Interactive comment on “Systematic bias of Tibetan Plateau snow cover in subseasonal-to-seasonal models” by Shuzhen Hu and Wenkai Li

Anonymous Referee #2

Received and published: 21 July 2020

This paper examines how a set of state-of-the-art subseasonal-to-seasonal (S2S) forecast systems predict the Tibetan Plateau-wide snow cover and surface temperature over the 1999-2010 period. The forecast systems from ECMWF, CMA and NCEP are used in the intercomparison. The evaluation of forecasted snow cover is made against the multi-instrument (IMS) snow cover data. A connection is drawn between the bias in snow cover, which increases systematically with lead time in all models, and a bias in surface temperature. Model experiments with the WRF model support the idea that the latter is induced by the snow excess through land-atmosphere coupling. Few papers have examined snow forecast on the S2S time scale on the Tibetan Plateau (TP), a region with well-known biases in snow and surface temperature. This is an innovative...
study that is well-worthy of publication in The Cryosphere. I however recommend a major revision of the paper, before it is in an acceptable form for publication.

MAJOR COMMENTS 1) Some brief description of how the three different models are initialised in terms of snow and land surface is needed, complementing the description of the land-surface models. What sort of snow analysis is used to initialise the different models? Isn’t ERA-5, which has strong snow biases over the TP, used to initialise the ECMWF S2S reforecasts? Which observations are used in these snow analyses, both globally and over the TP more specifically? While IMS is not used to initialise the ECMWF forecast model over the TP (see Orsolini Y. et al., 2019), is it used in the other systems? The quality of this initialisation would most certainly influence the snow forecasts at the S2S time scale. The initial snow values may be a key addition to bring in Fig 3. 2) Some more details about how the snow cover fraction conversion is derived in each model is needed. For example, a 100% snow cover may mean very different snow depth or snow water equivalent (the actual prognostic variables) in the different prediction systems. Also, I believe that IMS provides a binary information snow cover being 0 or 1, with the former case meaning that the fractional snow cover is below 50%. When aggregated to a 1-degree grid (L88), isn’t there a range or uncertainty in the IMS aggregated value (given that the observed fractional snow cover could actually be 0 or 50%)? Please clarify these points and the implications for forecast verification. 3) The authors consider the snow cover averaged over the entire Tibetan-Plateau, but they do not show any snow cover maps although there is considerable heterogeneity, as shown by the authors in previous publications. I wonder if there could be compensation effects between different geographical sub-regions, that could result in an agreement of the TPSC index between forecasts and IMS. Different prediction systems may have different regional biases. Showing such maps would re-assure the reader that the prediction systems capture the main climatological features of the snow distribution over the TP and its S2S variability (e.g., in subregions as in Li W. et al., IJOC, 2019). Could spatial pattern correlation between IMS and snow forecasts be helpful in this case? 4) The authors could also try to examine the snow-temperature coupling strength (as an
indicator of land-atmosphere interaction) in the respective prediction systems by correlating the local forecasted temperature and forecasted snow, as a function of lead time. See Diro and Lin (2020) or Li et al. (2019).

computational work, but it would add to the robustness of the conclusions. At least, a word of caution in the summary is warranted.

Typos / English L43: hydrologic cycle L28: radiative rather than radiant L108: the total variability L113: and the three different L143: the preceding week rather than the last week, seems more appropriate L152: accumulation, leading to a systematic TPSC bias. L168: growth is used for a declining variable. Either decrement, reduction or decline should be clearer. L173: real rate should be observed rate or rate derived from IMS. L173: indices or index L202: growth of SAT, rather than TPSC (it says 1.2 degree). L229: land-atmosphere interactions L256: Hence,