Math **INDUSTRY**

New techniques and technologies in data driven approaches to sustainability *

Symon Islam[†]   Sebastián Moraga[‡]   Edgar Pacheco[§]

Thomas Pender[¶]   Igor Pinheiro[‖]

Abstract

In this paper, we study several aspects of sustainability in agriculture. The need to provide information and projections to farmers and producers is addressed. Then we study several indicators of the environmental impacts on agricultural production.

1 Introduction

Sustainability is the greatest challenge facing the human race, and in the face of global warming coupled with population increase, the strain on vital sectors, like agriculture, is mounting at an accelerated pace. This project is focused on the use of open data to improve understanding, and, ideally, predictability, for the environmental impact from agriculture.

Agriculture is one of the most significant areas of economic activity for countries around the world, and it has a significant environmental impact; it is estimated that agriculture accounts for at least 10% of green house gas emissions in many countries. In nearly every country, there is an impact from agricultural practices through changing land use, water

[†]Department of Mathematics, Simon Fraser University, Vancouver, British Columbia, V5A 1S6, Canada. `symonyeal@gmail.com`.

[‡]Department of Mathematics, Simon Fraser University, Vancouver, British Columbia, V5A 1S6, Canada. `smoragas@sfu.ca`.

[§]Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta, T2N 1N4, Canada. `edgar.pachecocastan@ucalgary.ca`.

[¶]Department of Mathematics and Computer Science, University of Lethbridge, Lethbridge, Alberta, T1K 3M4, Canada. `thomas.pender@uleth.ca`.

[‖]Department of Mathematics, University of British Columbia, Vancouver, British Columbia, V6T 1Z2, Canada. `igor.ivpp@math.ubc.ca`.

consumption, use of fertilizer, GHGs and other emissions. From a local perspective agriculture creates different dynamics for indigenous plant and animal life in addition to creating different micro-climates. On a more broad scale, agriculture can put pressure on entire river systems and lead to changing weather patterns as atmospheric humidity and solar radiation emissions change.

The ideal outcome of this project is the design of a model and a system of action for environmental impacts that leverages open data and provides a complement to the current TheoryMesh system which is capturing operational impacts from farm activities. Invariably, the project and its outcome will be data driven.

There are many data sources available to investigate environmental changes and impacts due to agriculture. Furthermore, there is a proliferation of data sets that convey information about practices, agricultural production, and green house gas emissions. Combining data across data sources and interpreting the data in new ways could provide better insights on sustainability. Using machine learning techniques to creating models to describe these impacts could improve planning and shift practices to reflect longer term environmental impact. For example, technologies like Blockchain may provide a foundation to create an immutable data ledger for environmental impact while also leveraging smart contracts to take action on data when conditions are met.

For a broader view, the reader is encouraged to read The United Nations Sustainable Development Goals that describe a multi-faceted view of sustainability covering environment, economic and societal factors [6].

Finally, we note that tables and figures are reserved for an Appendix at the end of this document.

2 Models and results

In this section we discuss briefly the models used on the data collected in this research. The models we implemented are simple linear regresion models of the type

$$Y_i = \beta_0 + \sum_{j=1}^{k} \beta_j X_{i,j}, \tag{2.1}$$

for the observations $i \in \{1, 2, \ldots, N\}$, where N are total number of observations, and where there are k dependent variables .

Furthermore, in this study we use the dependence between two specific quantities using a correlation coefficient matrix, with the Pearson's correlation coefficients. This method is widely used to see linear correlation between variables that also uses a least squares fitting to the data obtained.

3   Problems

The issues alluded to above are broad, and they are difficult to quantify and to analyze; indeed, agricultural sustainability is a function of a large number of variables. In an effort to make the problem more tractable, we focused on two smaller problems.

The first problem arose in the following way. There is a lot of work being done on the broader, macro scale. However, it is interesting that not much work has been done on the local, micro level. For instance, a common complaint among contemporary farmers is that information cannot be compiled and disseminated in the same ways anymore. It used to be the case that a producer might say "Plant this crop on this day because that is what has always worked." This is no longer tenable with the changing climate and increased weather volatility.

All this is to say that the productivity of agricultural fields, what producers depend on, is being negatively affected, and new approaches to collecting, interpreting, and disseminating relevent information need to be produced.

As a second problem, we see that environment problems have been a real matter in the last century. Factors that could be correlated, for example greenhouse gases emission, humidity, or heat, can explain why certain aspects of area of production can affect others. The main issue here is try to identify such factors in Canada and the impact they may have on the environment and the surroundings if one or more of them change.

3.1   Productivity of Canadian Farms

There are many areas in the world that are experiencing the catastrophic effects of climate change. In the Middle East, many of the farmers who have had to rely on local tributaries to sustain their crops are having to to consider different ways of life as water levels of these rivers and lakes decrease and, ultimately, disappear. As these climate events unfold, the UN [5] warns of a proliferation of conflict over the increase in water scarcity.

For more northern regions, like Canada, the situation may develop differently. As the climate warms, areas to the north will continue to become more viable for a greater variety of agricultural practices and commodities. For instance, consider the length of the frost free days in the province of Alberta [1].

In each climate model, where we note that rcp26, rcp45, and rcp85 are, respectively, the low, mid, and high emisions models, and as indicated by Figure 1, the number of consecutive frost free days will invariably increase. With the increased length of the growing season, there will be more opportunities to grow a greater variety of crops in regions that have been hitherto difficult to cultivate.

There are other climate variables to consider as well. Observe the cumulative monthly mean temperature in Alberta over the consecutive frost free days shown in Figure 2 [1].

These climate projections per locale can be used by producers in order to project at what points they may profitably consider changes in the location of their business and the type of crops they choose to grow in these locations.

To further evince the increasing productivity of Canadian crop yields, and to provide specific crop data per locale for producers, we can consider data like that shown in Figure 3, which shows the change in barley yields in Alberta [4].

Assuming the yield to be a function of time and a number climate variables, say, the number of days with max temperature greater than 32 °C, the number of frost free days, the cumulative precipitation, and the cumulative monthly mean temperature, we endeavour to use the usual multilinear regression techniques and tools to exterpolate into the future (here we have the popular sklearn module).

We note the projections here are illustrative. The R^2 value for the projections under each climate model, with the variables given above, are each approximately 0.65. Presumably, by using a greater number of climate variables, and by incorporating other variables exogenous to our climate projections, greater accuracy can be attained.

Now, as is clear from Figure 4, the particular climate model we are assuming has a great effect on the projected yield. It is projections such as these that producers may find benificial in making informed choices in a time of increased volatility.

3.2   Indicators of Environmental Impact

In this work, we first stablished the main problems and overall challenges that food process and agriculture may face in the future. The main results showed in Section 2, asserting the existence of a linear correlation between variables from food production. With this in mind we have a better understanding of the indictors for sustainability. Where the preliminary results show why in certain cases, while the poultry production increases per capita some other factors decrease, like beef or pork production. The results of this research are shown in figures 5, 6, 7, 8, 9 and 10.

The previous research indicates there are many other factors to consider, and this is the starting of an opportunity for more investigation about the challenges of food production.

4   Conclusions

Next steps to consider are things such as the following: To look at the sustainability of the economic activities in Canada, and to develop a system to tackle problems as land use, water consumption and use of fertilizers.

The preliminary results obtained in the work are promising, as well the need for further improvements to open problems such as certification for the food suppliers and traceability of the products. These results could be reported in a future work.

References

[1] CLIMATE DATA CANADA Climate Data for a Resilient Canada
    DOI: https://climatedata.ca/

[2] D. MCINNES, Agri-food sustainability targets. A selected overview. DMci Strategies,
    October 2003.

[3] K. PARRIS, Sustainable management of water resources in agriculture. OECD Pub-
    lishing. France, 2010.

[4] STATISTICS CANADA Table 32-10-0359-01 Estimated areas, yield, production, average
    farm price and total farm value of principal field crops, in metric and imperial units
    DOI: https://doi.org/10.25318/3210035901-eng

[5] UNITED NATIONS ClimateChange
    DOI: https://www.un.org/climatechange?gclid=CjwKCAjwvuGJBhB1EiwACU1AidbTCITKc5CBRuumTV

[6] UNITED NATIONS THE 17 GOALS — Sustainable Development
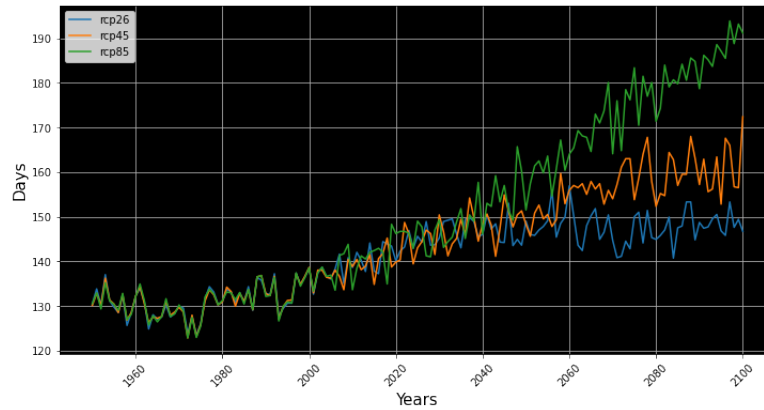   DOI: https://sdgs.un.org/goals

Appendix

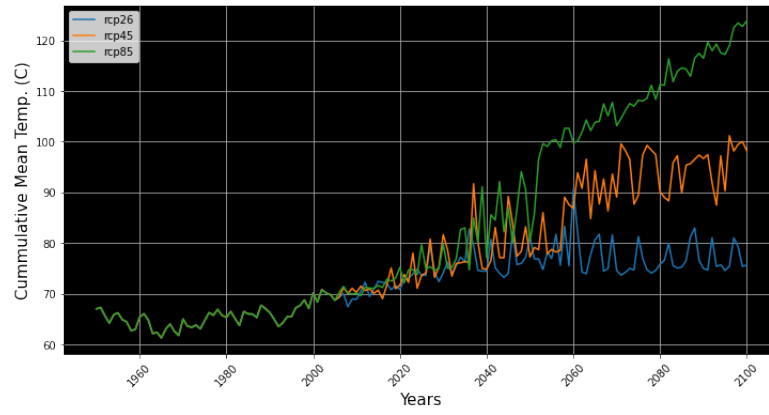Figure 1: Consecutive Frost Free Days in Alberta

Figure 2: cumulative Mean Temperatures in Alberta
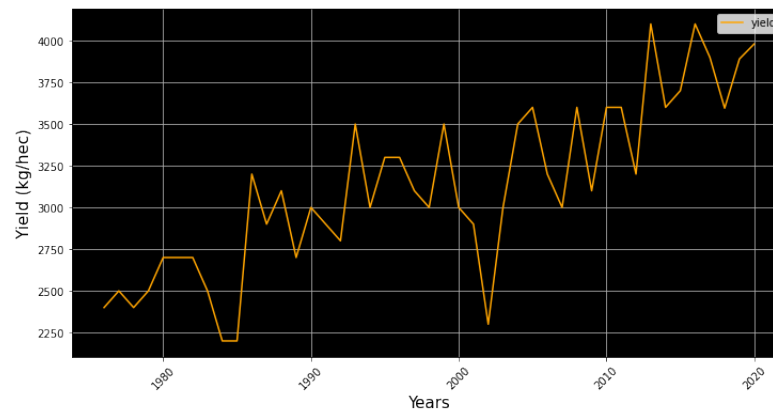
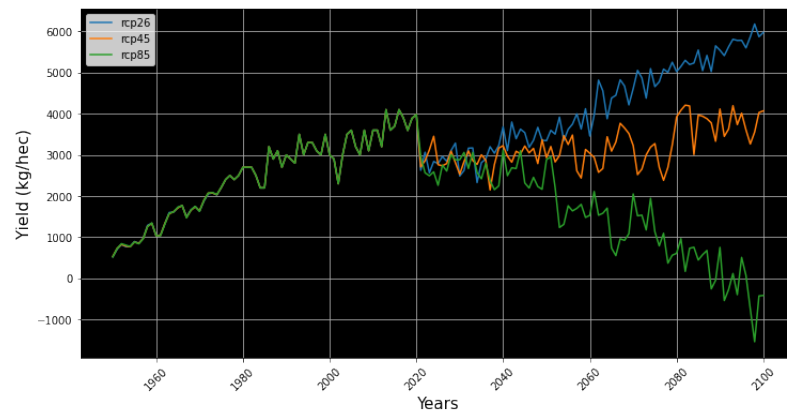

Figure 3: Barley Yields in Alberta
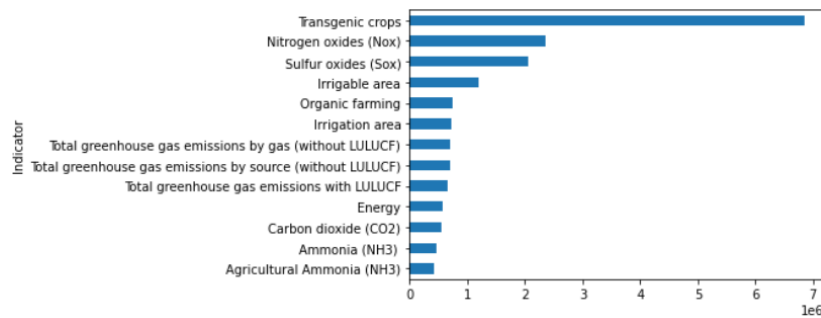
6

Figure 4: Projected Barley Yields in Alberta

Figure 5: Comparison of the different types of indicators values in Canada, from 1984 to 2017. Shows only the indicators over the mean value, where transgenic crops was the most used indicator.
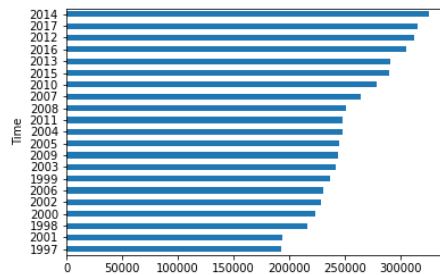

Figure 6: Shows the amount of indicators per capita from 1984 to 2017, starting around 2000000 to near 350000 units per capita. The year where more indicators were used was 2014 and 2017. Besides, the results shown are only the ones above the average within this time frame.
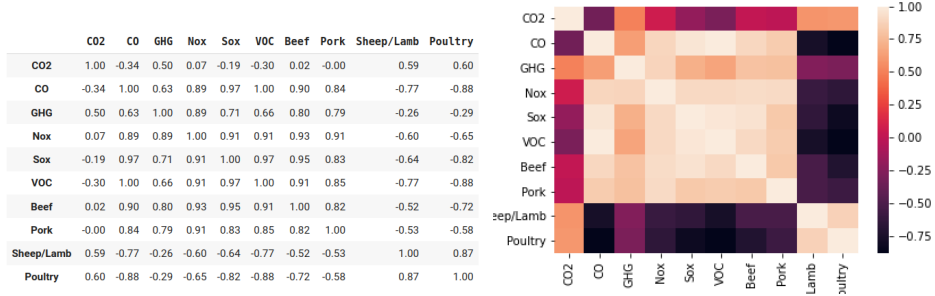
7

	CO2	CO	GHG	Nox	Sox	VOC	Beef	Pork	Sheep/Lamb	Poultry
CO2	1.00	-0.34	0.50	0.07	-0.19	-0.30	0.02	-0.00	0.59	0.60
CO	-0.34	1.00	0.63	0.89	0.97	1.00	0.90	0.84	-0.77	-0.88
GHG	0.50	0.63	1.00	0.89	0.71	0.66	0.80	0.79	-0.26	-0.29
Nox	0.07	0.89	0.89	1.00	0.91	0.91	0.93	0.91	-0.60	-0.65
Sox	-0.19	0.97	0.71	0.91	1.00	0.97	0.95	0.83	-0.64	-0.82
VOC	-0.30	1.00	0.66	0.91	0.97	1.00	0.91	0.85	-0.77	-0.88
Beef	0.02	0.90	0.80	0.93	0.95	0.91	1.00	0.82	-0.52	-0.72
Pork	-0.00	0.84	0.79	0.91	0.83	0.85	0.82	1.00	-0.53	-0.58
Sheep/Lamb	0.59	-0.77	-0.26	-0.60	-0.64	-0.77	-0.52	-0.53	1.00	0.87
Poultry	0.60	-0.88	-0.29	-0.65	-0.82	-0.88	-0.72	-0.58	0.87	1.00

Figure 7: We compare different factors from environment impact with a correlation matrix. (**Left**) Shows the pairwise correlation coefficients between various variables of the Canadian food production, such as GHG and farm supplies. (**Right**) Shows a visualization of the corresponding correlation matrix on the left.
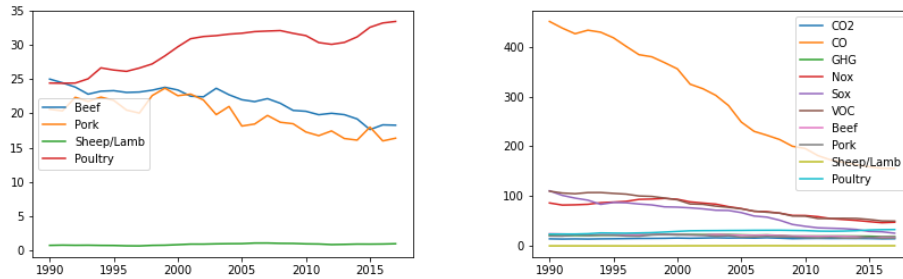

Figure 8: Comparison in time of food supplies production per-capita from 1990 to 2017. (**Left**) Shows the evolution of different products from farms, whereas we observe an increasing number of poultry but a decreasing number of other supplies as beef and pork. (**Right**) Shows the evolution of the farm products along GHG produced in the same time frame.
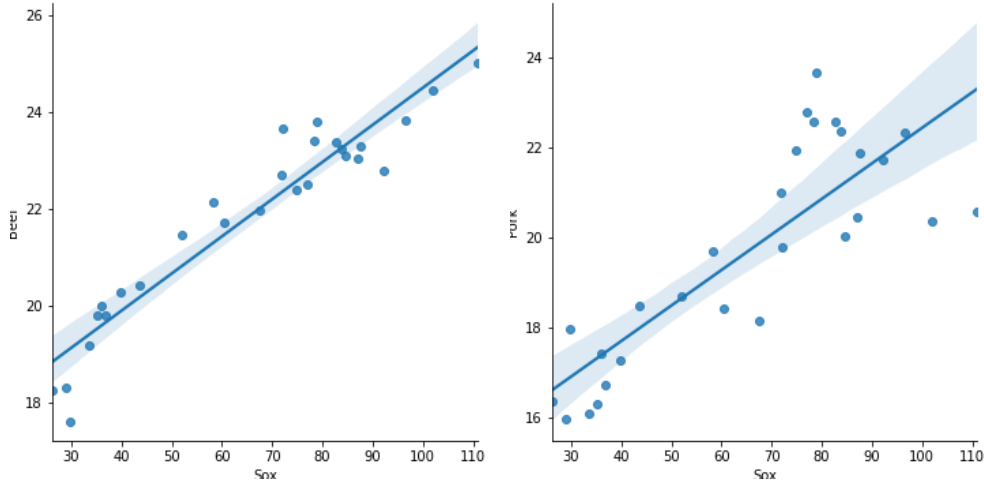
Figure 9: Shows a best-fit line for two variables and the respective confidence intervals. (**Left**) Comparison for a linear correlation between beef production per-capita and Sox gas emited. (**Right**) Comparison for a linear correlation between pork production per-capita and Sox gas emited.
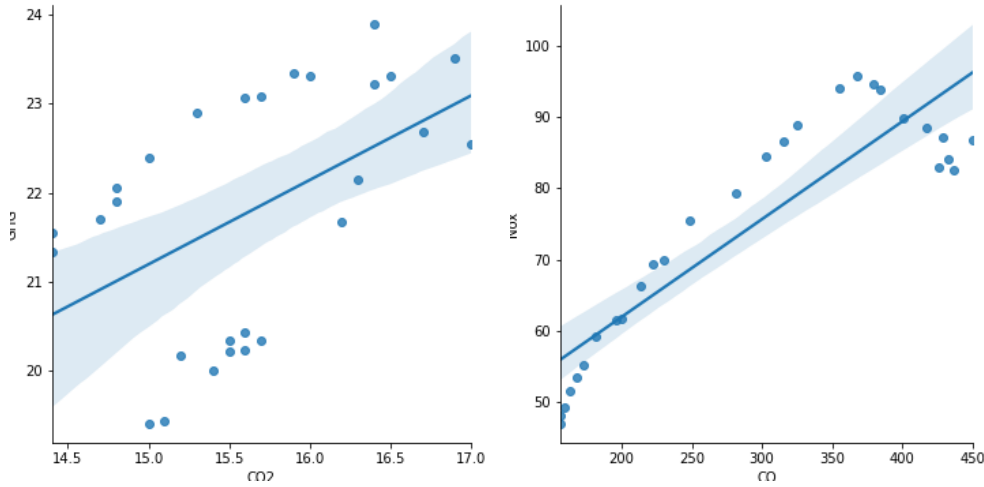

Figure 10: Shows a best-fit line for GHG and one specific gas with the respective confidence intervals. (**Left**) Comparison for a linear correlation between GHG production per-capita and CO2 emited. (**Right**) Comparison for a linear correlation between GHG production per-capita and CO emited.