

1 **Hybrid (dynamical-statistical) models to improve seasonal drought** 2 **prediction: application to regional meteorological droughts in China**

3 Zhiyong Wu^{1,2} · Hao Yin¹ · Hai He¹ · Yuan Li¹

4 ¹College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China

5 ²Yangtze Institute for Conservation and Development, Hohai University, Nanjing 210098, China

6 *Correspondence to:* Zhiyong Wu (wzyhhu@gmail.com)

7 **Abstract**

8 Accurate drought prediction is important for drought resistance and water resources management. However, the seasonal
9 drought prediction is of low accuracy for both dynamical and statistical models. In this study, we combined dynamical models
10 and machine learning to construct hybrid (dynamical-statistical) models. We used the random forest approach to identify
11 representative regions based on geopotential height, sea-level pressure, and 2-m temperature. The least absolute shrinkage and
12 selection operator (Lasso and an artificial neural network (ANN) were used to construct the statistical models, with atmospheric
13 variables as predictors and 3-month Standardized Precipitation Index (SPI3) as the predictand. The atmospheric variables
14 forecasted by the European Centre for Medium-Range Weather Forecasts (ECMWF) SEAS5 model were processed as
15 predictors to force the statistical models. The resulting hybrid models, constructed using dynamical models and machine
16 learning, were named as dynamic-Lasso ('D-Lasso') and dynamic-ANN ('D-ANN') separately. The results suggested that
17 prediction skills were improved by the hybrid models; compared to the best available dynamical model (UK Met Office), D-
18 ANN extends the forecast horizons by 6, 21, and 4 days in northern, eastern, and southern China, respectively. In spring and
19 summer, the correlation skills were also improved. The effective prediction of the atmospheric anomalies over the eastern and
20 southern Tibetan Plateau and the Northwest Pacific region was identified as the main contributor to successful seasonal drought
21 prediction. Overall, the hybrid models were able to predict drought processes effectively, and D-ANN outperformed the D-
22 Lasso in drought onset and persistence phases.

23 **Key words** Seasonal drought prediction · Hybrid models · Machine learning · China

24 **1 Introduction**

25 Droughts are natural disasters that occur under all climate regimes (Liu et al. 2016). Global chronically drought-prone areas
26 have increased significantly, from 16.19% between 1902 and 1949 to 41.09% between 1950 and 2008 (Wang et al. 2014).
27 Globally, an estimated 55 million people are affected by droughts every year, and as many as 700 million people are at risk of

28 being displaced as a result of drought by 2030 (World Health Organization 2019). According to the Food and Agriculture
29 Organization (FAO) of the United States, developing countries suffered losses of \$29 billion in agriculture due to droughts
30 between 2005 and 2015 (Conforti et al. 2018). Therefore, effective drought prediction is important to improve drought
31 resistance and water resources management, and reduce drought losses.

32 With the increase of our understanding of the climate system and computational capabilities, predicting droughts using general
33 circulation models (GCMs) has become a fundamental approach (Hao et al. 2018). Seasonal drought indices, such as the
34 Standardized Precipitation Index (SPI) and the Standardized Precipitation Evapotranspiration Index (SPEI), can be calculated
35 from GCM precipitation and temperature outputs (Lang et al. 2020; Ma et al. 2015; Ujeneza and Abiodun 2015). The advantage
36 of GCMs is that they are based on physical principles and can capture the nonlinear changes in the ocean, atmosphere, and on
37 land. In recent years, significant advances have been made in drought prediction, with the widespread use of post-processing
38 and multi-model ensembles (Li et al. 2020; Mo and Lyon 2015). Nevertheless, GCMs are limited by chaotic properties, initial
39 value errors, and model biases, which lead to uncertainties in the forecasted precipitation with a lead time of more than two
40 weeks (Li et al. 2017; Merryfield et al. 2020). Statistical methods usually first identify a range of predictors from historical
41 hydroclimatic observations (reanalysis) which are then input into statistical models to predict drought. Statistical methods are
42 less computationally demanding and usually provide a baseline level of skill (Kirtman et al. 2013). However, unclear physical
43 mechanisms and unstable statistical relationships in climate change are always the problems of statistical methods (Hao et al.
44 2018). To combine the advantages of dynamical and statistical models, many researchers have attempted to construct hybrid
45 (dynamical-statistical) models.

46 Hybrid models mainly involve two steps, first, calibration to correct the bias (and ensemble spread) of GCMs forecasts and,
47 second, merging forecasts from multiple sources (Madadgar et al. 2016; Ribeiro and Pires 2016; Yan et al. 2017). Previous
48 efforts have focused on combining forecasts from multiple dynamical and statistical models, which has achieved excellent
49 results (Hao et al. 2016; Madadgar et al. 2016; Wang et al. 2012). Nevertheless, some variables, such as geopotential height
50 and temperature, yield higher skill scores with long lead times and are often used to analyze atmospheric circulation anomalies
51 during droughts but are rarely used for prediction. For seasonal drought, large-scale circulation is an important mechanism,
52 and the role of atmospheric anomalies varies between droughts (Lhotka et al. 2020; Li et al. 2018b; Okumura et al. 2017;
53 Zhuang et al. 2020). Thus, some studies have aimed to extract information from atmospheric circulation patterns to predict
54 drought events. For example, Costa-Cabral et al. (2016) were able to project meteorological drought in California using the
55 North Pacific High (NPC) sea-level pressure anomaly; in northeastern Europe, more than 65% of drought events were detected
56 one month in advance by Lavaysse et al. (2018); and Richardson et al. (2020) explored the potential benefits of using weather
57 pattern predictions to improve sub-seasonal forecasts of meteorological drought. In these studies, the selection of an
58 appropriate method to obtain key information from atmospheric variables was critical.

59 With the explosive growth of observation (and reanalysis) data and the considerable increases in computing capabilities,
60 climate prediction with machine learning (ML) has become an active area of research (Reichstein et al. 2019). The most
61 significant advantage of ML is its ability to mine information from ‘big data’ for future prediction. In the past few years, ML

62 has helped improve climate predictions significantly as well as assist in interpreting the underlying mechanisms (Pan et al.
63 2020; Yoo-Geun et al. 2019), and this approach has also been used to improve drought prediction (Agana and Homaifar 2017;
64 Khan et al. 2020; Rezaeianzadeh and Tabari 2012; Zhang et al. 2019). However, drought mechanisms vary between different
65 regions and across temporal-spatial scales, meaning that further research is needed to determine whether ML is effective for
66 seasonal drought prediction.

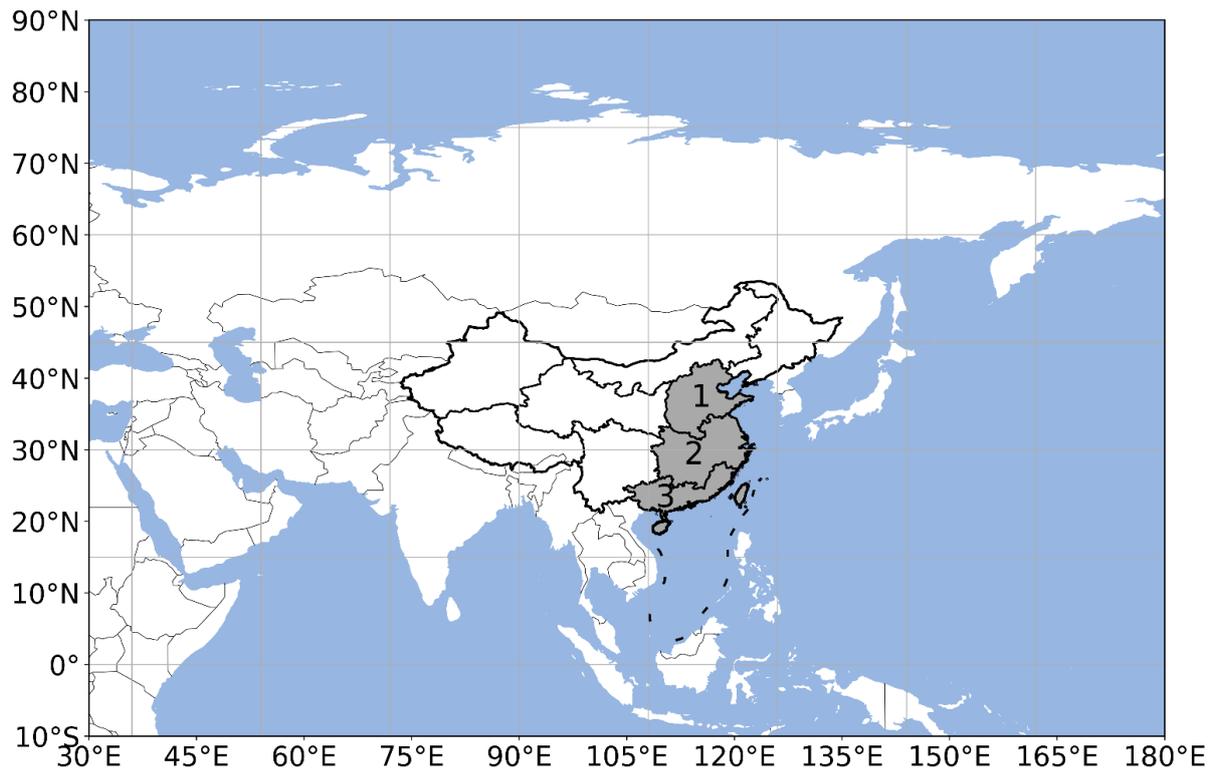
67 China has long been affected by drought due to its geographical location and climatic regime (Ayantobo et al. 2017), and in
68 recent years, extreme droughts have become increasingly frequent and severe in eastern and southern China (Chen et al. 2018;
69 Jin et al. 2013; Ma et al. 2020). For example, the great drought in southwestern China from September 2009 to March 2010
70 affected more than 69 million people and 6.6 million hectares of crops, with economic losses of more than 40 billion yuan
71 (China Meteorological Administration 2011). In 2013, drought in southern China affected more than 80 million people and
72 damaged 7.9 million hectares of crops (China Meteorological Administration 2014). While progress has been made in drought
73 prediction in China using dynamical models, statistical methods, and hybrid models (Xu et al. 2018a; Xu et al. 2018b; Zhang
74 et al. 2019; Zhu et al. 2020), there have been few attempts to employ atmospheric circulation datasets. Therefore, the
75 applicability of ML for seasonal drought prediction in different regions of China remains unknown.

76 In this study, we combine dynamical models and ML to construct hybrid models. We use a random forest (RF) approach to
77 identify typical regions from atmospheric circulation variables. The least absolute shrinkage and selection operator (Lasso)
78 and an artificial neural network (ANN) are used to extract statistical relationships between atmospheric variables in
79 representative regions and meteorological drought. The atmospheric variables forecasted by the dynamical models are then
80 processed to generate drought-prediction models. The remainder of this paper is structured as follows: Section 2 introduces
81 the study area and data; the ML models and predictor construction methods are described in Sect. 3; Section 4 presents the
82 prediction results; we will discuss the study limitations and directions for future work in Sect. 5; and the conclusions are
83 presented in Sect. 6.

84 **2 Study areas and data**

85 **2.1 Study areas**

86 Three of the nine drought regions of China (Fig. 1), namely Northern China (North), Eastern China (East), and Southern China
87 (South), were selected as the study regions, covering approximately 0.69, 0.91, and 0.37 million km², respectively. The East
88 Asian Monsoon climate greatly influences these regions, and precipitation shows large inter- and intra-annual variability (Cai
89 et al. 2017; Liu and Huang 2019; Zhang 2015).



90

91 **Fig. 1** The location of the drought regions of China and the coverage of atmospheric variables (30°-180°E, 10°S-90°N). The gray
 92 shaded areas are the study regions. 1-3 represent Northern, Eastern, and Southern China, respectively

93 **2.1 Data**

94 **2.2.1 Observations**

95 Observed daily precipitation amounts from the Climate Data Centre (CDC) of the National Meteorological Information Centre,
 96 China Meteorological Administration (CMA), were used, providing $0.5^\circ \times 0.5^\circ$ gridded data for the period 1979–2019
 97 (http://data.cma.cn/data/detail/dataCode/SURF_CLI_CHN_PRE_DAY_GRID_0.5.html). These data were initially used to
 98 calculate area-averaged precipitation over the three selected regions.

99 The following atmospheric circulation information was obtained from the European Centre for Medium-Range Weather
 100 Forecasts (ECMWF) ERA5 reanalysis dataset (<https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>) at a $2^\circ \times$
 101 2° resolution for the area 30–180° E, 10° S–90° N: geopotential heights at 200, 500, and 850 hPa (GH200, GH500, and
 102 GH850); 2-m temperature (T2M); and sea-level pressure (SLP).

103 **2.2.2 GCM hindcast data**

104 We focused on seasonal drought prediction for the next 90 days, corresponding to the ‘seasonal scale’ discussed throughout
 105 the paper. For this, data for the atmospheric variables (in 90-day blocks) were downloaded from the ECMWF SEAS5 hindcast

106 dataset for the period 1993–2016 (<https://cds.climate.copernicus.eu/cdsapp#!/dataset/seasonal-original-single-levels>),
 107 including GH200, GH500, GH850, T2M, and SLP. For comparison, we selected hindcast data from the ECMWF, UK Met
 108 Office (UKMO), and Météo France (<https://cds.climate.copernicus.eu/cdsapp#!/dataset/seasonal-original-single-levels>)
 109 hindcast precipitation (Table 1).

110 **Table 1 Data used in this study**

Data type	Variable name	Pressure level (hPa)	Abbreviation	Data source
Atmospheric variables	2-m temperature	surface	T2M	Reanalysis: ERA5 reanalysis
	Sea-level pressure	surface	SLP	
	Geopotential height	200/500/850	GH200/GH500/GH850	Hindcast: ECMWF SEAS5
Precipitation	Total precipitation	surface	TP	Observation: CMA Hindcast: ECMWF/UKMO/Météo France

111 3 Methodology

112 3.1 Daily-updated Standardized Precipitation Index

113 As a seasonal drought index, the SPI3 was calculated for the three study regions using area-averaged precipitation data.
 114 Traditionally, the SPI3 varies on a monthly scale, reflecting cumulative precipitation over the past three months; however, a
 115 monthly timescale does not well reflect the evolution of drought onset, persistence, and relief, presenting limitations for
 116 meeting the requirements for real-time monitoring and prediction for effective management. Therefore, we followed the World
 117 Meteorological Organization (WMO) recommendation to adjust the sliding time window of the SPI3 to 90 days (Svoboda et
 118 al. 2012). Thus, we obtained information on the daily variation of meteorological droughts while the mathematical treatment
 119 remained unchanged. Full calculation details are given by Liu et al. (2018).

120 3.2 Predictor construction

121 Hart and Grumm (2001) first used standardized anomalies to delineate weather-scale events, and this approach has been used
 122 in an increasing number of studies to identify atmospheric and oceanic anomalies (Duan et al. 2014; Liu et al. 2018; Liu et al.
 123 2017). Here, we calculated 90-day standardized anomalies (SA90) using the following equation:

$$124 \text{SA90} = \frac{X_{90} - \mu_{90}}{\sigma_{90}}, \quad (1)$$

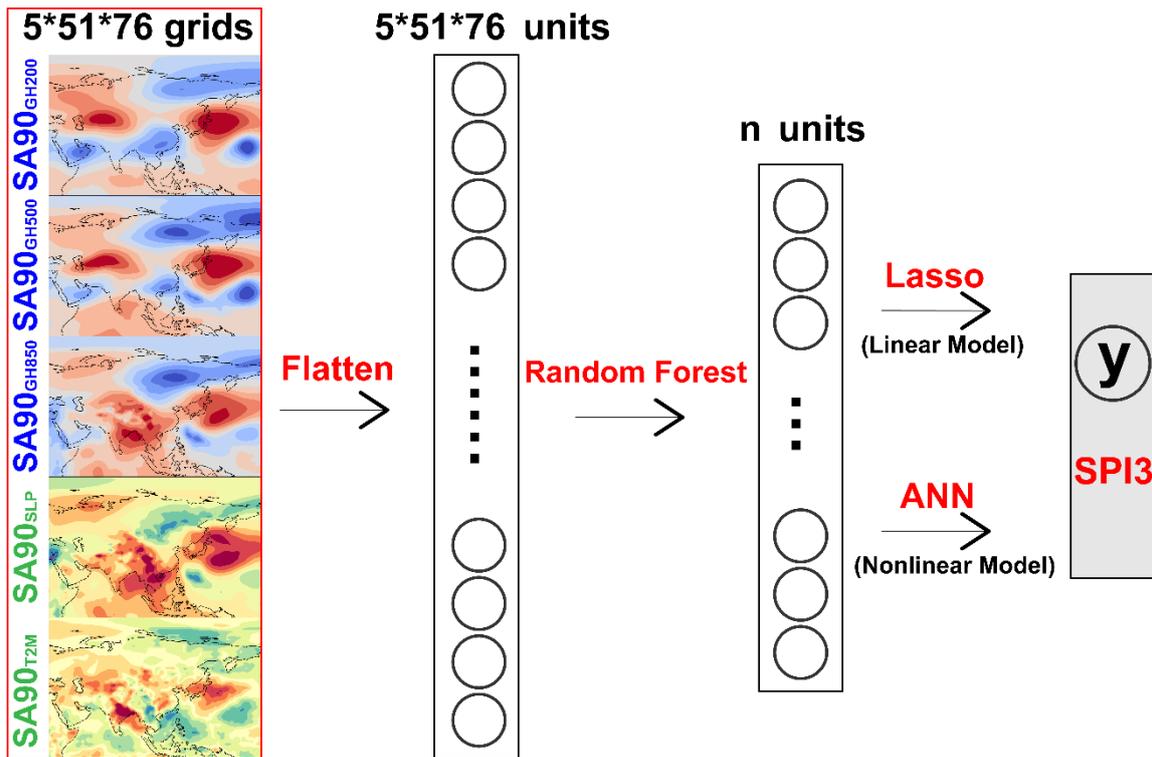
125 where X_{90} represents the mean values of atmospheric variables over the past 90 days; and μ_{90} and σ_{90} are the mean and
 126 standard deviation of the 90-day period, respectively. The climatic background period selected for this study was 1981–2010.
 127 For example, X_{90} for 1 April 2000 represented the 90-day average for the period from 3 January 2000 to 1 April 2000; and μ_{90}
 128 and σ_{90} are the 90-day average and standard deviation for the period from 3 January 1981 to 1 April 2010. The ERA5 reanalysis

129 dataset for the 40-year period between 1980 and 2019 was processed into SA90, from which we obtained the SA90 for all of
130 the considered atmospheric variables (i.e. $SA90_{GH200}$, $SA90_{GH500}$, $SA90_{GH850}$, $SA90_{T2M}$, and $SA90_{SLP}$).

131 3.3 Machine learning models

132 3.3.1 Construction of ML models

133 We constructed ML models for each of the three drought regions using 32 years of data from 1980 to 2011 for training and 8
134 years of data from 2012 to 2019 for validation. The models adopted the daily five-layer SA90 dataset as the predictor and the
135 SPI3 as the predictand. The model structure is shown in Fig. 2. First, we flattened the five-layer SA90 into a column of $5*51*61$
136 data points (the input neurons), extracted typical data through the RF, and then produced the output SPI3 data by forcing the
137 Lasso and ANN. The ML models (RF, Lasso, and ANN) are described in detail in Supplement 1–3.

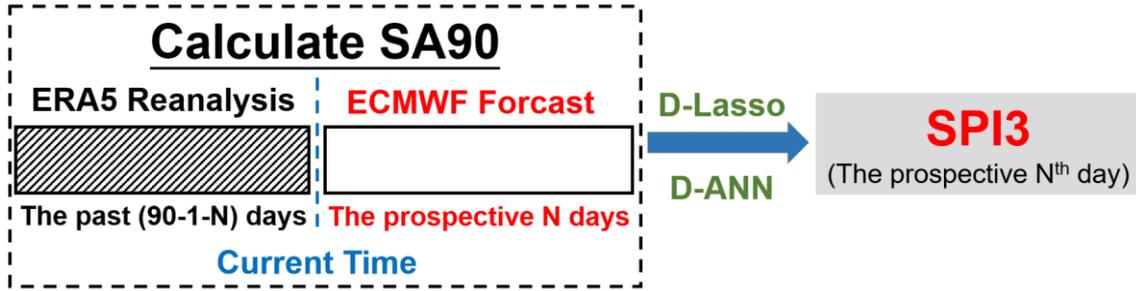


138
139 Fig. 2 ML model structure, where ANN contains a single layer and 16 nodes. $SA90_{GH200}$, $SA90_{GH500}$, $SA90_{GH850}$, $SA90_{T2M}$, and $SA90_{SLP}$
140 are the 90-day standardized anomalies of GH200, GH500, GH850, SLP and T2M, respectively. Lasso is a linear model, while ANN
141 is linear

142 3.3.2 Calculation of the prospective Nth-day SPI3

143 The ML models constructed in Sect. 3.3.1 incorporate the contemporaneous statistical relationships between the predictors and
144 the SPI3 and, therefore cannot predict the future. During the training and validation periods, the models were forced with
145 ERA5 data. For prediction, we first calculated SA90 using ERA5 reanalysis data for the past (90-1-N) days, ECMWF outputs

146 for the prospective N days, and then forced the ML models to predict the Nth-day SPI3, as shown in Fig. 3. Therefore, the
 147 drought prediction models we construct were hybrid (dynamical–statistical) models, which we name dynamic-Lasso (‘D-
 148 Lasso’) and dynamic-ANN (‘D-ANN’), respectively.



149
 150 Fig. 3 Schematic representation of the calculation process for the prospective Nth-day SPI3

151 3.3.3 GCM hindcast precipitation

152 The SPI3 calculated from the ECMWF, UKMO, and Météo France hindcast precipitation datasets were compared with the
 153 hybrid models. For prediction, we use observed precipitation for the past (90-1-N) days, GCM outputs for the prospective N
 154 days, and then calculated the Nth-day SPI3.

155 3.3.4 Evaluation metrics

156 Root-mean-square error (RMSE) was adopted as the loss function to describe the error between the model output and the
 157 observed data, and for the inverse calculation of the network residuals, as follows:

$$158 \text{ RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - x_i)^2}{n}}, \quad (2)$$

159 where x_i represents the observed value, y_i represents the predictand, and n represents the sample size.

160 The correlation coefficient (Corr) was used as the performance evaluation function for ML to measure the correlation between
 161 the model output and the expected values:

$$162 \text{ Corr} = 1 - \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (3)$$

163 where x_i is the observed value, y_i is the predictand, $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ (the sample mean), and analogously for \bar{y} ; and n represents
 164 the sample size.

165 **4 Results**

166 **4.1 Identification of droughts**

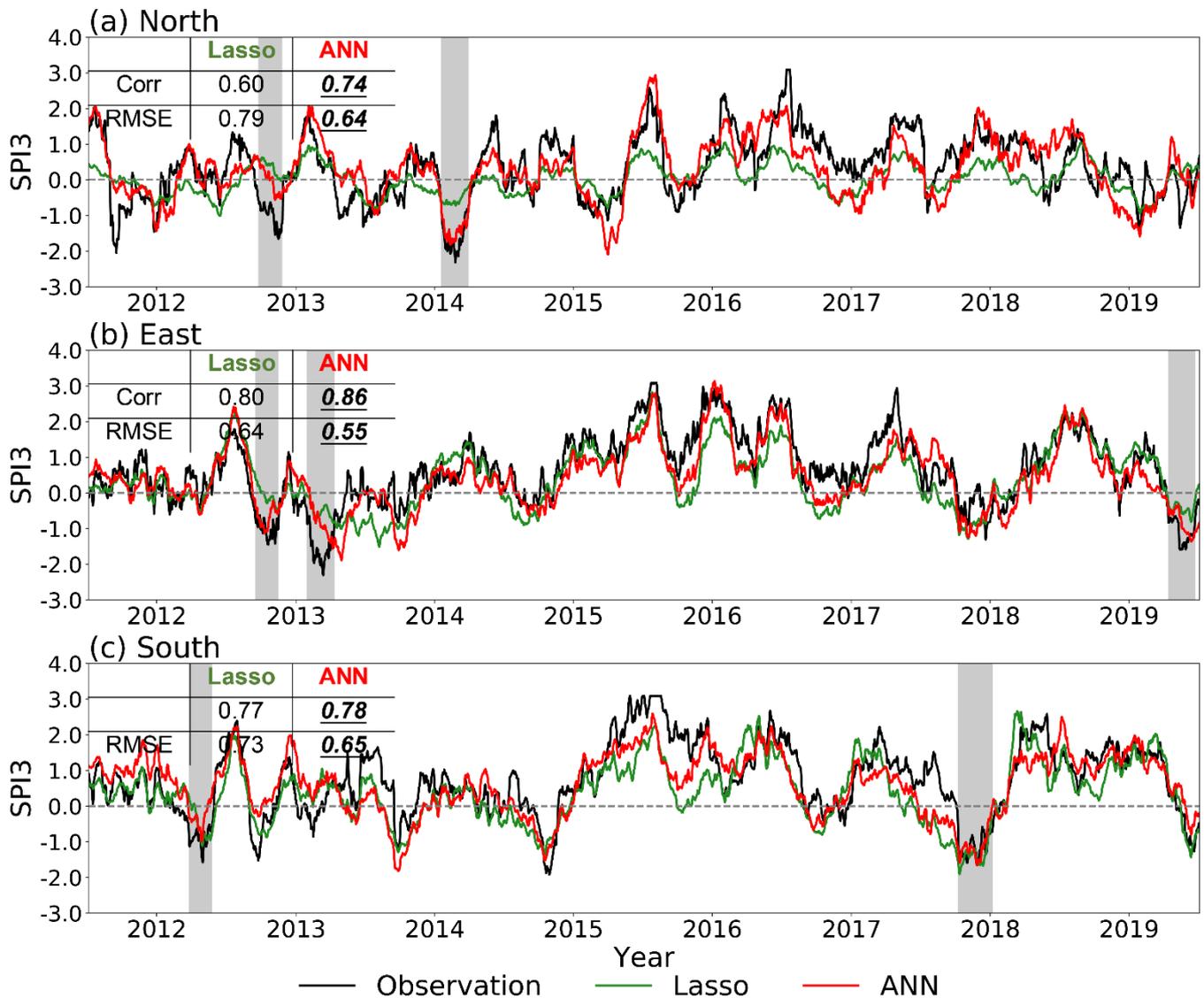
167 A drought event is defined as a number of consecutive days (> 60 days) with daily updated SPI3 values of < -0.5. Based on
 168 this approach, the droughts identified between 1979 and 2019 in the study regions are shown in Table 2.

169 **Table 2 Drought events identified in the three study regions between 1979 and 2019**

		Regions		
		North	East	South
Dates (Start– End)			1981/06/18–1981/10/06	1984/01/11–1984/04/17
			1984/01/12–1984/04/07	1985/05/20–1985/09/05
			1985/05/31–1985/09/23	1986/09/10–1986/11/17
		1981/05/03–1981/08/16	1986/02/20–1986/06/22	1988/06/07–1988/08/31
		1984/01/16–1984/05/23	1986/08/18–1986/11/14	1989/07/06–1990/01/10
		1986/03/23–1986/08/09	1988/06/23–1988/09/04	1991/04/03–1991/08/15
		1989/09/03–1989/11/11	1991/10/04–1992/01/04	1991/10/18–1991/12/26
		1991/09/09–1992/01/25	1992/07/26–1993/01/13	1992/09/12–1993/01/13
		1992/06/10–1992/09/22	1995/12/28–1996/03/25	1996/01/04–1996/03/27
		1995/03/07–1995/06/19	1998/10/26–1999/04/23	1996/11/14–1997/01/23
		1995/12/31–1996/05/02	2000/04/07–2000/06/22	1998/09/21–1998/12/01
		1997/06/11–1997/11/14	2001/04/05–2001/06/19	1999/01/25–1999/05/25
		1998/11/02–1999/04/10	2001/06/21–2001/08/30	2002/03/09–2002/06/17
		1999/07/14–1999/10/31	2001/09/01–2001/11/04	2003/03/25–2003/06/13
		2000/03/23–2000/06/26	2003/09/25–2003/12/10	2003/07/03–2004/02/04
		2001/04/12–2001/08/18	2004/03/07–2004/05/14	2004/05/22–2004/08/03
		2002/08/04–2002/12/05	2004/10/11–2005/01/08	2004/10/09–2005/02/05
		2008/12/23–2009/02/24	2007/05/06–2007/07/11	2005/11/12–2006/02/26
		2010/12/16–2011/02/27	2007/11/21–2008/01/27	2007/11/19–2008/01/30
		2011/05/26–2011/08/25	2009/09/27–2009/12/02	2009/01/31–2009/04/19
	2013/03/22–2013/05/25	2011/03/07–2011/08/04	2009/10/01–2009/12/01	
	2014/07/15–2014/09/27	2013/03/14–2013/05/15	2011/03/11–2011/10/24	
		2013/07/28–2013/10/09	2012/09/20–2012/11/21	
		2019/10/10–2019/12/21	2018/04/05–2018/07/06	

170 **4.2 Simulated SPI3 in validation periods**

171 According to Sect. 3, we constructed ML models based on data for the period 1980–2011. To test the robustness of the models,
 172 we forced the Lasso and ANN models for the period between 2012 and 2019 (Fig. 4). Both of these models' simulations were
 173 consistent with the observed SPI3 trends in all three regions, although ANN performed better overall, especially when
 174 simulating extreme values. This indicates that the ANN has a better information-extraction capability and better captures
 175 nonlinear atmospheric processes. In addition, the models performed much better for the east and south regions than the north
 176 study region, which may reflect the differing drought mechanisms in these regions.



177

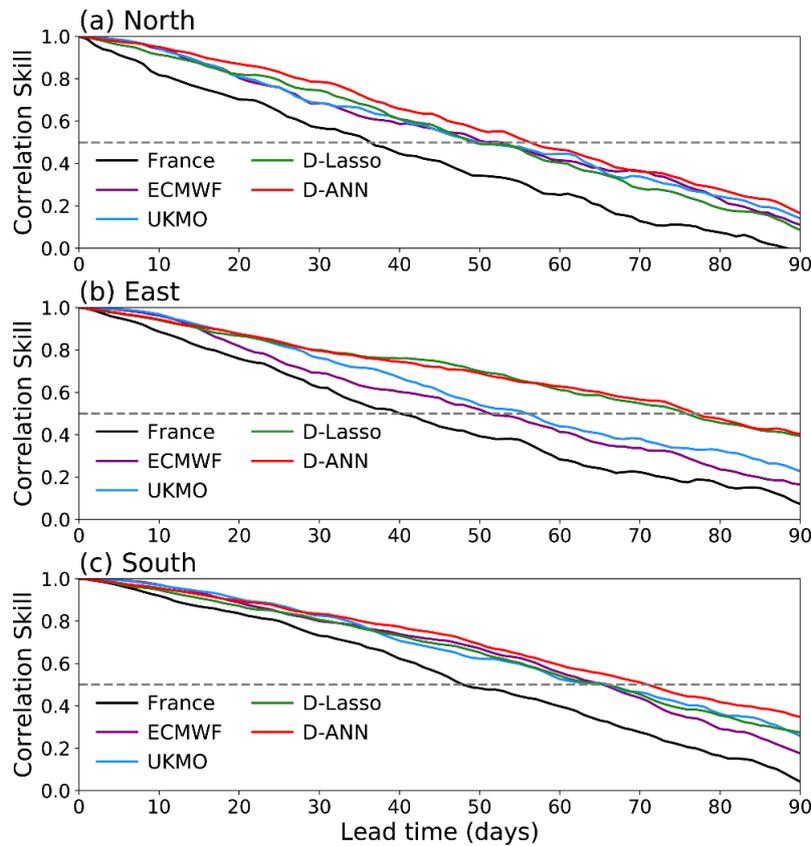
178 Fig. 4 Temporal trends in the observed and simulated SPI3 during the validation period (1 January 2012 to 31 December 2019). The
 179 black curve is the observed SPI3, and green and red curves are the Lasso and ANN simulations, respectively. The grey shading
 180 shows the identified drought events in Sect. 4.1

181 4.3 Seasonal drought prediction

182 Between 1993 and 2016, ECMWF hindcasts data for the first day of each month, as do D-Lasso and D-ANN. Therefore, we
 183 generated five sets of seasonal drought prediction products (SPI3) based on the Météo France, ECMWF, UKMO, D-Lasso,
 184 and D-ANN datasets.

185 Fig. 5 shows the correlation results of the SPI3 between 1993 and 2016. The strength of the correlations between all five
 186 models decreases as the lead time increases. At three weeks or less, the correlation between the ECMWF and UKMO outputs

187 is stronger than that of the hybrid models, indicating that dynamic models are more accurate in simulating precipitation with
 188 short lead times, which is consistent with other analyses (Lang et al. 2014; Li et al. 2017). However, when the lead time exceeds
 189 20 days, the hybrid models always outperformed the dynamic models; in particular, D-ANN showed the best performance for
 190 lead times above 30 days, which indicates that nonlinear hybrid models further improved drought predictions.
 191 Data correlations also differed between the study regions. For example, outputs from both the dynamical and hybrid models
 192 in the north region had lower correlation skills than the other two models. Among the three dynamical models, the UKMO
 193 model was best for seasonal drought prediction. Following Buizza and Leutbecher (2015), we determined the forecast lead
 194 times at which the correlation skills of the models were no longer higher than 0.5, defined as the forecast skill horizon. Based
 195 on this, D-ANN extended the forecast skill horizon by 6, 21, and 4 days in the north, east, and south regions, respectively,
 196 relative to the UKMO model. In particular, the correlation of D-ANN was higher than 0.5 with a 76-day lead time in the east
 197 region, while the correlations of the Météo France, ECMWF, and UKMO data were 0.18, 0.29, and 0.34, respectively.

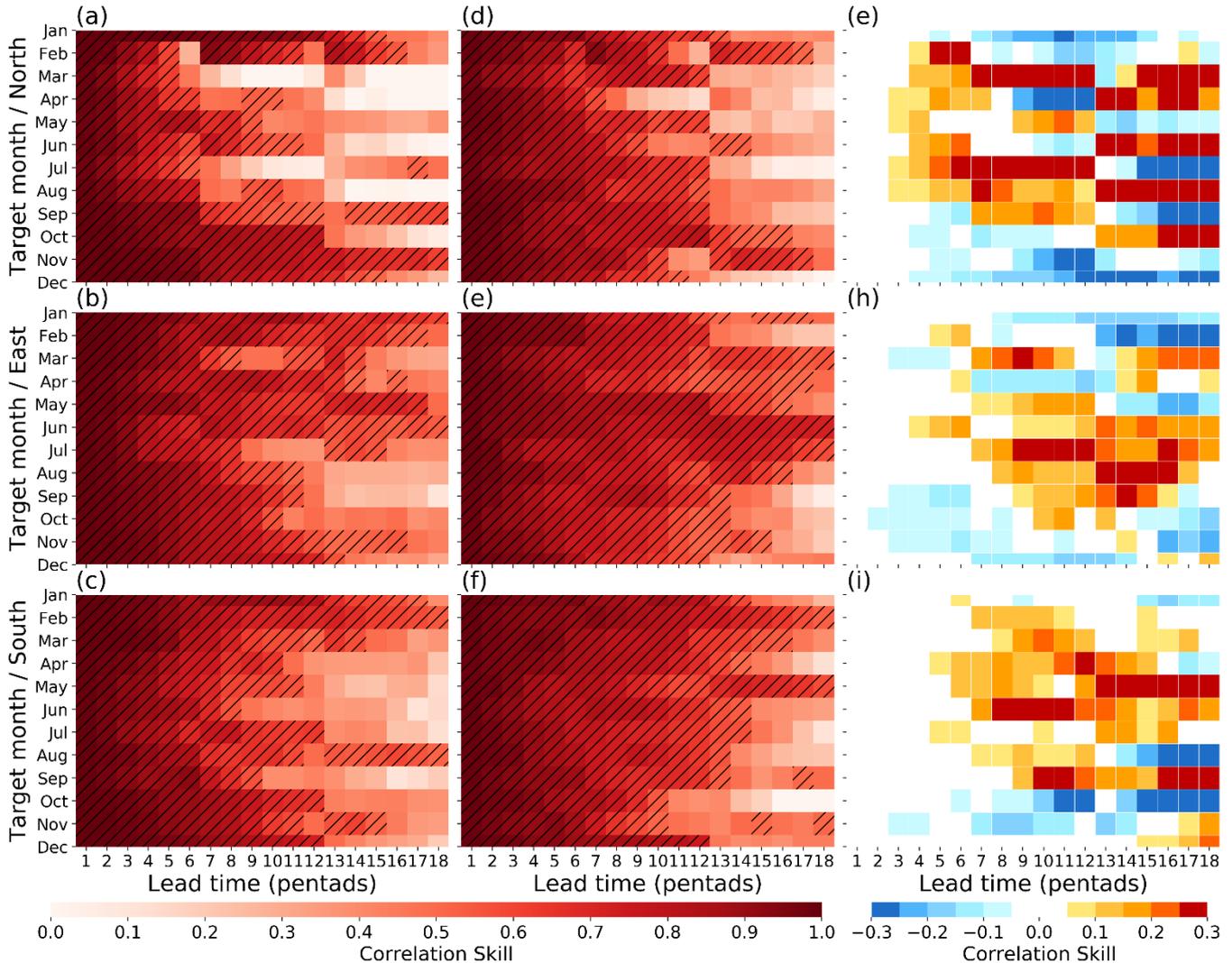


198

199 **Fig. 5** Correlation skills between five models predicting SPI3 at different lead times. Black, purple, blue, green, and red curves
 200 represent Météo France, ECMWF, UKMO, D-Lasso, and D-ANN outputs, respectively. The grey dashed lines indicate a
 201 correlation of 0.5

202 **4.4 Prediction skills across all months**

203 Figure 6 shows the SPI3 correlation skills for the UKMO and D-ANN models for all target months. Both underperform
 204 between September and December. To show the differences between the two models, the skills based on D-ANN were
 205 subtracted from the UKMO outputs (Fig. 6(e-i)). Compared to the UKMO outputs, D-ANN performed poorly from October
 206 to January but was more skillful across the spring and summer months. For lead times exceeding 30 days, D-ANN has a clear
 207 advantage. However, the correlation skills varied considerably among the regions, with higher skills in the east and south
 208 regions compared to the north.

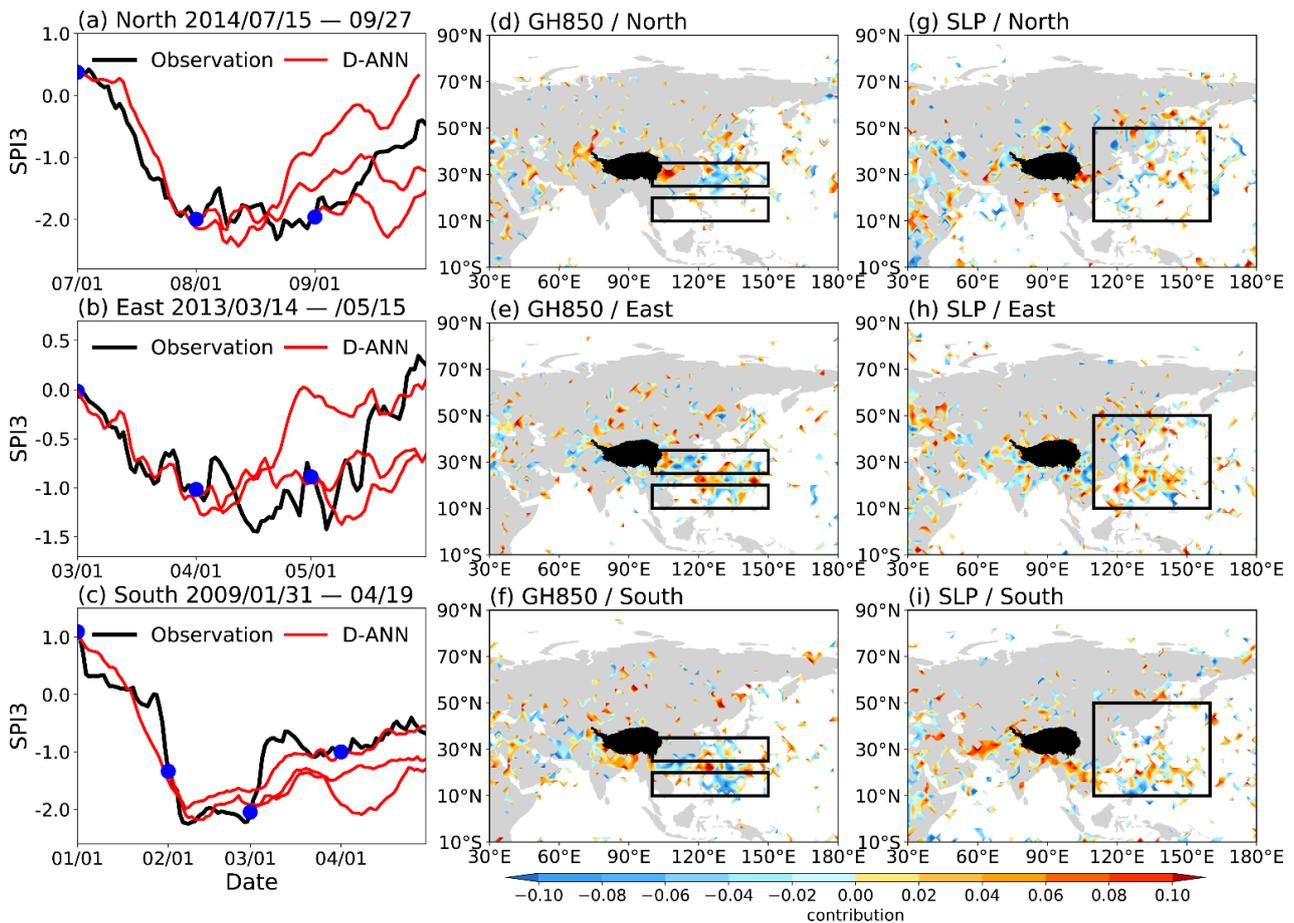


209
 210 **Fig. 6 Comparison of the correlation skills of UKMO and D-ANN models in different target months in the study regions: a–c are**
 211 **Corr(UKMO), d–f are Corr(D-ANN), and e–i are Corr(D-ANN) – Corr(UKMO). Shaded areas represent Corr > 0.5. Horizontal**
 212 **coordinates indicate different lead times (pentads); the first pentad includes the 1–5-day predictions, the second includes 6–10-day**
 213 **predictions, and so on. Vertical coordinates are target months**

214 **4.5 Interpretation of prediction mechanisms**

215 Fig. 7(a–c) shows that D-ANN can effectively predict seasonal droughts across the different regions. The color maps (Fig.
 216 7(d–e)) quantify the contribution of each grid predictor to the prediction, where positive and negative indicate that the predictor
 217 contributes to positive and negative to SPI3, respectively.

218 The representative regions identified by the RF for GH850 and SLP are mainly located in the east and south of the Tibetan
 219 Plateau and the Northwest Pacific. In addition, the representative regions in the north are sparser than in the east and south,
 220 which means the RF could not effectively extract key information from the atmospheric variables. Importantly, the East Asian
 221 Summer/Winter Monsoons are fundamental factors affecting droughts and floods in China (Ding et al. 2008; Wen et al. 2000;
 222 Yihui and Chan 2005; Zhou and Wu 2010). Interestingly, the representative regions were almost exclusively located in those
 223 areas where the CDC calculates the East Asian Summer/Winter Monsoons Indices (<http://cmdp.ncc-cma.net/Monitoring/monsoon.htm>). Further details of these indices are provided by Shi and Zhu (1996) and Zhang et al. (2003).
 224



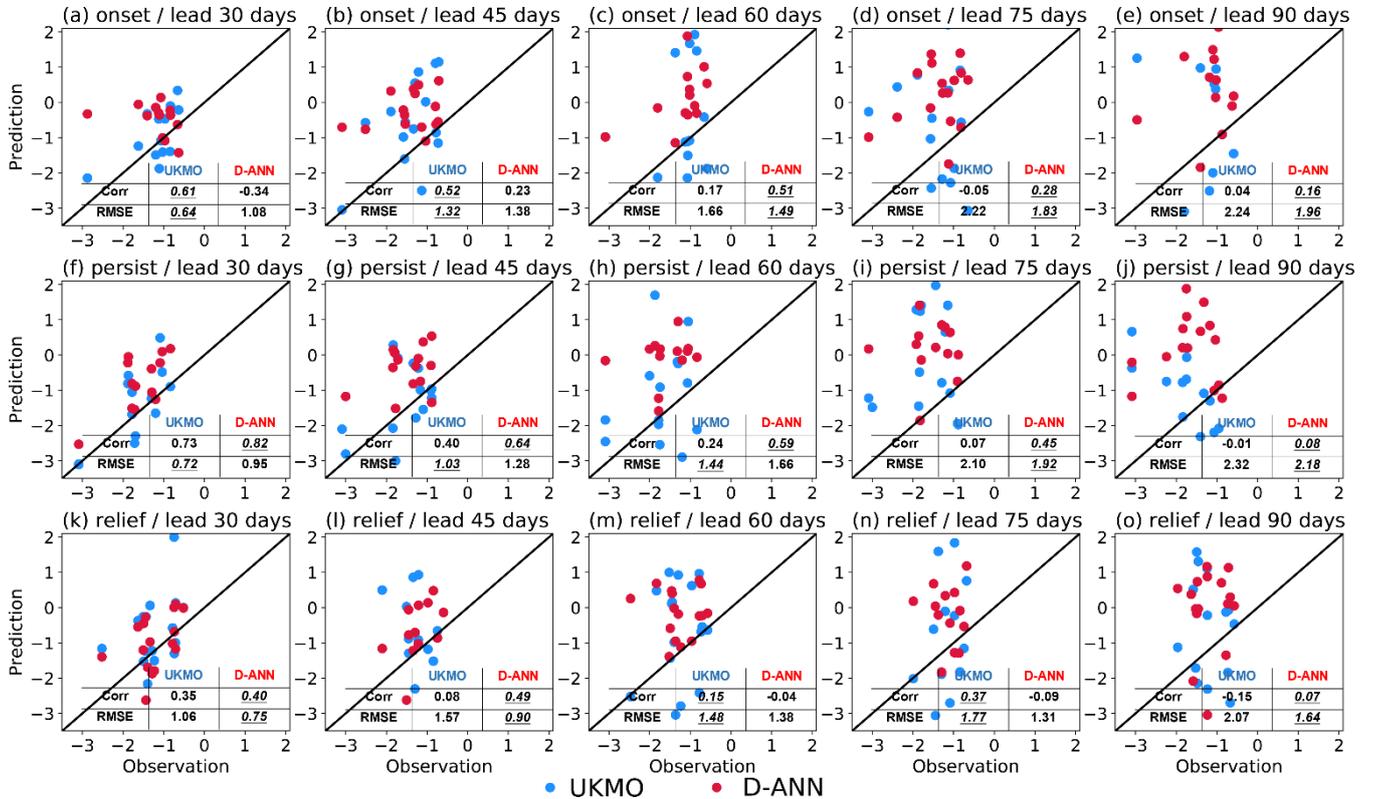
225
 226 **Fig. 7 Interpretations of the physical mechanisms of the D-ANN model: a–c, three droughts in the north, east, and south regions**
 227 **(black and red curves represent observations and D-ANN predictions, respectively, and blue dots are initial prediction moments);**
 228 **d–i, colors represent the contribution of the D-ANN model extracted from each grid, with positive values in red and negative**
 229 **values in blue; d–f, the two black boxes (10°–20° N, 100°–150° E and 25°–35° N, 100°–150° E) show the calculation areas of East**

230 Asian Summer Monsoon Index; g-i, the black box (10°–50° N, 110°–160° E) shows the calculation area of East Asian Winter
 231 Monsoon Index

232 **4.6 Predicting drought process**

233 We divided individual drought periods into three phases of onset, persistence, and relief. For example, for the drought from
 234 15 July 2014 to 27 September 2014, the onset, persistence, and relief periods correspond to 15 July–09 August, 10 August–3
 235 September, and 4 September–27 September, respectively. Fig. 8 shows the analysis of all droughts in the east study region
 236 from 1993 to 2016, and results for the north and south regions are included in Supplement 4.

237 The UKMO and D-ANN models showed different prediction abilities for all drought phases. Importantly, effective prediction
 238 of the onset phases of droughts is crucial for drought resistance. The UKMO model performed better for lead times < 45 days,
 239 while D-ANN performed better with lead times > 45 days. For the drought persistence phase, D-ANN always produced better
 240 predictions, with correlation values of 0.59 with a 60-day lead time, indicating that this model is better at predicting drought
 241 intensity. However, D-ANN predictions during the relief phases of droughts were poorer with lead times > 45 days, which
 242 likely reflects the requirement for large precipitation inputs during this phase.



243
 244 **Fig. 8 Predictions of drought phases with different lead times (days) in the east study region of China: UKMO and D-ANN outputs**
 245 **are shown in blue and red, respectively; a–e, drought onset phase; f–j, drought persistence phase; k–o, drought relief phase. Corr**
 246 **and RMSE represent correlation coefficient and root-mean-square error, respectively**

247 **5 Discussion**

248 **5.1 Network structure**

249 All of our results are based on the single-layer, 16-node ANN model described in Sect. 3. In this discussion, we focus on two
 250 important parameters in the ANN model, namely the number of layers and nodes.

251 **5.1.1 Hidden layer nodes: more or less?**

252 If hidden layer nodes are too few, it is difficult for a network to learn and acquire information-processing capabilities.
 253 Conversely, too many nodes in a hidden layer result in over-fitting (Zou et al. 2009). We trained the model by modifying the
 254 hidden nodes in the network while maintaining the number of layers at 1, the results of which are shown in Table 3. We found
 255 that when fewer than 16 nodes are used, model performance was compromised, and when more than 16 nodes are included,
 256 the model provides greater complexity but does not out-perform the validation dataset.

257 **Table 3 Performance of the ANN model with different numbers of nodes**

Regions	Metrics		Nodes				
			4	8	16	32	64
North	Corr	Training	0.84	0.95	0.97	0.95	0.98
		Validation	0.69	0.70	0.74	0.70	0.70
	RMSE	Training	0.58	0.33	0.25	0.31	0.22
		Validation	0.67	0.65	0.64	0.66	0.66
East	Corr	Training	0.83	0.95	0.96	0.97	0.93
		Validation	0.82	0.83	0.86	0.85	0.82
	RMSE	Training	0.60	0.34	0.31	0.24	0.37
		Validation	0.61	0.61	0.55	0.58	0.58
South	Corr	Training	0.94	0.94	0.96	0.95	0.93
		Validation	0.76	0.77	0.78	0.78	0.77
	RMSE	Training	0.36	0.35	0.28	0.32	0.39
		Validation	0.68	0.69	0.65	0.66	0.64

258 **5.1.2 Hidden layers: deep or shallow?**

259 The neural network learns from the data through the nodes in each layer, and the stacking of multiple layers allows the structure
 260 of complex applications to be learned (Pan 2019). However, it remains questionable a deeper network (i.e. with more layers)
 261 produced a stronger model (Tan 2020). To test this, keeping the number of nodes at 16, we varied the number of layers, as
 262 shown in Table 4. Based on this, we found that network performance did not improve as the number of layers was increased.
 263 From the analyses in Sec. 5.1.1 and Sec. 5.1.2, we found that the relationship between contemporaneous atmospheric
 264 circulation parameters and regional meteorological drought may not be overly complex, meaning that a single-layer, 16-node
 265 ANN model was able to effectively represent nonlinear relationships.

266 **Table 4 Performance of ANN models with different numbers of layers**

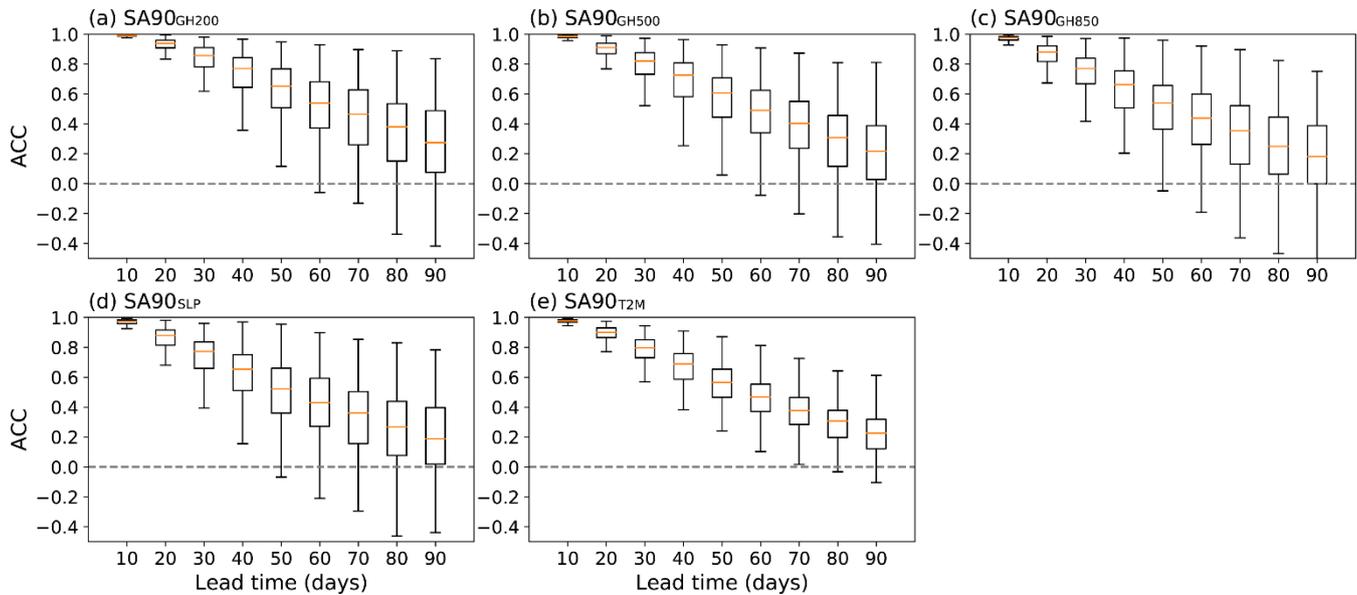
Regions	Metrics	Layers		
		1	2	3

North	Corr	Training	0.97	0.97	0.98
		Validation	0.74	0.70	0.73
	RMSE	Training	0.25	0.30	0.19
		Validation	0.64	0.65	0.63
East	Corr	Training	0.96	0.95	0.98
		Validation	0.86	0.77	0.82
	RMSE	Training	0.31	0.39	0.31
		Validation	0.55	0.67	0.69
South	Corr	Training	0.96	0.81	0.98
		Validation	0.78	0.65	0.75
	RMSE	Training	0.28	0.62	0.24
		Validation	0.65	0.78	0.72

267 5.2 Impact of forecast error from the ECMWF dataset

268 5.2.1 Anomaly correlation coefficients (ACC) of forecasted atmospheric variables

269 We used the forecasted atmospheric circulation dataset from the ECMWF as a predictor to force the hybrid models. Therefore,
 270 the models contained two kinds of error, first from the ML models and, second, from the GCM. This section focuses on the
 271 impact of the second error source. Fig. 9 shows that the ACC of all variables decreased with increasing lead time, which
 272 indicates that the error from the ECMWF dataset introduced a lot of uncertainty to the prediction. Indeed, forecasting skills
 273 are almost entirely lost for lead times above 50 days.



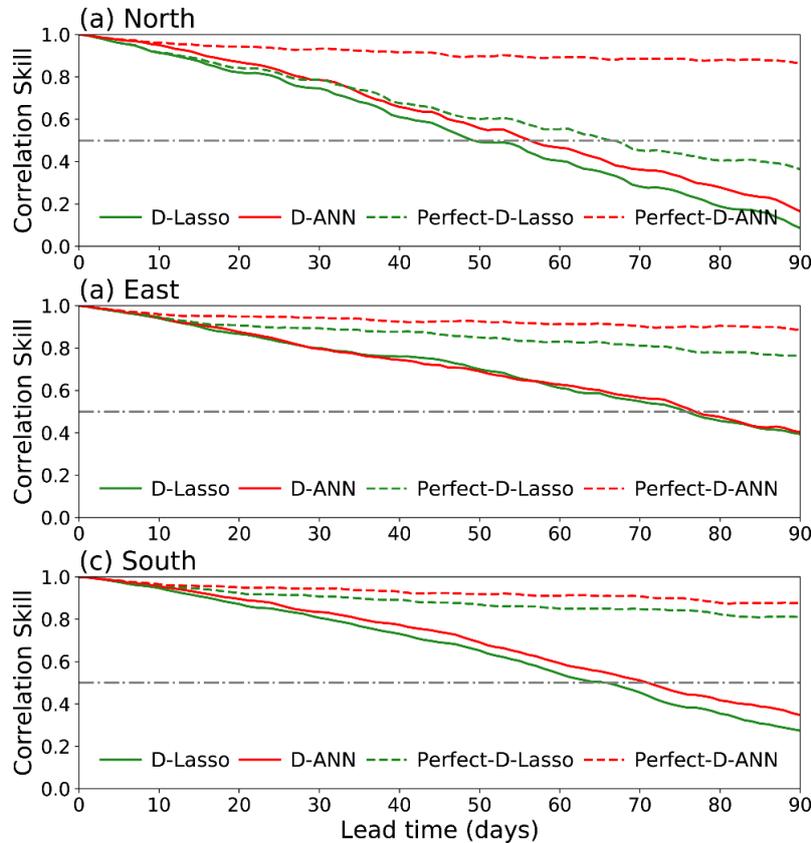
274

275 **Fig. 9** Boxplots of ACC for all atmospheric variables (30° – 180° E, 10° S– 90° N) with increasing lead times. The top to bottom five
 276 lines of a box plot represent the minimum, first quartile, median, third quartile, and maximum values of data, respectively

277 5.2.2 Perfect drought prediction

278 We replaced all of the atmospheric data for the prospective 90 days with reanalysis data to re-force the hybrid models, which
 279 we named ‘Perfect-D-Lasso’ and ‘Perfect-D-ANN’ (Fig. 10). The correlation skills of the Perfect hybrid models were almost

280 always above 0.8. Moreover, the performance of the models differed between the study regions, with the Perfect-D-ANN
 281 model performing best in the north region. Methods for post-processing GCM precipitation outputs have also been included
 282 in the dynamical-statistical prediction methods (Schepen et al. 2016), and so it is necessary to correct the forecasted
 283 atmospheric variables accordingly.



284
 285 **Fig. 10** Correlation skills of four models predicting the SPI3 with different lead times in the different study regions. The solid
 286 green curves represent D-Lasso model predictions; the solid red curves represent D-ANN model predictions; the dashed green
 287 curves represent Perfect-D-Lasso model predictions; and the dashed red curves represent the Perfect-D-ANN model predictions.
 288 The dashed grey lines indicate a correlation of 0.5

289 6 Conclusions

290 Seasonal drought prediction is important for drought resistance and water resources management. Here, we constructed
 291 seasonal drought prediction models with atmospheric variables (GH200, GH500, GH850, T2M, and SLP) as predictors and
 292 the SPI3 as the predictand. The resulting models were applied and evaluated for the prediction of meteorological droughts in
 293 north, east, and south regions of China. The main conclusions can be summarized as follows: (1) the daily-updated SPI3 was
 294 used to identify droughts for the last 40 years in the three drought regions; (2) the five-layer atmospheric variables were
 295 processed as model predictors; (3) the Lasso and ANN models were trained using the ERA5 reanalysis dataset for the period

296 1980–2011 and validated for the period 2012–2019, demonstrating good model robustness; (4) hybrid models (D-Lasso and
297 D-ANN) were constructed by combining the ECMWF and ML models; (5) month-by-month drought prediction was performed
298 for the period between 1993 and 2016. Compared to the dynamical models (Météo France, ECMWF, and UKMO), D-ANN
299 not only extended the forecast skill horizons in all regions but offers higher predictive abilities in spring and summer; (6) we
300 quantified the contribution of representative regions extracted using a RF approach, finding that these areas mainly correspond
301 to those areas where the CDC calculates the East Asian Summer/Winter Monsoon indices; and (7) the D-ANN model is more
302 skillful at predicting the onset and persistence phases of drought but performs comparatively poorly during the relief phase.
303 Important limitations of this study are that the selected predictors fail to incorporate the precursors of drought, such as sea
304 surface temperature, sea ice, and plateau snow cover (Li et al. 2018a; Schubert et al. 2016; Wang et al. 2017), and second, the
305 atmospheric variable outputs from the ECMWF model could not be adequately post-processed.

306 **Declarations**

307 **Competing interests** The authors declare no competing interests.

308

309 **Availability of data and material** All the datasets used in this paper are publicly accessible.

310

311 **Authors' contributions** ZW, with the help of HY and HH, designed the research. ZW and HH obtained the funding supporting
312 it. ZW conducted case analysis and wrote the original draft. HY, HH, and YL jointly reviewed the entire manuscript. All
313 authors finally read and approved the paper.

314

315 **Consent for publication** Not applicable

316 **Acknowledgments**

317 This work is supported by National Natural Science Foundation of China (grants No. 51779071); National Key Research and
318 Development Program of China (Grants No. 2017YFC1502403 and 2018YFC0407701).

319 **References**

- 320 Agana N, Homaifar A (2017) A Deep Learning Based Approach for Long-Term Drought Prediction. doi:10.1109/SECON.2017.7925314
321 Ayantobo OO, Li Y, Song S, Yao N (2017) Spatial comparability of drought characteristics and related return periods in mainland China
322 over 1961–2013 *Journal of Hydrology* 550:549-567 doi:<https://doi.org/10.1016/j.jhydrol.2017.05.019>
323 Buizza R, Leutbecher M (2015) The forecast skill horizon *Quarterly Journal of the Royal Meteorological Society* 141:3366-3382
324 Cai R, Tan H, Kontoyiannis H (2017) Robust surface warming in offshore China Seas and its relationship to the East Asian Monsoon wind
325 field and ocean forcing on interdecadal time scales *Journal of Climate* 30:8987-9005

- 326 Chen S, Zhang L, Liu X, Guo M, She D (2018) The Use of SPEI and TVDI to Assess Temporal-Spatial Variations in Drought Conditions
327 in the Middle and Lower Reaches of the Yangtze River Basin, China *Advances in Meteorology* 2018:9362041
328 doi:10.1155/2018/9362041
- 329 China Meteorological Administration (2011) *China Meteorological Disaster Yearbook (CMDY) China Meteorological Disaster Yearbook*
330 vol 978-7-5029-5446-8. Meteorological Press, Beijing, China
- 331 China Meteorological Administration (2014) *China Meteorological Disaster Yearbook (CMDY) China Meteorological Disaster Yearbook*
332 vol 978-7-5029-5556-4. Meteorological Press, Beijing, China
- 333 Conforti P, Ahmed S, Markova G (2018) Impact of disasters and crises on agriculture and food security, 2017
- 334 Costa-Cabral M, Rath JS, Mills WB, Roy SB, Bromirski PD, Milesi C (2016) Projecting and forecasting winter precipitation extremes and
335 meteorological drought in California using the North Pacific high sea level pressure anomaly *Journal of Climate* 29:5009-5026
- 336 Ding Y, Wang Z, Sun Y (2008) Inter-decadal variation of the summer precipitation in East China and its association with decreasing Asian
337 summer monsoon. Part I: Observed evidences *International Journal of Climatology: A Journal of the Royal Meteorological Society*
338 28:1139-1161
- 339 Duan W, He B, Takara K, Luo P, Nover D, Yamashiki Y, Huang W (2014) Anomalous atmospheric events leading to Kyushu's flash floods,
340 July 11–14, 2012 *Natural Hazards* 73:1255-1267
- 341 Hao Z, Hao F, Xia Y, Singh VP, Hong Y, Shen X, Ouyang W (2016) A statistical method for categorical drought prediction based on
342 NLDAS-2 *Journal of Applied Meteorology and Climatology* 55:1049-1061
- 343 Hao Z, Singh V, Xia Y (2018) Seasonal Drought Prediction: Advances, Challenges, and Future Prospects *Reviews of Geophysics* 56
344 doi:10.1002/2016rg000549
- 345 Hart RE, Grumm RH (2001) Using Normalized Climatological Anomalies to Rank Synoptic-Scale Events Objectively *Monthly Weather*
346 *Review* 129:2426-2442
- 347 Jin D, Guan Z, Tang W (2013) The Extreme Drought Event during Winter–Spring of 2011 in East China: Combined Influences of
348 Teleconnection in Midhigh Latitudes and Thermal Forcing in Maritime Continent Region *Journal of Climate* 26:8210-8222
349 doi:10.1175/JCLI-D-12-00652.1
- 350 Khan N, Sachindra DA, Shahid S, Ahmed K, Shiru MS, Nawaz N (2020) Prediction of droughts over Pakistan using machine learning
351 algorithms *Advances in Water Resources* 139:103562
- 352 Kirtman B, Anderson D, Brunet G, Kang I-S, Scaife AA, Smith D (2013) Prediction from weeks to decades. In: *Climate science for serving*
353 *society*. Springer, pp 205-235
- 354 Lang Y, Luo L, Ye A, Duan Q (2020) Do CFSv2 Seasonal Forecasts Help Improve the Forecast of Meteorological Drought over Mainland
355 China? *Water* 12:2010
- 356 Lang Y et al. (2014) Evaluating Skill of Seasonal Precipitation and Temperature Predictions of NCEP CFSv2 Forecasts over 17
357 Hydroclimatic Regions in China *Journal of Hydrometeorology* 15:1546-1559
- 358 Lavaysse C, Vogt J, Toreti A, Carrera ML, Pappenberger F (2018) On the use of weather regimes to forecast meteorological drought over
359 Europe *Nat Hazards Earth Syst Sci* 18:3297-3309 doi:10.5194/nhess-18-3297-2018
- 360 Lhotka O, Trnka M, Kyselý J, Markonis Y, Balek J, Možný M (2020) Atmospheric circulation as a factor contributing to increasing drought
361 severity in Central Europe *Journal of Geophysical Research: Atmospheres* 125:e2019JD032269
- 362 Li H, Chen H, Wang H, Sun J, Ma J (2018a) Can Barents Sea ice decline in spring enhance summer hot drought events over northern
363 China? *Journal of Climate* 31:4705-4725
- 364 Li Q, Yang S, Wu T, Liu X (2017) Subseasonal Dynamical Prediction of East Asian Cold Surges *Weather & Forecasting* 32
- 365 Li X, Li D, Li X, Chen L (2018b) Prolonged seasonal drought events over northern China and their possible causes *International Journal of*
366 *Climatology* 38:4802-4817
- 367 Li Y, Wu Z, He H, Wang QJ, Xu H, Lu G (2020) Post-processing sub-seasonal precipitation forecasts at various spatiotemporal scales across
368 China during boreal summer monsoon *Journal of Hydrology*:125742 doi:<https://doi.org/10.1016/j.jhydrol.2020.125742>
- 369 Liu Y, Huang R (2019) Linkages between the south and east asian monsoon water vapor transport during boreal summer *Journal of Climate*
370 32:4509-4524
- 371 Liu Z, Lu G, He H, He J (2018) A conceptual prediction model for seasonal drought processes using atmospheric and oceanic standardized
372 anomalies: application to regional drought processes in China *Hydrology and Earth System Sciences* 22:529-546
- 373 Liu Z, Lu G, He H, Wu Z, He J (2017) Anomalous features of water vapor transport during severe summer and early fall droughts in
374 Southwest China *Water* 9:244
- 375 Liu Z, Törnros T, Menzel L (2016) A probabilistic prediction network for hydrological drought identification and environmental flow
376 assessment *Water Resources Research* 52:6243-6262
- 377 Ma F, Yuan X, ye A (2015) Seasonal Drought Predictability and Forecast Skill over China *Journal of Geophysical Research: Atmospheres*
378 120:n/a-n/a doi:10.1002/2015JD023185
- 379 Ma S, Zhu C, Liu J (2020) Combined Impacts of Warm Central Equatorial Pacific Sea Surface Temperatures and Anthropogenic Warming
380 on the 2019 Severe Drought in East China *Advances in Atmospheric Sciences* 37:1149-1163 doi:10.1007/s00376-020-0077-8

381 Madadgar S, Aghakouchak A, Shukla S, Wood AW, Cheng L, Hsu K-L, Svoboda M (2016) A hybrid statistical-dynamical framework for
382 meteorological drought prediction: Application to the southwestern United States *Water Resources Research* 52:5095-5110

383 Merryfield WJ et al. (2020) Current and emerging developments in subseasonal to decadal prediction *Bulletin of the American*
384 *Meteorological Society* 101:E869-E896

385 Mo KC, Lyon B (2015) Global Meteorological Drought Prediction Using the North American Multi-Model Ensemble *Journal of*
386 *Hydrometeorology* 16:1409-1424 doi:10.1175/JHM-D-14-0192.1

387 Okumura YM, DiNezio P, Deser C (2017) Evolving impacts of multiyear La Niña events on atmospheric circulation and US drought
388 *Geophysical Research Letters* 44:11-614

389 Pan B (2019) Advancing Precipitation Prediction Using a Composite of Models and Data

390 Pan B, Anderson GJ, Goncalves A, Lucas DD, Bonfils CJW, Lee J (2020) Improving seasonal forecast using probabilistic deep learning
391 arXiv preprint arXiv:201014610

392 Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, Prabhat M (2019) Deep learning and process understanding for
393 data-driven Earth system science *Nature* 566:195 doi:10.1038/s41586-019-0912-1

394 Rezaeianzadeh M, Tabari H (2012) MLP-based drought forecasting in different climatic regions *Theoretical and Applied Climatology* 109
395 doi:10.1007/s00704-012-0592-3

396 Ribeiro AFS, Pires CAL (2016) Seasonal drought predictability in Portugal using statistical–dynamical techniques *Physics and Chemistry*
397 *of the Earth, Parts A/B/C* 94:155-166

398 Richardson D, Fowler H, Kilsby C, Neal R, Dankers R (2020) Improving sub-seasonal forecast skill of meteorological drought: a weather
399 pattern approach *Natural Hazards and Earth System Sciences* 20:107-124 doi:10.5194/nhess-20-107-2020

400 Schepen A, Wang QJ, Robertson DE (2016) Application to Post-processing of Meteorological Seasonal Forecasting. In: Duan Q,
401 Pappenberger F, Thielen J, Wood A, Cloke HL, Schaake JC (eds) *Handbook of Hydrometeorological Ensemble Forecasting*.
402 Springer Berlin Heidelberg, Berlin, Heidelberg, pp 1-29. doi:10.1007/978-3-642-40457-3_18-1

403 Schubert SD et al. (2016) Global meteorological drought: a synthesis of current understanding with a focus on SST drivers of precipitation
404 deficits *Journal of Climate* 29:3989-4019

405 Shi N, Zhu Q (1996) An abrupt change in the intensity of the East Asian summer monsoon index and its relationship with temperature and
406 precipitation over East China *International Journal of Climatology: A Journal of the Royal Meteorological Society* 16:757-764

407 Svoboda M, Hayes M, Wood D (2012) Standardized precipitation index user guide *World Meteorological Organization Geneva, Switzerland*

408 Tan T (2020) Are Deeper Networks Better? —A Case Study. <https://towardsdatascience.com/are-deeper-networks-better-a-case-study-6ee9bcb0725b>.

409

410 Ujeneza EL, Abiodun BJ (2015) Drought regimes in Southern Africa and how well GCMs simulate them *Climate Dynamics* 44:1595-1609

411 Wang C, Yang K, Li Y, Wu D, Bo Y (2017) Impacts of spatiotemporal anomalies of Tibetan Plateau snow cover on summer precipitation
412 in eastern China *Journal of Climate* 30:885-903

413 Wang Q et al. (2014) Temporal-spatial characteristics of severe drought events and their impact on agriculture on a global scale *Quaternary*
414 *International* 349:10-21

415 Wang QJ, Schepen A, Robertson DE (2012) Merging seasonal rainfall forecasts from multiple statistical models through Bayesian model
416 averaging *Journal of Climate* 25:5524-5537

417 Wen C, Graf HF, Ronghui H (2000) The interannual variability of East Asian Winter Monsoon and its relation to the summer monsoon
418 *Advances in Atmospheric Sciences* 17:48-60 doi:10.1007/s00376-000-0042-5

419 World Health Organization (2019) Drought. <https://www.who.int/health-topics/drought>.

420 Xu L, Chen N, Zhang X (2018a) A comparison of large-scale climate signals and the North American Multi-Model Ensemble (NMME) for
421 drought prediction in China *Journal of Hydrology* 557:378-390 doi:<https://doi.org/10.1016/j.jhydrol.2017.12.044>

422 Xu L, Chen N, Zhang X, Chen Z (2018b) An evaluation of statistical, NMME and hybrid models for drought prediction in China *Journal of*
423 *Hydrology* 566:235-249

424 Yan H, Moradkhani H, Zarekarizi M (2017) A probabilistic drought forecasting framework: A combined dynamical and statistical approach
425 *Journal of Hydrology* 548:291-304 doi:<https://doi.org/10.1016/j.jhydrol.2017.03.004>

426 Yihui D, Chan JCL (2005) The East Asian summer monsoon: an overview *Meteorology and Atmospheric Physics* 89:117-142
427 doi:10.1007/s00703-005-0125-z

428 Yoo-Geun H, Jeong-Hwan K, Jing-Jia L (2019) Deep learning for multi-year ENSO forecasts *Nature* 573

429 Zhang Q, Tao S, Chen L (2003) Interannual variation of East Asian Summer Monsoon index and East Asian atmospheric circulation *Acta*
430 *Meteorologica Sinica*:559-568

431 Zhang R (2015) Changes in East Asian summer monsoon and summer rainfall over eastern China during recent decades *Science Bulletin*
432 60:1222-1224 doi:10.1007/s11434-015-0824-x

433 Zhang R, Chen Z-Y, Xu L-J, Ou C-Q (2019) Meteorological drought forecasting based on a statistical model with machine learning
434 techniques in Shaanxi province, China *Science of The Total Environment* 665:338-346

435 Zhou LT, Wu R (2010) Respective impacts of the East Asian winter monsoon and ENSO on winter rainfall in China *Journal of Geophysical*
436 *Research: Atmospheres* 115

437 Zhu Q, Luo Y, Zhou D, Xu Y-P, Wang G, Tian Y (2020) Drought prediction using in situ and remote sensing products with SVM over the
438 Xiang River Basin, China Natural Hazards doi:10.1007/s11069-020-04394-x
439 Zhuang Y, Fu R, Wang H (2020) Large-Scale Atmospheric Circulation Patterns Associated with US Great Plains Warm Season Droughts
440 Revealed by Self-Organizing Maps Journal of Geophysical Research Atmospheres 125 doi:10.1029/2019JD031460
441 Zou W, Li Y, Tang A (2009) Effects of the number of hidden nodes used in a structured-based neural network on the reliability of image
442 classification Neural Computing and Applications 18:249-260 doi:10.1007/s00521-008-0177-3
443