acquisition	2,700	1,500 h	≈ 10 years	€
VAR annual costs	-	90(*)	-	-	€ year <sup>-1</sup>
Fixed initial investment cost	-	11,900	-	-	€/10 years

**Table A2.** The internal rate of return (IRR) for different annual sown areas (ha). Values expressed in %.

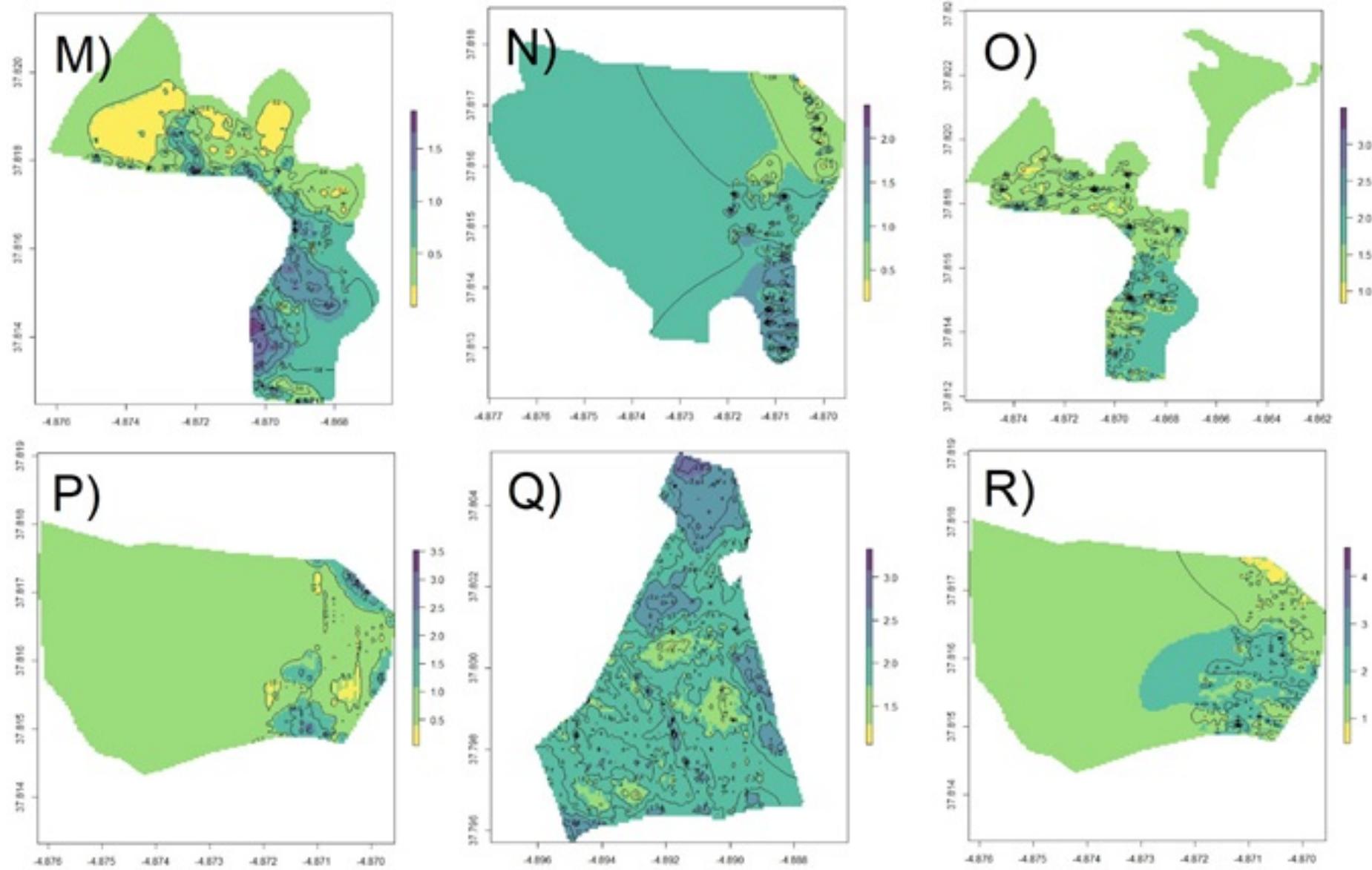
Area (ha)	50	100	250	500	1000	2000
S-1	-19%	-14%	-6%	1%	9%	18%
S-2	-19%	-14%	-6%	1%	9%	18%
S-3	-17%	-10%	2%	15%	36%	78%
S-4	-11%	-1%	23%	57%	122%	249%
S-5	-18%	-13%	-4%	4%	14%	30%
S-6	-11%	1%	26%	62%	132%	268%
S-7	-15%	-8%	8%	29%	67%	143%
S-8	-7%	7%	41%	91%	189%	377%
S-9	-19%	-14%	-6%	1%	9%	18%
S-10	-11%	-1%	23%	57%	122%	249%



**Figure A2.** Yield maps: A-F) Water-limited yield potential ( $Y_w$ ). Maps are ordered from season 2015/16 to 2020/21, respectively. Values are expressed in  $Mg\ DM\ ha^{-1}$ . Counter lines are represented by solid black lines, indication regions of equal yield level. Colours and scales are plot specific.



**Figure A2.** Yield maps: G-L) Actual yield ( $Y_a$ ). Maps are ordered from season 2015/16 to 2020/21, respectively. Values are expressed in  $Mg\text{ DM ha}^{-1}$ . Counter lines are represented by solid black lines, indication regions of equal yield level. Colours and scales are plot specific.



**Figure A2.** Yield maps: M-R) Yield gaps (YG). Maps are ordered from season 2015/16 to 2020/21, respectively. Values are expressed in Mg DM ha<sup>-1</sup>. Counter lines are represented by solid black lines, indication regions of equal yield level. Colours and scales are plot specific.



# Chapter 6

## Conclusions

The main conclusions from the research carried out in this PhD thesis are summarized below. For clarity, they appear according to each of the specific and general objectives raised in the introduction.

### First objective

The review of modelling approaches to simulate lateral water flows and their impact on yield (covered in Chapter two) revealed two major opportunities to simulate spatial water variations at crop field level:

- 1) The addition of spatial and continuous mechanisms to crop models.
- 2) The integration of lateral flows in current 'tipping-bucket' discrete approaches.

This thesis focused on the incorporation of both surface and subsurface lateral flows when simulating crop performance with a 'tipping-bucket' discrete approach (the AquaCrop model). In order to incorporate lateral inflows within AquaCrop simulations, an innovative approach was developed by distributing spatially the modelling scheme through a spatial segmentation that was defined by both crop (i.e., the incorporation of remote sensing estimations of canopy cover, as conducted in Chapter three) and topographic variables (i.e., the zonal segmentation as done in Chapters four and five).

### Second objective

Data assimilation methods were explored to document spatial variations within a field in Chapter three. Several practical advantages of using remote

sensing indicators (i.e., NDVI) for the assessment of canopy cover were identified and evaluated. It was shown that despite the overall uncertainty, they can be adopted with fair confidence in modelling applications, mostly in cases where the relative variations of predictions are prioritized over the absolute accuracy level. This represents substantial advantages for the use of crop simulation models within the context of precision agriculture.

The empirical models developed in Chapter three assist in the use of NDVI for determining canopy cover which improved crop growth estimates in subsequent experimental and modelling steps of this thesis. However, this must be seen as a complementary step to the impact of spatial variations of the water balance. Furthermore, the integration of lateral water movement, and the simulation of the ‘cause–effect’ relations between neighbouring cells/zones, is necessary to capture spatial compensations of yield determining factors, such as variations of harvest index or of root growth caused by differences in water availability.

### **Third objective**

This thesis carried out one of the first modelling studies conducted at commercial fields’ scale, supported by experimental data and delivering water lateral inflow patterns and their simulated impact on wheat yields over 30 years, leading to the following conclusions from Chapter 4:

**3.1)** Progress was made in the modelling process because lateral flows were experimentally measured in real field conditions, their magnitude and their temporal frequency were assessed, and these flows were integrated in the computation scheme of the water balance within the AquaCrop model.

**3.2)** The water accumulation process, represented in our analysis by the flow accumulation index, requires large scales to have implications in the water balance computation and in the yield simulation. According to our artificial neural network model (Chapter 4), the magnitude of lateral flows is determined by the rainfall amount, the runoff generated upstream (as well as at the same point), and the overall accumulation of flow generated upstream, which is determined by the catchment size. The accumulation of lateral flow governs its relevance in the modelling process. The LIF was segmented according to

different magnitude classes, by distinguishing between the zones where these flows are insignificant (high elevation zones) or important for the performance simulation (low elevation zones). For intermediate sites of the field (i.e., those located in medium elevation zones), no significant advances were achieved in the modelling process. For these cases, the effects of lateral flow over simulated yields remains unknown. The modelling process developed within this thesis allowed to capture the spatial variation between the upper and lower zones of a field but it did not achieve a complete and continuous distribution of the water balance throughout the field.

**3.3)** It was concluded that LIF contributed to yield variations in rainfed wheat production systems such as the one studied here. Simulated wheat yields varied (within field) an average of 16% due to LIF occurrence, and the net yield response to LIF in downslope areas averaged 383 kg grain yield (GY)  $\text{ha}^{-1}$  over 30 years. LIF impact on yield was mostly dependent on the year precipitation conditions.

#### **Fourth objective**

This thesis demonstrated that there is a relative (economic) advantage of VAR in rainfed wheat systems of undulating topography. The analysis contained in Chapter five suggested that there are economic opportunities for nitrogen management through VAR as a strategy for bridging yield gaps at intra-plot level, which are caused by lateral inflows from high to low parts of a field. However, this advantage would change greatly from year to year and from farm to farm. Both farm size and topographic structure (influencing the redistribution of water from high to low parts of the fields) impacted the dynamics of investment returns and considerable effects of scale were observed.

The economic viability of such a management system is profoundly affected by effects of scale, both in terms of the relationships between supplying (upstream) and the receiving (downstream) areas, and with respect to the relationships between costs and benefits associated with VAR adoption.

From the results of this thesis, the VAR system must be seen as an unfeasible strategy for typical local wheat fields and arable farms. In mean terms, the

areas of local rainfed wheat based systems are considerably lower than the minimum area required for a complete amortization of VAR technology. Under current price and policy conditions, the VAR system can represent a competitive advantage but only for considerably large farms which do not tend to be found in Europe.

VAR adoption would have an economic advantage in farms with an annual sown area greater than 567 ha year<sup>-1</sup> but this is considerably larger than typical European cereal farm sizes. However, it was observed that the profitability of adopting VAR is likely to respond largely to future market prices as the combination of further price increases and an additional payment on crop area could turn VAR technology economically viable for a much wider population of farmers. Under the most promising scenario, the adoption threshold could lower down to 46 ha year<sup>-1</sup>.

There is an opportunity associated with the evolution of the price levels for both grain and nitrogen, however, it is still necessary that the current price relationships are maintained to promote a relative advantage for farmers. Considerable inter-annual variation is expected and farmers might experience net financial losses in some specific years. It is considered that the economic results here presented are novel because this study focused on the additional advantage generated by VAR and not simply on the overall returns that could be unrelated to the technological transition itself.

## Main objective

This thesis demonstrated that spatial crop-water variation in rainfed wheat systems in Mediterranean conditions is a relevant process in determining spatial differences in yield, and it can be simulated by combining hydrologic and crop simulation models. Progress in the simulation of crop response to spatial variations delivers new opportunities for better crop production through site-specific management.

According to the simulations conducted, lateral inflow plays a considerable net contribution to yield variation over space, which reveals an agronomic and environmental opportunity to apply nitrogen with a variable application rate

(VAR). The agronomic implications and opportunities for VAR adoption depend on the proportional relations between water supplying and receiving zones and these must be assessed at both field and farm level. Although being of difficult amortization within the context of European arable farming, and under current conditions, there is a growing opportunity associated with the evolution of prices that will increase the relative advantage of implementing VAR as an agronomic strategy to deal with crop-water spatial variations.



# References

- AAEA. 2000. Commodity costs and returns estimation handbook. American Agricultural Economics Association.
- Abbate, P.E., Dardanelli, J.L., Cantarero, M.G., Maturano, M., Melchiori, R.J.M., Suero, E.E., 2004. Climatic and Water Availability Effects on Water-Use Efficiency in Wheat. Published in *Crop Sci.* 44, 474–483.
- Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E., Rasmussen, J., 1986. An introduction to the European hydrological system—Système hydrologique Europeen, "SHE", 1. History and philosophy of a physically-based, distributed modelling system. *J. Hydrol.* 87 (1–2), 45–59.
- Abrahamsen, P., Hansen, S., 2000. Daisy: an open soil-crop-atmosphere system model. *Environ. Model. Softw.* 15 (3), 313–330.
- Acutis, M., Scaglia, B., Confalonieri, R., 2012. Perfunctory analysis of variance in agronomy, and its consequences in experimental results interpretation. *Eur. J. Agron.* 43, 129–135.
- Acutis, M., Scaglia, B., Confalonieri, R., 2012. Perfunctory analysis of variance in agronomy, and its consequences in experimental results interpretation. *Eur. J. Agron.* 43, 129–135.
- Ahuja, L.R., Ma, L., Anapalli, S.S., 2019. Biophysical system models advance agricultural research and technology: some examples and further research needs. Bridging Among Disciplines by Synthesizing Soil and Plant Processes. *Advances in Agricultural Systems Modeling.* American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, Inc., Madison, WI.
- Ahuja, L.R., Ma, L., Anapalli, S.S., 2019. Biophysical System Models Advance Agricultural Research and Technology: Some Examples and Further

Research Needs, in: Bridging Among Disciplines by Synthesizing Soil and Plant Processes, Advances in Agricultural Systems Modeling. American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, Inc., Madison, WI.

Ahuja, L.R., Ma, L., Lascano, R.J., Saseendran, S.A., Fang, Q.X., Nielsen, D.C., Wang, E., Colaizzi, P.D., 2014. Syntheses of the current model applications for managing water and needs for experimental data and model improvements to enhance these applications. Practical Applications of Agricultural System Models to Optimize the Use of Limited Water. Advances in Agricultural Systems Modeling. American Society of Agronomy, Inc., Crop Science Society of America, Inc., Soil Science Society of America, Inc., Madison, WI, pp. 399–438.

Ahuja, L.R., Naney, J.W., Nielsen, D.R., 1984. Scaling soil water properties and infiltration modeling 1. *Soil Sci. Soc. Am. J.* 48 (5), 970–973.

Allen, L.H., Jr, Kakani, V.G., Vu, J.C.V., Boote, K.J., 2011. Elevated CO<sub>2</sub> increases water use efficiency by sustaining photosynthesis of water-limited maize and sorghum. *J. Plant Physiol.* 168, 1909–1918.

Allen, R.G., Pereira, L.S., 2009. Estimating crop coefficients from fraction of ground cover and height. *Irrig. Sci.* 28, 17–34.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., et al., 1998. Crop evapotranspiration- guidelines for computing crop water requirements-FAO irrigation and drainage paper 56. Fao Rome 300 (9), D05109.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., Others, 1998. Crop Evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56. Fao, Rome 300, D05109.

Allen, R.G., Pereira, L.S., Smith, M., Raes, D., Wright, J.L., 2005. FAO-56 dual crop coefficient method for estimating evaporation from soil and application extensions. *J. Irrig. Drain. Eng.* 131 (1), 2–13.

Allen, T.H., 1991. Investigation of curve number procedure. *J. Hydraul. Eng.* 117 (6), 725–737.

Andersen, E. 2017. The farming system component of European agricultural landscapes. *European Journal of Agronomy*, 82, 282–291.

ASCE, 1982. Task committee on glossary of hydraulic modeling terms modeling hydraulic phenomena. In: Viessman, W., Lewis, C.L. (Eds.), *Introduction to Hydrology*. ASCE, Prentice-Hall, USA, pp. 454–455.

Asseng, S., Ewert, F., Martre, P., Rötter, R.P., Lobell, D.B., Cammarano, D., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W., Reynolds, M.P., Alderman, P.D., Prasad, P.V.V., Aggarwal, P.K., Anothai, J., Basso, B., Biernath, C., Challinor, A.J., De Sanctis, G., Doltra, J., Fereres, E., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurrealde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Koehler, A.-K., Müller, C., Naresh Kumar, S., Nendel, C., O'Leary, G., Olesen, J.E., Palosuo, T., Priesack, E., Eyshi Rezaei, E., Ruane, A.C., Semenov, M.A., Shcherbak, I., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P.J., Waha, K., Wang, E., Wallach, D., Wolf, J., Zhao, Z., Zhu, Y., 2014. Rising temperatures reduce global wheat production. *Nat. Clim. Change* 5, 143.

Baddeley, A., Rubak, E., Turner, R., 2015. *Spatial Point Patterns: Methodology and Applications with R*. CRC Press.

Balafoutis, A.T., Beck, B., Fountas, S., Tsipopoulos, Z., Vangeyte, J., van der Wal, T., Soto-Embodas, I., Gómez-Barbero, M., Pedersen, S.M., 2017. Smart farming technologies – description, taxonomy and economic impact, in: *Progress in Precision Agriculture*. Springer International Publishing, Cham, pp. 21–77.

Baret, F., Guyot, G., 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sens. Environ.* 35, 161–173.

Basso, B., & Antle, J. (2020). Digital agriculture to design sustainable agricultural systems. *Nature Sustainability*, 3(4), 254–256.

Basso, B., Cammarano, D., De Vita, P., 2004. Remotely sensed vegetation indices: Theory and applications for crop management. *Rivista Italiana di Agrometeorologia* 1, 36–53.

Basso, B., Cammarano, D., Fiorentino, C., & Ritchie, J. T. 2013. Wheat yield response to spatially variable nitrogen fertilizer in Mediterranean environment. *European journal of Agronomy*, 51, 65–70.

Basso, B., Ritchie, J.T., Pierce, F.J., Braga, R.P., Jones, J.W., 2001. Spatial validation of crop models for precision agriculture. *Agric. Syst.* 68 (2), 97–112.

Batchelor, W.D., Basso, B., Paz, J.O., 2002. Examples of strategies to analyze

spatial and temporal yield variability using crop models. *Eur. J. Agron.* 18, 141–158.

Batte, M. T., & Ehsani, M. R. 2006. The economics of precision guidance with auto-boom control for farmer-owned agricultural sprayers. *Computers and Electronics in Agriculture*, 53(1), 28–44.

Beaudoin, N., Mary, B., Launay, M., Brisson, N., 2009. Conceptual Basis, Formalisations and Parameterization of the STICS Crop Model. Quae.

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M. L. & Bareth, G. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *International Journal of Applied Earth Observation and Geoinformation*, 39, 79–87.

Berni, J. A., Zarco-Tejada, P. J., Suárez, L., & Fereres, E. (2009). Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Transactions on geoscience and Remote Sensing*, 47(3), 722–738.

Beven, K. J., Kirkby, M. J., Freer, J. E., & Lamb, R. 2021. A history of TOPMODEL. *Hydrology and Earth System Sciences*, 25(2), 527–549.

Beza, E., Silva, J. V., Kooistra, L., & Reidsma, P. 2017. Review of yield gap explaining factors and opportunities for alternative data collection approaches. *European Journal of Agronomy*, 82, 206–222.

Boogaard, H., Wolf, J., Supit, I., Niemeyer, S., van Ittersum, M., 2013. A regional implementation of WOFOST for calculating yield gaps of autumn-sown wheat across the European Union. *Field Crops Res.* 143, 130–142.

Boogaard, H.L., De Wit, A.J.W., te Roller, J.A., Van Diepen, C.A., 2014. WOFOST CONTROL CENTRE 2.1; User's Guide for the WOFOST CONTROL CENTRE 2.1 and the Crop Growth Simulation Model WOFOST 7.1.7. Alterra, Wageningen, The Netherlands.

Booker, J.D., Lascano, R.J., Molling, C.C., Zartman, R.E., Acosta-Martínez, V., 2015. Temporal and spatial simulation of production-scale irrigated cotton systems. *Precis. Agric.* 16 (6), 630–653.

Boote, K.J., Jones, J.W., Pickering, N.B., 1996. Potential uses and limitations of crop models. *Agron. J.* 88, 704–716.

Boote, K.J., Jones, J.W., White, J.W., Asseng, S., Lizaso, J.I., 2013. Putting mechanisms into crop production models. *Plant Cell Environ.* 36 (9), 1658–1672.

Bouman, B.A.M., 1995. Crop modelling and remote sensing for yield prediction. *NJAS* 43, 143–161.

Bouman, B.A.M., van Kasteren, H.W.J., Uenk, D., 1992. Standard relations to estimate ground cover and LAI of agricultural crops from reflectance measurements. *Eur. J. Agron.* 1, 249–262.

Bouman, B.A.M., Van Keulen, H., Van Laar, H.H., Rabbinge, R., 1996. The 'school of de wit' crop growth simulation models: a pedigree and historical overview. *Agric. Syst.* 52 (2–3), 171–198.

Bouten, W., 1995. Soil water dynamics of the solling spruce stand, calculated with the FORHYD simulation package. *Ecol. Model.* 83 (1), 67–75.

Braden, H., 1985. Ein energiehaushalts-und verdunstungsmodell for wasser und stoff- haushaltsuntersuchungen landwirtschaftlich genutzer einzugsgebiete. *Mittelungen Deutsche Bodenkundliche Gesellschaft* 42 (Suppl), 294–299.

Bramley, R.G.V., 2009. Lessons from nearly 20 years of Precision Agriculture research, development, and adoption as a guide to its appropriate application. *Crop and Pasture Science*.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussière, F., Cabidoche, Y.M., Cellier, P., Debaeke, P., Gaudillère, J.P., Hénault, C., MarauX, F., Seguin, B., Sinoquet, H., 2003. An overview of the crop model stics. *Eur. J. Agron.* 18 (3), 309–332.

Brisson, N., Perrier, A., 1991. A semiempirical model of bare soil evaporation for crop simulation models. *Water Resour. Res.* 27 (5), 719–727.

Buchan, D.G., 2003. Richards' equation. In: Stewart, B.A., Howell, T.A. (Eds.), *Encyclopedia of Water Science*, pp. 809–811.

Buchan, D.G., Cameron, K.C., 2003. Darcy's law. In: Stewart, B.A., Howell, T.A. (Eds.), *Encyclopedia of Water Science*, pp. 109–112.

Bullock, D. S., & Lowenberg-DeBoer, J. 2007. Using spatial analysis to study

the values of variable rate technology and information. *Journal of Agricultural Economics*, 58(3), 517–535.

Byerlee, D., Edmeades, G., Fischer, T., 2014. Crop yields and global food security – will yield increase continue to feed the world? ACIAR.

Calera, A., Martínez, C., Melia, J., 2001. A procedure for obtaining green plant cover: Relation to NDVI in a case study for barley. *Int. J. Remote Sens.* 22, 3357–3362.

Campbell, G. S., & Norman, J. M. 1998. The light environment of plant canopies. In *An introduction to environmental biophysics* (pp. 247–278). Springer, New York, NY.

Campbell, G.S., 1985. *Soil Physics with BASIC: Transport Models for Soil-Plant Systems*. Elsevier.

Campbell, J.E., 1990. Dielectric properties and influence of conductivity in soils at one to fifty megahertz. *Soil Sci. Soc. Am. J.* 54, 332–341.

Campos, I., González-Gómez, L., Villo dre, J., Calera, M., Campoy, J., Jiménez, N., Plaza, C., Sánchez-Prieto, S., Calera, A., 2019. Mapping within-field variability in wheat yield and biomass using remote sensing vegetation indices. *Precis. Agric.* 20 (2), 214–236.

Campos, I., González-Gómez, L., Villo dre, J., Calera, M., Campoy, J., Jiménez, N., Plaza, C., Sánchez-Prieto, S., Calera, A., 2019. Mapping within-field variability in wheat yield and biomass using remote sensing vegetation indices. *Precis. Agric.* 20, 214–236.

Carberry, S.P., 2003. Crop development models. In: Stewart, B.A., Howell, T. (Eds.), *Encyclopedia of Water Science*. CRC Press, pp. 91–94.

Carlson, T.N., Gillies, R.R., Perry, E.M., 1994. A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover. *Remote Sens. Rev.* 9, 161–173.

Carlson, T.N., Ripley, D.A., 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sens. Environ.* 62, 241–252.

Casa, R., Silvestro, P.C., Yang, H., Pignatti, S., Pascucci, S., Yang, G., 2015. Development of farmland drought assessment tools based on the assimilation of

remotely sensed canopy biophysical variables into crop water response models, in: 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). ieeexplore.ieee.org, pp. 4005–4008.

Cassman, K. G. 1999. Ecological intensification of cereal production systems: yield potential, soil quality, and precision agriculture. *Proceedings of the National Academy of Sciences*, 96(11), 5952–5959.

Cassman, K. G., Dobermann, A., Walters, D. T., & Yang, H. 2003. Meeting cereal demand while protecting natural resources and improving environmental quality. *Annual Review of Environment and Resources*, 28(1), 315–358.

Cassman, K.G., 1999. Ecological intensification of cereal production systems: yield potential, soil quality, and precision agriculture. *Proc. Natl. Acad. Sci. U. S. A.* 96, 5952–5959.

Castellvi, F., Stockle, C.O., Perez, P.J., Ibanez, M., 2001. Comparison of methods for applying the Priestley-Taylor equation at a regional scale. *Hydrol. Process.* 15 (9), 1609–1620.

Chanzy, A., Chadoeuf, J., Gaudu, J.C., Mohrath, D., Richard, G., Bruckler, L., 1998. Soil moisture monitoring at the field scale using automatic capacitance probes. *Eur. J. Soil Sci.* 49, 637–648.

Chen, L., Young, M.H., 2006. Green-Ampt infiltration model for sloping surfaces. *Water Resour. Res.* 42 (7).

Chescheir, G.M., 2003. Drainage modeling. In: Stewart, B.A., Howell, T. (Eds.), *Encyclopedia of Water Science*. CRC Press, pp. 147–152.

Chianucci, F., Disperati, L., Guzzi, D., Bianchini, D., Nardino, V., Lastri, C., Rindinella, A., Corona, P., 2016. Estimation of canopy attributes in beech forests using true colour digital images from a small fixed-wing UAV. *Int. J. Appl. Earth Obs. Geoinf.* 47, 60–68.

Chowdhury, M. A. F., Meo, M. S., Uddin, A., & Haque, M. M. 2021. Asymmetric effect of energy price on commodity price: New evidence from NARDL and time frequency wavelet approaches. *Energy*, 231, 120934.

Ciha, A.J., 1984. Slope position and grain yield of soft white winter wheat. *Agron. J.* 76, 193–196.

Cihlar, J., Dobson, M.C., Schmugge, T., Hoogeboom, P., Janse, A.R.P., Baret, F., Guyot, G., Le Toan, T., Pampaloni, P. 1987. Review Article Procedures for the description of agricultural crops and soils in optical and microwave remote sensing studies. *International Journal of Remote Sensing*, 8(3), 427–439.

CNIG., 2019, Centro Nacional de Información Geográfica (CNIG). URL <http://centrodedescargas.cnig.es/CentroDescargas/index.jsp> (accessed 8.20.19).

Connor, D.J., Mínguez, M.I., 2012. Evolution not revolution of farming systems will best feed and green the world. *Global Food Security* 1, 106–113.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., & Böhner, J., 2015. System for automated geoscientific analyses (SAGA) v. 2.1. 4. *Geoscientific Model Development Discussions*, 8(2).

Cossani, C.M., Sadras, V.O., 2018. Water–nitrogen colimitation in grain crops. *Adv. Agron.* 150, 231–274.

Cronshey, R., 1986. Urban Hydrology for Small Watersheds. Technical report. US Dept. of Agriculture, Soil Conservation Service, Engineering Division.

de la Casa, A., Ovando, G., Bressanini, L., Martínez, J., Díaz, G., Miranda, C., 2018. Soybean crop coverage estimation from NDVI images with different spatial resolution to evaluate yield variability in a plot. *ISPRS J. Photogramm. Remote Sens.* 146, 531–547.

de la Casa, A.C., Ovando, G.G., Ravelo, A.C., Abril, E.G., Bergamaschi, H., 2014. Estimating maize ground cover using spectral data from Aqua-MODIS in Córdoba, Argentina. *Int. J. Remote Sens.* 35, 1295–1308.

De Veaux, R.D., Ungar, L.H., 1994. Multicollinearity: A tale of two nonparametric regressions, in: Selecting Models from Data. Springer New York, pp. 393–402.

de Vos, N.J., Rientjes, T.H.M., 2005. Constraints of artificial neural networks for rainfall-runoff modelling: trade-offs in hydrological state representation and model evaluation. *Hydrology and Earth System Sciences*. <https://doi.org/10.5194/hess-9-111-2005>.

de Wit, A., Boogaard, H., Fumagalli, D., Janssen, S., Knapen, R., van

Kraalingen, D., Supit, van der Wijngaart, R., van Diepen, K., 2018. 25 years of the WOFOST cropping systems model. *Agric. Syst.*

de Wit, C.T., 1965. Photosynthesis of Leaf Canopies. Technical Report 663. Pudoc,

De Wit, C.T., 1982. Simulation of living systems. *Simulation of Plant Growth and Crop Production* 3–8.

de Wit, C.T., van Keulen, H., 1987. Modelling production of field crops and its requirements. *Geoderma* 40, 253–265.

Deloitte, 2021. Deloitte's oil and gas price forecast. Deloitte prices forecast - <https://www2.deloitte.com/ca/en.html> (accessed 9.26.21).

Dente, L., Satalino, G., Mattia, F., Rinaldi, M., 2008. Assimilation of leaf area index derived from ASAR and MERIS data into CERES-Wheat model to map wheat yield. *Remote Sens. Environ.* 112, 1395–1407.

Derrien, M., Farki, B., Legléau, H., Sairouni, A., 1992. Vegetation cover mapping over France using NOAA-11/AVHRR. *Int. J. Remote Sens.* 13, 1787–1795.

Devia, G.K., Ganasri, B.P., Dwarakish, G.S., 2015. A review on hydrological models. *Aquat. Procedia* 4, 1001–1007.

DHI, 2017a. MIKE-SHE User Guide Manual. [http://manuals.mikepoweredbydhi.help/2017/MIKE\\_SHE.htm](http://manuals.mikepoweredbydhi.help/2017/MIKE_SHE.htm) (accessed 2019-3-NA).

DHI, 2017b. MIKE-SHE User Guide Manual V2. [http://manuals.mikepoweredbydhi.help/2017/MIKE\\_SHE.htm](http://manuals.mikepoweredbydhi.help/2017/MIKE_SHE.htm) (accessed 2019-3-NA).

Donatelli, M., Magarey, R.D., Bregaglio, S., Willocquet, L., Whish, J.P.M., Savary, S., 2017. Modelling the impacts of pests and diseases on agricultural systems. *Agric. Syst.* 155, 213–224.

Doraiswamy, P.C., Moulin, S., Cook, P.W., Stern, A., 2003. Crop Yield Assessment from Remote Sensing. *Photogrammetric Engineering & Remote Sensing* 69, 665–674.

Dorigo, W.A., Zurita-Milla, R., de Wit, A.J.W., Brazile, J., Singh, R., Schaepman,

M.E., 2007. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. *Int. J. Appl. Earth Obs. Geoinf.* 9, 165–193.

Drabik, D., & Peerlings, J. 2016. *Economics of Agribusiness*. Wageningen University.

Droogers, P., Bastiaanssen, W., 2000. Combining remote sensing and hydrological models to enhance spatial and temporal variability. *Remote Sensing and Hydrology 2000. IAHS-AISH Publ.* 267 574–579.

Duncan, W., Loomis, R., Williams, W., Hanau, R., et al., 1967. A model for simulating photosynthesis in plant communities. *Hilgardia* 38 (4), 181–205.

Dwarakish, G.S., Ganasri, B.P., 2015. Impact of land use change on hydrological systems: a review of current modeling approaches. *Cogent Geosci.* 1 (1), 1115691.

Eaton, A.K., Rouse, W.R., Lafleur, P.M., Marsh, P., Blanken, P.D., 2001. Surface energy balance of the western and central Canadian subarctic: variations in the energy balance among five major terrain types. *J. Clim.* 14 (17), 3692–3703.

Emerman, S.H., 1995. The tipping bucket equations as a model for macropore flow. *J. Hydrol.* 171 (1), 23–47.

Er-Raki, S., Chehbouni, A., Guemouria, N., Duchemin, B., Ezzahar, J., Hadria, R., 2007. Combining FAO-56 model and ground-based remote sensing to estimate water consumptions of wheat crops in a semi-arid region. *Agric. Water Manage.* 87 (1), 41–54.

Er-Raki, S., Chehbouni, A., Guemouria, N., Duchemin, B., Ezzahar, J., Hadria, R., 2007. Combining FAO-56 model and ground-based remote sensing to estimate water consumptions of wheat crops in a semi-arid region. *Agric. Water Manage.* 87, 41–54.

EUC., 2021. European Commission price dashboard  
[https://ec.europa.eu/info/food-farming-fisheries/farming/facts-and-figures/markets/prices/price-dashboard\\_en](https://ec.europa.eu/info/food-farming-fisheries/farming/facts-and-figures/markets/prices/price-dashboard_en) (accessed 18.10.21)

Evett, S.R., Schwartz, R.C., Tolk, J.A., Howell, T.A., 2009. Soil profile water content determination: Spatiotemporal variability of electromagnetic and

neutron probe sensors in access tubes. *Vadose Zone J.* 8, 926–941.

FAO., 2018. Yield gap analysis of field crops: Methods and case studies. Food & Agriculture Org.

FAO., 2021. Food Price Monitoring and Analysis Bulletin #8-2021

Feddes, R.A., Kowalik, P.J., Zaradny, H., 1978. Water uptake by plant roots. *Simulation of Field Water Use and Crop Yield.* pp. 16–30.

Feddes, R.A., Kowalik, P.J., Zaradny, H., 1978. Water uptake by plant roots. *Simulation of field water use and crop yield* 16–30.

FECA., 2021. Ayuda a los cultivos proteicos - importe unitario definitivo campaña 2020 <https://www.feca.gob.es/sites/default/files/Superficie Determinada IU DEFINITIVO CULTIVOS PROTEICOS-Ca 2020.pdf?token=eCCPTN25>. (accessed 18.01.22).

Finco, A., Bucci, G., Belletti, M., & Bentivoglio, D., 2021. The Economic Results of Investing in Precision Agriculture in Durum Wheat Production: A Case Study in Central Italy. *Agronomy*, 11(8), 1520.

Fischer, G., Van Velthuizen, H.T., Shah, M.M., et al., 2002. Global Agro-ecological Assessment for Agriculture in the 21st Century: Methodology and Results.

Fischer, R. A., 2015. Definitions and determination of crop yield, yield gaps, and of rates of change. *Field Crops Research*, 182, 9–18.

Fischer, R. A., & Connor, D. J., 2018. Issues for cropping and agricultural science in the next 20 years. *Field Crops Research*, 222, 121–142.

Fischer, R. A., Byerlee, D., & Edmeades, G., 2014. Crop yields and global food security. ACIAR: Canberra, ACT, 8–11. Fischer, R.A., Moreno Ramos, O.H., Ortiz Monasterio, I., Sayre, K.D., 2019. Yield response to plant density, row spacing and raised beds in low latitude spring wheat with ample soil resources: An update. *Field Crops Res.* 232, 95–105.

Florin, M. J., McBratney, A. B., & Whelan, B. M., 2009. Quantification and comparison of wheat yield variation across space and time. *European Journal of Agronomy*, 30(3), 212–219.

Franz, T.E., Pokal, S., Gibson, J.P., Zhou, Y., Gholizadeh, H., Tenorio, F.A.,

Rudnick, D., Heeren, D., McCabe, M., Ziliani, M., Jin, Z., Guan, K., Pan, M., Gates, J., Wardlow, B., 2020. The role of topography, soil, and remotely sensed vegetation condition towards predicting crop yield. *Field Crops Res.* 252, 107788.

French, R.J., Schultz, J.E., 1984. Water use efficiency of wheat in a Mediterranean-type environment. I. The relation between yield, water use and climate. *Aust. J. Agric. Res.* 35, 743–764.

Gamon, J.A., Field, C.B., Goulden, M.L., Griffin, K.L., Hartley, A.E., Joel, G., Penuelas, J., Valentini, R., 1995. Relationships Between NDVI, Canopy Structure, and Photosynthesis in Three Californian Vegetation Types. *Ecol. Appl.* 5, 28–41.

Gao, L., Li, D., 2014. A review of hydrological/water-quality models. *Front. Agric. Sci. Eng.* 1 (4), 267.

Gao, L., Wang, X., Johnson, B.A., Tian, Q., Wang, Y., Verrelst, J., Mu, X., Gu, X., 2020. Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: a review. *ISPRS J. Photogramm. Remote Sens.* 159, 364–377.

Gao, L., Wang, X., Johnson, B.A., Tian, Q., Wang, Y., Verrelst, J., Mu, X., Gu, X., 2020. Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: A review. *ISPRS J. Photogramm. Remote Sens.* 159, 364–377.

García-Ruiz, J.M., 2010. The effects of land uses on soil erosion in Spain: A review. *Catena* 81, 1–11.

García-Vila, M., Fereres, E., 2012. Combining the simulation crop model AquaCrop with an economic model for the optimization of irrigation management at farm level. *Eur. J. Agron.*

García-Vila, M., Fereres, E., Mateos, L., Orgaz, F., Steduto, P., 2009. Deficit irrigation optimization of cotton with AquaCrop. *Agron. J.* 101, 477–487.

Giraldez, J.V., Sposito, G., 1985. Infiltration in Swelling Soils. *Water Resour. Res.* 21, 33–44.

Gitelson, A.A., 2013. Remote estimation of crop fractional vegetation cover:

the use of noise equivalent as an indicator of performance of vegetation indices. *Int. J. Remote Sens.* 34, 6054–6066.

Gitelson, A.A., 2016. Remote Sensing Estimation of Crop Biophysical Characteristics at Various Scales. *Hyperspectral remote sensing of vegetation* 329.

Gitelson, A.A., Kaufman, Y.J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* 80, 76–87.

Glauber, J., & Laborde, D., 2022. How will Russia's invasion of Ukraine affect global food security?. *International Food Policy Research Institute*, 24.

Golden, H.E., Lane, C.R., Amatya, D.M., Bandilla, K.W., Raanan Kiperwas, H., Knights, C.D., Ssegane, H., 2014. Hydrologic connectivity between geographically isolated wetlands and surface water systems: a review of select modeling methods. *Environ. Model. Softw.* 53, 190–206.

Gómez, J.A., Sobrinho, T.A., Giráldez, J.V., Fereres, E., 2009. Soil management effects on runoff, erosion and soil properties in an olive grove of Southern Spain. *Soil Tillage Res.* 102, 5–13.

Goodwin, A.W., Lindsey, L.E., Harrison, S.K., Paul, P.A., 2018. Estimating Wheat Yield with Normalized Difference Vegetation Index and Fractional Green Canopy Cover. *Crop, Forage & Turfgrass Management* 4. <https://doi.org/10.2134/cftm2018.04.0026>

Grassini, P., van Bussel, L.G.J., Van Wart, J., Wolf, J., Claessens, L., Yang, H., Boogaard, H., de Groot, H., van Ittersum, M.K., Cassman, K.G., 2015. How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. *Field Crops Res.* 177, 49–63.

Green, W.H., Ampt, G.A., 1911. Studies on soil physics. *Soil Sci.* 4 (1), 1–24.

Griffin, T. W., 2006. Decision-making from on-farm experiments: spatial analysis of precision agriculture data.

Griffin, T. W., Shockley, J. M., Mark, T. B., Shannon, D. K., Clay, D. E., & Kitchen, N. R., 2018. Economics of precision farming. *Precision agriculture basics*, 1, 221-230.

Griffin, T.W., Fitzgerald, G.J., Lowenberg-DeBoer, J., Barnes, E.M., 2020.

Modeling local and global spatial correlation in field-scale experiments. *Agron.j.* 112, 2708–2721.

Günther, F., Fritsch, S., 2010. Neuralnet: Training of neural networks. *R J.* 2, 30.

Guilpart, N., Grassini, P., Sadras, V. O., Timsina, J., & Cassman, K. G. (2017). Estimating yield gaps at the cropping system level. *Field crops research*, 206, 21–32.

Guo, D., Zhao, R., Xing, X., Ma, X., 2019. Global sensitivity and uncertainty analysis of the AquaCrop model for maize under different irrigation and fertilizer management conditions. *Archives of Agronomy and Soil Science* 1–19.

Gutman, G., Ignatov, A., 1997. Satellite-derived green vegetation fraction for the use in numerical weather prediction models. *Adv. Space Res.* 19, 477–480.

Gutman, G., Ignatov, A., 1998. The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *Int. J. Remote Sens.* 19, 1533–1543.

Hallouin, T., Bruen, M., Christie, M., Bullock, C., Kelly-Quinn, M., 2018. Challenges in using hydrology and water quality models for assessing freshwater ecosystem services: a review. *Geosci. J.* 8 (2), 45.

Halvorson, G. A., & Doll, E. C., 1991. Topographic effects on spring wheat yields and water use. *Soil Science Society of America Journal*, 55(6), 1680–1685.

Han, C., Zhang, B., Chen, H., Wei, Z., Liu, Y., 2019. Spatially distributed crop model based on remote sensing. *Agric. Water Manage.* 218, 165–173.

Hansen, J.W., Jones, J.W., 2000. Scaling-up crop models for climate variability applications. *Agric. Syst.* 65 (1), 43–72.

Hansen, S., Abrahamsen, P., Petersen, C.T., Styczen, M., 2012. Daisy: model use, calibration, and validation. *Trans. ASABE* 55 (4), 1317–1333.

Hansen, S., Jensen, H.E., Nielsen, N.E., Svendsen, H., 1990. Daisy, A Soil Plant System Model. Danish Simulation Model for Transformation and Transport of Energy and Matter in the Soil Plant Atmosphere System. The National Agency for Environmental Protection, pp. 369.

Hargreaves, G.H., Samani, Z.A., 1982. Estimating potential

evapotranspiration. *J. Irrig. Drain. Div.* 108 (3), 225–230.

Hatfield, J. L., Gitelson, A. A., Schepers, J. S., & Walthall, C. L., 2008. Application of spectral remote sensing for agronomic decisions. *Agronomy Journal*, S-117.

Hawkins, R.H., 1973. Improved prediction of storm runoff in mountain watersheds. *J. Irrig. Drain. Div.* 99 (4), 519–523.

Hillel, D., 1980. *Fundamentals of Soil Physics*. Academic Press, New York.

Hochman, Z., van Rees, H., Carberry, P.S., Hunt, J.R., McCown, R.L., Gartmann, A., Holzworth, D., van Rees, S., Dalgliesh, N.P., Long, W., Peake, A.S., Poulton, P.L., McClelland, T., 2009. Re-inventing model-based decision support with Australian dryland farmers. 4. Yield Prophet helps farmers monitor and manage crops in a variable climate. *Crop Pasture Sci.* 60, 1057–1070.

Hoffman, G. J., & van Genuchten, M. T., 1983. "Soil properties and efficient water use: Water management for salinity control." Limitations to efficient water use in crop production (1983): 73–85.

Holzworth, D.P., Snow, V., Janssen, S., Athanasiadis, I.N., Donatelli, M., Hoogenboom, G., White, J.W., Thorburn, P., 2015. Agricultural production systems modelling and software: current status and future prospects. *Environ. Model. Softw.* 72, 276–286.

Hoogenboom, G., Porter, C.H., Shelia, V., Boote, K.J., Singh, U., White, J.W., Hunt, L.A., Ogoshi, R., Lizaso, J.I., Koo, J., et al., 2017. Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.7. DSSAT Foundation, Gainesville, FL. <https://DSSAT.net>.

Hsiao, T.C., 1973. Plant responses to water stress. *Annu. Rev. Plant Physiol.* 24 (1), 519–570.

Hsiao, T.C., 1993. Effects of Drought and Elevated CO<sub>2</sub> on Plant Water Use Efficiency and Productivity, in: *Interacting Stresses on Plants in a Changing Climate*. Springer Berlin Heidelberg, pp. 435–465.

Huang, J., Tian, L., Liang, S., Ma, H., Becker-Reshef, I., Huang, Y., ... & Wu, W., 2015. Improving winter wheat yield estimation by assimilation of the leaf area index from Landsat TM and MODIS data into the WOFOST model. *Agricultural*

and Forest Meteorology, 204, 106–121.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* 25, 295–309.

Hussein, F., Janat, M., Yakoub, A., 2011. Simulating cotton yield response to deficit irrigation with the FAO AquaCrop model. *Span. J. Agric. Res.* 9 (4), 1319–1330.

Imukova, K., Ingwersen, J., Streck, T., 2015. Determining the spatial and temporal dynamics of the green vegetation fraction of croplands using high-resolution RapidEye satellite images. *Agric. For. Meteorol.* 206, 113–123.

Jasinski, M.F., 1990. Sensitivity of the normalized difference vegetation index to subpixel canopy cover, soil albedo, and pixel scale. *Remote Sens. Environ.* 32, 169–187.

Jégo, G., Pattey, E., Liu, J., 2012. Using leaf area index, retrieved from optical imagery, in the STICS crop model for predicting yield and biomass of field crops. *Field Crops Res.* 131, 63–74.

Jensen, M.E., Burman, R.D., Allen, R.G., 1990. Evaporation and Irrigation Water Requirements. ASCE Manuals and Reports on Eng. Practices No. 70. Am. Soc. Civil Eng., New York, NY, pp. 978–990.

Jenson, S.K., Domingue, J.O., 1988. Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogramm. Eng. Remote Sens.* 54, 1593–1600.

Jia, Y., Shen, S., Niu, C., Qiu, Y., Wang, H., Liu, Y., 2011. Coupling crop growth and hydrologic models to predict crop yield with spatial analysis technologies. *JARS* 5 (1), 053537.

Jiang, Z., Huete, A.R., Chen, J., Chen, Y., Li, J., Yan, G., Zhang, X., 2006. Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. *Remote Sens. Environ.* 101, 366–378.

Jiang, Z., Huete, A.R., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.* 112, 3833–3845.

Jiménez-Muñoz, J.C., Sobrino, J.A., Plaza, A., Guanter, L., Moreno, J.,

Martinez, P., 2009. Comparison Between Fractional Vegetation Cover Retrievals from Vegetation Indices and Spectral Mixture Analysis: Case Study of PROBA/CHRIS Data Over an Agricultural Area. *Sensors* 9, 768–793.

Jin, X., Kumar, L., Li, Z., Feng, H., Xu, X., Yang, G., Wang, J., 2018. A review of data assimilation of remote sensing and crop models. *Eur. J. Agron.* 92, 141–152.

Jin, X., Li, Z., Feng, H., Ren, Z., & Li, S., 2020. Estimation of maize yield by assimilating biomass and canopy cover derived from hyperspectral data into the AquaCrop model. *Agricultural Water Management*, 227, 105846.

Jin, X., Li, Z., Nie, C., Xu, X., Feng, H., Guo, W., & Wang, J., 2018b. Parameter sensitivity analysis of the AquaCrop model based on extended fourier amplitude sensitivity under different agro-meteorological conditions and application. *Field Crops Research*, 226, 1–15.

Jin, X., Yang, G., Li, Z., Xu, X., Wang, J., Lan, Y., 2018a. Estimation of water productivity in winter wheat using the AquaCrop model with field hyperspectral data. *Precis. Agric.* 19, 1–17.

Johnson, L.F., Trout, T.J., 2012. Satellite NDVI assisted monitoring of vegetable crop evapotranspiration in California's San Joaquin Valley. *Remote Sensing* 4, 439–455.

Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., Munoz-Carpena, R., Porter, C.H., Rosenzweig, C., Wheeler, T.R., 2017a. Brief history of agricultural systems modeling. *Agric. Syst.* 155, 240–254.

Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., Munoz-Carpena, R., Porter, C.H., Rosenzweig, C., Wheeler, T.R., 2017b. Toward a new generation of agricultural system data, models, and knowledge products: state of agricultural systems science. *Agric. Syst.* 155, 269–288.

Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., Munoz-Carpena, R., Porter, C.H., Rosenzweig, C., Wheeler, T.R., 2017. Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural

systems science. *Agric. Syst.* 155, 269–288.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. *Eur. J. Agron.* 18 (3), 235–265.

Kamilaris, A., Kartakoullis, A., Prenafeta-Boldú, F.X., 2017. A review on the practice of big data analysis in agriculture. *Comput. Electron. Agric.* 143, 23–37.

Kauffeldt, A., Wetterhall, F., Pappenberger, F., Salamon, P., Thielen, J., 2016. Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environ. Model. Softw.* 75, 68–76.

Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. *Eur. J. Agron.* 18 (3), 267–288.

Kempenaar, C., Lokhorst, C., Bleumer, E.J.B., Veerkamp, R.F., 2016. Big Data analysis for smart farming: results of TO2 project in theme food security.

Kersebaum, K.C., 1989. Die Simulation der Stickstoff-Dynamik von Ackerböden. PhD thesis. Universität Hannover.

Khalfaoui, R., Baumöhl, E., Sarwar, S., & Výrost, T. (2021). Connectedness between energy and nonenergy commodity markets: Evidence from quantile coherency networks. *Resources Policy*, 74, 102318.

Kirby, M., Ahmad, M.D., Poulton, P., Zhu, Z., Lee, G., 2013. Review of Water, Crop Production and System Modelling Approaches for Food Security Studies in the Eastern Gangetic Plains. Accessed 2018-10-26. <https://www.researchgate.net/profile/Mobin-ud-DinAhmad/>.

Kirkegaard, J. A., & Hunt, J. R., 2010. Increasing productivity by matching farming system management and genotype in water-limited environments. *Journal of Experimental Botany*, 61(15), 4129–4143.

Kirkegaard, J.A., Lilley, J.M., Howe, G.N., Graham, J.M., 2007. Impact of subsoil water use on wheat yield. *Aust. J. Agric. Res.* 58, 303–315.

Klaij, M.C., Vachaud, G., 1992. Seasonal water balance of a sandy soil in Niger cropped with pearl millet, based on profile moisture measurements. *Agric. Water Manage.* 21, 313–330.

Koster, R.D., Suarez, M.J., 1994. The components of a 'SVAT' scheme and their effects on a GCM's hydrological cycle. *Adv. Water Resour.*

Kravchenko, A.N., Bullock, D.G., 2000. Correlation of Corn and Soybean Grain Yield with Topography and Soil Properties. *Agron. J.* 92, 75–83.

Kroes, J., Supit, I., Van Dam, J., Van Walsum, P., Mulder, M., 2017a. Impact of Capillary Rise and Recirculation on Crop Yields.

Kroes, J.G., van Dam, J.C., Bartholomeus, R.P., Groenendijk, P., Heinen, M., Hendriks, R.F.A., Mulder, H.M., Supit, I., van Walsum, P., 2017. SWAP Version 4, Theory Description and User Manual. Technical report. Wageningen University.

Krupnik, T.J., Andersson, J.A., Rusinamhodzi, L., Corbeels, M., Shennan, C., Gérard, B., 2019. Does size matters? A critical review of meta-analysis in Agronomy. *Exp. Agric.* 55, 200–229.

Lamb, J.A., Dowdy, R.H., Anderson, J.L., Rehm, G.W., 1997. Spatial and Temporal Stability of Corn Grain Yields. *J. Prod. Agric.* 10, 410–414.

Lasanta, T., García-Ruiz, J.M., Pérez-Rontomé, C., Sancho-Marcén, C., 2000. Runoff and sediment yield in a semi-arid environment: the effect of land management after farmland abandonment. *Catena* 38, 265–278.

Leroux, C., Jones, H., Clenet, A., Tisseyre, B., 2017. A new approach for zoning irregularly-spaced, within-field data. *Comput. Electron. Agric.* 141, 196–206.

Leroux, C., Jones, H., Taylor, J., Clenet, A., Tisseyre, B., 2018. A zone-based approach for processing and interpreting variability in multi-temporal yield data sets. *Comput. Electron. Agric.* 148, 299–308.

Li, K.Y., De Jong, R., Boisvert, J.B., 2001. An exponential root-water-uptake model with water stress compensation. *J. Hydrol.* 252 (1), 189–204.

Lobell, D. B., Cassman, K. G., & Field, C. B. (2009). Crop yield gaps: their importance, magnitudes, and causes. *Annual review of environment and*

resources, 34, 179–204.

Lobell, D.B., 2013. The use of satellite data for crop yield gap analysis. *Field Crops Res.* 143, 56–64.

Lobell, D.B., Azzari, G., 2017. Satellite detection of rising maize yield heterogeneity in the US midwest. *Environ. Res. Lett.* 12 (1), 014014.

Lobell, D.B., Cassman, K.G., Field, C.B., 2009. Crop yield gaps: their importance, magnitudes, and causes. *Annu. Rev. Environ. Resour.* 34 (1), 179–204.

Lobell, D.B., Thau, D., Seifert, C., Engle, E., Little, B., 2015. A scalable satellite-based crop yield mapper. *Remote Sens. Environ.* 164, 324–333.

Loomis, R.S., Connor, D.J., 1992. *Crop Ecology – Productivity and Management in Agricultural Systems*. Cambridge Press.

Loomis, R.S., Rabbinge, R., Ng, E., 1979. Explanatory models in crop physiology. *Annu. Rev. Plant Physiol.* 30 (1), 339–367.

Lorite, I.J., García-Vila, M., Santos, C., Ruiz-Ramos, M., Fereres, E., 2013. AquaData and AquaGIS: two computer utilities for temporal and spatial simulations of water-limited yield with AquaCrop. *Comput. Electron. Agric.* 96, 227–237.

Lovelace, R., Nowosad, J., & Muenchow, J., 2019. *Geocomputation with R*. CRC Press.

Lowenberg-DeBoer, J. M., & Erickson, B., 2019. Setting the record straight on precision agriculture adoption. *Agronomy Journal*.

Lukina, E.V., Stone, M.L., Raun, W.R., 1999. Estimating vegetation coverage in wheat using digital images. *J. Plant Nutr.* 22, 341–350.

Maestrini, B., Basso, B., 2018. Drivers of within-field spatial and temporal variability of crop yield across the US Midwest. *Sci. Rep.* 8, 14833.

Maestrini, B., Basso, B., 2018. Predicting spatial patterns of within-field crop yield variability. *Field Crops Res.*

Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling & Software* 15, 101–124.

Maina, F.Z., Siirila-Woodburn, E.R., 2020. The role of subsurface flow on evapotranspiration: A global sensitivity analysis. *Water Resour. Res.* 56. doi.org/10.1029/2019wr026612

Mateos, L., López-Cortijo, I., Sagardoy, J.A., 2002. SIMIS: the FAO decision support system for irrigation scheme management. *Agric. Water Manage.* 56 (3), 193–206.

McBratney, A., Whelan, B., Ancev, T., Bouma, J., 2005. Future directions of precision agriculture. *Precis. Agric.* 6 (1), 7–23.

McBratney, A., Whelan, B., Ancev, T., Bouma, J., 2005. Future Directions of Precision Agriculture. *Precis. Agric.* 6, 7–23.

McCutcheon, M.C., Farahani, H.J., Stednick, J.D., Buchleiter, G.W., Green, T.R., 2006. Effect of Soil Water on Apparent Soil Electrical Conductivity and Texture Relationships in a Dryland Field. *Biosystems Eng.* 94, 19–32.

Meng, J., Du, X., Wu, B., 2013. Generation of high spatial and temporal resolution NDVI and its application in crop biomass estimation. *International Journal of Digital Earth.*

Miller, M.P., Singer, M.J. & Nielsen, D.R., 1988. Spatial variability of wheat yield and soil properties on complex hills. *Soil Science Society of America Journal*, 52(4), pp.1133-1141.

Mirchandani, G., Cao, W., 1989. On hidden nodes for neural nets. *IEEE Transactions on Circuits and Systems* 36, 661–664.

Mohamed Sallah, A.-H., Tychon, B., Piccard, I., Gobin, A., Van Hoolst, R., Djaby, B., Wellens, J., 2019. Batch-processing of AquaCrop plug-in for rainfed maize using satellite derived Fractional Vegetation Cover data. *Agric. Water Manage.* 217, 346–355.

Mohsen, M.F.N., 1982. Some details of the Galerkin finite element method. *Appl. Math Model.* 6 (3), 165–170.

Moiling, C.C., Strikwerda, J.C., Norman, J.M., Rodgers, C.A., Wayne, R., Morgan, C.L.S., Diak, G.R., Mecikalski, J.R., 2005. Distributed runoff formulation designed for a precision agricultural landscape modeling system. *JAWRA J. Am. Water Resour. Assoc.* 41 (6), 1289–1313.

Montandon, L. M., & Small, E. E., 2008. The impact of soil reflectance on the quantification of the green vegetation fraction from NDVI. *Remote Sensing of Environment*, 112(4), 1835–1845.

Monteith, J.L., 1976. *Vegetation and the Atmosphere*. Academic Press, New York. Monteith, J.L., Unsworth, M.H., 1990. *Principles of Environmental Physics*, second ed. Edward Arnold, London 291pp.

Monzon, J.P., Calviño, P.A., Sadras, V.O., Zubiaurre, J.B., Andrade, F.H., 2018. Precision agriculture based on crop physiological principles improves whole-farm yield and profit: a case study. *Eur. J. Agron.* 99, 62–71.

Monzon, J.P., Calviño, P.A., Sadras, V.O., Zubiaurre, J.B., Andrade, F.H., 2018. Precision agriculture based on crop physiological principles improves whole-farm yield and profit: A case study. *Eur. J. Agron.* 99, 62–71.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE* 50, 885–900.

Mualem, Y., 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resour. Res.*

Mualem, Y., 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water resources research*, 12(3), 513–522.

Mulla, D. J., & Schepers, J. S., 1997. Key processes and properties for site-specific soil and crop management. The state of site specific management for agriculture, 1–18.

Murphy Jr., C.E., Knoerr, K.R., 1975. The evaporation of intercepted rainfall from a forest stand: an analysis by simulation. *Water Resour. Res.* 11 (2), 273–280.

Mwale, S. S., Azam-Ali, S. N., & Sparkes, D. L. 2005. Can the PR1 capacitance probe replace the neutron probe for routine soil-water measurement? *Soil use and management*, 21(3), 340–347.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I – A discussion of principles. *J. Hydrology* 10, 282–290.

Nendel, C., Berg, M., Kersebaum, K.C., Mirschel, W., Speck, X., Wegehenkel,

M., Wenkel, K.O., Wieland, R., 2011. The MONICA model: testing predictability for crop growth, soil moisture and nitrogen dynamics. *Ecol. Model.* 222 (9), 1614–1625.

Nielsen, D. C., & Halvorson, A. D., 1991. Nitrogen fertility influence on water stress and yield of winter wheat. *Agronomy journal*, 83(6), 1065–1070.

Nielsen, D., Biggar, J.W., Erh, K.T., 1973. Spatial variability of field-measured soil-water properties. *Hilgardia* 42, 215–259.

Nielsen, D.R., Alemi, M.H., 1989. Statistical opportunities for analyzing spatial and temporal heterogeneity of field soils. *Plant Soil* 115 (2), 285–296.

Nielsen, D.R., Van Genuchten, M.T., Jury, W.A., 1987. Monitoring and analyzing water and solute transport in the Vadose zone. *Proceedings of International Symposium on Groundwater Monitoring and Management* 23–28 hydrologie.org.

Nielsen, D.R., Wendroth, O., 2003. *Spatial and Temporal Statistics*. Schweizerbart'sche Verlagsbuchhandlung.

Padilla, F. L. M., Maas, S. J., González-Dugo, M. P., Mansilla, F., Rajan, N., Gavilán, P., & Domínguez, J., 2012. Monitoring regional wheat yield in Southern Spain using the GRAMI model and satellite imagery. *Field Crops Research*, 130, 145–154.

Paltineanu, I.C., Starr, J.L., 1997. Real-time soil water dynamics using multisensor capacitance probes: Laboratory calibration. *Soil Sci. Soc. Am. J.* 61, 1576–1585.

Panel, E. N. E., 2015. Nitrogen Use Efficiency (NUE) an indicator for the utilization of nitrogen in food systems. Wageningen University, Alterra, Wageningen, Netherlands.

Passioura, J.B., 1973. Sense and nonsense in crop simulation. *J. Aust. Inst. Agric. Sci.* 39 (3), 181–183.

Passioura, J.B., 1983. Roots and drought resistance. *Agric. Water Manage.* 7 (1), 265–280.

Passioura, J.B., 1996. Simulation models: science, snake oil, education, or engineering? *Agron. J.* 88, 690–694.

Passioura, J.B., 2002. Environmental biology and crop improvement. *Functional Plant Biology*, 29(5), 537–546.

Passioura, J.B., Angus, J.F., 2010. Chapter 2 - Improving Productivity of Crops in Water-Limited Environments, in: Sparks, D.L. (Ed.), *Advances in Agronomy*. Academic Press, pp. 37–75.

Pathak, H. S., Brown, P., & Best, T., 2019. A systematic literature review of the factors affecting the precision agriculture adoption process. *Precision Agriculture*, 20(6), 1292–1316.

Patrignani, A., Ochsner, T.E., 2015. Canopeo: A Powerful New Tool for Measuring Fractional Green Canopy Cover. *Agron. J.* 107, 2312–2320.

Pearl, J., 2019. The seven tools of causal inference, with reflections on machine learning. *Commun. ACM* 62, 54–60.

Pedersen, M. F., Gyldengren, J. G., Pedersen, S. M., Diamantopoulos, E., Gislum, R., & Styczen, M. E. 2021. A simulation of variable rate nitrogen application in winter wheat with soil and sensor information—An economic feasibility study. *Agricultural Systems*, 192, 103147.

Penman, H.L., 1948. Natural evaporation from open water, bare soil and grass. *Proc. R. Soc. Lond. A: Math. Phys. Sci.* 193 (1032), 120–145.

Penman, H.L., 1956. Estimating evaporation. *Trans. AGU* 37 (1), 43.

Penman, H.L., 1963. Vegetation and hydrology. *Soil Sci.* 96 (5), 357.

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J., Stenseth, N.C., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends Ecol. Evol.* 20, 503–510.

Philip, J.R., De Vries, D.A., 1957. Moisture movement in porous materials under temperature gradients. *Eos Trans. Amer. Geophys. Union* 38, 222–232.

Plant, R.E., 2001. Site-specific management: the application of information technology to crop production. *Comput. Electron. Agric.* 30, 9–29.

Poate, C.D., Daplyn, P.F., 1993. Data for Agrarian Development. CUP Archive.

Ponce, V.M., Hawkins, R.H., 1996. Runoff curve number: has it reached maturity? *J. Hydrol. Eng.* 1 (1), 11–19.

Prabhakara, K., Hively, W.D., McCarty, G.W., 2015. Evaluating the

relationship between biomass, percent groundcover and remote sensing indices across six winter cover crop fields in Maryland, United States. *Int. J. Appl. Earth Obs. Geoinf.* 39, 88–102.

Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon. Weather Rev.* 100 (2), 81–92.

Purevdorj, T.S., Tateishi, R., Ishiyama, T., Honda, Y., 1998. Relationships between percent vegetation cover and vegetation indices. *Int. J. Remote Sens.* 19, 3519–3535.

Qi, J., Marsett, R.C., Moran, M.S., Goodrich, D.C., Heilman, P., Kerr, Y.H., Dedieu, G., Chehbouni, A., Zhang, X.X., 2000. Spatial and temporal dynamics of vegetation in the San Pedro River basin area. *Agric. For. Meteorol.* 105, 55–68.

Quemada, M., Delgado, A., Mateos, L., & Villalobos, F. J., 2016. Nitrogen fertilization I: The nitrogen balance. In *Principles of agronomy for sustainable agriculture* (pp. 341–368). Springer

Quemada, M., Lassaletta, L., Jensen, L. S., Godinot, O., Brentrup, F., Buckley, C., ... & Oenema, O., 2020. Exploring nitrogen indicators of farm performance among farm types across several European case studies. *Agricultural Systems*, 177, 102689.

Raes, D., Steduto, P., Hsiao, T., Fereres, E., 2017. Chapter 3 calculation procedures. Reference Manual AquaCrop Version 6.0.

Raes, D., Steduto, P., Hsiao, T.C., Fereres, E., 2009a. *AquaCrop Reference Manual*. FAO, Land and Water Division, Rome, Italy.

Raes, D., Steduto, P., Hsiao, T.C., Fereres, E., 2009b. AquaCrop—the FAO crop model to simulate yield response to water: II. Main algorithms and software description. *Agron. J.* 101, 438–447.

Raes, D., Steduto, P., Hsiao, T.C., Fereres, E., 2012. Reference Manual AquaCrop (version 4.0). *AquaCrop*. <http://www.fao.org/nr/water/aquacrop.html>.

Rallison, R.E., 1980. Origin and evolution of the SCS runoff equation. *Symposium on Watershed Management 1980* 912–924.

Rattalino Edreira, J.I., Guilpart, N., Sadras, V., Cassman, K.G., van Ittersum, M.K., Schils, R.L.M., Grassini, P., 2018. Water productivity of rainfed maize and wheat: A local to global perspective. *Agric. For. Meteorol.* 259, 364–373.

Rattalino-Edreira, J. I., Mourtzinis, S., Conley, S. P., Roth, A. C., Ciampitti, I. A., Licht, M. A., Kandel, H., Kyveryga, P. M., Lindsey, L. E., Mueller, D. S. & Naeve, S. L., 2017. Assessing causes of yield gaps in agricultural areas with diversity in climate and soils. *Agricultural and forest meteorology*, 247, 170-180.

Rauff, K.O., Bello, R., 2015. A review of crop growth simulation models as tools for agricultural meteorology. *Agric. Sci. China* 6 (09), 1098.

Raun, W. R., Solie, J. B., Johnson, G. V., Stone, M. L., Mullen, R. W., Freeman, K. W., & Lukina, E. V. 2002. Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. *Agronomy Journal*, 94(4), 815-820.

Reitz, P., & Kutzbach, H. D., 1996. Investigations on a particular yield mapping system for combine harvesters. *Computers and electronics in agriculture*, 14(2-3), 137-150.

Richards, L.A., 1931. Capillary conduction of liquids through porous mediums. *Physics* 1 (5), 318–333.

Richardson, L.F., 1922. Weather Prediction by Numerical Process. Cambridge University Press.

Ritchie, J.T., 1972. Model for predicting evaporation from a row crop with incomplete cover. *Water Resour. Res.* 8 (5), 1204–1213.

Ritchie, J.T., 1981. Water dynamics in the soil-plant-atmosphere system. *Plant Soil* 58 (1–3), 81–96.

Ritchie, J.T., 1984. A user-orientated model of the soil water balance in wheat. In: Day, W., Atkin, R.K. (Eds.), *Wheat Growth and Modelling*. Springer US, Boston, MA, pp. 293–305.

Ritchie, J.T., 1998. Soil water balance and plant water stress. In: Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. (Eds.), *Understanding Options for Agricultural Production*. Springer Netherlands, Dordrecht, pp. 41–54.

Ritzema, H.P., 1994. Subsurface Flow to Drains. Technical Report 16.

Wageningen.

Roadknight, C.M., Balls, G.R., Mills, G.E., Palmer-Brown, D., 1997. Modeling complex environmental data. *IEEE Trans. Neural Netw.* 8, 852–862.

Robertson, M. J., Llewellyn, R. S., Mandel, R., Lawes, R., Bramley, R. G. V., Swift, L., Metz, N., O'Callaghan, C., 2012. Adoption of variable rate fertiliser application in the Australian grains industry: status, issues and prospects. *Precision Agriculture*, 13(2), 181–199.

Robertson, M. J., Lyle, G., & Bowden, J. W., 2008. Within-field variability of wheat yield and economic implications for spatially variable nutrient management. *Field Crops Research*, 105(3), 211–220.

Robertson, M., Isbister, B., Maling, I., Oliver, Y., Wong, M., Adams, M., & Tozer, P., 2007. Opportunities and constraints for managing within-field spatial variability in Western Australian grain production. *Field Crops Research*, 104(1-3), 60–67.

Rockström, J., Valentin, C., 1997. Hillslope dynamics of on-farm generation of surface water flows: The case of rain-fed cultivation of pearl millet on sandy soil in the Sahel. *Agricultural Water Management*.

Roggero, P.P., Trnka, M., Trombi, G., 2018. The response of process-based agro-ecosystem models to within-field variability in site conditions. *Field Crops Res.* 228, 1–19.

Rondeaux, G., Steven, M., Baret, F., 1996. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* 55, 95–107.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A.C., Müller, C., Arneth, A., Boote, K.J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T.A.M., Schmid, E., Stehfest, E., Yang, H., Jones, J.W., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proc. Natl. Acad. Sci. U. S. A.* 111 (9), 3268–3273.

RoTimi Ojo, E., Bullock, P. R., & Fitzmaurice, J., 2015. Field performance of five soil moisture instruments in heavy clay soils. *Soil Science Society of America Journal*, 79(1), 20–29.

Sadler, E.J., Russell, G., 1997. Modeling crop yield for site-specific

management. In: Pierce, F.J., Sadler, E.J. (Eds.), *The State of Site-Specific Management for Agriculture*.

Sadras, V., 2002. Interaction between rainfall and nitrogen fertilisation of wheat in environments prone to terminal drought: economic and environmental risk analysis. *Field Crops Research*, 77(2-3), 201–215.

Sadras, V., Alston, J., Aphalo, P., Connor, D., Denison, R.F., Fischer, T., Gray, R., Hayman, P., Kirkegaard, J., Kirchmann, H., Kropff, M., Lafitte, H.R., Langridge, P., Lenne, J., Mínguez, M.I., Passioura, J., Porter, J.R., Reeves, T., Rodriguez, D., Ryan, M., Villalobos, F.J. & Wood, D., 2020. Making science more effective for agriculture, in: *Advances in Agronomy*. Academic Press.

Sadras, V., Bongiovanni, R., 2004. Use of Lorenz curves and Gini coefficients to assess yield inequality within paddocks. *Field Crops Res.* 90, 303–310.

Sadras, V.O., Angus, J.F., 2006. Benchmarking water-use efficiency of rainfed wheat in dry environments. *Aust. J. Agric. Res.* 57, 847–856.

Sadras, V.O., Reynolds, M.P., de la Vega, A.J., Petrie, P.R., Robinson, R., 2009. Phenotypic plasticity of yield and phenology in wheat, sunflower and grapevine. *Field Crops Res.* 110, 242–250.

Sadras, V.O., Villalobos, F.J., Orgaz, F., Fereres, E., 2016. Effects of water stress on crop production. In: Villalobos, F.J., Fereres, E. (Eds.), *Principles of Agronomy for Sustainable Agriculture*. Springer International Publishing, Cham, pp. 189–204.

Saint-Venant, A.J.C.B., 1871. Theorie du mouvement non-permanent des eaux avec application aux crues des rivers et a l'introduction des mareas dans leur lit. *Acad. Sci. C. R.* 73 (99), 148–154.

Salvadore, E., Bronders, J., Batelaan, O., 2015. Hydrological modelling of urbanized catchments: a review and future directions. *J. Hydrol.* 529, 62–81.

Schaap, M.G., Leij, F.J., van Genuchten, M.T., 2001. rosetta: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of Hydrology*. [https://doi.org/10.1016/s0022-1694\(01\)00466-8](https://doi.org/10.1016/s0022-1694(01)00466-8)

Scheftic, W., Zeng, X., Broxton, P., Brunke, M., 2014. Intercomparison of Seven NDVI Products over the United States and Mexico. *Remote Sensing* 6,

1057–1084.

Schils, R., Olesen, J.E., Kersebaum, K.-C., Rijk, B., Oberforster, M., Kalyada, V., Khitrykau, M., Gobin, A., Kirchev, H., Manolova, V., Manolov, I., Trnka, M., Hlavinka, P., Palosuo, T., Peltonen-Sainio, P., Jauhiainen, L., Lorgeou, J., Marrou, H., Danalatos, N., Archontoulis, S., Fodor, N., Spink, J., Roggero, P.P., Bassu, S., Pulina, A., Seehusen, T., Uhlen, A.K., Żyłowska, K., Nieróbca, A., Kozyra, J., Silva, J.V., Maçãs, B.M., Coutinho, J., Ion, V., Takáč, J., Mínguez, M.I., Eckersten, H., Levy, L., Herrera, J.M., Hiltbrunner, J., Kryvobok, O., Kryvoshein, O., Sylvester-Bradley, R., Kindred, D., Topp, C.F.E., Boogaard, H., de Groot, H., Lesschen, J.P., van Bussel, L., Wolf, J., Zijlstra, M., van Loon, M.P., van Ittersum, M.K., 2018. Cereal yield gaps across Europe. *European Journal of Agronomy*, 101, 109–120.

Schimmelpfennig, D., 2016. Farm profits and adoption of precision agriculture. USDA.

Schmitter, P., Zwart, S.J., Danvi, A., Gbaguidi, F., 2015. Contributions of lateral flow and groundwater to the spatio-temporal variation of irrigated rice yields and water productivity in a West-African inland valley. *Agric. Water Manage.* 152, 286–298.

Schnitkey, G., Swanson, K., Paulson, N., & Zulauf, C., 2021. Weekly Farm Economics: Fertilizer Price Increases for 2021 Production.

Scopel, E., Muller, B., Arreola-Tostado, J.M., Chavez Guerra, E., MarauX, F., 1998. Quantifying and modeling the effects of a light crop residue mulch on the water balance: an application to rainfed maize in western Mexico. *Congrès Mondial des Sciences du sol*. 16. CIRAD-CA, El Grullo Jalisco.

Scott, S.P., 1983. Hysteretic effects on net infiltration. *Adv. Infiltr.* 163–170.

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? *Eur. J. Agron.* 94, 25–35.

Senthil Kumar, A.R., Sudheer, K.P., Jain, S.K., Agarwal, P.K., 2005. Rainfall-runoff modelling using artificial neural networks: comparison of network types. *Hydrological Processes: An International Journal* 19, 1277–1291.

Shaw, R.E., Meyer, W.S., McNeill, A., Tyerman, S.D., 2013. Waterlogging in Australian agricultural landscapes: a review of plant responses and crop

models. *Crop Pasture Sci.* 64 (6), 549–562.

Shu, Y., Li, H., Lei, Y., 2018. Modelling groundwater flow with MIKE SHE using conventional climate data and satellite data as model forcing in Haihe plain, China. *Water* 10 (10), 1295.

Shukla, M.B., Kok, R., Prasher, S.O., Clark, G., Lacroix, R., 1996. Use of artificial neural networks in transient drainage design. *Trans. ASAE* 39, 119–124.

Sida, T.S., Chamberlin, J., Ayalew, H., Kosmowski, F., Craufurd, P., 2021. Implications of intra-plot heterogeneity for yield estimation accuracy: Evidence from smallholder maize systems in Ethiopia. *Field Crops Res.* 267, 108147.

Silva, J. V., 2017. Using yield gap analysis to give sustainable intensification local meaning (Doctoral dissertation, Wageningen University and Research).

Silva, J. V., van Ittersum, M. K., ten Berge, H. F., Spätjens, L., Tenreiro, T. R., Anten, N. P., & Reidsma, P., 2021. Agronomic analysis of nitrogen performance indicators in intensive arable cropping systems: An appraisal of big data from commercial farms. *Field Crops Research*, 269, 108176.

Silva, J. V., Tenreiro, T.R., Spätjens, L., Anten, N.P.R., van Ittersum, M.K., Reidsma, P., 2020. Can big data explain yield variability and water productivity in intensive cropping systems? *Field Crops Res.* 255, 107828.

Silvestro, P.C., Pignatti, S., Pascucci, S., Yang, H., Li, Z., Yang, G., Huang, W., Casa, R., 2017. Estimating wheat yield in china at the field and district scale from the assimilation of satellite data into the aquacrop and simple algorithm for yield (SAFY) models. *Remote Sens.* 9 (5), 509.

Silvestro, P.C., Pignatti, S., Pascucci, S., Yang, H., Li, Z., Yang, G., Huang, W., Casa, R., 2017. Estimating Wheat Yield in China at the Field and District Scale from the Assimilation of Satellite Data into the Aquacrop and Simple Algorithm for Yield (SAFY) Models. *Remote Sensing* 9, 509.

Silvestro, P.C., Pignatti, S., Yang, H., Yang, G., Pascucci, S., Castaldi, F., Casa, R., 2017. Sensitivity analysis of the Aquacrop and SAFYE crop models for the assessment of water limited winter wheat yield in regional scale applications. *PLoS One* 12, e0187485.

Simunek, J., Hopmans, J.W., 2009. Modeling compensated root water and nutrient up- take. *Ecol. Model.* 220 (4), 505–521.

Simunek, J., Šejna, M., Saito, M., Sakai, M., van Genuchten, M.T., 2018a. HYDRUS-1D Version 4.17 – Manual. Technical report. Department of Environmental Science, University of California Riverside.

Šimůnek, J., Šejna, M., Saito, M., Sakai, M., van Genuchten, M.T., 2018. HYDRUS-1D Version 4.17 – Manual. Department of Environmental Science, University of California Riverside.

Simunek, J., Sejna, M., Van Genuchten, M.T., 1999. The HYDRUS-2D Software Package. International Ground Water Modeling Center.

Simunek, J., Šejna, M., van Genuchten, M.T., 2018b. New features of version 3 of the HYDRUS (2D/3D) computer software package. *J. Hydrol. Hydromech./Vodohospo. Cas.* 66 (2), 133–142.

Simunek, J., Sejna, M., Van Genuchten, M.T., Šimůnek, J., Šejna, M., Jacques, D., Šimůnek, J., Mallants, D., Saito, H., Sakai, M., 1998. HYDRUS-1D. Simulating the One-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-Saturated Media, Version 2.

Šimunek, J., Van Genuchten, M. T., & Šejna, M., 2012. HYDRUS: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4), 1263–1274.

Simunek, J., van Genuchten, M.T., 2008. Modeling nonequilibrium flow and transport processes using HYDRUS. *Vadose Zone J.* 7, 782–797.

Simunek, J., van Genuchten, M.T., Šejna, M., 2008. Development and applications of the HYDRUS and STANMOD software packages and related codes. *Vadose Zone J.* 7, 587–600.

Skaggs, R.W., Chescheir, G.M., 1999. Application of drainage simulation models. *Agricultural Drainage, Agronomy Monograph*. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI, pp. 537–564.

Smith, M.J., 2020. Getting value from artificial intelligence in agriculture. *Anim. Produc. Sci.* 60, 46–54.

Snyder, C. S., Bruulsema, T. W., Jensen, T. L., & Fixen, P. E., 2009. Review

of greenhouse gas emissions from crop production systems and fertilizer management effects. *Agriculture, Ecosystems & Environment*, 133(3-4), 247–266.

Soil Survey Staff, 1999. A basic system of soil classification for making and interpreting soil surveys. *Soil Taxonomy*. 2nd ed., USDA Agr.Hbk. 436, WA.

Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M., Xu, C., 2015. Global sensitivity analysis in hydrological modeling: review of concepts, methods, theoretical framework, and applications. *J. Hydrol.* 523, 739–757.

Sood, A., Smakhtin, V., 2015. Global hydrological models: a review. *Hydrol. Sci. J.* 60 (4), 549–565.

Soriano, M. A., Cabezas, J. M., Gómez, J. A., 2018. Soil water content and yield a vertisol in a rain-fed olive grove under four different soil management practices in a four year experiment. In *EGU General Assembly Conference Abstracts* (p. 5390).

Spiertz, H., 2014. Agricultural sciences in transition from 1800 to 2020: Exploring knowledge and creating impact. *Eur. J. Agron.* 59, 96–106.

Steduto, P., Hsiao, T. C., Fereres, E., & Raes, D., 2012. Crop yield response to water (FAO Irrigation and drainage paper 66). Rome: Food and Agriculture Organization of the United Nations.

Steduto, P., Hsiao, T.C., Fereres, E., 2007. On the conservative behavior of biomass water productivity. *Irrig. Sci.* 25, 189–207.

Steduto, P., Hsiao, T.C., Raes, D., Fereres, E., 2009. AquaCrop—The FAO Crop Model to Simulate Yield Response to Water: I. Concepts and Underlying Principles. *Agron. J.* 101, 426–437

Steven, Biscoe, P.V., Jaggard, K.W., Paruntu, J., 1986. Foliage cover and radiation interception. *Field Crops Res.* 13, 75–87.

Stewart, G., 2010. Meta-analysis in applied ecology. *Biol. Lett.* 6, 78–81.

Struik, P.C., 2016. Bridging the genotype–phenotype gap in 3D. *J. Expt. Bot.* 67 (15), 4427–4430.

Swinton, S. M., & Ahmad, M., 1996. Returns to farmer investments in precision agriculture equipment and services. In *Proceedings of the Third International Conference on Precision Agriculture* (pp. 1009–1018). Madison, WI, USA:

American Society of Agronomy, Crop Science Society of America, Soil Science Society of America.

Tarboton, D.G., Bras, R.L., Rodriguez-Iturbe, I., 1991. On the extraction of channel networks from digital elevation data. *Hydrol. Process.* 5, 81–100.

Team, R.C., 2000. R language Definition. R foundation for statistical computing, Vienna, Austria.

Tenreiro, T. R., García-Vila, M., Gómez, J. A., & Fereres, E., 2020b. Uncertainties associated with the delineation of management zones in precision agriculture. In EGU General Assembly Conference Abstracts (p. 5709).

Tenreiro, T. R., García-Vila, M., Gómez, J. A., Jimenez-Berni, J. A., & Fereres, E., 2020. Water modelling approaches and opportunities to simulate spatial water variations at crop field level. *Agricultural Water Management*, 240, 106254.

Tenreiro, T. R., García-Vila, M., Gómez, J. A., Jiménez-Berni, J. A., & Fereres, E., 2021. Using NDVI for the assessment of canopy cover in agricultural crops within modelling research. *Computers and Electronics in Agriculture*, 182, 106038.

Tenreiro, T. R., Jeřábek, J., Gómez, J. A., Zumr, D., Martínez, G., García-Vila, M., Fereres, E., 2022. Simulating water lateral inflow and its contribution to spatial variations of rainfed wheat yields. *European Journal of Agronomy*, 137, 126515.

Thomas, D.L., Smith, M.C., 2003. Hydrologic process modeling. In: Stewart, B.A., Howell, T. (Eds.), *Encyclopedia of Water Science*. CRC Press, pp. 418–420.

Thorp, K.R., Jaynes, D.B., Malone, R.W., 2008. Simulating the long-term performance of drainage water management across the Midwestern United States. *Trans. ASABE* 51 (3), 961–976.

Thorp, K.R., White, J.W., Porter, C.H., Hoogenboom, G., Nearing, G.S., French, A.N., 2012. Methodology to evaluate the performance of simulation models for alternative compiler and operating system configurations. *Comput. Electron. Agric.* 81, 62–71.

Tibshirani, R., Walther, G., Hastie, T., 2001. Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. Series B Stat. Methodol.* 63,

411–423.

Tinker, P.B., 1976. Roots and water-transport of water to plant roots in soil. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 273 (927), 445–461.

Todd, S.W., Hoffer, R.M., 1998. Responses of spectral indices to variations in vegetation cover and soil background. *Photogramm. Eng. Remote Sens.* 64, 915–922.

Tolk, J.A., Howell, T.A., Steiner, J.L., Krieg, D.R., Schneider, A.D., 1995. Role of transpiration suppression by evaporation of intercepted water in improving irrigation efficiency. *Irrig. Sci.* 16.

Toreti, A., Maiorano, A., De Sanctis, G., Webber, H., Ruane, A.C., Fumagalli, D., Ceglar, A., Niemeyer, S., Zampieri, M., 2018. Using reanalysis in crop monitoring and forecasting systems. *Agric. Syst.* <https://doi.org/10.1016/j.agrsy.2018.07.001>.

Torralba, M.A., 2013. Evaluación de la erosión hídrica en parcelas experimentales en campos agrícolas de secano mediterráneo (Doctoral dissertation, Universidad Complutense de Madrid).

Tozer, P. R., 2009. Uncertainty and investment in precision agriculture—Is it worth the money? *Agricultural systems*, 100(1–3), 80–87.

Trout, T.J., Johnson, L.F., Gartung, J., 2008. Remote sensing of canopy cover in horticultural crops. *HortScience* 43, 333–337.

USDA, 2021. Food Price Outlook, 2021 – Economic Research Service USDA. URL <https://www.ers.usda.gov/data-products/food-price-outlook/summary-findings/> (accessed 9.20.21).

van Dam, J.C., 2000. Field-scale Water Flow and Solute Transport: SWAP Model Concepts, Parameter Estimation and Case Studies. PhD thesis. [s.n.], S.l.

van Dam, J.C., Feddes, R.A., 2000. Numerical simulation of infiltration, evaporation and shallow groundwater levels with the Richards equation. *J. Hydrol.* 233 (1), 72–85.

van Dam, J.C., Huygen, J., Wesseling, J.G., Feddes, R.A., Kabat, P., Van Walsum, P.E.V., Groenendijk, P., Van Diepen, C.A., 1997. Theory of SWAP Version 2.0; Simulation of Water Flow, Solute Transport and Plant Growth in

the Soil-Water-Atmosphere-Plant Environment. Technical report. DLO Winand Staring Centre.

van der Keur, P., Hansen, S., Schelde, K., Thomsen, A., 2001. Modification of DAISY SVAT model for potential use of remotely sensed data. *Agric. For. Meteorol.* 106 (3), 215–231.

van Genuchten, M.T., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* 44 (5), 892–898.

van Genuchten, M.T., 1985. Convective-dispersive transport of solutes involved in sequential first-order decay reactions. *Comput. Geosci.* 11 (2), 129–147.

van Genuchten, M.T., 1987. A Numerical Model for Water and Solute Movement in and Below the Root Zone. United States Department of Agriculture Agricultural Research Service US.

van Ittersum, M.K., Leffelaar, P.A., van Keulen, H., Kropff, M.J., Bastiaans, L., Goudriaan, J., 2003. On approaches and applications of the Wageningen crop models. *Eur. J. Agron.* 18 (3), 201–234.

Van Leeuwen, W.J.D., Orr, B.J., Marsh, S.E., 2006. Multi-sensor NDVI data continuity: Uncertainties and implications for vegetation monitoring applications. *Remote Sens. Environ.*

Van Herwaarden, A. F., Angus, J. F., Richards, R. A., & Farquhar, G. D., 1998. 'Haying-off', the negative grain yield response of dryland wheat to nitrogen fertiliser II. Carbohydrate and protein dynamics. *Australian Journal of Agricultural Research*, 49(7), 1083-1094

van Van Diepen, C.A., Wolf, J., Van Keulen, H., Rappoldt, C., 1989. WOFOST: a simulation model of crop production. *Soil Use Manage.* 5 (1), 16–24.

Vanuytrecht, E., Raes, D., Willems, P., 2014. Global sensitivity analysis of yield output from the water productivity model. *Environmental Modelling & Software* 51, 323–332.

Verburg, K., 1996. Methodology in Soil Water and Solute Balance Modelling: An Evaluation of the APSIM-SoilWat and SWIMv2 Models. CSIRO Division of Soils.

- Verburg, K., Ross, P.J., Bristow, K.L., 1996. *Swimv2.1 User Manual*.
- Verger, A., Martínez, B., Camacho-de Coca, F., 2009. Accuracy assessment of fraction of vegetation cover and leaf area index estimates from pragmatic methods in a cropland area. *Journal of Remote Sensing*.
- Verhagen, A., Bootink, H.W.G., Bouma, J., 1995. Site-specific management: Balancing production and environmental requirements at farm level. *Agric. Syst.* 49, 369–384.
- Verhagen, J., Bouma, J., 1997. Modeling soil variability. In: *The State of Site-Specific Management for Agriculture*. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI, pp. 55–67.
- Viña, A., Gitelson, A.A., Guy-Robertson, A.L., Peng, Y., 2011. Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote Sens. Environ.* 115, 3468–3478.
- Viswanadham, Y., Silva Filho, V.P., André, R.G.B., 1991. The Priestley-Taylor parameter for the Amazon forest. *For. Ecol. Manage.* 38 (3), 211–225.
- von Hoyningen-Huene, J., 1981. Die Interzeption des Niederschlags in land- wirtschaftlichen Pflanzenbeständen. *Arbeitsbericht Deutscher Verband für Wasserwirtschaft und Kulturbau*, DVWK.
- Vos, R., Glauber, J., Hernández, M., & Laborde, D., 2022. 10. COVID-19 and food inflation scares. *COVID-19 and global food security: Two years later*, 64. Wageningen.
- Waldner, F., Horan, H., Chen, Y., & Hochman, Z., 2019. High temporal resolution of leaf area data improves empirical estimation of grain yield. *Scientific reports*, 9(1), 1–14.
- Wallor, E., Kersebaum, K.-C., Ventrella, D., Bindi, M., Cammarano, D., Coucheney, E., Gaiser, T., Garofalo, P., Giglio, L., Giola, P., Hoffmann, M.P., Iocola, I., Lana, M., Lewan, E., Maharjan, G.R., Moriondo, M., Mula, L., Nendel, C., Pohankova, E., Roggero, P.P., Trnka, M., Trombi, G., 2018. The response of process-based agro-ecosystem models to within-field variability in site conditions. *Field Crops Res.* 228, 1–19.

Wallor, E., Kersebaum, K.-C., Ventrella, D., Bindi, M., Cammarano, D., Coucheney, E., Gaiser, T., Garofalo, P., Giglio, L., Giola, P., Hoffmann, M.P., Iocola, I., Lana, M., Lewan, E., Maharjan, G.R., Moriondo, M., Mula, L., Nendel, C., Pohankova, E., Roggero, P.P., Trnka, M., Trombi, G., 2018. The response of process-based agro-ecosystem models to within-field variability in site conditions. *Field Crops Res.* 228, 1–19.

Wang, H., Shi, H., Wang, Y., 2015. The wetting of leaf surfaces and its ecological significances. In: Aliofkhazraei, M. (Ed.), *Wetting and Wettability*. InTech.

Wang, L., Liu, H., 2006. An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. *Int. J. Geogr. Inf. Sci.* 20, 193–213.

Ward, N.K., Maureira, F., Stöckle, C.O., Brooks, E.S., Painter, K.M., Yourek, M.A., Gasch, C.K., 2018. Simulating field-scale variability and precision management with a 3D hydrologic cropping systems model. *Precis. Agric.* 19 (2), 293–313.

Ward, N.K., Maureira, F., Stöckle, C.O., Brooks, E.S., Painter, K.M., Yourek, M.A., Gasch, C.K., 2018. Simulating field-scale variability and precision management with a 3D hydrologic cropping systems model. *Precis. Agric.* 19, 293–313.

Weiss, E., Marsh, S.E., Pfirman, E.S., 2001. Application of NOAA-AVHRR NDVI time-series data to assess changes in Saudi Arabia's rangelands. *Int. J. Remote Sens.*

Welsh, J. P., Wood, G. A., Godwin, R. J., Taylor, J. C., Earl, R., Blackmore, S., & Knight, S. M., 2003. Developing strategies for spatially variable nitrogen application in cereals, part II: wheat. *Biosystems Engineering*, 84(4), 495–511.

Wesseling, J. G., Elbers, J. A., Kabat, P., & Van Den Broek, B. J., 1991. SWATRE: Instructions for Input. Internal Note, Winand Staring Centre, Wageningen, The Netherlands. International Waterlogging and Salinity Research Institute, Lahore, Pakistan, 29.

Whelan, B. M., & McBratney, A. B., 2000. The “null hypothesis” of precision agriculture management. *Precision Agriculture*, 2(3), 265–279.

Whisler, F.D., Acock, B., Baker, D.N., Fye, R.E., Hodges, H.F., Lambert, J.R., Lemmon, H.E., McKinnion, J.M., Reddy, V.R., 1986. Crop simulation models in agronomic systems. *Adv. Agron.* 40 (1), 41–208.

Wickham, H., 2007. The ggplot package. URL: <https://cran.r-project.org/web/packages/ggplot2/index.html>

Wiegand, C.L., Richardson, A.J., Escobar, D.E., Gerbermann, A.H., 1991. Vegetation indices in crop assessments.

Willmott, C.J., 1981. On the validation of models. *Phys. Geogr.* 2, 184–194.

Wittich, K.P., Hansing, O., 1995. Area-averaged vegetative cover fraction estimated from satellite data. *Int. J. Biometeorol.* 38, 209–215.

Wolfert, S., Ge, L., Verdouw, C., Bogaardt, M.-J., 2017. Big Data in Smart Farming – A review. *Agric. Syst.* 153, 69–80.

Woodward, D.E., Hawkins, R.H., Jiang, R., Hjelmfelt, J.A.T., Van Mullem, J.A., Quan, Q.D., 2003. Runoff curve number method: examination of the initial abstraction ratio. World Water & Environmental Resources Congress 2003.

Xue, J., Su, B., 2017. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *Journal of Sensors* 2017. <https://doi.org/10.1155/2017/1353691>

Yang, J.M., Yang, J.Y., Liu, S., Hoogenboom, G., 2014. An evaluation of the statistical methods for testing the performance of crop models with observed data. *Agric. Syst.* 127, 81–89.

Yin, X., Struik, P.C., 2007. *Crop Systems Biology*. Springer, New York.

Zhou, H., Zhao, W.Z., 2019. Modeling soil water balance and irrigation strategies in a flood-irrigated wheat-maize rotation system. A case in dry climate, china. *Agric. Water Manage.* 221, 286–302.

Zumr, D., Dohnal, M., Hrnčíř, M., Císlarová, M., Vogel, T., Doležal, F., 2006. Simulation of soil water dynamics in structured heavy soils with respect to root water uptake. *Biologia* 61, S320–S323.

Zwart, S.J., Bastiaanssen, W.G.M., 2007. SEBAL for detecting spatial variation of water productivity and scope for improvement in eight irrigated wheat systems. *Agric. Water Manage.* 89 (3), 287–296.





# Summary

In sloping fields, rainfed crops experience different degrees of water stress caused by spatial variations in water and, consequently, yields also vary spatially within a field. This offers opportunities for precision agriculture through site-specific management. However, while significant advances have been accomplished in the engineering aspects of precision agriculture, such as increasing spatial resolution of data systems and automation, much less effort has been dedicated to the simulation of within field crop responses to spatial variations. Most studies on rainfed yield gaps ignore intra-plot variability, but if crop models are to be used in assisting site-specific management, they may greatly benefit from spatial water modelling approaches capable of accurately representing and simulating within-field variation of water-related processes.

This doctoral thesis represents a novel contribution to the agronomy of rainfed agricultural systems, evaluating the role played by water flows in areas of undulating topography in determining the spatial variations of wheat yield. The thesis has been carried out in chapters that are associated by following an integrative approach.

The thesis first reviewed some of the most widely adopted crop and hydrologic models and explored new opportunities for simulating spatial water variations at crop field level through the incorporation of lateral inflow at lower elevation zones of the field. From this standpoint, the spatial variations of yield gaps in rainfed wheat, caused by lateral flows from high to low areas, were assessed in Córdoba, Spain.

From an agronomic perspective, water lateral inflows (LIF) due to surface and subsurface runoff contribute to yield variations in rainfed wheat production systems such as the one studied here. The net contribution of these flows to spatial variations of rainfed potential yields showed to be relevant but highly irregular among years. Despite the inter-annual variability, typical of Mediterranean conditions, the occurrence of LIF caused simulated wheat yields to vary +16% from up to downslope areas of the field. Average crop yield ranged from 1.3 to 5.4 Mg grain yield (GY)  $\text{ha}^{-1}$ . The net yield responses to LIF, in downslope areas were on average 383 kg grain yield (GY)  $\text{ha}^{-1}$ , and the LIF marginal water productivity reached 24.6 ( $\pm 13.2$ ) kg GY  $\text{ha}^{-1} \text{ mm}^{-1}$  in years of maximum responsiveness. Such years of maximum responsiveness were associated with low rainfall during the vegetative stages of the crop in combination with LIF occurring at post-flowering stages. However, under field conditions, these differences were only visible in one of the two experimental years.

The economic implications associated with multiple scenarios of variable application rate of nitrogen were explored through a case

study and several recommendations were proposed. Both farm size (i.e., annual sown area) and topographic structure impacted the dynamics of investment returns. Under current policy-prices conditions, the adoption of variable application rate would have an economic advantage in farms similar to that of the case study with an annual sown area greater than  $567 \text{ ha year}^{-1}$ . Nevertheless, current trends on energy prices, transportation costs and impacts on both cereal prices and fertilizers costs enhance the viability of variable application rate adoption for a wider population of farm types. The profitability of adopting VAR improves under such scenarios and, in the absence of additional policy support, the minimum area for adoption of variable application rate decreases to a farm size range of  $68\text{--}177 \text{ ha year}^{-1}$ . The combination of price increases with the introduction of an additional subsidy on crop area could substantially lower the adoption threshold down to  $46 \text{ ha year}^{-1}$ , turning this technology economically viable for a much wider population of farmers.

## Resumen

En campos en pendiente, los cultivos de secano experimentan diferentes grados de estrés hídrico causados por variaciones espaciales de la humedad en el suelo, y los rendimientos varían espacialmente dentro del mismo campo. Esta variabilidad supone una oportunidad para la agricultura de precisión a través del manejo espacialmente variable. Sin embargo, si bien se han logrado avances significativos en los aspectos de la ingeniería de la variación espacial, como el aumento de la resolución espacial de los sistemas de datos y la automatización, se ha avanzado mucho menos en relación a la simulación de las respuestas de los cultivos a las variaciones espaciales de la humedad y los flujos hídricos. La mayoría de los estudios sobre las brechas de rendimiento de secano ignoran la variabilidad dentro de la parcela. Sin embargo, el uso de modelos de simulación de cultivos como medida de apoyo a los sistemas de gestión espacialmente variable, requiere que los enfoques de modelación espacial del agua sean capaces de representar y simular con precisión la variación dentro del campo de los factores relacionados con el agua disponible y la respuesta de los cultivos.

Esta tesis doctoral representa una nueva contribución a la agronomía de los sistemas agrícolas de secano, con énfasis en el papel que juegan los flujos de agua en zonas de topografía ondulada en la determinación de las variaciones espaciales del rendimiento del trigo. La tesis se ha desarrollado en capítulos que se complementan siguiendo un enfoque integrador.

La presente tesis doctoral revisó algunos de los modelos hidrológicos y de cultivo más ampliamente adoptados y exploró nuevas oportunidades para simular variaciones espaciales del agua a nivel de campo mediante la incorporación del flujo lateral de escorrentía superficial y sub-superficial en las zonas de menor elevación del campo. Desde este punto de vista, se evaluaron las variaciones espaciales de las brechas de rendimiento en trigo de secano, en Córdoba, España, que son causadas por flujos laterales de los puntos altos a los bajos.

Desde una perspectiva agronómica, las entradas laterales del agua

contribuyen a las variaciones de rendimiento en los sistemas de producción de trigo de secano como el que se ha estudiado en el ámbito de esta tesis. La contribución neta de estos flujos a las variaciones espaciales de los rendimientos potenciales de secano se mostró relevante pero altamente irregular entre diferentes años. A pesar de la variabilidad interanual, típica de las condiciones mediterráneas, la existencia de dichos flujos hizo que los rendimientos de trigo simulados variaran un +16% desde las áreas más elevadas de un campo hacia abajo. El rendimiento medio observado osciló entre 1.3 y 5.4 Mg de rendimiento de grano (GY)  $\text{ha}^{-1}$ . Las respuestas de rendimiento neto al flujo lateral, cuenca abajo, fueron en promedio 383 kg de rendimiento de grano (GY)  $\text{ha}^{-1}$ , y la productividad marginal de agua de LIF alcanzó 24.6 ( $\pm 13.2$ ) kg GY  $\text{ha}^{-1} \text{ mm}^{-1}$  en años de máxima capacidad de respuesta. Dichos años de máxima capacidad de respuesta se asociaron con bajas precipitaciones durante las etapas vegetativas del cultivo en combinación con flujos laterales en las etapas posteriores a la floración. En condiciones de campo, estas diferencias solo fueron visibles en uno de los dos años experimentales.

Las implicaciones económicas asociadas con múltiples escenarios de tasa de aplicación variable de nitrógeno se exploraron a través de un caso de estudio y se propusieron varias recomendaciones. Tanto el tamaño de la finca (el área sembrada anual) como la estructura topográfica afectaron la dinámica de los rendimientos de la inversión. Bajo las condiciones actuales de política agrícola, y de precios, la adopción de la tasa de aplicación variable tendría una ventaja económica en fincas similares a la del caso de estudio con un área sembrada anual superior a 567  $\text{ha año}^{-1}$ . Sin embargo, las tendencias actuales en los precios de la energía, los costes de transporte y los impactos tanto en los precios de los cereales como en los costes de los fertilizantes mejoran la viabilidad de la adopción de esta tecnología para una población más amplia de tipos de fincas. La rentabilidad de la adopción de aplicación variable de nitrógeno mejora bajo dichos escenarios y, en ausencia de apoyos adicionales, el área mínima para la adopción de aplicación variable disminuye hasta un rango de 68-177  $\text{ha año}^{-1}$  de área de siembra. La combinación de aumentos de precios con la introducción de un subsidio adicional asociado al

área de cultivo podría reducir sustancialmente el umbral de adopción hasta 46 ha año<sup>-1</sup>, lo que hace que la tecnología sea económicamente viable para una población mucho más amplia de agricultores.

# Acknowledgements

This thesis would not have been concluded without the fantastic support from my supervisors, I am tremendously grateful to the opportunity I had of being oriented by Professor Elías Fereres Castiel and Dr. José Alfonso Gómez Calero throughout these years. I doubtlessly thank you both for such an opportunity.

Particularly important is to acknowledge the enormous support and inspirational guidance received from Prof. Elías Fereres. Every day of this path, Professor Elías has been, since my first days as an under-graduate student of Agronomy in Córdoba back in 2014, a solid example of what it means to be a real scholar, a truly Professor, an excellent agronomist, and a great leader, guiding me and so many others with his personal example and his never ending willingness to be truly helpful and consequent in every opportunity of life to set ourselves at the service of others. The amount of learning received from Professor Elías goes extraordinarily beyond the academic life and I am hugely grateful for this. Gracias por todo, y siempre, Profesor!

Another crucial aspect for the conclusion of this thesis was the great opportunity to combine my academic investigation and theoretical path with an active collaboration with farmers, consultants and many others who take determinant roles within this passionate life sector that we call agriculture. I would like to dedicate my personal gratitude to the memory of Mr. Miguel Jimenez, a local farmer who collaborated directly with me during the first year of field experiments. A special thanks must also be addressed to the Cabrera family, truly inspiring local farmers and landowners who allowed me to conduct second year experiments on their fields and who were always open to receive me at their farm, to share information and to discuss points of interest that were determinant for the conduction of my thesis. I thank D. José María Cabrera, the

father, and his son José María Cabrera Millán for their great willingness to cooperate. Also important is to thank Eng. Juan José Herrero Carmona and Eng. Enrique Aranda Valera, from Cortijo La Reina, for the support given, by providing access to on-farm records and contributing with technical feedback to some important points of my investigation.

I must also acknowledge the technical support of Dr. Ignacio Puech, Dr. Juan Benavides, Dr. David Moldero, Mrs. Rafaela Gutierrez and Mr. Clemente Trujillo with field work. Important is also to thank Dr. João Vasco Silva, Prof. David Connor, Prof. Juan Vicente Giráldez Cervera, Prof. Francisco Avillez, Prof. José Pimentel Castro Coelho, Eng. Manuel Penteado and Dr. Francisco Orgaz Rosúa for contributing with active discussions and great personal examples to some important aspects of this thesis. I thank also all who have contributed to some chapters of this thesis, particularly important is to highlight the great contributions of Dr. Margarita García-Vila and Eng. Jakub Jeřábek.

A personal acknowledgment must be addressed to Eng. Carmen Ruz Ortiz from the department of agronomy, who have collaborated actively in the financial and operational management of my research and who is a great pillar of kindness, empathy, understanding and support for me and for so many others in the department of agronomy at IAS-CSIC. Muchas gracias Carmen!

Naturally, I would not have concluded my thesis without the daily support of my family and friends, especially my amazing parents and my beautiful wife. Gratitude to my father for the enormous and unique example that represents to me, in every aspect of my life. A great agronomist, an inspirational farmer and an incredible father, the big boss of the family! Gratitude to my mother for being always an example of joy and love, always beautiful and kind. Absolute gratitude to my wife, for everything! For her daily support, believing always in me, for carrying and being present, for loving our family, for being the best future mother in the world, and for delivering so many happiness to my life. I thank deeply God for all my family and friends. Without them, I would never have conducted my doctoral thesis. This thesis belongs primarily to them. Obrigado por tudo!

Finally, but more important than everything, this thesis is to You, because  
You are always there, always here!

Em tudo amar e servir!



# Curriculum Vitae

My name is Tomás Roquette Tenreiro and I was born on the 23<sup>th</sup> of June in 1992, in Lisbon, Portugal. I am an agronomist and a young-farmer. I have been mostly focused on applied research within the context of agricultural sciences, combining professional experiences in both the private and public context. In combination with scientific research, I have been also providing consultancy services mostly related to project planning and management, benchmarking, strategic studies and agricultural policy analysis. I am also a freelance writer of opinion articles, having contributed to 'Agora pensamos juntos', 'Agroportal', 'Observador', 'Sabado Online', 'Voz do Campo', 'Agroges' and 'Resource Wageningen Magazine'. Since 2020, I have been also collaborating as a reviewer for the scientific journals 'Agricultural Water Management' and 'Archives of Agronomy and Soil Science'.

I started my academic studies at the University of Lisbon, where I graduated on Agronomy & Agricultural Engineering in 2015. I spent the last year of my degree at the University of Córdoba, where I had the opportunity to meet Professor Elías Fereres and his research group at the department of agronomy.

In 2016, I moved to the Netherlands, where I studied a Master's degree on Crop Science at the University of Wageningen. During my time at Wageningen, I worked part-time on a dairy farm over a year, where I was responsible for the herd daily milking (1800 goats). In parallel, I conducted an internship at the Dutch company Agrovision, in Deventer, for 6 months, I collaborated as a freelance consultant with the start up 'Open-PD project' (an open network on plant pests and diseases), and I was a member of the International Catholic Community (ICC) of Wageningen, keeping an active participation at the local parish. I was also the manager of an academic consultancy project to the Louis

Bolk Institute, and for a very short-period, I visited the International Centre for Advanced Mediterranean Agronomic Studies (CIHEAM), in Montpellier, France.

In 2018, I moved back to Córdoba, where I started working as an intern at the agronomy department of the Institute for Sustainable Agriculture (CSIC). In late 2018, I was awarded with the second prize of the "IV Technical and scientific award for research and innovation on the olive sector", promoted by the Foundation of Caja Rural de Jaén, Spain. This award was conceived to the project titled "*Mapping cover crop dynamics in Mediterranean perennial cropping systems through remote sensing and machine learning methods*". Following that period, I started my PhD at the University of Córdoba. Since then, I have been combining several experiences, mostly working for international research projects at IAS-CSIC (i.e., EU-SustainFarm, FAO-NENA, EU-SHui), which allowed me to visit several countries over the last years (France, Austria, Italy, China and Germany). In 2019, I participated in the summer school on statistical analysis of spatial data in agricultural research, organized by the Italian Society of Agronomy and the companies ESRI-Italy & IBM.

During my PhD, I conducted field experiments in two different local farms, from which I would highlight a productive collaboration with the Company Cortijo La Reina, in Córdoba. More recently, I was a visiting scholar at the 'Higher Institute of Agronomy', in Lisbon, for a period of three months. Over the last years, I have combined research with consultancy and participation in several strategic studies, from which I would highlight the interesting cases of the Portuguese National Association of Fertilizer Industry (ANPIFERT), the Portuguese Association of Almond Nuts (APFS), the company AGRO.GES and our family company 'Anta de Cima' (farming, extensive livestock, grape and wine production, forestry). In 2021, I got married in Córdoba.