What are the benefits of this model?

1. It accounts for external sources of data (e.g., rainfall and water level in non-impacted sites) that might influence the focus variable (i.e., water level in impacted sites).
2. It coherently deals with gaps in the different data streams.
3. It generates predictions that fully account for all the available data and that appropriately represents sampling and parameter uncertainty.
4. A more parsimonious description of the variance-covariance matrix, potentially leading to more efficient estimates

One can think of this model as one for which we want to make predictions for the impacted sites based on control sites and other variables (e.g., rainfall). However, there are substantial gaps in several of these predictor variables.

**Factor model**

We create a matrix where each column corresponds to a time series of river level or precipitation. We assume that the response for location l at time t is given by:

Where is a design vector that accounts for within-year seasonal effects. We assume that . We can write this expression as:

where and

Assume that . Then, after marginalizing over these latent factors, we obtain:

where is a diagonal matrix with diagonal elements . Correlation between locations arise when we marginalize over because these guys are unique for every time point.

Notice the dimensions of these matrices and , where P is the number of factors. Our data are given in the matrix . Therefore, for this to make sense, p has to be smaller than the total number of “locations” N.

Our priors are

Where reflects the overall average for location l in a particular month. Because data were standardized, it is fine to have a variance of 1.

V is set to 3

for where . Notice that we have to define separately because that defines the overall precision, which is then increased exponentially by all the other ’s.

#-------------------------------------------------------------------

Notice that

#-------------------------------------------------------------------

**FCDs**

* For

Notice we can write

Therefore

Where is a diagonal matrix with elements .

Notice that might not necessarily contain all the elements if some of the are missing.

* For

Notice that might not necessarily contain all the elements if some of the are missing. Furthermore, notice that the dimensionality here depends on the non-missing ’s. Let’s call this quantity

* For

Notice that might not necessarily contain all the elements if some of the are missing.

* For
* For

Because , then:

* For

Because , then:

* For

Taking logs, this becomes

Sample this with Metropolis-Hastings

* For

Taking logs, this becomes

Sample this with Metropolis-Hastings

* For

Where is equal to 1 (remember that variables were standardized) and .

Notice that might not necessarily contain all the elements if some of the are missing.

#-------------------

Fix MH algorithm

#-------------------

Make predictions

Let’s partition this as and the mean and covariance matrix appropriately. Then:

where