Departamento de Economía agraria, Estadística y Gestión de empresas

Escuela Técnica Superior de Ingenieros Agrónomos

Universidad Politécnica de Madrid

Agricultural insurance policies in Spain and the EU:

An analysis of existent and new risk management tools focusing on indirect risk assessment, and asymmetric information

TESIS DOCTORAL

Alba CASTAÑEDA VERA

Ingeniera Agrónoma

MSc Plant Sciences

Master en Tecnología Agroambiental para una Agricultura Sostenible

Madrid 2016

Departamento de Economía agraria, Estadística y Gestión de empresas

Escuela Técnica Superior de Ingenieros Agrónomos

Universidad Politécnica de Madrid

Agricultural insurance policies in Spain and the EU:

**An analysis of existent and new risk management
tools focusing on indirect risk assessment, and
asymmetric information**

Memoria presentada por
Alba CASTAÑEDA VERA

Para la obtención del grado de Doctor por la Universidad
Politécnica de Madrid

Directores:

Dr. Alberto Garrido Colmenero, Doctor Ingeniero Agrónomo

Dra. M^a Inés Mínguez Tudela, Doctora Ingeniera Agrónomo

Tribunal nombrado por el Magfco y Excmo. Sr. Rector de la Universidad Politécnica de Madrid, el día _____.

Presidente:

Vocal:

Vocal:

Vocal:

Secretario:

Suplente:

Suplente:

Realizado el acto de defensa y la lectura de la tesis el día _____ en la E.T.S. de Ingenieros Agrónomos de la Universidad Politécnica de Madrid.

CALIFICACIÓN:

EL PRESIDENTE

LOS VOCALES

EL SECRETARIO

ALBERTO GARRIDO COLMENERO, catedrático del departamento de Economía agraria, Estadística y Gestión de empresas, e M^a INÉS MÍNGUEZ TUDELA, catedrática del departamento de Producción Agraria, ambos de la Universidad Politécnica de Madrid

CERTIFICAN

Que el trabajo presentado por ALBA CASTAÑEDA VERA, Ingeniera Agrónoma por la Universidad Politécnica de Madrid, constituye un trabajo original de investigación y que ha sido realizado bajo nuestra dirección.

Puesto que el mismo cumple con la normativa vigente en materia de Doctorado, autorizaos su presentación.

Madrid, de de 2016

Fdo.

Alberto Garrido Colmenero

M^a Inés Mínguez Tudela

SUMMARY

A sustainable agricultural insurance system provides benefits to both insurers and insured agents. Equilibrium in the loss ratio and risk pooling are fundamental to guarantee such sustainability. Achieving equilibrium in the loss ratio requires accurate risk estimation. An agricultural insurance risk pool is a concurrence of producers whose different production risks in time and space are combined. Pooling risks together allows the losses of the producers registering claims in a certain year and a certain area to be offset by total premiums paid by producers participating in the system either registering claims or not.

Agricultural insurance faces a number of market imperfections that end up in market failures, compromising loss ratio equilibrium and risk pooling. Common imperfections are information failure, asymmetric information and low insurance demand. This Thesis evaluates different aspects that hinder risk setting and that constitute disincentives to crop insurance demand in order to contribute with new insights that could make insurance systems more sustainable.

First, we address a central issue in agricultural risk assessment, the availability of data. For that, we adapted methodologies, some of them commonly implemented in other scientific and technical spheres, to be applied in crop insurance. In particular, we evaluate the opportunities for using index insurances, crop models and yield gap analysis. We conclude that the use of insurance historical data rather than using indirect methods is preferred to evaluate and calibrate insurance parameters when available. However, historical data may be insufficient to establish the real risk the insurance is covering. Then, the use of weather indexes or more complex crop models is useful to design new insurance packages or coverage. Weather indexes and crop models should be used with caution as the time and spatial dependence of the risk influences the scale at which models might have validity. The common lack of observed data seriously hinders testing simulation models and thus constitutes an important source of uncertainty. The higher the complexity of the cropping system, the harder it becomes the use of indirect methods for crop risk assessment.

Second, we evaluated the relative efficiency of alternative risk management instruments in enhancing and stabilizing farm revenue and the effectiveness of public support. Results reveal that direct payments and crop diversification are the most effective measures in decreasing income variability. Moreover, despite crop diversification does not requiring any

direct public support, it provides a significant decrease in income variability. Incentives to contracting protection tools are a good compromise for both improving farm resilience to yield and price variability and to restrain public expenditure, especially when used complementarily with direct payments and/or diversification. Lastly, we conclude that farmers' utility of contracting insurance is related not only to risk but also to its profitability.

Lastly, we study factors that influence to crop insurance demand. For that, the concept of yield gap was useful to build evidence about asymmetric information and to evaluate the relationships between insurance parameters and insurance demand. Results suggest that the accuracy in setting the insured yield is decisive in farmers' willingness to contract crop insurance, even more than premium subsidies. We propose that the development of new risk management tools based on mutual funds may potentially decrease insurance demand, decreasing the pool and compromising the system's sustainability. Alternatively, we propose evaluating the possibilities for extending coverture to farmers' revenue in order to keep or even increase risk pool is a sole and more efficient tool.

RESUMEN

Un sistema de seguros agrarios sostenible es beneficioso tanto para los aseguradores como para los asegurados. El equilibrio en el ratio de pérdidas y la agrupación de los riesgos son fundamentales para garantizar dicha sostenibilidad. Para alcanzar un equilibrio en el ratio de pérdidas es necesario poder estimar el riesgo de forma precisa. La agrupación de riesgos en el seguro agrario supone la concurrencia de productores cuyos riesgos de producción se combinan en tiempo y espacio permitiendo así que las indemnizaciones por pérdidas de unos productores en un determinado año y una determinada zona se compensen con el total de las primas pagadas por parte de todos los productores que participan en el sistema.

Sin embargo, el seguro agrario se enfrenta a una serie de imperfecciones que comprometen esa sostenibilidad, entre ellas la falta de información, la información asimétrica y una baja demanda del seguro. Esta tesis aborda estos aspectos con el objetivo de mejorar el conocimiento del sistema y alcanzar así sistemas de seguros más sostenibles. El contexto empírico de la tesis incluye dos sistemas productivos tradicionales de la Península Ibérica, el tomate de industria (cultivo hortícola de regadío) en la cuenca extremeña del Guadiana y el trigo de invierno (cereal de secano) en Castilla y León.

En primer lugar, se aborda un aspecto central en la estimación del riesgo agrario, la disponibilidad de datos. Para ello, se adaptan metodologías, algunas de ellas usadas habitualmente en otras esferas científicas y técnicas, para su implementación en el seguro de agrario. Concretamente, se analiza la implementación de seguros indexados, modelos de cultivo y el análisis de brechas de rendimientos. La tesis concluye al respecto que el uso de datos históricos del seguro siempre es preferible al uso de métodos indirectos para evaluar y calibrar parámetros de seguro, siempre que estos estén disponibles. Sin embargo, en ocasiones los datos históricos son inexistentes o insuficientes para establecer el riesgo real. Es en estos casos en los que el uso de índices climáticos sencillos o modelos de cultivo más complejos es útil para diseñar nuevas coberturas y líneas de seguros. En cualquier caso, estos métodos han de usarse con cautela, ya que la dependencia espacial y temporal del riesgo influye en la escala a la que estas estimaciones tienen validez. La habitual falta de datos observados, dificulta en gran medida la validación de los modelos, constituyendo una fuente de incertidumbre importante. Los resultados muestran cómo cuanto más complejo es el sistema de cultivo mayores dificultades se encuentran a la hora de implementar métodos para la estimación indirecta del riesgo.

En segundo lugar, se evalúa la eficiencia relativa de instrumentos alternativos de gestión del riesgo en la mejora y la estabilización de la renta de los productores y la eficacia del gasto público. Los resultados revelan que los pagos directos y la diversificación de cultivos son las medidas más eficaces para disminuir la variabilidad de la renta. Además, aunque la diversificación de cultivos no requiere ningún apoyo público directo, esta supone una disminución significativa de esa variabilidad. Los incentivos para contratar instrumentos de protección influyen positivamente tanto en la mejora de la resiliencia de la renta de las explotaciones a la variabilidad de rendimientos y precios como en la contención del gasto público, especialmente cuando se usan de forma complementaria junto con pagos directos y/o diversificación. Por último, la tesis concluye que la utilidad de los productores a la hora de contratar un seguro no es función únicamente del riesgo, sino también de la rentabilidad que ello supone.

Por último, se estudia qué factores influyen en la demanda del seguro. Para ello, el concepto de brecha de rendimiento, comúnmente usado en agronomía, se redefine para crear un indicador que busca evidenciar la presencia de información asimétrica. Este indicador se usa posteriormente para evaluar si existe dicha asimetría influye en la decisión de los productores de contratar el seguro, y en caso afirmativo, qué nivel de protección. Los resultados revelan que la presencia de asimetría en la información influye y que lo hace incluso con un mayor peso que el nivel de subvención de la prima. El análisis sugiere que el desarrollo de nuevas herramientas de gestión del riesgo basadas en fondos mutuales podría disminuir la demanda del seguro, disminuyendo la agrupación de los riesgos del seguro comprometiendo la sostenibilidad del sistema. Alternativamente, se propone evaluar las posibilidades de extender las coberturas del seguro a los ingresos de los productores para así mantener o incluso aumentar la agrupación del riesgo en una sola herramienta más eficaz.

Agradecimientos

No hubiera sido posible finalizar esta tesis sin la ayuda de todas aquellas personas que me han apoyado a lo largo de estos años, tanto profesional como personalmente. A través de las siguientes líneas, quiero agradecerles ese apoyo.

En primer lugar, me gustaría darles las gracias a mis directores de tesis, los profesores Alberto Garrido e Inés Mínguez. A Alberto Garrido por adentrarme en el mundo de los seguros agrarios, por las interesantes discusiones que han abierto nuevos enfoques, siempre de la mayor relevancia, y por su confianza en mí para abordarlos. A Inés Mínguez por su empeño en innovar y transversalizar que ha sido fundamental para dar sentido a este trabajo. Me siento muy afortunada por haber tenido la oportunidad de alcanzar el título de doctora de la mano de dos profesionales, ambos referencia en su campo de conocimiento. Quiero además agradecerles la gran libertad que me han dado a la hora de decidir las orientaciones de este trabajo y de formarme en las distintas materias en las que he necesitado profundizar, y sobre todo por su flexibilidad que me ha permitido compaginar en el último año finalizar esta tesis y preparar oposiciones.

I would like to express my gratitude to Peter Leffelaar from Wageningen University. I really enjoyed the always motivating discussions and appreciated his suggestions for going deeper and deeper in the research. My staying at PPS was, as always, very fruitful professionally and intellectually.

También me gustaría dar las gracias al resto de co-autores que han participado en la elaboración de los distintos capítulos de la tesis. Muchas gracias a Laura Barrios y Antonio Saa por el apoyo estadístico. Gracias a Carlos Cantero y a Jorge Álvaro por compartir con nosotros el fruto de su trabajo de tantos años, y por su implicación en nuestra publicación conjunta.

Los sectores agrarios en los que hemos trabajado son complejos, con multitud de factores y de actores implicados. No puedo dejar de dar las gracias a aquellos que me echaron una mano para comprender cada uno de los sectores abordados en esta tesis. En especial, gracias a Elena Ordiales del CTAEX por su ayuda para comprender el sector del tomate de industria en Extremadura y a M^a José Martínez Tineo de ENESA por sus aclaraciones sobre las particularidades de los seguros de cultivos herbáceos extensivos. Muchas gracias a David Nafría, Miriam Fernández y Alberto Gutiérrez por su apoyo en el trabajo de campo

en Castilla y León, para mí son un ejemplo por su forma de trabajar, su eficiencia y su compromiso.

A otros profesores de la ETSIA, Isabel Bardají, Miguel Quemada y Jon Lizaso por estar siempre dispuestos a resolver cualquier duda.

Por último, agradecer a mis compañeros de oficina y alrededores, Alicia, María-Alonso, Espe, Axel, Omar, Clara, Javi, María-Soldevilla y Juan por su amistad dentro y fuera de la escuela. María-Alonso y Espe, por desdibujar la línea que separa pasar el día con tus amigos y pasar el día trabajando, siempre me acordaré de las experiencias gastronómicas que nos hemos regalado a diario. A Juan, que (quizá a su pesar) ha sido y será mi asesor agrícola de confianza, por su punto de vista y por su humor castellano.

CONTENTS

1. Introduction.....	17
1. Introduction	18
1.1 Risk management in agriculture.....	18
1.2. Agricultural insurance system in Spain	19
1.3. Agricultural Insurance system sustainability and risk predictability	21
2. Research objective and aims	23
3. Outline and publications	26
4. Publications and statement of authorship	28
5. Research context.....	30
2. Assessment of insurance coverage and claims in rainfall related risks in processing tomato in Western Spain.....	33
1. Introduction	35
2. Materials and methods.....	36
2.1 Current damage appraisal for insurance payments	36
2.2. Processing tomato sector.....	37
2.3. Insurance database.....	40
2.4. Risk model.....	45
3. Results	47
3.2. Statistical analyses.....	48
3.3. Extrapolating risk.....	52
4. Discussion and conclusions	53
5. Acknowledgements	56
3. Selecting crop models for decision making in wheat insurance	57
1. Introduction	59
2. Materials and methods.....	60
2.1. Models	60
2.2. Aboveground biomass production.....	63
2.3. Canopy and crop development.....	65
2.4. Biomass partitioning and yield formation	68
2.5. Soil modules.....	68
2.6. Computation of water stress impact on biomass production	74
2.7. Experimental site and data.....	79
2.8. Models parameterization, calibration and comparison.....	81
3. Results and discussion	82
3.1. Models parameters and approaches	82
3.2. Critical remarks about the calibration and use of the models	101

4. Concluding remarks	103
5. Acknowledgements	105
6. Appendix A. Supplementary data	106
4. Evaluation of risk management tools for stabilising farm income under CAP 2014-2020	109
1. Introduction	111
2. Risk management instruments.....	113
3. Theoretical framework.....	115
4. Empirical application	116
5. Results	121
6. Discussion	126
7. Conclusions	129
8. Acknowledgement.....	130
5. Crop insurance demand in wheat production: focusing on yield gaps and asymmetric information	131
1. Introduction	133
2. Materials and methods.....	134
2.1 Winter cereal insurance	134
2.2 Study site	137
2.3 Actual, expected, insured and water limited yields.....	139
2.4 Yield gap analysis	145
2.5 Asymmetric information	146
2.6 Factors influencing insurance demand.....	148
3. Results	152
3.1. Wheat insurance demand and county characterization	152
3.2. CERES-Wheat model calibration and validation	152
3.3. Yields and yield gap analysis.....	153
3.4. Asymmetric information	155
3.5. Insurance demand, yield gaps and asymmetric information.....	156
4. Discussion and conclusions	158
5. Acknowledgements	161
6. Main results and conclusions	163
1. Summary of main results.....	164
2. Conclusions	167
3. Limitations and recommendations for future research.....	169
References	171

Chapter 1

1. Introduction

A. Castañeda-Vera^{1,2}

¹ Research Centre for the Management of Agricultural and Environmental Risks
(CEIGRAM), Universidad Politécnica de Madrid (UPM), Spain

² Agricultural Systems Group, Dpto. Producción Agraria, E.T.S. Ingenieros Agrónomos,
Universidad Politécnica de Madrid, Avda. Complutense s/n, 28040 Madrid, Spain

1. Introduction

1.1 Risk management in agriculture

Agriculture is a risky activity. Farm's vulnerability and exposure to hazards and uncertainty are very much related to farmers' decisions about resource allocation or production planning. Also uncontrollable risks, like weather variability and price volatility (both products prices and farm input prices), drive farmers' significant income fluctuation.

Agriculture is an intensively weather dependent activity. Consequences and severity of production variability ultimately depends on the socioeconomic situation of the affected region and the technical and financial aspects of the farms. In regions where subsistence agriculture plays a central role, problems of yield variability are associated to food insecurity, while in developed countries where food supply is not a central issue, problems are mainly related to farm economic resilience, inequalities in rural standards of living with respect to urban society and serious risks of farm activity abandonment.

In the European and Spanish contexts, the policy focus is on the effects on farm income. Income variability limits saving capacity and hinders the access to external financing to perform structural changes for farm resilience and adaptation to face all these risks. This increases farms' vulnerability and worsens the negotiating position of farmers within the food supply chain. It should be also mentioned the opportunity cost for farmers being devoted to farming instead of another activity that might have associated a higher utility including income level and stability, standards of life and lifestyle preferences.

A wide range of risk management strategies are implemented either privately or supported by agricultural policies. Risk management might just include individual decisions or involve third parties that share or assume part of the farmers' risks. Individual decisions comprise on-farm resources changing the production strategy (diversification, input intensification), the commercial strategy (vertical integration; forward contracting, storage, among others) or the use of farm benefits constituting farm financial reserves (savings, stabilisation accounts). Risk-sharing or risk-transferring strategies involve third parties that underwrite part of the risk in return of a monetary reward (insurance, mutual funds, hedging, forward contracts).

European Union's Common Agricultural Policy (CAP) has shifted from supporting product prices to supporting producers, mainly through direct payments. Market instruments remain now as safety nets (first CAP pillar), ensuring that intervention prices are only used in real crisis and exposing farmers to market signals, what has facilitated a better adaptation of supply to market demand. According to Cordier (2014), current measures on risk management in the first CAP pillar are weak, because safety nets are limited by low intervention prices and the relatively small financial reserves. Therefore, it is expected that risk management measures offered through the second CAP pillar or already implemented national schemes will gain weight in the future. Actually, future CAP (after 2020) development regarding risk management is already on the table. Bardají and Garrido, (2016) develop an overview of the last CAP reform 2014-20, evaluate its potential efficiency and propose recommendations to improve the effectivity of the policy to cope with income uncertainties.

In Spain, agricultural insurance has received major attention already from 1978 when its system was established by law. Such system was part of the political reforms towards economic stabilization in the political regime transition and towards adapting the agricultural policy to European standards before accessing the European Economic Community (EEC) in 1986. Currently, agricultural system is consolidated in Spain and, globally, it is considered an important stabilizing factor of more than the 40 % of the agricultural income in the agricultural sector in Spain (UNESPA, 2012).

1.2. Agricultural insurance system in Spain

Insurance provides an ex-post compensation for losses caused by natural hazards and it constitutes an alternative option to public compensation after catastrophes. National schemes already exist in the European Union (EU), as it is the case of Spain, but differ widely across Member States (Bardají and Garrido, 2016).

The current Spanish Agricultural Insurance System (SAIS) was established by law in 1978 with two main objectives. First, to establish feasible coverages, both technical and financially, that allowed farmers to cope with damages in their productions originated by catastrophic and uncontrollable risks; and secondly, to make available to the State an effective and economically viable instrument to support the sector with a more predictable

cost and, in return, to phase out the extraordinary ad-hoc payments devoted to farmers affected by losses due to insurable risks.

After accessing the EEC in 1986, State aids to agricultural insurance were considered compatible to the common market. Today, almost all agricultural productions are covered against most of the climatic risks through about fifty insurance lines. In 2014, 440947 farmers contracted policies for 11075 million € of insured capital of agricultural production (ENESA, 2016), comprising about 40% of the crops and 10% of the husbandry production. Most important claims between 2003 and 2008 were related to hailstorm events with 43% of the reported claims, followed by frosts (19.3 %), droughts (14.4 %), wind (8.4 %) and rainfall (7.4 %) (Ruiz-Zorrilla, 2010).

The SAIS is a mixed public-private system. A pool of private insurance companies, forming a co-insurance scheme (Agroseguro), provides coverage to all farmers that voluntarily opt to participate. National government coordinates the system, subsidizes insurance premium (often together with the regional governments) and reinsure the co-insurance through the public reinsurance entity, Consorcio de Compensación de Seguros (CCS). The Ministry of Agriculture, through the Agricultural Insurance National Entity (ENESA), performs studies to establish the conditions to which insurable risks can be actually insured. Premium subsidies are offered to stimulate farmers' contractual decisions. On average, public support reaches about 40% of the total premiums costs (Garrido et al., 2014). The optional nature of the insurance for farmers stimulates Agroseguro and ENESA to offer them attractive products.

Damage appraisal requires expertise and it is performed by loss adjusters contracted by Agroseguro. Indemnities are based on in-field evaluation following explicit monitoring procedures (Antón and Kimura, 2011), with the exception of coverage against drought in pastures for which a weather index is already implemented. While it is considered a costly system mainly due to the time spent in field visits, probably, it decreases the incidence of disagreements between the parts. This whole process is subject to quality control procedures independently performed by both Agroseguro and the CCS.

As in most insurance products, a system bonus-malus is applied. It makes the system more flexible favouring a better adaptation of risk premium to individual farmer's risk, instead of assigning the same risk to all farmers in the same area.

1.3. Agricultural Insurance System sustainability and risk predictability

A sustainable agricultural insurance system provides benefits to both insurers and insured agents. Insurance system benefit derives from total raised premiums after deducing indemnities for crop losses and administrative costs. However, insured farmers will only be willing to pay for insurance if premium cost is lower than the expected losses and, depending on their risk aversion, an extra amount of money merely to avoid the risk. This amount of money is called risk premium and it is defined by the well-known von Neumann-Morgenstern utility theorem. Equilibrium in the loss ratio and risk pooling are fundamental to guarantee such sustainability.

Loss ratio is the quotient between the indemnities paid by the insurance system and the total premiums paid by farmers and, therefore, depends on the hazards experimented each year. In 2014 and 2015, global loss ratio of the SAIS was 0.882 and 0.873, respectively. Achieving equilibrium in the loss ratio requires accurate risk estimation.

Risk pooling is fundamental to insurance. An agricultural insurance risk pool is a concurrence of producers whose different production risks in time and space are combined. Pooling risks together allows the losses of the producers registering claims in a certain year and a certain area to be offset by total premiums paid by producers participating in the system either registering claims or not.

Agricultural insurance faces a number of market imperfections that end up in market failures, compromising loss ratio equilibrium and risk pooling potential. Common imperfections are information failure, asymmetric information and low insurance demand.

In crop insurance, risk is defined as the probability of a claim multiplied by the amount of the loss. A claim is registered when actual yield is lower than the insured yield (deductibles considered), as a result of a hazard considered in the policy, and the difference between both yield levels is the loss to be indemnified. The calculated risk is used to set and update premiums, or to design new insurance policies. The accuracy with which the insurer quantifies the actual risk is highly dependent on the availability on high quality actual yield data.

Information failure occurs as records on observed yields are not always complete; then the information available for the insurer to calculate the probability and severity of a claim can

be biased. This gives rise to unbalanced loss ratios affecting the actuarial robustness of entire insurance systems.

Sometimes, lack of information can be due to the extension of coverage to new risks. In this situation no historical data exists on claims as no insurance has been yet offered. In these cases, the use of indirect methods and prediction tools might be of a great help. Crop models might be valuable tools to generate data on expected yields for risk assessment when no historical data is available. The problem is that sometimes model performance is hindered by a large uncertainty due to either a low level of complexity on plant physiology of soil hydrology, or, on the contrary, by a high demand on detail data for calibration, often unavailable.

Information failure may happen with historical data available as well. This is the case of asymmetric information that arises when the insurer has an imperfect knowledge of the insured risk and farmers' behaviour. Asymmetry of information commonly results in two behavioural responses: moral hazard and adverse selection. Moral hazard occurs when farmers modify their behaviour after contracting insurance in order to increase the probability of being indemnified (Goodwin, 1994), while adverse selection occurs when high-risk farmers contract insurance in greater numbers than low-risk farmers.

Usually, adverse selection means the insurer is unable to differentiate high-risk and low-risk farmers and charges them the same premium. In this situation, insurance becomes more profitable to high-risk farmers, while low-risk farmers will incline to drop out from insurance programmes, increasing initially loss ratios in the insurance system and premium rates for insured farmers in the medium and long term (e.g. Skees and Reeds, 1986; Skees, 2001).

Index insurance might solve these two classic problems of traditional insurance, lack of historical data and asymmetric information. Index insurance determines indemnities based on values obtained from an index that serves as a proxy for losses. Furthermore, index insurance does not require historical yield data and it is not susceptible to moral hazard and adverse selection (Barnett, 2004).

Therefore, risk management tools must be clearly defined in order to get technical and financially feasible coverage; otherwise, their scope in increasing farm income stability and public expenditure efficiency becomes difficult to estimate. In the case of publicly

supported agricultural insurance schemes, they aim at replacing disaster relief measures with a system that involve farmers' co-responsibility and allow foreseeable public expenses. Understanding factors driving farmers' insurance demand is fundamental to evaluate the system sustainability and to orient measures towards increasing farmers' insurance adoption (Shaik et al., 2008). The body of literature on crop insurance demand is vast. It is mainly focused on insurance subsidies (e.g. Babcock and Hart, 2005 and Donoghue, 2014), risk frequency (e.g. Coble et al., 1996), asymmetric information (e.g. Garrido and Zilberman, 2008 and Glauber et al., 2002) or farmers characteristics (Santeramo et al., 2016a). Low insurance demand is often due to the infrequency of events and the high premiums derived under the actuarial principles leading market-based insurance, which discourage farmers to pay the premiums every year when indemnities are seldom received (Mahul and Stutley, 2010). Moreover, such high premiums may make difficult meeting social objectives (Goodwin and Mahul, 2004), therefore justifying public support.

The availability of alternative instruments might affect insurance demand if they behave as substitutes. If new risk management instruments become more interesting to producers, they might exit the insurance system, diminishing demand and therefore risk pooling. This might be the case of the alternative approaches to risk management materialized after the last CAP reform: mutual funds and income stabilization tools (IST) cofinanced through the European Agricultural Fund for Rural Development (EAFRD). Farmers' willingness to adopt such instruments is relevant for policy-makers to design requirements and conditions of these instruments.

2. Research objective and aims

This thesis evaluates agricultural production risks and climate hazards, and assesses the performance of various existent and potentially new agricultural insurance policies. By looking at an unexpected increase in insurance claims in intensive horticulture (processing tomato) and at observed insurance penetration indices (insurance demand) in extensive arable crops (cereals), we investigate whether insurance's design can be improved to overcome the classical problem of asymmetric information. Furthermore, since CAP 2014-2020 offers new support policies of insurance and income stabilization tools, we bring the yields gap analysis into the realm of farms' income variability and evaluate the effects of diversification, an income stabilization tool (IST) and insurance, both from the farmer's

perspective and effectiveness of the public support. An underlying theme of the thesis, addressed in two of its core chapters, is the potential use of crop models to inform the design and parameterization of insurance policies.

The main objective was to evaluate different aspects that hinder risk setting and that constitute disincentives to crop insurance demand. We therefore attempt to contribute with new insights that could make the SAIS more sustainable.

We address the main objective of the thesis by answering the following research questions and specific objectives:

1. Are weather indexes and crop models accurate methods for indirect risk assessment?
 - a. We model rainfall-related risks in horticultural crops in terms of measurable and objective variables in Chapter 2.
 - b. We evaluate the possible impact of the extended coverage on the observed increase in claims in horticultural crops in Chapter 2.
 - c. We select the most suitable crop model for drought risk assessment in semi-arid regions in Spain in Chapter 3.
 - d. We assess the usefulness of crop models to be used within a crop insurance analysis and design in Chapters 3 and 5.
2. In view of the CAP 2014-2020 consideration of various policies supporting different risk management instruments, how do they compare against each other?
 - a. We quantify farm profitability and income stability offered by alternative risk management strategies provided at the European, national and private level in Chapter 4.
 - b. We evaluate the effectiveness of public support in stabilizing farm income in Chapter 4.
3. Are asymmetric information and the availability of alternative risk management instruments affecting crop insurance demand?
 - a. We conclude about the potential of several risk management tools to be adopted by farmers in Chapter 4.

- b. We quantify the gap between actual wheat yields and the yield data managed by the insurer to obtain indicators providing evidence of potential asymmetric information in wheat insurance in Chapter 5.
- c. We explore factors affecting the decision of farmers to contract insurance in Chapters 4 and 5.

The empirical context of this thesis includes two well-established productions in Spain, processing tomato in the Guadiana River basin (West Spain) and arable farms in Castilla y León (North-central Spain).

Spain is the fourth largest producer of processing tomato in the world, behind USA (California), China, and Italy, with 31900 ha in 2015 (ESYRCE, 2015). About 70% of the national production is harvested in the region of Extremadura, Western Spain (Macua-González et al., 2012). Processing tomato is grown following coordination agreement the processing industry to distribute the harvest over the two-month period when the processing industry is operating. Planting is also conditioned by the weather conditions because earlier planting dates are limited by frosts and late planting dates by excessive autumn rainfall. Main weather hazards for field-grown tomato are hail and extreme rainfall. Main weather hazards for field-grown tomato are hail and extreme rainfall. Risks related to rainfall have been covered by insurance since 1999. The recent extension of guarantees related to rainfall-related risks and the subsequent and unexpected large increase in claims make this region a suitable context for this thesis.

Spain is the fifth largest producer of cereals in the European Union, after France, Germany, Poland, and UK, with harvested production that ranged between 14 and 25.5 million Mg (6.2 Mha of crop) between 2000 and 2013 (Eurostat, 2013). About 47% of the cultivated land dedicated to arable crops (including cereal) is located in Castilla y León region, with about 2.9 Mha (ESYRCE, 2013). The most sensitive phenological stages to meteorological adversities are flowering (April-May, drought and heat stress) and grain filling (June-July, drought). In Spain, multi-peril insurance for cereal farmers has been offered since 1982 and has reached a relatively high penetration rate, reaching 63.9 % in 2015 (Machetti Bermejo, 2015). Arable farming represents the largest semi-arid Mediterranean agricultural system in terms of cultivated surface and insured area. In this

second empirical context, crop models performance is tested and potential factors affecting the decision of farmers on whether contracting insurance or not are explored.

3. Outline and publications

We used a combination of quantitative analytical methods to address the main objective of this thesis and answer the research questions. All chapters are illustrated with examples of Spanish agricultural systems comprising traditional irrigation and rainfed schemes, processing tomato in Extremenian Guadiana river basin (West Spain) and arable farming in the region of Castilla y León (northern Central Spain), respectively. Arable farming included common cases in the study area, thus wheat monoculture and a crop rotation of wheat, barley and dry peas.

This thesis is structured in seven chapters and the contents are as shown in Figure 1.

In this introductory chapter, we have presented the classical problems of agricultural insurance, focusing on the equilibrium in the loss ratio and the agricultural demand as main aspects that guarantee a sustainable insurance system. Two threats to loss ratio equilibrium are explored along the thesis: (i) no data available for risk assessment and (ii) asymmetric information. Moreover, it was compared the effectivity of alternative risk management instruments available that might behave as substitutes to crop insurance.

The intensity of agricultural risks is associated to uncertain biological, climatic, and socioeconomic variables that lead to an insufficient knowledge of the system. Such a limitation may have serious consequences for insurance, and might entail an inefficient use of public expenditure devoted to the system. In order to improve risk management, three methodologies for agricultural risk estimation and system's evaluation were assessed: (i) index insurance, (ii) crop models and (iii) yield gap analysis.

CAP reform for 2014-2020 introduced the possibility to use biological or weather indexes, defined as “*equivalent yield loss indexes established at farm, local, regional or national level*” (Article 17, Regulation (EU) 1305/2013), to estimate individual farmer production losses. The use of such indexes avoids problems of asymmetric information related to moral hazard because indemnities are paid for risks affecting at a regional level and not solely to

individual farmers. It also decreases the cost of the whole system by improving the current procedure of visual (and/or sampling) assessment by experts in the damage appraisal. Chapter 2 discusses the opportunities to insure processing tomato against rainfall risks through a Weather Index Insurance. The incidence of rainfall-related risks for processing tomato production in the Extremenian Guadiana River basin (West Spain) is modelled in terms of measurable and objective variables. The resulting model is then used to evaluate the possible impact of the extended coverage on an observed increase in claims.

Figure 1. Thesis structure.

Crop models are helpful for understanding the complexity of cropping systems as they comprise knowledge on physiological processes and plant interactions with the environment. This explains the increasing use of such models in decision-making, as for policy shaping and analysis or for farmer consultancy. There is a wide range of crop models available, but they differ in many aspects depending on the purposes for which they were

developed. The suitability of crop models for crop insurance analysis and design is studied in depth in Chapter 3. Four pre-selected crop models simulating wheat yield under rainfed growing conditions at the field scale are compared in terms of modelling approaches, process description and model outputs. The analysis is illustrated with the simulation of winter wheat (*Triticum aestivum* L.) over five growing seasons in Castilla y León.

Chapter 4 includes an analysis of the contribution of alternative risk management tools and strategies to stabilize agricultural income in arable farms and the willingness of farmers to adopt them. The analysis is illustrated with an average arable farm in the region of Castilla y León.

In Chapter 5, a novel method to measure asymmetric information is designed based on yield gaps, which we apply to wheat insurance again in the region of Castilla y León. Yield gap has been proven to be a valuable concept for assessing and understanding the ecological possibilities to meet food demand for an increasing population (van Ittersum et al., 2013). In this chapter, yield gaps are re-defined as the difference between insured and actual yields, indicating either a deficiency in the insurer's knowledge of the cropping system or advantages taken by the farmer based on wider information. Lastly, the role of asymmetric information on crop insurance demand is assessed.

Lastly in Chapter 6, we identify major conclusions, research contributions and limitations, and recommendations for future research.

4. Publications and statement of authorship

The four core chapters of the thesis are either published or under revision in scientific journals. All of them have been authored by a number of people.

Chapter 2 is published by the *European Journal of Agronomy* and participated in the general Assembly 2016 of the *European Geosciences Union* in Vienna in 2013.

Castañeda-Vera, A., Barrios, L., Garrido, A., Mínguez, M.I., 2014. Assessment of insurance coverage and claims in rainfall related risks in processing tomato in Western Spain. European Journal of Agronomy. 59, 39–48.

Castañeda-Vera, A.; Garrido, A.; Ruiz-Ramos, M.; Sánchez-Sánchez, E.; Mínguez, M.I. Quantitative assessment of current and future risks related rainfall in processing tomato in the Guadiana river basin (SW Spain). EGU General Assembly 2013, held 7-12 April, 2013 in Vienna, Austria, p.13631

Castañeda-Vera, A. worked on the database, performed the statistical analysis, interpreted data, wrote the manuscript and acted as corresponding author. Barrios, L. participated in data statistical analysis. Garrido, A. and Mínguez M.I. supervised development of work, helped in data interpretation and manuscript evaluation. Additionally, in the conference proceeding, Ruiz-Ramos, M.; Sánchez-Sánchez, E. participated in defining the work's objective.

Chapter 3 is also published by the *European Journal of Agronomy*.

Castañeda-Vera, A., Leffelaar, P.A., Álvaro-Fuentes, J., Cantero-Martínez, C., Mínguez, M.I., 2015. Selecting crop models for decision making in wheat insurance. European Journal of Agronomy. 68, 97–116.

Castañeda-Vera, A. worked on the literature review and models comparison, calibrated the models, interpreted data, wrote the manuscript and acted as corresponding author. Leffelaar, P. supervised the development of work and actively participated in the interpretation of the results. Álvaro-Fuentes, J. and Cantero-Martínez, C. provided data for models calibration. Mínguez, M.I. participated in defining the work's objective.

At the moment of printing this thesis, Chapter 4 has been accepted for being presented in the *EAAE seminar: prospects agricultural insurance in Europe* in Wageningen in October 2016 and it is being considered for the *Agricultural Finance Review* special issue. Castañeda-Vera, A. worked on the database and the mathematical model, interpreted data, wrote the manuscript and acted as corresponding author. Garrido, A. and Mínguez, M.I. supervised development of work, actively participated in the interpretation of the results and evaluated the manuscript.

Lastly, part of the work of Chapter 5 has participated in two conference proceedings. Castañeda-Vera, A. worked on the database, performed the yield gaps and statistical analysis, interpreted data and wrote the manuscript. Garrido, A. and Mínguez, M.I. supervised development of work, actively participated in the interpretation of the results and evaluated the manuscript.

Castañeda-Vera, A., Garrido, A. and Mínguez, M.I. Yield gap analysis and crop insurance. 5th International Symposium For Farming Systems Design. 7-10 September 2015, Montpellier, France

Mínguez, M.I, Castañeda-Vera, A., Garrido, A. Análisis de la brecha de rendimientos. Ciclo de Conferencias sobre el Seguro Agrario y Gestión de Riesgos en la Producción Agropecuaria. Ponencia nº 18, Madrid.

5. Research context

This thesis was completed within the context of the “Multiscale climate variability. Agricultural and economic impacts” (MULCLIVAR), funded by the Spanish Ministerio de Economía y Competitividad (MINECO) CGL2012-38923-C02-02. In this research project, the UPM team was coordinated by Professor M^a Inés Mínguez. My involvement in the project lasted four years, beginning in 2012. I was affiliated to the Department of Agricultural Economy, Statistics and Business Management in the School of Agricultural Engineering and at the Research Centre for the Management of Agricultural and Environmental Risks (CEIGRAM), with the collaboration of AgSystems research group, the three of them belonging to the Universidad Politécnica de Madrid (UPM).

In 2014, I spent four months as a visiting scholar at the Plant Production Systems group, at Wageningen University in the Netherlands, hosted by Professor Peter Leffelaar. The research was developed as part of the international PhD within the Agro-environmental Technology for Sustainable Agriculture (TAPAS) program.

MULCLIVAR is a joint project coordinated by TROPA-UCM-IGEO and AgSystems-CEIGRAM-UPM under the umbrella of Campus Moncloa. Its goal was to improve the understanding of impacts of climate variations at various timescales. MULCLIVAR is organized in two subprojects: the first one, subproject ODYN deals with the ocean influence on climate variability at interannual to multidecadal timescales with focus in the Euro-Atlantic and West African regions; while the second one, subproject ACER-Agro, examines the agro-economic impact of climate variability and change and adaptations on cropping systems in Spain. This thesis is part of the ACER-Agro subproject and contributed to understand the role of different public and private measures to smooth the negative impact of climate variability on farm economic stability, with a special attention to

crop insurance. Modelling approaches of a different level of complexity, linking climate variability and the farm economic performance, are explored in order to assess their suitability for improving risk management.

Chapter 2

2. Assessment of insurance coverage and claims in rainfall related risks in processing tomato in Western Spain

A. Castañeda-Vera^{1,2}, L. Barrios³, A. Garrido¹ and I. Mínguez^{1,2}

¹ Research Centre for the Management of Agricultural and Environmental Risks (CEIGRAM), Universidad Politécnica de Madrid (UPM), Spain

²Agricultural Systems Group, Dpto. Producción Agraria, E.T.S. Ingenieros Agrónomos, Universidad Politécnica de Madrid, Avda. Complutense s/n, 28040 Madrid, Spain

³Statistical Support Service, Consejo Superior de Investigaciones Científicas (CSIC)

European Journal of Agronomy 59 (2014): 39 - 48.

Abstract

An extension of guarantees related to rainfall-related risks in the insurance of processing tomato crops has been accompanied with a large increase in claims in Western Spain, suggesting that damages may have been underestimated in previous years. A data base was built by linking agricultural insurance records, meteorological data from local weather stations, and topographic data. The risk of rainfall-related damages in processing tomato in the Extremenian Guadiana river basin (W Spain) was studied using a logistic model. Risks during the growth of the crop and at harvesting were modelled separately. First, the risk related to rainfall was modelled as a function of meteorological, terrain and management variables. The resulting models were used to identify the variables responsible for rainfall-related damages, with a view to assess the potential impact of extending insurance coverage, and to develop an index to express the suitability of the cropping system for insurance. The analyses reveal that damages at different stages of crop development correspond to different hazards. The geographic dependence of the risk influences the scale at which the model might have validity, which together with the year dependency, the possibility of implementing index based insurances is questioned.

Keywords: processing tomato, risk, rainfall, crop insurance, index insurance, Extremadura, Spain.

1. Introduction

Agricultural insurance systems are plagued by numerous problems that threaten their economic performance and continuity. Poorly designed policies, the effect of asymmetric information (where insurance purchasers have more information than the insurance companies about actual risk and behaviour), unfair loss adjustment procedures, and biases in setting premiums are among the main factors that have caused agricultural insurance failures (IFAD/WFP, 2010; Mahul and Stutley, 2010; Skees et al., 2008). As in most branches of insurance, the loss or damage of the insured good must be a direct effect of an unambiguously measurable event, which in the case of cropping is the partial or complete loss of harvest caused by an observable climatic hazard or anomaly.

Since the passing of the agricultural insurance law in 1978 the aim in the Spanish insurance policy has been to provide a broad and extensive coverage to all agricultural and livestock production enterprises (Antón and Kimura, 2011). The practical objective is to stabilize income in the primary sector by providing protection against adverse climatic events. Risks covered for plant production are mainly related to climatic adversities. Most important claims between 2003 and 2008 were related to hailstorm events with 43% of the reported claims, followed by frosts (19.3 %), droughts (14.4 %), wind (8.4 %) and rainfall (7.4 %) (Ruiz-Zorrilla, 2010).

Processing tomato farmers have benefitted from expanded coverage since it was first offered in 1990. Additional risks related to rainfall have been progressively incorporated and the coverage has been extended to longer periods during crop development. It now includes damages caused by frost, hailstorms, persistent rainfall through the entire crop cycle and also floods caused by torrential rainfall. The insurance guarantees the production till mid-October and for a maximum of five months of the crop cycle. Damages from hailstorms are considered separately due to their distinct characteristics in frequency, scope, and type of damage. These expansions of coverage have been accompanied by a significant increase in claims in recent years. This suggests that damages related to heavy (over 40mm in 24h) or persistent rainfall (over a week with consecutive rainy days each with more than 0.2mm) may have been underestimated. However, since the damages covered by the current processing tomato insurance are insufficiently defined quantitatively in terms of meteorological variables, it is difficult to assess the frequency and magnitude of the crop losses.

If rainfall risk could be defined in terms of measurable weather variables, indemnities might be estimated by applying mathematical formulae to the observed weather data recorded in a given weather station, to which the local farmers purchasing the insurance policy would be associated. This could have a number of advantages. First, it might facilitate the valuation of current and future risk by using long weather data series and future climate projections. Second, it might simplify the payment of indemnities by improving the current system of visual (and/or sampling) assessment by experts, as it is currently done. Also, decrease the cost of damage appraisal, and as a consequence, overall the whole system. Claims under such a Weather Index Insurance (WII) would be indemnified based on values obtained from an index that serves as a proxy for losses. Success of such a system requires that the index for observed weather events is strongly correlated with crop losses on all farms; otherwise it would produce false positives and false negatives. This is one of the reasons why WII has been little adopted in all countries where it has been offered commercially (Binswanger-Mkhize, 2012).

The objective of this work was to define the incidence of rainfall-related risks for processing tomato production in the Extremenian Guadiana River basin of W Spain in terms of measurable and objective variables. First, the risk related to rainfall was modelled as a function of meteorological, terrain and management variables using data from the recent period of broadest insurance coverage. Second, the resulting model was applied to the previous period with more restricted coverage in order to evaluate the possible impact of the extended coverage on the observed increase in claims. Results are used to discuss the opportunities to insure processing tomato against rainfall risks through a WII. To our knowledge, this is the first attempt to check the potential of WII for horticultural crops.

2. Materials and methods

2.1 Current damage appraisal for insurance payments

Indemnities are based on in-field evaluation made by loss adjusters of *Agroseguro*, the pool of insurance companies, following explicit monitoring procedures (Antón and Kimura, 2011). The procedure has the following steps: (a) the farmer files a notification of the event considered to have caused the observed crop loss; (b) *Agroseguro* checks the occurrence of the event and arranges the visit of a loss adjuster to the farm; (c) the adjuster

gathers data, photos, and takes samples, recording all data in a tablet computer; (d) the adjuster either calculates the indemnity while at the farm or sends an offer within a few days; (e) the farmer can accept or reject the offered indemnity; (f-1) if it is accepted the payment is transferred within two months; (f-2) if it is rejected, *Agroseguro* reviews the adjustment and sends a new indemnity proposal; (g) if the farmer does not accept it, he/she must file a lawsuit. This whole process is subject to quality control procedures independently performed by both *Agroseguro* and the *Consorcio de Compensación de Seguros*, which is the public reinsurance entity in Spain.

2.2. Processing tomato sector

Spain is the fourth largest producer of processing tomato in the world, behind USA (California), China, and Italy. About 70% of the national production is harvested in the Guadiana River Basin in Extremadura, Western Spain (Macua-González et al., 2012). Yields average 80Mg ha⁻¹ fresh weight while best fields and farmers achieve 100 Mg ha⁻¹.

The study region is located in the middle of the production area (Figure 1) in the province of Badajoz, W Spain. Summers are warm and relatively dry while winters are mild and wet. The average monthly rainfall and mean maximum and minimum temperatures for the town of Badajoz obtained from AEMET (the Spanish Agency for Meteorology) are presented in Figure 2. The average altitude in the western area is 285m and the dominant soils have silty clay texture (Badajoz County). The eastern area has an altitude of 430m and the soils have a gradually greater content of sand (Don Benito County).

In the study area, production is organized through a technical committee that comprises tomato farmers, cooperatives and about fifteen tomato industries of different size that process the whole production of the region. This committee negotiates total cultivated area, prices, cultivars, quality standards and delivery calendars. These decisions are negotiated in order to facilitate interactions between farmers and industry, and the organization of the harvest season. Vertical coordination in the sector ensures that eligibility criteria to receive the subsidies from the Common Agricultural Policy are met. Homogeneity of production is based on a selection from the long list of available varieties as proposed by the technical committee. The cooperatives organise the planting and delivery calendars in order to share the risks of an early or late planting among all farmers

holding production quotas. Other decisions, including irrigation and fertilization, are taken by individual farmers. Input prices and adverse climate events remain uncontrollable.

Figure 1. Study region framed in the middle Guadiana river basin (W Spain).

Processing tomato is grown between April and October. Planting dates are scheduled and assigned to farmers to distribute the harvest over the two-month period when the processing industry is operating. Planting is also conditioned by the weather conditions because earliest planting dates are limited by frosts and latest by autumn rainfall. Harvest begins at the end of July and finishes at latest by early October.

Main weather hazards for field-grown tomato are hail and extreme rainfall. Tomato is very sensitive to high soil moisture content. Extreme rainfall events may cause flooding in heavy soils or flat areas. Although drip irrigation is gaining importance, furrow or flood irrigation remains common because land is levelled for flooded rice that is included in the crop sequences. Damages associated with waterlogging are hypoxia that inhibits respiration of roots and causes accumulation of endogenous ethylene. With rising temperatures, the effects of flooding lead to wilting and death of plants. The accumulation of surface water during warm days increases relative humidity within the crop encouraging the growth and

infection by bacterial and fungal diseases, such as powdery mildews (Guzman-Plazola et al., 2003).

Figure 2. Climodiagram corresponding to Badajoz. Average values for the period 1955-2011.

Proper drainage and levelling of fields and use of resistant varieties are among the management practices proposed to avoid problems related to waterlogging. Adverse weather exacerbates disease problems because phytosanitary treatments are not applied unless there is evidence of disease or unless rainfall is forecast during sensitive crop growth periods.

Risks related to rainfall have been covered by insurance since 1999. Initially, damages were indemnified only for crops that had reached fruit setting. The guarantee period was extended in 2010 to include damage occurring beforehand. In present insurance contracts, rainfall-related hazards are catalogued as “torrential rainfall”, “persistent rainfall”, or just “rainfall”, with no objective definition in terms of quantity or frequency. Crop loss is

related to fruit wilting or physical damage to vegetative tissue hampering crop recovery. Indemnities are only paid when damages affect more than 20% of the insured tomato yield, which is based on local expected yield considering the insurance records.

Risk is the product of hazard and vulnerability, whereas claim is the record of an incident after a farmer reports crop damage due to a weather event. In this paper, the risk during the growth of the crop is treated differently to that at harvesting. Only records that resulted in an indemnity payment are included as claims in this study.

2.3. Insurance database

Data of all processing tomato farmers insured in the region from 1990 to 2011 were made available by the agricultural insurance company (*Agroseguro*). Individual insurance records include information of field location, insurance contract date, insured yield and field size, cultivars used, risks suffered and date. Several filters were applied to the database in order to avoid duplicate records and typing mistakes. Duplicates that were identified by comparing farmer's and field's IDs for individual years were deleted. Lastly, fields insured before 2006 were also deleted because they did not contain transplanting date. *Agroseguro* protocols for ensuring data confidentiality were followed.

Figure 3 shows the frequency of claims after transplanting and the degree day summation by the crops ($^{\circ}\text{C d}$, base temperature of 10°C). Two periods can be differentiated, one between days 30 and 60 after transplanting, corresponding to 300-600 $^{\circ}\text{C d}$, and the second between days 100 and 130 after transplanting, corresponding to 1200 $^{\circ}\text{C d}$. Risks were identified according to these individual periods. Claims before day 90 after transplanting were coded as related to risks during crop growth (C_Growth), while claims registered in crops that had reached 90 days after transplanting were coded as related to risks during harvesting (C_Harvest).

The three municipalities, Badajoz, in the West, Mérida and Don Benito, in the East, were selected as the largest producers from 2006 and 2011. The corresponding weather stations were Bercial, Mérida, and Palazuelo, for which data are available in the System of Agroclimatic Information for Irrigated crops (SIAR). A database of 16353 observations was constructed with each entry corresponding to an insured field that were not, however,

homogenously distributed between the three counties. About two thirds of the data corresponded to fields located in Don Benito, followed by *ca.* one third in Badajoz, and lastly, Mérida, with the smallest number. The incidence of rainfall-related damages differed between years and county ranging from zero to 23.3% of the insured fields registering a claim (Table 1).

Figure 3. Claims depending on the days after transplanting and on the accumulated degree days of the crop.

The regression model includes continuous and categorical variables that represent different components of the system, viz. management, weather, topography. The management variable (TrD) to include the discussion on adaptation strategies; weather variables that explained risk occurrence would constitute the basis to design a WII; the significance of topographic variables would suggest the need to segment insurance conditions depending on the physical location of the fields; the County variable was included to evaluate the scale at which the model was valid; lastly, significance of the Year variable was oriented to assess the time-specificity of the model and, therefore, test the limitation of the cropping system to be insured by a WII.

Weather and topographic variables were included by cross-checking field data with meteorological data from local weather stations and a digital elevation model. In addition

new variables were calculated for slope and accumulation of runoff. These variables were calculated with the corresponding tool in the ESRI ArcMap 10.0 GIS over the Digital Elevation Model facilitated by the Service of Digital Cartography and Spatial Data Infrastructure belonging to the University of Extremadura. The variables are described in Table 2. Table 3 provides a description of the continuous independent variables.

Table1. Fields insured and percentage of claims during growth and at harvest per year and county.

Year	County								
	Badajoz			Mérida			Don Benito		
	n	C_Growth	C_Harvest	n	C_Growth	C_Harvest	n	C_Growth	C_Harvest
2006	580	0.00	0.00	76	3.95	0.00	790	0.38	0.00
2007	451	0.67	15.44	97	0.00	13.40	1405	0.07	1.94
2008	754	1.19	1.85	179	1.68	4.90	1335	3.52	0.00
2009	1275	0.00	0.00	216	0.00	0.00	2136	0.00	0.00
2010	1562	0.55	1.14	253	1.85	0.00	2623	1.64	0.86
2011	759	11.91	8.30	179	23.32	0.00	1683	2.74	0.21
Total	5381			1000			9972		

Source: Agroseguro.

Risk incidence varies between years, in part due to changes in the insurance coverage from 2010, so the risk model was fitted to three different datasets of both C_Growth and C_Harvest. The first dataset (*All*) included the insured fields in the three selected counties from 2006 to 2011; the second (*High*) included only the fields insured in the years for which indemnified fields exceeded 1% of the insured fields in at least one of the counties. In this set, greater losses were recorded during crop growth in 2008, 2010 and 2011 but at harvest in 2007 and 2011. The third dataset (*Extended*) included data from 2010 onwards, and therefore only fields insured with the extended risk coverages.

Table 2. Description of the dependent and independent variables included in the initial analysis.

Variable	Description
<i>C_Growth</i>	Claims during the growth of the crop
<i>C_Harvest</i>	Claims at harvesting
<i>Pr1</i>	Maximum daily precipitation between days (30 and 60) – (100 and 130) after transplanting
<i>Pr2</i>	Accumulated precipitation between days (30 and 60) – (100 and 130) after transplanting
<i>Pr3</i>	Variance in daily precipitation between days (30 and 60) – (100 and 130) after transplanting
<i>Pr4</i>	Maximum of consecutive days raining between days (30 and 60) – (100 and 130) after transplanting
<i>Pr5</i>	Maximum of the accumulated precipitation in five consecutive days between days (30 and 60) – (100 and 130) after transplanting
<i>Pr6</i>	Maximum daily precipitation in the crop growth period
<i>Pr7</i>	Accumulated precipitation in the crop growth period
<i>Te1</i>	Maximum of the maximum daily temperature reached between days (30 and 60) – (100 and 130) after transplanting
<i>Te2</i>	Minimum of the minimum daily temperature reached between days (30 and 60) – (100 and 130) after transplanting
<i>Te3</i>	Variance of the maximum daily temperature reached between days (30 and 60) – (100 and 130) after transplanting
<i>Te4</i>	Variance of the minimum daily temperature reached between days (30 and 60) – (100 and 130) after transplanting
<i>Te5</i>	Maximum of the moving average of the maximum temperature (5 d) between days (30 and 60) – (100 and 130) after transplanting
<i>Te6</i>	Minimum of the moving average of the minimum temperature (5 d) between days (30 and 60) – (100 and 130) after transplanting
<i>TrD</i>	Transplanting date (Day of the year)
<i>To1</i>	Average slope in the field
<i>To2</i>	Minimum slope in the field
<i>To3</i>	Average flow accumulation in the field
<i>To4</i>	Maximum flow accumulation in the field
<i>A</i>	$Pr4 \times Te5$
<i>B</i>	$Pr5 \times Te5$
<i>C</i>	$Pr5 / Pr4$
<i>D</i>	$Pr4 / Te3$
<i>E</i>	$Pr5 / Te3$
<i>County</i>	----
<i>Year</i>	----

Table 3. Descriptive statistics of continuous independent variables.

GROWTH						HARVESTING					
Variable	Obs	Mean	Std. Dev.	Min	Max	Variable	Obs	Mean	Std. Dev.	Min	Max
TrD	16353	114.1	14.0	91	151	TrD	14216	116.3	14.8	91.0	151.0
Pr1	16353	12.8	9.0	0.2	32.2	Pr1	14216	5.5	6.9	0.2	66.4
Pr2	16353	27.7	19.8	0.2	80.8	Pr2	14216	8.4	10.9	0.2	104.2
Pr3	16353	0.5	0.2	0.2	1.0	Pr3	14216	0.8	0.2	0.2	1.0
Pr4	16353	3.5	2.1	1.0	10.0	Pr4	14216	1.8	1.1	1.0	7.0
Pr5	16353	23.1	14.5	0.2	48.4	Pr5	14216	7.5	9.4	0.2	93.0
Pr6	16353	4.6	2.9	0.0	9.7	Pr6	14216	17.7	7.8	6.0	66.4
----	----	----	----	----	----	Pr7	14216	75.3	38.5	18.4	205.4
Te1	16353	35.9	2.4	23.4	41.8	Te1	14216	38.3	1.7	27.9	42.1
Te2	16353	9.4	2.1	1.7	15.8	Te2	14216	11.6	2.7	0.0	17.6
Te3	16353	4.1	0.8	0.1	6.2	Te3	14216	3.2	0.8	1.1	5.7
Te4	16353	2.5	0.6	0.2	4.4	Te4	14216	2.2	0.5	0.5	4.1
Te5	16353	34.3	2.5	23.0	40.3	Te5	14216	34.7	2.1	28.0	40.3
Te6	16353	17.6	2.0	11.3	24.4	Te6	14216	17.9	1.7	12.5	24.4
To1	16353	3.0	1.6	0.0	17.9	To1	14216	3.0	1.6	0.0	17.9
To2	16353	0.2	0.7	0.0	9.0	To2	14216	0.3	0.7	0.0	9.0
To3	16353	10.8	12.7	0.0	263.3	To3	14216	10.9	13.0	0.0	263.3
To4	16353	238.8	311.6	0.0	3108	To4	14216	235.8	314.2	0.0	3108
A	16353	116.4	65.7	27.3	361.2	A	14216	62.7	38.0	28.0	273
B	16353	775.9	482	7.2	1696.9	B	14216	266.1	334.1	5.6	3470.8
C	16353	7.1	4.9	0.2	40.0	C	14216	4.2	5.0	0.2	34.6
D	16353	0.9	0.6	0.2	7.1	D	14216	0.6	0.3	0.2	2.2
E	16353	5.8	3.8	0.1	38.6	E	14216	2.2	2.6	0.1	30.3

2.4. Risk model

A binary logistic regression model (also known as logit model) was used to investigate the influence of management, topography and meteorological variables on the occurrence of rainfall-related damages for individual fields. The model assumes a binomial distribution for the binary dependent variable and estimates the probability of a claim based on the independent variables of continuous or categorical nature. Indemnified fields are coded “1” and the remainder “0”.

The risk model is defined as:

$$\log \left[\frac{P(Y=1)}{1-P(Y=1)} \right] = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad [\text{Eq 1}]$$

where $P(Y = 1)$ and $1 - P(Y = 1)$ are the probabilities “indemnity” or “no indemnity”. It can be rewritten to predict the probability, between 0 and 1, of a claim for individual fields based on the independent variables (x_i 's) as

$$P(Y = 1) = \frac{e^{\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n}}{1 + e^{\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n}} \quad [\text{Eq 2}]$$

where the parameter α is the intercept in the model, and β_i are the regression coefficients for each covariate x_i .

The model was fitted with Stata v12 (StataCorp, 2011).

The inclusion of the independent variables in the model was performed using the *stepwise selection procedure*. New variables are added one at a time (forward steps). After each forward step, the log likelihood values of the included variables are reexamined and any non-significant variables are removed. This process continues until no variables are either added or removed from the model.

Jackknife post estimation (Efron, 1981) was performed in order to ensure that parameter estimation was robust. This procedure improves robustness generating alternative design-based standard errors for hypothesis testing and confidence intervals which reduce reliance on theoretical assumptions. The odd ratio was used to rate the independent variables that

better explain the dependent one. This ratio is interpreted as the effect of one unit of change in a variable when the other variables remain constant.

The indices of sensitivity and specificity were used to evaluate the accuracy of the models' fit. They are intrinsic qualities of a test and describe how well the models discern between true positives and true negatives. The sensitivity measures the power of the model to identify positives and the specificity measures the power to identify negatives:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad [\text{Eq 3}]$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad [\text{Eq 4}]$$

In this case, TP (true positives) is the number of “claims” correctly identified by the model and TN (true negatives) is the number of “no-claims” cases correctly identified by the model. FN (false negatives) is the number of “no-claims” identified as “claims” and FP (false positives) is the number of “claims” identified as “no-claims”.

Logit models output a probability of success. This probability is a continuous variable that is transformed into the binary variable *success-no success* using a cut-off value. The cut-off value is the probability at which an observation is to be classified either as positive or as negative. It influences the resulting number of true positives, true negatives, false positives, and false negatives after applying the model to a specific set of data. For a change in the cut-off selected, an increase in the sensitivity will result in a decrease in specificity and vice versa. The cut-off selection must reflect the best tradeoff between both indices, and therefore, obtain the optimum allocation of estimated positives and negatives to observed data. As a default, logit models often use a cut-off value of 0.5; for this analysis, we selected the cut-off values 0.2 and 0.8 to evaluate the significance of this parameter in the final classification of expected claims.

Table 4. Logit models for claims during the growth of the crop and at harvesting.

GROWTH	DS ¹		Variables (Significant variables in bold, $p \leq 0.05$)
	<i>All</i>	2006-2011	Te1 Pr3 Te6 TrD Te2 D To3 Pr2 B Pr5 Te3 Te5 Pr1 C E Te4 To1 Pr4 A
	<i>High</i>	2008 2010 2011	Te1 Pr3 Te6 TrD Te2 Pr2 D To3 Te5 C Pr1 Te3 B Pr5 Te4 E A Pr4
	<i>Extended</i>	2010-2011	<i>Te1 Pr3 D Te2 TrD E Te3 To3 Pr1 Pr2 C A Pr4 Te5 Pr5 B</i>
	Common variables		(Pr1 Pr2 Pr3 Pr4 Pr5) (Te1 Te2 Te3 Te5) (TrD) (A B D E)
HARVEST	DS		Variables (Significant variables in bold, $p \leq 0.05$)
	<i>All</i>	2006-2011	Pr1 Te6 TrD Te2 Te4 Pr7 C Te5 Pr2 D A Pr6 Pr5 Pr4 Te1 To2 To1 To3 Pr3
	<i>High</i>	2007 2011	<i>C TrD Te4 Te2 Te6 Pr2 Te5 Te3 A B Pr4 Pr3 Pr1 To2 D E To3 To1 Pr7 Pr6</i>
	<i>Extended</i>	2010-2011	<i>Pr1 Pr4 D A Te2 Te4 TrD Pr6 Te6 Te3 Pr3 B Pr7 E Te5 Pr5 Pr2</i>
	Common variables		(Pr1 Pr2 Pr6 Pr7) (Te1 Te2 Te4 Te5) (TrD) (B)

¹ DS: Datasets *All* (2006-2011); *High* (Years with higher proportion of damages); *Extended* (2010-2011)

3. Results

3.1. Statistical analyses

Models were fitted for each of the damages (C_Growth and C_Harvest) for the three datasets (Tables 4, 5 and 6). The categorical variables, County and Year, were included to evaluate their marginal contribution to the probability of claims and hence of the potential for development of WII that could only be based on weather variables. Fields located in Mérida in 2010 were not included for risks at harvesting as no claims were registered that year in that county.

Table 4 lists the variables included in the models according to their contribution to precision, the significant ones in bold. The common variables for both definitive models for each risk (C_Growth and C_Harvest), present after applying the stepwise selection procedure, are in italics. The weather variables related to variances in daily precipitation (Pr3) and temperatures (Te1, Te2, and Te6) were the first included in the stepwise selection procedure but transplanting date (TrD) was also significant. Topographic variables made a minor contribution to explain risk incidence. Models for dataset *Extended* included fewer variables than for databases *All* and *High*.

Table 5. Goodness of fit (R^2 statistic) with and without categorical variables included.

	Categorical variables effect on R^2						
	DS ¹	Model	n	None	County	Year	County & Year
GROWTH	<i>All</i>	2006-2011	16353	0.56	0.58	0.68	0.69
	<i>High</i>	2008 2010 2011	9327	0.66	0.67	0.71	0.71
	<i>Extended</i>	2010-2011	7059	0.70	0.71	0.71	0.72
	Categorical variables effect on R^2						
	DS	Model	n	None	County	Year	County & Year
HARVEST	<i>All</i>	2006-2011	15759	0.38	0.39	0.50	0.50
	<i>High</i>	2007 2011	4355	0.48	0.50	0.52	0.53
	<i>Extended</i>	2010-2011	6605	0.50	0.51	0.60	0.61

¹ DS: Datasets *All* (2006-2011); *High* (Years with higher proportion of damages); *Extended* (2010-2011)

Table 5 records the number of fields used to fit each of the models (n) and the goodness of fit (R^2) with and without the categorical variables, County and Year. Their inclusion substantially improves risk analysis of both C_Growth and C_Harvest models. Model *All* shows the largest improvement, with an increase of R^2 from 0.56 to 0.69 for risks during C_Growth, and from 0.38 to 0.50 for C_Harvest. In both cases, the largest effect is attributed to Year.

A classification of the sensitivity and specificity of the models is presented in Table 6. The fraction of correctly classified scenarios was greater when using a cut-off criterion of $p = 0.2$ compared to a cut-off of $p = 0.8$. In the models based on dataset *Extended* (years 2010 and 2011), a cut-off of $p = 0.2$ resulted in a sensitivity of 80% and a specificity over 98%.

Table 6. Sensitivity and specificity of the models the categorical variable County included.

	DS ¹	Claims (%) / cut-off 0.2		Claims (%) / cut-off 0.8	
		Sensitivity	Specificity	Sensitivity	Specificity
GROWTH	<i>All</i>	65.97	99.25	39.81	99.98
	<i>High</i>	75.12	98.21	54.50	100.00
	<i>Extended</i>	77.41	98.30	64.46	100.00
	DS	Claims (%) / cut-off 0.2		Claims (%) / cut-off 0.8	
		Sensitivity	Specificity	Sensitivity	Specificity
HARVEST	<i>All</i>	45.07	99.46	10.77	99.98
	<i>High</i>	70.80	98.18	28.32	100.00
	<i>Extended</i>	71.97	99.46	30.57	100.00

¹ DS: Datasets *All* (2006-2011); *High* (Years with higher proportion of damages); *Extended* (2010-2011).

Performance of the model for *Extended* dataset for claims during C_Growth with inclusion of the categorical variable County is presented in Table 7. All selected variables for rainfall

were significant. For variables with greater odd ratios, risk increases coincide with increases in variance of daily precipitation (Pr3) and increases in the maximum of accumulated precipitation within five consecutive days (Pr5). This suggests that claims are related to periods of intense and concentrated rainfall. The probability of damage diminished as transplanting date was delayed (management). The categorical variable County reveals that compared to the reference category, Badajoz, risk was significantly higher in Mérida and Don Benito.

Table 7. Parameters estimate for the logit model for dataset *Extended* explaining claims during the crop growth including the categorical variable County.

Variable	Coefficient	Standard error	p- value ²	Odd ratio	95% CI ¹
Intercept	127.439	28.9	***	--	101.9 / 198.4
Meteo					
Pr1	-0.5	0.1	***	0.6	-0.7 / -0.3
Pr2	0.1	0.0	***	1.1	0.1 / 0.2
Pr3	34.4	4.2	***	>1000	26.1 / 42.7
Pr4	-23.7	3.9	***	0.0	-31.4 / -16.1
Pr5	2.8	0.8	***	16.0	1.1 / 4.4
Te1	-1.4	0.8	*	0.2	-3.0 / 0.2
Te2	1.4	0.1	***	4.2	1.2 / 1.7
Te3	0.1	0.6	-	1.1	-1 / 1.3
Te5	-3.4	1.2	***	0.0	-5.7 / -1.0
A	0.7	0.1	***	2.0	0.5 / 0.9
B	-0.1	0.0	***	0.9	-0.1 / 0.0
D	-1.2	1.9	-	0.3	-5.0 / 2.6
E	-1.1	0.3	***	0.3	-1.7 / -0.4
Management					
TrD	-0.239	0.0	***	0.00	-0.3 / -0.2
County					
Badajoz	Reference	//	//	//	//
Mérida	4.790	1.7	***	34.1	0.5 / 6.5
Don Benito	5.323	1.7	***	180.2	2.0 / 8.4

¹ Coefficients confidence interval

² - p≥0.1* p<0.1; ** p<0.05; *** p<0.01

Performance of model for dataset *Extended* in explaining claims at harvesting and including the categorical variable County is presented in Table 8. Variables with higher odd ratios for daily precipitation, (Pr1, Pr6) and minimum temperature (Te2, Te4) suggest that claims coincide with heavy rainfall and high temperatures. In flat fields, prone to waterlogging, this is related to an increase in the relative humidity. Once again the probability of damage was greater when transplanting date was delayed. Coefficients of the categorical variable County that were estimated relative to Badajoz (the reference group) reveal that the number of claims was significantly higher in Don Benito. No conclusion could be drawn in Mérida because the number of claims was too small.

Table 8. Parameters estimate for the logit model for dataset *Extended* explaining claims at harvesting including the categorical variable County (Mérida not included).

Variable	Coefficient	Standard error	p- value ²	Odd ratio	95% CI ¹
Intercept	-50.3	-5.3	***	--	-68.8 / -31.8
Meteo					
Pr1	0.5	8.0	***	1.6	0.4 / 0.6
Pr2	-0.1	-3.0	***	0.9	-0.2 / 0.0
Pr5	-2.5	-7.5	***	0.1	-3.2 / -1.9
Pr6	0.2	1.3	-	1.2	-0.1 / 0.4
Te1	-0.2	-1.5	-	0.8	-0.4 / 0.1
Te2	1.5	10.5	***	4.5	1.2 / 1.8
Te4	4.8	9.0	***	116.2	3.7 / 5.8
Te5	-0.1	-0.4	-	0.9	-0.3 / 0.2
B	0.0	2.2	***	1.0	0.0 / 0.0
Management					
TrD	0.2	4.9	***	1.2	0.1 / 0.3
County					
Badajoz	Reference	//	//	//	
Don Benito	1.6	3.3	***	5.2	0.7 / 2.6

¹ Coefficients confidence interval

² - p≥0.1* p<0.1; ** p<0.05; *** p<0.01

3.2. Extrapolating risk

Figure 4 shows the application of the *Extended* statistical model, including the categorical variable County, to claims during crop growth and at harvesting between 2006 and 2011. Risk at harvesting was not estimated for Mérida since the model could not be validated for this county. Differences between observed and estimated risk differed between counties during this period. Overall, estimated risk during C_Growth was greater than the registered claims during 2006 to 2009. The overestimation was significant in 2008 in Badajoz (1.19% *vs.* 49.2%), in 2006 and 2008 in Mérida (3.95 *vs.* 28.94, and 1.68 *vs.* 35.75%, respectively), and in 2007 in Don Benito (0.07% *vs.* 31.10%). Estimated risk at harvest showed less differences with respect to registered claims, except for 2007 in Badajoz, where risk was underestimated (15.44% *vs.* 0.45%).

Figure 4. Estimated and observed field affected by rainfall during the growth of the crop.

4. Discussion and conclusions

The main objective of this paper was to assess the opportunities of protecting farmers of processing tomato against rainfall-related risks through a weather insurance index.

Claims registered during the growth of the crop (C_Growth) and those at harvesting (C_Harvest) were modelled separately to account for different crop vulnerabilities and different management decisions, even if in both cases the claims were. The first period corresponds to crop growth and development to flowering and early fruit formation, while the second period coincides with the end of fruit maturation up to harvesting (Onofri et al., 2009). Rao & Li, 2003 reviewed the physiological effects of flood duration and growth stage on tomato crops. They established greater sensitivity to flooding during flowering and that damages were related to ethylene accumulation and stomatal closure. At the end of crop development, flooding conditions hasten fruit maturation and decrease quality. Moreover, rainfall at harvesting prevents machinery from operating in fields at the optimal time, producing crop over maturation and consequently yield loss.

The differences in the most significant variables of the models, extreme temperatures and daily precipitation variance, during crop growth and at harvesting suggest that damages at different stages of crop development correspond indeed to different risks. Fields of processing tomato are more vulnerable when exposed to intense and concentrated rainfall during C_Growth, while at harvest they are threatened primarily by rainfall events combined with high temperature. The influence of high temperatures after rainfall events may be related to an increase in relative humidity that favours fungal diseases reducing crop growth and fruit quality (Guzman-Plazola et al., 2003). These conditions also accelerate fruit maturation and shorten the period the period the period that the farmer has before yield is lost. The transplanting date influences risks during both C_Growth and C_Harvest periods. Early transplanted plants are more exposed to damages during vegetative growth, while transplanting at later dates rises the risks at harvesting. Model results suggest, unsurprisingly, that shortening the cropping period by avoiding the earliest and latest transplanting dates would decrease risk. However, changes in harvest dates must be coordinated with the industry.

Agricultural risks can be characterised by their covariance, frequency and severity of damages (Hazell et al., 2010). The model improves substantially when categorical variables

are included, and the greatest R^2 was obtained with database *All*, for risks both during crop growth and risks at harvesting. This suggests that risk incidence is geographic- and time-specific, a finding that may explain why the Weather Index Insurance (WII) has found limited application for many crops and why it is little used in most places around the world where it is offered commercially (Binswanger-Mkhize, 2012).

Covariance is related to the degree at which claims are correlated across fields within a region. Meteorological data used in this analysis were obtained from regional stations rather than measured weather in individual fields. In a WII, weather variables operate as derivatives through which a farmer is indemnified depending on the weather at a certain moment rather than from directly measured damages. This means that there must be a trade-off between accuracy and practicality to avoid major discrepancies between actual losses and weather derivative indemnities; this is a common risk when using weather derivatives (Vedenov and Barnett, 2004). The significant influence of the categorical variable County, suggests that the risk under study has a medium to low covariate because the rainfall event impacts heterogeneously in the region. This geographic dependence suggests that extreme rainfall events cause fairly localized damage, influencing the scale at which the model might have validity. This, together with the year dependency, questions the potential of implementing index-based insurance for processing tomato rainfall extreme events.

A logit regression was selected to model the frequency rainfall-related risks in field tomato crops in order to identify the variables that significantly impact the probability of registering a claim. The models estimated in this paper explain the likelihood of a claim being registered for a tomato field as a function of weather, topography, transplanting date, and location. The objective of the model was not to establish cause – effect relationships for each independent variables but to measure the overall probability of a claim. Weather variables were then built over the periods with the highest probabilities of registering a claim, and not to the previous days of each of the observed claims. Therefore, in this work the models are used to study the concurrence of a weather event and the probability of registering a claim.

Although the coefficient of determination is moderately high (~ 0.75) in models for database *Extended* when it includes the categorical variable County, specificity approaches 100% and sensitivity exceeds 75% when using a cut-off of 0.2 ($p > 0.8$). That is to say that a

non-negligible proportion of observations could not be explained by the model. This might be associated to the low number of claims in proportion to the total number of insured fields and also to possible sparseness and/or multicollinearity of the data. Due to the unbalanced number of observations between counties, it could be expected that the model captures better the reality of the most represented ones.

Agricultural risks related to extreme climatic events have a low frequency and high severity (Binswanger-Mkhize, 2012). Therefore, the small number of successes is an inherent characteristic to the rationale of insurance. From a statistical point of view, insured crop claims might be then treated as “rare events”, as recommended by sound actuarial criterion (Seog, 2010). In our case, the range is between 0 and 35%, depending on the year and location. According to King & Zeng, (2001), it is difficult to model these cases for two main reasons. First, logistic models tend to underestimate the probability of occurrence of these events. Second, data collection becomes harder when success events are infrequent and so datasets often result a huge number of observations and relatively few and poorly measured explanatory variables.

Claims increased from 2010, in parallel to the extension of insurance coverage. The model fitted for database *Extended* (years 2010 and 2011) was used to evaluate how many claims would have been during 2006 to 2009 if the broadest insurance coverage scenario had been in place. The differences between simulated risk for the broadest coverage insurance and observed damages in 2006 to 2009 indicate that the extended guaranteed period does include actual threats that had been previously afforded by the farmer. From an actuarial point of view, risk assessment must accept that historical data may be insufficient to establish the real risk the insurance is covering. Moreover, the significant increase in indemnities in 2011 are not just related to an unusual year in terms of the precipitation pattern, but are also a consequence of the new integrated guarantees, as initial insurance coverage did not protect farmers when damages occurred before fruit setting.

Lastly, a WII requires an accurate model of the severity of damages (percentage over the insured yield) through objective and measurable variables. Logit model results point at a scale and year dependency of the risk. This suggests that data measured at the regional level are not able to fairly reproduce individual field risks. Modelling the severity of damages requires more detailed variables at field level, in terms of more detailed management

information of transplanting date, field history, and especially weather data measured in individual fields.

In this work, insurance data included mainly variables related to location, transplanting date, and expected yield. Meteorological and topographic variables were not available in the insurance database and therefore it had to be crosschecked using transplanting dates and location. The criterion of associating a weather station according to an administrative border obviates the field-to-field variances. This results in low sparseness and high probability of multicollinearity of the meteorological data for fields with similar transplanting dates. This limits the power of the model to explain risk incidence. The model was also limited in the number of field specific variables. There might be soil or management characteristics with a stronger influence on damage occurrence, which are absent in the insurance records datasets. The possibility of implementing Weather Index based Insurances (WII) is hampered by the significant dependency of the model on scale and year. Therefore, our work does not recommend the use of WII to crops that are very sensitive to weather anomalies and crop management.

5. Acknowledgements

This research was funded by the project MULCLIVAR, from the Spanish Ministerio de Economía y Competitividad (MINECO) CGL2012-38923-C02-02, A. Castañeda-Vera has a PhD grant from UPM. We thank *Agroseguro* and ENESA (Entidad Nacional de Seguros Agrarios) for the data used in the analyses, and Finca La Orden-Valdesequera, Cooperativa Casat (Don Benito, Badajoz) and the Centro Tecnológico Agroalimentario de Extremadura (CTAEX) for the clarifying information regarding processing tomato in Extremadura.

Chapter 3

3. Selecting crop models for decision making in wheat insurance

A. Castañeda-Vera^{1,2}, P.A. Leffelaar³, J. Álvaro-Fuentes⁴, C. Cantero-Martínez⁵ and I. Mínguez^{1,2}

¹ Research Centre for the Management of Agricultural and Environmental Risks (CEIGRAM), Universidad Politécnica de Madrid (UPM), Spain

² Agricultural Systems Group, Dpto. Producción Agraria, E.T.S. Ingenieros Agrónomos, Universidad Politécnica de Madrid, Avda. Complutense s/n, 28040 Madrid, Spain

³ Plant Production Systems, Wageningen University, Droevendaalsesteeg 1, 6708 PB, Wageningen, the Netherlands

⁴ Departamento de Suelo y Agua, Estación Experimental de Aula Dei, Consejo Superior de Investigaciones Científicas (EEAD-CSIC), POB 13034, 50080 Zaragoza, Spain

⁵ Departamento de Producción Vegetal y Ciencia Forestal, Unidad Asociada EEAD-CSIC, Agrotecnio, University of Lleida, Rovira Roure 191, 25198 Lleida, Spain

European Journal of Agronomy 68 (2015): 97 - 116.

Abstract

In crop insurance, the accuracy with which the insurer quantifies the actual risk is highly dependent on the availability on actual yield data. Crop models might be valuable tools to generate data on expected yields for risk assessment when no historical records are available. However, selecting a crop model for a specific objective, location and implementation scale is a difficult task. A look inside the different crop and soil modules to understand how the outputs are obtained might facilitate model choice. The objectives of this paper were (i) to assess the usefulness of crop models to be used within a crop insurance analysis and design and (ii) to select the most suitable crop model for drought risk assessment in semi-arid regions in Spain. For that purpose first, a pre-selection of crop models simulating wheat yield under rainfed growing conditions at the field scale was made, and second, four pre-selected models (Aquacrop, CERES-Wheat, CropSyst and WOFOST) were compared in terms of modelling approaches, process description and model outputs. Results suggest that the outputs of the four models for the simulation of winter wheat growth are comparable when water is not limiting, but differences are larger when simulating yields under rainfed conditions. These differences in rainfed yields are mainly related to the dissimilar simulated soil water availability. We concluded that for the simulation of winter wheat growth at field scale in such semi-arid conditions, CERES-Wheat and CropSyst are preferred. WOFOST is a satisfactory compromise between data availability and complexity when detail data on soil is limited. Aquacrop integrates physiological processes in some representative parameters, thus diminishing the number of input parameters, what is seen as an advantage when observed data is scarce. However, the high sensitivity of this model to low water availability limits its use in the region considered. Contrary to the use of ensembles of crop models, we endorse that efforts be concentrated on selecting, studying and calibrating the most adequate crop model on the basis of the objectives for which simulations will be used. The use of such complex methodologies as crop models is associated with numerous sources of uncertainty, although these models are the best tool available to get insight in these complex agronomic systems.

Keywords: Aquacrop, CERES-Wheat, CropSyst, WOFOST, Model choice, Rainfed semi-arid areas, Radiation use efficiency, Water deficit.

1. Introduction

Crop models are essential tools to understand the complexity of cropping systems since they compile knowledge on physiological processes and plant interactions with the environment, which are deemed crucial. In research, models are used to address research problems and interpret experimental results (Rinaldi et al., 2007), to evaluate the impact of alternative management strategies on production (Ventrella et al., 2012) and on the environment (Asseng et al., 1998a), to investigate crop production levels (van Ittersum et al., 2013) or to predict yields under changing climatic conditions (Asseng et al., 2013). Also in decision-making, crop models are increasingly used, for example, for policy shaping and analysis (e.g. CAP greening measures), farmer consultancy (e.g. Hunt et al., 2006) or risk management with early warning systems (e.g. Basso et al., 2013).

In crop insurance, risk is defined as the probability of registering a claim; in the case of drought insurance, a claim is registered when actual yield is lower than the insured yield. Consequently, risk depends on yield variability. The calculated risk is used to set and update premiums, or to design new insurance policies. The accuracy with which the insurer quantifies the actual risk is highly dependent on the availability on high quality actual yield data. However, records on observed yields are not always complete, and the information available for the insurer to calculate the probability and severity of a claim can be biased. This can give rise to unbalanced loss ratios, thereby affecting the actuarial robustness and sustainability of entire insurance systems. In these cases, crop models might be valuable tools to generate data on expected yields for risk assessment when no historical data is available.

However, the uncertainty associated with the use of crop models is large. Many crop and soil water models exist. They differ in aspects such as parameter requirements, time coefficient, simulation of the spatial scale or their either more physically or more empirically-based approach (e.g., Angulo et al., 2014; Kersebaum et al., 2007). Model complexity, the scale of application and the availability of data for model calibration and validation affect the reliability of simulations, especially when simulation purposes differ from those for which the selected model was designed (Kersebaum et al., 2007). In rainfed cropping systems in semi-arid areas, crops are highly dependent on soil moisture along the cropping cycle. The precision of soil water modules in simulating soil water dynamics and the capacity of the crop modules to translate the effects of water stress on crop canopy and biomass growth have an impact on the accuracy of simulated yields. Therefore, crop models should be used with caution, particularly when applied to more

resource-limited conditions and when used in decision support systems where environmental, social and economic assets are involved.

Asseng et al., (2013) found a larger uncertainty related to crop models than the one related to climate models when simulating under future climate projections, and that variation of simulated yield was larger for low-yielding environments, as subject to drought risk insurance. Moreover, Martre et al., (2014) studied two ensemble-based crop models and concluded that taking the mean or the median of the simulated values estimates better than any single crop model simulations.

Subsequently, a number of questions arise: (i) Which model(s) is/are preferred from the four investigated models? (ii) Is the use of several models better than the use a single model? (iii) Given the uncertainties associated to crop model simulations, is it worth to use crop models in crop insurance analysis and design? And what cautions should be taken when using crop model simulations for decision making regarding model(s) calibration and implementation?

The objectives of this paper were (i) to assess the usefulness of crop models to be used within a crop insurance analysis and design and (ii) to select the most suitable crop model for drought risk assessment in semi-arid regions in Spain. For that purpose first, a pre-selection of crop models simulating wheat yield under rainfed growing conditions at the field scale was made, and second, four pre-selected models were compared in terms of modelling approaches, process description and model outputs. Each pre-selected model calculates aboveground biomass accumulation and soil water balance using different approaches. Likewise, they use alternative forms to compute the effect of daily water stress on crop and biomass production. The four crop models were run for winter wheat (*Triticum aestivum* L.) over five growing seasons in NE Spain.

2. Materials and methods

2.1. Models

Four models were pre-selected from the 27 wheat simulation models included in the AgMIP wheat study (Asseng et al., 2013; Martre et al., 2014). The criteria to select them were having (i) a comparable structure in terms of submodules; (ii) different approaches to calculate the daily accumulation of biomass; (iii) different approaches to calculate the daily change in soil water content; and (iv) different approaches to calculate the penalties on crop growth due to water

deficit. The pre-selected models were: Aquacrop (Raes et al., 2009; Steduto et al., 2012), CSM-CROPSIM-CERES-Wheat (hereafter referred to as CERES-Wheat), available in the package Decision Support System for Agrotechnology Transfer (DSSAT) version 4.5 (Hoogenboom et al., 2012; Jones et al., 2003), CropSyst (Stöckle et al., 2003) and WOFOST (Boogaard et al., 2011; Supit et al., 1994).

With respect to the approach used to calculate the daily accumulation of biomass, Aquacrop calculates biomass production based on water availability through a transpirational water use efficiency coefficient (WUE, g biomass mm⁻¹); CERES-Wheat calculates biomass production rate based directly on radiation through a radiation use efficiency coefficient (RUE, g biomass MJ⁻¹); CropSyst combines the last two approaches, RUE and WUE; and, lastly, WOFOST calculates biomass production rate based on the net carbon assimilation by subtracting maintenance and respiration requirements from gross assimilation of CO₂ (Table 1).

The approaches used to calculate the daily change in soil water content were also different. Aquacrop uses a cascade approach (when no groundwater table is considered) that is computed on a 12 layers-subdivided soil; CERES-Wheat uses a cascade approach with upflow calculated based on soil water diffusivity with a user-defined soil subdivision; in CropSyst, the finite difference approach based on Richard's equation computed on a 20 layers-subdivided soil was selected; and, finally, WOFOST uses a cascade approach (when no groundwater table is considered) that is computed on a homogeneous soil of a single layer. The four models have a comparable structure. They have submodels for phenology and canopy development, for growth and biomass partitioning, and for a soil water balance.

The differences between the models are found in the detail of the different submodels. Main characteristics of the models are summarized in Tables 1 and 2, and most important equations and approaches are described in detail below.

Table 1. Main characteristics of Aquacrop, CERES-Wheat, CropSyst and WOFOST and the computation of water stress.

		AQUACROP	CERES-Wheat	CROPSYST	WOFOST
	Biomass accumulation	WUE ¹	RUE ¹	RUE and WUE	CO ₂ assimilation, maintenance and respiration
	Yield formation	AgB, HI ²	AgB, grains number and grain weight	AgB, HI	AgB, partitioning coefficients
	Roots modelling	Depth	Density and depth	Fraction per soil layer and depth	Depth
	Water uptake (Ta)	RAW ³ in the soil layer and soil depth	RAW and root density per layer	DWP ⁴ crop-soil and root fraction per layer	RAW in the root zone
Water deficit effect on...	... phenology	Early CC ⁵ senescence	No	Accelerate the accumulation of °C-days	No
	... AgB	$\downarrow Ta^7 \rightarrow \downarrow (d AgB) / dt$ $\downarrow CC \rightarrow \downarrow Ta \rightarrow \downarrow (d gB) / dt$	$\downarrow Ta \rightarrow \downarrow (d AgB) / dt$	$\downarrow Ta \rightarrow \downarrow (d AgB) / dt$	$\downarrow Ta \rightarrow \downarrow (d AgB) / dt$
	... partitioning	Modifies harvest index	Induces root growth Affects harvest index though sink size	Accelerates the accumulation of degree-days, hastening physiological maturity	No

¹ RUE: Radiation use efficiency; WUE: Water use efficiency² AgB: Aboveground biomass; HI: Harvest Index³ RAW: Readily Available Water, between pF 4.2 (–1500 kPa) and pF 2.5 (–33 kPa)⁴ DWP: Differences in water potential⁵ Canopy⁶ ↓: decrease, ↑: increase⁷ Ta: Actual transpiration; $\downarrow (d AgB) / dt$: Aboveground biomass production rate

2.2. Aboveground biomass production

CERES-Wheat calculates biomass production rate based directly on radiation through a radiation use efficiency coefficient (RUE, g biomass MJ⁻¹). Aquacrop calculates biomass production based on water availability through a transpirational water use efficiency coefficient (WUE, g biomass mm⁻¹). CropSyst combines the last two approaches, RUE and WUE. Lastly, WOFOST calculates biomass production rate based on the net carbon assimilation by subtracting maintenance and respiration requirements from gross assimilation of CO₂.

Aquacrop computes the daily aboveground biomass (AgB) production ($\frac{dAgB}{dt}$, kg ha⁻¹ day⁻¹) from the potential daily biomass production ($\frac{dAgB}{dt}_{POT}$, kg ha⁻¹ day⁻¹), the so-called water productivity coefficient (WP) (given as an input parameter), and the ratio of actual to reference evapotranspiration:

$$\frac{dAgB}{dt} = WP \times \left(\frac{Ta}{ET_0} \right) \quad [\text{Eq 1}]$$

$$\frac{Ta}{ET_0} = K_s \times CC^* \times K_{C,Tr} \quad [\text{Eq 2}]$$

$$CC^* = \min(1, 1.72CC - CC^2 + 0.3CC^3) \quad [\text{Eq 3}]$$

where Ta is the actual transpiration (mm day⁻¹), ET_0 is the reference evapotranspiration (mm day⁻¹), K_s is the water stress index (–), CC is the canopy cover (m² m⁻²), CC^* is the adjusted canopy cover for micro-advective effects (m² m⁻²) and $K_{C,Tr}$ (–) is the coefficient for maximum crop transpiration that corresponds with a well-watered soil and fully developed canopy ($CC=1$).

In CERES-Wheat, $\frac{dAgB}{dt}_{POT}$ is calculated from the intercepted photosynthetically active radiation ($IPAR$, MJ ha⁻¹ day⁻¹), that is previously calculated from the leaf area index (LAI , m² m⁻²), a light extinction coefficient (k , –) and PAR , and the radiation use efficiency (RUE , kg AgB dry matter MJ⁻¹) ($\frac{dAgB}{dt}_{IPAR}$, kg ha⁻¹ day⁻¹):

$$IPAR = PAR \times (1 - e^{-k \times LAI}) \quad [\text{Eq 4}]$$

$$\frac{dAgB}{dt}_{POT} = \frac{dAgB}{dt}_{IPAR} = RUE \times IPAR \quad [\text{Eq 5}]$$

where PAR is the photosynthetically active radiation ($\text{MJ ha}^{-1} \text{ day}^{-1}$) (estimated as half of the global radiation).

In CropSyst, the $\frac{dAgB}{dt}_{POT}$ is calculated as proposed by Tanner and Sinclair (1983) based on the minimum of (i) the radiation dependent biomass growth, based on the photosynthetically active radiation intercepted by the crop ($\frac{dAgB}{dt}_{IPAR}$, $\text{kg ha}^{-1} \text{ day}^{-1}$), and (ii) the transpiration dependent biomass growth ($\frac{dAgB}{dt}_{PT}$, $\text{kg ha}^{-1} \text{ day}^{-1}$), relying on the estimate of potential crop transpiration:

$$\frac{dAgB}{dt}_{POT} = \min\left(\frac{dAgB}{dt}_{IPAR}, \frac{dAgB}{dt}_{PT}\right) \quad [\text{Eq 6}]$$

$$\frac{dAgB}{dt}_{PT} = K_{BT} \times \frac{T_p}{VPD} \times 10^{-3} \quad [\text{Eq 7}]$$

$$T_p = ET_p \times (1 - e^{-k \times GAI}) \quad [\text{Eq 8}]$$

where RUE is the radiation use efficiency to produce AgB ($\text{kg dry matter MJ}^{-1}$), K_{BT} is a biomass-transpiration coefficient under given conditions of atmospheric vapour pressure deficit ($\text{Pa kg biomass (kg water)}^{-1}$), VPD is the vapour pressure deficit (kPa), T_p is the potential daily crop transpiration ($\text{kg water ha}^{-1} \text{ day}^{-1}$) and GAI is the green area index ($\text{m}^2 \text{ m}^{-2}$). CropSyst distinguishes LAI from GAI , in which just the green photosynthetically active fraction of the canopy is included.

WOFOST initiates crop growth from calculating the carbon assimilation (A_t , $\text{kg CH}_2\text{O ha(leaf)}^{-1} \text{ hr}^{-1}$):

$$A_t = A_{max} \times \left(1 - e^{-\epsilon \times \frac{IPAR}{A_{max}}}\right) \quad [\text{Eq 9}]$$

where A_{max} is the maximum daily gross assimilation ($\text{kg CH}_2\text{O ha(leaf)}^{-1} \text{ hr}^{-1}$), ϵ is the radiation use efficiency to produce assimilates in a single leaf ($\text{kg CH}_2\text{O ha(leaf)}^{-1} \text{ hr}^{-1} \text{ J}^{-1} \text{ m}^2 \text{ s}$) and $IPAR$ is here expressed in $\text{J m}^{-2} \text{ s}^{-1}$. The rate of assimilation by individual leaves is then transformed into the rate for the whole canopy using a three-point Gaussian integration method as a weighted average of the assimilate production rates at three levels within the canopy (Goudriaan, 1986). After extracting assimilates required for standing biomass maintenance (R_m , $\text{kg CH}_2\text{O ha(soil)}^{-1} \text{ hr}^{-1}$), the rate of AgB production is computed by multiplying the left assimilates by a conversion coefficient (C_e , $\text{kg biomass kg}^{-1} \text{ CH}_2\text{O}$):

$$\frac{dAgB}{dt} = 24 \times C_e (A_t \times LAI - R_m) \quad [\text{Eq 10}]$$

A_t is multiplied by LAI to convert the assimilation rate per ha of leaves into the rate per ha of soil and 24 transforms the hourly rate into a daily rate. C_e and R_m vary depending on the organ.

2.3. Canopy and crop development

Potential daily biomass production is affected by the temperature in all the models, as it is reliant on crop development (phenology). Phenology is simulated in all four models based on thermal time. CERES-Wheat, CropSyst and WOFOST include also the influence of day length. CERES-Wheat and CropSyst include the effect of vernalization. Crop development has a major influence on final yield, especially in CERES-Wheat and WOFOST in which biomass is partitioned to yield just after the grain filling period starts. Therefore, the length of the grain filling period (from anthesis to physiological maturity) has an enormous effect on final yield. Phenology can be accelerated by water and / or temperature stresses.

Potential daily biomass production is highly dependent on canopy development through LAI and CC , related to each other through the extinction coefficient (k) (Eq 11). CC is directly included in the computation of crop potential transpiration (Aquacrop) and intercepted radiation (CERES-Wheat, CropSyst and WOFOST).

$$CC = 1 - e^{-LAI \times k} \quad [\text{Eq 11}]$$

In Aquacrop, canopy development is calculated based on the actual canopy cover:

$$\begin{aligned} CC_t &= CC_o \times e^{CGC \times TT} & CC_t &\leq 0.5CC_{max} \\ CC_t &= CC_{max} - 0.25 \times \left(\frac{CC_{max}^2}{CC_o} \right) \times e^{-CGC \times TT} & CC_t &> 0.5CC_{max} \end{aligned} \quad [\text{Eq 12}]$$

and canopy decline:

$$CC_t = CC_{max} \times \left[1 - 0.005 \times \left(e^{\frac{CDC}{CC_{max}} \times TT} - 1 \right) \right] \quad [\text{Eq 13}]$$

where CC_o , CC_t and CC_{max} are the initial, actual and maximum canopy cover (fraction of soil covered by the crop), and CGC is the canopy relative growth rate ($^{\circ}\text{C-day}^{-1}$), CDC is the canopy decline rate (fraction soil covered $^{\circ}\text{C-day}^{-1}$), and TT is the accumulated thermal time ($^{\circ}\text{C-day}$).

The beginning of canopy decline is determined developmentally at a specified thermal time that is set through a crop parameter.

In CERES-Wheat, leaf area of a plant is calculated as the product of the rate of leaf appearance and the rate of expansion of growing leaves. In this model, only leaf blade area is considered and it is assumed that just one leaf at a time is expanding on the stem. LAI is calculated by multiplying the accumulated leaf area per plant and the plant density (PD , plants m^{-2}):

$$\frac{dLAI}{dt} = \left(\frac{dLeafAreaExpansion}{dt}_{plant} - \frac{dLeafAreaLoss}{dt}_{plant} \right) \times PD \quad [Eq\ 14]$$

where $\frac{dLeafAreaExpansion}{dt}_{plant}$ is the daily growth of leaf area per plant ($m^2\ plant^{-1}\ day^{-1}$), $\frac{dLeafAreaLoss}{dt}_{plant}$ is the daily senescence of leaf area ($m^2\ plant^{-1}\ day^{-1}$). The accumulation of leaf area per plant is sink-dependent, depending on the number of tillers and leaves:

$$\frac{dLeafAreaExpansion}{dt}_{plant} = \frac{dLeafAreaExpansion}{dt}_{MainStem} \times TN \quad [Eq\ 15]$$

where $\frac{dLeafAreaExpansion}{dt}_{MainStem}$ is the potential leaf area expansion of the main stem and TN is the number of tillers. The number of leaves is controlled by the thermal time accumulation and leaf appearance interval or phyllochron ($PHINT$, $^{\circ}C-day$). The first leaves of a single plant expand initially at a lower rate due to a low LAI and consequently are smaller:

$$\frac{dLeafAreaExpansion}{dt}_{MainStem} = \frac{dLeafAreaExpansion}{dt}_{ActualLeaf} \times \frac{DD}{PHINT} \quad [Eq\ 16]$$

$$\frac{dLeafAreaExpansion}{dt}_{ActualLeaf} = \frac{dLeafAreaExpansion}{dt}_{PreviousLeaf} \times (1 + \Delta LeafAreaExpansion) \quad [Eq\ 17]$$

where DD are the accumulated degree-days ($^{\circ}C-day$) and $\Delta LeafAreaExpansion$ is the increase of the potential leaf area of the new growing leaf with respect to the previous one (fraction). Actual daily leaf area growth is limited daily by the $\frac{dAgB}{dt}$ times a partitioning coefficient that depends on the development stage of the crop. The model assumes that each tiller can only support four green leaves. $\frac{dLeafAreaLoss}{dt}$ begins in the oldest leaf when the fourth leaf has been developed.

In CropSyst, the actual green area index as a function of time (GAI_t , $m^2 m^{-2}$) is computed from the actual leaf biomass production ($\frac{dLeaves}{dt}$, $kg m^{-2} d^{-1}$), the standing leaf biomass ($Leaves_t$, $kg m^{-2}$) and a partitioning coefficient of the biomass produced to leaves and stem ($LeafStemPart$, $m^2 kg^{-1}$):

$$\frac{dGAI}{dt} = \frac{\frac{dLeaves}{dt} \times SLA}{(LeafStemPart \times Leaves_t + 1)^2} \quad [Eq 18]$$

WOFOST computes LAI growth in two stages: (i) an earlier sink limited exponential growth dependent on the daily effective temperature (T_e) and a temperature dependent relative growth rate ($\frac{dLAI}{dt}_{growth,max}$, $(^\circ C - day)^{-1}$) and (ii) a source limited linear growth that starts once

$\frac{dLAI}{dt}_{growth,max}$ is reached:

$$\begin{cases} (i) & \frac{dLAI}{dt} = LAI_t \times \frac{dLAI}{dt}_{growth,max} \times T_e \\ (ii) & \frac{dLAI}{dt} = \frac{dLeaves_t}{dt} \times SLA \end{cases} \quad [Eq 19]$$

Canopy senescence can occur due to water stress ($\frac{dLeaves}{dt}_{WS}$), but also when LAI gets close to the maximum LAI ($\frac{dLeaves}{dt}_{HL}$):

$$\begin{cases} \frac{dLeaves}{dt}_{WS} = Leaves_t \times \left(1 - \frac{T_a}{T_p}\right) \times \frac{dLeaves}{dt}_{sen H_2O,max} \\ \frac{dLeaves}{dt}_{HL} = Leaves_t \times \frac{dLeaves}{dt}_{sen HL,max} \times \frac{LAI_t - LAI_{cr}}{LAI_{cr}} \end{cases} \quad [Eq 20]$$

where $\frac{dLeaves}{dt}_{sen H_2O,max}$ ($kg kg^{-1} day^{-1}$) and $\frac{dLeaves}{dt}_{sen HL,max}$ (day^{-1}) are the maximum relative death rate of leaves due to water stress ($1 - \frac{T_a}{T_p}$) and due to high LAI . The last one is assumed to be $0.03 day^{-1}$. LAI_t and LAI_{cr} are the actual and the critical LAI ($m^2 m^{-2}$). LAI_{cr} is the LAI required to intercept 95% of the incoming radiation and depends on the light extinction coefficient.

2.4. Biomass partitioning and yield formation

In CERES–Wheat, AgB and yield are sink and source limited. Sinks for vegetative biomass and yield are leaves and stems expansion and the number of grains, respectively. Biomass production is the source to fill those sinks and it is mainly driven by intercepted PAR. In WOFOST, the initial leaf area growth is sink-limited. Next to this phase, assimilates are distributed to different organs according to partitioning fractions that change with development stages following a source-limited approach. In Aquacrop and CropSyst, yield simulation depends on total AgB at physiological maturity and a harvest index. The unstressed harvest index is modified according to the water stress with a different intensity depending on the development stage at which the stress occurs.

2.5. Soil modules

Different models use alternative methods to calculate the water balance in the soil. They differ mainly in the calculation approach (cascade or Richards–Darcy), the complexity level with which the soil can be described in the profile, the soil hydraulic functions and the potential evapotranspiration equations used, and the computation of infiltration, soil evaporation, runoff, drainage and upflow (or capillary rise). Soil modules are often flexible as they offer alternative approaches at the user’s choice. Table 2 summarizes the equations used in the models and the characteristics of the soil modules as selected for this work. Most important equations defining the different approaches used in each of the studied models are described below.

Potential evapotranspiration

Reference evapotranspiration (ET_0) is calculated in the tested models from a variation of Penman’s equation (Penman, 1948), with the exception of Aquacrop, where it is provided as input data. CERES-Wheat and CropSyst offer two options (i) FAO56/ Penman–Monteith (Allen et al., 1998) and (ii) Priestley–Taylor (Priestley and Taylor, 1972). The first option was selected in both models. WOFOST calculates ET_0 based on Penman’s equation modified by Frère and Popov (Frère and Popov, 1979; Penman, 1956).

In low-humidity or more arid areas where the aerodynamic term of the Penman equation becomes more important than the energy term, ET_p is underestimated by Penman equation. In FAO56 method (as used in CERES-Wheat and CropSyst) this is corrected based on Doorenbos

and Pruitt, (1977). However, in WOFOST, this correction is included through the crop coefficient $CFET$ (mm mm^{-1}).

Infiltration, drainage and upflow

Soil water models can be categorized according to their complexity of process computations and how the soil profile is subdivided (Ranatunga et al., 2008). Most simple models follow a cascading approach: water is transferred from one layer to another from the surface towards the subsoil. Rainfall infiltrates or runs off. Infiltrated water is distributed immediately over the subsequent soil layers. Water content in a layer increases up to field capacity (FC) and when this level is reached, excess water is transferred to the next soil layer until all infiltrated water at that time step is distributed. Cascade-based models might have a single layer or multiple layers. More complex models are based on Richard's–Darcy. The soil profile is divided into elements separated by nodes, at which soil water potential and water content are defined.

In the cascade approach of CERES-Wheat, water exceeding field capacity in a soil layer infiltrates to the subsequent layer. This flow is limited by the soil water conductivity at saturation (K_{sat} , mm day^{-1}). Water flowing from the deeper layer is drained (Eq 21):

$$\frac{\partial \theta_t}{\partial t} = K_{sat}(\theta_t - \theta_{FC}) \quad [\text{Eq 21}]$$

where θ_t and θ_{FC} are the soil moisture content at time t and at field capacity ($\text{cm}^3 \text{cm}^{-3}$).

In Aquacrop, drainage is quantified according to Barrios-Gonzales, (1999), using an empirical drainage characteristic (τ , day^{-1}). The drainage characteristic (τ) is the fraction of drainable water ($\text{cm}^3 \text{cm}^{-3}$) lost from a fully saturated soil after one day of free drainage, in such a way that $1/\tau$ (days) is the time that it would take to deplete soil water from saturation to field capacity. As in Eq 22:

$$\frac{\partial \theta_i}{\partial t} = \tau(\theta_{SAT} - \theta_{FC}) \left[\left(\frac{e^{\theta_i - \theta_{FC}} - 1}{e^{\theta_{SAT} - \theta_{FC}} - 1} \right) \right] \quad [\text{Eq 22}]$$

$$\tau = 0.0866 \times K_{sat}^{0.35} \leq 1$$

where θ_i and θ_{SAT} are the soil moisture content at a specific layer i and at saturation ($\text{cm}^3 \text{cm}^{-3}$).

Table 2. Main characteristics of the soil modules as used in Aquacrop, CERES-Wheat, CropSyst and WOFOST.

	AQUACROP	CERES-Wheat	CROPSYST	WOFOST
Approach	Finite difference No upflow if no groundwater table	Cascade with upflow based on soil water diffusivity	Finite difference, based on Richard's equation	Cascade, no upflow if no groundwater table
Soil compartments	Subdivision with ΔZ increasing exponentially with depth, 12 nodes	User specified or subdivision with ΔZ increasing exponentially with depth	User specified (up to 20) or subdivision with ΔZ increasing exponentially with depth	One
Allows profile heterogeneity?	Yes	Yes	Yes	No
Potential ET	ET _o as input variable	$ET_o = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T + 273} (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_z)}$	$ET_o = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T + 273} (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_z)}$	$ET_o = \frac{\Delta}{\Delta + \gamma} R_{na} + \left(\frac{\gamma}{\Delta + \gamma}\right) E_{ac}$
Infiltration and drainage	$\frac{\partial \theta_t}{\partial t} = \tau(\theta_{SAT} - \theta_{FC}) \left[\left(\frac{e^{\theta_t - \theta_{FC}} - 1}{e^{SAT - \theta_{FC}} - 1} \right) \right]$ $\tau = 0.0866 \times K_{sat}^{0.35} \leq 1$	$\frac{\partial \theta_t}{\partial t} = K_{sat}(\theta_t - \theta_{FC})$	$\frac{\partial \theta_t}{\partial t} = K_{sat}(\theta_t - \theta_{FC})$	$\frac{\partial \theta_t}{\partial t} = K_{sat}(\theta_t - \theta_{FC})$ $\frac{\partial \theta_t}{\partial t} = P + I_e - E_w - \frac{SS_t}{\Delta t}$
Upflow	If a groundwater table is present: $\frac{\partial \theta_i}{\partial t} = e^{\frac{\ln(z)-b}{a}}$	$\frac{\partial \theta_i}{\partial t} = D(\theta) \times \frac{\theta_i - \theta_{i+1}}{(Z_i + Z_{i+1}) \times 0.5}$ $D(\theta) = \begin{cases} 0.88, \theta_i = LL \\ \max \left[100, 0.88 \times e^{0.35 \left(\frac{\theta_i}{2} + \frac{\theta_{i+1}}{2} \right)} \right], \theta_i > LL \end{cases}$	$\frac{\partial \theta_t}{\partial t} = \frac{\partial}{\partial t} \left[K(\theta) \left(\frac{\partial \phi}{\partial z} + 1 \right) \right]$	If a groundwater table is present: $\frac{\partial \theta}{\partial t}_{max} = \frac{W_{rz,fc} - W_r}{\Delta t}$ $W_{rz,fc} = RD \times \theta_{SAT} + \theta_{air}(z - RD) - \theta_{air}(z)$ $\theta_{air}(h) = \Delta \phi \frac{\theta_{air,-1} + 1.6 \theta_{air,0} + \theta_{air,1}}{3.6}$
Soil evaporation	$U_{limit} = 1000(\theta_{FC} - \theta_{air\ dry})Z_{evap}$ $\theta_{air\ dry} = 1/2 \theta_{PWP}$ Stage I: $E_a = (1 - CC^*)K_{ex}ET_o$ $K_{ex} = 1.1\text{ mm day}^{-1}$ Stage II: $E_a = K_r(1 - CC^*)K_{ex}ET_o$ $0 \leq K_r = \frac{e^{f_x W_{rel}} - 1}{e^{f_x} - 1} \leq 1$	U_{limit} is use specified $\theta_{air\ dry} = 1/2 \theta_{PWP}$ $E_{Eq} = R_n(0.00488 - 0.00437\alpha)(T_{mean} + 29)$ $E_p = \begin{cases} E_{Eq}[(T_{max} - 24) \times 0.05 + 1], T_{max} \geq 5^\circ C \\ E_{Eq} \times 0.01e^{0.18(T_{max} + 20)}, T_{max} < 5^\circ C \end{cases}$ $E_{soil} = \begin{cases} E_p(1 - 0.43LAI), LAI < 1 \\ E_{Eq} e^{-0.4LAI}, LAI \geq 1 \end{cases}$ Stage I: $E_a = E_{soil}$ Stage II: $E_a = E_{soil}t^{-0.5}$	$E_p = (1 - f_{Cres})(1 - f_{ccanopy})ET_p$ $\theta_{air\ dry} = 1/3 \theta_{PWP}$ Stage I: $E_a = E_p$ Stage II: $E_a = E_p \frac{\theta_t - \theta_{air\ dry}}{\theta_{FC} - \theta_{air\ dry}}$	$E_{p,bare} = ET_o e^{-k_{gb}LAI}$ $E_{w,max} = EO_w e^{-k_{gb}LAI}$ $E_{s,max} = EO_s e^{-k_{gb}LAI}$ $\begin{cases} E_a = E_{w,max} & \text{if } SS_t \geq 1\text{ cm} \\ E_a = E_{s,max} & \text{if } SS_t < 1\text{ cm and } \frac{\partial \theta_t}{\partial t} \geq 1\text{ cm day}^{-1} \\ E_a = E_{s,max}(\sqrt{D_{sitr}} - \sqrt{D_{sitr} - 1}) & \text{if } \frac{\partial \theta_t}{\partial t} < 1\text{ cm day}^{-1} \end{cases}$
Runoff	Curve number method	Curve number method	Curve number method	$Runoff = SS_t - \min(SS_t, SS_{max})$ $SS_t = SS_{t-1} + \left(P + I_e - E_w - \frac{\partial \theta_t}{\partial t} \right) \Delta t$

Δ : Slope of the saturation vapour pressure curve (kPa °C⁻¹)

R_{a} : net radiation at the crop surface (MJ m⁻² day⁻¹)

G : Soil heat flux density (MJ m⁻² day⁻¹)

γ : Psychometric constant at the sea level (kPa °C⁻¹)

T : Daily average temperature (°C)

$e_s - e_a$: vapour pressure deficit (kPa)

u_z : wind speed at z m

E_{ac} : Evaporative demand in equivalent evaporation (mm day⁻¹)

R_{net} : Net absorbed radiation in equivalent evaporation (mm day⁻¹)

$\theta_t, \theta_{\text{PWP}}, \theta_{\text{FC}}, \theta_{\text{SAT}}, \theta_{\text{airdry}}$: Soil moisture content at time t , at permanent wilting point, at field capacity, at saturation and at air dry (cm³ cm⁻³)

θ_{air}, h : Soil moisture content at time t , at selected matric head (h) (cm³ cm⁻³ kPa⁻¹)

$W_{\text{EZ}}, W_{\text{EZ}}$: Equilibrium and actual amount of water in the rooting zone (cm)

$\Delta\psi$: height of interval matric head (KPa)

τ : drainage characteristic (day⁻¹)

K_{sat} : Hydraulic conductivity at saturation (mm day⁻¹)

P : daily precipitation (mm day⁻¹)

I_e : effective irrigation (mm day⁻¹)

E_w : evaporation rate from a shaded water surface (mm day⁻¹)

a, b : parameters that depends on the soil layer textural class

z : ground water table depth (m)

SS_{max}, SS_t : maximum surface water storage and at time t (mm)

$D(\theta)$: water diffusivity (cm day⁻¹)

$EO_w, EO_s, E_{\text{p,bare}}$: potential evaporation rate from a shaded water surface, from a shaded soil surface and from a bare soil (mm day⁻¹)

$E_{w,\text{max}}, E_{s,\text{max}}$: maximum evaporation rate from a shaded water surface and from a shaded soil (mm day⁻¹)

Z_l, Z_{l+1} : soil depth of soil layer l and subsequent layer $l+1$ (m)

θ_l, θ_{l+1} : soil moisture content of soil layer l and subsequent layer $l+1$ (cm³ cm⁻³)

RD : actual rooting depth (m)

U_{limit} : maximum water to be evaporated in the stage I (mm)

Z_{evap} : upper soil layer depth (m)

E_a, E_p, E_{eq} : actual, potential and equilibrium soil evaporation rate (mm day⁻¹)

LAI : Leaf area index (m² m⁻²)

K_{ex} : maximum soil evaporation coefficient

CC^* : canopy cover adjusted by micro-advective effects (-)

f_s, W_{rel} : decline factor and relative wetness of the soil (-)

$f_{\text{Cress}}, f_{\text{Canopy}}$: fraction of incident radiation intercepted by the residues and by the canopy cover (0–1)

k_{gb} : extinction coefficient for solar radiation (-)

D_{dr} : days since last rain

K_r : relative water content of the upper soil layer

In the finite difference approach used in CropSyst, since water flow is based on differences in soil water potential, both downflow and upflow are instantaneously computed. Soil profile is divided into elements separated by nodes, at which soil water potential, water content, and root fractions are defined. This approach solves simultaneously for water transport in the soil and crop water uptake. Water content (θ , $\text{cm}^3 \text{ cm}^{-3}$) and the hydraulic conductivity are related to soil water potential at each node and the physical properties of the surrounding elements. The variation in volumetric water content ($\frac{\partial \theta}{\partial t}$, in $\text{cm}^3 \text{ cm}^{-3} \text{ day}^{-1}$) of a soil, both in time and space, is computed with the second order partial differential that is solved for each soil node by the Richards' equation (Eq 23):

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[K(\theta) \left(\frac{\partial \varphi}{\partial z} + 1 \right) \right] \quad [\text{Eq 23}]$$

where t is the time (days), z is the soil layer depth (m), $K(\theta)$ is the hydraulic conductivity as a function of soil water content (cm day^{-1}) and φ is the pressure head (m).

In WOFOST, soil is defined as a homogenous layer that increases at the same speed as rooting depth. Infiltrating water is automatically and homogeneously distributed in the whole soil profile, as in a cascade (Eq 1).

Soil evaporation

In Aquacrop, CERES-Wheat and CropSyst, actual soil evaporation (E_a) is simulated based on a two stage compartment. The first stage is the so-called energy limited stage where E_a is limited by the energy available at the soil surface and continues until a certain soil moisture content (U_{limit}) in the surface layer is reached (Ritchie, 1972). This threshold U_{limit} is determined in Aquacrop by the soil hydraulic characteristics and the depth of the surface layer (Philip, 1957) and can be modified by the user; in CERES-Wheat, U_{limit} is user specified and in CropSyst, U_{limit} is the permanent wilting point (θ_{pwp}) in the first soil layer (Table 2). In the second stage, each model computes E_a on different basis. In Aquacrop, E_a is computed in the second stage based on the relative water content of the upper layer (K_r) (Eq 24 and 25)

$$E_a = K_r (1 - CC^*) K_{ex} E T_o \quad [\text{Eq 42}]$$

$$0 \leq K_r = \frac{e^{fxW_{rel}} - 1}{e^{fx} - 1} \leq 1 \quad [\text{Eq 25}]$$

were K_{ex} is the maximum soil evaporation coefficient (-). In CERES- Wheat, E_a is proportional to the time by stage 2 started (eq 26).

$$E_a = E_{soil} t^{-0.5} \quad [\text{Eq 26}]$$

were E_{soil} is the potential soil evaporation (mm) (see Table 2 for equation). Lastly, in CropSyst, E_a in the stage 2 is computed as in eq 27:

$$E_a = E_p \frac{\theta_i - \theta_{air\ dry}}{\theta_{FC} - \theta_{air\ dry}} \quad [\text{Eq 27}]$$

were θ_{airdry} is the soil moisture content at air dry ($\text{cm}^3 \text{cm}^{-3}$).

In WOFOST, E_a depends on water surface storage and the infiltration rate

and the soil cover (LAI and extinction coefficient) (Eq 28 to 31).

$$E_{p,bare} = ET_o e^{-k_{gb} LAI} \quad [\text{Eq 28}]$$

$$\begin{cases} E_a = E_{w,max} & \text{if } SS_t \geq 1 \text{ cm} \\ E_a = E_{s,max} & \text{if } SS_t < 1 \text{ cm and } \frac{\partial \theta_i}{\partial t} \geq 1 \text{ cm day}^{-1} \\ E_a = E_{s,max} (\sqrt{D_{slr}} - \sqrt{D_{slr} - 1}) & \text{if } \frac{\partial \theta_i}{\partial t} < 1 \text{ cm day}^{-1} \end{cases} \quad [\text{Eq 29}]$$

$$E_{w,max} = EO_w e^{-k_{gb} LAI} \quad [\text{Eq 30}]$$

$$E_{s,max} = EO_s e^{-k_{gb} LAI} \quad [\text{Eq 31}]$$

were $E_{p,bare}$, EO_w and EO_s are the potential evaporation rate from a bare soil, from a shaded water surface and from a shaded soil surface (mm day^{-1}), k_{gb} is the extinction coefficient for global radiation (-), $E_{w,max}$ and $E_{s,max}$ are the maximum evaporation rate from a shaded water surface and from a shaded soil are the, SS_t is the surface water storage and at time t (mm) and D_{slr} are the days since last rain.

Runoff

In Aquacrop, CERES–Wheat, and CropSyst, runoff is calculated after a rainfall or irrigation event based on the USDA–SCS curve number approach (USDA, 2004):

$$RO = \begin{cases} 0, & P \leq 0.2S \\ \frac{(P-I_a)^2}{P-I_a+S}, & P > 0.2S \end{cases} \quad [\text{Eq 32}]$$

$$S = \frac{1000}{CN} - 10 \quad [\text{Eq 33}]$$

where RO is the runoff (mm day^{-1}), P is rainfall, I_a is water infiltrated or intercepted by the crop previous to runoff, S is the potential maximum soil moisture retention after runoff begins (mm day^{-1}) and CN is the empirical curve number value (mm^{-1}).

In WOFOST, runoff is based on the maximum surface storage ($SS_{\max}$) of the soil. $SS_{\max}$ is user specified (Goudriaan, 1977).

Tillage

Tillage is solely explicitly included in CERES-Wheat and CropSyst. Tillage is simulated in CERES-Wheat by adjusting the percentage of surface residues incorporated into the soil, the percentage of soil surface disturbed, the efficiency in soil mixing and the percentage of reduction in the hard pan (White et al., 2009). This computes changes in the bulk density, in the ease of roots to penetrate each of the soil layers or in K_{sat} . CropSyst simulates tillage effects through parameters such as the fraction of soil covered by residues, the fraction of fast cycling, slow cycling and lignified biomass and other parameters related to soil decomposition timing (Sommer et al., 2007). Tillage effects are simulated by changes in runoff, erosion, soil evaporation and the amount of water infiltrating into the soil.

2.6. Computation of water stress impact on biomass production

Besides the limitation to biomass growth determined by the potential transpiration of the crop, defined using the WUE coefficient as in the case of CropSyst and Aquacrop, biomass and yields are also reduced when they grow under water deficient conditions. This situation occurs when actual transpiration is lower than the atmospheric demand of water or potential transpiration due to low, easily available soil moisture. Actual transpiration is deemed to be equivalent to water uptake. Crop water uptake is driven by the differences in water potential in the soil and the roots. Aquacrop, CERES-Wheat and WOFOST simplify the process of computing water uptake directly from the easily available water (above the permanent wilting point). CropSyst calculates

water potential in the soil based on Campbell (1985). The effect of this water stress is translated into a reduction of biomass growth and yields in each of the studied models following different approaches. Water stress also affects other crop processes such as canopy expansion and partitioning of assimilated carbon. The effects of water stress effects are computed in each of the models as follows.

In Aquacrop, soil water stress is translated into biomass production which affects the expansion of the canopy cover ($K_{s,exp}$), induces both stomata closure ($K_{s,sto}$) and an early canopy senescence ($K_{s,sen}$) and reduces pollination ($K_{s,pol}$) (Raes et al., 2012).

Daily water-limited aboveground biomass production ($\frac{dAgB}{dt}_{adj}$) is affected by stomata closure ($K_{s,sto}$) reducing actual transpiration (Eq 34):

$$\frac{dAgB}{dt}_{adj} = WP \times \left(\frac{Ta}{ET_o} \right) \times K_{s,sto} \quad [\text{Eq 34}]$$

Water-limited canopy growth rate and canopy senescence (CGC_{adj} and CDC_{adj}) are affected by $K_{s,exp}$ and $K_{s,sen}$, respectively (Eq 35 and 36).

$$CGC_{adj} = CGC \times K_{s,exp} \quad [\text{Eq 35}]$$

$$CDC_{adj} = (1 - K_{s,sen}^8) \times CDC \quad [\text{Eq 36}]$$

Harvest index is affected by $K_{s,pol}$ for water deficit occurring during flowering (Eq 37), and by $K_{s,sto}$ when occurring after flowering (Eq 38).

$$HI_{adj} = HI_0 \times K_{s,pol} \quad [\text{Eq 37}]$$

$$HI_{adj} = HI_0 \times \sqrt[10]{K_{s,sto}} \times \left(1 - \frac{1-K_{s,sto}}{b} \right) \quad [\text{Eq 38}]$$

where b is a crop parameter that has a lower value as stronger the impact of water deficit on HI is.

The effect of water stress computed in the model is based on soil water content:

$$K_{s,x} = 1 - \frac{e^{S_{rel} \times f_{shape} - 1}}{e^{f_{shape} - 1}} \quad \text{for } K_{s,exp}, K_{s,sto} \text{ and } K_{s,sen} \quad [\text{Eq 39}]$$

$$K_s = 1 - S_{rel} \quad \text{for } K_{s,pol} \quad [\text{Eq 40}]$$

$$S_{rel} = \frac{\theta_i - \theta_{lower}}{\theta_{upper} - \theta_{lower}} \quad \text{for } \theta_{lower} \leq \theta_i \leq \theta_{upper} \quad [\text{Eq 41}]$$

$$\theta_{lower} = \theta_{FC} - (1 - p_{lower}) \times \theta_{WP} \quad [\text{Eq 42}]$$

$$\theta_{upper} = \theta_{FC} - (1 - p_{upper}) \times \theta_{WP} \quad [\text{Eq 43}]$$

where S_{rel} is the relative stress level, f_{shape} is a factor determining the convexity of the relation between water stress and available soil moisture; p_{upper} is the fraction of total available water, between wilting point (θ_{WP} , $\text{cm}^3 \text{ cm}^{-3}$) and field capacity (θ_{FC} , $\text{cm}^3 \text{ cm}^{-3}$), when water stress starts to affect plant growth and p_{lower} when water stress is at full strength. θ_{lower} equals θ_{WP} for $K_{s,sto}$ and $K_{s,sen}$. When there is no water stress, K_s is set at 1. Otherwise, K_s falls below 1 thus affecting crop transpiration, canopy growth or senescence rate and harvest index.

Water deficit in CERES-Wheat affects the rate of biomass production and canopy senescence (Ritchie, 1998). When potential water uptake decreases to a value lower than the potential transpiration rate, transpiration is reduced by a partial closure of the stomata and the potential biomass production is reduced by the same proportion as the actual to potential transpiration ($SWDF1$) (Eq 45). Leaf expansion, branching and tillering are physiological processes especially sensitive to the decrease in turgor pressure due to water deficit, even before stomata closure begins ($SWDF2$). In this sub-routine, turgor pressure is not directly calculated, but it is assumed that the effect of reduced turgor pressure on leaf expansion, branching and tillering begins when the root uptake (T_a) is lower than the 1.5 times the potential transpiration (T_p) (Eq 46). In CERES-Wheat biomass formation is sink limited, this has a double effect on final yields. Root water absorption is calculated based on the law of the minimum, where the flow rate from soil to roots is dominated by either soil resistance, root resistance or the atmospheric demand of water.

$$U_{CERES,i} = \min \left(q, \frac{2.67 \cdot 10^{-3} \times e^{62 \times (\theta_t - \theta_{LL})}}{6.68 - \ln(RD_i)} \right) \quad [\text{Eq 44}]$$

$$SWDF1 = \frac{T_a}{T_p} \quad [\text{Eq 45}]$$

$$SWDF2 = \frac{T_a}{1.5 \times T_p} \quad [\text{Eq 46}]$$

where $U_{CERES,i}$ is the daily root water absorption from soil layer i ($\text{cm}^3 \text{ cm}^{-1} \text{ root day}^{-1}$), θ_t and θ_{LL} are the soil moisture contents at time t and the lower limit ($\text{cm}^3 \text{ cm}^{-3}$), and RD_i is the root length density in soil layer i ($\text{cm cm}^{-3} \text{ soil}$). Root water absorption is restricted to a maximum value (q) of

$0.03 \text{ cm}^3 \text{ cm}^{-1} \text{ root day}^{-1}$, and if the total water uptake in the soil is higher than the atmospheric demand of water, q is restricted in every soil layer to equal the root water uptake to potential evapotranspiration. Water deficit during the vegetative growth of the crop promotes a shift in partitioning towards roots. This compensates water shortage with a more exhaustive exploration of the soil.

In CropSyst, biomass production rate is reduced in the same proportion as the ratio of actual to potential transpiration. Potential transpiration is calculated as a fraction of the atmospheric demand of water (ET_p) that depends on canopy cover or green area index:

$$T_p = ET_p \times e^{-k \times GAI} \quad [\text{Eq 47}]$$

where k is a canopy extinction coefficient for solar radiation and GAI is the green area index. Actual transpiration is equated to water uptake assuming no water storage in the plant. It is calculated as described in Stöckle and Jara (1998):

$$U_{CropSyst,i} = K_t \times \frac{C_{ri}}{1.5} \times (\varphi_{si} - \varphi_l) \quad [\text{Eq 48}]$$

$$C_{ri} = f_1 \times f_{cc} \times C_T \quad [\text{Eq 49}]$$

$$C_T = \frac{1.5 \times WU_{max}}{\varphi_{FC} - \varphi_{sto} \times K_t} \quad [\text{Eq 50}]$$

where $U_{CropSyst,i}$ is the water uptake for soil layer i ($\text{kg m}^{-2} \text{ day}^{-1}$), K_t is the unit conversion constant (seconds in a day), C_{ri} is the root conductance for the layer i (kg s m^{-4}), φ_l and φ_{si} are the soil layer and canopy average leaf water potential (J kg^{-1} equivalent to $\text{m}^2 \text{ s}^{-2}$), f_1 is the fraction of total root length in layer i (cm cm^{-1}), f_{cc} is the fraction of soil covered by the green canopy ($\text{m}^2 \text{ m}^{-2}$), WU_{max} is the maximum water uptake (mm day^{-1}); and φ_{FC} and φ_{sto} are the soil water potential at field capacity and the leaf water potential at the onset of stomata closure, respectively. Therefore, the larger the difference between water potential in the soil layer and the plant, the larger the water uptake. The total water uptake is the sum of the U_i from each layer. Water stress also affects crop development and the duration of the green area index in CropSyst. It accelerates the accumulation of degree-days under water stress conditions, as it is assumed that water deficit decreases crop transpiration. A decrease in crop transpiration results in an increase of the temperature inside the canopy resulting in an increase in the effect of the daily accumulation of degree-days.

In WOFOST, water stress reduces gross carbohydrates assimilation ($\text{kg CH}_2\text{O ha}^{-1} \text{ day}^{-1}$) in the same proportion as actual to potential transpiration. Actual transpiration depends on water uptake. Water uptake by the roots is computed taking into account that the resistance to moisture transport increases when soil moisture potential is high (close to wilting point) and decreases when low (close to or at field capacity). In the latter situation, a reduction of transpiration (R_{ws}) is assumed due to water stress. In WOFOST, this reduction in transpiration is calculated as:

$$R_{ws} = \frac{\theta_t - \theta_{WP}}{\theta_{WS} - \theta_{WP}} \quad \text{for } \theta_{WP} \leq \theta_t \leq \theta_{WS} \quad [\text{Eq 51}]$$

where θ_t and θ_{ws} are the soil moisture content at time t and a critical soil moisture ($\text{cm}^3 \text{ cm}^{-3}$), respectively. The last parameter is difficult to measure in trials, and in the model it is calculated as:

$$\theta_{ws} = (1 - p) \times (\theta_{FC} - \theta_{WP}) + \theta_{WP} \quad [\text{Eq 52}]$$

where p is the soil water depletion factor ($\text{cm}^3 \text{ cm}^{-3}$) (similar to easily available water concept in literature) that is a function of potential evapotranspiration rate for a closed canopy and a crop group number (*DEPNR*). This crop group number is a biological factor that determines the relationship between p and potential evapotranspiration (T_p). A high p -factor would (in the limit, maximum 1) give that $\theta_{ws} = \theta_{WP}$, while a low factor would give $\theta_{ws} = \theta_{FC}$. Assuming a homogenous soil with θ_{WP} at $0.095 \text{ cm}^3 \text{ cm}^{-3}$ and θ_{FC} at $0.235 \text{ cm}^3 \text{ cm}^{-3}$ (Table 3) and 6 mm of daily potential transpiration (T_p), the θ_{ws} for each group number can be calculated. Then, a drought sensitive crop belonging to *DEPNR* 1 (e.g. leaf vegetables), p equals 0.23 and θ_{ws} , 0.203; a crop with a medium drought sensitivity belongs to *DEPNR* 3 (e.g. potatoes), p equals 0.40 and θ_{ws} , 0.179; and a drought resistant crop belonging to *DEPNR* 5 (e.g. olive), p equals 0.60 and θ_{ws} , 0.151.

Another parameter indirectly controls the effect of water impact, previously described crop coefficient (*CFET*). *CFET* (mm mm^{-1}) multiplies the Penman evapotranspiration in such a way that values lower than 1 decrease ET_p , increasing the ratio ET_a/ET_p and decreasing the negative impact of water deficit, and it was intended to adapt the reference *ET* from reference grass to tall crops (Supit et al., 1994).

Table 3. Soil texture and hydraulic characteristics for soil defined with a heterogenous (per horizon) and a homogenous profile (as for WOFOST model).

<i>Texture definition</i>	<i>Horizons</i>	<i>Depth (cm)</i>	<i>Texture (%)</i>			<i>Moisture content (cm cm⁻³) at pF</i>		<i>Bulk density (g cm⁻³)</i>
			<i>Sand</i>	<i>Silt</i>	<i>Clay</i>	<i>(-1500kPa)</i>	<i>(-33kPa)</i>	
<i>per horizon</i>	<i>Ap 1</i>	0-3	21.3	57.6	21.1	0.080	0.240	1.46
	<i>Ap 2</i>	3-25	21.3	57.6	21.1	0.080	0.240	1.46
	<i>Bwy1</i>	25-84	13.8	60.6	25.6	0.100	0.230	1.45
	<i>Bwy2</i>	84-90	9.5	63.3	27.2	0.110	0.270	1.42
<i>Homogenous</i>	–	0-90	16.2	57.9	24.1	0.095	0.235	–

2.7. Experimental site and data

Observed data was obtained from a long term crop rotation experiment located in Agramunt, Lleida (41° 48'N, 1° 7' E, 330 m altitude) in the western area of Catalonia, Spain. The rotation included wheat and barley. Data from seasons in which wheat was grown, 2000-01, 2002-03, 2004-05, 2006-07 and 2011-12, was used. The climodiagram in Figure 1 depicts the most representative weather station; El Canós (41°41'28" N, 1°12'13" E, 429 m). Soil was Typic Xerofluvent (Soil Survey Staff, 1994), silty loam, 0.9 m deep (Table 3), with no groundwater impact on the rooting zone. Soil hydraulic characteristics were obtained in the laboratory using the Richards pressure membrane apparatus on disturbed samples.

The experiment was set up in 1990 and originally designed to evaluate the agronomic responses of the crops (growth, yield and its components, water productivity and nitrogen use efficiency) to different management systems based in the reduction of tillage (Cantero-Martínez et al., 2007). Tillage treatments consisted of: (i) mouldboard, intensive ploughing to a depth of 0.3 m, (ii) subsoiling to a depth of 0.5 m, (iii) minimum tillage consisted of a shallow ploughing with a cultivator pass to a depth of 0.15 m, and (iv) no tillage and direct drilling.

Cropping operations, apart from tillage, were similar in all areas. Crops were sown at planting densiting ranging between 400 and 500 seeds m² in 17 to 19 cm spaced rows. Fertilisation was performed according to crop requirements and soil tests. Pig slurry (15–20 m³ ha⁻²) was spread before sowing. N mineral fertiliser (50–60 kg N ha⁻²) was applied on the soil surface before tillering using ammonium nitrate (33.5%). After emergence and before tillering, grass and

broadleaf weeds were controlled by herbicide. No treatments were required for pests or diseases.

Only data for no tillage and minimum tillage treatments was selected for this work, as they appeared to be the most effective treatment in increasing yield under the driest conditions in the long term. Data on cultivars, sowing, and phenology, yield and aboveground biomass are shown in Table S1.

Figure 1. Climodiagram in El Canós (41°41'28" N, 1°12'18" E, 429 m). Mean monthly data between 2000 and 2013.

The experiment was a randomised block design, with 180 m² plot size (30 x 6) and treatments were repeated in three blocks. Plant data were collected in each plot at four phenological stages in the first season (2000-01): tillering, jointing, anthesis and physiological maturity; at three phenological stages in the second season (2002-03): jointing, anthesis and physiological maturity; and lastly, at two phenological stages in the remaining seasons (2004-05, 2006-07 and 2011-12):

anthesis and physiological maturity. Aboveground biomass was collected by removing all plants from three randomly selected samples of 0.5 m² in each plot. Samples were dried to constant weight in a forced-draft oven at 65 °C to obtain dry matter biomass (Cantero-Martínez et al., 2007). Soil water content was measured at four depths (0–25, 25–50, 50–75 and 75–90 cm) in the first four seasons and at three depths (0–33, 33–66 and 66–90 cm) in the last season before sowing, at mid-season and after harvest by the gravimetric method (Campbell and Mulla, 1990) by drying to constant weight in a forced-draft oven at 105 °C.

Treatments were repeated in three blocks. Soil water content measured before sowing was different in different blocks (Table S2) and therefore, they were modelled separately to assess the sensitivity of the four models to initial soil water content.

2.8. Models parameterization, calibration and comparison

The parameterization of the models was completed through an iterative process using the observed dates, measured crop growth variables, and default parameters available in the models' crop files. Degree days for different phenological stages were estimated from the base temperature, cutoff temperature and daily mean temperature. Other crop parameters were changed manually around the default parameter values and within a biologically logical range so that simulated yields fitted closer to field observations.

In the first step, parameters related to crop production with neither water or nutrient limitations nor pest and diseases were estimated. As no data to validate the simulated potential yields was available, parameters were selected in such a way that radiation use efficiency and daily potential biomass production were within a physiologically plausible range, and also comparable between models. To compare the models CERES-Wheat, CropSyst and WOFOST in terms of *RUE*, model equations were reproduced in a calculation sheet for a well-established crop during the vegetative growth (no grains yet); 3.5 t ha⁻¹ of AgB and a *LAI* of 3.4 were used as inputs. Aquacrop could not be compared in terms of *RUE*, as this model does not base biomass production on radiation but solely on the efficiency of transpired water.

In the second step, parameters for water limited production were calibrated for the trial in season 2000–01 and validated for trials in the remaining seasons. Plots treated with no tillage and with minimum tillage were not differentiated in terms of soil management in the simulation setting as

no data on soil changes after tillage was available. Models performance was evaluated using the root mean square error (*RMSE*). This statistic measures the distance between the simulated and the observed values in absolute terms (Eq 37). Larger values of *RMSE* indicate decreasing accuracy.

$$RMSE = \sqrt{\frac{\sum_1^n (S_i - O_i)^2}{n}} \quad [\text{Eq 37}]$$

The impact of water stress on most important crop processes was plotted against soil water availability. The effect on daily aboveground biomass production in each of the models was compared assuming a daily potential evapotranspiration of 6 mm and a daily biomass production of 0.2 t ha⁻¹. The effect on canopy growth and decline rates, and biomass partitioning into grains (harvest index) could only be plotted for Aquacrop no impact on these two processes is considered in the remaining models .

Weather and soil data was adapted to each of the model formats. Some weather variables coincide (*T*_{max}, *T*_{min} and Rainfall) and others are not demanded in all the models (net radiation, wind speed, *ET*_o and vapour pressure). Soil moisture content measured at -1500 kPa and at -33 kPa were used in the models as the field capacity and the permanent wilting point. The Penmann–Monteith equation was selected for *ET*_o calculation in CERES-Wheat and CropSyst, and the cascade approach for soil water dynamics was selected in CropSyst (no alternatives for Richard-Darcy approach is offered in the remaining models). In CERES-Wheat and CropSyst, soil layers were defined as the horizons shown in Table 3. For potential yield, the simulation option to simulate with no water limitation was used in CERES-Wheat and WOFOST; while an irrigation treatment to keep soil moisture at field capacity was used in Aquacrop and CropSyst.

3. Results and discussion

3.1. Models parameters and approaches

The different approaches used in the models, both in the soil and in the crop modules, make their comparison arduous in terms of soil hydraulic functions and crop parameter values. Models can be compared by (i) isolating computational approaches of a specific crop process; this enables identifying the origin of the different model's performance; or by (ii) comparing outputs, giving an idea on the results that can be expected by the model user. First, model's parameters

calibrated for potential production are compared in terms of radiation use efficiency and water productivity, and discussed. In the second step, soil water content ($\text{cm}^3 \text{ cm}^{-3}$) was simulated by each of the four models for a bare soil. As with crop potential production, no data to validate the simulated soil water content with no crop growing on it was available, but it helped in understanding the differences in yield and biomass simulation between models when simulated under water limiting conditions. Soil parameters were the same in each of the models, but for WOFOST in which *FC* and *PWP* were averaged as this model assumes a homogenous soil profile (Table 3). Initial soil water conditions were set at the 50% of the plant available water (average between *PWP* and *FC*). Lastly, the relationships defining the impact of water stress on a well-established crop, depending on available soil moisture ($\text{cm}^3 \text{ cm}^{-3}$) for the four models, are compared. Potential production simulated by the calibrated models, graphs of simulated water-limited versus observed production and soil water contents in time are also presented.

Potential crop biomass production

Crop parameters estimated for potential daily biomass production, canopy development and yield formation in all four studied models are presented in Table 4.

WOFOST is the most detailed model in terms of physiological processes integration as it simulates assimilate production, maintenance requirements of the standing biomass and respiration in the biomass production. All these processes are integrated in CERES-Wheat and CropSyst through a radiation use efficiency parameter.

In WOFOST, the light use efficiency to produce assimilates in a single leaf was set at $0.45 \text{ kg CH}_2\text{O ha (leaf)}^{-1} \text{ hr}^{-1} \text{ J}^{-1} \text{ m}^2 \text{ s}$ (Table 4). Assimilate production saturates at a maximum rate of assimilates of $40 \text{ kg CH}_2\text{O ha (leaf)}^{-1} \text{ hr}^{-1}$ (Table 4) (within the range proposed by Boons-Prins et al., 1993) and corresponding to about $59 \text{ kg CO}_2 \text{ ha (leaf)}^{-1} \text{ hr}^{-1}$ (considering a molecular weight of 30 for CH_2O and 44 for CO_2). These values fit closely to the experimentally determined photosynthesis functions shown in van Keulen and Seligman (1987).

The light extinction coefficient was 0.6 for CropSyst and WOFOST, and differed from the 0.85 set in CERES-Wheat (Table 4). This higher value is due to the fact that leaf area in CERES-Wheat excludes the sheath of the leaves, as authors assume that sheath compared to blade area exposed to light is a constant (Godwin et al., 1989). In CropSyst, the parameter light use efficiency was set at $2.4 \text{ g MJ (PAR)}^{-1}$, close to the default value of $2.3 \text{ g MJ (PAR)}^{-1}$ set at CERES-Wheat (Table 4). Sinclair and Muchow, (1999) reviewed the concept of RUE and

calculated it from numerous experiments in the literature for several crops. For wheat, intercepted PAR use efficiency ranged from 1.46 to 2.93 g MJ (PAR)⁻¹. The parameter K_{BT} in CropSyst corresponds to the transpiration use efficiency of a crop at a given daily vapour pressure deficit. A large variability for this parameter has been reported in the literature for wheat, ranging from 2.8 to 6.7 Pa (Kemanian et al., 2005). In this work, K_{BT} was set within the reported range, at 3.3 Pa (Table 4).

Potential daily AgB production of a well-established crop against the soil fraction covered by the canopy (Aquacrop) or the photosynthetically active radiation (CERES-Wheat, CropSyst and WOFOST) are contrasted in Figure 2. Total radiation use efficiency was compared for a well-established crop (3.5 Mg ha⁻¹ of AgB, 1.5 Mg ha⁻¹ of leaves and 0.4 Mg ha⁻¹ of roots, and a LAI of 3.4) in CERES-Wheat, CropSyst and WOFOST in Figure 2B. RUE was calculated using the linear part of the curve (and in CropSyst when AgB is solely *PAR* limited, Eq 6). It resulted in different values for the three models, being higher in CERES-Wheat, followed by CropSyst and lastly WOFOST with 2.51, 2.40 and 2.36 g AgB MJ⁻¹ of *PAR*, respectively (Figure 2B). Taken into account that roots were about 10% of the total biomass, the radiation use efficiency calculated for total biomass results in 2.79, 2.67 and 2.63 g total biomass MJ⁻¹ *PAR*. These results are within the range given by Sinclair and Muchow (1999) and described above. In CERES-Wheat, daily biomass production does not saturate at a certain *PAR* (as in CropSyst and WOFOST); biomass growth is sink limited and therefore controlled by a potential daily growth depending on the number of tillers, their developmental stages and the potential tiller weight (parameter G1 in Table 4). The CropSyst equation to compute AgB production from *IPAR* [Eq 5] does not saturate at high light levels; nevertheless, the equation that computes AgB production from T_p [Eq 7] forces biomass growth to saturate at a maximum rate. This maximum rate fluctuates at each time step as it includes the T_p and *VPD* that vary daily. In Figure 2B, this variation is shown in the curves for 0.5 and 1.0 kPa of *VPD*. Aquacrop uses a very different approach and therefore it was not possible to compare this model in terms of *RUE*, just the highest rate of daily biomass production in Aquacrop (180 kg dry matter ha⁻¹ day⁻¹ for a closed canopy) that is lower than the value in WOFOST for high *PAR* (200-250 kg dry matter ha⁻¹ day⁻¹ for a closed canopy and a *PAR* over 10 MJ m⁻² day⁻¹) (Figure 2). The highest rate of daily biomass production in CropSyst is defined by the vapor pressure deficit, for high *PAR*.

Table 4. Crop parameters in Aquacrop, CERES-Wheat, CropSyst and WOFOST crop files.

		AQUACROP	CERES-Wheat	CROPSYST	WOFOST
Biomass production	Extinction coefficient (k, –)	–	0.85	0.6	0.6
	Radiation use efficiency to produce aboveground dry matter (RUE, g MJ ⁻¹)	–	2.3	2.4	–
	ϵ , Radiation use efficiency to produce assimilates in a single leaf (kg ha ⁻¹ hr ⁻¹ J ⁻¹ m ² s)	–	–	–	0.45
	Maximum assimilation rate (kg CH ₂ O ha(leaf) ⁻¹ hr ⁻¹)	–	–	–	40.00
	Aboveground biomass production per meter of transpired water under given	–	–	3.3	–
	Water productivity (WP, g m ⁻²)	18	–	–	–
Crop development	LAI _o ($CC_o = 1 - e^{-0.6LAI_o}$)	0.006 (0.4 %)	0.4 (21.33 %)	0.011 (0.66 %)	0.1365 (7.9 %)
	LAI _{max} ($CC_{max} = 1 - e^{-0.6LAI_{max}}$)	3.16 (85 %)	3.50 (88 %)	–	–
	SLA at optimum temperature (m ² kg ⁻¹)	–	11.5	22.5	22.5
	GDD ¹ end canopy growth (°C–days)	1216	–	1205	–
	GDD ¹ begin senescence (°C–days)	1500	–	1500	–
	Days of optimum vernalizing temperature required for vernalization (P1V)	–	60	–	–
	Photoperiod response (% reduction in rate/10 h drop in pp) (P1D)	–	150	–	–
	GDD ¹ sowing–emergence (°C–days)	120	–	120	90
	GDD ¹ emergence–flowering (°C–days)	1200	–	1200	1200
	GDD ¹ anthesis–maturity (°C–days)	460	350	460	490
Yield formation	Kernel number per unit aboveground biomass weight at anthesis (#/g) (G1)	–	19.50	–	–
	Standard kernel weight under optimum conditions (mg) (G2)	–	30.50	–	–
	Standard, non-stressed mature tiller weight (including grain) (g dwt) (G3)	–	2.59	–	–
	DVS1	–	–	–	0.99
	Harvest index	0.45	–	0.5	–

¹ Thermal time is calculated for a base and cutoff temperature of 0°C and 26°C, respectively.

In all four models, canopy development has a crucial role in potential biomass production via transpired water in Aquacrop and intercepted radiation in CERES-Wheat, CropSyst and WOFOST (Table 1). Models were quite sensitive to initial conditions and just CC_0 in Aquacrop was modified (to 0.4%) (Table 4). LAI_0 differed between the remaining models. The lowest value corresponded to CropSyst (0.011), followed by Aquacrop (0.017), WOFOST (0.1367) and lastly by CERES-Wheat (0.4) (Table 4). In CERES-Wheat, canopy growth is sink-limited by the number of tillers and a maximum of four leaves per tiller and a LAI_{max} set internally in the model at 3.5. SLA was comparable in CropSyst and WOFOST, $22.5 \text{ m}^2 \text{ kg}^{-1}$, and it was very low in CERES, $11.5 \text{ m}^2 \text{ kg}^{-1}$, as it just refers to the blade of the leaves (Table 4). In Aquacrop, SLA is not included, but canopy development is driven by thermal time, based on a canopy growth rate and a canopy decline rate in % per $^{\circ}\text{C}$ -day up to CC_{max} , in this case 85% (Table 4).

Thermal time was estimated from phenology data, using sowing, anthesis and physiological maturity dates. In all models, base temperature was 0°C and cut-off temperature 26°C . Thermal time between sowing and emergence (120, 120 and 90°C -days in Aquacrop, CropSyst and WOFOST), between emergence and anthesis (1200°C -days in Aquacrop, CropSyst and WOFOST) and between anthesis and physiological maturity (460°C -days) was similar for all four models (Table 4); differences in WOFOST are related to the base temperature to compute thermal time accumulation before emergence. In CERES-Wheat, the time from sowing to anthesis is controlled by the variables $P1V$ (vernalization) and $P1D$ (photoperiod response). Despite thermal time was similar in all the models, anthesis and physiological maturity are reached in Aquacrop 12 and 11 days before expected (Table 5). A good fit of simulated to observed dates was obtained for 1309 and 477°C -days from emergence to anthesis and anthesis to maturity, respectively (not shown data). These values are substantially lower to those given by Boons-Prins et al. (1993) for Southern Europe, 1350 and 1000°C -days; especially the grain filling period.

Yield formation was computed also following alternative approaches. In Aquacrop and CropSyst, yield formation is based on a harvest index that was set to 0.45 and 0.5, respectively (Table 4). In CERES-Wheat, yield formation is sink-limited and determined by the weight of unstressed yield components. Kernel number at anthesis ($G1$) was 19.50 per g of canopy; the standard kernel weight under optimum conditions ($G2$) was 30.50 mg, and the non-stressed mature tiller weight including grain ($G3$), 2.59 g (Table 4). In WOFOST, the parameter $DVS1$ refers to the development stage at which assimilates start being accumulated to grain biomass. It was set to

0.99 (Table 4), just before flowering (at *DVS1*) to mimic translocation of assimilates from the vegetative biomass to the grains, a process that was not explicitly included in the model.

Models simulated similar yields under no limiting conditions, to 6 Mg ha⁻¹ (Table 5). Results are given as average yields for the calibration and the validation sets, respectively.

Figure 2. Aboveground growth rate of biomass (kg ha⁻¹ day⁻¹) against (A) soil covered by the canopy (%) and (B) Photosynthetically active radiation (MJ m⁻² day⁻¹) for a well-established crop (3.5 Mg ha⁻¹ of aboveground biomass, 1.5 Mg ha⁻¹ of leaves, 0.4 Mg ha⁻¹ of roots and a LAI of 3.4) following the approach of the four models under study: (A) in Aquacrop and (B) in CERES-Wheat, WOFOST and CropSyst (for vapour pressure deficit of 0.5 and 1.0 kPa).

Table 5. Model's performance for potential grain yield and aboveground biomass (Mg ha⁻¹) and phenology.

	Plots	N	Observed	Mean				RMSE			
				Aquacrop	CERES	CropSyst	WOFOST	Aquacrop	CERES	CropSyst	WOFOST
<i>Yield</i> (Mg ha ⁻¹)	Calib set ¹	6	–	6.17	6.48	6.41	6.29	–	–	–	–
	Valid set	24	–	6.45	5.80	5.00	5.83	–	–	–	–
<i>Biomass</i> (Mg ha ⁻¹)	Calib set	6	–	13.71	17.69	12.81	15.10	–	–	–	–
	Valid set	24	–	14.33	15.86	10.00	13.83	–	–	–	–
<i>Anthesis</i> (average day of the year)	Calib set	6	115.0	104.0	115.0	115.0	115.0	11.0	0.0	0.0	0.0
	Valid set	24	116.5	113.8	121.8	120.8	117.5	12.0	6.8	8.3	9.0
<i>Maturity</i> (average day of the year)	Calib set	6	147.0	140.0	147.0	147.0	147.0	7.0	0.0	0.0	0.0
	Valid set	24	152.0	142.0	150.5	148.8	147.0	11.3	5.0	7.3	8.5

¹ “Calib set”: season 2000-2001; “Valid set”: remaining seasons.

Soil water content in bare soil

Soil water content ($\text{cm}^3 \text{ cm}^{-3}$) was simulated for 1st January to 31st December 2001 with the four models assuming the soil left bare. Average soil water content simulations in the profile and soil water balance are shown in Figure 3 and Table 6, respectively. Different models show a rather dissimilar performance in terms of the evenness of soil water content changes and of the soil water balance.

Table 6. Soil water balance in Aquacrop, CERES-Wheat, CropSyst and WOFOST from 1st January to 31st December 2001.

	Change	Inputs	Outputs		
	ΔW (mm)	Rainfall (mm)	Soil evaporation (mm)	Drainage (mm)	Runoff (mm)
Aquacrop	+0.45	237.50	236.85	0.00	0.20
CERES-Wheat	-6.16	237.50	241.72	0.00	1.94
CropSyst	+27.90	237.50	224.94	-17.21	1.87
WOFOST	+17.50	237.50	220.00	0.00	0.00

Regarding the evenness of the changes in soil water content, WOFOST simulated the sharpest increases and decreases in soil water after rainfall events and in the dry periods, respectively, followed by CERES-Wheat. Aquacrop and Cropsyst simulated more even changes in soil water. This behaviour is the consequence of the number of layers in which the soil is subdivided, a single layer in WOFOST, 5 layers in CERES-Wheat, 12 in Aquacrop and, lastly, 16 in CropSyst (Table 2). Therefore, the more layers subdividing the soil profile, the more even the simulated soil water changes.

Rather dissimilar results were found in the soil water balance of different models (Table 6). CERES-Wheat has a negative water balance (-6.16 mm); while it was positive for the remaining models. The lowest increase in soil water was simulated by Aquacrop (+0.45 mm), followed by WOFOST (+17.50 mm) and CropSyst (+27.90 mm). These differences were expected to be due to soil water outputs (soil evaporation, deep drainage and runoff), as all models were inputted

with the same weather dataset (rainfall, the sole water input). Soil evaporation was the highest water output in all models. WOFOST simulated the lowest soil evaporation (220.00 mm), followed by CropSyst (224.94 mm), Aquacrop (236.85 mm) and CERES-Wheat (241.72 mm). Drainage was nul for all models but for CropSyst, where it was negative (-17.21 mm) despite no water table was considered. This is due to an artefact of the model when using the finite difference approach (as used in this work), the water content of the layer below the defined soil profile is assumed to be equal to the water content of the deepest simulated layer. Then, water upflow is allowed from subsoil to satisfy the daily change in gradient when high demand (Díaz-Ambrona et al., 2005). Finally, runoff was rather low for all models, being nul in WOFOST, 0.20 mm in Aquacrop, 1.87 mm in CropSyst and 1.94 mm in CERES-Wheat.

Figure 3. Soil water content ($\text{cm}^3 \text{cm}^{-3}$) simulated with Aquacrop, CERES-Wheat, CropSyst and WOFOST between 1st January and 31st December 2001.

Hence, the highest change in soil water in CropSyst is explained by the lower soil evaporation and the unexpected water input from the subsoil (negative deep drainage, Table 6). In WOFOST, soil water increase is related to the lower soil evaporation with respect to the remaining models. As no observed data on soil water outputs on the soil left bare was available, it was not possible to determine if actual soil evaporation was closer to the simulated by Aquacrop and CERES-Wheat (about 238mm) or to the simulated by WOFOST and CropSyst (about 220 mm); otherwise, soil evaporation could be calibrated by modifying U_{limit} in Aquacrop and CERES-Wheat (as it is user specified), or by modifying the superficial layer depth in CropSyst.

Water input from the subsoil simulated by CropSyst is an important fact to be considered. If a wet subsoil exists (as in more coarse-texture soils), using the finite-difference approach (as in CropSyst in this work), is likely to better simulate actual water dynamics in the soil; however, if no wet subsoil is present, using this approach, an overestimation of soil water content is expected (Díaz-Ambrona et al., 2005).

Sensitivity to soil water deficit

As reviewed by Saseendran et al., (2008), water stress might modify crop development rate, leaf initiation and expansion rate, photosynthesis, carbon allocation and partitioning, and root length and density in soil layers. All four models include the impact of water deficit on crop production but each to a different degree of detail (Table 1).

Parameters related to water stress are shown in Table 7. In Aquacrop, drought resistance is defined for canopy expansion, stomatal closure, flowering and canopy senescence. In CERES-Wheat, parameters related to water deficit effects are based on the ratio potential uptake/potential loss, on soil water content as a fraction of field capacity to permanent wilting point, and the acceleration of leaf area senescence by water shortage. Those parameters are fixed to a species scale and none of them were modified in the calibration. In CropSyst, leaf water potential at the onset of stomatal closure and at wilting point (-1300 and -2000 J kg^{-1}), maximum water uptake (10 mm day^{-1}) and leaf area duration sensitivity to water stress were used to calibrate water limited biomass production (Table 7). In WOFOST, two parameters define the effect of water deficit in water limited biomass production, a correction factor for transpiration rate (*CFET*) and a crop group number for water depletion (*DEPNR*). *CFET* was set at 1.1. *DEPNR* or group number was 3.5; as a moderate drought resistant crop (Table 7).

Table 7. Crop parameters related to water deficit as used in CropSyst, WOFOST and Aquacrop.

Parameter	Value
<i>Aquacrop</i>	
Canopy expansion sensitivity	$p_{upper}: 0.10; p_{lower}: 0.45; f_{shape}: 5$
Stomatal closure sensitivity	$p_{upper}: 0.55; f_{shape}: 6$
Early canopy senescence sensitivity	$p_{upper}: 0.58; f_{shape}: 6$
Sensitivity during flowering	$p_{upper}: 0.76$
Sensitivity after flowering (yield formation) (b)	3
<i>CropSyst</i>	
Leaf water potential at the onset of stomatal closure ($J\ kg^{-1}$)	-1300
Wilting leaf water potential ($J\ kg^{-1}$)	-2000
Maximum water uptake ($mm\ day^{-1}$)	10
Ratio unstressed crop ET/ ET_o	1 (0.8–1.4)
Leaf area duration sensitivity to water stress	2 (0-3)
Adjustment factor for phenological response to stress	1 (0-1)
<i>WOFOST</i>	
Correction factor transpiration rate	1.1
Crop group number for soil water depletion	3.5

Figure 4 shows the curves defining water stress impact on a well-established crop depending on available soil water ($cm^3\ cm^{-3}$) for the four models. Figure 4A includes the impact of water deficit on the daily aboveground biomass, assuming a potential production of $200\ kg\ ha^{-1}\ day^{-1}$ for CERES-Wheat (*SWDF1* and *SWDF2*), CropSyst and WOFOST. The lowest sensitivity to low soil water availability was found for Aquacrop, followed by CERES- Wheat *SWDF1* and CropSyst and WOFOST. Lastly, CERES- Wheat *SWDF2* (affecting leaf and stem extension growth and tillering, physiological processes especially sensitive to the decrease in turgor) showed a higher sensitivity for the same available water in the soil. As an overall picture, the sensitivity of *SWDF1* and CropSyst are comparable, while the water stress factors affecting different crop processes in Aquacrop, WOFOST and *SWDF2* differ from those and from each other. Figure 4B shows the canopy growth and decline ($\% \ ^\circ C\text{-}day^{-1}$) in Aquacrop. Maximum canopy growth and decline rates were obtain by dividing the maximum canopy cover (CC_{max}) by the thermal time for canopy expansion and for canopy decline. Canopy growth is the most sensitive process

according to the Aquacrop modelling approach (Figure 4B). For low water availability, lower than $0.063 \text{ cm}^3 \text{ cm}^{-3}$ (below the permanent wilting point), canopy stops expanding. This high sensitivity to low water availability explains that neither yield nor aboveground biomass was simulated by Aquacrop for most of the seasons (Table 8), those with the lowest initial soil water content before sowing (Table S2).

Figure 4C shows the negative effect of water deficit on harvest index (no positive effects resulted after calibration). These results suggest that wheat crop is more sensitive to water deficit during flowering and afterwards during the grain filling period.

Observed vs. simulated water limited yields, mid-season and at harvest aboveground biomass (Mg ha^{-1}) are presented in Table 8. The limitations of the Aquacrop model to simulate crop growth under the driest conditions prevented the possibility of using the calibration set of observations. The best fit in water limited yields was obtained with CropSyst model (0.26 and 1.25 Mg ha^{-1} , for the calibration and the validation set, respectively), followed by WOFOST (0.95 and 1.44 Mg ha^{-1}) and CERES-Wheat (1.18 and 1.97 Mg ha^{-1}). On the contrary, for aboveground biomass, the best fit was obtained with CERES-Wheat (4.02 and 5.90 Mg ha^{-1} , at mid-season and 5.54 and 5.76 Mg ha^{-1} , at harvest, for the calibration and the validation set, respectively), followed by CropSyst (5.24 and 6.47 Mg ha^{-1} , and 6.19 and 6.27 Mg ha^{-1}) and WOFOST (7.96 and 8.41 Mg ha^{-1} , and 12.89 and 8.65 Mg ha^{-1}). In this last model, aboveground biomass is clearly overestimated, and therefore a low harvest index is calculated.

The effect of water stress on plant development is highly specific to the species and to timing in relation to the growth stages. The inclusion of this effect in a model influences the accuracy in the prediction of the most important phenology dates, and it controls the length of crucial development stages as the grain filling period. WOFOST does not include the effect of water stress on plant development; this might explain the underestimation of the harvest index. Boogaard et al. (2013) found that growth reduction due to drought mainly occurred during the grain filling period in their simulations with WOFOST. On the contrary, the effect of water stress on crop development is directly included in other models. In Aquacrop, water stress lowers the canopy expansion rate (computed separately from biomass growth) and accelerates development (AQ_CSen) leading to earlier canopy senescence. In CERES-Wheat, a reduction in biomass production during the vegetative growth of the crop decreases sink strength, and consequently, the relative growth of both vegetative and storage organs are limited. This model also differentiates the strength of water deficit stress depending on the development stage of the crop,

being stronger during the exponential growth phase (*SWDF1* and *SWDF2* in Figure 4A). Crops also respond to water stress with respect to developmental rate (McMaster et al., 2005). In CropSyst, water stress accelerates thermal time accumulation and also modulates leaf area duration promoting earlier canopy senescence.

Plant response to water stress involves physiological changes as stomatal closure that tend to limit water consumption and which result in a reduction of both transpiration and photosynthesis (Farquhar and Sharkey, 1982). Most crop models use a relationship between readily water availability and the ratio of actual to potential transpiration to compute the effect of water deficit on crop growth (Saseendran et al., 2008). This is also the case in the four models studied in this paper. The reduction in the ratio actual to potential transpiration is translated in a lower daily biomass production in Aquacrop, CERES-Wheat and CropSyst and a reduction of daily assimilates production in WOFOST (Table 1).

Changes in the partitioning of carbon allocation and translocation of assimilates to different organs are plant adaptation mechanisms to changes in environmental conditions. Water stress during the vegetative growth of some crops involves a shift in partitioning of assimilates to roots; this helps roots to explore a larger soil volume for water (and nutrients) and increases water uptake to reduce stress (Asseng et al., 1998b). This effect is computed in CERES-Wheat. Flowering is a sensitive stage for many crops (Turner, 2004). Harvest Index (*HI*) is modified in Aquacrop and CropSyst according to water stress. Lastly, water deficits just before reaching maturity might have a positive effect on the *HI* (Geerts and Raes, 2009), this effect is targeted in Aquacrop in an empirical way.

All models implemented the response to water stress in a multiplicative form, based on penalties over the daily assimilates or AgB production, the development rate and the number of flowers pollinated. The difficulty to measure water stress in terms of the effect on individual processes in the crop is the reason for using this empirical approach that cannot be easily justified in terms of biological meaning.

Table 8. Model's performance for attainable grain yield, mid-season and aboveground biomass (Mg ha⁻¹).

	Plots	N	Observed	Mean				RMSE			
				Aquacrop ¹	CERES	CropSyst	WOFOST	Aquacrop	CERES	CropSyst	WOFOST
<i>Yield</i>	Calib set ²	6	2.62	–	1.45	2.62	3.55	–	1.17	0.25	0.93
(Mg ha⁻¹)	Valid set	24	2.55	3.24 (3.00)	2.43	3.08	2.00	1.01	1.13	1.15	1.01
<i>Mid-season AgB</i>	Calib set	24	3.71	–	3.98	4.66	7.81	–	1.41	1.41	4.10
<i>Biomass</i> (Mg ha⁻¹)	Valid set	54	5.81	6.47 (6.94)	5.93	6.33	7.63	2.49	2.99	2.24	2.23
<i>Final AgB</i>	Calib set	6	5.26	–	4.45	5.23	12.36	–	1.15	0.72	7.10
<i>Biomass</i> (Mg ha⁻¹)	Valid set	24	6.90	7.20 (8.64)	5.97	6.16	7.30	1.96	1.97	2.02	1.28

¹ Just includes season 2002-03 (all blocks) and season 2011-12 (block 1), as no crop growth was simulated in the remaining cases due to high sensitivity of Aquacrop to low water availability.

² “Calib set”: season 2000-01; “Valid set”: remaining seasons.

Figure 4. (A) Aboveground biomass ($\text{kg ha}^{-1} \text{ day}^{-1}$) in Aquacrop, CERES-Wheat (SWDF1: photosynthesis and transpiration, SWDF2: leaf and stem extension growth and tillering), CropSyst and WOFOST; (B) canopy growth and decline ($\% \text{ } ^\circ\text{C-day}^{-1}$) in Aquacrop and (C) harvest index (–) in Aquacrop versus the soil available water ($\theta_t - \theta_{WP}$). At available soil water $0 \text{ cm}^3 \text{ cm}^{-3}$ corresponds to the permanent wilting point at $pF=4.2$.

As an overall picture of rainfed crop simulations, rather dissimilar results were output by different models and no good fits were reached in general. *RMSE* in the validation set were higher than 1 Mg ha⁻¹ for water limited yields and mostly higher than 2 Mg ha⁻¹ for aboveground biomass. This might be explained by the substantial differences in soil water simulations (Tables 9 and 10 and Figure 5). Soil water in the profile was mostly overestimated in Aquacrop and WOFOST; while this was underestimated in CropSyst. These dissimilarities in soil water simulations are a joint result of the alternative approaches to calculate soil water dynamics used by different models (Table 2) and the differences in crop growth simulations that modifies soil water outputs (mainly through transpiration) affecting soil water content.

The alternative approaches include a different number of layers or sub compartments in the soil for computing the effect of this cascade on the soil water contents and a different detail on soil texture (heterogeneous vs. homogenous profile) (Table 2). WOFOST uses a single layer that grows in time with rooting depth and, therefore, it just considers a homogenous textured soil. The other models are more flexible in their structure and allow selecting the number of layers and detail on soil texture. This difference in soil texture leads to different pedotransfer functions.

Table 9. Model's performance for soil moisture content ($\text{cm}^3 \text{ cm}^{-3}$) at two soil depths (0-25 cm and 25-50cm) and total water available in the rooted soil profile (90cm) (mm) for Aquacrop, CERES-Wheat and CropSyst.

		Plots	N	Observed	Mean			RMSE		
					Aquacrop ¹	CERES	CropSyst	Aquacrop	CERES	CropSyst
Mid - season	$\theta_{0-25\text{cm}} (\text{cm}^3 \text{ cm}^{-3})$	Calib set ²	6	0.265	–	0.193	0.238	–	0.082	0.027
		Valid set	24	0.211	0.200 (0.255)	0.177	0.207	0.058	0.050	0.033
	$\theta_{25-50\text{cm}} (\text{cm}^3 \text{ cm}^{-3})$	Calib set	6	0.227	–	0.170	0.147	–	0.060	0.073
		Valid set	24	0.179	0.199 (0.219)	0.154	0.149	0.028	0.034	0.035
	Water profile (mm)	Calib set	6	207	–	137	135	–	69	71
		Valid set	24	169	196 (106)	133	143	19	40	27
Harvest	$\theta_{0-25\text{cm}} (\text{cm}^3 \text{ cm}^{-3})$	Calib set	6	0.098	–	0.162	0.187	–	0.073	0.089
		Valid set	24	0.083	0.146 (0.098)	0.125	0.161	0.087	0.077	0.083
	$\theta_{25-50\text{cm}} (\text{cm}^3 \text{ cm}^{-3})$	Calib set	6	0.095	–	0.089	0.099	–	0.012	0.011
		Valid set	24	0.110	0.159 (0.124)	0.123	0.101	0.059	0.037	0.024
	Water profile (mm)	Calib set	6	88	–	95	114	–	25	26
		Valid set	24	94	196 (106)	100	112	103	30	29

¹Just includes season 2002-03 (all blocks) and season 2011-12 (block 1), as no crop growth was simulated in the remaining cases due to high sensitivity of Aquacrop to low water availability.

²“Calib set”: season 2000-01; “Valid set”: remaining seasons.

Table 10. Model's performance for soil moisture content ($\text{cm}^3 \text{cm}^{-3}$) and total water available (mm) in the rooted depth for WOFOST with permanent wilting point (PWP) and field capacity (FC) at 0.095 and 0.235 $\text{cm}^3 \text{cm}^{-3}$, respectively. PWP and FC were calculated as the weighted average of the soil water volume at -1500 and -33 kPa, respectively, measured in the laboratory for each soil layer.

		Plots	N	Observed	WOFOST	
					Mean	RMSE
Mid - season	$\theta_{\text{profile}}^1 (\text{cm}^3 \text{cm}^{-3})$	Calib set ²	6	0.155	0.226	0.082
		Valid set	24	0.184	0.198	0.068
	Water profile (mm)	Calib set	6	207	127	80
		Valid set	24	179	170	12
Harvest	$\theta_{\text{profile}} (\text{cm}^3 \text{cm}^{-3})$	Calib set	6	0.066	0.128	0.065
		Valid set	24	0.109	0.137	0.050
	Water profile (mm)	Calib set	6	88	115	27
		Valid set	24	94	123	40

¹ Average soil moisture content in the rooted depth.

² "Calib set": season 2000-01; "Valid set": remaining seasons.

Transpiration is a key factor driving crop growth. Alternative root distribution modelling approaches (Table 1) leads to differences in simulated crop transpiration and yields in different models. The coincidence of higher soil moisture content with the larger root fraction, leads to a higher water uptake and thus to a higher crop transpiration and biomass production. Moreover, in a layered-soil model there may be higher water content in the top layers, enabling the roots to take up a higher amount of water. This is overlooked by models as WOFOST not simulating root distribution, just rooting depth in one layer assuming water to be homogeneously distributed.

All these interactions seriously hamper a completely separate test, unless data on soil water content on a bare soil is available.

Figure 5. Observed versus simulated water content in the potential rooting depth (90 cm) (mm), and soil moisture content at two depths (0-25cm and 25-50cm), for no tillage (black markers) and minimum tillage (hollow markers), at mid-season (cycles) and at harvest (triangles) for (A) Aquacrop, (B) CERES-Wheat, (C) CropSyst and (D) WOFOST. In WOFOST, the average soil moisture content in the profile was used, as no differentiation in layers is considered in this model.

3.2. Critical remarks about the calibration and use of the models

Crop models have a broad field of application in research and engineering. In research, models are used to better understand physiological processes of the crops or interactions between soil, crop and atmosphere. In engineering, crop models are often used to assist policy makers, to foresee the potential impact of changes in management or to quantify resource use requirements and efficiencies.

More careful attention is usually paid to the accuracy of each crop or soil parameter when using crop models within a research perspective, while from an engineering perspective focus is often more output-oriented. Nevertheless, the importance of restricting parameter values to their physiological and/or physical plausible range is evident. Model users solely concerned about model outputs might obtain good fits of model simulations to observed data with a set of parameters with no biological meaning, where an overestimation of some processes compensates the underestimation of others, e.g. too low a LAI might be masked by too high a light extinction coefficient or radiation use efficiency. This would result in the degeneration of the crop model being a very expensive statistical tool.

Despite the existence of both mechanistic and empirical models robust enough to be used for yield prediction and decision support in agriculture, mechanistic models are more adequate to understand the nature of the physiological processes (Yin and Struik, 2010) and to identify crop phenotypes or management strategies to improve yields. In either case, there will always be a certain degree of empiricism due to our limited knowledge of the system, and because the behaviour of the system as a whole (crop-soil-atmosphere) is not always directly explained based on the behaviour of isolated physiological processes (Yin and Struik, 2010). Empirical parameters might also combine in their value some of the errors of other parameters, especially when these are obtained by calibration to best fit model simulations to field observations.

The more a process is modelled with parameters lacking a clear unit, the harder it becomes to determine whether a parameter value is within a justifiable range. Parameters such as SLA or HI have a measurable, physical aspect, which in part depends on sampling expertise; on the contrary, parameters such as the f_{shape} in the curves of sensitivity in Aquacrop are abstract and cannot be measured, but can only be obtained by calibration.

Many uncertainties are also related to soil water simulation. For example, with respect to the assumptions that have to be made on the horizontal heterogeneity of the soils, the effect of

tillage, or the criterion used to fix the parameters for permanent wilting point (*PWP*) and field capacity (*FC*) required in soil modules. Tillage alters soil structure and soil organic matter content (Álvaro-Fuentes et al., 2013; Plaza-Bonilla et al., 2014; White et al., 2009); consequently, it modifies soil water dynamics. Observed data used within this work shows differences in soil water content between the two tillage treatments, being higher in plots with no tillage than in those with minimum tillage, especially in deeper layers ($\theta_{25-50\text{cm}}$) (Figure 5). However, model settings for soil management were not differentiated to simulate plots belonging to each of the treatments, as no data on the changes in soil water characteristics were available. The same *FC* and *PWP*, measured at -33 kPa and -1500 kPa using the Richards membranes apparatus, were used. Field capacity is an agronomic concept broadly used by agronomists and soil scientists with still no consensus on its definition. *FC* refers to the upper limit of the plant available water retained by the soil; it was defined by Veihmeyer and Hendrickson (1931) as the “amount of water remaining in soil after the excess water has drained away and after the rate of downward movement of water has materially decreased”, that is one to three days. Nevertheless, this definition is not rigorously defined in physical terms. Numerous operational definitions have been introduced, e.g. based on the matric potential and the drainage rate. Colman (1947) showed that an initially saturated sieved soil after three hours of internal drainage approaches a matric potential of -33 kPa. It is difficult to define when drainage rate has “materially decreased”; Nachabe (1998) assumed that soil water at *FC* is reached when drainage rate equals 0.05 mm day^{-1} . Anyway, any standardization brings together inaccuracies as it varies with soil texture, soil structure and the organic materials adhered to soil particles; e.g. the pressure values at *FC* have been reported from -10 to -100 kPa (Cassel and Nielsen, 1986).

It is often the case that no data on a number of variables is available, and the problem of data availability becomes worse as the scale of application is larger; e.g. the increasing heterogeneity of soils in a larger area. In a context with limited observed variables, even when constraining parameter values within their plausible range, model users cannot verify the correctness of the simulation; e.g. in a rainfed cropping system, no data on soil moisture content might lead to compensation of too low a water availability with a higher tolerance of the crop to water deficit and vice versa. This common lack of observed data seriously hampers testing simulation models and thus constitutes an important source of uncertainty. It is then recommended that trials are designed in such a way that most crucial processes in the cropping system are monitored to decrease the probability of incurring these compensating mechanisms. Data used in this work was originally collected to evaluate the effect of alternative tillage systems and therefore variables that would have been of interest regarding the objective of this work lack: e.g., data on plots

grown under non-limiting conditions and soil water content dynamics measured on bare soil. Data from plots irrigated according to ET_0 would have given a reference of potential yields and AgB (Table 5) and measured (not estimated) values for the parameters shown in Table 4 (such as RUE , WP , maximum LAI or CC and phenology under no limiting conditions). Soil water content measured on bare soil would have enabled the separate evaluation of the soil water modules, without the influence of a crop transpiring soil water.

The use of multi-crop model simulations is becoming common in climate change research (e.g. Asseng et al., 2013; Rötter et al., 2011). In this context, these authors propose that simulations of an ensemble of crop models are superior to a single model as errors would tend to cancel each other out. However, according to our experience, the uncertainties associated with a single model are multiplied with a multi-model approach, and averaging multiple model outputs does not contribute to our understanding of the crop-soil-atmosphere system. Data availability to validate model outputs, the degree of model's empiricism, and parameter measurability represent important sources of model uncertainty.

4. Concluding remarks

The different approaches used in the models, both in the soil and in the crop modules, make their comparison arduous in terms of soil hydraulic functions and crop parameter values. For that reason, models were compared by (i) an in-depth analysis of computational approaches of a specific soil and crop process, with the purpose of understanding the reasons of the different model's performance; and by (ii) contrasting model outputs, to orient results towards a more engineering or business use as it gives an idea on what will be expected by the model user. The results suggest that the outputs of the four models for the simulation of winter wheat growth are comparable when water is not limiting, but differences are larger when simulating yields under rainfed conditions. These differences in rainfed yields are mainly related to the dissimilar simulated soil water availability.

This paper answers the questions initially posed:

Which model(s) is/are preferred from the four investigated models?

For the simulation of winter wheat growth at field scale in such semi-arid conditions, CERES-Wheat and CropSyst are preferred. WOFOST is a satisfactory compromise between data availability and complexity when detail data on soil is limited. Aquacrop integrates physiological processes in some representative parameters, thus diminishing the number of input parameters, what is seen as an advantage when observed data is scarce. However, the high sensitivity of this model to low water availability limits its use in the region considered.

Is the use of several models better than the use of a single one?

In this paper, models were calibrated in such a way that input parameters (e.g. soil data, thermal time) and variables such as radiation use efficiency and daily potential biomass production were comparable between models. Though, differences between models' outputs are significant when simulating under water-limiting conditions. This suggests that the uncertainties associated with a single model are multiplied with a multi-model approach, and averaging multiple model outputs does not contribute to our understanding of the crop-soil-atmosphere system. Contrary to the use of ensembles of crop models, we endorse that efforts be concentrated on selecting, studying and calibrating the most adequate crop model on the basis of the objectives for which simulations will be used.

Given the uncertainties associated to crop model simulations, is it worth to use crop models in crop insurance analysis and design? If so, what cautions should be taken when using crop model simulations for decision making regarding model(s) calibration and implementation?

The use of such complex methodologies as crop models is associated with numerous sources of uncertainty, although these models are the best tool available to get insight in these complex agronomic systems.

The common lack of observed data seriously hampers testing simulation models and thus constitutes an important source of uncertainty. In order to assess the suitability of crop models for quantifying actual yield variability, a comparison of massive observed actual yields and

simulated yields, both on a regional scale, should be first conducted. Models should be preferably calibrated with data registered in trials specifically designed to obtain parameter values for the selected model.

5. Acknowledgements

This research was funded by the project MULCLIVAR, from the Spanish Ministerio de Economía y Competitividad (MINECO) CGL2012-38923-C02-02, A. Castañeda-Vera has a PhD grant from UPM. She visited Wageningen University from February to May 2014, where the first discussions on this paper started. We thank Joost Wolf (Wageningen University), Dirk Raes (KU Leuven University) and Tony Hunt (University of Guelph) for the clarifying discussions on WOFOST, Aquacrop and CERES-Wheat models, respectively.

6. Appendix A. Supplementary data

Table S1. Data on cultivars and sowing dates, and measured phenology dates, final yield and aboveground biomass and harvest index at the experimental trials.

<i>Treatment</i>	<i>Season</i>	<i>Block</i>	<i>Cultivar</i>	<i>Dates (day of the year)</i>			<i>Yield (kg ha⁻¹)</i>	<i>Biomass (kg ha⁻¹)</i>	<i>Harvest index</i>
				<i>Sowing</i>	<i>Anthesis</i>	<i>Maturity</i>			
<i>No tillage</i>	2000-01	1	<i>Soissons</i>	334	115	147	2858	5014	0.57
		2					3117	7480	0.42
		3					3398	6707	0.51
	2002-03	1	<i>Soissons</i>	310	120	150	4471	11521	0.39
		2					3155	8874	0.36
		3					3469	11618	0.30
	2004-05	1	<i>Anza</i>	320	110	145	3023	6818	0.44
		2					2033	5349	0.38
		3					2651	6466	0.41
	2006-07	1	<i>Soissons</i>	311	118	156	3743	8836	0.42
		2					3555	9160	0.39
		3					3695	10133	0.36
	2011-12	1	<i>Bokaro</i>	315	113	144	1538	5095	0.30
		2					904	3359	0.27
		3					681	2959	0.23
<i>Minimum tillage</i>	2000-01	1	<i>Soissons</i>	334	115	147	1922	4178	0.46
		2					2106	4000	0.53
		3					2313	4155	0.56
	2002-03	1	<i>Soissons</i>	310	120	150	3033	7220	0.42
		2					2734	8308	0.33
		3					2591	7826	0.33
	2004-05	1	<i>Anza</i>	320	110	145	1298	3893	0.33
		2					1952	5049	0.39
		3					1274	3412	0.37
	2006-07	1	<i>Soissons</i>	311	118	156	3181	9416	0.34
		2					3049	8120	0.38
		3					4546	11120	0.41
	2011-12	1	<i>Bokaro</i>	315	113	144	1674	4081	0.41
		2					1804	4214	0.43
		3					1061	2789	0.38

Table S2. Initial soil water contents measured at the experimental trials.

<i>Treatment</i>	<i>Season</i>	<i>Block</i>	<i>Date measurement (day of the year)</i>	$\theta_{0-25\text{cm}}$ (cm ³ cm ⁻³)	$\theta_{25-50\text{cm}}$ (cm ³ cm ⁻³)	<i>Initial soil water 0-90cm (mm)</i>
<i>No tillage</i>	2000-01	1	268	0.101	0.107	101
		2		0.100	0.074	79
		3		0.061	0.071	83
	2002-03	1	296	0.265	0.238	222
		2		0.232	0.182	167
		3		0.229	0.210	198
	2004-05	1	279	0.203	0.194	185
		2		0.174	0.166	150
		3		0.169	0.151	153
	2006-07	1	282	0.142	0.132	133
		2		0.145	0.113	111
		3		0.136	0.121	100
	2011-12	1	175	0.140	0.137	122
		2		0.143	0.094	95
		3		0.124	0.086	90
<i>Minimum tillage</i>	2000-01	1	268	0.087	0.077	96
		2		0.066	0.064	61
		3		0.062	0.060	67
	2002-03	1	296	0.233	0.168	159
		2		0.246	0.181	180
		3		0.200	0.175	161
	2004-05	1	279	0.123	0.130	127
		2		0.099	0.108	112
		3		0.114	0.126	129
	2006-07	1	282	0.126	0.112	91
		2		0.127	0.116	108
		3		0.113	0.101	103
	2011-12	1	175	0.137	0.100	123
		2		0.114	0.126	129
		3		0.083	0.062	67

Chapter 4

4. Evaluation of risk management tools for stabilising farm income under CAP 2014-2020

A. Castañeda-Vera^{1,2} and A. Garrido¹

¹ Research Centre for the Management of Agricultural and Environmental Risks (CEIGRAM), Universidad Politécnica de Madrid (UPM), Spain

² Agricultural Systems Group, Dpto. Producción Agraria, E.T.S. Ingenieros Agrónomos, Universidad Politécnica de Madrid, Avda. Complutense s/n, 28040 Madrid, Spain

Agricultural Finance Review special issue (under review).

Abstract

Guaranteeing farm income stability is an objective of the European Union's and the Spanish agricultural policies. Alternative risk management strategies with differing degree of public support, regulatory approaches and farmer's co-responsibility are presently available. In this paper, CAP direct payments, diversification, crop insurance and an Income Stabilisation Tool (IST) were compared considering (i) their effect on farm income and income stability, (ii) the expected farmers' willingness for adoption, and (iii) the efficiency of public expenditure invested in supporting them. The analysis is illustrated with an average arable farm in the region of Castilla y León (northern central Spain). Main conclusions point at direct payments and crop diversification as the most effective measures in decreasing income variability. Nevertheless, using crop insurance or an IST has potential for both improving farm resilience to yield and price variability and limiting public expenditure. The higher effectiveness of the IST in increasing farm income stability might result in a decrease of insurance demand if IST schemes are offered jointly with agro-insurance.

Keywords: income volatility, direct payments, diversification, income stabilisation fund, crop insurance, Spain.

1. Introduction

Farm income is volatile due to fluctuating yields and input and output prices. Yield variability is mainly related to weather, and pests and diseases; while prices of the main commodities are set internationally based on global annual production, demand changes and stocks, and other market factors, all of them hardly affected by individual farmer decisions (Garrido et al., 2016).

Guaranteeing stability and a fair standard of living to farmers are objectives of the Common Agricultural Policy (CAP), already established by the EC Treaty. CAP measures to reach such objectives have evolved to meet social demands towards a more environmentally respectful agriculture and restraining CAP spending, as well as to respond to external aspects as World Trade Organization (WTO) negotiations. CAP has shifted from supporting product prices to supporting producers, mainly through direct payments. Market instruments remain now as safety nets, ensuring that intervention prices are only used under real crisis scenarios; thus exposing farmers to significant variations of market prices.

For the eligible farmers, direct payments contribute to stability by ensuring a stable source of income, while other CAP measures reinforce farm income stabilisation, like crop diversification and risk management tools. Crop diversification is a requirement for receiving the greening payment, which accounts to one third of the direct payment check. Besides enhancing environmental effects, crop diversification might reduce income fluctuation, compensating the possible worse performance of a crop with a better performance of another in the same year. The use of alternative mixed public-private risk management tools is helpful to buffer the negative effects of such fluctuations. Such products involve the co-responsibility of farmers, lower public expenditure, and when designed within certain limits, they are considered not to distort markets (Meuwissen et al., 2008). In other developed countries, like the U.S. and Canada, many insurance products and income stabilization tools are designed and offered to farmers, with significant public support. In Europe, well-developed agricultural insurance systems already exist, as it is the case of Spain (Bardají and Garrido, 2016). This alternative approach has been materialized after the last CAP reform supporting crop insurance, mutual funds and income

stabilization tools (IST) through the European Agricultural Fund for Rural Development (EAFRD).

However, the actual success of such measures will not only depend on their ability to stabilise farm income but also on the willingness of farmers to adopt them. Some of the measures do not entail an active participation of beneficiaries (unconditional fraction of the direct payments), though others require a decision of farmers either by participating financially (insurance, saving accounts or mutual fund) or by modifying the farming strategy (diversification). Hence, it is relevant to ask whether farmers would be willing to adopt instruments that are available now and others upcoming in the next years.

Research about CAP's recent and new risk management instruments has focused on the extent to which such tools can decrease income volatility (e.g. Enjolras et al., 2014) or income inequalities between farms (e.g. Finger and El Benni, 2014; Mary et al., 2013) or in comparisons with already implemented tools. Also, research on risk management tools demand has looked at aspects as public incentives through subsidies or the presence of adverse selection (e.g. Donoghue, 2014; Garrido and Zilberman, 2008a; Santeramo et al., 2016; Shaik et al., 2008). Farmers' willingness to adopt such instruments, either by contracting them or by meeting cross-compliance to be beneficiaries, is relevant for policy-makers to design requirements and conditions of these instruments.

No evidence about the CAP 2014-2020 effects is yet available. This paper contributes to estimate them by analysing the role of alternative risk management tools and strategies to stabilize agricultural income in extensive arable farms in Spain and the potential willingness of farmers to adopt them. The analysis is illustrated with a representative arable farm in the region of Castilla y León (northern central Spain), the only Spanish region which has proposed to support Income Stabilisation Tool (IST) within its Rural Development Programmes (2014-2020), and a new tool for risk management proposed within the Rural Development policy. The aims of the study were to (i) quantify farm income and income stability offered by alternative risk management strategies provided at the European, national and private level, (ii) estimate the expected farmers' willingness to adopt the available instruments, and (iii) evaluate the efficiency of public expenditure invested in subsidizing them.

The paper is organised as follows. Section 2 defines the risk management instruments evaluated in this work. Section 3 lays out the theoretical model and Section 4 describes its

practical implementation. Section 5 presents the results that are discussed in Section 6. Lastly, Section 7 concludes the main findings.

2. Risk management instruments

Bielza et al. (2009) classified risk management tools in two groups, either concerning on-farm measures or risk sharing strategies. The first group involves on-farm resources changing the production strategy (diversification, input intensification, retention), the commercial strategy (vertical integration) or the use of farm benefits (stabilisation accounts). The second group involves third parties that assume part of the risk in return of a monetary reward (insurance or mutual funds).

Four risk management measures were considered within this work: CAP direct payments, crop diversification, crop insurance and an income stabilisation tool, designed as a mutual fund. Direct payments contribute to ensuring the long term viability of farms by providing a basic fixed income support, and therefore making them less vulnerable to fluctuations of income. Regulation (EU) No 1307/2013 sets out general guidelines for direct payments implementation, leaving a range of freedom to member states based on subsidiarity principle. Part of the direct payment is conditioned on the performance of farming practices enhancing the sustainable management of natural resources, the so-called, *greening* payment, being diversification one of the requisites. Diversification, through the introduction of more than a single crop in the farm, also contributes to farm income stability reducing yield and price fluctuation.

Direct and decoupled payments scheme presumably incentivises farmers to orient farming decisions to markets. This enhances competitiveness, but in the current context of increasing market and price liberalization, it also exposes farmers to market risks and crises. A number of risk management instruments are available to complement farmers' coping capacity with large income losses.

Regulation (EU) No 1305/2013 on support for rural development by the European Agricultural Fund for Rural Development (EAFRD) considers in its article 36 the possibility of supporting the premium of crop, animal, and plant insurance, mutual funds for adverse climatic events, animal and plant diseases, pest infestations and environmental incidents and income stabilisation tools (IST). However, a number of EU countries already implement similar instruments supporting them through State aids, as it is the case of

agricultural insurance in Spain. Using State aids comprises the advantages of being less restrictive in terms of loss coverage and subsidies, as they might be within the amber box in the WTO agriculture agreement.

Crop insurance is contracted by farmers to protect themselves against yield losses due to natural hazards such as hail, fire, crop damage caused by wild fauna, flood and excessive rainfall and other adversities impeding crop emergence or limiting crop growth (including drought). Multi-peril insurance for arable crops (including cereal, oilseed and protein crops) has been operating in Spain since 1982. It includes a number of insurance modules that differ based on the coverage, on individual fields or the whole farm, rainfed or irrigated, in the way indemnities are calculated when a loss occurs, and on the maximum insurable yield guaranteed.

Crop insurance premiums are proportional to insured yields, vary depending on the contracted module and deductible, and are subsidized by the national and regional governments. Updated maximum insurable yield per municipality is published every year in the Spanish Official Gazette (BOE). For farmers that have contracted the insurance in previous years, maximum insurable yield is based on his/her claims history.

IST works as a mutual fund protecting against low incomes. Regulation (EU) No 1305/2013 states in its article 39 that support to IST is allowed to contribute to up to 65% of the corresponding indemnity to farmers. Such allowance is only permitted when the IST compensates only a drop of income exceeding 30% of the average annual income of the individual farmer in the preceding three-year period (or a three-year average based on the preceding five-year period excluding the highest and lowest entry and payments), to fit with WTO regulations for been considered a green box measure of support. Additionally, such compensation shall compensate no more than 70% of the lost income in the year the producer becomes eligible to receive this assistance.

Article 39 in Regulation 1305/2013 specifies that income, for the purposes IST, *shall refer to the sum of revenues the farmer receives from the market, including any form of public support, deducting input cost*, therefore, to gross margin. Therefore, the IST was defined as follows. Each year, before the growing season, the farmer pays a premium to the mutual fund of the IST. After harvest selling, if income is lower than the 70% of the average income received by the farmer in the preceding five-year period, excluding the highest and lowest entry, the farmer is correspondingly indemnified.

3. Theoretical framework

Producers are expected to design their farming strategy based on the expected utility of different alternatives. Farm income and income stability are core variables driving the utility of a certain outcome.

Farm income (excluding fixed costs) results from deducing farming variable costs:

$$\tilde{\Pi}_t = \tilde{R}_t - \sum_i x_i C_i \quad [\text{Eq 1}]$$

where $\tilde{\Pi}_t$ is the expected farm income available for the farmer in the year t , x_i is the proportion of the crop i in the rotation (takes the value 1 when monoculture) and C_i are the direct costs associated to crop i . Average income was calculated assuming that farm strategy, either monoculture or crop rotation, does not change in time.

$\tilde{R}_t$ is the expected farm revenue in the year t , and it was calculated as follows:

$$\tilde{R}_t = \sum_i x_i (\tilde{p}_{it} \tilde{y}_{it}) + \tilde{T}_t + CAP \quad [\text{Eq 2}]$$

where $\tilde{p}_{it}$ and $\tilde{y}_{it}$ are the price and yield of crop i in year t , $\tilde{T}_t$ are the results of the protection tools in year t , including premiums and indemnities, and CAP is the direct payment received by the farmer. Three possibilities of CAP direct payments were considered depending on whether or not the farmer would be a beneficiary, and if so meeting or not meeting greening requirements.

Both x_i and C_i are assumed to be constant per crop. On the contrary, crop prices and yields vary from year to year as a consequence of climate variability and extreme events, and consequently, $\tilde{T}_t$ varies as well. $\tilde{T}_t$ is calculated differently depending on the protection option (with $\tilde{T}_t = 0$, when no protection tool is used). $\tilde{T}_t$ is calculated based on the protection tool, crop insurance or IST (Eq 3 and 4, respectively).

Contracting crop insurance implies paying a premium (P_{CI}) every year before sowing. Assuming farmer has been contracting insurance for years, insured yields ($Y_{ins,i}$) should be close to average observed yields. Insured yield was set at the average yield of the farmer in the preceding five-year period, excluding the highest and lowest entries. For arable crops insurance, farmers have the option of increasing the insured yield at midseason with a

complementary insurance payment if yield expectation exceeds that insured before sowing. In this work, for the sake of simplicity, this option was not modelled. Indemnities are received when actual yield falls below a certain proportion of the insured yield, a franchise (M):

$$\tilde{T}_t = \begin{cases} [M \times \sum_i p_{it}(Y_{ins,i} - \tilde{y}_{it})] - P_{CI}, & \sum_i \tilde{y}_{it} < M \times \sum_i Y_{ins,i} \\ -P_{CI}, & \sum_i \tilde{y}_{it} \geq M \times \sum_i Y_{ins,i} \end{cases} \quad [\text{Eq 3}]$$

where $Y_{ins,i}$ is the insured yield of crop i .

Contracting the IST also implies paying a premium (P_{IST}) every year before sowing. Indemnities are received when actual income falls below a certain proportion of the average income, a franchise (M):

$$\tilde{T}_t = \begin{cases} [(M \times \overline{\Pi}_5) - \Pi_t] - P_{IST}, & \Pi_t < M \times \overline{\Pi}_5 \\ -P_{IST}, & \Pi_t \geq M \times \overline{\Pi}_5 \end{cases} \quad [\text{Eq 4}]$$

where $\overline{\Pi}_5$ the average income in the last 5 years.

Therefore, both crop insurance and IST have a cost to the farmer (premium) and possible revenues (indemnities). IST and crop insurances premiums are paid independently from farm performance.

4. Empirical application

Two farm strategies scenarios were evaluated: a wheat monoculture (*Triticum aestivum* L.) and a crop rotation meeting CAP greening requirements including 40% wheat, 40% barley (*Hordeum vulgare* L.) and 20% dry beans (*Pisum sativum* L.).

Data on mean yields and market prices in Valladolid province (Castilla y León) between 1993 and 2015 are publicly available at the regional government website (JCyL, 2014) (Figures 1 and 2). Wheat prices, following the same pattern of international commodity prices, have experienced an increase and a high volatility from 2007 (FAO, 2011). Average costs were calculated per crop based on surveyed farms between 2010 and 2013 in

Valladolid province by ECREA resulting in 267, 237 and 105 € ha⁻¹ for wheat, barley and dry peas respectively (MAGRAMA, 2015).

Yield and prices were modelled with the software package @-Risk (Palisade Corporation, 2011) fitting data to the probability distribution functions with the lower value of the statistic chi-square. Asymmetric density functions non bounded on the right and truncated at zero in the left were selected, as used by many authors (Bielza et al., 2004).

Crop yields were modelled using 23 yearly observations (1993-2015). Wheat and dry peas were fitted to beta distributions and barley to a gamma. Several authors point at the significant decrease in yield variability when the aggregation scale of the data used in the analysis increases (Finger, 2012; Górski & Górski, 2003; Lobell, Ortiz-Monasterio, & Falcon, 2007; Marra & Schurle, 1994; Rudstrom, Popp, Manning, & Gbur, 2002). Marra and Schurle, (1994) concluded that standard deviation of yields should be corrected on about 0.1% for each 1% of difference between the mean areas of the scales considered. The average wheat area per farm in the study region is 58.9 ha (INE, 2014), and the agricultural rainfed area of the province, 495000 ha (JCyL, 2014). Using Marra and Schurle (1994) approximation, a rough estimation for a conversion factor to transform province-level to farm-level risks results in 10% increase in yield standard deviation from province to farm, that was applied to the crop yield distribution functions. Additionally, a probability of extreme event damage as hail or fire was considered and set at 4.6 % (Agroseguro, personal communication). Crop yields were found to be correlated at 0.94 for wheat-barley, 0.63 for wheat-peas and 0.65 for barley-peas. Such correlations were considered in the simulation model.

Lognormal distributions were used for prices, as in Bielza et al. (2004) following Goodwin et al. (2000). With the aim to evaluate the effectiveness of each of the considered risk management alternatives, and based on the sharply differentiated periods of price volatility, crop prices were modelled for two scenarios. The first scenario included data in the period 1993 to 2006 characterised by low and stable prices (*Low and stable* scenario). Data from 2007 to 2015, characterized by high and volatile prices, was used in the second scenario (*High and volatile* scenario) (Figure 3).

The franchise for crop insurance and the IST was 30 %, therefore farmers are indemnified only when the drop of yields and income exceeds a 30 % of the average yield and income, respectively. Indemnities compensate for 70 % of the yield or income lost in the year the

farmer becomes eligible to receive this assistance, for the crop insurance and the IST respectively. The average value of the basic CAP payments rights was set at 90.92 € ha⁻¹ (value for 2019) for direct payment beneficiaries and an increase in 51.7 % when meeting greening requirements, therefore 137.93 € ha⁻¹ (FEGA, 2016).

Figure 1. Wheat, barley and dry peas average yield (tonnes ha⁻¹) in Valladolid province between 1993 and 2015. Source: JCyL (2014).

Figure 2. Wheat, barley and dry peas average price (€ tonne⁻¹) in Valladolid province between 1993 and 2015. Source: JCyL (2014).

Crop insurance and IST premiums were calculated as the expected value of the indemnities using 2000 simulated yields and prices. Insurer's loading and commissions were assumed to be 20 % of the premium. The same loading was charged to the IST mutual fund as a safety measure for assuring the fund capacity to overcome crisis. Crop insurance premium is independent from prices and, therefore, it did not change in the two scenarios, it was calculated at 31.4 € ha⁻¹ for wheat monoculture and at 26.1 € ha⁻¹ for crop rotation. A 65% of the insurance premium is defrayed by the State. On the contrary, as IST indemnities depend on income, besides crop yields, IST premium also depends on crop prices and CAP direct payments. IST premium is not calculated over the accumulated indemnities, but on the 35%, as the remainder is defrayed by the European Commission. Results are shown in Figure 4.

Figure 3. Wheat, barley and peas price distributions in the *low and stable prices* scenario (L-S, representing the period 1993-2006) and the *high and volatile prices* scenario (H-V, representing the period 2007-2015).

Each of the combined alternatives (direct payments, diversification and protection tools) was evaluated for the two prices scenarios, using 2000 yield and prices simulations were analysed in 200 ten-year periods.

Figure 4. (A) Crop insurance premium (€ ha⁻¹) , total and state contribution, and (B) IST premium (€ ha⁻¹) for the beneficiaries and no beneficiaries of CAP direct payments, meeting and not meeting greening requirements, practising monoculture or a rotation and for the *low and stable prices* and *high and volatile prices* scenarios.

Farmers' evaluation for alternative farming strategies and risk protection tools were assessed by means of farmer's certainty equivalent. Farmers' certainty equivalents were calculated following Bielza et al., (2007), assuming to be directly proportional to expected income and inversely proportional to CV. Risk aversion was assumed to be constant independently from farmers' profitability, therefore using Constant Risk Aversion Coefficients (CARA).

Two CARA were used, a value of 0.3 to represent a lower risk averse and 0.6 for a higher risk adverse farmers.

$$CE = E(\tilde{\pi}) - \frac{\gamma}{2} \times \sigma(\tilde{\pi}) \quad [\text{Eq 6}]$$

where γ is the CARA coefficient.

Farm economic performance included farmers' profitability in the short term, expected annual income calculated as the mean of the simulated incomes, and in the long term (net present value considering ten years, NPV), low incomes (percentile 5 of annual incomes) and variability (annual income standard deviation and coefficient of variation, CV). NPV was calculated for two discount rates, 5 and 10%.

Additionally, the analysis was completed with an assessment of the increase of the expected and low incomes and certainty equivalent, and the decrease of the coefficient of variation for each alternative with respect to a reference per euro invested by the European Union or the State through subsidies. Such a reference was set as the alternative of no CAP payments, wheat monoculture and no protection tool contracted by the farmer. Each variable was calculated for each ten-year period; results are the mean of the 200 simulations.

5. Results

Results are presented as the main statistics of farm income, certainty equivalents and public cost efficiency. They refer to farmer's profitability in the short term (expected annual income) and in the long term (net present value, NPV), low incomes (percentile 5 of annual incomes) and variability (annual income standard deviation and coefficient of variation), all this reported in Table 1. Farmer's willingness to adopt these measures is evaluated based on the certainty equivalents for two levels of farmer risk aversion (Table 2). Lastly, public cost (PC, € ha⁻¹) and the efficiency of public support for each alternative are shown in Table 3 in terms of expected and low incomes, coefficient of variation and certainty equivalent for high risk averse farmers (CE_{0.6}).

Direct payments were the most effective measure in increasing farm expected and lower incomes. Farms receiving direct payments were significantly more profitable when meeting greening requirements (Table 1). The probability of having a negative income decreases for higher direct payments scenario, for crop rotation and when opting for a protection tool (either crop insurance or the IST) (Table 1).

Contracting protection tools (either crop insurance or the IST) were less effective in increasing farm expected and lower incomes than direct payments. This is due to the annual crop insurance and IST premium cost (Table 1). Regarding the protection tools, the IST was always more effective than crop insurance in terms of increasing farmer's profitability both in the long and in the short term (NPV and expected annual income). The same trend was found for the two discount rates used to calculate the NPV (Table 1).

Income variability was evaluated in absolute terms (standard deviation) or with respect to the mean income (CV). Direct payments were the most effective measure in decreasing annual income variability, followed by diversification. Results show that the decrease in income variability is caused particularly by a reduction of the left tail of income's distribution, while differences in the expected incomes are much lower (Table 1). Regarding protection options, the IST was more effective than crop insurance in increasing farm lower incomes and decreasing income variability (Table 1).

Direct payments are expected to be the measure that most satisfies farmers. Regarding protection tools, the IST is expected to be the preferred (Table 2).

Direct payments, crop insurance and IST involve public expenditure. Direct payments were clearly the most expensive (Table 3). A more interesting analysis arises from evaluating public expenditure of the instruments not in absolute terms but relative to the fulfilment of their objectives, thus increasing expected and low incomes, decreasing income variability and increasing farmer's willingness to adopt certain strategies, as shown in Table 3. Higher public expenditure efficiency was found for subsidizing protection tools than for direct payments, especially for increasing low incomes and decreasing variability per € of public expenditure. Regarding protection tools, higher public expenditure efficiency was found for subsidized IST than for crop insurance but for the most risky case, that is, for no CAP beneficiaries, practising monoculture in the *High and Volatile prices* scenario. Efficiency was always higher for farms diversifying (Table 3).

Public expenditure was efficient as well in potentially increasing farmer's willingness to contract crop insurance and IST, especially when farmers were not CAP beneficiaries and for more risk-averse farmers (Table 3).

Table 1. Expected income, Net present value for 10 years, percentile 5 of incomes (p5) and income standard deviation (St dev, € ha-1) and coefficient of variation, (CV, –) for the beneficiaries and no beneficiaries of CAP direct payments, meeting and not meeting greening requirements, practising monoculture or a rotation and using crop insurance, an income stabilization tool (IST) or no protection option for two prices scenarios (low and stable and high and volatile).

Direct payments	Strategy	Protection option	Scenario LOW and STABLE						Scenario HIGH and VOLATILE					
			Exp ³	NPV _{0.05} ⁴	NPV _{0.10}	p5	St dev	CV	Exp	NPV _{0.05}	NPV _{0.10}	p5	St dev	CV
None	Monoculture	None	94	726	579	-110	134	2.22	231	1786	1424	-54	199	1.17
		Crop insurance	102	786	626	-39	106	1.19	247	1906	1519	43	163	0.71
		IST	127	978	779	39	75	0.58	265	2048	1632	123	130	0.50
	Rotation	None	84	652	520	-56	102	1.54	202	1566	1250	0	155	0.81
		Crop insurance	90	699	557	-18	86	1.03	215	1660	1324	51	136	0.65
		IST	117	904	720	44	67	0.60	238	1838	1465	118	115	0.49
	Monoculture	None	185	1428	1138	-19	134	1.96	322	2488	1982	37	199	0.65
		Crop insurance	193	1488	1185	52	106	0.59	337	2608	2077	134	163	0.49
		IST	213	1645	1310	115	83	0.40	353	2727	2172	199	138	0.39
Basic	Rotation	None	175	1354	1079	35	102	0.60	293	2268	1808	91	155	0.54
		Crop insurance	181	1401	1116	73	86	0.48	306	2362	1882	142	136	0.45
		IST	199	1537	1223	115	75	0.38	321	2484	1979	190	123	0.38
	Monoculture ²	None	232	1791	1426	28	134	0.62	369	2851	2271	84	199	0.56
		Crop insurance	240	1851	1474	99	106	0.46	384	2971	2366	181	163	0.43
		IST	255	1972	1570	151	87	0.35	395	3057	2435	235	142	0.36
	Rotation	None	222	1717	1368	82	102	0.47	340	2631	2097	138	155	0.46
		Crop insurance	228	1764	1405	120	86	0.38	353	2725	2171	189	136	0.39
		IST	241	1865	1485	151	79	0.33	364	2814	2242	227	126	0.35

¹ Income refers to gross margin, as it only includes variable costs

² Only if farms area lower than 30 ha

³ Exp: Expected income; NPV_{0.05} and NPV_{0.10}: Net present value for 10 years at discount rates at 5% and 10%

⁴ Net present value for discount rate at a 5% (VPN_{0.05}) and 10% (VPN_{0.05}).

Table 2. Certainty equivalent (CE, € ha⁻¹) for two coefficients of absolute risk aversion (CARA) for the beneficiaries and no beneficiaries of CAP direct payments, meeting and not meeting greening requirements, practising monoculture or a rotation and using crop insurance, an income stabilization tool (IST) or no protection option for two prices scenarios (low and stable and high and volatile).

Direct payments	Strategy	Protection option	Scenario LOW and STABLE		Scenario HIGH and VOLATILE	
			CE _{0.3}	CE _{0.6}	CE _{0.3}	CE _{0.6}
None	Monoculture	None	54	13	171	112
		Crop insurance	70	38	198	149
		IST ¹	104	82	226	187
	Rotation	None	54	23	156	109
		Crop insurance	65	39	174	133
		IST ¹	97	77	203	169
Basic	Monoculture	None	145	104	262	202
		Crop insurance	161	129	289	240
		IST ¹	188	163	311	270
	Rotation	None	145	114	247	200
		Crop insurance	156	130	265	224
		IST ¹	176	154	284	248
Basic + Greening	Monoculture ¹	None	192	151	309	249
		Crop insurance	208	176	336	287
		IST ¹	229	203	353	310
	Rotation	None	192	161	294	247
		Crop insurance	203	177	312	271
		IST ¹	218	194	326	288

¹ Only if farms area lower than 30 ha

² Certainty equivalent for a Coefficient of Relative Risk Aversion of 0.3 (CE_{0.3}) and 0.6 (CE_{0.6}).

Table 3. Public expenditure (PC, € ha⁻¹ yr⁻¹) and public expenditure efficiency (€ €⁻¹) with respect to the expected income, income stability and certainty equivalent (CE), for the beneficiaries and no beneficiaries of CAP direct payments, meeting and not meeting greening requirements, practising monoculture or a rotation and using crop insurance, an income stabilization tool (IST) or no protection option for two prices scenarios (low and stable and high and volatile).

			Public expenditure efficiency ³ (€ € ⁻¹)										
Direct payments	Strategy	Protection option	Scenario LOW and STABLE					Scenario HIGH and VOLATILE					
			PE ²	Expected	Low	CV	CE _{0.6}	PE ²	Expected	Low	CV	CE _{0.6}	
None	Monoculture	None (Ref 1)	0	--	--	--	--	0	--	--	--	--	
		Crop insurance	20	0.4	3.5	-0.050	1.2	20	0.8	4.8	-0.022	1.8	
		IST	33	1.0	4.5	-0.049	2.1	39	0.9	4.5	-0.017	1.9	
	Rotation	None	0	--	--	--	--	0	--	--	--	--	
		Crop insurance	17	-0.2	5.4	-0.070	1.5	17	-1.0	6.2	-0.030	1.3	
		ISA	23	1.0	6.8	-0.072	2.8	26	0.3	6.6	-0.026	2.2	
	Basic	Monoculture	None	91	1.0	1.0	-0.003	1.0	91	1.0	1.0	-0.006	1.0
			Crop insurance	111	0.9	1.5	-0.015	1.0	111	1.0	1.7	-0.006	1.2
			IST	117	1.0	1.9	-0.016	1.3	124	1.0	2.0	-0.006	1.3
	Rotation	None	91	0.9	1.6	-0.018	1.1	91	0.7	1.6	-0.007	1.0	
		Crop insurance	108	0.8	1.7	-0.016	1.1	108	0.7	1.8	-0.007	1.0	
		IST	107	1.0	2.1	-0.017	1.3	111	0.8	2.2	-0.007	1.2	
Basic + Greening	Monoculture ¹	None	138	1.0	1.0	-0.012	1.0	138	1.0	1.0	-0.004	1.0	
		Crop insurance	158	0.9	1.3	-0.011	1.0	158	1.0	1.5	-0.005	1.1	
		IST	161	1.0	1.6	-0.012	1.2	168	1.0	1.7	-0.005	1.2	
	Rotation	None	138	0.9	1.4	-0.013	1.1	138	0.8	1.4	-0.005	1.0	
		Crop insurance	155	0.9	1.5	-0.012	1.1	155	0.8	1.6	-0.005	1.0	
		IST	151	1.0	1.7	-0.013	1.2	155	0.9	1.8	-0.005	1.1	

¹ Only if farms area lower than 30 ha

² PE: Public expenditure (insurance premium + CAP direct payments) (€ ha⁻¹ yr⁻¹)

³ EI: Δ Expected income /PC, CV: Δ Coefficient of Variation /PC, CE_{0.4}: Δ Certainty equivalent for a Coefficient of Relative Risk Aversion of 0.4/PC and CE_{0.6}: Δ Certainty equivalent for a Coefficient of Relative Risk Aversion of 0.6/PC, all with respect to Ref 1 are referred to alternative 1, no CAP beneficiary, monoculture and no protection system.

Farmer's profitability and low income were higher in the *high and volatile prices* scenario. For this scenario, the income standard deviation was higher but the coefficient of variation was lower than for the *low and stable prices* scenario. The efficiency of protection tools in decreasing income variability was higher in the *low and stable prices* scenario (between 12 and 85 %) than in the *high and volatile prices* scenario (between 39 and 70 %). Higher decreases were found for the riskier alternatives, that is, no CAP beneficiaries practicing monoculture (Figure 5).

Figure 5. Decrease in the income coefficient of variation (%) with respect to a reference (No CAP, monoculture and no protection tool) for the beneficiaries and no beneficiaries of CAP direct payments, meeting and not meeting greening requirements, practising monoculture or a rotation and using crop insurance, the Income Stabilisation Tool (IST) or no protection option for (A) the *low and stable prices* and (B) *high and volatile prices* scenarios.

6. Discussion

Direct payments were the most effective measure in terms of increasing farms expected and lower incomes. This is in agreement with conclusions drawn about the income effects of direct support in the post 2003 scheme, as the one performed by Agrosynergie (2011). Such evaluation was performed at regional scale for the EU-27 covering macroeconomics, based on regional data, and microeconomics, on individual farm data for seven types of farming, economic farm size, farm location and types of farm organization. Agrosynergie

(2011) concludes that decoupled payments provide a positive and robust contribution to increasing and stabilizing income.

Farms receiving direct payments were significantly more profitable when meeting greening requirements. Therefore, direct payments also have a clear effect of lowering farm risk, even when such measure is not directly intended to support farm risk management (as it increases farm expected income), shortening the left-tail of the income distribution; in the case of the CAP 2014-2020, also indirectly promoting the adoption of risk management measures, namely crop diversification, to meet greening requirements.

Crop diversification is a requirement for being eligible to greening payments. It does not require any direct public support but affords a significant decrease in income variability. Morales et al. (2008) found farm diversification to be a clear substitute to insurance and futures and option markets, being diversified farmers those with a lesser need to contract such risk sharing tools. Similarly, Santeramo et al., (2016) found insurance demand to be negatively correlated with crop diversification, and concluded that it constitutes itself a form of insurance. Similar conclusions are drawn in this work, where farms diversifying had lower income variability, therefore having a lower need to contract protection tools.

Contracting protection tools contributed significantly to decreasing income's coefficient of variation. This result is in agreement with previous studies on the potential effects of income stabilization tools on income stability in the sense that it contributes to decreasing income variability (Finger and El Benni, 2014; Mary et al. 2013). Mary et al. (2013) evaluated the potential implementation of the IST introduced by the CAP 2014-2020 using a farm household model calibrated to French cereal farms and found a high efficiency of the tool, being the decrease in income coefficient of variation of more than 35 %. In this work, decreases in income coefficient of variation when using the IST were between 73 and 85 % in the *low and stable prices* scenario and between 57 and 70 % in the *high and volatile prices* scenario (Figure 5). The higher effectiveness in increasing farm income stability of the IST with respect to crop insurance is due to the direct protection offered by IST against low incomes. Instead, crop insurance only protects against yield losses.

Individual decisions on farm strategies as practicing crop rotation or contracting risk management tools are complementary to CAP direct payments for achieving the CAP objectives of stabilizing farmers' income. This suggests that incentives to contracting protection tools as in the rural development regulation of the CAP reform 2014-2020 are a

good compromise for both improving farm resilience to yield and price variability and restraining public expenditure. Likewise, sharing risk management responsibility with farmers, and conditioning aids to the performance of risk reducing measures (as diversification, crop insurance or an IST), even when supporting such measures (crop insurance premiums and IST indemnities), are essential to develop responsibility and involvement of farmers and other stakeholders (Cordier, 2014).

Public support measures, and the corresponding expenditure, might behave as compensations or as incentives. In the first case, public measures as direct payments aim at compensating market failures or at protecting strategic activities with difficulties to survive due to high opportunity costs or to competitive disadvantages. In the second case, measures aim at incentivizing the adoption of certain practices, thereby driving agricultural activity towards the objective stated in a particular policy. Subsidies to protection tools (either for crop insurance premiums or to IST indemnities) and cross-compliance of direct payments are under this category. The success of CAP measures towards reducing farm income risk will be related to the level of adoption of such practices. It is then relevant to question whether farmers are willing to adopt protection tools or not, which is to say whether public expenditure actually incentivizes farmers to adopt the available risk management strategies. The certainty equivalent is an indicator of the expected willingness of the farmers to adopt the proposed tools, and therefore, the increase in certainty equivalent per public euro invested in subsidizing such tools might be interpreted as the efficiency of public expenditure in incentivizing such behaviour in farmers.

Cereal insurance in Spain has a relatively high penetration rate, reaching 63.9 % in 2015 (Machetti Bermejo, 2015). This supports the results suggesting that crop insurance always resulted in higher certainty equivalent than not contracting any protection system, especially for more risk averse farmers. As it has not yet implemented in this region, no observed data is available to contrast the results obtained for IST. However, results suggest that farmers might be expected to be more willing to contract IST than not contracting any protection tool and even than contracting crop insurance. Therefore, the implementation of IST schemes could derive in a decrease in insurance demand.

The robustness of these protection systems is highly dependent on the accuracy with which the insurer quantifies the actual yield and overall risk, and additionally the actual annual costs, in the case of the IST. However, records on individual actual yields and input costs

are solely available to the farmer, resulting in an important source of information asymmetries (moral hazard and adverse selection). A number of methodologies for yield and costs estimation is being developed and implemented in order to avoid moral hazard in crop, revenue and income insurances. That is the case of indexes and crop models for yield assessment (Castañeda-Vera et al., 2015; Escribano Rodríguez et al., 2014) or scoring grids fixed by experts for costs estimation (Bardají and Garrido, 2016).

Contrary to the view of direct payments as behaving as risk preventive measures, some authors suggest a possible negative effect of farm vulnerability, becoming sometimes in risk-enhancing tools or even promoting moral hazard. El Benni et al. (2012) suggests that Swiss agricultural policy evolution from market-based support to direct payments has decreased farm income variability, behaving direct payments as insurance for farmers making them more willing to take risk from crop production. A similar result was found by Enjolras et al., (2014) for French farmers. This might question the efficiency of such structural policies regarding risk management.

A differential effect of CAP measures might be expected in different farm types and location. Conclusions of this work are therefore valid for the area and farming strategies considered, but might change for other regions with different climate (yield variability) and cultivated crops. A wider study performing a similar analysis at a national and regional scale, including common crops and farm strategies and using historical or simulated data for yields and prices, would be useful to evaluate the equilibrium of the efficiency of horizontal measures such as those considered within the CAP 2014-2020 in different European countries and regions.

7. Conclusions

Direct payments were the most effective measure in increasing farm expected and lower incomes and income stability, especially when meeting greening requirements. Crop diversification does not require any direct public support but provides itself a significant decrease in income variability.

The IST was found to be highly efficient in decreasing the income coefficient of variation especially in the scenario where prices were lower and more stable and for the riskier

alternatives (no CAP beneficiary and practising monoculture); thus for no CAP beneficiaries practicing monoculture. The effectiveness in increasing certainty equivalent was higher for the IST than for crop insurance. Therefore, the implementation of IST schemes could result in a decrease in insurance demand, as the instruments are substitutes.

Subsidized protection tools are a good compromise for both improving farm resilience to yield and price variability and to restrain public expenditure as they achieve high level of income stability to a much lower cost than subsidizing direct payments.

The relative efficiency between risk management instruments remained unchanged in both prices scenarios. This suggests the validity of the conclusions of this paper even under the great uncertainty of rapidly changing economic and climatic environment that characterized agricultural sector performance in the last decade.

8. Acknowledgement

This research was funded by the project MULCLIVAR, from the Spanish Ministerio de Economía y Competitividad (MINECO) CGL2012-38923-C02-02, A. Castañeda-Vera has a PhD grant from UPM. We thank Isabel Bardají (Universidad Politécnica de Madrid) for the clarifying discussions on public support to risk management tools in the EU and Spain.

Chapter 5

5. Crop insurance demand in wheat production: focusing on yield gaps and asymmetric information

A. Castañeda-Vera^{1,2}, A. Saa^{1,3}, I. Mínguez^{1,2} and A. Garrido¹

¹ Research Centre for the Management of Agricultural and Environmental Risks (CEIGRAM),
Universidad Politécnica de Madrid (UPM), Spain

² Agricultural Systems Group, Dpto. Producción Agraria, E.T.S. Ingenieros Agrónomos,
Universidad Politécnica de Madrid, Avda. Complutense s/n, 28040 Madrid, Spain

³ Resources Valuation Group, Dpto. Producción Agraria, E.T.S. Ingenieros Agrónomos,
Universidad Politécnica de Madrid, Avda. Complutense s/n, 28040 Madrid, Spain

Agricultural Journal of Agronomy (under review)

Abstract

Analysis of yield gaps and asymmetric information were conducted in the context of crop insurance to evaluate their influence on insurance demand by Spanish wheat producers. The analysis includes simulated yield using a validated crop model, CERES-Wheat previously selected among others, whose suitability to estimate actual risk when no historical data are available was assessed.. Results suggest that the accuracy in setting the insured yield is decisive in farmers' willingness to contract crop insurance under the wider coverage. Historical insurance data, when available, provide a more robust technical basis to evaluate and calibrate insurance parameters than simulated data, using crop models. Nevertheless, the use of crop models might be useful in designing new insurance packages when no historical data is available or to evaluate scenarios of expected changes. In that case, it is suggested that yield gaps be estimated and considered when using simulated attainable yields.

Keywords: risk management, rainfed wheat, crop insurance penetration rate, crop models, Spain.

1. Introduction

The uptake of agricultural insurance is slow because, by definition, it provides financial compensation only for infrequent events, which discourage farmers to pay the premiums every year when indemnities are seldom received. Moreover, agricultural insurance faces a number of market imperfections that end up in market failures, compromising the system's sustainability or the branch when it is offered privately. This justifies governments' intervention in promoting of insurance demand through the implementation of public incentives, such as premium subsidies, and a significant degree of market intervention, guidance and overseeing (Mahul and Stutley, 2010).

Asymmetric information is a common cause of insurance market failure. It occurs when insured farmers have more information than the insurer about their actual risk and behaviour. Asymmetry of information commonly results in two behavioural responses: moral hazard and adverse selection. Moral hazard occurs when farmer's expected indemnity under insurance is larger than under optimal uninsured conditions (Coble et al., 1997). This means that farmers modify their behaviour after contracting insurance in order to increase the probability of being indemnified (Goodwin, 1994) or reduce the effort to escape risk once they are covered for it. Adverse selection happens when high-risk farmers contract insurance in greater numbers than low-risk farmers; and the insurer, being unable to differentiate them, is forced to average out the risks of both types in calculating the premium. In these situations, the insurer does not have adequate information to calculate the unbiased probability and severity of claims. The result can be unbalanced loss ratios, thereby affecting the actuarial robustness and sustainability of entire insurance systems. Optimization of public interventions requires evaluations *-ex ante-* to support decision making, and *-ex post-* to assess impacts, the latter using methods such as indicators to quantify the degree of market failure and intervention effects.

Yield gaps are defined in agronomy as the differences between attainable yields based on climate and soil conditions and actual farmers' yields; they are usually defined with reference to some specified spatial and temporal scale (Lobell et al., 2009). Quantifying such gaps is useful to identify suboptimal crop management and the opportunity for the agronomic improvement of productivity (e.g. Affholder et al., 2012; Lobell et al., 2009; van Wart et al., 2013). Yield gap has been proven to be a valuable concept for assessing and understanding the ecological possibilities to meet food demand for an increasing population (van Ittersum et al., 2013). Nevertheless, the use of this concept is not limited to the biophysical aspects of agronomy, but may also be used

to support research on socioeconomic aspects of agricultural production. Examples are seen in the limits to technology adoption due to either technical constraints or for economic issues arising from market conditions (Godfray et al., 2010). In the context of crop insurance, yield gaps might be also defined as the difference between insured and actual yields. Such a gap indicates either a deficiency in the insurer's knowledge of the cropping system or advantages taken by the farmer based on wider information. Either situation is a case of asymmetric information.

Problems of asymmetric information are related to the difficulties associated with measuring actual farmers' risks; and lack of data on actual yields is the main cause of this asymmetry (Mahul and Stutley, 2010). Crop models are increasingly used in decision-making, for example, for policy shaping and analysis (e.g. CAP greening measures), farms' consultancies (e.g. Hunt et al., 2006), and risk management with early warning systems (e.g. Basso et al., 2013). Crop models might also be valuable to generate data on expected yields for risk analyses and failure studies in the crop insurance market when no historical data are available.

This paper presents analyses of yield gaps and asymmetric information in the context of crop insurance, and the possible influence of those in the decision of a farmer to contract insurance. A novel method to measure asymmetric information was designed based on yield gaps and applied to wheat insurance in the region of Castilla y León (northern central Spain). The analysis includes simulated yield data using the CERES-Wheat crop model in order to assess the suitability of such models to estimate actual yield risk when no historical data are available. The aims of the study were to (i) quantify the gap between actual wheat yields and the yield data managed by the insurer, (ii) obtain indicators providing evidence of potential asymmetric information in wheat insurance, (iii) explore asymmetric information and yield gap-related factors affecting the decision of farmers to contract insurance, and (iv) evaluate the use of a crop model within the context of crop insurance.

2. Materials and methods

2.1. Winter cereal insurance

In Spain, multi-peril insurance for cereal farmers has been operating since 1982. The database for this study pertains to crop seasons 2010-11 and 2011-12 and includes four insurance modules (P, 1, 2 and S) that were available to cereal farmers. Module S has not been offered since 2013. The

modules differ based on the coverage, on individual fields or the whole farm, rainfed or irrigated, and in the way indemnities are calculated when a loss occurs, and on the maximum insurable yield guaranteed (Table 1). Two types of risks are defined. Type A risks include hail, fire, crop damage caused by wild fauna, flood and excessive rainfall. Type B risks include adversities impeding crop emergence (no-emergence risks) or limiting crop growth (including drought).

For irrigated crops, losses associated to both risk types A and B are indemnified per plot if the final yield is lower than the insured yield, the latter having been freely set by the farmer. Under rainfed conditions, however, losses associated to type B risks can be indemnified per plot (Module 2) or per farm (Modules 1 and S). In the latter case the farmer is indemnified only if the average yield of all farm plots is lower than the insured yield. Low yields in one farm plot may be compensated by high yields on another plot. Module S limited insured yield to a maximum officially assigned by municipalities by the Spanish Agency of Agricultural Insurance (ENESA, initials in Spanish). Modules 1 and 2 limit insured yield to a maximum insurable yield calculated for individual farmers based on their insurance history records. Lastly, farmers might select a deductible, most of them opting for the lowest of 30%.

To simplify the analysis, the four modules (P, 1, 2 and S) were combined into two options based on the risks covered. The first option, called *Basic*, comprises module P that guarantees farmers' production with coverage for type A risks. The second option, called *Extended*, that includes modules 1, 2 and S, provides coverage for risks of both type A and type B. Farmers choosing the *Extended* option may increase insured yield at midseason with a complementary insurance payment if yield expectation exceeds that insured before sowing.

Crop insurance premiums are proportional to insured yields and vary depending on the contracted module and deductible. They are subsidized by the national and regional governments. The subsidy varies depending on the option contracted (base), the farmer's characteristics (age, gender and others), and whether the contract is a renewal or is contracted through a farmer association. The lowest base subsidy is for module P (option *Basic*) increasing for modules S, 2 and 1 (option *Extended*) as 1, 8, 17 and 22% of the premium cost. Additional subsidies based on the farmer's characteristics (young farmers, gender and others), and whether it is contracted collectively through a farmer association are also different for different modules, in such a way that Module P (option *Basic*) has the lower maximum assignable subsidy followed by Module S, Module 2, and lastly Module 1 (option *Extended*) with 16, 33, 52 and 57 % of the premium cost (Table 1).

Table 1. Wheat insurance options: modules and characteristics.

<i>Option</i> ¹	<i>Module</i> ²	<i>Risks</i> ³	<i>Regime</i>	<i>Complement</i> ⁴	<i>for rainfed wheat & type B risks</i>		<i>Subsidies (%)</i>					
					<i>Loss</i>	<i>Max</i>	<i>Young/</i>		<i>Certified</i>			
					<i>indemnity</i> ⁵	<i>insurable yield</i> ⁶	<i>Base</i>	<i>women</i>	<i>seeds</i> ⁷	<i>Collective</i>	<i>Renewal</i>	<i>Up to...</i>
Basic	P	A	Rainfed and irrigated	No	-	-	1	6	-	5	4	16
Extended	S	A and B	Rainfed	Yes	Per farm	Zonal	8	16	-	5	4	33
	2	A and B	Rainfed and irrigated	Yes	Per plot	Yields history	17	16	5	5	9	52
	1	A and B	Rainfed and irrigated	Yes	Per farm	Yields history	22	16	5	5	9	57

¹ Option *Basic* corresponds to Module P, while Module 1, Module 2 and Module S were aggregated in option *Extended*, as the same risks are covered in the three of them.

² Modules P, S, 2 and 1 as denominated in the arable crops insurance, they were listed according to the base subsidies from the lowest to the highest.

³ Risk covered, types: *A* - Hail, fire, wild fauna, flood and excessive rainfall; *B* - No emergence and any other climatic adversity.

⁴ Farmers choosing the *Extended* option might increase the insured yield at midseason if their yield expectative is higher than the one insured before sowing.

⁵ *Per farm*: the farmer is indemnified only if the average final yield of the whole farm is lower than the insured yield; *Per plot*: the farmer is indemnified for each plot in which the final yield is lower than the insured yield.

⁶ *Zonal*: maximum insurable yield officially established on a geographic basis; *yields history*: maximum insurable yield adjusted based on individual farmers registered past yields.

⁷ Subsidies associated to practices that reduce risk. In the case of cereals, this subsidy is applicable for farmers using certified seeds.

The proportion of young farmers, women, associated farmers and other characteristics, giving rise to additional subsidies, is still low. In this work, it was assumed that farmers of different characteristics (young or not, men and women, associated farmers and farmers renewing their contracts) are distributed similarly in all counties in the study region.

2.2. Study site

Spain is the fifth largest producer of cereals in Europe, after France, Germany, Poland, and UK, with harvested production that ranged between 14 and 25.5 million Mg (6.2 Mha of crop) between 2000 and 2013 (Eurostat, 2013). About 47% of the cultivated land dedicated to arable crops (including cereal) is located in Castilla y León region, with about 2.9 Mha (ESYRCE, 2013).

Figure 1. Winter wheat insurance demand per county and insured wheat area (1000 ha) per insurance module in 2011 and 2012.

Castilla y León is located in northern central Spain (Figure 1). Climate is continental Mediterranean with warm and dry summers and cold and wet winters. It is a high plateau (830 masl average) around the Douro river basin surrounded by mountain ranges. There is a clear gradient of temperature and rainfall from North-East (cooler and wetter) to South-West (warmer and drier).

The most commonly cultivated cereals are barley (*Hordeum vulgare* sp.) and bread wheat (*Triticum aestivum* sp.), with 54% and 30% of total regional production, respectively. Wheat is mainly grown without irrigation (JCyL, 2014) so that yields respond to the high rainfall variability (Figure 2). Average rainfed yields during 2000–2013 varied between 2.0 and 4.5 Mg ha⁻¹ depending on soils and climate (JCyL, 2014). Crops of winter wheat are sown during October–December and spring wheat during February–March, in both cases depending on rainfall and soil moisture. Wheat is harvested between late June and mid-August. The most sensitive phenological stages to meteorological adversities are flowering (April–May, late frosts, drought and heat stress) and grain filling (June–July, heat and drought).

Figure 2. Annual rainfall corresponding to Valladolid (1960-2010). Source: AEMET.

2.3. Actual, expected, insured and water limited yields

Expected yields

Actual yields (Y_a) were obtained from the ESYRCE database (*Encuesta sobre Superficies y Rendimientos de Cultivos*). This survey has been conducted every year since 1995 by the Ministry of Agriculture in coordination with local administrations. Each year a sample of farms is selected from a 1km x 1km grid (MAGRAMA, 2014). Surveys are performed at the farm level and include information on cultivated area, average yield per crop, crop varieties, management practices, farm infrastructure and others. The original database comprised 121,309 yield observations between 1995 and 2014, distributed across the study region. After aggregation at municipality level, the database for the present analysis contained 7879 yield observations.

Expected yields (Y_{exp}) prior to sowing and for each of the counties were estimated assuming a linear trend from the seasonal adjusted Y_a series for the next year. Y_{exp} includes spatial variability of the mean expected yield in 531 municipalities across the study region, in 2011 and 2012.

Insurance-related yields

Data from all municipalities in Castilla y León with at least one insured farmer in 2011 or 2012 was made available by ENESA for 107,709 insurance policies over two years, contracted in 452 municipalities. Each record includes information on the number of sold policies, total insured production and area, and insured yield per crop insurance module. These records were aggregated by county, year, and insurance module.

The zonal maximum insurable yield (per municipality) for wheat grown under rainfed conditions for each municipality is published every year in the Spanish Official Gazette (BOE). The ones used for this study were published in BOE No 217, 02.09.2013, at the Order AAA/1629/2013. These data were aggregated by county to define the average zonal maximum insurable yield (Y_{insZ}).

Three other insurance-related yields were defined for the insurance options *Basic* and *Extended*, the latter including the complementary insurance of the option, as described in the previous section:

- *YinsB* is the average yield insured in the option *Basic*, the yield expected before sowing that is freely set by the farmer.
- *YinsE* is the average yield insured in the option *Extended* (Module 1, Module 2 and Module S), yield insured by the farmers before sowing and is limited to an individual maximum insurable yield calculated on the basis of the farm's yields and insurance history.
- At midseason, farmers have the option of contracting a complementary insurance policy, in which the insured yield is adjusted according to farmer's yield expectations at that time of season (*YinsEC*). This option is offered to farmers to permit them revising their contracted insured yields in cases, for example, of abnormally high spring rainfall.

YinsZ is defined at county level (as it does not vary from a year to another) and *YinsB*, *YinsE* and *YinsEC* are defined per year and per county.

YinsZ includes spatial variability, including 531 municipalities all along the study region. *YinsB*, *YinsE* and *YinsEC* includes spatial and temporal variability (2 years) of insured yield data aggregated at municipality level using the mean insured yield in each of the insurance options in the same 531 municipalities.

Attainable rainfed yields

Attainable rainfed yields (*Y_w*) were simulated using the crop model CERES–Wheat (Godwin et al., 1989). CERES–Wheat as available in the package Decision Support System for Agrotechnology Transfer (DSSAT) version 4.5 (Hoogenboom et al., 2012; Jones et al., 2003). CERES–Wheat has been widely applied at the regional scale (e.g. Bannayan, Crout, & Hoogenboom, 2003; Zhao et al., 2011) and preferred for the simulation of winter wheat growth at field scale in semi-arid conditions in Spain after a detailed comparison (Castañeda-Vera et al., 2015).

CERES-Wheat was first calibrated and validated using published data from the trials conducted by the Agricultural Research Program of Castilla y León between 2004 and 2010 (Table 2) and following the same approach as in Castañeda-Vera et al., (2015). The experimental sites are widely spread throughout the grain production areas and are representative of its climate and soil variability. The objective of these trials was to monitor the attainable rainfed production of different wheat cultivars. Therefore, they are assumed to have been grown in optimal growing conditions, being water availability the solely limiting factor (depending on rainfall). Crop data belonging to the cultivar Marius (*Triticum aestivum* cv. Marius) were selected to calibrate the crop

model. This cultivar is commonly used by farmers in this region; moreover, it has been used as a test cultivar, and thus it was sown in each of the trials every year. Crop data was limited due to the different purpose of the trials with respect to what it concerns to this work. Crop development data such as emergence, anthesis and physiological maturity dates, and yield were available (Table 2). Sowing dates and plant density were also provided for all the sites.

Soil texture and depth maps were obtained from the ITACyL (Agro-Technological Institute in Castilla y León) soil portal (<http://suelos.itacyl.es/descarga>) to simulated rainfed yields. Three soil textures (sandy, loam and clayey, Figure 3-A) and two depth-types (1.0 m and 0.3 m, Figure 3-B) were defined (Table 3). Three maps with polygons defining depth, soil texture and climate zones were built using ESRI ArcMap 10.0 GIS and combined into a single map. Different climate zones were selected assuming that accumulated rainfall and evapotranspiration are main driving variables for crop growth under rainfed conditions in such a semi-arid region.

Accumulated rainfall and evapotranspiration are aggregated in Thornthwaite's aridity index (Thornthwaite, 1948). The map based on Thornthwaite's aridity index in the Agroclimatic Atlas of Castilla y León (Álvarez-Arias et al., 2013) was used to assign a climate zone to each sub-area (Figure 3-C). Weather data assigned to each climate zone was obtained from the stations belonging to the System of Agroclimatic Information for Irrigated crops (SiAR) with data between 2000 and 2013 (Table 4). The new map contained polygons assigned with a number defining a single combination of soil texture-depth-climate zone (Figure 3-D).

Models' statistical performance was evaluated using the root mean square error (*RMSE*). This statistic measures the distance between the simulated and the observed values in absolute terms (Eq 1):

$$RMSE = \sqrt{\frac{\sum_i^n (S_i - O_i)^2}{n}} \quad [\text{Eq 1}]$$

where S_i and O_i are the simulated and observed yields, respectively, and n is the number of observations used. Larger values of *RMSE* indicate decreasing accuracy.

Table 2. Trials seasons, anthesis and physiological maturity (days after sowing, DAS), and grain yield (Mg ha⁻¹) per location.

Location	Season	Anthesis (DAS)	Ph. Maturity (DAS)	Yield (Mg ha⁻¹)
Barca (Soria)	2005-06 / 2008-09	179	213	1.7
Becerril (Palencia)	2004-05 / 2005-06 / 2007-08 / 2008-09 / 2009-10	185	—	2.9
Castrillo de la Guareña (Zamora)	2005-06	164	197	2.1
Fuentepiñel (Segovia)	2004-05 / 2005-06 / 2007-08 / 2008-09 / 2009-10	189	222	5.8
Fuentes de Año (Palencia)	2005-06 / 2007-08 / 2008-09	180	221	2.5
Osorno (Palencia)	2005-06 / 2007-08	182	—	2.6
Palencia de Negrilla (Salamanca)	2004-05 / 2005-06	184	—	2.6
Soto de Cerrato (Palencia)	2008-09	190		1.5
Tobar (Burgos)	2004-05 / 2005-06	189	216	4.2

Table 3. Soil data used for the simulation of water limited yields using CERES-Wheat.

Soil type	Depth (m)	Clay (%)	Silt (%)	PWP ¹ (cm ³ cm ⁻³)	FC ¹ (cm ³ cm ⁻³)	SAT ¹ (cm ³ cm ⁻³)	Ksat ¹ (cm h ⁻¹)	Bulk density (g cm ⁻³)	Coarse fraction (%)	Organic matter (%)
Sandy	0.05	19	7	0.131	0.197	0.369	2.59	1.59	2	0.45
	0.30	19	7	0.123	0.184	0.352	2.59	1.61	5	0.3
	1.00	21	8	0.139	0.207	0.374	0.43	1.6	0	0.3
Loam	0.05	25	31	0.186	0.327	0.457	1.32	1.32	2	1.44
	0.30	25	31	0.186	0.327	0.457	1.32	1.32	2	1.44
	1.00	30	24	0.187	0.297	0.411	0.43	1.44	4	0.7
Clayey	0.05	40	32	0.248	0.394	0.453	0.06	1.23	9	1.69
	0.30	40	32	0.248	0.394	0.453	0.06	1.23	9	1.69
	1.00	52	31	0.289	0.425	0.456	0.06	1.28	6	0.8

¹ PWP: Permanent Wilting Point; FC: Field capacity; SAT: Saturation; Ksat: Hydraulic conductivity at saturation.

Table 4. Minimum, average and maximum temperature and annual accumulated rainfall per weather zone.

Weather	Stations	Tmin (°C)	Tmean (°C)	Tmax (°C)	Rainfall (mm)
1	Torrecilla de la Orden (Va) Navas de Arévalo (Áv) Tordesillas (Va) Villalarbo (Za)	5.5	7.4	20.3	303
2	Fuentes de Nava (Pa) Almazán (So) Gomezserracín (Sg) San Esteban de Gormaz (So) Villoldo (Pa) Nava de la Asunción (Sg) Encina de Esgueva (Va)	4.5	7.5	19.3	409
3	Ciudad Rodrigo (Sa) Herrera de Pisuergra (Pa) Zotes del Páramo (Le) Villaluenga de la Vega (Pa)	5.2	6.9	19.0	427
4	Cerraledo (Le) Mansilla Mayor (Le) Valle de Valdelucio (Bu)	5.1	6.9	18.9	556

The model was run for each combination of soil texture-depth-climate zone in the region. CERES-Wheat simulates wheat biomass production in dry matter basis. Simulated yields were converted into fresh matter assuming 13% humidity in order to make them comparable with actual, expected and insured yields. Average simulated rainfed yields given in fresh matter basis (Y_w) were assigned to the corresponding polygon in the map.

Y_w includes spatial and temporal variability of simulated attainable rainfed yield data aggregated at municipality level using the mean yield in 531 municipalities all along the study region and per year (2000-2013).

2.4. Yield gap analysis

Five yield gaps ($GapZ$, $GapB$, $GapE$, $GapEC$, $GapW$) were defined by the following yield differences:

$$GapZ = Y_{insZ} - Y_{exp} \quad [Eq\ 2]$$

$$GapB = Y_{insB} - Y_{exp} \quad [Eq\ 3]$$

$$GapE = Y_{insE} - Y_{exp} \quad [Eq\ 4]$$

$$GapEC = Y_{insEC} - Y_{exp} \quad [Eq\ 5]$$

$$GapW = Y_w - Y_{exp} \quad [Eq\ 6]$$

where Y_{exp} is the county average expected yield; Y_{insZ} is the county average zonal maximum insurable yield; Y_{insB} is the county average insured yield in the option *Basic* per year, and Y_{insE} , Y_{insEC} are the county average insured yield each year in the option *Extended* before and after contracting the complementary insurance, respectively (Table 1). Lastly, Y_w is the county average simulated attainable rainfed yield.

Y_{insZ} is considered to have a perfect fit when equal to Y_a ; in that case, the insurance system would be considered to have a perfect knowledge of the actual average yield in the region. Therefore, $GapZ$ was assessed as a measure of the rightness of the insurance parameter Y_{insZ} in such a way that the lower $GapZ$, the better the parameter fit. $GapB$, $GapE$ and $GapEC$ indicate the distance between insured and expected yields; they were evaluated as preliminary indicators of asymmetric information for each of the insurance options, reflecting farmer's information advantage with respect to the insurer. $GapE$ and $GapEC$ were also contrasted in order to evaluate how farmers increase insured yield when uncertainty decreases (insured before sowing vs. insured at midseason). Lastly, $GapW$ provides an insight into how similar or dissimilar simulated rainfed yields are from expected yields. $GapW$ helps evaluating the appropriateness of using crop models simulated data in crop insurance parameterization when limited observed data are available.

Figure 3. (A) Soil texture, (B) soil depth, (C) Thornthwaite aridity index and (D) soil texture-depth-climate zones in the study region (Castilla y León, northern central Spain). Adapted from the Agroclimatic Atlas of Castilla-y-León (Álvarez-Arias et al., 2013).

2.5. Asymmetric information

An asymmetric information indicator was obtained as the difference between the probability of indemnity based on the risk expected by the insurer, in which possible asymmetric information is assumed, and the one based on the actual risk, in which no asymmetric information is assumed. The procedure to calculate the asymmetric information indicator is illustrated in Figure 4. The probability of indemnity was calculated as the cumulative probability of the actual yields being lower than a threshold yield (y_{thres}) (Eq 7, Figure 4-A).

$$\text{Probability of indemnity} = \int_0^{y_{thres}} f(y) dy \quad [\text{Eq 7}]$$

A probability density function of actual yields, $f(Ya)$, left-truncated at zero, was fitted for each of the counties using the software package @-Risk (Palisade Corporation, 2011), selecting the distribution with the lowest Akaike Information Criterion (AIC) statistic. This threshold yield (y_{thres}) was defined as the insured yield (Y_{insB} , Y_{insE} or Y_{insEC}) after subtracting a 30% deductible, as this is the most

common deductible level selected by the farmers; and therefore, the risk assumed by the insurer. To compute the probability of indemnity based on the actual risk (non-asymmetric information), y_{thres} was calculated from the average actual yield $f(Ya)$ ($\bar{Y}_a$) (Figure 4-B). For the insurance options *Basic* and *Extended*, y_{thres} was calculated from the average insured yields (Y_{insB} , Y_{insEC}) (shown in Figure 4-C for Y_{insEC}).

Lastly, the asymmetric information indicator was defined for the insurance options *Basic* ($AsymB$) and *Extended* before ($AsymE$) and after contracting the complementary insurance ($AsymEC$), as shown in Eq 8, Eq 9 and Eq 10 and represented in Figure 4-D.

$$AsymB = \int_0^{0.7 \times Y_{insB}} f(Ya) dy - \int_0^{0.7 \times \bar{Y}_a} f(Ya) dy \quad [Eq 8]$$

$$AsymE = \int_0^{0.7 \times Y_{insE}} f(Ya) dy - \int_0^{0.7 \times \bar{Y}_a} f(Ya) dy \quad [Eq 9]$$

$$AsymEC = \int_0^{0.7 \times Y_{insEC}} f(Ya) dy - \int_0^{0.7 \times \bar{Y}_a} f(Ya) dy \quad [Eq 10]$$

where $AsymB$ and $AsymEC$ are the differences between the probability of indemnity based on the insured yields ($y_{thres} = 0.7 \times Y_{insB}$; and $y_{thres} = 0.7 \times Y_{insEC}$) and the one based on observed actual yields ($y_{thres} = 0.7 \times \bar{Y}_a$).

Lastly, the asymmetric information indicator was defined for the insurance options *Basic* ($AsymB$) and *Extended* before ($AsymE$) and after contracting the complementary insurance ($AsymEC$), as shown in Eq 8, Eq 9 and Eq 10 and represented in Figure 4-D.

$$AsymB = \int_0^{0.7 \times Y_{insB}} f(Ya) dy - \int_0^{0.7 \times \bar{Y}_a} f(Ya) dy \quad [Eq 8]$$

$$AsymE = \int_0^{0.7 \times Y_{insE}} f(Ya) dy - \int_0^{0.7 \times \bar{Y}_a} f(Ya) dy \quad [Eq 9]$$

$$AsymEC = \int_0^{0.7 \times Y_{insEC}} f(Ya) dy - \int_0^{0.7 \times \bar{Y}_a} f(Ya) dy \quad [Eq 10]$$

where $AsymB$ and $AsymEC$ are the differences between the probability of indemnity based on the insured yields ($y_{thres} = 0.7 \times Y_{insB}$; and $y_{thres} = 0.7 \times Y_{insEC}$) and the one based on observed actual yields ($y_{thres} = 0.7 \times \bar{Y}_a$).

Figure 4. Procedure to calculate the asymmetric information indicator: (A) the probability of indemnity calculated as the cumulative probability of yield being lower than a threshold yield (y_{thres}), (B) the probability of indemnity with no asymmetric information where y_{thres} is the average actual yield after subtracting a 30% deductible ($0.7 \times \bar{Y}_a$), (C) the probability of indemnity with asymmetric information for the insurance option Extended including the complementary insurance (AsymEC), where y_{thres} is the average insured yield in insurance option EC after subtracting a 30% deductible ($0.7 \times \bar{Y}_{insEC}$), and lastly, (D) the asymmetric information indicator calculated as the difference between (C) and (B).

2.6. Factors influencing insurance demand

To address wheat production- yields gaps- and asymmetric information- related factors influencing crop insurance demand, a statistical analysis was performed. In this work, insurance demand at the county level is thus an aggregation of all farmers' individual decisions about whether insuring or not. Such decisions are taken before sowing, though farmers are allowed to modify the insured yield at mid-season, only if the contract option *Extended* (Y_{insEC}).

A factor analysis was performed in order to understand variables relationship. Moreover, it was useful to constrain explanatory information to non-correlated variables to be used later on in a regression model exploring the influence on insurance demand. Variables used within the analysis are shown in Table 5.

For that, in a first step, the suitability of data for factor analysis was evaluated using the Kaiser-Meyer-Olkin (KMO) test. The value of KMO was 0.743, enough to conclude that data was suitable for such analysis (Kaiser, 1974). Second, the number of extracted factors was determined using the eigenvalues. We found that restraining factors with eigenvalue higher than 1 was too restrictive for such a model with a high number of variables. Therefore, we used the alternative approach of creating a scree plot (graphing the eigenvalues of all factors listing them in decreasing order of their eigenvalue) and restraining the number of factors above the inflection point (Cattell, 1966). The resulting number of factors was 4. Lastly, factor loadings were computed by an iterated principal factors' algorithm until convergence and rotated using the varimax rotation method.

Factor analysis allowed selecting a single variable from each group (the one with the higher loading). Selected variables were used as independent variables in two linear regression models fitted to investigate the impact of yield gaps and county wheat production characteristics on the decision of farmers in a given county and year to insure their wheat production ($Insured_{wheat}$) and the proportion of insured area under the option basic ($Insured_B / Insured_{wheat}$). The variable $Insured_{wheat}$ was calculated as the fraction of total wheat cultivated area (JCyL, 2014), that was actually insured in any of the insurance modules, per county and year. It thus represents an index of insurance penetration measured at the county level. Models were fitted with Stata v12 (StataCorp, 2011).

Table 5. Notation, description and descriptive statistics of yields and yield gaps used within the yield gap analysis and the asymmetric information and insurance demand model for winter wheat in Castilla y León. Data is aggregated from the original data base of observations per county.

Variable	Definition	N	mean	min	p10	p50	p90	max
$Area_{wheat}$ (1000 ha)¹	Area cultivated with wheat in a given year and municipality	79	0.50	0.03	0.17	0.43	0.99	1.43
$Rainfed$ (fraction)	Fraction of the wheat area cultivated under rainfed conditions	79	0.91	0.30	0.78	0.96	1.00	1.00
Y_a (Mg ha ⁻¹)	Actual rainfed yield	79	2.76	1.29	2.18	2.75	3.47	4.83
CV_{Y_a} (-)	Coefficient of variation of actual yields under rainfed conditions (1995-2013)	79	0.09	0.02	0.03	0.06	0.18	0.51
$Skewness_{Y_a}$ (-)	Skewness of actual yields under rainfed conditions (1995-2013)	79	0.08	-1.61	-1.11	-0.13	1.94	2.61
Y_{exp} (Mg ha⁻¹)	Expected rainfed yield	79	2.77	1.28	2.20	2.74	3.45	4.80
Y_{insZ} (Mg ha ⁻¹)	Zonal maximum insurable yield	79	1.96	1.10	1.48	2.00	2.38	2.70
$GapZ$ (Mg ha⁻¹)	Gap between Y_{insZ} and the expected yield under rainfed	79	-0.85	-3.21	-1.41	-0.80	-0.37	-0.04
Y_w (Mg ha ⁻¹)	Simulated attainable rainfed yield	79	3.78	2.43	2.99	3.80	4.42	4.77
$GapW$ (Mg ha ⁻¹)	Gap between Y_w and Y_{exp}	79	1.01	-1.37	-0.36	1.04	2.13	3.14
Y_{insB} (Mg ha ⁻¹)	Average insured yield in option <i>Basic</i>	76	3.54	2.22	2.89	3.40	4.16	6.61
$GapB$ (Mg ha ⁻¹)	Gap between Y_{insB} and Y_{exp}	76	0.75	-0.76	-0.16	0.67	1.44	4.34
Y_{insE} (Mg ha ⁻¹)	Average insured yield in option <i>Extended</i> before sowing	79	2.22	1.17	1.65	2.23	2.62	2.90
$GapE$ (Mg ha ⁻¹) ¹	Gap between Y_{insE} and Y_{exp}	79	-0.56	-2.59	-1.09	-0.51	-0.03	0.31
Y_{insEC} (Mg ha ⁻¹)	Average insured yield in option <i>Extended</i> including the complementary insurance	79	3.63	1.82	3.29	3.68	3.94	4.11
$GapEC$ (Mg ha ⁻¹)	Gap between Y_{insEC} and Y_{exp}	79	0.86	-1.61	0.41	0.90	1.34	1.68
$Insured_{wheat}$ (-)	Fraction of the area cultivated with wheat insured in a given year and county	79	0.26	0.02	0.17	0.25	0.39	0.52
$Insured_B$ (-)	Fraction of the area cultivated with wheat insured in option <i>Basic</i> in a given year and county	76	0.10	0.00	0.01	0.09	0.19	0.29
$Insured_E$ (-)	Fraction of the area cultivated with wheat insured in option <i>Extended</i> in a given year and county	79	0.17	0.01	0.05	0.15	0.37	0.52
$AsymB$ (-)	Asymmetric information indicator for option <i>Basic</i>	75	0.17	-0.10	-0.03	0.17	0.33	0.74
$AsymE$ (-)	Asymmetric information indicator for option <i>Extended</i>	78	-0.08	-0.26	-0.13	-0.08	-0.01	0.04
$AsymEC$ (-)	Asymmetric information indicator for option <i>Extended</i> including complementary insurance	78	0.22	-0.10	0.07	0.23	0.34	0.50

¹ Variables included in the insurance demand models (in bold).

Table 6. Total and winter wheat cultivated area (1000 ha), total and per farmer winter wheat insured area (1000 ha and ha farmer-1), fraction of winter wheat cultivated in rainfed conditions and the fraction of winter wheat insured per province in Castilla y León. Source: ENESA database and JCyL, (2014).

Province	Cropping area (1000 ha)	WW ¹ area (1000 ha)		Rainfed WW (% WW area)		BASIC				EXTENDED			
						WW insured area (1000 ha)		Insured _{wheat} (%)		WW insured area (1000 ha)		Insured _{wheat} (%)	
		2011	2012	2011	2012	2011	2012	2011	2012	2011	2012	2011	2012
Ávila	187.1	33.9	32.3	92	92	3.2	4.5	10	14	3.1	2.5	9	8
Burgos	617.9	218.6	206.7	96	95	13.6	17.0	6	8	56.1	42.9	26	21
León	302.5	49.8	53.4	41	46	7.4	9.5	15	18	5.1	4.4	10	8
Palencia	465.1	98.5	114.1	80	81	7.9	10.3	8	9	19.2	17.0	20	15
Salamanca	283.8	63.9	64.9	92	94	6.3	6.7	10	10	8.7	9.6	14	15
Segovia	265.9	62.3	66.1	96	96	4.1	5.5	7	8	9.9	10.9	16	16
Soria	346.3	89.3	113.7	97	96	1.1	3.0	1	3	37.4	38.5	42	34
Valladolid	571.2	62.7	72.2	80	80	7.5	13.3	12	18	10.9	7.2	17	10
Zamora	486.9	55.9	61.8	82	83	5.2	8.3	9	13	10.0	7.7	18	12

¹ WW: Winter wheat

² Average in the study region

3. Results

3.1. Wheat insurance demand and county characterization

Figure 1 and Table 6 show the fraction of the total area cultivated with wheat that was insured in 2011 and 2012 ($Insured_{wheat}$ as described in Table 5), and the total area insured in each of the insurance modules ($Area_{wheat}$). The highest insurance demand was found in the eastern counties (Soria province), with up to 60% of the area cultivated with wheat in 2011. Module 2 gathers most of the insured area, with more than 0.3 Mha in both studied years, followed by Module P. Figure 5 shows that insurance demand for each of the options offered to farmers, *Basic* and *Extended*, were negatively correlated and clearly substitutes from the farmers' perspective.

Figure 5. Fraction of the area cultivated with wheat insured under each of the options offered to farmers, *Basic* and *Extended*.

3.2. CERES-Wheat model calibration and validation

The calibrated crop parameters reproduce a wheat cultivar with high vernalization requirements (60 days), and intermediate photoperiod sensitivity (105 % reduction in rate per 10 h drop in photoperiod), a short grain filling period (400 °C-days), and an intermediate phyllochron for

wheat (95°C-days). Parameters related to potential biomass production were 19.5 kernels per unit canopy weight ($\# \text{ g}^{-1}$), 30.5 kg grain weight, and 2.59 g per standard, non-stressed mature tiller, including grain (Table 7). Table 7 shows observed and simulated data and the root mean square error (*RMSE*) for phenology and grain yield for cultivar Marius. For calibration and validation, *RMSE* for anthesis and physiological maturity were 2.6 and 6.5, and 11.5 and 9.9 days, respectively. For grain yield, *RMSE* were 0.4 and 0.9 Mg ha^{-1} .

Table 7. Calibration and validation of CERES-Wheat for anthesis and physiological maturity (days after sowing, DAS), and grain yield (Mg ha^{-1}).

	N	Anthesis (DAS)			Ph. Maturity (DAS)			Grain yield (Mg ha^{-1})		
		Obs	Sim	RMSE	Obs	Sim	RMSE	Obs	Sim	RMSE
Calibration	7	183.4	184.3	2.6	213.0	215.5	11.5	2.5	2.4	0.4
Validation	17	183.8	182.3	6.5	217.9	218.2	9.9	3.5	3.4	0.9

P1V:60 days; P1D: 105 % reduction in rate per 10 h drop in photoperiod; P5:400 °C-days
G1:19.5 $\# \text{ g}^{-1}$; G2: 30.5 kg; G3:2.59 g
PHINT: 95 °C-days

3.3. Yields and yield gap analysis

Figures 5 and 6 show the kernel density estimation of yields and yield gaps including all counties defined within this work. In both cases, they were estimated using Stata v12 and the Epanechnikov kernel. The y-axis has the units and dimensions of the reciprocal of the variable in the x-axis. Thus, density is not measured on a probability scale, and therefore it might exceed 1.

Figure 6 shows the kernel density estimation of the zonal maximum insurable yield (Y_{insZ}), expected yield (Y_{exp}), and water limited yields (Y_w) (Figure 6-A), and the $GapZ$ and $GapW$ gaps, (Figure 6-B). The lowest yields were Y_{insZ} , followed by Y_{exp} and lastly, Y_w , with the mean at 1.96, 2.77 and 3.78 Mg ha^{-1} , respectively (Table 5, Figure 6-A). The largest variability was found for Y_w , followed by Y_{exp} and lastly Y_{insZ} (Figure 6-A). $GapZ$ had the mean at -0.85 Mg ha^{-1} and a low variability (percentiles 10 and 90, being -1.45 and -0.37 Mg ha^{-1} , respectively), while $GapW$ had a mean of 1.01 Mg ha^{-1} and a high variability (percentiles 10 and 90, -0.36 and 2.13 Mg ha^{-1} , respectively) (Table 5 and Figure 6-B).

Figure 6. Kernel density estimation of yields and yield gaps: (A) Zonal maximum insurable yields (solid line, Y_{insZ}), expected yields (dashed line, Y_{exp}) and water-limited yields (dotted line, Y_w); and (B) maximum insurable yields to expected yields gaps (solid line, $GapZ$) and water-limited to expected yields gaps (dotted line, $GapW$).

Figure 7. Kernel density estimation of insured yields and yield gaps: (A) expected yields (dashed-dot-dot line, Y_{exp}) and farmer's insured yields before sowing in option Basic (dotted line, Y_{expB}), before sowing in option Extended (solid line, Y_{insE}), and adjusted at mid-season in option Extended (dashed line, Y_{insEC}), and (B) gaps between farmer's insured yields before sowing in option Basic and expected yields (dashed line, $GapB$), before sowing in option Extended and expected yields (solid line, $GapE$), and the adjusted yields at mid-season in option Extended to expected yields (dotted line, $GapEC$).

Figure 7-A represents the kernel density estimation of expected yields (Y_{exp}), and the farmers' insured yield before sowing with option *Basic* (Y_{insB}), before sowing in the option *Extended* (Y_{insE}), and at midseason insured with option *Extended* and adjusted with the complementary insurance (Y_{insEC}). Figure 7-B shows the gaps $GapB$, $GapE$, and $GapEC$. The lowest yields were Y_{insE} , followed by Y_{exp} , Y_{insB} and lastly, Y_{insEC} , with the mean at 2.22, 2.77, 3.54 and 3.63 Mg ha⁻¹, respectively (Table 5, Figure 7-A). $GapE$ was the lowest, and had a negative mean at -0.56 Mg ha⁻¹, followed by $GapB$ with a positive mean at 0.75 Mg ha⁻¹, and lastly $GapEC$ with the mean at 0.86 Mg ha⁻¹ (Table 5).

3.4. Asymmetric information

To calculate the probabilities of indemnity (Eq 7) and the asymmetric information indicators (Eq 8, 9 and 10), probability density distribution functions were fitted for actual yields, $f(Ya)$, in each of the counties in the study region. Out of the 38 distribution functions fitted, 12 were Weibull, 11 Triangular, 4 Gamma, 7 BetaGeneral, 2 Logistic and lastly, 2 Logistic distributions. Such distributions have been proposed in earlier literature for crop yield modelling in agricultural economics (Atwood et al., 2003; Gallagher, 1987; Nelson and Preckel, 1989; Sherrick et al., 2004; Tolhurst and Ker, 2015). Results of these density fits are available from the authors upon request.

Figure 8 shows the kernel density estimation of the probability of indemnity (Figure 8-A) and the asymmetric information indicator (Figure 8-B). Again, the kernel density was estimated using Stata12 and the Epanechnikov kernel. The lowest probability of indemnity was found when using the average insured yield for the option *Extended* before sowing ($y_{thres} = Y_{insE}$), even lower than when no asymmetric information was considered ($y_{thres} = \overline{Ya}$). The probability of indemnity when using the average insured yield was higher for the option *Extended* including the complementary insurance (at mid-season, $y_{thres} = Y_{insEC}$) than for the option *Basic* ($y_{thres} = Y_{insB}$) (Figure 8-A). The probability of indemnity when no asymmetric information was considered ($y_{thres} = \overline{Ya}$) showed the lowest variability. Consequently, the asymmetric information indicator was higher for the option *Extended* including the complementary insurance ($y_{thres} = Y_{insEC}$) than for the option *Basic* ($y_{thres} = Y_{insB}$), and lastly and option *Extended* before sowing ($y_{thres} = Y_{insE}$) with a mean of 0.22, 0.17 and -0.08 (-), and a high variability with a range between percentiles 10 and 90 of 0.07 and 0.34, -0.03, and 0.33, and -0.13 and -0.01 (-), respectively (Table 5, Figure 8-A).

Figure 8. Kernel density estimation of the probability of indemnity and the asymmetric information indicator: (A) Frequency distribution of the probability of indemnity being y_{thres} calculated from average actual yield (solid line, Y_a), from the farmers' insured yields in option extended adjusted at mid-season (dashed grey line, Y_{insEC}) and from the farmers' insured yields in option Basic before sowing insured (dotted line, Y_{insB}); and (B) Frequency distribution of the asymmetric information indicator for farmers' insured yields in option Basic before sowing (dotted line, $AsymB$), and the asymmetric information indicator for farmers' insured yields in option Extended adjusted at including the complementary insurance (dashed grey line, $AsymEC$).

3.5. Insurance demand, yield gaps and asymmetric information

Table 8 shows the four factors including loadings and the percentage of variance of each of them, once rotated. Position of relevant variables determines the factors influencing the demand of wheat insurance in the study region. As a result, 49.5% of the variance is explained by factor 1 variables, defined as "Expectations" in Table 8 and Figure 9. Factor 2 variables "Yield asymmetric information" explains 19.5%, factor 3 variables "Yield variability" 17.5 % and lastly, factor 4 "Commercial effort", 13.5%.

Factor 1 grouped a bunch of variables ($GapZ$, $GapW$ and $AsymEC$) indicating a high correlation between them. Higher yield gaps between expected yields and maximum insurable yield ($GapZ$) and attainable rainfed yields ($GapW$) are then correlated with higher asymmetric information for the insurance option Extended when including the complementary insurance ($AsymEC$). This makes sense as when the farmer sets Y_{insEC} , the certainty on final yields is higher at mid-season. $GapZ$ was selected as the representative variable of factor 1 to be included in the regression models.

Table 8. Factor analysis model including loadings and the percentage of variance of the rotated factors.

	Factor 1	Factor 2	Factor 3	Factor 4
Variables	<i>Expectations</i>	<i>Asymmetric information</i>	<i>Yield variability</i>	<i>Commercial effort</i>
$Area_{wheat}$ (1000 ha)				0.81
$Rainfed$ (fraction)				
CV_{Ya} (-)			0.80	
$Skewness_{Ya}$ (-)				
$GapZ$ (Mg ha ⁻¹)	0.86			
$GapW$ (Mg ha ⁻¹)	0.69			
$AsymB$ (-)		0.90		
$AsymE$ (-)				
$AsymEC$ (-)	0.86			
Proportion (%)	49.5	19.5	17.5	13.5

Blanks represent $abs(\text{loading}) < .5$

KMO test value=0.73

Figure 9. Factor analysis results for crop insurance demand.

Two regression models were fitted using the most representative variables of each of the four factors to explain the demand of wheat insurance and the prevalence of the insurance option *Basic* ($Insured_{wheat}$ and $Insured_B / Insured_{wheat}$). Explanatory variables were, therefore, $Area_{wheat}$, $GapZ$, CV_{Ya} and $AsymB$. Table 9 shows the model parameters estimates of each of the explanatory variables and the significance of their contribution to explain the dependent variable, number of observations (N), and the coefficient of determination (R^2_{adj}).

The results show that wheat insurance demand ($Insured_{wheat}$) was higher in counties with higher wheat cultivated area ($Area_{wheat}$), higher $GapZ$ and lower asymmetric information for contracts in option *Basic* ($AsymB$). At the same time, counties with a higher $AsymB$ tend to have a higher proportion of insured area under the option *Basic*. Greater interest in contracting crop insurance would be expected in those counties with higher yield variability (higher CV_{Ya}) and in which crop yield distributions are skewed to the left (lower $Skewness_{Ya}$), with yields closer to the maximum observed more frequently than very low yields. However, model results discarded CV_{Ya} as an influencing factor on insurance demand.

Table 9. Parameter estimates in the regression models for the proportion of the wheat cultivated area that was insured ($Insured_{wheat}$) and the proportion of insured area in option *Basic* with respect to the total insured area in 2011 and 2012 in Castilla y León. Observations are per county and year.

Model	Model – explained variable	
	$Insured_{wheat}$	$Insured_B \text{ option Basic} / Insured_{wheat}$
Explanatory variables		
$Area_{wheat}$ (1000 ha)	0.08 ^{***1}	---
$GapZ$ (Mg ha ⁻¹)	0.03 [*]	---
CV_{Ya} (-)	---	---
$AsymB$ (-)	-0.28 ^{***}	0.78 ^{***}
N	72	72
R²_{adj}	0.30	0.17

¹ --- Not significant; * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

4. Discussion and conclusions

An analysis was conducted to identify the factors explaining insurance demand at the county level, looking in detail at the role of wheat production, yields gaps and asymmetric information-related factors. Yield gaps and wheat production in the study region were explored to assess their influence on farmer's decisions to contract insurance. In this section, both the results and the methods used are discussed.

Are subsidies driving farmer's agricultural insurance choice?

Premium subsidies are the most common public intervention to incentive crop insurance demand (Babcock and Hart, 2005; Claassen et al., 2005; Garrido and Zilberman, 2008b). Garrido

and Zilberman (2008) characterized economic factors influencing the demand of agricultural insurance using individual data from farmers insuring their crops in seven counties all along Spain. They concluded that premium subsidies were the leading incentive to contract crop insurance. However, in the case of wheat insurance in Castilla y León, and for option *Extended* (the one with the largest guarantees), most of the insured area was gathered by Module 2, despite subsidies in Module 1 being higher than in Module 2. The second most selected insurance alternative was Module P, the option *Basic*, with the lowest premium subsidy (Figure 1 and Table 1). Both in Module 2 and Module P, loss compensation is evaluated per plot, and the farmer is indemnified for each plot in which final yield is lower than insured yield. This suggests that the accuracy in setting insured yield is decisive in the willingness of the farmers of contracting insurance.

Wheat insurance demand models

The long tradition of cereal insurance in Spain, with more than 30 years of experience, has favoured strong commercial networks, especially in those counties with large arable crops orientation. This explains that insurance demand was higher in counties with larger areas cultivated with winter wheat (Table 9). Moreover, higher insurance demand is expected in vast regions with low farmers' density where commercial efforts by insurance companies are higher. Such an effort was assumed to be more profitable in areas with large arable crops orientation and thus with larger surface cultivated with wheat ($Area_{wheat}$); one successful sale would yield a higher selling commission. Moreover, regions with larger surface devoted to wheat production are considered as more crop specialized areas. Crop specialization, and thus lower income sources, increases farm revenue risk as farms become dependent on the performance of a single crop, becoming crop insurance an interesting alternative risk management tool (Cabas et al., 2008; Nieuwoudt et al., 1985).

Zonal maximum insurable yields were 0.85 Mg ha^{-1} lower than expected yields ($GapZ$, Table 5). This might explain why Module S is no longer offered, in favour of Module 1 and Module 2, in which maximum insurable yield is based on farmers' insurance history.

Information asymmetry indicators ($AsymB$, $AsymE$ and $AsymEC$) were also included in the analysis as a measure of farmers' information advantages. Results show a higher insurance demand related to a lower $AsymB$, but a higher proportion of the insured area under the option

Basic for higher $AsymB$. This suggests that regions with a better knowledge on the system (that is, lower asymmetric information and better adjusted insurance parameters), farmers have a higher tendency to contract insurance, and to do it under a larger coverage (option Extended). This result can also be the other way round: increasing insurance demand might help decreasing asymmetric information as the probability of participation of less risky producers increases, mitigating adverse selection (Shaik et al., 2008). Concurrently, efforts towards decreasing asymmetric information might be expected to enhance crop insurance demand; the lower the asymmetric information, the less biased the available data for insurance parameter calibration. Properly calibrated insurance parameters would make insuring more attractive to a higher number of farmers.

Is the concept of yield gap useful for asymmetric information assessment in agricultural insurance?

Yield gaps were defined in this work as the difference between the different concepts of “yield” used in a crop insurance context and actual yields. Such a difference was interpreted (both) as a deficiency in the insurer’s knowledge of the production system and as the information advantages of the farmer; therefore, as an indicator of asymmetric information. The concept of yield gap was useful to design an approach to quantify asymmetric information and to evaluate the relationships between insurance parameters (as $YinsZ$) and insurance demand.

Earlier research found risk aversion not being the only (and not even the main) cause of crop insurance participation, but also actuarial incentives (as subsidies) and asymmetric information would encourage to uptake (Just et al., 1999). Moreover, Babcock (2015) concluded that farmers regard crop insurance not only as a risk management tool but also as an income support. Thus, within this analysis, it was assumed that farmer chooses between the two prospects- insurance and no insurance – on the basis of expected profitability of each of the choices. The higher the $YinsZ$ (closer to $Yexp$), the higher the probability of harvested yield being lower than $YinsZ$; therefore, the probability of being indemnified increases and contracting an insurance becomes more profitable for the farmer. Consequently, and despite the later regression analysis discarded it, higher insurance demand could have been expected in counties with higher $GapZ$.

Are crop models useful for crop insurance assessment?

GapW was the highest yield gap found among the gaps included in this analysis. This result suggests that, as expected, it is more accurate to calibrate insurance parameters based on actual reported yields instead of attainable yields (or simulated water limited yields), and therefore on historical data rather than using crop models; otherwise farmers' yields would be overestimated.

Mean *GapW* in Castilla y León was 1.01 Mg ha^{-1} . This differs from those reported in Boogaard et al., (2013), within an analysis for the whole of Europe using the crop model WOFOST. They calculated *Yw* between 5 and 6 Mg ha^{-1} and *GapW* between 3 and 4 Mg ha^{-1} in Castilla y León. Therefore, mean *Yw* was about 1 Mg ha^{-1} higher and mean *Ya* was about 1.5 Mg ha^{-1} lower than reported in this work. These differences might be related to the accuracy in calibration and the scale of application. Inaccuracies of the models to simulate certain physiological processes and differences in the cultivars used in different regions might result in significant errors when the scale of application is large. Moreover, the orography and soil variability are difficult to include in the analysis. This is the case when calibrating models in a specific region, and then, results are extrapolated to a larger area with a low resolution in soil maps and weather data. Inaccuracies are expected to be higher in such a heterogeneous region as the Iberian Peninsula.

Nevertheless, the use of crop models might be useful in designing new insurance packages when no historical data is available or to evaluate scenarios of expected changes, such as under expected climate change. In that case, it is suggested that yield gaps be estimated and considered when using simulated attainable yields.

5. Acknowledgements

This research was funded by the project MULCLIVAR, from the Spanish Ministerio de Economía y Competitividad (MINECO) CGL2012-38923-C02-02, A. Castañeda-Vera has a PhD grant from UPM. We thank David A. Nafría-García, Miriam M. Fernández-Sánchez, and Alberto Gutiérrez-García (ITACyL) for the cooperation in soil data gathering, for clarifications on Castilla-y-León cereals farming and production regions and technical help. We also thank M^a José Martínez-Tineo (ENESA) for the clarifying discussions on cereals insurance, and the Spanish Ministry of Agriculture, Food and Environment (MAGRAMA) for meeting our request to use ESYRCE data.

Chapter 6

6. Main results and conclusions

A. Castañeda-Vera^{1,2}

¹Research Centre for the Management of Agricultural and Environmental Risks (CEIGRAM), Universidad Politécnica de Madrid (UPM), Spain

²Agricultural Systems Group, Dpto. Producción Agraria, E.T.S. Ingenieros Agrónomos, Universidad Politécnica de Madrid, Avda. Complutense s/n, 28040 Madrid, Spain

The overall objective of this thesis was to evaluate different aspects that hinder risk setting and that constitute disincentives to crop insurance demand. This chapter brings together the discussions of preceding chapters to answer the research questions initially posed and analyses general implications of the results and methodological issues. In section 1, the overall contributions of this thesis are discussed answering to the research questions. Section 2 summarizes the thesis's major conclusions and lastly, section 3 presents a few ideas for further research.

1. Summary of main results

Are weather indexes and crop models accurate methods for indirect risk assessment?

A strong component of this thesis was the adaptation of methodologies commonly implemented in other scientific and technical spheres to be applied in crop insurance. Methodologies involve mainly quantitative approaches and integrate agronomic, climate and socioeconomic variables at local and regional scale. Such methodologies are implemented to address a central issue in agricultural risk assessment, the availability of data (Goodwin and Mahul, 2004). Chapters 2, 3 and 5 evaluate the suitability of weather indexes and crop models for indirectly assess risk when historical data is missing.

A good fit of insurance parameters requires accurate data for risk estimation. We explored the consequences of lack of information in Chapter 2. In regions with a large tradition on a certain insurance line, databases on historical records on claims are broader to better elicit actual risks, what improves insurance premiums settings. However, there are situations in which no data are available for insurance calibration. This is the case of new coverage or new insurance lines. Chapter 2 demonstrates how unexpected increases in claims and indemnities might be associated not only to unusual years in terms of weather but also to underestimated risks due to the lack of accurate information.

Barnett (2004) points at the advantages of using index insurance. He argues that it is not susceptible to moral hazard and adverse selection, the two possible consequences of asymmetric information. Furthermore, it is claimed that it has a lower potential for error and lower transaction costs. However, index insurances have a major limitation; the fact

that farm yield is not always perfectly correlated with the index. Index insurances are already implemented in Spain for losses in natural pastures (Escribano Rodríguez et al. 2014). They are a cheaper substitute to current loss appraisal but they are limited to systemic risks, affecting homogenously in the region, and to more simple risks, where a clear relationship between risk and damage can be deduced (Bardají and Garrido, 2016). However, for more complex risks where a greater number of meteorological and managerial variables are involved, indexes implementation becomes complicated. That is the case of rainfall-related risks in processing tomato as studied in Chapter 2, where the low sparseness and high probability of multicollinearity of the meteorological data limits the power of the model to explain risk incidence. Moreover, the possibility of implementing weather index insurance is hampered by the significant dependency of the model on scale and year. Such dependency results in a high heterogeneity of the risk impact. Rainfall-related crop losses are dissimilar in the region. This makes difficult to describe the risk with the minimum accuracy and simplicity required to make this method applicable for commercial ends.

In Chapter 3, options for using crop models to generate data on expected yields for risk assessment are evaluated. The Chapter concludes that the use of such complex methodologies is the best tool available to get insight in these complex agronomic systems when no historical data exists or to evaluate scenarios of expected changes, such as under climate change. However, crop models are associated with numerous sources of uncertainty, and therefore they should be used with caution. The same conclusion can be drawn from Chapter 5, which also suggests the estimation of yield gaps to be considered when using simulated attainable yields.

In view of the CAP 2014-2020 consideration of various policies supporting different risk management instruments, how do they compare against each other?

Chapter 4 evaluates the relative importance of crop insurance with respect to the available risk management tools currently offered in the CAP 2014-2020. We found that direct payments are the most effective in increasing farm income stability. However, when used complementarily, protection options as crop insurance and Income Stabilisation Tools (IST) improved farm resilience to yield and price variability and restrained public expenditure.

Previous experiences in Spain show low participation of farmers in risk management instruments different than insurance, as the case of schemes based on mutual funds (e.g. potato farmers in Álava) or future markets (e.g. citrus). Therefore, before delivering a new product, a deep study on the availability of enough data to guarantee a proper functioning of the instrument, as well as an assessment on the expected farmers' insurance adoption is a need.

Such need on farmers' insurance adoption is pointed at in Chapter 4. It emphasizes the importance of understanding aspects affecting farmers' decision-making to properly evaluate insurance demand. Therefore, we examine farmers' rationale behind contracting insurance in Chapter 5. Actual average insurance penetration rates (about 40 % of the crop production is insured) support the conclusion that the system counts on farmers' trust. However, penetration rates vary significantly between risks, sectors and regions. We explored which factors influence farmers' decisions with respect to contracting insurance or not. We confirm that farmers' utility of contracting insurance is related not only to risk but also to its profitability. That is, they expect from the insurance not only to cover part of the income lost due to weather adversities, but also to get an extra income, as they insured higher yields than the actual expected yields.

Are asymmetric information and the availability of alternative risk management instruments affecting crop insurance demand?

Another threat to crop insurance associated to misinformation is the presence of asymmetric information. Such situation may lower the demand of crop insurance by low risk farmers, remaining higher risk farmers in the system. If no data are available to discriminate high-risk from low-risk farmers, this may lead to adverse selection in the case insurer is not able to modify crop insurance parameters (as premium or maximum insurable yield) to readapt them to the new farmers' pool risk, and so to potential problems of insurance market failure. In Chapter 5, we found the presence of asymmetric information in wheat insurance. Moreover, we found it influences farmer's decision on whether contracting insurance or not, as results showed a higher demand in regions with a better fit of insurance parameters (as maximum insurable yields).

Yield gap has been widely used to assess and understand the ecological possibilities to meet food demand for an increasing population (van Ittersum et al., 2013). Yield gaps are defined in agronomy as the differences between attainable yields based on climate and soil conditions and actual farmers' yields, and are usually defined with reference to some specified spatial and temporal scale (Lobell et al., 2009). In Chapter 5, we redefine yield gaps as the difference between the different concepts of "yield" used in a crop insurance context and actual yields. We interpreted such difference (both) as a deficiency in the insurer's knowledge of the production system and as the information advantages of the farmer; therefore, as an indicator of asymmetric information. The concept of yield gap was useful to design an approach to build evidence about the presence of asymmetric information and to evaluate the relationships between insurance parameters and insurance demand.

An interesting result in Chapter 4 was that crop insurance and IST are substitutive instruments, suggesting that the implementation of IST lead to a decrease in insurance demand. The introduction of such new protection schemes could potentially displace farmers that are currently participating in the system, decreasing the pool potential and, therefore, compromising its sustainability. Therefore, the introduction of new risk management schemes should be preceded by *ex-ante* evaluations to avoid unbalances in the system.

2. Conclusions

This thesis integrates methods and tools to evaluate aspects threatening Agricultural Insurance sustainability. Such sustainability refers to the maintenance of the system in the long term and it requires that both companies providing insurance and insured farmers benefit from it. In this attempt we evaluated threats for accurate risk assessment and factors decreasing farmers' insurance demand that may result in unbalances in the loss ratio. Major conclusions are classified in three sets answering the three research questions initially posed.

The first set of conclusions pertains to methods for indirect risk assessment, which is a key step in insurance design when historical crop loss data is unavailable.

- The use of insurance historical data rather than using indirect methods is preferred to evaluate and calibrate insurance parameters when available.
- However, historical data may be insufficient to establish the real risk the insurance is covering. Therefore, the use of weather indexes or more complex crop models is useful to design new insurance packages or coverage.
- Weather indexes and crop models should be used with caution. The time and spatial dependence of the risk influences the scale at which models might have validity. The common lack of observed data seriously hinders testing simulation models and thus constitutes an important source of uncertainty.
- The higher the complexity of the cropping system, the harder it becomes using indirect methods for crop risk assessment.

The second set of conclusions refers to the relative efficiency of alternative risk management instruments in enhancing and stabilizing farm income and the effectiveness of public support.

- Direct payments and crop diversification are the most effective policy measures in decreasing income variability.
- Crop diversification does not require any direct public support, but provides a significant decrease in income variability.
- Incentives for contracting protection tools are a good compromise for both improving farm resilience to yield and price variability and to restrain public expenditure, especially when used complementarily with direct payments and/or diversification.
- Farmers' utility of contracting insurance is related not only to risk reduction, but also to its financial profitability.

Lastly, the third set of conclusions relates to factors affecting crop insurance demand.

- The concept of yield gap is useful to build evidence about asymmetric information and to evaluate the relationships between insurance parameters and insurance demand.
- The accuracy in setting the insured yield is decisive in farmers' willingness to contract crop insurance, probably more than premium subsidies.

- The development of new risk management tools based on mutual funds may potentially decrease insurance demand, decreasing the pool and compromising the system's sustainability.

3. Limitations and recommendations for future research

This thesis represents an attempt to contribute with new insights that could make agricultural insurance systems more sustainable. In the research process we found some limitations that suggest paths for future research lines.

First, the lack of observed data adapted to simulation models structure seriously hampers such models testing. This constitutes an important source of uncertainty to determine whether simulated data is actually suitable for risk assessment.

A common recommendation is that models should be preferably calibrated with data registered in trials specifically designed to obtain parameter values for the selected models. This recommendation is valid for initial steps of evaluation; however, to implement crop models with commercial ends in crop insurance, we consider it is more pragmatic building new models that optimize the trade-off between the level of detail of the data available or achievable and the accuracy required in this specific application; that is, risk estimation and indirect quantification of actual losses.

Second, we assessed the potential of alternative risk management instruments, defined as mutual funds in the regulation, offered in the last CAP reform for 2014-2020. We found these mutual funds (IST) may behave as substitutes of crop insurance, potentially decreasing the pool of farmers in the already established SAIS. This, together with the difficulties found in previous experiences to introduce mutual funds, suggests that the increase of protection to a revenue or income level could have a higher acceptance when designed within an insurance scheme. That is, designing insurance lines covering either revenue or income instead of only yields, and being offered alternatively to farmers together with current lines, keeping farmers in the system or even attracting new farmers drawn by such wider coverage.

Third, the assessment presented in this thesis attempts to integrate farmer perception of risk and its relative weight on utility (e.g., contracting or not insurance and between

alternative risk protection tools). Different assumptions for the proposed models and tools could be further studied as well by the participation of farmers to provide realism into the consequences of information imperfections and policy making process for new risk management instruments.

References

- Affholder, F., Poeydebat, C., Corbeels, M., Scopel, E., Titttonell, P., 2012. The yield gap of major food crops in family agriculture in the tropics: Assessment and analysis through field surveys and modelling. *F. Crop. Res.* 143, 106–118.
- Agrosynergie, 2011. Evaluation of income effects of direct support. Brussels.
- Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration: guidelines for computing crop water requirements. Rome.
- Álvarez-Arias, M.V., Cubero-Jiménez, D., Fernández-Sánchez, M., Garrido-del Pozo, N., Gutiérrez-García, A., Nafría-García, D. alfonso, Villarino-Barrera, I., 2013. Atlas Agroclimático de Castilla y León. Agencia Estatal de Meteorología and Instituto Tecnológico Agrario de Castilla y León, Valladolid.
- Álvaro-Fuentes, J., Morell, F.J., Madejón, E., Lampurlanés, J., Arrúe, J.L., Cantero-Martínez, C., 2013. Soil biochemical properties in a semiarid Mediterranean agroecosystem as affected by long-term tillage and N fertilization. *Soil Tillage Res.* 129, 69–74.
- Angulo, C., Gaiser, T., Rötter, R.P., Børgesen, C.D., Hlavinka, P., Trnka, M., Ewert, F., 2014. “Fingerprints” of four crop models as affected by soil input data aggregation. *Eur. J. Agron.* 61, 35–48.
- Antón, J., Kimura, S., 2011. Risk Management in Agriculture in Spain (No. 43), OCDE Food, Agriculture and Fisheries.
- Asseng, S., Anderson, G.C., Dunin, F.X., Fillery, I.R.P., Dolling, P.J., Keating, B.A., 1998a. Use of the APSIM wheat model to predict yield, drainage, and NO₃- leaching for a deep sand. *Aust. J. Agric. Res.* 49, 363–378.
- Asseng, S., Ewert, F., Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P.J., Rötter, R.P., Cammarano, D., Brisson, N., Basso, B., Martre, P., Aggarwal, P.K., Angulo, C., Bertuzzi, P., Biernath, C., Challinor, A.J., Doltra, J., Gayler, S., Goldberg, R., Grant, R., Heng, L., Hooker, J., Hunt, L.A., Ingwersen, J., Izaurralde, R.C., Kersebaum, K.C., Müller, C., Naresh Kumar, S., Nendel, C., O’Leary, G., Olesen, J.E., Osborne, T.M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M.A., Shcherbak, I., Steduto, P., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., Wallach, D., White, J.W., Williams, J.R.,

- Wolf, J., 2013. Uncertainty in simulating wheat yields under climate change. *Nat. Clim. Chang.* 3, 827–832.
- Asseng, S., Ritchie, J.T., Smucker, A.J.M., Robertson, M.J., 1998b. Root growth and water uptake during water deficit and recovering in wheat 265–273.
- Atwood, J., Shaik, S., Watts, M., 2003. Are Crop Yields Normally Distributed? A Reexamination. *Am. J. Agric. Econ.* 85, 888 – 901.
- Babcock, B.A., 2015. Using Prospect Theory to Explain Anomalous Crop Insurance Coverage Choice, in: Allied Social Science Association. American Economic Association, Boston, MA.
- Babcock, B.A., Hart, C.E., 2005. Influence of the premium subsidy on farmers' crop insurance coverage decisions (No. 05-WP 393). Ames, Iowa.
- Bannayan, M., Crout, N.M.J., Hoogenboom, G., 2003. Application of the CERES-Wheat Model for Within-Season Prediction of Winter Wheat Yield in the United Kingdom. *Agron* 95, 114–125.
- Bardají, I., Garrido, A., 2016. State of play of risk management tools implemented by MS during the period 2014-2020: national and European frameworks. European Union Parliament, Brussels.
- Barnett, B.J., 2004. Agricultural index insurance products: strengths and limitations, in: *Agricultural Outlook Forum*. USDA, p. 16.
- Barrios-Gonzales, J.M., 1999. Comparative study of three approaches in the computation of soil water balance.
- Basso, B., Cammarano, D., Carfagna, E., 2013. Review of Crop Yield Forecasting Methods and Early Warning Systems Abstract, in: First Meeting of the Scientific Advisory Committee of the Global Strategy to Improve Agricultural and Rural Statistics. FAO Headquarters, Rome, Italy, pp. 1–56.
- Bielza, M., Conte, C.G., Gallego, F.J., Stroblmair, J., Catenaro, R., Dittmann, C., 2009. Risk Management and Agricultural Insurance Schemes in Europe. Joint Research Centre European Commission, Luxembourg.
- Bielza, M., Garrido, A., Sumpsi, J.M., 2004. Revenue Insurance as an Income Stabilization Policy: An Application to the Spanish Olive Oil Sector. *Cah. d'Economie Sociol. Rural.* 5–27.

- Bielza, M., Garrido, A., Sumpsi, J.M., 2007. Finding optimal price risk management instruments: The case of the Spanish potato sector. *Agric. Econ.* 36, 67–78.
- Boogaard, C.A., Van Diepen, R.P., Rötter, J.M.C.A., Cabrera H.H., Van Laar, H.L., De Wit, H.L., Te Roller, J.A., 2011. WOFOST Control Centre 1.8 and WOFOST 7.1.3. Wageningen.
- Boogaard, H., Wolf, J., Supit, I., Niemeyer, S., van Ittersum, M., 2013a. A regional implementation of WOFOST for calculating yield gaps of autumn-sown wheat across the European Union. *F. Crop. Res.* 143, 130–142.
- Boogaard, H., Wolf, J., Supit, I., Niemeyer, S., van Ittersum, M., 2013b. A regional implementation of WOFOST for calculating yield gaps of autumn-sown wheat across the European Union. *F. Crop. Res.* 143, 130–142.
- Boons-Prins, E.R., Koning, G.H.J. de, Diepen, C.A. van, Penning de Vries, F.W.T., 1993. Crop specific simulation parameters for yield forecasting across the European Community (No. 32), Simulation Reports CABO-TT. Wageningen, the Netherlands.
- Cabas, J.H., Leiva, A.J., Weersink, A., 2008. Modeling Exit and Entry of Farmers in a Crop Insurance Program. *Agric. Resour. Econ. Rev.* 1, 92–105.
- Campbell, G.S., 1985. *Soil Physics with BASIC: Transport Models for Soil-Plant Systems*. Elsevier, Amsterdam.
- Campbell, G.S., Mulla, D.J., 1990. Measurement of soil water content and potential. *Irrig. Agric. Crop.* 127–141.
- Cantero-Martínez, C., Angás, P., Lampurlanés, J., 2007. Long-term yield and water use efficiency under various tillage systems in Mediterranean rainfed conditions. *Ann. Appl. Biol.* 150, 293–305.
- Cassel, D.K., Nielsen, D.R., 1986. Field capacity and available water capacity, in: Klute, A. (Ed.), *Methods of Soil Analysis. Part 1: Physical and Mineralogical Methods*. American Society of Agronomy and Soil Science Society of America, Madison, pp. 901–926.
- Castañeda-Vera, A., Leffelaar, P.A., Álvaro-Fuentes, J., Cantero-Martínez, C., Mínguez, M.I., 2015. Selecting crop models for decision making in wheat insurance. *Eur. J. Agron.* 68, 97–116.
- Cattell, R.B., 1966. The scree test for the number of factors. *Multivariate Behav. Res.* 1, 245–276.

- Claassen, R., Lubowski, R.N., Roberts, M.J., 2005. Extent, Location and Characteristics of Land Cropped Due to Insurance Subsidies, in: American Agricultural Economics Association Annual Meeting. Providence, Rhode Island, p. 27.
- Coble, K.H., Knight, T.O., Pope, R.D., Williams, J.R., 1996. Modeling Farm-Level Crop Insurance Demand with Panel Data. *Am. J. Agric. Econ.* 78, 439.
- Coble, K.H., Knight, T.O., Pope, R.D., Williams, J.R., 1997. An Expected-Indemnity Approach to the Measurement of Moral Hazard in Crop Insurance. *Am. J. Agric. Econ.* 79, 216–226.
- Colman, E.A., 1947. A laboratory procedure for determining the field capacity of soils. *Soil Sci.* 277.
- Cordier, J., 2014. Comparative study of risk management tools supported by the 2014 Farm Bill and the CAP 2014-2020, Agriculture and Rural Development.
- Díaz-Ambrona, C.G.H., O’Leary, G.J., Sadras, V.O., O’Connell, M.G., Connor, D.J., 2005. Environmental risk analysis of farming systems in a semi-arid environment: effect of rotations and management practices on deep drainage. *F. Crop. Res.* 94, 257–271.
- Donoghue, E.J.O., 2014. The Effects of Premium Subsidies on Demand for Crop Insurance.
- Doorenbos, J., Pruitt, W.D., 1977. Guidelines for predicting crop water requirements.
- Efron, B., 1981. Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods. *Biometrika* 68, 589–599.
- El Benni, N., Finger, R., Mann, S., 2012. Price volatility and farm income stabilization. Modelling Outcomes and Assessing Market and Policy Based Responses, in: 123rd European Association of Agricultural Economics. Dublin, pp. 1–16.
- ENESA, 2016. Informe de contratación del seguro agrario. Madrid.
- Enjolras, G., Capitanio, F., Aubert, M., Adinolfi, F., 2014. Direct payments, crop insurance and the volatility of farm income. Some evidence in France and in Italy. *New Medit* 13, 31–40.
- Escribano Rodríguez, J.A., Hernández Díaz-Ambrona, C.G., Tarquis Alfonso, A.M., 2014. Selección de índices de vegetación para la estimación de la producción herbácea en dehesas. *Pastos* 44, 6–18.
- ESYRCE, 2013. Encuesta sobre superficies y rendimientos de cultivo. Madrid.

- ESYRCE, 2015. Encuesta sobre superficies y rendimientos de cultivos. Madrid.
- Eurostat, 2013. Agriculture, forestry and fishery statistics. Luxemburg.
- FAO, 2011. The State of Food Insecurity in the World. How does international price volatility affect domestic economies and food security? Rome.
- Farquhar, G.D., Sharkey, T.D., 1982. Stomatal Conductance and Photosynthesis. *Annu. Rev. Plant Physiol.* 33, 317–345.
- FEGA, 2016. Nota informativa sobre el establecimiento de los valores medios regionales definitivos y otros parámetros utilizados en el proceso de asignación de los derechos de pago básico. Madrid.
- Finger, R., 2012. Biases in Farm-Level Yield Risk Analysis due to Data Aggregation. *Ger. J. Agric. Econ.* 61, 30–43.
- Finger, R., El Benni, N., 2014. A Note on the Effects of the Income Stabilisation Tool on Income Inequality in Agriculture. *J. Agric. Econ.* 65, 739–745.
- Frère, M., Popov, G.S., 1979. Agrometeorological crop monitoring and forecasting (No. 17), Plant Production and Protection. Rome.
- Gallagher, P., 1987. U.S. soybean yields: Estimation and forecasting with nonsymmetric disturbances. *Am. J. Agric. Econ.* 69, 796–803.
- Garrido, A., Arroyo, C., Eiriz, G., 2014. Los seguros agrarios y el apoyo a la gestión del riesgo, in: *Reflexiones En Torno a La PAC*. Cajamar Caja Rural, pp. 203–236.
- Garrido, A., Brümmer, B., M'Barek, R., Meuwissen, M., Morales-Opazo, C. (Eds.), 2016. *Agricultural Markets Instability. Revisiting the Recent Food Crises*. Routledge.
- Garrido, A., Zilberman, D., 2008a. Revisiting the demand for agricultural insurance: the case of Spain. *Agric. Financ. Rev.* 68, 43–66.
- Garrido, A., Zilberman, D., 2008b. Revisiting the demand of agricultural insurance: The case of Spain. *Agric. Financ. Rev.* 68, 43–66.
- Geerts, S., Raes, D., 2009. Deficit irrigation as an on-farm strategy to maximize crop water productivity in dry areas. *Agric. Water Manag.* 96, 1275–1284.
- Glauber, J.W., Collins, K.J., Barry, P.J., 2002. Crop insurance, disaster assistance, and the role of the federal government in providing catastrophic risk protection. *Agric. Financ. Rev.* 62, 81 – 101.
- Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty,

- J., Robinson, S., Thomas, S.M., Toulmin, C., 2010. Food Security : The Challenge of feeding 9 billion people. *Science* (80-.). 327, 812–818.
- Godwin, D., Ritchie, J., Singh, U., Hunt, L., 1989. A User's Guide to CERES Wheat - V2.10. Alabama.
- Goodwin, B.K., 1994. Premium rate determination in the Federal Crop Insurance Program: What do averages have to say about risk? *J. Agric. Resour. Econ.* 19, 382–395.
- Goodwin, B.K., Mahul, O., 2004. Risk Modeling Concepts Relating to the Design and Rating of Agricultural Insurance Contracts (No. 3392).
- Goodwin, B.K., Roberts, M.C., Coble, K.H., 2000. Measurement of Price Risk in Revenue Insurance: Implications of Distributional Assumptions. *J. Agric. Resour. Econ.* 25, 195–214.
- Górski, T., Górka, K., 2003. The effects of scale on crop yield variability. *Agric. Syst.* 78, 425–434.
- Goudriaan, J., 1977. Crop micrometeorology : a simulation study. Centre for Agricultural Publishing and Documentation, Wageningen, the Netherlands.
- Guzman-Plazola, R.A., Davis, R.M., Marois, J.J., 2003. Effects of relative humidity and high temperature on spore germination and development of tomato powdery mildew (*Leveillula taurica*). *Crop Prot.* 22, 1157–1168.
- Hoogenboom, G., Jones, J.W., Wilkens, P.W., Porter, C.H., Boote, K.J., Hunt, L.A., Singh, U., Lizaso, J.L., White, J.W., Uryasev, O., Royce, F.S., Ogoshi, R., Gijsman, A.J., Tsuji, G.Y., Koo, J., 2012. Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.5 [CD-ROM].
- Hunt, J., Rees, H. Van, Hochman, Z., Carberry, P., Holzworth, D., Dalgliesh, N., Poulton, P., Rees, S. Van, Huth, N., Peake, A., 2006. Yield Prophet ® : An online crop simulation service, in: *Ground Breaking Stuff. Proceedings of the 13th ASA Conference, Perth, Western Australia*, pp. 1–5.
- IFAD/WFP, 2010. The Potential for Scale and Sustainability in Weather Index Insurance. Rome.
- INE, 2014. Encuesta sobre la Estructura de las Explotaciones Agrícolas 2013. Notas de prensa.
- JCyL, 2014. Anuario de estadística agraria de Castilla y León [WWW Document]. Anu.

- estadística. URL
http://www.jcyl.es/web/jcyl/AgriculturaGanaderia/es/Plantilla100/1284228463984/_/_/_
- Jones, J., Hoogenboom, G., Porter, C., Boote, K., Batchelor, W., Hunt, L., Wilkens, P., Singh, U., Gijsman, A., Ritchie, J., 2003. The DSSAT cropping system model. *Eur. J. Agron.* 18, 235–265.
- Just, R.E., Calvin, L., Quiggin, J., 1999. Adverse Selection in Crop Insurance : Actuarial and Asymmetric Information Incentives. *Am. J. Agric. Econ.* 81, 834–849.
- Kaiser, H.B., 1974. An index of factor simplicity. *Psychometrika* 39, 31–36.
- Kemanian, A.R., Stöckle, C.O., Huggins, D.R., 2005. Transpiration-use efficiency of barley. *Agric. For. Meteorol.* 130, 1–11.
- Kersebaum, K.C., Hecker, J., Mirschel, W., Wegehenkel, M., 2007. Modelling water and nutrient dynamics in soil – crop systems : a comparison of simulation models applied on common data sets, in: *Modelling Water and Nutrient Dynamics in Soil-Crop Systems*. Springer, Dordrecht, pp. 1–17.
- King, G., Zeng, L., 2001. Logistic Regression in Rare Events Data. *Polit. Anal.* 9, 137–163.
- Lobell, D.B., Cassman, K.G., Field, C.B., 2009. Crop Yield Gaps : Their Importance , Magnitudes , and Causes Crop Yield Gaps : Their Importance , Magnitudes , and Causes. *Annu. Rev. Environ. Resour.*
- Lobell, D.B., Ortiz-Monasterio, J.I., Falcon, W.P., 2007. Yield uncertainty at the field scale evaluated with multi-year satellite data. *Agric. Syst.* 92, 76–90.
- Machetti Bermejo, I., 2015. El sistema de seguros agrarios combinados en España [WWW Document]. Consorc. Compens. Seguros - Rev. Digit. URL
<http://www.conorsegurosdigital.com/es/numero-02/portada-02/el-sistema-de-seguros-agrarios-combinados-en-espana> (accessed 2.11.15).
- Macua-González, J.I., Lahoz-García, I., Calvillo-Ruiz, S., Bonzal-Yaguas, J.M., 2012. Variedades de tomate para industria: Resultados de la campaña 2012.
- MAGRAMA, 2014. Metodología de la Encuesta anual sobre superficies y rendimientos de cultivos en España. Madrid.
- MAGRAMA, 2015. Estudios de Costes y Rentas de las Explotaciones Agrarias (ECREA) [WWW Document]. URL

- http://www.magrama.gob.es/es/ministerio/servicios/analisis-y-prospectiva/ECREA_Informes-Agricolas.aspx (accessed 10.6.15).
- Mahul, O., Stutley, C.J., 2010. Government Support to Agricultural Insurance. Challenges and Options for Developing Countries (No. 53881). Washington. D.C.
- Marra, M.C., Schurle, B.W., 1994. Kansas Wheat Yield Risk Measures and Aggregation : A Meta-Analysis Approach. *J. Agric. Resour. Econ.* 19, 69–77.
- Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J.W., Rötter, R.P., Boote, K.J., Ruane, A.C., Thorburn, P.J., Cammarano, D., Hatfield, J.L., Rosenzweig, C., Aggarwal, P.K., Angulo, C., Basso, B., Bertuzzi, P., Biernath, C., Brisson, N., Challinor, A.J., Doltra, J., Gayler, S., Goldberg, R., Grant, R.F., Heng, L., Hooker, J., Hunt, L. a, Ingwersen, J., Izaurralde, R.C., Kersebaum, K.C., Müller, C., Kumar, S.N., Nendel, C., O’Leary, G., Olesen, J.E., Osborne, T.M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M., Shcherbak, I., Steduto, P., Stöckle, C.O., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., White, J.W., Wolf, J., 2014. Multimodel ensembles of wheat growth: many models are better than one. *Glob. Chang. Biol.* 21, 911–925.
- Mary, S., Santini, F., Boulanger, P., 2013. An Ex-Ante Assessment of CAP Income Stabilisation Payments using a Farm Household Model, in: 87th Annual Conference of the Agricultural Economics Society. Warwick, United Kingdom, pp. 1–27.
- McMaster, G.S., Wilhelm, W.W., Frank, A.B., 2005. Developmental sequences for simulating crop phenology for water-limiting conditions. *Aust. J. Agric. Res.* 56, 1277–1288.
- Meuwissen, M.P.M., Asseldonk, M.A.P.M. van, Huirne, R.B.M., 2008. Income stabilisation in European agriculture. Design and economic impact of risk management tools. Wageningen academics publishers.
- Morales, C., Garrido, A., Pálinkás, P., 2008. Risks Perceptions and Risk Management Instruments in the European Union : do farmers have a clear idea of what they need? Business 1–8.
- Nachabe, M.H., 1998. Refining the definition of field capacity in the literature. *J. Irrig. Drain. Eng.* 230–232.
- Nelson, C.H., Preckel, P.V., 1989. The conditional beta distribution as a stochastic production function. *Am. J. Agric. Econ.* 71, 370–378.
- Nieuwoudt, W.L., Johnson, S.R., Womack, A.W., Bullock, J.B., 1985. The demand of

- crop insurance (No. 1985-16), Agricultural Economics Report.
- Onofri, A., Becaffi, C., Benincasa, P., Guiducci, M., Tei, F., 2009. Is CropSyst adequate for management-oriented simulation of growth and yield of processing tomato? *J. Appl. Hortic.* 11, 17–22.
- Palisade Corporation, 2011. @RISK: A Hands-On Tutorial.
- Penman, H.L., 1948. Natural evaporation from open water, bare soil and grass. *Proc. R. Sociaty London* 193, 120–145.
- Penman, H.L., 1956. Evaporation: An introductory survey. *Netherlands J. Agric. Sci.* 9–29.
- Philip, J.R., 1957. Evaporation, and moisture and heat fields in the soil. *J. Meteorol.* 14, 354–366.
- Plaza-Bonilla, D., Álvaro-Fuentes, J., Cantero-Martínez, C., 2014. Identifying soil organic carbon fractions sensitive to agricultural management practices. *Soil Tillage Res.* 139, 19–22.
- 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, 81–92.
- Raes, D., Steduto, P., Hsiao, T.C., Fereres, E., 2009. 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. Calculation procedures, in: *Aquacrop v4.0 Reference Manual*. Rome, p. 130.
- Ranatunga, K., Nation, E., Barratt, D., 2008. Review of soil water models and their applications in Australia. *Environ. Model. Softw.* 23, 1182–1206.
- Rao, R., Li, Y., 2003. Management of flooding effects on growth of vegetable and selected field crops. *Horttechnology* 13, 610–616.
- Rinaldi, M., Ventrella, D., Gagliano, C., 2007. Comparison of nitrogen and irrigation strategies in tomato using CROPGRO model. A case study from Southern Italy. *Agric. Water Manag.* 87, 91–105.
- Ritchie, J.T., 1972. Model for predicting evaporation from a row crop with incomplete cover. *Water Resour. Res.* 8, 1204–1213.
- Rötter, R.P., Carter, T.R., Olesen, J.E., Porter, J.R., 2011. Crop–climate models need an overhaul. *Nat. Clim. Chang.* 1, 175–177.

- Rudstrom, M., Popp, M., Manning, P., Gbur, E., 2002. Data Aggregation Issues for Crop Yield Risk Analysis. *Can. J. Agric. Econ.* 50, 185–200.
- Ruiz-Zorrilla, J., 2010. La experiencia española en el aseguramiento y gestión de riesgos., in: *La Gestión de Riesgos Y Crisis En El Seguro Agrario*. Ministerio de Medio Ambiente y Medio Rural y Marino, pp. 1–48.
- Santeramo, F.G., Adinolfi, F., Capitano, F., Goodwin, B.K., 2016a. Farmer Participation, Entry and Exit decisions in the Italian Crop Insurance Program. *J. Agric. Econ.*
- Santeramo, F.G., Adinolfi, F., Capitano, F., Goodwin, B.K., 2016b. Farmer Participation, Entry and Exit decisions in the Italian Crop Insurance Program (No. 69508), Munich Personal RePEc Archive.
- Saseendran, S.S., Ahuja, L.R., Ma, L., Boote, K.J., Hoogenboom, G., 2008. Current Water Deficit Stress Simulations in Selected Agricultural System Models, in: *Response of Crops to Limited Water: Understanding and Modeling Water Stress Effects on Plant Growth Processes*. pp. 1–38.
- Seog, S.H., 2010. *The economics of risk and insurance*. Wiley-Blackwell.
- Shaik, S., Coble, K.H., Knight, T.O., Baquet, A.E., Patrick, G.F., 2008. Crop Revenue and Yield Insurance Demand : A Subjective Probability Approach. *J. Agric. Appl. Econ.* 757–766.
- Sherrick, B.J., Zanini, F.C., Schnitkey, G.D., Irwin, S.H., 2004. Crop Insurance Valuation Under Alternative Yield Distributions by. *Am. J. Agric. Econ.* 86, 406 – 419.
- Sinclair, T.R., Muchow, R.C., 1999. Radiation Use Efficiency. *Adv. Agron.* 65, 215–265.
- Skees, J.R., 2001. The bad harvest. *Regulation* 24, 16–21.
- Skees, J.R., Barnett, B.J., Collier, B., 2008. *Agricultural Insurance: Background and Context for Climate Adaptation Discussions*, Global AgRisk, inc. Lexington. Lexington, KY.
- Skees, J.R., Reeds, M.R., 1986. Rate Making for Farm-Level Crop Insurance: Implications for Adverse Selection. *Am. J. Agric. Econ.* 68, 653–659.
- Soil Survey Staff, 1994. *Keys to Soil Taxonomy*. Washington. D.C., USA.
- Sommer, R., Wall, P.C., Govaerts, B., 2007. Model-based assessment of maize cropping under conventional and conservation agriculture in highland Mexico. *Soil Tillage Res.* 94, 83–100.
- Steduto, P., Hsiao, T.C., Fereres, E., Raes, D., 2012. Crop yield response to water (No.

- 66), FAO Irrigation and drainage. Rome.
- Stöckle, C.O., Donatelli, M., Nelson, R., 2003. CropSyst , a cropping systems simulation model. *Eur. J. Agron.* 18, 289–307.
- Supit, I., Hooijer, A.A., van Diepen, C.A., 1994. System description of the WOFOST 6.0 crop simulation model implemented in CGMS. Volume 1: Theory and Algorithms (No. EUR 15956 EN), An Agricultural information system for the European community. Luxembourg.
- Thorntwaite, C.W., 1948. An Approach toward a Rational Classification of Climate. *Geogr. Rev.* 38, 55–94.
- Tolhurst, T., Ker, A., 2015. On Technological Change in Crop Yields. *Am. J. Agric. Econ.* 97, 137–158.
- Turner, N.C., 2004. Agronomic options for improving rainfall-use efficiency of crops in dryland farming systems. *J. Exp. Bot.* 55, 2413–25.
- UNESPA, 2012. Memoria social del seguro español.
- USDA, 2004. Estimation of Direct Runoff from Storm Rainfall, in: National Engineering Handbook. U.S. Department of Agriculture (USDA), pp. 1–51.
- van Ittersum, M.K., Cassman, K.G., Grassini, P., Wolf, J., Tittonell, P., Hochman, Z., 2013. Yield gap analysis with local to global relevance—A review. *F. Crop. Res.* 143, 4–17.
- van Keulen, H., Seligman, N.G., 1987. Simulation of water use. nitrogen nutrition and growth of a spring wheat crop. Centre for Agricultural Publishing and Documentation, Wageningen, the Netherlands.
- van Wart, J., Kersebaum, K.C., Peng, S., Milner, M., Cassman, K.G., 2013. Estimating crop yield potential at regional to national scales. *F. Crop. Res.* 143, 34–43.
- Vedenov, D. V., Barnett, B.J., 2004. Efficiency of Weather Derivatives as Primary Crop Insurance Instruments. *J. Agric. Resour. Econ.* 29, 387–403.
- Veihmeyer, F.J., Hendrickson, A.H., 1931. The moisture equivalent as a measure of the field capacity of soils. *Soil Sci.* 181–194.
- Ventrella, D., Charfeddine, M., Giglio, L., Castellini, M., 2012. Application of DSSAT models for an agronomic adaptation strategy under climate change in Southern of Italy: optimum sowing and transplanting time for winter durum wheat and tomato.

- Ital. J. Agron. 7, 109–115.
- White, J.W., Jones, J.W., Porter, C., McMaster, G.S., Sommer, R., 2009. Issues of spatial and temporal scale in modeling the effects of field operations on soil properties. *Oper. Res.* 10, 279–299.
- Yin, X., Struik, P.C., 2010. Modelling the crop: from system dynamics to systems biology. *J. Exp. Bot.* 61, 2171–83.
- Zhao, H., Gao, G., Yan, X., Zhang, Q., Hou, M., Zhu, Y., Tian, Z., 2011. Risk assessment of agricultural drought using the CERES-Wheat model: a case study of Henan Plain, China. *Clim. Res.* 50, 247–256.