Forks in the Garden Paths

There are two statements of the research hypothesis of the study. In the "research summary,' published in the print version of Science, Sabo et al. (2018:1270) "… hypothesized that high fisheries yields are driven by measurable attributes of hydrologic variability, and that these relationships can be used to design and implement future flow regimes that improve fisheries yield through control of impending hydropower operations". In the full article published online, Sabo et al. "… tested the historic role of multiple variance components of the LMB hydrograph in driving harvest of fishes (total annual harvest in kilograms, standardized by effort) from the bag net or "Dai" fishery on the Tonle Sap River"(p 1). The second formulation is more specific, but either opens the gate to the garden of forking paths, since these research hypotheses leave many choices to the investigators: for example, how should the variance components of the hydrograph, or measurable attributes of hydrologic variability, be determined? Which of these should be used in the analysis? How far should the catch data be reduced? Since there are several possibilities, which gaging station record should be selected to represent the "variance components of the LMB hydrograph?" We discuss several of these choices below.

Using the record at Stung Treng:

Sabo et al. related the Dai fishery to the gage record at Stung Treng, but several alternatives are available (Halls et al. 2013). Perhaps Sabo et al. used the Stung Treng record because it is long, and so was useful for another aspect of their study (looking for the effects of upstream dams on the flow regime). However, it seems an odd choice. Stung Treng is a potential dam site several hundred kilometers upstream from the Tonle Sap-Mekong confluence. The Tonle Sap connects the Mekong to a large natural lake and adjacent floodplains that support large populations of many fishes (Figure __). As described in Dietsch et al. (2014):

> In Cambodia, the Lower Basin is characterized by a wide and flat flood plain and contains the Tonlé Sap River and the Tonlé Sap Lake (also called the Great Lake). The Mekong, Tonlé Sap River, and Tonlé Sap Lake form a unique hydrologic system in which the Tonlé Sap River reverses flow twice a year depending on high or low stage of the Mekong River. When stage is low in the Mekong River, the Tonlé Sap River flows to the Mekong River. Conversely, when stage is high in the Mekong River, the Tonlé Sap River flows to Tonlé Sap Lake. Tonlé Sap Lake serves as a flood-storage reservoir for the Mekong River. The Tonlé Sap Basin has a drainage area of approximately 84,000 km² and Tonlé Sap Lake is the largest fresh- water lake in Southeast Asia (Kummu and Sarkkula, 2008). Tonlé Sap Lake serves as an important agricultural and fisheries production area for Cambodia.

The Dai fishery targets fish migrating down the Tonle Sap to the Mekong as the Great Lake drains, using what are essentially stationary trawls with mouths facing into the current (Halls et al 2013). Although flow in the Mekong clearly affects conditions in the Great Lake, flow variation at Stung Treng is attenuated by distance, and the Great Lake also receives ~40% of its inflows from tributaries and rainfall (Halls et al. 2012). Stage in the lake lags stage at Stung Treng, and short-term variation in the Stung Treng stage is filtered out before it reaches the Great Lake (Figure 4).

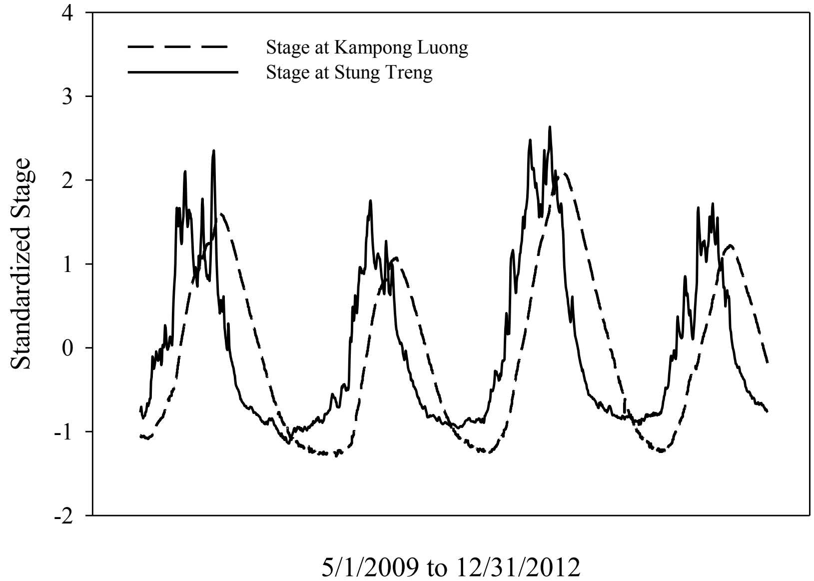

Figure 4. Normalized stage of the Mekong River at Stung Treng (solid line) and the Great Lake at Kampong Luong (dashed line). Units are standard deviations.

Sabo et al. may have elected not to use the Kampong Luong data because of gaps in the record. However, these are not extensive, and estimating the missing Kampong Luong data from the Stung Treng data or other and closer gage records in the area seems the better alternative, especially given the objective of testing "… the historic role of multiple variance components of the LMB hydrograph in driving harvest of fishes (total annual harvest in kilograms, standardized by effort) from the bag net or "Dai" fishery on the Tonle Sap River."

Using the stage record from Stung Treng:

Sabo et al. (2013) analyzed data on the stage of the Mekong River at Stung Trengt (J. Sabo, pers. comm., 1/18), rather than discharge. This seems odd, since the stage depends on the local channel geometry as well the discharge, which is the more relevant variable. Although the stage-discharge relationship for Stung Treng is not strongly non-linear, using the discharge record for the analyses instead of the stage record would have given somewhat different results (the current rating curve for Stung Treng is: Discharge = $2180.085(\text{Stage}-1.234)^{1.360}$). For example, the mean discharge is not the discharge corresponding to the mean stage, so flood pulse extent defined in terms of discharge would be somewhat different from flood pulse extent defined in terms of stage. Similarly, the characteristic signal for the discharge record would be somewhat different from the characteristic signal for the stage record, and so would NAA and the other variables defined in terms of it. (Confusingly, the article refers to discharge 25 times, and stage only 9 times,) Choice of a 20 year stage record:

From the DFFT analysis, Sabo et al. (p. 2) "…extracted the characteristic signal for the [stage] time series spanning 1993-2012, (consistent with the record of fishery yield, 1996-2012)." Time lags are not involved in the modeling, except for the 18 month window for the definition of the NAA component variables, so this seems arbitrary, and the stage record for 1995-2012 would seem more consistent with the fishery record. Probably Sabo et al. used 20 years because they also used 20 year intervals in the other part of their study, but the characteristic signal for 1995-2012) would have been somewhat different, and therefore so would the values of the variables.

Selecting variables for the analysis.

Because some pairs of the set of nine NAA shape component variables exhibited too much collinearity and some were not significantly related to CUPE, the set was reduced to 6, which were used in the MARSS modeling. Presumably, either of the collinear variables in a pair could have been selected, and if the variables were not exactly collinear, the choice would affect the results. They also choose to model only main effects, and not interactions. Apparently, some variables were selected based on statistical significance, despite their small effect sizes (Figure 2). One of these choices, high HSAF (from multiple storm peaks), seems implausible as a "driver" of the fishery, in light of the attenuation of the peaks between Sturng Treng and the Great Lake (Figure 4). Other choices may be perfectly reasonable, but regardless, they are choices made in light of the data.

Reducing the data to catch per row of Dais per year:

For the MARSS analysis, catch data from individual Dai were combined by location, as described by Sabo et al. (2017:2):

> Flow-fish relationships were developed using catch (biomass in kilograms) and effort data (Dai days, equal to number of fishing days across all Dais in a given Dai row and year) collected by the Inland Fisheries Research and Development Institute of Cambodia and maintained by the Mekong River Commission. The data set includes monthly species-specific harvest, allowing the calculation of catch per unit effort (CPUE) for 64 Dais in 14 locations on the Tonle Sap River. We pooled catch data within location for each year, culminating in a data set of 14 time series of total CPUE, each 17 years long.

How the data were reduced matters, because the harvest data are estimates based on samples collected with a complex stratified sampling plan that varied over time, as described in detail in Halls et al. (2013). Focusing on one of the complications, the Dai are arranged in 14 rows (locations), with one to seven Dai in each row. Some of the Dai regularly catch more fish than others, so the sampling is stratified with more intensive sampling of the 18 more productive Dai, and less intensive sampling of the other 46. This stratification allows for more accurate estimates of the total catch for a given level of monitoring effort, but the result is that the row data are of uneven quality, because the fraction of Dai in the high yield stratum in each row varies from all to none. The MARSS model estimates a separate observation error for each row for each year (Equation 2 in Sabo et al. 2017), but Sabo et al. do not describe any way to account for the varying expected accuracy of the data for each row.