UNIVERSIDAD DE LA REPÚBLICA
FACULTAD DE AGRONOMÍA

ESTIMACIÓN Y ANÁLISIS ESPACIOTEMPORAL DEL RENDIMIENTO Y LAS BRECHAS DE RENDIMIENTO DE SOJA EN URUGUAY

por

Gonzalo Darío RIZZO AYPHASSORHO

TESIS presentada como uno de los
requisitos para obtener el título de
Magíster en Ciencias Agrarias
opción Ciencias Vegetales

MONTEVIDEO
URUGUAY
Agosto 2018

Tesis aprobada por el tribunal integrado por PhD Santiago Dogliotti, PhD Andrés Berger, Dr. Sebastián Mazzilli, el 29 de octubre de 2018. Autor: Ing. Agr. Gonzalo Rizzo. Director Ing. Agr. Oswaldo Ernst, Co-directora Dra. Ing. Agr. Mónica Cadenazzi.

AGRADECIMIENTOS

Agradezco enormemente a Oswaldo Ernst, por el apoyo y la confianza depositada en mí a lo largo de todo este proyecto. Su capacidad de guiarme, y de al mismo tiempo, permitirme explorar libremente nuevos caminos me han enriquecido en conocimientos y experiencias extremadamente valiosas para mi formación. Su gran generosidad y visión me han impulsado a crecer como profesional y como persona.

A Patricio Grassini por su apoyo constante y gran disposición para responder mis inquietudes e impulsarme en mi formación.

A Juan Ignacio Rattalino Edreira por los momentos compartidos, sus consejos y estar dispuesto a compartir y desarrollar conocimiento conmigo.

A Juan Cardellino por brindarme generosamente el conocimiento que tantas veces fui a buscar.

A FUCREA por aportarnos generosamente la información de sus productores. A Deborah Gasso y Cristina Capurro por permitirnos acceder a la valiosísima información de sus experimentos.

A la Agencia Nacional de Investigación e Innovación (ANII) por otorgarme la beca de estudio y a la Facultad de Agronomía por concederme la posibilidad de priorizar mi formación como investigador y docente.

TABLA DE CONTENIDO

RESUMEN

Los objetivos de esta tesis fueron (i) estimar el rendimiento potencial sin limitantes (Yp), el rendimiento potencial limitado por agua (Yw) y la brecha de rendimiento para soja de primera y de segunda en Uruguay, (ii) analizar su variabilidad a través de las zonas climáticas, tipos de suelos y del tiempo, (iii) investigar el concepto de brecha de rendimiento impuesta por el sistema (CSY_g), y (iv) dilucidar las variables que explican el efecto año. Yp y Yw fueron estimados usando modelos de simulación de cultivos validados para las condiciones del país, con bases de datos de clima de largo plazo de buena calidad, datos de los tipos de suelos y las prácticas de manejo dominantes. Yg fue definido como la diferencia entre Yw y rendimiento actual (Ya). Yw y Yg fueron estimados para sitios específicos ubicados en la zona agrícola dominante, y luego estos resultados se escalaron a nivel país en base a la distribución espacial del área de cultivo y a las zonas climáticas. El Yw promedio a nivel nacional de Uruguay fue estimado en 3,6 y 3,4 Mg ha^{-1} para soja de primera y soja de segunda. Yg se estimó en un 40% del Yw de soja de primera y 50% del Yw de soja de segunda. El tipo de suelo capturó el 5 y el 1% de la variabilidad total para soja de primera y de segunda. El efecto año fue la mayor fuente de variabilidad (≈75%) en ambos sistemas de cultivo de soja. Para soja de primera, el agua almacenada en el suelo a la siembra, el déficit hídrico durante el período crítico, la disponibilidad de agua durante la estación de crecimiento y la radiación solar durante el período crítico resultaron las variables de mayor importancia. En soja de segunda las cuatro variables más importantes fueron el déficit hídrico durante el período crítico, el agua almacenada en el suelo a inicio del período de fijación de vainas, la radiación solar durante el período crítico y la disponibilidad de agua durante el ciclo del cultivo. La clasificación en función del riesgo mostró que los suelos en la zona principalmente agrícola y limitada extensión en el este pueden alcanzar el Yw a escala nacional en >70% de los años.

Palabras clave: variabilidad del rendimiento, modelos de simulación de cultivos, rendimiento potencial en secano, soja de siembra tardía

ESTIMATION AND SPATIO-TEMPORAL ANALYSIS OF YIELD AND YIELD GAP OF SOYBEAN IN URUGUAY

SUMMARY

The aims of the present study were (i) to estimate full-season and double-cropped soybean non-limiting yield potential (Yp), water-limited yield potential (Yw) and yield gap (Yg) in Uruguay, (ii) to analyze their variability among climatic zones, soil types and years, (iii) to investigate cropping system-imposed yield gap (CSY_g), and (iv) to elucidate variables explaining year effect. Yp and Yw were estimated using well calibrated crop model coupled with long-term good quality weather data, soil types and dominant management practices. Yg was estimated as difference between Yw and actual yield (Ya). First, Yw and Yg were estimated for specific locations in major crop producing areas, and then upscaled to country level based on spatial distribution of crop area and climate zones. Uruguay national average Yw was estimated on 3.6 and 3.4 Mg ha^{-1} for full-season and double-cropped soybean. Yg represented 40% of full-season soybean Yw and 50% of the double-cropped Yw. Soil type captured 5 and 1% of total variability, for full-season and double-cropped soybean. Year effect resulted the most important source of variability (≈75%) in both soybean cropping systems. For full-season soybean, soil available water at sowing date, water deficit during critical period, water availability during cropping season and solar radiation during critical period resulted the top four most important variables. For double-cropped soybean top four most important variables were water deficit during critical period, soil available water at beginning of pod setting, solar radiation during critical period and water availability along cropping season. Risk-based classification showed that some site in the main soybean producing area and limited extensions on eastern Uruguay, can reach national level Yw in more than 70% of years.

Keywords: yield variability, crop-model, water-limiting yield, late-sown soybean

1. INTRODUCCIÓN

1.1. MARCO GENERAL

Se estima que hacia 2050 la población mundial rondará los 9,5 billones de personas (Cassman, Dobermann, Walters, y Yang, 2003; Foley et al., 2011; Godfray et al., 2010) y para que ni una de ellas vea restringidas sus necesidades básicas nutricionales es necesario incrementar la producción total y equilibrar la distribución de estos a nivel global. Para dar respuesta a la primera parte de este dilema, surge la intensificación ecológica de la agricultura (Cassman, 1999; Tittonell, 2014), que plantea el desafío central de incrementar la producción por unidad de superficie, sin alterar el funcionamiento de los ecosistemas, asegurando sus servicios para las generaciones venideras. Generar avance substancial en el sentido de este paradigma, implica como paso inicial, medir la capacidad productiva de cada ambiente. En este marco se vuelve indispensable contar con evaluaciones de brechas de rendimiento.

Durante la primera década del siglo XXI, se desarrolló el fenómeno expansivo de la superficie agrícola de la región Pampeana de Sudamérica, más importante de la historia (Viglizzo et al., 2011). Dicho proceso, fue fundamentalmente desencadenado por el incremento sostenido de los precios de los *commodities* a nivel mundial. El Uruguay, siguió el comportamiento regional, pasando de una superficie agrícola de 400 mil hectáreas en 2003 a una superficie de 2 millones en 10 años, que marcó un récord histórico. El proceso de expansión fue principalmente dominado por el cultivo de soja, cuya superficie se multiplico por 110 durante el mencionado período. Sin embargo, los rendimientos a nivel nacional reportados por DIEA (2008, 2016) para este cultivo, se mantuvieron bajos (media 2,1 Mg ha^{-1}) y prácticamente sin experimentar cambio (b = 0,01; R^2 = 0). La incorporación de suelos con nula o baja capacidad agrícola, ha sido manejada como una de las causas de este estancamiento de los rendimientos medios del país.

Si bien hasta el presente, no se han realizado estimaciones de las brechas de rendimiento para el cultivo de soja en el Uruguay, los resultados reportados provenientes de experimentos de campo en situaciones de secano y riego sugieren que los rendimientos potenciales de este cultivo se ubican en torno a 6,2 Mg ha^{-1} (Giménez, 2017). A su vez, análisis de bases de registros de productores han reportado que el grupo de chacras de rendimientos más altos del set de datos estudiado (3,1 a 4,7 Mg ha^{-1}) estuvo asociado a suelos definidos como de alta aptitud agrícola, mientras que el grupo de chacras de rendimientos más bajos (1,5 a 2,1 Mg ha^{-1}) mantenían una fuerte asociación con los suelos de menor aptitud agrícola (Gonzalez, 2013).

1.2. NIVELES DE RENDIMIENTO Y BRECHA DE RENDIMIENTO

El rendimiento potencial no limitado ha sido definido como el rendimiento de un cultivar cuando es cultivado en un ambiente al cual se encuentra adaptado, con nutrientes y agua no limitantes, y libre de plagas, malezas y enfermedades (Evans, 1993; Evans y Fisher, 1999; van Ittersum y Rabbinge, 1997). En estas óptimas condiciones, el crecimiento del cultivo es determinado por la radiación solar, la temperatura, la concentración atmosférica de CO2, y las prácticas de manejo que influencian la duración del ciclo del cultivo y la intercepción de luz, tales como la fecha de siembra, el grupo de madurez y la densidad de plantas. En sistemas de secano, donde el agua almacenada en el suelo a la siembra junto con las precipitaciones durante el ciclo del cultivo, no son suficientes para cubrir los requerimientos hídricos del cultivo, el rendimiento potencial limitado por agua (Yw) es definido por el aporte de agua y su distribución durante la estación de crecimiento, y por las propiedades del suelo que afectan el balance de agua para el cultivo, y la pendiente del terreno (Lobell, Cassman, y Field, 2009; van Ittersum et al., 2013). La diferencia entre Yp (o Yw en condiciones de secano) y el rendimiento promedio de los productores (Ya), es definido como la brecha de rendimiento (Yg). Cerrar esta brecha de rendimiento mediante el ajuste de las prácticas de manejo que aplican los productores, genera oportunidades para incrementar la producción de granos, sin la necesidad de convertir áreas naturales en sistemas de cultivos (Cassman et al., 2003; van Ittersum et al., 2013).

En general, las evaluaciones de brechas de rendimiento se han realizado con el fin de evaluar el potencial incremento de los rendimientos de los cultivos que se puede obtener por adoptar las mejores prácticas de manejo disponibles. Muy escasos son los estudios que han intentado evaluar la brecha de rendimiento del sistema ya sea medido como kg ha^{-1} de grano (Liang et al., 2011) o como GJ ha^{-1} yr^{-1} (Guilpart, Grassini, Sadras, Timsina, y Cassman, 2017). En estos casos, el objetivo fue el de medir la brecha entre el rendimiento potencial del sistema de cultivos y el rendimiento actual de dicho sistema. También se ha empleado para estudiar alternativas de sistemas, con el propósito de identificar el sistema con los rendimientos potenciales más elevados (Guilpart, Grassini, Sadras, et al., 2017). En el presente trabajo, presentamos un enfoque novedoso, empleando el análisis de brechas de rendimiento como herramienta para la estimación de la brecha de rendimiento que el sistema de cultivos le impone al cultivo de soja de segunda (SYg), respecto del cultivo de soja de primera. Dado que un productor tiene la posibilidad de decidir entre un abanico de posibles sistemas a implementar (por ejemplo, sistemas de monocultivo o sistemas de secuencias de cultivos donde alternan cultivos de inviernos y cultivos de verano). La elección de un sistema donde la soja es sembrada luego de la cosecha de un cereal de invierno implica que la fecha de siembra del cultivo de soja queda subordinada a la fecha de cosecha del cultivo invernal, y por lo tanto, se abre una brecha entre el rendimiento obtenido por la soja de primera y el rendimiento de la soja de segunda.

1.3. PROTOCOLO DEL ATLAS MUNDIAL DE BRECHAS DE RENDIMIENTO

El Global Yield Gap Atlas (www.yieldgap.org) es el proyecto de evaluación de brechas de rendimiento más importante a nivel mundial, liderado por la Universidad de Nebraska-Lincoln y la Universidad de Wageningen. El mismo, tiene el objetivo de desarrollar un atlas mundial de brechas de rendimiento de los principales cultivos que hacen a la alimentación global.

El Global Yield Gap Atlas propone hacer estimaciones de brechas de rendimientos a través de varias escalas espaciales, desde sitio-específicos dentro de las zonas más importantes para la producción de cierto cultivo (i.e. puntos en localidades con gran proporción del área del cultivo y con una zona buffer asociada), a nivel de zonas climáticas (CZs, definidas por grados días, estacionalidad de la temperatura y un índice de aridez), a grandes unidades administrativas dentro de un país (departamentos, provincias o estados), hasta a nivel de promedio nacional. Por lo tanto, es fundamental identificar las CZs y los sitios específicos y su zona buffer asociada dentro de estas CZs, que sean representativos de cómo ese cultivo es producido en términos de clima, suelos y sistema de cultivo. La información de sistema de cultivo que es necesaria para seguir este protocolo refiere a la proporción que representa de la superficie total del cultivo en cuestión, la intensidad de cultivo y algunos aspectos del manejo (por ejemplo, fecha de siembra, densidad de plantas, grupo de madurez), a cada una de esas diferentes escalas a la que se trabaja. Esta metodología se fundamenta en resultados de algunos artículos recientes publicados por van Ittersum et al (2013) y van Wart et al (2013). Los sitios específicos (o puntos) son definidos como localidades con datos climáticos disponibles. La zona buffer de un sitio seleccionado con datos climáticos encierra un área dentro de 100 km de radio centrado en la estación meteorológica y cortado por los bordes de la CZ a la que pertenece.

Dentro de esa zona buffer, se recolectan los datos de la combinación tipos de suelo dominantes × sistema de cultivos para un régimen hídrico dado (tanto secano como irrigado, o ambos si es que los dos se encuentran de forma significativa). Para cada zona buffer, Yp y/o Yw son estimados por simulaciones usando los datos climáticos, la información de los tipos de suelos y de los sistemas de cultivo como una entrada en un modelo de simulación de cultivos. El método de escalado pasa desde zona buffer (si es que hay más de una zona buffer dentro de una CZ), a las CZs, a nivel subnacional, hasta nivel nacional. Este enfoque requiere flexibilidad de las fuentes de datos de clima ya que los puntos seleccionados con datos climáticos deben estar bien ubicados dentro de la CZ en las principales zonas productoras del cultivo. En los casos en los que no se dispone con estaciones meteorológicas con al menos 10 años de datos medidos de buena calidad, se cuenta con segunda opción con la de datos generados para 20 años a partir de un mínimo de tres años de datos medidos. La construcción de bases de datos híbridas (una parte de datos observados y una parte de datos generados a partir de datos de una estación meteorológica cercana que quizás solo cuenta con datos de precipitaciones y/o temperatura) es la opción que se maneja en tercer lugar en este protocolo. Como última opción se plantea el uso de datos derivados de datos grillados. Debido a que se requieren datos detallados de los sistemas de cultivos y tipos de suelos de cada localidad, unos de los objetivos del protocolo de selección es el de minimizar el número de puntos (y sus asociadas zonas buffer) necesarios para obtener estimaciones robustas de Yp y/o Yw dentro de un país.

En este protocolo, los datos de clima, suelo y sistema de cultivo se consideran igualmente importantes para capturar la variabilidad dentro de una zona climática. Para las estimaciones de brechas de rendimientos, los datos de rendimientos actuales de chacras son también muy importantes. La selección de las CZs y los sitios específicos con datos climáticos son el punto de partida de este protocolo para minimizar el número de localidades de las cuales se necesitará otros datos esenciales, y al mismo tiempo lograr una adecuada cobertura del área de producción del cultivo para asegurar una evaluación a través de un rango representativo de sistemas de cultivos y suelos.

La distribución espacial del área de cultivo en general es derivada de la base de datos SPAM (You, Wood, Wood-Sichra, y Wu, 2014). SPAM contiene datos grillados (resolución de 5 arc minutos, aproximadamente 10 × 10 km al ecuador) del área de cultivo alrededor del año 2000 para los 20 mayores cultivos básicos, regímenes hídricos (secano y regado), y para la agricultura de secano el área de los cultivos es desagregada por el nivel de insumos de los sistemas de cultivo (subsistencia, bajo y alto insumo). Si hay estadísticas nacionales del área de producción de los cultivos disponibles, mapas actualizados del área de cosecha de los cultivos pueden ser generados para los países donde el área de cultivos se ha expandido recientemente (por ejemplo, Uruguay, Argentina y Brasil).

El protocolo desarrollado por el Global Yield Gap Atlas, ha sido aplicado ampliamente alrededor del mundo para la determinación y análisis de la brecha de rendimiento de diversos cultivos. En Argentina, fue aplicado para evaluar la brecha de rendimiento de los cultivos de soja, trigo y maíz, y estudiar el efecto del fenómeno ENSO (El Niño South Oscilation) (Aramburu Merlos et al., 2015). En Brasil, fue aplicado para la determinación de la brecha de rendimiento de la caña de azúcar y analizar la capacidad del país de alcanzar la demanda proyectada por esta materia prima en escenarios futuros (Marin, Martha, Cassman, y Grassini, 2016). En Australia, fue aplicado para estudiar la brecha de rendimiento del trigo y contrastar los resultados obtenidos mediante la aplicación del protocolo del Global Yield Gap Atlas y un método que requiere una riqueza mayor de datos (habiéndolo permitido la alta densidad de datos con la que cuenta este país), siendo la robusticidad de ambos métodos (por la similitud de sus resultados) un resultado notable el de este trabajo. En Estados Unidos ha sido aplicado para medir la brecha de rendimiento de arroz (Espe et al., 2016). En Bangladesh, fue aplicado para estudiar la capacidad de este país de cubrir su propia demanda (dada su alta población) de trigo, maíz y arroz para el 2030 y el 2050, cerrando sus actuales brechas de rendimiento. En África, ha sido intensamente aplicado con el fin de analizar la soberanía alimentaria de este continente, ya que es donde se espera el mayor crecimiento demográfico durante la primera mitad del siglo XXI (van Ittersum et al., 2016; van Oort et al., 2015) y los trabajos más recientes han aplicado el protocolo para evaluar y mapear la profundidad radicular en los suelos de África Sub-Sahariana (Guilpart, Grassini, van Wart, et al., 2017; Leenaars et al., 2018). Además, se he empleado como marco de trabajo para estimar brechas de rendimiento a nivel de sistema de cultivos (Guilpart, Grassini, Sadras, et al., 2017).

1.4. RELEVANCIA A NIVEL LOCAL Y GLOBAL

Determinar las brechas de rendimiento de diferentes cultivos es la primera tarea que realizar cuando se busca ir hacia sistemas más intensivos, ya que permite trazar un plan para cerrar las brechas y lograr alcanzar ciertos objetivos productivos. La determinación de brechas de rendimiento en sistemas de doble cultivo es incipiente, y en ninguno de los casos se han determinado las brechas de rendimiento que se impone por un sistema de cultivo a un cultivo en particular. Sistemas de doble cultivos en los que alternan cereales de invierno con cultivos de verano (soja o maíz), se encuentran ocupando algunas de las regiones más productoras de granos del mundo, como ser las pampas sudamericanas (South American Pampas) y las planicies del norte de China (Hebei Plain and North China Plain), dada la benevolencia del clima de estas vastas regiones.

A nivel local, esta tesis es la primera determinación de las brechas de rendimiento de soja de primera y soja de segunda. Conocer los rendimientos potenciales sin limitantes y limitado solo por agua, permite conocer los rendimientos que se pueden alcanzar en las diferentes zonas del país. Entender cuáles son las variables que explican la variabilidad de los mismo, nos permitiría buscar soluciones tendientes a maximizar las chances de alcanzar los máximos rendimientos, llevando a sistemas más intensivos y rentables.

1.5. OBJETIVOS

1.5.1. Objetivos generales

El objetivo general de este trabajo fue la estimación de la brecha de rendimiento del cultivo de soja de primera y segunda del Uruguay, y determinar las variables que definen la variabilidad de los rendimientos potenciales.

1.5.2. Objetivos específicos

1. Cuantificar y mapear Yp, Yw, Ya y Yg para soja de primera y soja de segunda separadamente.
2. Calcular índice de limitacón hídrica y su relación con el Yw y la brecha de rendimiento.
3. Estimar la brecha de rendimiento impuesta por el sistema de cultivo.
4. Cuantificar el efecto zona climática, tipo de suelo y año en el Yw para soja de primera y segunda.
5. Identificar las variables de oferta de recurso de mayor importancia y definir su comportamiento.
6. Delimitar zonas de riesgo.

1.6. HIPÓTESIS DE TRABAJO

Hipótesis 1: El Yp, Yw y Ya son variables espacialmente y dependientes del sistema de cultivo, generando un Yg que puede ser mapeado.

Hipótesis 2: El sistema de cultivo le impone una brecha de rendimiento al cultivo de segunda.

Hipótesis 3: La variabilidad temporal del Yw puede atribuirse al efecto combinado del sistema de cultivo-zona climática-suelo-año, sobre la oferta de recursos y la captura de precipitaciones.

Hipótesis 4: El riesgo de producción de soja es diferente según zonas y sistema de cultivo.

2. QUANTIFYING SYSTEM-IMPOSED YIELD GAP FOR FUL-SEASON AND DOUBLE-CROPPED SOYBEAN IN URUGUAY[1]

Gonzalo Rizzo[a], Oswaldo Ernst[a]

[a] Departamento de Producción Vegetal, Facultad de Agronomía, Estación Experimental Mario Alberto Cassinoni, Universidad de la República, Paysandú, 60000, Uruguay

2.1 ABSTRACT

Favorable weather conditions prevailing on most South American Pampas (Uruguay included) allows to grow more than one crop per year. Indeed, increase crop intensity per unit area has been highlighted as an opportunity to increase crop production in a high food demanding world scenario. The aim of the present study was (i) to estimate full-season and double-cropped soybean Yp, Yw and Yg, (ii) to analyze how Yp, Yw and Yg varied among climatic zones and years, and (iii) to investigate cropping system-imposed yield gap (CSY_g) for soybean using Uruguay as a proof of concept. Yw was estimated by means of well calibrated crop model coupled with long-term good quality weather data, soil types and dominant management practices. Yg was estimated as the difference between Yw and actual yield (Ya). First, Yw and Yg was estimated for specific locations in major crop producing areas, and then upscaled to country level based on spatial distribution of crop area and climate zones. From the studied period national level Yg represented 40% of Yw for full-season soybean and 50% for double-cropped soybean. Uruguay national average Yw was estimated on 3.6 and 3.4 Mg ha^{-1} for full-season and double-cropped soybean, respectively. Estimated Yw and Yg were highly variable among years, and variable among climatic zones. Full-season soybean resulted in the highest maximum Yw, however double-cropped soybean resulted in a higher Yw

---

[1] For publication on Field Crops Research journal.

than full-season soybean in 35% of the site-year cases, indicating that diversify between cropping systems could help to reduce risk. We found that $CSY_g^{weather}$ was responsible for the 60% of the sum of three horizontal yield gaps. These results highlight the viability of increasing crop intensity as measure to enhance grain crop production in this main grain producing region.

Keywords: South American Pampas, crop-model, Global Yield Gap Atlas, water-limiting yield, late-sown soybean

2.2. INTRODUCTION

To cope with food demand by the 9 billion people world, as forecasted by 2050, crop production needs to be increased 60% (Alexandratos and Bruinsma, 2012). Although, production can be increased by expansion of current crop area, higher yields or both (Bruinsma, 2009), to follow an ecological intensification of agriculture paradigm (Cassman, 1999; Tittonell, 2014) supposes to increase production to high levels without causing environmental damage. In this scenario, to increase yield per unit area through reductions of yield gaps (Yg) seems to be the most suitable option for major agricultural regions (Fischer et al., 2014). Increasing yield potential (Yp) is another way to increase global food production, although this comes as a promising option for those regions that have efficiently closed their Yg (Fischer et al., 2014).

Yp is defined as the yield of a cultivar when grown in an environment to which it is adapted, with nutrients and water non-limiting and with pests, weeds, and diseases effectively controlled (Evans, 1993; Evans and Fisher, 1999; van Ittersum and Rabbinge, 1997). In these optimal conditions, crop growth is determined by solar radiation, temperature, atmospheric CO_2 concentration, crop physiological attributes governing light interception, conversion into biomass, and partition into the harvestable organs. In rainfed cropping systems where water supply from stored soil water at sowing and in-season rainfall is not enough to meet fully crop water demand, water-limited yield potential (Yw) is also determined by water supply amount and its distribution during the growing season, and by soil properties influencing the crop water balance, such as rootable soil depth, available-water holding capacity, and terrain slope (van Ittersum et al., 2013). In rainfed conditions, Yg is defined as the difference between Yw and actual farmer yield (van Ittersum et al., 2013).

Previous studies on yield gap analysis have been done for many countries around the world, because knowing the Yg is the first step to identified bottlenecks where more research or politics measures are needed to bring Ya to higher levels (and thus, to close Yg). Aramburu Merlos et al. (2015) estimated Argentinian Yw and Yg for soybean, wheat and maize following GYGA protocols (www.yieldgap.org). They estimated soybean Yw about 3.9 Mg ha^{-1} and its Yg at national level of 1.3 Mg ha^{-1}, varying across the country and being strongly dependent on year to year variation on water supply. However, while this research account for the double-cropped soybean yield into a double cropping sequence, the result were weighted into the whole soybean results, and thus the CS_{Yg} was not analyzed.

Given the amount and distribution of the annual rainfall and the extension of the frost-free period, almost in the entire South American Pampas, it is possible to cultivate a double-crop sequence in a year (*i.e.*, a winter crop followed by a late summer crop). Double crop sequences suppose a higher resource capture compared to full-season crops (Caviglia et al., 2004; Rao and Willey, 1983; Van Opstal et al., 2011), that return an increase on total grain production per unit of land and time. Nevertheless, the impact of the cropping system design on Yp, Yw and Yg of double-cropped soybean compared to those for full-season soybean crop have not been previously reported. When summer crops are sown following barley or wheat crop in a double cropping system (double-cropped soybean, *i.e.* soybean sown after a winter cereal crop), management practices such as sowing date and cycle duration from the previous crop impose additional limitations to subsequent crop's Yp or Yw.

Under such conditions, CS_{Yg} can be defined has the difference between full-season soybean Yp (Yw or Ya) and double-cropped soybean Yp (Yw or Ya).

Till present, estimations of cropping-systems yield gaps have been focused on total amount of grain yield (Liang et al., 2011) or units of energy (Guilpart et al., 2017) produced by the sequence of crops. These approaches are useful to compare between different cropping systems and evaluate alternatives based on total cropping system yield (either grain or energy). However, they do not allow to quantify losses or gains on each specific crop grain yield due to changes in crop sequence (*i.e.* comparison between soybean cropped after a winter cereal crop or after a winter fallow). Our work focuses on the CS_{Yg} caused by farmer's decision when planning their cropping systems and choosing between full-season or double-cropped soybean. In Uruguay, more than 40% of soybean is sown following a winter crop (wheat or barley) integrating double annual cropping systems. This provides an opportunity to evaluate the impact of agricultural intensification on soybean Yg. To our knowledge, no previous study has quantified this CS_{Yg} component on a full-season crop.

The aim of this study was to assess full-season and double-cropped soybean for Uruguay, using for this purpose validated crop simulation models, coupled with high-quality weather, soil, and crop management data, following the protocols of the GYGA project (Grassini et al., 2015; van Bussel et al., 2015; www.yieldgap.org/methods). Specific objectives were: (i) to estimate full-season and double-cropped soybean Yp, Yw and Yg, (ii) to analyze how Yp, Yw and Yg varied among climatic zones and years, and (iii) to investigate CS_{Yg} for soybean using Uruguay as a proof of concept.

2.3. MATERIALS AND METHODS

2.3.1. Study area and site selection

Data on soybean harvested area was retrieved for each enumeration area (*i.e.* the smallest administrative unit in Uruguay) from the 2011 National Agricultural Census (DIEA, 2011). Only rainfed soybean crops were accounted for in the present study as irrigated area accounts for less than one percent of area sown with soybean.

Weather data sources selection and quality control was done based on the Global Yield Gap Atlas protocol (Grassini et al., 2015). Daily maximum and minimum air temperature, precipitation and relative humidity were derived from EEMAC (Mario A. Cassinoni Experimental Station from Faculty of Agronomy), INIA (National Institute of Agricultural Research) and INUMET (Uruguayan Weather Institute) weather stations. Measured daily incident solar radiation was not available for the INUMET and INIA weather station, hence data from NASA-POWER were used as source of daily incident solar radiation. Previous works evaluating the NASA-POWER solar radiation data have reported a very good agreement with measured solar radiation data in those regions with flat topography (van Wart et al., 2013a; White et al., 2011). Same behavior have been found in Argentina (Aramburu Merlos et al., 2015) and Uruguay (*n*= 15,419 daily observations). By combing temperature and precipitations from EEMAC, INIA and INUMET, and the solar radiation from NASA-POWER a complete data base of weather records for 2000-2016 was obtained. This number of years is consider suitable for robust estimation of Yw and its variability (Grassini et al., 2015).

Reference weather stations (RWS) used in this study, were selected based on the harvested area within a buffer zone area of 100 km radius centered on each RWS and clipped by the climate zone (CZ) were the RWS were located (van Bussel et al., 2015). Each CZ corresponds to a particular combination of annual growing degree days (AGDD; base temperature 0°C), aridity index (AI, values increase for more humid conditions, and decrease with more arid conditions) and temperature seasonality (TS) (van Wart et al., 2013b). RWS were iteratively selected starting with the one with largest harvested area coverage until reaching 50% of crop-harvested area and more than 70% coverage by the CZ where RWS were located. The three CZ used in this study presented the same range of AGDD (5990-7111) and TS (3833-8355), being different on their AI range: 8686-10181 (CZ 6702), 6589-7785 (CZ 6502) and 7786-8685 (CZ 6602) (Fig. 1).

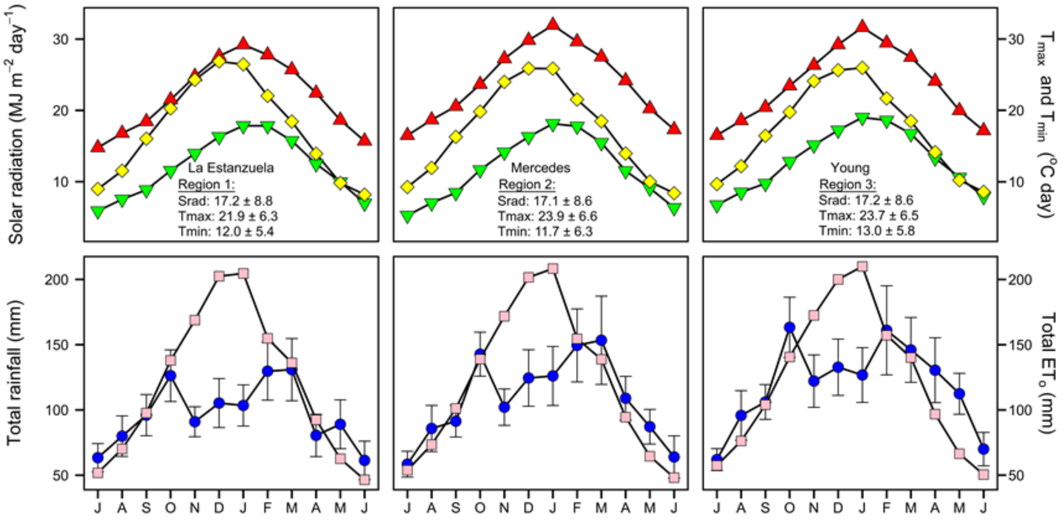

Fig. 1. *Monthly average incoming solar radiation (in yellow), maximum (Tmax, in red) and minimum temperature (Tmin, in green), total rainfall (in blue) and total grass-based reference evapotranspiration based on long-term (2000-2016) (ET_o, in pink) weather data at station located in La Estanzuela (south-west, Region 1), Mercedes (central-west, Region 2) and Young (north-west, Region 3). Error bars indicate ± SD and are shown only for rainfall. Average (± SD) solar radiation (Srad) in MJ m^{-2} d^{-1}, maximum (Tmax) and minimum temperature (Tmin) in °C for each region are shown in the top panels.*

Soils cultivated with soybean into each RWS were identified based on data provided by Natural Resources Directory of Uruguayan Ministry of Livestock, Agriculture and Fisheries (www.mgap.gub.uy/unidad-organizativa/direccion-general-de-recursos-naturales). Soils covering the 90% of the soybean harvested area from each RWS (10 to 14 soils series) were selected based on the CONEAT map (i.e. most detailed national-level soil map; 1:40,000). Functional soil properties required to run crop simulation models (e.g., field capacity and permanent wilting point) were derived from soil series descriptions following Ritchie and Crum (1989), after the revisions made by (Gijsman et al., 2002). Maximum rooting depth was allowed to reach the maximum soil depth of each soil, ranging from 0.8 to 1.3 m.

2.3.2. Potential and water-limited yields estimation

Potential and water-limited yields were estimated through simulations using CROPGRO-Soybean model embedded in DSSAT v4.6 (Hoogenboom et al., 2015; Jones et al., 2003). Genetic coefficients were derived from Mercau (2007) and Monzon (2012, 2007), and unpublished data from well-managed experiments. The model performance to simulate Yp and Yw was assessed by comparison of model simulated yield against measured yields from well-managed rainfed and irrigated experiments that explore a wide range of sowing dates, sites, years and water availability (Fig. 2).

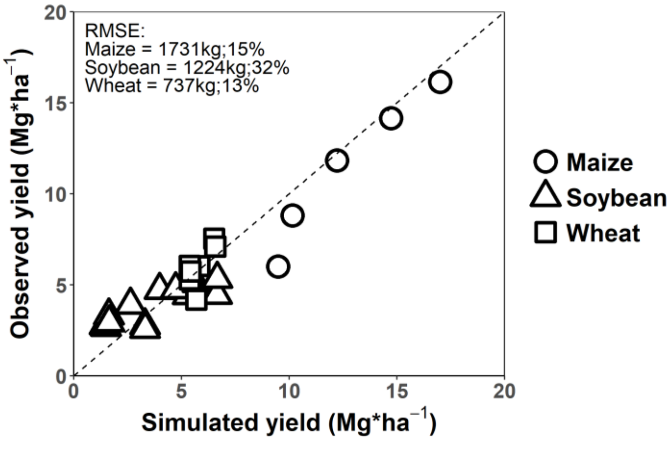

Fig. 2. *Model performance assessment comparing simulated and observed crop yield data. Dashed line represents x=y. The root mean square error expressed in Mg ha^{-1}, and coefficient of variation expressed in percentage. Genetic coefficients were derived from Mercau* (2007) *and Monzon* (2012, 2007), *and unpublished data from well-managed experiments.*

2.3.3. Simulated cropping systems and management

Actual yields (Ya) at subnational level and cropping system management practices data (*e.g.* sowing date, cultivars, plan population density and crop sequence) are not publicly available in Uruguay. Hence, Ya and most popular management practices were retrieved from 2009-2016 agronomic records from farmers of the Uruguayan Federation of Regional Consortia of Agricultural Experimentation (FUCREA using its Spanish acronym, www.fucrea.org/), and derived for each RWS based on farm locations. Retrieved information include: average planting dates, dominant cultivar name and maturity, and actual and optimal plant population density (Table 1). The provided data were subsequently corroborated by national experts. Dominant crop rotations and proportion of each of them to the total harvested area was calculated based on cropping sequences reported in the Planes de Uso de Suelos (Information provided by Dirección General de Recursos Naturales del Ministerio de Ganadería Agricultura y Pesca, www.mgap.gub.uy/unidad-organizativa/direccion-general-de-recursos-naturales). Following Aramburu Merlos et al. (2015), to account for the differential water availability of each soil at sowing of each crop through the studied period, we assumed 50% of the plant available soil water in the first year of the time series. Simulated crop sequences were: (i) 1-year cover crop/soybean (*i.e.* continuous soybean), (ii) 2-year cover crop/soybean-wheat/ double-cropped soybean, and (iii) 3-year cover crop/soybean-wheat/double-cropped maize-soybean except for Treinta y Tres. Results from the simulated crop sequences where weighted based on its contribution to total crop sequence harvested area into each buffer-RWS. Atmospheric CO_2 concentration was set constant at 400 ppm.

Table 1. *Soybean (full-season and double crop) management practices, as retrieved from 2009-2016 agronomic records from farmers of the Uruguayan Federation of Regional Consortia of Agricultural Experimentation (FUCREA, www.fucrea.org), applied to estimate water-limited and potential yields at each reference weather station (RWS).*

	Maturity group		Plant density (plant m^{-2})		Sowing date	
RWS	Full-season	Double	Full-season	Double	Full-season	Double
La Estanzuela	VI	VI	35	35	10-Nov	12-Dic
Trinidad	VI	VI	35	35	10-Nov	14-Dic
Mercedes	VI	VI	35	35	10-Nov	12-Dic
Young	VI	VI	35	35	10-Nov	02-Dic
Treinta y Tres	VI	VI	35	35	19-Nov	24-Dic

2.3.4. Upscaling method

Following van Bussel et al. (2015), each simulated cropping system by soil type combination was weighted by the crop specific contribution to the harvested area within each RWS to estimate average Yw and Yp. Averages were calculated separately for full-season soybean (*i.e.*, a full-season soybean crop) and double-cropped soybean (*i.e.*, a soybean crop sowed immediately after harvest of a winter cereal crop). Annual Ya was calculated based on Ya reported for the CREA zone where RWS was located. Yp, Yw and Ya were upscaled to CZ and country levels, based on the relative contribution of each RWS to total crop-specific harvested area. Yg was calculated as the difference between Yw and Ya, across all spatial scales (*i.e.*, RWS, CZ and country). Yp and Yw were used to estimate calculate the water limitation index defined by Aramburu Merlos et al. (2015).

2.3.5 Boundary-function analysis

Total in season rainfall effect on maximum Yw for full-season and double-cropped soybean separately, were determined by boundary-function analysis. Boundary-function was calculated by means of 90th quantile regression as follows:

$$Yw_{90} = a_{pot}\{1 - \exp\left[\frac{-(X - c_{pot})}{b_{pot}}\right]\} \qquad \text{Eq. (2)}$$

Where Yw_{90} is the Yw at 90^{th} quantile, X represent the total in season rainfall, parameter a_{pot} is the absolute maximum Yw, b_{pot} is a measure of the potential response of Yw to total in season rainfall increments, and c_{pot} represent the minimum total in season rainfall for Yw concretion (i.e. threshold value). Data were analyzed using the *quantreg* package (Koenker et al., 2018) of R software (R Core Team, 2017).

2.3.6 Estimation of CS_{Yg}

CS_{Yg} was estimated in two directions. Horizontally, as de difference between full-season and double-cropped soybean at each yield level (*i.e.* Yp, Yw and Ya), to determine, respectively, CS_{Yg} due to change on solar radiation interception and thermal conditions, water availability and dynamics, and lack of fine tune of management practices.

$$CSY_g^{weather} = Y_p^{sj1} - Y_p^{sj2} \qquad \text{Eq. (3)}$$

where $CSY_g^{weather}$ is the CS_{Yg} due to change on solar radiation interception and thermal conditions, $Y_p^{sj_1}$ is the full-season soybean Yp and $Y_p^{sj_2}$ is double-cropped soybean Yp.

$$CSY_g^{weather-water} = Y_w^{sj1} - Y_w^{sj2} \qquad \text{Eq. (4)}$$

where $CSY_g^{weather-water}$ is the CS_{Yg} due to change on water availability and its dynamics, $Y_w^{sj_1}$ is the full-season soybean Yw and $Y_w^{sj_2}$ is double-cropped soybean Yw.

$$CSY_g^{weather-water-manage} = Y_a^{sj1} - Y_a^{sj2} \qquad \text{Eq. (5)}$$

where $CSY_g^{weather-water-manage}$ is the CS_{Yg} due to lack of fine tune of management practices, $Y_a^{sj_1}$ is the full-season soybean Ya and $Y_a^{sj_2}$ is double-cropped soybean Ya.

And diagonally, taking full-season soybean immediately upper yield level as reference, to determine the combined CS_{Yg}, as follows:

$$CSY_g^{weather-water} = Y_p^{sj1} - Y_w^{sj2} \qquad \text{Eq. (6)}$$

where $CSY_g^{weather-water}$ refers to the combined CS_{Yg} due to change on solar radiation interception, thermal conditions and water availability and dynamics; $Y_p^{sj_1}$ is the full-season soybean Yp; and $Y_w^{sj_2}$ is the double-cropped soybean Yw.

$$CSY_g^{weather-water-manage} = Y_w^{sj1} - Y_a^{sj2} \qquad \text{Eq. (7)}$$

where $CSY_g^{weather-water-manage}$ refers to the combined CS_{Yg} due to change on water availability and dynamics and the lack of fine tune on management practices; $Y_w^{sj_1}$ is the full-season soybean Yw; and $Y_a^{sj_2}$ is the double-cropped soybean Ya.

2.4. RESULTS

2.4.1. Selected weather stations and crop area coverage

Harvested soybean area in Uruguay averaged *ca.* 1.1 Mha during 2010-2016. Spatial distribution of soybean area was highly concentrated in the west-littoral, where most agricultural suitable land are placed (Fig. 3). Given the concentration of agriculture in this country, just five RWS where sufficient to cover 71% of national soybean harvested area. Furthermore, the three CZ into which the RWS where located, accounted for 97% of the national harvested area for soybean.

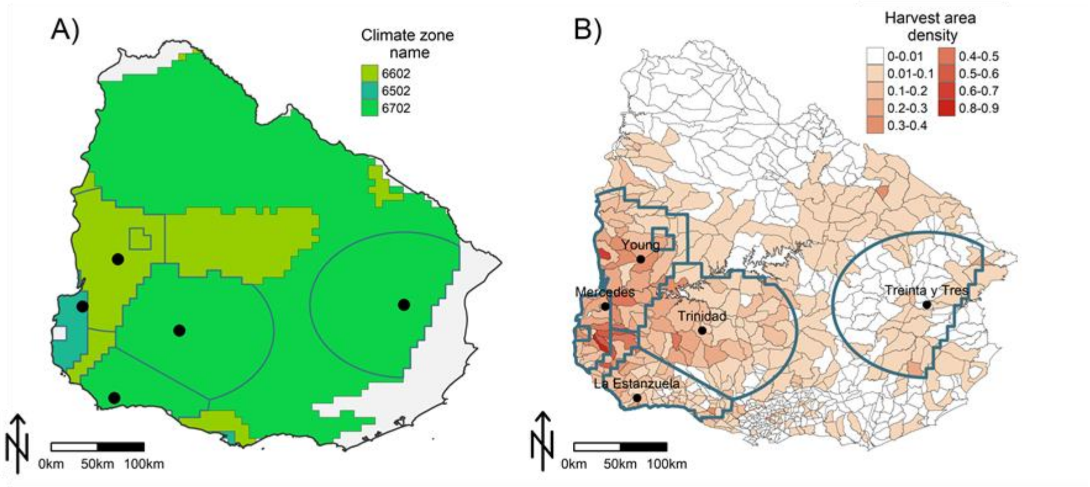

Fig. 3. *Maps of Uruguay showing (A) selected climate zones (designated by its original numeration), reference weather station (filled points) and its buffer zones (line delimited areas around weather station); and (B) soybean harvested area density per enumeration-area (as proportion of the total enumeration-area area), from the 2011 agricultural census* (DIEA, 2011).

2.4.2 Spatial and temporal variation in yield potential

At national level, average Yp was 6.6 and 5.7 Mg ha^{-1} for full-season and double-cropped soybean. The highest Yp for full-season soybean and double-cropped soybean was found in the CZ 6602 (6.7 and 5.8 Mg ha^{-1}, respectively). CZ 6502 showed a Yp of 6.5 and 5.7 Mg ha^{-1} for full-season and second crop soybean respectively (Fig. 3). Lowest Yp for full-season and double cropped soybean was found in the CZ 6702 (6.5 and 5.6 Mg ha^{-1}, respectively). Double cropped soybean Yp was consistently lower than full-season soybean, with an average difference across CZ of 13%. However, Yp showed low inter annual variation (CV lower than 10%) both for full-season soybean and double-cropped soybean.

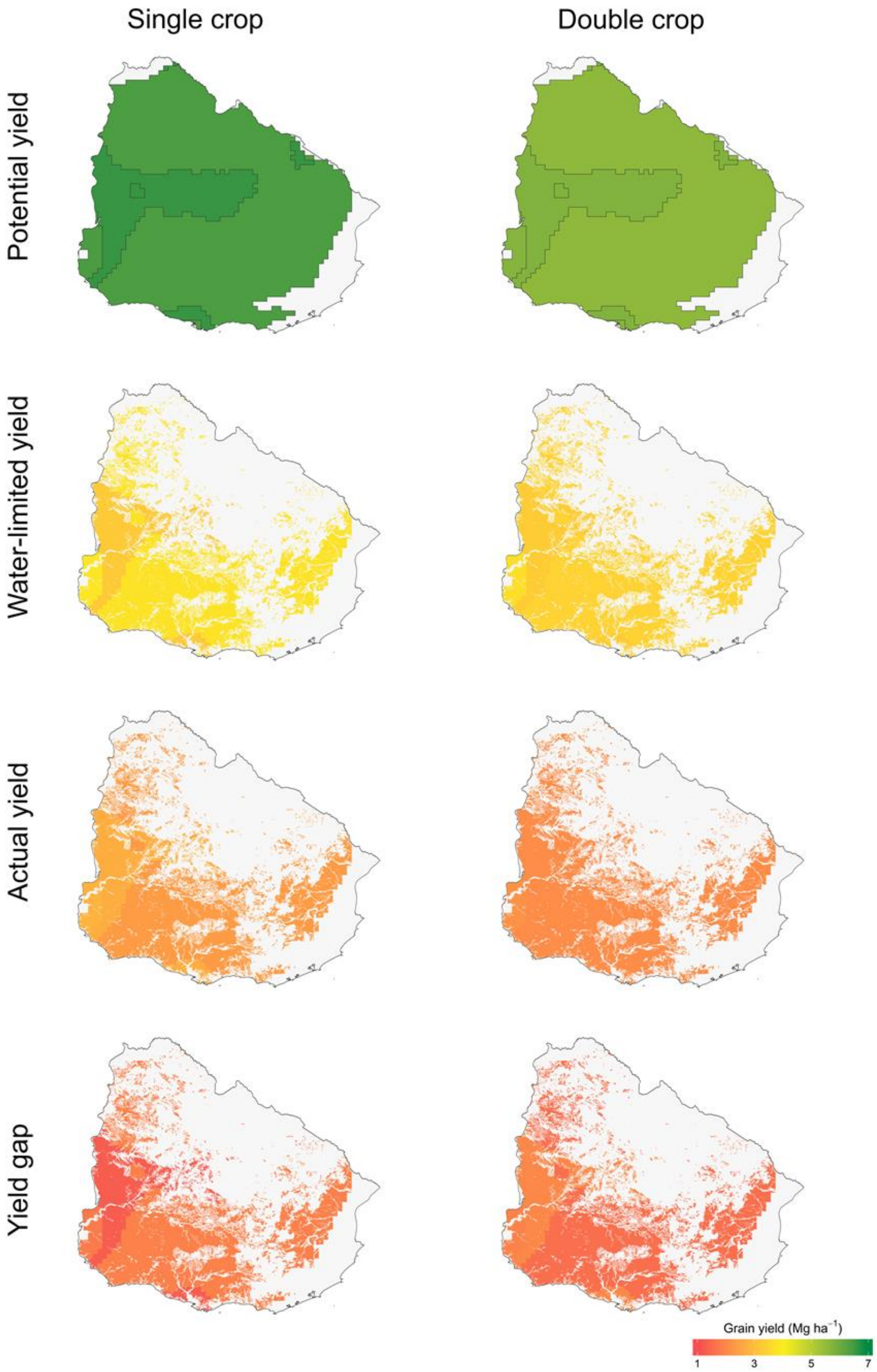

Fig. 4. *Potential yield, water-limited yield, actual yield and yield gap (Mg ha*$^{-1}$*) at climate zone level, for full-season (left column) and double-cropped soybean (right column). Colored area into each climate zone refers to soils accounted in simulations.*

2.4.3 Spatial and temporal variation in water-limited yield

National average Yw was 3.6 and 3.4 Mg ha^{-1} for full-season and double-cropped soybean, respectively. Full-season soybean Yw was highest in the southwest and decreased toward the northwest from 3.8 Mg ha^{-1} in the CZ 6502 to 3.2 Mg ha^{-1} in the CZ 6602 (Fig. 4). Difference between double-cropped and full-season soybean vary inconsistently across CZ. While double cropped soybean Yw in CZ 6702 was 34% lower than full-season soybean, in the CZ6602 the result was opposite being Yw for second crop soybean 5% higher than full-season crop soybean.

Differences were found on temporal variation through the CZ but being consistent between both soybean cropping systems. CZ 6602 showed the highest temporal variability (CV equal to 64 and 61% for full-season and double-cropped soybean) and CZ 6702 resulted in the lowest temporal variation (CV = 42%, for full-season and double-cropped soybean) (Fig. 5). A negative correlation was found between Yw and interannual-variability (r = -0.61, p = 0.059).

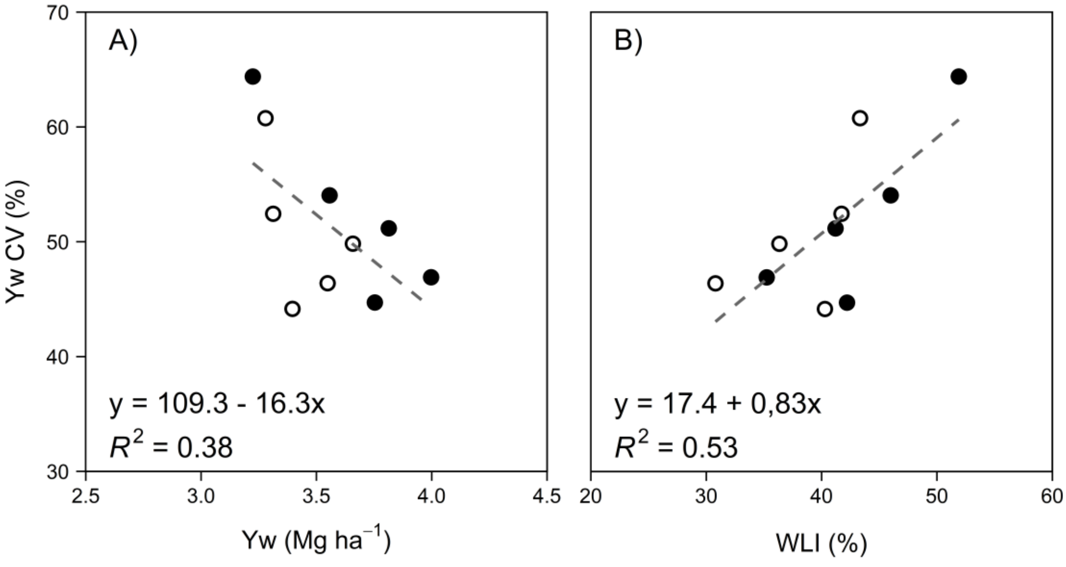

Fig. 5. *Coefficient of variation (CV, in %) of water limited yield potential (Yw) calculated for each reference weather station, as function of (A) Yw (Mg ha^{-1}) and (B) water limitation index (WLI, i.e., difference between yield potential and water-*

limited yield, expressed as percentage of yield potential), for full-season (empty circles) and double-cropped soybean (filled circles).

2.4.4. Producer-reported yield and its variability

National average Ya calculated by upscaling method resulted in 2.5 and 2.0 Mg ha^{-1} for full-season and double cropped soybean respectively. The agreement between these results and the values reported by the Uruguayan Ministry of Agriculture (MGAP by its Spanish acronym), confirm the robustness of the applied method to upscale the result from RWS to larger geographical areas (Fig. 6). Full-season soybean showed consistently higher Ya (ranging from 2.3 to 2.7 Mg ha^{-1} in CZ 6702 and 6502 respectively) and less stable (CV ca. 11%) than double cropped soybean Ya (2.0 to 2.1 Mg ha^{-1}, CV 5 to 11% for CZ 6702 and 6502 respectively) (Fig. 3).

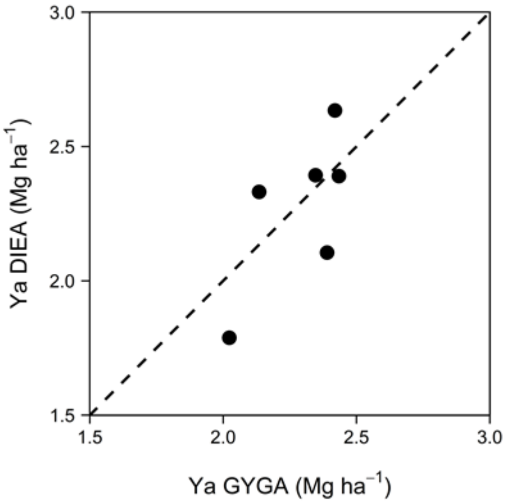

Fig. 6. *Soybean national average actual yield (Ya, Mg ha^{-1}) reported by Direction of Agricultural Statistics (DIEA, 2017) as function of national average Ya estimated following bottom-up upscaling method defined by Global Yield Gap Atlas Protocol, for the period 2010-2016.*

2.4.5. Comparison between Yw for full-season and double-cropped soybean

The maximum Yw at high rainfall (i.e. parameter *a*) was smaller for double-cropped soybean (5.8 Mg ha^{-1}) than for full-season soybean (6.7 Mg ha^{-1}) (Fig. 7). The response of Yw to increments in rainfall (i.e. parameter *b*) was 0.1 Mg ha^{-1} mm^{-1} higher in full-season than in double-cropped soybean. Parameter *c* (i.e. threshold value for soybean yield concretion) was higher for double-cropped soybean (236 mm) than for full-season soybean (207 mm). Yw for double-cropped soybean was higher than Yw for full-season soybean in 35% RWS-year situations, averaging 0.5 Mg ha^{-1} higher Yw than full-season soybean, being the same value that average gain when Yw full-season soybean was higher than double-cropped soybean.

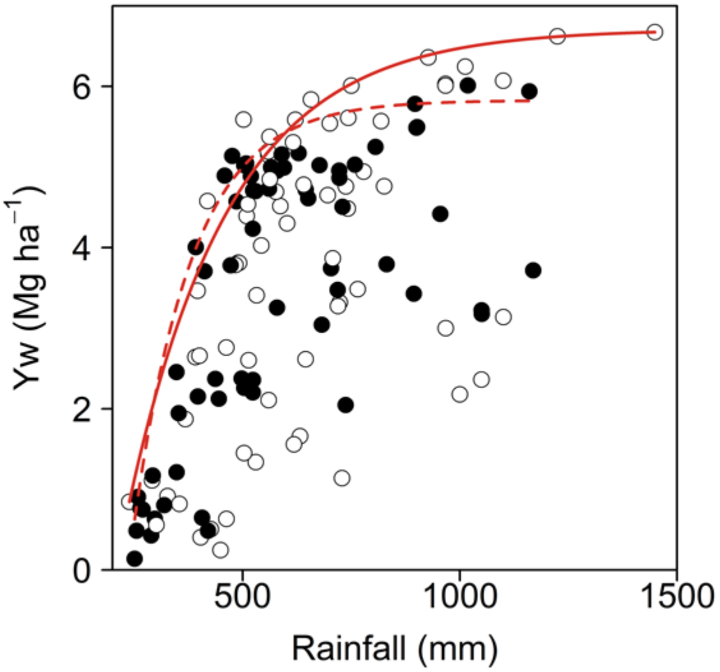

Fig. 7. *Water-limited soybean yield (Mg ha^{-1}) versus total rainfall (from planting to harvest, mm), for full-season (empty circles) and double-cropped soybean (filled circles). Each datapoint corresponds to a RWS-year combination, considering 2003-2015 period and RWS shown in figure 1. Separate polynomial models were*

fitted to the 90th percentiles for full-season (solid) and double-cropped soybean (dashed line).

2.4.6. Quantifying CS_{Yg} on soybean

We estimated that double-cropped soybean had a loss of 0.9 Mg ha^{-1} of its Yp, which were explained by the change on the environmental variables to which the crop was exposed (solar radiation and thermal conditions), representing 28% of the combined $CSY_g^{weather}$ (Fig. 8). We found that full-season soybean had an Yg due to water of 3.0 Mg ha^{-1} while for double-cropped soybean it was equal to 2.3 Mg ha^{-1}, being 0.2 Mg ha^{-1} explained by the change on the water availability and/or its dynamic by the change on the crop system. CSY_g explained by management was estimated on 0.5 Mg ha^{-1}.

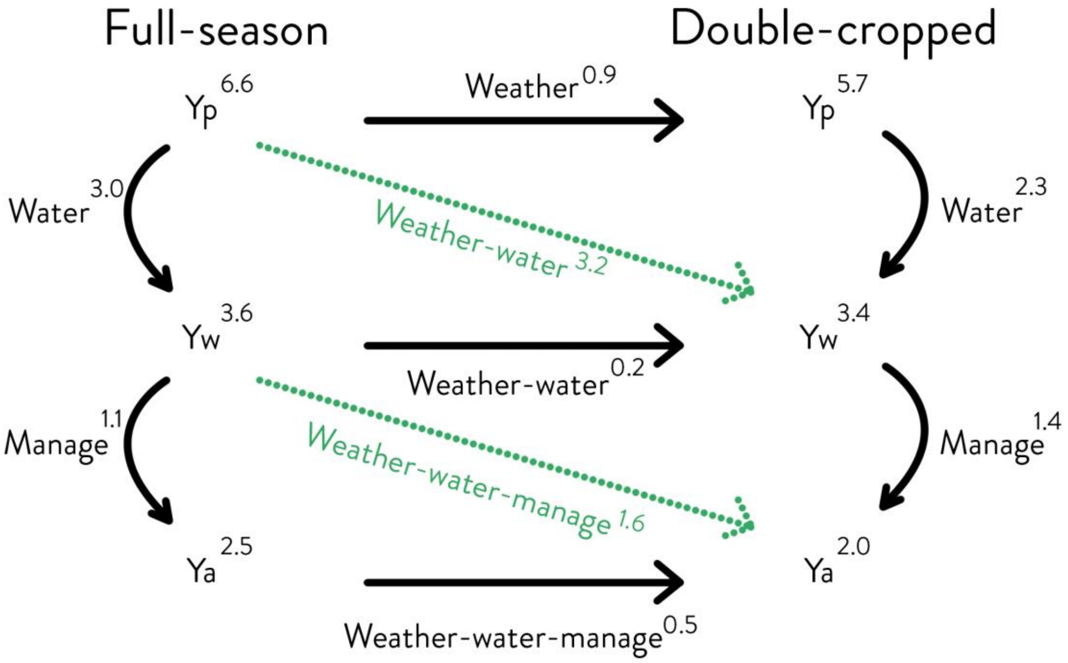

Fig. 8. *Full-season and double-cropped soybean non-limited yield potential (Yp), water-limited yield potential (Yw) and actual yield (Ya) results. Schematic representation of its relationships and variables explaining the movement through*

each level (black arrows). Green dotted arrows show the copping system-imposed yield gap between full-season soybean and double-cropped soybean.

2.5. DISCUSSION

In the present study, we found spatial variability on Yp, Yw, Ya and Yg across the country at CZ level. The spatial variation due to CZ effect was higher in full-season soybean than in double-cropped soybean and notably consistent through all analyzed yield levels. Full-season soybean Yp estimated in our analysis compares well with those quantified by Giménez (2017) from field experiments under non-limiting water conditions. The estimates of Yw and Yg shown in the present study were consistent with those reported by Aramburu Merlos et al. (2015) for soybean in Argentinian climate zones applying the same methodology. Indeed, when analyzed regionally, we found that shared climate zone showed a similar Yw pattern ranging from 4.0 Mg ha^{-1} in Argentinian side to 3.8 Mg ha^{-1} in Uruguayan side of the same climate zone, which highlight the spatial consistency of the applied method. Being the AI the only variable, from those used by van Wart et al. (2013b) to define climate zones, that varyed across the analyzed climate zones, it seems to be related to solar radiation and thermal conditions.

From the boundary function analyses, we found that the maximum Yw for full-season and double-cropped soybean was 0.1 Mg ha^{-1} higher to its average at national level Yp. This result indicates that on seasons with the most favorable conditions, even in rainfed systems, soybean Yp could be reached by Yw. In addition, double-cropped soybean has a higher rainfall threshold for grain yield setting, probably explained by the scarce amount of stored soil water at wheat harvest (Andrade and Satorre, 2015; Hamblin and Tennant, 1987) and the short fallow period before double-cropped soybean sowing date. Due to its higher Yw and its lower grain yield setting threshold, full-season soybean has a higher rate of increment on Yw per millimeter of increment of rainfall than double-cropped soybean. The lower response on grain yield to rainfall on double-cropped soybean could be also explained because part of the rainfall amount was used to refill soil profile. Another remarkable result of this study was that although double-cropped soybean had lower Yp than full-season soybean, its Yw was higher than Yw for full-season soybean on 35% of the years. In those regions with so variable in-season rainfall, to grow soybean in a full-season soybean cropping system it is not guarantee of the highest yields. Thus, to diversify cropping systems is the best tool we have to deal with that uncertainty even with the seasonal ENSO forecasts (Bert et al., 2006).

The lower Yp for double-cropped soybean than Yp for full-season soybean could be explain by its late sowing date that reduced capture of solar radiation and its use efficiency (Calviño et al., 2003; Monzon et al., 2014). Double-cropped soybean Ya also resulted lower than full-season soybean, but the difference between full-season and double-cropped soybean were higher in Ya than in Yw (0.5 Mg ha^{-1} and 0.2 Mg ha^{-1} for Ya and Yw, respectively), which was explained by the differences on crop management practices (apart from sowing date) that farmers apply to each soybean cropping system. For example, double-cropped soybean is usually underfertilized and exposed to low-phosphorus environments (Hoffman Berasain, 2013). To the high phosphorus demand of the soybean crop (Salvagiotti et al., 2013, 2004) and low content in Uruguayan soils (Hernández et al., 1995), when soybean is grown as second crop, the immobilization induced by stubble decomposition is added, increasing the phosphorus deficit for this crop (Singh and Singh, 1994). Poor control measures of the yield reducing factors, affecting mainly to double-cropped soybean, could explain not only the gap between full-season and double-cropped soybean, but also the national Yg of each one separately. However, to elucidate the Yg causes require a detailed management practices database analysis, combined with crop simulation models and appropriate spatial framework.

Temporal variation on soybean Yw was another remarkable result of our work, being low Yw associated with high interannual variability and *vice versa*. This temporal variation resulted higher than those reported for Argentina (Aramburu Merlos et al., 2015), may be explained by the high variability of the rainfall and the lower water holding capacity of the Uruguayan soils. Soil maps in Uruguay were descripted in an edaphological sense, but no data on the real rooting depth for those soils has been published till the present. May be, soybean crops (among other) can explore a deeper zone of the soil beyond the descripted soil depth, because they have the potential ability to grow root up to 2 m (Dardanelli et al., 1997; Ordóñez et al., 2018).

Although full-season soybean had a higher maximum Yw than double-cropped one, by the possibility to explore higher radiation environment and the higher amount of water stored in soil at the beginning of the cropping season, this behavior is not consistent from year to year (Andrade and Satorre, 2015; Monzon et al., 2007). In this study, double-cropped soybean resulted in a higher Yw than full-season soybean in 35% of the cases, with the same average gain than full-season soybean. These finding highlight the implications that the cropping sequence design have on yields and shows clearly that in rainfed condition there are years in which double-cropped soybean is better than full-season soybean. Double-cropped soybean became more competitive than full-season soybean if the previous cereal crop yields are considered. Increasing crop sequence intensity has been proposed as promissory measure to increase global grain production given the benevolent weather conditions of some regions like those on South American Pampas, that allow to grow more than one crop per year.

We found that $CSY_g^{weather}$ was responsible for the 60% of the total horizontal yield gap (*i.e.* sum of $CSY_g^{weather}$, $CSY_g^{weather-water}$ and $CSY_g^{weather-water-manage}$), showing clearly that delaying sowing date generates a dramatical fall of double-cropped Yp (Fig. 8). We believe that to create more intensified cropping systems it is necessary to close this cropping-systems imposed yield gap, by maximizing the capture and use-efficiency of resources in an environment of reduced supply. Some alternatives to do so, could be the following: (i) improve spatial arrange by reducing inter-row distance, (ii) increase plant population, (iii) improve plan architecture to capture higher amount of incident solar radiation, (iv) select cultivars more adapted to dose radiation and thermal conditions, and (v) identify the best maturity group for each system. The cropping system imposed a small yield gap due to change on soil-stored water amount and/or its dynamic. This small yield gap explained by water together with inter-year variability on in-season rainfall explain that 35% of years double-cropped soybean yields were higher than full-season soybean. However, into each cropping system the Yg explained by water was the highest, showing that water was the most important limiting factor. $CSY_g^{weather-water-manage}$ was relatively high, and yield gap explained by management into each cropping system was higher in double-cropped soybean compared with full-season soybean. Being full-season and double-cropped soybean two different crops growing under different environments, they should be managed accordingly to close Yg by management, but there is a dearth of knowledge on the double-cropped soybean best management practices.

Surprisingly, for rainfall amount over 750 mm we found a wide distance between datapoints and boundary line, even being our study based on simulated data. The inefficiencies in this scenario can be attributed to shallow soils, low capacity of water storage, distribution of the rainfall, and as suggested by Andrade and Satorre (2015), less favorable thermal conditions and solar radiation. There is a dearth of knowledge around the causes of those inefficiencies and proposals on how to predict those situations and neither on how to manage it. We believe that the following three areas should be prioritized in the research agenda: (i) identifying the agronomical variables related to inefficiencies; (ii) understanding the underlying processes behind that variables, since the variability on Yw drives the variability on Ya; (iii) development of management practices strategies so farmers bring their yields closest to potential of its environments, with a significant reduction of the risk.

Agriculture has become one of the most important productive sector in Uruguay, as consequence of its recent expansion and intensification, producing on average 2.4 Mt soybean grain annually, with high inter annual variability (CV 25%) (DIEA, 2017). Soybean is the third most exported product, accounting for 13% of the total exports (3,2 Mt during 2017; OPYPA, 2017). By closing national average Yg calculated for the 2005–2016 period (40 and 50% of Yw for full-season and double-cropped soybean, respectively) to an attainable level of 20% of Yw, Uruguay could have potentially produced an extra 0.5 Mt of soybean on the existing soybean harvested area. If this increase in production achieved through closing Yg were exported, which is very likely due to the low domestic demand, it would have represented an increase of 225 million dollars per year in the national income (21% of Uruguay total agriculture gross domestic product).

2.6. CONCLUSIONS

The present yield gap analysis shows that Uruguay had the potential to increase its full-season and double-cropped soybean grain yields by a respective 1.5 and 1.7 Mg ha^{-1}. By closing it yield gap to an attainable level, Uruguay could have a substantial increase on the national income, without expanding actual cropland area. Magnitude and variability of Yw and Yg varied across the soybean cropped area. Double-cropped soybean resulted with highest Yw than full-season soybean on 35% of the RWS-year situations, with the same average marginal gain for both soybean types. Solar radiation and thermal conditions were responsible for the 60% of the total horizontal cropping system-imposed yield gap.

2.7. REFERENCES

Alexandratos, N., Bruinsma, J., 2012. World Agriculture towards 2030/2050: the 2012 revision 154.

Andrade, J.F., Satorre, E.H., 2015. Full-season and double crop systems in the Argentine Pampas: Environmental determinants of annual grain yield. Field Crops Res. 177, 137–147.

Aramburu Merlos, F., Monzon, J.P., Mercau, J.L., Taboada, M.A., Andrade, F.H., Hall, A.J., Jobbagy, E., Cassman, K.G., Grassini, P., 2015. Potential for crop production increase in Argentina through closure of existing yield gaps. Field Crops Res.

Bert, F.E., Satorre, E.H., Toranzo, F.R., Podestá, G.P., 2006. Climatic information and decision-making in maize crop production systems of the Argentinean Pampas. Agric. Syst. 88, 180–204.

Bruinsma, J., 2009. The resource outlook to 2050: by how much do land, water and crop yields need to increase by 2050?, Expert Meeting on How to Feed the World in 2050. FAO, Rome.

Calviño, P.A., Sadras, V.O., Andrade, F.H., 2003. Development, growth and yield of late-sown soybean in the southern Pampas. Eur. J. Agron. 19, 265–275.

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

Caviglia, O.., Sadras, V.., Andrade, F.., 2004. Intensification of agriculture in the south-eastern Pampas. Field Crops Res. 87, 117–129.

Dardanelli, J.L., Bachmeier, O.A., Sereno, R., Gil, R., 1997. Rooting depth and soil water extraction patterns of different crops in a silty loam Haplustoll. Field Crops Res. 54, 29–38.

DIEA, 2017. Anuario Estadístico Agropecuario 2017. DIEA-MGAP. http://www.mgap.gub.uy/unidad-organizativa/oficina-de-programacion-y-politicas-agropecuarias/publicaciones/anuarios-diea/anuario-2011. Last checked August 2018.

DIEA, 2011. Census 2011. DIEA-MGAP. http://www.mgap.gub.uy/. Last checked Febrary 2018.

Evans, L.T., 1993. Crop Evolution, Adaptation and Yield. Cambridge University Press, Cambridge, UK.

Evans, L.T., Fisher, R.A., 1999. Yield potential: Its definition, measurement, and significance. Crop Sci. 39, 1544–1551.

Fischer, T., Byerlee, D., Edmeades, G., 2014. Crop yields and global food security: will yield increase continue to feed the world? ACIAR Monograph No. 158. Australian Centre for International Agricultural Research. Canberra. xxii–634.

Gijsman, A.., Jagtap, S.., Jones, J.., 2002. Wading through a swamp of complete confusion: how to choose a method for estimating soil water retention parameters for crop models. Eur. J. Agron. 18, 77–106.

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

Guilpart, N., Grassini, P., Sadras, V.O., Timsina, J., Cassman, K.G., 2017. Estimating yield gaps at the cropping system level. Field Crops Res. 206, 21–32.

Hamblin, A., Tennant, D., 1987. Root length density and water uptake in cereals and grain legumes: how well are they correlated. Aust. J. Agric. Res. 38, 513.

Hernández, J., Otegui, O., Zamalvide, J.P., Agronomía, U. de la R. (Uruguay) F. de, 1995. Formas y contenidos de fósforo en algunos suelos del Uruguay, Boletín de investigación. Universidad de la República, Facultad de Agronomía.

Hoffman Berasain, E.M., 2013. Criterios y estrategias de fertilización que apunten a asegurar el suministro de fósforo en cultivos de soja de segunda. Informaciones Agronómicas de Hispanoamérica. 11 (Septiembre 2013), 2–7.

Hoogenboom, G., Porter, C.H., Sheila, V., Boote, K.J., Singh, U., White, J.W., Hunt, L.A., Ogoshi, R., Lizaso, J.I., Koo, J., Asseng, S., Singels, A., Moreno, L.P., Jones, J.W., 2015. Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.7 (https://DSSAT.net). DSSAT Foundation, Gainesville, Florida, USA.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. Eur. J. Agron. 18, 235–265.

Koenker, R., code), S.P. (Contributions to C.Q., code), P.T.N. (Contributions to S.Q., code), A.Z. (Contributions to dynrq code essentially identical to his dynlm, code), P.G. (Contributions to nlrq, advice), B.D.R. (Initial (2001) R. port from S. (to my everlasting shame--how could I. have been so slow to adopt R.

and for numerous other suggestions and useful, 2018. quantreg: Quantile Regression.

Liang, W., Carberry, P., Wang, G., Lü, R., Lü, H., Xia, A., 2011. Quantifying the yield gap in wheat–maize cropping systems of the Hebei Plain, China. Field Crops Res. 124, 180–185.

Mercau, J.L., Dardanelli, J.L., Collino, D.J., Andriani, J.M., Irigoyen, A., Satorre, E.H., 2007. Predicting on-farm soybean yields in the pampas using CROPGRO-soybean. Field Crops Res. 100, 200–209.

Monzon, J.P., Mercau, J.L., Andrade, J.F., Caviglia, O.P., Cerrudo, A.G., Cirilo, A.G., Vega, C.R.C., Andrade, F.H., Calviño, P.A., 2014. Maize–soybean intensification alternatives for the Pampas. Field Crops Res. 162, 48–59.

Monzon, J.P., Sadras, V.O., Abbate, P.A., Caviglia, O.P., 2007. Modelling management strategies for wheat–soybean double crops in the south-eastern Pampas. Field Crops Res. 101, 44–52.

Monzon, J.P., Sadras, V.O., Andrade, F.H., 2012. Modelled yield and water use efficiency of maize in response to crop management and Southern Oscillation Index in a soil-climate transect in Argentina. Field Crops Res. 130, 8–18.

Ordóñez, R.A., Castellano, M.J., Hatfield, J.L., Helmers, M.J., Licht, M.A., Liebman, M., Dietzel, R., Martinez-Feria, R., Iqbal, J., Puntel, L.A., Córdova, S.C., Togliatti, K., Wright, E.E., Archontoulis, S.V., 2018. Maize and soybean root front velocity and maximum depth in Iowa, USA. Field Crops Res. 215, 122–131.

OPYPA, 2017. Oleaginosos y derivados: situación y perspectivas. http://www.mgap.gub.uy/unidad-organizativa/oficina-de-programacion-y-politicas-agropecuarias/estadisticas-y-documentos/30-20. Last checked August 2018.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rao, M.R., Willey, R.W., 1983. Effects of Genotype in Cereal/Pigeonpea Intercropping on the Alfisols of the Semi-Arid Tropics of India. Exp. Agric. 19, 67.

Ritchie, J.T., Crum, J., 1989. Converting soil survey characterization data into IBSNAT crop model input, in: Symposium Organized by the International Society of Soil Science (ISSS), Wageningen (Netherlands), 22-26 Aug 1988. Pudoc.

Salvagiotti, F., Barraco, M., Dignani, D., Sanchez, H., Bono, A., Vallone, P., Gerster, G., Galarza, C., Montoya, J., Gudelj, V.J., 2013. Plant stand, nodulation and seed yield in soybean as affected by phosphate fertilizer placement, source and application method. Eur. J. Agron. 51, 25–33.

Salvagiotti, F., G, G., Bacigaluppo, S., C, G., N, G., V, G., P, V., Juárez, M., 2004. Efectos residuales y directos de fósforo y azufre en el rendimiento de soja de segunda. Cienc. Suelo 22, 92–101.

Singh, H., Singh, K.P., 1994. Nitrogen and phosphorus availability and mineralization in dryland reduced tillage cultivation: Effects of residue placement and chemical fertilizer. Soil Biol. Biochem. 26, 695–702.

Tittonell, P., 2014. Ecological intensification of agriculture — sustainable by nature. Curr. Opin. Environ. Sustain. 8, 53–61.

van Bussel, L.G.J., Grassini, P., Van Wart, J., Wolf, J., Claessens, L., Yang, H., Boogaard, H., de Groot, H., Saito, K., Cassman, K.G., Van Ittersum, M.K., 2015. From field to atlas: Upscaling of location-specific yield gap estimates. Field Crops Res. 177, 98–108.

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. Field Crops Res. 143, 4–17.

van Ittersum, M.K., Rabbinge, R., 1997. Concepts in production ecology for analysis and quantification of agricultural input-output combinations. Field Crops Res. 52, 197–208.

Van Opstal, N.V., Caviglia, O.P., Melchiori, R.J.M., 2011. Water and solar radiation productivity of double-crops in a humid temperate area. Aust. J. Crop Sci. 5, 1760.

van Wart, J., Grassini, P., Cassman, K.G., 2013a. Impact of derived global weather data on simulated crop yields. Glob. Change Biol. 19, 3822–3834.

van Wart, J., van Bussel, L.G.J., Wolf, J., Licker, R., Grassini, P., Nelson, A., Boogaard, H., Gerber, J., Mueller, N., Claessens, L., van Ittersum, M.K., Cassman, K.G., 2013b. Use of agro-climatic zones to upscale simulated crop yield potential. Field Crops Res. 143, 44–55.

White, J.W., Hoogenboom, G., Wilkens, P.W., Stackhouse, P.W., Hoel, J.M., 2011. Evaluation of Satellite-Based, Modeled-Derived Daily Solar Radiation Data for the Continental United States. Agron. J. 103, 1242–1251.

3. ASSESSING SOIL TYPES AND CLIMATIC ZONES INFLUENCE ON WATER-LIMITED YIELD OF FULL-SEASON AND DOUBLE-CROPPED SOYBEAN USING A STATISTICAL APPROACH[2]

Gonzalo Rizzo[a], Jorge Franco[b], Oswaldo Ernst[a]

[a] Departamento de Producción Vegetal, Facultad de Agronomía, Estación Experimental Mario Alberto Cassinoni, Universidad de la República, Paysandú, 60000, Uruguay

[b] Departamento de Biometría, Estadística y Computación, Facultad de Agronomía, Estación Experimental Mario Alberto Cassinoni, Universidad de la República, Paysandú, 60000, Uruguay

3.1. ABSTRACT

To close actual yield gaps supposes to identify it causes, but also estimates the magnitude of its effect on crop yields and understands its behavior. In the present study we developed a statistical scheme that allow us to identify the main sources of water-limited yield potential variability and elucidate its behavior. We apply our methodology in the Uruguayan full-season and double-cropped soybean cropping systems. Water-limited yield potentials were estimated using well validated crop simulation models across three climate zones and the soybean-cultivated soil types into those climate zones, using the most popular management practices for each site. Mixed models were used to identify the sources of variability, and random forest analysis was applied to identify the most important variables explaining the year effect and its behavior. Although climate zone only account for 1% of the total water-limited yield potential variability, there were significant differences between them from 0.2 to 0.5 Mg ha^{-1}. Soil type analyzed as a nested effect into the climate

---

[2] For publication on Field Crops Research journal.

zone captured 5 and 1% of the total variability, for full-season and double-cropped soybean respectively, and with significative differences between soil types. Year effect resulted the main source of variability on both soybean cropping systems (71 and 74% of total variability for full-season and double-cropped soybean respectively). For full-season soybean, soil water available at sowing date, water deficit during critical period, water availability during cropping season and solar radiation during critical period resulted the four most important variables. For double-cropped soybean the four most important variables were water deficit during critical period, soil water available at beginning of pod setting, solar radiation during critical period and water availability during cropping season. Risk classification showed the presence of few sites where there is low probability to reach water limited yield potential estimated at national level. Sites where it is likely to reach national level Yw in more than 70% of the years were present in the main soybean producing region and limited extensions on east of Uruguay. In the main soybean producing areas the probability to reach the estimated Yw at national level was in the 50-70% range.

Keywords: yield variability, crop simulation model, second crop, advanced statistics

3.2. INTRODUCTION

To guarantee future food security in a 9 billion people world, with limited availability of land suitable for crop production and low environmental footprint of agriculture (Cassman et al., 2003; Foley et al., 2011; Godfray et al., 2010), we are challenged to design new cropping systems under an ecological intensification of agriculture paradigm (Cassman, 1999; Tittonell, 2014). This paradigm suppose sustaining yield increases in the existing irrigated systems and favorable dryland systems, but also taking advantage of opportunities for greater cropping intensity in vast agricultural regions (Fischer et al., 2014).

Benefic effects of the double cropping systems have been previously reported such as increasing annual land productivity and stubble inputs (Amthor, 2000; Andrade et al., 2015; Calviño and Monzon, 2009; Caviglia et al., 2004; Francis and Smith, 1985; Graß et al., 2013), but also double crop yields are usually more stable than those of a full-season crops (Andrade and Satorre, 2015; Graß et al., 2013). However, double cropped soybean has lower yield potential due to the typical delayed sowing date of second crops (Andrade et al., 2015). Furthermore, double cropped soybean resulted less efficient in using water available than soybean as full-season crop and is usually nutritionally affected by the previous crop (Andrade et al., 2015).

The aim of this research was to analyze how the soil type and the climatic zone affect grain yield of full-season and double-cropped soybean. To achieve this objective, we focused on the following goals: (i) creating a statistical framework to assess the significance and composition of variance explained by the effects of soil type and climatic zone on grain yield, (ii) elucidate key drivers explaining these effects on full-season and double cropped soybean grain yield using a crop simulation model.

3.3. MATERIALS AND METHODS

3.3.1. Study area and site selection

The study area corresponds to soybean harvested area in Uruguay. Data was retrieved for each enumeration area (*i.e.* the smallest administrative unit in Uruguay) from the 2011 National Agricultural Census (DIEA, 2011).

Weather data sources selection and quality control followed the Global Yield Gap Atlas protocol (Grassini et al., 2015). Maximum and minimum air temperature, precipitation and relative humidity were derived from EEMAC (Mario A. Cassinoni Experimental Station from Faculty of Agronomy of the Univesidad de la República), INIA (National Institute of Agricultural Research) and INUMET (Uruguayan

Weather Institute) weather stations. Measured daily incident solar radiation was not available for the INUMET and INIA weather station, hence data from NASA-POWER (www.power.larc.nasa.gov) were used as source of daily incident solar radiation.

Reference weather stations (RWS) used in this study, were selected based on the harvested area within a buffer zone area of 100 km radius centered on each RWS and clipped by the climate zone (CZ) were the RWS were located (van Bussel et al., 2015). Each CZ corresponds to a particular combination of annual growing degree days (AGDD; base temperature 0°C), aridity index (AI, values increase for more humid conditions, and decrease with more arid conditions) and temperature seasonality (TS) (van Wart et al., 2013). RWS were iteratively selected starting with the one with largest harvested area coverage until reaching 50% of crop-harvested area and more than 70% coverage by the CZ where RWS were located. The three CZ used in this study presented the same range of AGDD (5990-7111) and TS (3833-8355), being different on their AI range: 8686-10181 (CZ 6702; *i.e.* CZ's name), 6589-7785 (CZ 6502) and 7786-8685 (CZ 6602).

Soils cultivated with soybean into each RWS were identified based on data provided by Natural Resources Directory of Uruguayan Ministry of Livestock, Agriculture and Fisheries (www.mgap.gub.uy/unidad-organizativa/direccion-general-de-recursos-naturales). Soils covering the 90% of the soybean harvested area from each RWS (10 to 14 soils series) were selected based on the CONEAT map (*i.e.* most detailed national-level soil map; 1:40,000). Functional soil properties required to run crop simulation models (e.g., field capacity and permanent wilting point) were derived from soil series descriptions following Ritchie and Crum (1989), after the revisions made by (Gijsman et al., 2002). Maximum rooting depth was allowed to reach the maximum soil depth of each soil, ranging from 0.6 to 2.3 m.

3.3.2. Water-limited yield potential estimation

Water-limited yields were estimated through simulations using CROPGRO-Soybean model embedded in DSSAT v4.6 (Hoogenboom et al., 2015; Jones et al., 2003). Genetic coefficients were derived from Mercau (2007) and Monzon (2012, 2007), and unpublished data from well-managed experiments. The model performance to simulate Yp and Yw was assesed by comparison of model simulated yield against measured yields from well-managed rainfed and irrigated experiments that explore a wide range of sowing dates, sites, years and water availability (chapter 2 of this thesis)

3.3.3. Simulated cropping systems and management

Actual yields (Ya) at subnational level and cropping system management practices data (*e.g.* sowing date, cultivars, plant density and crop sequence) are not publicly available in Uruguay. Hence, Ya and most popular manage practices were retrieved from 2009-2016 agronomic records from farmers of the Uruguayan Federation of Regional Consortia of Agricultural Experimentation (FUCREA using its Spanish acronym, www.fucrea.org/), and derived for each RWS based on farm locations. Retrieved information include: average planting dates, dominant cultivar name and maturity, and actual and optimal plant population density (Table 1). The provided data were subsequently corroborated by national experts. Dominant crop rotations and proportion of each of them to the total harvested area was calculated based on cropping sequences reported in the Planes de Uso de Suelos (Information provided by DGRN Ministerio de Ganadería Agricultura y Pesca; www.mgap.gub.uy/unidad-ejecutora/direccion-general-de-recursos-naturales).

Following Aramburu Merlos et al. (2015), to account for the differential water availability of each soil at sowing of each crop through the studied period, we assumed 50% of the plant available soil water in the first year of the time series. Simulated crop sequences were: (i) 1-year cover crop/soybean (*i.e.* continuous soybean), (ii) 2-year cover crop/soybean-wheat/soybean double crop, and (iii) 3-year cover crop/soybean-wheat/maize double crop-soybean except for Treinta y Tres. Atmospheric CO2 concentration was set constant at 400 ppm.

Table 1. *Soybean (full-season and double crop) management practices, as retrieved from 2009-2016 agronomic records from farmers of the Uruguayan Federation of Regional Consortia of Agricultural Experimentation (FUCREA, www.fucrea.org), applied to estimate water-limited and potential yields at each reference weather station (RWS).*

	Maturity group		Plant density (plant m-2)		Sowing date	
RWS	Full-season	Double	Full-season	Double	Full-season	Double
La Estanzuela	VI	VI	35	35	10-Nov	12-Dic
Trinidad	VI	VI	35	35	10-Nov	14-Dic
Mercedes	VI	VI	35	35	10-Nov	12-Dic
Young	VI	VI	35	35	10-Nov	02-Dic
Treinta y Tres	VI	VI	35	35	19-Nov	24-Dic

3.3.4. Data analysis

After performing simulations, a total of 3835 complete data sets for both soybean cropping system were ready for use in the analysis. Each field was a result of the combination of soybean cropping system (full-season and double-cropped soybean), cropping sequence (2-3; as descripted on section 3.3.3), RWS (5 selected RWS), soil type (10-18 selected soil types) and year (13 years). Not all RWS presented the same number of cropping sequence (*i.e.* in Treinta y Tres only two cropping sequences were found), and each RWS varied in the number of soils (Supplementary Table 1).

Linear nested models were fitted separately for full-season and double-cropped soybean cropping system, to assess the soil type, CZ and year effects on Yw. Into each model, the soil types were treated as nested fixed effects into CZ effect, because soil types are present into each CZ covering a relative low portion of it, and therefore, each soil type add a positive or negative effect to CZ effect to which it belongs. Soil type effect was nested into CZ in the statistical model and year was represented as a fixed effect, being the three variables considered as factors. To estimate the variance composition, the linear model was transcripted to its homolog random effect model.

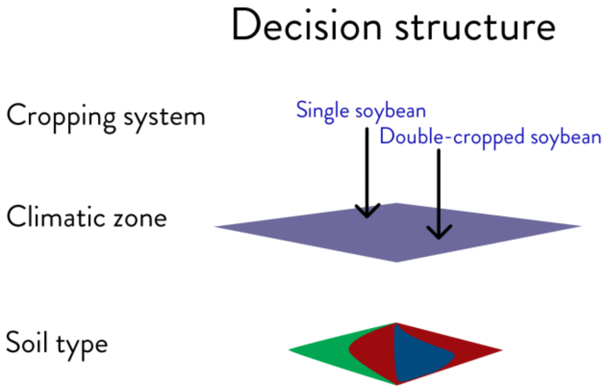

Fig. 1. *Graphical representation of the decision structure. Cropping system decision are options that farmers can choose every year and even into the cropping season. Soil types could be chosen only before the crop is sowed, but climatic zone where it is placed is not a possible decision. Thus, soil type must be referenced to the climatic zone governing its climate.*

Differences between CZ and differences between soil types were evaluated for each soybean cropping system separately, by means of contrasts using fixed effect model. Contrasts were performed in R (R Core Team, 2017) using *emmeans* package (Lenth, 2018). Statistical differences ($P < 0.05$) between means were determined by a *t*-tests.

Relative importance of climatic and ecophysiological attributes explaining year effect were assessed with random forest (RF) algorithm (Breiman, 2001) implemented in the *randomForest* package (Liaw and Wiener, 2002) in R software environment (R Core Team, 2017). RF is a machine-learning technique method for classification and regression that construct a large set of decision trees with each tree trained by selecting a random set of variables and a random sample from the training data set (Breiman, 2001). Contribution of agricultural practices, climatic, echophysiological and edaphoclimatic attributes were analyzed separately for full-season and double-cropped soybean. In this analysis, we evaluated the effect of 23 variables (Table 2). Sowing date and crop sequences were inputs for the model, while the others were created or captured from the outputs of the model.

After identifying the four key variables behind year effect, partial dependence plots were used to show (graphically) the marginal effect of a full-season variables in random forest estimates of the Yw (Friedman, 2001). Partial dependence plot on a predictor variable provides the average trend of that variable (integrating all other predictors in the model). RF is not a tool for traditional statistics, it does not compute *P* values, or regression coefficients, or confidence intervals. RF identify variables that strongly affects yield variation, but this does not mean that there are significant differences in yields. Partial dependence plots were analyzed with segmented models to improve its interpretation by defining breaking points and slopes for each segment. Segmented models were constructed in R environment using *segmented* package (Muggeo, 2008, 2003).

Table 2. *Variables related to management practices, climatic, ecophysiology and edaphoclimatic, assessed in the random forest analysis.*

Type	Variable	Unit
Agricultural practice	Crop sequence	
	Sowing date	Day of year
Climatic	Cumulated precipitations from sowing date to R3	mm
	Total solar radiation during critical period	MJ m^{-2}
	Total solar radiation during cropping season	MJ m-2
	Maximum air temperature during cropping season	°C
	Maximum air temperature during critical period	°C
	Minimum air temperature during cropping season	°C
	Minimum air temperature during critical period	°C
	Cumulated precipitations during critical period	mm
	Cumulated precipitations during cropping season	mm
	Water deficit during cropping season	mm
	Water deficit during critical period	mm
	Precipitation deficit during cropping season	mm
	Precipitation deficit during critical period	mm
Ecophysiological	Emergency date	Day of year
	Maturity date	Day of year
	Cropping season duration	Day of year
	Critical period duration	Days
Edaphoclimatic	Soil available water at R3	mm
	Soil available water at sowing date	mm
	Available water during critical period	mm

	Available water during cropping season	mm

3.3.5. Risk definition

Risk was defined for full-season and double-cropped soybean as the probability to reach its estimated Yw at national level by previous Uruguay yield gap assessment (3.6 and 3.4 Mg ha^{-1} for full-season and double-cropped soybean respectively, see Section 2). Estimation of the probability was made as following:

$$Risk_{S*X}\ (\%) = \frac{\sum_{i=1}^{N}(if\ Yw_i > Yw_{nac}\ ,\ then\ Yw_i = 1\ else,\ Yw_i = 0)}{N} \times 100 \qquad \text{Eq. (1)}$$

where $Risk_{S*RWS}$ is the estimated risk expressed as percentage for a given soil type *S* located into the RWS *X*, Yw_i is the estimated Yw mean for the year *i*, Yw_{nac} is the estimated Yw at national level, and N is the total number of estimated Yw_i.

Finally, risk estimations were classified into the following four classes: < 30%, 30 – 50%, 50 – 70%, and > 70%. The results of the classification were mapped.

3.4. RESULTS

3.4.1. Overall model

The fitted model explained 76 and 77% of the total water-limited yield potential variance for both, full-season and double-cropped soybean, respectively (Table 3). These models were highly significant (p-value < 0.05, for either full-season and double-cropped soybean). Nested effect of soil type into CZ and year effect were significant (p-value < 0.05) in both soybean cropping systems. Year effect explained the 71% of the variance in the model for full-season soybean, while for double-cropped soybean it explained 74% of the total variance of the model (Table 3). Soil type nested into CZ represented 5 and 1% of model variance for full-season and double-cropped soybean respectively. CZ explained only 1% of model variance for both soybean cropping systems.

Table 3. *Percentage of the total variance of the model explained by each parameter.*

Parameter	Full-season	Double-cropped
Year	71%	74%
CZ	1%	1%
Soil type nested in CZ	5%	1%
Residual	23%	24%

3.4.2. Climatic zone effect into full-season and double cropped soybean water-limited yield potential

Full-season soybean estimated means were 3.7, 3.9 and 3.4 Mg ha^{-1} for CZ1, CZ2 and CZ3, respectively. Contrasts performed for the three CZ in full-season soybean were statistically significant (p-value < 0.01, for the three contrasts) (Table 4). Estimated means for double-cropped soybean were 3.3, 3.6 and 3.3 for CZ1, CZ2 and CZ3, respectively. CZ2 showed the highest estimated yield, being statistically different from CZ1 and CZ3 (p-value < 0.05).

Table 4. *Full-season and double-cropped soybean contrasts between estimated means for each climate zone (CZ). Results were averaged over soil types and years. Tukey method was used for comparison.*

Soybean	Contrast	Estimate	SE	DF	T ratio	p-value
Full-season	CZ1 - CZ2	-0.272	0.048	2999	-5.686	<0.01
	CZ1 - CZ3	0.246	0.043	2999	5.789	<0.01
	CZ2 - CZ3	0.518	0.054	2999	9.660	<0.01
Double-cropped	CZ1 - CZ2	-0.373	0.087	750	-4.297	<0.01
	CZ1 - CZ3	-0.065	0.076	750	-0.855	0.67

	CZ2 - CZ3	0.308	0.101	750	3.057	<0.01

3.4.3. Nested effect of soil type into climate zone on water-limited yield potential for full-season and double-cropped soybean

Estimated means for soil types nested into climate zone ranged from 2.9 (soil type G10.2 in CZ1 and soil type 10.15, 03.40, 11.3 and 11.5 from CZ3) to 5.2 Mg ha^{-1} (soil type 03.51 in CZ2) for full-season soybean (Fig. 2 y 3). Estimated means for double-cropped soybean resulted in a lower and narrower range, from 2.7 (soil type G10.2 in CZ1) to 4.2 Mg ha^{-1} (soil type 03.51 in CZ2 and soil type 11.4 in CZ3) (Fig. 2). For full-season soybean 108 contrasts between soils types belonging to same CZ resulted in significant difference (p-value < 0.05). However, only one significant difference (p-value < 0.05) between soil types from the same CZ were found in double-cropped soybean (soil type 5.02b was 1 Mg ha^{-1} higher than soil type 12.11, p-value = 0.02) and four contrasts resulted in a significant difference for different soil types from different CZs.

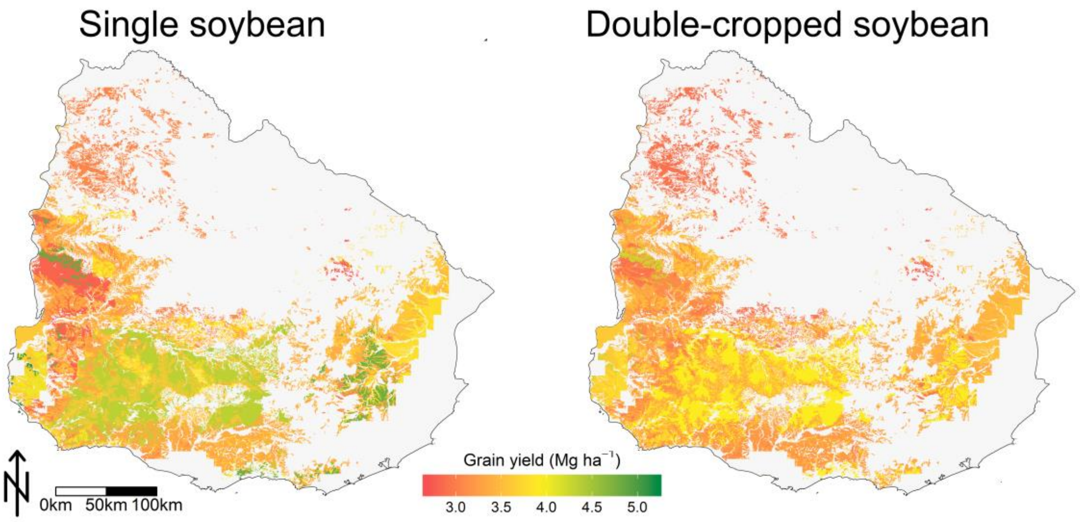

Fig. 2. *Water limited yield (Mg ha^{-1}) for soil type nested into climate zone.*

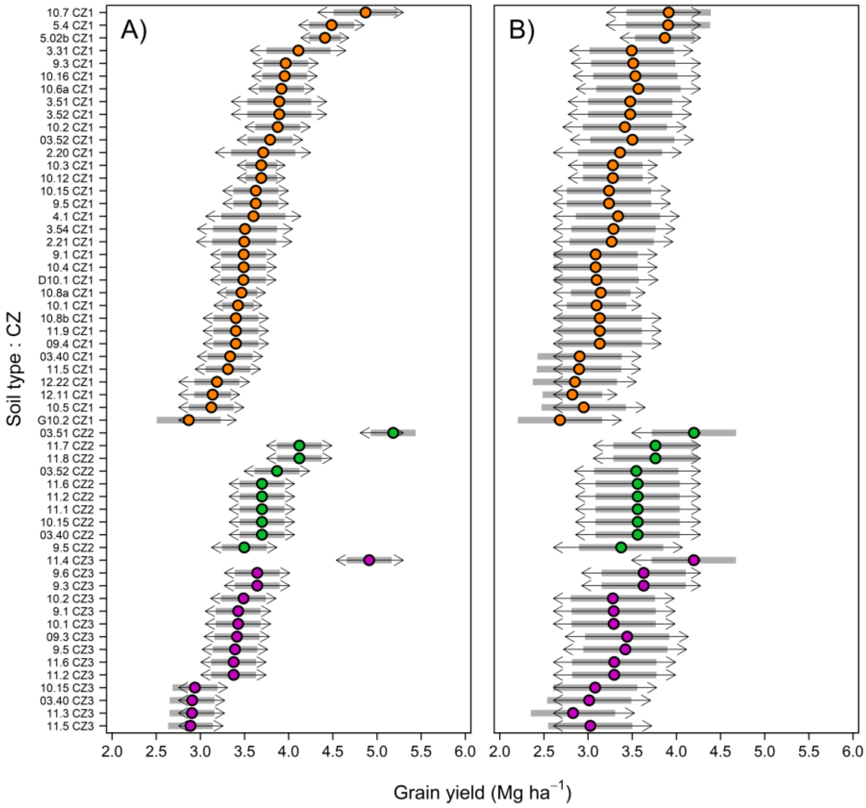

Fig. 3. *Estimated water-limited yield (Mg ha^{-1}) for soil type nested into climate zone for (A) full-season and (B) double-cropped soybean. Gray bars are confidence intervals for estimated marginal means (EMMs), and black arrow range shows confidence intervals for the comparison among them. When arrow from one mean overlap an arrow from another group, the difference is not significant, based on Tukey (α = 0.05).*

3.4.4. Elucidating year effect into full-season and double cropped soybean

The RF model using agricultural practices, climatic, echophysiological and edaphoclimatic attributes for full-season soybean explained 99.8% of the total variance, while for double-cropped soybean the RF model with the same attributes caught 99.2% of the variance. Variables explaining grain yield ranked differently on its importance for each soybean cropping systems (Fig. 4). For full-season soybean the five most important variables were: i) soil available water at sowing date (SWXD_SD, in mm; 68% IncMSE), ii) water deficit during critical period (WDPC, in mm; 64% IncMSE), iii) water available on the cropping season (AWSEASON, in mm; 49% IncMSE), iv) total solar radiation during critical period (RADPC, in MJ m^{-2}; 48% IncMSE), and v) available water during critical period (AWPC, in mm; 45% IncMSE) (Fig. 5A). For double-cropped soybean top five on variables importance were: i) water deficit during critical period (WDPC, in mm; 51% IncMSE), ii) soil available water at beginning of critical period (SWXD_R3, in mm; 44% IncMSE), iii) total solar radiation during critical period (RADPC, in MJ m^{-2}; 41% IncMSE), iv) available water during critical period (AWPC, in mm; 35% IncMSE), and v) soil available water at sowing date (SWXD_SD, in mm; 32% IncMSE).

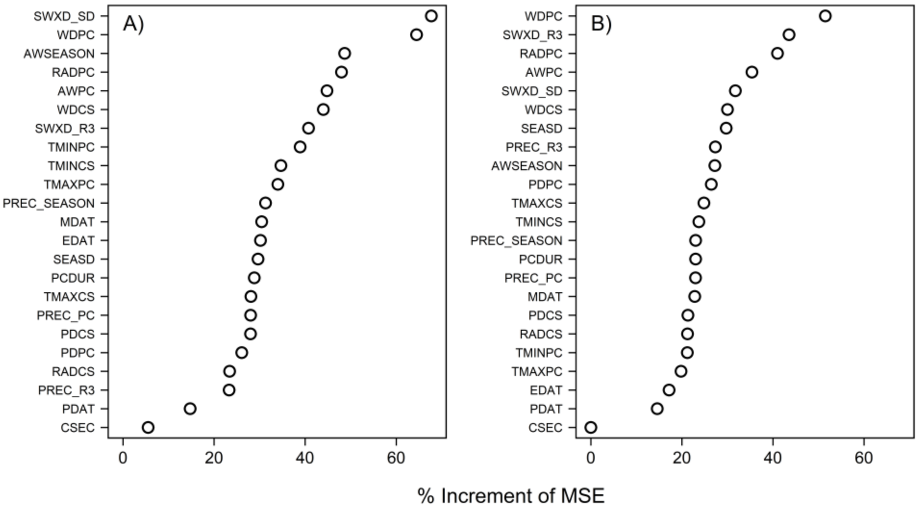

Fig. 4. *Main agricultural practices, climatic, echophysiological and edaphoclimatic attributes ranked by the percentage increase in mean squared error (higher % increment in mean squared error, higher variable importance), generated by random forest analysis for (A) full-season and (B) double-cropped soybean.*

Partial dependence plots are useful to show the marginal effect of each variable, but usually the relationships become hard to explain, in this case, segmented linear regressions can help to analyze the behavior of this relationships. In full-season soybean case, soil available water at sowing date showed a very low variation on grain yield (lower than 0.2 Mg ha^{-1}), but with a linear increase from 100 to 250 mm lower than 0.001 Mg ha^{-1} mm^{-1} (indicating that soils with higher water-storage capacity helps to increase soybean yields) (Fig. 5A). Yw was negatively related to water deficit during the critical period. Grain yield dropped rapidly with increasing water deficit in critical period and then leveled off. Cropping season water balance showed a linear relation with grain yield from 1000 to 1200 mm. Total solar radiation during critical period showed a linear relationship with grain yield. In double-cropped soybean water deficit during the critical period had a similar behavior than in full-season soybean, but with a dramatic drop from 0 to 50 mm (b = - 0.029), and then leveled off. Soil water stored at R3 showed a linear and positive increase until 120 mm and then it stabilized. Incident solar radiation in critical period of double-cropped soybean showed a steep jump on grain yield from 3.0 Mg ha^{-1} to 3.6 Mg ha^{-1} when moved from 950 to 970 MJ m^{-2}. Yw had a rapidly increase from zero to 220 mm of available water during the cropping season and then leveled off.

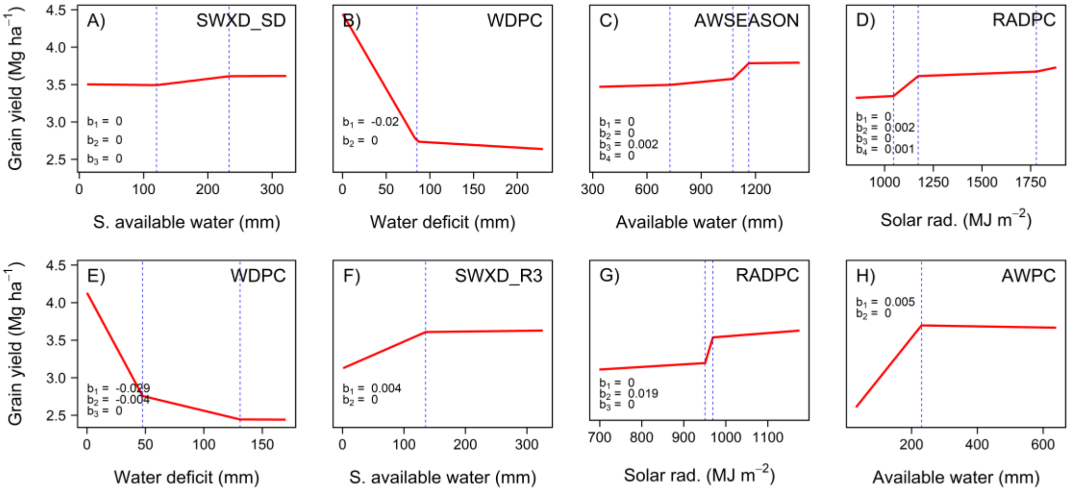

Fig. 5. *Partial dependence plot representing the marginal effect of four most important variables for full-season (A-D) and double-cropped soybean (E-H). The most important variables for full-season soybean from random forest (measured as percentage of increment of mean squared error) were: water deficit during critical period (i.e. WDPC = ETo – ETc, in mm), soil water available at sowing date (SWXD_SD, in mm), total solar radiation during critical period (RADPC, in MJ m^{-2}) and cropping season water balance (i.e soil water available at sowing date, plus rainfall during cropping season, minus soil water available at harvest date;*

AWSEASON in mm). While for double-cropped soybean most important variables were the same as full-season soybean except soil water available at sowing date, that was substituted by soil water available at the beginning of critical period (SWXD_R3, in mm).

Interaction between water deficit during critical period and total solar radiation during critical period showed a nonlinear relationship of the grain yield, but the effect of each of these variables was approximately the same for each value of the other variable (Fig. 6). Thus, the effect of the two variables are approximately additive.

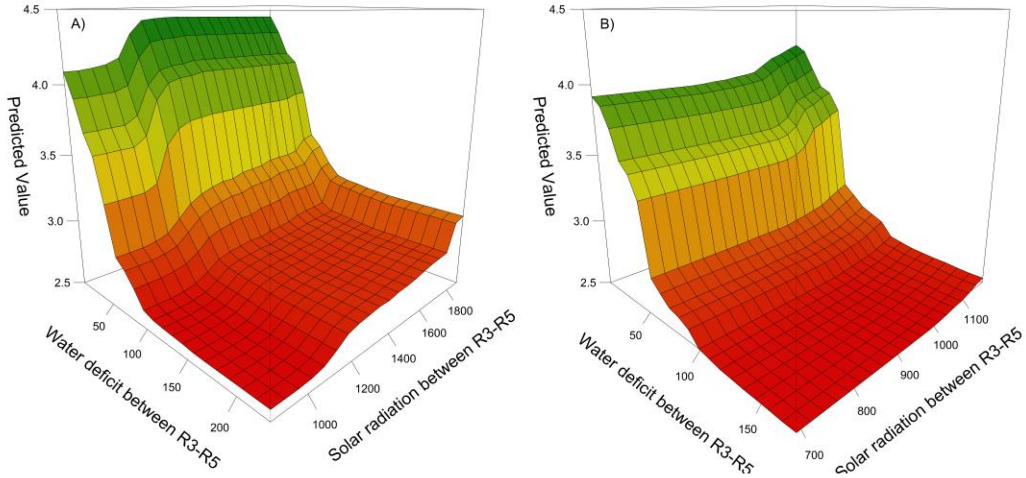

Fig. 6. *Bivariate partial dependence plot for grain yields (Mg ha^{-1}) for full-season (A) and double-cropped soybean (B). Interaction between water deficit during critical period (ETo – ETc; in mm) and cumulated solar radiation during critical period (MJ m^{-2}), estimated by the partial dependence estimates by random forest model.*

3.4.5. Risk based zones definition

Across Uruguay we found that dominant locations where soybean is grown has the possibility to reach the estimated Yw at national level over 50% of years (classified between 50-70% risk, light green zones Fig. 7). Sites that yields higher than long-term national Yw in less than one third of the year (red zones, Fig. 7), were present only for full-season soybean cropping system and covering insignificant soybean harvested areas. However, sites that usually (more than 70% of years) exceed national Yw covered larger areas and some of them were presented on the main soybean producing region for both soybean cropping systems, representing ≈ 5% of total Uruguay soybean harvested area.

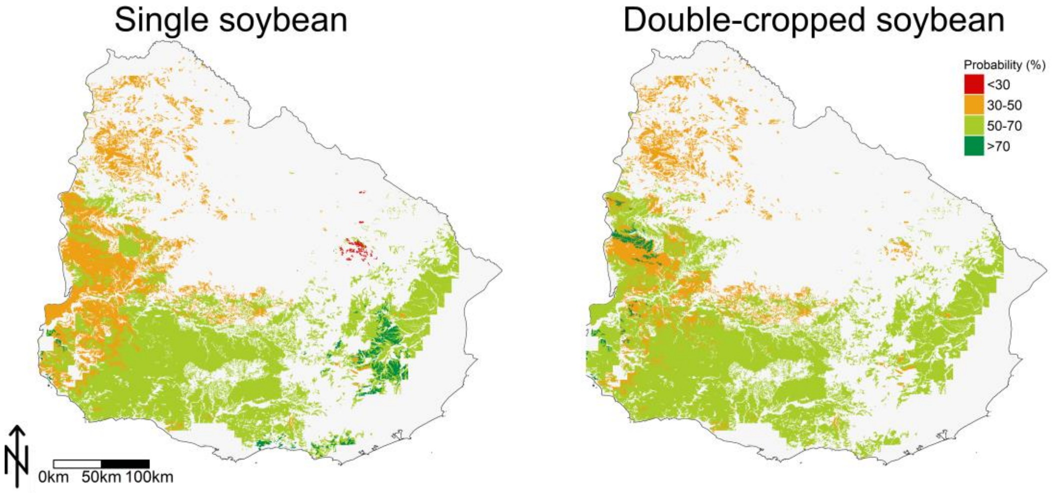

Fig. 7. *Zones defined by percentage of years that soybean grain yield exceeds long-term average national Yw.*

3.5. DISCUSION

In the present research we developed a novel statistical approach to estimate full-season and double-cropped soybean Yw and deepen on the understanding of its causes, using Uruguayan soybean cropping systems as case of study. Crop yield variation across climate zone has been previously reported (Aramburu Merlos et al., 2015; Mourtzinis et al., 2018; Niang et al., 2017; Rattalino Edreira et al., 2017), but here we estimate that climate zones were responsible for just 1% of full-season and double-cropped Yw variability (Table 3) having each climate zone a different effect on full-season soybean Yw. However, in double-cropped soybean the differences between climate zones became less clear. Lower yield potential due to the delayed sowing date in double-cropped soybean shift reproductive growth to less favorable conditions including shorter, cooler days, less solar radiation and higher probability of frost damage (Calviño et al., 2003; Egli and Bruening, 2000) explaining the narrower differences between climatic zones.

Considering the number of soils necessary to cover the 90% of soybean harvested area, allows us to account for a wide range of soils types where soybean is sown in Uruguay. Here, we found that soil type accounted for five percentage of the variability on full-season soybean Yw and only for one percentage on double-cropped soybean Yw, but even generated significant differences in Yw for both soybean cropping systems. Although, previous studies have stated the importance of edaphic attributes on crops Yw and Yg (Aramburu Merlos et al., 2015; Niang et al., 2017), this is the first time that significant differences on soybean Yw due to soil types are reported. Furthermore, the significant differences found on soybean Yw attributed to soil type nested in CZ suggest the needs for more research on possible technologies to be applied to groups of similar soil types. These would lead to better strategies to reduce risk and maximize profit in each site. Analyze across regions, with an appropriate quantification of the soil effect could be helpful to better understanding the actual cropping systems situations, and identification of sites where more research/extension is needed to efficiently reach the estimated ceiling yields.

Although place where the field is located (climate zone and soil type) define a portion of Yw magnitude and could locate it in a position with higher comparative advantages, our study showed that the grater source of variability is the year effect, which explain almost the 75% of soybean Yw variability (Table 4). Some earlier studies considering a similar number of years have reported temporal variation as an important source of variability on soybean yields (Aramburu Merlos et al., 2015; Grassini et al., 2014). Understanding how the variables explaining this effect operates, we can design more adaptative management practices to get the highest yields given the year weather conditions. Random forest analysis resulted a powerful tool for this analysis allowing to rank variables by its importance and the understand the overall effect of each variable in the full-season and double-cropped soybean Yw (Carranza and Laborte, 2015; Mutanga et al., 2012; Vincenzi et al., 2011). Our analysis showed that variables related to soil attributes and water were the most important, just being solar radiation in the critical period into the top four variables on importance. In full-season soybean, water stored in soil at sowing was the most important variable but explaining a variation lower than 0.5 Mg ha^{-1} of soybean Yw. Nevertheless, it has an important effect of the soil type because only very deep soil with an appropriate texture will be able to store a high amount of water that independize the soybean crop from rainfall. In double-cropped soybean water stored in soil at sowing did not appear as an important variable indicating that was not possible to store enough water in the soil at sowing, likely due to short period between harvest of previous crop and sowing of double-cropped soybean (Andrade and Satorre, 2015). In turn, water stored in soil at R_3 appeared in the second place of importance, so positive water balance before critical period in deep soil could generate a water budget that lead to higher double-cropped soybean yields. Unsurprisingly, water deficit during critical period was the variable capturing the wider range of soybean yield in both cropping systems (from 2.8 to 4.5 Mg ha^{-1} in full-season and 2.5 to 4.0 Mg ha^{-1} in double-cropped soybean). In full-season soybean water deficit in critical period showed a linear decrease on yield when moved from zero to 100 mm of water deficit. In double-cropped soybean there was a steep jump on yield when the water stress moved from zero to 50 mm.

Bivariate partial dependence plot is useful to understand complex interactions (Cutler et al., 2007), in this study we used it to investigate possible interactions between water deficit during critical period and solar radiation during critical period (Fig. 5). In full-season soybean the effects are almost additive, and there is not additional effect over the 1200 MJ m^2. In double-cropped soybean water deficit is the most important variable explaining yield variation with not change through solar radiation conditions.

Another main result of our study is a site-specific risk zonation based on the probability of reaching the estimated long-term national Yw and classified in four categories (Fig. 8). We found that sites into extreme categories (probability to reach national level Yw at least 30% or beyond 70% of the seasons), cover a small soybean harvested area portion. In the major area where soybean is grown we find sites where 50 to 70% of the seasons it is possible to reach national level Yw, and others where it can only be reached in 30 to 50% of the years. In a previous yield gap assessment for soybean in Uruguay, a yield gap of 40 to 50% of Yw was estimated for full-season and double-cropped soybean respectively (chapter 2 of this thesis). Thus, given that probabilistically the main soybean producing areas from Uruguay are near to the estimated soybean Yw at national-level, an attainable level of 80% of this Yw should be reached by removing nutritional limitations and reducing factor (diseases, weeds and pests).

3.6. CONCLUSIONS

In our study, we were able to attribute each portion of the total variance to its causes by means of a novel approach combining modern statistical methods. But maybe, the most important result is the crystal-clear explanations of the applied statistical methods' outputs. We found different variables related to each soybean cropping system, that help to understand the behavior of these crops Yw. Our work was based on simulated soybean cropping systems, and perhaps more lights could be shined by applying our methods on empirical data. The risk zonation developed here was simple and allow us to identify specific sites with low to high risk.

3.7. REFERENCES

Amthor, J.S., 2000. The McCree-de Wit-Penning de Vries-Thornley respiration paradigms: 30 years later. Ann. Bot. 86, 1–20.

Andrade, J.F., Poggio, S.L., Ermácora, M., Satorre, E.H., 2015. Productivity and resource use in intensified cropping systems in the Rolling Pampa, Argentina. Eur. J. Agron. 67, 37–51.

Andrade, J.F., Satorre, E.H., 2015. Single and double crop systems in the Argentine Pampas: Environmental determinants of annual grain yield. Field Crops Res. 177, 137–147.

Aramburu Merlos, F., Monzon, J.P., Mercau, J.L., Taboada, M.A., Andrade, F.H., Hall, A.J., Jobbagy, E., Cassman, K.G., Grassini, P., 2015. Potential for crop production increase in Argentina through closure of existing yield gaps. Field Crops Res.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.

Calviño, P., Monzon, J., 2009. Farming Systems of Argentina: Yield Constraints and Risk Management, in: Sadras, V.O., Calderini, D.F. (Eds.), Crop Physiology: Applications for Genetic Improvement and Agronomy. Elsevier, Amsterdam, pp. 55–70.

Calviño, P.A., Sadras, V.O., Andrade, F.H., 2003. Development, growth and yield of late-sown soybean in the southern Pampas. Eur. J. Agron. 19, 265–275.

Carranza, E.J.M., Laborte, A.G., 2015. Random forest predictive modeling of mineral prospectivity with small number of prospects and data with missing values in Abra (Philippines). Comput. Geosci. 74, 60–70.

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

Cassman, K.G., Dobermann, A., Walters, D.T., Yang, H., 2003. Meeting cereal demand while protecting natural resources and improving environmental quality. Annu. Rev. Environ. Resour. 28, 315–358.

Caviglia, O.., Sadras, V.., Andrade, F.., 2004. Intensification of agriculture in the south-eastern Pampas. Field Crops Res. 87, 117–129.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88, 2783–2792.

DIEA, 2011. Census 2011. DIEA-MGAP. http://www.mgap.gub.uy/. Last checked Febrary 2018.

Egli, D.B., Bruening, W.P., 2000. Potential of early-maturing soybean cultivars in late plantings. Agron. J. 92, 532–537.

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

Foley, J.A., Ramankutty, N., Brauman, K.A., Cassidy, E.S., Gerber, J.S., Johnston, M., Mueller, N.D., O'Connell, C., Ray, D.K., West, P.C., Balzer, C., Bennett, E.M., Carpenter, S.R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., Tilman, D., Zaks, D.P.M., 2011. Solutions for a cultivated planet. Nature 478, 337–342.

Francis, C.A., Smith, M.E., 1985. Variety development for multiple cropping systems. Crit. Rev. Plant Sci. 3, 133–168.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232.

Gijsman, A.., Jagtap, S.., Jones, J.., 2002. Wading through a swamp of complete confusion: how to choose a method for estimating soil water retention parameters for crop models. Eur. J. Agron. 18, 77–106.

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 327, 812–818.

Graß, R., Heuser, F., Stülpnagel, R., Piepho, H.-P., Wachendorf, M., 2013. Energy crop production in double-cropping systems: Results from an experiment at seven sites. Eur. J. Agron. 51, 120–129.

Grassini, P., Torrion, J.A., Cassman, K.G., Yang, H.S., Specht, J.E., 2014. Drivers of spatial and temporal variation in soybean yield and irrigation requirements in the western US Corn Belt. Field Crops Res. 163, 32–46.

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

Hoogenboom, G., Porter, C.H., Sheila, V., Boote, K.J., Singh, U., White, J.W., Hunt, L.A., Ogoshi, R., Lizaso, J.I., Koo, J., Asseng, S., Singels, A., Moreno, L.P., Jones, J.W., 2015. Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.7 (https://DSSAT.net). DSSAT Foundation, Gainesville, Florida, USA.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. Eur. J. Agron. 18, 235–265.

Lenth, R., 2018. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.2.2. https://CRAN.R-project.org/package=emmeans.

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 2, 18–22.

Mercau, J.L., Dardanelli, J.L., Collino, D.J., Andriani, J.M., Irigoyen, A., Satorre, E.H., 2007. Predicting on-farm soybean yields in the pampas using CROPGRO-soybean. Field Crops Res. 100, 200–209.

Monzon, J.P., Sadras, V.O., Abbate, P.A., Caviglia, O.P., 2007. Modelling management strategies for wheat–soybean double crops in the south-eastern Pampas. Field Crops Res. 101, 44–52.

Monzon, J.P., Sadras, V.O., Andrade, F.H., 2012. Modelled yield and water use efficiency of maize in response to crop management and Southern Oscillation Index in a soil-climate transect in Argentina. Field Crops Res. 130, 8–18.

Mourtzinis, S., Rattalino Edreira, J.I., Grassini, P., Roth, A.C., Casteel, S.N., Ciampitti, I.A., Kandel, H.J., Kyveryga, P.M., Licht, M.A., Lindsey, L.E., Mueller, D.S., Nafziger, E.D., Naeve, S.L., Stanley, J., Staton, M.J., Conley, S.P., 2018. Sifting and winnowing: Analysis of farmer field data for soybean in the US North-Central region. Field Crops Res. 221, 130–141.

Muggeo, V.M.R., 2008. segmented: an R Package to Fit Regression Models with Broken-Line Relationships. R News 8, 20–25.

Muggeo, V.M.R., 2003. Estimating regression models with unknown break-points. Stat. Med. 22, 3055–3071.

Mutanga, O., Adam, E., Cho, M.A., 2012. High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. Int. J. Appl. Earth Obs. Geoinformation 18, 399–406.

Niang, A., Becker, M., Ewert, F., Dieng, I., Gaiser, T., Tanaka, A., Senthilkumar, K., Rodenburg, J., Johnson, J.-M., Akakpo, C., Segda, Z., Gbakatchetche, H., Jaiteh, F., Bam, R.K., Dogbe, W., Keita, S., Kamissoko, N., Mossi, I.M., Bakare, O.S., Cissé, M., Baggie, I., Ablede, K.A., Saito, K., 2017. Variability and determinants of yields in rice production systems of West Africa. Field Crops Res. 207, 1–12.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rattalino Edreira, J.I., Mourtzinis, S., Conley, S.P., Roth, A.C., Ciampitti, I.A., Licht, M.A., Kandel, H., Kyveryga, P.M., Lindsey, L.E., Mueller, D.S., Naeve, S.L., Nafziger, E., Specht, J.E., Stanley, J., Staton, M.J., Grassini, P., 2017. Assessing causes of yield gaps in agricultural areas with diversity in climate and soils. Agric. For. Meteorol. 247, 170–180.

Ritchie, J.T., Crum, J., 1989. Converting soil survey characterization data into IBSNAT crop model input, in: Symposium Organized by the International Society of Soil Science (ISSS), Wageningen (Netherlands), 22-26 Aug 1988. Pudoc.

Tittonell, P., 2014. Ecological intensification of agriculture — sustainable by nature. Curr. Opin. Environ. Sustain. 8, 53–61.

van Bussel, L.G.J., Grassini, P., Van Wart, J., Wolf, J., Claessens, L., Yang, H., Boogaard, H., de Groot, H., Saito, K., Cassman, K.G., Van Ittersum, M.K., 2015. From field to atlas: Upscaling of location-specific yield gap estimates. Field Crops Res. 177, 98–108.

van Wart, J., van Bussel, L.G.J., Wolf, J., Licker, R., Grassini, P., Nelson, A., Boogaard, H., Gerber, J., Mueller, N., Claessens, L., van Ittersum, M.K., Cassman, K.G., 2013. Use of agro-climatic zones to upscale simulated crop yield potential. Field Crops Res. 143, 44–55.

Vincenzi, S., Zucchetta, M., Franzoi, P., Pellizzato, M., Pranovi, F., De Leo, G.A., Torricelli, P., 2011. Application of a Random Forest algorithm to predict spatial distribution of the potential yield of Ruditapes philippinarum in the Venice lagoon, Italy. Ecol. Model. 222, 1471–1478.

Supplementary Table 1. Soil type, relative contribution to crop harvested area within reference weather station (RWS) buffer, and main soil characteristics.

RWS	CONEAT	Profile	Soil type	Soil weight (%)	Depth (m)	Topsoil Texture	Subsoil texture
La Estanzuela	5.02b	M25-05	Typic Argiudol	20	1.4	Loam	Silty clay
	10.8b	N26-01	Typic Argiudol	17	0.9	Clay loam	Clay
	10.12	M24-04	Typic Argiudol	15	1.2	Clay	Clay
	10.3	M24-04	Typic Argiudol	15	1.2	Clay	Clay
	10.5	L27-10	Typic Argiudol	10	0.9	Loam	Clay
	10.6a	L28-03	Typic Argiudol	6	1.2	Clay loam	Silty clay
	10.8a	L27-14	Typic Argiudol	5	1	Clay loam	Clay
	10.1	N21-02	Typic Argiudol	5	1.1	Clay loam	Clay
	11.9	N26-01	Typic Argiudol	4	0.9	Clay loam	Clay
	09.4	N26-01	Typic Argiudol	2	0.9	Clay loam	Clay
	03.52	P17-08	Alfic Argiudol	2	1.1	Clay loam	Clay loam
Mercedes	11.7	P23-04	Typic Argiudol	32	1.4	Loam	Clay
	11.2	P19-03	Typic Argiudol	22	1.1	Loam	Silty clay
	11.6	P19-03	Typic Argiudol	14	1.1	Loam	Silty clay
	03.52	P17-08	Alfic Argiudol	8	1.1	Clay loam	Clay loam
	03.40	P19-03	Typic Argiudol	7	1.1	Loam	Silty clay
	03.51	O20-04	Typic Argiudol	5	2	Loam	Clay
	11.1	P19-03	Typic Argiudol	4	1.1	Loam	Silty clay
	11.8	P23-04	Typic Argiudol	4	1.4	Loam	Clay
	10.15	P19-03	Typic Argiudol	3	1.1	Loam	Silty clay
	9.5	N20-01	Typic Argiudol	2	1	Clay loam	Sandy clay
Treinta y Tres	10.7	USDA 13	Typic Argiudol	40	1.9	Loam	Clay loam
	4.1	D19-13	Alfic Argiudol	17	1	Loam	Clay loam
	2.21	E20-15	Abruptic Argiudol	11	1	Loam	Clay
	3.52	D21-57	Alfic Argiudol	9	1.2	Clay	Silty clay loam
	3.51	D21-57	Alfic Argiudol	6	1.2	Clay	Silty clay loam
	2.20	F23-19	Typic Argiudol	6	1.1	Loam	Clay loam
	3.54	E20-29	Abruptic Argiudol	4	1	Clay	Clay

	G10.2	F16-10	Typic Hapludol	4	0.8	Loam	Clay
	3.31	C21-51	Typic Endoaqualf	2	1.2	Loam	Silty clay loam
	12.11	M13-04	Typic Argiudol	1	0.9	Clay loam	Clay
Trinidad	5.02b	M25-05	Typic Argiudol	18	1.4	Loam	Silty clay
	10.12	M24-04	Typic Argiudol	17	1.2	Clay	Clay
	10.3	M24-04	Typic Argiudol	15	1.2	Clay	Clay
	10.1	N19-04	Typic Argiudol	11	0.8	Clay loam	Sandy clay loam
	10.16	K22-01	Typic Argiudol	7	1.2	Clay loam	Clay
	10.2	J20-03	Typic Argiudol	6	1.1	Loam	Clay loam
	5.4	M25-05	Typic Argiudol	4	1.4	Loam	Silty clay
	9.1	N19-04	Typic Argiudol	4	0.8	Clay loam	Sandy clay loam
	11.5	N21-03	Typic Argiudol	4	0.9	Loam	Clay
	12.11	M13-04	Typic Argiudol	2	0.9	Clay loam	Clay
	12.22	M13-04	Typic Argiudol	2	0.9	Clay loam	Clay
	10.4	N19-04	Typic Argiudol	2	0.8	Clay loam	Sandy clay loam
	10.15	N20-01	Typic Argiudol	2	1	Clay loam	Sandy clay
	9.5	N20-01	Typic Argiudol	2	1	Clay loam	Sandy clay
	D10.1	H20-04	Typic Argiudol	1	0.8	Loam	Clay loam
	10.8a	L27-14	Typic Argiudol	1	1	Clay loam	Clay
	9.3	O11-03	Typic Argiudol	1	1.1	Clay	Sandy clay loam
	03.40	O17-04	Typic Argiudol	1	0.8	Loam	Clay
Young	11.5	N21-03	Typic Argiudol	19	0.9	Loam	Clay
	10.1	N19-04	Typic Argiudol	9	0.8	Clay loam	Sandy clay loam
	10.15	O17-06	Typic Argiudol	9	0.8	Clay	Clay
	11.3	P15-02	Typic Argiudol	9	0.6	Clay	Clay loam
	11.4	P17-06	Petrocalcic Paleudol	9	2.3	Loam	Sandy loam
	11.6	P19-03	Typic Argiudol	9	1.1	Loam	Silty clay

10.2	O17-03	Typic Hapludol	6	1.4	Loam	Clay
03.40	O17-04	Typic Argiudol	6	0.8	Loam	Clay
11.2	P19-03	Typic Argiudol	6	1.1	Loam	Silty clay
9.1	N19-04	Typic Argiudol	4	0.8	Clay loam	Sandy clay loam
9.3	O11-03	Typic Argiudol	4	1.1	Clay	Sandy clay loam
9.5	N20-01	Typic Argiudol	3	1	Clay loam	Sandy clay
09.3	O18-03	Typic Argiudol	3	1	Loam	Sandy clay
9.6	O11-03	Typic Argiudol	2	1.1	Clay	Sandy clay loam

4. DISCUSIÓN GENERAL Y CONCLUSIONES

Al momento de iniciar esta tesis se comenzaba a implementar el protocolo desarrollado por el proyecto del Global Yield Gap Atlas en varios países de la región y el mundo. A su vez, disponíamos de muchos otros trabajos enfocados a la determinación de las brechas de rendimiento de soja (y otros cultivos) usando distintas metodologías. Sin embargo, en Uruguay se habían realizado varios ensayos de campo tendientes a cuantificar la respuesta al riego y los rendimientos potenciales, que daban idea de los rendimientos alcanzables, pero no se contaba con evaluaciones de las brechas de rendimiento del cultivo de soja. Por ende, este proyecto de maestría surge con el objetivo de generar de información original que sea de utilidad para todos los actores relacionados a la producción y comercialización de soja, y sea un paso importante hacia la intensificación ecológica de los sistemas agrícolas del país.

Para dar respuesta a los objetivos planteados en esta tesis, se aplicó el protocolo del Global Yield Gap Atlas para realizar las estimaciones de Yp, Yw, Ya y Yg a través de varias escalas, desde la zona buffer de cada estación meteorológica de referencia, pasando por escala de zona climática, hasta nivel nacional. Luego de estimados los diferentes niveles de rendimientos, se investigó el efecto que tienen los tipos de suelos, las diferentes zonas climáticas y el año. Para ellos se aplicaron diversos métodos estadísticos: (i) modelos lineales generales, para determinar la magnitud de cada uno de los efectos, (ii) modelos mixtos, con el fin de estimar la proporción de la varianza total que fue explicada por cada uno de los componentes, (iii) análisis con *random forest*, con el cual se estimó la importancia relativa de un conjunto de 23 variables (entre las cuales se encontraban, prácticas de manejo, climáticas, edafoclimáticas y ecofisiológicas), y (iv) modelos segmentados, para mejorar la interpretación de los resultados de los gráficos de dependencia parcial producidos por el *random forest*. Finalmente, en esta tesis se realizó una zonificación del riesgo a partir de las estimaciones de Yw. Los resultados fueron mapeados, para que la información generada sea fácilmente comprendida y pueda ser empleado por consultores, técnicos, productores, plantas de acopio, compañías abocadas a la comercialización, agentes aseguradores, políticos y agentes de los ministerios, entre otros.

4.1. APORTES ORIGINALES AL CONOCIMIENTO CIENTÍFICO

En esta tesis, hemos determinado que a nivel nacional el rendimiento potencial de soja en secano en Uruguay es de 3,6 Mg ha^{-1} para soja de primera y de 3,4 Mg ha^{-1} para soja de segunda. La brecha de rendimiento estimada a esta misma escala y para el período 2003-2015 resultó de un 40% para soja de primera y de un 50% para soja de segunda. Cerrar estas brechas de rendimiento, implicaría un incremento de 500.000 toneladas de grano de soja, en la superficie de soja actual (DIEA, 2016). Cuando se lo mira en el contexto de una región ecológica, resulta notable la consistencia espacial con los resultados publicados por Aramburu Merlos et al. (2015) para la Argentina. La variabilidad temporal constatada en nuestra tesis es coincidente con datos nacionales previamente publicados (Giménez, 2017; Giménez y García Petillo, 2011), lo que también fue constatado en Argentina (Aramburu Merlos et al., 2015). La estimación del índice de limitación hídrica (WLI, por su sigla en inglés), permite evaluar el grado en el que un ambiente se encuentra limitado por agua. Nuestras estimaciones del WLI a la escala de estaciones meteorológicas de referencia sugieren que los diferentes sitios tomados en cuenta para nuestro estudio difirieron en cuanto a grado de limitación por agua. A su vez, encontramos que ambientes donde la limitación hídrica es baja están siendo manejados ineficientemente y se dejan de aprovechar la mayor oferta de recursos que ellos presentan. Creemos que estas mayores brechas de rendimiento en los sitios con mejores aptitudes hídricas son explicadas por la ponderación del riesgo que hacen los productores, dada la baja capacidad predictiva de los escenarios climáticos para intervalos de 2-3 meses.

Hasta el presente, los estudios a nivel de sistema de cultivos básicamente se han enfocado en el rendimiento total del sistema, ya sea medido como rendimiento (Henderson et al., 2016; Hochman, Prestwidge, y Carberry, 2014; Liang et al., 2011) o bien como transformando el rendimiento en grano a su equivalente en unidades de energía (Guilpart, Grassini, Sadras, et al., 2017). Sin embargo, estos estudios no se han detenido a cuantificar el efecto que las decisiones de los productores al elegir uno u otro sistema de cultivo (es decir, soja de primera o soja de segunda), en el rendimiento de la soja de segunda (la cual presenta su fecha de siembra relegada a la cosecha del cultivo de invierno). En nuestro estudio, encontramos que el 60% de la del total de la brecha impuesta por el sistema de cultivo medido de forma horizontal, estuvo explicado por la caída en el Yp al retrasar la fecha de siembra (Andrade y Satorre, 2015; Monzon, Sadras, Abbate, y Caviglia, 2007). La brecha de rendimiento impuesta por el sistema debido al cambio en la oferta inicial y dinámica del agua sólo fue de 0,2 Mg ha^{-1}, lo cual es coincidente con el resultado de que el 35% de los años el Yw de la soja de segunda haya sido mayor al Yw de la soja de primera.

Las CZs han sido previamente reportadas como una fuente de variabilidad importante de los rendimientos (Aramburu Merlos et al., 2015; Mourtzinis et al., 2018; Niang et al., 2017; Rattalino Edreira et al., 2017), sin embargo, en esta tesis encontramos que las mismas solo fueron responsables del 1% de la variabilidad temporal de los Yw tanto de soja de primera como de soja de segunda. Aunque, representaron baja proporción de la varianza total, las CZs generaron efectos significativos en los Yw de ambos sistemas de cultivo de soja. Los tipos de suelo son otra de las grandes fuentes de variabilidad de los Yw, que en nuestro estudio representaron el 5% de la varianza total en el caso del Yw de soja de primera y 1% para el Yw de la soja de segunda. Mediante análisis de contrastes se determinó que los suelos pertenecientes a los Grupos CONEAT 03.51 anidado en la zona climática II y 10.7 anidado en la zona climática I fueron los suelos que permitieron alcanzar los mayores Yw para soja de primera, y los Grupos CONEAT 03.51 anidado en la zona climática II y 11.4 anidado en la zona climática III lo fueron para la soja de segunda.

Nuestros resultados indican que la fuente de variabilidad más importante es el año, siendo responsable del 71% y 74% de la varianza total de soja de primera y de segunda respectivamente. La importancia de la variabilidad temporal en los cultivos de verano es un resultado simplemente coincidente con algunos trabajos previamente publicados (Aramburu Merlos et al., 2015; Grassini, Torrion, Cassman, Yang, y Specht, 2014), la contribución más importante de este artículo al conocimiento científico es la estrategia planteada para identificar las variables que generan la variabilidad a través del tiempo (descripta detalladamente en la sección 3.3 de esta tesis), pero sobre todo, para identificar la forma en que influyen en el rendimiento. Las variables relacionadas a las propiedades del suelo, el balance de agua, junto con las variables que describen el estrés hídrico y la radiación solar durante el período crítico fueron las variables posicionadas en los primeros lugares del ranquin de importancia basado en el incremento del cuadrado medio del error (%IncMSE, medida de la importancia dentro del paquete *randomForest* (Liaw y Wiener, 2002)). Este algoritmo, colocó en el primer lugar de importancia dentro de las variables para soja de primera a la cantidad de agua almacenada en el suelo al momento de la siembra. En el caso de soja de segunda, aparece en segundo lugar de importancia, la cantidad de agua acumulada en el suelo al inicio del estadio R3, ya que el período entre la cosecha del cereal de invierno y la siembra de soja no permite generar una reserva de agua importante (Andrade y Satorre, 2015). Ambas variables están muy relacionadas a la capacidad de almacenaje de agua que poseen los suelos. Además, la medida de importancia, indica que pasar de que sean pocos los kilos de grano de soja que se agregan al rendimiento final, estos kilos agregados son muy estables y se concretan sistemáticamente. En los sistemas de cultivos en secano la variable que explica el mayor rango de rendimiento en grano es el estrés hídrico durante el período crítico del cultivo de soja (Grassini et al., 2014; Rattalino Edreira et al., 2017) y esta variable resultó en las primeras posiciones de nuestro análisis y su efecto fue estimado como el de mayor pendiente en los gráficos de dependencia parcial (b_1 = -0,02 para soja de primera y b_1 = -0,03 para soja de segunda).

Finalmente, otro resultado a destacar de esta tesis es el enfoque desarrollado para descomponer la varianza total entre sus componentes, combinando métodos estadísticos clásicos con métodos estadísticos avanzados, y llegando a resultados de sencilla interpretación. Este esquema de trabajo puede ser fácilmente actualizado y/o aplicado a bases de datos empíricas una vez que las mismas se encuentren disponibles. Si bien, el análisis *random forest* es considerado una caja negra, en esta tesis logramos entender claramente el comportamiento de las variables de entrada. Consideramos a estas herramientas como muy promisorias para el análisis de interacciones complejas en grandes bases de datos, que luego pueden ser trabajadas independientemente para profundizar en el conocimiento de los procesos subyacentes.

4.2. LINEAMIENTOS PARA INVESTIGACIONES FUTURAS

A pesar de los avances logrados en la cuantificación de las brechas de rendimientos de los cultivos de soja de primera y de segunda en el Uruguay, del entendimiento de las variables que explican el efecto año en el Yw de los sistemas de cultivo de soja y de la delimitación de zonas de riesgo, numerosos aspectos de interés no han sido totalmente dilucidados y al mismo tiempo surgen nuevos interrogantes a resolver en investigaciones futuras. Algunos de estos aspectos se plantean a continuación.

Trabajar con modelos de simulación implica conocer el funcionamiento de los cultivos y que la herramienta a emplear este calibrada, con el fin de representar de la manera más fiel posible el crecimiento y desarrollo de los cultivos a estudiar. Pero además, requiere de utilizar la información más precisa posible de los sistemas que se van a simular, para que estas simulaciones sean representativas de la realidad. En Uruguay, existe un vacío de información acerca de la máxima profundidad de enraizamiento que pueden alcanzar los cultivos. Si bien existen mapas muy precisos de los tipos de suelo, e incluso contamos con descripción de perfiles modales que entre otras propiedades detallan la profundidad total del perfil (www.mgap.gub.uy/unidad-organizativa/direccion-general-de-recursos-naturales/tramites-y-servicios/biblioteca-digital), no existe información públicamente disponible que permita conocer la máxima profundidad que puede ser explorada por las raíces de los cultivos en esos suelos. Nuestro trabajo puede ser fácilmente actualizado una vez que esta información esté disponible. Pero creemos importante resaltar, la necesidad de conocer con mayor precisión la profundidad de nuestro recurso suelo.

Nuestro trabajo de tesis se centró en la determinación y la evaluación de las brechas de rendimiento de los dos sistemas de cultivo de soja presentes en el país, aplicando un protocolo ampliamente publicado y enmarcado en el proyecto de evaluación de brechas de rendimiento más importante a nivel mundial. Esto implicó el uso de modelos de simulación de cultivos, para los cuales se validaron coeficientes genéticos para las condiciones de nuestro país. Información de más ensayos en condiciones óptimas que permitan explorar los rendimientos potenciales y cuyo objetivo sea el de calibrar un modelo de simulación de cultivos, puede permitir evaluar con mayor precisión los coeficientes genéticos empleados para la simulación. Una vez que esta información se encuentre disponible, puede ser empleada para actualizar este trabajo en la medida de que implique una mejora de las estimaciones de las brechas de rendimiento de nuestro país.

Esta tesis constituye un paso inicial hacia la intensificación ecológica de la agricultura, ya que provee información robusta aplicando una metodología transparente y ampliamente reconocida, de los rendimientos potenciales (Yp y Yw) la amplitud de las brechas de rendimiento. Un paso subsiguiente sería el de combinar la información generada con datos reales de chacras, para profundizar el análisis de las causas de estas brechas y ya no a una escala de zona buffer de una estación meteorológica de referencia, sino a la escala de la chacra. Las herramientas para aplicar este tipo de análisis se encuentran disponibles y creemos que indudablemente es un paso que debe ser dado en el menor tiempo posible.

Finalmente, creemos esta información que es generada con el propósito de contribuir a la mejora de los sistemas agrícolas de nuestro país, informando a todos los actores comprometidos con la producción de granos, todos estos esfuerzos son en vano si no logramos que sea efectivamente transmitida a la sociedad. Hoy sobran las herramientas para acortar distancias entre las Universidades y los centros de investigación y los productores o tomadores de decisiones, por lo tanto, debemos hacer el mejor uso posible de las mismas para que el conocimiento generado llegue a los usuarios y por ende repercuta positivamente en la producción nacional.

5. BIBLIOGRAFÍA

Andrade, J. F., y Satorre, E. H. (2015). Single and double crop systems in the Argentine Pampas: Environmental determinants of annual grain yield. *Field Crops Research*, *177*, 137–147.

Aramburu Merlos, F., Monzon, J. P., Mercau, J. L., Taboada, M. A., Andrade, F. H., Hall, A. J., … Grassini, P. (2015). Potential for crop production increase in Argentina through closure of existing yield gaps. *Field Crops Research*.

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

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

DIEA. (2008). *Anuario estadístico agropecuario*. [en línea]. Consultado 1 de agosto de 2017. Disponible en: http://www2.mgap.gub.uy/portal/page.aspx?2,diea,diea-anuario-2008,O,es,0,

DIEA. (2016). *Anuario estadístico agropecuario*. [en línea]. Consultado 1 de agosto de 2017. Disponible en: http://www.mgap.gub.uy/unidad-ejecutora/oficina-de-programacion-y-politicas-agropecuarias/publicaciones/anuarios-diea/anuario2016

Espe, M. B., Cassman, K. G., Yang, H., Guilpart, N., Grassini, P., Van Wart, J., … Linquist, B. A. (2016). Yield gap analysis of US rice production systems shows opportunities for improvement. *Field Crops Research*, *196*, 276–283.

Evans, L. T. (1993). *Crop Evolution, Adaptation and Yield*. Cambridge, UK: Cambridge University Press.

Evans, L. T., y Fisher, R. A. (1999). Yield potential: Its definition, measurement, and significance. *Crop Science*, *39*(6), 1544–1551.

Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., … Zaks, D. P. M. (2011). Solutions for a cultivated planet. *Nature*, *478*(7369), 337–342.

Giménez, L. (2017). Respuesta del maíz y la soja a diferentes disponibilidades hídricas en distintas etapas de desarrollo. *Agrociencia Uruguay*, *21*(2), 77–90.

Giménez, Luis, y García Petillo, M. (2011). Evapotranspiración de cultivos de verano para dos regiones climáticamente contrastantes de Uruguay. *Agrociencia Uruguay*, *15*(2), 100–108.

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., … Toulmin, C. (2010). Food Security: The Challenge of Feeding 9 Billion People. *Science*, *327*(5967), 812–818.

Gonzalez J. (2013). *Análisis de las variables ambientales y de manejo vinculadas a las variaciones de rendimiento en grano en cultivos agrícolas de secano, a partir de datos de chacras de agricultores CREA*. [Tesis para la obtención del título de Ingeniero Agrónomo]. Universidad de la República. Facultad de Agronomía. Uruguay. 386 p.

Grassini, P., Torrion, J. A., Cassman, K. G., Yang, H. S., y Specht, J. E. (2014). Drivers of spatial and temporal variation in soybean yield and irrigation requirements in the western US Corn Belt. *Field Crops Research*, *163*, 32–46.

Guilpart, N., Grassini, P., Sadras, V. O., Timsina, J., y Cassman, K. G. (2017). Estimating yield gaps at the cropping system level. *Field Crops Research*, *206*, 21–32.

Guilpart, N., Grassini, P., van Wart, J., Yang, H., van Ittersum, M. K., van Bussel, L. G. J., … Cassman, K. G. (2017). Rooting for food security in Sub-Saharan Africa. *Environmental Research Letters*, *12*(11), 114036.

Henderson, B., Godde, C., Medina-Hidalgo, D., van Wijk, M., Silvestri, S., Douxchamps, S., … Herrero, M. (2016). Closing system-wide yield gaps to increase food production and mitigate GHGs among mixed crop–livestock smallholders in Sub-Saharan Africa. *Agricultural Systems*, *143*, 106–113.

Hochman, Z., Prestwidge, D., y Carberry, P. S. (2014). Crop sequences in Australia's northern grain zone are less agronomically efficient than implied by the sum of their parts. *Agricultural Systems*, *129*, 124–132.

Leenaars, J. G. B., Claessens, L., Heuvelink, G. B. M., Hengl, T., Ruiperez González, M., van Bussel, L. G. J., … Cassman, K. G. (2018). Mapping rootable depth and root zone plant-available water holding capacity of the soil of sub-Saharan Africa. *Geoderma*, *324*, 18–36.

Liang, W., Carberry, P., Wang, G., Lü, R., Lü, H., y Xia, A. (2011). Quantifying the yield gap in wheat–maize cropping systems of the Hebei Plain, China. *Field Crops Research*, *124*(2), 180–185.

Liaw, A., y Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*, 18–22.

Lobell, D. B., Cassman, K. G., y Field, C. B. (2009). Crop yield gaps: their importance, magnitudes, and causes. *Annual Review of Environment and Resources, 34, 179–204.*

Marin, F. R., Martha, G. B., Cassman, K. G., y Grassini, P. (2016). Prospects for increasing sugarcane and bioethanol production on existing crop area in Brasil. *Bioscience*, *66*(4), 307–316.

Monzon, J. P., Sadras, V. O., Abbate, P. A., y Caviglia, O. P. (2007). Modelling management strategies for wheat–soybean double crops in the south-eastern Pampas. *Field Crops Research*, *101*(1), 44–52.

Mourtzinis, S., Rattalino Edreira, J. I., Grassini, P., Roth, A. C., Casteel, S. N., Ciampitti, I. A., … Conley, S. P. (2018). Sifting and winnowing: Analysis of farmer field data for soybean in the US North-Central region. *Field Crops Research*, *221*, 130–141.

Niang, A., Becker, M., Ewert, F., Dieng, I., Gaiser, T., Tanaka, A., … Saito, K. (2017). Variability and determinants of yields in rice production systems of West Africa. *Field Crops Research*, *207*, 1–12.

Rattalino Edreira, J. I., Mourtzinis, S., Conley, S. P., Roth, A. C., Ciampitti, I. A., Licht, M. A., … Grassini, P. (2017). Assessing causes of yield gaps in agricultural areas with diversity in climate and soils. *Agricultural and Forest Meteorology*, *247*, 170–180.

Tittonell, P. (2014). Ecological intensification of agriculture — sustainable by nature. *Current Opinion in Environmental Sustainability*, *8*, 53–61.

van Ittersum, Martin K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., y Hochman, Z. (2013). Yield gap analysis with local to global relevance—A review. *Field Crops Research*, *143*, 4–17.

van Ittersum, Martin K., van Bussel, L. G. J., Wolf, J., Grassini, P., van Wart, J., Guilpart, N., … Cassman, K. G. (2016). Can sub-Saharan Africa feed itself? *Proceedings of the National Academy of Sciences*, *113*(52), 14964–14969.

van Ittersum, M.K., y Rabbinge, R. (1997). Concepts in production ecology for analysis and quantification of agricultural input-output combinations. *Field Crops Research*, *52*, 197–208.

van Oort, P. A. J., Saito, K., Tanaka, A., Amovin-Assagba, E., Van Bussel, L. G. J., van Wart, J., … Wopereis, M. C. S. (2015). Assessment of rice self-sufficiency in 2025 in eight African countries. *Global Food Security*, *5*, 39–49.

van Wart, J., Grassini, P., y Cassman, K. G. (2013). Impact of derived global weather data on simulated crop yields. *Global Change Biology*, *19*, 3822–3834.

Viglizzo, E. F., Frank, F. C., Carreño, L. V., Jobbágy, E. G., Pereyra, H., Clatt, J., … Ricard, M. F. (2011). Ecological and environmental footprint of 50 years of agricultural expansion in Argentina: FOOTPRINT OF ARGENTINE AGRICULTURE. *Global Change Biology*, *17*(2), 959–973.

You, L., Wood, S., Wood-Sichra, U., y Wu, W. (2014). Generating global crop distribution maps: From census to grid. *Agricultural Systems*, *127*, 53–60.