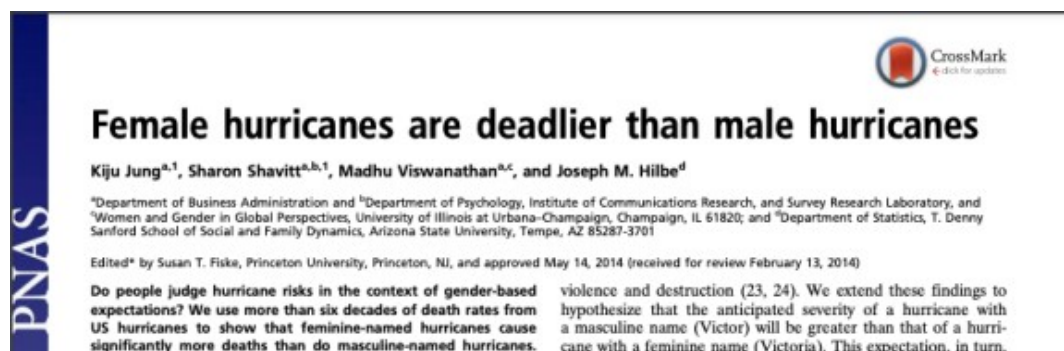Do you remember the notorious hurricane / himmicane study (Jung et al., PNAS, 2014)?

At the time, there was a heavy backlash against the study, and probably rightly so, as the statistical analysis turns out to be highly unstable against a change of the regression formula. You can find some links here. Over the years, however, I have found that this study has at least one virtue: it is an excellent example for teaching students about the importance of selecting the right functional relationship when running an analysis, and that substantial "dark" uncertainty can arise from these researcher degrees of freedom.
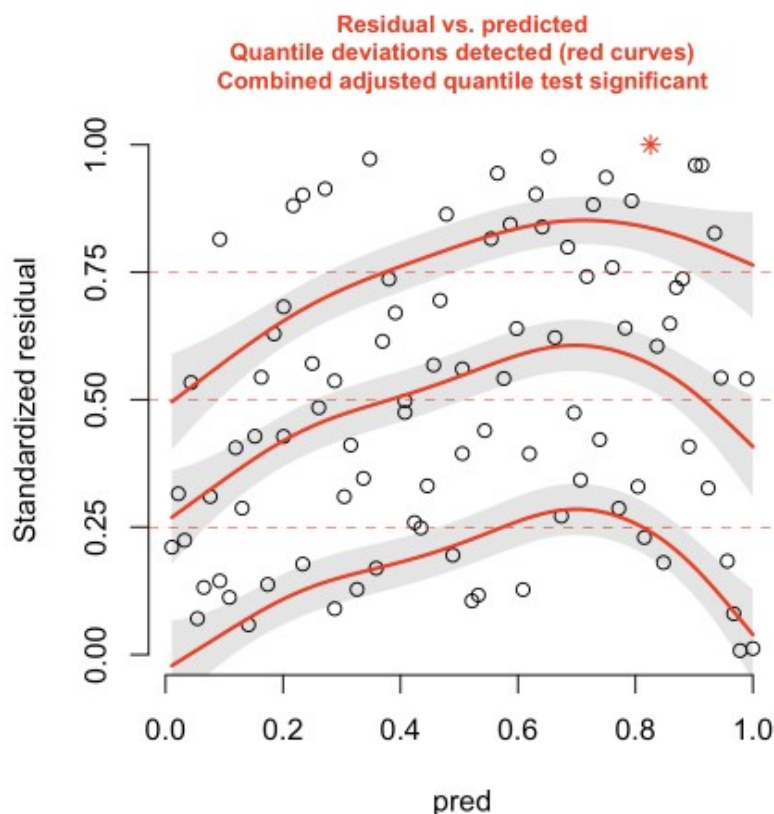
The reason why the hurricanes make such an excellent pedagogical example is that, as I point out here, the effect of femininity is highly unstable and depends strongly on which predictors you select, presumably because of high collinearity, the considered interaction(s) and the unbalanced femininity / mortality distribution.

In the stats course that I just finished teaching, I gave the students the task to re-analyze the hurricane data, which also led me to run some DHARMa residuals checks on the original negative binomial model fitted by Jung et al. Here is the residual analysis of the model with DHARMa, for technical reasons fit with glmmTMB and not with mgcv (as in the original study). The main DHARMa residual plot shows a kind of funky pattern, but those are not flagged as significant by the tests:
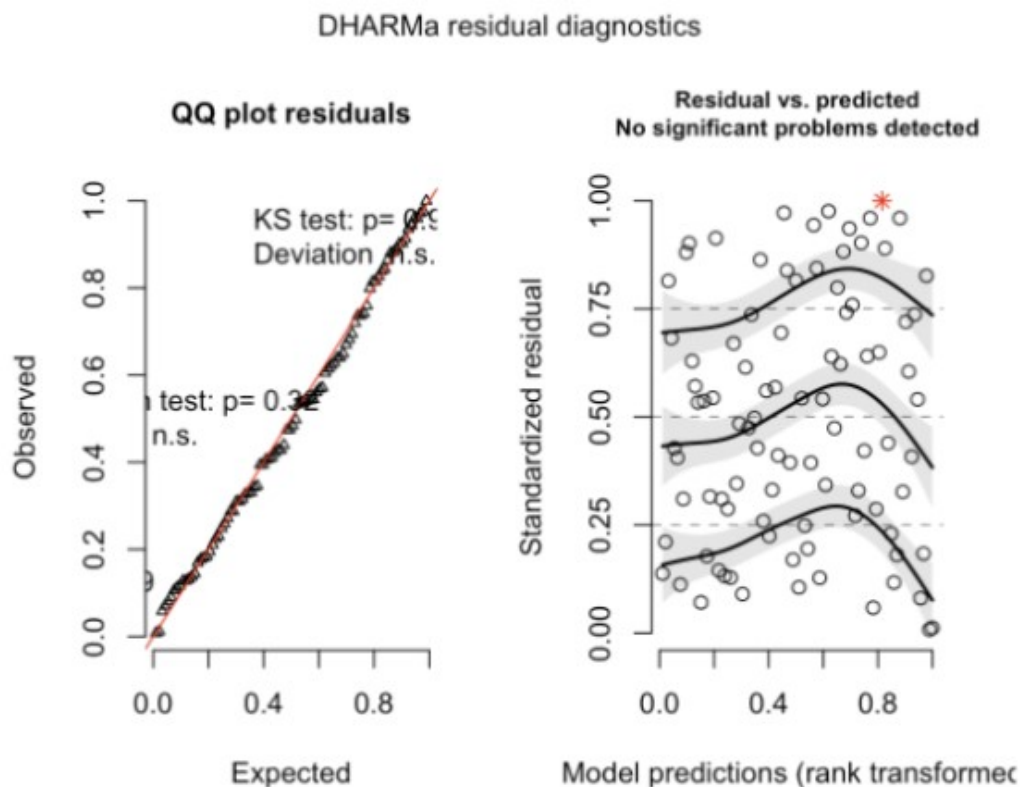

If we plot residuals against NDAM, however, we get a clear and very significant misfit. The original model

is obviously not acceptable, and the student teams that did the re-analysis practically all spotted this immediately. Serves as a reminder of how efficient systematic residual checks for GLMMs are. In the defense of the authors: DHARMa was not available at the time, although this pattern was also visible in standard Pearson residuals, as pointed out by Bob O'Hara at the time.

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant

We also find (light) spatial autocorrelation, but with a negative lag 1. One may speculate that this could arise if people are more careful after a particularly deadly hurricane in the last year, but it is equally possible that this is a fluke / false positve.

Residuals vs. time   **Autocorrelation**
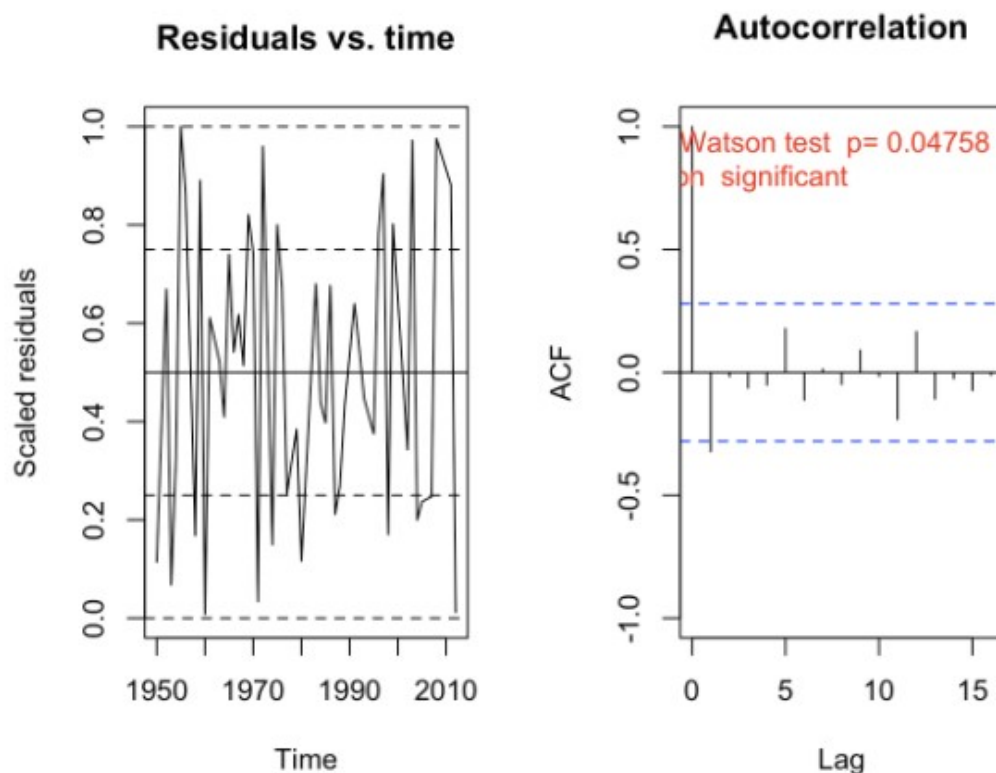
If you want to repeat the residual analysis, here's the code. The data is conveniently stored by PNAS.

```
library(gdata)
Data = read.xls("http://www.pnas.org/content/suppl/2014/05/30/1402786111.DCSupplemental/
pnas.1402786111.sd01.xlsx&quot;,
nrows = 92, as.is = TRUE)
library(glmmTMB)
originalModelGAM = glmmTMB(alldeaths ~ scale(MasFem) *
(scale(Minpressure_Updated.2014) + scale(NDAM)),
data = Data, family = nbinom2)
# Residual checks with DHARMa
library(DHARMa)
res <- simulateResiduals(originalModelGAM)
plot(res)
# no significant deviation in the general plot, but try this
# which was highlighted by https://www.theguardian.com/science/grrlscientist/2014/jun/04
/hurricane-gender-name-bias-sexism-statistics
plotResiduals(res, Data$NDAM)
# we also find temporal autocorrelation
res2 = recalculateResiduals(res, group = Data$Year)
testTemporalAutocorrelation(res2, time = unique(Data$Year))
```