

Thoughts on an improved temperature model

Lizzie, so far

August 3, 2020

Contents

1 What Lizzie suggests we do

I think the sine curve is not so bad. If we want to improve we could consider a spline (or something similar) **to capture weather fronts**, which is our main issue (to me, at least, from looking at the residuals). I would also like to **add a baseline long-term trend** since most locations are getting warmer with climate change. Even if we don't see it here I'd be interested in having it in the model soon-ish as we would want it in most models.

2 Overview of options considered

1. Fourier: I love Fourier, but I am not sure what we'll get from this beyond the sine curve for climate data here.
2. Autoregressive: Maybe, but I worry these will require a lot of decisions we're not sure of—the lag may vary over time, the timing of seasons varies ... they feel not flexible enough to capture weather fronts.
3. Sine curve: Seems pretty good as a first pass, but looks to miss weather fronts.
4. Splines: this is just a catch-all term for smoothing $f(x)$ s. They seem to usually chop up the time-series to create some bends. I have seen a lot of climate folks use thin-plate splines, but I am still trying to figure out why.
5. GAMs: This is just a generalized additive model where the linear predictor uses some smoothing $f(x)$ s.
6. ...

3 Emails and other notes

3.1 Christy Rollinson email (July 2020)

With the temperature modeling are you predicting past or future? I have a couple approaches I've used for gap filling and spatio-temporally downscaling daily and subfamily meteorology

data with uncertainty, but they're probably overkill. I've been using splines in GAMs a lot for this because of the nice tradeoff between being fairly computationally quick, ability to capture "surprises" and a variety of shapes among sites with the same function, and ability to robustly estimate uncertainty.

3.2 Ben Cook email (July 2020)

As for your query, I guess it depends on exactly what you want to do. By "model the temperature data", do you mean you want to generate a new synthetic time series (say, daily data for one year) sampled from the existing data? One approach to do this would be to, using the multiple years of data you have, generate an average daily climatology, and. then convert each year to ANOMALIES. Once could then resample from the anomalies (e.g., using an autoregressive approach) and then superimpose these new synthetic anomalies back on to the climatology. You would probably want to fit a different autoregressive model to each season (DJF, MAM, JJA, SON).

3.3 From Ailene (July 2020)

Lizzie chatted with Ailene in late July. Ailene was very pro-splines and just sent some links.

Splines are the heart of the brms non-linear modeling (see here).

Link about GAMs (in brms).

Splines in Stan.

3.4 Random references from Lizzie

A Spline Model of Climate for the Western United States for monthly data, from 2006

For daily rainfall but seems to have a good review.

Another article with a semi-useful intro.

A 2019 thin-plate spline paper

3.5 From Iñaki email (July 2020)

The only paper I can propose to read is the chapter we wrote with Isabelle about phenology models (see attached file). There is an important review of different models used to simulate phenology stages but it was quite focused on dormancy (endo and eco) phases. There is also

this one from Rebaudo et al. (focused on insects).

Usually we tested last years (with Amber and other colleagues): GDD, Sigmoid, Richardson, Wang, Normal (also called Chuine), triangle. There are other interesting as Logan (1, 2, 3), Anderson, Sine. But I did not tested them for plants and they are very similar to other models.

Lizzie notes: the two files are in the phenmodels/temperatre refs folder, but I am not sure they include temperature models.