ResidueGas DELIVERABLE NO. 4.3

Decision support tool for quantifying effects of crop residue management effects on on the GHG balance at regional scale

March 2021

Authors:
KIT: Christian Werner, Klaus Butterbach-Bahl, Clemens Scheer, Edwin Haas

This report is a publicly accessible deliverable of the ResidueGas project. The present work has been carried out within the project 'Improved estimation and mitigation of nitrous oxide emissions and soil carbon storage from crop residues', which is funded in the frame of the ERA-NET FACCE ERA-GAS. FACCE ERA-GAS has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 696356.

This report may be downloaded from the internet and copied, provided that it is not changed and that it is properly referenced. It may be cited as:

Werner, C., Butterbach-Bahl, K., Scheer, C., Haas, E., 2021. Decision support tool for quantifying effects of crop residue management effects on on the GHG balance at regional scale. ResidueGas deliverable 4.3. April 2021.

Table of Contents

1. Summary **2**

2. Introduction **3**

3. Materials and methods **4**

4. Results and discussion **5**

4.1 ResidueGas web app 5

4.2 Existing Web tools for soil carbon dynamics and greenhouse gas emissions 6

5. Conclusions **8**

6. References **9**

1. Summary

A decision support system was developed to assess estimates of the impact of residues management on soil carbon dynamics and nitrous oxide (N_2O) emissions from arable soils for EU-27. The model is based on a random forest machine learning model to identify all drivers for the prediction of target components such as soil organic carbon dynamics and soil N_2O emissions. The random forest model classifies these drivers accordingly to predict the target components. A web App for data exploration was developed.

2. Introduction

In EU-27, around 3.7% of the overall greenhouse gas emissions can be attributed to the cultivation of agricultural soils (Source: European Environment Agency, 2015). Intensification and the use of synthetic nitrogen (N) fertilizers in the second half of the last century has put the EU arable farming under pressure. Under the threat of climate change and global warming, the maintenance of soil organic carbon (SOC) in terrestrial ecosystems and specially in EU arable land is critical for mitigating climate change, long-term productivity and food security for Europe.

Simulation models for biosphere-atmosphere exchange, greenhouse gas (GHG) emissions and SOC dynamics are valuable tools in predicting the impacts of climate change on GHG emissions, carbon (C) storage and developing management strategies for the mitigation of climate change and global warming; however, their utility is generally reduced due to need for specific data.

The CERES-EGC and the LandscapeDNDC (Haas et al., 2013) terrestrial ecosystem models are very complex process based representations of C and N cycle dynamics in terrestrial ecosystems. Both models have been applied in the ResidueGas project to assess the C and N cycling in European arable land cultivation on a spatial resolution of 0.25° x 0.25° degrees from 1950 – 2100 for two regional downscaled climate change projections RCP 4.5 and RCP 8.5. The simulation results contain SOC dynamics of topsoil (30cm) and soil based N_2O emissions under the most common local representation of arable land management. This dataset enables to examine the impact of land and residues management and climate change on SOC storage and losses as well as soil N_2O emissions.

The results were used to train a random forest model (Breiman 2001) to predict SOC change and N_2O emissions based on soil physicochemical properties, climate drivers, and agricultural management. This model acts as a surrogate to the computationally expensive mechanistic models by approximating the simulated C and N dynamics predicted by those models. Finally, a dynamic web App was developed to explore the outcome of the dataset and enabling users to assess the effect of various soil, climate and management conditions on SOC storage and N_2O emissions.

3. Materials and methods

From the work package 4.2 of the ResidueGas project, 16 EU wide inventory simulations for soil carbon and N dynamics and associated N_2O emissions (CERES-EGC and LandscapeDNDC, RCP4.5 and RCP8.5, four residue management scenarios: Baseline, Exported, Surface and Tillage) were available to train two random forest regression models (predicting annual SOC storage change and soil N_2O emissions, respectively). The random forest models are trained on soil properties, climatic conditions and arable management practices and the resulting target values annual SOC change and soil N_2O emissions.

In particular the following covariates were given to the model: model selection, residue management practice, soil physicochemical properties (pH, texture, bulk density), crop type, synthetic and organic N fertilizer rate, CO_2 levels and annual temperature. In total, 16 training datasets simulated with LandscapeDNDC containing 8800 grid cells and annual values for the years 2000 to 2100 were used.

The models were trained as random forest regressors (Python 3.8, Scikit-Learn 0.24.1).
The web App to explore site and management conditions effects on SOC dynamics and soil N_2O emissions was build using Streamlit 0.67 (https://streamlit.io). The training of a random forest model for a problem with over 394 Millions degrees of freedom (16 simulations times 8800 grid cells times 100 years times 16 crops times 2 result quantities) requires average computing time of less than a hour on a modern 12 core server.

A prototype of the ResidueGase web App was set up on a public available internet server under the IP http://195.37.187.130:10206.

4. Results and discussion

4.1 ResidueGas web app

Figure 1 shows the ResidueGas GHG and SOC calculator Web App for decision making. The App offers different selections of the inventory datasets like residue management selection (Baseline, Exported, Surface, Tillage), the 14 different crop types simulated in the EU wide inventory (pulses, sugar beet, food corn, oat, potato, rapeseed, rye, silage corn, soy bean, spring wheat, spring barley, sunflower, winter barley and winter wheat), synthetic N fertilization rate (Urea) and organic N fertilization rate per year [kg-N ha^{-1}], annual average temperature and CO_2 levels, soil physicochemical properties as SOC content [%], bulk density, soil pH, and soil texture (% of sand, silt and clay).

Based on the selection, the random forest model will estimate annual change in SOC content (ΔSOC), and annual soil N_2O emissions [kg-N_2O-N ha^{-1} yr^{-1}].

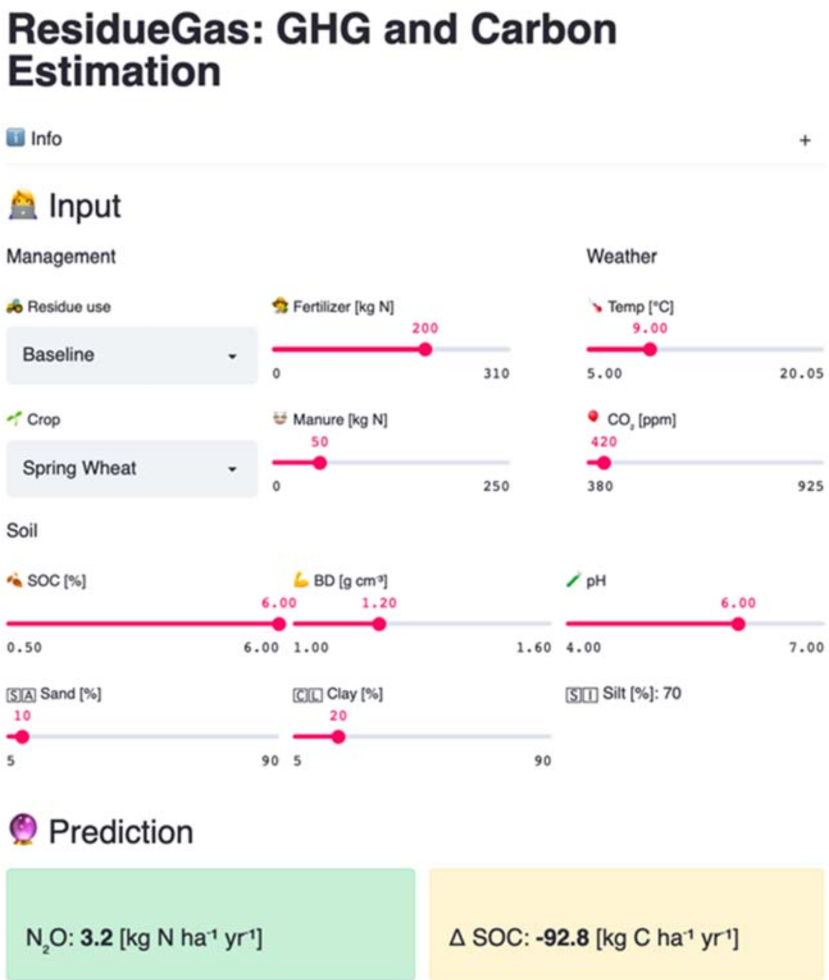

Figure 1 ResidueGas Explorer App Internet Interface, http://195.37.187.130:10206

The web App explores all inventory data from the 16 individual inventory simulations across the EU-27 including different climatic gradients from Mediterranean Malta to Finland, from costal Ireland to continental Hungary, intensity gradients from high intensity arable land cultivation with N fertilization of more than 350 kg-N ha^{-1} yr^{-1} to subsidence based agriculture in Romania and Bulgaria with total N fertilization of less than 30 kg-N ha^{-1} yr^{-1}.

The random forest model enables the prediction of SOC changes and soil N_2O emissions for a wide range of site, climate, and arable management conditions. Machine learning random forest models were recently developed as surrogates for process based models (Saha et al., 2021). Their prediction capabilities based on the large amount of training data have improved drastically in recent years. With the availability of large datasets of observed soil N_2O emissions (Dorich et al., 2021) sophisticated machine learning surrogates for the process based agro-ecosystem models will become available in the near future. Up to now, the available models are purely academic and have not been implemented for any stakeholder use so far.

4.2 Existing Web tools for soil carbon dynamics and greenhouse gas emissions

Cotton Greenhouse Gas Calculator
The Farming Enterprise Greenhouse Gas Emissions Calculator has been especially developed from Queensland University of Technology for use by cotton farmers. It uses site specific soil and climate input data. The methodology to estimate soil carbon dynamics and greenhouse gas emissions is based in IPCC Tier 1 approach. (https://research.qut.edu.au/sae/resources/calculators/cotton/)

Farming Enterprise Greenhouse Gas Emissions Calculator
The Farming Enterprise Greenhouse Gas Emissions Calculator has been developed from Queensland University of Technology to provide an estimate of farm-based emissions in Queensland. The calculator is based on regional input data on soil and climate and estimates a farm scale soil carbon and greenhouse gas balance. (https://research.qut.edu.au/sae/resources/calculators/farming-enterprise-greenhouse-gas-emissions-calculator/)

Cool Farm Tool
The Cool Farm Tool is an online greenhouse gas calculator that is free for growers to help them measure the carbon footprint of crop and livestock products. Its methodology is based on empirical emission factors. (http://coolfarmtool.org/CoolFarmTool/)

SOCRATES online model
The SOCRATES model is a simple process based representation of SOC dynamics in terrestrial ecosystems, which requires minimal data inputs and specifically designed to examine the impact of land use and land use change on soil carbon storage. It also contains a simple yield calculator. It is well advanced over the former simple calulators but does not contain N_2O emissions. (https://research.qut.edu.au/socrates/input)

US Cropland Greenhouse Gas Calculator
The calculator from Michigan State University was created to help farmers, extension educators, agencies, policymakers, and others learn about greenhouse gas emissions from field

crop agriculture in order to make informed decisions about crop management and environmental stewardship. It includes soil and climate data for all districts in the US. The methodology is based on IPCC Tier I or II. (http://surf.kbs.msu.edu)

The Susalps App

The Susalps app is a web representation of the LandscapeDNDC model enabling to estimate soil carbon dynamics changes and N_2O emissions from alpine grasslands in Germany. This is the most advanced tool. It enables to e.g. differentiate effects of different management options (like 5 versus 6 grassland cuts) on SOC dynamics and N_2O emission. But the tool is quite complex to use and the simulations will be performed online. (https://dss.susalps.de/demo/)

More web based online greenhouse gas tools are available from USDA: Animal Agriculture Tools (https://www.climatehubs.usda.gov/archive/content/animal-agriculture-tools-0.html).

5. Conclusions

Compared to the above mentioned existing tools the ResidueGas web App is a tool that can provide estimates of SOC dynamics and soil N_2O emissions with low input data requirements. Nevertheless, the model can only explore existing data in the inventory simulations to identify and classify its drivers. There might be parameter combinations of drivers not being represented at all, which can lead to biases. Therefore, random sets of site simulations should assist the dataset training to ensure that the manifold covers all possible combinations of the parameter space.

At present, the web App enables the selection of the simulation model. For the final application, the both datasets will be harmonized and merged to reduce selection dependency. Further work will consider the application and training of other types of machine learning and artificial intelligence (AI) approaches such as neuronal networks, which enable to interpolate between the data and extrapolate out of the parameter space. Such artificial intelligence approaches enable predictions between the residue management scenarios instead of classifying one of them by the random forest approach. Ultimately, AI approaches will enable predictions of SOC dynamics and soil N_2O emission maps for EU-27 (as done for regionalization of CO_2 fluxes by Reitz et al. (2021) or for ecosystem services by Willock et al. (2018)) for any given percentages of residues remaining on, or exported off the field, incorporated or left on the surface, as been produced by the 16 inventory simulations with the regional process based models described in deliverable report 4.2.

The prediction quality of the machine learning model depends on the quality of the inventory simulations with the process based agro-ecosystem models. The comparison of the simulated inventories have shown uncertainties related to structural differences in the two models. Further uncertainties can be assigned to the input data used for the inventory simulations. The physico-chemical soil data used to initialize the models contain some uncertainty related to its data acquisition and to the spatial aggregation onto the simulation grid. The regional EU-27 dataset of arable land cultivation and fertilization contributes to the overall modelling uncertainty. The management data is partly based on data acquisition from the year 2000 which may be outdated, e.g. in the arable regions in East-Europe which have experienced a severe land use change resulting from agricultural intensification (Kümmerle et al., 2016)

6. References

Breiman, L. (2001) Random Forests. Machine Learning 45, 5–32 https://doi.org/10.1023/A:1010933404324.

Dorich, C. D., Conant, R. T., Albanito, F., Butterbach-Bahl, K., Grace, P., Scheer, C., … van der Weerden, T. J. (2020). Improving N2O emission estimates with the global N2O database. Current Opinion in Environmental Sustainability, 47, 13–20. https://doi.org/https://doi.org/10.1016/j.cosust.2020.04.006

EU Environmental Agency (2015), Agri-environmental indicator - greenhouse gas emissions Statistics Explained, https://ec.europa.eu/eurostat/statistics-explained/pdf-scache/16817.pdf

Gabrielle, B., Laville, P., Duval, O., Nicoullaud, B., Germon, J. C., and Hénault, C. (2006), Process-based modeling of nitrous oxide emissions from wheat-cropped soils at the subregional scale, Global Biogeochem. Cycles, 20, GB4018, doi:10.1029/2006GB002686.

Haas, E., Klatt, S., Fröhlich, A., Kraft, P., Werner, C., Kiese, R., … Butterbach-Bahl, K. (2013). LandscapeDNDC: A process model for simulation of biosphere-atmosphere-hydrosphere exchange processes at site and regional scale. Landscape Ecology, 28(4), 615–636. https://doi.org/10.1007/s10980-012-9772-x

Kuemmerle, T., Levers, C., Erb, K., Estel, S., Jepsen, M. R., Müller, D., … Reenberg, A. (2016). Hotspots of land use change in Europe. Environmental Research Letters, 11(6), 64020. https://doi.org/10.1088/1748-9326/11/6/064020

Pedregosa, F., Varoquaux G., Gramfort A., et al., (2011) scikit-learn, Machine Learning in Python, Journal of Machine Learning Research, Vol. 12, 2825—2830

Reitz O, Graf A, Schmidt M, Ketzler G, Leuchner M, (2021) Upscaling net ecosystem exchange over heterogeneous landscapes with machine learning. *Journal of Geophysical Research: Biogeosciences*, 126, https://doi.org/10.1029/2020JG005814

Saha, D., Basso, B., & Robertson, G. P. (2021). Machine learning improves predictions of agricultural nitrous oxide (N_2O) emissions from intensively managed cropping systems. Environmental Research Letters, 16(2), 24004. https://doi.org/10.1088/1748-9326/abd2f3

Simon Willcock, Javier Martínez-López, Danny A.P. Hooftman, Kenneth J. Bagstad, Stefano Balbi, Alessia Marzo, Carlo Prato, Saverio Sciandrello, Giovanni Signorello, Brian Voigt, Ferdinando Villa, James M. Bullock, Ioannis N. Athanasiadis, (2018), Machine learning for ecosystem services, Ecosystem Services, Volume 33, Part B, https://doi.org/10.1016/j.ecoser.2018.04.004