Supplementary Material

1 SUPPLEMENTARY TABLES

Table S1. Characteristics of the datasets used.

Dataset	Variable	Resolution	Source
CenTrends v1.0	Precipitation	0.1°	Funk et al. (2015)
GLEAM v3.5	SMroot SMsurf	0.25°	Martens et al. (2017); Miralles et al. (2010)
NOAA CDR v5	NDVI LAI	0.05°	Vermote (2019b) Vermote (2019a)
Berkeley	LST	1°	Rohde and Hausfather (2020)
COBE-SST2	SST	1°	Hirahara et al. (2014)
NCEP-NCAR CDAS-1	U200 U850	2.5°	Kalnay et al. (1996)
NOAA and KNMI	Climate indices (Table S2)	/	`https://psl.noaa.gov/data/climateindices/` `https://climexp.knmi.nl/selectindex.cgi`
ECMWF SEAS5	Precipitation hindcast	1°	(Johnson et al., 2019)

Table S2. Selected ocean–atmosphere oscillation indices and their abbreviations.

Climate Index	Abbreviation
Antarctic Oscillation	AAO
Atlantic Meridional Mode	AMM
Atlantic Multidecadal Oscillation	AMO
Arctic Oscillation Index	AO
El Niño–Southern Oscillation (Niño3.4)	ENSO
Indian Ocean Dipole	IOD
Madden–Julian Oscillation	MJO 1-10
Pacific Decadal Oscillation	PDO
Pacific Meridional Mode	PMM
Southern Annular Mode	SAM
Southern Oscillation Index	SOI
North Atlantic Oscillation	NAO
East Atlantic Pattern	EA
West Pacific Pattern	WP
East Pacific/ North Pacific Pattern	EP.NP
Pacific/North American Pattern	PNA
East Atlantic/West Russia Pattern	EA.WR
Scandinavian Pattern	SCA
Tropical/Northern Hemisphere Pattern	TNH
Polar/Eurasia Pattern	POL

Table S3. The KGE components value (r, α and β) results of nested cross-validation for all months and lead-times (1–3 months). The best values of A1 are indicated in blue and the best values of A2 are indicated in bold. For comparison, the SEAS5 KGE composition is additionally shown.

		KGE r		KGE α		KGE β	
		A1	A2	A1	A2	A1	A2
March							
L1	Lasso	0.48±0.52	-0.02±0.40	0.91±0.52	**0.95±0.81**	1.28±0.76	1.28± 0.83
	Ridge	0.48±0.58	0.01±0.47	0.78±0.51	0.87±0.67	1.13±0.50	1.33±0.90
	RF	0.54±0.18	**0.29±0.47**	0.56±0.47	0.69±0.61	1.34±0.60	1.33±0.77
	NN	0.53±0.49	-0.22±0.17	0.22±0.11	0.13±0.05	0.89±0.47	**0.76±0.29**
	SEAS5	0.73		0.59		1.09	
L2	Lasso	0.26±0.41	-0.02±0.10	1.10±1.11	1.47±1.80	1.28±1.08	1.25±1.61
	Ridge	0.34±0.54	-0.10±0.39	**1.10±0.96**	**0.97±0.86**	1.24±0.98	1.18± 1.39
	RF	0.25±0.35	**0.00±0.33**	0.39±0.35	0.41±0.35	0.98±0.45	**1.03±0.56**
	NN	0.37±0.38	-0.03±0.30	0.55±0.79	0.15±0.08	0.91±0.85	0.85±0.45
	SEAS5	0.18		0.25		0.76	
L3	Lasso	0.24±0.23	**-0.14±0.28**	1.81±2.48	**0.98±1.15**	0.71±0.36	1.41±0.89
	Ridge	0.24±0.29	-0.23±0.30	1.04±1.03	1.13±1.45	1.21±0.89	1.35±0.95
	RF	0.13±0.44	-0.24±0.42	0.51±0.56	0.63±0.58	1.24±0.94	1.33±0.91
	NN	-0.10±0.47	-0.16±0.32	0.46±0.26	0.10±0.10	1.48±0.97	**0.77±0.38**
	SEAS5	0.04		0.21		0.67	
April							
L1	Lasso	0.44±0.40	-0.10±0.23	0.54±0.11	**0.60±0.23**	0.99±0.12	1.10±0.11
	Ridge	0.52±0.40	-0.14±0.28	0.50±0.12	0.40±0.12	1.00±0.14	1.03±0.17
	RF	0.25±0.37	**0.12±0.29**	0.31±0.07	0.30±0.09	1.03±0.17	**1.00±0.18**
	NN	0.40±0.28	0.06±0.13	0.29±0.12	0.24±0.06	0.94±0.18	0.85±0.19
	SEAS5	0.50		0.57		0.89	
L2	Lasso	0.43±0.19	-0.25±0.25	0.59±0.17	**0.62±0.20**	1.06±0.08	1.04±0.10
	Ridge	0.64±0.18	-0.20±0.38	0.55±0.11	0.48±0.13	1.03±0.12	1.09±0.17
	RF	0.38±0.13	-0.24±0.42	0.36±0.09	0.31±0.08	0.96±0.10	1.03±0.18
	NN	0.60±0.32	**0.26±0.35**	0.45±0.14	0.30±0.06	0.92±0.18	**1.01±0.30**
	SEAS5	0.32		0.29		0.73	
L3	Lasso	0.37±0.34	0.06±0.16	0.56±0.18	**0.62±0.13**	1.03±0.15	1.05±0.16
	Ridge	0.60±0.20	**0.17±0.15**	0.59±0.07	0.42±0.07	1.03±0.12	1.07±0.16
	RF	0.56±0.20	-0.10±0.33	0.25±0.05	0.29±0.08	1.01±0.14	1.03±0.18
	NN	0.32±0.51	-0.11±0.30	0.26±0.11	0.18±0.06	0.98±0.25	**1.02±0.26**
	SEAS5	-0.15		0.25		0.67	
May							
L1	Lasso	0.38±0.36	0.00±0.33	0.62±0.17	0.48±0.16	1.10±0.11	1.00±0.27
	Ridge	0.71±0.16	0.31±0.40	0.64±0.17	**0.53±0.18**	1.04±0.06	1.01±0.29
	RF	0.60±0.23	**0.37±0.16**	0.33±0.11	0.38±0.12	1.00±0.17	**1.00±0.21**
	NN	0.64±0.20	0.17±0.38	0.35±0.05	0.21±0.08	0.99±0.18	0.99±0.32
	SEAS5	0.80		0.58		0.93	
L2	Lasso	0.44±0.29	0.05±0.19	**0.53±0.17**	0.53±0.12	1.06±0.23	1.03±0.34
	Ridge	0.27±0.26	0.00±0.34	0.50±0.17	0.49±0.17	1.05±0.21	1.04±0.33
	RF	0.34±0.15	**0.27±0.20**	0.43±0.05	0.35±0.07	1.06±0.22	0.98±0.27
	NN	0.48±0.23	0.20±0.33	0.35±0.19	0.25±0.12	0.93±0.30	0.99±0.33
	SEAS5	0.61		0.23		0.89	
L3	Lasso	0.21±0.32	-0.30±0.3	0.51±0.23	**0.60±0.20**	1.07±0.28	0.98±0.44
	Ridge	0.42±0.31	-0.25±0.37	0.54±0.16	0.50±0.12	1.15±0.09	1.05±0.34
	RF	0.49±0.23	-0.05±0.21	0.28±0.02	0.32±0.14	1.08±0.24	1.08±0.30
	NN	0.39±0.18	**0.19±0.23**	0.25±0.04	0.16±0.04	0.91±0.24	**0.99±0.32**
	SEAS5	0.65		0.25		0.93	

2 SUPPLEMENTARY FIGURES

Figure 1a. March.

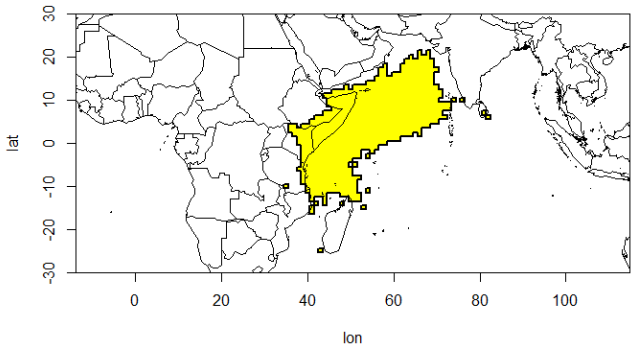

Figure 1b. April.

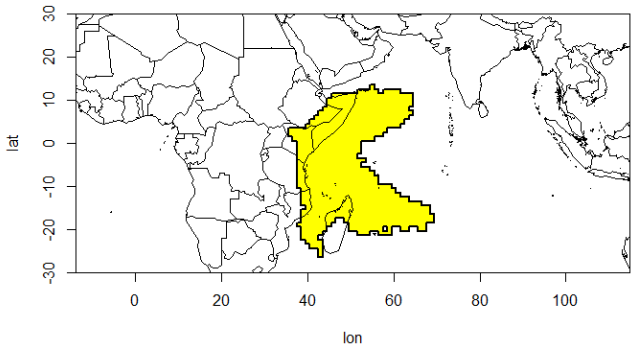

Figure 1c. May.

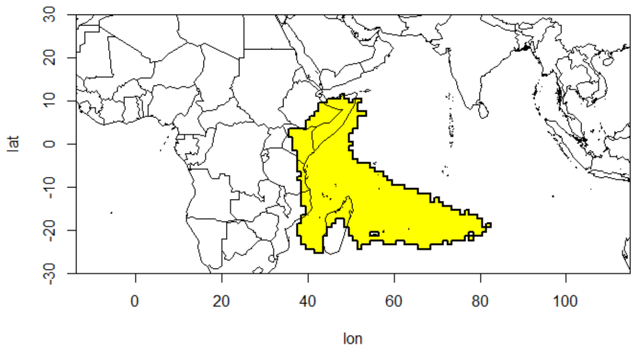

Figure 1. Regions sourcing moisture during the long rains corresponding to the 90th quantile of precipitation (based on FLEXPART).

3 SUPPLEMENTARY MATERIAL: PSEUDOCODES

Algorithm 1: Nested cross-validation procedure using grid search with correlation-based predictor selection done on all available data (1981–2014), prior to cross-validation procedure.

```
Load D; dataset containing the input features (i.e. predictors) X (with correlations over
  whole period 1981–2014) and the output feature y (rainfall);
Define F_1 and F_2; F_1 is the number of outer folds and F_2 is the number of inner folds;
Define P; contains n different sets of hyperparameters;
Define K; contains m values for k in the univariate feature selection method;
Define M; the model assessed;
Split D in F_1 folds;
for i = 1 to F_1 folds do
    Set fold i as D_{i,test};
    Set remaining folds as D_{i,train};
    Split D_{i,train} in F_2 folds;
    for j = 1 to F_2 folds do
        for feat = 1 to m do
            Select K_{feat} features from the D_{i,train} input dataset; D_{i,k,train};
            Set fold j as D_{j,k,tune};
            Set remaining folds as D_{j,k,train};
            for p=1 to n do
                Set the hyperparameters to the ones defined in P_p;
                Train M on D_{j,k,train} with the hyperparameter set;
                Evaluate model performance with D_{j,k,tune};
            end
        end
    end
    Select optimal hyperparameter set p' and optimal feature selection number k' where
      average model performance over the F_2 folds is best;
    Train M with optimal parameter set P_{p'} on D_{i,k',train};
    Evaluate model performance with D_{i,k',test};
end
```

Algorithm 2: Nested cross-validation procedure using grid search with correlation-based predictor selection in the cross-validation procedure (on training data).

Load D_a; dataset containing the physics-guided input features X_a and the output feature y (rainfall);
Define F_1 and F_2; F_1 is the number of outer folds and F_2 is the number of inner folds;
Define P; contains n different sets of hyperparameters;
Define K; contains m values for k in the univariate feature selection method;
Define M; the model assessed;
Split D_a in F_1 folds;
for $i = 1$ *to* F_1 *folds* **do**
    Define D_b; dataset containing significantly correlated input data of SSTs, winds (200 mb and 850 mb) and climate indices (method: Section 3.4.1), only based on correlations with y for data of all years except those in fold i;
    Merge D_a and D_b as full input dataset D;
    Set fold i as $D_{i,test}$;
    Set remaining folds as $D_{i,train}$;
    Split $D_{i,train}$ in F_2 folds;
    for $j = 1$ *to* F_2 *folds* **do**
        for $feat = 1$ *to* m **do**
            Select K_{feat} features from the $D_{i,train}$ input dataset; $D_{i,k,train}$;
            Set fold j as $D_{j,k,tune}$;
            Set remaining folds as $D_{j,k,train}$;
            for $p = 1$ *to* n **do**
                Set the hyperparameters to the ones defined in P_p;
                Train M on $D_{j,k,train}$ with the hyperparameter set;
                Evaluate model performance with $D_{j,k,tune}$;
            end
        end
    end
    Select optimal hyperparameter set p' and optimal feature selection number k' where average model performance over the F_2 folds is best;
    Train M with optimal parameter set $P_{p'}$ on $D_{i,k',train}$;
    Evaluate model performance with $D_{i,k',test}$;
end

4 SUPPLEMENTARY MATERIAL: HYPERPARAMETER TUNING AND ADDITIONAL FEATURE SELECTION

4.1 Hyperparameter tuning

Hyperparameters are parameters that have to be optimized based on the input dataset. In this study, hyperparameters of the different methods (ridge, lasso, random forest and neural network) were tuned in the inner cross-validation loop with a grid search algorithm. For this, a discrete parameter space had to be defined after which all possible combinations were evaluated (Probst et al., 2019). Tuning was done in the inner cross-validation loop (Figure 2) which overcomes the problem of the parameters overfitting to the training data. The performance measure that was used for optimization is the R-squared metric.

For both the lasso and ridge regression, the only hyperparameter that needed tuning is λ which is a measure of the amount of shrinkage: a larger λ leads to a greater amount of shrinkage (Hastie et al., 2009). In this study λ was varied over 50 values in a uniformly-spaced grid with as minimum and maximum values 10^{-4} and 10 respectively.

The random forest has multiple hyperparameters that can be tuned. Probst et al. (2019) stated that default values often work well and additionally declared that it is not completely clear which hyperparameters should be tuned routinely. They stated that the effect of tuning is much smaller than for other algorithms. Therefore, it was opted to use default values of all parameters except the two most important ones (number of trees and number of subsetted predictors considered when making a split, according to Clauwaert and Waegeman (2022)) over which a grid search was done (Table S4).

Table S4. Random forest grid description of hyperparameters that were tuned.

Hyperparameter	Minimum value	Maximum value	Step size
Number of trees	10	100	5
Number of predictors considered at split	1	10	1

Lastly, a single-layer neural network was used. Neural networks have a lot of hyperparameters to tune. As tuning all of them would be computationally expensive, some of them were fixed; the fixed hyperparameters are shown in Table S5 and the tuned ones in Table S6. To decide the values of the fixed parameters, exploratory studies on the data were done by trail and error on simple test–train splits and by consulting literature. Both an early stopping mechanism[1] and dropout regularisation were added to avoid overfitting. The maximum number of neurons that was used in the hidden layer was dependent on the number of predictors.

Table S5. Fixed single-layer neural network hyperparameters.

Hyperparameter	Fixed Value/Method
Optimizer	Adam
Batch size	15
Maximum epochs	150
Dropout	0.2

Table S6. Values over which single-layer neural network hyperparameters were tuned.

Hyperparameter	Grid
Number of neurons	min: 8, max: number of predictors, step size: 8
Learning rate	[0.0005; 0.001; 0.005]
Activation function	[ReLu; Sigmoid]

4.2 Additional feature selection

Next to the main two predictor selection approaches (A1 and A2), there was some additional feature selection done in the inner cross-validation loop for both approaches as the number of predictors can become large as it depends on the number of significantly correlated clusters of SSTs and winds. The number of predictors can then also be seen as parameter of regularisation, therefore tuning the number

[1] After ten epochs with no improvement on validation loss, training was stopped.

of predictors in the inner cross-validation loop could improve the performance results without overfitting. Especially for data with a high number of predictors compared to observations, some feature selection can help improve forecasts (Hastie et al., 2009). A univariate selection method was used where Pearson correlation was done between the target (monthly rainfall) and the predictors, and subsequently ranked the predictors accordingly, followed by the selection of the K best predictors. In the inner loop, the number of correlation-based predictors (K) that were selected as input in the model was thus tuned. In this additional feature selection, K varies over different values; in this study, it was opted to go from selecting the best 20% of the significantly correlated data to selecting 100% in steps of 20%.

REFERENCES

Clauwaert, J. and Waegeman, W. (2022). *Predictive modelling course: Tree methods (PClab nr. 7)*

Funk, C., Nicholson, S. E., Landsfeld, M., Klotter, D., Peterson, P., and Harrison, L. (2015). The Centennial Trends Greater Horn of Africa precipitation dataset. *Scientific data* 2, 1–17

Hastie, T., Tibshirani, R., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, vol. 2 (Springer)

Hirahara, S., Ishii, M., and Fukuda, Y. (2014). Centennial-scale sea surface temperature analysis and its uncertainty. *Journal of Climate* 27, 57–75

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., et al. (2019). SEAS5: the new ECMWF seasonal forecast system. *Geoscientific Model Development* 12, 1087–1117

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American meteorological Society* 77, 437–471

Martens, B., Miralles, D. G., Lievens, H., Van Der Schalie, R., De Jeu, R. A., Fernández-Prieto, D., et al. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development* 10, 1903–1925

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J. (2010). Global land-surface evaporation estimated from satellite-based observations. *Hydrol. Earth Syst. Sci. Discuss* 7, 8479–8519

Probst, P., Wright, M. N., and Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 9, e1301

Rohde, R. A. and Hausfather, Z. (2020). The Berkeley Earth land/ocean temperature record. *Earth System Science Data* 12, 3469–3479

[Dataset] Vermote, E. (2019a). NOAA Climate Data Record (CDR) of AVHRR Leaf Area Index (LAI) and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), Version 5

[Dataset] Vermote, E. (2019b). NOAA Climate Data Record (CDR) of AVHRR Normalized Difference Vegetation Index (NDVI), Version 5.