Crop-yielding Prediction using Neural Network for Stochastic Differential Equation Parameters Estimation

Dr. Ayman Elgharabawy
Biological Data Science Institute (BDSI)
The Australian National University
Canberra ACT 2600 Australia
ayman.gh@anu.edu.au
+61 466054162

Abstract

Crops-yielding prediction is a multi-variant random dynamic and computationally expensive problem. This project employs a collection of stochastic differential equations (SDE) in Itô form using an artificial neural network (ANN) to predict the parameters of the stochastic differential equation of the chaotic time series of wheat yielding daily, monthly, and yearly. We aim to build a novel and unique stochastic wheat model that gives the best prediction based on SDE parameter estimation. The project will define new SDE parameters that can accurately be estimated in SDE using a neural network under specific noise level regimes to predict the crop yielding. The dataset used is the European Commission's MARS Crop Yield Forecasting System (MCYFS) including spring barley, potato, grain maize, and sunflower. The historical data are in daily and yearly scales.

Research Objective

High-performance wheat-yielding prediction model for large-scale data by using the stochastic approach with machine learning in parameter estimation, which reduces the computational time compared to a traditional and deep neural network that has many overfitting issues due to unbalanced data.

Research Output

In conducting this study, we will accomplish the following specific deliverables:

- Novel Crop SDE model that can predict crop yielding and a specific number of parameters with new drift and diffusion functions.

Introduction

Crop yielding prediction is a complex problem. It is considered a multi-variant random dynamic system because it has hundreds of factors based on weather, soil type, genotype, and phenotype variations. Those variations include almost thousands of factors. Due to data scaling, this large data multi-variant is computationally expensive using traditional or deep machine learning because it has thousands of genotype, phenotype, and environment variable features. These invariant and hidden factors are similar to stock exchange rate prediction.

Research Method

Parameter Estimation

Given a one-dimension time-homogeneous SDE:

$$dX = \mu(X;\theta)dt + g(X;\theta)dW \tag{1}$$

the task is to estimate the parameter θ from a sample of (N+1) observations $X_0, X_1, ..., X_n$ of the process at known times $t_0, t_1, ..., t_n$. In the statement of equation (1), dW is the differential of the Wiener process (Brownian motion), $\mu(X;\theta)$ is the instantaneous drift, and $g(X;\theta)$ is the instantaneous diffusion.

Crop SDE Model

The SDE model will propose new functions for the drift and diffusion SDE parts. The output model is evaluated by Euler and Stratonovich methods.

Research Activities

Activity	**Timeline**
Implementing the SDE model based on small testing data using Python	1-3 months
Tuning the model by adding parameters and functions	1-3 months
Implementing NN that changes the parameters based on the output performance	2-3 months
Optimize the ANN hyperparameters, the SDE functions, and number of yielding on real datasets	1-2 months

Table 1: Activity and timeline

Skills Required

Advanced math, Stochastic Differential Equations.

References

- Zhi Xie, Don Kulasiri, Sandhya Samarasinghe, and Channa Rajanayaka. The estimation of parameters for stochastic differential equations using neural networks. Inverse Problems in Science and Engineering, 15:629–641, 2007.
- Geng Yuan, Hiroaki Mukaidani, and Tadashi Shima. Uncertainty quantification in neural networks using stochastic differential equations. 2023 62nd Annual Conference of the Society of Instrument and Control Engineers (SICE), pages 1543–1548, 2023.
- Aldi Eka Wahyu Widianto, Endah R.M. Putri, Imam Mukhlash, and Mohammad Iqbal. European high-dimensional option pricing using backward stochastic differential equation-based convolutional neural network. Proceedings of the 2023 6th International Conference on Mathematics and Statistics, 2023.
- Dilli Paudel, Hendrik Boogaard, Allard de Wit, Sander Janssen, Sjoukje A. Osinga, Christos Pylianidis, and Ioannis N. Athanasiadis. Machine learning for large-scale crop yield forecasting. Agricultural Systems, 2020.