

Understanding and predicting pesticide use on golf courses using deep machine learning

Interim report

July 2021

Guillaume Grégoire, Ph.D.
Assistant professor, Département de phytologie

Executive summary

This project was initiated in January 2021 and aims to use artificial intelligence to predict pesticide use on golf courses based on different inputs. To do so, we used a database containing 14 000 pesticide applications made on 379 golf courses in Québec during the 2003-2017 period. We developed a model that used Treated area, Administrative Region, Golf course ID, Number of holes, and Year as input variables and actual active ingredient rate (AAIR) as the output variable. Due to the nature of the data, we determined that a hybrid model combining three techniques: Support Vector Machine (SVM), Random Forest (RF) and Grasshopper Optimization Algorithm (GOA) was the most accurate for AAIR forecasting. In the next few months, we will add weather data for each golf course to the model.

Full interim report

This project was initiated in January 2021, after some delays caused by the COVID-19 pandemic. We recruited a graduate student, Isa Ebtehaj, to work on this project as part of his Ph.D. thesis. Data on pesticide use from 2003-2017 was obtained from the Ministère de l'Environnement et de la Lutte contre les changements climatiques (MELCC) in December 2020. This Excel database contains all pesticide applications made on golf courses in Québec (379 courses) during this period, which represents close to 14 000 applications. For each application, this information is provided in the database:

- Administrative region
- Golf course ID
- Number of holes
- Year
- Pesticide type (herbicide, fungicide, insecticide, rodenticide, growth regulator)
- Commercial name
- Federal class
- Registration number
- Quantity applied (L or Kg)
- Treated area
- Active ingredient

- Active ingredient concentration
- Active ingredient total (i.e. Quantity applied X Active ingredient concentration)
- Actual active ingredient rate (i.e. AAIR = Active ingredient total / Treated area)

The first step of the project was to implement and train an artificial intelligence (AI) model that could predict pesticide use (i.e. output or dependent variables) based on a given set of data (i.e. input or independent variables). The Treated area, Region, golf course ID, number of holes, and year were used as the input variables and actual active ingredient rate (AAIR) was the output variable.

Due to the significant difference between the minimum and maximum of the independent input variables as well as the dependent output variable, we determined that a hybrid model was required for AAIR forecasting. Thus, two machine learning categories including tree-based and non-tree-based techniques were checked to find the optimal model for AAIR forecasting. The developed hybrid model combines three approaches in machine learning: Support Vector Machine (SVM), Random Forest (RF) and Grasshopper Optimization Algorithm (GOA). The comparison of the developed hybrid RF-SVM-GOA method with tree-based techniques including M5P, Random Tree (RT), Reduced Error Pruning (REP) Tree, Random Forest (RF), and non-tree-based techniques including Generalized Structure of Group Method of Data Handling (GSGMDH) and Evolutionary Polynomial Regression (EPR), RF-SVM-GOA outperformed other employed machine learning-based techniques.



Faculté des sciences de l'agriculture et de l'alimentation
Département de phytologie

We are currently in the process of writing a scientific paper based on this new hybrid machine-learning approach. We will also present our work to the Québec Plant protection society symposium on artificial intelligence on September 16th, 2021. The next step will be to add historic meteorological data (min temperature, max temperature, average temperature and precipitations) based on each golf course location to the model in order to add a climate component to the pesticide use forecasting.