Methods –

Data collection – We used historical data collected by O.A. Stevens and others to create a dataset of first flowering days (FFD) for 25 flowering plant species. The observations were made at a tallgrass prairie site in Clay county Minnesota that has been a Nature Conservancy preserve since 1975. Individual data points represent the day of the year on which a given plant species was observed flowering at the site although all species were not observed in all years. The Stevens dataset represents continuous data from 1910 to 1961 and subsequent observations are from 2012 through 2020. Thus, there is a 52 year gap in data at the end of the past century. The plant species analyzed were limited to those that met a series of minimum data requirements. The 25 species chosen had a minimum of five years of data and at least one observation prior to 1962 and one after.

In order to quantify different environmental variables related to annual climate patterns, we used daily climate data collected in Fargo,North Dakota, USA, as part of the National Atmospheric and Oceanic Administration (NOAA) National Climatic Data Center (NCDC) observing network (<http://www.ncdc.noaa.gov/oa/ncdc.html>). The climate data collection site (46 ° 56 ’ N, 96 ° 49 ’ W) is located at the Fargo International Airport 32 km west of the flowering observation site. The dataset includes daily estimates of maximum and minimum temperature, snowpack (0=bareground) and snowfall beginning in 1942. However, snowpack data is unavailable for 1997 through 2004. As a result we were able to analyze data for 29 years (1942-1961 and 2012-2020).

We used the raw climate data to calculate four variables regarding seasonal patterns of temperature or winter snowfall. Accumulated Degree Growing Unit (AGDU) is an annual estimate of the accumulation of warm temperatures over the first three months of a year and is intended to quantify the relative warmth of spring for a given year. AGDU was calculated as the day of the year on which the sum of the growing units of the year exceeded 300. Growing units were defined as a daily measure of the difference between the average temperature and 35˚ F; units were set to zero if the average temperature was below 35. Thus, in years with warmer spring months the AGDU value will be lower.

Three different precipitation variables were calculated. The winter snowfall amount for a given year (TSNOW) was calculated as the sum of snowfall over the first 90 days. A second variable associated with snowfall was the Date of Bare Ground (DOBG) or the day of the year when snowpack first reached zero. A couple records indicated a short period, one to two days, of snowpack late in the season which were excluded for a more realistic representation of first bare ground. The third correlate of winter snowfall was Snowpack on Day X (SPDX), a variable designed to estimate the extent of snowpack just prior to the growing season. In order to calculate SPDX for each year we used model selection to identify which day in March represented the optimal day for best predicting the first flowering day (FFD) based on snowpack. The most predictive day was determined separately for each plant species. Once the most predictive day in March was determined for a given species by running individual linear regressions and choosing the model with the lowest AIC, snowpack on that day each year was assigned as the SPDX value. Larger SPDX values indicate greater snowpack on a selected day in March per species.

Statistical Analyses -

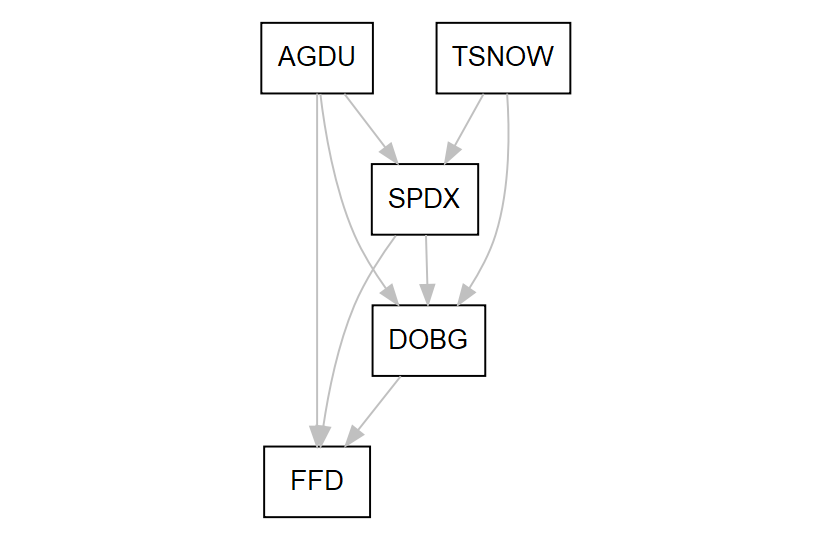
The relationships between the precipitation, temperature, and flowering date variables were examined using structural equation models for each species based on path analysis. Since many of the species had few first flowering date records, we applied maximum likelihood estimation to ensure the model was the most probable for the given data. Both direct and indirect effects were included in the model. The structural equations were built using the *lavaan* function in R.

Maybe - We built models excluding one variable at a time. Goodness of fit parameters were used to select the model that was most appropriate for the data.

1. Initial full model development

A. Assigned extrinsic and intrinsic variable as follows..

The exogenous variables were AGDU and TSNOW. The endogenous variables were DOBG, SPDX, and FFD. The model included direct and indirect effects between variables, variances within variables, and covariances between the exogenous variables.



B. Used lavaan to calculate x,y and z in each of the species

2. Model Selection

A. Goal: choose the best model for analysis of relationships among temperature, snow and FFD

B. For each species: use lavaan to run reduced models – no SPDX and no DBG (others?); compare AIC values; choose model with lowest AIC