Replies to Email Query

2025-03-26

Answers (with help from chatGPT)

1. Yes, Equation (3) defines a model where the left-hand side is expressed as cloglog(π), which is the complementary log-log transformation of π. This transformation is used because the probability π is modeled as a function of covariates.

If you are obtaining values of r similar to those in Figure 3 by considering the left-hand side of Equation (3) as π rather than cloglog(π), it suggests a possible misinterpretation. The correct interpretation should involve taking the inverse of the complementary log-log function to retrieve π from the regression output. This means:

$$\pi = 1 - e^{-e^{(\alpha_0+\alpha_1 SOI+\alpha_2 NAO)}}$$

You should apply the inverse cloglog transformation to correctly interpret the predicted probabilities.

2. Yes, the binomial parameter p can be calculated using the equation:

$$\lambda = r(1+p)$$

where: λ is the expected annual hurricane frequency, r is the expected annual cluster rate, p represents the probability that a cluster contains two hurricanes rather than just one.

Implication for p. If both λ and r are functions of climate covariates (e.g., NAO and SOI), then p is also dependent on these covariates:

$$p = \frac{\lambda}{r} - 1$$

This equation suggests that p is not necessarily constant. Instead, it varies with the ratio of hurricane frequency to cluster rate, which could change depending on environmental conditions. We also mention that by comparing the coefficients in Equations (4) and (5), a regression model for p can be derived, further supporting the idea that p is influenced by covariates. If you assume that p is constant, you might be oversimplifying the clustering behavior of hurricanes. Instead, modeling p as a function of climate variables could provide a more accurate representation of hurricane occurrence patterns.

3. You're on the right track in considering $P(H=0|r,p)$. In the cluster model, the total number of hurricanes in a given year is given by:

$$H = N + M$$

where: $N \sim \text{Poisson}(r)$ and where $M \sim \text{Binomial}(N,p)$ is the number of extra hurricanes from a cluster of size 2.

For the special case $k=0$ (i.e., no hurricanes in a year), this means there must be no clusters at all. That is, $N=0$, because if at least one cluster existed, there would be at least one hurricane.

Since N follows a Poisson distribution, the probability of no clusters occurring is:

$$P(N=0|r) = \frac{e^{-r}r^0}{0!} = e^{-r}$$

Thus, the probability that there are no hurricanes in a given year is:

$$P(H=0|r,p) = P(N=0) = e^{-r}$$

Which matches your intuition. The expression involving $k - i$ and i does not apply in this case because there are no terms to sum with $N = 0$. The binomial part of the model (which governs cluster sizes) only comes into play when at least one cluster is present.

So, in summary, your initial thought was correct:

$$P(H = 0|r, p) = e^{-r}$$

The undefined expression issue does not arise because the summation formula is only used for $k > 0$ while for $k = 0$ we directly use the Poisson probability of zero clusters.

4. Here is the approach for generating estimated annual hurricane counts for X years using our cluster model.

Step 1: Since the number of clusters per year, N follows a Poisson distribution with rate r $N \sim$ Poisson(r). For each year, generate a random draw from this Poisson distribution to determine the number of clusters that occur in that year.

Step 2: Each cluster contains either one or two hurricanes, governed by a Binomial process $M \sim$ Binomial(N, p) where N is the number of clusters from step 1 and M is the number of additional hurricanes from clusters that contain two hurricanes (each cluster contributes at least one hurricane, and an additional one with probability p).

Step 3: Then the total number of hurricanes in a given year is $H = N + M$

Step 4: Repeat for X number of years