



Predicción de índice verde en chacras agrícolas

Christopher Quincke Ferreira

Guillermo Ettlin Alonso

Proyecto de Grado presentado como parte de los requisitos necesarios para la obtención del título de Grado en Ingeniería en Computación de la Facultad de Ingeniería de la Universidad de la República.

Tutor del Proyecto:

Ph.D. Ing. Pablo Rodríguez Bocca

Usuario responsable:

Ph.D. Ing. Agr. Andrés Berger

Montevideo – Uruguay

Julio de 2017

(Dedicatoria) A mi familia por todo el apoyo y aliento durante estos años, especialmente a mis padres por el gran esfuerzo de mantenerme en Montevideo mientras fue necesario. A mis compañeros y amigos por haber hecho que estos años de estudio hayan sido una gran experiencia.

Y a mi novia, que siempre está ahí para alentarme y ayudarme en todo lo que se presenta.

Guillermo Ettlin Alonso

A toda mi familia y amigos por el apoyo incondicional a lo largo de todos estos años de estudio.

En especial a mis padres por el esfuerzo que implicó mantenerme en Montevideo, a mis hermanos y primos por siempre estar y a mis abuelos por ser grandes consejeros.

Christopher Quincke Ferreira

Agradecimientos

Quisieramos agradecer a nuestro Tutor, Dr. Ing. Pablo Rodríguez Bocca, quien nos acompañó y guió en este largo proceso. También agradecemos al INIA, en especial al Ing. Agr. MSc, PhD Andrés G. Berger, que nos recibió y ayudó desde el comienzo, cuando este proyecto no era más que una idea.

Además, debemos agradecer a la Empresa Agronegocios del Plata(ADP) por brindarnos los datos relativos a las chacras y sus zafras necesarios para esta investigación.

RESUMEN

En las últimas décadas, como consecuencia de una gran demanda mundial de alimentos y de la exigencia de una mayor competitividad, la agricultura de precisión ha sido impulsada cobrando una gran importancia y desarrollándose a pasos agigantados. Uno de los desafíos más importantes dentro de la agricultura de precisión, es la posibilidad de predecir el rendimiento de los cultivos. Siendo la agricultura para nuestro país un componente importante de la economía, la investigación en esta área es primordial para mantener la competitividad. En nuestro país, el principal abanderado para realizar este tipo de investigaciones es el INIA (Instituto Nacional de Investigación Agropecuaria). En este proyecto, dada la creciente importancia que tiene el control, monitoreo y predicción de rendimientos de cultivos, se realiza una investigación interdisciplinaria contando con la colaboración y el apoyo del INIA, en la que se intenta predecir el índice verde de cultivos de soja ubicados en el territorio nacional. Para ello, se utiliza un simulador de crecimiento de cultivos provisto por el INIA, un conjunto de datos relativos a las zonas anteriores de soja sobre las chacras en estudio e imágenes de los satélites MODIS y Landsat. El simulador de crecimiento de cultivos provisto por el INIA, recibe como entrada un conjunto de parámetros relativos a las características iniciales del suelo, las características del cultivo y las condiciones meteorológicas del lugar donde se cultiva. La predicción del índice verde mediante este simulador consiste en, a partir de los datos relativos a las zonas anteriores de soja sobre las chacras en estudio, de las imágenes satelitales y datos meteorológicos, obtener los valores de los parámetros relativos a las características iniciales del suelo y así poder utilizarlo para la predicción de futuras zonas. Palabras claves:

NDVI, producción, modelo de crecimiento agrícola, imágenes satelitales, procesamiento de imágenes.

Tabla de contenidos

1	Introducción	1
2	Estado del arte	4
2.1	Introducción	4
2.2	¿Qué son los sensores remotos?	5
2.3	¿Qué es el NDVI?	8
2.4	¿Qué es el LAI?	9
2.5	Relación entre propiedades espectrales y propiedades de cultivos	10
2.5.1	Relación entre NDVI y LAI	11
2.5.2	Relación entre NDVI y rendimiento de cultivos	12
2.6	Modelos de crecimiento	13
2.6.1	Tipos de modelos	13
2.6.2	Reseña histórica de modelos y sus características	15
2.6.3	Modelo utilizado en nuestro proyecto	18
2.7	Investigaciones que utilizan NDVI, rendimiento de cultivos y simuladores de crecimiento de cultivos	19
3	Descripción del problema	23
3.1	Introducción	23
3.2	Objetivos	24
3.3	Fuentes de datos y su calidad	24
3.3.1	Imágenes satelitales	25
3.3.2	Información de zonas	27
3.3.3	Datos meteorológicos	29
3.4	Forma de evaluar la solución	29
3.5	Explicación del problema a partir de una instancia	30
3.5.1	Datos de la chacra	31
3.5.2	Datos meteorológicos	32

3.5.3 Ejemplo de solución	34
4 Solución del problema	36
4.1 Introducción	36
4.2 Obtención de las imágenes satelitales y procesamiento	36
4.3 Extracción de datos del INIA	37
4.4 Caracterización de las variedades de soja	39
4.5 Obtención de series de NDVI a partir del cruzamiento de datos y las imágenes satelitales	40
4.6 Creación de estimadores <i>baseline</i>	44
4.7 Calibración del simulador de crecimiento de cultivos	50
4.8 Simulación del crecimiento de cultivos	53
4.9 Evaluación de la calidad de la estimación con RMSE	53
5 Pruebas	55
5.1 Introducción	55
5.2 Chacras Utilizadas	55
5.3 Consideraciones para realizar la simulación	57
5.4 Estadísticas globales sobre los resultados	58
5.5 Estadísticas por chacra sobre los resultados	62
5.6 Posibles mejoras	64
6 Conclusiones	69
Referencias bibliográficas	72
Glosario	75
Apéndices	76
Apéndice 1 Chacras utilizadas en el proyecto	77
Apéndice 2 Resultados detallados	83

Capítulo 1

Introducción

La predicción del rendimiento de los cultivos es uno de los tópicos más importantes dentro de la agricultura de precisión. Siendo la agricultura para nuestro país un componente importante de la economía, la investigación en esta área es primordial para mantener la competitividad. Además, industrias que utilizan ciertos cultivos como suministros están interesados en sus rendimientos, debido a que su precio depende en gran medida de ello. Adicionalmente, organizaciones que brindan ayuda humanitaria se basan en información temprana sobre la producción de cultivos para planificar la respuesta de emergencia e intervenciones de ayuda alimentaria. En este proyecto, dada la creciente importancia que tiene el control, monitoreo y predicción de rendimientos de cultivos, se realiza una investigación interdisciplinaria contando con la colaboración y el apoyo del INIA, en la que se intenta predecir el índice verde de cultivos de soja ubicados en el territorio nacional. Para ello, se utiliza un simulador de crecimiento de cultivos provisto por el INIA, un conjunto de datos relativos a las zonas anteriores de soja sobre las chacras en estudio e imágenes de los satélites MODIS y Landsat. El simulador de crecimiento de cultivos provisto por el INIA, recibe como entrada un conjunto de parámetros relativos a las características iniciales del suelo, las características del cultivo y las condiciones meteorológicas del lugar donde se cultiva. La predicción del índice verde mediante este simulador consiste en, a partir de los datos relativos a las zonas anteriores de soja sobre las chacras en estudio, de las imágenes satelitales y datos meteorológicos, obtener los valores de los parámetros relativos a las características iniciales del suelo y así poder utilizarlo para la predicción de futuras zonas. Con tal fin, se utilizan los datos de todas las zonas menos de

la última para calibrar el simulador y los datos de la última zafra para evaluar el resultado de la predicción. Para ello, se toma el RMSE (Root Mean Square Deviation) como medida de error de la predicción y se crea un estimador *baseline* para comparar el error de la predicción con el error obtenido mediante un método de estimación sencillo. Como resultado de esta investigación, se demuestra que a partir de un conjunto de datos (contorno de la chacra (polígono vectorial), fecha de siembra y cosecha al menos 4 zafras, datos climáticos para las zafras, y características de la variedad de semilla sembrada), se logra calibrar el simulador obteniendo buenas predicciones de índice verde para futuras zafras.

A lo largo de este documento se expone la investigación realizada en el proyecto de grado *Predicción de Índice Verde en Chacras Agrícolas*, para la obtención del título de grado de Ingeniero en Computación de la Facultad de Ingeniería de la UdelaR. El documento consta de los siguientes capítulos:

1. Introducción.
2. Estado del arte.
3. Descripción del problema.
4. Solución del problema.
5. Pruebas.
6. Conclusiones.

Previo a interiorizarse en la investigación que se lleva a cabo en este proyecto, es necesario conocer algunos conceptos básicos relativos a la agronomía, modelos de crecimiento de cultivos y sensores remotos. Con tal fin, en el Capítulo 2, se da un marco teórico a la investigación. Primero que nada, se introducen algunos conceptos que serán claves para comprender la investigación. Posteriormente, se comentan algunas investigaciones que se relacionan con este proyecto y que le dan sentido al mismo. Además, se hace una reseña histórica sobre los distintos modelos de crecimientos de cultivos, destacando las características de cada tipo de modelo de crecimiento y evaluando la aplicabilidad de los modelos en la práctica de acuerdo a su complejidad. La importancia de este capítulo, no solo radica en comprender conceptos teóricos necesarios para entender este proyecto, sino que también es importante entender el potencial que puede tener poder predecir el índice verde de un cultivo. En el Capítulo 3, luego de contextualizar la investigación, se explica el problema a resolver en este proyecto. Antes que nada, se deja en claro el objetivo del proyecto, en este caso,

poder predecir el NDVI de chacras agrícolas donde se siembra soja. Luego, se comenta acerca de los insumos con los que se cuenta, mencionando cuales son los datos sobre las chacras y sobre las siembras históricas que el INIA provee, y sobre la obtención de otros insumos como lo son las imágenes satelitales y los datos climáticos. En el Capítulo 4, se define la solución del problema. Ello implica, detallar las herramientas utilizadas, los procedimientos seguidos, los desafíos surgidos y la forma en que fueron resueltos. Asimismo, se define como se califica que tan buena es una solución mediante el RMSE (Root Mean Square Deviation). Por otra parte, con el objetivo de poder comparar los resultados de la estimación obtenida con otro método de estimación, se define un estimador sencillo al que se denomina *baseline*. Más tarde, en el Capítulo 5, se dan datos sobre las chacras utilizadas para la investigación. También se observan algunas consideraciones tomadas a la hora de la codificación teniendo en cuenta cuestiones de performance. Adicionalmente, se muestran los resultados obtenidos para cada una de las chacras estudiadas. Finalmente, en el Capítulo 6, se muestran los resultados obtenidos a nivel global, intentando darle explicación a los mismos. Por otra parte, se exponen posibles mejoras a esta investigación y otras investigaciones que pueden surgir de esta, teniendo en cuenta los resultados obtenidos, falencias del simulador utilizado y avances tecnológicos que a priori, podrían derivar en mayor precisión en las estimaciones.

Capítulo 2

Estado del arte

2.1. Introducción

Existen varios métodos de predicción de rendimiento de cultivos. El método más tradicional es la evaluación del estado de los mismos por parte de expertos en la materia. Las observaciones y mediciones se realizan a lo largo de toda la zafra para el cultivo en estudio y a partir de los datos obtenidos se puede pronosticar el rendimiento mediante métodos de regresión (utilizando datos históricos), o por el conocimiento de expertos. Otros métodos menos tradicionales utilizados para pronosticar el rendimiento de los cultivos son el uso de sensores y modelos de simulación de cultivos. Como los cambios en la vitalidad de los cultivos, la densidad y la productividad afectan a las propiedades ópticas de las imágenes satelitales tomadas sobre estos, el desarrollo de los cultivos ha sido monitorizado mediante sensores desde los inicios de los mismos. Dentro de los sensores se distinguen dos tipos: sensores *in situ* y sensores remotos. *In situ*, es una expresión latina que significa en *el sitio* o en *el lugar*, y que suele utilizarse para designar un fenómeno observado en el lugar¹. Por lo tanto, los sensores *in situ*, son aquellos que realizan una medición en el mismo lugar en que se produce el fenómeno medido. Los sensores remotos, se definen como sistemas de detección y medida a distancia, generalmente empleados desde aeronaves o satélites, con los que se obtiene información meteorológica, oceanográfica y sobre la cubierta vegetal². Las imágenes satelitales pueden desempeñar un papel importante en el suministro de información sobre el tipo

¹https://en.wikipedia.org/wiki/In_situ

²http://www.etisig.catamarca.gov.ar/index.php?option=com_content&view=article&id=86&Itemid=103

de cultivo, las condiciones y el rendimiento, desde un nivel de agregación de una pequeña chacra hasta áreas geográficas extensas como países o continentes. Actualmente, esta información es requerida y utilizada en muchas áreas del conocimiento con distintos fines. Por ejemplo, compañías agrícolas, agencias de seguros y juntas agrícolas internacionales requieren mapas de cultivo para realizar inventarios sobre los cultivos sembrados a lo largo del tiempo en distintos puntos geográficos. Industrias que utilizan ciertos cultivos como suministros y agencias gubernamentales están interesados en sus rendimientos y en la superficie dedicada a su producción, ya que los precios de los mismos a nivel mundial dependen en gran medida de ello. Además, organizaciones que brindan ayuda humanitaria a nivel internacional se basan en información temprana y fiable sobre la producción de cultivos con el fin de organizar la respuesta de emergencia y las intervenciones de ayuda alimentaria.

Por otro lado, con el fin de modelar el desarrollo de los cultivos agrícolas matemáticamente y simplificando lo que realmente ocurre en la naturaleza, se puede considerar el desarrollo de los mismos como una función que depende de un conjunto de variables. Entre estas variables podemos distinguir variables meteorológicas, como lo son, la luz solar, temperatura y precipitaciones. También juegan un rol muy importante variables relativas a las características del suelo, la genética de las plantas y las prácticas agrícolas. Actualmente, existen diversos sistemas informáticos que simulan el crecimiento de los cultivos modelandolos de esta forma. La complejidad de los modelos es muy variada. Algunos son muy sencillos, con no más de 7 parámetros de entrada para describir el desarrollo de un cultivo en particular [1], otros modelan procesos en detalle como la fotosíntesis foliar, y por tanto requieren de muchos parámetros de entrada. La dificultad en la utilización de estos modelos radica en que algunos requieren muchos datos de entrada que son difíciles de obtener, y además pueden requerir la calibración de ciertos parámetros para cada región donde se utilizará el modelo. Ésto hace difícil aplicarlos en algunas situaciones reales debido a la imposibilidad de recopilar datos de entrada y/o calibrar parámetros.

2.2. ¿Qué son los sensores remotos?

La detección mediante sensores remotos (teledetección), se define como la ciencia de la adquisición de información acerca de un objeto a través de los

datos obtenidos por un dispositivo que no está en contacto con el objeto en estudio³. Los datos se pueden obtener desde diversas plataformas, tales como satélites, aviones, vehículos no tripulados y radiómetros de mano. Pueden ser recogidos por diferentes dispositivos como sensores, cámaras digitales y grabadoras de vídeo. Los instrumentos utilizados para medir radiación electromagnética se denominan sensores. Los sensores remotos recogen datos mediante la detección de la energía que se refleja desde la Tierra. Pueden clasificarse en pasivos y activos. Los sensores pasivos responden a los estímulos externos. Por ejemplo, las imágenes satelitales, registran la energía natural que es reflejada o emitida desde la superficie de la Tierra. La fuente más común de la radiación detectada por los sensores pasivos es el reflejo de la luz del sol. Por el contrario, los sensores activos utilizan estímulos internos para recopilar datos sobre la Tierra. Por ejemplo, emitiendo un rayo láser a distancia desde el sistema de detección hacia la superficie de la Tierra y midiendo el tiempo que demora el láser en reflejar de nuevo el sensor del sistema de detección. Los sensores remotos tienen una amplia gama de aplicaciones en muchas áreas. Algunas de ellas son⁴:

- Aplicaciones costeras: monitor de cambios en la costa, monitor del transporte de sedimentos, y mapa de características de la costa. Los datos pueden ser utilizados para la cartografía costera y la prevención de la erosión.
- Aplicaciones oceánicas: Monitor de la circulación oceánica y sistemas de corrientes, miden la temperatura del océano y la altura de olas. Los datos pueden usarse para entender mejor los océanos y cómo gestionar mejor los recursos que estos proveen.
- Evaluación del peligro: monitoreo de huracanes, terremotos, erosión e inundaciones. Los datos pueden ser utilizados para evaluar el impacto de un desastre natural y crear estrategias de preparación para un evento de este tipo.
- Gestión de recursos naturales: Supervisar el uso del suelo y hábitats de vida silvestre. Los datos pueden ser utilizados para reducir al mínimo el daño que el crecimiento urbano tiene sobre el medio ambiente y ayudar a decidir cómo proteger mejor los recursos naturales.

³http://www.esa.int/SPECIALS/Eduspace_ES/SEM01U3FEXF_0.html

⁴<http://oceanservice.noaa.gov/facts/remotesensing.html>

- Aplicaciones agronómicas: como los cambios en la vitalidad de los cultivos, la densidad y la productividad afectan a las propiedades ópticas de las imágenes satelitales tomadas sobre estos, el desarrollo de los cultivos puede ser monitorizado mediante sensores remotos.

En esta investigación nos resultan de interés las aplicaciones agronómicas de los sensores remotos. En particular, aquellos datos que se pueden obtener de las plantas mediante sensores remotos que tienen relación con el desarrollo y crecimiento de las mismas. Mas específicamente, las imágenes satelitales MODIS y Landsat son los sensores remotos utilizados.

MODIS⁵ (*Moderate Resolution Imaging Spectroradiometer*) es un sensor de 36 canales visibles a infrarrojos térmicos. Se lanzó por primera vez el 18 de diciembre de 1999 en el satélite Terra, el buque insignia del programa *Earth Observation System* (EOS) de la NASA. Un segundo MODIS fue desplegado en el satélite Aqua el 4 de mayo de 2002.

El Programa Landsat⁶ es una serie de misiones satelitales de observación de la Tierra administradas conjuntamente por la NASA y el Servicio Geológico de los Estados Unidos. Los instrumentos a bordo de los satélites Landsat han adquirido millones de imágenes a lo largo de las misiones y los datos son un recurso valioso para la investigación y las aplicaciones del cambio global en la agricultura, la silvicultura, la geología, la planificación regional y la educación.

De las longitudes de onda de la radiación solar, la clorofila de las plantas absorbe la azul y la roja, mientras la verde es reflejada. La reflexión de la radiación visible, es realizada principalmente por los pigmentos de las hojas, mientras que el infrarrojo cercano (NIR) es reflejado por la estructura interna del Mesófilo de las hojas. La radiación NIR pasa a través de la primer capa de la hoja; cuando alcanza el Mesófilo y las cavidades internas de las hojas se dispersa, tanto hacia arriba (lo que se conoce como radiación reflejada) y hacia abajo (transmite radiación). La Figura 2.1 ilustra este proceso. El comportamiento de la reflectancia de la radiación NIR es también una función del índice de área de hoja (LAI), grosor de la hoja, el aire interno de la hoja y el agua contenida. La disminución relativa de la reflectancia es más alta en el espectro visible que en la radiación NIR, a causa de los efectos de la transmisión de la radiación NIR a través de las hojas y la absorción del rojo y azul por parte de la clorofila.

⁵<https://nsidc.org/data/modis/>

⁶<https://landsat.usgs.gov/>

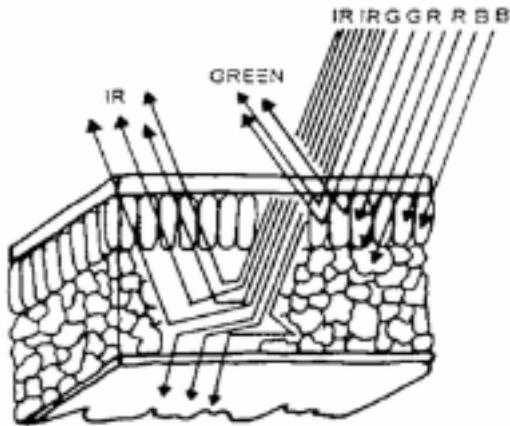


Figura 2.1: Sección de una hoja e interacciones entre la estructura foliar y la radiación solar. Imagen tomada de [2].

El reflejo de estas dos longitudes de onda pueden ser detectadas por un sensor posicionado sobre el cultivo. Como resultado, por lo explicado anteriormente, los cultivos saludables mostrarán altos valores de reflectancia en la radiación NIR y valores bajos de reflectancia en el espectro visible. Para el espectro visible, la reflectancia de las hojas es menor que la del suelo, mientras que para el NIR, la reflectancia de las hojas es mayor a la del suelo. Esta observación justifica la importancia de la medición de la reflectancia de estas longitudes de ondas en aplicaciones agronómicas.

2.3. ¿Qué es el NDVI?

Los índices de vegetación (VIS), son combinaciones matemáticas o relaciones de bandas espectrales principalmente entre las bandas de color rojo, verde e infrarrojo; que están diseñados para encontrar las relaciones funcionales entre las características de los cultivos. Los índices de vegetación están fuertemente modulados por la interacción de la radiación solar con la fotosíntesis de cultivos y por lo tanto son indicativos de la dinámica de las propiedades biofísicas relacionadas con el estado de los cultivos. Uno de los índices de vegetación que puede ser obtenido mediante sensores remotos y que es de vital importancia en aplicaciones agronómicas, es el Índice de Vegetación de Diferencia Normalizada (Normalized Difference Vegetation Index, NDVI) [3]. El NDVI, es un índice usado para estimar la cantidad, calidad y desarrollo de la vegetación

con base a la medición, por medio de sensores remotos instalados comúnmente desde una plataforma espacial, de la intensidad de la radiación de ciertas bandas del espectro electromagnético que la vegetación refleja. Debido a que los primeros instrumentos usados para la observación terrestre, tales como el Satélite de Tecnología de Recursos Terrestres (Earth-Resources Technology Satellite, ERTS) desarrollado por la Administración Nacional de Aeronáutica y del Espacio de los Estados Unidos (NASA) y el Radiómetro Avanzado de Muy Alta Resolución (Advanced Very High Resolution Radiometer, AVHRR) de la Administración Nacional Oceánica y Atmosférica de los Estados Unidos (NOAA), adquirían datos en el rojo e infrarrojo cercano, fue natural hacer uso de las marcadas diferencias en la reflexión vegetal para determinar su distribución espacial en las imágenes satelitales. El índice de vegetación de diferencia normalizada, se calcula a partir de estas medidas individuales de la siguiente manera:

$$NDVI = \frac{IRCercano - Rojo}{IRCercano + Rojo} \quad (2.1)$$

en donde las variables ROJO y IRCercano están definidas por las medidas de reflexión espectral adquiridas en las bandas del rojo e infrarrojo cercano, respectivamente. Estas reflexiones espectrales son en sí cocientes de la radiación reflejada sobre la radiación entrante en cada banda espectral individual; por tanto, éstos toman valores en un rango de 0,0 a 1,0. El NDVI varía como consecuencia entre -1,0 y +1,0.

2.4. ¿Qué es el LAI?

El índice de área foliar (Leaf Area Index, LAI) es la cantidad de área foliar por unidad de área de suelo (m^2/m^2)⁷. Se puede definir tanto para plantas individuales como para un conjunto de ellas. Hay una serie de razones por las cuales hay un creciente interés en cuantificar el LAI. Monitorear la distribución y cambios en el LAI es importante para evaluar el crecimiento y vigor de la vegetación en el planeta. También es utilizado como parámetro en modelos climáticos, y además, relaciona la biosfera y la atmósfera a través de diversos procesos como la fotosíntesis, respiración y transpiración de las plantas.

⁷<https://disc.gsfc.nasa.gov/hydrology/data-holdings/nldas-parameters/leaf-area-index>

2.5. Relación entre propiedades espectrales y propiedades de cultivos

La relación entre las propiedades espectrales de los cultivos y su rendimiento ha sido reconocida desde los primeros experimentos de campo de espectrometría [4]. Se ha estudiado el uso de datos espectrales extensivamente mediante el uso de imágenes satelitales a partir del lanzamiento del primer satélite civil de observación de la Tierra (Landsat-1) en el año 1972. Sin embargo, sólo a partir de la creciente disponibilidad de imágenes satelitales de baja resolución de la serie de satélites meteorológicos NOAA (National Oceanic and Atmospheric Administration) - AVHRR (Advanced Very High Resolution Radiometer) a principios de los años 80, se han podido realizar análisis similares extendidos a grandes áreas. Gracias a su amplitud de barrido, los sistemas de baja resolución tienen una mejor visión sinóptica y frecuencia temporal en comparación con sensores de alta resolución. Cada imagen abarca una anchura de hasta 3.000 km, de tal manera que toda la superficie de la Tierra se escanea todos los días. La desventaja intrínseca de estos sensores es, por supuesto, su baja resolución espacial, con tamaños de cada píxel de aproximadamente 1 a 2 kilómetros cuadrados. Como consecuencia, el espectro registrado en cada pixel es información mixta de varios tipos de superficie. Esto complica seriamente la interpretación (y validación) de la señal, así como la fiabilidad de los productos de información derivada [4]. Los estudios [5] [6] demostraron que las propiedades de reflectanciapectral de los cultivos, y en particular, las combinaciones de las reflectancias de las bandas rojo e infrarrojo cercano (utilizadas para calcular los distintos índices verdes), son muy útiles para la monitorización de los cultivos. Entre los distintos índices verdes basados en estas dos bandas espectrales, el índice de vegetación de diferencia normalizada ó NDVI (por sus siglas en inglés), propuesto por Deering en 1978 [7], se ha convertido en el indicador más popular para estudiar la salud de la vegetación y su evolución. Además, la investigación en monitoreo de la vegetación ha demostrado que el NDVI está estrechamente relacionado con el índice de área de hoja (LAI) y también es una medida indirecta de la productividad primaria a través de su relación casi lineal con la fotosíntesis -Fraction of Absorbed Photosynthetically Active Radiation. (FAPAR)-.

2.5.1. Relación entre NDVI y LAI

Para obtener el LAI existen métodos directos e indirectos. Entre los métodos directos se encuentran herramientas como la LAI-2200C de LI-COR⁸, a través de las cuales se pueden obtener muy buenas mediciones del LAI. Sin embargo, estos métodos requieren de mucho tiempo, trabajo y dinero. También existen métodos más simples que consisten en recolectar las hojas de una planta, medir su área, y dividirla por la extensión de suelo que ocupa. Siendo por lo tanto un método destructivo. Es por estos motivos que por lo general estos métodos son utilizados como referencia para métodos indirectos. Una forma de medición indirecta del LAI es mediante los datos obtenidos por teledetección. Si bien esta forma de medición requiere lidiar con el efecto de las condiciones atmosféricas y la contaminación, proporciona datos actualizados y amplía el área de estudio. Existen diversas investigaciones que buscan relacionar el NDVI con el LAI. Tal es el caso de [8] cuyo objetivo principal es investigar las relación entre el LAI y dos tipos de índices verdes (NDVI y EVI) basados en datos in situ en el bosque de Takayama, Japón. Para ello, se compararon los datos LAI in situ con datos de NDVI y EVI (Enhanced Vegetation Index) in situ con el objetivo de obtener ecuaciones de regresión que relacionen dichos valores. Luego, estas ecuaciones se aplicaron al NDVI y EVI obtenidos mediante imágenes satelitales MODIS para validar la relación existente. En la Figura 2.2 se muestra una comparación entre el LAI obtenido mediante mediciones in situ contra el LAI estimado basado en imágenes MODIS NDVI de esta investigación. Las cruces y círculos en blanco son mediciones de LAI in situ, mientras los círculos negros son estimaciones del LAI basado en imágenes satelitales MODIS NDVI.

⁸https://www.licor.com/env/pdf/area_meters/brochure_LAI-2200C_2014-web.pdf

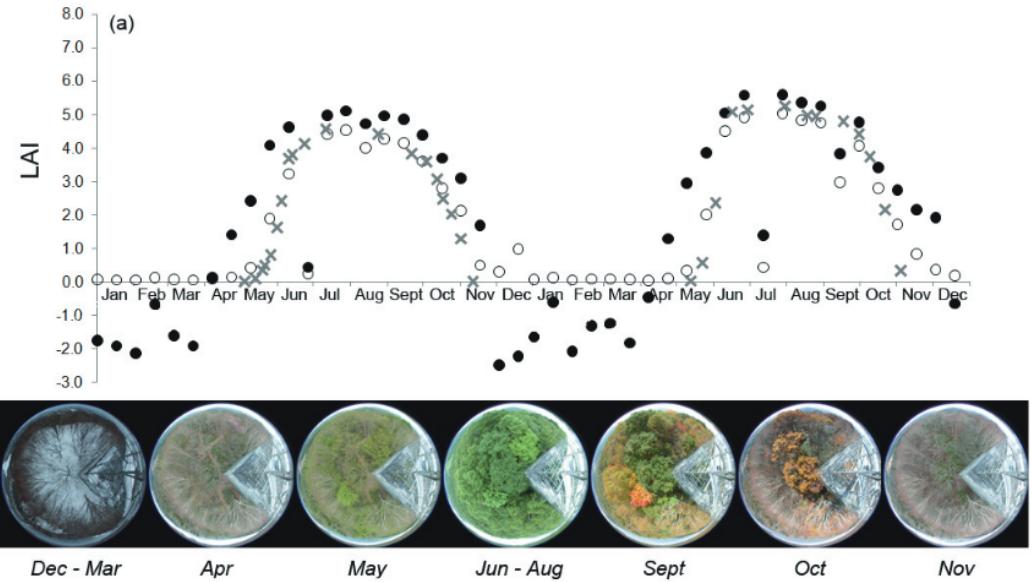


Figura 2.2: LAI obtenido mediante mediciones in situ vs LAI estimado basado en imágenes MODIS NDVI. Las cruces y círculos en blanco son mediciones de LAI in situ, mientras los círculos negros son estimaciones del LAI basado en imágenes satelitales MODIS NDVI. Imagen tomada de [8].

2.5.2. Relación entre NDVI y rendimiento de cultivos

Existen diversas investigaciones que relacionan el índice verde con el rendimiento de un determinado cultivo. Algunas de ellas, parten de la relación antes mencionada entre los índices de vegetación y la FAPAR, ya que el rendimiento de muchos cultivos se determina principalmente por la actividad fotosintética de plantas agrícolas en determinados períodos antes de su cosecha. Otras investigaciones, buscan la relación existente entre los índices de vegetación y la Biomasa. Tal es el caso de [9]. En dicha investigación, en primer lugar se obtuvo experimentalmente una relación entre el NDVI y la Biomasa para el trigo. Posteriormente, se determinó una fórmula para calcular el Índice de cosecha a partir del NDVI. Finalmente, se calcula la productividad a partir de la Biomasa y el Índice de cosecha. La relación entre el NDVI y la Biomasa fue descrita mediante la siguiente ecuación:

$$Y = 58,618 \cdot x^2 + 91,116 \cdot x + 26,429 \quad (2.2)$$

donde Y es la Biomasa y x es la sumatoria del NDVI.

La Biomasa de un cultivo es insuficiente para estimar la productividad, ya

que es sólo masa seca, en este caso, de trigo. Una variable que juega un rol muy importante a la hora de determinar el Índice de cosecha, es el volumen de agua que el cultivo recibe durante la etapa de llenado de granos. El NDVI reacciona de acuerdo al volumen de agua recibido durante este período. Por ejemplo, si el volumen de agua que el cultivo recibe excede o es menor a lo necesario, las hojas se ponen amarillas y marchitas, consecuentemente el NDVI decrece. Considerando lo antedicho, se puede utilizar el NDVI para estimar el Índice de cosecha. En [9] se determina una estimación del Índice de cosecha mediante la siguiente ecuación:

$$HINDVI = HImax - HIrange \cdot \left(1 - \frac{\sum NDVI_{post}}{\sum NDVI_{pre}}\right) \quad (2.3)$$

donde:

- HINDVI es el Índice de cosecha.
- HImax es el Índice de cosecha óptimo (para el caso estudiado es 0.48).
- HIrange es el intervalo en el que varía el Índice de cosecha.

Finalmente se estima el rendimiento del trigo mediante la siguiente ecuación: $WY = AGB \cdot HINDVI$.

Los resultado obtenidos en [9] fueron muy buenos, obteniendo errores de 2 %, 9.3 % y 3.9 % entre lo cosechado y lo estimado para los años 2008, 2009 y 2010 respectivamente.

2.6. Modelos de crecimiento

2.6.1. Tipos de modelos

Diversos tipos de modelos de crecimiento se han desarrollado a lo largo de los años, los cuales pueden ser clasificados en tres amplios grupos. [2] [10]

Los modelos **estadísticos** dependen de la disponibilidad de datos de rendimiento históricos de grandes áreas, para identificar tendencias. Este enfoque se basa en examinar los datos y encontrar una ecuación o un conjunto de ellas que se ajusten a los mismos. Si bien son mejores que las estimaciones empíricas, carecen de la precisión necesaria para algunos usos, y no son sensibles a variaciones particulares de algunos períodos -tales como sequías o inundaciones-.

Los modelos **mecanicistas** intentan utilizar los mecanismos fundamentales de procesos de la planta y del suelo para simular los resultados específicos. Implican cálculos detallados y computacionalmente costosos. Pueden incluir por ejemplo la simulación de procesos fotosintéticos tales como la intercepción de luz, la absorción de dióxido de carbono (CO₂), la respiración y la producción de biomasa en diferentes órganos de la planta, y la pérdida de CO₂ durante la respiración. Se requiere una cantidad importante de información para poder utilizar estos modelos. Los modelos mecanicistas rara vez se utilizan con fines de resolución de problemas; más bien se utilizan con fines académicos para obtener una mejor comprensión de los procesos e interacciones específicas.

Los modelos **funcionales** utilizan métodos simplificados para simular procesos complejos. En algunos casos, los modelos mecanicistas pueden proporcionar información útil que se puede simplificar en funciones empíricas. Por ejemplo, muchos de los modelos funcionales utilizan la radiación solar diaria como la cantidad de energía disponible para la fotosíntesis. La energía interceptada por el cultivo se aproxima utilizando la información tomada del LAI de la planta, para aproximarse a la producción de biomasa utilizando un concepto simple de la eficiencia en el uso de radiación, por ejemplo, la biomasa producida por unidad de radiación interceptada. Si bien estas funciones son más simples, por lo general producen resultados razonables en comparación con las mediciones de campo. La evapotranspiración también se simula utilizando sólo datos diarios del tiempo mediante la incorporación de conceptos similares a los utilizados para la simulación de la producción de biomasa. Ésta es utilizada para mantener un balance de agua en el suelo, para así aproximar cuándo un déficit o un exceso hídrico afecta la producción de biomasa. La mayoría de los modelos funcionales utilizan conceptos de capacidad para describir la cantidad de agua disponible para las plantas en lugar del uso de conceptos de tasa instantánea de la física del suelo. Un límite inferior y superior de la capacidad de agua se define como entrada, y las entradas de agua y salidas en el suelo proporcionan la retroalimentación para determinar la disponibilidad de agua para las plantas. Estos modelos por lo general se ejecutan en incrementos de tiempo diarios mediante el uso de las entradas diarias de precipitaciones, temperatura, radiación, y riego. Utilizan muchos menos datos de entrada en comparación con los modelos mecanicistas, por lo que son más sencillos y útiles para aquellos que no están familiarizados con los procesos biofísicos implicados en las simulaciones. Este tipo de modelos, probados ade-

cuadamente, pueden proporcionar un nivel adecuado de detalle necesario para evaluar varias cuestiones que afectan a la producción de cultivos.

2.6.2. Reseña histórica de modelos y sus características

Los primeros trabajos en modelado de tasas de fotosíntesis se realizaron en la década de los 60 [11]. Los resultados obtenidos a partir de estos modelos se utilizaron entre otros, para estimar el potencial de producción de alimentos para algunas áreas del mundo y para proporcionar indicaciones para el manejo de cultivos. En 1970, se desarrolló el modelo ELCROS (ELementary CROp growth Simulator, [12]). Este simulador incluía un modelo estático de fotosíntesis y la respiración del cultivo era tomada como una fracción fija por día de la biomasa, más una cantidad proporcional a la tasa de crecimiento. Además, añadió un equilibrio funcional entre la raíz y el crecimiento de brotes. Más tarde, en 1978 surgió el modelo BACROS (BAsic CROp growth Simulator, [13]), con la capacidad de predecir la fotosíntesis, la transpiración y la producción de cultivos con pasos de tiempo de una hora o más corto. A medida que el interés pasó de entender el funcionamiento de los cultivos a usos específicos, se desarrollaron modelos más simples que podían ser utilizados en diversos tipos de cultivos. Estos modelos se distinguieron de acuerdo con tres niveles de producción: potencial, limitada en agua, y limitada en nitrógeno. Los modelos de producción potencial, como SUCROS1 (Simple and Universal CROp growth Simulator, [14]), suponen que los rendimientos de los cultivos son determinados por factores climáticos (temperatura y radiación). En los de producción limitada en agua, tales como SUCROS2 [15], se comparan las tasas reales y potenciales de transpiración, y su relación, definida como indicador de estrés, es la base para estimar la fotosíntesis real. En el tercer nivel de producción, se simula la influencia del estrés de nitrógeno pero con un enfoque menos mecánico que en los otros dos niveles de producción (por ejemplo SWHEAT, [16]). El modelado mecánico de los procesos relacionados con el nitrógeno es importante, especialmente si se van a aplicar modelos para evaluar los rasgos de la calidad de los cultivos, particularmente el contenido de proteínas de las semillas. En la Figura 2.3 se muestra gráficamente la evolución de los distintos modelos de crecimiento de cultivos a lo largo del tiempo, enfocada en los modelos de la “Escuela de De Wit” [17].

Otro conjunto de modelos ampliamente utilizados, comenzaron a desarro-

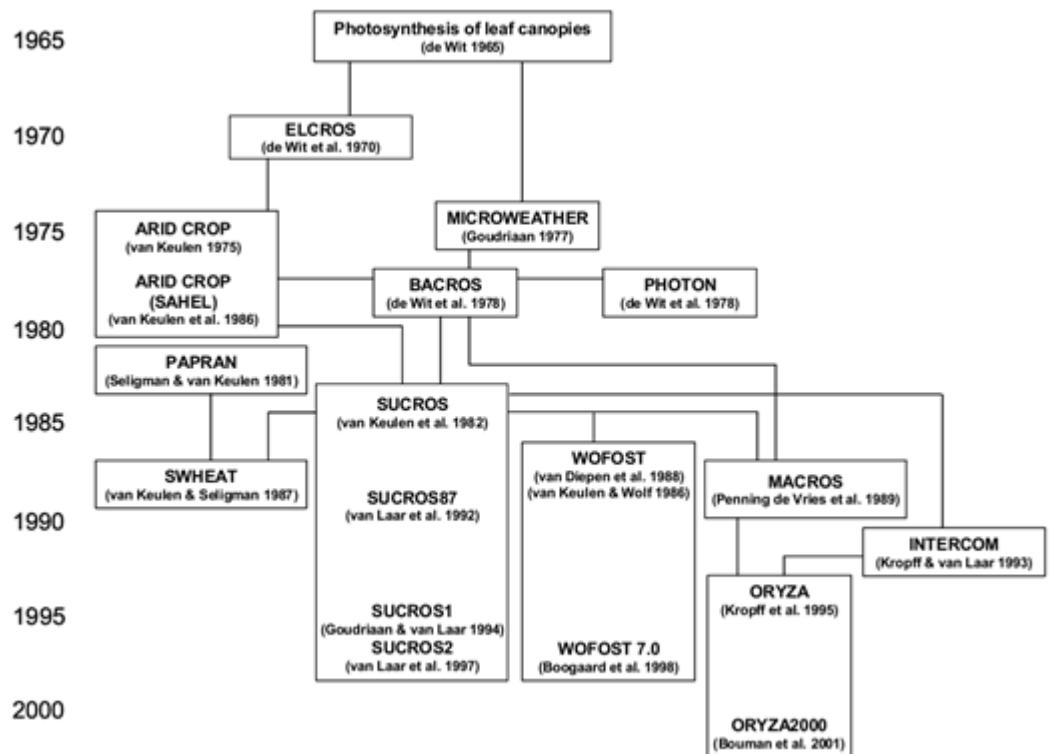


Figura 2.3: Evolución de modelos de crecimiento de cultivo de la “Escuela de De Wit” en donde se observa la transformación de los modelos, desde modelos complejos pensados para comprender el funcionamiento de los cultivos hacia modelos más simples pensados para usos específicos , 1965-2000. Imagen tomada de [11]

llarse en 1982, por el proyecto IBSNAT (International Benchmark Sites Network for Agro-technology Transfer, [10]), de la Universidad de Hawaii, bajo contrato con la Agencia de los Estados Unidos para el Desarrollo Internacional (USAID). Desde el comienzo, IBSNAT fue un intento de demostrar la efectividad de entender las opciones a través del análisis de sistemas y simulaciones para el beneficio final de los agricultores en todo el mundo. En este proyecto, muchas instituciones de investigación, universidades e investigadores de todo el mundo dedicaron una enorme cantidad de tiempo y recursos y se centraron en la producción de un “sistema de apoyo a la toma de decisiones” capaz de simular los riesgos y consecuencias de la elección de distintas alternativas. La definición de la cantidad mínima de datos requerida para la ejecución de simulaciones, la evaluación de resultados y la prueba y aplicación del producto en problemas agrícolas globales que requieren simulaciones de rendimiento específicas en cada lugar. El producto principal del IBSNAT fue un Sistema de Apoyo a la Decisión para la Transferencia de Agro-Tecnología (DSSAT, por sus siglas en inglés, [10]). DSSAT ha sido utilizado como una herramienta de investigación y enseñanza. Como herramienta de investigación, su papel para derivar recomendaciones sobre manejo de cultivos e investigar temas ambientales y de sostenibilidad es incomparable. Los productos DSSAT permiten a los usuarios hacer coincidir los requisitos biológicos de los cultivos con las características físicas de la tierra para proporcionarles opciones de manejo para una mejor planificación del uso de la tierra. También existen usos comerciales de la herramienta, enfocándose los rendimientos para maximizar la rentabilidad. El paquete consta de:

1. sistema de gestión de base de datos para el suelo, el clima, los coeficientes genéticos y los insumos de gestión;
2. modelos de simulación de cultivos;
3. programas de utilidad;
4. programas de generación de tiempo;
5. programa de evaluación de estrategias para evaluar opciones incluyendo elección de variedad, fecha de siembra, densidad de población de plantas, espaciamiento de hileras, tipo de suelo, riego, aplicación de fertilizante, condiciones iniciales en rendimientos, estrés hídrico en las etapas de desarrollo vegetativo o reproductivo y rendimientos netos.

Uno de los modelos más utilizados en el INIA es un desarrollo hecho por

Xinyou Yin y H.H. van Laar. Llamado GECROS (Genotype-by-Environment interaction on CROp growth Simulator), fue presentado en el libro [11], en 2005. El Ing. Agr. Andrés Berger desarrolló una adaptación de este modelo al lenguaje Python⁹ y lo aplicó a datos locales.

Las características principales de GECROS son:

- componentes acoplados directamente para el nitrógeno foliar, Conductancia estomatal, fotosíntesis, transpiración y Senescencia foliar;
- simulación de producción de proteínas de semilla en relación a la disponibilidad de nitrógeno;
- aplicable a la mayoría de los cultivos agrícolas y a cualquier situación de producción (ya sea potencial, o limitada en agua, o limitada por nitrógeno) que estén libres de malas hierbas, plagas y enfermedades.

2.6.3. Modelo utilizado en nuestro proyecto

El modelo utilizado es una adaptación hecha por el INIA para soja de un modelo presentado por G. S. Campbell y R. Díaz en 1988 [18]. Este modelo fue creado con el objetivo de estimar la disponibilidad de agua en la tierra y predecir la evapotranspiración de cultivos, ya que ésta está fuertemente relacionada con la materia seca que producen y el rendimiento. Se decidió la utilización de este modelo a instancias del Ing. Agr. Andrés Berger, ya que según su experiencia, con él se pueden obtener muy buenos resultados teniendo en cuenta los pocos parámetros necesarios para su operación. En particular, es un modelo que especialmente útil para cultivos altamente dependientes del agua, como lo es la soja. Además, ha sido utilizado ampliamente por el INIA, validando sus resultados tanto para cultivos de maíz como de soja. Si utilizamos la clasificación anteriormente presentada, ubicamos este modelo en el grupo de los modelos funcionales, ya que si bien contiene modelado de múltiples procesos involucrados en el crecimiento del cultivo, no realiza cálculos detallados sobre el funcionamiento de estos conceptos, tal como lo haría un modelo mecanicista. Las partes principales del modelo son: intercepción de precipitaciones por el área foliar, escurrimiento de las precipitaciones, filtrado de precipitaciones en el suelo, potencial de evapotranspiración, evaporación en la superficie del suelo, transpiración, crecimiento de raíces y captación de agua de raíces. Los

⁹<https://www.python.org/>

intervalos de simulación son diarios, y los principales parámetros de entrada son:

- datos meteorológicos: mm de precipitaciones, temperatura máxima, temperatura mínima (por día).
- fechas de plantación, emergencia, floración y maduración.
- capacidad de campo (field capacity) y punto de marchitez permanente (permanent wilting point).
- coeficiente de extinción de luz.
- proporción materia seca / agua (cuánto se produce de materia seca por litro de agua evaporada). Típicamente denominado eficiencia del uso del agua (water use efficiency)
- profundidad máxima de enraizamiento.

2.7. Investigaciones que utilizan NDVI, rendimiento de cultivos y simuladores de crecimiento de cultivos

Como se mencionó en la introducción de este capítulo, existen múltiples investigaciones que intentan predecir alguna(s) variable(s) relativa(s) al rendimiento de un cultivo mediante la utilización de simuladores de crecimiento de cultivos y datos tomados de distintos sensores remotos. En esta sección comentaremos algunas de ellas. La primer investigación a comentar es [19]. Este estudio se realizó en el norte de la provincia de Buenos Aires, Argentina. El área de estudio se puede dividir en función del tipo de suelo en dos subzonas homogéneas: una subzona con menor contenido de arcilla en el suroeste y una subzona con mayor contenido de arcilla en el noreste (ver Figura 2.4). Se estudiaron nueve zonas de cultivo (2003-2011) de trigo. En los primeros cinco años, se calibró y validó un modelo empírico con rendimientos de trigo observados en el campo y los datos del producto MODIS NDVI, mientras que en los otros cuatro años el modelo calibrado se aplicó mediante mapas de rendimiento y comparándolos con los rendimientos oficiales. Como conclusión, se encontró un valor de coeficiente de determinación de 0,75 después de relacionar la estimación del modelo empírico calibrado con los rendimientos oficiales. El coeficiente de determinación es el cuadrado del coeficiente de correlación de Pearson, y da

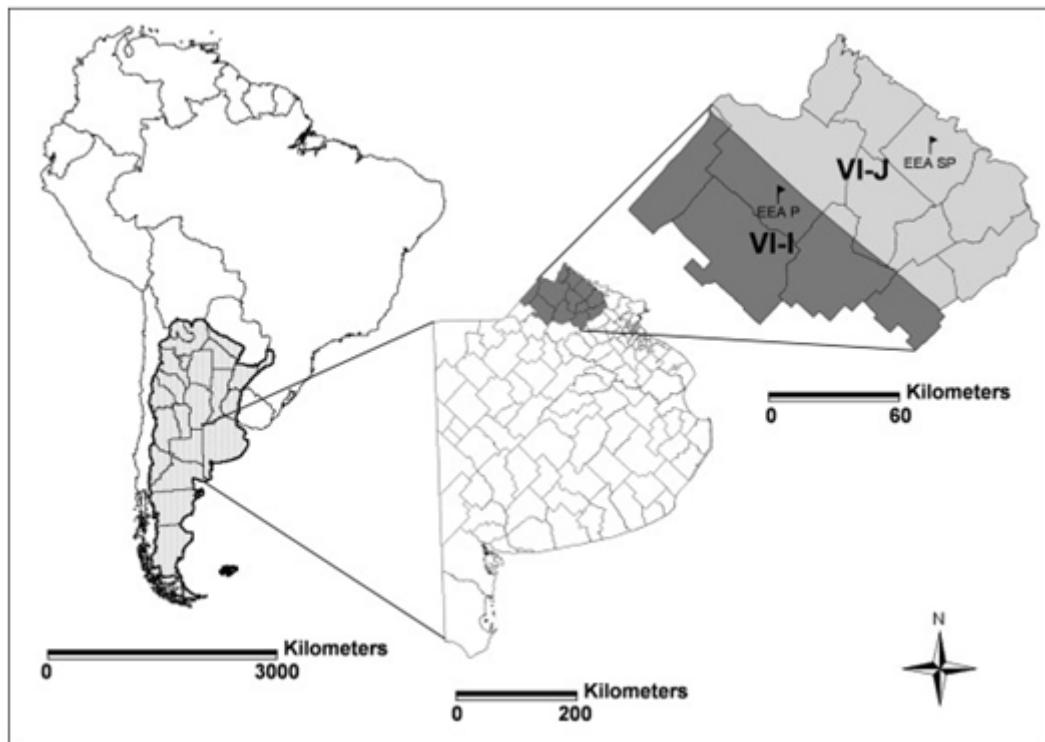


Figura 2.4: Zona VI-I con menor contenido de arcilla y Zona VI-J con mayor contenido de arcilla. Imagen tomada de [19].

la proporción de variación de la variable Y (variable predecida) que es explicada por la variable X (variable predictor o explicativa). Si la proporción es igual a 0, significa que la variable predictor no tiene capacidad predictiva de la variable a predecir (Y). Cuanto mayor sea la proporción, mejor sería la predicción. Si llegara a ser igual a 1 la variable predictor explicaría toda la variación de Y, y las predicciones no tendrían error.¹⁰ A través del modelo calibrado, el NDVI percibió la variabilidad temporal y espacial en el área de estudio. Fue posible hacer comparaciones objetivas entre zonas y temporadas de cultivo. El modelo calibrado fue una buena herramienta para predecir el rendimiento del trigo en las zonas estudiadas. La variación interanual, observada en la descripción de la relación entre el NDVI y el rendimiento de los datos analizados, indica la posibilidad de considerar condiciones agrometeorológicas para obtener precisión en la estimación del rendimiento. Esto requiere más investigación, especialmente, con el fin de utilizar el número mínimo de insumos o datos disponibles. El modelo desarrollado en este estudio debe ser probado en otras regiones húmedas de trigo para confirmar los resultados encontrados en el norte de Buenos Aires.

¹⁰https://es.wikipedia.org/wiki/Coeficiente_de_determinaci%C3%B3n

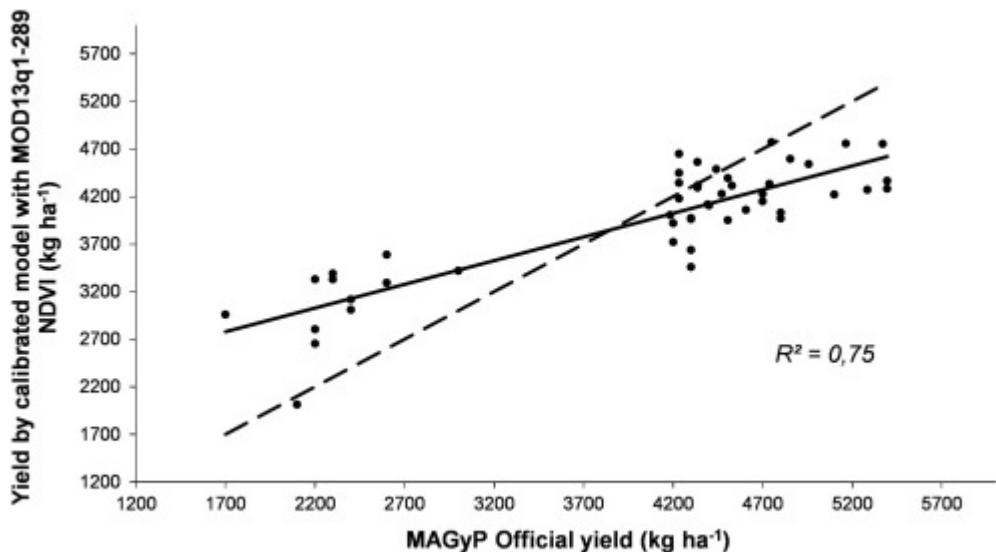


Figura 2.5: Relación entre la estimación del modelo empírico calibrado y los rendimientos oficiales de MAGyP para el rendimiento promedio de trigo para cada departamento en cuatro temporadas de crecimiento (2008-2011). Número de departamentos: 12. Imagen tomada de [19].

El estudio "Managing Wheat From Space" [20] exploró las relaciones entre el NDVI tomado de las imágenes MODIS, las observaciones de la fracción cubierta de verde (FGrC) tanto simulado como medido, el LAI y la biomasa para el trigo de tierras secas en Australia. Se muestrearon 37 lotes obteniendo datos in situ de la biomasa en la Antesis, la fracción cubierta de verde (FGrC), el NDVI (desde un sensor óptico) desde el 2003 hasta el 2006 y para el 2012. Además se utilizó el NDVI del satélite MODIS para el mismo período. Las mediciones de FGrC y NDVI de 2012 se ajustaron a MODIS NDVI, resultando en relaciones lineales positivas cuando los valores de MODIS NDVI fueron menores a 0.80. El AGB medido también se relacionó positiva y linealmente con MODIS NDVI, dando como resultado un Coeficiente de Determinación global de 0,81 y un Error Cuadrático Medio (RMSE, Root Mean Square Error) de 1397 kg / ha. El RMSE es una medida de uso frecuente de las diferencias entre valores (valores muestrales y poblacionales) predichos por un modelo o un estimador y los valores realmente observados. Si \hat{Y} es un vector de n predicciones y Y es el vector de los valores reales, entonces: $RMSE = 1/n \sum_{i=1}^n (\hat{Y}_i - Y_i)^2$. El RMSE representa la desviación estándar de la muestra de las diferencias entre

los valores predichos y los valores observados. Estas diferencias individuales se llaman residuos cuando los cálculos se realizan sobre la muestra de datos que se utilizó para la estimación, y se llaman errores de predicción cuando se calcula fuera de la muestra.¹¹ En la investigación, se realizaron simulaciones de crecimiento de cultivos para catorce lotes desde 2003 hasta 2006, y seis lotes en 2012.

¹¹https://en.wikipedia.org/wiki/Root-mean-square_deviation

Capítulo 3

Descripción del problema

3.1. Introducción

Actualmente en el Instituto Nacional de Investigación Agropecuaria (La Estanzuela) se utilizan modelos de simulación de crecimiento de cultivos como parte de los servicios de asesoramiento a productores. Estos modelos reciben como entrada un conjunto de parámetros relativos a las características iniciales del suelo, las características del cultivo, y de las condiciones climáticas donde se cultiva. En base a este conjunto de parámetros, más una serie de escenarios posibles de tiempo que se plantean, es posible estimar el crecimiento/rendimiento que tendrá el cultivo. Según el Ing. Agr. Andrés Berger del INIA los resultados de estas predicciones son muy buenos si se cuenta con una cantidad de parámetros suficiente y de buena calidad. Sin embargo, no siempre es posible obtener información suficiente para calibrar estos modelos, y para hacerlo se necesita mucho trabajo manual, incluyendo pruebas de “ensayo y error”. El modelo utilizado en este proyecto, puede usarse para distintos cultivos mediante una calibración adecuada de algunos parámetros para cada uno de ellos. En este caso, considerando que el INIA cuenta con una calibración adecuada del modelo para simular el crecimiento de la soja en nuestro país y que dicha calibración ha sido testeada obteniendo buenos resultados, se utilizará el modelo con la parametrización calibrada para la simulación del crecimiento de la soja y se trabajará siempre con este cultivo. Pero, es bueno tener presente que esta investigación puede extenderse a otros cultivos utilizando el mismo simulador adecuadamente calibrado para el cultivo sobre el que se quiera trabajar.

3.2. Objetivos

Como se comentó en la Sección 2.5.2, existe una relación entre el NDVI y el rendimiento de un determinado cultivo y existen diversas investigaciones que intentan obtener dicha relación para un cultivo en un punto geográfico determinado. Por tanto, si pudiéramos predecir el NDVI para una zafra futura, podríamos potencialmente predecir el rendimiento del cultivo. Dicho esto, el objetivo principal de esta investigación es, a partir de imágenes satelitales y del simulador de crecimiento de soja provisto por el INIA, poder predecir el NDVI de la soja para un conjunto de chacras sobre las que se poseen datos históricos de cinco años de siembras de soja. En cuanto a las imágenes satelitales a utilizar, al momento de realizar esta investigación existen dos fuentes de datos posibles con distintas características: LANDSAT y MODIS. Otro objetivo del proyecto es investigar la posibilidad de utilizar ambas fuentes de datos, considerando que los datos que proveen, a priori, podrían complementarse. En secciones posteriores se explican las características principales de cada una de las fuentes de imágenes satelitales y de los datos históricos que se poseen sobre las chacras en las que se sembró soja.

3.3. Fuentes de datos y su calidad

Con el objetivo de que la solución a alcanzar sea aplicable en la mayor cantidad de casos posibles, se partió de la base de pocos requerimientos de datos previos, ya que muchas veces no es posible acceder a datos suficientes o con la calidad necesaria para la aplicación de modelos. Por lo tanto, se definió que sería necesario contar con información de al menos cinco zafras en cada chacra a estudiar, conteniendo los datos de:

- fecha de siembra.
- fecha de cosecha.
- variedad de semilla (determina tiempo a floración y madurez).

Además, se necesita de las chacras un polígono georeferenciado(shapefile), y datos diarios meteorológicos en la zona durante las zafras, incluyendo:

- precipitaciones (mm/m^2).
- temperatura mínima.
- temperatura máxima.

3.3.1. Imágenes satelitales

Debido a la imposibilidad de obtener mediciones del índice verde *in situ*, en especial para zafras anteriores, la posibilidad de contar con imágenes satelitales es fundamental para realizar el monitoreo de la evolución de los cultivos. El monitoreo por imágenes satelitales, a pesar de ser menos exacto, tiene la ventaja de no ser destructivo para el cultivo, y brinda la posibilidad de contar con información más frecuentemente que con análisis *in situ*. Al momento de comenzar esta investigación, se evaluaron dos fuentes de imágenes satelitales: Landsat y MODIS. Los Landsat son una serie de satélites construidos y puestos en órbita por EE.UU. para la observación en alta resolución de la superficie terrestre. Los satélites de la serie Landsat que estuvieron en órbita durante el período de tiempo sobre el que se poseen datos de zafras de soja son: el Landsat 7, operativo desde 1999 hasta 2013 y el Landsat 8, operativo desde 2013 hasta la actualidad. Para las bandas necesarias en la construcción del NDVI (Infrarrojo Cercano y Rojo), Landsat 7 posee una resolución de 30 metros y una frecuencia de una imagen cada 16 días. El tamaño aproximado de cada imagen es de 170 km norte-sur por 183 km este-oeste. Debido a un fallo en un corrector (SLC, Scan Line Corrector), las imágenes obtenidas luego del 30 de mayo del 2003 tienen faltantes de información, quedando utilizables aproximadamente un 78 % de los píxeles. Si bien existen técnicas para llenar los huecos en las imágenes, hacen uso de imágenes de días o años anteriores, por lo cual la información recuperada difícilmente sea de utilidad para el objetivo de esta investigación.

Landsat 8, posee también una resolución de 30m y una frecuencia de una imagen cada 16 días, pero con un corrimiento de 8 días con respecto a las obtenidas por Landsat 7. El tamaño de cada imagen es de 190 km norte-sur por 180 km este-oeste. Aunque Landsat 8 fue creado para ser consistente con los anteriores, debido a la diferencia en sensores utilizados las bandas que componen el NDVI no son iguales, por lo que la continuidad temporal del índice no es la ideal. En la Tabla 3.1 se muestran estas diferencias. En el artículo [21] se llega a la conclusión de que si bien las diferencias son pequeñas, podrían tener un impacto significativo dependiendo del uso que se le dé a los datos.

MODIS (Moderate Resolution Imaging Spectroradiometer, o Espectrorradiómetro de imágenes de media resolución en español) es un instrumento

¹<https://landsat.usgs.gov/landsat-8-18-data-users-handbook>

Landsat-7 ETM+ Bands (μm)			Landsat-8 OLI and TIRS Bands (μm)		
			30 m Coastal/Aerosol	0.435 - 0.451	Band 1
Band 1	30 m Blue	0.441 - 0.514	30 m Blue	0.452 - 0.512	Band 2
Band 2	30 m Green	0.519 - 0.601	30 m Green	0.533 - 0.590	Band 3
Band 3	30 m Red	0.631 - 0.692	30 m Red	0.636 - 0.673	Band 4
Band 4	30 m NIR	0.772 - 0.898	30 m NIR	0.851 - 0.879	Band 5
Band 5	30 m SWIR-1	1.547 - 1.749	30 m SWIR-1	1.566 - 1.651	Band 6
Band 6	60 m TIR	10.31 - 12.36	<i>100 m TIR-1</i>	<i>10.60 - 11.19</i>	Band 10
			<i>100 m TIR-2</i>	<i>11.50 - 12.51</i>	Band 11
Band 7	30 m SWIR-2	2.064 - 2.345	30 m SWIR-2	2.107 - 2.294	Band 7
Band 8	15 m Pan	0.515 - 0.896	15 m Pan	0.503 - 0.676	Band 8
			30 m Cirrus	1.363 - 1.384	Band 9

Figura 3.1: Comparación de bandas espectrales ETM+ (Landsat 7) con bandas espectrales OLI y TIRS (Landsat 8). Tabla tomada del manual de usuario de Landsat 8¹.

científico presente en dos satélites: Terra² (puesto en órbita en diciembre de 1999) y Aqua³ (puesto en órbita en mayo de 2002). Ambos orbitan la tierra 16 veces por día, obteniendo imágenes de cada punto en la tierra una vez por día. El producto NDVI disponibilizado tiene una resolución temporal de 16 días para cada satélite, debido a que se compone con varias imágenes, descartando píxeles de mala calidad (por ejemplo a causa de la nubosidad) y eligiendo el valor del píxel más próximo al nadir de entre los dos valores de NDVI más alto⁴. Cada imagen posee un tamaño de 10 grados de latitud y longitud, representado en una grilla (raster) de 4800 filas por 4800 columnas, dando una resolución de 250 metros. Debido a que ambos satélites están sincronizados y poseen el mismo sensor, los productos NDVI de ambos (MOD13Q1⁵ y MYD13Q1⁶) pueden ser combinados obteniendo una mayor resolución temporal (8 días) con imágenes de buena calidad, sin nubes ni defectos.

²<https://terra.nasa.gov/about>

³<http://aqua.nasa.gov/content/about-aqua>

⁴<https://modis.gsfc.nasa.gov/data/dataproducts/mod13.php>

⁵https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13q1

⁶https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/myd13q1

3.3.2. Información de zafras

Para poder estudiar la evolución del NDVI en las distintas chacras, es necesario conocer algunos datos. Primero que nada, se requiere conocer el período de tiempo sobre el que se debe estudiar la evolución del NDVI para cada chacra/lote. Para ello, es fundamental contar con las fechas de siembra y de cosecha de soja para cada chacra/lote. Además, como se plantea obtener la evolución del NDVI a partir de imágenes satelitales, es necesario poder georeferenciar las chacras/lotes a estudiar. Estos datos fueron provistos por el INIA a través de los siguientes archivos:

- dos planillas de cálculo, uno con datos de la siembra y otro con datos de la cosecha.
- Un archivo pdf, que muestra la distribución y el tamaño de los lotes dentro de una chacra.
- Un archivo shapefile que permite georeferenciar los lotes de las chacras.

A continuación se describen los datos recibidos en cada uno de estos archivos y se muestran ejemplos de los mismos.

Los archivos con datos de siembra y cosecha, contienen los siguientes datos:

- Fecha de siembra: fecha en la que se sembró la soja.
- Fecha de cosecha: fecha en la que se cosechó la soja.
- Campo: nombre de la chacra en la que se realizó la siembra.
- Lote: identificador del lote dentro del campo en el que se realizó la siembra.
- Ha: cantidad de hectáreas sembradas.
- Dosis x Ha: indica la dosis por hectárea de soja sembrada.
- Cultivo: puede tomar dos valores SOJA ó SOJA II. Cuando es SOJA, indica que el cultivo es de primera, es decir, que no fue precedido por un cultivo invernal. Por el contrario, cuando es SOJA II indica que el cultivo es de segunda, o sea, que fue precedido por un cultivo invernal.
- Zona: indica la zona donde se encuentra la chacra en la que se realizó la siembra.
- Artículo: nombre comercial de la variedad de soja sembrada.

Las dos planillas de cálculo poseen varias hojas y en cada hoja se indican los datos de siembra (en el archivo correspondiente a los datos relativos a la

Fecha Aplicación	Campo	Lote	Articulo	Ha	Dosis x Ha	Cultivo	Zona
03/01/2010	8001-CUFRE	806,2	SOJA ADP 6126 (SEMILLA)	39,5	0,099	SOJA II	COLONIA
05/01/2010	8001-CUFRE	803,2	SOJA ADP SPS 600	70	0,071	SOJA II	COLONIA
05/01/2010	8001-CUFRE	803,2	SOJA ADP 6126 (SEMILLA)	47	0,143	SOJA II	COLONIA
05/01/2010	8001-CUFRE	808,2	SOJA ADP 6126 (SEMILLA)	70	0,1	SOJA II	COLONIA
10/01/2010	8006-EL ENCUENTRO	1,2	SOJA ADP 6126 (SEMILLA)	131	0,107	SOJA II	COLONIA
10/01/2010	8006-EL ENCUENTRO	1,2	SOJA ADP 6126 (SEMILLA)	33	0,1	SOJA II	COLONIA
10/01/2010	8006-EL ENCUENTRO	4,2	SOJA ADP 6126 (SEMILLA)	100	0,1	SOJA II	COLONIA
15/01/2010	8006-EL ENCUENTRO	5,2	SOJA ADP 6126 (SEMILLA)	171	0,105	SOJA II	COLONIA

Tabla 3.1: Siembra zafra 2009/2010

Fecha Aplicación	Campo	Lote	Articulo	Ha	Cultivo	Zona
16/05/2010	8001-CUFRE	803,2	COSECHA	98,5	SOJA II	COLONIA
30/05/2010	8001-CUFRE	808,2	COSECHA	70	SOJA II	COLONIA
30/05/2010	8001-CUFRE	806,2	COSECHA	39,5	SOJA II	COLONIA
31/05/2010	8001-CUFRE	803,2	COSECHA	19	SOJA II	COLONIA
01/06/2010	8006-EL ENCUENTRO	5,2	COSECHA	171	SOJA II	COLONIA
05/06/2010	8006-EL ENCUENTRO	4,2	COSECHA	100	SOJA II	COLONIA
07/06/2010	8006-EL ENCUENTRO	1,2	COSECHA	164	SOJA II	COLONIA

Tabla 3.2: Ejemplo de datos de cosecha de la zafra 2009/2010

siembra) y datos de cosecha (en el archivo correspondiente a los datos relativos a la cosecha) de todas las chacras para una zafra. Por ejemplo, para la zafra 2009/2010 el archivo con datos de siembra posee una hoja con los datos de la Tabla 3.1.

Por ejemplo, el archivo con datos de cosecha posee una hoja para la temporada 2009/2010 con los datos de la Tabla 3.2.

Además de estas planillas, se posee para cada chacra un archivo que muestra la distribución y el tamaño de los lotes dentro de una chacra. Un ejemplo de estos archivos se presenta en la Figura 3.2a y 3.2b. En el Apéndice 1 se detallan todas las chacras utilizadas en este proyecto y sus detalles geográficos.

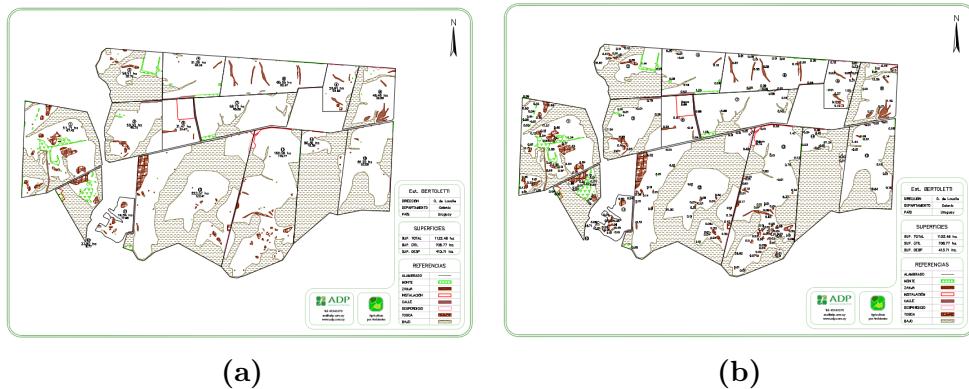


Figura 3.2: Informacion geográfica de la chacra Bertoletti.

3.3.3. Datos meteorológicos

Los datos meteorológicos son utilizados por el simulador de crecimiento para estimar la evolución del cultivo, a partir de la cual se estimará el NDVI. Para cada día de todas las zafra estudiadas, se requiere medidas meteorológicas básicas (precipitaciones, temperatura mínima, temperatura máxima, radiación solar) en cada punto geográfico de interés. Ninguna de las chacras estudiadas cuenta con sensores meteorológicos, por lo que es necesario estimar esta información a partir de sensores remotos. Esta estimación no es trivial, y para lograr buena precisión se requiere interpolar distintas fuentes de información (algunas de las cuales no son públicas). Realizar este complejo proceso se encuentra fuera del alcance del proyecto, por lo que los datos meteorológicos utilizados fueron provistos por el proyecto “Agronómico” de la ANII⁷, a cargo del Dr. Ing. Pablo Rodríguez-Bocca.

3.4. Forma de evaluar la solución

Como se comentó en la Sección 3.2, el objetivo principal de esta investigación es, a partir de imágenes satélitales y del simulador de crecimiento de soja provisto por el INIA, poder predecir el NDVI de la soja para un conjunto de chacras sobre las que se poseen datos históricos de cinco años de siembras de soja. Para poder evaluar y comparar las soluciones obtenidas, es necesario elegir una forma de evaluación de las mismas. En esta investigación, el resultado de una solución de una instancia del problema, consistirá de una estimación

⁷PPI_X_2014_1_14246

del NDVI de un lote de una chacra para una zafra. Dicho resultado, deberá ser comparado con el NDVI “real” obtenido a partir de imágenes satelitales. La forma elegida de evaluar la solución obtenida es a través del RMSE (Root Mean Square Error). Además de medir el error proveniente de la estimación, es necesario poder comparar dicho error con el error proveniente de algún otro método de estimación sencillo. Para ello se construyó un estimador *baseline* que se explica en la Sección 4.6.

3.5. Explicación del problema a partir de una instancia

A modo de clarificar lo expuesto en este capítulo, a continuación se explica el problema a resolver y la forma de evaluación de la solución a partir de una instancia del mismo. Tomemos como ejemplo, el lote 37 de la chacra “ALVAREZ OLANO”. Como se comentó anteriormente para dicho lote se cuenta con los siguientes datos:

- dos planillas de cálculo, uno con datos de la siembra y otro con datos de la cosecha.
- un archivo, que muestra la distribución y el tamaño de los lotes dentro de una chacra.
- un archivo shapefile que permite georeferenciar los lotes de las chacras.

En las Tablas 3.3 y 3.4 se muestran estos datos para el lote a estudiar.

Zafra	Fecha Aplicación (siembra)	Campo	Lote	Artículo	Ha	Dosis X Ha	Cultivo	Zona
2009/2010	14/12/2009	ALVAREZ OLANO	37	SOJA ADP 4910 (SEMILLA)	78	0,094	SOJA II	DOLORES
2010/2011	07/10/2010	ALVAREZ OLANO	37	SOJA ADP DON MARIO 6202	78	0,077	SOJA II	DOLORES
2012/2013	01/11/2012	ALVAREZ OLANO	37	SOJA BIO-SOJA 6.5 (BASICA)	78	0,081	SOJA	DOLORES
2013/2014	24/11/2013	ALVAREZ OLANO	37	SOJA BIO 6.5 (C1)	78	0,085	SOJA II	DOLORES
2014/2015	12/11/2014	ALVAREZ OLANO	37	SOJA ADP 5909	78	0,078	SOJA II	DOLORES

Tabla 3.3: Tabla de con datos de siembra de chacra “ALVAREZ OLANO”

Zafra	Fecha Aplicación (cosecha)	Campo	Lote	Artículo	Ha	Cultivo	Zona
2009/2010	12/04/2010	ALVAREZ OLANO	37	COSECHA	78	SOJA II	DOLORES
2010/2011	05/04/2011	ALVAREZ OLANO	37	COSECHA	78	SOJA II	DOLORES
2012/2013	21/05/2013	ALVAREZ OLANO	37	COSECHA	76	SOJA	DOLORES
2013/2014	20/04/2014	ALVAREZ OLANO	37	COSECHA	78	SOJA II	DOLORES
2014/2015	29/04/2015	ALVAREZ OLANO	37	COSECHA	78	SOJA II	DOLORES

Tabla 3.4: Tabla de con datos de cosecha de chacra “ALVAREZ OLANO”

3.5.1. Datos de la chacra

En las Figuras 3.3 y 3.4 puede apreciarse la información geográfica de la chacra “ALVAREZ OLANO”

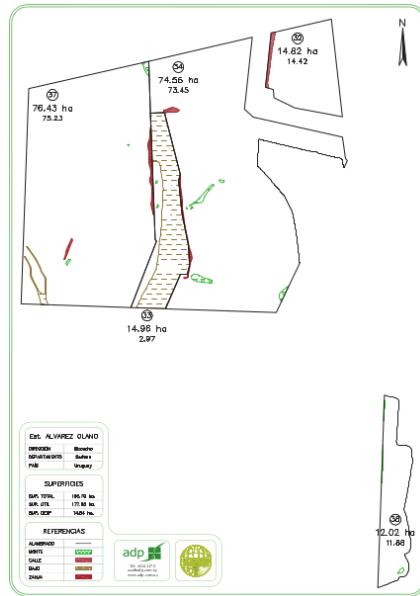


Figura 3.3: Informacion geográfica de la chacra Alvarez Olano.



Figura 3.4: En la imagen se observa el archivo shapefile correspondiente la chacra Alvarez Olano abierto mediante la herramienta Google Earth Pro. El lote 37 de la chacra está pintado de violeta.

3.5.2. Datos meteorológicos

El simulador de crecimiento de cultivo requiere de un grupo reducido de datos meteorológicos para cada día. Ellos son:

- Temperatura máxima y mínima.
- Precipitaciones.
- Radiación solar.

En la Tabla 3.5, se muestran a modo de ejemplo estos datos para los meses de Enero y Febrero de 2010. Vale destacar que estos mismos datos se poseen para todas los días de las zafras en estudio, solo se expone un conjunto reducido de datos para que el documento no sea tan extenso sin propósito alguno.

Fecha	Temperatura máxima (°C)	Temperatura mínima (°C)	Precipitaciones (mm)	Radiación solar (W/m2)
2010-01-01	28.3	14.6	0	23.9
2010-01-02	30.9	14.3	0	29.2
2010-01-03	31.2	19.5	1	27.3
2010-01-04	29.4	19.5	2.9	20.9
2010-01-05	31.5	20.8	0.9	27.6
2010-01-06	24.7	15.5	0	21.9
2010-01-07	26	12.1	0	25.9
2010-01-08	28.9	14.8	0	29
2010-01-09	31.7	19.4	0	25.6
2010-01-10	30.8	21.9	2.1	21.5
2010-01-11	31.1	19	19.2	29.6
2010-01-12	24.6	17.2	0	21
2010-01-13	27.3	13.5	0	30.8
2010-01-14	29.7	15.2	0	30.6
2010-01-15	30.6	18.7	0	26.9
2010-01-16	25.5	18.3	0	17
2010-01-17	33.1	17.2	0	30.2
2010-01-18	32	19.1	10.7	17.6
2010-01-19	24.3	17.7	0	13.5
2010-01-20	28.5	15.9	0	30.3
2010-01-21	31.4	14.1	0	30.2
2010-01-22	30.3	19	0	28.6
2010-01-23	30.6	19.3	0	27.3
2010-01-24	30.9	18.4	0	29.8
2010-01-25	32.1	20	0	28.9
2010-01-26	30.8	19.8	0	29
2010-01-27	31.7	19.5	0	28.8
2010-01-28	32	20.3	0	28.9
2010-01-29	32.6	21	0	29.2
2010-01-30	33.6	21.8	0	21
2010-01-31	26.7	21.3	4.4	12
2010-02-01	32.5	21	0	21.4
2010-02-02	28.2	22	2.1	12
2010-02-03	30.6	21.2	63.1	13.8
2010-02-04	26.7	19.6	0	11.5
2010-02-05	25.1	19.2	77.3	8.9
2010-02-06	26.8	20.4	0	12.2
2010-02-07	26.2	20.1	23.3	14.4
2010-02-08	28.7	17.5	0	25
2010-02-09	31.4	16.3	0	28.5
2010-02-10	31.1	19.3	0	25.6
2010-02-11	30.8	20.6	0	25.5
2010-02-12	31.1	21.7	0.5	26.7
2010-02-13	33.9	22.3	0.7	18.5
2010-02-14	25.4	19.1	0	21.1
2010-02-15	26.9	17.8	3.3	10.4
2010-02-16	24	18.1	2.6	13.5
2010-02-17	26.6	16.7	0	22.3
2010-02-18	29.7	18	0	27
2010-02-19	32.3	21.8	10.9	17.8
2010-02-20	26.9	22.1	5	8.2
2010-02-21	25.7	21.7	0	9
2010-02-22	27.9	21.5	5.2	10.5
2010-02-23	24.2	15.8	0	18
2010-02-24	21.7	14.1	0	25.1
2010-02-25	21.3	11.2	0	25.4
2010-02-26	24.3	11	0	24
2010-02-27	25.5	15.3	0	24.5
2010-02-28	25.6	15.2	0	23.3

Tabla 3.5: Ejemplo de datos del tiempo para una chacra.

3.5.3. Ejemplo de solución

Luego de presentados los datos que se poseen para este lote, el objetivo es, a partir de los datos de las zafras 2009/2010, 2010/2011, 2012/2013 y 2013/2014 y los datos sobre el NDVI obtenido de las imágenes satelitales (MODIS y Landsat), poder calibrar algunos de los parámetros de entrada del simulador de crecimiento de cultivos provisto por el INIA. Luego de calibrados los parámetros de entrada del simulador, “ejecutar” una simulación para la zafra 2014/2015, comparar los resultados obtenidos de NDVI con el simulador con los “reales” (extraídos de las imágenes satelitales) y cualificar la solución obtenida mediante RMSE.

Consideraremos una celda dentro del lote 37 de la chacra ALVAREZ OLANO, entonces la Figura 3.5 muestra la gráfica de la evolución del NDVI para todas las zafras menos la última de dicha celda (2009/2010, 2010/2011, 2012/2013, 2013/2014). A partir de estas series, se debe calibrar el simulador de crecimiento de cultivos para posteriormente simular el resultado de la última zafra (2014/2015).

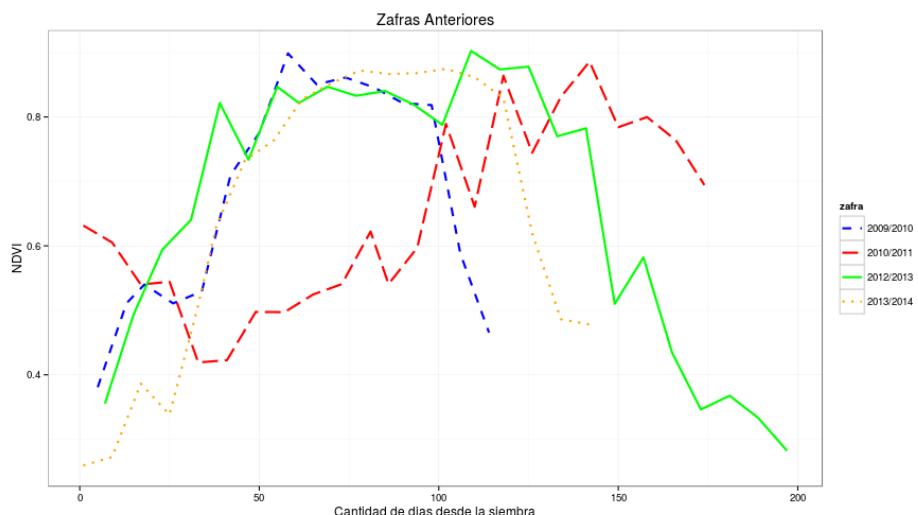


Figura 3.5: En la imagen se observa las series de NDVI de todas las zafras menos la última para una celda del lote 37 de la chacra Alvarez Olano.

En la Figura 3.6 se muestra el resultado de la ejecución de la simulación para la última zafra (2014/2015) luego de calibrado el simulador de crecimiento de cultivos. Además de la estimación (curva), se observan los datos de NDVI reales para dicha zafra extraídos de imágenes satelitales.

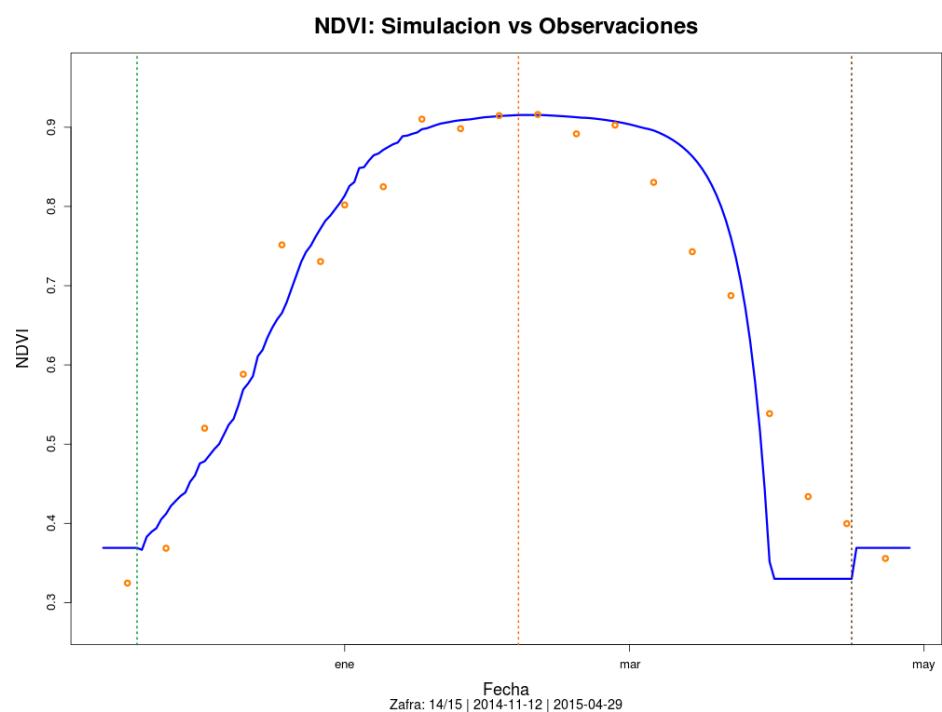


Figura 3.6: En la imagen se observa la estimación de la última zafra (curva) comparada con los datos reales extraídos de imágenes satelitales (puntos).

Capítulo 4

Solución del problema

4.1. Introducción

En este capítulo se describe la solución planteada para el problema. Para ello, enumeraremos los pasos seguidos y describiremos cómo fueron realizados, las herramientas utilizadas, los inconvenientes y desafíos que surgieron, y la forma en que fueron resueltos. La solución del problema planteada, puede describirse como el siguiente conjunto de pasos a seguir:

1. Obtención de las imágenes satelitales y procesamiento.
2. Extracción de datos del INIA.
3. Caracterización de las variedades de soja.
4. Cruzamiento de datos.
5. Obtención de series de NDVI a partir del cruzamiento de datos y las imágenes satelitales.
6. Creación de estimadores *baseline*.
7. Calibración del simulador de crecimiento de cultivos en zafras conocidas.
8. Simulación del crecimiento de cultivo para la zafra a estimar.
9. Evaluación de la calidad de la estimación con RMSE.

4.2. Obtención de las imágenes satelitales y procesamiento

Como se mencionó en la Sección 3.3, las fuentes de imágenes satelitales utilizadas en este proyecto son MODIS y Landsat. Las imágenes de ambos

satélites pueden ser descargadas desde Internet libremente. En este caso, las imágenes de MODIS fueron descargadas desde la web de la Universidad de Recursos Naturales de Viena¹, mientras las imágenes de Landsat fueron descargadas desde Servicio Geológico de los Estados Unidos (USGS)², a través de la herramienta EarthExplorer³.

Si bien ambas fuentes de datos son de libre acceso, la facilidad para conseguirlas no es igual. Mientras en el caso de MODIS se pueden descargar desde un repositorio FTP, las de Landsat deben ser encargadas en EarthExplorer eligiendo diversos parámetros tales como la zona de interés, el rango de fechas, y el producto deseado (en este caso Landsat 7 y 8). Luego, se debe ingresar el listado de imágenes a descargar en otra parte de la herramienta, que las dejará disponibles para su descarga en un software propio llamado Bulk Download Application⁴, que se utiliza para descargar las imágenes automáticamente.

Para el procesamiento de las imágenes y la extracción de los datos necesarios para el estudio se utilizó el lenguaje de programación R. Para la manipulación de las imágenes satelitales se utilizó el paquete “raster”⁵ y para el manejo de shapefiles el paquete “rgdal”⁶.

En el caso de MODIS, el producto descargado ya está procesado, y consiste de imágenes (una cada 8 días, como se mencionó en la Sección 3.3.1) en formato GeoTIFF (.tif) en la que cada píxel tiene el valor NDVI.

En cambio, el producto que se obtiene en el caso de Landsat, contiene múltiples imágenes, una para cada banda (1 al 8). Además, se le deben aplicar correcciones atmosféricas y de nubes utilizando los metadatos de las imágenes⁷⁸.

4.3. Extracción de datos del INIA

Con el objetivo de que los datos provistos por el INIA sean procesados por un sistema informático, es necesario transformarlos desde archivos Excel a otro formato cuyo procesamiento por parte de un sistema informático sea

¹<http://ivfl-info.boku.ac.at/index.php/eo-data-processing>

²<https://landsat.usgs.gov/>

³<https://earthexplorer.usgs.gov/>

⁴<https://earthexplorer.usgs.gov/bulk>

⁵<https://cran.rstudio.com/web/packages/raster/raster.pdf>

⁶<https://cran.r-project.org/web/packages/rgdal/rgdal.pdf>

⁷<https://landsat.usgs.gov/using-usgs-landsat-8-product>

⁸<https://landsat.usgs.gov/landsat-surface-reflectance-high-level-data-products>

más sencillo. Por ello, convertimos los archivos recibidos desde el INIA que se mencionan en 3.5 con formato Excel a archivos de tipo CSV (Comma Separated Values). El resultado de dicha transformación es una archivo CSV que posee las siguientes columnas:

- Zafra: indíca los años correspondientes a la zafra (Ejemplo. 09/10 identifica la zafra llevada a cabo entre los años 2009 y 2010).
- Chacra: nombre de la chacra.
- Lote: identificador del lote.
- Archivo: ruta al archivo que contiene la información geográfica de la chacra (shapefile).
- Variedad: Variedad de la soja sembrada (nombre comercial de la variedad de la soja).
- Dosis x ha: Dosis por hectárea de semillas sembradas.
- Tipo de siembra: Indica si la siembra es de primera o segunda.
- Fecha Aplicación: fecha en la que se realizó la siembra.
- Fecha Cosecha: fecha en la que se realizó la cosecha.
- Archivo Meteorológico: indica la ruta al archivo donde se encuentran los datos meteorológicos para esa chacra.

Por otro lado, teniendo en cuenta que en este tipo de investigaciones la calidad de los datos es un factor de vital importancia, al transformar los datos revisamos manualmente la coherencia de los mismo. En este caso, consideramos que la verificación de la calidad de los datos implica chequear los siguientes puntos sobre los mismos:

- Fechas de siembra y cosecha: las fechas de cosecha debe ser posterior a la de siembra y el intervalo transcurrido entre estas fechas debe ser razonable (de 3 a 6 meses).
- Lote: el número de lote debe coincidir con algún número de lote de los que se encuentran en el pdf.
- Ha: La cantidad de hectáreas cosechadas debe ser coherente respecto a la cantidad de hectáreas sembradas. Además, la cantidad de hectáreas sembradas debe ser menor o igual a la cantidad de hectáreas que el lote posee (dato extraído del pdf).

Luego de realizado este control sobre los datos se constataron varios errores. Inicialmente se contaba con datos de siembras y cosechas sobre 11 chacras que

suman un total de 5175,34 hectáreas, de las cuales 4316,82 hectáreas son útiles para ser sembradas. Luego de realizadas algunas validaciones sobre la calidad de los datos y de intentar relacionar los datos de siembras y cosechas con la información geográfica, nos quedamos con datos de siembras y cosechas relativos a 7 chacras. La superficie total de estas 7 chacras es de 2553,72 hectáreas, mientras la superficie útil para sembrar de estas chacras es de 1944,11 hectáreas. El principal problema encontrado respecto a la calidad de los datos fue relacionar la información de las siembras y las cosechas recibidas en archivos excel con la información geográfica. En muchos casos, el identificador del lote en la información de siembras y cosechas (archivos excel) no coincide con el identificador del lote en los archivos que proveen información geográfica (shapefiles y pdfs). Para algunos casos se pudo relacionar esta información en base a la superficie sembrada y la superficie total de cada lote. En otros casos no fue posible obtener esta relación e incluso los datos relativos a la cantidad de hectáreas sembradas no eran coherentes con los tamaños de los lotes. Lamentablemente, con el fin de mantener datos fiables se decidió descartar estas chacras.

4.4. Caracterización de las variedades de soja

Debido a que el simulador requiere como parámetros de entrada las fechas de floración y la de madurez, y estos datos no están disponibles, es necesario derivarlos a partir de la fecha de siembra y la variedad de semilla sembrada. Para ello se recurre a estudios realizados tanto por los vendedores de las semillas como por el propio INIA. Los datos proporcionados por los vendedores son de carácter comercial, por lo que en la medida de lo posible, se intenta utilizar la información proveniente de estudios de INIA e INASE. Estos estudios, se realizan en campos experimentales bajo ambientes controlados, siguiendo protocolos que dan confiabilidad a los resultados. De acuerdo a [22]:

“La Evaluación Nacional de Cultivares es realizada bajo la responsabilidad del Instituto Nacional de Semillas (INASE) con el objetivo de proveer información objetiva y confiable sobre el comportamiento de los cultivares de las distintas especies de importancia agrícola a nivel nacional, requisito necesario para la inscripción de los mismos en el Registro Nacional de Cultivares. Al presente, esta información es generada a través de un convenio con el Instituto Nacional de Investigación Agropecuaria (INIA). La evaluación se realiza siguiendo Protocolos elaborados por un comité técnico de trabajo multidiscipli-

nario e interinstitucional (INASE-INIA), siendo sometidos a consideración del Grupo de Trabajo Técnico en Evaluación (GTTE) correspondiente, en el que están representados los diversos sectores especializados. Estos Protocolos son revisados y actualizados periódicamente para responder a cambios en las necesidades de técnicos y productores que reflejan la dinámica en las tecnologías de producción agrícola del Uruguay.”

De acuerdo a las fuentes consultadas, se obtiene una tabla de días a emergencia, floración y maduración para cada variedad (Tabla 4.1). Además, dado que estos datos no sólo dependen de la variedad de la semilla, sino también de la época de siembra, se establecen tres variantes de estos datos para cada variedad, agregando días si se trata de una siembra temprana (antes de 5 de noviembre), o restando si es una siembra tardía (después de 20 de diciembre).

Variedad	Emergencia.0	Floracion.0	Maduracion.0	Emergencia.1	Floracion.1	Maduracion.1	Emergencia.2	Floracion.2	Maduracion.2
SOJA ADP 6126 (SEMILLA)	8	86	168	7	76	158	6	66	148
SOJA ADP 4910 (SEMILLA)	8	64	142	7	54	132	6	44	122
SOJA ADP DON MARIO 6202	8	98	178	7	88	168	6	78	158
SOJA ADP DON MARIO 7.0i	8	84	170	7	74	160	6	64	150
SOJA BIO-SOJA 6.5 (BASICA)	8	78	169	7	68	159	6	58	149
SOJA BIO 6.5 (C1)	8	78	169	7	68	159	6	58	149
SOJA ADP 5909	8	89	158	7	79	148	6	69	138
SOJA ADP 5009	8	66	165	7	56	155	6	46	145
SOJA NIDERA 5009 C1	8	66	165	7	56	155	6	46	145
SOJA NIDERA 5909 F	8	89	158	7	79	148	6	69	138
SOJA TMG 7161 RR	8	68	132	7	58	122	6	48	112
SOJA ADP CARDINAL 590	8	87	166	7	77	156	6	67	146
SOJA ADP D.MARIO 6.8i	8	88	178	7	78	168	6	68	158

Tabla 4.1: Fechas de emergencia, floración y maduración de variedades de soja utilizadas en este estudio.

4.5. Obtención de series de NDVI a partir del cruzamiento de datos y las imágenes satelitales

El producto objetivo de esta etapa de la solución es una tabla (persistida en un archivo .csv) en la que cada fila contiene:

- Chacra.
- Lote.

Chacra	Lote	Fecha	Celda	NDVI
ALVAREZ OLANO	37	2009-12-11	2	0,4279
ALVAREZ OLANO	37	2009-12-19	2	0,5595
ALVAREZ OLANO	37	2009-12-27	2	0,6082
ALVAREZ OLANO	37	2010-01-01	2	0,5700
ALVAREZ OLANO	37	2010-01-09	2	0,6082
ALVAREZ OLANO	37	2010-01-17	2	0,5654
ALVAREZ OLANO	37	2010-01-25	2	0,7225
ALVAREZ OLANO	37	2010-02-02	2	0,7905
ALVAREZ OLANO	37	2010-02-10	2	0,7568
ALVAREZ OLANO	37	2010-02-18	2	0,8278
ALVAREZ OLANO	37	2010-02-26	2	0,8310
ALVAREZ OLANO	37	2010-03-06	2	0,8339
ALVAREZ OLANO	37	2010-03-14	2	0,7951
ALVAREZ OLANO	37	2010-03-22	2	0,8149
ALVAREZ OLANO	37	2010-03-30	2	0,5431
ALVAREZ OLANO	37	2010-04-07	2	0,4889
ALVAREZ OLANO	37	2010-04-15	2	0,3738

Tabla 4.2: Imágenes satelitales procesadas para zafra 2009/2010 en un punto.

- Celda (píxel en el raster).
- Fecha.
- NDVI.

Estos datos son considerados como las observaciones más cercanas a la realidad, y son los utilizados para compararlos con las simulaciones, como se explicará en la Sección 4.7.

La obtención de estos datos consiste en un algoritmo simple que utiliza funciones espaciales para el cruzamiento de los shapefile de las chacras con las imágenes satelitales. El Algoritmo 1 muestra un pseudocódigo de dicho proceso.

En principio, se intentó utilizar ambas fuentes para la obtención de estos datos, ya que podrían complementarse de buena manera, siendo las imágenes MODIS las aportantes de densidad temporal y las Landsat de mayor resolución. Esto no fue posible debido a la diferencia en los valores de NDVI arrojados por ambos, y por la baja resolución temporal de las imágenes Landsat, tal como se puede apreciar en la Figura 4.1 (en una zafra de 4 meses se tendría como máximo 8 imágenes Landsat, en comparación con 16 imágenes MODIS).

Además, como se explicó en la sección 3.3.1, para los períodos utilizados en el estudio, ocurren otros dos fenómenos: el fallo en el sensor en Landsat 7 que deja solo el 78 % de los píxeles utilizables, como se muestra en la Figura 4.2, y el cambio de satélites entre Landsat 7 y Landsat 8, agregando otro factor de error.

Algorithm 1 Obtención de series de NDVI a partir de archivos shapefile de las chacras y de imágenes satelitales.

- 1: Cargar shapefiles de chacras de interés.
 - 2: Cargar listado de imágenes disponibles.
 - 3: **for** cada imagen y chacras **do**
 - 4: Comparar proyecciones de chacra e imagen, si son distintos reproyectar chacra.
 - 5: **for** cada lote **do**
 - 6: Crop⁹: recorte de imagen al tamaño de shapefile.
 - 7: Mask¹⁰ descartar pixeles del raster que no estén incluidos en el polígono del lote.
 - 8: Extract¹¹ extraer valores de la imagen resultante, identificando las celdas por número y filtrando aquellas que estén en al menos un 75 % incluidas dentro del polígono.
 - 9: Aregar resultado al archivo csv.
 - 10: **end for**
 - 11: **end for**
-



Figura 4.1: Comparación de evolución NDVI extraído de imágenes Landsat vs Modis en un punto central del lote 37 de chacra Alvarez Olano. Se puede apreciar la diferencia de resolución temporal, dado que existen 6 datos Landsat y 21 datos MODIS.

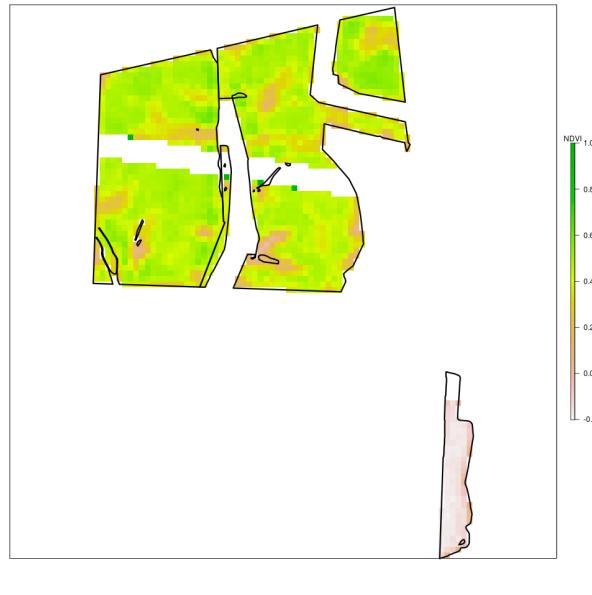


Figura 4.2: Ejemplo de imagen NDVI Landsat recortada con shapefile de una chacra, dejando solo los píxeles que tienen su centro en el interior de la misma. Se puede apreciar la falta de píxeles en el centro de la misma.

Existen algunas técnicas que podrían solucionar este problema, por ejemplo la que se plantea en [23], fusionando los datos de ambas fuentes y logrando muy buenos resultados. Sin embargo, debido al tiempo necesario para aplicar estas técnicas y a que la resolución de este problema no fue considerado como un objetivo de este estudio, se toma la decisión de utilizar solamente las series extraídas de las imágenes MODIS. Esta decisión implica que se tiene una menor resolución en las imágenes satelitales, y que debido a que el tamaño de las celdas es relativamente grande, los datos en los límites de las chacras sean descartados o de mala calidad. Un ejemplo de esto puede verse en la Figura 4.3.

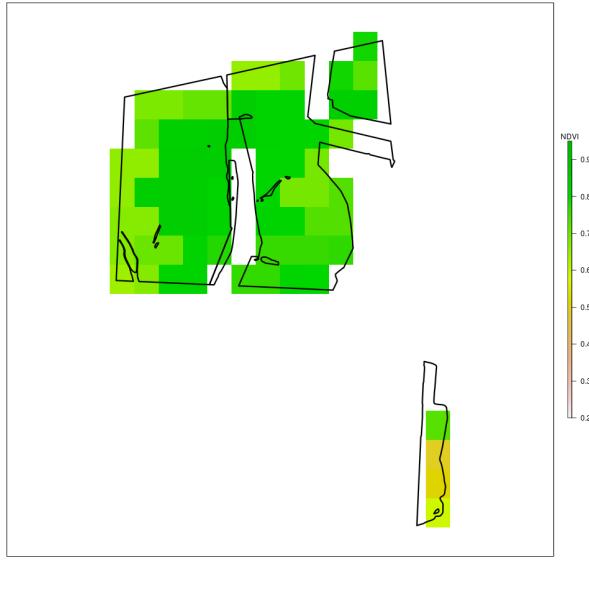


Figura 4.3: Ejemplo de imagen NDVI MODIS recortada con shapefile de una chacra, dejando solo los píxeles que tienen su centro en el interior de la misma.

4.6. Creación de estimadores *baseline*

A la hora de determinar qué tan buena es una estimación, no sólo alcanza con cuantificar el error proveniente de la solución generada, sino que es necesario contar con algún otro estimador (no muy complejo), a modo de poder comparar los errores obtenidos en las soluciones de ambos estimadores y tener un contexto que permita cualificar los errores. A partir de aquí, a este estimador le llamaremos: estimador *baseline*. Con este fin, para cada celda de cada lote, se crearon dos estimadores *baseline* utilizando distintas metodologías y, dado que nos interesa comparar el error de la estimación que surge del simulador de crecimiento de cultivos con la mejor estimación posible, se seleccionó el mejor de ellos como estimador *baseline* de esa celda. La primer metodología consiste en la construcción de un estimador *baseline* de la forma de un polinomio de tercer grado, mientras la segunda, consiste en la construcción de un estimador *baseline* de la forma de una función de doble logística. A continuación, se explica en qué consiste cada metodología de construcción de estimadores y se muestran ejemplos. La primer metodología, consiste en la construcción de un polinomio de tercer grado de la forma $NDVI(t) = at^3 + bt^2 + ct + d$, que tiene como valor de retorno el NDVI y donde la variable t indica la cantidad de días transcurridos desde la siembra. O sea, la función *baseline* permite conocer la

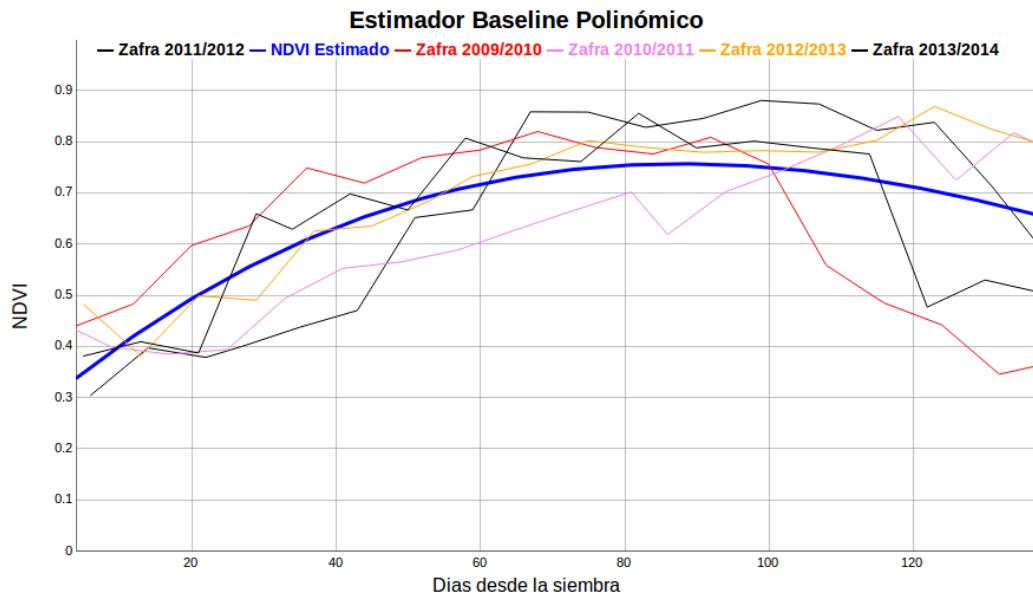


Figura 4.4: Evolución del NDVI del estimador *baseline* polinómico para una celda del lote 3 de la chacra “CUFRE” y evolución del NDVI de las zafras a partir de las que el estimador fue construido.

evolución del NDVI en el tiempo desde la siembra hasta la cosecha para una celda de un lote en particular. La construcción del estimador, se basa en los datos históricos del NDVI obtenidos de las imágenes satelitales. Con el objetivo de determinar los valores de los coeficientes a, b, c y d de la función *baseline* $NDVI(t)$; a partir de las series históricas de NDVI sobre las que se poseen datos, construimos una nueva serie cuya curva minimiza el error respecto a las curvas de las series históricas. O sea, la curva resultante podría verse como un “promedio” de las curvas correspondientes a las series históricas de NDVI. Para ello utilizamos la función `lm`¹² de R. Dicha función, se utiliza para ajustar modelos lineales. Puede utilizarse para realizar análisis de regresión, de estrato único de varianza y análisis de covarianza. En las gráficas de las Figuras 4.4 y 4.5 se muestran ejemplos de cómo queda este estimador para un par de celdas.

La segunda metodología de construcción del estimador *baseline*, se basa en [24] donde se propone una función de doble logística para modelar cómo varía el NDVI a lo largo del tiempo. La función propuesta se presenta en la Ecuación (4.1).

¹²<https://stat.ethz.ch/R-manual/R-devel/library/stats/html/lm.html>

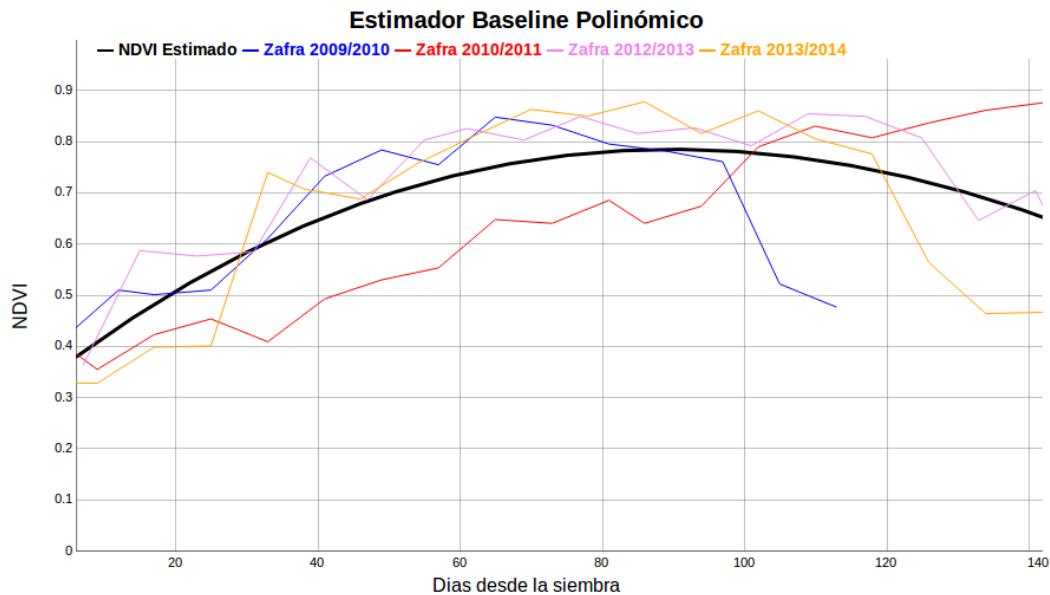


Figura 4.5: Evolución del NDVI del estimador *baseline* polinómico para una celda del lote 34 de la chacra “ALVAREZ OLANO” y evolución del NDVI de las zafras a partir de las que el estimador fue construido.

$$NDVI(t) = wNDVI + (mNDVI - wNDVI) \times \left(\frac{1}{1 + \exp(-mS \times (t - S))} + \frac{1}{1 + \exp(mA \times (t - A))} - 1 \right) \quad (4.1)$$

donde:

- $wNDVI$: corresponde al NDVI durante el invierno (NDVI mínimo).
- $mNDVI$: corresponde al NDVI máximo.
- S : es el punto de inflexión donde la curva crece.
- A : punto de inflexión donde la curva decrece.
- mS : es la tasa de incremento en el punto de inflexión donde la curva crece.
- mA : es la tasa de decrecimiento en el punto de inflexión donde la curva decrece.

En la Figura 4.6 se muestra graficamente como se modela la evolución del NDVI a lo largo de una zafra, mostrando el significado de cada uno de los parámetros de la ecuación 4.1. En este caso, para obtener los valores de los

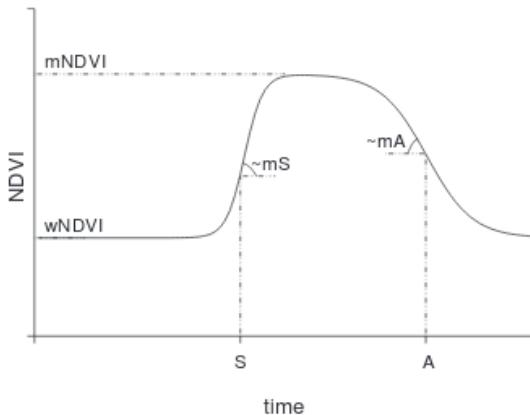


Figura 4.6: Función de doble logística que modela la evolución del NDVI a lo largo de una zafra

parámetros se utiliza la función *nls*¹³ de R. Esta función determina las estimaciones de los mínimos cuadrados no lineales (ponderadas) de los parámetros de un modelo no lineal. Además del modelo y la serie que se quiere ajustar, la función *nls* requiere como parámetro un vector semilla con valores iniciales de los parámetros. Para la construcción de este estimador *baseline*, utilizando la función *nls* de R, se ajusta el modelo para cada una de las series históricas sobre las que se poseen datos, obteniendo como resultado valores de los parámetros antes mencionados (*wNDVI*, *mNDVI*, *S*, *A*, *mS*, *mA*) para cada zafra de la celda en estudio. Luego de obtenidos los valores de los parámetros para todas las zafras históricas, la función *baseline* se obtiene calculando el promedio de los valores de cada uno de los parámetros.

Con mayor formalidad, supongamos que se poseen N zafras históricas, para cada zafra $i : 1..N$ sean:

- $wNDVI^i$: el valor del parámetro *wNDVI* obtenido para la serie histórica correspondiente a la zafra i .
- $mNDVI^i$: el valor del parámetro *mNDVI* obtenido para la serie histórica correspondiente a la zafra i .
- S^i : el valor del parámetro *S* obtenido para la serie histórica correspondiente a la zafra i .
- A^i : el valor del parámetro *A* obtenido para la serie histórica correspondiente a la zafra i .

¹³<https://stat.ethz.ch/R-manual/R-devel/library/stats/html/nls.html>

- mS^i : el valor del parámetro mS obtenido para la serie histórica correspondiente a la zafra i .
- mA^i : el valor del parámetro mA obtenido para la serie histórica correspondiente a la zafra i .

y:

- $wNDVI^r$: Promedio($wNDVI^i$) $i \in 1..N$
- $mNDVI^r$: Promedio($mNDVI^i$) $i \in 1..N$
- S^r : Promedio(S^i) $i \in 1..N$
- A^r : Promedio(A^i) $i \in 1..N$
- mS^r : Promedio(mS^i) $i \in 1..N$
- mA^r : Promedio(mA^i) $i \in 1..N$

entonces el estimador *baseline* de doble logística tiene la siguiente fórmula:

$$NDVI(t) = wNDVI^r + (mNDVI^r - wNDVI^r) \times \left(\frac{1}{1 + \exp(-mS^r \times (t - S^r))} + \frac{1}{1 + \exp(mA^r \times (t - A^r))} - 1 \right) \quad (4.2)$$

En las gráficas de las Figuras 4.7 y 4.8 se muestra cómo queda este estimador para un par de celdas.

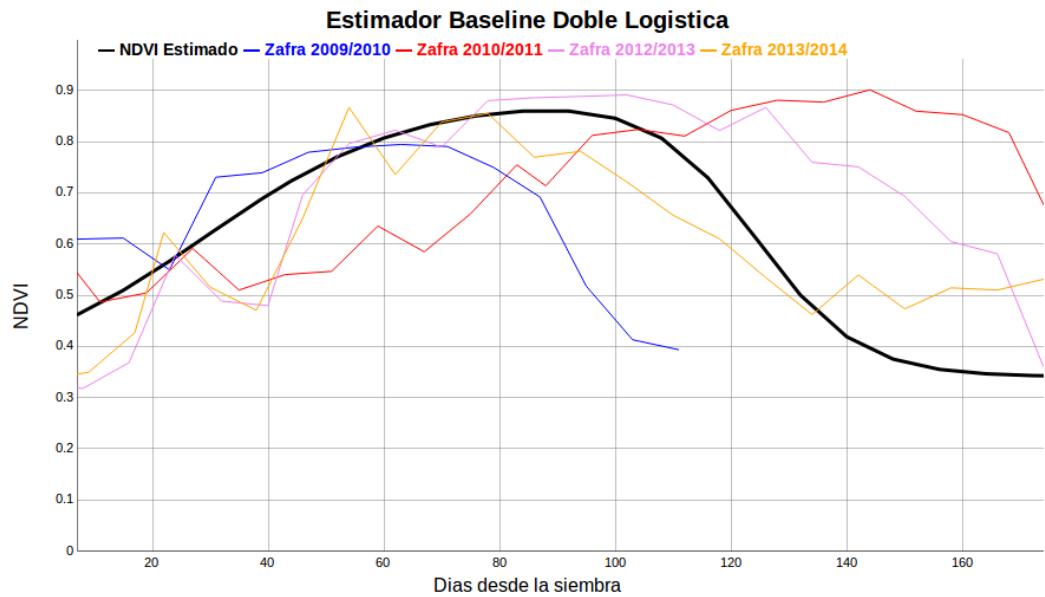


Figura 4.7: Evolución del NDVI del estimador *baseline* de doble logística para una celda del lote 1 de la chacra “POCHELLU” y evolución del NDVI de las zafras a partir de las que el estimador fue construido.

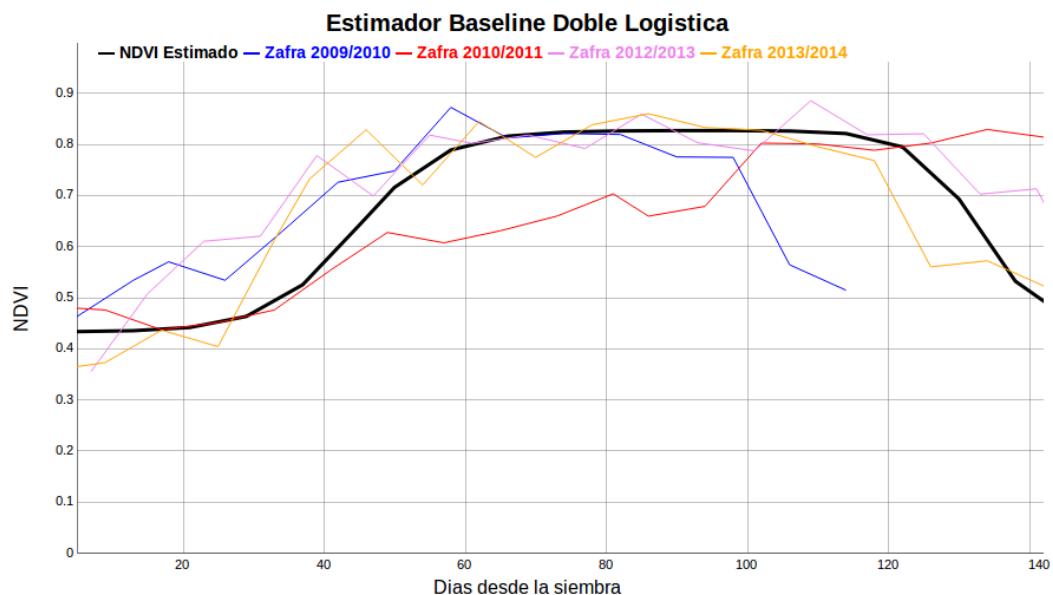


Figura 4.8: Evolución del NDVI del estimador *baseline* de doble logística para una celda del lote 37 de la chacra “ALVAREZ OLANO” y evolución del NDVI de las zafras a partir de las que el estimador fue construido.

4.7. Calibración del simulador de crecimiento de cultivos

Luego de contar tanto con las entradas necesarias para la ejecución del simulador, como con los datos suficientes para comparar los resultados con el objetivo, se procede a la calibración del mismo. La calibración del modelo se realiza para la unidad más pequeña de datos, es decir, para cada píxel de las imágenes satelitales correspondientes a un cuadrado de 250 metros de lado. El procedimiento consiste en utilizar la información de las primeras $N - 1$ zafras para calibración, dejando la restante para validación. Por ejemplo, para los puntos del lote 37 de la chacra Alvarez Olano, se utilizan las zafras 2009/2010, 2010/2011, 2012/2013, 2013/2014 para la calibración y la 2014/2015 para validación. En cuanto a los parámetros del modelo a optimizar, se tomó la decisión a partir de lo sugerido por el Ing. Agr. Andrés Berger, teniendo en cuenta la sensibilidad del modelo a dichos parámetros, y descartando aquellos que no debían ser optimizados ya que no están relacionados a la calidad del suelo. Dichos parámetros son:

- capacidad de campo (field capacity) y punto de marchitez permanente (permanent wilting point).
- profundidad máxima de enraizamiento.

Tanto para éstos parámetros como los restantes se parte de valores iniciales proporcionados por el Ing. Agr. Andrés Berger, que además fueron puestos a prueba, variando los mismos dentro del rango aceptable y simulando todas las zafras en todos los puntos, obteniendo un RMSE general. Con este proceso, se validó que efectivamente dichos valores minimizaran la diferencia entre las simulaciones y las observaciones a nivel general, y que por lo tanto fueran buenos valores iniciales para las optimizaciones.

En la Figura 4.9 se puede ver el análisis realizado al parámetro kc (coeficiente de extinción de luz), para este parámetro se decidió utilizar el valor 0,3. Este valor se deja fijo para todas las simulaciones, ya que no depende de una zafra en particular, sino que es dependiente del cultivo.

También se realizaron pruebas para obtener valores iniciales en los parámetros a optimizar, de manera de que el algoritmo de optimización partiera cerca de valores ideales. Como se puede observar en la imagen Figura 4.10, para el parámetro $rdmax$ se eligió el valor 1,5 como valor inicial.

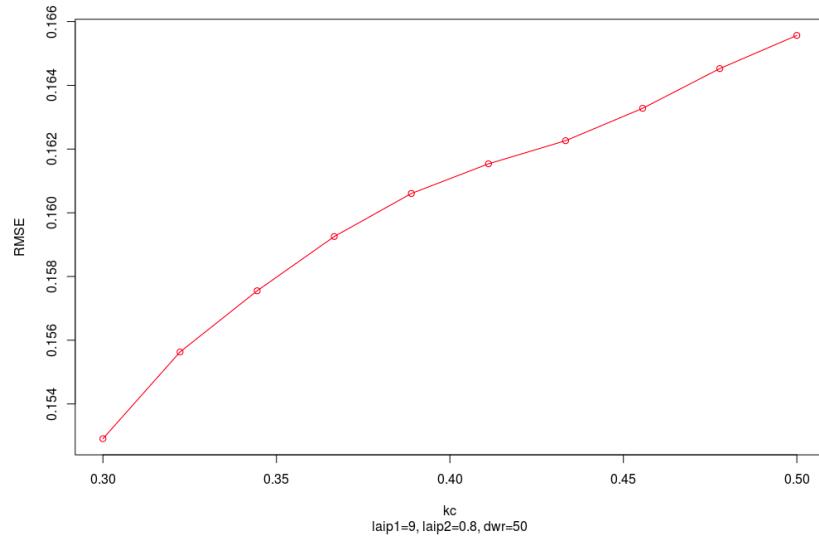


Figura 4.9: Gráfica mostrando el RMSE de las simulaciones vs. observaciones a medida que se cambia el valor del parámetro kc (coeficiente de extinción de luz).

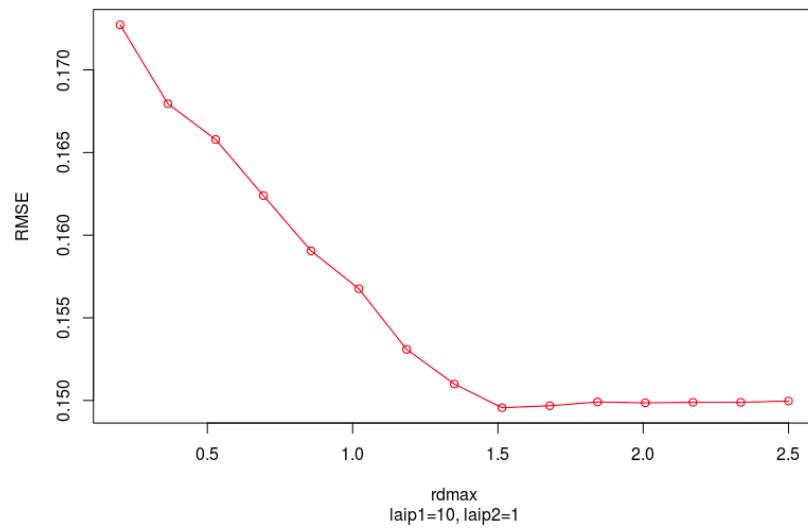


Figura 4.10: Gráfica mostrando la evolución del RMSE de las simulaciones vs. observaciones a medida que se varía el parámetro $rdmax$ (profundidad máxima de enraizamiento).

Dado que los algoritmos de optimización funcionan maximizando o minimizando una función, para la optimización de los parámetros se recurrió a creación de una función auxiliar, que consiste en comparar los resultados de las simulaciones con las observaciones, utilizando RMSE. Por lo tanto, el objetivo es minimizar la diferencia entre ambas. Vale destacar que dado que la función a minimizar es una función auxiliar que realiza una diferencia entre curvas calculando RMSE, no puede calcularse un gradiente, por lo que los métodos que requieren uno no pueden ser utilizados. Además, se debe utilizar un método que pueda restringir los parámetros a optimizar, ya que estos solo pueden tomar valores dentro de rangos definidos. Para lograr este objetivo, R posee múltiples librerías de optimización, que fueron evaluadas teniendo en cuenta velocidad de convergencia y comparando resultados. Los métodos puestos a prueba fueron:

- *optim*¹⁴ Librería de optimización de propósito general que engloba diversos métodos, por defecto utiliza el método Nelder Meal (1965). Para el caso de optimización con restricciones, utiliza el método L-BFGS-B introducido en Byrd et. al. (1995).
- *nlminb*¹⁵ Librería de optimización con y sin restricciones que utiliza PORT, un conjunto de rutinas desarrolladas por AT&T Bell Telephone Laboratories, Inc.¹⁶
- *spg*¹⁷ Función incluida en la librería BB, permite la optimización con restricciones, y si bien el algoritmo utiliza gradientes, es posible utilizarla sin uno, en cuyo caso es calculado numéricamente.

Para compararlos, se utilizaron tres métricas: cantidad de veces en las que converge, tiempo promedio de ejecución y RMSE promedio de la solución encontrada.

Teniendo en cuenta que se calibra el modelo para cada punto de cada lote, se comparan los resultados de 429 ejecuciones de estos métodos.

Cantidad de convergencias:

optim	155
nlminb	228
spg	429

¹⁴<https://stat.ethz.ch/R-manual/R-devel/library/stats/html/optim.html>

¹⁵<https://stat.ethz.ch/R-manual/R-devel/library/stats/html/nlminb.html>

¹⁶<http://ms.mcmaster.ca/~bolker/misc/port.pdf>

¹⁷<https://cran.r-project.org/web/packages/BB/BB.pdf>

Tiempo promedio de ejecución (segundos):

optim	222
nlinb	125
spg	179

RMSE promedio:

optim	0,1194
nlinb	0,1187
spg	0,1184

Con estos datos, se tomó la decisión de utilizar la función *spg*, ya que supera a las demás en todos los ítems de comparación.

4.8. Simulación del crecimiento de cultivos

Luego de calibrado el modelo de crecimiento de cultivos a partir de los datos de las zafras históricas y los datos climáticos, para cada celda de cada lote, se procede a realizar una simulación para la última zafra sobre la que se poseen datos históricos utilizando dicho modelo. Supongamos que se poseen N zafras con datos históricos, entonces se utilizan los datos de $N - 1$ zafras históricas para calibrar el modelo y posteriormente se realiza una simulación con el modelo calibrado para la zafra N . De esta forma, podemos comparar la serie de NDVI obtenida a partir del simulador de crecimiento de cultivos calibrado con la serie real de NDVI (obtenida de las imágenes satelitales) para la zafra N . Además, recordemos que el estimador *baseline* también se construye utilizando $N - 1$ zafras y se simula la zafra N . Por ende, también se compara la serie obtenida a partir del simulador de crecimiento de cultivos calibrado con el estimador *baseline*.

4.9. Evaluación de la calidad de la estimación con RMSE

Luego de realizada la simulación, resta evaluar que tan buena es la estimación obtenida. La evaluación de la estimación implica, por un lado, calcular el RMSE de la estimación realizada respecto a los datos reales (la serie de NDVI

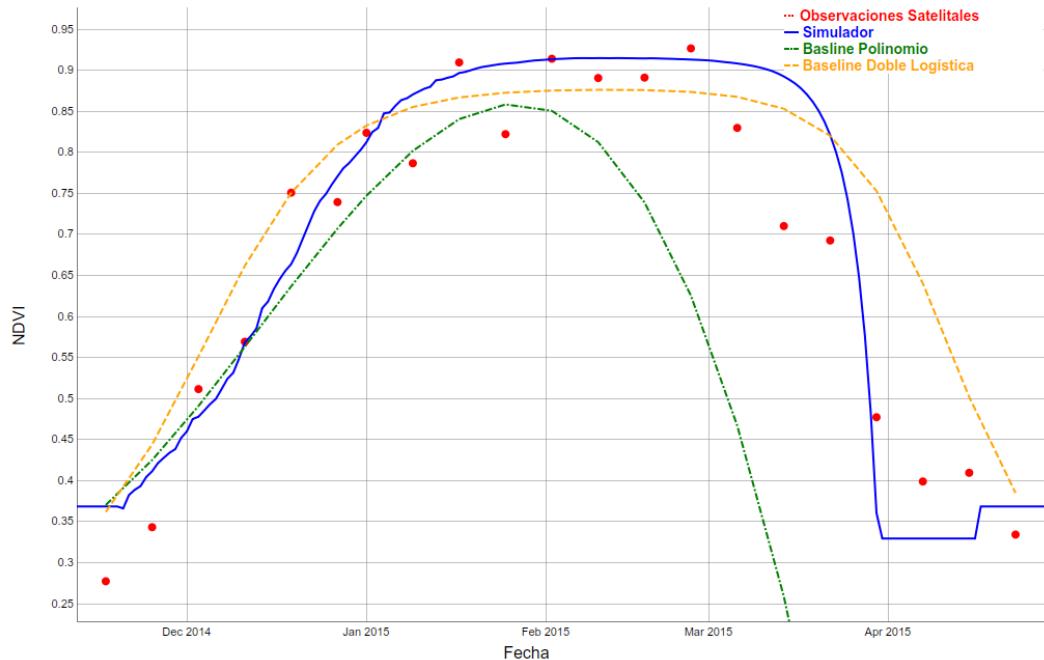


Figura 4.11: Comparación de resultados arrojados por los tres métodos con las observaciones satelitales para la celda 17 del lote 37 de la chacra “ALVAREZ OLANO”, en la zafra 14/15. Los puntos rojos son los valores NDVI a lo largo de la zafra, mientras que la línea azul continua representa el resultado del simulador, la anaranjada a rayas el resultado del estimador doble logística, y la verde a rayas y puntos el del estimador polinómico.

obtenida de las imágenes satelitales MODIS para la zafra en estudio), por otro lado, calcular el RMSE de la estimación del algoritmo *baseline* respecto a los datos reales (la serie de NDVI obtenida de las imágenes satelitales MODIS para la zafra en estudio). De esta forma, se cuantifica el error de la estimación respecto a los datos reales y también se compara este error con el error proveniente de aplicar el algoritmo *baseline*. Como ejemplo, si se toma la celda 17 del lote 37 de la chacra “ALVAREZ OLANO”, se obtiene un RMSE de 0,0678 para el simulador, un RMSE de 0,0785 para el *baseline* doble logística, y un RMSE de 0,1983 para el *baseline* polinómico. La representación gráfica de estos resultados puede verse en la Figura 4.11.

En el siguiente Capítulo aplicaremos este procedimiento a todas las chacras y evaluaremos los resultados.

Capítulo 5

Pruebas

5.1. Introducción

En este capítulo, se describen las pruebas realizadas y sus resultados, incluyendo detalles sobre la implementación y sobre los datos utilizados.

5.2. Chacras Utilizadas

Previo a comentar sobre las simulaciones realizadas y sus resultados, mostraremos información de interés sobre las chacras utilizadas en la simulación. La descripción de los datos utilizados se hace en el Capítulo 3, aquí se muestran datos generales sobre las chacras. Por un detalle de cada chacra, ver el Apéndice 1.

Las chacras utilizadas están situadas en los departamentos de Colonia, San José y Soriano, tal como se puede apreciar en la Figura 5.1. Esto se debe principalmente al hecho de que los datos fueron proporcionados por el Ing. Agr. Andrés Berger, quien trabaja en la Estación Experimental del INIA “La Estanzuela”, ubicada en el departamento de Colonia. Estas 7 chacras, están divididas en 38 lotes, totalizando 2553,72 hectáreas, de las cuales se contabilizan 1945,11 hectáreas útiles (Ver Tabla 5.1).

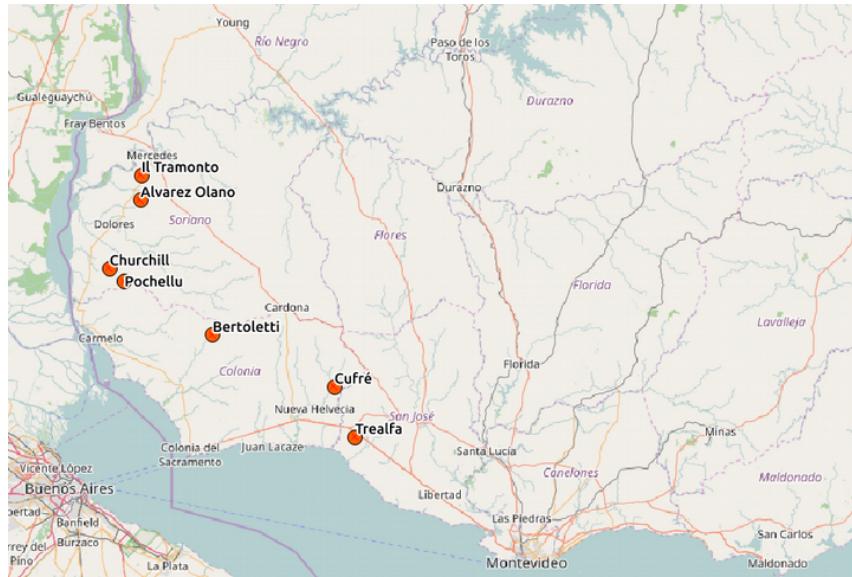


Figura 5.1: Ubicación geográfica de las chacras utilizadas en el proyecto

Nombre	Departamento	Lotes	Superficie total (Ha)	Superficie útil (Ha)	Celdas
ALVAREZ OLANO	Soriano	5	192,79	177,95	50
BERTOLETTI	Colonia	15	1122,48	708,77	106
CHURCHILL	Soriano	1	155,72	149,84	49
CUFRÉ	Colonia	5	239,25	233,20	25
IL TRAMONTO	Soriano	1	386,18	266,33	82
POCHELLU	Colonia/Soriano	3	232,22	202,07	62
TREALFA	San José	8	225,08	206,95	52
Total		38	2553,72	1945,11	426

Tabla 5.1: Nombre, Departamento, superficie y cantidad de celdas de las chacras utilizadas en el proyecto.

5.3. Consideraciones para realizar la simulación

Para el período de estudio (2009-2015) se utilizan 310 imágenes MODIS, totalizando 4,1 G. Estas imágenes sirven para generar los datos con los que se comparan las soluciones. Luego de procesadas se cuenta con 137.330 observaciones, siendo una observación el valor del NDVI en un día de una celda en una chacra (explicado en el Capítulo 4). Considerando que el simulador de crecimiento de cultivos utilizado realiza iteraciones diarias, que se calibra para cada una de las 426 celdas disponibles (ver Tabla 5.1), que la zafra promedio tiene una duración de 161 días, y que se tiene un promedio de 5 zafras por celda, se estima un total de 342.930 corridas del simulador. Además, al estimar un promedio de 20 corridas necesarias para alcanzar la convergencia del optimizador, se llega a un estimado de 6.858.600 de ejecuciones.

Teniendo en cuenta estas estimaciones, y previendo que el mismo código puede ser reutilizado posteriormente aún con un mayor volumen de datos, se tomaron algunos recaudos a la hora de codificar. Con el objetivo de utilizar la mayor cantidad de recursos de hardware disponibles, se emplearon las librerías de R `doParallel`¹ y `foreach`² que permiten explotar la paralelización provista por el procesador del hardware utilizado.

El código paralelizado consiste en el cálculo del estimador *baseline*, la calibración del simulador de crecimiento de cultivos y la ejecución del simulador de crecimiento de cultivos calibrado para la última zafra en estudio. La paralelización se realiza para todos los píxeles dentro de un lote. Las librerías `parallel` y `foreach` son utilizadas en conjunto, la librería `parallel` proporciona un contexto a través del cual mediante la librería `foreach`, se pueden ejecutar loops en paralelo. A nivel de código, primero incluimos las librerías utilizadas de la siguiente manera:

```
library(foreach)
library(doParallel)
```

La inicialización del contexto multihilo en R se realiza de la siguiente forma:

```
cl<-makeCluster(NUMERO_CORES_A_UTILIZAR)
registerDoParallel(cl)
```

¹<https://cran.r-project.org/web/packages/doParallel/doParallel.pdf>

²<https://cran.r-project.org/web/packages/foreach/foreach.pdf>

Donde NUMERO_CORES_A_UTILIZAR debe ser sustituido por la cantidad de cores con los que se quiere paralelizar (depende de los recursos de hardware disponibles).

Posteriormente para ejecutar un loop en varios hilos utilizamos la librería foreach de la siguiente forma:

```
foreach (i=1:cantidadPixel .combine='rbind',  
.packages=c("Metrics", "doParallel"), .inorder=FALSE) %dopar % {  
    // Código a ejecutar en paralelo  
}
```

Finalmente, se debe ejecutar la siguiente sentencia:

```
stopCluster(c1)
```

Para la ejecución de las tareas paralelizadas antes mencionadas, se utilizó el siguiente hardware:

- Procesador: 3rd Generation Intel® CoreTM i73630QM (6MB Cache, up to 3.4GHz w/ Turbo Boost 2.0)
- Memoria RAM: 8GB (2 X 4GB) Dual Channel DDR3 at 1600MHz

Codificando de esta forma y utilizando el hardware mencionado, el tiempo de ejecución de estas tareas para todos los píxeles de todas las chacras estudiadas, paralelizadas en 6 hilos, fue de aproximadamente 327 minutos (5 horas y 27 minutos).

5.4. Estadísticas globales sobre los resultados

A continuación, se muestran los resultados obtenidos a nivel global, sin discriminar los resultados de cada lote ni de cada chacra. En el Apéndice 2 pueden verse los detalles por chacra, lote y celda.

Se analiza el promedio de error (RMSE), para todas las celdas (píxeles del satélite MODIS) estudiadas, contrastando el error de los algoritmos *baseline* (descriptos en Capítulo 4) contra el error del simulador de crecimiento de cultivos. Con los objetivos de simplificar la interpretación de los resultados, y de ser exigentes con los resultados, se utiliza el “mejor” algoritmo en cada caso (el que arroje menor RMSE) para comparar, llamándolo simplemente *baseline*. También se muestra la desviación estándar de cada método de estimación de NDVI.

	Prom. Error	Desv. Estándar
Simulador	0,0912	0,0203
Baseline	0,0980	0,0409

Tabla 5.2: Promedio de error y desviación estándar de todos los puntos.

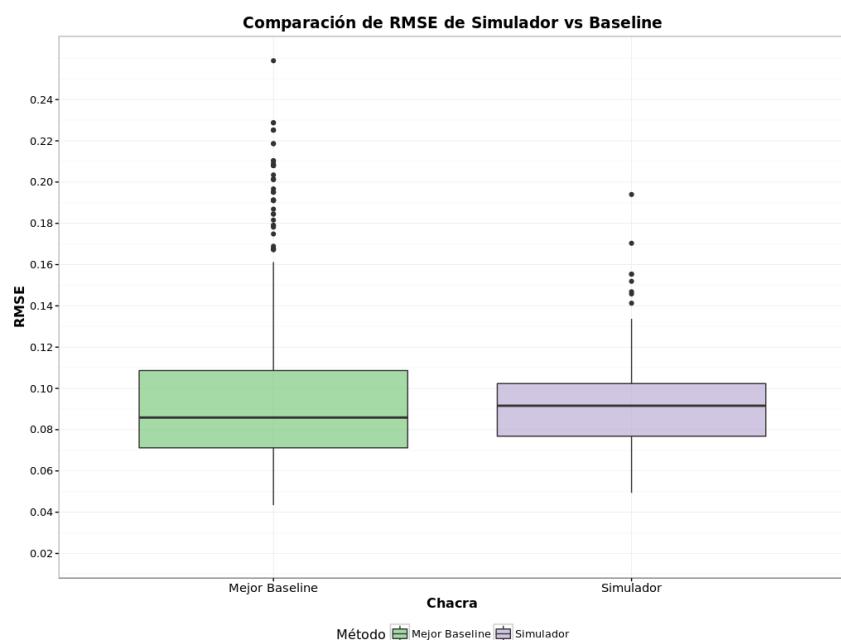


Figura 5.2: Comparación de resultados simulador vs *baseline* en todos los puntos.

Como se puede observar en la Tabla 5.2 y la Figura 5.2, tanto los errores del algoritmo *baseline* como del simulador de crecimiento de cultivos se encuentran por debajo del 10 %, siendo el error del simulador de crecimiento de cultivos un poco inferior al error del algoritmo *baseline*. Por otro lado, la desviación estándar del simulador de crecimiento de cultivos es muy inferior (casi la mitad) al estimador *baseline*. Esto es un punto a favor del simulador de crecimiento de cultivos, considerando que, implica que en general, el simulador de crecimiento de cultivos está más cerca del promedio de la serie real de NDVI que el estimador *baseline*.

Como se comentó en el Capítulo 4, en este estudio se consideraron aquellas celdas (píxeles del satélite MODIS) tales que más del 75 % de su superficie se encuentre dentro de algún lote de las chacras estudiadas. El hecho de que no toda la celda se encuentre dentro del lote puede ser una fuente de error, dado que la superficie de la celda que no se encuentra totalmente incluído en el lote puede alterar el NDVI (en esta superficie podría haber un camino rural, un arroyo, un monte u otro tipo de superficie que afecte de alguna forma el NDVI de esa celda). Por ello es de interés conocer el resultado entre los píxeles que se encuentran totalmente contenidos dentro de un lote y los que no. Para esto, desglosamos el resultado considerando solo aquellas celdas cuya superficie se encuentra incluida en un 95 % o más dentro del lote (268 píxeles, 64 % del total). Los resultados que se pueden apreciar en la Tabla 5.3 y la Figura 5.3, confirman que al limitar el estudio a puntos menos influenciados por posibles fuentes de error, se da una dinámica inversa en los estimadores. El promedio de error del simulador de crecimiento de cultivos decrece levemente, mientras el error del estimador *baseline* crece, superando el 10 %. Lo mismo ocurre con la desviación estándar, disminuye para el simulador de crecimiento de cultivos y aumenta para el estimador *baseline*.

	Prom. Error	Desv. Estándar
Simulador	0,0905	0,0207
Baseline	0,1010	0,0440

Tabla 5.3: Promedio de error y desviación estándar en celdas contenidas en un 95 % o más en las chacras.

Finalmente, desglosamos los resultados observando sólo aquellos puntos 100 % incluídos dentro de algún lote de las chacras estudiadas (185 píxeles, 44 % del total), confirmando la tendencia observada para aquellos píxeles con

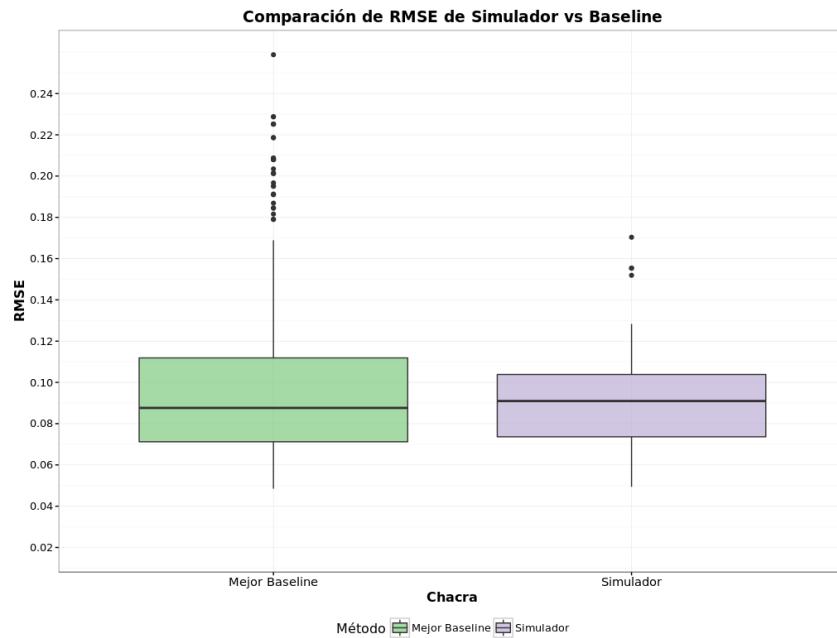


Figura 5.3: Comparación de resultados simulador vs *baseline* en celdas contenidas en un 95 % o más en las chacras.

más de un 95 % de la superficie incluida dentro de un lote. Como se puede apreciar en la Tabla 5.4 y la Figura 5.4, el error del simulador de crecimiento de cultivos nuevamente decrece, ubicándose por debajo del 9 %, mientras el error del estimador *baseline* aumenta nuevamente. Sin embargo, la desviación estándar aumenta para los dos.

	Prom. Error	Desv. Estándar
Simulador	0,0889	0,0198
Baseline	0,1030	0,0462

Tabla 5.4: Promedio de error y desviación estándar en puntos contenidos en un 100 % en las chacras.

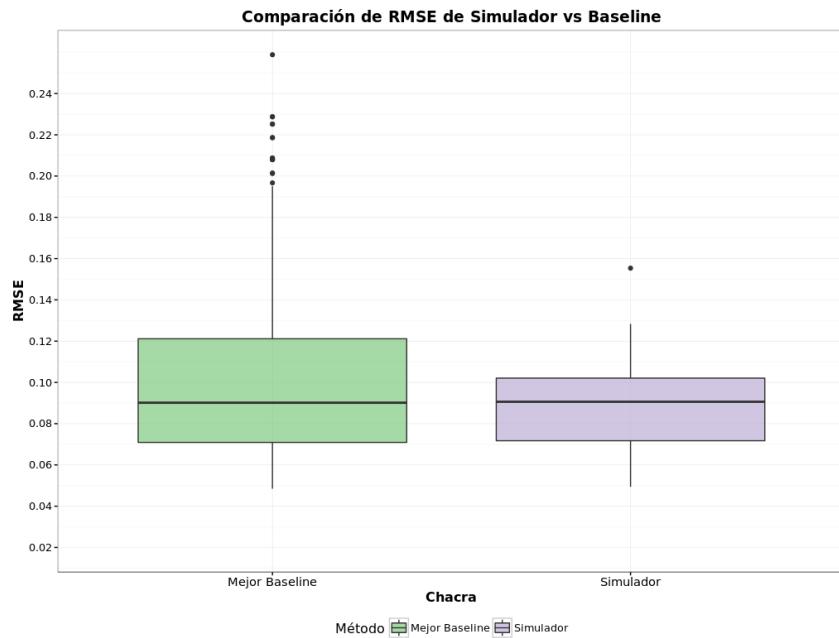


Figura 5.4: Comparación de resultados simulador vs *baseline* en puntos contenidos en un 100 % en las chacras.

5.5. Estadísticas por chacra sobre los resultados

Si bien analizando la desviación estándar de los resultados generales puede deducirse que el error no tiene grandes variaciones, resulta interesante analizar el comportamiento de estos indicadores diferenciando por chacra. En la Figura 5.5 y Tabla 5.5 se pueden apreciar las diferencias entre chacras, donde por ejemplo vemos que la chacra “POCHELLU” es la que tiene mejores resultados (RMSE de 0,0722), y la chacra “CHURCHILL” los peores (RMSE de 0,1168).

También es de interés observar las diferencias en los parámetros calibrados entre chacras, dado que éstos dan una referencia sobre la calidad del suelo de las mismas. En la Tabla 5.6 se muestra el promedio de cada parámetro calibrado para las chacras, mientras que en la Figura 5.6 se puede apreciar gráficamente la diferencia entre chacras de los valores calibrados en el parámetro Profundidad máxima de enraizamiento (rdmax). Los valores detallados de la calibración de los parámetros pueden verse en el Apéndice 2.

Chacra	#Celdas mejor Baseline	#Celdas mejor Simulador	Prom. Error Baseline	Prom. Error Simulador
ALVAREZ OLANO	29	21	0,0790	0,0856
BERTOLETTI	92	13	0,0770	0,0924
CHURCHILL	5	44	0,1786	0,1168
CUFRE	14	11	0,0815	0,0839
IL TRAMONTO	23	56	0,1102	0,0937
POCHELLU	28	34	0,0853	0,0722
TREALFA	26	23	0,0864	0,0921
Total	217	202	0,0980	0,0912

Tabla 5.5: Cantidad de celdas en las que simulador y *baseline* son mejores, junto con el promedio de error (RMSE) de cada estimador, discriminado por chacra.

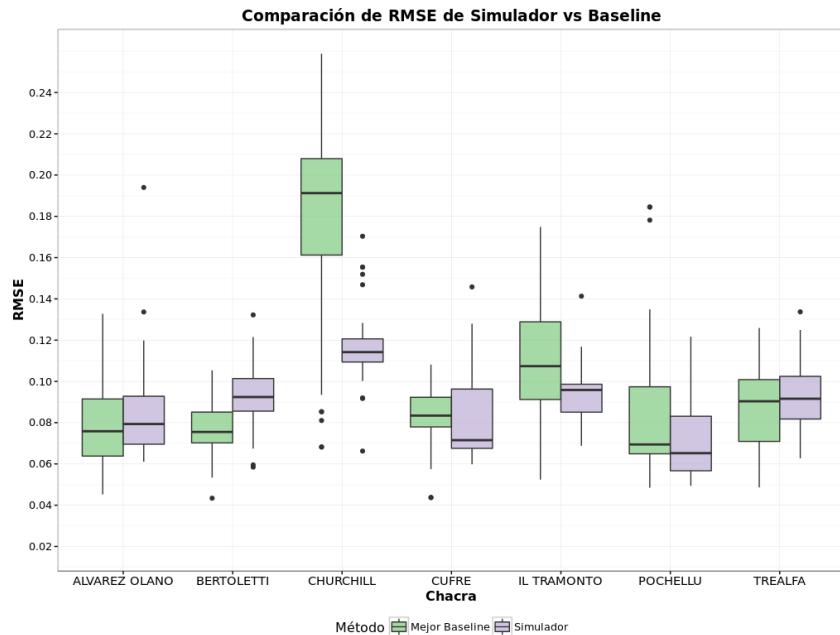


Figura 5.5: Comparación de resultados simulador vs *baseline* discriminado por chacra.

Chacra	Promedio - rdmax	Promedio - fc	Promedio - pwp
ALVAREZ OLANO	1,0912	0,3146	0,1293
BERTOLETTI	0,2150	0,2177	0,0982
CHURCHILL	1,6582	0,3818	0,1509
CUFRE	0,4687	0,2345	0,1036
IL TRAMONTO	0,8859	0,2609	0,1121
POCHELLU	1,6946	0,3837	0,1515
TREALFA	1,0495	0,3344	0,1357
General	0,9465	0,2958	0,1233

Tabla 5.6: Promedio de los parámetros calibrados diferenciados por chacra. Profundidad máxima de enraizamiento (rdmax), capacidad de campo (field capacity - fc), y punto de marchitez permanente (permanent wilting point - pwp)

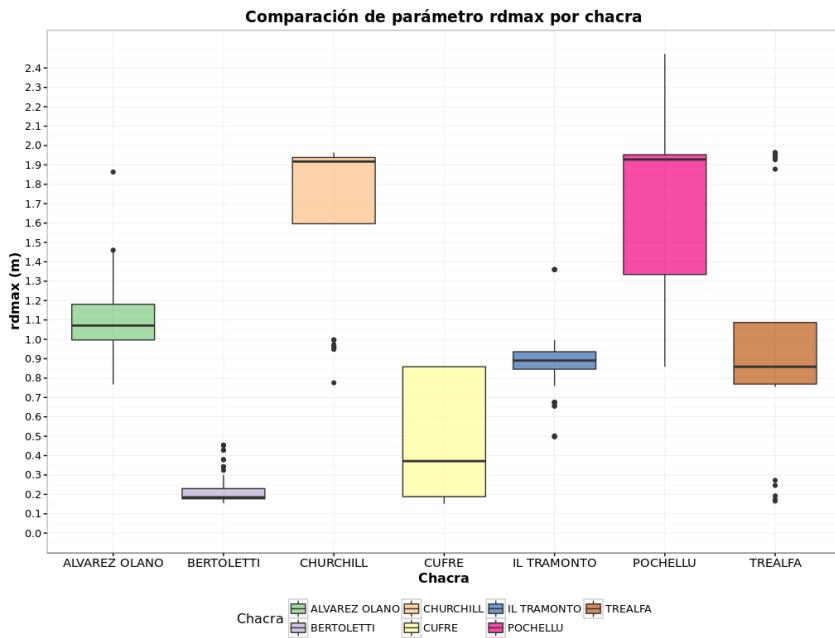


Figura 5.6: Comparación entre chacras de resultados de calibración del parámetro de profundidad máxima de enraizamiento (rdmax).

5.6. Posibles mejoras

Al observar los gráficos de las simulaciones, se puede apreciar que la mayor parte del error proviene de la última parte de la zafra, en el período posterior a la floración hasta la madurez. Las Figuras 5.7 y 5.8 son ejemplos de simulaciones que son muy buenas en el período de crecimiento, pero no lo son al final. En la primera se puede apreciar que el simulador continúa con un NDVI elevado luego de la floración (línea punteada roja) para luego caer abruptamente, mientras que debería tener un decrecimiento más suave. En la segunda, nuevamente se puede apreciar una caída demasiado abrupta.

Considerando que esto es una carencia del simulador, se realizaron pruebas excluyendo los últimos 5 puntos (últimas 5 observaciones satelitales, representando 40 días). Tal como lo evidencian los datos mostrados en la Tabla 5.7 y la Figura 5.9, los resultados mejoran considerablemente. Mientras el error y la desviación estándar del *baseline* casi no sufren alteraciones, el RMSE del simulador disminuye desde 9,1 % hasta 7,7 %, manteniendo prácticamente incambiada la desviación estándar.

Al igual que en la Sección 5.4, se realiza el desglose de los resultados diferenciando aquellas celdas (pixeles del satélite MODIS) que se encuentran

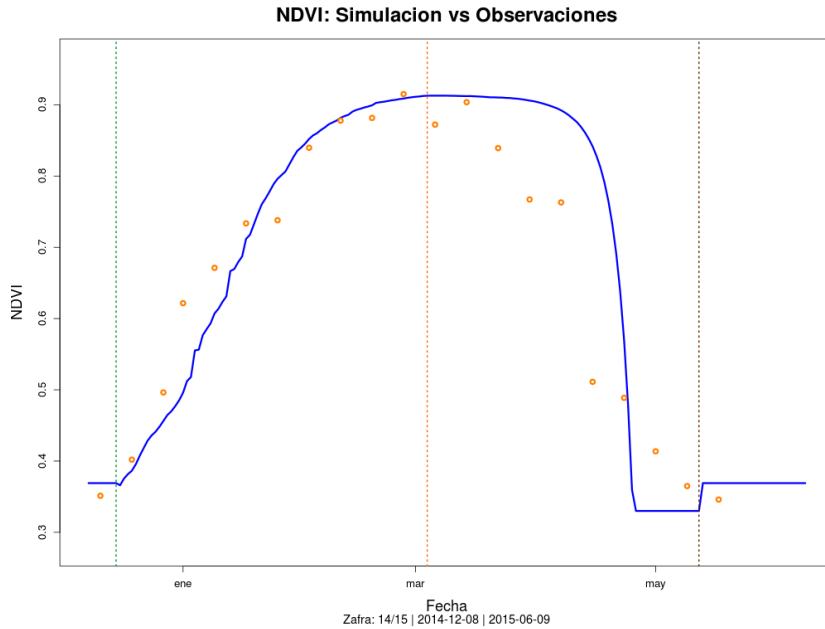


Figura 5.7: Simulación de zafra 14/15 en lote 5 de chacra Trealfa.

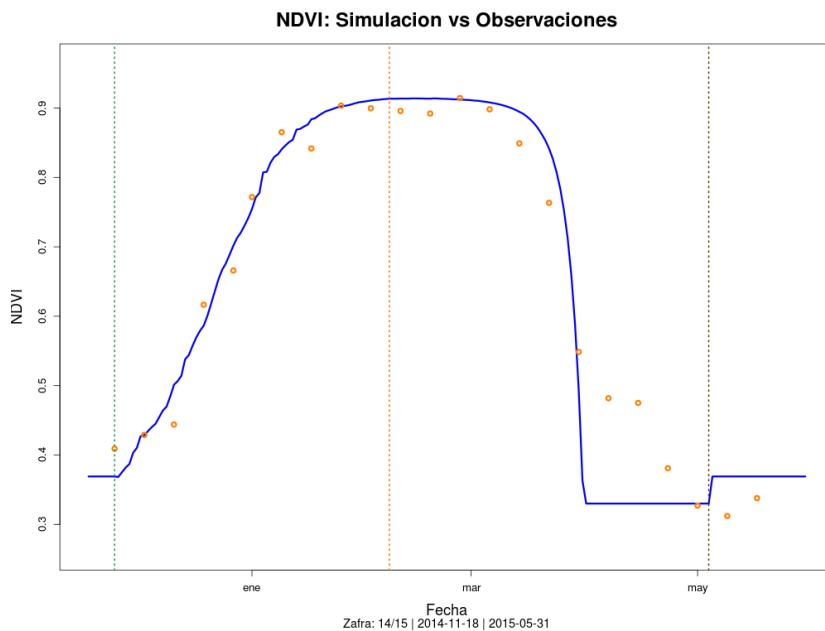


Figura 5.8: Simulación de zafra 14/15 en lote 2 de chacra Pochellu.

	Prom. Error	Desv. Estándar
Simulador	0,0769	0,0227
Baseline	0,0960	0,0418

Tabla 5.7: Promedio de error y desviación estándar de todos los puntos, sin considerar últimas 5 observaciones.

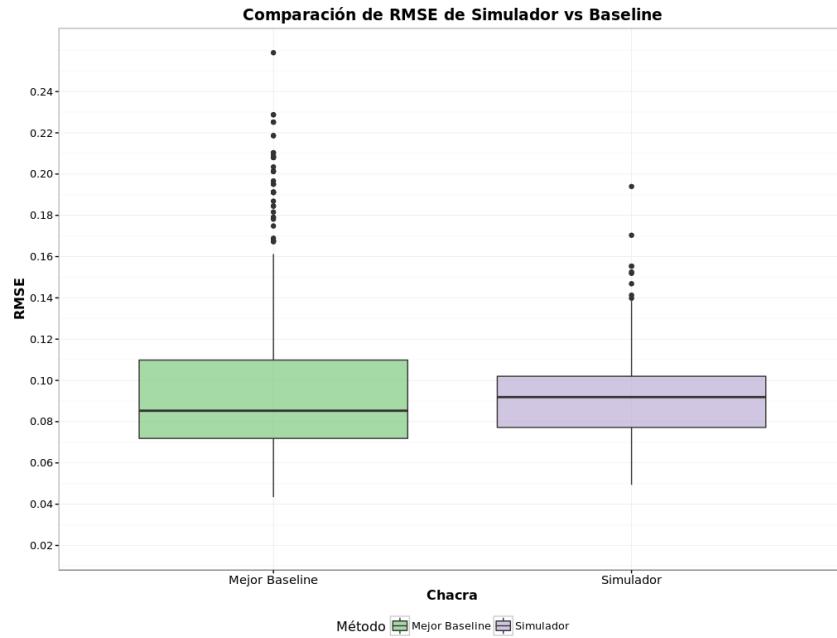


Figura 5.9: Comparación de resultados simulador vs *baseline* en todos los puntos, sin considerar últimas 5 observaciones.

totalmente incluídas en las chacras de los que no lo están. En la Tabla 5.8 y la Figura 5.10 se pueden apreciar los resultados obtenidos al utilizar sólo los puntos que están contenidos en un 95 % o más en las chacras. Ambos estimadores registran una pequeña baja de 0,1 % en el error, mientras que en la desviación estándar el simulador disminuye en 0,1 %, el *baseline* aumenta en el mismo valor.

	Prom. Error	Desv. Estándar
Simulador	0,0759	0,0216
Baseline	0,0966	0,0429

Tabla 5.8: Comparación de resultados simulador vs *baseline* en puntos contenidos en un 95 % o más en las chacras, sin considerar últimas 5 observaciones.

Finalmente, al considerar sólo los puntos totalmente contenidos en las chacras, los resultados mostrados en la Tabla 5.9 y Figura 5.11 repiten la tendencia: tanto el simulador como el *baseline* bajan en un 0,1 % el error. Con respecto a la desviación estándar, el simulador vuelve a bajar casi un 0,1 %, mientras que el *baseline* prácticamente mantiene su valor.

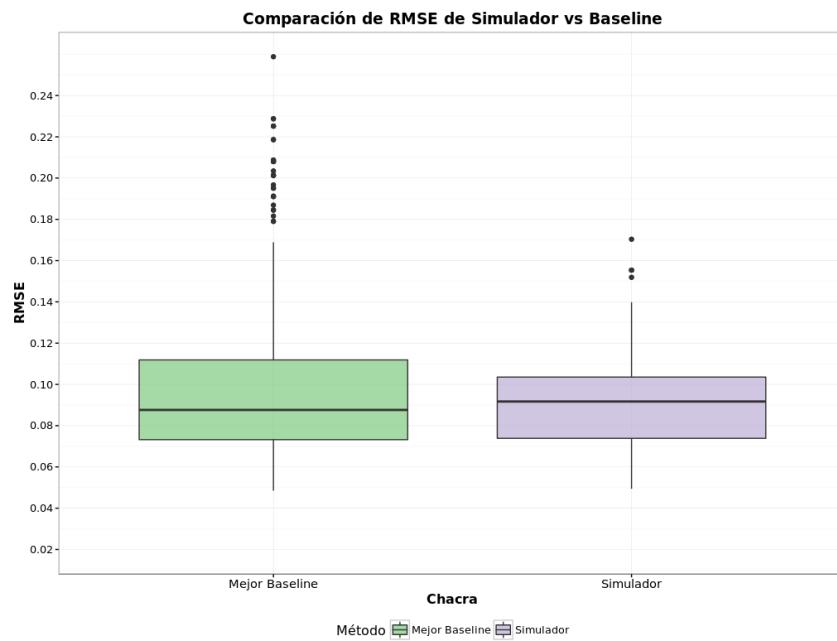


Figura 5.10: Comparación de resultados simulador vs *baseline* en puntos contenidos en un 95 % o más en las chacras, sin considerar últimas 5 observaciones.

	Prom. Error	Desv. Estándar
Simulador	0,0749	0,0208
Baseline	0,0957	0,0430

Tabla 5.9: Comparación de resultados simulador vs *baseline* en puntos contenidos en un 100 % en las chacras, sin considerar últimas 5 observaciones.

Chacra	Promedio Simulador	Desv. Estándar Simulador	Promedio Baseline	Desv. Estándar Baseline
CHURCHILL	0,0892	0,0269	0,1533	0,0496
POCHELLU	0,0710	0,0188	0,0675	0,0293
ALVAREZ OLANO	0,0633	0,0300	0,0738	0,0390
IL TRAMONTO	0,0857	0,0153	0,1269	0,0349
CUFRE	0,0725	0,0255	0,0708	0,0108
BERTOLETTI	0,0683	0,0147	0,0906	0,0134
TREALFA	0,0922	0,0169	0,0723	0,0146
Total	0,0769	0,0227	0,0960	0,0418

Tabla 5.10: Promedio de error y desviación estándar discriminado por chacra, sin tener en cuenta últimas 5 observaciones.

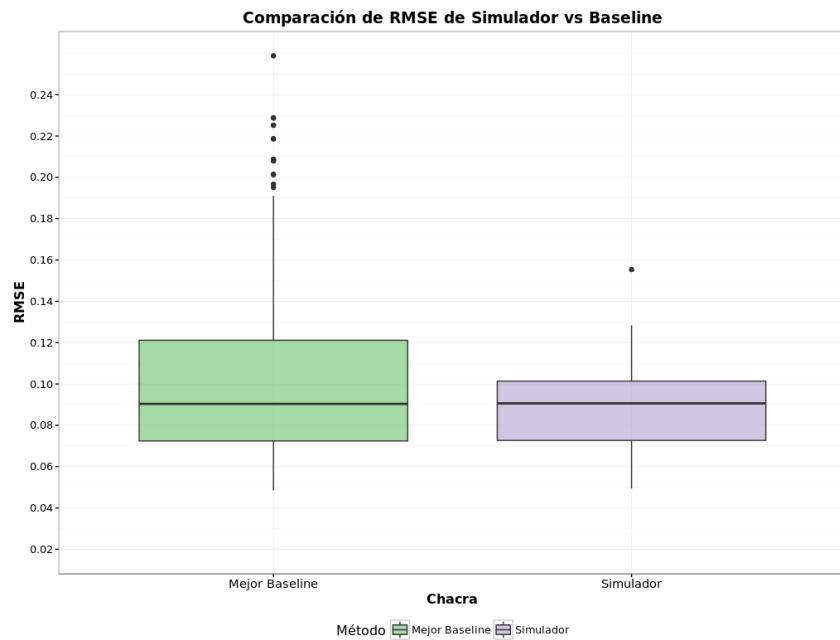


Figura 5.11: Comparación de resultados simulador vs *baseline* en puntos contenidos en un 100 % en las chacras, sin considerar últimas 5 observaciones.

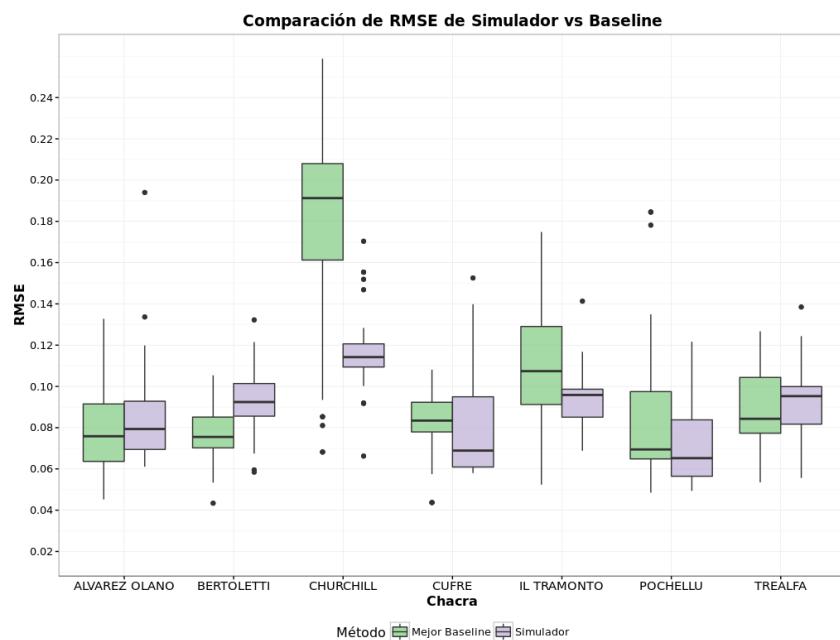


Figura 5.12: Comparación de resultados simulador vs *baseline* discriminado por chacra, sin tener en cuenta últimas 5 observaciones.

Capítulo 6

Conclusiones

En este trabajo se explican algunas de las razones por las que es de utilidad contar con un buen predictor del NDVI, ya que esto permitiría contar con una buena estimación del rendimiento de los cultivos. Una buena predicción puede ayudar a los productores para mejorar su planificación y mejorar el rendimiento de sus cultivos, así como podría ayudar a los gobiernos a monitorear el cumplimiento de restricciones sobre uso del suelo y rotación de cultivos. Otro posible uso, sería el de estimar los rendimientos a nivel global o regional, y de esta manera poder determinar los precios de una manera más exacta.

En el Capítulo 2 se muestra la relación que tienen el NDVI con el LAI, así como el NDVI y el rendimiento. Esta relación es importante dado que los simuladores de crecimiento de cultivos que el INIA utiliza y ha validado (pyGECROS y la adaptación de G. S. Campbell y R. Díaz -la utilizada en este trabajo-) retornan como resultado una estimación del LAI, mientras que de las imágenes satelitales utilizadas, se deriva el NDVI.

Asimismo, en el Capítulo 3, se describen los datos necesarios para completar esta tarea y el origen de éstos. Es de destacar que se ha llegado a una solución que es aplicable a la realidad y no sólo en el ámbito científico, ya que no se requiere de muchos datos iniciales, o de difícil obtención. En resumen, se puede asegurar que teniendo el contorno de la chacra (polígono vectorial), fecha de siembra y cosecha al menos 4 zonas, datos climáticos para las zonas, y características de la variedad de semilla plantada en las mismas, se puede calibrar el simulador. Con dicha calibración, se puede llegar a buenos resultados en siguientes simulaciones, aunque como lo indica la lógica, estarán fuertemente ligadas a los datos climáticos, que en caso de ser una incógnita (zonas futuras),

no serán más que suposiciones o pronósticos.

Para lograr el objetivo, se propone una solución que incluye el preprocesamiento de los datos de zafra, el procesamiento de las imágenes satelitales, la calibración del simulador y la evaluación de los resultados, comparándolos con otros dos estimadores (llamados *baselines*) construidos en el proyecto (ver Sección 4.6). Con respecto a las imágenes satelitales, se hace uso solamente de las MODIS y no las Landsat por problemas de calidad y cortes temporales (Ver Sección 4.5).

Como forma de evaluar la solución, se utiliza el Error Cuadrático Medio (RMSE). En este trabajo, se realizan pruebas en 7 chacras, que representan aproximadamente 1945 hectáreas. El error arrojado por el simulador (promedio general de todas las celdas) es de 0,0912 (9,1 %); mientras que el del *baseline* es de 0,0983 (9,8 %). Con estos resultados, podemos afirmar que se ha cumplido el objetivo de dar solución del problema planteado, y que se obtiene un error debajo del 10 %, lo que en el área de estudio se puede considerar como un buen resultado.

Sin embargo, se debe tener en cuenta que para llegar a un resultado aceptable, se debe realizar mucho trabajo con los datos, a pesar de que el simulador no requiera tantos. Existen diversas fuentes de error, entre las cuales están: transformación de LAI a NDVI, posibles imágenes con datos erróneos, errores en datos meteorológicos, errores en el simulador, calidad de los datos de las zafras.

Luego de realizado este proyecto y considerando algunos aspectos que comentaremos en esta sección, pensamos que a partir de esta investigación pueden surgir otras investigaciones similares que tomen como referencia lo realizado en este estudio. Uno de los aspectos a considerar que nos llevan a creer que a partir de esta investigación se pueden realizar otras similares, es que, como se comentó en la Sección 4.1 del Capítulo 4, el simulador de crecimiento de cultivos puede utilizarse con otros cultivos si se calibra adecuadamente. Por tanto, una de las posibles investigaciones que pueden derivarse de esta, sería una similar pero utilizando otro cultivo que no sea soja. En la Sección 2.6, mencionamos que existen varios simuladores de crecimiento de cultivos, y en particular se nombró al GECROS, simulador que el INIA utiliza, con adaptación en el lenguaje Python llamada pyGECROS. Teniendo en cuenta esto, también podría realizarse una investigación similar pero utilizando éste u otros simuladores cuyos resultados estén probados en nuestro país.

Asimismo, debemos considerar otras variantes que pueden realizarse sobre este estudio que surgen de utilizar otras fuentes de datos. El 23 de Junio del año 2015 se puso en órbita un nuevo satélite llamado sentinel-2A. El 7 de Marzo del año 2017 se puso en órbita otro satélite llamado sentinel-2B. Ambos forman parte de una misión de observación terrestre desarrollada por la Agencia Espacial Europea llamada sentinel-2. La misión principal de estos satélites es proporcionar información para las prácticas agrícolas y forestales y para ayudar a gestionar la seguridad alimentaria. La resolución de estos satélites es de 10 metros, muy superior a la que ofrece el satélite MODIS, de 250 metros. Y la frecuencia con la que toma imágenes de un mismo punto geográfico es de 5 días¹. Estos satélites siguen una política de datos abiertos. Por tanto, además de las variantes que pueden surgir de la utilización de otro simulador de crecimiento de cultivos y/o trabajar con otro cultivo, también podría realizarse la misma investigación pero utilizando otra fuente de datos de mayor precisión para medir el NDVI, como podrían ser los datos del satélite Sentinel. Esta investigación podría realizarse dentro de algunos años dado que es necesario acumular imágenes de cierta cantidad de años y estos satélites son muy recientes. Un proyecto que puede dar continuidad a la investigación realizada, sería desarrollar una herramienta que utilice lo expuesto en esta investigación para calibrar el simulador de crecimiento de cultivo, y posteriormente, permita realizar distintas simulaciones bajo distintos supuestos climáticos, utilizando por ejemplo la fecha de siembra o la variedad de semilla como variables sobre las cuales optimizar.

¹http://esamultimedia.esa.int/docs/S2-Data_Sheet.pdf

Referencias bibliográficas

- [1] Bondeau A, Smith PC, Zaehle S, Schaphoff S, Lucht W, Cramer W, et al. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Blackwell Publishing Ltd.; 2007.
- [2] Basso B, Cammarano D, Carfagna E. Review of crop yield forecasting methods and early warning systems. In: Proceedings of the First Meeting of the Scientific Advisory Committee of the Global Strategy to Improve Agricultural and Rural Statistics, FAO Headquarters, Rome, Italy; 2013. p. 18–19.
- [3] Panda SS, Ames DP, Panigrahi S. Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques. *Remote Sensing*. 2010 Mar;2(3):673–696. Available from: <http://dx.doi.org/10.3390/rs2030673>.
- [4] Rembold F, Atzberger C, Savin I, Rojas O. Using Low Resolution Satellite Imagery for Yield Prediction and Yield Anomaly Detection. *Remote Sensing*. 2013 Apr;5(4):1704–1733. Available from: <http://dx.doi.org/10.3390/rs5041704>.
- [5] Tucker CJ. Red and photographic infrared linear combinations for monitoring vegetation. *Remote sensing of Environment*. 1979;8(2):127–150.
- [6] Tucker CJ, Holben BN, Elgin Jr JH, McMurtrey III J, et al. Relationship of spectral data to grain yield variation. *Photogrammetric Engineering and Remote Sensing*. 1980;46(5):657–666.
- [7] Deering DW. Rangeland reflectance characteristics measured by aircraft and spacecraft sensors. *Dissertation Abstracts International*; 1978.

- [8] Potithep S, Nasahara NK, Muraoka H, Nagai S, Suzuki R. What is the actual relationship between LAI and VI in a deciduous broadleaf forest? In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science. vol. XXXVIII; 2010. .
- [9] Zhang H, Chen H, Zhou G. The Model of Wheat Yield Forecast Based on Modis-Ndvi - a Case Study of Xinxiang. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences. 2012 Jul;p. 25–28.
- [10] Rauff KO, Bello R. A Review of Crop Growth Simulation Models as Tools for Agricultural Meteorology. Agricultural Sciences. 2015 September;6:1098–1105.
- [11] Yin X, van Laar HH. Crop Systems Dynamics. The Netherlands: Wageningen Academic Publishers; 2005. Available from: <http://www.wageningenacademic.com/doi/abs/10.3920/978-90-8686-539-0>.
- [12] de Wit C, Brouwer R, De Vries FP. The simulation of photosynthetic systems. In: Proceedings of the IBP/PP technical meeting, Trebon, 14-21 September 1969; 1970. p. 47–70.
- [13] CT dWea. Simulation of assimilation, respiration and transpiration of Crops. PUDOC, editor. Simulation Monographs. Centre for Agricultural Publishing and Documentation, Wageningen, the Netherlands; 1978. Available from: <http://edepot.wur.nl/167486>.
- [14] De Vries FP, Van Laar H. Simulation of growth processes and the model BACROS. In: Simulation of plant growth and crop production. Pudoc; 1982. p. 114–135.
- [15] van Laar HH, Goudriaan J, van Keulen H. SUCROS97: Simulation of crop growth for potential and water-limited production situations. As applied to spring wheat. vol. 14 of Quantitive Approaches in Systems Analysis; 1997. Available from: <http://edepot.wur.nl/4426>.
- [16] van Keulen H, Seligman NG. Simulation of water use, nitrogen nutrition and growth of a spring wheat crop. Pudoc, editor. Simulation Monographs; 1987.

- [17] Bouman BAM, van Keulen H, van Laar HH, Rabbinge R. The ‘School of de Wit’ crop growth simulation models: A pedigree and historical overview. *Agricultural Systems*. 1996;52(2):171 – 198. Available from: <http://www.sciencedirect.com/science/article/pii/0308521X9600011X>.
- [18] Campbell GS, Diaz R. Simplified soil-water balance models to predict crop transpiration. In: Bidinger FR, Johansen C, editors. *Drought Research Priorities for the Dryland Tropics*. International Crops Research Institute for the Semi-Arid Tropics; 1988. p. 15–26.
- [19] Lopresti MF, Bella CMD, Degioanni AJ. Relationship between MODIS-NDVI data and wheat yield: A case study in Northern Buenos Aires province, Argentina. *Information Processing in Agriculture 2*. 2015;.
- [20] Perry EM, Morse-McNabb EM, Nuttall JG. Managing Wheat From Space: Linking MODIS NDVI and Crop Models for Predicting Australian Dryland Wheat Biomass. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2014;7.
- [21] Roya DP, Kovalsky V, Zhang HK, Vermote EF, Yana L, S S Kumara AE. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Geospatial Science Center of Excellence, South Dakota State University Brookings, SD 57007, USA*. 2015;.
- [22] INIA-INASE. Resultados experimentales de la Evaluación Nacional de Cultivares de Soja, Período 2012. *Evaluación Nacional de Cultivares*. 2012; Available from: http://www.inia.org.uy/convenio_inase_inia/Evaluacion_CV/Ano2012/PubSoja2012.pdf.
- [23] Yang D, Su H, Yong Y. MODIS-Landsat Data Fusion for Estimating Vegetation Dynamics - A Case Study for Two Ranches in Southwestern Texas. In: In Proceedings of the 1st Int. Electron. Conf. Remote Sens.. vol. 1. Sciforum Electronic Conference Series; 2015. .
- [24] Atzbergerb C, Høgdac KA, Johansenc B, Skidmoreb AK. Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *International Institute for Geo-Information Science and Earth Observation (ITC)*, PO Box 6, 7500 AA Enschede, The Netherlands. 2005;.

Glosario

Antesis “es el período de florescencia o floración de las plantas con flores; estrictamente, es el tiempo de expansión de una flor hasta que está completamente desarrollada y en estado funcional, durante el cual ocurre el proceso de polinización, si bien es frecuentemente usado para designar el período de floración en sí; el acto de florecer.” Wikipedia. [19](#)

Biomasa Masa total de los organismos en una determinada área. [12](#)

Conductancia estomatal Es la medida de la velocidad de paso del dióxido de carbono (CO₂) que entra, o vapor de agua que sale a través de los estomas de una hoja. Los estomas son pequeños poros en la parte superior e inferior de una hoja que son responsables de absorber y expulsar el CO₂ y la humedad desde y hacia el aire exterior. [18](#)

Índice de cosecha Fracción entre la masa de la semilla y la masa total de la planta. (Masa de la semilla / Masa de la planta). [12](#), [13](#)

Mesófilo Es un término botánico que designa el tejido que se encuentra entre las epidermis del haz y del envés de las hojas. [7](#)

Senescencia Es el proceso de envejecimiento en las plantas. Las plantas pueden tener envejecimiento inducido por estrés o por edad. [18](#)

APÉNDICES

Apéndice 1

Chacras utilizadas en el proyecto

Chacra Alvarez Olano

Ubicada en el departamento de Soriano, a poco más de 15km de la ciudad de Dolores y a 20 km de la ciudad de Mercedes. La chacra cuenta con 5 lotes (ver Figuras 1.1a y 1.1b) descriptos en la Tabla 1.1.

Lote	Superficie Total (Ha)	Superficie útil (Ha)
32	14,82	14,42
33	14,96	2,97
34	74,56	73,45
37	76,43	75,23
38	12,02	11,88
TOTAL	192,79	177,95

Tabla 1.1: Lotes de la chacra Alvarez Olano y sus superficies



(a) Ubicación de la chacra Alvarez Olano en el departamento de Soriano. **(b)** Lotes de la chacra Alvarez Olano.

Figura 1.1: Imágenes satelitales de chacra Alvarez Olano mostrando su ubicación y sus lotes.

Chacra Bertoletti

Ubicada en el departamento de Colonia, a 10km de la ciudad de Ombués de Lavalle. La chacra cuenta con 15 lotes (ver Figuras 1.2a y 1.2b) descriptos en la Tabla 1.2.

Lote	Superficie Total (Ha)	Superficie útil (Ha)
1	111,71	67,45
2	59,51	25,70
3	60,33	40,75
4	51,28	46,38
5	31,47	24,37
6	65,25	62,37
7	68,70	58,80
8	26,91	24,69
9	48,46	32,66
10	33,82	8,77
11	19,34	17,98
12	222,57	122,41
13	160,56	110,77
14	80,36	12,59
15	81,22	53,06
TOTAL	1122,48	708,77

Tabla 1.2: Lotes de la chacra Bertoletti y sus superficies



(a) Ubicación de la chacra Bertoletti en el departamento de Colonia. **(b)** Lotes de la chacra Bertoletti.

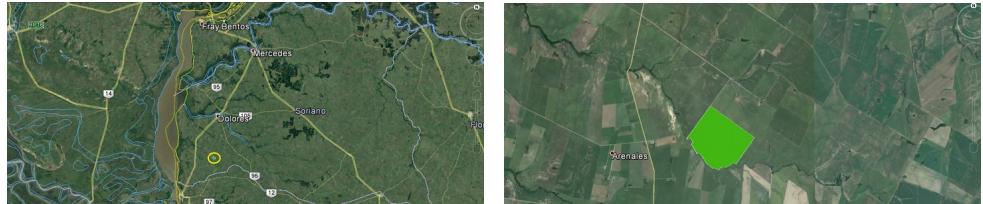
Figura 1.2: Imágenes satelitales de chacra Bertoletti mostrando su ubicación y sus lotes.

Chacra Churchill

Ubicada en el departamento de Soriano, a 21km de la ciudad de Dolores y a 11km de la localidad de Agraciada. La chacra cuenta con un sólo lote (ver Figuras 1.3a y 1.3b, y Tabla 1.3).

Lote	Superficie Total (Ha)	Superficie útil (Ha)
1	155,72	149,84
TOTAL	155,72	149,84

Tabla 1.3: Lotes de la chacra Churchill y sus superficies



(a) Ubicación de la chacra Churchill en el departamento de Soriano. (b) Lotes de la chacra Churchill.

Figura 1.3: Imágenes satelitales de chacra Churchill mostrando su ubicación y sus lotes.

Chacra Cufré

Ubicada en el departamento de Colonia, a sólo 2km de la localidad de Cufré en el límite departamental con San José. La chacra cuenta 5 lotes (ver Figuras 1.4a y 1.4b) descriptos en la Tabla 1.4.

Lote	Superficie Total (Ha)	Superficie útil (Ha)
1	46,49	46,14
2	51,93	48,44
3	47,81	47,37
4	66,57	65,55
5	26,45	25,70
Total	239,25	233,20

Tabla 1.4: Lotes de la chacra Cufré y sus superficies



(a) Ubicación de la chacra Cufré en el departamento de Colonia.

(b) Lotes de la chacra Cufré.

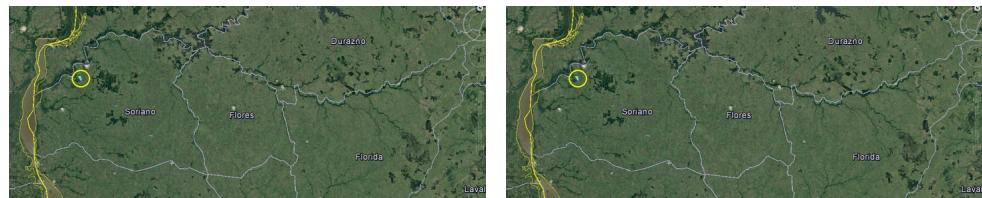
Figura 1.4: Imágenes satelitales de chacra Cufré mostrando su ubicación y sus lotes.

Chacra Il Tramonto

Ubicada en el departamento de Soriano, a solo 10 km de la ciudad de Mercedes. La chacra cuenta con un solo lote (ver Figuras 1.5a y 1.5b, y Tabla 1.5).

Lote	Superficie Total (Ha)	Superficie útil (Ha)
1	386,18	266,33
Total	386,18	266,33

Tabla 1.5: Lotes de la chacra Il Tramonto y sus superficies.



(a) Ubicación de la chacra Il Tramonto en el departamento de Soriano.

(b) Lotes de la chacra Cufré.

Figura 1.5: Imágenes satelitales de chacra Il Tramonto mostrando su ubicación y sus lotes.

Chacra Pochellu

Parte de la chacra se encuentra en el departamento de Colonia y parte en el departamento de Soriano. Se encuentra a 11km de la localidad de Agraciada y a 27km de la ciudad de Nueva Palmira. La chacra cuenta con 3 lotes (ver Figuras 1.6a y 1.6b) descriptos en la Tabla 1.6.

Lote	Superficie Total (Ha)	Superficie útil (Ha)
1	65,06	58,44
2	109,85	87,60
3	57,31	56,03
Total	232,22	202,07

Tabla 1.6: Lotes de la chacra Pochellu y sus superficies



(a) Ubicación de la chacra Pochellu en el límite de los departamentos de Colonia y Soriano.

(b) Lotes de la chacra Pochellu.

Figura 1.6: Imágenes satelitales de chacra Pochellu mostrando su ubicación y sus lotes.

Chacra Trealfa

Ubicada en el sur oeste del departamento de San José, a sólo 5km de la ciudad de Ecilda Paullier. La chacra cuenta con 8 lotes (ver Figuras 1.7a y 1.7b, y Tabla 1.7).

Lote	Superficie Total (Ha)	Superficie útil (Ha)
1	24,27	24,27
2	16,84	14,88
3	7,42	7,42
4	18,04	18,01
5	65,13	65,07
6	5,22	5,22
7-8	11,04	11,04
9	15,68	15,68
10	45,36	45,36
Total	225,08	206,95

Tabla 1.7: Lotes de la chacra Trealfa y sus superficies.



(a) Ubicación de la chacra Trealfa en el departamento de San José.

(b) Lotes de la chacra Trealfa.

Figura 1.7: Imágenes satelitales de chacra Trealfa mostrando su ubicación y sus lotes.

Apéndice 2

Resultados detallados

Chacra Alvarez Olano

A continuación se muestran los resultados obtenidos de la simulación y de los algoritmos *baseline* para el establecimiento Alvarez Olano.

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
32	4	93 %	1,1312	0,2688	0,1146	0,2039	0,0637	0,0801	DOBLE LOGÍSTICA
32	5	100 %	1,1781	0,2792	0,1179	0,1557	0,1299	0,1178	SIMULADOR
32	8	97 %	1,4547	0,2414	0,1058	0,2446	0,0928	0,1124	DOBLE LOGÍSTICA
34	2	83 %	0,8950	0,3005	0,1248	0,1496	0,1101	0,1002	SIMULADOR
34	3	99 %	0,8609	0,2958	0,1233	0,1497	0,0923	0,0826	SIMULADOR
34	8	93 %	0,9967	0,3361	0,1362	0,1904	0,0803	0,0725	SIMULADOR
34	9	100 %	1,1389	0,3039	0,1259	0,1936	0,0486	0,0693	DOBLE LOGÍSTICA
34	10	100 %	1,1389	0,3039	0,1259	0,1936	0,0486	0,0693	DOBLE LOGÍSTICA
34	16	100 %	1,1889	0,2943	0,1228	0,1581	0,0733	0,0758	DOBLE LOGÍSTICA
34	17	100 %	1,0745	0,3205	0,1312	0,2032	0,0526	0,0682	DOBLE LOGÍSTICA
34	18	81 %	1,0745	0,3205	0,1312	0,2032	0,0526	0,0682	DOBLE LOGÍSTICA
34	23	100 %	1,0678	0,3216	0,1315	0,2045	0,1119	0,0859	SIMULADOR
34	24	100 %	1,0678	0,3216	0,1315	0,2045	0,1119	0,0859	SIMULADOR
34	30	94 %	1,0678	0,3216	0,1315	0,2045	0,1119	0,0859	SIMULADOR
34	31	100 %	1,0016	0,3241	0,1324	0,2137	0,0582	0,0732	DOBLE LOGÍSTICA
34	32	100 %	1,0016	0,3241	0,1324	0,2137	0,0582	0,0732	DOBLE LOGÍSTICA
34	37	94 %	1,1888	0,3081	0,1272	0,1674	0,0685	0,0987	DOBLE LOGÍSTICA
34	38	100 %	1,1814	0,3225	0,1318	0,2024	0,0504	0,0699	DOBLE LOGÍSTICA
34	39	100 %	1,3173	0,2733	0,1160	0,2153	0,0642	0,0928	DOBLE LOGÍSTICA
34	40	87 %	1,3173	0,2733	0,1160	0,2153	0,0642	0,0928	DOBLE LOGÍSTICA
34	45	100 %	1,0380	0,3304	0,1344	0,1795	0,0673	0,0861	DOBLE LOGÍSTICA
34	46	100 %	1,1169	0,3104	0,1280	0,2254	0,0712	0,0694	SIMULADOR
34	47	91 %	0,9967	0,3349	0,1358	0,2217	0,0613	0,0856	DOBLE LOGÍSTICA
34	51	79 %	0,9659	0,3545	0,1421	0,1416	0,0776	0,1337	DOBLE LOGÍSTICA
34	52	80 %	0,9077	0,3608	0,1441	0,1707	0,0682	0,0771	DOBLE LOGÍSTICA

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
34	53	82 %	0,9967	0,3497	0,1406	0,1942	0,0738	0,0892	DOBLE LOGÍSTICA
37	2	92 %	0,8967	0,3540	0,1419	0,1891	0,0624	0,1001	DOBLE LOGÍSTICA
37	3	100 %	0,8967	0,3540	0,1419	0,1891	0,0624	0,1001	DOBLE LOGÍSTICA
37	4	100 %	0,9463	0,3349	0,1358	0,1897	0,0767	0,0786	DOBLE LOGÍSTICA
37	5	80 %	0,9463	0,3349	0,1358	0,1897	0,0767	0,0786	DOBLE LOGÍSTICA
37	7	100 %	0,9946	0,3735	0,1482	0,1688	0,1121	0,0910	SIMULADOR
37	8	100 %	1,1794	0,3094	0,1276	0,1808	0,0725	0,0611	SIMULADOR
37	9	100 %	1,1794	0,3094	0,1276	0,1808	0,0725	0,0611	SIMULADOR
37	10	82 %	1,0422	0,3346	0,1357	0,1499	0,0452	0,0784	DOBLE LOGÍSTICA
37	12	100 %	1,0062	0,3298	0,1342	0,2145	0,0628	0,1198	DOBLE LOGÍSTICA
37	13	100 %	1,0798	0,3191	0,1307	0,1983	0,0785	0,0678	SIMULADOR
37	14	100 %	1,1897	0,3075	0,1270	0,1685	0,0915	0,0665	SIMULADOR
37	15	80 %	1,1897	0,3075	0,1270	0,1685	0,0915	0,0665	SIMULADOR
37	17	100 %	1,0798	0,3191	0,1307	0,1983	0,0785	0,0678	SIMULADOR
37	18	100 %	1,0798	0,3191	0,1307	0,1983	0,0785	0,0678	SIMULADOR
37	19	100 %	1,1897	0,3075	0,1270	0,1685	0,0915	0,0665	SIMULADOR
37	20	87 %	1,0611	0,3245	0,1325	0,1549	0,0717	0,0757	DOBLE LOGÍSTICA
37	22	99 %	1,0288	0,3117	0,1284	0,2166	0,1118	0,1096	SIMULADOR
37	23	99 %	1,1900	0,3082	0,1273	0,1717	0,0758	0,0727	SIMULADOR
37	24	100 %	1,1900	0,3082	0,1273	0,1717	0,0758	0,0727	SIMULADOR
37	25	97 %	1,8632	0,2247	0,1005	0,1801	0,1247	0,0819	SIMULADOR
37	27	97 %	0,8681	0,3691	0,1468	0,1714	0,0787	0,0987	DOBLE LOGÍSTICA
37	28	100 %	0,8681	0,3691	0,1468	0,1714	0,0787	0,0987	DOBLE LOGÍSTICA
37	29	100 %	1,4600	0,2567	0,1107	0,1279	0,0822	0,0889	DOBLE LOGÍSTICA
38	3	83 %	0,7673	0,2464	0,1074	0,1800	0,1328	0,1940	DOBLE LOGÍSTICA

Tabla 2.1: Resultados para la chacra Alvarez Olano

Los algoritmos *baseline*, tienen un menor RMSE en 29 celdas del establecimiento, mientras el simulador de crecimiento de cultivos tuvo un menor RMSE en 21 celdas del establecimiento. El promedio de RMSE para todas las celdas del establecimiento los algoritmos *baseline* es de 0,0790; mientras el promedio de RMSE para todas las celdas del establecimiento del simulador de crecimiento de cultivos es de 0,0856.

Chacra Bertoletti

A continuación se muestran los resultados obtenidos de la simulación y de los algoritmos *baseline* para el establecimiento Bertoletti.

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
2	2	81 %	0,2739	0,1500	0,0765	0,0690	0,0632	0,0871	DOBLE LOGÍSTICA
3	4	94 %	0,1828	0,2815	0,1187	0,0645	0,0827	0,0781	POLINOMIO
3	5	100 %	0,1828	0,2815	0,1187	0,0645	0,0827	0,0781	POLINOMIO
3	7	97 %	0,1744	0,2751	0,1166	0,0788	0,0749	0,0675	SIMULADOR
3	8	100 %	0,1744	0,2751	0,1166	0,0788	0,0749	0,0675	SIMULADOR
3	9	100 %	0,1934	0,2738	0,1162	0,0831	0,0597	0,0739	DOBLE LOGÍSTICA
3	12	97 %	0,1744	0,2751	0,1166	0,0788	0,0749	0,0675	SIMULADOR
3	13	100 %	0,1934	0,2738	0,1162	0,0831	0,0597	0,0739	DOBLE LOGÍSTICA
3	14	83 %	0,1934	0,2738	0,1162	0,0831	0,0597	0,0739	DOBLE LOGÍSTICA
3	16	78 %	0,1793	0,1518	0,0771	0,0434	0,0650	0,0852	POLINOMIO
3	18	80 %	0,1549	0,1700	0,0829	0,0726	0,0613	0,0595	SIMULADOR
4	2	92 %	0,2026	0,3097	0,1277	0,0999	0,0890	0,0969	DOBLE LOGÍSTICA
4	3	91 %	0,2026	0,3097	0,1277	0,0999	0,0890	0,0969	DOBLE LOGÍSTICA
4	4	89 %	0,1868	0,3644	0,1453	0,0916	0,1035	0,1008	POLINOMIO
4	5	80 %	0,1802	0,3412	0,1378	0,0889	0,0887	0,0900	DOBLE LOGÍSTICA
4	8	100 %	0,2026	0,3097	0,1277	0,0999	0,0890	0,0969	DOBLE LOGÍSTICA
4	9	100 %	0,1868	0,3644	0,1453	0,0916	0,1035	0,1008	POLINOMIO
4	10	100 %	0,1802	0,3412	0,1378	0,0889	0,0887	0,0900	DOBLE LOGÍSTICA
4	11	100 %	0,1802	0,3412	0,1378	0,0889	0,0887	0,0900	DOBLE LOGÍSTICA
4	12	98 %	0,1788	0,3730	0,1480	0,0834	0,1076	0,1197	POLINOMIO
4	15	91 %	0,1877	0,3019	0,1252	0,0916	0,0779	0,0758	SIMULADOR
4	16	97 %	0,4525	0,1503	0,0766	0,1047	0,0593	0,1036	DOBLE LOGÍSTICA
4	17	100 %	0,4278	0,1500	0,0765	0,1011	0,0681	0,0997	DOBLE LOGÍSTICA
4	18	91 %	0,4278	0,1500	0,0765	0,1011	0,0681	0,0997	DOBLE LOGÍSTICA
5	6	94 %	0,3792	0,1500	0,0765	0,0750	0,0757	0,0844	POLINOMIO
5	7	76 %	0,3428	0,1500	0,0765	0,0890	0,0832	0,0856	DOBLE LOGÍSTICA
5	10	96 %	0,3792	0,1500	0,0765	0,0750	0,0757	0,0844	POLINOMIO
5	11	99 %	0,3428	0,1500	0,0765	0,0890	0,0832	0,0856	DOBLE LOGÍSTICA
6	12	99 %	0,2277	0,1796	0,0860	0,0749	NO CONV.	0,0856	POLINOMIO
6	13	96 %	0,2658	0,1718	0,0835	0,0950	0,1814	0,0967	POLINOMIO
6	14	86 %	0,2658	0,1718	0,0835	0,0950	0,1814	0,0967	POLINOMIO
6	15	97 %	0,1952	0,2459	0,1073	0,1090	0,0681	0,0914	DOBLE LOGÍSTICA
6	16	100 %	0,1963	0,2609	0,1121	0,1042	0,0729	0,0947	DOBLE LOGÍSTICA
6	17	100 %	0,1963	0,2609	0,1121	0,1042	0,0729	0,0947	DOBLE LOGÍSTICA
6	18	84 %	0,3012	0,1500	0,0765	0,1153	0,0712	0,1034	DOBLE LOGÍSTICA
6	19	100 %	0,3012	0,1500	0,0765	0,1153	0,0712	0,1034	DOBLE LOGÍSTICA
6	20	96 %	0,1815	0,1500	0,0765	0,0936	0,0757	0,0879	DOBLE LOGÍSTICA
6	21	99 %	0,2997	0,1500	0,0765	0,1008	0,0705	0,1056	DOBLE LOGÍSTICA
6	23	97 %	0,1892	0,3030	0,1256	0,0995	0,0755	0,1071	DOBLE LOGÍSTICA
6	24	99 %	0,2569	0,2129	0,0967	0,0874	0,0667	0,1022	DOBLE LOGÍSTICA
6	25	91 %	0,2569	0,2129	0,0967	0,0874	0,0667	0,1022	DOBLE LOGÍSTICA
6	26	100 %	0,1932	0,2717	0,1155	0,0957	0,0711	0,0949	DOBLE LOGÍSTICA
6	27	100 %	0,2826	0,1500	0,0765	0,0990	0,0653	0,0919	DOBLE LOGÍSTICA
6	28	80 %	0,2826	0,1500	0,0765	0,0990	0,0653	0,0919	DOBLE LOGÍSTICA
6	29	80 %	0,1817	0,3578	0,1432	0,0992	0,0660	0,0901	DOBLE LOGÍSTICA
7	9	77 %	0,1785	0,2603	0,1119	0,0946	0,0889	0,0975	DOBLE LOGÍSTICA
7	12	100 %	0,1664	0,1963	0,0914	0,0930	0,1077	0,0904	SIMULADOR
7	13	99 %	0,1679	0,2853	0,1199	0,0854	0,0852	0,0775	SIMULADOR
7	14	100 %	0,1768	0,2625	0,1126	0,0905	0,0754	0,0717	SIMULADOR

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
7	15	95 %	0,1768	0,2625	0,1126	0,0905	0,0754	0,0717	SIMULADOR
7	16	94 %	0,1776	0,2725	0,1158	0,1002	0,0978	0,0905	SIMULADOR
7	17	100 %	0,1930	0,2704	0,1151	0,0979	0,0671	0,0924	DOBLE LOGÍSTICA
7	18	100 %	0,1930	0,2704	0,1151	0,0979	0,0671	0,0924	DOBLE LOGÍSTICA
7	19	91 %	0,1712	0,2163	0,0978	0,0921	0,0689	0,0720	DOBLE LOGÍSTICA
7	22	100 %	0,1885	0,1781	0,0855	0,0762	NO CONV.	0,0908	POLINOMIO
7	23	91 %	0,1885	0,1781	0,0855	0,0762	NO CONV.	0,0908	POLINOMIO
7	26	100 %	0,1731	0,1537	0,0777	0,0784	0,0615	0,0709	DOBLE LOGÍSTICA
7	27	98 %	0,1806	0,1522	0,0772	0,0925	0,0947	0,0876	SIMULADOR
7	28	84 %	0,2373	0,1500	0,0765	0,0822	0,0810	0,0971	DOBLE LOGÍSTICA
7	29	86 %	0,2373	0,1500	0,0765	0,0822	0,0810	0,0971	DOBLE LOGÍSTICA
8	1	95 %	0,1827	0,1500	0,0765	0,1053	0,0760	0,1018	DOBLE LOGÍSTICA
8	2	97 %	0,1827	0,1500	0,0765	0,1053	0,0760	0,1018	DOBLE LOGÍSTICA
8	3	79 %	0,1784	0,1507	0,0767	0,0991	0,0975	0,1006	DOBLE LOGÍSTICA
8	4	98 %	0,2297	0,1860	0,0880	0,0820	0,0702	0,0932	DOBLE LOGÍSTICA
8	5	94 %	0,1836	0,1800	0,0861	0,0888	0,0610	0,0949	DOBLE LOGÍSTICA
8	7	92 %	0,1848	0,1833	0,0872	0,0736	0,0719	0,0885	DOBLE LOGÍSTICA
8	8	95 %	0,1848	0,1833	0,0872	0,0736	0,0719	0,0885	DOBLE LOGÍSTICA
11	7	89 %	0,1849	0,2923	0,1221	0,0851	0,0951	0,0759	SIMULADOR
11	11	92 %	0,2690	0,2125	0,0966	0,0832	0,0807	0,0988	DOBLE LOGÍSTICA
11	14	79 %	0,2690	0,2125	0,0966	0,0832	0,0807	0,0988	DOBLE LOGÍSTICA
12	4	76 %	0,1601	0,2744	0,1164	0,0901	0,0730	0,0848	DOBLE LOGÍSTICA
12	5	100 %	0,1600	0,3200	0,1310	0,1053	0,1059	0,0585	SIMULADOR
13	4	99 %	0,2373	0,1500	0,0765	0,0822	0,0810	0,0971	DOBLE LOGÍSTICA
13	5	99 %	0,1644	0,2165	0,0978	0,0720	NO CONV.	0,1006	POLINOMIO
13	6	76 %	0,1644	0,2165	0,0978	0,0720	NO CONV.	0,1006	POLINOMIO
13	10	94 %	0,1656	0,1854	0,0879	0,0897	0,0746	0,0831	DOBLE LOGÍSTICA
13	11	94 %	0,1835	0,1502	0,0766	0,0965	0,0824	0,0884	DOBLE LOGÍSTICA
13	12	100 %	0,1835	0,1502	0,0766	0,0965	0,0824	0,0884	DOBLE LOGÍSTICA
13	13	100 %	0,1800	0,2962	0,1234	0,1011	0,0942	0,1014	DOBLE LOGÍSTICA
13	18	92 %	0,1690	0,1500	0,0765	0,0721	0,0628	0,0855	DOBLE LOGÍSTICA
13	19	100 %	0,1777	0,1966	0,0915	0,0968	0,0947	0,1067	DOBLE LOGÍSTICA
13	20	87 %	0,1777	0,1966	0,0915	0,0968	0,0947	0,1067	DOBLE LOGÍSTICA
13	25	97 %	0,1805	0,2007	0,0928	0,0803	0,0986	0,0863	POLINOMIO
13	26	100 %	0,1849	0,2639	0,1130	0,0745	0,0976	0,1028	POLINOMIO
13	30	100 %	0,1805	0,2007	0,0928	0,0803	0,0986	0,0863	POLINOMIO
13	31	100 %	0,1805	0,2007	0,0928	0,0803	0,0986	0,0863	POLINOMIO
13	32	98 %	0,1849	0,2639	0,1130	0,0745	0,0976	0,1028	POLINOMIO
13	33	100 %	0,1849	0,2639	0,1130	0,0745	0,0976	0,1028	POLINOMIO
13	34	99 %	0,1842	0,1585	0,0792	0,0894	0,1073	0,1215	POLINOMIO
13	37	96 %	0,2549	0,1500	0,0765	0,0862	0,0839	0,1169	DOBLE LOGÍSTICA
13	38	100 %	0,4536	0,1500	0,0765	0,0868	0,0881	0,1116	POLINOMIO
13	39	100 %	0,4536	0,1500	0,0765	0,0868	0,0881	0,1116	POLINOMIO
13	40	100 %	0,2206	0,1511	0,0769	0,0852	0,0844	0,1101	DOBLE LOGÍSTICA
13	44	93 %	0,3244	0,1500	0,0765	0,0884	0,0995	0,1128	POLINOMIO
13	45	82 %	0,1681	0,2528	0,1095	0,0869	0,1076	0,1068	POLINOMIO
13	46	93 %	0,1583	0,2017	0,0931	0,0815	0,6368	0,1040	POLINOMIO
13	53	95 %	0,1648	0,1862	0,0881	0,0781	0,0835	0,0879	POLINOMIO
13	60	92 %	0,1691	0,1500	0,0765	0,0756	0,0844	0,0935	POLINOMIO

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
13	61	82 %	0,1691	0,1500	0,0765	0,0756	0,0844	0,0935	POLINOMIO
13	65	97 %	0,1622	0,2608	0,1120	0,0668	0,0821	0,0799	POLINOMIO
13	66	96 %	0,1622	0,2608	0,1120	0,0668	0,0821	0,0799	POLINOMIO
13	71	99 %	0,1643	0,1566	0,0786	0,0601	0,0534	0,1059	DOBLE LOGÍSTICA
13	72	90 %	0,1628	0,1815	0,0866	0,0719	0,0822	0,0725	POLINOMIO
14	5	99 %	0,1683	0,3049	0,1262	0,0876	0,1045	0,1094	POLINOMIO
14	6	79 %	0,1694	0,1500	0,0765	0,0749	0,1249	0,1322	POLINOMIO

Tabla 2.2: Resultados para la chacra Bertoletti

Los algoritmos *baseline* tienen un menor RMSE en 92 celdas del establecimiento, mientras el simulador de crecimiento de cultivos tuvo un menor RMSE en 13 celdas. El promedio de RMSE para todas las celdas del establecimiento de los algoritmos *baseline* es de 0,0770, mientras el promedio de RMSE para todas las celdas del establecimiento del simulador de crecimiento de cultivos es de 0,0924.

Chacra Churchill

A continuación se muestran los resultados obtenidos de la simulación y de los algoritmos *baseline* para el establecimiento Churchill.

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
1	4	86 %	0,7757	0,3129	0,1288	0,1888	0,0935	0,0662	SIMULADOR
1	15	78 %	0,9624	0,3069	0,1268	0,2514	0,1518	0,0919	SIMULADOR
1	16	100 %	0,9624	0,3069	0,1268	0,2514	0,1518	0,0919	SIMULADOR
1	17	77 %	0,9970	0,4000	0,1567	0,2776	0,1913	0,1087	SIMULADOR
1	18	83 %	0,9970	0,4000	0,1567	0,2776	0,1913	0,1087	SIMULADOR
1	27	94 %	0,9624	0,3069	0,1268	0,2514	0,1518	0,0919	SIMULADOR
1	28	98 %	0,9970	0,4000	0,1567	0,2776	0,1913	0,1087	SIMULADOR
1	30	97 %	1,5965	0,4000	0,1567	0,2255	0,1816	0,1037	SIMULADOR
1	31	100 %	0,9612	0,2927	0,1223	0,2063	0,1341	0,1182	SIMULADOR
1	32	99 %	0,9612	0,2927	0,1223	0,2063	0,1341	0,1182	SIMULADOR
1	38	100 %	1,9455	0,4000	0,1567	0,2382	0,1790	0,1002	SIMULADOR
1	39	100 %	1,9455	0,4000	0,1567	0,2382	0,1790	0,1002	SIMULADOR
1	41	94 %	1,9045	0,4000	0,1567	0,2812	0,1673	0,1121	SIMULADOR
1	42	95 %	1,9045	0,4000	0,1567	0,2812	0,1673	0,1121	SIMULADOR
1	43	93 %	1,9018	0,4000	0,1567	0,2854	0,2104	0,1129	SIMULADOR
1	44	94 %	1,6987	0,4000	0,1567	0,2351	0,1612	0,1049	SIMULADOR
1	45	100 %	1,6987	0,4000	0,1567	0,2351	0,1612	0,1049	SIMULADOR
1	46	90 %	0,9728	0,3031	0,1256	0,1967	0,1588	0,1240	SIMULADOR
1	50	100 %	1,9633	0,4000	0,1567	0,2402	0,0811	0,1283	DOBLE LOGÍSTICA

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
1	51	100 %	1,9450	0,4000	0,1567	0,2752	0,1869	0,1094	SIMULADOR
1	52	100 %	1,9305	0,4000	0,1567	0,3188	0,2087	0,1133	SIMULADOR
1	53	100 %	1,9305	0,4000	0,1567	0,3188	0,2087	0,1133	SIMULADOR
1	54	100 %	1,9292	0,4000	0,1567	0,3222	0,2186	0,1151	SIMULADOR
1	55	100 %	1,9292	0,4000	0,1567	0,3222	0,2186	0,1151	SIMULADOR
1	56	100 %	1,9173	0,4000	0,1567	0,3043	0,1967	0,1133	SIMULADOR
1	57	100 %	1,9023	0,4000	0,1567	0,2871	0,2012	0,1204	SIMULADOR
1	58	99 %	1,9023	0,4000	0,1567	0,2871	0,2012	0,1204	SIMULADOR
1	59	99 %	1,7068	0,4000	0,1567	0,2587	0,1689	0,1519	SIMULADOR
1	62	98 %	1,9633	0,4000	0,1567	0,2071	0,0853	0,1554	DOBLE LOGÍSTICA
1	63	100 %	1,9633	0,4000	0,1567	0,2071	0,0853	0,1554	DOBLE LOGÍSTICA
1	64	100 %	1,9512	0,4000	0,1567	0,2931	0,2588	0,1128	SIMULADOR
1	65	100 %	1,9381	0,4000	0,1567	0,3151	0,1910	0,1143	SIMULADOR
1	66	100 %	1,9381	0,4000	0,1567	0,3151	0,1910	0,1143	SIMULADOR
1	67	100 %	1,9247	0,4000	0,1567	0,3209	0,2287	0,1130	SIMULADOR
1	68	100 %	1,9247	0,4000	0,1567	0,3209	0,2287	0,1130	SIMULADOR
1	69	100 %	1,9181	0,4000	0,1567	0,3127	0,2015	0,1164	SIMULADOR
1	70	88 %	1,9381	0,4000	0,1567	0,3000	0,2102	0,1468	SIMULADOR
1	75	88 %	0,9490	0,2927	0,1223	0,1250	0,0682	0,1251	DOBLE LOGÍSTICA
1	76	100 %	0,9490	0,2927	0,1223	0,1250	0,0682	0,1251	DOBLE LOGÍSTICA
1	77	100 %	1,9451	0,4000	0,1567	0,2252	0,2379	0,1206	SIMULADOR
1	78	100 %	1,9305	0,4000	0,1567	0,3094	0,2080	0,1142	SIMULADOR
1	79	100 %	1,9305	0,4000	0,1567	0,3094	0,2080	0,1142	SIMULADOR
1	80	96 %	1,7084	0,4000	0,1567	0,2911	0,1951	0,1251	SIMULADOR
1	81	95 %	1,9396	0,4000	0,1567	0,2517	0,2034	0,1704	SIMULADOR
1	88	96 %	1,9451	0,4000	0,1567	0,2252	0,2379	0,1206	SIMULADOR
1	89	100 %	1,9451	0,4000	0,1567	0,2252	0,2379	0,1206	SIMULADOR
1	90	100 %	1,9305	0,4000	0,1567	0,3094	0,2080	0,1142	SIMULADOR
1	91	100 %	1,7084	0,4000	0,1567	0,2911	0,1951	0,1251	SIMULADOR
1	92	93 %	1,7084	0,4000	0,1567	0,2911	0,1951	0,1251	SIMULADOR

Tabla 2.3: Resultados para la chacra Churchill

Los algoritmos *baseline* tienen un menor RMSE en 5 celdas del establecimiento, mientras el simulador de crecimiento de cultivos tiene un menor RMSE en 44 celdas. El promedio de RMSE para todas las celdas del establecimiento de los algoritmos *baseline* es de 0,1786; mientras el promedio de RMSE para todas las celdas del establecimiento del simulador de crecimiento de cultivos es de 0,1168.

Chacra Cufré

A continuación se muestran los resultados obtenidos de la simulación y de los algoritmos *baseline* para el establecimiento Cufré.

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
3	6	100 %	0,8586	0,3564	0,1427	0,0923	0,1427	0,0598	SIMULADOR
3	7	79 %	0,8586	0,3564	0,1427	0,0923	0,1427	0,0598	SIMULADOR
3	9	88 %	0,2531	0,1500	0,0765	0,0908	0,0799	0,0948	DOBLE LOGÍSTICA
3	10	100 %	0,2531	0,1500	0,0765	0,0908	0,0799	0,0948	DOBLE LOGÍSTICA
3	11	100 %	0,8586	0,3209	0,1313	0,0957	0,1065	0,0687	SIMULADOR
3	13	95 %	0,8586	0,3564	0,1427	0,0923	0,1427	0,0598	SIMULADOR
3	14	91 %	0,2533	0,1500	0,0765	0,0897	0,1161	0,0937	POLINOMIO
3	16	100 %	0,1924	0,1500	0,0765	0,0856	0,1030	0,1187	POLINOMIO
3	17	87 %	0,3717	0,1544	0,0779	0,1064	0,1154	0,0824	SIMULADOR
3	19	93 %	0,1818	0,3586	0,1434	0,0994	0,0818	0,1039	DOBLE LOGÍSTICA
3	20	99 %	0,1717	0,4000	0,1567	0,1027	0,0590	0,0680	DOBLE LOGÍSTICA
3	21	100 %	0,1717	0,4000	0,1567	0,1027	0,0590	0,0680	DOBLE LOGÍSTICA
3	23	98 %	0,1520	0,1500	0,0765	0,0707	0,0641	0,1280	DOBLE LOGÍSTICA
3	30	80 %	0,1724	0,1508	0,0768	0,0575	0,0752	0,1458	POLINOMIO
4	4	100 %	0,7671	0,2662	0,1138	0,0779	0,1081	0,1033	POLINOMIO
4	9	98 %	0,8586	0,2922	0,1221	0,0974	0,1225	0,0675	SIMULADOR
4	10	77 %	0,8586	0,2922	0,1221	0,0974	0,1225	0,0675	SIMULADOR
4	14	98 %	0,8611	0,2927	0,1223	0,1081	0,1157	0,1165	POLINOMIO
4	19	100 %	0,5133	0,1500	0,0765	0,0921	0,0834	0,0610	SIMULADOR
4	20	96 %	0,5133	0,1500	0,0765	0,0921	0,0834	0,0610	SIMULADOR
4	23	80 %	0,5294	0,1500	0,0765	0,0916	0,0957	0,0963	POLINOMIO
4	24	80 %	0,5133	0,1500	0,0765	0,0921	0,0834	0,0610	SIMULADOR
4	25	80 %	0,3203	0,1500	0,0765	0,0997	0,0929	0,0743	SIMULADOR
5	2	90 %	0,1881	0,1830	0,0871	0,0753	0,0437	0,0715	DOBLE LOGÍSTICA
5	3	84 %	0,1881	0,1830	0,0871	0,0753	0,0437	0,0715	DOBLE LOGÍSTICA

Tabla 2.4: Resultados para la chacra Cufré

Los algoritmos *baseline* tienen un menor RMSE en 14 celdas del establecimiento, mientras el simulador de crecimiento de cultivos tiene un menor RMSE en 11 celdas. El promedio de RMSE para todas las celdas del establecimiento de los algoritmos *baseline* es de 0,0815; mientras el promedio de RMSE para todas las celdas del establecimiento del simulador de crecimiento de cultivos es de 0,0839.

Chacra II Tramonto

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
1	52	86 %	0,8503	0,2314	0,1026	0,1503	0,0912	0,1045	DOBLE LOGÍSTICA
1	53	100 %	0,8422	0,2442	0,1067	0,1729	0,1678	0,0852	SIMULADOR
1	54	100 %	0,8998	0,2992	0,1244	0,1620	0,1049	0,0688	SIMULADOR

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
1	55	100 %	0,8998	0,2992	0,1244	0,1620	0,1049	0,0688	SIMULADOR
1	56	100 %	0,9095	0,2702	0,1150	0,1628	0,0666	0,0736	DOBLE LOGÍSTICA
1	57	100 %	0,9095	0,2702	0,1150	0,1628	0,0666	0,0736	DOBLE LOGÍSTICA
1	58	85 %	0,9915	0,2127	0,0966	0,1350	0,0524	0,0869	DOBLE LOGÍSTICA
1	60	100 %	0,9148	0,2722	0,1157	0,1257	0,1011	0,0990	SIMULADOR
1	66	100 %	0,8838	0,2735	0,1161	0,1377	0,0607	0,0795	DOBLE LOGÍSTICA
1	67	97 %	0,8786	0,2756	0,1168	0,1469	0,0877	0,0909	DOBLE LOGÍSTICA
1	68	77 %	0,8441	0,2230	0,0999	0,1428	0,1074	0,1004	SIMULADOR
1	69	86 %	0,8441	0,2230	0,0999	0,1428	0,1074	0,1004	SIMULADOR
1	70	100 %	0,9837	0,2034	0,0936	0,1425	0,0797	0,0931	DOBLE LOGÍSTICA
1	71	100 %	0,9816	0,2034	0,0936	0,1523	0,1317	0,0827	SIMULADOR
1	72	100 %	0,9816	0,2034	0,0936	0,1523	0,1317	0,0827	SIMULADOR
1	73	100 %	0,9882	0,2038	0,0938	0,1552	0,0900	0,0797	SIMULADOR
1	74	94 %	0,9967	0,2154	0,0975	0,1537	0,1292	0,0913	SIMULADOR
1	76	100 %	0,9087	0,2799	0,1182	0,1535	0,1212	0,0848	SIMULADOR
1	77	92 %	0,9087	0,2799	0,1182	0,1535	0,1212	0,0848	SIMULADOR
1	83	100 %	0,8921	0,2844	0,1196	0,1416	0,0820	0,0839	DOBLE LOGÍSTICA
1	84	87 %	0,8921	0,2844	0,1196	0,1416	0,0820	0,0839	DOBLE LOGÍSTICA
1	85	94 %	1,3600	0,1500	0,0765	0,1648	0,0821	0,0973	DOBLE LOGÍSTICA
1	86	92 %	1,3600	0,1500	0,0765	0,1648	0,0821	0,0973	DOBLE LOGÍSTICA
1	87	100 %	0,8261	0,2072	0,0949	0,1425	0,0912	0,0969	DOBLE LOGÍSTICA
1	88	100 %	0,9755	0,2082	0,0952	0,1552	0,0992	0,0963	SIMULADOR
1	89	100 %	0,9755	0,2082	0,0952	0,1552	0,0992	0,0963	SIMULADOR
1	90	100 %	0,9454	0,2198	0,0989	0,1646	0,1080	0,0815	SIMULADOR
1	92	100 %	0,9069	0,2975	0,1238	0,1507	0,0971	0,0829	SIMULADOR
1	93	100 %	0,9205	0,2767	0,1172	0,1504	0,1097	0,0842	SIMULADOR
1	100	77 %	1,3600	0,1500	0,0765	0,1648	0,0821	0,0973	DOBLE LOGÍSTICA
1	101	100 %	1,3600	0,1500	0,0765	0,1648	0,0821	0,0973	DOBLE LOGÍSTICA
1	102	96 %	0,8261	0,2072	0,0949	0,1425	0,0912	0,0969	DOBLE LOGÍSTICA
1	103	100 %	0,8261	0,2072	0,0949	0,1425	0,0912	0,0969	DOBLE LOGÍSTICA
1	104	100 %	0,9755	0,2082	0,0952	0,1552	0,0992	0,0963	SIMULADOR
1	105	100 %	0,9454	0,2198	0,0989	0,1646	0,1080	0,0815	SIMULADOR
1	106	82 %	0,9454	0,2198	0,0989	0,1646	0,1080	0,0815	SIMULADOR
1	108	100 %	0,9205	0,2767	0,1172	0,1504	0,1097	0,0842	SIMULADOR
1	109	98 %	0,9205	0,2767	0,1172	0,1504	0,1097	0,0842	SIMULADOR
1	117	100 %	0,6755	0,3073	0,1270	0,1831	0,0904	0,0959	DOBLE LOGÍSTICA
1	118	100 %	0,6755	0,3073	0,1270	0,1831	0,0904	0,0959	DOBLE LOGÍSTICA
1	119	100 %	0,8655	0,2815	0,1187	0,1880	0,1008	0,0993	SIMULADOR
1	120	100 %	0,8904	0,2714	0,1154	0,1597	0,1335	0,0964	SIMULADOR
1	121	100 %	0,8904	0,2714	0,1154	0,1597	0,1335	0,0964	SIMULADOR
1	122	87 %	0,9080	0,2973	0,1237	0,1583	0,0982	0,0937	SIMULADOR
1	133	85 %	0,8904	0,2901	0,1214	0,1467	0,0973	0,0987	DOBLE LOGÍSTICA
1	134	100 %	0,4970	0,3978	0,1560	0,1541	0,1258	0,1168	SIMULADOR
1	135	100 %	0,4970	0,3978	0,1560	0,1541	0,1258	0,1168	SIMULADOR
1	136	100 %	0,5019	0,3921	0,1542	0,1733	0,1255	0,1137	SIMULADOR
1	137	100 %	0,6717	0,3277	0,1335	0,1525	0,0928	0,0914	SIMULADOR
1	138	93 %	0,6717	0,3277	0,1335	0,1525	0,0928	0,0914	SIMULADOR
1	153	100 %	0,6555	0,3088	0,1275	0,1674	0,1117	0,1021	SIMULADOR
1	167	97 %	0,7596	0,2617	0,1123	0,1428	0,0976	0,1050	DOBLE LOGÍSTICA

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
1	168	77 %	0,6555	0,3088	0,1275	0,1674	0,1117	0,1021	SIMULADOR
1	170	93 %	0,8018	0,2568	0,1108	0,1341	0,0969	0,0908	SIMULADOR
1	171	93 %	0,6749	0,2914	0,1219	0,1252	0,1184	0,1413	DOBLE LOGÍSTICA
1	184	99 %	0,7729	0,2770	0,1173	0,1742	0,1240	0,1116	SIMULADOR
1	185	99 %	0,8521	0,2583	0,1112	0,1581	0,1247	0,0958	SIMULADOR
1	186	100 %	0,8557	0,2411	0,1057	0,1777	0,0985	0,0986	DOBLE LOGÍSTICA
1	187	99 %	0,8557	0,2411	0,1057	0,1777	0,0985	0,0986	DOBLE LOGÍSTICA
1	201	100 %	0,8978	0,2277	0,1014	0,1534	0,1974	0,1080	SIMULADOR
1	202	97 %	0,8668	0,2759	0,1169	0,1786	0,1676	0,1054	SIMULADOR
1	203	100 %	0,8481	0,2151	0,0974	0,1569	0,1286	0,0981	SIMULADOR
1	218	100 %	0,8711	0,2820	0,1189	0,1447	0,1616	0,1017	SIMULADOR
1	219	100 %	0,8830	0,2745	0,1164	0,1646	0,0902	0,0851	SIMULADOR
1	234	77 %	0,8830	0,2745	0,1164	0,1646	0,0902	0,0851	SIMULADOR
1	235	78 %	0,8830	0,2745	0,1164	0,1646	0,0902	0,0851	SIMULADOR
1	251	78 %	0,8898	0,2847	0,1197	0,1539	0,1633	0,1007	SIMULADOR
1	252	91 %	0,8898	0,2847	0,1197	0,1539	0,1633	0,1007	SIMULADOR
1	268	90 %	0,9060	0,2815	0,1187	0,1753	0,1440	0,0934	SIMULADOR
1	270	90 %	0,9579	0,2913	0,1218	0,1533	0,0980	0,0925	SIMULADOR
1	272	97 %	0,9633	0,2822	0,1189	0,1721	0,1099	0,0978	SIMULADOR
1	285	82 %	0,9082	0,2750	0,1166	0,1925	0,1748	0,0890	SIMULADOR
1	286	100 %	0,9361	0,2714	0,1155	0,1859	0,1420	0,0906	SIMULADOR
1	287	88 %	0,9361	0,2714	0,1155	0,1859	0,1420	0,0906	SIMULADOR
1	288	76 %	0,8738	0,2790	0,1179	0,1938	0,1395	0,0963	SIMULADOR
1	302	97 %	0,9361	0,2714	0,1155	0,1859	0,1420	0,0906	SIMULADOR
1	303	100 %	0,8738	0,2790	0,1179	0,1938	0,1395	0,0963	SIMULADOR
1	304	77 %	0,8738	0,2790	0,1179	0,1938	0,1395	0,0963	SIMULADOR
1	319	99 %	0,8387	0,2343	0,1035	0,1829	0,1379	0,1035	SIMULADOR

Tabla 2.5: Resultados para la chacra Il Tramonto

Los algoritmos *baseline* tienen un menor RMSE en 23 celdas del establecimiento, mientras el simulador de crecimiento de cultivos tiene un menor RMSE en 56 celdas. El promedio de RMSE para todas las celdas del establecimiento de los algoritmos *baseline* es de 0,1102; mientras el promedio de RMSE para todas las celdas del establecimiento del simulador de crecimiento de cultivos es de 0,0937.

Chacra Pochellu

A continuación se muestran los resultados obtenidos de la simulación y de los algoritmos *baseline* para el establecimiento Pochellu.

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
1	3	86 %	1,3352	0,4000	0,1567	0,2142	0,1782	0,0797	SIMULADOR
1	6	100 %	1,9372	0,4000	0,1567	0,2138	0,0655	0,0781	DOBLE LOGÍSTICA
1	7	80 %	1,9372	0,4000	0,1567	0,2138	0,0655	0,0781	DOBLE LOGÍSTICA
1	10	99 %	1,9560	0,4000	0,1567	0,2290	0,0637	0,0643	DOBLE LOGÍSTICA
1	11	100 %	1,9560	0,4000	0,1567	0,2290	0,0637	0,0643	DOBLE LOGÍSTICA
1	12	100 %	1,9191	0,4000	0,1567	0,2967	0,1285	0,0619	SIMULADOR
1	13	85 %	2,4729	0,3923	0,1542	0,2621	0,0490	0,0652	DOBLE LOGÍSTICA
1	14	79 %	2,4729	0,3923	0,1542	0,2621	0,0490	0,0652	DOBLE LOGÍSTICA
1	15	100 %	2,4664	0,3931	0,1545	0,2776	0,0553	0,0811	DOBLE LOGÍSTICA
1	16	93 %	2,4664	0,3931	0,1545	0,2776	0,0553	0,0811	DOBLE LOGÍSTICA
1	18	100 %	1,0588	0,3261	0,1330	0,2675	0,1261	0,0920	SIMULADOR
1	19	100 %	1,9119	0,1922	0,0901	0,2689	0,1349	0,0555	SIMULADOR
1	20	100 %	1,9119	0,1922	0,0901	0,2689	0,1349	0,0555	SIMULADOR
1	21	100 %	1,9511	0,4000	0,1567	0,3288	0,1269	0,0539	SIMULADOR
1	23	94 %	1,9537	0,4000	0,1567	0,2907	0,1234	0,0598	SIMULADOR
1	26	95 %	1,9119	0,1922	0,0901	0,2689	0,1349	0,0555	SIMULADOR
1	27	94 %	1,9119	0,1922	0,0901	0,2689	0,1349	0,0555	SIMULADOR
2	4	80 %	1,9613	0,4000	0,1567	0,0857	0,0702	0,0598	SIMULADOR
2	5	99 %	1,9396	0,4000	0,1567	0,0871	0,0806	0,0848	DOBLE LOGÍSTICA
2	6	100 %	1,9396	0,4000	0,1567	0,0871	0,0806	0,0848	DOBLE LOGÍSTICA
2	14	97 %	1,9577	0,4000	0,1567	0,0787	0,0558	0,0698	DOBLE LOGÍSTICA
2	15	100 %	1,9577	0,4000	0,1567	0,0787	0,0558	0,0698	DOBLE LOGÍSTICA
2	16	100 %	1,9526	0,4000	0,1567	0,0843	0,0579	0,0505	SIMULADOR
2	17	100 %	1,9526	0,4000	0,1567	0,0843	0,0579	0,0505	SIMULADOR
2	18	100 %	1,9574	0,4000	0,1567	0,0845	0,0649	0,0494	SIMULADOR
2	19	100 %	1,9452	0,4000	0,1567	0,0735	0,0660	0,0768	DOBLE LOGÍSTICA
2	20	94 %	1,9452	0,4000	0,1567	0,0735	0,0660	0,0768	DOBLE LOGÍSTICA
2	21	79 %	1,9477	0,4000	0,1567	0,0675	0,0851	0,1115	POLINOMIO
2	27	84 %	1,9631	0,4000	0,1567	0,0971	0,1049	0,0838	SIMULADOR
2	28	100 %	1,9374	0,4000	0,1567	0,0832	0,0685	0,0563	SIMULADOR
2	29	100 %	1,9374	0,4000	0,1567	0,0832	0,0685	0,0563	SIMULADOR
2	30	100 %	1,9629	0,4000	0,1567	0,0906	0,0686	0,0540	SIMULADOR
2	31	99 %	1,9518	0,4000	0,1567	0,0891	0,0681	0,0637	SIMULADOR
2	32	100 %	1,9518	0,4000	0,1567	0,0891	0,0681	0,0637	SIMULADOR
2	35	91 %	1,9456	0,4000	0,1567	0,0529	0,0777	0,1216	POLINOMIO
2	41	100 %	0,8586	0,4000	0,1567	0,1025	0,0975	0,0963	SIMULADOR
2	42	98 %	1,9572	0,4000	0,1567	0,1001	0,0812	0,0576	SIMULADOR
2	49	83 %	1,8615	0,4000	0,1567	0,0744	0,0660	0,1011	DOBLE LOGÍSTICA
2	61	94 %	1,8615	0,4000	0,1567	0,0744	0,0686	0,1011	DOBLE LOGÍSTICA
2	63	89 %	1,9370	0,4000	0,1567	0,0780	0,0950	0,1056	POLINOMIO
2	75	77 %	1,8710	0,4000	0,1567	0,0506	0,0681	0,0994	POLINOMIO
2	77	82 %	1,9603	0,4000	0,1567	0,0713	0,0942	0,0652	SIMULADOR
2	88	98 %	1,2375	0,4000	0,1567	0,0555	0,0908	0,0948	POLINOMIO
2	89	76 %	1,2302	0,4000	0,1567	0,0714	0,1194	0,0768	POLINOMIO
3	5	83 %	1,3627	0,4000	0,1567	0,2212	0,1256	0,0565	SIMULADOR
3	6	98 %	1,3627	0,4000	0,1567	0,2212	0,1256	0,0565	SIMULADOR
3	10	88 %	1,3345	0,4000	0,1567	0,2257	0,0898	0,0574	SIMULADOR
3	11	98 %	1,2975	0,4000	0,1567	0,3078	0,0770	0,0550	SIMULADOR
3	12	100 %	1,2975	0,4000	0,1567	0,3078	0,0770	0,0550	SIMULADOR

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
3	13	100 %	1,3627	0,4000	0,1567	0,2212	0,1256	0,0565	SIMULADOR
3	14	100 %	1,0012	0,4000	0,1567	0,2212	0,0924	0,0907	SIMULADOR
3	18	84 %	1,2146	0,4000	0,1567	0,1876	0,0678	0,0654	SIMULADOR
3	19	100 %	1,2146	0,4000	0,1567	0,1876	0,0678	0,0654	SIMULADOR
3	20	100 %	1,3400	0,4000	0,1567	0,2715	0,0485	0,0582	DOBLE LOGÍSTICA
3	21	100 %	1,3400	0,4000	0,1567	0,2715	0,0485	0,0582	DOBLE LOGÍSTICA
3	22	100 %	1,3515	0,4000	0,1567	0,2517	0,0650	0,0713	DOBLE LOGÍSTICA
3	23	85 %	1,0004	0,4000	0,1567	0,2497	0,0782	0,0866	DOBLE LOGÍSTICA
3	27	79 %	1,0057	0,3792	0,1500	0,2831	0,0939	0,1149	DOBLE LOGÍSTICA
3	28	100 %	1,0057	0,3792	0,1500	0,2831	0,0939	0,1149	DOBLE LOGÍSTICA
3	29	100 %	1,1635	0,3682	0,1465	0,2333	0,0982	0,0642	SIMULADOR
3	30	100 %	0,9992	0,4000	0,1567	0,2178	0,1845	0,0608	SIMULADOR
3	31	100 %	0,9992	0,4000	0,1567	0,2178	0,1845	0,0608	SIMULADOR

Tabla 2.6: Resultados para la chacra Pochellu

Los algoritmos *baseline* tienen un menor RMSE en 26 celdas del establecimiento, mientras el simulador de crecimiento de cultivos tiene un menor RMSE en 23 celdas. El promedio de RMSE para todas las celdas del establecimiento de los algoritmos *baseline* es de 0,0864; mientras el promedio de RMSE para todas las celdas del establecimiento del simulador de crecimiento de cultivos es de 0,0928.

Chacra Trealfa

A continuación se muestran los resultados obtenidos de la simulación y de los algoritmos *baseline* para el establecimiento Trealfa.

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
1	2	88 %	0,1656	0,1501	0,0766	0,1196	NO CONV.	0,1249	POLINOMIO
1	3	100 %	0,1733	0,1563	0,0785	0,0921	0,1105	0,1153	POLINOMIO
1	6	100 %	0,2460	0,1500	0,0765	0,1057	0,1327	0,0790	SIMULADOR
1	7	79 %	0,8586	0,3041	0,1259	0,0913	0,0759	0,0627	SIMULADOR
1	10	100 %	0,8037	0,4000	0,1567	0,0930	0,0778	0,0818	DOBLE LOGÍSTICA
2	7	95 %	0,8586	0,2966	0,1235	0,0998	0,1162	0,0857	SIMULADOR
2	8	96 %	0,8586	0,3063	0,1267	0,1037	0,0740	0,0739	SIMULADOR
2	9	84 %	0,8586	0,3063	0,1267	0,1037	0,0740	0,0739	SIMULADOR
3	5	82 %	0,9216	0,1751	0,0845	0,1259	0,1688	0,0938	SIMULADOR
4	1	92 %	1,9270	0,3997	0,1566	0,0897	0,0737	0,0999	DOBLE LOGÍSTICA
4	2	100 %	1,9631	0,3926	0,1543	0,1088	0,0904	0,1042	DOBLE LOGÍSTICA
4	3	97 %	1,9631	0,3926	0,1543	0,1088	0,0904	0,1042	DOBLE LOGÍSTICA
4	4	100 %	0,8586	0,3017	0,1252	0,1060	0,0770	0,1088	DOBLE LOGÍSTICA

Lote	Celda	% en chacra	rdmax	fc	pwp	RMSE Polinomio	RMSE Doble Logística	RMSE Simulador	Mejor Estimador
4	5	97 %	0,7668	0,3998	0,1566	0,1017	NO CONV.	0,0902	SIMULADOR
4	7	93 %	0,7668	0,3998	0,1566	0,1017	NO CONV.	0,0902	SIMULADOR
5	10	100 %	0,8586	0,3127	0,1287	0,1028	0,0671	0,1073	DOBLE LOGÍSTICA
5	14	95 %	0,8586	0,3127	0,1287	0,1028	0,0671	0,1073	DOBLE LOGÍSTICA
5	15	77 %	0,8995	0,2475	0,1078	0,0957	0,0796	0,1106	DOBLE LOGÍSTICA
5	18	100 %	0,7638	0,3782	0,1497	0,1058	0,0897	0,1016	DOBLE LOGÍSTICA
5	19	100 %	0,8586	0,3053	0,1263	0,0932	0,0709	0,1051	DOBLE LOGÍSTICA
5	22	100 %	0,8586	0,4000	0,1567	0,0891	0,0503	0,0842	DOBLE LOGÍSTICA
5	23	100 %	0,8586	0,4000	0,1567	0,0891	0,0503	0,0842	DOBLE LOGÍSTICA
5	24	100 %	0,7702	0,4000	0,1567	0,1047	0,0673	0,0937	DOBLE LOGÍSTICA
5	27	100 %	1,8781	0,4000	0,1567	0,0859	0,1300	0,0837	SIMULADOR
5	28	100 %	0,8586	0,4000	0,1567	0,0901	0,0655	0,0707	DOBLE LOGÍSTICA
5	29	98 %	0,8586	0,4000	0,1567	0,0901	0,0655	0,0707	DOBLE LOGÍSTICA
5	30	92 %	0,7681	0,4000	0,1567	0,1009	0,1470	0,0974	SIMULADOR
5	32	100 %	0,8586	0,4000	0,1567	0,0901	0,0655	0,0707	DOBLE LOGÍSTICA
5	33	100 %	0,8586	0,4000	0,1567	0,0901	0,0655	0,0707	DOBLE LOGÍSTICA
5	34	100 %	0,7681	0,4000	0,1567	0,1009	0,1470	0,0974	SIMULADOR
5	37	99 %	1,9396	0,3999	0,1567	0,0822	0,0486	0,0839	DOBLE LOGÍSTICA
5	38	100 %	1,9602	0,4000	0,1567	0,1072	0,0957	0,1156	DOBLE LOGÍSTICA
5	39	83 %	1,9602	0,4000	0,1567	0,1072	0,0957	0,1156	DOBLE LOGÍSTICA
5	42	97 %	1,9612	0,3984	0,1562	0,0942	0,0663	0,0798	DOBLE LOGÍSTICA
5	43	91 %	1,9626	0,3965	0,1556	0,0954	0,2995	0,1024	POLINOMIO
9	3	79 %	0,7558	0,2899	0,1214	0,0970	0,1209	0,0944	SIMULADOR
10	4	91 %	0,7551	0,2993	0,1244	0,0928	0,0937	0,0892	SIMULADOR
10	7	89 %	0,2722	0,1581	0,0791	0,1148	0,0981	0,1337	DOBLE LOGÍSTICA
10	8	100 %	0,7691	0,4000	0,1567	0,1048	0,1190	0,0833	SIMULADOR
10	9	100 %	1,0871	0,4000	0,1567	0,1024	0,0981	0,0916	SIMULADOR
10	12	100 %	0,8586	0,3028	0,1255	0,1248	0,1092	0,0879	SIMULADOR
10	13	100 %	0,8586	0,3028	0,1255	0,1248	0,1092	0,0879	SIMULADOR
10	14	88 %	0,8604	0,3120	0,1285	0,1123	0,0837	0,0740	SIMULADOR
10	18	97 %	0,8604	0,3120	0,1285	0,1123	0,0837	0,0740	SIMULADOR
7-8	5	88 %	1,9434	0,3745	0,1485	0,1036	0,1691	0,0958	SIMULADOR
7-8	6	95 %	1,9434	0,3745	0,1485	0,1036	0,1691	0,0958	SIMULADOR
7-8	8	100 %	1,9434	0,3745	0,1485	0,1036	0,1691	0,0958	SIMULADOR
7-8	9	94 %	0,9761	0,1768	0,0851	0,1044	0,1333	0,0988	SIMULADOR
7-8	11	84 %	0,1924	0,2285	0,1017	0,0937	0,0686	0,0719	DOBLE LOGÍSTICA

Tabla 2.7: Resultados para la chacra Trealfa

Los algoritmos *baseline* tienen un menor RMSE en 24 celdas del establecimiento, mientras el simulador de crecimiento de cultivos tiene un menor RMSE en 25 celdas. El promedio de RMSE para todas las celdas del establecimiento los algoritmos *baseline* es de 0,0889, mientras el promedio de RMSE para todas las celdas del establecimiento del simulador de crecimiento de cultivos es de 0,0928.