

DEFINICIÓN DE ZONAS HOMOGÉNEAS DE DAÑO CAUSADO POR GRANIZO EN CULTIVOS AGRÍCOLAS UTILIZANDO DATOS DE SENSORES REMOTOS

DEFINITION OF HOMOGENOUS DAMAGE ZONES CAUSED BY HAIL IN AGRICULTURAL CROPS USING REMOTE SENSING DATA

Leandro Sosa-Avaro^{a,*}, Ana Justel^b, Iñigo Molina^a

^a Departamento de Ingeniería Topográfica y Cartográfica, Escuela Técnica Superior de Ingenieros en Topografía, Geodesia y Cartografía, Universidad Politécnica de Madrid, Campus Sur, A-3, Km 7, 28031 Madrid, España. leandroleonelsosa@gmail.com; inigo.molina@upm.es

^b Departamento de Matemáticas, Facultad de Ciencias, Universidad Autónoma de Madrid, Campus de Cantoblanco, Madrid 28049, España. ana.justel@uam.es

Abstract:

The frequency and intensity of extreme worldwide weather events have increased in recent decades, causing enormous economic losses. In Argentina, in the 2017-2018 agricultural season, more than 200 million dollars were allocated to protect soy, corn and wheat crops against hail. After a storm, a field survey by an expert is required to estimate the yield losses caused by hail, without prior information on the affected regions or their intensity. The precision in the estimation of the damage depends to a great extent on the identification of Homogeneous Damage Zones (HDZ) within the plot, which is further used to quantify the total damage. Currently, HDZ are delimited using in-situ visual techniques. This research suggests developing an algorithm to define the HDZ applying Machine Learning techniques to vegetation indices derived from Sentinel 1 and 2 data. For this purpose, 5 microwave indices (DPDD, IDPDD, VDDPI, MPDI and DPSVI) and 5 spectral indices (NDVI, EVI, SAVI, AVI and NPCRI) were tested. The most sensitive indices to changes were selected, for both microwave and optical signals, and were in turn integrated into the damage detection model. A K-Means ($k = 3$) machine learning algorithm was used to define the classes. To validate the algorithm, 38 storms that occurred between 2017 and 2020 were analyzed in 91 soybean, wheat and corn plots located in the Argentine Pampean plain. The One-Way ANOVA model ($p < 0.05$) was applied to each plot. The selected indices were DPSVI and NPCRI. HDZ were correctly detected in 66.67%, 78.13% and 72.70% of the analyzed cases, for corn, wheat and soybean crops, respectively. It is concluded that the designed algorithm allows defining efficiently HDZ caused by hail, giving transparency and precision to the work of the expert and reducing time consuming field surveys.

Key words: remote sensing, machine learning, algorithm, hail, agriculture

Resumen:

La frecuencia e intensidad de los eventos meteorológicos extremos, a nivel mundial, se han incrementado en las últimas décadas, provocando enormes pérdidas económicas. En Argentina, en la campaña agrícola 2017-2018, se destinaron más de 200 millones de dólares para proteger los cultivos soja, maíz y trigo contra granizo. Luego de una tormenta, un perito visita el campo para estimar las mermas de rendimiento causadas por granizo, sin información previa de las regiones afectadas ni su intensidad. La precisión en la estimación del daño depende en gran medida de la identificación de Zonas Homogéneas de Daños (ZHD) dentro de la parcela para ponderar el daño total. Actualmente, las ZHD se delimitan a campo con técnicas visuales. Se propone desarrollar un algoritmo para definir las ZHD aplicando técnicas *Machine Learning* a índices de vegetación calculados con datos *Sentinel 1* y *2*. Se procesaron y compararon 5 índices de microondas (DPDD, IDPDD, VDDPI, MPDI y DPSVI) y 5 espectrales (NDVI, EVI, SAVI, AVI y NPCRI) y se seleccionó el más sensible a los cambios para cada tipo de señal; además, se incorporaron como variable de entrada al modelo las derivadas de ambos índices. Para definir las clases se empleó *K-Means* ($k = 3$). Para validar el algoritmo se analizaron 38 tormentas ocurridas entre los años 2017 y 2020 en 91 parcelas de soja, trigo y maíz ubicadas en la llanura pampeana argentina. Se aplicó a cada parcela el modelo *One-Way ANOVA* ($p < 0.05$). Los índices seleccionados fueron DPSVI y NPCRI. Se detectaron correctamente ZHD en un 66,67%, 78,13% y 72,70% de los casos analizados, para los cultivos de maíz, trigo y soja, respectivamente. Se concluye que el algoritmo permite definir en forma eficiente ZHD causados por granizo dando transparencia y precisión a la labor del perito y disminuyendo el tiempo de sus tareas a campo.

Palabras clave: sensores remotos, aprendizaje automático, algoritmo, granizo, agricultura

* Corresponding Author: Leandro Sosa-Avaro, leandroleonelsosa@gmail.com

1. Introducción

La frecuencia e intensidad de los eventos meteorológicos extremos, a nivel mundial, se han incrementado en las últimas décadas, provocando enormes pérdidas económicas (Halsnæs *et al.* 2018). Los productores agropecuarios atenúan las consecuencias de las catástrofes implementando estrategias de gestión dentro de su establecimiento pero ciertos riesgos, como el granizo, exceden su capacidad financiera de actuación. En estos casos, transfieren el riesgo a empresas especializadas en administrarlos; a través de seguros agropecuarios. Cuando ocurre una tormenta de granizo, en parcelas aseguradas, un perito realiza una estimación del daño a campo. La precisión de la evaluación depende en gran medida de la detección de Zonas Homogéneas de Daños (ZHD) que permiten extrapolar los datos evaluados en las estaciones de muestreo, a la superficie total de la parcela dañada. Actualmente, las superficies de estas zonas se estiman con técnicas visuales y el responsable de la evaluación no posee ningún conocimiento de la región afectada dentro de la parcela, ni de la intensidad del daño, al momento de llegar al sitio. Una alternativa es utilizar sensores remotos para detectar las ZHD.

En la presente investigación se utilizaron sensores activos de microondas, específicamente Radars de Apertura Sintética (SAR) y pasivos, del tipo multiespectral para definir en forma automática las ZHD. La señal de microondas es sensible a los cambios producidos en la arquitectura del dosel y en su configuración geométrica, tanto por el normal crecimiento de las plantas, como por la ocurrencia de una tormenta de granizo. Diversos autores (Huang *et al.* 2016; Kim *et al.* 2012; Kim and van Zyl 2009; Periasamy 2018; Sahadevan *et al.* 2013) proponen que es posible utilizar índices calculados con datos SAR, para estimar parámetros biofísicos del cultivo, suavizando las características no deseadas contenidas en la señal retrodispersada. Por su lado, los sensores ópticos captan la luz solar reflejada por los objetos y pueden identificar características de los cultivos con alta precisión. Mientras que la atmósfera no interfiere la señal microonda, en las imágenes ópticas si es importante corregir la presencia de brillos solares intermedios y detectar nubes y sus sombras. Las ventajas propias de cada sensor permiten inferir que la utilización combinada de ambos tipos de señales es lo más adecuado para detectar cambios en parámetros biofísicos de los cultivos.

A través de una formulación aritmética que integren dos o más rangos de longitudes de ondas en la señal multiespectral o tipos de polarizaciones en la señal de microondas, se generan índices de vegetación para describir variables biofísicas de los cultivos. La principal ventaja de emplear índices de vegetación es su simplicidad y su principal crítica es que se utiliza un número reducido de variables que no suelen reflejar la complejidad de las condiciones reales (Verrelst *et al.* 2015). Para minimizar esta desventaja es preciso seleccionar los índices más sensibles a los cambios que manifieste la variable de interés permitiendo discriminar grupos de plantas que presentan diferente grado de alteración.

A pesar de existir más de 30 años de datos satelitales de observación de la Tierra, hasta donde llega nuestro conocimiento, no existen investigaciones que usen datos SAR para la caracterización del daño que produce el

granizo en cultivos. Con datos multiespectrales se han analizado algunos casos locales y esporádicos (Bentley *et al.* 2002; Capellades *et al.* 2009; Chandler *et al.* 2003) que presentan limitaciones prácticas para su aplicación regional y operativa (De Leeuw *et al.* 2014).

En este marco, desarrollar un algoritmo que permita definir ZHD en cultivos agrícolas afectados por el fenómeno meteorológico granizo siguiendo una metodología estándar y precisa, aplicada a datos satelitales, supone una clara contribución aportando objetividad y transparencia al proceso de evaluación del daño.

2. Materiales y métodos

2.1. Area de estudio

El estudio se realizó en la llanura Pampeana. El clima de esta región es templado al sur y subtropical húmedo al norte. La temperatura media anual es de 17 grados y el régimen de precipitaciones, monzónico. La llanura contiene las provincias argentinas: Buenos Aires, Entre Ríos, Santa Fe, Córdoba, La Pampa y San Luis. En la actualidad, es la región más industrializada del país, traccionada principalmente por la producción de soja, maíz y trigo, y sus derivados. Dentro de ella, las regiones de mayor ocurrencia de granizo se ubican en las cercanías de las Sierras de Córdoba, Tandil y la Ventana; y en la región costera (Sierra *et al.* 1993). Históricamente la intensidad del daño es mayor en cultivos de invierno como el trigo debido a que las tormentas intensas ocurren principalmente a final de primavera y comienzos de verano, momento en que el trigo se encuentra con sus granos formados y expuestos (Saluzzi and Nuñez 1975). La superficie sembrada en la campaña agrícola 2019-2020 fue de 16,9, 9,5 y 6,9 millones de hectáreas para los cultivos soja, maíz y trigo (MAGYP 2021). Para la misma campaña, la prima emitida por todo el mercado asegurador en Argentina, para asegurar los cultivos de soja, maíz y trigo contra el riesgo granizo, fue de 230 millones de US\$ (ME 2021).

2.2. Datos y preprocesamientos

En el presente estudio se emplearon datos *Sentinel 1* y 2 del programa *Copernicus* (Agencia Espacial Europea y Agencia Europea de Medio Ambiente, UE) y evaluaciones de daño realizadas in situ en 44 parcelas de soja, 32 de trigo y 15 de maíz, en tormentas de granizo ocurridas entre los años 2017 y 2020.

De la misión satelital *Sentinel 1* se utilizaron imágenes *Ground Range Detected* procesadas a partir de productos *Single Look Complex* utilizando *Sentinel-1 Toolbox*. En la mayor parte del territorio argentino existe solapamiento de las escenas observadas por los satélites *Sentinel 1A* y *1B* y sus ángulos de incidencia difieren considerablemente en un mismo punto de observación. Para evitar esto, cuando se solaparon los datos de las plataformas A y B, se seleccionó aquella que primero capturó datos después de la tormenta de granizo. La resolución espacial del producto utilizado es de 10 metros. De la misión *Sentinel 2*, se utilizaron productos con nivel de procesamiento 1C; sus datos representan la reflectancia en la parte superior de la atmósfera. La resolución temporal de *Sentinel 2* es de 5 días o de 2 y 3 días en sectores con solapamiento de las órbitas. La

resolución espacial de las bandas utilizadas es de 10 metros.

Para cada parcela evaluada se emplean los datos satelitales registrados por ambas misiones dentro de los límites de su geometría y de una ventana temporal de 60 días antes y después de la fecha de ocurrencia de la tormenta. Cuando la ventana incluye alguna fecha anterior a la siembra del cultivo o finaliza después de la cosecha, la serie se trunca para evitar incorporar información de otros cultivos. Los datos satelitales fueron consultados y procesados en la plataforma *Google Earth Engine*.

Para evaluar el daño in situ sobre cada cultivo se siguieron las normas establecidas por el United States Department of Agriculture y la Federal Crop Insurance Corporation (USDA and FCIC, 2019a, 2019b, 2016). En cada una de las 91 parcelas estudiadas se realizaron como mínimo diez estaciones de muestreo, alcanzando en total 1174 observaciones (646 en soja, 153 en maíz y 375 en trigo). Para cada estación de muestreo se documentó la latitud, longitud y el porcentaje de daño estimado.

Para eliminar el moteado de la señal de microondas se utilizaron filtros morfológicos, calculando para cada píxel la mediana sobre un kernel circular de 15 píxeles de radio (Mlynarczuk and Porzycka-Strzelczyk 2018). Para filtrar las nubes y sus sombras y corregir otras interferencias de la atmósfera, optamos por un método temporal de ejecución rápida como es el de regresiones lineales móviles, que se aplica de manera independiente a cada serie temporal de los píxeles (El Hajj *et al.* 2008). En las regresiones, se utilizó una ventana temporal móvil de 15 días para garantizar la eliminación de las interferencias, pero sin perjudicar los valores por degradación de los datos a causa de un excesivo suavizado.

2.3. Selección de índices de vegetación

Como base para definir las variables que se emplean en la clasificación de píxeles en ZHD se consideran los índices de vegetación más informativos sobre el estado de los cultivos que se pueden calcular con datos procedentes de las misiones satelitales *Sentinel 1* y *2*. Se seleccionó uno por cada tipo de señal y se los denominó IS_1 e IS_2 .

Se consideraron los siguientes índices de microondas:

- Distancia Diagonal de Polarización Dual (DPDD) (Periasamy 2018): representa suelo sin vegetación con valores pequeños y, a medida que aumenta la biomasa, aumenta el valor del índice hasta alcanzar valores máximos en presencia de vegetación muy abundante.
- Distancia Diagonal Inversa de Polarización Dual (IDPDD) (Periasamy 2018): separa mejor que DPDD los píxeles con vegetación de los que tienen suelo descubierto, pero presenta un cierto grado de incertidumbre para diferenciar vegetación y cuerpos de agua.
- Índice de Despolarización Dual Vertical (VDDPI) (Periasamy 2018): Este índice cuantifica la biomasa midiendo la relación entre la potencia polarizada total y la copolarizada.

- Índice Diferenciado de Polarización Microonda (MPDI) (Becker and Choudhury 1988): Este índice es especialmente sensible a la rugosidad de la superficie y la humedad del suelo, características que se alteran luego de una tormenta de granizo.
- *Dual Polarization SAR Vegetation Index* (DPSVI) (Periasamy 2018): este índice es una combinación de VDDPI y IDPDD, que conserva las ventajas de ambos índices, al tiempo que separa la vegetación de los cuerpos de agua.

Entre los índices espectrales se consideraron:

- Índice de Vegetación de Diferencia Normalizada (NDVI) (Gitelson and Merzlyak 1997): caracteriza la cantidad de clorofila que contiene la vegetación. Valores altos de NDVI corresponden a áreas que reflejan más en el espectro del infrarrojo cercano y esto se corresponde con una vegetación más densa y saludable. NDVI se satura en condiciones de alta biomasa y es muy sensible a las variaciones de fondo del dosel (Huete 1988).
- Índice de Vegetación Mejorado (EVI) (Liu and Huete 1995): EVI tiene mejor comportamiento que NDVI en situaciones con alto rendimiento de biomasa y responde mejor a las variaciones estructurales del dosel tales como: el índice de área foliar, el tipo de dosel, la fisonomía de la planta o la arquitectura del dosel (Gao *et al.* 2000). EVI requiere una correcta y minuciosa eliminación previa del ruido externo y del sensor para estimar reflectancias de superficie (Liu and Huete 1995).
- Índice de Vegetación Ajustado al Suelo (SAVI) (Huete 1988): deriva de la ecuación de NDVI y agrega un coeficiente para corregir la influencia del brillo del suelo en áreas donde la cobertura vegetativa es baja. Esta corrección es importante en doseles abiertos donde la señal de fondo ejerce efecto sobre las reflectancia registradas.
- Índice de vegetación avanzado (AVI) (Roy *et al.* 1996): utiliza, al igual que NDVI las reflectancia del rojo e infrarrojo cercano. Sin embargo, al utilizar un grado de potencia de la respuesta infrarroja es más sensible a las pequeñas diferencias de densidad del dosel (Roy *et al.* 1996).
- Índice de Proporción de Pigmentos de Clorofila Normalizado (NPCRI) (Merzlyak *et al.* 1999): Este índice es sensible a la relación de los contenidos de carotenoides y clorofilas presentes en la planta. Es especialmente útil para caracterizar la senescencia de las hojas y la maduración de los frutos. Además, es capaz de diferenciar cultivo senescente y suelo sin vegetación.

Para seleccionar los índices IS_1 e IS_2 se consideraron los datos de la tormenta de granizo del 10 de febrero de 2019, ocurrida en las cercanías de la localidad de Lozada, provincia de Córdoba, Argentina. El análisis se realizó en una parcela de 90 hectáreas de soja (en adelante denominada parcela modelo) en estado fenológico R3 según escala de Fehr (Fehr and Caviness 1977). Al momento de la tormenta, el cultivo cubría totalmente el suelo y presentaba una estructura homogénea en toda su superficie. Luego de la misma se

registraron regiones sin daño, otras con destrucción total del cultivo y otras con valores intermedios de daño.

Se seleccionaron todas las imágenes *Sentinel 1* y *2* que contenían la parcela modelo y fueron capturadas dentro de la ventana de tiempo que abarca desde el 12 de diciembre de 2018 (60 días antes de la tormenta) hasta el 11 de abril de 2019 (60 días después de la tormenta). Se recortó cada imagen utilizando la geometría de la parcela. Con las imágenes recortadas de la misión *Sentinel 1* se constituyó el conjunto de datos S1 y con las imágenes *Sentinel 2* el conjunto S2. Se pre-procesaron los conjuntos de datos según se describe en la sección 2.2. Para cada conjunto de datos se calcularon, por píxel de cada imagen, los 5 índices de vegetación correspondientes a cada señal. Se calcularon los desvíos estándares de los valores del índice por imagen. Luego se seleccionó por índice el valor máximo de desvío estándar antes de la tormenta (Max_{ant}) y el máximo después de la tormenta (Max_{post}). Se calcularon los cocientes de estos dos valores y se seleccionó el Is_1 y el Is_2 que obtuvo el menor valor. De esta manera, se seleccionó el índice de microondas y el índice espectral que presentó la menor dispersión antes de la tormenta respecto a la dispersión registrada luego de la misma.

2.4. Algoritmo de aprendizaje automático

Se propone un algoritmo de aprendizaje automático no supervisado para determinar las ZHD dentro de una misma parcela a partir de imágenes satelitales *Sentinel 1* y *2*. Los pasos que sigue el algoritmo son: (1) selección y pre-procesado de datos satelitales para eliminar el moteado y la presencia de nubes y otros fenómenos atmosféricos; (2) cálculo de los índices de vegetación más sensibles a los daños causados por el granizo como base para la construcción de la matriz de datos para la clasificación; (3) preparación de las variables de clasificación y construcción de la matriz de datos; y (4) clasificación no supervisada de píxeles en ZHD.

Se espera que el algoritmo clasifique los píxeles en ZHD. Los píxeles de la matriz de datos siempre pertenecen a una misma parcela y por tanto, representan una misma especie de cultivo y variedad, sembrada en la misma fecha. La clasificación de ZHD tendrá en cuenta la evolución del cultivo en el período previo y posterior a la tormenta. Las clases deberían agrupar plantas que se comportan de manera similar, para las variables biofísicas estimadas por los índices, antes y después de la tormenta.

Se realizó una interpolación lineal de los valores del índice espectral estimando su valor para cada fecha de la captura de las imágenes SAR. Esto permitió contener en la matriz la misma cantidad de datos provenientes de la señal microondas que espectrales, igualando así la representatividad de ambas señales. Además, de esta forma las imágenes multiespectrales tienen equidistancias temporales, permitiendo así calcular las tasas de cambios de los valores de los índices.

Se calcularon las tasas de cambio como las derivadas de los índices diferenciando la serie, restando a la observación de un momento la del instante anterior, en la misma ubicación. La derivada de la primera imagen no es posible calcularla debido a la ausencia de datos previos, por lo tanto se elimina de la matriz de datos luego del cálculo de las derivadas. Así, la matriz de datos para una

parcela tendrá tantas filas como píxeles y un número de variables que dependen del momento de la tormenta. El máximo número de variables serán 36 dadas por 4 variables (Is_1 , Is_2 y sus respectivas derivadas) con 9 observaciones (4 anteriores a la tormenta y 5 posteriores) cada una de ellas.

Para definir las ZHD se consideró adecuado utilizar K-means (Wong and Hartigan 1979) por ser un método simple, fácil de implementar, rápido y eficiente en términos de costo computacional para grandes conjuntos de datos (Arthur and Vassilvitskii 2006). Para la detección de ZHD se consideró apropiado agrupar en 3 clases los píxeles de las parcelas. Se utilizó la función de distancia Euclídea para medir las cercanías a centroides. En cualquier otra aplicación del algoritmo, este parámetro se puede modificar para integrar otros objetivos en la clasificación. El algoritmo se programó en Python 3, utilizando las librerías pandas, geopandas, ee, eeconvert y datetime.

2.5. Validación del algoritmo con datos in situ

Para validar la calidad de las agrupaciones que proporciona el algoritmo que hemos presentado se utiliza información externa, no empleada para establecer las ZHD en cada parcela. Aunque para ninguna de las parcelas que se han podido analizar se tiene la información real sobre la verdadera ZHD a la que pertenece cada uno de los píxeles, si disponemos de las mediciones del daño realizadas in situ en algunos puntos.

Para cada parcela se identificaron los píxeles con información *in situ*. Estos píxeles son una submuestra de las tres ZHD identificadas con el algoritmo, para la que se conoce el porcentaje de daño real medido. Para validar la calidad de las agrupaciones de los píxeles en ZHD, comparamos las medias de los porcentajes de daño calculadas con los datos de cada una de las tres zonas. Si el algoritmo es capaz de identificar correctamente las ZHD, el porcentaje de daño real medio debe ser diferente entre cada una de las zonas definidas. Se probó la homogeneidad entre ZHD utilizando test *one-way* ANOVA. Las diferencias se consideran estadísticamente significativas si el valor de $p < 0,05$.

3. Resultados y discusión

3.1. Selección de los índices de vegetación

Los cinco índices de microondas fueron sensibles a los cambios provocados por el granizo en los píxeles de la parcela. El índice de microondas seleccionado como atributo de clasificación (Is_1) fue DPSVI, que obtuvo el menor cociente (Max_{ant}/Max_{post}), igual a 0.1535. El modelo DPSVI se basa en el patrón de diagrama de dispersión construido entre el coeficiente de retrodispersión de las polarizaciones VV y VH y destaca el aporte de la polarización cruzada. Los 3 parámetros contemplados en la ecuación ($\sigma_{vv(i)}$, $\sigma_{vh(i)}$ y $\sigma_{vv(max)}$) son altamente influyentes en la caracterización de la biomasa (Periasamy 2018). Para un mismo cultivo, cuando la producción de biomasa aumenta, se obtienen mayores valores de retrodispersión en polarización cruzada (Chauhan and Srivastava 2016). Las señales de microondas de la banda C penetran a través de las hojas y experimentan una dispersión múltiple dentro de las

pequeñas ramas del cultivo, siendo eficiente la señal para estimar biomasa en cultivos (Ghasemi *et al.* 2011). La capacidad de penetración en la superficie de la frecuencia de la banda C es de alrededor de 1 m, y se satura a alrededor de 60–70 toneladas/ha (Nizalapur *et al.* 2010). Esto explicaría su buen comportamiento frente a los cambios generados por la precipitación de granizo en la parcela modelo.

Todos los índices espectrales, al igual que los de microondas, fueron sensibles al daño producido por la tormenta de granizo. Al comparar el cociente entre el desvío estándar máximo antes y después para cada índice espectral observamos que NPCRI tuvo un comportamiento notablemente superior al resto ($Max_{ant}/Max_{post} = 0.1483$). Esto se debería a su capacidad para diferenciar suelo sin cobertura, vegetación saludable y senescente. NPCRI es un cociente normalizado de los valores de reflectancia del rojo y el azul. El granizo cuando impacta con el dosel destruye hojas y tallos, provocando su senescencia. Los tejidos vegetales experimentan cambios de color como resultado de cambios en el contenido de pigmentos y sus proporciones relativas. Los cambios en el color de los tejidos de las plantas senescentes están relacionados principalmente con una creciente degradación de la clorofila y el aumento relativo de carotenoides. Exteriormente se visualiza un cambio de color de verde a amarillo primero y luego con la necrosis un viraje a marrón. Los carotenoides y las clorofilas exhiben una absorción fuerte en el rango espectral del azul, valores de reflectancia utilizados para el cálculo del NPCRI (Merzlyak *et al.* 1999). La capacidad del NPCRI para diferenciar estos componentes explicaría su mejor comportamiento.

3.2. Validación del algoritmo

Para las parcelas con cultivo de maíz se encontró que en el 66,67% de los casos (10 de 15 analizadas) existe evidencia significativa de que el porcentaje medio de daño no es igual en las tres ZHD encontradas con el algoritmo. Para el cultivo de trigo se halló que en un 78,13% de las parcelas (25 de 32 analizadas) el daño medio por ZHD no fue igual y, finalmente, para soja el porcentaje de parcelas fue de 72,70% (32 de 44 analizadas).

Distinguimos algunas situaciones particulares entre las parcelas que no se ha encontrado evidencia significativa de que los daños sufridos por los cultivos sean diferentes

en las tres ZHD definidas por el algoritmo. Se han identificado 6 parcelas con daños muy similares en toda su extensión. En estas situaciones asumir que hay 3 agrupaciones diferentes fuerza al algoritmo de *k-means* a hacer una separación artificial de los píxeles. En estos casos no hay diferencias entre las medias de los daños, no por una mala clasificación, sino porque debería haber existido una sola clase. Por otro lado, tampoco se ha encontrado evidencia significativa de que los daños sean diferentes entre las ZHD en 8 parcelas con alta heterogeneidad en el estado de la vegetación antes de la tormenta. Son ejemplos las parcelas con excesos hídricos en partes de su superficie. Las diferencias entre los daños provocados por el granizo resultaron menos evidentes que las diferencias existentes en el cultivo antes de la tormenta.

Sin considerar estas 14 parcelas, el resultado global de la validación asciende a 87,01%. Este nivel de ajuste permite utilizar el algoritmo con absoluta confianza. De esta manera, se brinda una herramienta objetiva para detectar zonas de daño con características similares, permitiendo realizar un ajuste del daño más preciso y justo para las partes que intervienen en el contrato de seguro: asegurado, asegurador y re-asegurador.

4. Conclusiones

En este estudio se ha propuesto un algoritmo para determinar de manera automática áreas homogéneas de daños causados por tormentas de granizo dentro de la extensión de una parcela usando datos satelitales de las misiones *Sentinel 1* y *2*. La validación global mostró que en el 87,01% de los casos analizados existe evidencia significativa de que el porcentaje medio de daño no es igual en las tres ZHD determinadas con el algoritmo. La metodología de trabajo que hemos propuesto proporciona transparencia a la definición de las ZHD y mejora el alcance de la valoración de los daños ya que los sensores remotos tienen la capacidad de muestrear toda la superficie de la parcela. Así se evita tener que extrapolar las escasas observaciones puntuales que realizan los peritos en el campo. En futuras extensiones del algoritmo sería interesante adaptar la selección de los índices a cada tipo de cultivo y sus estados fenológicos, teniendo en cuenta también los diferentes tipos de suelo de la región donde se cultiva.

References

- ARTHUR, D., and VASSILVITSKII, S., 2006. k-means++: The Advantages of Careful Seeding. *Chem. Eng.*, **8**, pp. 26–28.
- BECKER, F., and CHOUDHURY, B.J., 1988. Relative sensitivity of normalized difference vegetation Index (NDVI) and microwave polarization difference Index (MPDI) for vegetation and desertification monitoring. *Remote Sens. Environ.* **24**, pp. 297–311. DOI: 10.1016/0034-4257(88)90031-4
- BENTLEY, M.L., MOTE, T.L., and THEBPANYA, P., 2002. Using Landsat to Identify Thunderstorm Damage in Agricultural Regions. *Bulletin of the American Meteorological Society*, **83**(3), pp. 363-376
- CAPELLADES, M.A., REIGBER, S., and KUNZE, M., 2009. Storm damage assessment support service in the U.S. corn belt using RapidEye satellite imagery. In *Remote Sensing for Agriculture, Ecosystems, and Hydrology XI(7472)*, p. 747208. International Society for Optics and Photonics. DOI: 10.1117/12.830393
- CHANDLER, O., APAN, A., PULLINGER, R., and BULLEN, K., 2003. Quantifying Hail Damage for Crop Loss Assessment: Techniques Using Remote Sensing and Geographic Information Systems. In *Proceedings of the 11th Australasian Remote Sensing and Photogrammetry Conference (ARSPC 2002)*, pp. 412-421. Causal Productions.

- CHAUHAN, S., and SRIVASTAVA, H.S., 2016. Comparative evaluation of the sensitivity of multi-polarized SAR and optical data for various land cover classes. *Int. J. Adv. Remote Sens. GIS Geogr.*, **4**(1), pp. 1-14.
- DE LEEUW, J., VRIELING, A., SHEE, A., ATZBERGER, C., HADGU, K.M., BIRADAR, C.M., KEAH, H., and TURVEY, C., 2014. The potential and uptake of remote sensing in insurance: A review. *Remote Sens.*, **6**, pp. 10888–10912. DOI: 10.3390/rs61110888
- EL HAJJ, M., BÉGUÉ, A., LAFRANCE, B., HAGOLLE, O., DEDIEU, G., and RUMEAU, M., 2008. Relative Radiometric Normalization and Atmospheric Correction of a SPOT 5 Time Series. *Sensors (Basel)*, **8**. DOI: 10.3390/s8042774
- FEHR, W.R., and CAVINESS, C.E., 1977. *Stages of soybean development*. Pp. 1–12.
- GAO, X., HUETE, A.R., NI, W., and MIURA, T., 2000. Optical-biophysical relationships of vegetation spectra without background contamination. *Remote Sens. Environ.*, **74**, pp. 609–620. DOI: 10.1016/S0034-4257(00)00150-4
- GHASEMI, N., SAHEBI, M.R., and MOHAMMADZADEH, A., 2011. A review on biomass estimation methods using synthetic aperture radar data. *Int. J. Ofgeomatics Geosci.*, **1**, pp. 776–788.
- GITELSON, A.A., and MERZLYAK, M.N., 1997. Remote estimation of chlorophyll content in higher plant leaves. Remote estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing*, **18**(12), pp. 2691-2697.
- HALSNÆS, K., LARSEN, M.A.D., and KASPERSEN, P.S., 2018. Climate change risks for severe storms in developing countries in the context of poverty and inequality in Cambodia. *Nat. Hazards*, **94**, pp. 261–278. DOI: 10.1007/s11069-018-3387-8
- HUANG, Y., WALKER, J.P., GAO, Y., WU, X., and MONERRIS, A., 2016. Estimation of Vegetation Water Content From the Radar Vegetation Index at L-Band. *IEEE Trans. Geosci. Remote Sens.*, **54**, pp. 981–989. DOI: 10.1109/TGRS.2015.2471803
- HUETE, A. R., 1988. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.*, **25**, pp. 295–309. DOI: 10.1016/0034-4257(88)90106-X
- KIM, Y., JACKSON, T., BINDLISH, R., LEE, H., and HONG, S., 2012. Radar Vegetation Index for Estimating the Vegetation Water Content of Rice and Soybean. *IEEE Geosci. Remote Sens. Lett.*, **9**, pp. 564–568. DOI: 10.1109/LGRS.2011.2174772
- KIM, Y., and VAN ZYL, J., 2009. A time-series approach to estimate soil moisture using polarimetric radar data. *IEEE Transactions on Geoscience and Remote Sensing*, **47**(8), pp. 2519-2527.
- LIU, H.Q., and HUETE, A., 1995. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Trans. Geosci. Remote Sens.*, **33**, pp. 457–465.
- MAGYP. 2021. Estimaciones Agrícolas (magyp.gob.ar). Ministerio de Agricultura, Ganadería y Pesca | Presidencia de la Nación, Argentina. Available: <http://datosestimaciones.magyp.gob.ar/> [7/18, 2021].
- ME. 2021. Ministerio de Economía, Superintendencia de Seguros de la Nación | Argentina.gob.ar. Available: <https://www.argentina.gob.ar/superintendencia-de-seguros> [7/18, 2021].
- MERZLYAK, M.N., GITELSON, A.A., CHIVKUNOVA, O.B., and RAKITIN, V.Y., 1999. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiol. Plant.*, **106**, pp. 135–141. DOI: 10.1034/j.1399-3054.1999.106119.x
- MLYNARCZUK, M., and PORZYCKA-STRZELCZYK, S., 2018. Speckle filtering in SAR images using morphological filters. *International Multidisciplinary Scientific GeoConference: SGEM*, **18**(2.3), pp. 175-182. DOI: 10.5593/sgem2018/2.3/S10.023
- NIZALAPUR, V., JHA, C., and MADUGUNDU, R., 2010. Estimation of above ground biomass in Indian tropical forested area using multi-frequency DLR-ESAR data. *Int. J. Geomatics Geosci.*, **1**, pp. 67–178.
- PERIASAMY, S., 2018. Significance of dual polarimetric synthetic aperture radar in biomass retrieval: An attempt on Sentinel-1. *Remote Sens. Environ.*, **217**, pp. 537–549. DOI: 10.1016/j.rse.2018.09.003
- ROY, P.S., SHARMA, K.P., and JAIN, A., 1996. Stratification of density in dry deciduous forest using satellite remote sensing digital data - An approach based on spectral indices. *Journal of biosciences*, **21**(5), pp. 723-734. DOI: 10.1007/BF02703148
- SAHADEVAN, D.K., SITIRAJU, S., and SHARMA, J., 2013. Radar Vegetation Index as an Alternative to NDVI for Monitoring of Soyabean and Cotton. In *Proceedings of the XXXIII INCA International Congress (Indian Cartographer)*, Jodhpur, India, pp. 19-21)
- SALUZZI, M., and NUÑEZ, J.M., 1975. Comportamiento de granizadas sobre diversas áreas cultivadas del país. *Geoacta* **7**, pp. 77–90.
- SIERRA, E.M., BELTRAN, A.B., and MAIO, S., 1993. Peligrosidad del granizo para los cereales en la region pampeana. *Rev. la Fac. Agron.*

USDA, FCIC, 2019a. Corn loss adjustment standards handbook, 25150.

USDA, FCIC, 2019b. Soybean Loss Adjustment Standards, 25440.

USDA, FCIC, 2016. Small grains adjustment standards, 25430.

VERRELST, J., CAMPS-VALLS, G., MUÑOZ-MARÍ, J., RIVERA, J.P., VEROUSTRAETE, F., CLEVERS, J.G.P.W., and MORENO, J., 2015. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties - A review. *ISPRS J. Photogramm. Remote Sens.*, **108**, pp. 273–290. DOI: 10.1016/j.isprsjprs.2015.05.005

WONG, M.A., and HARTIGAN, J.A., 1979. Algorithm as 136: A k-means clustering algorithm. *J. R. Stat. Soc. Ser. C Applied Stat.*, **28**, pp. 100–108.