1. Data and Methods
   1. Ethnographic Data

Our starting sample is the widely-used Standard Cross-Cultural Sample (SCCS) of 186, mostly preindustrial societies, constructed to maximize historical independence of cases (Murdock and White, 1969). These SCCS sample gives a rounded latitude and longitude for each case. If the ethnographer focused on a community, the latitude and longitude might be specific to that community. If the ethnographer was generalizing to the whole society (often the case with smaller societies) a centroid might be given. However, because we wanted to identify the closest weather station, we examined maps within the ethnography and consulted Google Earth to try to get a more precise location that would match the topography mentioned in the ethnography. This was especially important in mountainous or coastal regions where major differences in topography might be close-by. Weather stations that were not within XXX (I believe it was 222 km) were excluded.

For our present study we used only those societies that were more reliably coded on hazards (a 6 or less on the Ember and Ember (1992b) reliability scales). There were 98 societies from the 186 that met this criterion. When matching up weather stations to enthographic stations, there are not always stations available, and sometimes there are multiple stations, so the the numbers vary between 65 for temperature indices and 96 for wet precipitation indices and 93 for dry precipitation indices. There also additions cases (post-fixed with “s) where multiple weather stations match up against a particular SCCS site.

* 1. Weather Station Data

Daily precipitation, minimum temperature, and maximum temperature were obtained from weather stations in the Global Historical Climatology Network (GHCN) (https://www.ncdc.noaa.gov/ghcn). In most cases, the weather stations started to record these variables in the 1950s, while our ethnographic floods occurred prior to these instrumental measurements. Because there are significant gaps in the daily data, we developed a simple gap-filling procedure. For precipitation the nine GHCN stations nearest the SCCS latitude and longitude (see above) were chosen, while for minimum and maximum temperature, the four closest stations were chosen. The weather station closest to the SCCS site was the primary station to be gap-filled based on the other chosen stations. Only periods where at least 2 of the stations shared common temporal coverage were used, and those stations with greater than 50% missing values were ignored. A linear relationship was used for commonly shared periods to fill in the gaps. For gaps where none of the 4 or 9 stations had data, the gap was filled with the corresponding monthly/seasonal/annual data for the entire dataset of the primary station. Finally, all stations with less than 10 years of data were eliminated.

* 1. Gridded Data

The monthly temperature (T), precipitation (P) and Penman-Monteith potential evapotranspiration (PET\_pm) data were taken from the Climate Research Unit (CRU), which ranges from 1901 to 2014 with a spatial resolution of 0.5° × 0.5° (CRU3.23: http://catalogue.ceda.ac.uk/uuid/5dca9487dc614711a3a933e44a933ad3.) [Harris et al., 2014]. The ethnographic time periods are overlapping but are not equivalent to these years. These variables were used to calculate the P-E drought index, which describes the deficit of soil water. The Palmer Drought Severity Index (PDSI) [Palmer, 1965] is another drought index that contains information about antecedent and present soil moisture. We used the CRU high resolution (0.5° × 0.5°) monthly PDSI global dataset [Van der Schrier, 2013] from 1901 to 2009.

PDSI is derived from a two-layer soil water model accounting for antecedent moisture conditions [Palmer, 1965]. Larger negative values relate to more severe droughts, such that < -2.0 is moderate drought), < -3.0 is severe drought), and -4.0 is extreme drought [Van der Schrier, 2013; Table 3].

The P-E index was developed by Kuang-Yu Chang (personal communication) to describe droughts. It does not incorporate the antecedent moisture conditions like PDSI. The P-E index defines the soil water deficit as a function of P and PET:

(1)

in which every site has a maximum P-PET every month. The lower the index value, the drier the soil. We develop two versions of the P-E index, one derived from the Thornthwaite [Willmott et al., 1985] PET (PET\_th) and the other derived from the Pennmen Monteith PET (PET\_pm).

We ranked the P-E index values in descending order, and determined the values for specified percentile thresholds based on biome (i.e. 60%, 70%, 80%, 85%, 90%, 95%, and 98%; Table 1). Additionally, we had to downscale the monthly indices to annual indices. First, we focused on the months within the growing season which were included in the ethnographic data.

**Table 1.** Drought indices based on gridded CRU data, with their time and the thresholds for defining droughts.

|  |  |  |
| --- | --- | --- |
| **Drought Index** | **Time** | **Threshold** |
| PDSI | 1901-2009 | Numeric values |
| P-E\_th | 1901-2014 | Percentiles |
| P-E\_pm | 1901-2014 | Percentiles |

We also developed three drought indices (R05PS, R01PS, CDD) based on daily precipitation data from weather stations (Table 2), analogous to the extreme precipitation indices used for wet conditions below.

1. Precipitation indices

Thirteen extreme precipitation indices [Jiang et al, 2016; Table 2] were derived from the daily precipitation of GHCN weather stations (Table 2).

**Table 2.** Indices derived from daily precipitation from GHCN weather stations.

|  |  |  |
| --- | --- | --- |
| **Index** | **Definition** | **Wet/Dry** |
| R10S | Number of days within a season with precipitation > 10 mm/day | Wet |
| R20S | Number of days within a season with precipitation > 20 mm/day | Wet |
| R95PS | Sum of daily precipitation > 95th percentile for the season | Wet |
| R99PS | Sum of daily precipitation > 99th percentile for the season | Wet |
| RX1S | Seasonal maximum 1-day precipitation | Wet |
| RX5S | Seasonal maximum 5-day precipitation | Wet |
| SDIIS | Total precipitation divided by total number of days with precipitation > 1mm/day | Wet |
| CWD | number of continuous rainy days / total number of days | Wet |
| R05PS | Sum of daily precipitation < 5th percentile for the season | Dry |
| R01PS | Sum of daily precipitation < 1st percentile for the season | Dry |
| CDD | number of continuous no precipitation days / total number of days | Dry |
| \*PRCPTOTS | Total seasonal precipitation | Wet |
| \*CoefVar | Coefficient of Variation: standard deviation daily precipitation/mean daily precipitation [I thought that there were different coefficients as well--intra-annual and inter-annual) | Wet |

\* non-extreme index

For each of these precipitation indices, we derived the standard deviation (e.g. >1σ or 2σ) of each index value from the mean value of the index in each season (i.e. DJF, JJA, MAM, SON) for each year. The largest standard deviation within the growing season represents the standard deviation for the year.

The consecutive day indices (CWD, CDD) were based on growing seasons for each society, as well as the month prior and following the growing season. This was done for the four most abundant crops at each SCCS site. For hunter-gatherer or pastoral societies, we instead calculated these indices for each season (like the other indices). The Coefficient of Variation was calculated for both inter-annual and intra-annual variations.

1. Temperature indices

Besides drought and flooding events, crops are also subject to stress due to extreme cold spells. We therefore used the daily minimum and maximum temperatures to calculate two cold-spell indices (Table 3).

**Table 3.** Indices derived from daily minimum and maximum temperature from GHCN weather stations.

|  |  |  |
| --- | --- | --- |
| **Index** | **Definition** | **Cold/Hot** |
| FD | Number of Frost Days: Number of days when Tmin < 0oC | Cold |
| CSDI | Cold Spell Duration Index: number of consecutive days with Tmin < 10th percentile (based on entire time period at each site)/ total number of days | Cold |
| \*Tmin | Minimum daily temperature | Cold |
| \*Tmax | Maximum daily temperature | Hot |

\* non-extreme index

1. Predictability

Predictability, constancy, and contingency were calculated for all the indices based on the weather station data. Predictability is based upon the Colwell index (Colwell, 1974), and is variation among successive periods in the pattern of a periodic phenomenon, composed of contingency (seasonality) and constancy (inter-annual variability). Predictability is high if the tail end of precipitation has a high recurrent pattern in seasonality (i.e. large contingency) or inter-annual variability (large constancy). For most of the indices, the contingency was based on bins of meteorological seasons, whereas for consecutive day indices of agricultural societies, the contingency was based upon pre, post, and during the growing season. Also for consecutive day indices (CWD, CDD), although the range can be from 0 to 1, the maximum value varies from site-to-site, so the predictability bins were based on site-specific percentile ranges. An evaluation of the uncertainties associated with predictability is available in Jiang (2016).

References:

Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister (2014), Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset, *International Journal of Climatology,* 34: 623-642, doi: 10.1002/joc.3711.

Jiang, M., B.S. Felzer, and D. Sahagian, 2016. Predictability of extreme precipitation events in the conterminous U.S., 1949-2010. *Journal of Climate*. 29, 2621-2633.

Jiang, 2016. On the History and Evidence of the Colwell Index in Quantifying Environmental Predictability, and Its Applications in Characterizing Precipitation Predictability in the Conterminous United States. Available from ProQuest Dissertations & Theses Global. (1796990214). Retrieved from https://search.proquest.com/docview/1796990214?accountid=36155

Palmer, W. C., (1965), Meteorological drought. *Research Paper 45, U.S. Dept. of Commerce*, 58 pp.

Van der Schrier, G., J. Barichivich, K. R. Briffa, and P.D. Jones (2013), A scPDSI-based global data set of dry and wet spells for 1901-2009, *Journal of Geophysical Research: Atmospheres,* 118:4025-4048, doi:10.1002/jgrd.50355.