Reconciling weather data with insurance data

Author: Hanna Płotka, Risk Analyst, Agri & Specialty Risk Modelling, SCOR Global P&C

Abstract

The results of agricultural insurance schemes are highly weather dependent, but incorporating weather data into pricing is often very difficult. Many challenges must be overcome before the design and risk assessment of a sustainable product is achieved. One of the main issues to address is the complexity of weather datasets, which makes it nearly impossible for non-expert users to deal with them. Another issue is obtaining a methodology for bringing weather data into the (re)insurance framework.

Here, we present a methodology embedded in a SCOR in-house tool which aims to address both of these challenges, allowing non-expert users to gain a quantitative understanding of the impact of weather on agricultural exposure. This allows not only the design of parametric weather products, but also assists in the pricing of other more standard (re)insurance products, especially in regions for which loss experience data is scarce or unavailable.

The tool allows users to examine weather data easily, regardless of the underlying data source or format. Meteorological variables, such as precipitation or temperature, coming from different sources may be examined and compared. Additionally, weather indices such as the Continuous Days Temperature Index may be calculated to estimate return periods of extreme events. These indices may also be combined with exposure and loss history information to reconcile a purely "weather" view with a (re)insurance risk view.

We end with showing an example of the application of the methodology for the case of freeze events affecting crops in the most sensitive stage of the growing cycle.

Introduction

A thorough understanding of the risks present is the deciding factor of the long-term success (or failure) of many agricultural insurance schemes worldwide. These risks may be split into a range of categories: "human factors", such as farm management practices or the choice of plant varieties in a given region, "natural resource factors" such as the soil types or the availability of water, as well as "weather and climate factors", to name a few. The chaotic nature of weather systems results in high year-to-year variability of weather conditions and thus large uncertainties in agricultural production and insurance pay-outs. Most agricultural insurance schemes concentrate on protecting against adverse weather conditions due to the large uncertainties caused by weather conditions (agriculture is a so-called "open roof industry").

One way of looking at weather events which cause agricultural losses is to divide them into two categories based on their geographic extent. Following this classification, the first type of weather events generally do not affect large geographic areas, although they may vary in frequency and intensity from year to year. An example of such a weather event are hailstorms. The other type of weather events may affect very large geographic regions, in some extreme cases even several countries simultaneously, and which additionally may (although not must) be persistent in time. These may be called "systemic risks", and their examples include droughts, such as the Great Chinese Famine

of 1958 – 1961, or frost events such as those which affected large parts of central and southern Europe in spring 2016.

Understanding the frequency and intensity of events with large geographic extent is crucial for designing sustainable insurance schemes. This includes understanding not only the 1-in-100 years events, but also those 1-in-10 or 1-in-25 years ones which may still cause considerable damage. Often, trying to understand these events is made difficult by a lack of data – both detailed data related to agricultural production and losses through history, adjusted to reflect present-day agricultural farming practices, crop varieties and exposure, as well as of meteorological data measuring the extent and intensity of weather events. In this paper we aim to examine the usefulness and limitations of the data available, and present a methodology for combining the different available views.

The paper is structured as follows. We first give an overview of the types of weather and agricultural insurance data available and discuss their limitations. We next give an overview of the methodology applied in a SCOR in-house tool, STRATUS, for reconciling weather and insurance data, making most of the combined view of the risk. We show an example of the application of the tool, and finally end with a few conclusions and outlook.

Data available for agricultural insurance

When assessing the risk associated with a given agricultural insurance cover it is beneficial to examine data from a variety of different perspectives. One possibility is to view the data from a purely insurance perspective, assuming that what has happened in the past is representative of what will occur in the future. Another possibility is to view data from a pure weather perspective, under the assumption that extreme (geographically wide-spread) weather events are going to directly affect agricultural production and thus the insurance programme. Below we discuss both of these methods, and how to bridge the gap between them.

Insurance data

Viewing historical results is a common first step in assessing the performance of agricultural insurance schemes. These historical results first need to be de-trended and restated to reflect present-day conditions. De-trending needs to be done to account for factors such as evolving farm size and management practices, the introduction of new technologies which open new opportunities for e.g. precision agriculture, or the ever-improving plant varieties which have the capacity for producing higher yields and have greater (or smaller) resistance to damage. Restating needs to consider the changes in exposure due to e.g. crop rotation, changes in land-use, or changes in policy structure and underwriting. This is a challenging task on its own and for a thorough analysis considerable expertise is required and is not dealt with in this paper.

A challenge for the assessment of insurance schemes is often the lack of insurance data. This could be because an insurance scheme is new on the market, or because an existing scheme has been restructured to such an extent that historical results are not representative even after restating and de-trending (in particular because they e.g. introduce covers for additional perils). Additionally, only in recent years the technological advances in both computational power and data storage as well as geographic information systems have made it possible for data to be stored at a granularity which allows detailed analysis.

Without a long enough history of insurance data it is nearly impossible to draw conclusions about return periods of extreme loss events, almost always resulting from extreme weather conditions, which would stress-test the scheme. That is why it is beneficial to turn to weather data to help in the evaluation of the risk, and to help with the estimation of the probable maximum loss (PML).

Weather data

There are several types of weather data which may be used for analysis. These may either come from free, publicly available sources (e.g. governmental agencies such as the US NOAA) or from sources with restricted access, for the use of which a fee needs to be paid or which are available for academic use only. These may be divided into either observational data or model output data. Observational data is data which comes from directly measuring atmospheric conditions. This may be done by either remotely sensing the atmosphere by the use of instruments such as satellites, or by on-ground measurements, for example at weather stations. Satellite measurements allow for an almost complete three-dimensional daily global picture of the state of the atmosphere. The Earth's atmosphere has been monitored with satellites since 1979, and thus presents a long-term continuous view of atmospheric conditions. The limitation is that satellite data prior to 1979 is not available. Weather station data may offer a longer history, but is limited to measurements coming from a single location, and may have biases resulting from such factors as changing the measurement instrument, changing the location of the measurement, or human error in recording the measurement, to name a few. Sophisticated methodologies exist for interpolating the measurements of multiple weather stations onto a grid giving a more complete regional view, see e.g. Haylock et al. (2008), but one must keep in mind that the availability and quality of station data is very region-dependent. Europe has a very large network of weather stations, whereas many regions with developing markets, for example in Africa or South America, have very limited data available.

For regions with limited data availability, reanalysis weather datasets may be used. These are datasets which present a complete, global view of the state of the atmosphere by initialising a Global Circulation Model (representing the dynamics of the atmosphere) with observational data coming from a range of possible sources such as satellite measurements and station data. For more information, see for example Dee et al. (2014), Trenberth et al. (2008) and many others. Reanalysis datasets allow examining regions where observational data may not be readily available, while making sure that data is consistent and does not violate any laws of physics.

There are several problems associated with using weather data for understanding agricultural insurance risk. Weather data comes from a variety of different sources, and in various formats. It requires considerable knowledge of the underlying data formats to prepare it for analysis. Also, knowledge of meteorology itself is only a first necessary step to link meteorological "measured" variables, such as temperature or precipitation, to agricultural production and losses. This linking can be done by calculating various indices, which should represent the peril of interest, such as drought. Depending on the index chosen, one or more perils can be represented simultaneously. This requires knowledge of the loss driver in crop production and hence in the insurance scheme.

STRATUS – a SCOR in-house tool

Overview

At SCOR, we have developed a tool, STRATUS, which allows the linking of agricultural insurance data with weather data. STRATUS aims to extract and process meteorological data for risk assessment in agriculture and other weather-related reinsurance lines of business. It consists of two basic modules: a data extraction module and a data analysis module. The data extraction module allows a user to extract available meteorological variables, or indices calculated from available meteorological data, for a region and time period in question (e.g. data for the winter months December and January in the Netherlands, for the years 1979-2016). This data can either by plotted on a map or saved in a user-friendly format, such as Excel, for further analysis. A screenshot of this module is shown in figure 1.
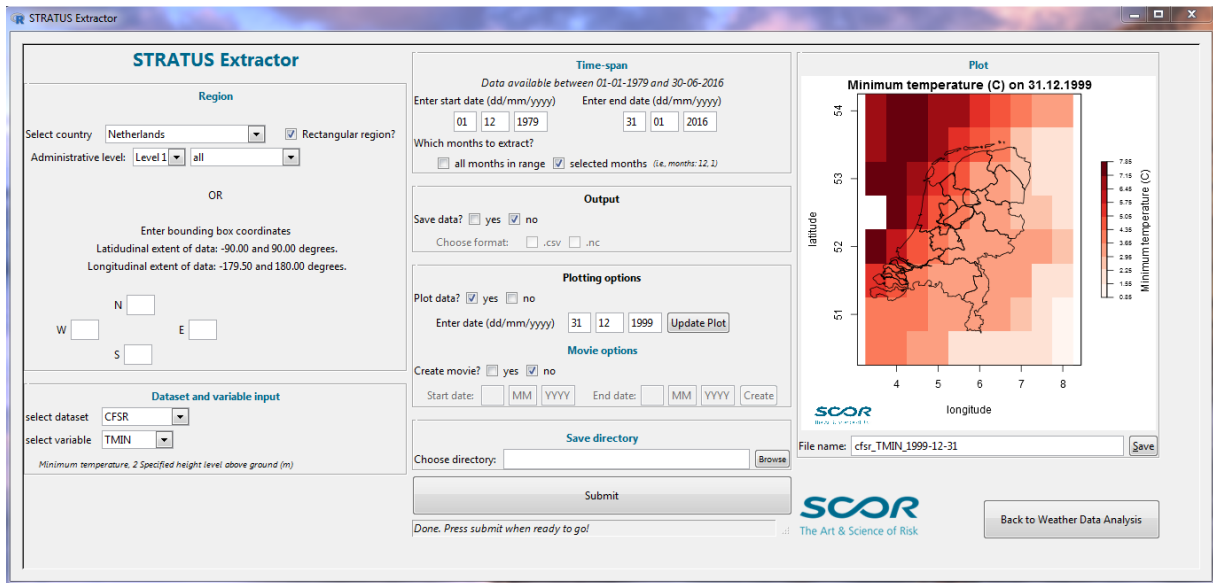
Figure 1 An example of how the data extraction module of STRATUS may be used for viewing what the minimum temperature in the Netherlands was on December 31st, 1999. Data provider: Climate Forecast System Reanalysis (CFSR) product created by the US National Center for Environmental Prediction (NCEP), see Saha et al. (2010) and Saha et al. (2014) for more details.

The data analysis module allows the user to calculate a weather severity index for a selected region and time period. The user also enters the exposure information of the insurance portfolio in question to calculate an index that depends on weather conditions as well as on client exposure. This is to ensure that regions in which there is little or no exposure (or values at risk) do not contribute towards the overall result as much as regions with high exposure. STRATUS additionally allows the calculated index to be compared with loss data, to ensure that an index is representative of the conditions on the field, and it estimates return periods for the most severe events. STRATUS is currently able to handle the following perils: drought, excessive rain, and heat stress (including frost), and has global weather data coverage.

The steps involved in using STRATUS for analysis are best demonstrated with an example.

Example – freeze events

Consider a hypothetical agricultural insurance scheme in Switzerland. The scheme covers various perils and crops throughout the country, but is a relatively new cover with only 7 years of loss history. Since its introduction, the scheme has twice suffered very large losses coming from early spring freeze events affecting flowering fruit trees. We are interested in knowing the return periods of such events to assess whether the scheme is sustainable in design. The on-field experts are estimating that such severe spring events are rare, with return periods of around 15 years, despite the empirical return period being 3.5 years.

As a first step, a restating of the available 7 years of historical insurance data is performed outside of STRATUS. In this step, we account only for the present-day exposure and losses coming only from freeze events on fruit trees. We then enter and plot the exposure into STRATUS. At this stage we can also load the restated loss data into the tool. A screenshot of the STRATUS analysis module is shown in figure 2.

Next, the weather dataset and weather severity index must be selected. We are interested in a larger geographical region, and so it is insufficient to consider station data which give measurements at a single point. Instead, the CFSR reanalysis dataset is used, which contains gridded data. We consider the consecutive days temperature index (Klein Tank et al. 2009) which is designed to reflect the amplitude and persistence of heat and cold conditions. The idea behind the index is that a crop will

likely suffer considerable damage if it is exposed to hot or cold conditions for a long enough period without interruption.
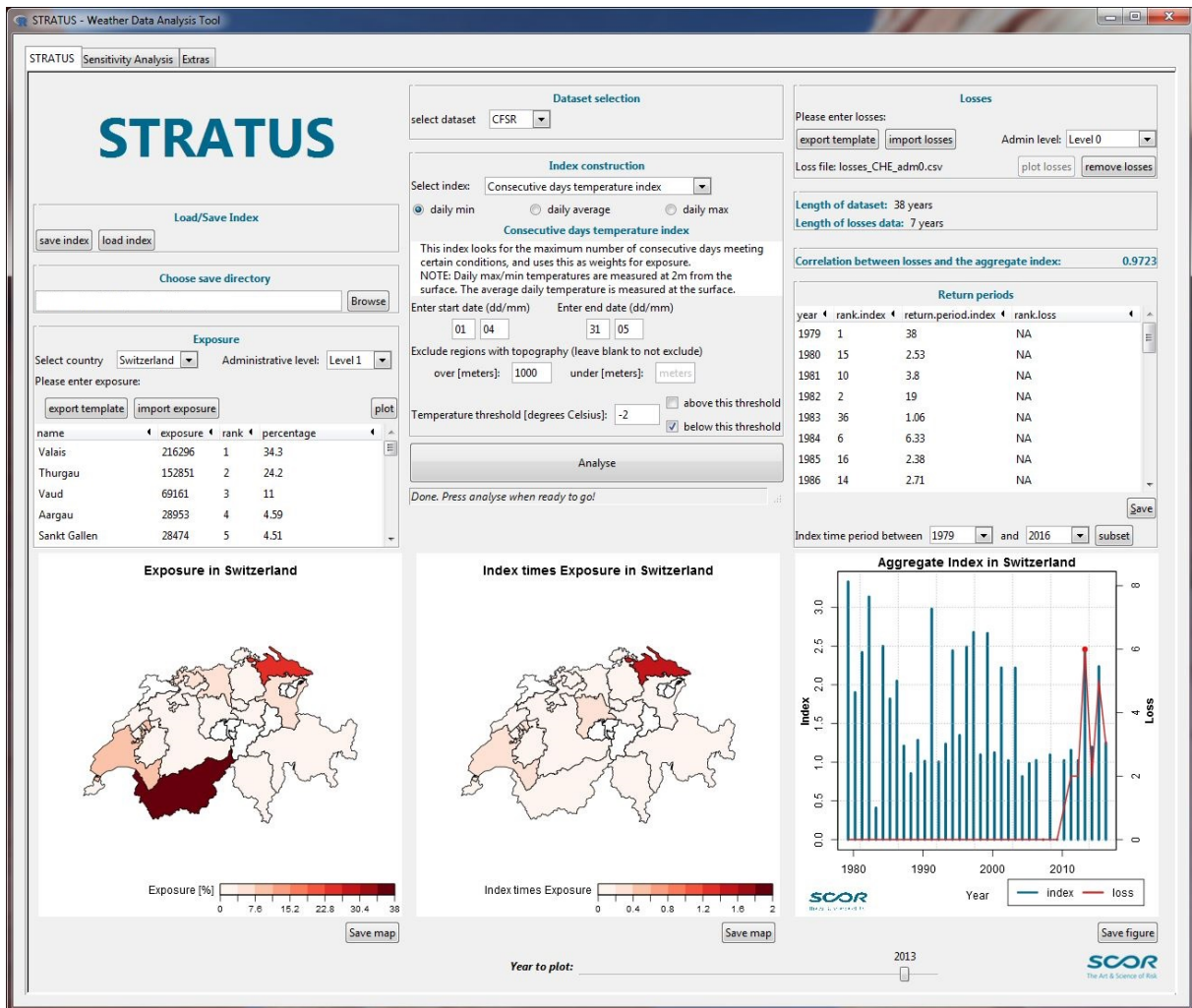

Figure 2 An example of the STRATUS analysis module. A hypothetical exposure of an insurance scheme in Switzerland is plotted.

The time period of interest is spring when the fruit trees are flowering. The exact dates vary from year to year, but from discussions with on-field experts and a review of literature, the critical time period was determined to be the period between 1st April and 31st May. Large parts of Switzerland are mountainous regions which are unsuitable for fruit tree orchards. Therefore we do not wish to consider weather data in regions at an altitude over 1000m and filter any data in these regions from our analysis. Finally, we need to define the critical temperature below which damage to the flowers of fruit trees occurs. These flowers are very sensitive to freeze events so temperatures below 0°C cause damage. However, as we are using a gridded dataset, temperature extremes are smoothed (Haylock et al. 2008), and so the threshold needs to be adjusted to -2°C to account for this.

With the above information, analysis may be performed as follows. A value for the index is calculated at each grid point by using the CFSR daily minimum 2-m air temperature. This gives one value per season per grid point. The index is then aggregated to find an average value of the index for each administrative region (at the same granularity at which exposure data was provided, see figure 2). In this case, we calculate an average value of the index per canton per season in Switzerland (in a real case one would need to drill down to a lower level geographical entity). This spatial data is then weighted by the exposure in each canton, so that in cantons with no exposure, the index is set to zero,

and the value of the index in cantons with exposure is weighted depending on their contribution to the overall exposure. For example, say the value of the index in 2013 was x in the canton of Valais, which has 34.3% of the overall exposure. We then set the value of the index in Valais in 2013 to 0.343x. We show an example of the graphical results of this step in the left panel of figure 3.

Since the performance of an insurance scheme is usually assessed on a yearly basis, we then aggregate the weighted index values in order to be able to compare the weather severity from year to year. This is done by summing the weighted index values over all regions to obtain one value per season, as shown in the middle panel of figure 3. This allows for a comparison of different seasons – in our example, the years 2013 and 2015 indeed had higher index values than those in the other years for which insurance data is available. The correlation between the aggregate index and the losses (available for only 7 years) is 97.2%, leading us to believe that the index is representative of the performance of the scheme for this peril and crop. Furthermore, STRATUS calculates empirical return periods for each year, meaning that the return period of the most severe event is set equal to the length of the weather dataset, the second most severe event to the length of the dataset divided by two, and so on. In this example, we see that our two most severe loss events of 2013 and 2015 both have a return period between 3 and 5 years, which is much more frequent than the 15 year estimate by on-field experts.
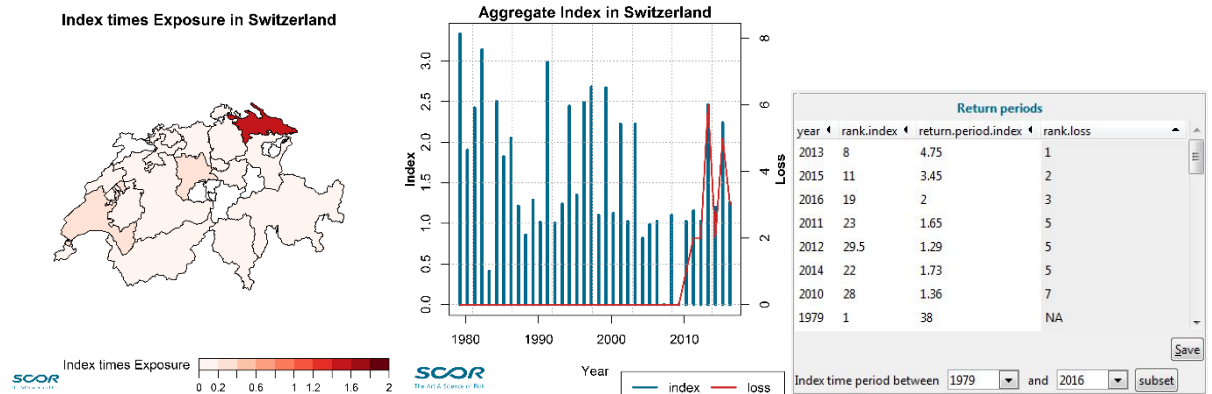

Figure 3 Left: An example of the result of weighing the average index value per canton in 2013 with the exposure in each canton. Middle: An example of the aggregate index through time plotted alongside the losses. Right: An estimate of the return periods.

Conclusions and outlook

In this paper we have discussed the limitations of examining agricultural insurance data from a pure insurance or a pure weather perspective. Doing so gives an incomplete view of the risk, especially when only very short experience data is available, when a cover for a new peril is introduced, or when an insurance scheme is completely restructured making past experience data unrepresentative of the current conditions. Additionally, from an insurance perspective it is important to consider only those regions in which exposure is present – regions in which a peril had been particularly severe but which contain no exposure do not contribute to the overall losses.

We have above presented a tool with an embedded methodology for addressing these issues, and for allowing a more in-depth view of the risk. Of course, many factors play a role in agricultural production. It is not necessarily the case that an insurance loss occurs when STRATUS calculates high values of a weather index, and contrarily, a scheme may not be completely loss-free even when the weather index has low values. For example, using the scenario from before, additional weather factors can be considered: one could look at temperatures and rainfall conditions in the spring to be able to better estimate the flowering window variability from year to year, and hence more precisely choose the risk period. However, the results of the analysis done in STRATUS may be a first and important step for

analysing insurance programmes, and the extraction module of the tool can also be used for obtaining data for further analysis.

Note that the example presented in this paper is based on the fruit tree exposure in Switzerland as found in Caloz & Boehlen (2015) and on purely fictional loss data. The consecutive days temperature index is calculated using actual weather data over Switzerland. To the author's knowledge, a pure spring frost insurance cover does not exist in Switzerland.

Acknowledgements

References

1. Caloz, M. and Boehlen, D. 2015: Obst- und Tafeltraubenanlagen der Schweiz 2015. Fachbereich Pflanzliche Produkte, 30. November 2015, Bundesamt für Landwirtschaft BLW.
2. Dee, D. P., Balmaseda, M., Balsamo, G., Engelen, R., Simmons, A. J., J.-N. Thepaut. 2014: Toward a consistent reanalysis of the climate system. Bull. Amer. Meteor. Soc., 95(8), 1235-1248.
3. Haylock, M.R., N. Hofstra, A.M.G. Klein Tank, E.J. Klok, P.D. Jones, M. New. 2008: A European daily high-resolution gridded dataset of surface temperature and precipitation. J. Geophys. Res (Atmospheres), 113, D20119, doi:10.1029/2008JD10201
4. Klein Tank, A. M., Zwiers, F. W., X. Zhang. 2009: Guidelines on Analysis of extremes in a changing climate in support of informed decisions for adaptation. Geneva: WMO.
5. Saha, S., & al. 2010: The NCEP Climate Forecast System Reanalysis. Bull. Amer. Meteor. Soc., 91, 1015–1057, doi: 10.1175/2010BAMS3001.1.
6. Saha, S., & al. 2014: The NCEP Climate Forecast System Version 2. J. Climate, 27, 2185–2208, doi: 10.1175/JCLI-D-12-00823.1.
7. Trenberth, K. E., T. Koike, K. Onogi. 2008: Progress and prospects in reanalysis. Eos Trans. AGU, 89(26), 234–235, doi:10.1029/2008EO260002.