The Cryosphere Discuss., 8, C3018–C3025, 2015
www.the-cryosphere-discuss.net/8/C3018/2015/

The Cryosphere
Discussions
Open Access

Interactive comment on "Verification of forecasted winter precipitation in complex terrain" *by* M. Schirmer and B. Jamieson

M. Schirmer and B. Jamieson

michael.schirmer@usask.ca

Received and published: 4 February 2015

We thank the anonymous reviewer for these very helpful, detailed and constructive comments. They helped to improve this manuscript significantly. The reviewer's comments are quoted in citation marks.

"General comments:

1) In this study, continuous modelled precipitation time series are made of successive 12-h forecasts (+6h to +18h using two initiation times per day). The authors should precise which initiation times are considered (00Z and 12Z?). An interesting aspect would be to consider time series made of successive 24-h forecast from one initiation time per day. This would allow the authors to build two continuous time series. Using

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

24-h precipitation forecast is more relevant for an avalanche or flood forecaster than using successive 12-h forecasts. Indeed, they generally need to take decision based on the forecast for the next 24 hours (see the example of road closure P 5745). What is the impact on performance measures? Does it change the economic values of a forecast?"

The reviewer is completely right that successive 24 hour forecasts are more meaningful for the purpose to verify 24 hour precipitation sums. For example for GEMLAM, we are building successive 12-hour blocks including the forecast hours +7 h to +18 h, and switching to the next initiation time afterwards using the forecast hours +7 h to +18 h as well. This means that a decision maker would have the quality presented by our analysis only up to 18 hours in advance. But we are communicating results in a 24 hour precipitation sum, for which earlier time steps were filled with a previous initiation time. This is now more clearly stated the manuscript. In an ideal way the forecast hours +7 h to +30 h would be have been used (excluding the first six hours because of spinup issues). However, there are some arguments for keeping our procedure. The main argument is that 'Snotel' data is available in both a 1 hour and a daily dataset. The daily dataset is quality checked prior to downloading (email communication with staff of the USDA NRCS National Water and Climate Center, 11 June 2014). These daily dataset covers the time window from 8 UTC to 8 UTC. To match successive 24 hour forecasts with these daily observations, forecast hours of up to +37 h need to be included for GEMLAM initialized at 18 UTC. It can be assumed that the longer the forecast the lower the quality of the model. For the 6 UTC initiation time the corresponding forecast hours would be even worse. This is not our aim, we wanted to verify the short-term forecast. Only with hourly observation data, the ideal set of forecast hours (+7 h to +30 h) can be considered, but we would lose the quality control on the observation side for the 'Snotel' data set. The effect of switching to the hourly 'Snotel' data set cannot be tested in an easy way, because a thorough data quality procedure for many stations and for two years need to be done. Our procedure to generate the time series is very similar as for the operational forecast product we deliver to the Canadian

Avalanche Centre (see also Bellaire et al., 2011, 2013; Bellaire and Jamieson, 2013) and we were operationally interested in the quality precipitation generated in this way. In regions without weather stations available, our way of generating the time series are especially useful: In forecast areas with weather stations, our filling procedure with a previous initiation time could have been replaced by using observations. But there are many regions for which no weather stations are available, and for which decision makers are interested in a short-term forecast (up to 18 hours), which is communicated in a 24 hour sum. Keeping the 24 hour sum is useful, since it is a common measure and decision makers are used to it. Also, rather irrelevant timing differences between model and station are not effecting the performance measures in a daily sum. This is now more clearly mentioned in the manuscript. We suggest to keep our procedure but to communicate the limitations in a clearer way (focus on short term forecast of up to 18 hours, but included in a 24 hour precipitation sum, and mentioning the reasons). Additionally we provided for those stations recording in hourly data (all except of 'Snotel') an analysis that uses the ideal successive 24 hour forecasts including the forecast hours +7 h to +30 h and reporting the decrease in quality compared to the original version that uses only forecasts of up to 18 hours. This decrease is small compared to the differences between the two NWP models of different resolution, which delivers an argument for keeping our procedure as well.

"2) The elevation difference between actual and model terrain height is a key parameter when evaluating NWP models in complex terrain. It is only mentioned in the text (P5733 l. 23 to P 5734 l. 3). A figure summarizing the differences between actual and model terrain height at different horizontal resolutions would help the reader to quantify the importance of theses differences. To handle these differences, the authors corrected the modelled data (including precipitation) for elevation differences following Liston and Elder (2006). The impact of the correction must be clearly quantified, especially since precipitations are corrected based on a factor that varies seasonally (Eq. 33 and Tab 1 in Liston and Elder (2006)). The text mentions that "these corrections increased the performance of the model" (p 5734, l. 1). To what extent are they improved? Are modelfriendly scores similar when considering for evaluation only stations with an absolute value of difference between actual and model terrain height lower than a given value (100 m or 200 m for example)? At 2.5 km grid spacing, the number of stations should be sufficient to compute relevant statistics."

We added a figure showing the elevation differences between model and weather stations. We also added a more detailed description of the effect of elevation corrections on the performance measures and included an evaluation without corrections considering only the stations with a small elevation difference (see newly added section 3.5 in the results section). We thank the reviewer for this very constructive comment.

"3) The authors use the "daily new snow amount" to evaluate the quality of forecasted precipitation. The term "daily new snow amount (HTN)" should be more precisely quantified. Indeed, it is usually defined as: "Height of new snow is the depth in centimetres of freshly fallen snow that accumulated on a snow board during a standard observing period of 24 hours." (Fierz et al, 2009). In this study, the height of new snow has not the same definition and refers to a difference of snow depth between 24 hours. It includes the settling of new snow under its own weight and the settling of the underlying snow layers. The author uses SNOWPACK to account for the settling processes. A more accurate description of the use of SNOWPACK would be very helpful. For example, through a subsection describing the use of a detailed snowpack model to evaluate daily new snow amount: (i) Which atmospheric forcings are used to drive SNOWPACK? (ii) Is SNOWPACK run continuously from the beginning of the winter? (iii) What are the main limitations of the method: settling, density of fresh snow, melting, wind-induced erosion, : : : (partially discussed P 5743 l. 5-11)."

We included more detailed definition of the term "new snow amount", including the fact that it is also dependent on the settling of the underlying snow and not only on the settling due to its own weight. We also added a more detailed description of SNOWPACK, input parameters, model setup and discuss main limitations, which are to our opinion the parameterizations for new snow density and settling derived in Switzerland. This was discussed in P 5743 l. 5-11. Melting may be a marginal factor for stations over 1500 m between November and March. Moreover, melting would only result in a problem if a positive snow depth change caused by precipitation would coincide with a melt event. In all other cases negative snow depth changes due to settling or melt will be pooled under the category 'no precipitation'. Erosion due to saltation, suspension and sublimation processes effect both measurement systems and were thus discussed in section 3.2 (now 3.3). We added in the model description that SNOWPACK was used without its snow drift mode and thus mentioned processes were not accounted for in the model. This is reasonable since modelled wind speeds from the NWP models were not verified here and would have been a large source of error (see also Vionnet et al., 2014). The stations are regularly located in rather wind sheltered, representative areas as discussed in section 3.2 (now 3.3).

"Specific comments:

1) Title: The name of the paper is questionable since it also contains an evaluation of output from a precipitation analysis system. Outputs from this system are not "forecasted precipitation". Therefore the name of the paper should be modified. Maybe "Verification of analysed and forecasted winter precipitation in complex terrain"." We appreciate the helpful comments of the reviewer, and are very happy to change our title to this one, which reflects the content of this paper better. "2) P 5728 l. 19 to 26: This paragraph is rather unclear and should be reformulated to focus more on the importance of a good estimation of winter precipitation in complex terrain and why NWP models are relevant for this estimation. Maybe split this paragraph into two paragraphs." This paragraph was split in two parts. The second part was reformulated and is now better focusing on the relevancy of forecasted winter precipitation in complex terrain. "3) P 5729: Clearly define the terms "high-resolution" versus "low-resolution" since the meaning of these expressions differs from one community to another." We added a definition ("a few kilometres grid size") for high-resolution models which was used in Rotach et al. (2009). When referring to Weusthoff et al. (2010), we changed

“low-resolution” into “lower resolution models”, and provide a range of grid sizes used for both categories in their study. “5) P 5733 l. 5 to 14: Include a short description of physical parameterizations in the NWP models involve in the generation of precipitation (cloud microphysical scheme, convection scheme : : :). This will help the users from other models to know what is implemented in GEM.” We added a description of the two NWP models and included citations describing the models in more detail. “6) P 5735, l.6-17: Precise over which hours are considered to compute observed daily accumulation (HN and HNW)? Same question for simulated daily accumulation (P 5734, l. 7-9) ?” This is clarified in the new version of the manuscript. “7) P 5737 l. 19: Eq (4) to (8) must be coherent. In Tab. 1, the variables a, b, c and d refer to numbers of events while in Eq (4) a, b and c refer to the relative frequency of the different outcomes contained in the contingency table (a/n, b/n and c/n with n being the total numbers of observations).” We thank the reviewer for detecting this typo! “8)P 5738, l. 17-19: the authors mention the analysis of model performance as a function of difference between station and model elevation. However, the results of this analysis do not appear in the paper (see General comment 2).” We added the analysis as suggested by the reviewer (see General comment 2). “9) P 5739, l. 5-10: A potential explanation could also be the settling of new snow. Steinkolger et al (2009) reports settling rates reaching 10 cm/day for freshly fallen snow.” We made clearer what we meant with “differences in units” and added settling as an explanation. “10) P 5740, l.11-13: The poorer performances of CaPA in mountainous terrain in wintertime is not only associated with the quality of data entering the analysis system. It is also associated with the fact that correlation functions do not account for elevation and the number of stations entering the analysis may not be optimal in mountainous terrain.” We thank the reviewer again for these helpful comments and added these explanations. “11) P 5743 l 25: Present the analysis of economic value in a separated subsection to clarify the paper and reduce the size of current section 3.1.” Done as suggested. “12) P 5747 l. 2-3: no dependency was found with elevation at the scale of western Canada and NW US. What about potential elevation dependencies at the scale of a mountain range with a sufficient number of stations? In the US, it appears on Fig. 7 that you may have a sufficient number of stations in some mountain ranges to carry out such analysis." Stations in interesting clusters, for example around Mt. Rainier, WA, or Old Baldy, MT, do not show a necessary elevation distribution nor a windward/leeward distribution. Thus, we concluded that our stations, even in the US, are to our opinion not distributed in such a way that the mountain range scale can be investigated. "13) P 5751 l. 14-17: the evaluation of a regional climate model (RCM) in complex terrain is not the main topic of this paper focusing on the evaluation of NWP system to forecast daily winter precipitation in complex terrain. The configuration of the NWP model may have evolve during the evaluation period and this evolution period covers only 2 winters (contrary to Ikeda et al. (2010) who studied for example four winters). I recommend the authors to remove the mention to RCM throughout the paper (at the end of the introduction and in the conclusion)." The presented analysis in two winters showed clearly an improvement in winter precipitation from the lower to the higher resolution model. This should suggest how RCM should be configured (e.g. resolution, cloud and precipitation microphysics) to capture winter precipitation in complex terrain. This is why we think this two year analysis of NWP models allows to suggest implications on the design of RCM. We therefore would like to include this link in the conclusion and in the introduction from this analysis to RCM.

"Technical comments Text Abstract: mention that this study is focusing on winter precipitation earlier in the Abstract."

Changed as suggested.

"P 5730 l. 1: replace "used in our study" by "evaluated in our study" since no specific GEM simulation has been carried out in this study."

Changed as suggested.

"P 5731, l. 29: use "Mahfouf" instead of "Mahfoufh"."

TCD

8, C3018–C3025, 2015

Changed as suggested.

“P 5737, l. 18: “: : : based on the empirical : : :””

Changed as suggested.

“P 5747, l.23: add “turbulent suspension” as a process not resolved at the scale of current NWP models.” Changed as suggested.

Interactive comment on The Cryosphere Discuss., 8, 5727, 2014.