Hydrology and Earth System Sciences Discussions

Interactive comment on "Dynamically vs. empirically downscaled medium-range precipitation forecasts" *by* G. Bürger

Anonymous Referee #1

Received and published: 19 June 2009

General Comments The paper presents and compares two precipitation forecast systems (forecast model and downscaling) over three basins in the German Alps. The paper is well organized and well written. I do think that some additional material is needed to better argument that the comparison of the two systems is equivalent to the comparison of the two downscaling methods. I also do think that the results should be more analyzed because I think that the skill is low in general. See my specific comments for more details.

Specific comments The title and the introduction state that the goal of the paper is to compare two downscaling methods: empirical or dynamical. However, when the systems are presented, it becomes clear that the author will in fact compare two systems

and not two downscaling methods. The reader has to wait for the discussion section for the author to explain how in fact he/she thinks it is equivalent to compare two downscaling methods. I think that an earlier comment, at the beginning of the results section for example, would help clarify this point earlier.

On the same subject, the author relies on literature ("published comparisons", page 3527, line 26) to justify his argument that difference between the two forecast models (before downscaling) are not significant and therefore the comparison of the two systems is equivalent to comparing the two downscaling methods. But the author also mentions that those publications diverge on their conclusions of which model is more skillful than the other (do they use the same scores?), and that it is strongly dependent on time and location. I think that the author should be more convinced about this "un-significance of differences between the two models" as this statement is the basis for the topic of the paper. I would encourage the author to make his own comparison of the raw forecast (output from the models), and use the same skill score he uses in his analysis for downscaled forecasts. Then he can briefly compare his comparisons to published comparisons but most important, he can justify in a stronger way that comparing the two systems is equivalent to comparing the two downscaling methods.

In the results sections, page 3525, the rate of misses are very high (53 missed for 59 events for GME/LM, even 33 misses for IFM/EDS)? Similarly, Gilbert skill scores seem very low (0.21 and 0.06). What would be the skill if climatology was used? Or persistence? Do other systems report scores that low? The skill is positive but the system will be skillful if they improve upon persistence or climatology. The scope of the paper is to compare the two systems but one additional conclusion might be that none of them improve upon persistence or climatology.

In the cost loss analysis, page 3526, it seems to me that the conclusions are very sensitive to the loss over cost ratio (L/C). Is a L/C ratio of 10 typical? What would be the conclusions if loss was in fact 100 times more expensive than cost?

Technical corrections Page 3519, line 29: explain IFS acronym Page 3522, line 24: explain EDS acronym Page 3525, line 6: I would specify that the reduced variability is shown by a flat cdf to be clearer. Page 3525, line 6: "Compare to observations, too few strong events are simulated by GME/LM". I would put this statement in the next paragraph, once "strong events" are defined. Page 3527, line 17: one system forecasts only up to 2 days in advance, the other one 5. The sentence is then misleading because it states that both systems go to 5 days.

———————————————————

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 6, 3517, 2009.