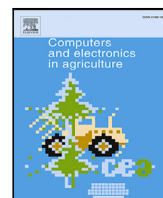




Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Premonition Net, a multi-timeline transformer network architecture towards strawberry tabletop yield forecasting

George Onoufriou^{a,*}, Marc Hanheide^a, Georgios Leontidis^b^a School of Computer Science, University of Lincoln, Lincoln, LN67TS, Lincolnshire, UK^b Interdisciplinary Centre for Data and AI and School of Natural and Computing Sciences, University of Aberdeen, Aberdeen, AB24 3FX, Aberdeenshire, UK

ARTICLE INFO

Keywords:

Deep learning
Strawberry yield forecasting
Time series
Transformers

ABSTRACT

Yield forecasting is a critical first step necessary for yield optimisation, with important consequences for the broader food supply chain, procurement, price-negotiation, logistics, and supply. However yield forecasting is notoriously difficult, and oft-inaccurate. Premonition Net is a multi-timeline, time sequence ingesting approach towards processing the past, the present, and premonitions of the future. We show how this structure combined with transformers attains critical yield forecasting proficiency towards improving food security, lowering prices, and reducing waste. We find data availability to be a continued difficulty however using our premonition network and our own collected data we attain yield forecasts 3 weeks ahead with a testing set RMSE loss of 0.08 across our latest season.

1. Introduction

Precise and accurate yield forecasting is a key component in Fresh Produce (FP) Supply Chain Management (FSCM), since it plays a critical role in price negotiations, logistics, and scheduling. In particular accurate yield estimates, are required a minimum of 3 weeks ahead (in the strawberry domain) which we call the horizon (Fig. 1), so that adequate time can be given to bidding, labour timetabling, logistics, and procurement. However, forecasting FP is incredibly difficult especially over a 3-week horizon where any number of variabilities can exist such as environmental fluctuations. Dealing with the latter would necessitate some weather forecasting to be considered, which is a problem in its own right. Instead, we show how good yield forecasting can be and improve upon current practices while allowing for future works to delve specifically into weather forecasting.

Yield forecasting is difficult in particular due to the lack of data needed to develop forecasting models. Such data is mostly non-existent, or incredibly difficult to attain, largely because of the difficulty around data collection, the perceived sensitivity with which this data is held, and the lack of clear benefits to the digital collection of such data. We also see resistance to the positive dynamic impetus of modernisation requiring a departure from growers' previous fixed practices.

FP optimisation is of global strategic importance since horticulture and agriculture are some of the biggest producers of greenhouse gasses, such that there can be a significant benefit to optimising production

or minimising waste. In the UK our government has committed to reducing greenhouse gasses to net 0 by 2050, and agriculture has been expressly named as a key contributor of greenhouse gasses in the United Nations Climate Change Conference 2021 (COP21). Inaccurate forecasting or more specifically under/overestimation leads to food waste and destruction costs or importing of FP from abroad. Assuming the cause of this discrepancy/variability is adverse weather conditions, then those same weather conditions will have affected geographically approximate growing sites. In the UK climate discrepancies usually mean fruit must be imported from abroad, given our size, to meet any given procurement contract, as all the neighbouring growing sites will have suffered the same adverse environmental conditions and thus under-production.

Other works (outlined with more detail in Section 1.1) have sought to solve the lack of data availability in agriculture using satellite/remote-sensing data, using various machine learning, statistical, and some deep learning techniques. In this paper, we show how we can collect data at some scale but with local/ high granularity, including fruit images, weather conditions, and irrigation data locally. Here we shall focus specifically on strawberry yields of strawberry tabletop and how we can predict them. We exemplify this approach at our Riseholme strawberry tabletop/polytunnel growing site and employ this data to create accurate forecasts with this 3-week horizon/ window to meet the needs of the bidding and procurement process. We do all

* Corresponding author.

E-mail address: gonoufriou@lincoln.ac.uk (G. Onoufriou).

<https://doi.org/10.1016/j.compag.2023.107784>

Received 15 November 2022; Received in revised form 6 March 2023; Accepted 16 March 2023

Available online 23 March 2023

0168-1699/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

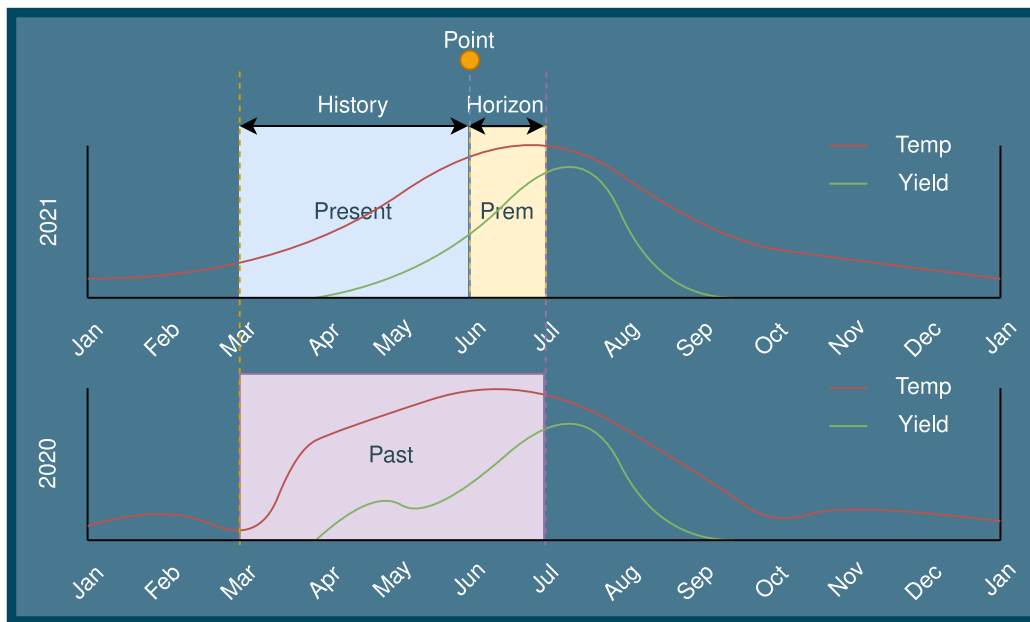


Fig. 1. Past (purple-pink), present (blue) and premonition (yellow) timelines/ windows overlaid on a depiction/rough reference of strawberry yields through the years of 2020 and 2021 along with temperature. Depicting the point of prediction relative to (at the seam of) horizon and history.

this in collaboration with Berry Gardens Growers (BGG), one of the UK's largest soft, and stone fruit producers, and with their direction on industry standards to keep as close to the typical expectations as reasonably possible. We also have fortnightly visits by agronomists to ensure we are growing the strawberries satisfactorily.

We use this data in various neural network architectures in Section 2 and evaluate their performance in Section 4, since the literature would suggest that deep learning approaches are the most performant even for FP. Of these new architectures, we showcase our Premonition Network which seeks to improve upon current tabular/sequence prediction approaches using all three forms of context, the past, the present, and the premonition of the future. We use the past to learn the overarching distribution, we use the present to set some scale and granularity, and we use the premonition for variability from the standard distribution.

1.1. Related work

There are relatively few works in strawberry yield prediction using deep learning, instead, the majority focus on statistical machine learning, and almost none refer to privacy considerations (Maskey et al., 2019; van der Velde and Nisini, 2019; Jafari et al., 2020; Bouras et al., 2021; Gastli et al., 2021; Hopf et al., 2022; Paudel et al., 2021; Zhu et al., 2022; Bali and Singla, 2022). However, several papers have stressed that a lack of data availability (Pearson et al., 2019; Durrant et al., 2021, 2022), or more specifically a high expense of acquisition significantly hinders the smooth application of state-of-the-art neural networks towards the creation of powerful forecasting models (Chen et al., 2019; Maskey et al., 2019; Jafari et al., 2020; Gastli et al., 2021; Nassar et al., 2020). Many of the aforementioned papers largely choose to tackle this lack of data by using satellite imagery although in some cases they use the California strawberry commission data paired with the California strawberry commission irrigation management information system (CIMIS). Unfortunately, the data mentioned in these papers are behind multiple walls, and the CIMIS data is currently unavailable from the original source, so while we were able to find an excerpt of the CIMIS data elsewhere we were unable to find the full dataset making it very difficult to compare to.

Many different proposals for methods of predicting/forecasting yield (generically) exist, some using classical machine learning (e.g. Paudel et al., 2021) others such as those by Nassar et al. (2020) use

neural networks in their specific case a mixture of CNN, LSTMs, GRUs and some attention heads. However, all emphasise the need for better forecasting systems as demand increases and supply decreases due to global factors such as (but not limited to) COVID-19 and the Russia-Ukraine war. Current yield forecasting methods are highly archaic, often times they can be as simple as forecasting the average of the last few years' yields, or simple linear models based on heat hours. One such example is the European Commission's MARS crop forecasting system (MCYFS) which has purportedly seen no improvement in its forecasting performance since 2006 and uses no machine learning. Lastly, the work by Paudel et al. (2021) shows that machine learning can already at the very least match (at the start of the season) or beat existing large-scale traditional crop yield forecasting systems such as the aforementioned MCYFS system.

The MCYFS system from 2006 to 2015 has a median MAE of 0.379, 0.368, 0.570 in soft wheat durum wheat and grain maize (van der Velde and Nisini, 2019). The most performant forecasts for this system appear to be sunflower yields at 0.162 MAE. However the assessment carried out by van der Velde does not state over what period these yield predictions are made specifically whether that be a few weeks, days or months ahead making this also a difficult comparison to make. It is also apparent that forecasting is becoming increasingly difficult with the higher degree of variability in climate conditions as the performance of this largely static forecasting system seems to be in slow decline (van der Velde and Nisini, 2019).

More sophisticated and bespoke machine learning for strawberry tabletop forecasting has still a long way to go, with only a handful of applications appearing in the literature. However, as previously stated data is incredibly difficult to attain in this domain. Nassar et al. (2020) appears to show how the compound deep learning models outperform standalone deep learning models and traditional machine learning models. Nevertheless, as with much work in this space, it is difficult to garner any concrete comparable statistics. From one of their diagrams (14) we believe we can see their most performant model to produce an MAE loss of roughly 0.14 or 14% MAPE. They call this model Attention-ConvLSTM2D. While we do not have access to the same data as they have, we have seen even simple GRU models attain similar performance in our strawberry tabletop. However, we believe we can improve this performance on our own data by means of attention as their paper

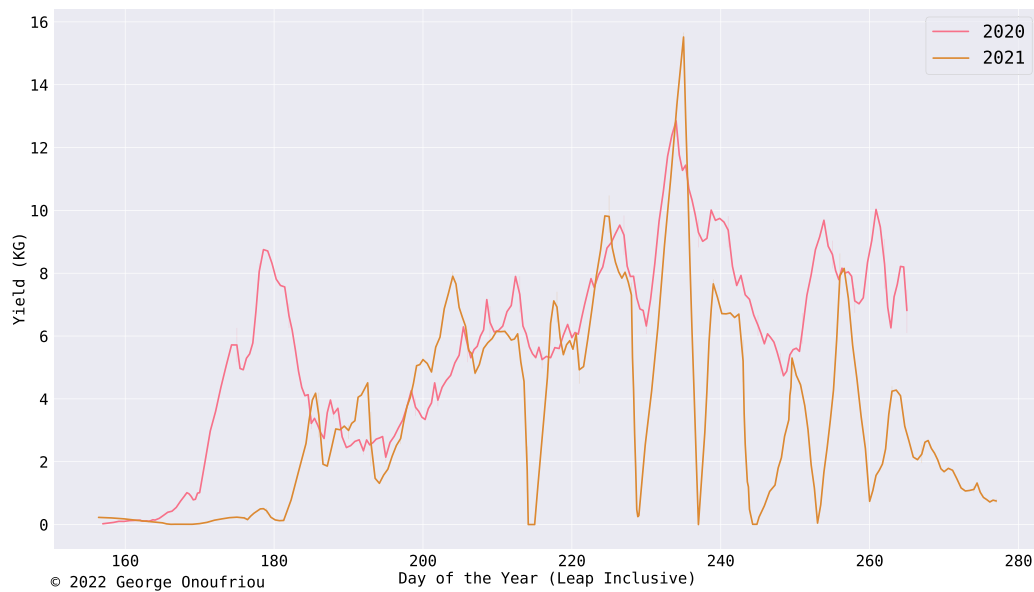


Fig. 2. Seven day rolling average line-plot of the strawberry yields of both the 2020 and 2021 seasons.

would also suggest, but instead of standalone attention heads we intend to use a much more complex and performant transformer model.

Transformers as proposed by Vaswani et al. (2017) are state-of-the-art neural network components for sequence-to-sequence problems. Strawberry yield prediction is such a problem thus we are keen to implement and use them in this scenario, having used other methods to varying degrees of success in the past (Onoufriou et al., 2020, 2021a). We also note that in contrast to our previous techniques transformers and their attention heads can help focus the neural network into parts of the data that are most important thus reducing the need for quite as much data compared to equivalently complex neural networks.

In short, yield forecasting is essential for improving on food security, and sustainable development (Zhu et al., 2022). Yield estimation is difficult due to a lack of data availability and thus a lack of research using modern data-hungry techniques in this domain (Chen et al., 2019; Maskey et al., 2019; Jafari et al., 2020; Gastli et al., 2021; Nassar et al., 2020). Most attempt to solve this data shortfall by using remote sensing, or by using a select few difficult-to-attain datasets like the California commissions data (Jafari et al., 2020; Nassar et al., 2020; Zhu et al., 2022). Few works have applied modern deep learning/neural networks successfully to agriculture, especially strawberries, the majority use either old neural network forms or do not use neural networks at all.

2. Material and methods

We have collected 3 years of strawberry tabletop data at our Riseholme campus. This data comprises 2 polytunnels, each with 5 rows of strawberry tabletop, each tabletop being 20 m long. Thus in total, we had 200 m of strawberry tabletop over any single season. Over these rows we had two different June bearing varieties at any one time from Driscoll's Zara, Katrina, and Malling Centenary. Fig. 3 shows the two varieties chosen for the 2021 growing season from the aforementioned three, as can be seen, their performance while similar, differ in that Katrina is expected to output more total yields in any given picking session on average. The data capture devices we employed for this strawberry tabletop was:

- Irrigation data from the tabletop irrigation system. This includes features describing the nutrients, moisture levels, soil temperature, input irrigation, and irrigation runoff. With a sample rate of 1 sample per 2 min.

- Environmental data from a central weathervane which collected information about: Temperature, humidity, wind direction, wind speed, solar radiance, and precipitation. With a sample rate of 1 sample per 15 min.
- Yield weight and quality data from our strawberry picking team. With a sample rate of 2 full picks per row per week.

2.1. Data wrangling

One of the biggest challenges when working with any time-series dataset is to ensure synchronicity. Since all 3 data sources are sampled at different sometimes overlapping intervals it was necessary to re-sample the datasets to achieve synchronisation. We opted to synchronise over the 15 min intervals to match the weathervane data. We later downsampled the synchronised data to a much more manageable 4-hour interval when fed into our MTT.

One of the other challenges when working with any data is missing or unrepresentative samples. Unfortunately in real-world scenarios we always expect to capture some missing or inaccurate data, especially when humans are necessarily involved in the process. We chose to use a forward-fill strategy whereby any missing values are filled with the last known values. The only features not forward-filled are ones that are sampled too infrequently to be able to reasonably forward-fill them. This means any missing values in yields for instance (which are collected bi-weekly) are removed as we cannot reasonably infer them from neighbouring values.

Now that we have a regular dataset with no missing values we can begin example extraction as per Fig. 1. We create hopping windows that end on/are aligned to observed yield outcomes in the current/predicted-for year. The window lengths we chose are 21 days for the premonition, 12 weeks for the present and the cumulative period for both combined in the previous year as the past. This way we have information on adverse weather forecasts, current strawberry performance, and performance of strawberries at the same site last year. We then create time sequences using expected date ranges. the historic data and when we have specific outcomes for fruit yields. This meant we roughly formed 2 examples for every week in the growing season. We then further split this data by row into training (2, 3, 4, 6, 7, 8, 10), and testing (1, 5, 9) sets, while further subdividing the training set into training and validation using k-fold cross-validation where $k = B_i$ with a batch size of $B_s = 32$ which resulted in $B_i = 10$ batches. We held out the two final shuffled batches as a per-epoch validation set. We split in

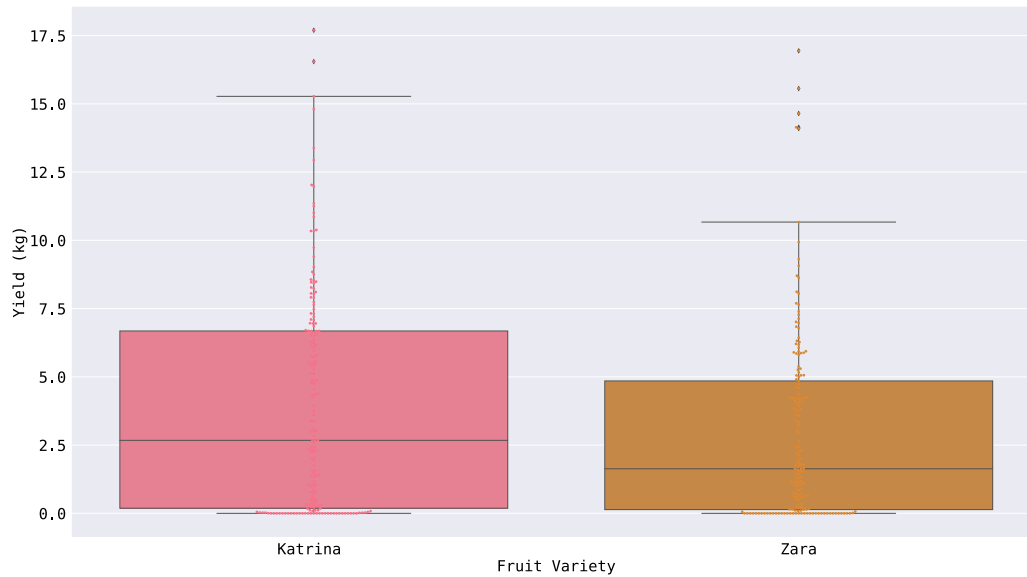


Fig. 3. Yield performance of the Katrina and Zara strawberry varieties over the 2021 growing season.

this manner to ensure there is no overlap between training and testing sequences, and it enables us to have a full multi-year view since there are not enough years of data with which to hold out.

Finally, we normalised our dataset feature-wise using a basic linear transformation Eq. (1).

$$x_i^{(t)} = (b - a) \frac{x_i^{(t)} - \min(x_i)}{\max(x_i) - \min(x_i)} + a \quad (1)$$

Where the desired normalised feature value for x_i at timestep t post normalisation $x_i^{(t)}$ is in $[a, b]$. We chose our range to be $[-1, 1]$. We inverted our results to real values using the inversion Eq. (2).

$$x_i^{(t)} = (x_i^{(t)} - a) \frac{(\max(x_i) - \min(x_i))}{b - a} + \min(x) \quad (2)$$

2.2. Architecture

As can be seen in Fig. 4, our MTT consists of 3 differently parameterised transformers merged together using a dense layer. Thus our architecture is comprised of 3 encoders, 3 decoders and a dense layer.

2.2.1. Encoder and decoder

As is standard for transformer networks it is necessary to decide upon some form of positional encoding (Vaswani et al., 2017). In our case we use a standard fixed positional encoding where even positions are encoded using Eq. (3) and odd positions are encoded using Eq. (4).

$$PE_{pos,2i} = \sin\left(\frac{pos}{10000^{2i/D}}\right) \quad (3)$$

$$PE_{pos,2i+1} = \cos\left(\frac{pos}{10000^{2i/D}}\right) \quad (4)$$

This positional encoding for each odd and even position is then added to the feature vector to allow the neural network some context into the order of inputs. There was no need to form a tokenised input embedding since we already have a distinct feature space described in our feature vector directly from the tabular sequences.

2.2.2. Dense

The dense layer is a simple linear layer with enough weights to form the weighted sum of the inputs and concatenate them into a singular value output in Eq. (5)

$$\hat{y} = \sum a_r W_r + \sum a_n W_n + \sum a_f W_f \quad (5)$$

Towards gathering data we employed our own data collection pipeline on our Riseholme strawberry tabletop site, the respective yields of this site can be seen in Fig. 2. All the following data is streamed into MongoDB and accessed using aggregation pipelines to help speed up the transformation process.

2.2.3. Weight initialisation

For weight initialisation, we used the default pytorch Kaiming uniform initialisation as defined in Algorithm 1 for leaky-ReLU (Nair and Hinton, 2010; Radford et al., 2015).

Algorithm 1 Kaiming uniform weight initialisation using leaky-ReLU with the fan-in method. where a : (default 0 for ReLU, or -0.01 for leaky-ReLU) is the negative slope of the rectifier used after this layer. W : a randomised weight matrix with mean 0 and variance 1 (shape e.g (64, 32)) $mode$: is a flag which represents a different value for the fan whether the method being used is for feedforward or backpropagation (e.g if $mode = fanin$ then $fan = 64$ else $fan = 32$ given previous example W matrix).

```

function KAIMING_UNIFORM_WEIGHT_INIT( $a, W, d$ )
  if  $mode = fanin$  then
     $fan = dim(W, 0)$ 
  else
     $fan = dim(W, 1)$ 
   $std = \sqrt{\frac{2}{(1+a^2) \times fan}}$ 
  return  $W \star std$ 

```

2.2.4. Loss function

We chose to use the Mean Squared Error (MSE) as our loss function where $MSE = \frac{\sum_{i=0}^{N-1} (y - \hat{y})^2}{N}$. This allows us to exponentially penalise large more errors than small errors on our continuous yield forecast. We in particular seek to reduce the networks tolerance for larger single errors as these would mean even if the total error was the same, being particularly peaked in one prediction would result in the growers having to import fruit that particular week. We would much rather be consistently out by a known amount than having almost perfect performance one week and then large errors the next.

As is commonly the case we use Adaptive moment (ADAM Kingma and Ba, 2014) as our neural network optimiser as it has been shown to be more performant than just first order or second order moments

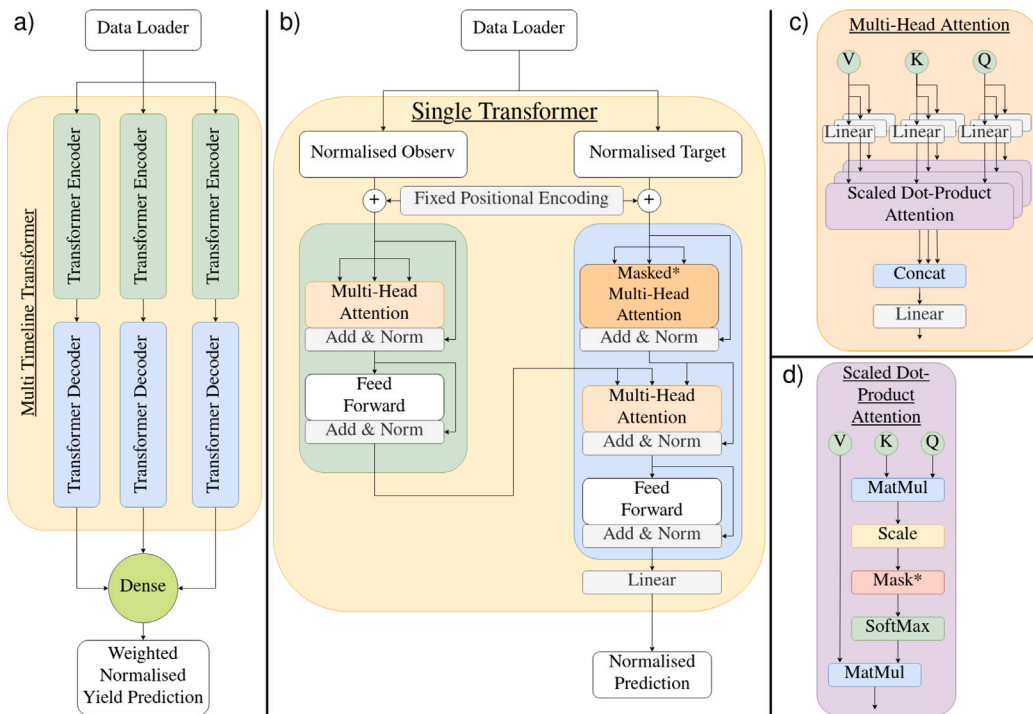


Fig. 4. (a) Multi Timeline Transformer (MTT) architecture whereby three single transformers that each process different data streams, are merged by a learned dense layer to weigh their significance. (b) A full single transformer architecture comprised of fixed positional encoding, encoder, decoder, and linear layers notably missing Softmax. (c) Multi-head attention mechanism with query, key, and value matrices. This is a sub-components of transformer encoder and decoders with optional masks to maintain the temporal blindness when processing all the data simultaneously. (d) Scaled dot-product attention showing the various matrix operations necessary to compute. this is a sub-component of multi-head attention.

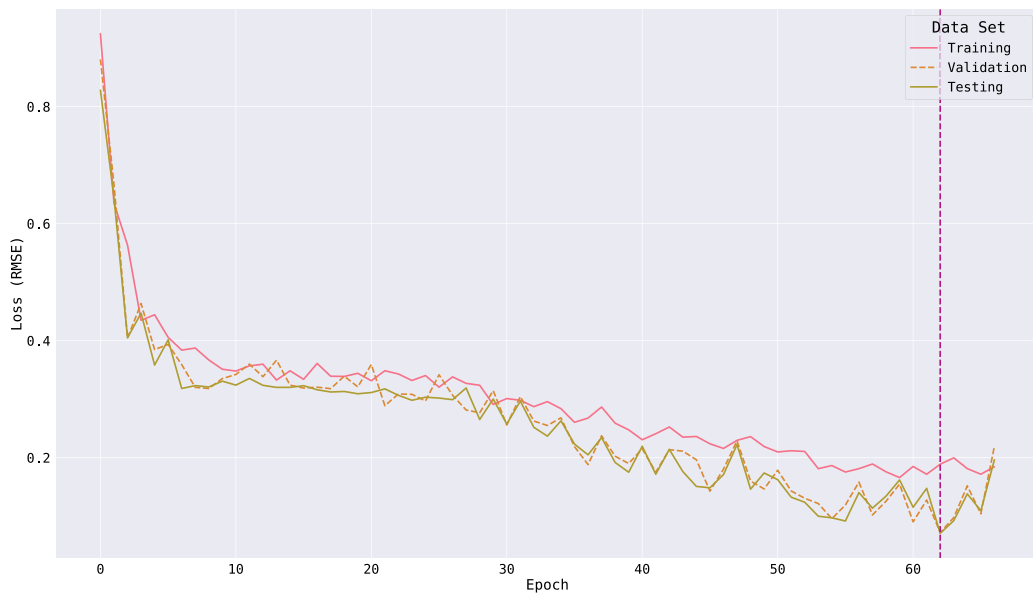


Fig. 5. Three timeline transformer loss training, validation and testing sets, per epoch of training. Beyond 62 epochs (pink vertical line) validation and testing loss steeply increases again. It should be noted that we plot root mean squared error (RMSE) against epoch while our loss function was MSE, this is so the loss is converted back into human-readable proportions.

and is by and large the defacto standard. We calculated our first order moments $m_t = \beta_1 * m_{t-1} + (1 - \beta_1) * g_t$, $\hat{m}_t = \frac{m_t}{1 - \beta_1^t}$ and second order moments.

2.3. Models

We primarily focused on two different types of model. One holistic model that learned from all of the training rows using random subsets

for training and validation (Fig. 5). Then we also attempted to create smaller weaker predictors as an ensemble only trained on a smaller set of the training data to each other as an ensemble to attain simple certainty metrics, which we deem would be invaluable towards building trust in the models and enabling re-investigation of uncertain scenarios. We split the training data used into 3-row sets of tabletop for each ensemble member. Each ensemble member is equivalent to the base MTT, including weight initialisation, loss function, and optimiser.



Fig. 6. Ordered forecasts of single MTT compared to ground truth with a horizon of 3 weeks and a history of 12 weeks.

Table 1

Expected errors by forecasting source. All models are from our previous work trialling different methods on the same dataset.

Forecaster	Expected error
Grower	25% ^a
Agronomist	17% ^a
Recurrent Neural Network (RNN)	21%
Long-Short Term Memory network (LSTM)	38%
Gated Recurrent Network (GRU)	16%
Multi-Timeline Transformer (MTT)	8%
Ensemble of MTT (average)	27%
Ensemble of MTT (median)	30%

^aThese are estimates and may not be representative of any grower or agronomist specifically but are instead ballpark figures for illustration based on our information from our industry partners.

Overall this means there was a one-row overlap between the first–second and second–third MTT. The results of our two current attempted approaches along with our past approaches and expected forecasting performance of growers and agronomists can be seen in Table 1.

3. Results

As can be seen in Table 1 our primary MTT that can forecast three weeks ahead within 8% RMSE is a large improvement over current capabilities as forecasts by agronomists tend to not only vary wildly from agronomist to agronomists (14 to 30%), rely on specialist human presence, and are less accurate than our current model. However, a large caveat is that our model was created with intensive/high-quality environmental and yield data, on a small site compared to the typical industrial settings.

The results shown in Table 1 and Fig. 6 are a significant step forward in the prediction of strawberry yields, however, there are some weaknesses to our approach and the yield outcomes. Firstly our ensemble is significantly underperforming especially since a single predictor trained on the whole dataset beats the ensemble significantly. This is likely due to data, with almost three times the parameters, we suspect that we require more training data to learn adequately, yet they receive 1/3 of the total training data each. However, as time progresses and more data becomes available to us over more seasons, we believe this ensemble will outperform the single MTT while enabling ensemble-based certainty estimation. Secondly and most difficult is the data itself.

While we are fortunate to have access to our Riseholme campus and the strawberry tabletop site, there is still a lack of data available for use. This relatively small site means we likely have not learned some of the more complex variances present on larger sites where the sensors' immediate environment might be significantly different to another area on the growing site some distance away meaning the data in such scenarios might be significantly less representative of the conditions experienced by the strawberries.

4. Discussion

Our strawberry dataset while covering 200 m of strawberries is still limited. Commercial sites in comparison have hectares of such crops, meaning our 200 m is not as representative of larger sites with more intra-crop variability. However, as previously mentioned data availability is scarce making it practically very difficult to collect hectares of data, not least due to actual or perceived data sensitivity by the respective growers. In spite of this, while there may need to be some adjustments to account for more intra-crop variability of these larger sites our neural networks perform well given the data availability. While the sites are smaller and easier to learn, they also have fewer data to do so, which we believe to be a fair trade-off with no loss in difficulty between their larger sites and our smaller site.

We have a high level of intra-crop variability with our dataset in the similarity between rows. Largely while there is inter-row variance there is still a risk of overfitting since even if the neural network cannot see row 1, for instance, it may be able to relate the yields of row 1 from previously trained/known yields of row 2. We would have ideally liked to have split by time, and claimed one whole season as a completely separate testing set with none of those rows being trained on. However, due to the reality of strawberry seasonality and that there are only so many seasons with which it was possible to collect data, we had to split in such a way as to give the neural networks context for at least two seasons from start to finish. This is only necessary since the current methods of strawberry prediction in industry are largely based on the occurrences of the last season. As such we attempted to base our methods on existing techniques, and intuitively the performance of the strawberries last year will be related to the current season's performance unless some large shift in methods between the seasons occurs.

Fig. 3 shows a significant number of zero/near-zero values. This is due to the slow start at the beginning of every season as shown in Fig. 2.

In our data collection, we still recorded fruitless strawberry picking sessions to account for some strawberry plant varieties producing for longer periods in the growing season, whereas others started later. This is significant as the total berries one would expect to harvest over the season is affected. In particular, for the 2021 season, we experienced a very slow start to our season with very low yields when compared to the 2020 season. Later in the season, we may also experience zero/near-zero values, these are difficult to distinguish from actual low values and bad picking sessions. One way that we might have made such assumptions is by assuming the harvesting effect causes all temporally adjacent picks of the same row to have diminishing returns.

Our MTT used an interval of 4 h despite our data being synchronised over 15-minute intervals. This was a tradeoff between data density (thus model complexity), and data availability. Since we only had a finite number of concrete outcomes that we observed we had to limit the complexity and weights of the model so that it could train its fewer weights with what limited data we had for concrete observations. In contrast, if we had used a data density of 15-minute intervals we would have had to have significantly larger weight matrices being backpropagated from the limited number of observed yield values. If, however, we found ourselves with large hectare scale datasets with many more observed outcomes we could tune the model to be more complex to leverage this data, allowing the model to understand much more complex relationships like the aforementioned expected intra-crop variability.

It may also be noted that we use a simple missing data imputation algorithm strategy namely forward-fill which involves filling missing values with the last known value. This was chosen as we mostly only incurred individual or relatively sparsely missing data. In larger sites one might expect to find entire regions that have some data unavailability for some time, meaning more advanced data-filling strategies may be necessary under such conditions. However, in our site, since the missing values were relatively sparse, the forward fill strategy is sufficient to allow us to leverage data in spite of any missing observations or features. The only notable exception is that of yield values. Since yield values were recorded sparsely a single missing value represents a much larger significance. Thus any such missing values are excluded entirely. Thankfully we had very few such missing values.

Due to the data scarcity, we used fixed positional encoding as opposed to learned encoding. This means the gradients would not be shared with the learned positional encoding. This is sufficient since in the original transformer paper (Vaswani et al., 2017) fixed positional encoding and learned positional encoding result in similar performance.

Finally, we chose to use a tri-transformer architecture merged using a dense fully connected layer. We did this to allow the neural network to train separate contextualising units for each potential timeline. This way we can easily conceptualise the timelines as follows. The pasts purpose is to have a broad view of the relationship between the features and the expected outcomes. This is important as we want to ensure the network has context for how yields are expected to outcome given past scenarios. The present serves to contextualise how this current specific season or crop is performing such that it can later be related to what has happened in the past. The future timeline/transformer is to add mitigations and adverse effects, such that high expected fluctuations can be considered at the merging layer.

5. Conclusions

In this study, we propose a new multi-timeline transformer architecture that outperforms many other forms of neural networks (including CNNs, RNNs, LSTMs, and GRUs) in forecasting strawberry yields, even on small datasets. Multi-timeline transformers are very capable of learning from the past, the present, and the premonition of the future, even when these use similar approaches to human forecasters. There has been little work in forecasting strawberries using state-of-the-art deep learning methods, and all of the works that do exist struggle with

data availability. Data is clearly the principal problem, as we need more data to develop good machine learning models. Therefore, we need to encourage the sharing and collection of more and better data so that more impactful research can be performed. With more data, we can properly test ensemble models and similarly data-hungry models.

Regarding our future work, one key area would involve implementing certainty metrics that do not require the use of ensembles so that we can keep the neural network parameters down. This would reduce the necessary data to train more complex models. We also seek to make transformers that are abelian compatible such that we can use some of the previously proposed fully homomorphically encrypted (FHE Gentry and Halevi, 2010) deep learning methods with these currently incompatible but performant transformers (Onoufriou et al., 2021b,a).

Lastly, we seek to find ways in which to make our data available for wider use, currently that is not possible due to contractual constraints which were necessary to enable us to collect this data with industrial varieties in the first instance. However, we seek to remedy this in future.

CRedit authorship contribution statement

George Onoufriou: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Marc Hanheide:** Resources, Project administration, Funding acquisition. **Georgios Leontidis:** Conceptualization, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

Funding statement: This research was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) studentship, United Kingdom [grant numbers 2155898, BB/S507453/1].

Special Thanks: The authors would also like to thank the [Lincoln Centre for Autonomous Systems \(L-CAS\)](#) for their help and support with the autonomous robotic [Thorvald](#) systems, and the availability of tools and resources such as the Riseholme strawberry tabletop itself.

References

- Bali, N., Singla, A., 2022. Emerging trends in machine learning to predict crop yield and study its influential factors: a survey. *Arch. Comput. Methods Eng.* 29 (1), 95–112.
- Bouras, E.H., Jarlan, L., Er-Raki, S., Balaghi, R., Amazirh, A., Richard, B., Khabba, S., 2021. Cereal yield forecasting with satellite drought-based indices, weather data and regional climate indices using machine learning in Morocco. *Remote Sens.* 13 (16), 3101.
- Chen, Y., Lee, W.S., Gan, H., Peres, N., Fraisse, C., Zhang, Y., He, Y., 2019. Strawberry yield prediction based on a deep neural network using high-resolution aerial orthoimages. *Remote Sens.* 11 (13), 1584.
- Durrant, A., Markovic, M., Matthews, D., May, D., Enright, J., Leontidis, G., 2022. The role of cross-silo federated learning in facilitating data sharing in the agri-food sector. *Comput. Electron. Agric.* 193, 106648.
- Durrant, A., Markovic, M., Matthews, D., May, D., Leontidis, G., Enright, J., 2021. How might technology rise to the challenge of data sharing in agri-food? *Global Food Secur.* 28, 100493.

- Gastli, M.S., Nassar, L., Karray, F., 2021. Deep learning models for Strawberry yield and price forecasting using satellite images. In: 2021 IEEE International Conference on Systems, Man, and Cybernetics. SMC, IEEE, pp. 1790–1796.
- Gentry, C., Halevi, S., 2010. Implementing Gentry's Fully-Homomorphic Encryption Scheme. Cryptology ePrint Archive, Report 2010/520, <https://eprint.iacr.org/2010/520>.
- Hopf, A., Boote, K.J., Oh, J., Guan, Z., Agehara, S., Shelia, V., Whitaker, V.M., Asseng, S., Zhao, X., Hoogenboom, G., 2022. Development and improvement of the CROPGRO-Strawberry model. *Sci. Hort.* 291, 110538.
- Jafari, F., Ponnambalam, K., Mousavi, J., Karray, F., 2020. Yield forecast of California strawberry: Time-series models vs. ML tools. In: 2020 IEEE International Conference on Systems, Man, and Cybernetics. SMC, IEEE, pp. 3594–3598.
- Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- Maskey, M.L., Pathak, T.B., Dara, S.K., 2019. Weather based strawberry yield forecasts at field scale using statistical and machine learning models. *Atmosphere* 10 (7), 378.
- Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted boltzmann machines. In: *Icml*.
- Nassar, L., Okwuchi, I.E., Saad, M., Karray, F., Ponnambalam, K., Agrawal, P., 2020. Prediction of Strawberry yield and farm price utilizing deep learning. In: 2020 International Joint Conference on Neural Networks. IJCNN, pp. 1–7. <http://dx.doi.org/10.1109/IJCNN48605.2020.9206998>.
- Onoufriou, G., Hanheide, M., Leontidis, G., 2020. The augmented agronomist pipeline and time series forecasting. <http://dx.doi.org/10.31256/Qm1Fu7L>.
- Onoufriou, G., Hanheide, M., Leontidis, G., 2021a. EDLaas; fully homomorphic encryption over neural network graphs. <http://dx.doi.org/10.48550/ARXIV.2110.13638>, URL: <https://arxiv.org/abs/2110.13638>.
- Onoufriou, G., Mayfield, P., Leontidis, G., 2021b. Fully homomorphically encrypted deep learning as a service. *Mach. Learn. Knowl. Extr.* 3 (4), 819–834.
- Paudel, D., Boogaard, H., de Wit, A., Janssen, S., Osinga, S., Pylianidis, C., Athanasiadis, I.N., 2021. Machine learning for large-scale crop yield forecasting. *Agricult. Sys.* 187, 103016.
- Pearson, S., May, D., Leontidis, G., Swainson, M., Brewer, S., Bidaut, L., Frey, J.G., Parr, G., Maull, R., Zisman, A., 2019. Are distributed ledger technologies the panacea for food traceability? *Global Food Secur.* 20, 145–149.
- Radford, A., Metz, L., Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint [arXiv:1511.06434](https://arxiv.org/abs/1511.06434).
- van der Velde, M., Nisini, L., 2019. Performance of the MARS-crop yield forecasting system for the European Union: Assessing accuracy, in-season, and year-to-year improvements from 1993 to 2015. *Agricult. Sys.* 168, 203–212. <http://dx.doi.org/10.1016/j.agsy.2018.06.009>, URL: <https://www.sciencedirect.com/science/article/pii/S0308521X18300179>.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. *Adv. Neural Inf. Process. Syst.* 30.
- Zhu, Y., Wu, S., Qin, M., Fu, Z., Gao, Y., Wang, Y., Du, Z., 2022. A deep learning crop model for adaptive yield estimation in large areas. *Int. J. Appl. Earth Obs. Geoinf.* 110, 102828.

George Onoufriou is a Ph.D. student at the University of Lincoln, studying encrypted deep learning as part of a funded collaborative training partnership/ studentship. This studentship comes from the British Biotechnology and Biological Sciences Research Council in collaboration with Berry Gardens Growers, towards fruit yield forecasting. George is a cypherpunk, keenly interested in deep learning, privacy/ cryptography, (Arch) Linux, and open-source software in general. He is also founder of Deep Cypher Ltd., which seeks to enable and conduct encrypted deep learning as a service.

Prof. Marc Hanheide is a Professor of Intelligent Robotics & Interactive Systems in the School of Computer Science at the University of Lincoln, UK, and the director of the University's cross-disciplinary research centre in Robotics, the Lincoln Centre for Autonomous Systems (L-CAS). He received the Diploma in computer science from Bielefeld University, Germany, in 2001 and the Ph.D. degree (Dr.-Ing.) also in computer science from Bielefeld University in 2006. From 2006 to 2009 he held a position as a senior researcher in the Applied Computer Science Group. From 2009 until 2011, he was a research fellow at the School of Computer Science at the University of Birmingham, UK. Marc Hanheide is a PI in many national and international research projects, funded by H2020, EPSRC, InnovateUK, DFG, industry partners, and others, as well as the director of the EPSRC Centre for Doctoral Training (CDT) in Doctoral Training in Agri-Food Robotics (AgriFoRwArdS). The STRANDS, ILIAD, RASberry, and NCNR projects are among the bigger projects he is or was involved with. He features regularly in public media, has published more than 100 peer-reviewed articles, and is actively engaged in promoting the public understanding of science through appearances in dedicated events, media appearances, and public lectures.

Dr. Georgios Leontidis is the Director of the Interdisciplinary Centre for Data and AI, and a Reader in Machine Learning at the University of Aberdeen, UK, leading the Machine Learning group. Before joining the University of Aberdeen, Georgios was a Senior Lecturer at the University of Lincoln, UK. His expertise lies within the areas of Capsule Neural Networks, Domain Adaptation, and Self-Supervised Learning. Georgios has applied his research in different settings, including agri-food, gas turbines, nuclear reactors, environment, retail refrigeration systems, and healthcare. Georgios has been leading several projects, funded by industry, such as Siemens Energy and Tesco, EU-H2020, and UKRI. He is also a TEDx speaker and has featured in public media.