One more idea about normalizing the GPI. What really matters here is not how many storms are implied by each model in the control runs, but rather how different is the GPI in the LGM simulation (or Pliocene or Holocene) from the control. So some information along the lines of "CCSM4 showed a 10% increase in the LGM GPI compared to the control, while MRI showed a decrease of 5%" could be useful. And for this, the value of 'b' or whether to include 'b' at all in the definition becomes irrelevant.