Dear Editor,
Thank you for your constructive comments to the discussion paper and for the opportunity to submit the revised paper. We have answered all the comments from reviewers in the authors' responses and revised the manuscript accordingly. The revised paper is significantly improved as a result of addressing these comments as we now have assessed the sensitivity to our findings to all factors that could possibly be impacting the results.

As the reviewers and the editor suggested, we have restructured the manuscript into the standard IMRAD format. We have merged the introduction and background sections as suggested and updated the text to make the introduction shorter and more succinct. After reading the review, we think that the description of data (including the source and methods to generate the data) used in the study was not clear enough. It can be noted that the gridded temperature and wind speed were generated for all of Norway by the Norwegian Meteorological Institute and we used this data in our study. In this revision, we tried to better describe all the data sources used within the subsection "Data" under the section "Materials and Methods".

The methodology and the context of the study as well as main results remain unchanged. We have added a separate discussion section to the paper. As the reviewers and editor mentioned, we agree that the justification for the approaches used in the study was limited in the discussion paper even though we feel we described the methodology clearly. In the revision, we have used the Discussion section to provide proper justification for the approaches used and to discuss the uncertainties and limitations in the study.

Reviewer 2 pointed out that the back calculation of reflectivities from the hourly radar precipitation data originally based on accumulated data of 7.5 minutes would only be correct if the precipitation is even within the hour. Unfortunately, the reflectivity data used to produce the radar precipitation rates (SRI product) used in the study are not stored in the production process, and therefore not available at the Norwegian Meteorological Institute (met.no) (Elo 2018, Personal communication). However, Plan Position Indicator (PPI) of the lowest elevation beam from Hurum radar with the original short time resolution is available from met.no and in the revised paper we have repeated our computations for the comparison of the proposed nonparametric radar precipitation estimation with radar precipitation rates computed using separate equations for snow and rain.

In the revised manuscript, we have used the PPI data to redistribute the hourly data (SRI) by assuming precipitation intensity distribution within each hour is as same for SRI as PPI. The redistributed precipitation rates were then converted to reflectivities and these data are used for the analysis as in the original manuscript. It should be noted that there is uncertainty in how accurately the redistributed intensity distribution of SRI represents the original distribution, however, this exercise at least used a possible realistic distribution.

We believe that we have revised the manuscript as described in the authors' responses to reviewers and in the case of using different Z-R relationship for snow and rain, the availability of the PPI data made it possible to improve the analysis beyond what is discussed in the response to the reviewer2.

According to HESS requirements, this submission consists of the revised manuscript and a point-by-point reply to the comments to the three reviewers and a marked-up manuscript version showing the changes made. In case, if you need any clarifications or further details, please feel free to contact us. We hope that Hydrology and Earth System Science will find it an interesting contribution.

With thanks,

Sincerely,
Kuganesan Sivasubramaniam
Ashish Sharma
Knut Alfredsen

Response to the review of hess-2018-0351

RC1: Responses to S.R. Fassnacht (Referee 1)

The authors wish to thank the reviewer for his constructive comments and corrections to the discussion paper. In the following, we have responded to each of the comments from the reviewer and showed the page and line numbers of the revised manuscript if any changes. The comment from the reviewer (RC) is in italic font while the author comment (AC) and changes in the manuscript (CM) are in blue normal font.

This is an interesting paper that should give us some improved insight into using weather radar to estimate rain and snow fall. This is very relevant in higher latitudes and in mountain environments were snow is important. To date, there has been limited work in using weather radar for snowfall estimation in a hydrological context. The methods presented herein could be used in many locations. However, the writing is unclear, and I got lost at times. I suggest that the authors revisit their objectives and ensure that the paper addresses these. Also, the Discussion is essentially missing as the work is not put into context of the few other relevant studies. Below I outline restructuring and a problem with the Methods/Data.

Equation 4 uses air temperature and relative humidity to estimate the phase of the precipitation from Koistinen et al. (2004) and used by Saltikoff et al. (2015) for Finland. However, air temperature at the gauge is used, and this is not correct. Fassnacht et al. (1999; 2001) lapsed the air temperature up to the radar measurement height. There can be 5 to 10 degrees Celsius difference between the temperature at the height (2 m above the ground) and the radar height (1 km stated on page 8 line 15). At minimum this should be discussed?

AC: As it is mentioned, we adopted the operational method from the Finnish Meteorological Institute as presented by Koistinen et al. (2004) and Saltikoff et al. (2015). In their phase equation, near surface temperature (2m above ground) is defined and we did follow their method as defined in the papers. We do agree with the reviewer that air temperature at radar measurement height can be different from gauge height and hence the estimated phase can be different. This has already been discussed on p17, l12-13 “Further, our phase classification is at gauge level, and represents near surface conditions. The phase of the precipitation can be different at the elevation where the radar measures the reflectivity.”

Air temperature can be lapsed to the radar measurement height to estimate the phase of precipitation. Fassnacht et al. (1999) and Fassnacht et al. (2001) assumed the temperature lapse rate to be zero in their studies as winter lapse rate is often zero in mid latitude areas. For the Nordic region, Tveito and Førland (1999) showed that vertical lapse rate varies with season and location. Further, Tveito et al. (2000) found that local terrain conditions have greater influence in local temperature gradient during winter. Due to the occurrence of inversions, lapse rate can deviate substantially from the standard (-6.5° C/km) during the winter months and it can be as low as -1.2° C/km (Tveito et al., 2000, Tveito and Førland, 1999). The estimated temperature at radar measurement height and hence the probability of liquid phase (Plp) are therefore highly uncertain. The measurements of phase information at radar measurement height with the use of dual polarized radars can be a useful data source for further investigation.

After receiving the reviewer’s comment, we investigated the sensitivity of our results to the use of a lapse rate. The air temperature at the radar measurement height (1 km) was computed using a standard moist adiabatic lapse rate (-6.5° C/km) as used in Nordic meteorological studies (Tveito et al., 2000). We estimated the probability of liquid precipitation (Plp) in Eq. (4) by using the lapsed temperature at radar measurement height while assuming the relative humidity unchanged. The estimated Plp was used to classify the precipitation phase and repeated the work as presented in section 5.5. Our results (not included in the revised paper) showed that new classification did not improve RMSE compared to the use of near surface

phase classification. We attribute this to the considerable uncertainty associated with the use of the lapse rate as noted by others (Al-Sakka et al., 2013, Tveito et al., 2000). Further, equation (Eq. (4)) is developed and tested for surface phase classification and it incorporates relative humidity as a variable. We assumed relative humidity at the radar measurement height similar to the surface value. This is also a potential error source in the computation.
We, therefore want to keep the Finnish Meteorological Institute's operational method of surface phase estimation to classify the precipitation as the method of choice in the paper. This is both in operational use and developed for the Nordic area which gives us some confidence in the method. We discuss this issue in the revised manuscript.
CM: p18, l11-20

What about a split sample approach of calibration and evaluation for the partial weights and k-nn approach? Also, the partial weight for radar precipitation was shown to vary from 0.4 to 1 (Figure 2), so why was a single average (mean) used in the k-nn prediction model. Is this approach not robust enough to have a different partial weight, or perhaps a gridded partial weight? It is stated that there is no spatial pattern in the partial weights, but an interpolated residual type approach could be used (e.g., Fassnacht et al., 2003 among others).
AC: As described on p12, l15-17, a split sample test was done to verify the results obtained from the leave one out cross validation (LOOCV) approach and presented in the paper.
Partial weights did not show any spatial pattern that would allow us to generate an informed specification of the weights that could be applied over the study region. Further, the RMSE estimated at gauge locations with the single average partial weight for the study area (as we presented in this paper) showed a strong resemblance with the RMSE estimated by using the partial weight estimated from the five nearest gauges. Hence, we decided to use a single average partial weight to present in this paper.
As the reviewer mentioned, it is possible to use gridded partial weight or an interpolated residual type approach. However, we found that the gain in RMSE is not significant for the effort of using gridded or residual type partial weights. The added complexity of gridding the partial weights does not add significant information to the analysis and we therefore recommend using an average value in the computations for this study region.
CM: p17, l17-19

The paper does need restructuring and rewriting. At present I get lost in where I am in the text, regardless of the "foreshadowing" sentences that appear at the end of various sections.
1) At the end of the Introduction, the paper should tell the reader specific objectives that were investigated or research questions that were answered.
AC: As per the reviewer's suggestions we have added specific objectives at the end of the introduction.
CM: p4, l1-8

2) Some of the material in the Background is repeated from the Introduction. For example, the three paragraphs in section 2.1 (Radar precipitation estimation in cold climates) mostly in the Introduction. Either reduce the Introduction or merge the Background with the Introduction to remove the repetition. I suggest the latter and to consider adding sections to the Introduction (e.g., 1.1. Weather radar use for hydrology, 1.2. Radar precipitation estimation in cold climates, 1.3. Nonparametric Radar rainfall estimates).
AC: The Introduction and Background sections are integrated into a shorter and more specific Introduction to the paper as suggested by the reviewer.
CM: p1-p4, section 1

3) There are methods presented in the Study Area and Data section. These two sections 3. Methods and 4. Study Area and Data need to be revisited to put all the methods together. I suggest a brief section first on Study Area, then a section on Data and Methods, describing the data first, then the methods used.

AC: As reviewer suggested, we have added a section "Materials and Methods" where a brief subsection first on Study area, and then Data followed by Methodology.
CM: From p4, l9 to p9, l15

4) The Results and Discussion are combined, and the Discussion is thus limited. I recommend that the Results and Discussion sections should be presented separately, or that the Discussion be much more in depth. There are only three citations in the entire Results and Discussion section, while numerous useful citations are presented in the Introduction and Background sections. There is no Discussion that put this work into context; the Results and Discussion only presents how do these results compare to the findings of Fassnacht et al. (1999), Koistinen et al. (2004) and Saltikoff et al. (2015).
AC: As reviewer suggested, we have added a separate Discussion section. In this section, we provided proper justification for the approaches used and discussed the uncertainties and limitations in the study.
CM: From p16, l1 to p18, l34

5) At the end of the Summary and Conclusions, it is stated that "while this study uses data for one weather radar in arriving at its conclusions, preliminary analysis suggests the problems noted here to be generic." If there are additional "preliminary" results from some should be presented. This statement is important but is hanging.
AC / CM: This was based on initial evaluation using X-band radar data, but since these data are not yet finally processed by the Norwegian Meteorological Institute, we decided to remove this statement.

6) In various locations throughout the text, sentences are added that foreshadow the next or subsequent sections. These are not necessary and should be removed. The meshing of meteorological data with the radar data on a 1 x 1 km grid, including the interpolation of the station data is confusing. This is the four paragraphs on page 8 line 23 through page 9 line 21. This section needs to be rewritten, as it is unclear what is done. Perhaps a table could be added that describes the four datasets (T, RH, wind, and radar). From the text, I assume that temperature and RH data have been gridded at a 1-km resolution from the station data: T using the Optimal Interpolation in a Bayesian setting (Lussana et al., 2016) and RH using the nearest neighbor. What is the Lussana et al. (2016) method? While wind data are available, they are downscaled from a 10 km numerical model dataset. What numerical model dataset is used? This section needs to provide more details - it does not have to be much longer, the methods just need to be clarified. Also, the gauge precipitation (Pgauge) data are used as point measurements meshed with the gridded data; this would also be included in the aforementioned table. These data are revisited in section 4.1.
AC: We have removed the foreshadowing sentences where appropriate.

Regarding the other comments, the work presented in this paper uses 68-gauge locations and not grid locations. Since all gauges do not have air temperature, wind speed and relative humidity measurements, we used hourly gridded (1km x 1km) datasets from the Norwegian Meteorological Institute to generate air temperature and wind speed time series at the gauge locations. However, we do not have access to hourly gridded relative humidity (RH) data for the study area, but we do have measurements from 25 gauge locations. We used these gauge measurements where they were available at the location of the precipitation gauge, and for those gauges with missing RH we used measured RH from the closest gauge.

The hourly gridded (1 km x 1 km) air temperature and wind speed datasets were generated at the Norwegian Meteorological Institute. Lussana et al. (2016) spatially interpolated the past observed air temperature records from meteorological stations to develop the hourly gridded temperature dataset for Norway. They used the Optimal Interpolation method in a Bayesian setting (Lussana et al., 2016). Norwegian Meteorological Institute derived the hourly gridded (1 km x 1 km) wind speed dataset by statistical downscaling from the 10 km numerical dataset, “NORA10” (documented in Reistad et al. (2011)) and “AROME” 2.5 km (documented in Müller et al. (2017)) using a local quantile regression (Lussana 2018, personal communication, 18 July).

As reviewer suggested, the following table (Table I) is added to describe the datasets used in the study and we rewrite the text to make this section clearer.

Table I: Different dataset used in this study and their source and spatial distribution

Description	Gauge Precipitation	Radar precipitation	Air Temperature	Wind Speed	Relative Humidity
Spatial Distribution (Gridded / Gauge Locations)	At gauge locations	Gridded 1kmx1km	Gridded 1kmx1km	Gridded 1kmx1km	At gauge locations
Data Source	Gauge measurements	Radar measurements	Gauge measurements (spatially interpolated)	Downscaled from the numerical model (“NORA10" and AROME)	Gauge measurements

CM: From p5, l3 to p7, l28, section 2.2

The font size is too small in most figures and the text is often grey. This makes the figures difficult to read. This should be addressed throughout.
AC / CM: Font size is increased in the figures.

Specific Comments:
Page 1, line 1: I suggest saying “In colder climates ...”
AC / CM: The sentence has been deleted in the revised manuscript.

p1, l1 and l2, p2 l1, etc.: be consistent with “form of precipitation” and “state of precipitation.” I suggest calling it “phase of precipitation” throughout the text.
AC / CM: Text is updated and the term “phase of precipitation” is used throughout the manuscript.

p1, l5: “estimate” or “adjust”?
AC / CM: The sentence has been deleted in the revised manuscript.

p1, l11: usually “catch error” is called “undercatch”
AC / CM: The term “undercatch” is used throughout the manuscript.

p1, l15: do you mean gauge air “temperature” or temperature at the radar measurement height?
AC: In this study, we used the air temperature at the gauge level. Text is updated to state this.
CM: p1, l16

p1 l15 and subsequently: to be more specific, use “warmer” than instead of “above” when referring to air temperatures. “Above” implies an altitude above the ground, which is typically associated with a colder air temperature. p2, l25: use “colder than” instead of “below,” etc.
AC: Text is reworded.
CM: p1, l16

p1, l15-16: the end of the sentence "which indicates that the partial dependence of precipitation on air temperature is most important for colder climates alone" is unclear. Please reword.
AC: Text is reworded.
CM: p1, l15-17

p1, l22: should "2010b" be "2010a?" Check this, as (Villarini and Krajewski, 2010a) has not appeared yet.
AC: It is corrected.
CM: p1, l22

p2, l13: why "Conventionally?" *Use another word so that the reader does not confuse radar types, such as "the original way" (i.e., conventional), Doppler, dual-polar, multiwavelength.*
AC: "Standard approach" can be a right word here. Text is reworded.
CM: p2, l9

p2, l18: since Canada is mentioned here (Crozier et al., 1991) could be add to the citation list on line 20
AC: The citation (Crozier et al., 1991) is added.
CM: p2, l18

p2, l25: add an "s" to "quarter"
AC: The sentence has been deleted in the revised manuscript.

p3, l2: "different temperatures cause different shapes of crystals." For solid precipitation, i.e., snow, the degree of super-saturation also affects the crystal shape.
AC: We acknowledge that degree of supersaturation also affects the crystal shape. Text is updated.
CM: p3, l4

p3, l5: what is meant by "multiple snow types?" Does this imply shapes? If so, state this explicitly.
AC: Snow type refers not only the shape of the snowflakes but also the particle density (Saltikoff et al., 2015). Text is updated.
CM: p3, l7-8

p3, l8: in many cases the correlation between probability of snow and temperature is an "'S' shaped structure," (see Fassnacht et al., 2001 for a summary illustration), but a simpler linear relation has also been used (e.g., Fassnacht et al., 2013).
AC: We agree with the reviewer that a simple linear relation has also been used but we mentioned the general pattern. Text is updated to include linear relation information.
CM: p3, l11

p3, l9-10: be specific with "the dielectric property of solid particles (ice) is very different from liquid particles (water)." "Very different" is vague.
AC: The text is reworded.
CM: p3, l12

p3, l12: reverse the order of the Hasan et al. (2016) references. You present 2016b before 2016a.
AC: It is corrected.
CM: p3, l29

p3, l13: change the word "Historical"
AC: Historical is changed with "past observed".
CM: p3, l31

p3, l25-29: delete the sentences in the rest of the paragraph starting with "the rest of the paper is structured as follows." You do not need to tell what the sections of the paper are, that reads like the end of a thesis. Instead, give us specific objectives to investigate or research questions that are answered.
AC: As per the reviewer's suggestion, we delete the sentence and the text is updated to list specific objectives of the work. Also see response above.
CM: p4, l1-8

p3, l28: The results should be presented, then there should be a separate Discussion section.
AC: As reviewer suggested, we added a separate Discussion section.
CM: From p16, l1 to p18, l34

p4, l27: there has also been some work on phase discrimination using multiple radar wavelengths (e.g., Al-Sakka et al., 2013).
AC: Thank you for pointing to a relevant work. We use and cite this paper in the revised manuscript.
CM: p18, l17 and l22

p4 l29 to p5 l3: This paragraph can be reduced to 1-2 sentences, as this information is generally known.
AC: We update the paragraph to make it more succinct.
CM: p3, l15-19

p5, l12-13: please reconsider "nonparametric approaches ... weakness is that the method is sensitive to outliers." I am not sure that this is correct. Parametric approaches tend to be sensitive to outliers.
AC: Nonparametric approaches result in "local" biases as a result of outliers but the effect on global attributes is limited. On the other hand, parametric alternatives can be impacted globally due to biases in the estimated parameters. We reword the sentence in the revised manuscript.
CM: p3, l28-29

p5, l18-20: these two foreshadowing equations are not necessary.
AC / CM: The two foreshadowing sentences have been removed in the revised manuscript.

p7, l4: "classification of precipitation phase at gauge level" is good, but don't we need the phase of precipitation at the radar height to select the appropriate radar Z-R equation? Although this is what Koistinen et al. (2004) and Saltikoff et al. (2015), it doesn't necessarily make it correct.
AC: We agree with the reviewer that phase of precipitation at radar measurement height can be different from estimated phase at gauge height. We responded above in detail.

Figure 1: a) I assume that the "length of the observations" is the number of hours with precipitation? b) I also assume that the hypsometry curve is cumulative % of stations below the specified elevation. Please be specific. c) the font size is small and difficult to read. Enlarge and also don't use grey. d) are the red names local cities? Are they important? If so, move them so they are legible.
AC: a) Yes, it is the number of hours with precipitation b) Yes, elevation of gauge locations, c) Font size is increased to make them readable, d) Yes, the cities' names are not important, and the figure and text are updated.
CM: p4, Fig. 1

p7, l12-13: What is the "accumulated hourly radar precipitation rate product?" Is this accumulated from sub-hourly to yield an hourly total, or is the hourly data added?
AC: This is accumulated from sub-hourly to yield an hourly total.
CM: p5, l24-25

p7, l14: tell us how many gauges in the "a relatively dense network of precipitation gauges."
AC: We have changed the text in the revised manuscript.
CM: p5, l27

p8, l26: instead of "are" use past tense through the methods.
AC / CM: It is corrected.

p8, l34: reword the last sentence "However, we used data from all available precipitation gauges for this study." Perhaps state something about the total number of gauge hours of data used (likely in the order of 100,000 gauge-hours).
AC: Nearly 103000 total gauge hours were used in this study. Text is updated with total gauge hours.
CM: p7, l14

p9, l2: provide a source for the "gridded hourly temperature and wind speed dataset"
AC: The source for gridded hourly temperature is observations from the Norwegian meteorological stations (refer p9, l7-8.)
The source for gridded wind speed data is the "NORA10" (documented in Reistad et al. (2011)) and "AROME" 2.5 km (documented in Müller et al. (2017)).
Text is updated with source for gridded wind speed data. See response to comment 6) above for more detail.
CM: p6, Table 1

p9, l6: delete "more details on the procedure adopted for catch correction are provided in the next sub-section."
AC/ CM: This sentence is deleted.

p9, l10: change "resulted" to "resulting"
AC: "resulted" is replaced with "resulting".
CM: p6, l11-12

p9, l29: how little is "intensities below 0.1 mmh-1 contributes little?"
AC: The percentage (quantity) is not mentioned in the cited work (Engeland et al., 2014). However, analysis of the data used in this study showed that intensities lower than 0.1 mmh^{-1} and greater than 0.05 mmh^{-1} are nearly 10 % of the total data above 0.05 mmh^{-1}.
CM: p7, l8-10

p10, l2-8: this is background. It could be moved to earlier in the text, as this is the methods/data section. Tell us what was done. This sentence could also be deleted.
AC / CM: As reviewer suggested, we have deleted the sentence.

p10, l4: the word "Nordic" is not necessary here, as it could also be relevant in southern environments
AC / CM: We agree. "Nordic" has been deleted in the revised manuscript.

p10, l4-5: the end of the sentence is redundant "due to large catch errors for snow." In could state that it is "due to high wind conditions." It is wind that causes undercatch, not "large catch errors"
AC: Text is reworded.
CM: p7, l15-16

p10, l7: "Wolff et al., 2015" is not in the citation list
AC: References is updated with "Wolff et al. (2015)"
CM: p22, l35

p10, l9 or previous: what type of precipitation gauge and what type of shield are used? This is very important information to assess the degree of undercatch and the error associated with the undercatch correction.
AC: "The gauges in the study site consists of "Geonor" type and tipping bucket. Both types are with an Alter wind shield. This information is added to the manuscript.
CM: p5, l27-28

p10, l9-17: throughout this paragraph it is "undercatch" correction. This should be consistent, as there are other errors.
AC: We agree. Text is updated.
CM: p7, l20-28

p10, l13-14: "It was found that correlation between the corrected precipitation by using measured wind speed data (15-gauge locations) and gridded data are over 0.97..." Does this mean the correlation undercatch correction using gauge wind speed versus the downscaled gridded wind speed?
AC: Yes, the correlation between undercatch corrected precipitation using gauge wind speed versus undercatch corrected precipitation using the downscaled gridded wind speed. This was done to verify that gridded wind would provide a realistic correction compared to wind measured at the site.

p10, l14: there are only 15-gauge locations with wind speed measurements. How are the other 53 precipitation gauges corrected for undercatch? From my comment above (p10, l13-14), I assume that the downscaled gridded wind speed was used to correct for undercatch at all 68 precipitation gauges. This is not clear.
AC: The downscaled gridded wind speed was used to correct for undercatch at all 68 precipitation gauges. To control the result of correcting with gridded wind speed, we compared the corrected precipitation using gridded wind speed with the 15 locations where we had wind speed measurements at the gauge site.
CM: p7, l23-26

p10, l19-22: delete this paragraph. We know what you are going to do Figure 2: add tick marks to the y-axis.
AC: The foreshadowing paragraph is deleted. Tick marks are added to the y-axis.
CM: p10, Fig. 2

p10, l23 through p11: Are the "Partial weight of predictors" constant over time, i.e., is there a specific value for station that does not change?
AC: No, we did not use a fixed value for any given station. The value varies with the data.

Table 1: As this is a summary of Figure 2, this table could be converted to two horizontals box and whisker plots on Figure 1. The reader doesn't really care about the specific numbers, just the range.
AC: As the reviewer suggested we have plotted box and whisker plot (refer Fig. I in Appendix). However, we feel Table is a better representation of the information we wish to convey and we prefer to keep the Table 1 in the revised manuscript.

p12, l1 and elsewhere: the word "prediction" implies that this is for the future. I suggest using "estimation" throughout.
AC: "Prediction is replaced with "estimation".
CM: p10, l4

Figure 3: the caption is confusing; break into two sentences. Also, are these "length of the data (circle size)" the same as in Figure 1? If so, then don't display again here, unless this is relevant later?
AC: The caption is reworded.
Yes, Length of the data (circle size) is same on both Fig. 1 and Fig. 3. Fig.3 is updated.
CM: p12, Fig.3

p13, l4-7: these three sentences could be reduced to a small histogram of RMSE reductions that is added to figure 3.
AC: As the reviewer suggested, we have plotted a histogram (refer Fig. II in Appendix). However, we prefer to keep these sentences for simplicity in the revised manuscript as well.

p12 to 15: why was a single average (mean) partial weight (beta P = 0.68) used in the k-nn prediction model, when it was shown (Figure 2 and Table 1) that the partial weight varies from 0.4 to 1?
AC: We responded above in detail.
However, the use of station specific partial weight can be more precise when sufficient past observed data are available at each gauge location.

p15 section 5.3: all this text except for the last sentence is background or methods and should be moved to an appropriate location earlier in the paper.
AC: We think that the first paragraph of section 5.3 present information on the data that is relevant for the understanding of the low intensity analysis and figure 5. We would therefore rather keep this section as it is in the paper.

p15, l15: does this "still significant" has a statistical meaning? If so, explain how. If not don't use the word significant.
AC: Yes, the results are statistically significant as the RMSE was estimated using leave one out cross validation (LOOCV). We clarify this in the revised manuscript.
CM: p13, l6-8

p15, section 5.4: what is the range of the "Temperature Classes?" You only discuss T > 10C. What about T < 10C? The point of this section is unclear.
AC: The results for colder and warmer than 10° C has been added and the paragraph has been updated.
CM: p14, l1-5, and Fig.2 in the Supplementary material

Figure 6: what are the dots above the solid and mixed phase?
AC: The dots symbolise outliers, values outside 1.5 * IQR which is represented by the whiskers. An explanation is added to the figure caption.
CM: p16, Fig. 6

p18, section 5.6: these statements seem to be hanging. Can you present specifics?
AC: We did a test on uncorrected gauge precipitation data (not corrected for wind induced undercatch) during an early phase of the study and found that air temperature works as a covariate also there. For this assessment, we used the uncorrected gauge precipitation at 88 gauge locations. More than 80 % of the precipitation gauge locations in the study area showed clear improvement. The intent of section 5.6 was to make this point clearly. However, in order to avoid lengthening the paper with more results, we decided to remove section 5.6 in the revised manuscript and merge the statements into other sections.
CM: p18, l25-29

p18, l5: what does "the use of temperature as an additional predictor variable is having consistent impact" mean? The words "consistent impact" are not clear.

AC: We mean by “consistent impact” that the study with uncorrected gauge precipitation as described above resulted in similar results (resulted in partial weight for air temperature and improvement in RMSE with air temperature as an additional predictor) as like the study with corrected gauge precipitation.
CM: p18, l25-29

References

Al-Sakka, H., Boumahmoud, A.-A., Fradon, B., Frasier, S. J. & Tabary, P. 2013. A New Fuzzy Logic Hydrometeor Classification Scheme Applied to the French X-, C-, and S-Band Polarimetric Radars. *Journal of Applied Meteorology and Climatology,* 52**,** 2328-2344.

Crozier, C., Joe, P., Scott, J., Herscovitch, H. & Nichols, T. 1991. The king city operational doppler radar: Development, all-season applications and forecasting. *Atmosphere-Ocean,* 29**,** 479-516.

Fassnacht, S., Soulis, E. & Kouwen, N. 1999. Algorithm application to improve weather radar snowfall estimates for winter hydrologic modelling. *Hydrological processes,* 13**,** 3017-3039.

Fassnacht, S. R., Kouwen, N. & Soulis, E. D. 2001. Surface temperature adjustments to improve weather radar representation of multi-temporal winter precipitation accumulations. *Journal of Hydrology,* 253**,** 148-168.

Koistinen, J., Michelson, D. B., Hohti, H. & Peura, M. 2004. Operational Measurement of Precipitation in Cold Climates. *In:* MEISCHNER, P. (ed.) *Weather Radar: Principles and Advanced Applications.* Berlin, Heidelberg: Springer Berlin Heidelberg.

Lussana, C., Ole, E. T. & Francesco, U. 2016. seNorge v2.0: an observational gridded dataset of temperature for Norway. Norway: Norwegian Meteorological Institute.

Müller, M., Homleid, M., Ivarsson, K.-I., Køltzow, M. a. Ø., Lindskog, M., Midtbø, K. H., Andrae, U., Aspelien, T., Berggren, L., Bjørge, D., Dahlgren, P., Kristiansen, J., Randriamampianina, R., Ridal, M. & Vignes, O. 2017. AROME-MetCoOp: A Nordic Convective-Scale Operational Weather Prediction Model. *Weather and Forecasting,* 32**,** 609-627.

Reistad, M., Breivik, Ø., Haakenstad, H., Aarnes, O. J., Furevik, B. R. & Bidlot, J.-R. 2011. A high-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea, and the Barents Sea. *Journal of Geophysical Research: Oceans,* 116.

Saltikoff, E., Lopez, P., Taskinen, A. & Pulkkinen, S. 2015. Comparison of quantitative snowfall estimates from weather radar, rain gauges and a numerical weather prediction model. *Boreal Env. Res,* 20**,** 667-678.

Tveito, O., Førland, E., Heino, R., Hanssen-Bauer, I., Alexandersson, H., Dahlström, B., Drebs, A., Kern-Hansen, C., Jónsson, T. & Vaarby Laursen, E. 2000. Nordic temperature maps. *DNMI report,* 9.

Tveito, O. E. & Førland, E. J. 1999. Mapping temperatures in Norway applying terrain information, geostatistics and GIS. *Norsk Geografisk Tidsskrift-Norwegian Journal of Geography,* 53**,** 202-212.

Wolff, M. A., Isaksen, K., Petersen-Øverleir, A., Ødemark, K., Reitan, T. & Brækkan, R. 2015. Derivation of a new continuous adjustment function for correcting wind-induced loss of solid precipitation: results of a Norwegian field study. *Hydrol. Earth Syst. Sci.,* 19**,** 951-967.

Appendix

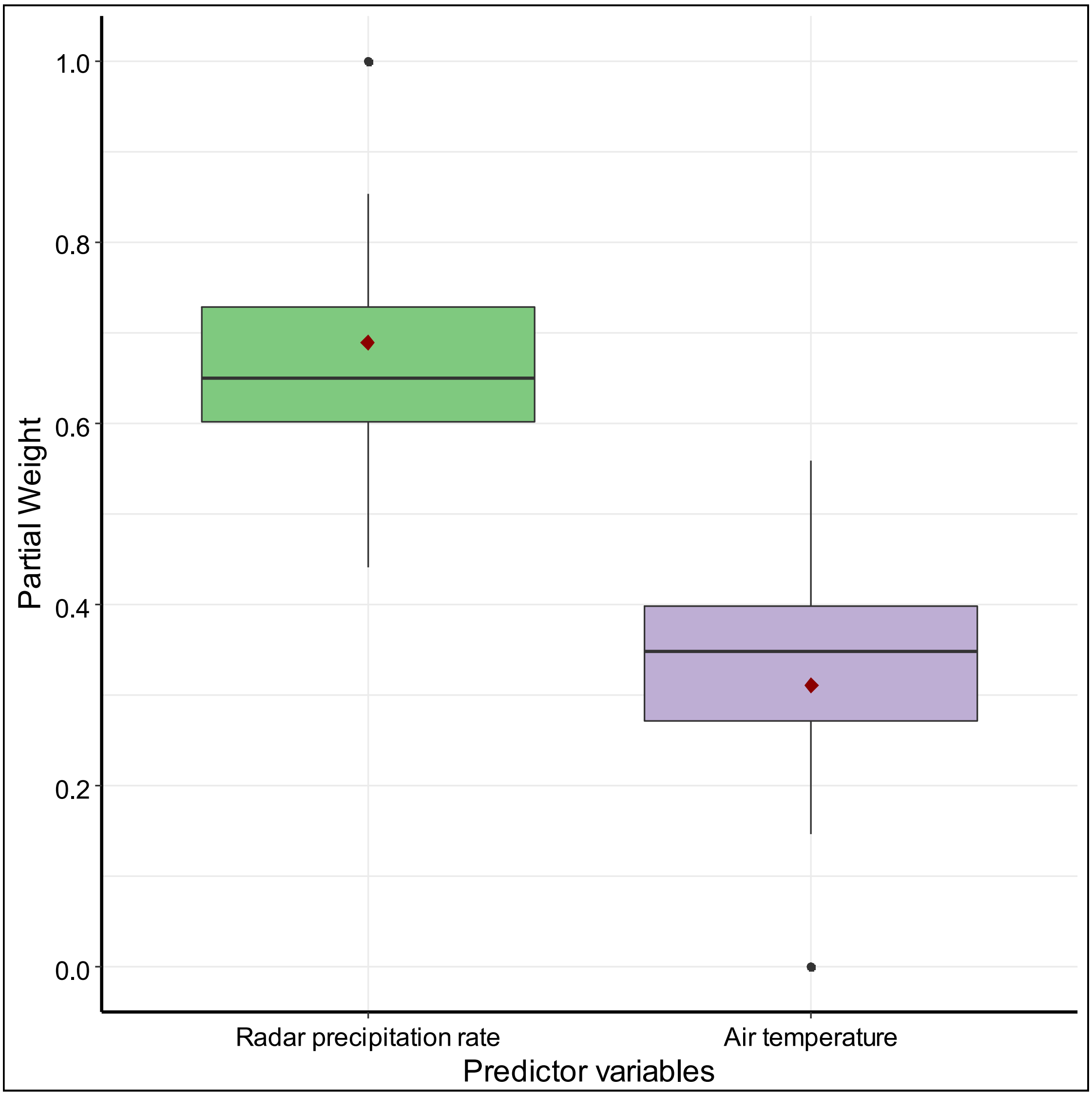

Figure I. Box and whisker plot of estimated partial weight of predictor variables (Radar precipitation rate and air temperature) at 68 gauge locations in the study area. The summation of partial weights is equal to 1.

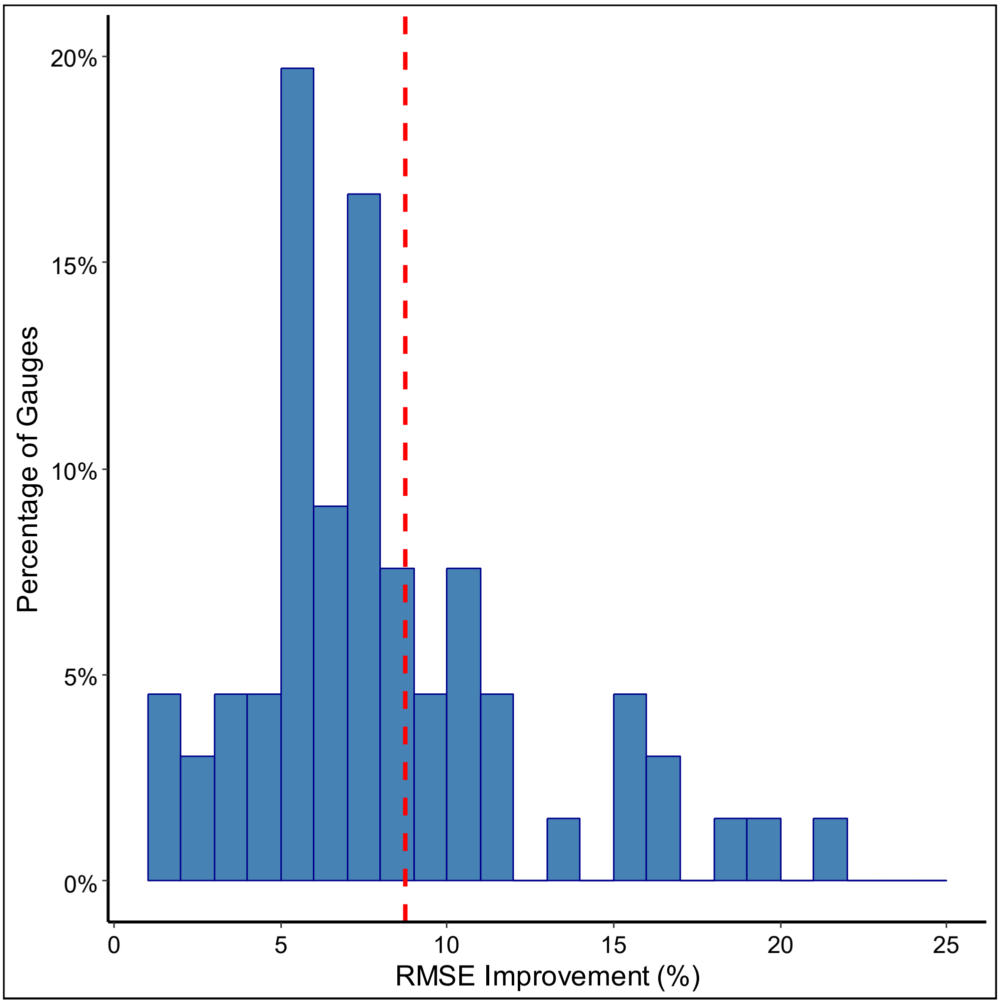

Figure II. The Percentage of precipitation gauge locations against percentage improvement in RMSE with air temperature as an additional predictor variable at those gauge locations and the mean RMSE improvement percentage (red dash line) for gauge locations (68 gauges) in the study area.

Response to the review of hess-2018-0351

RC2: Responses to Anonymous Referee 2

The authors wish to thank the reviewer for his constructive comments and corrections to the discussion paper. In the following, we have responded to each of the comments from the reviewer and showed the page and line numbers of the revised manuscript if any changes. The comment from the reviewer (RC) is in italic font while the author comment (AC) and changes in the manuscript (CM) are in blue normal font.

General comments:

The authors apply the non-parametric k - nearest neighbour method (k-nn) to estimate radar precipitation from gridded surface observations of rainfall and temperature for the Oslo region in Norway. They show that utilising temperature as second predictor variable reduces the root mean squared error significantly compared to a k-nn model without temperature and compared to the original procedure using a constant Z-R relationship or separate snow/rain Z-R relationships.

The application of this method for radar rainfall estimation including temperature is novel and of interest not only for readers living in regions with colder climates. The research is done systematically and quite carefully. The paper is written well and clear in structure. However, there are three major points and some minor things which need attention before the paper can be published. One main point are the lengthy introduction and background sections which could be shortened. A second important point concerns the method to estimate the partial weights. It becomes not clear, that this method is really providing optimal weights. And, third, there seems to be an issue with the back-calculation of Z using the inverse Z-R relationship on a different time resolution as for the original forward calculation. Detailed information about this and the minor things are given below.

Detailed comments:

1. Sections 1 and 2: Both sections together cover almost 4 pages and represent the introduction with the state of the art. This is quite lengthy. The introduction is very general; the background is more focussed on the topic at hand. I would suggest to shorten these parts especially the introduction significantly and may be use the background as introduction.

AC: We have merged the Introduction and background sections as suggested and updated the text to make the introduction more succinct (nearly 2 pages now).

CM: p1 to p4, section 1

2. Eq. 1: As predictor R(t) is used. Why not using Z(t) as predictor? For R(t) already a (wrong) Z-R-relationship has been applied, introducing great uncertainty. If a linear relationship is required a log-log transformation of Z(t) and Rest(t) could be applied beforehand. This needs at least to be discussed.

AC: In the methodology presented in the paper, reflectivity (dBZ) could in principle be used instead of radar precipitation rate as shown by Hasan et al. (2016) for the univariate case. As we do not have access to the reflectivity (Z(t)) data from the Hurum radar for this study, we had to use the hourly radar precipitation rate which is available from the Norwegian Meteorological Institute. While the reflectivity can be back-calculated by inverting the algorithm that was used operationally, we feel this may add additional uncertainty and would not matter given the regression algorithm being used is nonparametric. Further, it can be noted that one key purpose of this work is to see how we can improve the radar precipitation rate data available to us as a finished product (hourly Surface Rainfall Intensity (SRI) product) from the meteorological institute.

The text is updated in the revised manuscript to clarify this.

CM: In the Discussion, from p16, l9 to p17, l2

3. Fig. 1: The units for observation length and elevation are missing. Also, the text of the legend is tiny and hard to read.
AC: The units are added, and the font size of the text is increased in Fig.1.
CM: p4, Fig. 1

4. Section 5.1: It is not clear if the estimation of the partial weights using partial information correlation (PIC) is really beneficial or even optimal. In order to prove the merit of PIC I would suggest to test two additional cases a) equal weights for P and T and b) using simple linear partial correlations. The performance for the latter two cases measured by RMSE should be worse than by PIC weighing.
AC: The partial informational correlation (PIC) provides a generic measure of statistical dependence of predictors of a general linear or nonlinear system. Estimation of partial weight using PIC shows the partial dependence of radar precipitation estimation on air temperature.
Earlier papers have shown that the estimated PIC and weights collapse to what would be estimated using a linear regression model if the system is linear (Mehrotra and Sharma, 2006, Sharma and Mehrotra, 2014, Sharma et al., 2016). As the system here is nonlinear, our approach of using PIC to estimate partial weights appears more justified.
After receiving the reviewer's comment, we tested our approach using equal weights. We found that the gain in RMSE was not significant with the use of PIC based partial weight compared to equal weights, but the mean error was reduced when we used partial weight estimated using PIC. The manuscript is updated to discuss this.
CM: p17, l10-17

5. Fig. 4: This bar plot is not easy to read. I would suggest to use box-whisker plots instead.
AC: As the reviewer suggested we have added box and whiskers plot to the revised manuscript. However, bar plot presents the results at each gauge location compared to a lumped box plot which we find interesting to report so we would like to add the bar plot to the supplementary information.
CM: p13, Fig. 4 in the revised manuscript and Fig. 1 in the Supplementary material.

6. Page 16, line 1: The back-calculation of Z from R using a non-linear relationship on hourly data gives an estimated average Z value for each hour. This estimate can be quite different from the observed average Z value if the rainfall distribution within the hour is not unique. In the forward calculation the Z-R relationship is applied on 7.5 min Z values to calculate 7.5 minute rainfall intensities. Because of the non-linearity of the Z-R relationship a simple back calculation on a different time step than the one the original calculation was applied is not possible. For non-linear functions f is E[f(x)] < > f[E(x)].
AC: We do agree with the reviewer that back calculated reflectivity on a different time step (hour) than the original calculation is not same as the average value unless the precipitation is even within the hour and we fully acknowledge that this introduces uncertainties in the results. As mentioned above, we do not have access to reflectivity data (or precipitation rates with original short time resolution). In order to compare our proposed nonparametric radar precipitation estimation with radar precipitation rates computed using separate equations for snow and rain, we decided to back calculate the reflectivity from the data available to us.

We do have an update to the above statement. The hourly SRI product is based on corrected reflectivities with a time resolution of 15 minutes (before June 2013) and 7.5 minutes (after June 2013). The VPR corrected reflectivity data used to produce the accumulated hourly radar precipitation (SRI product) used in the study are not stored in the production process, and therefore not available at met.no (Elo 2018, Personal communication). However, the Plan Position Indicator (PPI) of the lowest elevation beam from Hurum radar is available with the original short time resolution from met.no. We have therefore used the PPI data to redistribute the hourly data used in the study by assuming precipitation intensity distribution within each hour is the same in the SRI data as in the PPI dataset. The redistributed precipitation rates with original time resolution (15 or 7.5 minutes depending on the year) were then converted to reflectivities using an inverse of the Marshall and Palmer equation.

The back calculated reflectivity was converted to precipitation rate using the separate snow and rain equation according to the computed phase. The precipitation rates were then accumulated to hourly time resolution and we did the comparison as in the original manuscript.

Here, it should be noted that there is uncertainty in how accurately the redistributed intensity distribution of SRI represents the original distribution, however, this exercise at least used a possible realistic distribution.

In the revised manuscript, we present this procedure of back calculation using PPI data and we discuss the limitation in the comparison.

CM: p5, l7-8 in section 2.2
From p14, l12 to p15, l11 in section 3.5
p18, l2-15 in section 4

References

Hasan, M. M., Sharma, A., Johnson, F., Mariethoz, G. & Seed, A. 2016. Merging radar and in situ rainfall measurements: An assessment of different combination algorithms. *Water Resources Research,* 52**,** 8384-8398.

Mehrotra, R. & Sharma, A. 2006. Conditional resampling of hydrologic time series using multiple predictor variables: A K-nearest neighbour approach. *Advances in water resources,* 29**,** 987-999.

Sharma, A. & Mehrotra, R. 2014. An information theoretic alternative to model a natural system using observational information alone. *Water Resources Research,* 50**,** 650-660.

Sharma, A., Mehrotra, R., Li, J. & Jha, S. 2016. A programming tool for nonparametric system prediction using Partial Informational Correlation and Partial Weights. *Environmental Modelling & Software,* 83**,** 271-275.

Response to the review of hess-2018-0351

RC3: Responses to G. Ravazzani (Referee 3)

The authors wish to thank the reviewer for his constructive comments and corrections to the discussion paper. In the following, we have responded to each of the comments from the reviewer and showed the page and line numbers of the revised manuscript if any changes. The comment from the reviewer (RC) is in italic font while the author comment (AC) and changes in the manuscript (CM) are in blue normal font.

General comments:

In this paper, a non-parametric method is applied to estimate radar precipitation considering both rainfall and temperature. The use of radar for precipitation estimation is an interesting topic. Many papers have been presented about this topic, but the specific problem authors deal in this paper is how to assess solid precipitation in cold regions. The solution they propose is of interest for cold climates in northern Europe, of course, but I suppose it could be extended to other areas where solid precipitation occurs.

Specific comments:

Authors used 68 rain gauges in this study that are clustered around urban areas. Do authors think that this uneven distribution may affect results? In other terms, is the location of raingauges relevant for the application of the proposed procedure?

AC: The method used is independent of the gauge locations, and the computed estimates are ascertained for each gauge individually.

P 9 L 14 "The gridded hourly wind speed datasets are derived from a statistical downscaling of a 10 km numerical model dataset onto a 1 km grid". Did authors verify how the method is sensible to the specific realization of the statistical downscaling?

AC: The Norwegian Meteorological Institute derived the hourly gridded (1 km x 1 km) wind speed dataset by statistical downscaling from the 10 km numerical dataset, "NORA10" and we used this in this study. We did not evaluate their method of downscaling. However, as described on p10, l12-14 (discussion paper), to control the result of correcting with gridded wind speed, we compared the corrected precipitation using gridded wind speed with the 15 locations where we have wind speed measurements at the gauge site.

CM: p6, l14-16 and p7, l23-26

Authors apply correction to gauge precipitation to consider wind induced underestimation. Gauge precipitation is affected by several sources of uncertainty. Wind is of course relevant, but another systematic error is related to the calibration of raingauges that causes underestimation for high rainfall intensity and overestimation for low rainfall intensity. Further uncertainty arises when solid precipitation has to be measured. How did authors deal with these errors? Are they already managed by the meteorological institute?

AC: Norwegian Meteorological Institute manages the calibration of raingauges and takes necessary measures to reduce the uncertainty that arises when solid precipitation has to be measured. Further, data from the raingauges are gone through routine quality control before being released to the public through the database. However, the meteorological institute does not do wind induced undercatch correction for their precipitation data.

CM: p5, l29-32

Section 5.6 is very short compared to the rest of the paper and I did not fully understand what is the intention of authors. I think they should better explain this part or remove it.

AC: we did a test on uncorrected gauge precipitation data (not corrected for wind induced undercatch) showing that temperature works as a covariate also there. The intent of section 5.6 was to make this point clearly. However, in order to avoid lengthening the paper with more results, we decided to remove section 5.6 (discussion paper) in the revised manuscript and merge the above statement into other sections.
CM: p18, l25-29

Technical corrections:
P.4 L. 6 The Finnish Meteorological Institute
AC: It is corrected.
CM: p2, l22

Estimating Radar Precipitation in Cold Climates: The role of Air Temperature within a Nonparametric Framework

Kuganesan Sivasubramaniam[1], Ashish Sharma[2], and Knut Alfredsen[1]

[1]Department of Civil and Environmental Engineering, Norwegian University of Science and Technology, 7491 Trondheim, Norway

[2]School of Civil and Environmental Engineering, University of New South Wales, Sydney, NSW2052, Australia

Correspondence to: Kuganesan Sivasubramaniam (kuganesan.sivasubramaniam@ntnu.no)

Abstract. The use of ground based precipitation measurements in radar precipitation estimation is well known in radar hydrology. However, the approach of using gauged precipitation and near surface air temperature observations to improve radar precipitation estimates in cold climates is much less common. In cold climates, precipitation is in the form of ~~precipitation (~~snow or rain or a mixture of ~~snow and rain) results in uncertainty in radar precipitation estimation. Estimation often proceeds without distinguishing the state of~~ the two phases. Air temperature is intrinsic to the phase of the precipitation and could therefore be a possible covariate in the models used to ascertain radar precipitation ~~which is known to impact the radar reflectivity – precipitation relationship~~estimates. In the present study, we investigate the use of air temperature within a non-parametric predictive framework to improve radar precipitation estimation for cold climates. ~~Compared to radar reflectivity - gauge relationships, this approach uses gauge precipitation and air temperature observations to estimate radar precipitation.~~ A nonparametric predictive model is constructed with radar precipitation rate and air temperature as predictor variables, and gauge precipitation as an observed response using a k-nearest neighbour (k-nn) regression estimator. The relative importance of the two predictors is ascertained using an information theory-based ~~rationale~~weighting. Four years (2011-2015) of hourly radar precipitation rate from the Norwegian national radar network over the Oslo region, hourly gauged precipitation from 68 gauges, and gridded observational air temperature were used to formulate the predictive model and hence make our investigation possible. Gauged precipitation data were corrected for wind induced ~~catch error~~ undercatch before using them as true observed response. The predictive model with air temperature as an added covariate reduces root mean squared error (RMSE) by up to 15 % compared to the model that uses radar precipitation rate as the sole predictor. More than 80 % of gauge locations in the study area showed improvement with the new method. Further, the associated impact of air temperature became insignificant at more than 85 % of gauge locations when the ~~temperature was above~~ near surface air temperature was warmer than 10° C, which indicates that the partial dependence of precipitation on air temperature is most ~~important for colder climates alone~~useful for colder temperatures.

1 Introduction

Hydrological applications require accurate precipitation estimates at the catchment scale ~~. Use of point precipitation gauges often proves inadequate in representing the spatio-temporal variability in the precipitation field~~ (Beven, 2012; Kirchner, 2009).

Weather radars provide quantitative precipitation estimates over a large area with high spatial and temporal resolution. However, weather radars measure the precipitation rate indirectly, using the energy scattered back by hydrometeors in the volume illuminated by a transmitted electromagnetic beam (Villarini and Krajewski, 2010a). The backscattered energy is measured as reflectivity which is used to estimate precipitation ~~. This measured reflectivity depends on many factors such as size, shape, orientation (if non-spherical), state and concentration of particles in the radar illuminated volume in the atmosphere along with their dielectric properties (Hong and Gourley, 2015; Joss et al., 1990).~~ (Hong and Gourley, 2015).

The nature of radar precipitation measurements is subject to many sources of error. ~~These errors occur during the sampling or measurement of reflectivity as well as in the process of converting the reflectivity (Z) to precipitation rates (R) (Chumchean et al., 2006).~~ Some of the known errors in the reflectivity measurement are ground clutter, beam blocking, anomalous propagation, bright band, hail, and attenuation (Berne and Krajewski, 2013; Chumchean et al., 2003). During the conversion, the use of inappropriate ~~Z-R~~ Z - R relationship leads to ~~Z-R~~ Z - R conversion error. Due to the presence of such significant errors (both random and systematic), radar data are still not widely used in hydrological applications ~~as broadly and efficiently as they could be~~ (Berne and Krajewski, 2013; Chumchean et al., 2003). Many studies (e.g., Abdella, 2016; Villarini et al., 2008; Ciach et al., 2007; Chumchean et al., 2006) have focused on estimating these errors in order to improve quantitative radar precipitation estimates; however, some of the underlying physical processes are still not understood well enough to allow significant advances (Villarini and Krajewski, 2010b).

~~Conventionally~~In the standard approach, radar measurements of reflectivity (Z) are converted into precipitation rate (R) using the parametric Z - R relationship derived by Marshall and Palmer (1948) in the form of a power law, $Z = aR^b$. The variability of the power law parameters (a and b) is related to a number of factors including the drop size distribution (DSD) of hydrometeors. Drop size distribution varies in time and space as well as for the type and the phase of precipitation ~~(Chumchean et al., 2008; Joss et al., 1990; Uijlenhoet, 2001). The~~ (Chumchean et al., 2008; Joss et al., 1990; Uijlenhoet, 2001; Wilson and Brandes, 1979).

In cold climates, precipitation occurs in the form of snow or rain or a mixture of snow and rain. Several studies (e.g., Battan, 1973; Sekhon and Srivastava, 1970; Marshall and Gunn, 1952) have investigated the Z - R relationship ~~is not unique and hence, we depend on empirical relationships instead (Wilson and Brandes, 1979)~~regarding the precipitation phase and proposed different parameter sets. Most radar ~~systems~~ operations in cold climate countries (e.g., Canada and Finland~~etc.~~) use two sets of Z - R relations, one for rain and one for snow, often calibrated in situ to measure a water equivalent radar reflectivity factor ~~(the dielectric constant for water is used) (Koistinen et al., 2004)~~ (Koistinen et al., 2004; Crozier et al., 1991; Smith, 1984). However, the Norwegian radars and European radar project OPERA have used a single Z - R relationship ~~(~~(Marshall and Palmer (1948) relation for rain ~~)~~ ($Z = 200R^{1.6}$)) throughout the year. The use of the single reflectivity-precipitation relationship can result in phase dependent bias in radar precipitation estimation.

~~In cold climates, quantitative radar precipitation estimates are often formulated for warmer summer months and / or specific storm events (Berne and Krajewski, 2013; Saltikoff et al., 2015). Norway and adjacent countries in northern Europe experience cold temperatures below +10 degrees Celsius for nearly three quarter of the year. Continuous runoff simulation is required for water resources management applications such as hydropower production planning, design and operation~~

~~of water infrastructure, flood forecasting and ecological assessments (Hailegeorgis et al., 2016). Today, precipitation runoff models mostly use gauge precipitation measurements for continuous simulation. A continuous timeseries of radar-based precipitation estimates and the reduced reliance on traditional precipitation gauges are of great interest to hydrologists given the spatio-temporal detail that is offered. Further, a single radar can measure precipitation over many small catchments with its extended spatial coverage, which otherwise remain ungauged without any ground precipitation measurement (Berne and Krajewski, 2013). An example of~~ The Finnish Meteorological Institute devised two equations for rain ($Z = 316R^{1.5}$) and snow ($Z_e = 100S^2$) for operational use (Saltikoff et al., 2015). Here Z_e represents the equivalent radar reflectivity factor of snow. For the use of phase dependent reflectivity-precipitation (Z - R) relationship, the precipitation phase of the radar pixel must be estimated. Air temperature has traditionally been used to determine the phase of the ~~usefulness of radar precipitation data for continuous simulation was presented by Fassnacht et al. (1999). If a radar based continuous precipitationtime series is to be generated with the objective of use in hydrological modelling in boreal regions, the effect of varying precipitation phase on radar~~ precipitation ~~estimates must be considered~~(Al-Sakka et al., 2013). The Finnish Meteorological Institute uses temperature and humidity observations from synoptic stations to estimate the precipitation phase and uses that information to apply a different parameter set for rain or snow (Koistinen et al., 2004; Saltikoff et al., 2015). However, Saltikoff et al. (2000) reported that real time phase dependent adjustment of two different parameter sets does not improve the snowfall estimates significantly. To account for varying precipitation phase (multiple snow types and mixture of snow and rain), many parameter sets could be required. Moreover, the precipitation phase changes rapidly even within the single winter storm and hence, operationally, switching between different parameter sets can be a challenging task (Koistinen et al., 2004; Saltikoff et al., 2015).

Fassnacht et al. (2001, 1999) demonstrate the use of surface air temperature to estimate the fraction of snow content in mixed precipitation and use it to adjust the radar estimates for mixed precipitation. It was reported that this adjustment improved the accumulated snow estimates in Ontario, Canada. Further Fassnacht et al. (1999) showed that the adjusted radar data provided more realistic precipitation estimates for precipitation-runoff models than corrected gauge precipitation data.

Starting from its origin and throughout its entire journey, the rain drop or snow crystal is shaped by temperature. During the formation and growth of cloud droplets, different temperatures and the degree of super saturation cause different shapes of crystals to form, and then the crystals start to fall. The falling crystals are then characterised by the temperature of the air through which they fall. As a result, the air temperature determines the final properties and the phase of the hydrometeor that reaches the ground surface (Fassnacht et al., 2001). Further, studies showed that there are multiple snow types ~~and~~ with different shapes and densities and they vary in time, based partially on temperature (Saltikoff et al., 2015). Many studies (Auer Jr, 1974; Kienzle, 2008; Killingtveit, 1976; Rohrer, 1989) examined the relationship between the precipitation phases (snow, rain and mixture of snow and rain) and temperature. The probability of occurrence of snowfall versus temperature shows generally an approximately ‘S’ shaped structure ~~in~~ and in some cases linear relation (Fassnacht et al., 2013) in these studies. ~~Additionally, as mentioned earlier, measured reflectivity depends on dielectric properties of hydrometeors. The dielectric~~ Further, the dielectric property of solid particles (ice) is ~~very different from~~ not the same as liquid particles (water) and moreover, it varies with temperature (Joss et al., 1990). These imply that temperature is intrinsic to both the phase of precipitation and the ensuing conversion of reflectivity into the incident ground precipitation.

~~Hasan et al. (2016a) presented a nonparametric approach to estimate ground rainfall using radar reflectivity as a univariate predictor variable in a tropical setting. Historical radar reflectivity and rain were used in formulating the nonparametric model. In the present work, we tested their method in the context of Norwegian radar precipitation estimation and found it to be sub-optimal especially for the colder precipitation events. Norway is a cold climate country located in the northern high latitudes, experiencing different phases of precipitation. Norwegian radars use a single Z - R relationship throughout the year and the phase of the precipitation is not considered. If phase dependent Z - R relationships were used, they will still require an added algorithm to establish the dominant phase for the event. Based on the discussion in the previous paragraph, the intuition is that air temperature observations can be used together with observed gauge precipitation to adjust radar precipitation in cold climates. Moreover, the nonparametric approach of Hasan et al. (2016a) can be extended to allow use of a bivariate predictor vector with temperature as an additional predictor variable. This forms the basis for the investigation reported here.~~

~~This study set out to investigate the use of air temperature as an additional predictor in the radar precipitation estimation with the objective of improving quantitative radar precipitation estimation for cold climates. Compared to traditional radar-gauge adjustment, the proposed method is based on nonparametric approach using gauge precipitation and air temperature observations to adjust the radar precipitation. The rest of the paper is structured as follows. The following section reviews radar precipitation estimation in cold climates as well as nonparametric methods and their use in radar precipitation estimation. The methods and tools used to formulate the nonparametric model are presented in Section 2.3. Section 2.1 describes the study area and data used to test the method. The results from the study are discussed in Section **??**. Finally, summary and conclusions are presented in Section 5.~~

2 ~~Background~~

1.1 ~~Radar precipitation estimation in cold climates~~

~~In cold climates, precipitation occurs in the form of snow or rain or a mixture of snow and rain. As mentioned earlier, radar operations in cold climates often use two sets of parameters to convert radar reflectivity Z, into precipitation intensity (rain(R) or snow(S), $\mathrm{mm\,h^{-1}}$). Several studies (Battan, 1973; Marshall and Gunn, 1952; Sekhon and Srivastava, 1970) have investigated and then proposed different parameter sets (coefficients "a" and "b" in the power law equation relating reflectivity to precipitation) for rain and snow. The parameter set proposed by Sekhon and Srivastava (1970) has been used as a standard for snow, just as the work by Marshall and Palmer (1948) has been used widely for rain (Fassnacht et al., 2001; Saltikoff et al., 2015) .~~

~~The Finish Meteorological Institute operationally uses their own equations for rain ($Z = 316R^{1.5}$) and snow ($Z_e = 100S^2$) (Saltikoff et al., 2015). Here Z_e represents the equivalent radar reflectivity factor of snow and it is different from Z because the radar signal processing uses the dielectric constant of liquid (water) instead of dielectric constant of solid (ice) for snow. Zhang et al. (2016) used the equation $Z_e = 75S^2$ for the NEXt Generation Radar network (NEXRAD) in the United States~~

~~which offers similarities to the equation by the Finish meteorological institute for snow. However, Saltikoff et al. (2000) reported that real time phase dependent adjustment of two different parameter sets does not improve the snowfall estimate significantly. To account for varying precipitation phase (multiple snow types and mixture of snow and rain), many parameter sets could be required. Moreover, the precipitation phase changes rapidly even within the single winter storm and hence, operationally, switching between different parameter sets can be a challenging task (Koistinen et al., 2004; Saltikoff et al., 2015) .~~

~~For the use of phase dependent reflectivity-precipitation (Z - R) relationship, the precipitation phase of the radar pixel must be estimated. Earlier, weather radar operations in cold climates switched between summer and winter Z - R relationships according to calendar date. However, this is obviously uncertain. As mentioned in the introduction, air temperature can be used to determine the phase (whether snow or rain) of the precipitation. The Finnish Meteorological Institute uses temperature and humidity observations from synoptic stations to estimate the precipitation phase and uses that information to apply a different parameter set for rain or snow (Koistinen et al., 2004; Saltikoff et al., 2015). Fassnacht et al. (2001) demonstrate the use of surface air temperature to estimate the fraction of snow content in mixed precipitation and use it to adjust the radar estimate for mixed precipitation. It is reported that these adjustments improve the accumulated snow estimates in Ontario, Canada. Observations from dual polarised weather radars can also be used to classify precipitation phases (Ryzhkov and Zrnic, 1998) . However, many radars use a single polarity and moreover, even from dual polarised radars, data on phase information are not readily available to end users to help refine their estimation algorithms. Operational use of dual polarised radars in hydrometeor classification has progressed significantly; however, the classification for high latitude winter storms is still challenging (Chandrasekar et al., 2013).~~

1.1 ~~Nonparametric Radar rainfall estimates~~

Parametric (or regression type) and nonparametric approaches (nearest neighbour and kernel density estimation) have been used to build predictive models for a range of applications. ~~When sufficient data are available, nonparametric approaches are efficient alternatives for specifying an underlying model as compared to parametric approaches. Nearest neighbour and kernel density estimation are amongst the most commonly used nonparametric methods. The simplicity of nonparametric approaches have made them attractive for use in hydrology and other sciences (Mehrotra and Sharma, 2006).~~ A key advantage of nonparametric approaches is that less rigid assumptions about the distribution of the observed data are needed (Silverman, 1986) and hence no major assumptions about the process being modelled are required to construct the complete predictive system ~~(Sharma and Mehrotra, 2014). Due to the availability of sufficient radar precipitation rate observations, nonparametric methods provide an attractive basis for assessing the hypotheses posed here.~~ (Mehrotra and Sharma, 2006; Sharma and Mehrotra, 2014) .

Ciach et al. (2007) used a nonparametric kernel regression to model radar rainfall uncertainty. They described the relation between true rainfall and radar-rainfall as the product of a systematic distortion function along with a random component and presented procedures to identify the two components. The distortion function could account for systematic biases which can be mathematically defined as a conditional expectation function, while the random component accounts for random errors in radar rainfall estimation. Villarini et al. (2008) estimated the conditional expectation function (distortion function) using both nonparametric (similar to ~~Villarini et al. (2008)~~Ciach et al. (2007)) and copula-based methods and compared the difference in performance between the two approaches using different quality metrics. It was found that performance of the nonparametric method was comparable with the copula-regression estimate and even outperformed when Nash Sutcliffe Efficiency (NSE) was used as a quality metric. The strength of nonparametric approaches is the ability to adapt to the data locally and the weakness is that the method ~~is sensitive to outliers and to large variability of data at the smallest (sub hourly) time scales. Hasan et al. (2016a) used~~ results in “local” biases as a result of outliers (Villarini et al., 2008). Hasan et al. (2016a) presented a kernel based nonparametric ~~method for radar rainfall estimation. In their approach , expected ground rainfall was estimated for a given reflectivity using a kernel-based conditional probability distribution. However, none of the methods above considered an additional covariate as air temperature as proposed in this study.~~ approach to estimate ground rainfall using radar reflectivity as a univariate predictor variable in a tropical setting. Past observed radar reflectivity and gauged rainfall were used in formulating the nonparametric model.

2 ~~Methodology~~

~~This section describes the methods used to formulate a nonparametric predictive modelwith incident air temperature and radar precipitation rate as the two predictors for radar precipitation~~ In this study, the hypothesis is that near surface air temperature observations can help improve radar precipitation estimates in cold climates. ~~A description of how the incident air temperature is incorporated as a covariate in the nonparametric radar precipitation estimation approach is presented next.~~

1.1 ~~Radar precipitation estimation~~

~~The proposed radar precipitation estimation algorithm consists of two steps. The first step quantifies the partial dependence of precipitation on radar precipitation rate and incident air temperature . The second step then uses the identified predictors in a non-parametric setting to estimate the precipitationresponse. Gauge precipitation is used as a ground reference or true precipitation~~ Here, the nonparametric approach of Hasan et al. (2016a) can be extended to allow use of a bivariate predictor vector with air temperature as an additional predictor variable to precipitation. This forms the basis for the investigation reported in this study.

~~The conditional estimation of precipitation using the two covariates can be described as follows:~~

$$\cancel{R_{est}(t)|\left[R(t),T(t)\right]}$$

~~Here, $(R_{est}(t))$ is the estimated ground precipitation from a given pair of radar rain rate $(R(t))$ and incident air temperature $(T(t))$ values at a given geographical location in the two-dimensional space (x, y) and time, t.~~

~~The conditional estimation in Eq. (1) uses two covariates, in contrast to Hasan et al. (2016b, a) where a nonparametric kernel regression estimator using a single covariate (R(t)) was adopted. Readers are referred to (Mehrotra and Sharma, 2006; Sharma and Mehrotra, 2014; Sharma et al., 2016) for further details on the nonparametric~~

~~modelling framework used in this work. This study uses the k-nearest neighbour (k-nn) regression estimator as the nonparametric predictive model. This model can be expressed as:~~

$$E\Big(R_{est}(t)\big|\big[R(t),T(t)\big]\Big) = \sum_{k=1}^{K}\frac{\frac{g_k}{k}}{\sum_{j=1}^{K}\frac{1}{j}}$$

~~Where k denotes the number of observed pairs of radar precipitation rate and temperature considered "similar" to the current conditioning vector $[R,T]$. Similarity here is defined on the basis of a weighted Euclidean distance that is further explained below. $E(.)$ denotes the expectation operator, in the absence of which the uncertainty about the expected value can be computed. The term g_k represents the observed gauge precipitation corresponding to k^{th} neighbour of the conditioning vector. K is a maximum number of neighbours permissible and it is an important parameter in the k-nearest neighbour method. In the present study, K is taken as equal to the square root of the sample size as suggested by Lall and Sharma (1996).~~

~~The order of each neighbour is ascertained based on a weighted Euclidean distance metric, written as :~~

$$\xi_i^2 = \left(\frac{\beta_R(R-r_i)}{s_R}\right)^2 + \left(\frac{\beta_T(T-t_i)}{s_T}\right)^2$$

~~Here, ξ_i is the distance of the conditioning vector $[R,T]$ to the i^{th} data point (r_i,t_i) in a two-dimensional space. s_R~~ This study set out to investigate the use of air temperature as an additional predictor in the radar precipitation estimation with the objective of improving quantitative radar precipitation estimation for cold climates. Compared to traditional radar-gauge adjustment, the proposed method is based on nonparametric approach using gauge precipitation and ~~s_T are sample standard deviations of the radar precipitation rate and temperature , and β_R~~ air temperature observations to adjust the radar precipitation. The precipitation estimates using a nonparametric model with temperature as a covariate is compared to a model without temperature and ~~β_T are partial weights denoting the relative importance each conditioning variable has on the ensuing response respectively (Sharma and Mehrotra, 2014). The sample standard deviations are used to standardise the predictor variables to make them independent of their measurement scale, while the partial weights allow elimination of a predictor variable if not relevant to the prediction being made. Readers are referred to Sharma and Mehrotra (2014) for the informational theory rationale that allows for the estimation of these partial weights, and the NPRED, R package ((Sharma et al., 2016), downloadable from http://www.hydrology.unsw.edu.au/download/software/npred) that enables their estimation for any sample data set.~~

1.1 ~~Model evaluation criteria~~

~~A number of metrics have been used in literature to evaluate and compare the performance of models (Hasan et al., 2016a; Villarini et al., 2008). The root mean square error (RMSE) is commonly used as a performance measure and it provides the overall skill measure of a predictive model (Hasan et al., 2016a). We used primarily RMSE~~

~~as a quality metric to evaluate the performance of the proposed model. Mean absolute error (MAE) and mean error (ME) were used as additional quality metrics. Definition of RMSE, MAE and ME can be found in the literature (e.g., Hasan et al., 2016a; Villarini et al., 2008).~~

1.1 ~~Determination of phase~~

~~In order to assess the usefulness of the proposed approach, it was compared against an alternate approach where the precipitation phase for first ascertained, followed by the application of different Z-R relationships for snow and rain . For the classification of precipitation phase at gauge level, we adopted the method from Finnish Meteorological Institute which is used operationally in Finland for phase classification (Koistinen et al., 2004; Saltikoff et al., 2015):~~

$$P_{lp} = \frac{1}{1+e^{22-2.7T-0.2H}}$$

~~Here, P_{lp} represents the probability of liquid precipitation , $T\ (^\circ C)$ the air temperature, and H (%) the relative humidity at a height of 2 m. If $P_{lp} < 0.2$,~~ to the original precipitation rates using a constant Z - R relationship. In addition, precipitation ~~is considered as solid and if $P_{lp} > 0.8$, precipitation is considered as liquid. For the case of $0.2 \leq P_{lp} \leq 0.8$, precipitation is considered as mixed (Koistinen et al., 2004; Saltikoff et al., 2015)~~rates using separate rain $(Z - R)$ or snow $(Z_e - S)$ relationships are back calculated from the original precipitation rates and are compared to the nonparametric estimates. Further, we investigate if improvements in precipitation estimates varies with temperature ranges and if the method is dependent on the precipitation intensities.

2 ~~Study area~~ Materials and ~~data~~methods

2.1 Study area

The proposed nonparametric predictive model using radar precipitation rate and air temperature as covariates was tested on radar data over the Oslo region in Norway. The radar data used in the current research is ~~the accumulated~~ an hourly radar precipitation rate product generated from the national weather radar network of Norway. The present study area is limited to the 50 km radius of radar range from Hurum radar station as shown in Fig. 1~~where a relatively dense network of precipitation gauges are available~~. The Hurum radar is located at 59.63° N latitude and 10.56° E longitude and it is about 30 km from Oslo, the capital city of Norway and it ~~is~~ has been in operation since November 2010. Data for the period from January 2011 to May 2015 were used for this study.

2.2 Data

The Norwegian Meteorological Institute (met.no) operates nine C-band Doppler weather radar installations which covers the entire land surface of Norway. The sensitive C-band installations with smaller wavelengths (4 - 8 cm) are placed in the Nordic to detect snowfall and clear air echoes (Koistinen et al., 2004). The wave length of the Hurum radar is 5.319 cm. The Norwegian

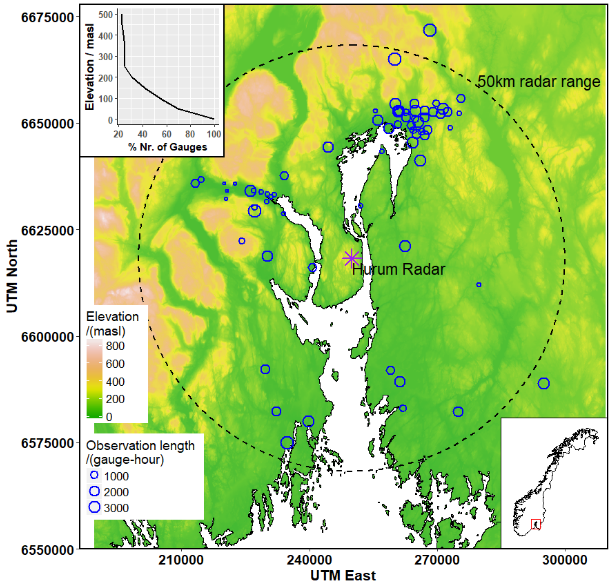

Figure 1. Precipitation gauge locations (blue circles) and length of the observations at each precipitation gauge location (size of the circles) and radar station (purple star mark) overlaid on topography of the study area, Oslo region of Norway. Hypsometric distribution (cumulative percentage of gauges below the specified elevation) ~~distribution~~ of the gauges is on the top left corner.

radar network scans the atmosphere with a 7.5 minute temporal resolution~~.~~; however, the temporal resolution was 15 minutes until June 2013. The met.no processes the raw radar volume scan from the radar stations. The data goes through extensive quality control and data transformations before the radar products are distributed to end users (Elo, 2012). The met.no performs a routine that removes clutter and other noise (non-meteorological echo) from the radar scan first. Then it reconstructs the gap in the data caused by clutter. The processing algorithm segments the volumetric radar reflectivity data as convective or stratiform precipitation type and it computes the Vertical Profile of Reflectivity (VPR) depending on precipitation types. VPRs of convective and stratiform precipitation types are distinctly different (Abdella, 2016; Chumchean et al., 2008). Bright band effect and non-uniform vertical profile of reflectivity are major sources of uncertainties in radar precipitation estimation in high latitude regions (Abdella, 2016; Joss et al., 1990; Koistinen et al., 2004; Koistinen and Pohjola, 2014). The radar data are corrected for bright band effects that appear in the VPR.

After the processing, the met.no generates and distributes various radar products. One of the radar precipitation rate products available for the public to use in hydrological applications is the Surface Rainfall Intensity (SRI). The SRI product uses the lowest Plan Position Indicator (PPI) and projects the aloft reflectivity data down to a reference height (1 km) near to the ground. The projection method is known as VPR correction that takes the vertical variability of reflectivity and bright band effect into account (Elo, 2012). The VPR corrected reflectivity is transformed from polar to Cartesian coordinate system with $1\ \mathrm{km} \times 1\ \mathrm{km}$ spatial resolution and the mosaic of nine weather radar imageries is merged to single SRI product covering the entire Norway. Finally, the reflectivity is converted to precipitation rate by using parametric Z - R relationship $(Z = 200R^{1.6})$

derived by Marshall and Palmer (1948) and the precipitation rate is accumulated to the temporal resolution desired (hourly in this case). The accumulated hourly SRI product was used in this study. It can be noted that the ~~Norwegian meteorological institute~~ met.no uses the single Z - R relationship (Marshall-Palmer for rain) for all seasons throughout the year.

~~Data for the period from January 2011 to May 2015 were used for this study . A spatial subset of accumulated hourly radar precipitation rate with 1 $\mathrm{km} \times 1\ \mathrm{km}$ spatial resolution~~ Within the study area, there are 68 precipitation gauges with available hourly precipitation data for the study~~area was downloaded from the met. no's "thredds" server (http://thredds.~~. The gauges in the study site consists of Geonor weighing gauges and tipping bucket gauges. Both types are with an Alter wind shield. The met.no ~~/). The data are in netCDF file format and~~ manages the calibration of gauges and takes necessary measures to reduce the uncertainty that arises when solid precipitation has to be measured. Further, data from the gauges are gone through routine quality control before being released to the public. However, the ~~gridded array is in Universal Transverse Mercator (UTM) 33 projected coordinate system. The hourly precipitation measurements from precipitation gauges are downloaded from the met. no's web portal for accessing meteorological data for Norway, "eKlima" (http://eklima.~~met.no ~~). Within the study area, 88 precipitation gauges are in operation with hourly observation, however only 68 gauges are available during the period from 2011 to 2015.~~ does not do wind induced undercatch correction for the precipitation data.

The precipitation gauges' locations ~~(68 gauges)~~ used in the study are shown in Fig. 1 overlaid on the topography of the study area. As shown in Fig. 1, precipitation gauges are not evenly distributed. The urban areas ~~(Oslo, Drammen, Lillestrom and Tonsberg)~~ are densely gauged (Nearly 0.25 gauges/km^2 near Oslo and approximately 0.1 gauges/km^2 near other major cities) and rest of the area is sparsely gauged with hourly observation. Further, the precipitation data from precipitation gauges come with varying length because some gauges ~~are~~ have been in operation since 2013 or later and some gauges have a number of missing values during their operation. ~~However, we used data from all available precipitation gauges for this study.~~

Some of the gauging stations are equipped with hourly temperature and other meteorological measurements (including wind speed and relative humidity). ~~For this study, we used gridded hourly temperature and wind speed dataset with 1 $\mathrm{km} \times 1\ \mathrm{km}$ grid resolution. The data are available from the Norwegian meteorological institute. The gridded wind speed data is available until May 2015. Even though, radar precipitation rates~~

In addition to precipitation and air temperature data~~are available from January 2011 to date, due to the unavailability of wind speed data for catch correction of gauge precipitation , the study period is limited to four years (January 2011 - May 2015). More details on the procedure adopted for catch correction are provided in the next sub-section~~, wind speed and relative humidity data were also required for this study. The wind speed was used for undercatch correction of precipitation gauges and relative humidity was used in the precipitation phase computation. Table 1 describes the datasets used in the study and the source and the spatial distribution of each dataset.

~~The gridded temperature dataset for Norway is spatially interpolated based on the historical air temperature observations from Norwegian meteorological stations . The interpolation is based on Optimal Interpolationin a Bayesian setting (Lussana et al., 2016)~~

The hourly gridded (1 km x 1 km) air temperature and wind speed datasets were generated by met.no. Lussana et al. (2016) spatially interpolated the past observed air temperature records from meteorological stations to develop the hourly gridded

Table 1. Different datasets used in the study and their source and spatial distribution.

Description	Gauge precipitation	Radar precipitation	Air temperature	Wind speed	Relative humidity
Spatial Distribution	Gauge locations	Gridded (1 km x 1 km)	Gridded (1 km x 1 km)	Gridded (1 km x 1 km)	Gauge locations
Data Source	Gauge	Radar	Gauge (interpolated)	NORA10 and AROME	Gauge

temperature dataset for Norway using Optimal Interpolation. In this three-dimensional spatial interpolation, the elevation of each grid point ~~is~~ was obtained from a high-resolution digital elevation model and the real elevation of stations ~~stored as metadata~~ were used. The ~~resulted~~ resulting interpolated air temperature is on the regular grid which is 2 ~~-~~m above the ground terrain elevation. For further details of the interpolation method, readers are referred to ~~the Norwegian meteorological institute'~~met.no's report by Lussana et al. (2016). ~~This gridded temperature data with an hourly temporal resolution was used to derive temperature time series for the precipitation gauge locations.~~

The ~~gridded hourly wind speed datasets are derived from a statistical downscaling of a~~ met.no derived an hourly gridded wind speed dataset by statistical downscaling from the 10 ~~km numerical model datasetonto a 1 km grid (same grid as the hourly gridded air temperature). Gridded~~ km numerical dataset, "NORA10" (Reistad et al., 2011) combined with data from the "AROME" 2.5 km numerical dataset (Müller et al., 2017) using a local quantile regression (Lussana 2018, personal communication). The 1 km $\times$ 1 km grid of the wind speed data is the same as for temperature. Even though, radar precipitation rates and air temperature data are available from January 2011 to date, the unavailability of wind speed data ~~was used to correct wind induced under-catch of precipitation gauges.~~ for undercatch correction after 2015 limited the study period to four years (January 2011 - May 2015).

Hourly measured relative humidity data ~~is available at 25 gauge~~ are available at 25-gauge locations within the study area. ~~Relative humidity data together with air temperature were used to compute the phase of the precipitation at gauge level in this study.~~ Spatial variation of relative humidity is relatively small within 50 - 100 km distances ~~and hence simple interpolation techniques can be used~~ (Beek, 1991). It can be noted that nearest gauge with relative humidity measurement is less than 50 ~~-~~km for most gauges in this study ~~. In this study,~~ and data from the nearest gauge was used for gauge locations ~~with missing relative humidity data, relative humidity data available from nearest gauge were used.~~without humidity measurements.

The datasets were downloaded and prepared for the study as follows. A spatial subset of hourly radar precipitation rate, air temperature and wind speed data with 1 km $\times$ 1 km spatial resolution for the study area was downloaded from met.no's "thredds" server (http://thredds.met.no/). The data are in netCDF file format in UTM33N projection. The hourly precipitation measurements from 68 precipitation gauges and relative humidity measurements of 25 gauges were downloaded from met.no's web portal for accessing meteorological data for Norway, "eKlima" (http://eklima.met.no).

As precipitation gauge locations and radar precipitation rate grids are in the same ~~UTM33~~ UTM33N coordinate system, they were simply overlaid and the radar pixel of 1 km^2 overlapping each precipitation gauge was located. One location near Oslo has three precipitation gauges within a 1 km $\times$ 1 km pixel. Except for that, all pixels consist of a single gauge. The pixel value (radar precipitation rate) for each hour was extracted and continuous hourly time series of radar precipitation rates for all gauges were generated. Similarly, time series of air temperature and wind speed at gauge locations were derived from the gridded temperature and wind speed data respectively.

The precipitation intensities in the study area ~~(high latitudes)~~ is relatively low. An analysis of statistical properties of precipitation rates in mid Norway showed that intensities less than $1.76\ \mathrm{mm\,h^{-1}}$ contributes to 50 % of the total precipitation volume while less than $6\ \mathrm{mm\,h^{-1}}$ contributes to 88 % (Engeland et al., 2014). Further, the same study found that precipitation intensities below $0.1\ \mathrm{mm\,h^{-1}}$ contributes little to the total precipitation and might be treated as zero precipitation. In addition, an analysis of the data used in this study showed that intensities between $0.05\ \mathrm{mm\,h^{-1}}$ and $0.1\ \mathrm{mm\,h^{-1}}$ are nearly 10 % of the total data above $0.05\ \mathrm{mm\,h^{-1}}$. Timesteps with gauge precipitation or radar precipitation rate less than $0.1\ \mathrm{mm\,h^{-1}}$ were therefore removed in this study. Finally, an observed dataset of hourly gauge precipitation and corresponding radar precipitation rate and air temperature for those hourly timesteps were prepared for all precipitation gauge locations. The length of the dataset (number of gauge-hours) at each gauge location used in this study is shown with the size of the circles in Fig. 1. It can be noted that nearly 103000 total gauge-hours were used for the study.

2.3 ~~Catch correction for precipitation gauges~~

~~Accuracy of precipitation gauge measurement is essential to achieve better results from water balance calculations, hydrological modelling and calibration of remote sensing algorithms.~~ Solid precipitation exhibits significant under-catch in windy conditions. Consideration of ~~catch errors~~ undercatch is more important in high latitude ~~Nordic~~ and mountainous regions due to ~~large catch errors for snow.~~ high wind conditions. A Field study in Norway showed that precipitation gauges, even with wind shield, catch 80 % of true precipitation at wind speeds of $2\ \mathrm{m\,s^{-1}}$, 40 % at $5\ \mathrm{m\,s^{-1}}$, and only 20 % at $7\ \mathrm{m\,s^{-1}}$ for solid precipitation at temperatures equal or below $-2°\ C$ ~~(Wolff et al. , 2015).~~ (Wolff et al., 2015). As this study uses gauge observation as a ground observed truth, corrected gauge observation is required for a reliable outcome from the investigation.

We corrected gauge precipitation for wind induced ~~under catch~~ undercatch by using the Nordic precipitation correction model (Førland et al., 1996). The Nordic model classifies the precipitation phase using air temperature and uses different equations for solid and liquid precipitation and a average value of the two equations ~~was~~ is used for mixed precipitation. The correction equations use wind speed and air temperature at each gauge location. ~~As mentioned above, gridded hourly~~ To verify whether the gridded wind speed data ~~was used for aerodynamic correction~~used in this study would provide a realistic correction, we compared it with the corrected precipitation using measured wind speed at 15 gauge locations. It was found that correlation between the corrected precipitation by using measured wind speed data (15-gauge locations) and gridded data are over 0.97 for all ~~those~~ 15 gauge locations. Based on the ~~catch error~~ undercatch computations in this study, the mean correction factor of hourly precipitation (corrected precipitation/observed precipitation) is 1.61 for solid and 1.14 for liquid precipitation while median are 1.53 and 1.11 for solid and liquid precipitation respectively. ~~Corrected gauge precipitation was used as true observed~~

2.3 Methodology

2.3.1 Radar precipitation estimation

The proposed radar precipitation estimation algorithm consists of two steps. The first step quantifies the partial dependence of precipitation on radar precipitation rate and incident air temperature. The second step then uses the identified predictors in a non-parametric setting to estimate the precipitation response. Gauge precipitation is used as a ground reference or true precipitation in this study.

3 ~~Results and Discussion~~

~~The performance of nonparametric radar precipitation estimation using air temperature as an additional covariate is presented in this section. The bivariate model is compared with the bench mark of the univariate nonparametric model where radar precipitation is used as the sole predictor. We tested the proposed method for a number of criteria and the results are presented below~~The conditional estimation of precipitation using the two covariates can be described as follows:

$$R_{est}(t)|\left[R(t), T(t)\right] \tag{1}$$

Here, $(R_{est}(t))$ is the estimated ground precipitation from a given pair of radar precipitation rate $(R(t))$ and incident air temperature $(T(t))$ values at a given geographical location in the two-dimensional space (x, y) and time, t.

The conditional estimation in Eq. (1) uses two covariates, in contrast to Hasan et al. (2016a, b) where a nonparametric kernel regression estimator using a single covariate (R(t)) was adopted. Readers are referred to (Mehrotra and Sharma, 2006; Sharma and Mehrotra, 2014; Sharma et al., 2016) for further details on the nonparametric modelling framework used in this work. This study uses the k-nearest neighbour (k-nn) regression estimator as the nonparametric predictive model. This model can be expressed as:

$$E\Big(R_{est}(t)|\left[R(t), T(t)\right]\Big) = \sum_{k=1}^{K} \frac{\frac{g_k}{k}}{\sum_{j=1}^{K}\frac{1}{j}} \tag{2}$$

Where k denotes the number of observed pairs of radar precipitation rate and air temperature considered “similar” to the current conditioning vector $[R, T]$. Similarity here is defined on the basis of a weighted Euclidean distance that is further explained below. $E(.)$ denotes the expectation operator, in the absence of which the uncertainty about the expected value can be computed. The term g_k represents the observed gauge precipitation corresponding to k^{th} neighbour of the conditioning vector. K is a maximum number of neighbours permissible and it is an important parameter in the k-nearest neighbour method. In the present study, K is taken as equal to the square root of the sample size as suggested by Lall and Sharma (1996).

The order of each neighbour is ascertained based on a weighted Euclidean distance metric, written as:

$$\xi_i^2 = \left(\frac{\beta_R(R - r_i)}{s_R}\right)^2 + \left(\frac{\beta_T(T - t_i)}{s_T}\right)^2 \tag{3}$$

Here, ξ_i is the distance of the conditioning vector $[R,T]$ to the i^{th} data point (r_i,t_i) in a two-dimensional space. s_R and s_T are sample standard deviations of the radar precipitation rate and temperature, and β_R and β_T are partial weights denoting the relative importance each conditioning variable has on the ensuing response respectively (Sharma and Mehrotra, 2014). The sample standard deviations are used to standardise the predictor variables to make them independent of their measurement scale, while the partial weights allow elimination of a predictor variable if not relevant to the prediction being made. Readers are referred to Sharma and Mehrotra (2014) for the informational theory rationale and partial informational correlation (PIC) that allows for the estimation of these partial weights, and the NPRED, R package ((Sharma et al., 2016), downloadable from http://www.hydrology.unsw.edu.au/download/software/npred) that enables their estimation for any sample data set.

2.3.2 Model evaluation criteria

A number of metrics have been used in literature to evaluate and compare the performance of models (Hasan et al., 2016a; Villarini et al., 2008). The root mean square error (RMSE) is commonly used as a performance measure and it provides the overall skill measure of a predictive model (Hasan et al., 2016a). We used primarily RMSE as a quality metric to evaluate the performance of the proposed model. Mean absolute error (MAE) and mean error (ME) were used as additional quality metrics. Definition of RMSE, MAE and ME can be found in the literature (e.g., Hasan et al., 2016a; Villarini et al., 2008).

2.3.3 Determination of phase

In order to assess the usefulness of the proposed approach, it was compared against an alternate approach where the precipitation phase was first ascertained, followed by the application of different Z-R relationships for snow and rain. For the classification of precipitation phase at gauge level, we adopted the method from Finnish Meteorological Institute which is used operationally in Finland for phase classification (Koistinen et al., 2004; Saltikoff et al., 2015):

$$P_{lp} = \frac{1}{1 + e^{22 - 2.7T - 0.2H}} \tag{4}$$

Here, P_{lp} represents the probability of liquid precipitation, T $(^\circ C)$ the air temperature, and H (%) the relative humidity at a height of 2 m. If $P_{lp} < 0.2$, precipitation is considered as solid and if $P_{lp} > 0.8$, precipitation is considered as liquid. For the case of $0.2 \leq P_{lp} \leq 0.8$, precipitation is considered as mixed (Koistinen et al., 2004; Saltikoff et al., 2015).

3 Results

3.1 Partial weight of predictors

For each precipitation gauge location, we estimated the partial weights associated with radar precipitation rate and incident air temperature using the observed hourly radar precipitation rate and air temperature and the corresponding gauge precipitation data.

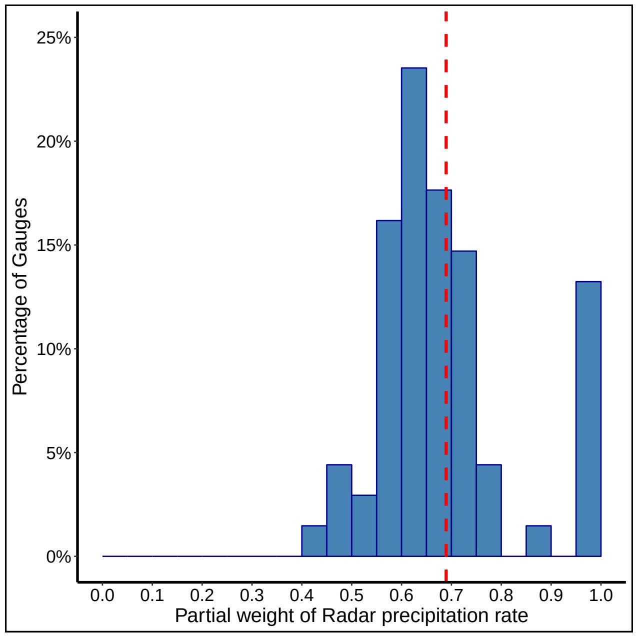

Figure 2. The Percentage of precipitation gauge locations against estimated partial weight of radar precipitation rate (β_R) at those gauge locations and the mean partial weight (red dash line) for gauge locations (68 gauges) in the study area. Partial weights provide a measure of relative importance of predictor variables on the response (refer Eq. (3)) and the summation of partial weights $(\beta_R + \beta_T)$ is equal to 1.

Figure 2 shows the histogram of partial weight of radar precipitation rate (β_R) computed for the 68 precipitation gauge locations in the study area~~of 50 km radius from the Radar station as shown in Fig. 1.~~. It is noted that the summation of partial weights of radar precipitation rate (β_R) and air temperature (β_T) is scaled to 1. Hence, the partial weight associated with air temperature (β_T) is equal to $1 - \beta_R$. Looking at Fig. 2, almost 87 % of the gauge locations resulted in non-zero partial weight for air temperature $(\beta_T > 0)$. In these locations, radar precipitation estimation partially depends on air temperature. It can be seen that partial weight of radar precipitation rate (β_R) is equal to 1 for nearly 13 % of the gauge locations and the partial weight associated with air temperature (β_T) is therefore zero. There, the bivariate problem collapsed into a univariate problem with radar precipitation rate as a single predictor.

Table 2 shows the summary statistics of computed partial weights among the precipitation gauge locations in the study area. It can be seen that the partial weight associated with air temperature is in the range of mean $+/-$ 0.1 for more than 70 %

Table 2. Summary statistics of computed partial weights for radar precipitation rate and air temperature in the study area.

Partial Weight	Mean	1st Quartile	3rd Quartile	15th Percentile	85th Percentile
Radar precipitation rate (β_R)	0.68	0.60	0.73	0.57	0.79
Air temperature (β_T)	0.32	0.40	0.27	0.43	0.21

of gauge locations. The gauge locations which resulted in associated partial weight for air temperature ($\beta_T > 0$) are spread throughout the study area. However, we have not found a clear pattern of spatial variation in the estimated partial weights~~at gauge locations within the study area~~.

3.2 Performance of k-nn prediction model

The k-nearest neighbour regression based estimator was used to ~~predict~~ estimate precipitation at each gauge location. The observed dataset and the computed partial weights of predictors were used with the NPRED k-nn regression tool to specify the ~~model~~proposed model with radar precipitation rate and air temperature as two predictors (knn-RT). For comparison, a reference model using the k-nn regression estimator ~~but~~ with radar precipitation rate as a single predictor variable (~~hourly radar precipitation rate~~knn-R) was also developed.

We calculated the k-nn regression estimate of expected response by using the leave-one-out cross-validation (LOOCV) procedure, whereby leaving out one observed response value (gauge precipitation) from the regression and estimating the expected response value for that observed response. This ensures the modelled outcomes represent the results that will be obtained using a new or independent data set. The improvement in radar precipitation estimation with the use of air temperature as an additional covariate is measured as a percentage reduction in RMSE compared to the reference model.

All the gauge locations with an associated partial weight of air temperature ($\beta_T > 0$) show an improvement in radar precipitation estimation. The mean improvement in RMSE is 9 % ~~while the median is 7.5 % and it is~~ and the improvement is more than 5 % for 80 % of the gauge locations where ~~air temperature was identified as an additional covariate~~β_T is greater than zero. It can be noted that partial weight for each gauge location was calculated independently using the data from that specific location and then the RMSE was estimated by LOOCV estimated using the entire data at that gauge location. However, a split sample test was done to verify the results, where two-thirds of the data were used to estimate partial weight and one-third of the data were used to estimate RMSE for each gauge location. The split sample test gave similar results~~as before~~.

We also examined the spatial cross-validation of computed partial weights. First, a single average partial weight was calculated by taking the arithmetic mean of ~~partial weights of~~ the partial weights for all gauge locations ~~which were computed independently at each gauge location and~~ presented in Fig. 2 and Table 2. This single average value of partial weight (0.68, 0.32) was used with the predictive models to estimate radar precipitation and the improvement in RMSE was estimated. Then, for each gauge location, an average partial weight was calculated by leaving that gauge out and adopting the mean partial weight from five nearest gauges. The k-nn prediction model was again re-specified for each gauge location using the computed average partial weight of the 5 nearest gauges. The ~~results of~~ percentage improvement in RMSE obtained by this method showed a strong resemblance to the results with a single mean value of partial weight~~for the study area~~. It is possible, therefore, that a regional or nearest neighbour average value of partial weight can be used for ungauged locations. As with the partial weight, the improvement in RMSE at gauge locations does not ~~clearly show any systematic~~ show any pattern of spatial variation.

Based on above examinations, the spatial variation of station specific partial ~~weight~~ weights can be discarded and a single average value adopted. Hence, in the results that follow, we use a single average partial weight computed for the study area. As shown in Table 2, the mean value of partial weight for the radar precipitation rate is 0.68 and ~~air temperature is 0.32. The~~ 0.32 for air temperature. The proposed k-nn regression prediction model with radar ~~rain~~ precipitation rate and air temperature as two predictors at each gauge location was specified with this single average partial weight.

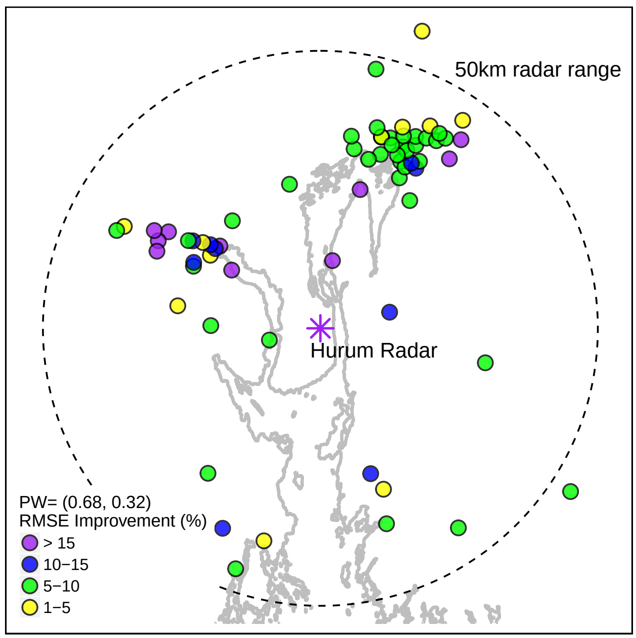

Figure 3. ~~Percentage~~ The percentage of improvement in RMSE at each gauge locations (colour scale) for predictive model with radar precipitation rate and air temperature as two predictors with the singe average partial weight ($\beta_R = 0.68$ and $\beta_T = 0.32$) compared to radar precipitation rate as a single predictor~~and length of the data (circle size)~~, ~~which are used in the predictive model,~~ overlaid on the coastline of the study area.

Figure 3 shows the percentage improvement in RMSE for the proposed model with ~~radar precipitation and air temperature as two predictors with~~ the single average partial weight of (0.68, 0.32) compared to the reference model~~with radar precipitation rate as a single predictor~~. The precipitation gauges' locations are shown by circles and ~~their sizes are proportional to the length of the data used in the nonparametric predictive models at each gauge location. A~~ a filled discrete colour scale is used to show percentage improvement in RMSE. All the gauge locations show improvement in RMSE with the use of temperature as an additional covariate ~~comparing with~~ compared to the reference model~~of radar precipitation as a single predictor~~. Looking at

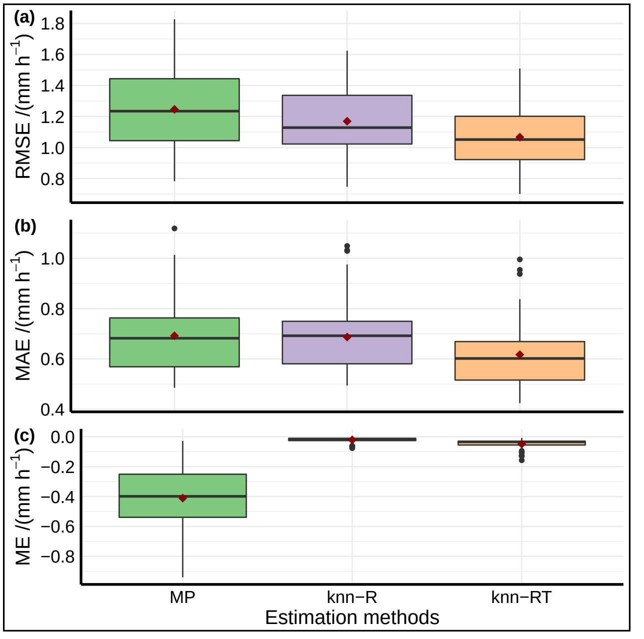

Figure 4. ~~Bar~~ Box plot representing three quality metrics (RMSE, MAE and ME) estimated at gauge locations for the original data (MP) and for the two nonparametric models (knn-R and knn-RT). Here, knn-R denotes the nonparametric model with radar precipitation rate as a single predictor, while knn-RT denotes the nonparametric model with radar precipitation rate and air temperature as two predictors with fixed partial weight of (0.68, 0.32). The values outside $1.5 * IQR$ are represented by the whiskers.

Fig. 3, the majority of gauge locations have a green colour and the improvement is between 5 - 10 % ~~on~~ at those locations. ~~Mean~~ The mean value of improvement is 8.5 % ~~while the median is 7 %~~. Over 80 % of the gauge locations in the study area show more than 5 % improvement in RMSE while nearly 15 % show more than 15.0 % improvement. As discussed earlier and as seen in Fig. 3, this study did not find any pattern of spatial variation in the results. However, ~~this spatial plot clearly shows that~~ the spatial plot shows the improvement in RMSE with the use of temperature as an additional predictor is spread throughout the study area.

In addition to RMSE, we computed MAE and ME for the proposed model and the reference model ~~with radar precipitation as a single predictor~~ at gauge locations. The above quality metrics were also computed for the original data of radar precipitation rates for comparison.

Figure 4 shows the summary of computed quality metrics for the two predictive models (knn-R and knn-RT) and the original data of radar precipitation rates ~~.~~ (MP). A bar plot representing these three quality metrics at each individual gauge location is available in the supplementary material (Supplementary Fig. 1). Looking at Fig. 4, the mean error of the original data (~~denoted as~~ MP) was negative for almost all gauge locations. This shows the under estimation of radar precipitation compared to precipitation measured by the gauges. Both nonparametric predictive models reduce the mean error considerably and bring it to near zero while they reduce the RMSE and MAE significantly for almost all gauge locations. It can be seen from ~~the~~

Fig. 4 (a) and (b) that the predictive model with radar precipitation as a single predictor (knn-R) reduces the RMSE and MAE. The proposed predictive model with radar precipitation and air temperature as two predictors (knn-RT) further reduces both RMSE and MAE and improves the radar precipitation estimation for most of the gauge locations.

~~Although the main focus of this paper is to investigate the benefit of using temperature as an additional covariate in radar precipitation estimation, the results of the nonparametric radar precipitation estimation in this study are comparable with the results of Hasan et al. (2016a), although in a different setting. They tested their nonparametric method of radar rainfall estimation (radar reflectivity as a single predictor) in Sydney, Australia and they have reported 10 % improvement in RMSE compared to the traditional parametric Z - R relationship. In our study, k-nearest neighbour nonparametric method with radar precipitation rate as a single predictor resulted in a mean 6 % reduction in RMSE. The bivariate model with air temperature as an additional predictor resulted in a mean 14 % reduction in RMSE compared to the original radar precipitation rate data derived using a parametric equation~~ ~~$(Z = 200R^{1.6})$.~~

~~The above results demonstrate the usefulness of air temperature as an additional predictor variable in deriving radar precipitation in cold climates. Some further investigations of when this improvement can be expected to be most are presented next.~~

3.3 Performance for different threshold intensities

~~This~~ The study used the precipitation intensities of radar precipitation and gauge precipitation equal or above $0.1\ \mathrm{mmh}^{-1}$. As described in Sect. ~~2.1~~2.2, precipitation intensities are relatively low in this region, consistent with intensities in cold climates. ~~A data analysis~~ An analysis of the data used in this study showed that intensities are lower than $0.5\ \mathrm{mm\,h}^{-1}$ for around 60 % of the observations and only 5 % of the data have either gauge or radar precipitation rates above $2.0\ \mathrm{mm\,h}^{-1}$.

To investigate whether very low intensities dominate the results presented earlier, we tested our proposed model for a range of intensities for both gauge and radar precipitation. Figure 5 shows the box plot of RMSE values estimated at gauge locations for threshold intensities $0.1\ \mathrm{mm\,h}^{-1}$, $0.5\ \mathrm{mm\,h}^{-1}$ and $2.0\ \mathrm{mm\,h}^{-1}$. Looking at Fig. 5, the improvement with the use of air temperature as an additional covariate is still ~~significant for more severe intensities as well~~seen over the intensity threshold. The results are statistically significant as the RMSE was estimated using leave one out cross validation (LOOCV) and are not impacted by the complexity of the model used.

3.4 Variation with Temperature Classes

For each gauge location, we also estimated partial weights for different temperature classes. ~~Partial informational correlation~~ The Partial Informational Correlation (PIC) and hence the partial weight was found to vary with temperature ~~class. For temperature above~~ classes. For temperatures warmer than 10° C, ~~more than 85 %~~ most of the gauge locations were estimated as having zero partial weight for air temperature while those locations resulted in non-zero partial weight ($\beta_T > 0$) for temperatures colder than 10° C. It is therefore likely that radar precipitation estimation depends on air temperature ~~for colder climates dominantly. The presence of hail may be the reason for a few precipitation gauge locations still exhibiting non-zero partial weight for air temperature above 10 C.~~ mainly in colder temperatures.

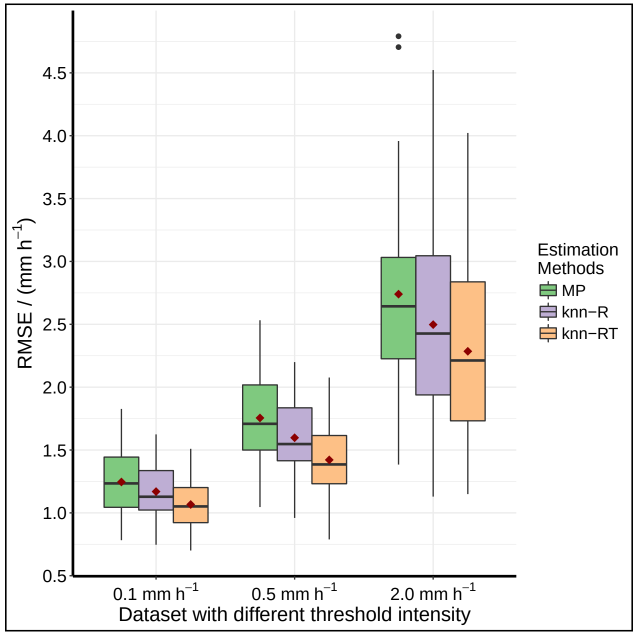

Figure 5. Box plot of $RMSE$ $(\mathrm{mm\,h^{-1}})$ values estimated at gauge locations for the original data (MP) and the two nonparametric models (knn-R and knn-RT) using data with intensities of radar precipitation rate and gauge precipitation greater than or equal $0.1\ \mathrm{mm\,h^{-1}}$, $0.5\ \mathrm{mm\,h^{-1}}$ and $2.0\ \mathrm{mm\,h^{-1}}$. Mean value of RMSE for each model by red diamond point. Here, knn-R - nonparametric model with radar precipitation rate as single predictor and knn-RT -nonparametric model with radar precipitation rate and air temperature as two predictors with the partial weight of (0.68, 0.32). The values outside $1.5 * IQR$ are represented by the whiskers.

Further, we estimated RMSE for the ~~dataset above~~ datasets with temperatures colder and warmer than 10° C for each gauge location using the proposed model with the average partial weight (0.68, 0.32) and ~~estimated the improvement compared to the reference modelwith radar precipitation rate as a single predictor. Nearly 70 % of gauge locations still showed improvement in RMSE~~ the reference model. The proposed model reduces the RMSE significantly for temperatures colder than 10° C; however, the ~~improvement is insignificant when the air temperature is above~~ performance is nearly as same as the reference model for temperatures warmer than 10° C (Supplementary Fig. 2). This shows that the use of air temperature as an additional covariate is most useful for the temperatures colder than 10° C.

3.5 Separate parametric equations for ~~rain and~~ snow and rain as a benchmark

As ~~we discussed in Sect. ??~~discussed earlier, the switch between a snow and rain Z - R relation is fast becoming a standard for weather radar operations in cold climates. We compared the proposed nonparametric radar precipitation estimation ~~models~~ model with radar precipitation estimation by using two different parametric Z - R relationships, one for snow and other for rain. In this study, we used the radar snow equation of Finish Meteorological Institute ($Z_e = 100S^2$) while keeping the Marshall and Palmer equation ($Z = 200R^{1.6}$) for rain.

~~For this investigation, we converted the original radar precipitation rates back to reflectivity using inversion~~ The analysis reported so far in the paper is based on the accumulated hourly radar precipitation rate product available from met.no. The evaluation using separate parametric equations for snow and rain as a bench mark requires radar reflectivity data to recompute radar precipitation rate using separate Z-R relationships for snow and rain. The reflectivity data used to produce the accumulated hourly radar precipitation rate (SRI product) used in the study are not stored in the production process, and therefore not available at met.no. As mentioned previously, the hourly product is based on corrected reflectivities with a time resolution of 15 minutes (before 2013) and 7.5 minutes (after 2013). These are then accumulated to the final hourly product. However, the Plan Position Indicator (PPI) of the lowest elevation beam from Hurum radar is available from met.no.

To back calculate reflectivites with original short time resolution based on the available hourly radar precipitation rate, it was assumed that the precipitation intensity distribution in each hour is the same for both the SRI and the PPI product, and that the hourly precipitation rates (SRI) therefore could be distributed within the hour using the intensity distribution of the PPI data. This procedure then gives us a series of precipitation rates with a time resolution of either 15 or 7.5 minutes depending on the year. The estimated precipitation rates were then converted to reflectivities using an inverse of the Marshall and Palmer equation ($R = (Z/200)^{1/1.6}$). ~~We estimated also~~

We estimated the probability of liquid precipitation (P_{lp}) using Eq. (4) ~~in order to classify and hence apply~~ and applied two different Z - R relationships to compute the precipitation rate according to the precipitation phase. Hourly air temperature and relative humidity at each gauge location were used ~~in this study for the estimation of~~ to estimate the probability of liquid precipitation (P_{lp}). Data were classified as solid or liquid or mixed precipitation using the computed hourly value of probability of liquid precipitation (P_{lp}). The back calculated reflectivity was converted to precipitation rates using the snow equation ($Z_e = 100S^2$) for solid phase and the rain equation ($Z = 200R^{1.6}$) for liquid phase. A weighted combination of solid and liquid was used for mixed precipitation by using the value of P_{lp} as recommended by Koistinen et al. (2004); Saltikoff et al. (2015). The precipitation rates were then accumulated to hourly time resolution. The Precipitation rates estimated by the two equations as described above is denoted by FMIMP for the further analysis.

For each gauge location, RMSE was calculated for the estimated radar precipitation rates by two equations (FMIMP). Here ~~wind~~ undercatch corrected gauge precipitation was used as a true observed value. RMSE of FMIMP is compared with the RMSE of original radar precipitation rates (MP) and the ~~two nonparametric predictive models (knn-R and~~ proposed nonparametric predictive model (knnRT). Figure 6 shows the box plot comparison of RMSE values in $\mathrm{mm\,h^{-1}}$ estimated at gauge locations for entire data and phase classes separately.

Looking at Fig. 6, the use of two equation (FMIMP) with the snow equation for solid and partially for mixed phase reduces the RMSE for solid and mixed precipitation phase classes and hence the RMSE of entire dataset compared to the original precipitation rates estimated by Marshall and Palmer equation (MP). The application of a different equation for snow reduces the phase dependent bias in the Norwegian radar precipitation estimation. The average reduction in RMSE at gauge locations is 6 % of RMSE value of the original radar precipitation rates. However, it can be seen in Fig. 6 that the use of different equations for snow and rain does not ~~reduces~~ reduce the RMSE to the level of the nonparametric approach (knn-RT). Comparing ~~FMMP~~ FMIMP and knn-RT, there is a further reduction of nearly 10 % in RMSE.

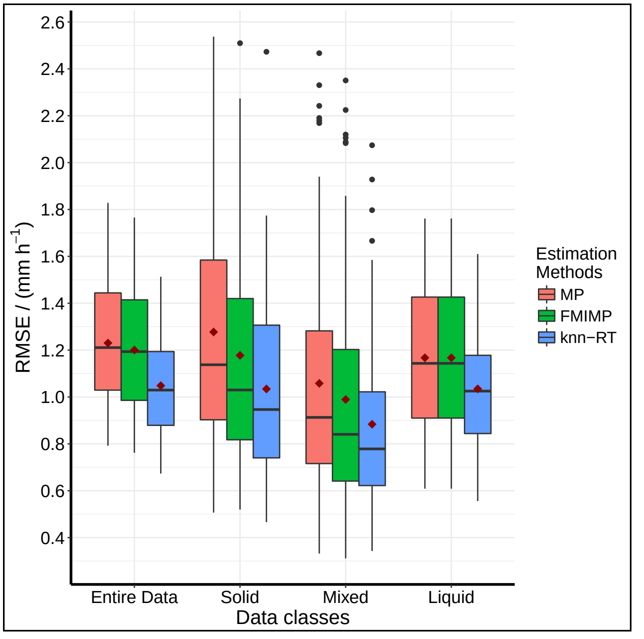

Figure 6. Box plot of comparison of $RMSE$ $(mm\ h^{-1})$ estimated at gauge locations for the original precipitation rates by Marshall and Palmer equation (MP) and precipitation rates estimated by different equation for snow and rain (FMIMP) and for the nonparametric model (knn-RT). RMSE values shown for entire data and separately for solid, mixed and liquid phase classes. Mean value of RMSE for each model by red diamond point. Here knn-RT - nonparametric model with radar precipitation rate and air temperature as two predictors with the partial weight of (0.68, 0.32). The values outside $1.5 * IQR$ are represented by the whiskers.

~~It~~

4 Discussion

In colder climates, the phase dependent uncertainties in radar precipitation estimation have hampered the extensive use of radar precipitation in hydrological applications (Berne and Krajewski, 2013; Saltikoff et al., 2015). To improve the quantitative radar precipitation estimates for hydrological applications, the study assessed the relevance of air temperature as an additional factor in the computation of radar precipitation in cold climates. In this paper, we show that using near surface air temperature as a second predictor variable in a nonparametric k-nearest neighbour (k-nn) method reduces the root mean squared error significantly compared to a k-nn model with radar precipitation rate as a single predictor and to the original hourly radar precipitation rates.

Despite phase dependent bias, accumulated radar precipitation rate products (e.g., met.no and OPERA) derived using a single Z-R relationship have been distributed to end users (Elo, 2012; Michelson et al., 2012). A key objective of the current study is to improve the hourly radar precipitation rates available to the public as a finished product (SRI product) from met.no that covers the entirety of Norway. However, the findings from this study can be helpful not only in Norway but also in a number of places where accumulated hourly product using a single Z - R relationship is applied. It can be noted that reflectivity data (dBZ) could be used instead of radar precipitation rate in the methodology presented in the paper if such data are available as shown by Hasan et al. (2016a).

A nonparametric framework was used for the investigation posed in the paper. Earlier studies (Hasan et al., 2016a, b) reported that given the availability of large amount of radar data, nonparametric approaches produce more reliable radar rainfall estimates compared to a traditional parametric Z - R relationship. In these studies, the nonparametric model used the radar reflectivity as a single predictor. This is the first study to our knowledge that considered the air temperature as an additional covariate in the radar precipitation estimation and the approach provided a clear improvement in the estimation. However, the improvement was significant for temperatures colder than 10° C. This appears mostly due to the different phase of precipitation in colder temperatures (including the presence of hail).

Partial informational correlation (PIC) based partial weights were used first to assess the partial dependence of radar precipitation estimation on air temperature, and then the weights were used with the k-nearest neighbour (k-nn) model. A simple k-nn approach is to use an equal weight for predictor variables or weights estimated using a simple linear partial correlation. Mehrotra and Sharma (2006) argue that the approach of assuming both predictor variables are equally important can result in increased bias and predictive uncertainty. Moreover, earlier studies (Mehrotra and Sharma, 2006; Sharma and Mehrotra, 2014; Sharma et al., 2016) have shown that the estimated PIC and weights collapse to what would be estimated using a linear regression model if the system is linear. As the system here is nonlinear, the use of the PIC to estimate partial weights seems to be the best approach. The study used a single average partial weight for the study area. If needed, it can be possible to use gridded partial weight with the k-nn model. However, we found that the gain in RMSE is not significant for the effort of added complexity of gridding the partial weights.

Although the main focus of this paper is to investigate the benefit of using air temperature as an additional covariate in radar precipitation estimation, the results of the nonparametric method of radar precipitation estimation found are comparable with the results of Hasan et al. (2016a). They tested their kernel based nonparametric method of radar rainfall estimation (radar reflectivity as a single predictor) in Sydney, Australia and reported a 10 % improvement in RMSE compared to the traditional parametric Z - R relationship. In this study, the k-nearest neighbour nonparametric method with radar precipitation rate as a single predictor resulted in a mean reduction in RMSE of 6 %. The proposed bivariate k-nn model with air temperature as an additional predictor resulted in a mean reduction in RMSE of 14 % compared to the original radar precipitation rates data.

The near surface air temperature, also together with relative humidity or wet-bulb temperature, has been used to estimate the dominant phase of precipitation in the selection of Z - R relationships (Koistinen et al., 2004; Saltikoff et al., 2015, 2000) . Fassnacht et al. (1999) and Fassnacht et al. (2001) reported the use of near surface air temperature to adjust the radar precipitation estimation and the benefit of the adjustment for hydrological applications. However, their approach was to use the temperature to estimate the probability of snow and use that information for the adjustment of radar precipitation. Further, the method was limited to mixed precipitation only, while the work presented here adjusts precipitation rate (it could be rain or snow or a mixture thereof) by using the k-nearest neighbour approach with near surface air temperature as a covariate.

The performance of the proposed k-nn method with temperature as a covariate was assessed primarily using a k-nn model without temperature and original radar precipitation rates derived by a single Z - R relationship as bench marks. As most cold climate radar operations use two separate equations for snow and rain, the study compared the nonparametric estimates with the precipitation rates estimated by two equations. First, reflectivities were back calculated in order to apply two equations. For this, we used PPI precipitation rates to distribute the VPR corrected SRI precipitation rates by assuming both have same intensity distribution within each hour. While there is uncertainty in how accurately the redistributed intensity distribution of SRI represents the original distribution, this exercise at least used a possible realistic distribution. Secondly, it should be noted that the phase classification used in this ~~study~~ evaluation is a model-based classification even though it is used operationally. The estimated phase can differ from actual observed phase at gauge level. Observations from disdrometers can provide a more accurate phase information at gauge level. ~~However, disdrometers are not available everywhere.~~ Even if a few disdrometers were located within the study region, their representativeness in space and time would be limited (Saltikoff et al., 2015). Further, our phase classification is at gauge level, and represents near surface conditions. The phase of the precipitation can be different at the elevation where the radar measures the reflectivity. ~~The measurement of phase information~~

Air temperature can be lapsed to the radar measurement height to estimate the phase of precipitation. Fassnacht et al. (1999, 2001) assumed the temperature lapse rate to be zero in their studies since winter lapse rates is often zero in mid latitude areas. For the Nordic region, Tveito and Førland (1999) showed that the vertical lapse rate varies with season and location. Further, Tveito and Førland (1999) found that local terrain conditions have greater influence in local temperature gradient during winter. Due to the occurrence of inversions, lapse rate can deviate substantially from the standard ($-6.5°\,\mathrm{C\,km^{-1}}$) during the winter months and it can be as low as $-1.2°\,\mathrm{C\,km^{-1}}$ (Tveito et al., 2000; Tveito and Førland, 1999). The estimated temperature at radar measurement height and hence the probability of liquid phase (P_{lp}) are therefore highly uncertain (Al-Sakka et al., 2013; Tveito et al., 2000). We, therefore, use the Finnish Meteorological Institute's operational method of near surface phase estimation to classify the precipitation as the method of choice for the evaluation as this method is both in operational use and developed for the Nordic area (Koistinen and Saltikoff, 1998; Gjertsen and Ødegaard, 2005; Saltikoff et al., 2015). The measurements of phase information at radar measurement height with the use of dual polarized radars can be a useful data source (Ryzhkov and Zrnic, 1998; Chandrasekar et al., 2013; Al-Sakka et al., 2013) for further investigation. However, many radars use a single polarity and moreover, even from dual polarised radars, data on phase information are not readily available to end users to help refine their estimation algorithms.

4.1 ~~Uncorrected gauge precipitation as an observed response~~

~~We tested the proposed method with measured gauge precipitation without wind induced catch correction. The~~ The study used the undercatch corrected gauged precipitation as a ground truth. We did a test on uncorrected gauge precipitation ~~was used there as an observed response in the model. For this investigation, we used six years of data from 88 precipitation gauges in the study area. Even though the wind induced catch error is making the observations less reliable ,~~ data (not corrected for wind induced undercatch) during an early phase of the ~~use of temperature as an additional predictor variable is having consistent impact as~~

~~with the results presented earlier using corrected gauge precipitation~~ study and found that air temperature as a covariate lead to improved RMSE in radar precipitation estimates also with uncorrected precipitation. It is often challenging to get reliable wind speed measurements for an operational real time radar precipitation estimation, and this finding implies that the method can also be used with uncorrected gauge precipitation to adjust the radar precipitation rates.

~~The purpose of the current study was to improve the quantitative radar precipitation estimates~~ The improved precipitation rates obtained through the nonparametric estimation of radar precipitation can be a data source for hydrological applications. ~~For this objective, this study assessed the relevance of temperature as an additional factor in the computation of radar precipitation for cold regions and climates.~~The spatial detail of the radar precipitation could solve issues related to precipitation representativity for hydrological modelling (Smith et al., 2004; Kirchner, 2009; Hailegeorgis et al., 2016). For many hydrological applications, short duration precipitation is needed and extending the study to sub-hourly time resolution and multiple radar bands (e.g., X band, S band) would be an interesting continuation to this work.

5 Conclusions

While parametric phase dependent Z-R relationships adjusted with gauged precipitation have been discussed extensively in the literature, this ~~is the first investigation to our knowledge that evaluates the use of~~ study extends current work with air temperature as a covariate in the radar precipitation adjustment~~and~~ , further presents a procedure whereby precipitation can be estimated in ~~cold climates. The proposed nonparametric bivariate model was evaluated using different quality metrics and tested for a number of criteria.~~ colder climates.

~~The key findings from this study are the following: The use of air temperature as an additional predictor variable in a nonparametric model improved the estimation of radar precipitation significantly. While this appears mostly due to the different phase of precipitation in colder temperatures (including the presence of hail), the proper use of temperature as a covariate can assist in better quantification of precipitation when knowledge of precipitation phase is not available. Care must be taken to use appropriate techniques to estimate precipitation when including temperature as a covariate. In the present study, use was made of a nonparametric technique which allowed for databased relationships to be formed. When equivalent data (ground precipitation especially) is not available, parametric equivalents will be needed instead. More work is needed to determine the best parametric relationship that could be adopted in such a situation.~~ An improvement of 15 % in the root mean squared error was ~~noted using the simple nonparametric approach adopted when including~~ obtained using a simple nonparametric method with air temperature as an additional covariate. More than 80 % of the locations ~~data was available for exhibited clear improvements in estimates.~~ showed improvement when temperature was used in the nonparametric model. The improvement was independent of precipitation intensities. However, the temperature effect became insignificant when air temperature was warmer than 10° C.

~~While this study uses data for one weather radar in arriving at its conclusions, preliminary analysis suggests the problems noted here to be generic.~~ Given the importance of weather radars as a means of precipitation measurement, and their ability to observe in remote regions in a continuous setting, the above finding ~~has considerable implications for ongoing operations~~ could be important for using radar precipitation data for hydrological applications especially in cold climates.

Code and data availability. Radar precipitation rate data used in the study are available in the Norwegian Meteorological Institute's (met.no) thredds server (http://thredds.met.no/thredds/catalog/remotesensingradaraccr/catalog.html). Precipitation observations from precipitation gauges, other meteorological measurements (wind speed and relative humidity) and gauges' meta information can be obtained from met.no's web portal "eKlima" (http://eklima.met.no). Access to the web portal is available upon request. Gridded observational hourly air temperature data and gridded wind speed data are available in the met.no's thredds server (http://thredds.met.no/thredds/catalog.html). NPRED programming tool, which is used for computation in the study, is available as R package and it can be downloadable from the following link as follows: http://www.hydrology.unsw.edu.au/download/software/npred

Competing interests. The authors declare that there are no competing interests.

Acknowledgements. The authors gratefully acknowledge the Norwegian meteorological institute (met.no) for providing radar ~~rain~~ precipitation rate, gauge precipitation and ~~air temperature~~ meteorological data for this study. The authors would particularly like to thank Christoffer Artturi Elo and Cristian Lussana at met.no for assisting to get the radar precipitation rate and gridded meteorological data. A great appreciation goes to Water Research Centre, University of New South Wales (UNSW), Sydney, Australia for hosting the first author for research practicum. The authors acknowledge the Norwegian Research Council and Norconsult for funding this research work under the Industrial Ph.D. scheme (Project No.: 255852/O30).

References

Abdella, Y.: Quantitative estimation of precipitation from radar measurements: analysis and tool development, Ph.D. thesis, Norwegian University for Science and Technology, Trondheim, 2016.

Al-Sakka, H., Boumahmoud, A.-A., Fradon, B., Frasier, S. J., and Tabary, P.: A new fuzzy logic hydrometeor classification scheme applied to the French X-, C-, and S-band polarimetric radars, Journal of Applied Meteorology and Climatology, 52, 2328–2344, 2013.

Auer Jr, A. H.: The rain versus snow threshold temperatures, Weatherwise, 27, 67–67, 1974.

Battan, L. J.: Radar observation of the atmosphere, University of Chicago Press, Chicago, 1973.

Beek, E.: Spatial interpolation of daily meteorological data; thoretical evaluation of available techniques, Tech. rep., DLO The Winand Staring Centre., Wageningen,The Netherlands, 1991.

Berne, A. and Krajewski, W. F.: Radar for hydrology: Unfulfilled promise or unrecognized potential?, Advances in Water Resources, 51, 357–366, 2013.

Beven, K.: Data for Rainfall–Runoff Modelling, in: Rainfall-Runoff Modelling: The Primer, chap. 3, pp. 51–82, John Wiley & Sons, Ltd, Chichester, UK, 2 edn., https://doi.org/10.1002/9781119951001.ch3, 2012.

Chandrasekar, V., Keränen, R., Lim, S., and Moisseev, D.: Recent advances in classification of observations from dual polarization weather radars, Atmospheric Research, 119, 97–111, 2013.

Chumchean, S., Sharma, A., and Seed, A.: Radar rainfall error variance and its impact on radar rainfall calibration, Physics and Chemistry of the Earth, Parts A/B/C, 28, 27–39, 2003.

Chumchean, S., Sharma, A., and Seed, A.: An integrated approach to error correction for real-time radar-rainfall estimation, Journal of Atmospheric and Oceanic Technology, 23, 67–79, 2006.

Chumchean, S., Seed, A., and Sharma, A.: An operational approach for classifying storms in real-time radar rainfall estimation, Journal of hydrology, 363, 1–17, 2008.

Ciach, G. J., Krajewski, W. F., and Villarini, G.: Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD data, Journal of Hydrometeorology, 8, 1325–1347, 2007.

Crozier, C., Joe, P., Scott, J., Herscovitch, H., and Nichols, T.: The king city operational doppler radar: Development, all-season applications and forecasting, Atmosphere-Ocean, 29, 479–516, 1991.

Elo, C. A.: Correcting and quantifying radar data, Tech. Rep. 2, Norwegian Meteorological Institute, Norway, 2012.

Engeland, K., Sultan Abdella, Y., and Lepioufle, J.-M.: Statistical properties of rain rates in mid Norway as seen by a vertically looking micro rain radar (MRR), in: EGU General Assembly Conference Abstracts, vol. 16, 2014.

Fassnacht, S., Soulis, E., and Kouwen, N.: Algorithm application to improve weather radar snowfall estimates for winter hydrologic modelling, Hydrological Processes, 13, 3017–3039, 1999.

Fassnacht, S., Kouwen, N., and Soulis, E.: Surface temperature adjustments to improve weather radar representation of multi-temporal winter precipitation accumulations, Journal of Hydrology, 253, 148–168, 2001.

Fassnacht, S., Venable, N., Khishigbayar, J., and Cherry, M.: The probability of precipitation as snow derived from daily air temperature for high elevation areas of Colorado, United States, in: Cold and Mountain Region Hydrological Systems Under Climate Change: Towards Improved Projections (Proceedings of symposium H02, IAHS-IAPSO-IASPEI Assembly, Gothenburg, Sweden, July 2013), 360, pp. 65–70, IAHS, 2013.

Førland, E., Allerup, P., Dahlström, B., Elomaa, E., Jónsson, T., Madsen, H., Perälä, J., Rissanen, P., Vedin, H., and Vejen, F.: Manual for operational correction of Nordic precipitation data, Tech. Rep. 24/96, Norwegian Meteorological Institute,DNMI, Oslo, Norway, 1996.

Gjertsen, U. and Ødegaard, V.: The water phase of precipitation—a comparison between observed, estimated and predicted values, Atmospheric Research, 77, 218–231, 2005.

Hailegeorgis, T. T., Alfredsen, K., Abdella, Y. S., and Kolberg, S.: Evaluation of storage–discharge relationships and recession analysis-based distributed hourly runoff simulation in large-scale, mountainous and snow-influenced catchment, Hydrological Sciences Journal, 61, 2872–2886, 2016.

Hasan, M. M., Sharma, A., Mariethoz, G., Johnson, F., and Seed, A.: Improving radar rainfall estimation by merging point rainfall measurements within a model combination framework, Advances in Water Resources, 97, 205–218, 2016a.

Hasan, M. M., Sharma, A., Johnson, F., Mariethoz, G., and Seed, A.: Merging radar and in situ rainfall measurements: An assessment of different combination algorithms, Water Resources Research, 52, 8384–8398, 2016b.

Hong, Y. and Gourley, Jonathan, J.: Radar Hydrology: Principles, Models, and Applications, CRC Press, Boca Raton, 2015.

Joss, J., Waldvogel, A., and Collier, C.: Precipitation measurement and hydrology, in: Radar in meteorology: Battan Memorial and 40th Anniversary Radar Meteorology Conference, pp. 577–606, American Meteorological Society, 1990.

Kienzle, S. W.: A new temperature based method to separate rain and snow, Hydrological Processes, 22, 5067–5085, 2008.

Killingtveit, Å.: Water balance study in Sagelva research basin (In Norwegian: En studie av vannbalansen i Sagelva hydrologiske forskningsfelt), Ph.D. thesis, Institutt for vassbygging, Norges tekniske høgskole, Trondheim, 1976.

Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward, Water Resources Research, 45, 2009.

Koistinen, J. and Pohjola, H.: Estimation of ground-level reflectivity factor in operational weather radar networks using VPR-based correction ensembles, Journal of Applied Meteorology and Climatology, 53, 2394–2411, 2014.

Koistinen, J. and Saltikoff, E.: Experience of customer products of accumulated snow, sleet and rain, in: Advanced Weather Radar Systems, COST75, International seminar, Locarno, Switzerland, pp. 397–406, Office for Official Publications of the European Communities, Luxembourg, 1998.

Koistinen, J., Michelson, D. B., Hohti, H., and Peura, M.: Operational measurement of precipitation in cold climates, in: Weather Radar, pp. 78–114, Springer, 2004.

Lall, U. and Sharma, A.: A nearest neighbor bootstrap for resampling hydrologic time series, Water Resources Research, 32, 679–693, 1996.

Lussana, C., Ole, E. T., and Francesco, U.: seNorge v2.0: an observational gridded dataset of temperature for Norway, Tech. Rep. 14, Norwegian Meteorological Institute, Norway, 2016.

Marshall, J. and Gunn, K.: Measurement of snow parameters by radar, Journal of Meteorology, 9, 322–327, 1952.

Marshall, J. S. and Palmer, W. M. K.: THE DISTRIBUTION OF RAINDROPS WITH SIZE, Journal of Meteorology, 5, 165–166, 1948.

Mehrotra, R. and Sharma, A.: Conditional resampling of hydrologic time series using multiple predictor variables: A K-nearest neighbour approach, Advances in water resources, 29, 987–999, 2006.

Michelson, D., Koistinen, J., Peltonen, T., Szturc, J., and Rasmussen, M. R.: Advanced weather radar networking with BALTRAD+, in: Proc. 7th European Conference on Radar in Meteorology and Hydrology ERAD, 2012.

Müller, M., Homleid, M., Ivarsson, K.-I., Køltzow, M. A., Lindskog, M., Midtbø, K. H., Andrae, U., Aspelien, T., Berggren, L., Bjørge, D., et al.: AROME-MetCoOp: A Nordic convective-scale operational weather prediction model, Weather and Forecasting, 32, 609–627, 2017.

Reistad, M., Breivik, Ø., Haakenstad, H., Aarnes, O. J., Furevik, B. R., and Bidlot, J.-R.: A high-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea, and the Barents Sea, Journal of Geophysical Research: Oceans, 116, 2011.

Rohrer, M.: Determination of the transition air temperature from snow to rain and intensity of precipitation, in: WMO IASH ETH International Workshop on Precipitation Measurement, pp. 475–582, 1989.

Ryzhkov, A. and Zrnic, D.: Discrimination between rain and snow with a polarimetric radar, Journal of Applied Meteorology, 37, 1228–1240, 1998.

Saltikoff, E., Koistinen, J., and Hohti, H.: Experience of real time spatial adjustment of the ZR relation according to water phase of hydrometeors, Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, 25, 1017–1020, 2000.

Saltikoff, E., Lopez, P., Taskinen, A., and Pulkkinen, S.: Comparison of quantitative snowfall estimates from weather radar, rain gauges and a numerical weather prediction model, Boreal Env. Res, 20, 667–678, 2015.

Sekhon, R. and Srivastava, R.: Snow size spectra and radar reflectivity, Journal of the Atmospheric Sciences, 27, 299–307, 1970.

Sharma, A. and Mehrotra, R.: An information theoretic alternative to model a natural system using observational information alone, Water Resources Research, 50, 650–660, 2014.

Sharma, A., Mehrotra, R., Li, J., and Jha, S.: A programming tool for nonparametric system prediction using Partial Informational Correlation and Partial Weights, Environmental Modelling & Software, 83, 271–275, 2016.

Silverman, B. W.: Density estimation for statistics and data analysis, vol. 26, CRC press, 1986.

Smith, M. B., Seo, D.-J., Koren, V. I., Reed, S. M., Zhang, Z., Duan, Q., Moreda, F., and Cong, S.: The distributed model intercomparison project (DMIP): motivation and experiment design, Journal of Hydrology, 298, 4–26, 2004.

Smith, P. L.: Equivalent radar reflectivity factors for snow and ice particles, Journal of Climate and Applied Meteorology, 23, 1258–1260, 1984.

Tveito, O., Førland, E., Heino, R., Hanssen-Bauer, I., Alexandersson, H., Dahlström, B., Drebs, A., Kern-Hansen, C., Jónsson, T., Vaarby Laursen, E., et al.: Nordic temperature maps, Tech. Rep. 9, Norwegian Meteorological Institute, Norway, 2000.

Tveito, O. E. and Førland, E. J.: Mapping temperatures in Norway applying terrain information, geostatistics and GIS, Norsk Geografisk Tidsskrift-Norwegian Journal of Geography, 53, 202–212, 1999.

Uijlenhoet, R.: Raindrop size distribution and radar reflectivity-rain rate relationships for radar hydrology, Hydrology and Earth System Sciences, 5, 615–627, 2001.

Villarini, G. and Krajewski, W. F.: Sensitivity studies of the models of radar-rainfall uncertainties, Journal of Applied Meteorology and Climatology, pp. 288–309, 2010a.

Villarini, G. and Krajewski, W. F.: Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall, Surveys in Geophysics, pp. 107–129, 2010b.

Villarini, G., Serinaldi, F., and Krajewski, W. F.: Modeling radar-rainfall estimation uncertainties using parametric and non-parametric approaches, Advances in Water Resources, 31, 1674 – 1686, 2008.

Wilson, J. W. and Brandes, E. A.: Radar measurement of rainfall—A summary, Bulletin of the American Meteorological Society, 60, 1048–1058, 1979.

Wolff, M., Isaksen, K., Petersen-Øverleir, A., Ødemark, K., Reitan, T., and Brækkan, R.: Derivation of a new continuous adjustment function for correcting wind-induced loss of solid precipitation: results of a Norwegian field study, Hydrology and Earth System Sciences, 19, 951–967, 2015.

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cocks, S., Martinaitis, S., et al.: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities, Bulletin of the American Meteorological Society, 97, 621–638, 2016.