AIDI 1002

Project EDA and SOW V2

Gavin Rowsell

Using the historical weather data from 2019 for the typical 8-week growing period (July to August), can we create a model to predict temperature within a greenhouse during daytime (6AM to 9PM)?

Due to the lack of available data for temperature and condition measurements within a greenhouse, a function was designed through my own research and industry insight from a fellow colleague to provide realistic internal temperature of a diffused glass, 30 feet by 96 feet greenhouse without any control systems in place. Two models can be created with this data each with their own uses. The first will be helpful to those deciding whether to invest in a greenhouse to grow their desired plants however it does not explore the usefulness of machine learning effectively due to the use of the synthetic data in model creation. The second will provide a better example of the benefits of machine learning and will be useful to those who already have a greenhouse they are maintaining with various control systems.

1st model uses:

If you have a specific crop you would like to grow, this model to predict the temperature of a small-scale greenhouse. By checking average temperatures and extreme values for the desired growing period one can get an idea of either what season to grow their plant or what control systems they may have to invest in to better maintain their plants. Maybe you want to know if investing in a small greenhouse will be beneficial for you. By using the average temperature in your area for a certain time period you can see how hot or cold the greenhouse will get and compare that to the ideal temperature range for your specific plants. For example, if you wanted to grow tomatoes which grow best at night in temperatures of 20 degrees Celsius and best during the day at 26 degrees Celsius, you can use this model to see what the high and low points inside a non controlled greenhouse will be in your area. Maybe in your location, the predicted night temperature stays within the desired range however the day temperature is too hot. The model has shown you that some ventilation will be needed if you want to grow during the months you input.

2nd model uses:

For those who already have a greenhouse with control systems the second model can be useful. The additional feature of this model is the inside temperature 1 hour ago column. This addition of this feature allows users of the model to input hourly forecast data (which has become more and more accurate) to step forward into the later hours of the day and predict the temperature of their greenhouse depending on changes they make to its current temperature. As an owner lets consider that you have some experience with your heating and cooling systems, and you know the range of temperatures you can use your systems to achieve. You can then use that temperature along with the hourly weather data you will have access to, to predict the future conditions of your greenhouse. Let’s imagine a hypothetical scenario, you have a greenhouse that you are growing tomatoes in and you would like to keep the temperature from going above 26 degrees Celsius. It is currently 10 AM and you realize the weather today is going to climb quite high and it is very sunny out thus you are worried about the temperature of your greenhouse. Using the first model with the hourly forecast data you realise that without intervention the temperature of your greenhouse will exceed the desirable range for your crop at 1PM. You may think: “No problem, I can open my vents around 1 to keep my temperature down” however in the forecast the wind is supposed to pick up considerably at that time as well. You don’t want to risk damaging your ventilation system or have dust and other foreign substances blowing into your greenhouse so you can’t open your vents at 1PM. By using this second model, you are able to step forward through the hours and realise that if you can get your temperature down to 23 degrees by 12 Noon, your temperature will no longer exceed the maximum value of 26 degrees at 1 PM even though you have to turn your vents off due to the high winds. The uncovered relationship between the previous hour internal temperature along with the outside conditions to the future temperature of the greenhouse can help one manage their temperature while considering other factors. Using this model, one can create a system superior to a traditional reactive system (a system that activates when the temperature moves outside a set point).

As a proof of concept, the model 2 discussed above is a demonstration of the usefulness of machine learning to apply data in order to improve the decision-making process in complex situations such as the management of the micro or macro climate of a greenhouse. With more accessibility to data, not only would the accuracy of a proposed model improve but you may be able to make even more valuable predictions such as the compensation for over and under shoot of large control systems which was described in the initial SOW.

Below is the method developed to generate realistic greenhouse temperatures for the dataset which will be used in the model’s creation.

Assumptions:

Diffused glass material, 30 feet wide and 96 feet long greenhouse oriented with the longer side pointing north to south. This orientation is assumed so the wind direction plays a role in how the wind speed will affect the temperature (wind hitting a larger surface area will cool the surface more).

<u>Effects of various features on inside temp:</u>

Outside temp – Use as base, offset applied according to the other features

Time – Used with outside temp, the greatest increase will be ~ 10 C at noon add noise for UV index, no effect on temperature before 6 and after 8

Wind Speed – Greenhouses on windy hills will typically need ~20% more heat (greater material heat loss)

Condition – Indicates the strength of the time of day factor when considering sun heating add noise for different cloud cover amounts.

Method to generate inside temp feature:

insideTemp = outTemp + (rand(9, 11)/ (((12 - time)^2)/4) + 1) * sqrt((1/Con) + ((outTemp * 1 – sqrt(windspeed* windDir/10) /rand(30, 50))

The rand functions will add a reasonable amount of noise to the final value. Temperatures are in Celsius, wind speed is in Mph, and Time given as an hour between 6 and 21 inclusive (6 AM to 9 PM). The condition value (Con) ranges from 1 – 4.5 depending on cloud cover (1 = clear sky, 4.5 = dark sky). The wind direction ranges from 1 – 3.5 and again is used to account for the surface area the wind will be hitting.

Sources:

Weather Data: https://www.wunderground.com/history/daily/ca/windsor/CYQG/date/2019-7-1

Greenhouse Temp Research Links:

https://extension.uga.edu/publications/detail.html?number=B792&title=Greenhouses:%20Heating,%20Cooling%20and%20Ventilation#:~:text=One%20air%20exchange%20per%20minute,of%20about%205%20degrees%20F.

https://www.sciencedirect.com/science/article/pii/S0895717707002130

https://www.researchgate.net/profile/T_Boulard/publication/274541763_Effect_of_Roof_and_Side_Opening_Combinations_on_the_Ventilation_of_a_Greenhouse_Using_Computer_Simulation/links/55eec63e08ae199d47bf2a19/Effect-of-Roof-and-Side-Opening-Combinations-on-the-Ventilation-of-a-Greenhouse-Using-Computer-Simulation.pdf