# Chapter 19 Farm Weather Insurance Assessment



Antonella Catucci, Alessia Tricomi, Laura De Vendictis, Savvas Rogotis, and Nikolaos Marianos

Abstract The pilot aimed to develop services supporting both the risk and the damage assessment in the agro-insurance domain. It is based on the use of remotely sensed data, integrated with meteorological data, and adopts machine learning and artificial intelligence tools. Netherlands and Greece have been selected as pilot areas . In the Netherlands, the pilot was focused on potato crops for the identification of areas with higher risk, based on the historical analysis of heavy rains. In addition, it covered automated detection of potato parcels with anomalous behaviours (damage assessment) from satellite data, meteorological parameters and soil characteristics. In Greece, the pilot worked with 7 annual crops of high economic interest to the national agricultural sector. The crops have been modelled exploiting the last 3year NDVI measurements to identify their deviations from the normal crop health behaviour for an early identification of affected parcels in case of adverse events. The models were successfully tested on a flooding event that occurred in 2019 in the Komotini region. Even though the proposed methodologies should be tested over larger areas and compared against a larger validation dataset, the results already now demonstrate how to reduce the operating costs of damage assessors through a more precise and automatic risk assessment. Additionally, the identification of parameters that most affect the crop yield could transform the insurance industry through index-based solutions allowing to dramatically cut costs.

A. Catucci (⊠) · A. Tricomi · L. De Vendictis e-GEOS, Rome, Italy e-mail: antonella.catucci@e-geos.it

S. Rogotis · N. Marianos NEUROPUBLIC SA, Methonis 6 and Spiliotopoulou, 18545 Piraeus, Greece e-mail: s\_rogotis@neuropublic.gr

N. Marianos e-mail: n\_marianos@neuropublic.gr

© The Author(s) 2021 C. Södergård et al. (eds.), *Big Data in Bioeconomy*, https://doi.org/10.1007/978-3-030-71069-9\_19

#### **19.1** Introduction, Motivation and Goals

Agricultural insurance protects against loss or damages to crops or livestock. It has a great potential to provide a value to farmers and their communities, both by protecting farmers when shocks occur and by encouraging greater investment in crops. This concept is particularly evident if considering current challenges related to climate change effects and increase of world population. However, in practice, insurance effectiveness has often been constrained by the difficulty of designing optimal products and by demand constraints. The objective of the pilot is the provision and assessment of services for the agriculture insurance market in selected areas based on the Copernicus satellite data series, also integrated with meteorological data, and other ground available data by using big data methods and AI tools.

Among the relevant needs of the insurances operating in agriculture, there are: the more consolidated procedures of damage assessment by means of earth observation techniques and the most promising evaluation of risk parameters down to parcel level.

For the **risk assessment phase**, the integrated usage of historical meteorological series and satellite derived indices, supported by proper modelling, allow to tune EO-based parameters in support to the risk estimation phase. The availability of this information allows a better estimation of potential risky areas and then a more accurate pricing and designing of insurance products. These advantages could positively impact the increase of insurance penetration. Moreover, the definition of key parameters related to the field lost by using machine learning-based approaches has the potential to support the design of innovative insurance products (such as parametric insurance) that are very promising for farmer protection.

For **damage assessment**, the operational adoption of remotely sensed data allows optimization and tuning of new insurance products based on objective parameters. This could imply a strong reduction of ground surveys, with positive impact on insurance costs and reduction of premium to be paid by the farmers.

#### 19.2 Pilot Set-Up

The pilot included trial stages in two different areas of interest: the Netherlands and Greece.

In the Netherlands, the pilot has been realized considering potato crop that is particularly relevant for the national market. The pilot included the generation of different products to enable the detection of parcels with anomalous behaviours and the identification of the most influencing parameters of high impact on crop yield. Some examples of products are introduced here:

• Weather-based risk map that is intended to show occurrences of extreme weather events, heavy rains in particular, in order to identify areas with possible high damage frequency.

• **Intra-field analysis** that is aimed to detect the growth homogeneity and evidencing irregular areas within the parcel, providing an indicator of field anomalies.

Different partners have been involved in the pilot activities. Copernicus satellite data (both optical and SAR) and services have been provided by e-GEOS, the provision of machine learning technology by EXUS, meteorological data and services from MEEO. The involvement of end-users and the provision of local agronomic knowledge have been assured by NBAdvice.

In Greece, the pilot worked with annual crops (e.g. tomato, maize, cotton) of high economic interest to the Greek agricultural sector, in several regions of Northern Greece and in particular in Evros, Komotini and Thessaly. The pilot evaluated incidents like floods and heatwaves that fall under the definition of the climate-related systemic perils. The pilot effectively demonstrated how big data enabled technologies and services dedicated for the agriculture insurance market can eliminate the need for on-the-spot checks for damage assessment and promote rapid payouts. The role of field-level data has been revealed as their collection, and monitoring is important in order to determine if critical/disastrous conditions are present (heat waves, excessive rains and high winds). Field-level data can be seen as the "starting point" of the damage assessment methodology, followed within the Greek pilot case. Moreover, regional statistics deriving from this data can serve as a baseline for the agricultural insurance products.

NP led the activities for the execution of the full lifecycle of this pilot case with technical support from FRAUNHOFER and CSEM. Moreover, a major Greek insurance company, INTERAMERICAN, was actively engaged in the pilot activities, bringing critical insights and its long-standing expertise into fine-tuning and shaping the technological tools to be offered to the agriculture insurance market.

The goal of this particular pilot case was to enable a better management of the damage assessment process (reduction of the required time) and to support other processes of the insurance companies.

#### **19.3** Technology Used

#### 19.3.1 Technology Pipeline

For the **trial stage in the Netherlands**, the pipeline has been composed of three main logical steps (Fig. 19.1):

**Data Preparation**: a set of data has been collected and properly pre-processed in order to get them ready for the processing phase. In particular, the following datasets have been considered: Sentinel-2 optical data, Sentinel-1 SAR data (soil moisture), Proba-V data, weather data including main parameters influencing crop growth (land



Fig. 19.1 Overview of the three main components of the pipeline for the trial in the Netherlands

surface temperature, 24-h precipitation accumulations, humidity, evapotranspiration) and crop data (crop type, parcel boundaries and location, soil type).

**Processing Engine**: the processing step includes different approaches implemented by means of proprietary algorithms that allows the extraction of relevant information that can be used by insurance companies and risk managers. In particular, the three main components are:

- classification and correlation extraction based on machine learning methods
- inter-field anomaly detection and intra-field algorithms
- risk analysis tools.

The processing engine is composed of different blocks that are part of the DataBio shared architecture.

**Visualization**: the visualization phase has been realised by components that are part of the DataBio architecture.

For the **trial stage in Greece**, a set of data collection, processing and visualization components has been used so as to technically support the pilot activities. More specifically the following technological components should be acknowledged:

In terms of **Data Collection**, a set of heterogeneous data is required in different spatial and temporal resolutions to provide services to the insurance companies. Moreover, historical data is critical for shaping insurance products and conducting effective assessments. Data abundancy holds the key for creating sound insurance products/tools. To collect all this data several data collection modules are used:

• In-situ telemetric stations provided by NP, so-called gaiatrons, that collect weather data,

• Modules for the collection, pre-processing of earth observation products, the extraction of higher-level products and assignment of vegetation indices at parcel level.

#### In terms of **Data Processing**:

- GAIABus DataSmart Machine Learning Subcomponent (NP): The specific component supports: EO data preparation and handling functionalities. It also supports multi-temporal object-based monitoring and modelling for damage assessment.
- GAIABus DataSmart Real-Time Streaming Subcomponent (NP): This component supports:
  - Real-time data stream monitoring for NP's gaiatrons installed in the pilot sites,
  - Real-time validation of data,
  - Real-time parsing and cross-checking.
- Neural Network Suite (CSEM): this component was used as a machine learning crop identification system for the detection of crop discrepancies that might derive from reported weather-related catastrophic events.
- Georocket, Geotoolbox and SmartVis3D (FRAUNHOFER): This component has a dual role: It is a back-end system for big data preparation, handling fast querying and spatial aggregations (data courtesy of NP), as well as a front-end application for interactive data visualization and analytics.

In terms of **Data Visualization**, the main component in this category is Neurocode (NP). Neurocode allowed the creation of the main pilot UIs in order to be used by the end-users (insurance companies). An additional DataBio component providing information visualization functionalities is Georocket (FRAUNHOFER).

### 19.3.2 Reflection on Technology Use

In the Netherlands, an **historical risk map** was generated based on SPOT-VGT/Proba-V 1 km fAPAR data from 2000 to 2017 (Fig. 19.2). The index was defined as the sum of fAPAR over the growing season. The risk map allows to detect zones with a higher damage frequency in the past. This technology seems to be effective to generate and to give an overview of the risk in a selected area. Nevertheless, more accurate datasets can be used to analyse more in depth the situation.

In addition, **weather-based risk maps** were produced to complement the historical risk map. The weather risk maps are intended to show the occurrence of extreme weather events in the past and are aimed to investigate if a reliable correlation between damages occurred to the crops and extreme weather events (heavy rains, in particular) occurs. The main goal was to define damage patterns and to zoom in on areas with a high damage frequency. At the end, eight different risk maps were calculated, one per threshold provided by end-users. Moreover, starting from the list of dates related



Fig. 19.2 Map classifying the Netherlands territory in terms of number of years with damages

to damage claims and provided by the insurance companies for the years 2015–2018, the extraction of precipitation values (with the respective location coordinates) has been performed, in order to find further locations (in addition to those provided by the insurance company) where heavy rain events have occurred (see Fig. 19.3).

As concerning the detection of parcels with **anomalous behaviours** and identification of **influencing parameters**, the following approach was considered.

The dataset was split according to the different types of potato, and each group was clustered using satellite data, meteorological measurements and soil characteristics with a monthly aggregation.

Fig. 19.3 Map of precipitation extracted from KNMI dataset on date 30/08/2015. Yellow points: locations provided by the insurance company—blue points: further locations with 24-h precipitation values above the 50 mm threshold





Fig. 19.4 Cumulative distribution function of mean normalized difference vegetation index (NDVI) grouped by cluster and type of potato

The clustering-based service has proved to be a very useful technique to identify parcels with anomalous behaviour and allows to consider in a single analysis all the variables that can affect the growth and the yield of a crop (Fig. 19.4). Unfortunately, it was not possible to validate the results due to lack of data from insurances but the approach seems to be very promising. Moreover, the performed activity reveals that temperature is a factor with high impact on NDVI of potatoes. See Fig. 19.5 where the first plot shows the average NDVI trends of parcels belonging to different clusters. The second one is related to the average temperature recorded over the area defined by the "blue" cluster, characterized by higher temperatures and lower NDVI values in the peak period, and over the area defined by the "red" cluster, characterized by lower temperatures and higher values of NDVI in the peak period.

Lastly, the **intra-field analysis** was performed over areas with a high presence of potatoes. The scope of the analysis has been to analyse each parcel to detect the growth homogeneity and evidencing irregular areas, providing an indicator of field anomalies. In order to resume the approach, a brief description of the intra-field analysis follows.

After creating an inner buffer in order to avoid border effects, the extraction of temporal profile at parcel level was performed. Some filters were applied in order to exclude parcels that were not cultivated or areas with high percentage of cloud coverage. Then, the observation that corresponded to the maximum growth stage of the crop was identified. At the end, each parcel was classified at pixel level according to statistical thresholds. See Fig. 19.6.



Fig. 19.5 NDVI and temperature trends over parcels belonging to different clusters

Intra-field service is extremely effective in detecting soil anomalies that do not allow crops to grow homogeneously within parcels. This service provides an indicator of soil goodness: texture and depth, for instance, have consequences on water consumption and on regular growth.

In Greece, crop type and area tailored crop models have been created for the whole Greek arable area making use of EO-derived NDVI measurements that have proven to be suitable for assessing plant health. In total, for each one of the 55 Sentinel-2 tiles that cover the whole Greek arable land, 7 major arable crops for the local agri-food sector were modelled and namely wheat, maize, maize silage, potato, tomato, cotton and rice ( $55 \times 7 = 385$  models in total). The models were developed exploiting multi-year NDVI measurements from the available last three (3) cultivating periods and instead of using sample statistics (few objects of interest but many observations referring to them), population statistic methods (large number of objects of interest)



Fig. 19.6 Areas with anomalous growth within a parcel (in red and orange)

but with few observations referring to them) were employed instead in order to identify NDVI anomalies. As sound insurance models are typically created using large multi-year historical records (~30 years), this approach is ideal for deriving robust estimates for setting anomaly thresholds (exploiting the space-time cube to have enough degrees of freedom). The goal is to detect deviations in NDVI measurements in respect to what is considered normal crop health behaviour for a specific time instance. Thereby, each crop model consists of 36 NDVI probability distributions that refer to all decades of the year. By adjusting these high and low thresholds (part of the strategy of the insurance company), it is evident that measurements found at the distribution extremes can be spotted and flagged as anomalies. Typically, insurance companies are looking for negative anomalies (below 15%) that provide strong indications of a disastrous incident (Fig. 19.7).

The figures (Table 19.1) graphically depict three different crop models created using the aforementioned procedure.

The effectiveness of the proposed monitoring methodology was tested against a flooding event (11/7/2019) in Komotini that affected cotton farmers in the region and led to significant crop losses (Fig. 19.8).

Initially, Gaiatron measurements confirmed that flooding conditions were present at the area as a result of increased volumes of rainfalls. This proves that the region might have been affected by the systemic risk and should be more thoroughly examined (Fig. 19.9).

This triggered an EO-based crop monitoring approach that captures the impact of the peril to crop's health. After only 2 weeks, the approach identified statistically significant differences compared to the respective crop model that indicates damages at field level. This validates the initial hypothesis that floods were responsible for severely affecting the region's crop health and consequently proves that the established methodology can be a powerful tool for early identification of



**Fig. 19.7** Crop NDVI probability distribution referring to a single decade of the year (wheat-Larisa region-2nd decade of February). Anomalies can be found at the distribution extremes



Table 19.1 Crop models of cotton, maize and wheat

(continued)



Table 19.1 (continued)

Wheat model in Larisa region (T34SFJ tile) by decade (horizontal axis). Light green threshold indicates lower 15% extremes while dark green threshold indicates upper 85% extremes of the probability distribution. Red line is presenting a single parcel status for the whole 2018 with its NDVI measurements staying within "normal" ranges for the critical cultivating periods



Fig. 19.8 Aftermath of the floods in Komotini region (11/7/2019)



Fig. 19.9 Rainfall volume (mm) in the Komotini region

potentially affected/damaged parcels, crop types and areas. The findings have been presented both to the insurance company and the farmers in order to show how these technologies can bridge the gap among the farming and the insurance world (Fig. 19.10).

By mapping the outcome of the followed damage assessment procedures on top of a map, it is evident that high-level assumptions can be made. This involves the risk at which the insurance company is exposed to and prioritizing the work that needs to



Fig. 19.10 Parcel monitoring at Komotini region (cotton) showing negative anomaly (deviation) for two consecutive decades just after the disastrous incident

be conducted by field damage evaluators (until now prioritization is not data-driven) that are advised to begin with parcels exhibiting higher damage estimates and steadily move to those with lower ones (Fig. 19.11).



Fig. 19.11 High-level overview of the affected area, color coded with the output of the followed damage assessment procedures

#### **19.4 Business Value and Impact**

#### **19.4.1** Business Impact of the Pilot

#### **Business Impact of the Pilot—Netherlands**

Results are promising in terms of general procedures and methods. These need to be tested over larger areas and compared with validation data provided by the final users (insurance). The data availability is a crucial challenge for this market considering the very restricted dissemination level of the information and the high competitive level. In fact, the insurance companies are not interested in supporting the development of products that can be available also for their competitors. To overcome these potential limitations, a set-up phase of the service in an operative environment is necessary in close cooperation with the insurance company involved. This collaboration has the potential to transform the tested methods into operative services filling the existing gap between prototype development and final product.

In order to analyse the benefit of the tested technology for the insurance industry (risk estimation also by means of machine learning), it is important to define the three levers of value in insurance market:

- 1. Sell More
- 2. Manage Risk Better
- 3. Cost Less to Operate.

The activity performed in the pilot impacts essentially the point "Cost Less to Operate". One clear way to reduce operating costs in insurance is to add information and increase automation to complex decision-making processes, such as underwriting. To keep processing costs in check, many insurance carriers have a goal to increase the data available in support to a more precise and automatic risk evaluation in support of the underwriting. In fact, the use of decision management technologies like risk maps, machine learning and artificial intelligence can reduce the time spent to analyse each contract and focus team members on higher value activities. Moreover, the identification of parameters that most affect the crop yield performed in the pilot can support an innovative insurance typology called "parametric insurance". This particular insurance typology is revolutionizing the insurance industry allowing to dramatically cut operative costs removing the in-field direct controls.

The first step in building a parametric product is determining the correlation between the crop losses and a particular index representative of the climate event associated with the loss. The activity performed in the pilot by using a machine learning approach is to identify the most important parameter affecting the crop yield that can be the basis for a parametric or index-based insurance.

Quantifying the potential impact of the proposed solution for the insurance industry is a complex issue considering the work necessary to transform the methodology in an operative service. Just to provide some business projection, it can be considered that direct European agricultural insurance premiums in 2016 were 2.15 m  $\in$  (estimated by Munich RE) (Fig. 19.12).



Fig. 19.12 Premium value distribution in Europe estimated by Munich RE

It can be considered that around 70% of this amount is spent by insurances to reimburse damages and the remaining 30% is used to pay internal costs and reinsurances. Considering this dimension and considering the row and very preliminary estimation obtained by the pilot, it is possible to assume that the cost that can be saved by using EO-based services in support of risk assessment is around 2% of the total cost used by the insurance to pay internal costs. Table 19.2 summarizes the potential available market for these services in Europe.

#### **Business Impact of the Pilot—Greece**

There is a constantly increasing need for agricultural insurance services, due to the adverse effects of climate change and the lack of sufficient compensation frameworks. From their side, insurance companies with offerings for the agricultural sector need to have precise and reliable systems that will facilitate the damage evaluation processes and will ensure swift and fair compensation to those who actually deserve it, thus

	Size by revenue	Market share	
Available market in Europe	2150 m€ 30% * 2% = 12.9 m€	100%	Potential cost reduction by using downstream services supporting the insurance industry is assumed the 2% of the total insurances income (30% of the insurance premium)

Table 19.2 Market projection in Europe

Market projections Market segment: insurance allowing follow-up/reactive measures to be undertaken and supporting food security in general.

In the two trial periods of DataBio, tailored agri-insurance tools and services have been developed with and for the agri-insurance companies that perform EO-based damage assessment at parcel level and target towards evolving to next-generation index-based insurance solutions. The pilot results clearly show that data-driven services can facilitate the work of the insurance companies, offering tools that were previously unavailable and were responsible for severe bottlenecks in their day-to-day activities including:

- long wait for official damage evaluation reports,
- dependence on the human factor,
- difficulties in prioritizing work after receiving several compensation claims.

## 19.4.2 Business Impact of the Technology on General Level

The remote sensing literature offers numerous examples proposing earth observation techniques to support insurance, for example in the assessment of damage from fire and hail [1, 2]. To date, however, few operational applications of remote sensing for insurance exist and are operative. Many scientific papers claiming potential applications of remote sensing [3–5], typically stress the technical possibilities, but without considering and proving its contribution in terms of "value" for the insurer. The discrepancy between the perceived potential and the actual uptake by the industry is probably the result of two main reasons:

- technological solutions not adequate and too expensive, in relation to the valued added
- over-optimistic assumptions by the remote sensing community, regarding the industry's readiness to adopt the information by remote sensing.

Despite this situation, EO can still play a central role in supporting the insurance market in agriculture trying to design services that can really bring value to the users. This is the case of supporting in field verification and parametric insurance products (innovative insurance products). The present pilot investigates these services demonstrating the potential and opening up the route for new collaboration with users.

# **19.5** How-to-Guideline for Practice When and How to Use the Technology

As said, the methodology needs to have a pre-operational set-up phase in close collaboration with the insurance company. In fact, the developed method can be

applied to different areas and crops but only if an adequate training set of data related to occurred losses are available.

In Greece, the proposed solution is based on mature technologies and high-quality data, in order to ensure high accuracy and quality for the designed tools and services. EO-based methodologies were used in order to extract useful information from EO products for:

- damage assessment targeting towards a faster and more objective claims monitoring approach just after the disaster,
- the adverse selection problem. Through the use of high-quality data, it will be possible to identify the underlying risks associated with a given agricultural parcel, thus supporting the everyday work of an underwriter,
- large-scale insurance product/risk monitoring, that will allow the insurer to assess/monitor the risk at which the insurance company is exposed to from a higher level.

More and more insurance companies are interested in entering the agricultural market, which exhibits high value, due to its vulnerability to extreme weather phenomena. However, before they integrate such technology- and data-driven tools, they need to be persuaded that these tools will help them reduce operational costs by minimizing the human intervention and ensuring high quality of services. The involvement of one of the largest insurance companies in Greece in this pilot case (INTERAMERICAN) helps in bringing the proposed solution closer to the market, and with their precious feedback, it will be more easily available for commercial exploitation.

#### **19.6 Summary and Conclusion**

The objective for the pilot was to find useful services for the insurance to gain more insight about the risk and the impact of heavy rain events for crops. In the Netherlands, for instance, potato-crops are very sensitive to heavy rain, which may cause flooding of the field (due to lack of runoff) and saturation of the soil. This may cause the loss of the potato yield in just a few days. Areas of greater risk can be charged with higher costs for the farmer. The investigated correlation among precipitation and losses can support the identification of index for parametric insurance products. Moreover, instead of just raising the premium, the intention of the pilot was to be able to create awareness and incentives for farmers to prevent losses. Therefore, the services serve multiple purposes. Weather is an important factor in crop insurance, because it represents a critical aspect influencing yield. The analysis of the long-term precipitation, categorized in threshold values, for intense rain events, gave insight in the areas with higher risk. In the pilot, the relation between one single event and the potential yield loss has been analysed. For this purpose, an annotated set of data, where actual losses were determined, was necessary. Because of the privacy issues related to sharing the damage data, the location of damaged fields in the Netherlands

could not be pinpointed precisely enough for correlation to the EO data. Without the details about historical events, this relationship could not be determined. In Greece, where a massive flood event occurred, impacts have been identified by analysing NDVI anomalies for the most common crop types. During the pilot activities, we realized that a service, based on the alert that a heavy rain event took place, would be useful for gaining insight about the impact on other locations. Additionally, in order to find the most limiting aspect in the crop development, we created a dataset based on the Sentinel-2 raster size to combine NDVI with SAR, precipitation (cumulative), temperature and soil type. The developed methodology, however, is valuable for further analysis, not limited to insurance topics and can be extended to other crops in support to risk assessment and also for design of new insurance products such as parametric insurance.

#### References

- Young, F. R., Apan, A., & Chandler, O. (2004). Crop hail damage: Insurance loss assessment using remote sensing. In *Mapping and Resource Management: Proceedings of RSPSoc2004*.
- Peters, A. J., Griffin, S. C., Viña, A., & Ji, L. (2000). Use of remotely sensed data for assessing crop hail damage. *PE&RS, Photogrammetric Engineering & Remote Sensing*, 66(11), 1349– 1355.
- Turvey, C. G., & McLaurin, M. K. (2012). Applicability of the normalized difference vegetation index (NDVI) in index-based crop insurance design. *Weather, Climate, and Society, 4*, 271–284. https://doi.org/10.1175/WCAS-D-11-00059.1
- De Leeuw, J., Vrieling, A., Shee, A., Atzberger, C., Hadgu, K. M., Biradar, C. M., Keah, H., & Turvey, C. (2014). The potential and uptake of remote sensing in insurance: A review. *Remote Sensing*, 6(11), 10888–10912.
- Savin, I. Yu., & Kozubenko, I. S. (2018). Possibilities of satellite data usage in agricultural insurances. *RUDN Journal of Agronomy and Animal Industries*, 13(4), 336–343.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

