A machine learning pipeline for climate impacts: crop models versus deep learners

Joseph Gallear[1 2], Andrew J Challinor[1], Netta Cohen[2], Anthony G Cohn[2], Julia Chatterton[3]

Contact: eejg@leeds.ac.uk

[1]University of Leeds, School of earth & Environment, [2]University of Leeds School of Computing, [3]Safety & Environmental Assurance Centre, Unilever R&D

UNIVERSITY OF LEEDS

Aims:

- Which machine learning frameworks best outperform the GLAM crop model?
- How much training data is required for machine learning frameworks to outperform crop modelling?
- How complex does a machine learning framework need to be to outperform a crop model?
- Can deep learning be successfully applied to global gridded weather crop yield relationships?

Background:

- Machine learning is defined as "A computer learning from some experience E with respect to some task T and some performance measure P, if its performance on T as measured by P, improves with experience E" – (Thomas Mitchell, 1997).
- Machine learning frameworks tested so far include random forest, feed forward and 1 dimensional convolutional neural nets, and nearest neighbors regression.

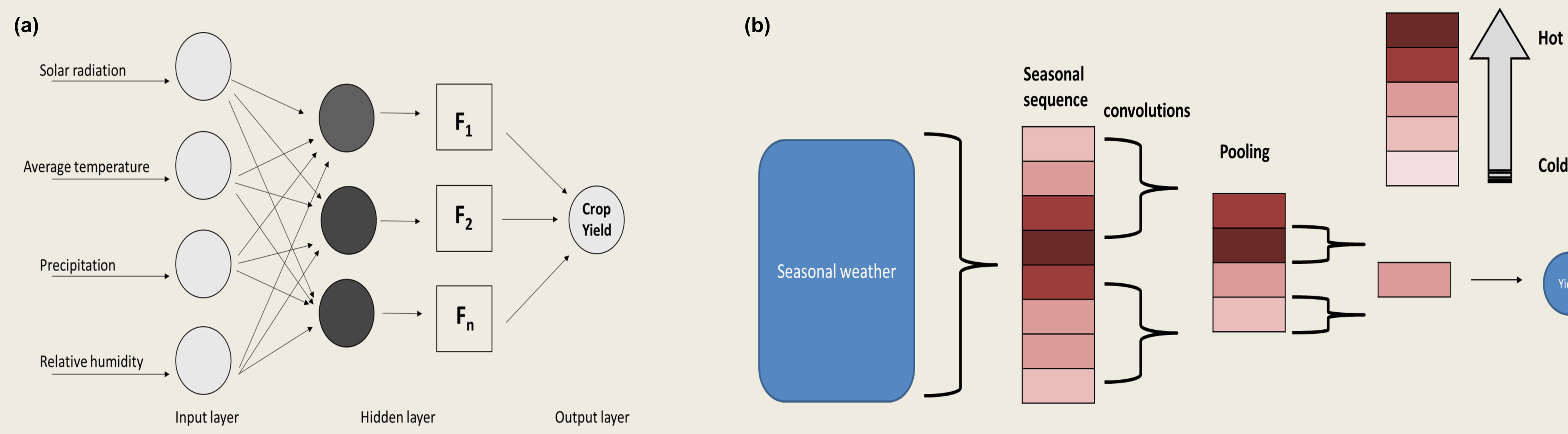


Fig 1. (a) A schematic of a feed forward neural network with fully connected layers as opposed to **(b)** a method of training a neural network by incorporating each growing season as an ordered sequence using convolutional layers.

Methods outline:

- Figure 2 displays how the study is progressing with each preprocessing step attempting to build a fair comparison between the GLAM crop model and machine learning methods.
- The data set used is described in Watson et al. (2015)
- Model complexity varies according to number of parameters, and dimensionality.
- Depending on final results, data will be removed or added using surrogate data to determine how much data is required for machine learning to outperform the GLAM crop model.

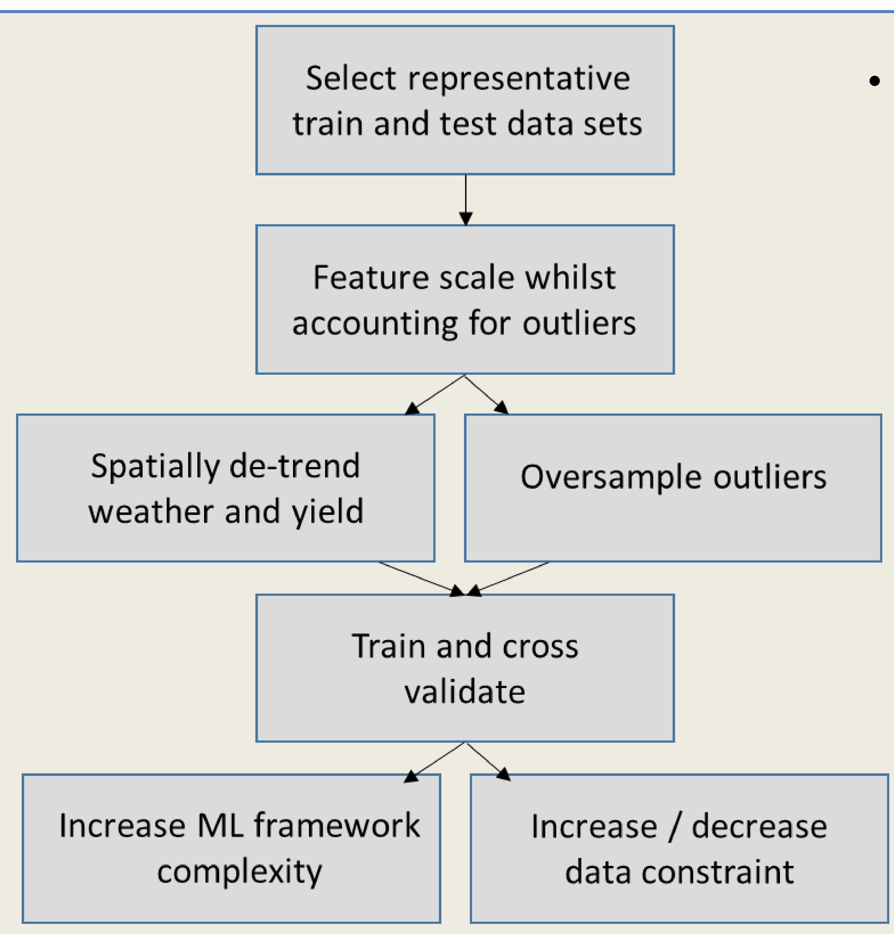


Fig 2. Flow chart detailing the sequential study methodology.

- Input weather data is paired with yield data at the department scale, meaning that a 1:1 input to target yield ratio is gained (Figure 3).

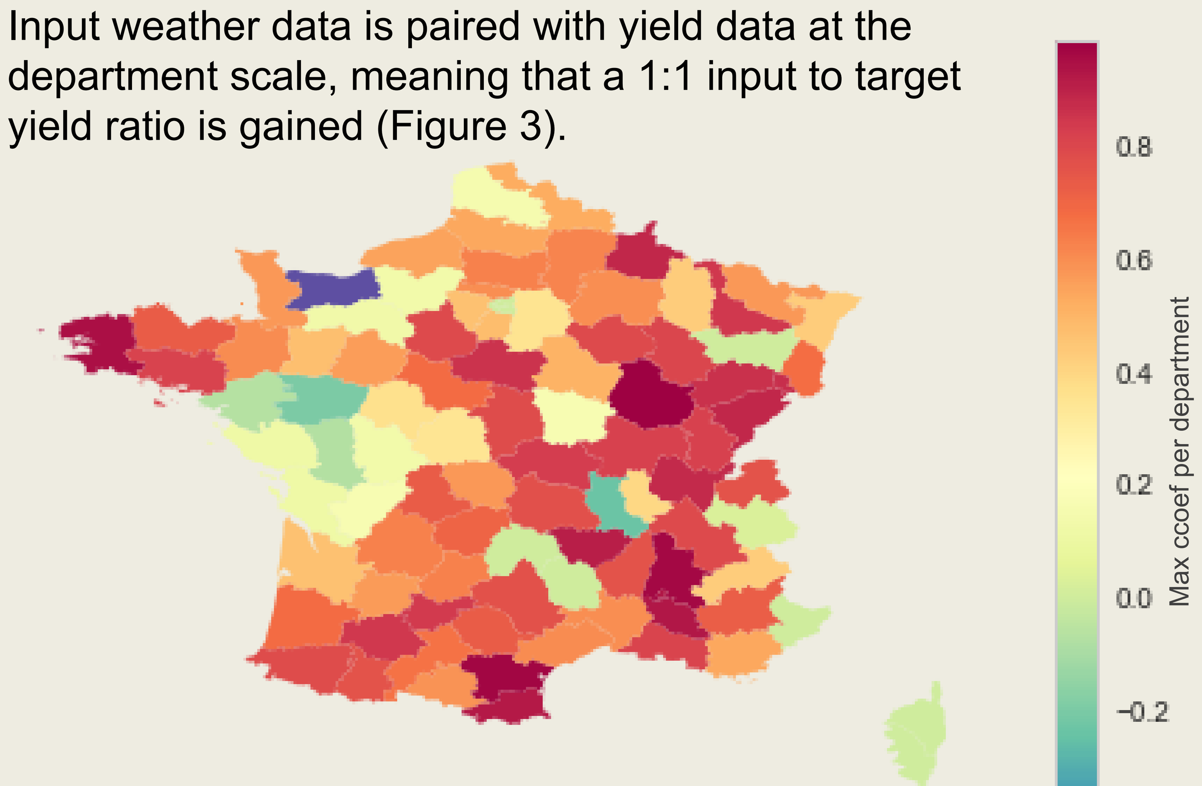


Fig 3. Maximum GLAM correlation per department was used as training data. This is to ensure the weather data was used per department which best reproduced the department level yield data

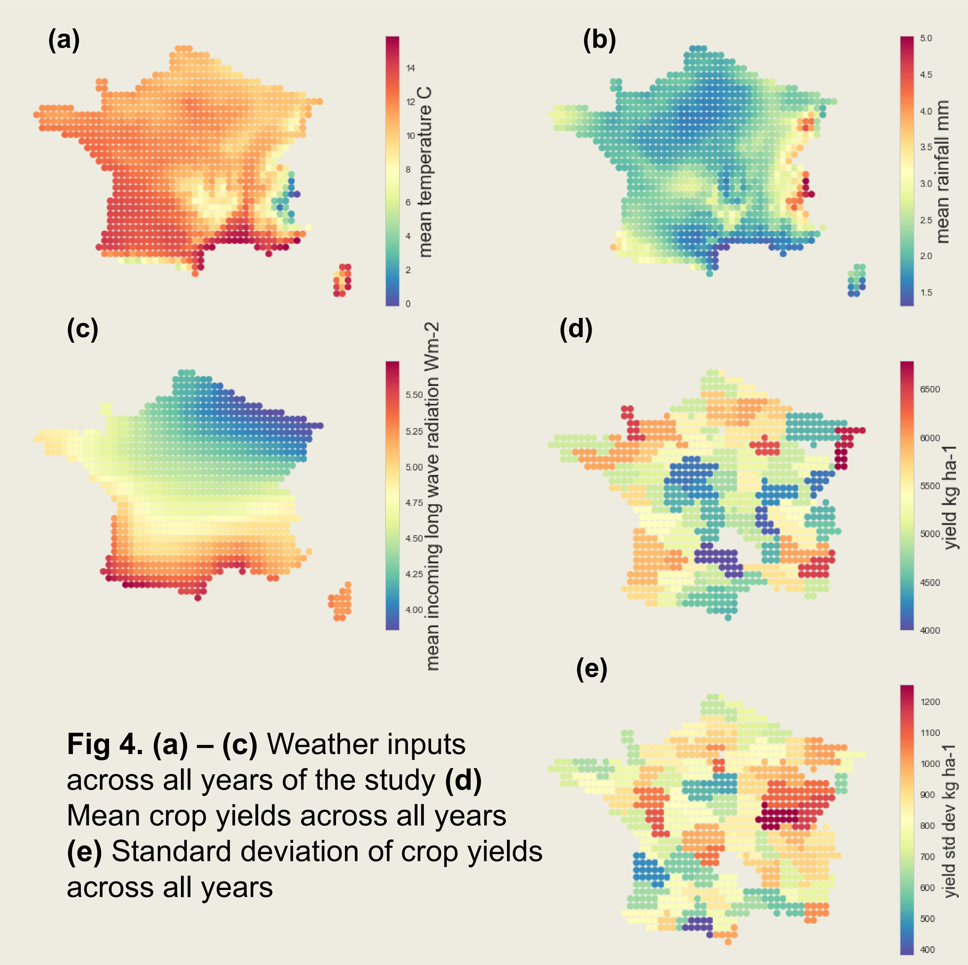


Fig 4. (a) – (c) Weather inputs across all years of the study **(d)** Mean crop yields across all years **(e)** Standard deviation of crop yields across all years

GLAM results:

- GLAM results are taken from a previous study (Watson et al. (2015).
- GLAM model skill varied on a department basis

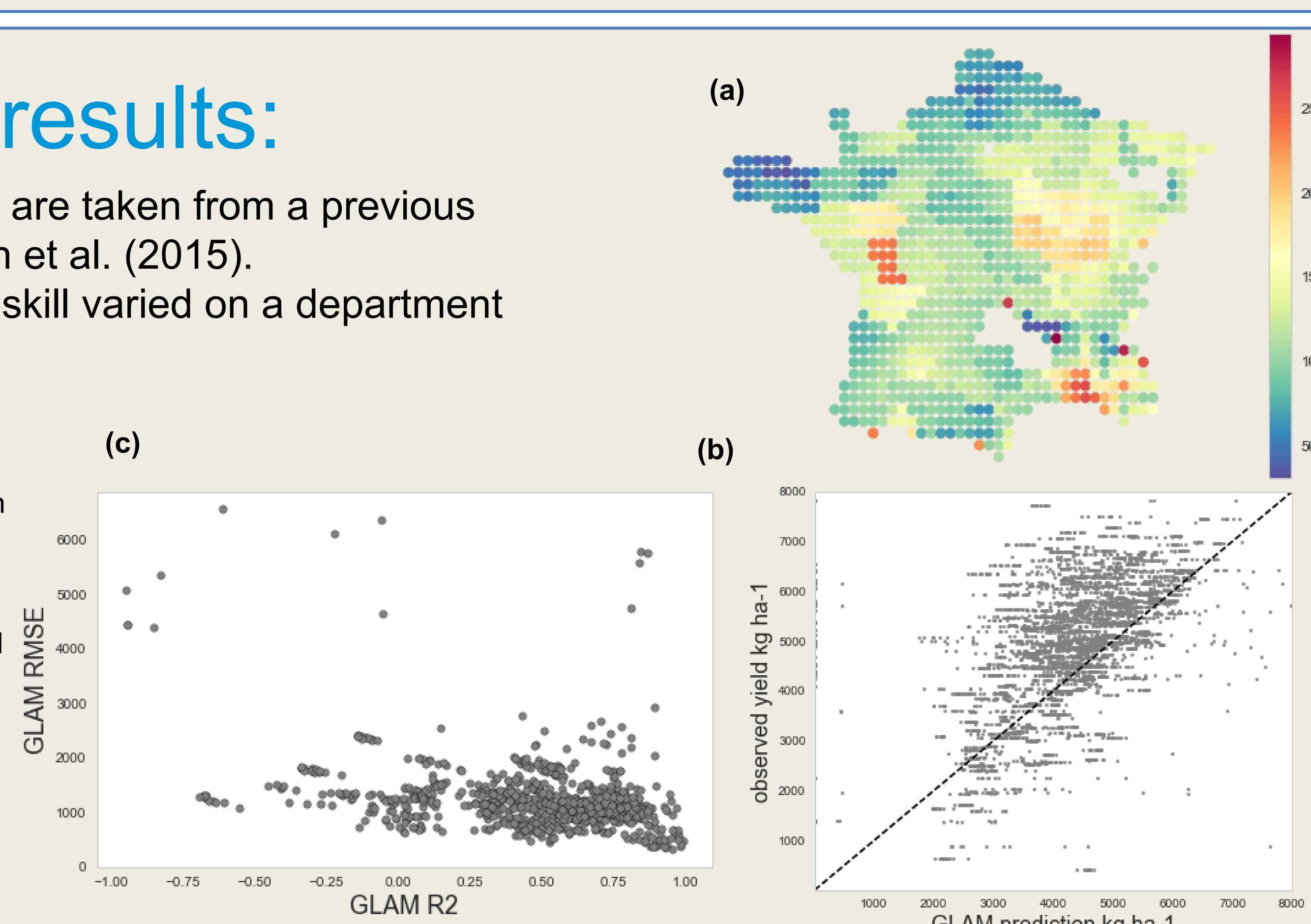


Fig 5. Results of the GLAM crop model adapted from Watson et al. (2015) **(a)** Shows GLAM skill (RMSE) spatially using data re-gridded to the department scale. **(b)** shows a model fit of GLAM predictions against observed data. **(c)** displays the ccoef against RMSE.

Models comparison:

- Random forest best produced an overall mean error on par with GLAM however the model overfit and therefore poorly reproduced year to year variability
- Feed forward and convolutional neural networks and nearest neighbours (KN) regression better predicted variation in yield, however have not yet outperformed the crop model in either case.

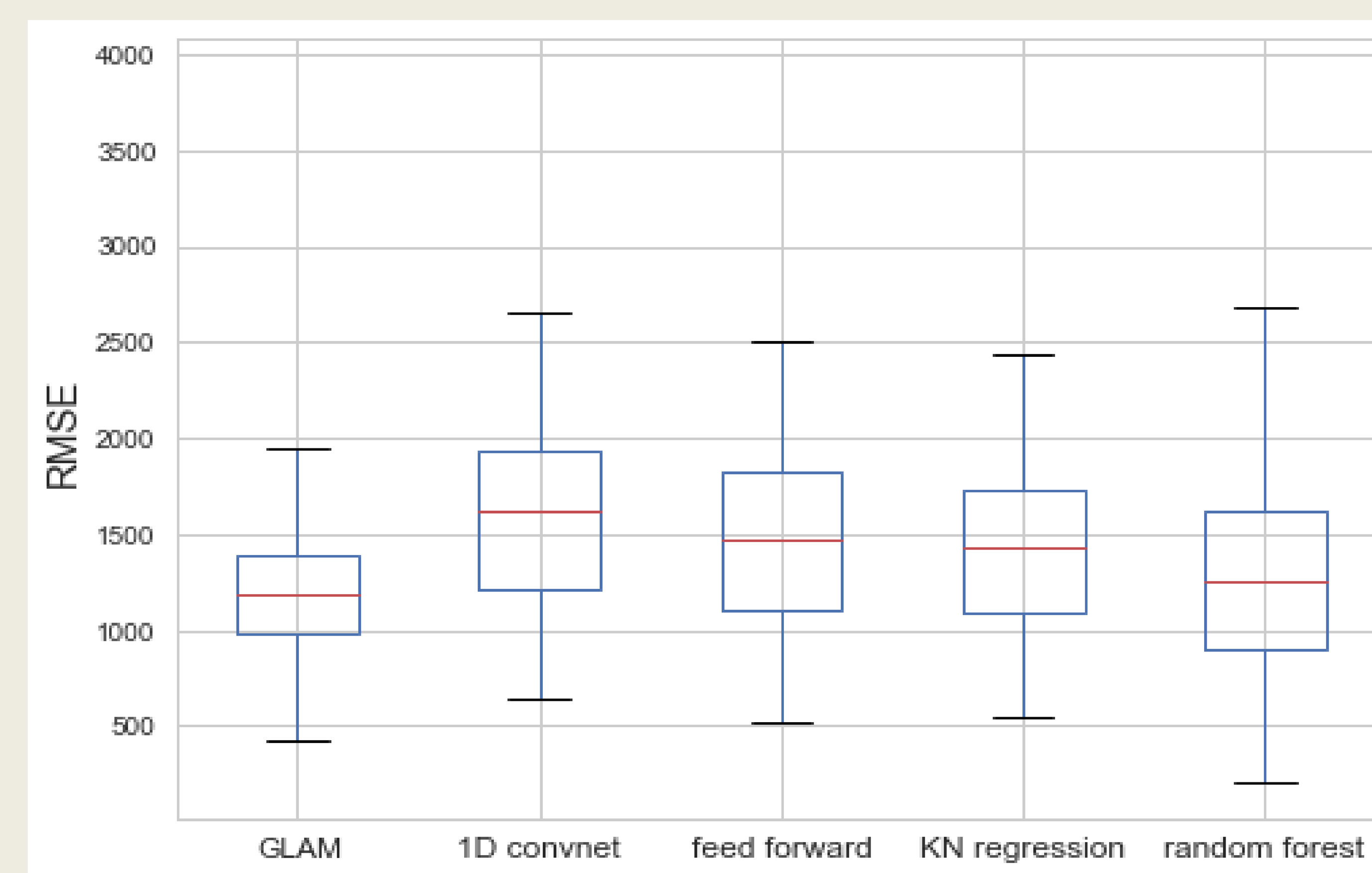
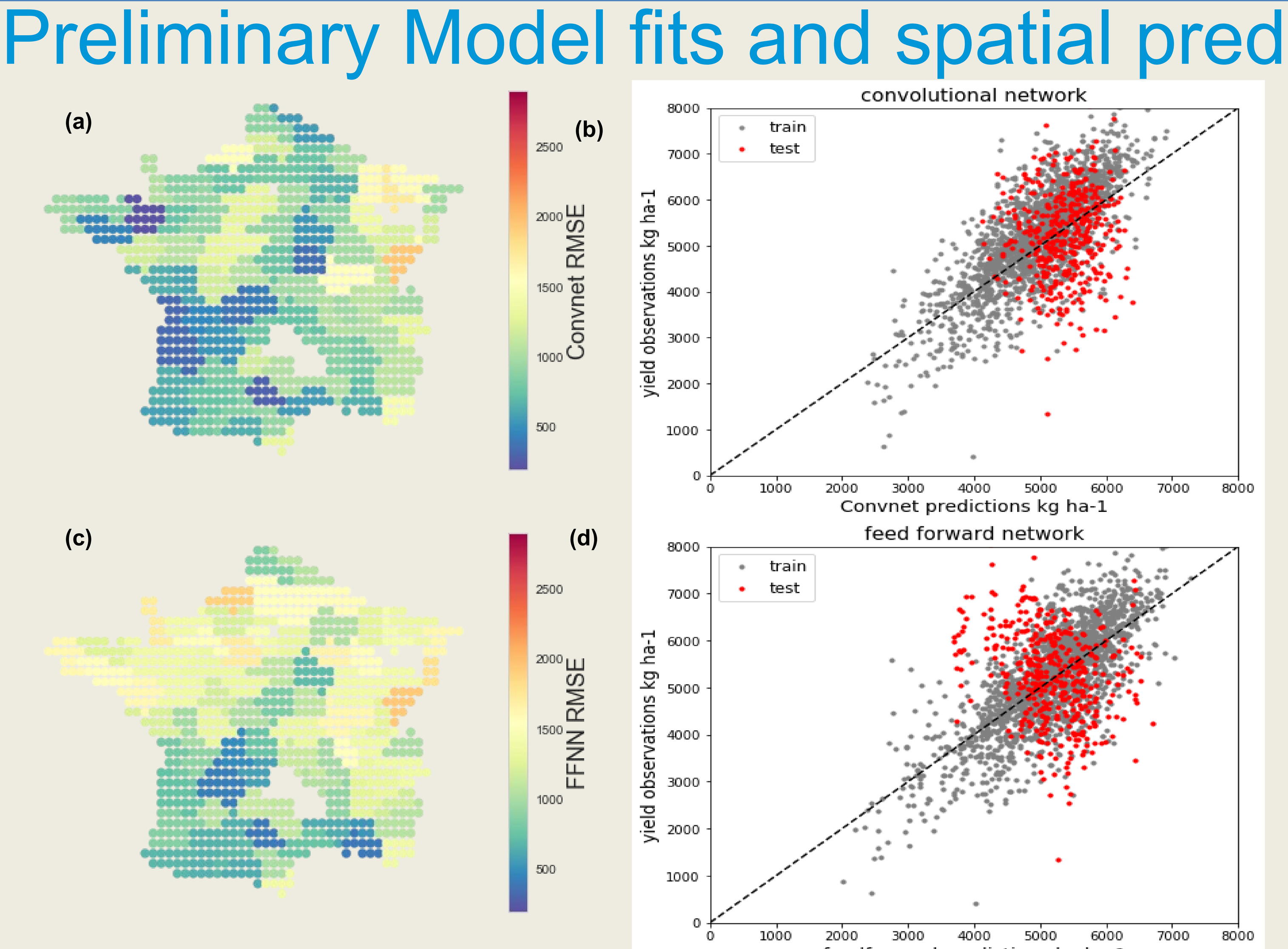


Fig 6. Boxplots of RMSE from GLAM along with a 1-dimensional convolutional neural network, a dense feed forward network, nearest neighbours (KN) regression and random forest model. A comparison was made with GLAM on the last 5 years of the dataset. Last 5 years had a large number of outliers.

Preliminary Model fits and spatial predictions:



- Spatially model skill varies with the standard deviation of each department (Figure 7 a & c).
- Model skill also varied greatest from GLAM in areas where the GLAM yield gap parameter was needed to correct for non-weather spatial factors
- Model fits (Figure 7 b & d) were largely affected by the combination of training and testing years as well as choice of feature scaling technique. This is likely due to low initial data quantity.
- Model fit is displayed where test years contain a low number of outlier values.

Fig 7. RMSE **(a)** and model fit **(b)** of a convolutional neural network, and dense feed forward neural network **(c) (d).** The convolutional network uses 1 dimensional convolutions to read each growing season as a sequence, whereas the dense feed forward network uses unordered daily values. In model fits **(b)** & **(d)** grey points indicate model fit on the training years with red points denoting the first 5 year test period. Fit on first five years is shown here to display model fit with minimal outliers in test set.

Next steps:

To build a fair comparison with GLAM, The following steps will be incorporated into the methodology:

1. Spatially de-trend the weather data by clustering based on soil properties and homogenous weather (Figure 8 displays the principal components of the data)
2. Use an oversampling technique such as SMOTE-R to reduce overfitting
3. Incorporate LSTM cells into NN models
4. test Bayesian model frameworks
5. Test auto encoding methods for pre-processing

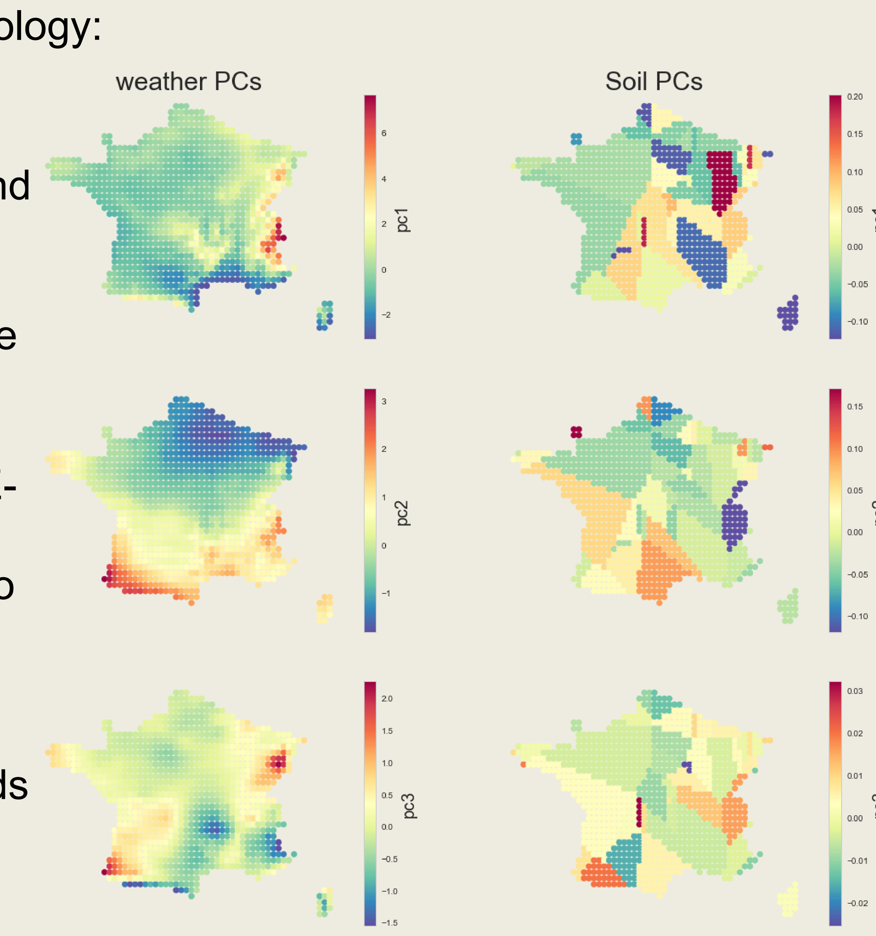


Fig 8. 3 weather and soil principal components showing spatial patterns of each of the major components of variation.

Acknowledgements

Special thanks to James Franke, Christoph Müller & Alex Ruane for providing access to the GGCMI phase 2 dataset which facilitated discussions which led to this work.

References

Elavarasan D., Vincent D R., Sharma V., Zomaya A Y., Srinivasan K., (2018)., Forecasting yield by integrating agrarian factors and machine learning models: A survey., Computers and electronics in agriculture., 155., [No issue no.]: 257-252

Mitchell T, (1997), Machine learning, New York, McGraw-Hill

Watson J., Challinor A J., Fricker T E., Ferro C A T., (2015)., Comparing the effects of calibration and climate errors on a statistical crop model and a process-based crop Model., Climatic Change., **132**., (1): 93-109

Natural Environment Research Council