Descriptions of each R script and the order in which they should be executed.

**Data formatting:**

1. **phenocam\_metadata.R**: Download metadata for all PhenoCam grassland sites.
2. **download\_format\_pheno\_files.R:** Download and format all grassland site data (transition dates, environmental data) for use in phenor.
3. **Format\_SM\_Mosaic.R:** Add soil moisture to dataset.
4. **Replace\_Daymet\_precip.R:** Replace Daymet precip with SEGA or Ameriflux precip for Southwestern sites.
5. **split\_sites\_KG.R:** Divide sites by their Koppen-Geiger classification.

**Model fitting:**

1. **precip\_phenology\_models.R**: Fit all original and precipitation models to data.
2. **SM\_phenology\_models.R**: Fit soil moisture models to data.

*Within “Monsoon” folder: code for fitting each model on the supercomputer. We ran the code to fit the models 25x in parallel and then used different scripts to combine and organize the output.*

1. **looped\_models\_Monsoon\_precip.R**: Fit all original and precipitation models on the supercomputer. Output is a separate “.rds” file containing the model fits each time the code is executed (25x).
2. **looped\_models\_Monsoon\_SM.R**: Fit all soil moisture models on the supercomputer. Output is a separate “.rds” file containing the model fits each time the code is executed (25x).
3. **Read\_output\_files.R**: Combine all 25 model fits (output) from the supercomputer code above (works for both precip and SM model output). Creates results tables (lists of AIC/RMSE) and scatterplots/histograms to visualize spread of model fits.
4. **Combine\_results.R**: Pull out the best run for each model. Combine KG regional results and calculate AIC and RMSE of the combined dataset. Calculate residuals for use in dendrogram.
5. **validation\_loop\_Monsoon.R:** Leave-one-site out validation executed on the supercomputer. Output is a separate “.rds” file containing the validation fit each time the code is executed (25x).
6. **Read\_validation\_output\_files.R**: Combine all 25 validation runs (output) from the supercomputer code above. Calculate statistics from the validation runs (RMSE mean, sd, minimum). Combine results from separate validation runs for each KG group to determine overall validation RMSE. Extra code to extract the optimized parameters from each run to see the spread in values.

**Data visualization:**

1. **Site\_map\_KG.R**: Create site map with 2-letter Koppen-Geiger classification as background.
2. **Dendrogram\_figure.R**: Create dendrogram of models based on their residuals.
3. **Pretty\_figures.R**: Create all the figures used in the manuscript.
4. **Compare\_models\_ArrowPlots.R**: Extra code to create additional arrow plots comparing different model fits.

**Future projections:**

1. **Format\_FutureData.R:** Format future climate data for use in phenor. Use phenor and the best-fit parameters for the top model to predict spring transition dates into the future for a single site.
2. **FutureProj\_MannKendall:** Perform Mann-Kendall test and calculate Sen’s slopes for projected spring transition dates. Plot results of all 4 sites together.