**Automated breakdown analysis of time series for developments in Dutch agriculture**

Paul van Leeuwena & Jamal Roskamb

a Wageningen Economic Research, email: [paul2.vanleeuwen@wur.nl](mailto:paul2.vanleeuwen@wur.nl)

b Wageningen Economic Research, email: [jamal.roskam@wur.nl](mailto:jamal.roskam@wur.nl)

**Introduction**The FADN (Farm Accountancy Data Network) is a European instrument for evaluating the income of agricultural holdings and the impacts of the Common Agricultural Policy. A wide variety of data is collected within the framework of the FADN. Hence, the FADN is a main basis for much research carried out within Wageningen Economic Research. Many of the data analyses are performed manually (via spreadsheets). Such an approach (in which analyses are performed manually) is time-consuming, subjective to the bias of the analyst, prone to man-made errors, and only first-level causes are identified. It has therefore been proposed to set up an automated breakdown analysis of time series. Such an analysis entail the following advantages:

* less man hours are needed to perform the analyses because the manual work is now automatically executed;
* more objective and less prone to errors because the same automated method can always be performed in the same way;
* deeper analyses can be more easily performed because causes at a deeper level can automatically be identified which is beneficially for sector experts, policymakers and farmers.

After consulting various sector experts within Wageningen Economic Research, the product French-fries potatoes (in Dutch: frietaardappelen) was selected for the development of an automated analysis. The problem in this sector is that costs are not always easy to allocate. Several products can be grown and harvested on a field every year. In addition, there are differences in the labour intensity with regard to the harvest of these products. This causes problems when allocating, for example, machine costs. Despite these challenges, French-fries potatoes was chosen as a case study.

**Objective**

The aim of this project was to develop an automated breakdown analysis of time series for French-fries potatoes that can be extended towards products from both the same sector and other sectors. The emphasis in this research was on the development of the instrument to perform an analysis instead of a correct data analysis (this means that data quality was of secondary importance).

**Data and methodology**

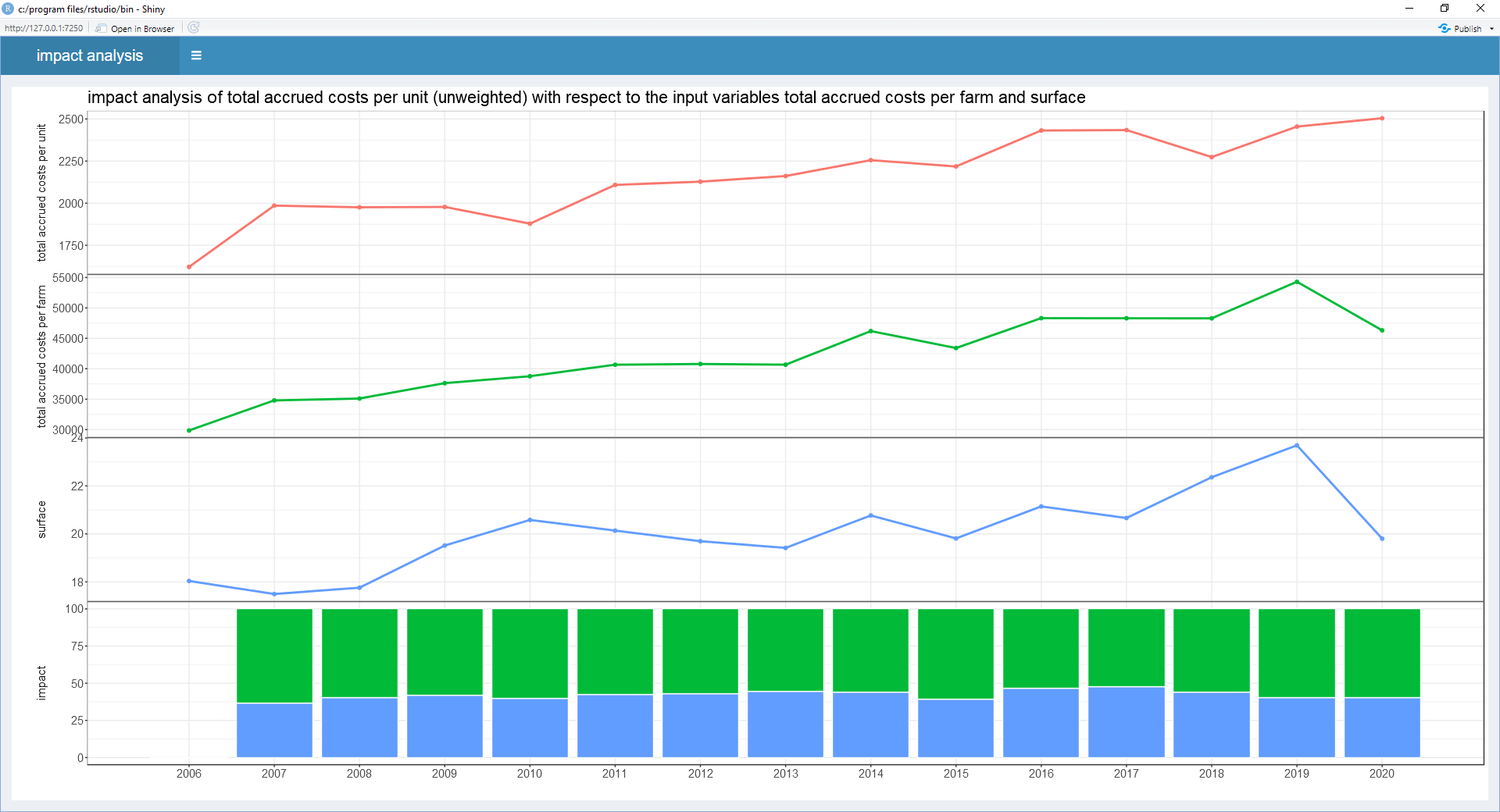
The project started with a breakdown analysis of the most important cost items for the production of French-fries potatoes. Initially a start was made with fictitious data, where an example was used to show the principle of effects of changes in price and quantity on the revenue.

The method consists of relating the change of the variable of interest, e.g. total accrued costs per unit, to the change of the underlying variables that make up the variable of interest, in this case total accrued costs per farm and surface. As the relationship between the variable of interest and the underlying variables is known, the impact of a change in one of the underlying variables on the change of the variable of interest can be exactly determined. See the attached model report for more details.

After elaborating the concept, we switched to the "real" figures. The existing classification of costs as used in the Dutch FADN was aligned as much as possible. Once all cost items were identified, costs were assigned to the production of the French-fries potatoes. The focus in the analysis was on the direct costs. Initially, the idea was to include both direct costs and indirect costs in the analysis, but budget and time of colleagues were limiting factors here. The allocation of indirect costs requires more sector knowledge from the experts in the organization.

**Results**

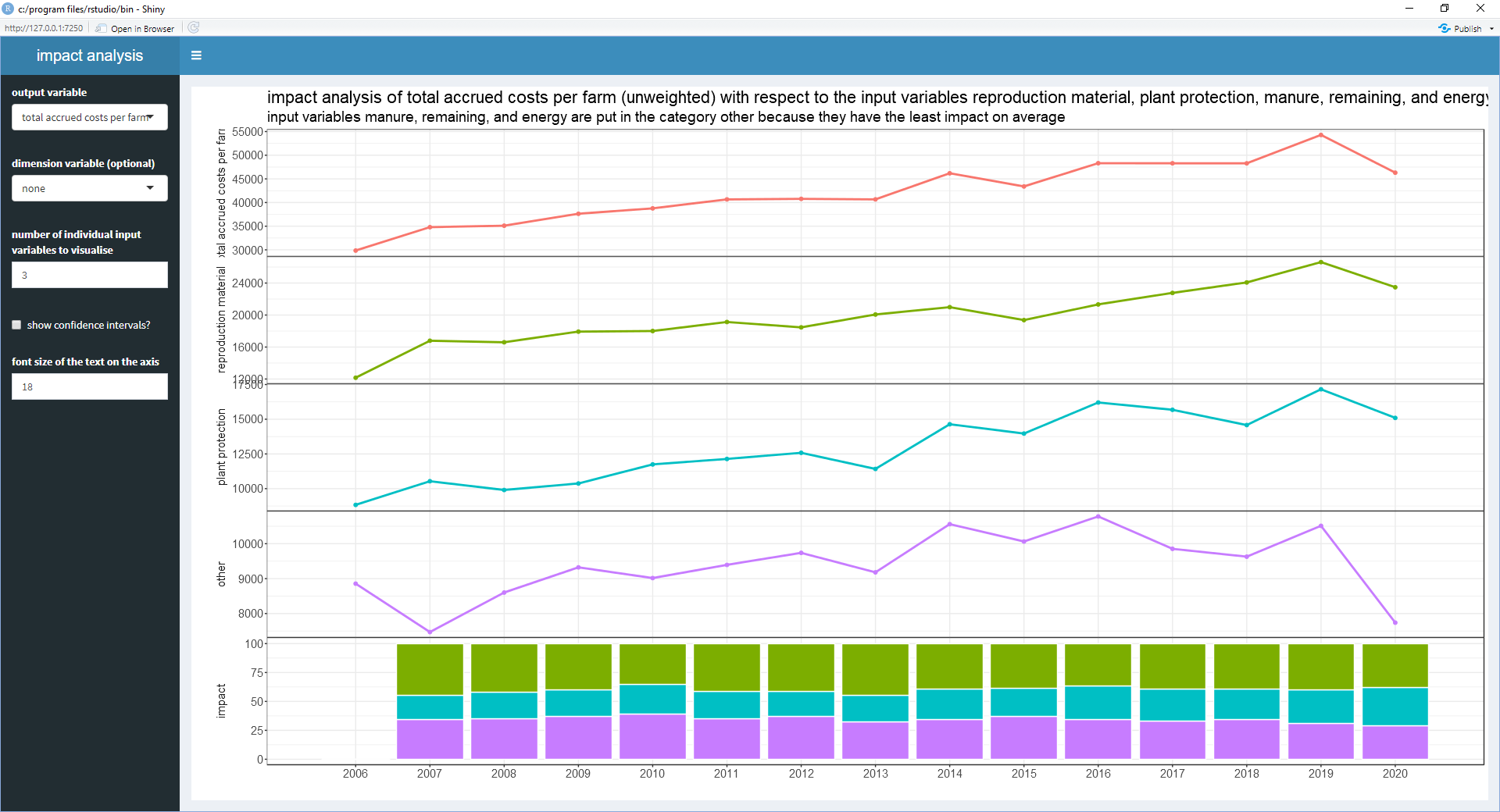
The R Shiny app enables a user to investigate and rank the impact of a change in one of the underlying variables on the change of the variable of interest. Screenshots of the tool are shown below.



The screenshot above has the total accrued costs per unit on top, in red, and the of the underlying variables

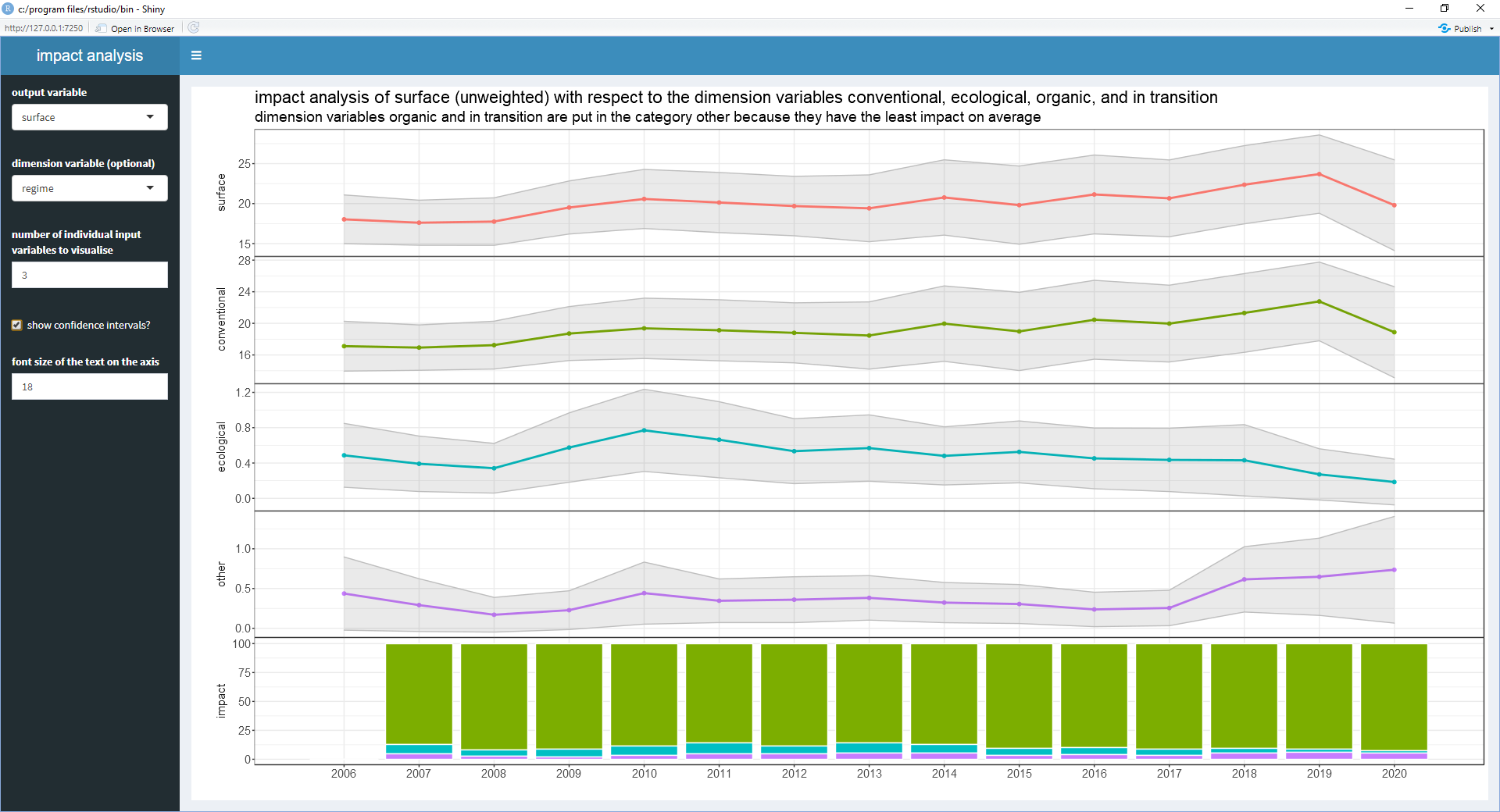
that make up the variable of interest, in this case total accrued costs per farm and surface, in green and blue. The bottom panel visualizes the impact of the underlying variables with the total accrued costs per farm as the variable with the most impact.

A further investigation is possible by drilling down to total accrued costs per farm as the next variable of interest, see the screenshot below.



The left panel displays the options the user has: select the output variable (the variable of interest), an (optional) dimension variable, the maximum number of input variables to visualize (the remaining ones are merged into a category other), whether to show 95%-confidence intervals, and the font size.

Another screenshot below displays the dimension variables enabled.



**Future outlook**

The results of this project have been well received within Wageningen Economic Research. Colleagues are enthusiastic about the developed tool and want to further develop and extent the tool. Currently, Wageningen Economic Research is developing a Data Warehouse system. Colleagues have expressed the wish to further develop the tool and include it in the Data Warehouse system. The possibilities for the coming year are currently being further explored, with the emphasis on obtaining new financial resources.