# Studentship Project: Annual Progress Report 10/2020 to 03/2025

| | | | |
|---|---|---|---|
| **Student Name:** | **Jack Stevenson** | **AHDB Project Number:** | **SF/TF 170/a** |
| **Project Title:** | **Precision Agriculture: Using AI and Expert models to predict strawberry yields with respect to Powdery Mildew** | | |
| **Lead Partner:** | University of Lincoln/NIAB EMR | | |
| **Supervisor:** | Professor Stefanos Kollias, Dr. Mark Else | | |
| **Start Date:** | 24/09/2020 | **End Date:** | 25/03/2025 |

## 1. Project aims and objectives

Project Aims:

- Map variation of temperature and humidity using multiple observation signals, with the final goal to utilise minimal observation signals to map to different spatial resolutions.
- Use the developed variation model to estimate yield and Powdery Mildew Disease risk.

Current Objectives:

- Utilize different signals from the raw data to build a simple function of tunnel variation and fit that to a different tunnel design/setup.
- Use this estimated state to perform state forecasts for the different signals to the growing system
- Apply these state forecasts to models of disease and yield.

## 2. Key messages emerging from the project

It is likely that we can accurately estimate the future temperature and humidity of a polytunnel using Artificial Intelligence. We have developed more complex models which are able to utilise both temperature and humidity data from polytunnels to estimate future temperatures and humidities to a validation accuracy of approximately 50% within a boundary of $0.05°C$ and 5% relative humidity.

## 3. Summary of results from the reporting year

From the previous year, we identified that the data collection aspect of this project was of key importance, as machine learning models require high quality data to provide the most accurate results. Whilst plans were put in place for data collection this past year, a break in studies was taken during the main data collection period, so whilst some data was collected, and the method used was more reliable, we can neither confirm nor deny the effectiveness of these approaches as part of the data collection exercise. Likewise, the live deployment of the model was also not implemented, so we cannot comment on it at this time.

**Architecture Development**

The first part of this year's work involved utilising multiple different model architectures on a single quantity to identify which architecture will enable the most accurate multiday forecast.

The three architectures used are a convolutional based feature extraction, a transformer-based feature extraction and transfer learning using the Resnet model as a backbone.

The convolutional approach utilises 5 blocks of layers in the form of [Convolutional, Convolutional, Skip and Add, Dropout, Convolutional] to extract features from the dataset. The transformer approach also utilises 5 blocks of layers which implement the transformer architecture utilising normalisation, multi-head attention, dropout and convolutional layers. We also implement these architectures using Dense layers, as a point of contrast between them. Finally, the ResNet approach uses the ResNet50 Architecture to first extract information from the data, which is then used by a final layer to generate predictions.

Once we have designed and implemented these models, we then train them to compare how well they can predict and generalise the task. To do this we use a normalised mean square error, calculated as

$$NMSE = \frac{\sum \frac{(T^2 - P^2)}{T^2}}{N}$$

Equation 1: Normalised Mean Square Error

The results of this can be seen in figure 1a, which shows the NMSE for each model as a group, when comparing the results of training the models on different amounts of looking ahead, what we will be referring to as a time skip. E.g. a time skip of 1 means 1 time step into the future, whereas a time skip of 12 represents 12-time steps into the future.

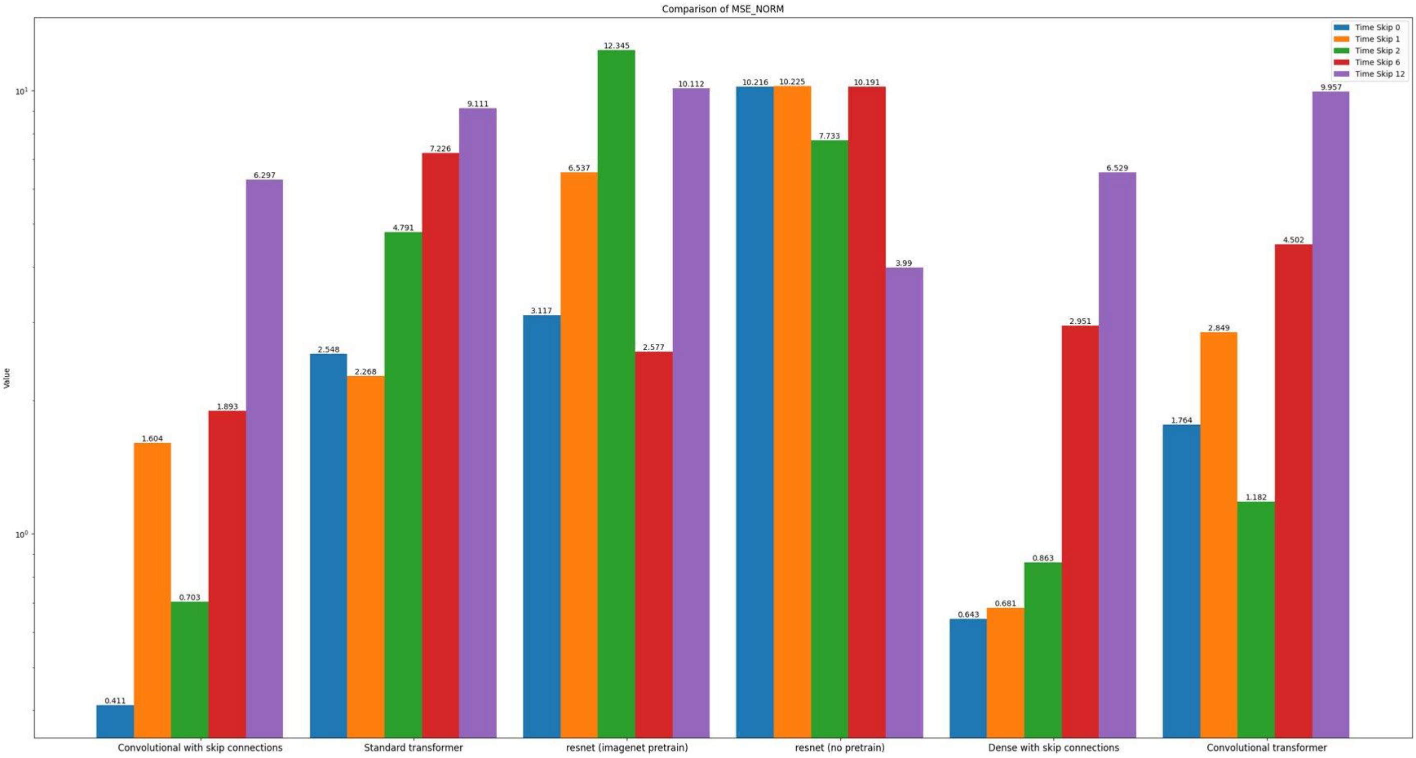


*Figure 1A: Comparison of NMSE during training for each model developed. A smaller NMSE means a better model.*

From figure 1A, we can see that all models produce progressively worse NMSE values as we look further ahead into the future. However, there are some architectures, such as the Dense based network with skip connections showing promise in all time skips, whereas other architectures only show promise in one or two unique time skips, such as both convolutional based architectures performing better on average than all of the other architectures when using a time skip of size 2.

Figure 1B shows the same models but utilises the MSE metric instead of NMSE, and generally supports the trends shown in Figure 1A, suggesting that we should be using Dense models with skip connections as the primary model architecture to build on top of.

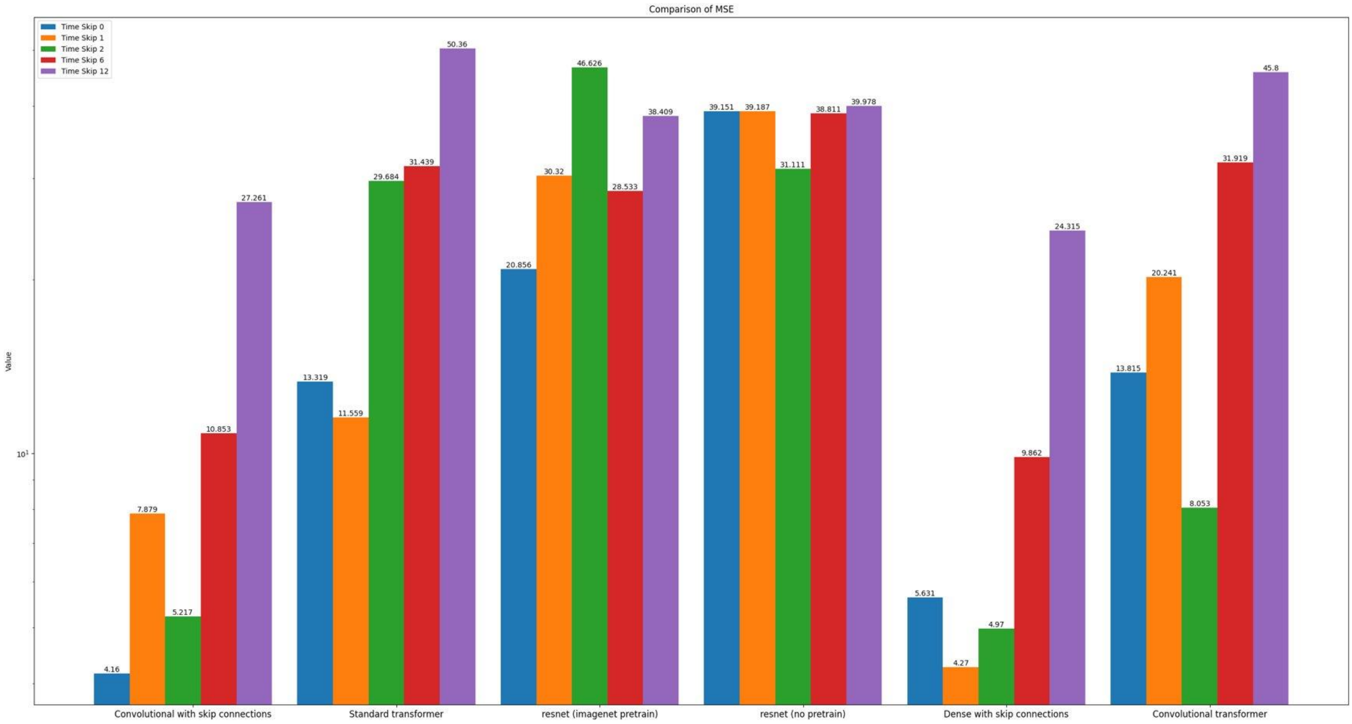

*Figure 1B: Comparison of model architectures using the MSE metric during training*

There are 2 parts to utilising an AI model, the training, which we have discussed above, and testing or the validation of the models. Figure 1C shows the predictions of the convolution model with skip connections (Block 1 on figures 1A and 1B) for time skips 0, 1, and 6 for temperature observations, and comparing them to the ground truth observations for approximately 500-time steps. Figure 1D shows the same information, but for the standard transformer model (Block 2 on figures 1A and 1B). We chose to use time skip 0 as part of our work is to be able to estimate the tunnel environment at different spatial locations given what we know about the history of the tunnel environment, then time skips 1 and 6 were used to provide variance on near predictions and far predictions. Both models can capture the trends shown in the data at all time steps tested, however they tend to overestimate the predicted value, with an upper bound of up to 23% error over all time steps.

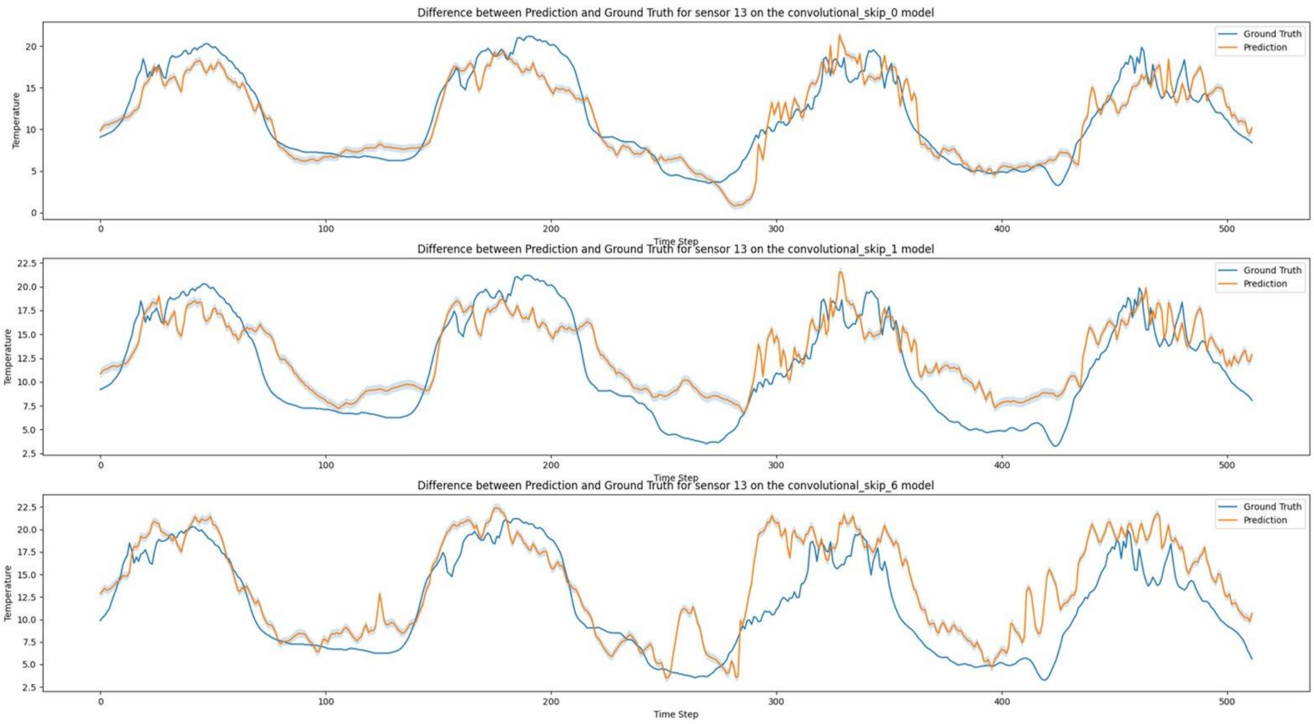


*Figure 1C: Comparison of predictions (orange) and ground truth (blue) for the convolutional skip model for the temperature at time skips 0, 1, and 6.*

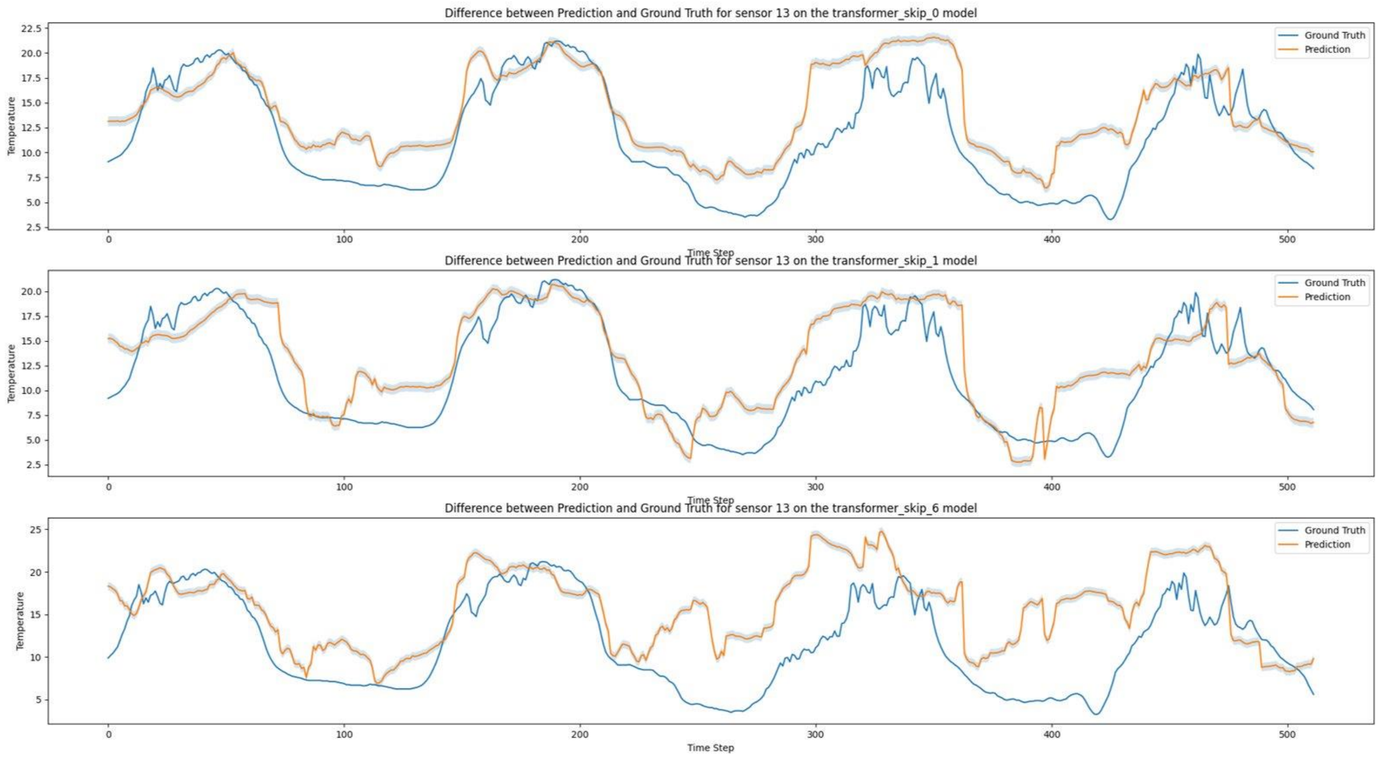


Figure 1D: *Comparison of predictions (orange) and ground truth (blue) for the Dense transformer model for the temperature at time skips 0, 1, and 6.*

When calculating the accuracy of these models, we use 2 metrics, an accuracy that respects a tolerance (Equation 2) of ±0.5 degrees C and Standard Accuracy (Equation 3). This results in table 1E, which show the accuracies for each graph above. Both metrics shows that the convolutional model performs better on average, and that future predictions can be of good accuracy no matter how far into the future it is.

$$Acc_{tolerance}(G,P,T) \; = \; \frac{\sum (G-T)<P<(G+T)}{N} \text{ (Equation 2)}$$

$$Acc_{standard}\;(G,P) = \frac{\sum \frac{(G-P)^2}{G^2}}{N} \text{ (Equation 3)}$$

Equations 2 and 3: Parameter G refers to the Ground Truth, Parameter P refers to the model predictions, Parameter T refers to the accepted Tolerance and N is the number of samples.

| Model | Number of steps skipped | Accuracy with Tolerance | Standard Accuracy |
|---|---|---|---|
| Convolutional | 0 | 21% | 5% |
| | 1 | 10% | 20% |
| | 6 | 12% | 23% |
| Transformer | 0 | 10% | 31% |
| | 1 | 13% | 28% |
| | 6 | 07% | 24% |

*Table 1E: Table showing the standard accuracy and tolerant accuracy of the models for each of the graphs in figures 1C and 1D*

**Combining Models**

It should be clear that while the currently developed models are able to provide some prediction, they do not provide very accurate information for growers to make informed decisions. Therefore, we have taken into consideration the application of ensemble learning. Ensemble learning uses multiple models to perform the same task, and then performs an aggregation function on top of the individual predictions. Multiple model machine learning uses this intuition to generate multiple models, of which none can solve the complete task, but some smaller subcomponent of the task which can enable the model to express more complete information as part of its predictions.

We use this idea to train a model to predict whether a given temperature and relative humidity (TRH) pair is observed in the daytime or nighttime (DNP). This is then used as an accompaniment to the main model, adding an additional feature to the training data that the model can then learn from. Caution does need to be exercised as many machine learning models are susceptive to a garbage-in garbage-out frame of reference.

We then train the DNP on the data to obtain a prediction for whether a TRH pair occurs in the daytime or nighttime. We then implement a custom ANN layer which calls the DNP model on each time step of the training data. This prediction is then concatenated with the original input to provide an extra feature the model can use.

In Figure 2A, we compare the NMSE of a baseline model and the concatenated model. Both models utilise the same architecture, aside from the first layer of the concatenated model, which implements the above procedure. We can see that the concatenation of these extra features lowers the overall NMSE of the model.

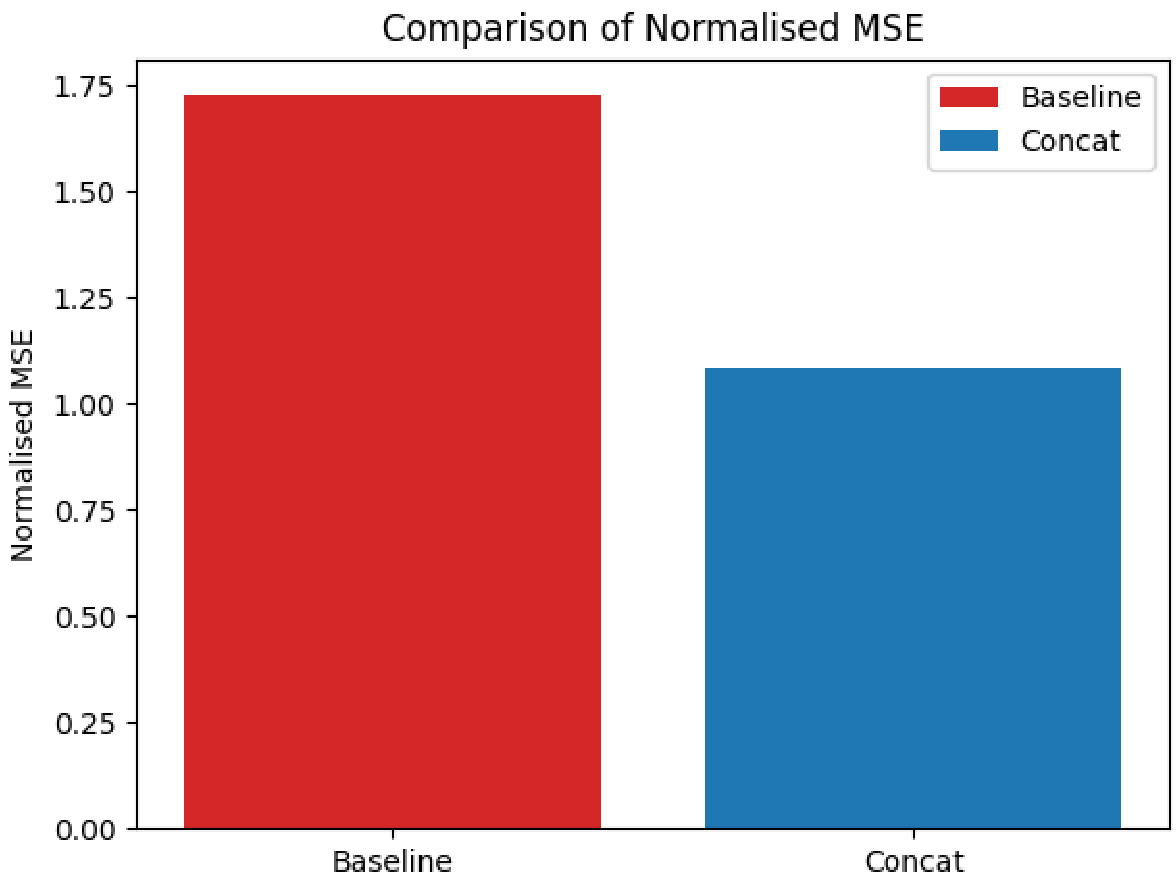


*Figure 2A: Comparison of NMSE for the Baseline and Concatenated models. A smaller NMSE is better.*

**4. Key issues to be addressed in the next year**

We will be focusing on 2 main points of action in 2024. The first point of action is to improve our model in model approach: i) to achieve better accuracy and ii) enable it to be used in the framework of both Temporal domain and Spatial domain prediction. This can be done, by imputation of (sensors') missing values, and application of the trained models in spatio-temporal setting. Imputation of missing values will be attained by using Singular Value Decomposition or Fourier transformation, with removal of noisy values and inverse computation. Furthermore, the model will be iterated upon to provide an ensemble of approaches, including concatenation of features, imputation of missing values and regression, all targeting to improve model performance. The second point of action is using our developed system on edge devices to predict what the future environmental state of the polytunnel will be, in both short-term accuracy and long-term predictions.

**5. Outputs relating to the project**
*(events, press articles, conference posters or presentations, scientific papers):*

| Output | Detail |
|---|---|
| | |

| | |
|---|---|
| | |
| | |

## 6. Partners (if applicable)

| | |
|---|---|
| **Scientific partners** | |
| **Industry partners** | |
| **Government sponsor** | |