Deep Learning for Drought and Vegetation Health Modelling

Thomas Lees, Gabriel Tseng, Steven Reece, Jian Peng, Alex Hernandez-Garcia, Clement Atzberger, Simon Dadson

UNIVERSITY OF OXFORD

Key Messages:

1. We trained a **recurrent neural network** to forecast a drought index (VCI) one month ahead.

2. We achieve **state of the art performance** when compared across four Arid Districts in Kenya.

3. We **interpret** what the models are learning by using:
   a. Clustering analysis to show how the model represents pixel-similarity.
   b. DeepLIFT to measure the contribution of each feature to a prediction.

Model Performance — Clustering — Model Interpretability — Discussion & Conclusions

Slides online

Deep Learning for Drought and Vegetation Health Modelling

Thomas Lees*, Gabriel Tseng**, Steven Reece*, Jian Peng*, Alex Hernandez-Garcia^, Clement Atzberger^^, Simon Dadson*

* University of Oxford
** Okra Solar
^ University of Onnasbruck
^^ University of Natural Resources and Life Sciences (BOKU)

* I’m interactive! Use the navigation slide by clicking **this** button.

Navigation

Use this slide to navigate quickly to your areas of interest!

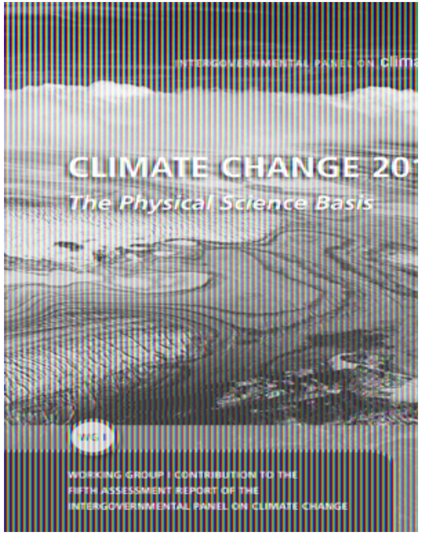
Click here to return to this page!

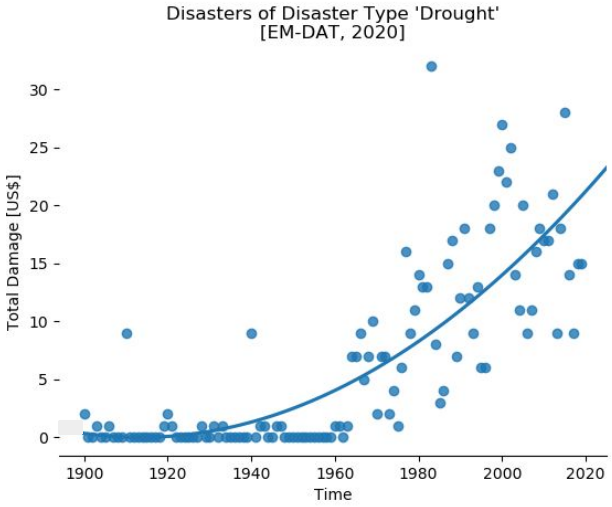
Motivation

Thematic Focus

Agricultural drought is a pressing global problem. The IPCC have “*medium confidence*” that drought frequency and intensity have increased 1980-2013.

$29 billion in losses to developing world agriculture between 2005 and 2015



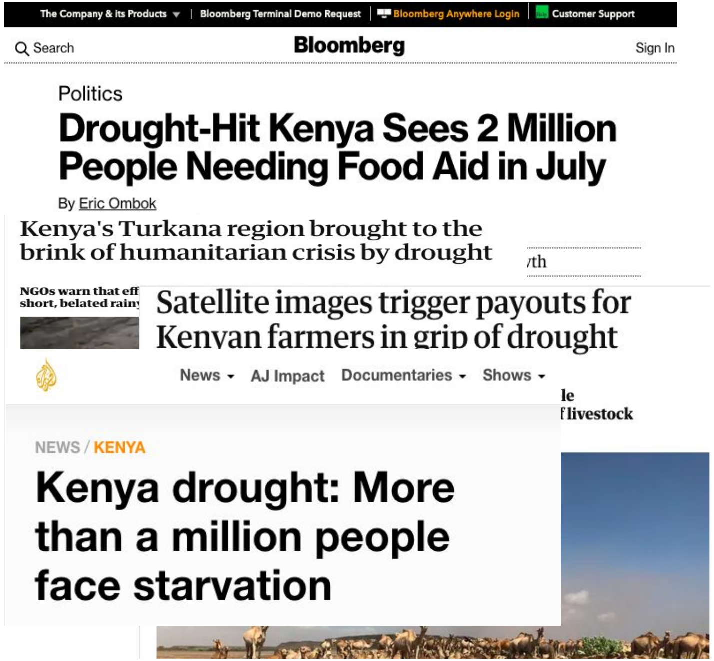


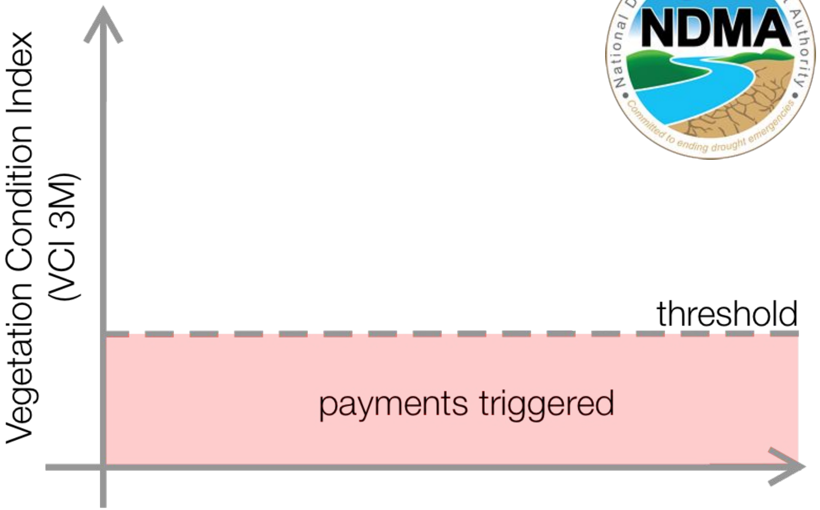
Under future climate change we expect droughts to get worse.

Motivation

Geographical Focus

Kenya distributes emergency funds using a vegetation index through the National Drought Management Authority (NDMA).





It's important to make good predictions to minimise the damage caused by drought, allowing the NDMA to respond in a timely manner.

Proposal

Proposal

Data-Driven Methods

Physical models are the gold-standard but under-developed in vegetation health modelling.

Hydrological & Meteorological Data

$$X_{t-n} = \{x_{static}, x_{dynamic,t-n}\}$$

"Information on certain parameters is not only difficult to access … [we also] lack understanding of the different physical processes. This has led to the widespread use and development of **data-driven models over process-based models**"

Anshuka et. al. 2019

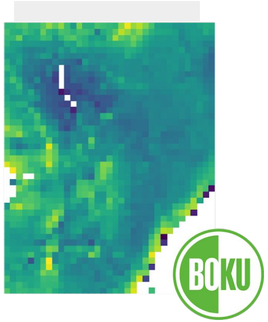
Regression Model

$$y_t = f(X_{t-1,...,t-n})$$

learning a **vegetation health** model

Satellite derived Drought Indicator

$$y_t \in [0, 100]$$

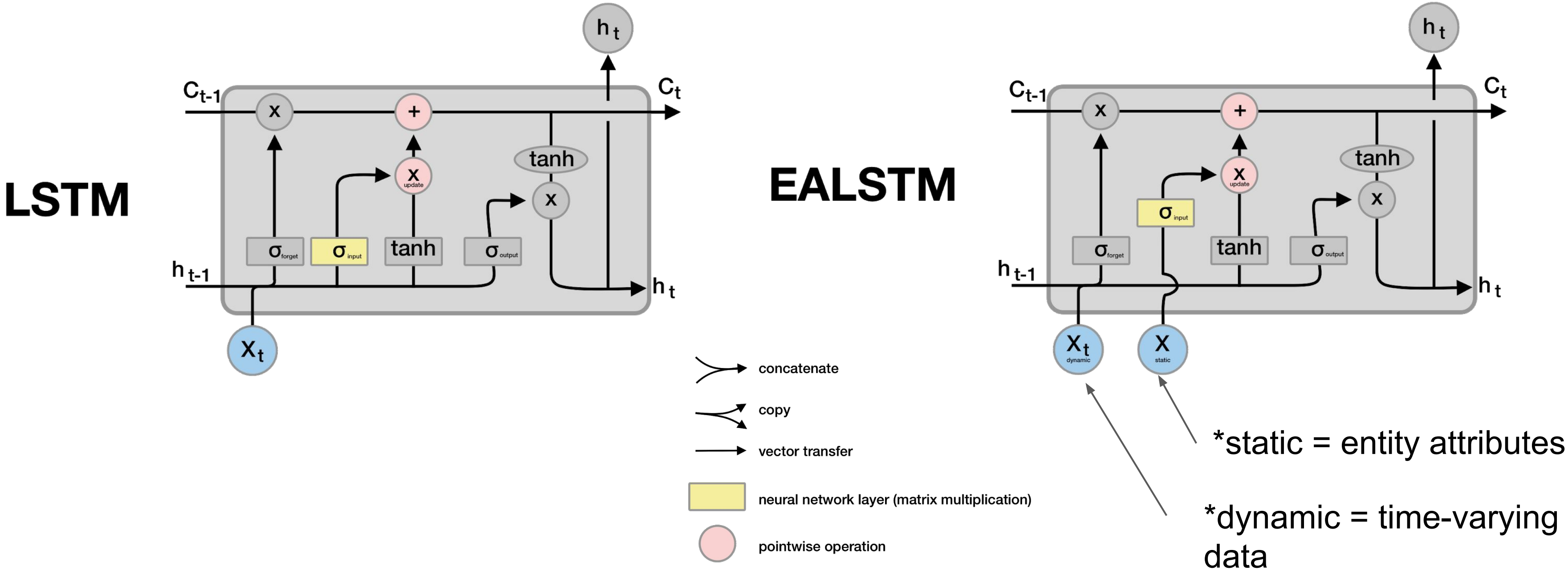


We use machine learning methods to predict the vegetation index.

Proposal

Entity Aware LSTM

The Entity Aware LSTM was used by Kratzert et. al. (2019) to model rainfall-runoff processes in the USA.

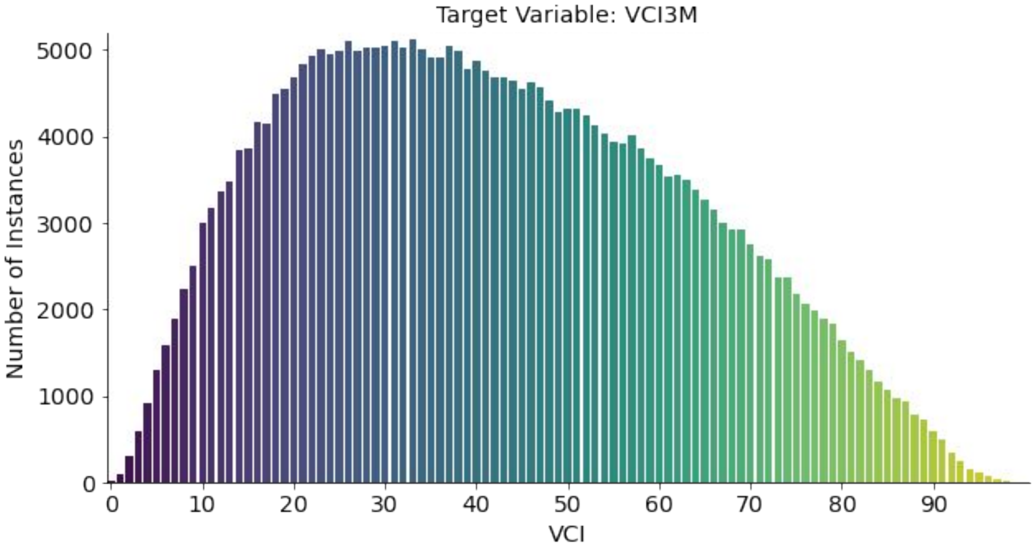


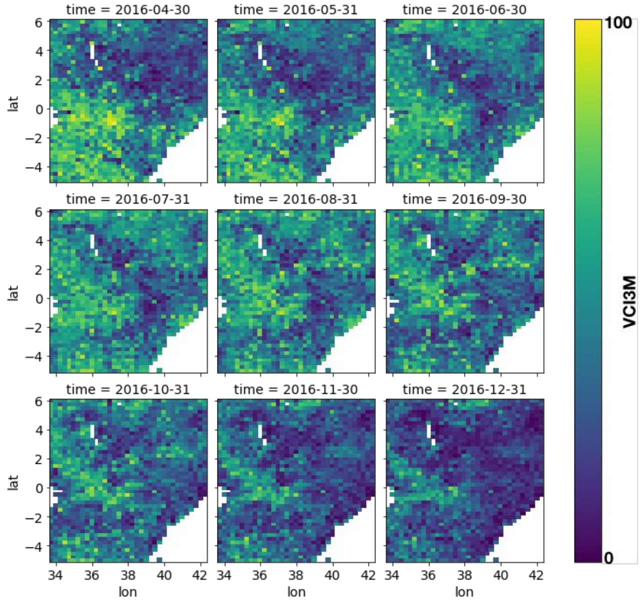
The EA-LSTM captures the idea that dynamical relationships will differ for different *entities* (locations) depending on static attributes.

Data

The 3 monthly mean Vegetation Condition Index (Klisch and Atzberger 2016) is used by the NDMA* in Kenya for monitoring drought conditions. It is derived from MODIS NDVI.

Target data





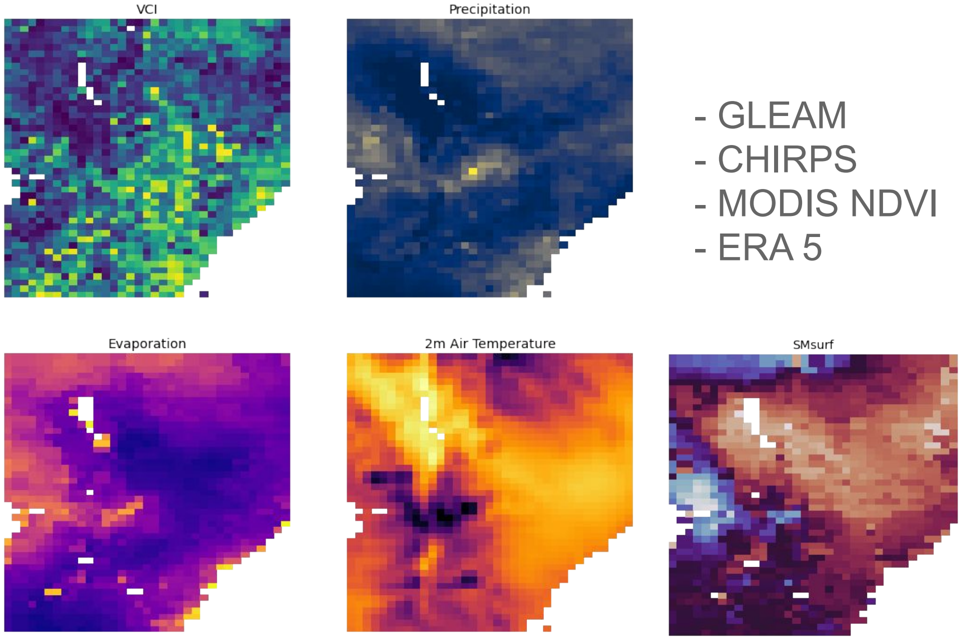
We predict vegetation health, a proxy for drought stress used in an operational context in Kenya.

*NDMA = National Drought Management Authority

Data

The dynamic data is made up of variables that vary over space AND time.

Dynamic Data



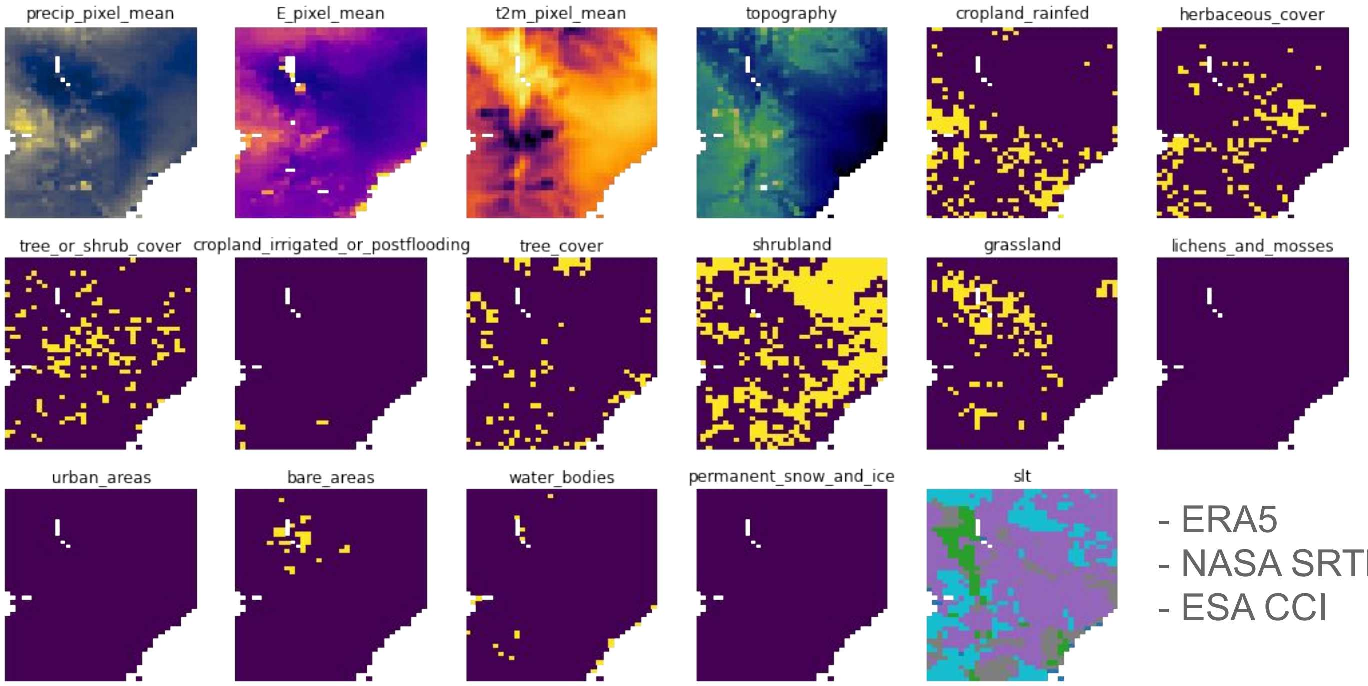
- GLEAM
- CHIRPS
- MODIS NDVI
- ERA 5

We use the previous 3 months of data to predict the next month Vegetation Condition Index.

Data

The static data is made up of variables that vary over space but NOT time.

Static Data



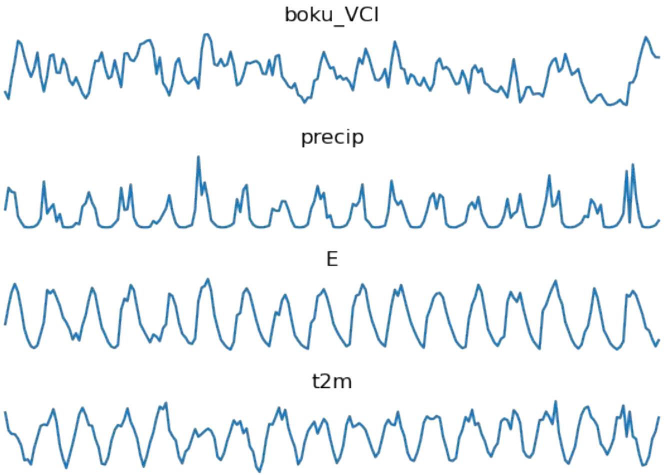
Note: These purple and yellow fields are boolean one-hot-encoded values for land cover classes.

- ERA5
- NASA SRTM
- ESA CCI

We include various static attributes including: land cover classes, topography, soil types and spatial aggregations of the dynamic variables.

Data

Time-Series



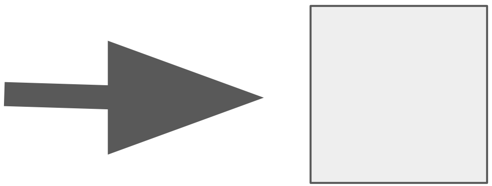
+ pixel attributes

X_dynamic

X_static

The forcing data comes from the previous 3 months (X_dynamic) and the static pixel attributes (X_static) to make a prediction of VCI. All data is **normalised** to have a standard deviation of 1 and a mean of 0.

y

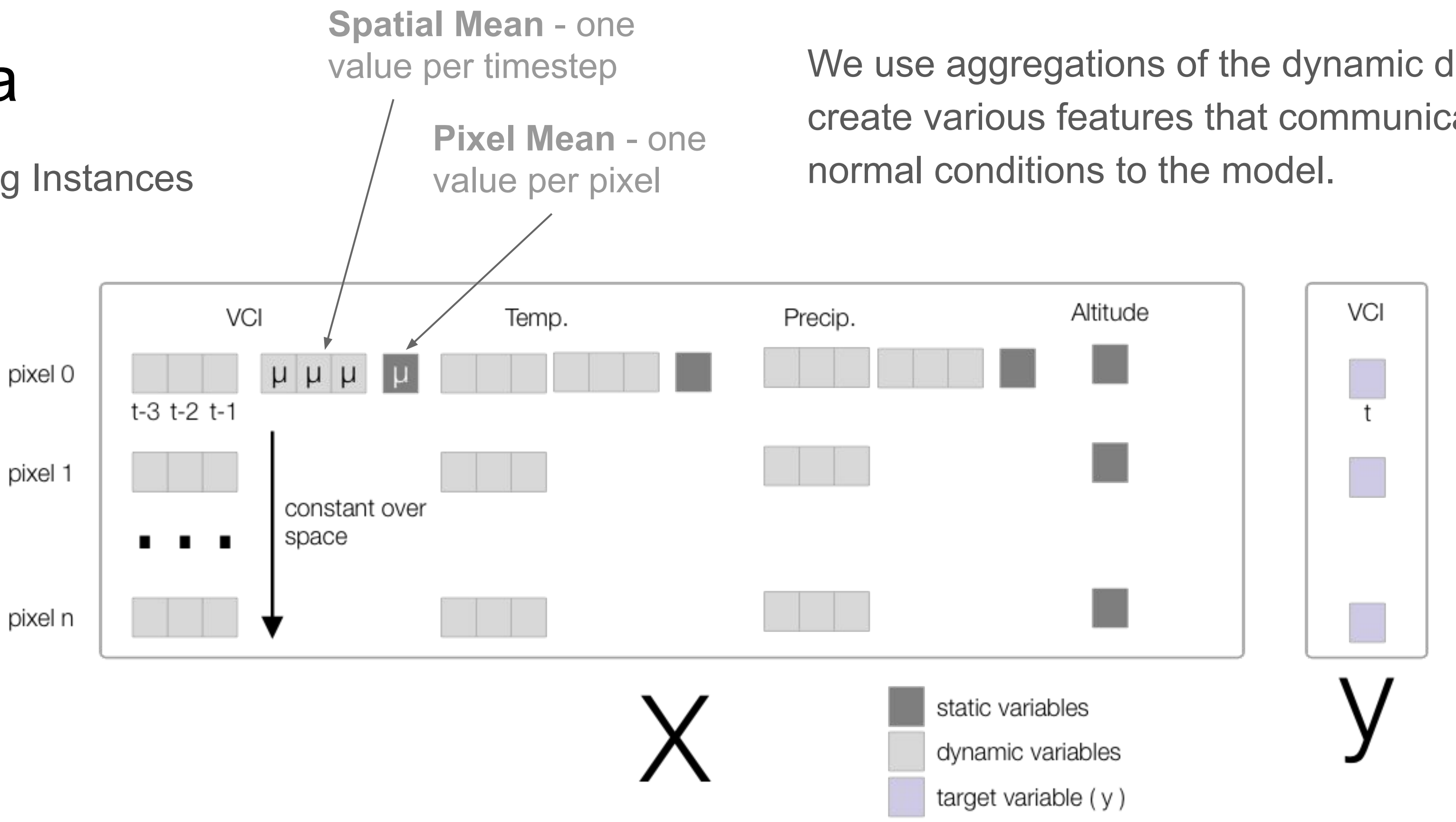


We force the model with time-varying and time-invariant data to make a prediction of a scalar value (y) one month ahead.

Data

Training Instances

We use aggregations of the dynamic data to create various features that communicate normal conditions to the model.



We treat each pixel as an independent observation of VCI to create instances of X, y pairs for model training and testing.

Model Performance

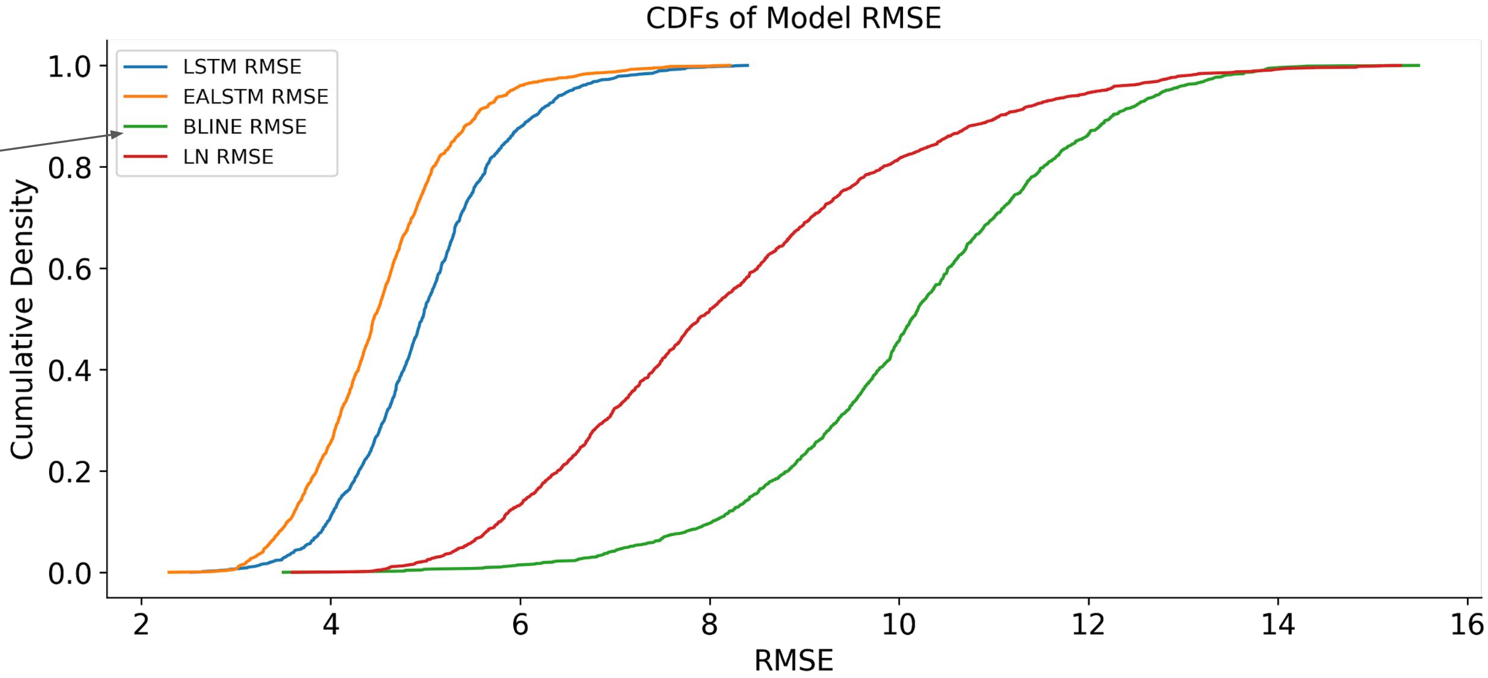
Results

Are we justified in using relatively complex models?

Comparison of Methods

The baseline model (BLINE) is a Persistence model.

"Predict no change from the previous timestep".

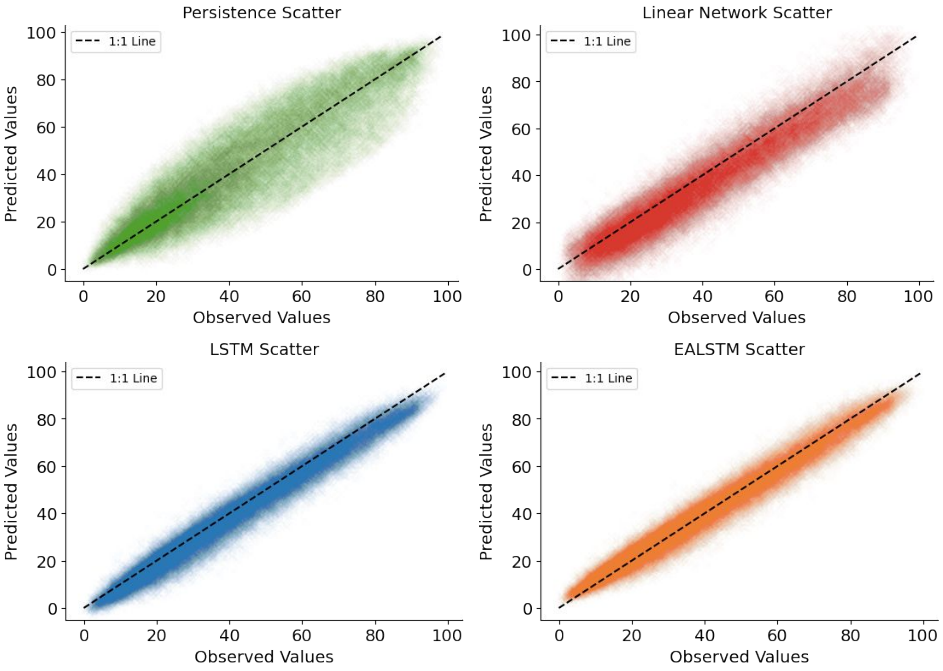


The LSTM and EALSTM perform similarly with the EALSTM slightly outperforming the LSTM.

Results

Comparison of Methods

We plot the Observed VCI3M values against the Predicted VCI3M values for each model to get an idea of overall model performance



Slight underprediction of values across the whole distribution

Slight overprediction of low values, underprediction of high values

The EALSTM outperforms the other models across the distribution.

Results

Geographical Errors

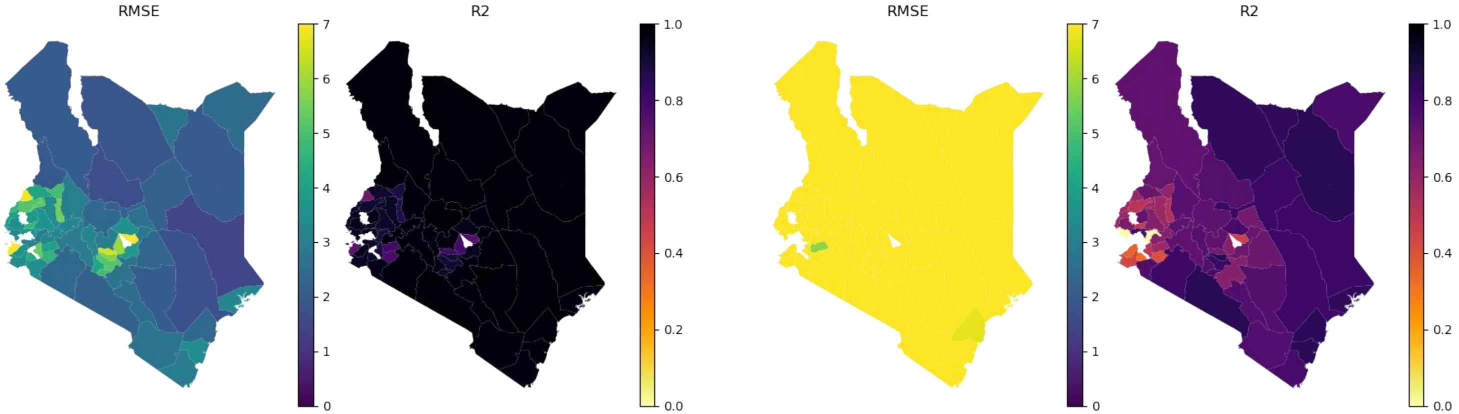
We calculate the mean values for each timestep for every pixel in each district of Kenya..

Darker colours reflect better performance.

Lighter colors reflect worse performance.

EALSTM

Persistence Baseline

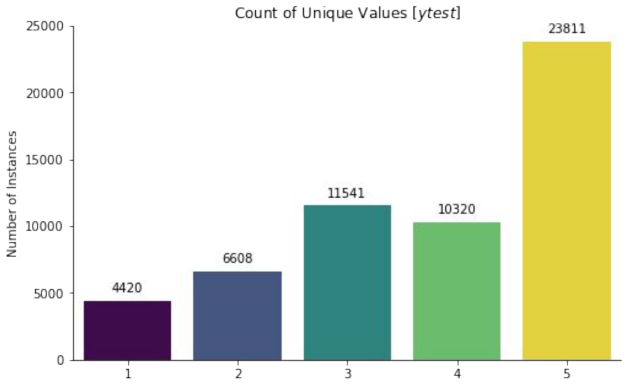


We perform well in the arid and semi-arid counties characterised by subsistence farming and pastoralists. We perform worst in the highly productive regions.

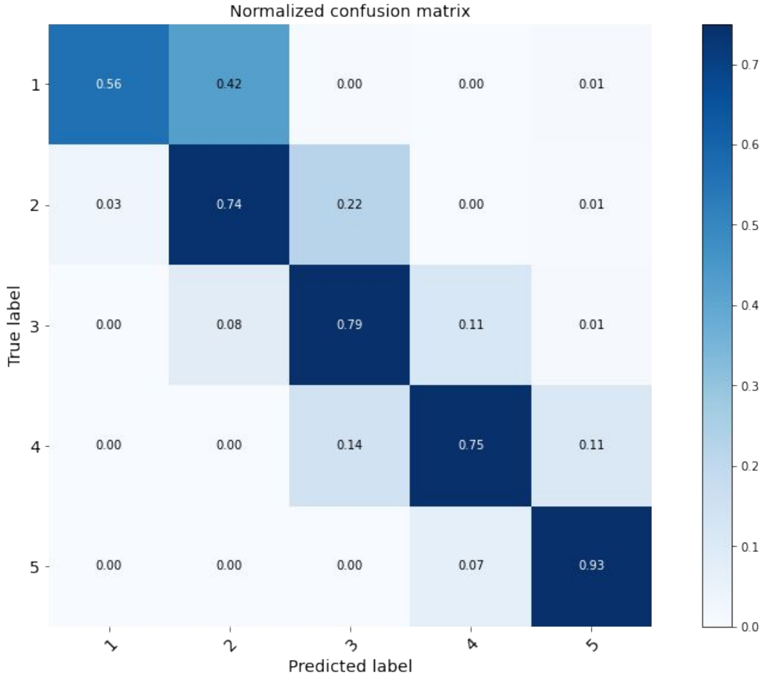
Results

Extremes

VCI3M Limits	Description	Value
0 <= x <10	Extreme vegetation deficit	1
10 <= x <20	Severe vegetation deficit	2
20 <= x <35	Moderate vegetation deficit	3
35 <= x <50	Normal vegetation conditions	4
50 <= x <=100	Above normal vegetation conditions	5



How does the EALSTM perform for the extreme conditions?



We perform well for most drought classes but the models could be improved for the most extreme drought conditions (Class 1).

Results

State-of-the-art

Adede et. al. (2019) use an ensemble of 111 linear neural networks or 111 support vector regression models to predict VCI3M in each district one month ahead.

District	Adede (2019)	Persistence	LSTM	EALSTM
Mandera	0.94	0.66	0.93	**0.95**
Marsabit	0.94	0.74	0.95	**0.96**
Turkana	0.91	0.74	0.94	**0.95**
Wajir	0.96	0.72	0.96	**0.97**

*Note: All values reported in the table are R2 Values (Coefficient of Determination)

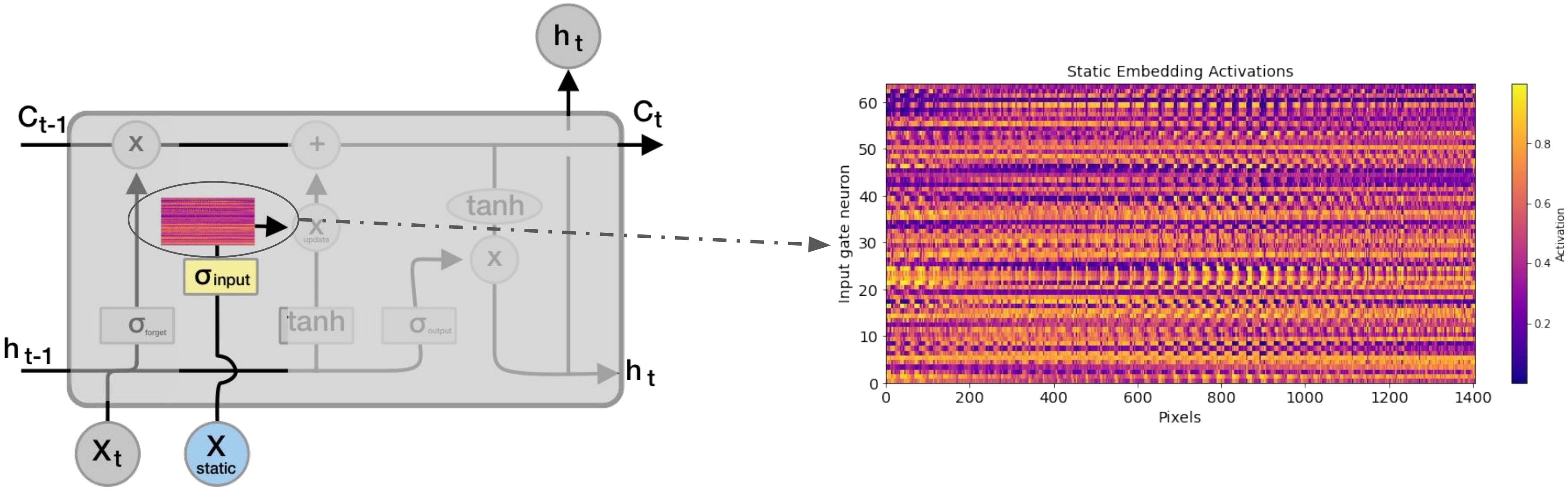
We are competitive with the Adede models and produce our forecasts at a much higher spatial resolution.

Clustering

Method

Clustering

The EALSTM passes the static data through a sigmoid layer, returning a vector of values unique to each pixel (*activations*).

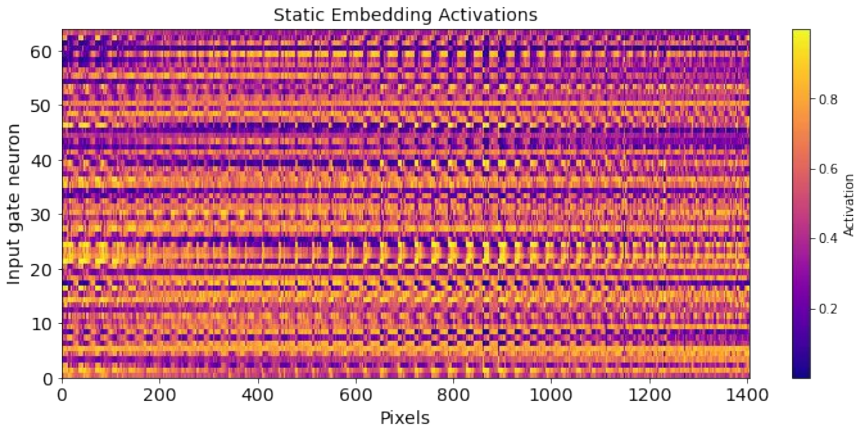


The EALSTM gives us the ability to extract how the model learns to group pixels with similar vegetation health behaviours.

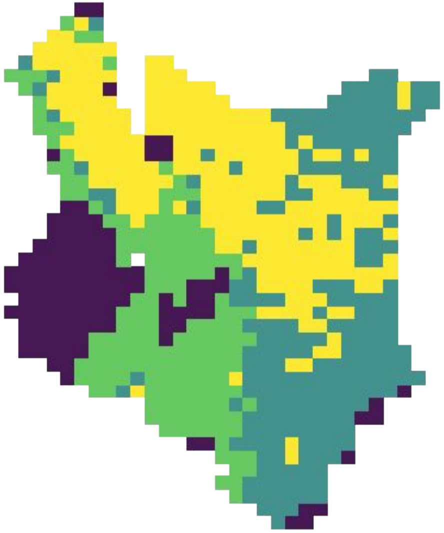
Method

Clustering

The trained model maps hydro-meteorological variables to vegetation health. This mapping is conditioned on the pixel-attributes (*X_static*).



K-means clustering

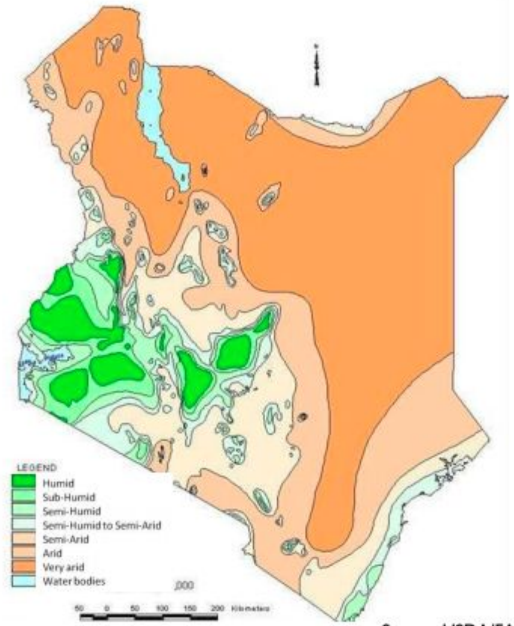


We clustered the static embedding, representing the parts of the network the model was learning to utilize to group similar behaviours.

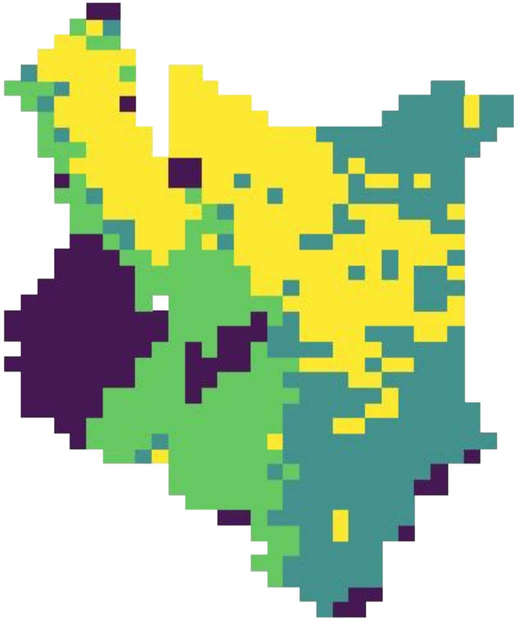
Results

Embeddings

The agro-ecological zones map on the left shows an expert delineation of vegetation regimes.



Source: USDA/FAS



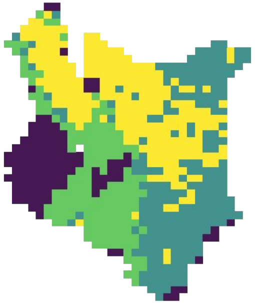
We have a region delineating the Turkana Channel, a dry and hot area of Kenya. A region for Lake Victoria and the coast, reflecting more moist areas. The Highlands Region delineating pixels of high topographic complexity. Finally a warm and relatively dry Eastern Region.

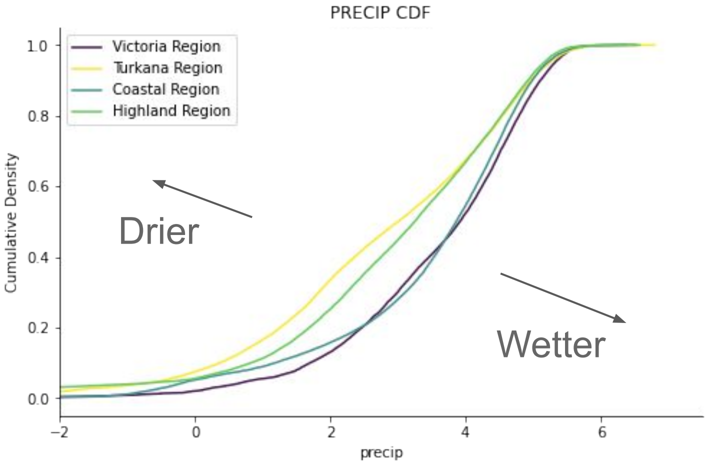
Visual inspection suggests the model learns groupings of pixels that are physically realistic.

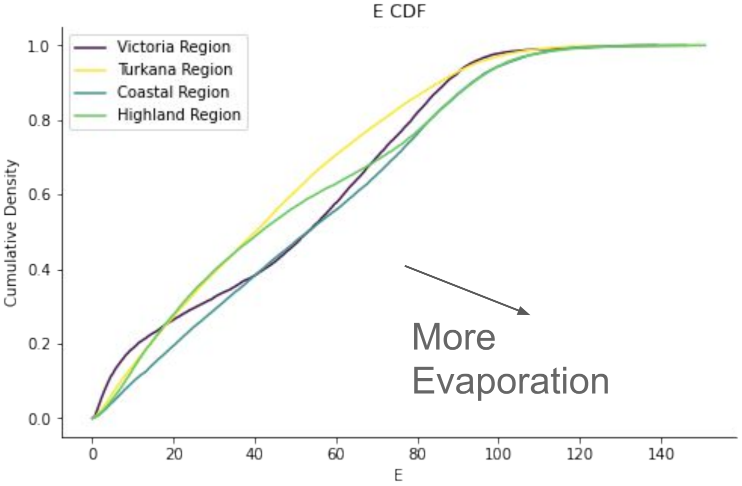
Results

Embeddings

We can look at the distribution of the dynamic variables in the different locations to interpret what the clusters mean.







The clusters are currently poorly delineated in terms of the characteristics of the regions. Interpreting these clusters requires further analysis.

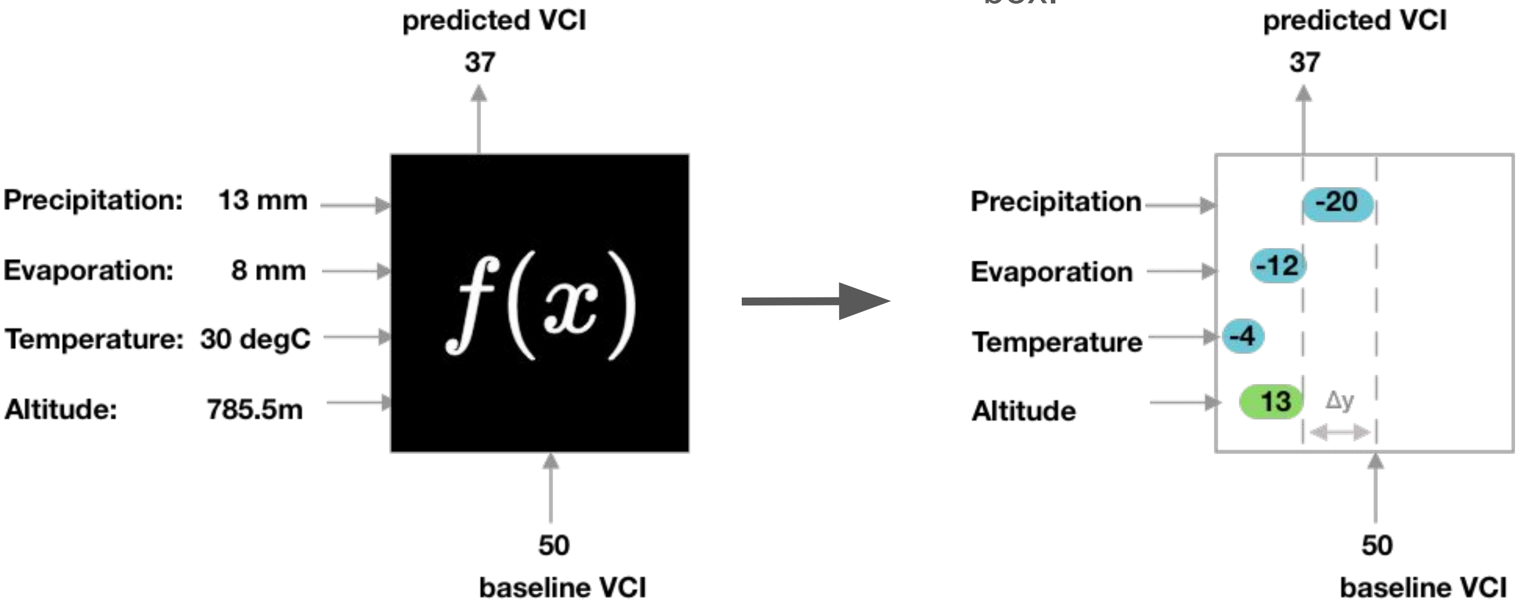
Model Interpretability

(Preliminary Results)

Method

DeepLIFT

DeepLIFT calculates feature importances by comparing the activation of layers between the input being explained and a baseline (e.g. the mean of the dataset). Opening the black box.



Δy represents the change in the y value.

Baseline - Predicted

We use DeepLIFT to approximate Shapley Values, determining the **instance-wise*** contribution of a feature to the model’s prediction.

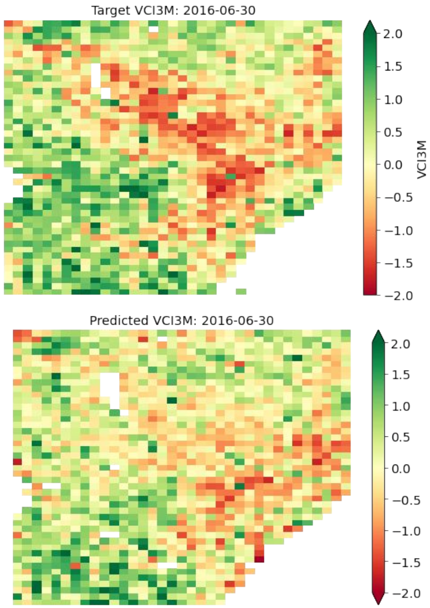
Method

DeepLIFT calculates the contribution of each feature for every individual prediction! These have a size of effect and direction of effect.

Local Feature Importance

Target (y)

Predicted (y_hat)

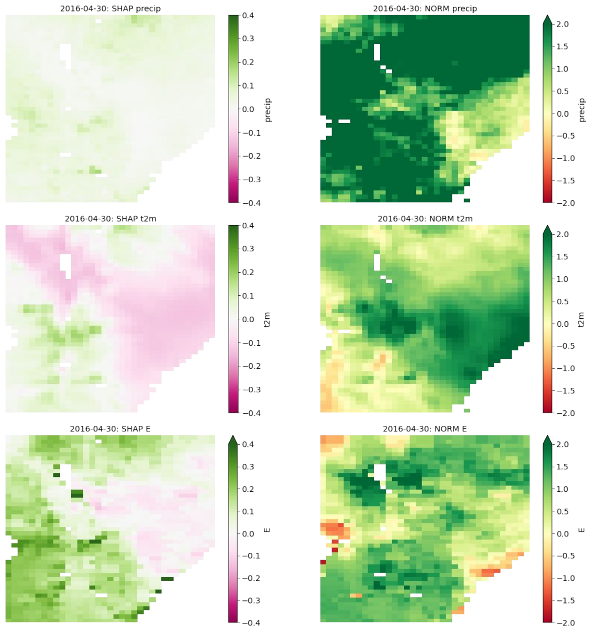


Feature Importance (SHAP)

Dark shows a feature is important. Light shows little contribution.

Green shows a positive contribution.

Pink shows a negative contribution.



Raw Value (NORM)

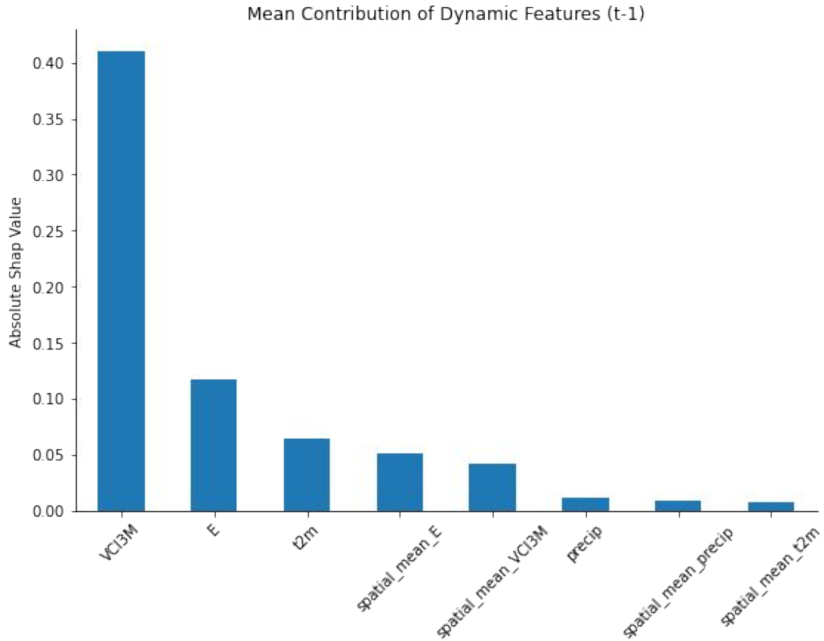
The normalised raw value allows us to determine the direction of the relationship.

The local feature importance tells us why the model made a particular prediction for a particular pixel-time (*instance*).

Results

Global Feature Importance

We can take a mean of all the individual feature contributions across all instances (X-y pairs). This gives us a way of measuring global feature importance.



Note: The **spatial mean** values are the mean values for the whole field. Giving the model information on how the neighbouring values vary in that timestep.

As we are predicting a 3 monthly moving average temporal autocorrelation is high. Therefore, the autoregressive feature comes out as the most important.

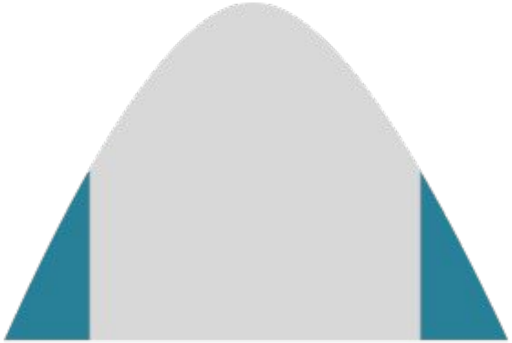
Discussion and Conclusions

Improvements

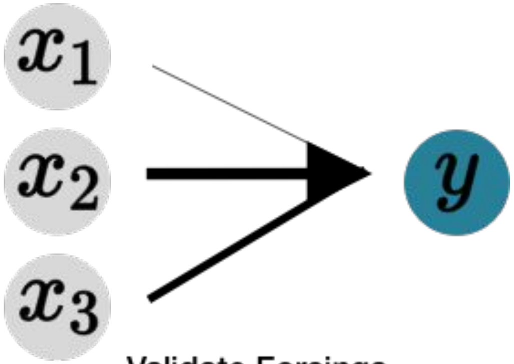
Requirements for Operational Models

Forecasts need a measure of certainty, to be accurate at predicting drought events, and should reflect known physical relationships.

Uncertainty Quantification



Predict the Extremes



Validate Forcings

Further work needs to be done to quantify uncertainty, improve prediction of the most extreme droughts and to explore the patterns the model learns.

Future Questions

We have identified a number of questions to improve the analysis. Do you have any thoughts?

How can we incorporate information about spatial autocorrelation? Perhaps using a convolutional structure to feed in 2D arrays of pixel values (as an image) rather than each pixel independently.

How can we better interpret the output of clustering the static embedding? We likely need more static variables to justify increasing the dimensionality of the data from 19 to 64.

How can we provide uncertainty estimates with our forecast? Two possibilities. 1) Use an ensemble of models, trained with different data or with different hyperparams. 2) Predict the Vegetation Deficit Index classes (1-5) directly and use the certainty of the class prediction.

How can we innovate using methods and visualisations from the field of interpretable machine learning? We have begun using DeepLIFT for understanding the contribution of features. Is this enough? Does this make sense for hydrological and climate applications?

How can we test model robustness? Withhold particular data points from the training data and observe how the error changes. Add noise to the input data.

Conclusion

Overview

We tested predictions of vegetation health 1 month ahead using machine learning methods driven by hydro-meteorological variables.

1. The Entity Aware LSTM **accurately predicts** a drought index (VCI3M) one month ahead.
2. We perform well for most drought classes, however, **performance can be improved for the most extreme droughts**.
3. Clustering analysis of the static embedding in the EALSTM offers a method to interpret which pixels the model decides should share behaviours.
4. The autoregressive feature contributes the most information to the predictions.
5. We need a measure of uncertainty for the forecasts to be operationally useful.

Machine learning methods allow us to make accurate predictions of a drought index in Kenya, providing timely information to improve outcomes.

Other Materials

More information and previous presentations of similar work undertaken in this area.

Links

ICLR 2020 Presentation Materials

Machine Learning Pipeline for Drought Prediction

Documentation for the Pipeline we developed

ECMWF Summer of Weather Code 2019 Presentation

These slides online

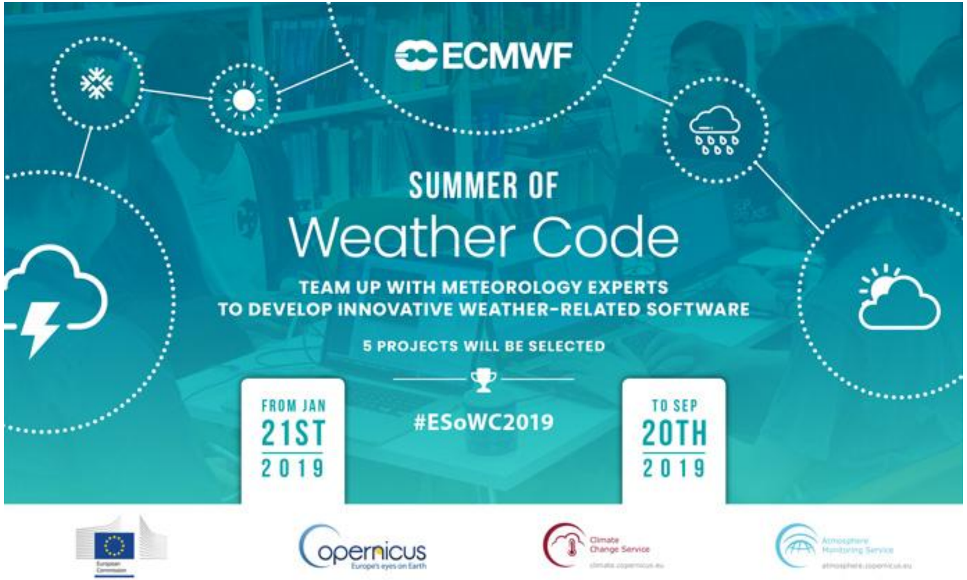
thomas.lees@chch.ox.ac.uk

@tommylees112

tommylees112

Please follow the links above and feel free to reach out to us on any medium.

Acknowledgements



Much of this work was completed as part of the ECMWF Summer of Weather Code 2019. Thank you to ECMWF for their support.