W O R L D   M E T E O R O L O G I C A L   O R G A N I Z A T I O N

COMMISSION FOR INSTRUMENTS
AND METHODS OF OBSERVATION

PROJECT TEAM AND (REDUCED) INTERNATIONAL ORGANIZING COMMITTEE FOR THE WMO SOLID PRECIPITATION INTERCOMPARISON EXPERIMENT

Fifth Session

Sodankylä, Finland
19 – 23 May 2014

FINAL REPORT

CONTENTS

EXECUTIVE SUMMARY

This report provides a summary of the meeting of the Project Team and (reduced) Fifth session of the International Organizing Committee (IOC) of the WMO Solid Precipitation Intercomparison Experiment (SPICE) that was held in Sodankylä, Finland from 19 to 23 May 2014.

The meeting reviewed the status of the experiment on all sites and of the data transfer to the data archive. The meeting also reviewed the methodologies developed for the data analysis, both of the reference as well as for the instruments under test. The meeting agreed on the way forward for continuing the data analysis.

AGENDA

1. ORGANIZATION OF THE SESSION

1.1 Opening of the Session

1.2 Adoption of the Agenda

1.3 Working Arrangements for the Session

2. REPORT OF THE CHAIRPERSON

3. SITE COMMISSIONING OVERVIEW

4. REPORT ON REFERENCE

5. STRATEGY TO ACHIEVE THE PROJECT OBJECTIVES

6. OTHER BUSINESS

7. DRAFT REPORT OF THE SESSION

8. CLOSURE OF THE SESSION

GENERAL SUMMARY

1. ORGANIZATION OF THE SESSION

1.1 Opening of the Session

1.1.1 The meeting of the Project Team and (Reduced) International Organizing Committee (IOC) for the WMO Solid Precipitation Intercomparison Experiment (SPICE), Fifth Session, was opened on Monday, 19 May 2014 at 8:30, by Ms Rodica Nitu, the IOC Chairperson and SPICE Project Leader. The list of participants is given in Annex I.

1.1.2 Mr Osmo Aulamo, welcomed the participants to the Arctic Research Centre of the Finnish Meteorological Institute (FMI). He made a brief presentation of the FMI Arctic Research Centre and Sodankylä Geophysical Observatory.

1.2 Adoption of the Agenda

The meeting adopted the Agenda as reproduced at the beginning of this report.

1.3 Working Arrangements for the Session

The working hours and tentative timetable for the meeting were agreed upon.

2. REPORT OF THE CHAIRPERSON

2.1 Ms Rodica Nitu, the SPICE Project Leader and Chairperson of the IOC, informed the meeting that the CIMO Management Group (MG) had positively received her proposal to extend the project by one year and to support a data analyst provided it could be mainly funded by the CIMO Trust Fund. The CIMO MG recommended not to include new participants in the project at this stage, but to concentrate on the data analysis to ensure the final report could be published in 2016. The meeting welcomed the support of the CIMO MG for the extension of the project.

2.2 The meeting requested all site managers to clarify the implications of the project prolongation to the temporary import arrangements for the instruments on their sites and to make appropriate arrangements with their customs authorities, if required.

2.3 Ms Nitu presented the progress achieved towards acquiring a comprehensive data set, the progress made on the report on the Configuration of the SPICE Working Field Reference System, an overview of the last two measuring seasons, including successes, challenges, interactions with instrument providers, items of interest for the future, lessons learned and plans of the sites for 2014/15. She also explained her expectations from the meeting, in view of finalizing a detailed workplan (prioritizing activities) that would enable the project team to complete the data analysis by end of 2015 and publication of the final report by 2016. She recommended to now focus on the analysis of all the SPICE data rather than focussing longer on the evaluation of the references, recognizing that this aspect might have to be reopened at a later stage, depending on the experiences and findings arising from the rest of the data analysis. An extract from her report is provided in Annex II.

2.4 Ms Nitu stressed the need to communicate the team's approach to the computation of the references, so as to obtain the endorsement of the stakeholder community before addressing the results of the instruments. Many persons have expressed to her their interest for the approach and results related to the field working references. The report on the field working references is expected to be published in the second half of 2014.

2.5 Ms Nitu noted that she had been impressed by the strong engagement of all the project team throughout the year, and for the contributions provided towards meeting the project objectives to date. She stressed that the aim of the meeting is to plan the work ahead in order to ensure that a significant part of the data analysis will be achieved by the time of the next meeting and to enable the finalization and publication of the project report by 2016.

3. SITE COMMISSIONING OVERVIEW

3.1 The meeting was informed about the status of completion of the commissioning protocol by all sites. The available commissioning reports are available on the SPICE website (http://www.wmo.int/pages/prog/www/IMOP/intercomparisons/SPICE/SPICE.html). The meeting recognized that the completion of those reports required a significant work by the site managers, but that they were essential in demonstrating the readiness of sites to produce high quality data and to demonstrate the quality of the experiment. The meeting noted that the project leader had requested that in case a change had to be made in a commissioning report, this needed to be reflected in a tracking table inserted at the beginning of the document.

3.2 The meeting recalled that the tracking of changes on the sites after their commissioning remained an issue and that a Site Change Tracking Sheet had been developed to describe modifications, extraordinary maintenance operations, instruments failures, etc. that would be relevant for the data analysis. The meeting requested that all sites maintain this Site Change Tracking Sheet up-to-date, including all maintenance activities, including those performed by instrument providers.

3.3 The meeting was also informed about the details related to the reference installed on each site, and was presented with a comprehensive list of all the instruments included in the experiment and of their configuration. Overall, more than 80 precipitation gauges are being tested, including 14 different operating principles and 10 different shield configurations. 30 instruments for the measurement of snow on the ground are also included in the experiment.

3.4 More than ten site Commissioning reports are published or about to be published. The meeting was concerned that some sites had not completed their commission report to date and urged all sites to urgently complete them to provide the needed confidence in the conduct of the intercomparison. The situation is as follows:

- The site manager of Marshall could unfortunately not attend the meeting, but a very advanced draft version of the commissioning reports from Marhsall was received. Francesco Sabatini was requested to liaise with Marshall to ensure the completion of the Marshall commissioning report.
- The report of Col de Porte is almost ready for publication.
- Arkady Koldaev presented the status of the sites of Volga River and Valdai and informed that these sites would submit the final documents in the near future.
- The Hala Gasienicowa (Poland) site manager has already provided a first set of information that still needs to be arranged according to the SPICE commissioning report template. Some improvements have been recently done to the site configuration. Maciej Karzynski agreed to finalize the Commissioning Protocol using the standard format by June 2014. This site has a specific relevance for Snow on the Ground (SoG) observations, since it has long historical data series.
- The Tapado (Chile) site manager, Shelley MacDonnell, provided her commissioning report at an early stage. Unfortunately, all subsequent communication attempts with her have failed. The IOC will consider other options to contact representatives of this site and encourage them to send the information required, the site update and the data.
- The italian sites (Forni Glacier – Italy and Pyramid EV-K2_CNR Lab. - Nepal) approved at the SPICE IOC-4 session (2013) as S4 sites of interest for the SoG, provided a draft version of their commissioning reports. They are completing the final version of their commissioning reports which are expected to be ready for June 2014.
- No draft reports were received from the sites of Joetsu and Rikubetu (Japan) and Gochang (Rep. of Korea).

3.5 The meeting recalled that the maintenance and calibration of the instruments needed to be carried strictly according to the instrument user manual and additional guidance received by the instrument providers, if applicable. The meeting requested the site managers to ensure they follow those practices, contact the manufacturer in case of doubt, and document their activities in the Site Change Tracking Sheet. The meeting further requested them to ensure that required maintenance and calibration activities be organized in the summer months to ensure that all instruments will be fully operational at the onset of the winter.

3.6 The meeting recalled that any modification of the site (incl. instrument setup) requires prior agreement by the IOC and requested all site managers to inform the project leader of any requirement/proposals for changes in their site configuration.

3.7 Yuri Melnichuk presented the configuration of the Valdai bush gauge, which is actually composed of 3 individual Tretyakov gauges with Teytakov shields. One of them is surrounded by a wooden single-fence of 4 m diameter. The values of these 3 gauges are averaged to derive the "bush-gauge" value.

3.8 The meeting recommended to consider installing an automatic gauge in the Valdai bush to compare its performance with the traditional bush gauge.

3.9 Arkady Koldaev presented results from the Valdai and Volga sites. The meeting noted that the shape of the DFIR-fence of Valdai did not correspond to the recommendation of SPICE.

3.10 The meeting also recalled that the IOC had noted that the field working reference (combined R1 and R2, and R3) systems of Volga site were not configured according to the IOC SPICE recommendations and that it would therefore not be possible to link the precipitation amount measurements results from the Volga site with those from any other site in SPICE. The IOC had informed the Volga site manager of these concerns and recommended to modify the configuration of the site.

3.11 Ms Antonella Senese of the University of Milan, on behalf of the Italian site managers, presented the update and the improvements performed at these sites. The Forni Glacier site (Italian Alps, Ortles-Cevedale Group) was equipped with different sensors measuring the snow on the ground on 6th May 2014. The automatic measurements are performed by two different sonic rangers (every 60' by Campbell SR50 and every 10' by Sommer USH-8) and a snow pillow (every 10' by Park Mechanical). The manual observations are carried out by snow pits (every month, according to the AINEVA protocol, see www.AINEVA.it), by 4 graduated stakes at the corners of the snow pillow (photographed every 60' by an automatic camera). The snow surface temperature can be estimated by the outgoing longwave radiation (measured every 30' by a Kipp&Zonen net radiometer). For the next winter (2014/2015), the site team is planning to install other graduated stakes close to the two sonic rangers. The meteorological parameters and the energy fluxes measured by the automatic weather station constitute the ancillary measurements: i) air temperature, ii) relative humidity, iii) wind speed, iv) wind direction, v) incoming and reflected solar radiation, vi) incoming and outgoing longwave radiation, vii) liquid precipitation, viii) atmospheric pressure, and ix) icing occurrence (deducted by the air temperature and the relative humidity).

3.12 Regarding the choice of the sensors to be installed at the surface of the Forni Glacier, the main issue is the energy supply, which is represented only by solar panels and lead gel battery. Secondly, the glacier is a dynamic body and the ice surface is not smooth. Finally, the Forni Glacier is considered as a Site of Community Importance (SCI, code IT2040014) and it is located in a wide natural protected area (the Stelvio National Park), thus also requiring a deep analysis of the possible expected impacts of instruments and devices before their installation.

3.13 The Pyramid International Laboratory-Observatory (Lobuche, SoluKhumbu, Nepal) is established in the framework of the collaboration between Ev-K2-CNR and Nepal Academy of Science & Technology-NAST. In May two Sommer USH-8 sensors (automatic measurements) and graduated rods (manual observations) were installed. The ancillary measurements are i) precipitation occurrence/rain, ii) atmospheric pressure, iii) air temperature, iv) relative humidity, v) wind speed/direction at 5 m height, vi) net radiation (short- and longwave), vii) soil temperature (-5 cm and -20 cm), vii) soil moisture, and viii) soil heat flux. Photography and video equipment are also available for recording and archival of site conditions. An important support is represented by the constant presence of local staff.

3.14 Ms Senese indicated that the two site commissioning protocols will be sent by June/July 2014. The data measured in Italy and in Nepal will be transferred to NCAR in September 2014.

3.15 Ms Rodica Nitu offered to work with Gochang, to support them in finalizing their commissioning report.

3.16 The meeting was invited to comment on the configuration of the second DFIR of Bratt's lake (West DFIR) in which an unheated Geonor is presently mounted and to consider whether it would be beneficial for SPICE if it were replaced by a Pluvio[2] for the 2014/2015 winter. As a result the Brat's Lake site manager was asked, if possible, to install a Pluvio2 gauge, thus replacing the unheated Geonor. This would allow the comparison between R2 references using each of the two WG recommended.

4. REPORT ON REFERENCE

4.1 The Data Analysis Team (DAT) prepared a draft report on the field reference for precipitation amount describing the concepts that are proposed to be used to derive the reference data.

4.2 Mareile Wolff presented the current status on the SPICE Report on the Field Reference for Precipitation Amount (SPICE REF). Content on most topics have been provided. The input will now need to be streamlined, identifying redundancies and probable gaps. Working towards the next version, some re-organizing will be done to achieve a consistent and logical structure of the document. It will then be internally reviewed. It is aimed to have a draft ready before TECO-2014 and CIMO-16 in St. Petersburg (7-16 July 2014). A presentation of this report is planned at TECO-2014. Final publication is planned for September 2014. All interested parties, including instrument providers, will then be invited to comment on it and to communicate potential concerns to the project leader and chair of the data analysis team, so that they could be addressed.

4.3 The reference report also includes as Annexes, 2 page summaries describing the references used at each site. All site managers who have not provided their input for those annexes were requested by the meeting to provide it to the Project Leader by 15 June 2014.

5. STRATEGY TO ACHIEVE THE PROJECT OBJECTIVES

Summary of proposed strategies to achieve the project objectives

5.1 The meeting was presented with the progress made in developing the methodologies that would be considered for use for the data analysis and towards meeting as many of the project objectives as possible. Some of those proposals will have to be refined, and further discussed and endorsed by the team, while others are almost finalized.

5.2 Major Emanuele Vuerich presented the procedures and methods that were used for the WMO Field Intercomparison on Rainfall Intensity Gauges (WMO FI-RI) held from October 2007 to April 2009 in Vigna di Valle (Italy). The intercomparison report was published as IOM Report No. 99 and is available at http://www.wmo.int/pages/prog/www/IMOP/publications/IOM-99_FI-RI.pdf. In particular Mj Vuerich presented the methodology used for the derivation of the field reference rainfall intensity (RI) as composition of working reference gauges, the event selection, the determination of the uncertainty of the reference and the achievable uncertainty of gauges under test. He also showed the procedures adopted for quality assurance, including the field calibrations by means of a portable device, inspection and maintenance of gauges, the use of an automatic quality control (AQC) developed for both raingauges and ancillary instruments and the organization of a participants-local staff meeting during the campaign for strengthening their involvement and their support to the proper operation of their instruments. He also displayed the list of participating raingauges and showed that the majority of those models are now installed in SPICE sites, concluding that the RI intercomparison could represent a valuable source of information for SPICE procedures and analysis methodology.

5.3 Mr Mike Earle presented the data processing and quality control methodology for the derivation of reference datasets. Current recommendations for the processing and quality control (QC) of reference datasets from Geonor T-200B3 and OTT $Pluvio^2$ gauges and the approach taken to establish the associated methodology are detailed elsewhere, in the report from the SPICE IOC-4 meeting held in Davos, Switzerland and the WMO-SPICE Reference Report (currently in preparation). The current procedure has been developed primarily using 6 s data from Geonor gauges; additional testing is required to ensure filtering reduces noise to a comparable level in $Pluvio^2$ data, and to establish appropriate filter parameters for 1 min data. Another issue to be addressed is the specific method of averaging data from the three wires of a given Geonor gauge to produce a single, representative dataset.

5.4 In general, the methodology recommended for processing reference data can be applied to data from gauges under test, provided the necessary gauge-specific parameters and threshold values are defined. Additional methods have been identified for potential application, noting that manual intervention may be required for cases in which gauge performance has been significantly compromised. Additional details are provided in Annex III. The meeting agreed that this methodology could possibly be used beyond SPICE for implementation in operational networks. The meeting recommended considering to develop guidelines for operational use by network managers at the time of the completion of SPICE.

5.5 Ms Audrey Reverdin presented the event selection methodology. In order to analyze the site data sets, precipitation events must be identified. Because of the wide diversity of SPICE site climatologies and to achieve comparable site data sets, a uniform method is required. Following the proposal for event selection presented at the SPICE IOC-4 meeting in Davos, a refinement of the methodology with more precise steps including partly tested thresholds has been developed. Starting from the quality controlled reference data sets, 30-min events are selected through a 1-min based procedure if they fulfill several conditions, among which having sufficient accumulation during a 30-min window and being preferably selected from two independent sensors, e.g. a reference gauge for accumulation and an optical precipitation detector for occurrence of precipitation. As an output, an event file is created with all selected events listed, along with their characteristics and related parameters used for further analysis. The detailed methodology together with rationales for choices of thresholds and preliminary results is provided in Annex IV.

5.6 Prof. GyuWon Lee presented an assessment of observation uncertainties, error uncertainties, error modeling, catch ratio for solid precipitation accumulation and snow on the ground data based on measurement performed at Gochang and CARE. Uncertainty in snow measurement was quantified for manual and automatic observations from the CARE and Gochang sites. Standard statistical measures and two methods to quantify instrumental uncertainties were used: 1) equation of error propagation and 2) error modeling. The uncertainty in manual measurements of snow depth highly relies on quantization of measurement and can reach to 0.3 cm. The uncertainty in automatic measurement ranges from 0.5 cm to 3.2 cm. The random uncertainty of snowfall measurement varies from gauge types and the bias is also categorized in terms of types of windshields. The catch ratio is modeled by temperature-dependent linear functions and single/multiple sigmoid functions with Bayesian estimation theory. In general, the sigmoid function provides better performance than the linear function with more flexibility in terms of temperature and opens a new way of investigating multiple data sets from different SPICE sites. More details on the method presented by Prof. Lee are available in Annex V. This method will be tested and considered for use for the analysis of the SPICE data by the data analysis team.

5.7 Roy Rasmussen and Bruce Baker presented their proposal on how to apply the data of the R3 references (consisting of a pair of similar automated gauges, one shielded and one unshielded) to allow comparisons of all sites, including those without a DFIR. The key assumption behind the two gauge configuration reference is that the transfer function of an unshielded gauge is sufficiently different than an Alter shielded gauge and that the nature of this difference can be used to determine the appropriate transfer function to a DFIR for each site. The concept has been explained with the data from the Marshall site. Application and further evaluation of the method with data from other sites is in progress. This work will be described in the SPICE report on references.

5.8 A method using the minimum ratio between the unshielded gauge and single Alter gauge as a function of wind speed was devised. The meeting recommended that they further test this method considering not binning the data and/or providing more perspective on the underlying data points per bin.

5.9 The meeting agreed that this method will have to be further tested by the data analysis team, towards its possible application throughout the SPICE data analysis to compare the results of the different sites between them.

5.10 John Kochendorfer presented a methodology for assessing the uncertainty of instruments in field environment. Uncertainty in reference precipitation gauge measurements can be quantified by comparing like reference gauges to each other. Developments in the methodology used to do this and limits to the scope of such an approach were presented. Using 5 years of half-hour precipitation data from the Marshall, CO, USA testbed five Geonor gauges were compared to each other. It was found that in rain the effect of the wind shield on half-hour precipitation was undetectable, and measurements from the five different gauges could be treated as identical despite the fact that they were recorded within different wind shields. Relative uncertainty was quantified by calculating a scatter index (s_x), as the standard deviation of the gauge accumulation for every 30-min period of rain with more than 0.25 mm of precipitation. The average s_x for all five years of rain measurements was 0.13 mm. The s_x were also normalized by the total mean precipitation every half-hour ($\bar{x}$) to examine the percent random error in precipitation measurements, and was shown to increase as the total amount of 30-min precipitation accumulation decreased.

$$s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

Scatter index (s_x), where x_i is the individual half-hour precipitation measurements from each of the gauges, $\bar{x}$ is the mean of all five half-hourly measurements, and n is the number of gauges.

5.11 The meeting recommended that this method be further tested on the Pluvio^2 and manual gauge data.

5.12 Prof. Daqing Yang presented an overview of the current results on the assessment of R0 vs R1 and R1 vs R2 reference systems. The analysis of recent data showed a lower catch efficiency of the DFIR with respect to the bush gauge by approximately 5% with respect to data obtained during the first WMO Solid Precipitation Intercomparison (SPICE-1). The results from Caribou Creek, which uses an automatic gauge in a bush are quite different. More details are available in Annex VI.

5.13 The meeting recognized that beside the different methods used in Caribou Creek and Valdai, the type and density of the trees surrounding the gauges is also different. The meeting recommended that the Valdai site considers installing an automatic gauge in their bush to be able to better compare the results and the influence of the type of bush.

5.14 The meeting noted that the comparison between the bush gauge and the DFIR stopped at wind speed values of around 9 m/s. The meeting agreed that a cautious approach would have to be applied when recommending corrections to be applied to data from sites that are experiencing higher wing speeds.

5.15 Ms Mareile Wolff presented an overview over methodologies explored to date for assessing catch efficiency and derive adjustments by different teams, including some results. The document is based on reports, journal articles and input from the Col de Porte site team (France). It is attached in the Annex VII.

5.16 The meeting recognized that some of these methods could be considered for the evaluation of the SPICE data. Some of them might require to be adapted for example to account for shorter averaging/reporting time that are feasible nowadays.

5.17 John Kochendorfer made a presentation on determining the best time interval for deriving the catch efficiency. Because catch efficiency is evaluated as a ratio rather than a difference or a sum, the length of the time period over which the precipitation has accumulated must be chosen with some care. Rather than relying upon statistical analysis to inform this choice, a broader approach that takes into account the physical constraints involved in measuring precipitation is preferred. The time period must be long enough to calculate a mean wind speed that is representative of the entire site, acknowledging that some of the SPICE precipitation gauges are separated by a distance of more than 200 m. Because wind speed varies with turbulence, the period must be long enough to represent the average wind speed accurately rather than the effects of a small number of eddies. The ideal time period should also be long enough to allow for the accurate measurement of relatively low precipitation rates. At the same time, a time period that is too long will be subject to mesoscale, frontal, and diurnal changes in the wind speed and precipitation type, and such changes within an individual precipitation measurement period are undesirable for the creation of transfer functions. A shorter time period also provides more measurement periods for evaluation. Figures describing the prevailing scales of atmospheric motion were presented along with some analysis of Marshall snow data on the minimum threshold. The group concluded that it would continue to use 30-min time periods in the development of the event selection, while acknowledging that the choice of time period may change as the result of continued analysis.

5.18 The meeting recommended that the DAT tests the impact of using various averaging interval, and consider adopting a reporting interval that is widely used, if appropriate. The meeting further recommended that the methodologies that will be used for SPICE be also applicable (in real-time) to operational networks.

5.19 Samuel Buisan presented a summary of data fields and processing of Pluvio^2 data output and how they are used for SPICE and in operational applications. Pluvio^2 are available on a number of SPICE testsites. Some of the SPICE participants also use these instruments in their operational networks. The positive experience of SPICE participants, as well as the problems they encountered with these instruments were presented to the meeting.

5.20 Pluvio^2 offers a wide range of measurement outputs, in real time and non-real time, that are described in the manual. However, raw data is not available and measurement output comes only from processed data from OTT algorithms. The sampling frequency used in operational networks is 10 minutes and 1 minute and only some of the Pluvio^2 outputs are recorded on national archives. The sampling frequency which is used for SPICE project is 1 minute. Within SPICE, all measurement outputs are considered useful and are archived in the central SPICE database for further data analysis.

5.21 Yves-Alain Roulet presented topics relevant for tracking during SPICE tests to account among other on the robustness of the sensor tested. Challenges and relevant issues on all aspects concerning SPICE test sites have been assessed using site reports provided by the site managers. Among them, the following elements were found to have significant importance for several sites.

- Siting: Local phenomena or influences from surroundings can have large influence on the homogeneity of the measurements within a site. This has to be considered by DAT.
- Data collection: Data transfer to NCAR is still pending for several sites. Support is needed for some of them.
- Geonor: Some sites reported issues with noises on the data (due to vibration). Heating algorithm was also found to be an issue (see also 5.22).
- OTT Pluvio^2: Evaporation in non-negligible quantities, even with oil layer in the bucket, has been reported. Phantom accumulation is also a common problem (parameter “Accumulated NRT”). OTT will be asked to provide support in solving this issue.
- Heating configuration (see also 5.22): Issue to define one configuration for all climates, as too low heating may result in snow capping and high heating in evaporation loss.
- Anti-freeze and oil (see also 5.23): Use of anti-freeze mixture is recommended to prevent freezing in the gauge bucket. Oil layer is recommended to prevent anti-freeze and water evaporation. But inappropriate anti-freeze mixture and/or oil type can result in undesired

effects, such as non-realistic accumulation (hygroscopic effect), oil as separator between anti-freeze and incoming precipitation (oil viscosity too high).

5.22 Yves-Alain Roulet presented on behalf of Craig Smith a summary of the heating configurations used in various regimes and the results and experiences made to date within SPICE. Several sites provided information on their heating configurations. The need for specific configurations, varying in some aspects from the requirements that were defined during SPICE IOC-4 (Davos, 2013), has been recognized. In particular, the lower threshold for air temperature was set at -30°C in some locations. This was done to prevent the heaters from operating at very low temperatures and putting unnecessary stress on the 12 V power supplies that were only rated to -20°C. Geonor heater was reported to be not sufficient to keep the rim temperature at +2°C when air temperature drops below -5°C (Bratt's Lake and Caribou Creek). A solution for doubling the voltage will be evaluated. For the OTT Pluvio^2, some capping events were observed in Sodankyla. The issue has been solved by using the manufacturer's algorithm.

5.23 Yves-Alain Roulet presented on behalf of Jeff Hoover an assessment of antifreeze and oils highlighting which of their characteristics are relevant for their performance. This work is presently being prepared for publication in a scientific journal. A summary of the points investigated in the analysis is provided in Annex VIII.

5.24 Samuel Morin presented the plan for the analysis of the snow on the ground (SoG) data that is provided in Annex IX.

5.25 The meeting recognized that the SoG sites for SPICE have different target configurations for automated snow height sensors (natural grass, plastic mats, artificial grass, concrete etc.) and this corresponds to different national observation strategies and practices. Crossing the influence of weather conditions, snow height sensor type and brand, and target configurations, is beyond the scope of SPICE SoG given the number of sites and instruments available. Therefore, the meeting recommended that:

(1) to maximize the consistency between the two SPICE observation winters (2013-2014 and 2014-2015), sites keep the same configuration for the upcoming snow season 2014-2015.

(2) the SoG team reviews existing status of target configurations in the largest possible number of national networks (within and outside SPICE), and reviews the body of knowledge addressing the impact of target configuration (national reports, publications, etc.).

(3) where and when possible, report on the behavior of existing snow targets used during the 2013/2014 and 2014/2015 SPICE observation winters, using photographic or observer information to note accumulation or melt differential from the surrounding area.

5.26 The NCAR data archive facility is the primary data repository for SPICE data including for SoG data. Site managers are urged to transfer SoG data to NCAR (including manual data) so that the data can be analyzed by SoG analysis teams. The meeting recommended that all further SoG data analysis is carried out on the basis of NCAR data repository (see other NCAR-data-related actions)

5.27 Recognizing that NCAR role should focus on data archival primarily, the meeting proposed that SoG quality control (QC) be carried out outside of the NCAR platform. The SoG team will review automated QC approaches based on thresholds and elaborate on manual QC complements; implementation of QC will be carried out by Craig Smith.

5.28 For the SoG data analysis, emphasis should be placed on data analysis approaches spanning the largest possible number of site. The coordination of SoG data analysis will be taken care of by SoG team leader(s) (Craig Smith and Samuel Morin). It was recognized that GyuWon Lee has carried out work on CARE and Gochang data and that he may add additional sites to his analysis. Other data analysis will be taken care of by individual groups with as many as possible interactions. The meeting appreciated the interest of a number of persons to take part in this analysis, who include Craig Smith, Daqing Yang, Barry Goodison (Canada), Samuel Morin

(France), GyuWon Lee (Rep. of Korea), Timo Laine, Osmo Aulamo, Leena Leppanen, Niina Puttonen (Finland), Antonella Senese (Italy), Yves-Alain Roulet, and Audrey Reverdin (Switzerland).

5.29 For SoG sites featuring at least a R3 precipitation reference, the event selection will be based on the designated precipitation gauge measurements (and computed on the NCAR platform). For SoG sites without a precipitation reference, there is no strong need to define precipitation events. These sites will be used to assess the comparability of SoG sensors regardless whether precipitation is ongoing or not. Of course, ancillary information from the sites (albedo, variation of snow height, etc.) can be measured to assess whether precipitation is ongoing or not if need be.

5.30 For sites willing to do so, additional SoG manual measurements can be performed (density of fresh snow, snow pit measurements including grain size etc.) and contribute to specific analysis points including whether snow precipitation can be inferred from SoG measurements.

5.31 Christian Zammit made a presentation on current solutions and configurations at the remote SPICE sites. Contributions were sought from remote SPICE site managers in regards to: i) configuration of the gauge; ii) power source design; and iii) failure identification and remediation. Contributions were received from some sites only. The answers regarding the gauge configuration are mainly related to the fact that the sites are located either on outcrop or on boulders. In addition consent conditions associated with the establishment of the site limits what can be done at a site (Mueller Hut). In term of power supply all responding sites, but Mueller Hut that is fully battery operated, are connected to the grid. It is to note that all those sites have battery pack back up for data loggers and DC power instrument. All the sites rely on real time communication system to assess/identify any potential failure of any instrument through real time post-processing. At some sites, webcam is used to assess the existence of physical damage on site, or automatic alarm systems are in place to cater for instrument failure (Weissfluhjoch). This is completed by the use of network of observers associated with the station. For failure remediation process, for most of the remote sites the only solution is to go on site to assess the issue with replacement gauges. However this requires access to the site, availability of personal as well as appropriate budget.

5.32 The meeting recognized that the team would have to develop recommendations for the operation of instruments at remote sites, which would have to take into account the experience made on all sites. The meeting therefore encouraged the sites of Italy and Nepal to contribute their experience to this topic and to actively contribute to it. The meeting appreciated the offer of Samuel Morin to contribute the experience of the Col de Porte automatic stations that is fully running on batteries.

5.23. Following the proposal made during SPICE IOC-4 (see SPICE IOC-4 Final report sec. 7.17), Major Emanuele Vuerich presented a field calibrator recently developed by the WMO-CIMO Lead Centre in Italy. Its aim is to calibrate catching type gauges by reference low intensities similar to snowfall intensities. It is based on a double-syringe pumping system and provides intensities in the range 0.5 mm/h – 40 mm/h for gauges of different collector sizes (the range can be extended up to 190 mm/h if needed). Mj Vuerich explained how it works, showed its performance in terms of repeatability and uncertainty and proved the suitability of this device for field calibrations at low intensities.

5.33 The meeting was pleased that this system would be tested at the CARE site and compared with the results from traditional calibration methods towards developing recommendations for the practical calibration of precipitation gauges.

Breakout sessions

5.34 Break-out sessions were organized to address several topics as follows:

- Derivation of reference datasets (Pluvio/Geonor issues, timescales,…),
- Data availability and derivation of event data (defining the event file table, how to get all data at NCAR, derivation of non-reference data,…),

- Methods of assessment (Uncertainty, correction factor vs catch ratio, precipitation type and particles microphysics, …),
- Operational aspects of the intercomparison (provider contact, consistency, shadowing effects, …).

5.35 Each of the beak-out groups reported to the meeting. A summary of the recommendations made by the meeting are listed below. Additional details are also provided in the workplan provided in Annex X.

Report of breakout session on derivation of reference datasets

5.36 The group recommended using the Accumulated NRT data from Pluvio^2 gauges as an additional 'quality controlled' data stream for input to event selection algorithm. Results will be compared with those using the Bucket RT data (current practice), and differences quantified.

5.37 Simple arithmetic averaging of data from the three wires of a given Geonor gauge will be used to generate the first iteration of reference datasets. Further testing will be conducted in the interest of using more advanced methods (weighted averaging, majority voting) in later iterations.

5.38 To establish a 'common ground' for subsequent precipitation event selection and analysis, the noise in reference datasets for Geonor and Pluvio^2 gauges should be of similar magnitude. Testing of different filter widths will be conducted using 6 s and 1 min datasets for both gauge types, and the residual noise will be compared between gauge types. This work may prompt revisions to the recommended data processing approach for reference datasets.

5.39 The time resolution of reference datasets should reflect the specific application (e.g. event selection, research, forecasting, satellite validation, climate). Reference datasets with a resolution of 1 min will be generated to serve as the basis for all applications; these datasets can be aggregated/averaged to longer time intervals, as required. Aggregation/averaging to longer time intervals can also be considered as a means of reducing uncertainty/noise in reference datasets.

Report of breakout session on data availability, data transfer to NCAR, and derivation of event data

5.40 In order to perform the data analysis, it is critical to have all the data stored in a central respository, using agreed formats. A gap analysis of the data available at NCAR will be performed. As some sites have not been able to transfer the data to NCAR yet, or are experiencing problems with the data transfer, the meeting decided to provide support to these groups and assigned persons in charge of helping each of them with the aim of transferring all data to NCAR by 1 Sept. 2014.

5.41 The meeting recommended that NCAR prepares webpages for all sites that do not have one yet and handles the incoming data. The site managers are responsible to provide to NCAR the information needed to establish their webpage and to validate the data available on the NCAR archive.

5.42 The meeting requested the site managers to indicate which precipitation detector/disdrometer is to be used for the derivation of the reference data and which wind sensors are to be used for gauge height wind and 10 m wind.

5.43 For the derivation of the event data, there is a need to document the implemented data aggregation for all instruments under test related to the derivation of the accumulation during events. The data aggregation for the sensors under test will lead to events files that will be implemented at NCAR. The meeting recognized that it would be valuable to include in the data analysis the ability to detect and assess false reports from the sensors under test

5.44 For the (manual and automatic) quality control (qc) assessment, the break-out group recommended to:

- Develop and implement qc procedures for sensors under test data, and to define fields/thresholds for automatic qc,

- Sites to store and share the information on site events that could affect the data and would trigger flags. These could be used for data analysis and manual qc.
- Assess the possibility of storing the data/site logs (site changes tracking) at NCAR, to be uploaded by Site Managers at least once per season, at the end of the season.
- Expand the site data logs (site changing tracking) to include events that would affect the data and upload site changing tracking log to NCAR.

5.45 It is expected that links from maintenance records to data will be established. NCAR will develop a proposal on how to upload maintenance information and modification of the data format including these quality control flags, for consideration by the IOC.

Report of breakout session on methods of assessment

5.46 The break-out group recommended developing transfer functions towards DFIR for catch ratio dependent on wind and temperature for selected gauge configurations including data from all sites. Two different approaches for deriving those transfer functions will be tested, one of them being based on the Bayesian statistics. The results from the two methods will be compared. Dependencies on other variables, such as intensities, will be searched for. The same functions will be applied to all sites to look for site-specific biases and to possibly group sites according to observed differences. These grouping will be compared to those obtained from the R3 analysis.

5.47 The break-out group recommended describing statistics from the event files (how many events, total snow amount, wind, per winter per site; average catch ratio per gauge per site per winter). The group also recommended to assess the type of QC that need to be applied before the event assessment and to compare the gauge accumulation from event file vs gauge accumulation from beginning and end of the season.

5.48 The derivation of the precipitation type using some combination of sensors (disdrometers, snowfall, radiation, ….) will be investigated. This will be tested using manual observations from some sites. It was recognized that transfer functions solely based on precipitation type (and not on temperature) might be more accurate, but will be working for just a few stations as necessary data are not available everywhere.

5.49 The major problem for the linkage between different reference type R0-R1 and R1-R2 is the availability of data. The meeting therefore recommended that the site of Valday consider installing an automatic gauge in the bush. The performance of the two automatic gauges in the bush at Caribou Creek will be investigated and a method similar to that used in Valdai will be tested. For the R1-R2 analysis, all available data need to be analysed, including some measurement that were recorded prior to SPICE, such as in Jokioinen and Bratt’s Lake. For the R3 analysis, the break-out group recommended that more data be analysed with the technique proposed to date, preferably with and without wind-binning and that site-categories be identified based on similarities in the results.

5.50 Further work will be performed to assess gauge uncertainties. Multiple gauges of same configuration at one site (basically references) will be compared using different methods/approaches.

Report of breakout session on operational aspects of the intercomparison

5.51 The break-out group recommended that a registry of problems (and associated solutions) experienced with the instruments under test be established in the form of a shared google-doc. It would be accessible to all site managers, so that they could enter relevant information and seek guidance on problems experienced.

5.52 For problems that are appearing at many sites, the break-out group recommended that one person/site be tasked to contact the manufacturer, and to report on recommended solution to concerned site managers and the IOC, so that appropriate decisions can be made if changes of configuration/software version were recommended by the manufacturer.

5.53 The meeting recommended that all manufacturers be informed by the project lead of the prolongation of the experiment, of the possibility to obtain precipitation data, to encourage them to liaise with sites (look at data, visit sites) and to inform the project leader on any problem that they are facing concerning the instruments they provided.

5.54 The meeting also recommended that the project leader informs all site managers on the need to closely liaise with instrument providers, to recall the publication guidelines (inform provider in case of reporting results from instruments under test) and that an appropriate time for feedback from manufacturers on intended publications would be typically 3 weeks. The meeting recommended that a short version of the SPICE disclaimer be developed for use in abstract that are published without a full publication. The meeting recommended that site managers increase their communication with the instrument provider to ensure they are confident with the extension of the project and with the quality of the data produced by their instruments.

NCAR data archive

5.55 The meeting agreed that it is crucial that all the data be available at the NCAR archive. It is important to understand which data is missing on the archive and why, so that the data analysis can proceed.

5.56 The meeting recommended that NCAR focuses first on ingesting all the data. Relevant algorithms (QC, event selection, …) should first be developed off-line and implemented in NCAR once they are mature enough.

5.57 Roy Rasmussen reiterated that NCAR has the mean to host all the data (raw and QC'd), but also recognized that it would be valuable to mirror the NCAR site at another location.

5.58 The meeting requested that all manual observations be also transmitted to the NCAR data archive using the procedure for reporting manual measurements, both for precipitation amount and for snow on the ground, so that that the data analysis team will have all data to carry out the evaluation.

5.59 The meeting agreed that all the data analysis should be carried out using the QC'd data retrieved from NCAR to ensure that the data was properly QCed and that all groups are working with the same type of data.

5.60 The meeting agreed that the focus of the work should now be placed on analyzing the data from all gauges rather than on refining further the QC methodology and investigating further the differences in/specificities of the reference instruments used for the experiment. In view of ensuring a timely publication of the final results of the experiment, the meeting also agreed to concentrate the analysis on 30-min data for the moment and that other sampling intervals (10-min, 1-hour) would only be considered later on in case of need.

Some organizational aspects

5.61 The meeting recalled that the instruments provided for the experiment have to be operated and maintained strictly according to the instrument user guide.

5.62 The meeting encouraged the site managers to regularly communicate with the instrument providers to ensure they are confident with the quality of the data from their instruments and that they are in agreement with the extension of the intercomparison.

5.63 The meeting recommended that all sites interested in continuing the experiment beyond 2015 start making necessary bilateral arrangements to enable them to continue operating relevant instruments provided by manufacturers on their site.

5.64 Yves-Alain Roulet informed the meeting that he would be discussing the problem of phantom accumulation in Pluvio^2 with the manufacturers. The phantom accumulation was observed in the absence of rain in some gauges. The meeting invited Mr Roulet to inform the team on the recommendations of the manufacturer and to make recommendation to the team on whether specific actions (such as maybe a change of software) should be considered for the instruments under test, and also for the Pluvio^2 that are used as reference

Development of a future workplan

5.65 Based on the discussions above, the meeting developed a detailed workplan (provided in Annex X) for continuing the data analysis of the SPICE data during 2014-15, putting now a much larger focus on the analysis of the instruments under test.

5.66 It was agreed that a similar approach would be used as last year, distributing the work among the team members and in particular among the DAT members. A combination of small focused teleconferences, full DAT teleconferences, and SPICE team teleconferences will be used to advance the work in an efficient manner. Also some small dedicated working group meetings will be envisaged in case of need and opportunities to progress some specific aspects, while a full SPICE IOC meeting is likely to take place in Q2 2015.

6. OTHER BUSINESS

6.1 A visit of the Sodankylä site took place on Wednesday 21 May 2014. The meeting welcomed this opportunity to visit a SPICE testsite and the measuring facilities available. Following this visit, the meeting made recommendations to improve the configuration of the site.

6.2 Ms Leena Leppanen presented the variety of snow measurements performed in Sodankylä and specific data quality checks and data analysis carried out with these data. A summary of her presentation is provided in Annex XI.

6.3 Ms Niina Puttonen presented the Sodankylä SPICE site and some preliminary results. A summary of the site configuration is provided in Annex XII.

Publications

6.4 The meeting recalled that it had encouraged the publication of results from SPICE sites in scientific journal, both prior and after the publication of the final report. The meeting noted that publications in scientific journals would help in disseminating the results from SPICE beyond the WMO community and that they were complementary to the SPICE final report. Also, the SPICE Final Report could be building on publications of partial results of the experiment. In this context, the meeting supported the proposal to organize a journal special issue, such as Atmospheric Measurement Techniques. The meeting welcomed the offer from Samuel Morin to coordinate the establishment of such a special issue dedicated to SPICE.

6.5 The meeting was informed that at the forthcoming International Union of Geodesy and Geophysics (IUGG) General Assembly from June 22 to July 2, 2015 in Prague, the International Association for Hydrological Sciences (IAHS), the International Association of Meteorological and Atmospheric Sciences (IAMAS) and the International Association of Cryospheric Sciences (IACS) will likely convene a Joint Symposium entitled: "JH2 Precipitation measurements, instrumentation and statistics at all scales". The meeting encouraged SPICE team members to contribute to this session. The deadline for abstract submission is 31 January 2015.

6.6 The meeting agreed that a short version of the disclaimer should be used when abstracts discussing SPICE data are published without an associated full publication. The meeting requested that the following short disclaimer be included in such abstracts: “The data presented in this work were (optional: partly) obtained as part of the World Meteorological Organization's Solid Precipitation Intercomparison Experiment (SPICE). Analyses and views described are those of the author(s) and do not necessarily represent the official outcome of SPICE.”

Resources

6.7 With consideration to the concern of SPICE-IOC expressed for the efforts required for data analysis and the need of resources, Italy proposed to the meeting and IOC the availability of a PhD student, Mr Roberto Azzoni, to contribute to SPICE-DAT and help address the issues related to the data analysis. He works at the University of Milan and in close cooperation with Ms Antonella Senese. The University and the national committee EvK2-CNR (both operating SPICE sites of Forni Glacier-Italy and Pyramid Observatory-Nepal) will support him. This contribution could be made available for other aspects of SPICE data analysis additionally to those related to Forni Glacier and Pyramid Observatory-Nepal.

6.8 The meeting welcomed the offer from representatives of Finland, Italy and Switzerland to contribute additional support to the project through the involvement of some of their staff members and students for the evaluation of the SPICE data. The meeting noted the need to identify how these resources could best contribute in meeting the overall objectives of SPICE, including in making best use of the data collected on the SPICE site of those participants as they have the best insight in the potential of their site's data. This will be achieved through discussion between the project team, DAT chair and interested parties.

7. DRAFT REPORT OF THE SESSION

The meeting reviewed the draft report of the session and decided to finalize it and to approve it by correspondence.

8. CLOSURE OF THE SESSION

The session closed on Friday, 24 May 2014 at 17:30 hours.

ANNEX I

LIST OF PARTICIPANTS

Ms Rodica NITU ***Chair, IOC-SPICE*** ***SPICE Project Leader*** ***Site manager of CARE (Canada)***	Environment Canada 4905 Dufferin St. TORONTO, ON, M3H 5T4 **Canada** Tel.: +1 416 739 4133 Fax: +1 416 739 5721 rodica.nitu@ec.gc.ca
Mr Osmo AULAMO ***Site Manager of Sodankylä (Finland)***	Finnish meteorological Institute Tähteläntie 62 99600 Sodankylä **Finland** Tel.: +358 40 5181541 osmo.aulamo@fmi.fi
Mr Shane BILISH ***Site Representative of Guthega Dam (Australia)***	Snowy Hydro Limited PO Box 332 Cooma NSW 2630 **Australia** Tel.: +61 2 6453 2658 Fax: +61 2 6453 2052 Shane.Bilish@snowyhydro.com.au
Mr Samuel BUISAN ***Site manager of the Aramon – Formigal site (Spain)***	State Meteorological Agency (AEMET) Aragon Regional Office Pase del Canal 17 50007 Zaragoza **Spain** Tel.: +34 976 696959 sbuisans@aemet.es
Mr Michael EARLE	Environment Canada 45 Aldeney Dr. Queens Square Dartmouth, NS B2W 0A8 **Canada** Tel.: +1 902 426 4477 Michael.earle@ec.gc.ca
Ms Henna-Reetta HANNULA ***Site Representative of Sodankylä (Finland)***	Finnish Meteorological Institute Arctic Research Centre Tähteläntic 62 99600 Sodankylä **Finland** Tel.: +358 406707112 Henna-reetta.hannula@fmi.fi
Ms HeeJin IN ***Site Manager of Gochang Observatory (Republic of Korea)***	Korea Meteorological Administration 61 16-gil Yeouidaebang-ro, Dongjak-gu SEOUL 156-720 **Republic of Korea** Tel.: +82 2 836 2385 Fax: +82 2 836 2386 hjin@korea.kr

Mr Maciej KARZYNSKI ***Site Manager of Hala Gasnienicowa (Poland)***	Institute of Meteorology and Water Management Department of the Observation and Measurement Service Section of Nivology National Research Institute Cracow's Branch 14 Piotra Borowego Str. 30-215 Cracow **Poland** Tel.: +48 12 6398132 Fax: +48 12 6398201 maciej.karzynski@imgw.pl
Ms HyeLim KIM ***Site Representative of Gochang Observatory (Republic of Korea)***	Korea Meteorological Administration 61 16-gil Yeouidaebang-ro, Dongjak-gu SEOUL 156-720 **Republic of Korea** Tel.: +82 2 836 2385 Fax: +82 2 836 2386 Kimhl@korea.kr
Dr John KOCHENDORFER ***Invited Expert***	NOAA / ATDD 456 South Illinois Avenue Oak Ridge, TN 37830 **United States** tel.: +1 865 576 1238 fax: +1 865 576 1327 john.kochendorfer@noaa.gov
Dr Arkadiy KOLDAEV ***Site Representative of Volga River (Russian Federation)***	Research and Production Association Roshydromet "TYPHOON" 3, Pervomayskaya Street 141700 DOLGOPRUDNY Moscow Region **Russian Federation** Tel.: +79260328665 avk425@mail.ru
Mr Timo LAINE ***Site Representative of Sodankylä (Finland)***	Finnish Meteorological Institute **Finland** timo.laine@fmi.fi
Prof. GyuWon LEE ***Invited Expert***	Kyungpook National University 80 Daehak-ro, Bugku, Daegu Korea (RoK), 702-701 **Republic of Korea** Tel.: +82 2 836 2385 Fax: +82 2 836 2386 gyuwon@kma.ac.kr
Ms Leena LEPPANEN ***Site Representative of Sodankylä (Finland)***	Finnish Meteorological Institute Arctic Research Centre Tähteläntic 62 99600 Sodankylä **Finland** Tel.: +358 406707133 Leena.leppanen@fmi.fi

Mr Yves LEJEUNE ***Site Representative of Col de Porte (France)***	Météo-France Centre National de Recherches Météorologiques Centre d'Etudes de la Neige CNRM-GAME/CEN 1441, rue de la piscine 38400 St Martin d'Hères **France** tel: +(33) 4 76 63 79 24 fax: +(33) 4 76 51 53 46 yves.lejeune@meteo.fr
Mr Yuri MELNICHUK	ROSHYDROMET State Hydrological Institute Valdai Branch Pobedystr. 2 175400 Valdai Novgorod Region **Russian Federation** Tel.: +7 964 593 52 78 yurimelnichuk@yandex.ru
Mr Samuel MORIN ***Site Manager of Col de Porte (France)***	Météo-France Centre National de Recherches Météorologiques Centre d'Etudes de la Neige CNRM-GAME/CEN 1441, rue de la piscine 38400 St Martin d'Hères **France** Tel: +(33) 4 76 63 79 03 Fax: +(33) 4 76 51 53 46 samuel.morin@meteo.fr
Mr Antti POIKONEN ***Site Representative of Sodankylä (Finland)***	Observation Services Finnish Meteorological Institute PO Box 503 00101 Helsinki **Finland** Tel: +358 29 539 5747 Fax: +358 29 539 5703 timo.laine@fmi.fi
Ms Niina PUTTONEN ***Site Representative of Sodankylä (Finland)***	Finnish Meteorological Institute Aviation and Military Weather Service PO Box 8178 (Rovaniemi Airport) 96101 Rovaniemi **Finland** Tel.: +358 503493931 Niina-puttonen@fmi.fi
Ms Audrey REVERDIN ***Site Representative of Weissfluhjoch (Switzerland)***	Météo Suisse Station Aérologique Case postale 316 CH-1530 PAYERNE **Switzerland** Tel.: +41 26 662 62 82 audrey.reverdin@meteoswiss.ch

Dr Yves-Alain ROULET ***Member, IOC-SPICE*** ***Site Manager of Weissfluhjoch (Switzerland)***	Météo Suisse Station Aérologique Case postale 316 CH-1530 PAYERNE **Switzerland** Tel.: +41 26 662 6258 Fax: +41 26 662 6212 yves-alain.roulet@meteoswiss.ch
Mr Francesco SABATINI ***Member, IOC-SPICE***	CNR-IBIMET Institute of Biometeorology Via Giovanni Caproni, 8 50145 Florence **Italy** tel.: +(39) 055 303 3711 tel.: +(39) 055 522 6029 fax: +(39) 055 308 910 f.sabatini@ibimet.cnr.it
Ms Antonella SENESE ***Site Representative of Forni Glacier (Italy) and Pyramid International Laboratory Observatory (Nepal)***	University of Milan Earthe Department “A. Desio” Via Mangiagalli 34 20133 Milan **Italy** Antonella.senese@unimi.it
Capt. Emanuele VUERICH ***Chair, CIMO ET-II*** ***Invited Expert***	Italian Met Service – Air Force Centre of Meteorological Experimentations Via Braccianese Claudia, km 20,100 00062 Vigna di Valle (Rome) **Italy** tel.: +39 06 99 88 7702 fax: +39 06 99 87 297 vuerich@meteoam.it
Dr Mareile WOLFF ***Site Manager of Haukeliseter (Norway)***	Norwegian Meteorological Institute P.O. Box 43 Blindern Henrik Mohns plass 1 0313 Oslo **Norway** Tel.: +47 2296 3185 Fax: +47 2296 3050 Mareile.wolff@met.no
Dr Daqing YANG ***Site Manager of Caribou Creek (Canada)***	National Hydrology Research Centre (NHRC) Environment Canada 11 Innovation Boulevard Saskatoon, SK S7N 3H5, **Canada** tel: +(1 306) 975 6483 fax: +(1 306) 975 5143 daqing.yang@ec.gc.ca
Mr Christian ZAMMIT	National Institute of Water and Atmospheric Research Iokyle Street Riccarton 8011 **New Zealand** christian.zammit@niwa.co.nz

REMOTE PARTICIPATION FOR PARTS OF THE MEETING BY TELECONFERENCE	
Bruce BAKER ***Vice-President of CIMO*** ***Member, IOC-SPICE***	NOAA / ATDD 456 South Illinois Avenue OAK RIDGE, TN 37830 **United States** Tel.: +1 865 576 1248 Fax: +1 865 576 1327 bruce.bakernoaa.gov
Mr Kohei HONDA ***Site Representative of Joetsu and Rikubetsu (Japan)***	Japan Meteorological Agency 1-3-4 Otemachi Chiyodaku Tokyo 100-8122, **Japan** Tel: +81-3-3211-6019 kohei_honda@met.kishou.go.jp
Paul JOE ***Invited Expert***	Environment Canada 4905 Dufferin Street Toronto Ontario M3H 5T4 **Canada** Tel.: +1 416 739 4884 Fax: +1 416 739 4211 paul.joe@ec.gc.ca
Roy RASMUSSEN ***SPICE Project Co-Leader*** ***Site Manager of Marshall (USA)***	National Center for Atmospheric Research 3450 Mitchell Lane Boulder, CO 80301 **United States** Tel.: (1 303) 497-8430 Fax: (1 303) 497-8401 rasmus@ucar.edu
Craig SMITH ***Site Manager of Bratt's Lake (Canada)***	Environment Canada 11 Innovation Boulevard Saskatoon, SK S7N 3H5, **Canada** Craig.Smith@ec.gc.ca
Floortje HEUVEL ***Site Representative of Weissfluhjoch (Switzerland)***	Météo Suisse Station Aérologique Case postale 316 CH-1530 PAYERNE **Switzerland** floortje.heuvel@meteoswiss.ch
WMO SECRETARIAT 7 bis, avenue de la Paix Case postale 2300 CH 1211 Geneva 2 Switzerland	**IMOP website** http://www.wmo.int/pages/prog/www/IMOP/IMOP-home.html
Ms Isabelle RÜEDI *Head, Instruments and Methods of Observation Unit* *WMO Observing Systems Division*	Tel.: +41 22 730 8278 Fax: +41 22 730 8021 E-mail: IRuedi@wmo.int

---

ANNEX II

REPORT OF THE CHAIR

Following the SPICE-4 meeting in Davos, Switzerland, in June 2014, the project has made significant progress in acquiring a comprehensive data set at its participating sites, and advanced the assessment of methodologies for data processing and analysis, with a primary focus on the derivation of the reference data set.

Project Extension

The project team has asked the CIMO Management for support for continuing SPICE for an additional season, for the winter of 2015, which would further contribute to establishing the data foundation which will allow for a comprehensive assessment of the results.

At its meeting in March 2014, the CIMO Management Group considered positively the arguments presented for an additional season, the winter of 2014/15. As part of the decision, CIMO Management requested the following:

- the SPICE formal field tests be completed in 2015 in the Northern and Southern Hemispheres.
- the Final Report of the Intercomparison be ready for publication in 2016.
- no new participants be included in SPICE, following the 2013 acceptance of 5 new sites and 5 additional instruments.
- Financial resources supporting data analysis would depend on the availability of funds in the CIMO Trust Fund, as a result of contributions made by Members.

Acquiring a comprehensive data set

Site	Commissioning of sites	In SPICE since	precipitation (PA)/ snow on ground (SoG)	References	Instruments from Providers (Yes/No)	Data available to project (e.g. via NCAR) as of May 2014
Guthega Dam (Aus)	Y	2012	PA	R3	N	N
Bratt's Lake (CA)	Y	2012	PA/SoG	R2;R3	Y	12/13; 13/14
CARE (CA)	Y	2012	PA/SoG	R1;R2;R3, MANUAL SoG	Y	12/13; 13/14
Caribou Creek (CA)	Y	2012	PA/SoG	R0a; R2, R3	Y	12/13; 13/14
Tapado (Chile)	N	2012	PA	R3	Y	N
Sodankylä (FI)	Y	2012	PA/SoG	R2;R3	Y	12/13; 13/14
Col de Port (FR)	Y	2013	PA/SoG	R3	Y	Partially 13/14
Forni Glacier (It)	Y	2013	SoG	SoG	N	N
Joetsu (Jp)	In progress	2013	PA	R2;R3	N	N
Rikubetsu (Jp)	In progress	2013	PA	R2;R3	N	N
Gochang (KR)	In progress	2013	PA/SoG	R2;R3	N	N
Pyramid (Nepal)	In progress	2013	SoG	SoG	N	N
Mueller Hut (NZ)	Y	2012	PA/SoG	R3	N	N
Haukeliseter (NO)	Y	2012	PA	R2;R3	Y	partially
Hala (PL)	In progress	2012	SoG	SoG	Y	partially
Valdai (Rus)	Y	2012	PA	R0;R1	N	12/13; 13/14
Volga (Rus)	In progress	2013	PA/SoG	R1	N	N

Weissfluhjoch (CH)	Y	2012	PA/SoG	R2;R3	Y	partially
Marshall (USA)	In progress	2012	PA/SoG		Y	12/13; 13/14

2013: inclusion additional SPICE sites

The IOC-SPICE-4 accepted 5 additional sites for participation in the Intercomparison, based on the fact that each represent environments of interest for the measurement of solid precipitation and snow on ground, and strong indication was given of the rigour of the experiments. These are Col de Porte (France) organised by Meteo France, Gochang (Republic of Korea) organized by Korean Meteorological Administration, ARAMON-Formigal (Spain) organized by Spanish State Meteorological Agency (AEMET), Forni Glacier (Italy) and Pyramid International Observatory (Nepal) organized by EVK2CNR (Italy).

2013: inclusion of additional Instruments

Several additional instruments have been added in 2013 to the intercomparison. Given their operating principles, these instruments are considered of interest to the scientific and operational community. These are:

- PWD 53/PWD33 (Vaisala): Sodankyla
- PWD 52 (Vaisala): Sodankyla
- FS11P: Sodankyla (1)
- TPS3100 hotplate (Yankee): Bratt's Lake (1), Sodankyla (1) and Haukeliseter (1)
- FROS-D (Univ. Colorado): Weissfluhoch (1) and CARE (1)
- ANS410/H (Eigenbrodt GmbH): Marshall (1)
- SMH30 (Jenoptik): Col de Porte (2), in addition to the 2012 submission
- SR50AH (Campbell Scientific Canada): Col de Porte (2), in addition to the 2012 submission
- CS725 (Campbell Scientific Canada): Caribou Creek (1), in addition to the 2012 submission.

Report on the Configuration of the SPICE Working Field Reference System

The report on the SPICE Field Working Reference System is the first major deliverable of SPICE. The report will introduce the configuration of the Field Working Reference Systems agreed by the SPICE IOC, and implemented on the participating sites. The reference datasets derived from the Field Working Reference System(s) will be used as basis for comparison and reporting of the performance of all instruments under test. These results will be used for the derivation and verification of the transfer functions to be developed to account for gauge undercatch, and for the characterization of the instruments under test.

The analysis methodology for SPICE is under development, and it will build on the work conducted for the derivation of the reference dataset.

The SPICE reference reports has been developed with significant contribution from a large number of project team members, a proof of the commitment and engagement of the team members.

The report is expected to be published by WMO in the second half of 2014, and it should be distributed through all available means, to ensure that diverse feedback is received and other communities of experts have the chance to comment on the SPICE approaches to references. The goal is that, at the time of the Final Report, the issue of reference will not hinder the focus on the project results.

It is noted, however, that the preparation of this report has taken significant project resources. Moving forward, we need to strike the right balance for increasing the efficiency in using the available resources. It is recognized that the experts participating in SPICE are key contributors to the programs of their organizations and they have many other responsibilities to fulfill.

Seasons 2012/13 and 2013/14 overview,

Based on input received from: Guthega Dam, Weissfluhjoch, Haukeliseter, Sodankylä, CARE, Bratt's Lake, Caribou Creek, Joetsu, Rikubetsu, Mueller Hut.

Successes:

1. Full experiments being run on almost all sites with minimum data loss.
2. Collaboration with other organizations is evolving: Col de Porte, Formigal, Gochang, Weissfluhjoch, towards organizing and running the tests, resolving issues, and data analysis.
3. Newly accepted sites have been operational for the 13/14 season, and the representatives have been actively engaged in the project.
4. Increased effort on the planning of the SoG component of SPICE, including the expansion of SoG experiments at CARE, Col de Porte, Gochang.
5. Haukeliseter: Final report and paper of the national wind correction of precipitation measurements with a resulting adjustment equation based on three winter data sets from Haukeliseter. The data set includes a very large amount of high-wind cases, which extends the validity of the adjustment function.
6. Resolution of some field issues: e.g. Guthega Dam.
7. Active collaboration and engagement between teams.

Challenges:

1. Heater configuration: in cold conditions the heaters are not always able to always maintain the temperature of the rim at 2 °C (at 50 W, controlled power supply), while when using the Geonor heaters (200 W continuous), it appears to result in over heating.
2. Use of oils and antifreeze mix tailored to the climate conditions.

Interactions with the Instrument Providers:

Sites have made available the data from instruments provided by Instruments Providers, to a large degree. To date the interaction with the Instrument Providers on the review of data of their instruments has not been consistent.

More sustained engagement with manufacturers/IPs is desired, to ensure that the instruments are operated effectively throughout the intercomparison. The project team needs to assess the impact of project extension on the temporary import arrangements for the instruments included in the test.

Some manufacturers have visited the SPICE sites, reviewing the installation and operation of their instruments. CAE visited CARE and Marshall in 2013. Sodankylä was visited by Vaisala in 2013, EML 2013. Belfort Instruments visited Weissfluhjoch regarding the installation, upgrade and calibration of their gauge, in 2012.

Concerns:

Items of Interest for the future:

- Use of disdrometers data : Potentially compare between the disdrometer data and the 3D ultra-sonic anemometer.
- Impact on the data quality of shield mounted on pole (vibration).
- Pluvio^2 behavior under specific weather conditions, and understanding of its data.
- A weighing gauge with higher resolution in precipitation rate for weaker snowfall rates than the ability of the present instruments are frequently occurred in cold regions.
- Further assessment of heating impact and the impact of evaporation

SITE Lessons learned:

- The connection between international SPICE and local partners needs to be improved.
- Quality AND quantity of anti-freeze and oil is decisive for a good balance between evaporation prevention and easy transmission of precipitation through the oil layer.
- Web Camera worked well in cold winter and dark night.
- Wind sensors should be mounted as undisturbed as possible (top of the mast, even for gauge-height wind sensors) in a distance to everything else. Installation at gauge is not recommended, as wind measurements are highly affected by the construction.
- It takes a lot of effort to keep things going; still, there is no perfect installation.

Plans for 14/15:

Most of the sites plan to continue the work in the configuration of the 2013/14 season. Additionally, a few sites plan new developments:

1. Formigal:
    a. Installation of a DFAR.
2. Col de Porte:
    a. Implementing a temperature control for the rim heater.
    b. Using the same antifreeze mixture on all weighing gauges.
    c. Replace the Tretyakov shield of the OTT Pluvio^2, perhaps with a single alter shield.
3. Guthega Dam:
    a. R3 gauge heaters temperature adjusted to operate for T_{rim} < 2 °C.
4. Weissfluhjoch:
    a. Replace the unheated OTT Pluvio^2 gauge (R3 reference) with a heated version.
    b. Change the oil (currently linseed oil).
    c. Joint project with Swiss Federal Polytechnical School and SLF for a field campaign during winter 2014-2015: deployment of an X-band radar scanning above Weissfluhjoch.
    d. New instruments will be installed at Weissfluhjoch SPICE site: a 2D Video Disdrometer and a Parsivel: it will permit the comparison of particle type characterization methods.
    e. Potential purchase and installation of a MASC (Multi-Angle Snowflake Camera).
5. Bratt's Lake:
    a. Install Yankee Hotplate sensor.
    b. Disdrometer data available for the entire 2014/2015 winter.
6. Haukeliseter:
    a. Complete installation of the hotplate.
7. CARE:
    a. Install (potentially) a 3rd DFIR with a Pluvio^2 (R2P)

Plans for data analysis from a national perspective

1. Weissfluhjoch:
    a. No data analysis from a national perspective, except some analysis on Pluvio^2 accumulation problems, with strong implications on our operational network.
    b. Data analysis for SPICE focused on hydrometeor types: joint project with EPFL and SLF.
2. Haukeliseter:
    a. Active participation in the data analysis team.
    b. If resources allow, analysis of the precipitation detectors and sensors.
3. Formigal:
    a. To consolidate the site for SPICE project and future projects.
    b. To compare reference gauges and DFAR results with instrumentation used on automatic weather stations on AEMET observing network.
    c. Installation of new emerging technologies (disdrometers).
4. Guthega Dam:
    a. Existing study with Monash University using data from Guthega Dam. SPICE data (i.e. from new R3 reference gauges) not currently included.
    b. Potential for some additional work to be done as part of a PhD position looking at the use of precipitation data in the forecasting of snowmelt and streamflow.
5. Col de Porte:
    a. Develop a common framework to engage multiple groups interested in SPICE results.
6. Bratt's Lake:

   a. Assess and develop transfer functions for an un-heated Pluvio[2] in a single Alter shield.

7. Caribou creek:
   a. Work with DAT on reference gauge assessment, DFIR vs. Bush gauge.
   b. Combine our data with those from Bratt's lake and CARE.
   c. Gauge specific data analyses of the merged data.
   d. Snow depth and SWE data analyses and comparison with gauge measurements.
8. CARE:
   a. Develop configuration of Double Alter shields and assess against R1 and R2.
   b. Develop specifications for snow depth sensors and SD targets based on the knowledge base developed with the 13/14 and 14/15 tests.
   c. Develop recommendations for the deployment of non-catchment type sensors.
   d. Better understand the impact of heating of WG; develop recommendations for operational use.

SPICE-5: Meeting desired outcomes

The SPICE-5 meeting is expected to lead to the following outcomes:

- Define an overall Work plan for June 2014-June 2015.
- Define a detailed work plan June-Oct 2014.
- Identification of resources assuming specific tasks.
- Data analysis plan to take into account the diversity of objectives and instrument types (catchment, non-catchment, SoG), and address the development of transfer functions and linking of results for similar gauges operated on different sites.
- Analysis of the similarities and differences between gauges used as part of the Field Working Reference System.
- Clarity of engagement and outcomes from all sites.

ANNEX III

DATA PROCESSING AND QUALITY CONTROL METHODOLOGY FOR THE DERIVATION OF REFERENCE DATASETS

Mike Earle, Audrey Reverdin, Mareile Wolff, Craig Smith, Samuel Morin, and Rodica Nitu

Version history

Version	Date	Notes
1	May 7, 2014	Initial draft prepared by Mike Earle with input from Audrey Reverdin and Craig Smith

1. Development of processing and quality control methodology: approach and methods tested

Data filtering

- Initial testing of methods using sample site datasets from Geonor gauges in R2 and R3 configurations, with 6 s and 1 min temporal resolution.
- **Qualitative assessment** of filter performance.
- Identified max/min filter (range check) as an effective method for removing outliers; ‘jump’ filter as an effective method for identifying and removing data spikes, and identifying and flagging potential baseline shifts (e.g. due to ‘dumps’ of accumulated precipitation from the orifice into the bucket, or due to bucket emptying).
- Methods tested for removal of high-frequency noise: moving average, Savitzky-Golay (3rd order polynomial fit), Gaussian filter in time domain (moving window), Gaussian filter in frequency domain (employing fast Fourier-transforms to generate periodogram, allowing for identification and removal of high-frequency noise).
- Gaussian filter in time domain shown to be effective; decision to test further alongside moving average filter, which was both effective and widely-applied, historically.
- Artificial datasets generated for known precipitation rates to allow for **quantitative assessment** of filter performance using moving average and Gaussian methods; true precipitation signal known for artificial events.
- Tested moving average and Gaussian approaches with various parameters (filter width, standard deviation of Gaussian distribution) using artificial datasets for rates covering range of expected conditions, between 0.6 mm/hr (light precipitation) to 30 mm/hr (heavy precipitation).
- Used root-mean-square error (RMSE) to assess filter performance and make recommendations for processing approach.
- Current recommendations for filtering approach based on this assessment are outlined in Section 2.

Compensating for temperature effects

- Diurnal solar radiation/temperature variations impact precipitation amounts reported by Geonor and Pluvio^2 gauges.
- Influence for Geonor gauges generally characterized by a decrease of 0.1 mm for every 10 degree increase in temperature. Adjustment functions can be derived from daily accumulation-temperature relationships in non-precipitating conditions; however, variation of relationship with specific transducer, bucket amount, incident solar radiation (cloud cover, solar elevation and azimuth), and the fact that temperature is not recorded at the sensing element complicate the adoption of these functions for continuous temperature compensation in the processing of Geonor reference datasets.
- Hysteresis has been observed in the accumulation-temperature relationship for Pluvio^2 gauges, precluding the determination of adjustment functions as described above.
- Given the difficulty in compensating for temperature effects, it was decided that this would not be part of the processing methodology for reference datasets; rather, the emphasis was placed on characterizing the temperature during precipitation events (allowing for sorting of events by temperature behavior in subsequent analysis).
- It should be noted that the Pluvio^2 firmware applies static temperature compensation to processed outputs from these gauges, which will be considered as part of the intercomparison.

Averaging Geonor data

- To facilitate subsequent analysis, the accumulated precipitation measurements from the three transducers in a given Geonor gauge must be averaged to generate a single, ‘representative,’ gauge output.
- Averaging methods considered: simple arithmetic averaging; weighted averaging based on noise, in which noisy wires are weighted less when computing averages; and majority voting, in which pairs of transducers with consistent outputs are considered to be ‘correct’ (majority rules), and used to justify the exclusion of data from a third wire with output differing beyond a set threshold.
- Key consideration: the load (weight of bucket and accumulated precipitation) is shared among the three transducers; hence, giving less weight to, or excluding, a given wire may bias the averaged output from the ‘true’ value.
- For this reason, arithmetic averaging is currently recommended, with a flag to identify potential transducer performance issues (e.g. noise, offsets in magnitude, differences in response to precipitation).
- In cases of severe transducer performance issues, the other averaging methods, or manual intervention, may be required. Additional testing on site data is required to establish guidelines for implementing procedures beyond the simple arithmetic approach.

2. Processing of reference datasets from Geonor and Pluvio^2 gauges

Recommended approach

- Max/min filter (range check), with maximum thresholds corresponding to bucket capacity for each gauge.

- Jump filter to remove spikes exceeding threshold value (TBD); data flagged if number of consecutive jumps exceeds another threshold value (also TBD), indicating a potential shift in baseline.

- For Geonor and Pluvio^2 precipitation data with 6 s temporal resolution, a moving average or Gaussian filter with a width of 2 minutes is recommended to mitigate the influence of high-frequency noise; for the latter, the standard deviation of the Gaussian distribution is recommended to be either equal to, or one half of, the filter width.

- For Geonor and Pluvio^2 precipitation data with 1 minute temporal resolution, a moving average or Gaussian filter with a width of 8 minutes is recommended to mitigate the influence of high-frequency noise; for the latter, the standard deviation of the Gaussian distribution is recommended to be either equal to, or one half of, the filter width.

- Arithmetic averaging of precipitation data from three transducers of a given Geonor gauge, with flag(s) to indicate potential performance issues.

Caveats

- Filter methodology has been tested and developed primarily using 6 s Geonor datasets. Subsequent testing on 6 s Pluvio^2 datasets has indicated that the recommended filter width may not be optimal, given the lower frequency of noise observed; however, the 2 minute filter width was still recommended for these data to maintain the same time response as the Geonor approach.

- Limited testing has been conducted using 1 minute datasets for both reference gauges; current recommended filter width not well-established, and should be used with caution.

- No methods are currently employed to mitigate/compensate for the effects of temperature or evaporation.

3. Applicability of processing and quality control methods to other gauge data

- Max/min filter (range check) can be broadly applied to data from all gauges, and requires only the maximum and minimum value thresholds for implementation.

- Similarly, the jump filter can be broadly applied to other gauge data, provided threshold values for the maximum increase in a given parameter per data point (jump threshold) and number of subsequent ‘jumps’ indicating a baseline shift (if applicable; for example, due to system resets) are defined.

- For any gauge data subject to high-frequency noise, moving average or Gaussian filters can be applied. Prior to implementation, tests should be conducted on site datasets to establish a rough guideline for the filter parameters (filter width, standard deviation of Gaussian distribution).

- For gauges with multiple sensing elements/transducers, similar averaging methods can be used. Arithmetic averaging is the most straightforward approach, but majority voting is an established method of averaging/quality control for systems with independent, redundant, sensing elements.

4. Additional data processing and quality control procedures for consideration

- Characterization of missing data and values removed by filtering to facilitate diagnosis of gauge and/or site issues.

- Application of gauge-specific corrections and offsets for processing of snow on the ground (SoG) data. The SoG analysis plan indicates that offsets are to be applied for temperature and gauge geometry, as well as offsets for zero snow depth and zero snow water equivalent (SWE). The offsets can be applied on an ongoing basis, while the offsets are determined at the end of a measurement season, and so can only be applied retroactively.

- Temperature artefact compensation in weighing gauge data by matching increases in bucket weight with corresponding decreases in bucket weight. This method has been shown to be effective in situations without appreciable evaporation, in which decreases in bucket weight can predominate. As decreases in temperature are not necessarily balanced by corresponding increases within a given dataset, this method should be viewed as a means of mitigating the influence of temperature-induced variations, rather than fully compensating for, or correcting, these variations.

- Ideally, all quality control methods and processing could be standardized and automated to ensure that all gauge data are processed in the same way, using the same set of rules; however, given the variability of the natural phenomena being measured, and the limitations of the measuring devices and site installations involved, it is difficult to define rules for processing data in all potential scenarios. For this reason, manual intervention may be required in some cases. To the extent possible, manual processing should be governed by well-defined criteria in order to ensure consistent application across datasets. These criteria will necessarily need to be defined at a later stage of analysis; if the rules for when and how to intervene were clear at this stage, they could be coded and automated. Development of these rules/criteria would help to inform future iterations of data processing and quality control methods.

ANNEX IV

EVENT SELECTION METHODOLOGY
(DEVELOPED FOR THE REFERENCE)

Audrey Reverdin, Craig Smith, Mareile Wolff, Mike Earle, Samuel Morin, Barry Goodison

Version history

Document	Version	Date	Notes
Event selection for reference gauges data	1	May 11, 2014	Initial draft prepared by Audrey Reverdin with input from Mareile Wolff and Mike Earle
Event selection for SoG data	1	April 10, 2014	Initial draft prepared by Craig Smith with input from Samuel Morin and Barry Goodison
Whole document	2	May 14, 2014	Revision following comments from Mike Earle and Yves-Alain Roulet

Table of Contents

1. Description of Concept

- In order to analyze the site data sets, precipitation events must be identified. Precipitation events can be highly variable, and the different site climatologies add to this diversity. Existing detection methods are very individual and are not very well documented. In order to achieve comparable site data sets, a uniform method is required.
- Pre-analyses conducted using datasets from several sites have shown that false detected precipitation events will add a lot of noise to the data set, confounding subsequent analysis. In order to analyse the nature of the precipitation, and possibly relationships to meteorological parameters, a relatively low false-alarm rate is required for the precipitation event identification process.
- Starting from the quality controlled reference data sets, the following methodology for the aggregation of precipitation events was developed in order to create analysis-ready (level 3) datasets, which are as comparable as possible across sites.

2. Chosen Algorithm

2.1 Algorithm description

- The event selection algorithm enables the selection of precipitation events using quality-controlled data from two instruments:
    - *A precipitation detector* : indicates if precipitation occurs or not (yes/no output).
    - *A reference gauge* : collects precipitation and gives the corresponding accumulation.
- The algorithm creates an output file in which all of the events selected through the process are listed, along with their characteristics and related parameters used for further analysis.
- The algorithm is based on 1 min data. It has been decided to average 6 s data before using them for event selection.
- The flowchart in **Figure 1** illustrates the different steps of the event selection. It is composed of three columns:

  Column 1 : uses data from both the precipitation detector and reference gauge.

  Column 2 : describes the method without implementing data from a precipitation detector (sensor not available, not working for some reason, not stable, in maintenance or under revision); events selected using this approach are based only on data from the accumulation gauge.

  Column 3 : indicates to go further through the algorithm process.

  Columns 1 and 2 indicate different ways for the data to pass through, depending on whether or not there is a sensitive and reliable precipitation detector. The specific approach followed for a given event will be tracked in the output file by means of a data flag.
- The event selection algorithm follows three main steps :
    - **First step – Starting Point**

      The first step searches for possible starting points of precipitation events. Starting points are detected when sufficient accumulation occurs in a ten-minute period after a selected data point. The procedure has the following features:

      1) *Rate check over 10 minute moving window*

         Check if there is an accumulation in the reference gauge over a ten minute period. A potential starting point is identified if the accumulation is greater than, or equal to, a threshold value (currently defined as 0.1 mm). If insufficient accumulation is recorded (less than the threshold value), the 10 min window moves one minute forward.

      2) *1st minute check*

If sufficient accumulation is detected, the first minute of the 10 min period is checked for recorded precipitation. If either the precipitation sensor is positive or a positive accumulation in the gauge is observed for that minute, then the minute is marked as a possible starting point.

Note: Both methods (precipitation detector or accumulation gauge) are set equal and only one is required to identify a potential starting point. Tests have shown that requiring a logical "and" for these two conditions will identify significantly less starting points. Falsely identified starting points will be screened out in the following "event check" step (see below), and are therefore not a problem.

- **Second step – Event Selection**

  The second step looks at the 30 minutes following a potential starting point. To be selected as an event, the 30 min window has to fulfill the following two conditions :

  1) *Net precipitation duration sufficiently long*

     The number of minutes during which precipitation is detected has to be more than 60% of the window time, i.e. more than 18 minutes. The precipitation duration is primarily calculated based on precipitation detector data (first column) by looking at the number of "YES" cases that happened during the 30 minute period. In cases where precipitation detector data are not available, the duration check can be satisfied if the number of 1 minute reference gauge data points with increasing accumulation exceeds the same 60% frequency threshold (second column).

  2) *Accumulation of reference gauge sufficient*

     The total accumulation in the reference gauge during the 30 minute period must meet or exceed a defined threshold. This threshold rate has been set to 0.25 mm over 30 minutes when a reliable precipitation detector is available (first column), and to 0.5 mm over 30 minutes if such an instrument is not available (second column).

     A more strict rate check has been imposed in cases when independent validation by a precipitation detector is not available, to help ensure the veracity of selected events.

  If these two required conditions are fulfilled, the window is set as a 30 min event; if not, the algorithm goes to the next potential starting point.

Note: The possibility to identify an event based solely on accumulation data (second column) has been introduced in order to allow the algorithm to run in cases where a precipitation detector is not available or reliable. However, to keep track of the specific method by which events are selected, events selected in this manner are flagged. This flag will appear in the output file to caution the analyst that it's maybe not a real event, or at least that it has been selected without independent validation by a precipitation detector. On this basis, events that aren't flagged are judged to be more probable and reliable.

- **Third step – Event parameters**

  When an event is identified, the algorithm calculates several different parameters to characterize the event in detail for further analysis.

*Note: Among the list of required parameters (see **Figure 1**), the net precipitation duration of the event is calculated both with the precipitation detector data and the reference gauge accumulation data. The initial idea was to have the net duration of the event for cases without a reliable precipitation detector, but reporting both allows for a consistency check between the results from each approach.*

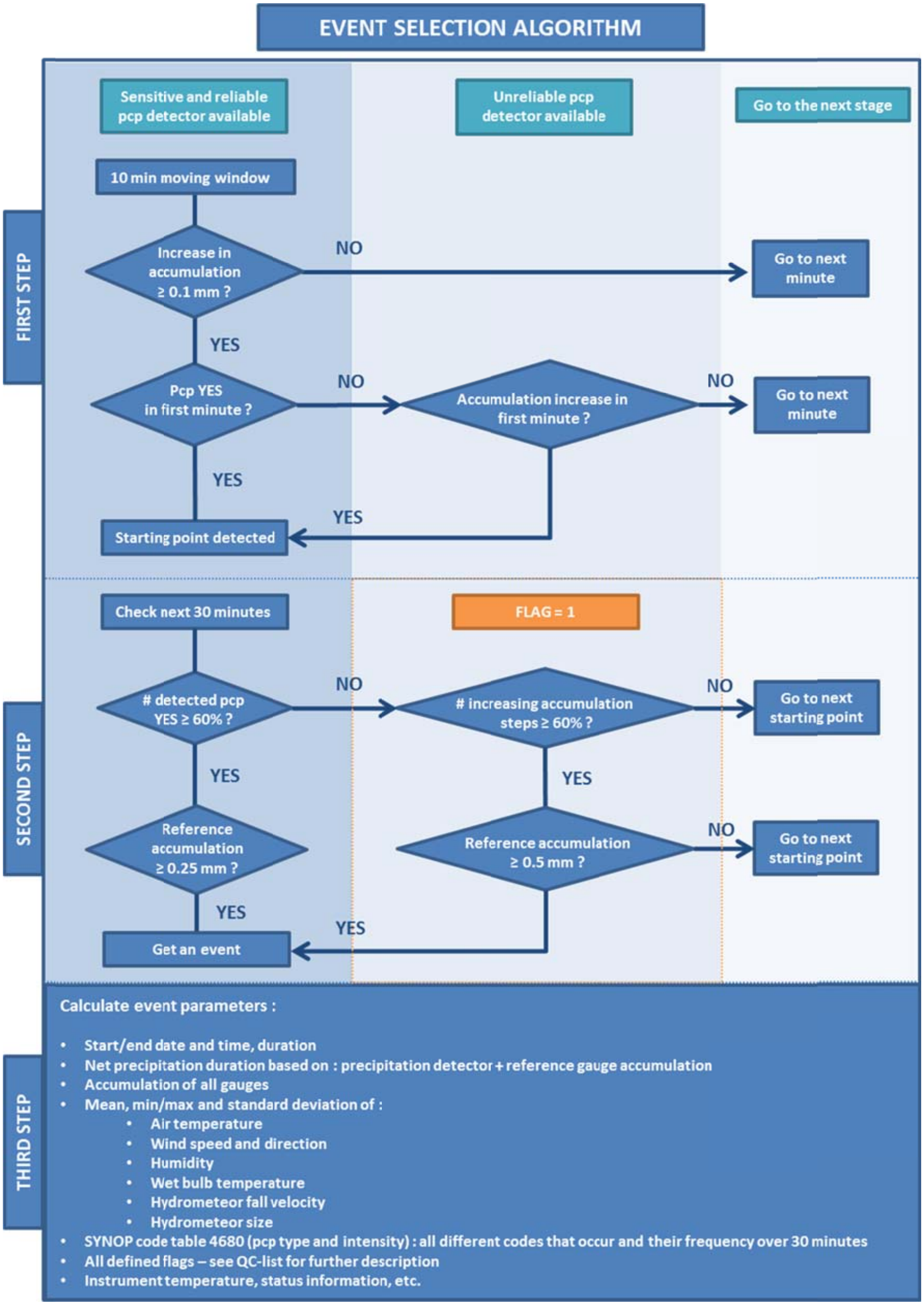


Figure 1 : Event selection algorithm flowchart.

2.2 Thresholds and periods selected for algorithm

The following thresholds and periods to be used in the algorithm were selected following discussions of the SPICE Data Analysis Team (DAT) during the 4th SPICE meeting in Davos (Switzerland). These thresholds/periods generally represent best estimates, and will be tested and refined further throughout the analysis.

First step :

- *10 minute window*

  Ten minute windows were chosen because teams in Norway and in the US used to apply 10 min running averages to the data as a very simple short-wave-filter, which reduces the noise such that typical accumulation amounts are clearly detectable. It is expected that this will also be valid for the filtering process now applied to the precipitation data. Furthermore, the 10 min period will allow for a more precise detection of the actual starting point of a precipitation event than the use of the complete event period (30 minutes). However, the 10 minutes is rather conservative and is only based on experiences from unfiltered Geonor measurements in Norway and USA; no tests have been performed using Pluvio[2] data.

- *Rate check : accumulation ≥ 0.1 mm*

  A precipitation event requires an accumulation greater than zero. A threshold value of 0.1 was selected because residual noise in the data could potentially lead to a false starting point if the threshold was only set to zero.

Second step :

- *30 minute window*

  This fixed period of time is needed to compare events from different sites. A period of 30 minutes is short enough to allow for stable conditions (temperature, wind speed, etc.) during the event. Furthermore, analysis has shown that selecting 30 minute periods optimize the balance between the number and significance of events selected. It might be possible to change the period between 10 minutes and 1 hour, but moving to 12 or 24 hours periods is not reasonable for the current algorithm, and would require a different process.

- *Net precipitation duration ≥ 60% of time*

  It has been decided within the SPICE project that the precipitation does not necessarily need to be continuous during the event period, but should still occur over a 'significant' portion of the event time. The choice of 60% of the time was set as a starting point, and needs to be further evaluated.

- *Rate checks : accumulation ≥ 0.25/0.5 mm*

  The selection of a threshold accumulation of 0.25 mm over 30 minutes is based on previous experience in characterizing events by participating sites/analysts, with threshold rates of 0.2 – 0.3 mm/30min typically employed. The threshold should be sufficiently high to distinguish 'real' events from measurement artefacts, but not so long as to significantly reduce the number of events identified.

  A higher threshold rate of 0.5mm/30min was set for the cases where precipitation detector data are not reliable, in an attempt to compensate for not having independent validation of events.

Note: The first goal of DAT in defining such thresholds is to select 'real' events, meaning events that have sufficiently high accumulation and last sufficiently long to make sure they are reliable, and not very light or spurious events (that are more difficult to characterize) or measurement artefacts (e.g. artificial accumulation due to temperature effects).

2.3 Inputs : Reference gauges as "event-selectors" and precipitation detector

- The event selection needs to be done with data from the most accurate and reliable gauge on site to ensure the best quality of the selection. It is thus appropriate to choose the on-site reference gauge as the dedicated *event-selector* :
  - For S2 sites, the event-selector should be the reference gauge in the DFIR (R2 reference). In cases of sites with multiple DFIR-references, only one will be used for the event selection.
  - For S3 sites, the shielded gauge of the R3 reference will be used as the event-selector (as it should collect more snow than the unshielded gauge).

Note: The choice of the event-selector should follow these recommendations, but the DAT suggests to keep the possibility to change the event-selector, if necessary, throughout the duration of the project.

- The precipitation detector used to select an event should be an optical precipitation detector as defined during the 4th SPICE meeting in Davos (IOC-SPICE-4 Report, ANNEX IV, p. 4).
  - For S2 sites, it is located near the reference gauge within the DFIR, between the Alter shield and the inner wooden fence.
  - At sites without a DFIR-fence it should be at a wind-protected place or shielded by a suitable shield.

2.4 Outputs : Parameters in event file

- The list of parameters in the output event file should be as consistent as possible for all sites to facilitate comparative analysis; however, since no two sites have identical equipment or sensor setups, some adaptation is required. The DAT decided to start with a parameter list common for all sites (checking what is available on all sites), followed by the individual site parameters. All possible parameters and their determination for the event file should be gathered in a separate file. This parameter files (eventually one per SPICE site) will be used to generate event lists based on each site's available instrumentation.
- The list of parameters the DAT agreed to start with includes :
  - Parameters characterizing the event :
    - Start and end date and time of the event
    - Duration of the event (should be 30 minutes)
  - Parameters used to select the event :
    - Reference gauge accumulation
    - Net precipitation duration (from precipitation detector and accumulation increases)
  - Parameters from other instruments during the event :
    - Accumulation of all gauges on site
    - Air temperature
    - Wind speed and direction
    - Humidity
    - Wet bulb temperature
    - Hydrometeor fall velocity
    - Hydrometeor size
    - SYNOP codes and their frequency
    - Housekeeping parameters : instrument temperature, status information, etc.

Note: Of course, this list should be flexible as the DAT gains new experiences during the analysis and might want to look at additional parameters.

- In addition to these parameters, two additional columns should be included in the output event file for flags :
  - Event selection flag : indicates if the event has been chosen with reliable precipitation detector data or not (value of 0 or 1, respectively);
  - QC flag(s) : indicates any potential shifts in the measurement baseline and/or gauge performance issue over the 30 minute period.

Note: The goal of these flags is to inform the event file user of events which may be less reliable and/or impact subsequent analysis.

- Some of the above listed parameters (air temperature, wet bulb temperature, hydrometeor fall velocity and size distributions, SYNOP codes) have been chosen to describe the precipitation type during an event. The named parameters are those that are in principle available on all sites.

2.5 Preliminary tests results

- Different tests have to be performed on the event selection algorithm to determine the best values to choose on the various steps of the whole algorithm process and to assure the good quality of event selection. Iterative process have to be made to get the optimal values for accumulation rate and net precipitation duration thresholds as well as for time windows.
- Although the DAT aims to test all thresholds in a consistent and systematic way, some preliminary testing took place on few algorithm components :
  - *Accumulation rate threshold test*

    The accumulation rate threshold in the second step defined as 0.25mm/30min has been changed into 0.2mm/30min to assess the difference in terms of the number of selected events. The result shows (see **Figure 2**) that the lower threshold better characterize the entirety of the event (start and end of the event taken into account) while the upper threshold tend to miss part of the event if the accumulation rate is low. As a consequence, the number of 30 min events selected with the lower threshold is greater than for the upper threshold. The DAT concluded that a 0.2mm/30min could be therefore used for the event selection, but further tests need to be done in order to know the balance between the capacity to well describe the event and the chance to select too light accumulation that would not originate from a real event.

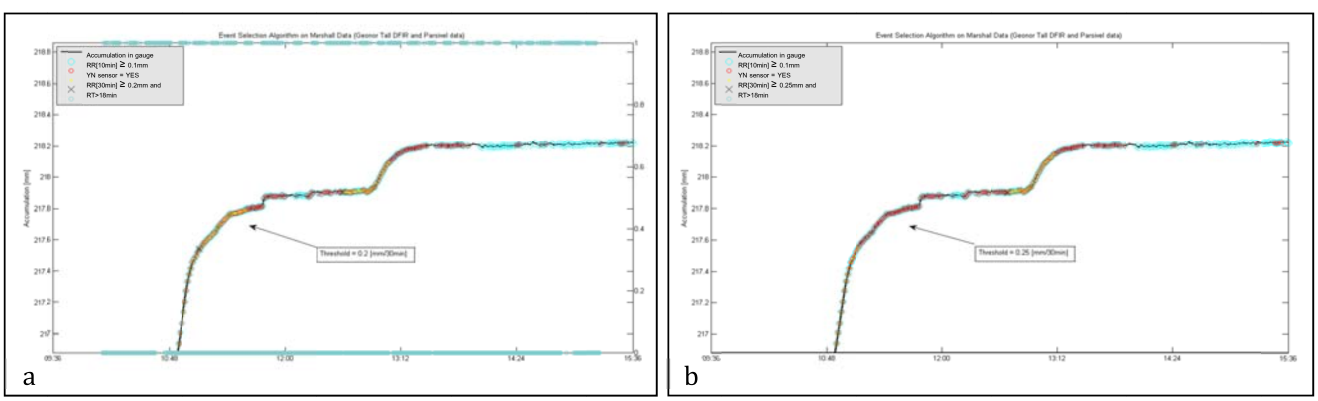


Figure 2: Two event selection results on the same dataset with an accumulation rate threshold of (a) 0.2 mm/30min and (b) 0.25 mm/30min on the second step of the algorithm.

- *Flag test*

  The original idea of flagging data (for unreliable precipitation detector) comes from some data of Bratt's Lake site where three of their Geonor gauges did accumulate significant precipitation quantities, while the YES/NO sensor didn't record the whole event (see **Figure 3.a**). Without the second column of the flowchart (see **Figure 1**), the algorithm wouldn't have recorded any event as the 60% of net precipitation duration based on precipitation detector would not be fulfilled. With the possibility to select events only based on accumulation data, events are then recorded and flagged (see **Figure 3.b**).

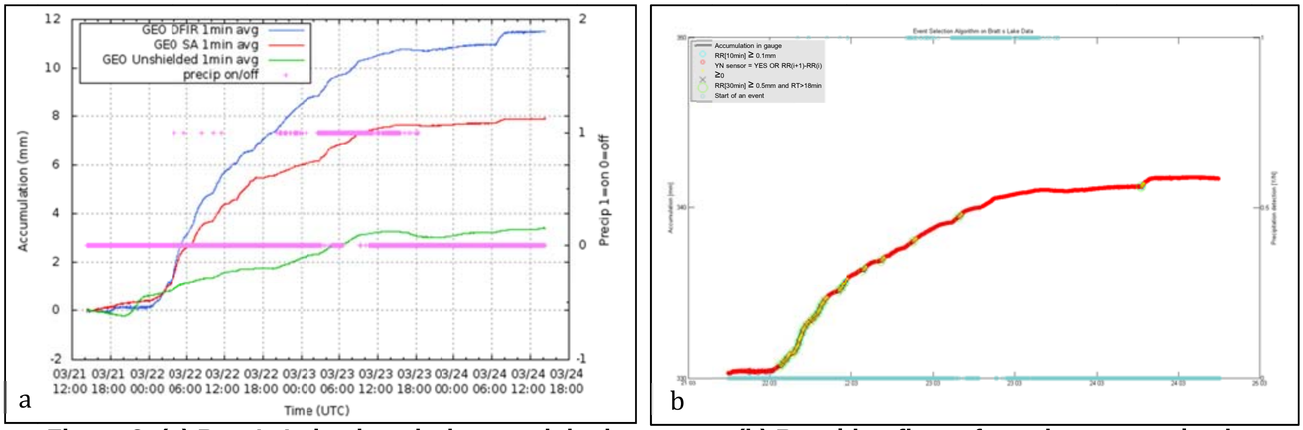


Figure 3: (a) Bratt's Lake data during precipitation event ; (b) Resulting figure from the event selection algorithm with flagged events.

- *Event-selectors test*

  In order to test stability of the algorithm and to assess differences between several potential event-selectors on one site, the algorithm has been applied on Marshall data of four gauges placed within DFIRs (three Geonor gauges, one Pluvio^2 gauge). Over one month of data, the algorithm retrieves globally the same amount of events for the four gauges (see **Figure 4**). This leads to the conclusion that either gauge could be taken as the event-selector.

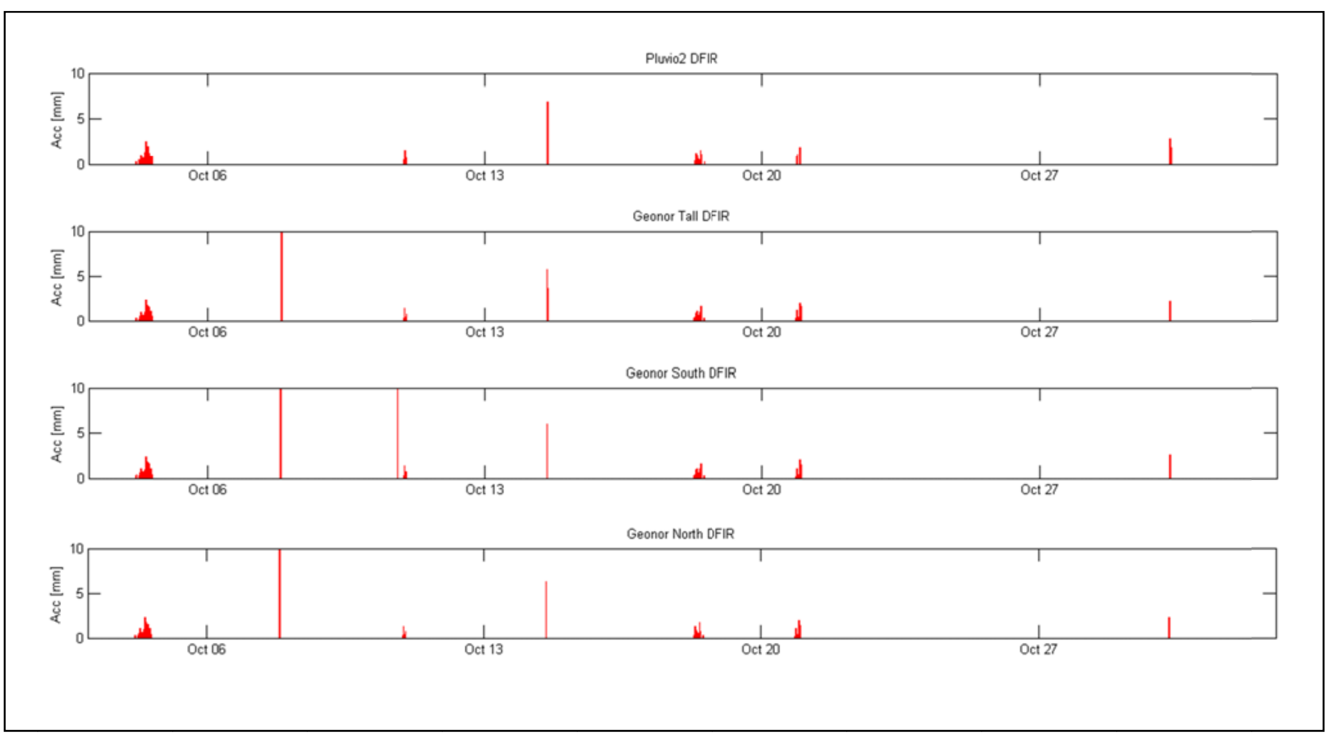


Figure 4: Events selected for four different configurations of DFIR at Marshall site.

3. Limitations and Possible Alternatives

- *Limitation of event selection methodology :*

  The main purpose of the described event identification process is to find clear precipitation data in order to understand and describe physical relationships between the precipitation loss due to wind speed, direction, precipitation type and possibly other meteorological parameters. In order to make this process applicable at all stations (for comparability), some generalizations had to be made. The set limits for duration of an event and minimum rates are chosen to be rather high in order to make sure that only real precipitation events are caught at all sites. The drawback of this approach is that not all events will be identified completely; the less intense on-/offset of an event might be missed.

- *Fuzzy Logic as a possible alternative* :

  The DAT is aware of alternative techniques to identify precipitation events. For example, a fuzzy logic approach does not consider precipitation in terms of a definite Yes or No, but rather, gives a quantified probability of precipitation occurrence. Capturing additionally the cases where precipitation was not only 100% certain, but quite likely, will undoubtedly bear lot of valuable information. At this stage of the analysis, though, it was decided to start with a rather simple process in order to focus analysis resources on working with the produced data. Based on the further analysis of the precipitation events, a possible review of the selection process will be evaluated.

- *Outlook on data analysis :*

  After describing the physics/statistics of the event-only dataset and the derivation of one or more transfer functions, an evaluation of a more operational data set has to be performed. At that stage an evaluation of the “maybe”-cases will be performed.

4. Event Identification for Snow-on-Ground (SoG) Data Analysis

Craig Smith, Samuel Morin, Barry Goodison

- *Limitation of using solely SoG data for event selection :*

  It is recognized that event identification of SoG events using only SoG instrumentation would be problematic. The accumulation of snow on the ground is dynamic by nature, with small accumulations forming, melting, and re-forming, often more dependent on ground temperature (for shallow snowpack) and snowpack properties (compaction) than solely precipitation accumulation. Wind redistribution can also cause issues with snow amount increasing or decreasing under the sensor without actual precipitation. However, it is still necessary to identify and delineate snowfall events to meet some of the SoG objectives.

- *Use of gauge event selection algorithm for event identification :*

  We propose to use the gauge identification of snowfall events (see **Figure 1**) as a starting point for event identification for SoG purposes. Using gauge (and ancillary) data and the gauge event selection process, snowfall event start and end times can be identified for use in the SoG analysis. This will be a useful starting point for intercomparisons between gauge and SoG instrumentation. These can also serve as a starting point for determining the quantitative and temporal thresholds of the SoG instrumentation.

- *Proposed thresholds and strategy :*

  The proposal is to start with SoG measurement thresholds of 1 cm depth at a temporal resolution of 1 hour. Analysis will refine these thresholds and this process : A SoG accumulation event will begin following the initiation of a gauge snowfall event (as indicated in **Figure 1**) AND upon the exceedance of the minimum depth threshold using SoG measurements averaged over a 1 hour window.

ANNEX V

UNCERTAINTY ANALYSIS OF SOLID PRECIPITATION MEASUREMENTS

GyuWon Lee, Yong-gu Kim, Jeong-Eun Lee
Kyungpook National University, Korea (ROK)

The uncertainty can be originated from many different reasons and can be divided into different terms such as instrumental uncertainty and observational uncertainty. This short document describes two methods in quantifying these uncertainties: error propagation and simple statistical error modeling.

1. Error propagation

The uncertainty of solid precipitation measurement is first calculated by the standard statistics such as the Bias Error (BE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Bias Removed Root Mean Square Error (BRRMSE) that are defined as the following equations:

$$\text{Bias Error (BE)} = \frac{1}{N}\sum(y-x)$$

$$\text{Mean Absolute Error (MAE} = \frac{1}{N}\sum|y-x|$$

$$\text{Root Mean Square Error (RMSE)} = \left[\frac{1}{N}\sum(y-x)^2\right]^{0.5}$$

$$\text{Bias Removed RMSE (BRRMSE)} = \left[\frac{1}{N}\sum(y-x-BE)^2\right]^{0.5}$$

where x and y are snow depth or snowfall rate and the N is the number of data for a given pair. The NBE, NMAE, NRMSE, and NBRRMSE are the normalized value of BE, MAE, RMSE, BRRMSE divided by the average of x.

The error propagation equation is used to quantify the uncertainty of manual snow depth measurements, automatic snow depth sensors, and recording gauges. For given measurements from two sensors(x and y), the variance of the difference ($y=x_1-x_2$) of two measurements can be written in the followings:

$$\sigma_y{}^2 = \sigma_{x_1}{}^2\left(\frac{\partial y}{\partial x_1}\right)^2 + \sigma_{x_2}{}^2\left(\frac{\partial y}{\partial x_2}\right)^2 + 2\sigma_{x_1x_2}{}^2$$

where x_1 and x_2 are either snow depth or snowfall rate. The $\sigma_{x_1}{}^2$ and $\sigma_{x_2}{}^2$ are the variance of x_1 and x_2. The $\sigma_{x_1x_2}{}^2$ is the covariance of x_1 and x_2. It is reasonable to neglect the covariance for random noise of two same type instruments and the $\sigma_{x_1}{}^2$ and $\sigma_{x_2}{}^2$ should be identical foe the same type instruments. The bias (BE) can exist even for the same type of two instruments. Finally, we can derive the following uncertainty

$$\sigma_{x_1} = \sigma_{x_2} = \sqrt{\frac{\frac{1}{n}\sum_n y^2 - BE^2}{2}}$$

The uncertainty of each instrument is derived from the difference of two measurements of the same type after removing possible bias.

We can extend this analysis into *n* instruments of the same type. The following is an example of five Geonor gauges (g_i where $i=1$ to 5).

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} {\sigma_{g_1}}^2 \\ {\sigma_{g_2}}^2 \\ {\sigma_{g_3}}^2 \\ {\sigma_{g_4}}^2 \\ {\sigma_{g_5}}^2 \end{bmatrix} = \begin{bmatrix} {\sigma_{g_1-g_2}}^2 \\ {\sigma_{g_1-g_3}}^2 \\ {\sigma_{g_1-g_4}}^2 \\ {\sigma_{g_1-g_5}}^2 \\ {\sigma_{g_2-g_3}}^2 \\ {\sigma_{g_2-g_4}}^2 \\ {\sigma_{g_2-g_5}}^2 \\ {\sigma_{g_3-g_4}}^2 \\ {\sigma_{g_3-g_5}}^2 \\ {\sigma_{g_4-g_5}}^2 \end{bmatrix}$$

The right term is known by the difference of the pairs. Thus this matrix can be easily solve by the least square fitting to derive the uncertainty of individual sensor σ_{g_i}.

2. Statistical error modeling

For given i gauges with j wind shield, precipitation measurements can be describe as below:

$$R_{ij}(t) = \mu_t + \alpha_i + \beta_j + \varepsilon_{ij}(t),$$

where $R_{ij}(t)$ is observed precipitation intensity (mmh^{-1}) at time t at gauge type i and wind shield type j. The μ_t is reference precipitation amount at time t, which is assumed to be known. The α_i and β_j are considered as bias due to the effect of gauge types ($i = 1,2,3,4$) and the effect of wind shield types, respectively. The $\varepsilon_{ij}(t)$ is the instrumental uncertainty (or random noise), $\epsilon_{ij}(t) \sim \mathrm{N}(0, \sigma_\varepsilon^2)$.

The parameters α_i and β_j can be estimated by minimizing $\sum_{i=1}^{4}\sum_{j=1}^{4}\sum_{t=1}^{T}\left(R_{ij}(t) - \mu_t - \alpha_i - \beta_j\right)^2$.

$$\hat{\alpha}_i = \frac{1}{4T}\sum_{j=1}^{4}\sum_{t=1}^{T}\left(R_{ij}(t) - \mu_t\right)$$

$$\hat{\beta}_j = \frac{1}{4T}\sum_{i=1}^{4}\sum_{t=1}^{T}\left(R_{ij}(t) - \mu_t\right)$$

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{16T - 4 - 4 + 1}\sum_{i=1}^{4}\sum_{j=1}^{4}\sum_{t=1}^{T}\left(R_{ij}(t) - \mu_t - \hat{\alpha}_i - \hat{\beta}_j\right)^2$$

In addition, the mean difference among gauge types i and i' ($= \alpha_i - \alpha_{i'}$) can be estimated by $\hat{\alpha}_i - \hat{\alpha}_{i'}$. The mean difference between wind shield types j and j' ($= \beta_j - \beta_{j'}$) can be estimated by $\hat{\beta}_j - \hat{\beta}_{j'}$.

ANNEX VI

SUMMARY OF R0 VS. R1 AND R1 VS. R2 ANALYSIS

Daqing Yang, Kai Wong, Craig Smith

1. Problem Statement

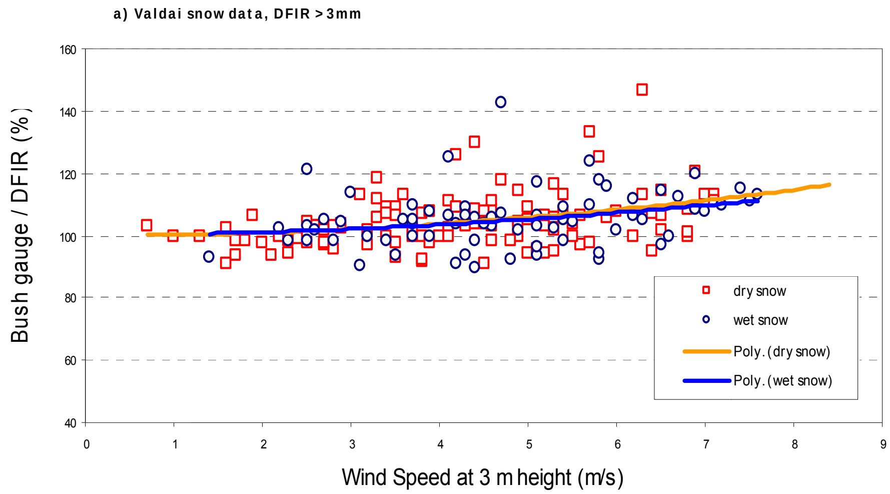
During the previous solid precipitation intercomparison, the project reference gauge (termed the reference standard) was a manually measured Tretyakov gauge inside a large octagonal double fence (currently called the R1 reference). This configuration was called the Double Fence Intercomparison Reference (DFIR). It was validated against many years of manually observed Tretyakov gauges sheltered by bushes (currently called the R0 reference). The current intercomparison requires an automated reference (or an R2 reference). Prior to changing the older reference configuration and recommending a new configuration, a link must be established between them. The following sections serve to establish this link by reviewing historical intercomparisons between the R0 and R1 and the R1 and R2 configurations and by using new data collected at CARE during the current SPICE intercomparison.

2. R0 vs. R1 (Daqing Yang, Craig Smith)

Past data and analyses show that the Bush Gauge at Valdai systematically catches more precipitation (snow and mixed precipitation) than the DFIR and wind speed during the storm affects gauge catch. For instance, the Bush Gauge measures 20-50% more snow over a 12-hour period than the DFIR for wind speeds of 6-7 m/s (**Fig. 0a**). The correction of the DFIR for wind induced loss, thus, is necessary in order to best represent true precipitation. It is important to point out that this error is not a constant loss at all wind speeds; it changes with wind speed and precipitation type.

In comparison to previous analyses (Yang et al., 1993; Goodison et al 1998), recent work (Yang et al., 2013) produces similar but more reasonable results; which suggest lower snow undercatch by the DFIR relative to the Bush Gauge by 3-6%. This means that the DFIR performance is better than what we previously documented in the past WMO intercomparison (Goodison et al., 1998). This result will affect the evaluation of national precipitation gauges against the DFIR. More effort is needed to quantify this impact through field data collections and additional data analyses at selected WMO test sites.

It has been noted that the intercomparison data between the bush and DFIR gauges came only from the Valdai station in Russia. The Valdai data are valuable to represent the climate and snow conditions near that site. There is, however, a need to further test the reference (gauge) systems in a broader range of climate regimes. As part of Canada’s contribution to the WMO SPICE project, a test site has been set up in the southern Canadian Boreal forest SK (the Caribou Creek) to compare the DFIR and bush systems with automatic gauges. Data collection has begun since February of 2013 and will continue for some years. Preliminary analysis of hourly data snowfall suggests that DFIR measured, on average, 10% more snow than the bush gauge, and catch ratio (DFIR/BUSH) did not change much with wind speed up to 5 m/s **(Fig. 0b**). This result is very different from the outcome of the Valdai site. Our effort continues to collect more data at the Caribou Creek site and refine the analysis, so as to better assess both references for the WMO SPICE project and other gauge intercomparison studies in the broader northern regions.



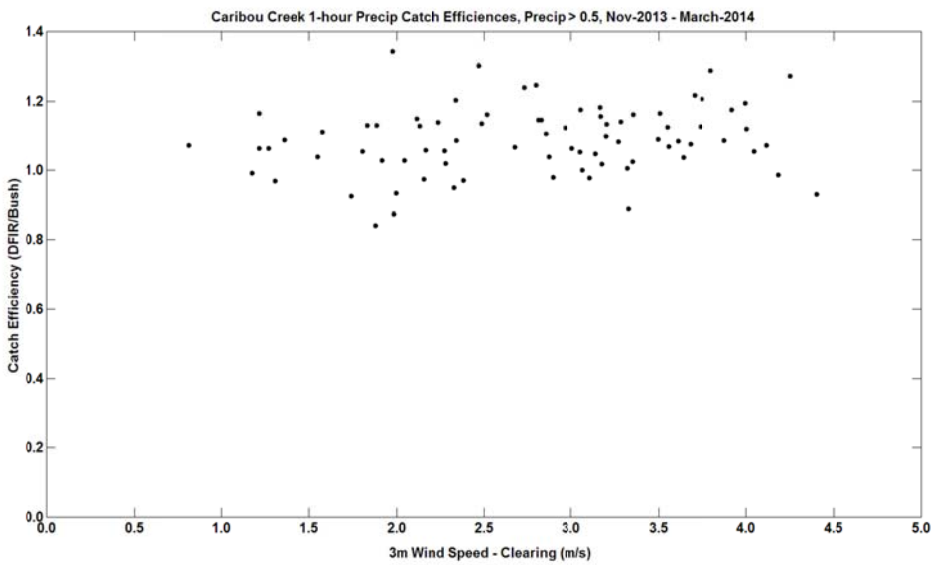


Fig. 0: Catch ratio vs. wind speed for a) Valdai/Russia, and b) Caribou Creek/Canada.

3. R1 vs. R2 Historical (Craig Smith)

R1 vs. R2 (unheated) intercomparisons were obtained during the previous WMO Solid Precipitation Intercomparison at Jokioinen, Finland between 1988 and 1993. Observations of the manual gauge were made twice daily by the Finnish Meteorological Institute, including present weather observations and precipitation type. **Error! Reference source not found.**~~Figure 1~~ compares the observations of the DFIR and the accumulated Geonor-DF for observation periods when both gauges measured at least 1 mm and the maximum temperature during the period did not exceed -2°C. Previous analysis (not shown here) has suggested that scatter in gauge intercomparisons and CE-Wind Speed relationships increase during warmer snowfall events so this analysis has been limited to cold snow events. Figure 2 shows that there is virtually no impact on the relative catch between the R1 and R2 at Jokioinen due to wind.

Further data (post WMO Intercomparison) is available from FMI for Jokioinen and will be incorporated into future analysis.

From 2003 to 2011, Environment Canada intercompared an R1 with an R2 (unheated) at the Bratt's Lake intercomparison facility. Manual observations of precipitation were made twice daily with a Tretyakov gauge inside a DFIR. Concurrently, manual observations of present weather and precipitation type were also made. Manual DFIR observations were then compared to the R2 over the same observation period. Data filtered by precipitation amount (> 1 mm) and temperature (< -2°C) are shown in Figure 3 and 4.

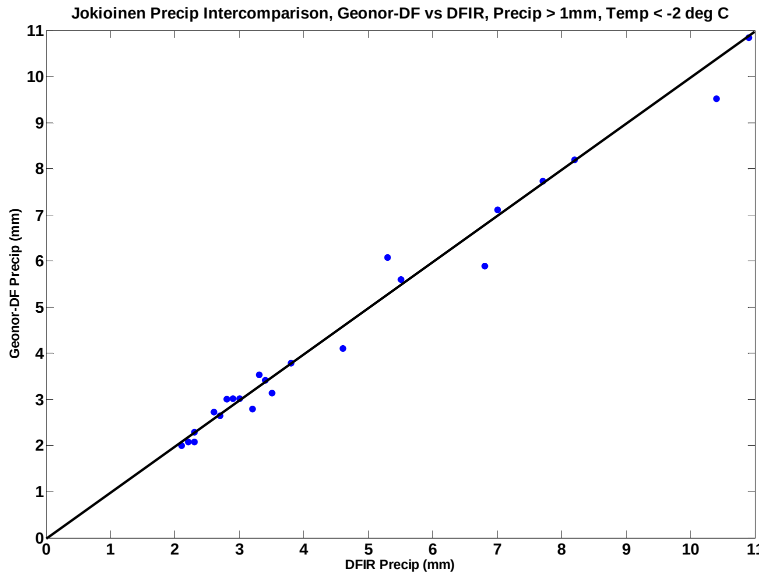


Figure 2~~1~~: Comparison between the Geonor-DF (R2) and the DFIR (R1) with a 1:1 relationship (black line). Data is for daily measurements of snowfall greater than 1mm during periods when the temperature did not exceed -2°C.

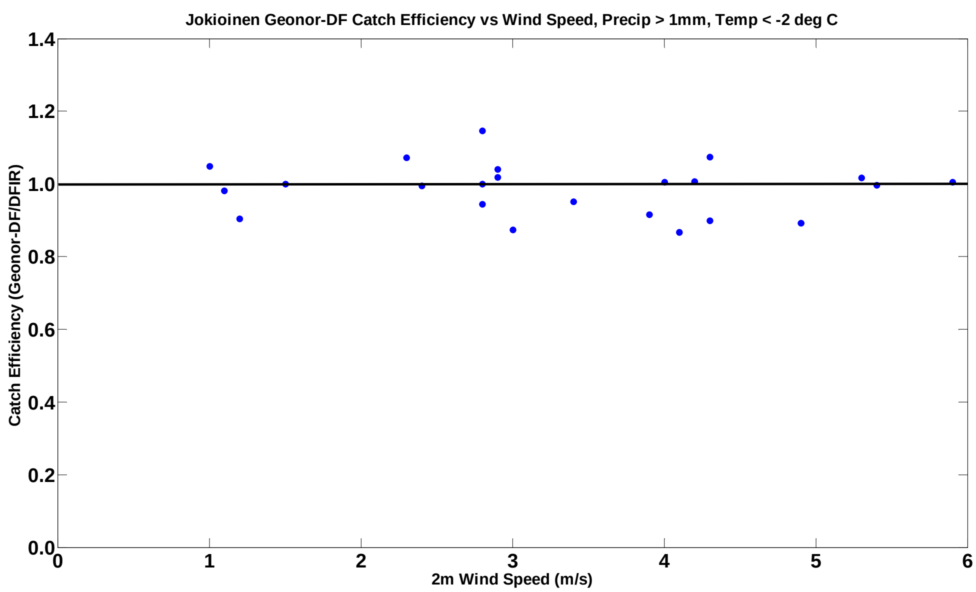


Figure 3~~2~~: Catch efficiency – wind speed relationship for the Jokioinen double fence Geonor (Geonor-DF) as compared to the DFIR for daily precipitation amounts greater than 1 mm while temperatures did not exceed -2°C. DFIR is adjusted for wetting loss but not wind bias.

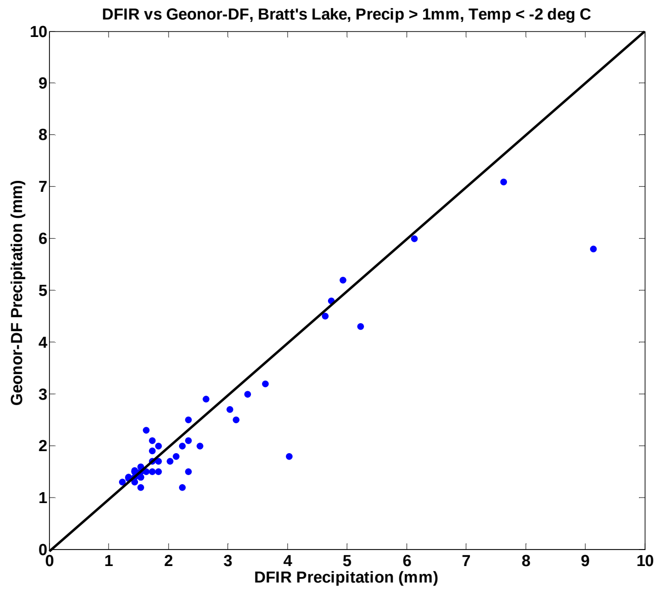


Figure 3: Comparison between the Geonor-DF (R2) and the DFIR (R1) with a 1:1 relationship (black line). Data is for daily measurements of snowfall greater than 1 mm during periods when the temperature did not exceed -2°C.

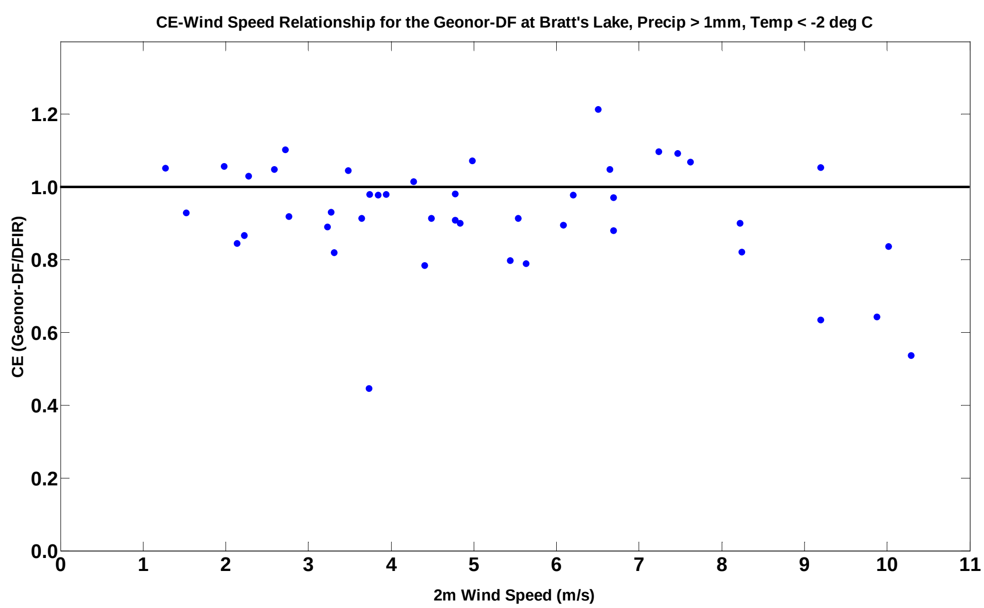


Figure 4: Catch efficiency – wind speed relationship for the Bratt's Lake double fence Geonor (Geonor-DF) as compared to the DFIR for daily precipitation amounts greater than 1mm while temperatures did not exceed -2°C. DFIR is adjusted for wetting loss but not wind bias.

As with the Jokioinen data, the Bratt's Lake data shows very little relative difference between the two configurations with increasing wind speed, with the possible exception of very high wind speeds. Unfortunately, blowing snow at higher wind speeds makes these data somewhat suspect.

4. R1 vs. R2 During WMO-SPICE (Kai Wong)

During the 2012/2013 and 2013/2014 Northern Hemisphere SPICE winters, a R1 vs. R2 (heated) intercomparison was made at the Environment Canada's CARE (Egbert) SPICE site. The manual Tretyakov DFIR was observed once daily and compared to the same period observed by the automated R2. The data was filtered by temperature (< -2 °C) and amount (> 1mm). Figure 5 shows the scatter plot of the intercomparison and Figure 6 shows the relationship between catch efficiency of the R2 and wind speed.

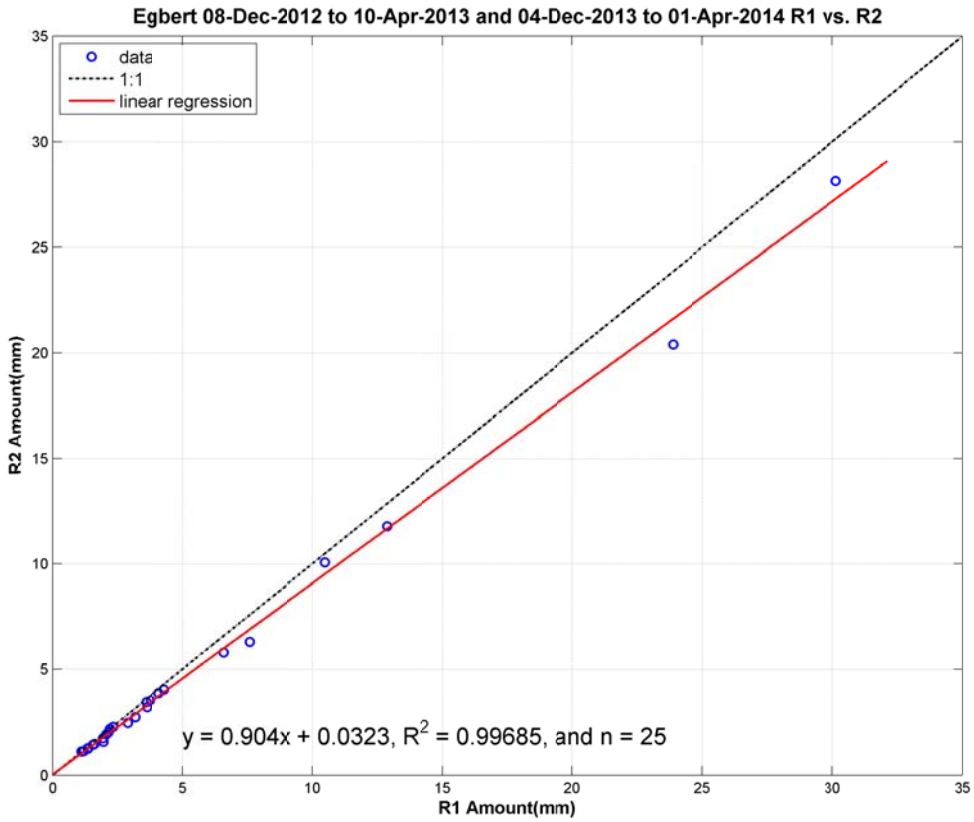


Figure 5: scatter plot of R2 vs. R1 at CARE for 2012/2013 and 2013/2014 when both gauges measure greater than 1 mm

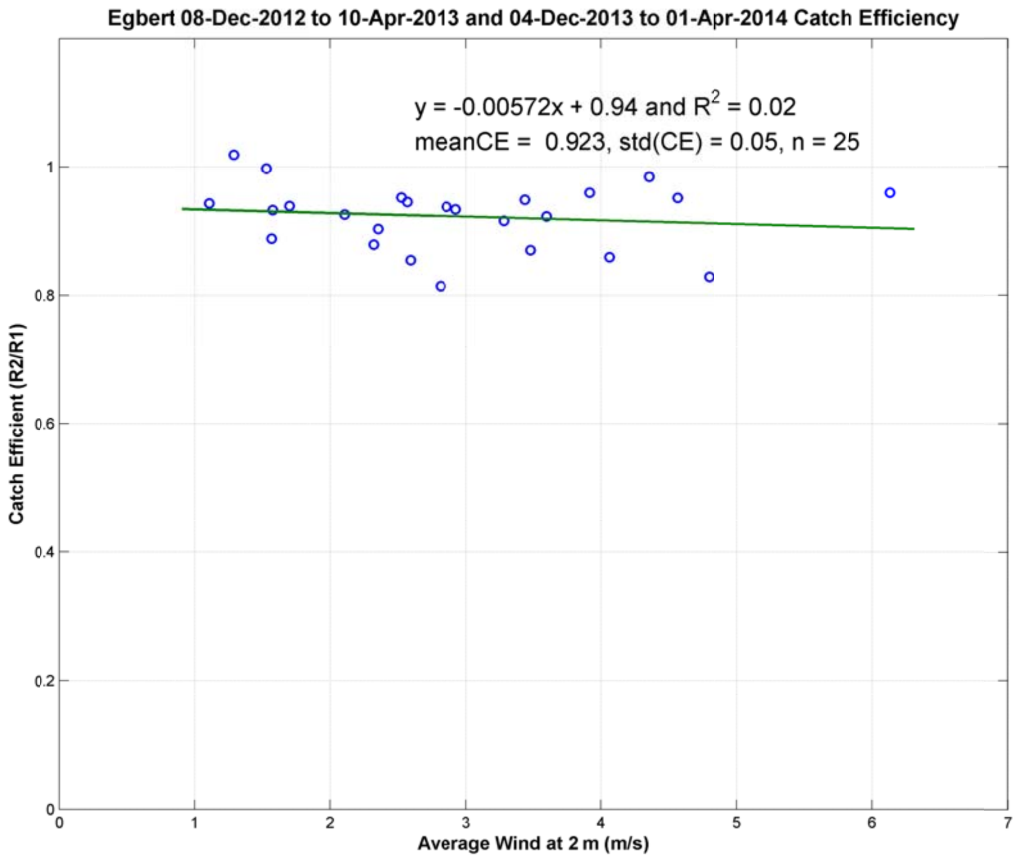


Figure 6: Catch Efficiency vs. wind Speed for events with 1 mm or more accumulations. The mean and standard deviation of the CE are also included

Although the slope of the regression line shown in Figure 6 is not significantly different than zero, the systematic under catch of the R2 (as compared to the R1), which is approximately 8% on average, is evident. This cause of this offset is being examined with the calibration of the R2 already being ruled out. Since the manual gauge is not heated while the R2 is heated, the effects of heating should be examined more closely.

ANNEX VII

METHODOLOGIES EXPLORED TO DATE FOR ASSESSING CATCH EFFICIENCY AND DERIVE ADJUSTMENTS BY DIFFERENT TEAMS, INCLUDING PAST RESULTS

Prepared by Mareile Wolff (prepared 14 May 2014, updated 16 May 2014 and 23.05.2014) with contributions from Y. Lejeune, S. Morin.

In this document, methods used for deriving catch ratio adjustments are collected. Text passages are often copied directly from the references.

Content

1. WMO Solid Precipitation Measurement Intercomparison – final report by Goodison et al., 1998

In the report a common analysis of all data per gauge is presented as well as the individual country reports are listed at the end, partly providing interesting variations of the chosen common analysis method.

> *The detailed analyzing procedure is described in the* ***sixth meeting report of the IOC*** *that will be electronically available soon. Here is the analysis-recipe:*
>
> […] The final step is to determine the catch ratio for each precipitation gauge as a function of wind speed at gauge height and shelter–height air temperature. Wind speed should be the primary effect with air temperature as a secondary variable. These relationships need not be linear. The equation may include linear, exponential and power terms. [..]
>
> […] only events in which the measured precipitation in the DFIR equals or exceeds 3 mm are to be included […]
>
> [….] blowing snow events are to be eliminated.

It is the duty of the researchers from each nation to develop this catch ratio for the national precipitation gauge of their own country.

And actually the meeting report (as well as the final report) contained an extended report/paper on site-data analysis from each country.

Average catch ratios per gauge

Goodison et al., 1998: chapter 4.4.1 (p. 29, example Tretyakov) (more details in WMO/CIMO(1993))

Table 4.4.1 Summary of the Intercomparison of shielded Tretyakov gauge against the DFIR at 11 WMO sites.

	Snow				Snow/Rain				Rain/Snow				Rain			
Station	Event	Ws	DFIR	Tret Catch	Event	Ws	DFIR	Tret Catch	Event	Ws	DFIR	Tret Catch	Eve nt	Ws	DFIR	Tret Catch
	(day)	(m/s)	(mm)	(%)	(day)	(m/s)	(mm)	(%)	(day)	(m/s	(mm)	(%)	(day)	(m/s)	(mm)	(%)
Valdai	304	4.1	1181.7	63.1	85	4.6	584.9	71.2	75	4.5	489.7	86.3	230	3.8	1259.2	91.4
Reynolds	50	2.5	105.6	84.4	27	3.8	71.4	88.5	8	4.4	29.3	85.4	40	2.7	206.4	92.0
Danville	157	1.5	1036.2	91.6	21	1.0	999.5	95.0	18	1.4	348.7	94.5	30	1.0	446.3	94.3
Jokioinen	334	2.6	740.9	67.2	149	3.1	405.6	72.5	131	2.9	414.3	84.5	567	2.5	1694.4	86.6
Harzgerode	42	3.0	112.7	72.2	53	3.9	110.2	78.5	127	4.2	538.8	82.4	172	4.2	475.3	81.3
Bismarck	32	3.3	94.6	65.4	16	3.1	53.3	67.8	-	-	-	-	3	3.3	9.3	71.6
Joseni	94	1.1	194.0	85.8	14	1.3	39.8	92.9	11	2.2	53.6	86.9	34	1.2	85.0	90.6
Parg	65	1.0	486.9	91.0	16	1.2	250.1	90.3	31	1.5	550.8	90.7	141	1.6	1573.8	88.2
Trent U	76	2.0	262.0	81.1	31	2.0	172.3	90.6	20	2.3	219.4	95.0	80	1.9	581.9	95.0
Regina	117	3.5	199.1	59.4	36	4.3	76.9	63.1	-	-	-	-	5	3.9	5.1	97.4
Kortright	107	2.5	274.7	83.1	25	2.7	198.4	85.3	1	4.2	31.9	91.8	64	2.3	342.6	90.0

"Investigation of the mean catch ratio of all observations at each site versus mean wind speed at gauge height during precipitation days shows that there is a general dependence of the mean catch ratio (CR, %) on mean wind speed (V) for snow only and snow mixed with rain:

CRsnow=101.9-10.3*V (n=11, r^2=0.79)

CRsnow+rain=100.4-6.7*V (n=11, r^2=0.6)

For rain only and rain mixed with snow, there is no significant correlation between mean catch ratio and mean wind speed for precipitation days."

"To avoid the over-adjustment or under-adjustment of the wind-induced errors, a constant catch ratio (..) is not recommended for any gauge in any season. Instead the relation of daily or event gauge catch as a function of corresponding daily mean or event mean wind speed should be applied to the gauge measured daily or for an event, since studies show that gauge catch varies by individual precipitation event."

- ➔ Nice for an overview
- ➔ Could be done for daily, hourly (operational) and 30 min-events (and other time-intervals) to get a feeling for differences caused by different time scales

Catch ratio versus wind speed – regression analysis per gauge

(Goodison et al., 1998, chapter 4.4.2 (example Tretyakov))

"Regression analysis of all daily gauge catch ratios as a function of the daily wind speed at gauge height and daily air temperature gave best-fit regression equations for the different types of precipitation.

[…]

A similar analysis of data for individual sites […] could result in different equations. It is critical, however, that a representative range of wind speed be sampled so that any derived equation can be applied to a range of sites with differing wind speeds.

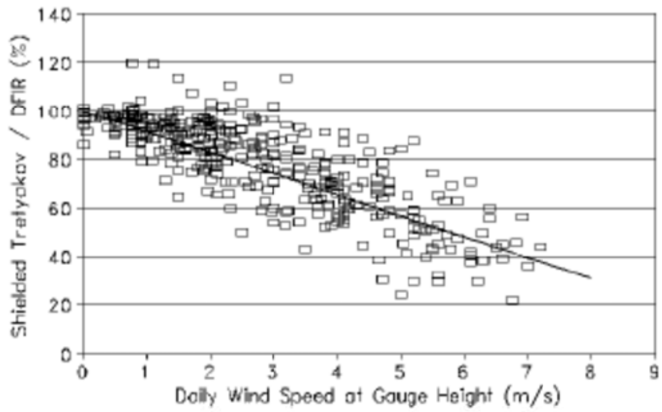


Figure 4.4.1 Shielded Tretyakov gauge catch (% of DFIR) of daily snow (DFIR > 3.0 mm) versus daily mean wind speed at the gauge orifice height at 11 WMO Intercomparison sites, Tmax = -10°C for the curve.

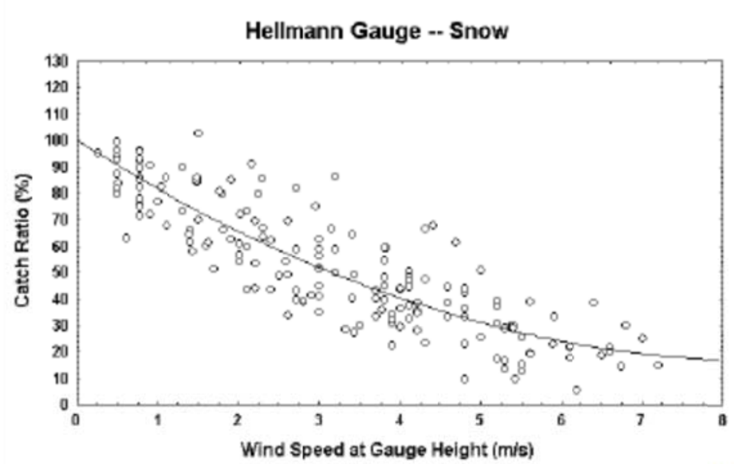


Figure 4.5.2 Catch ratio of snow as a function of wind speed at gauge height of the Hellmann gauge.

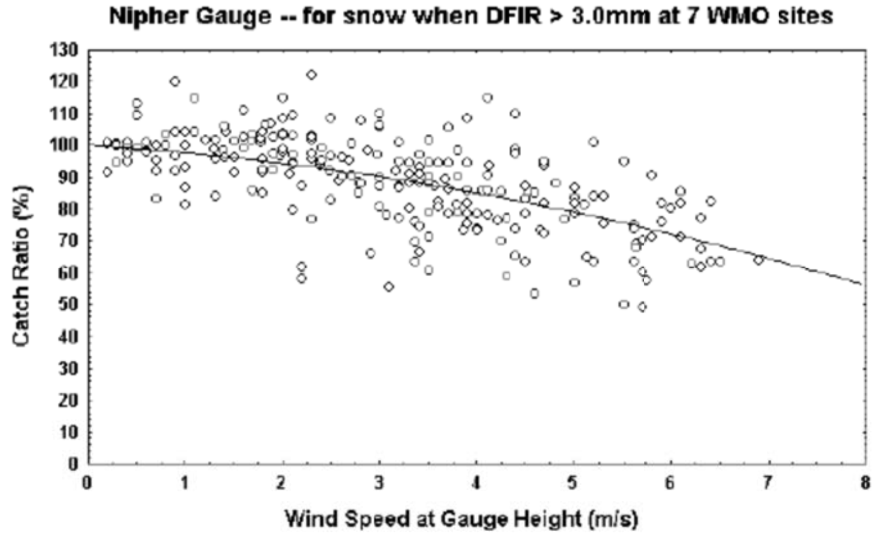
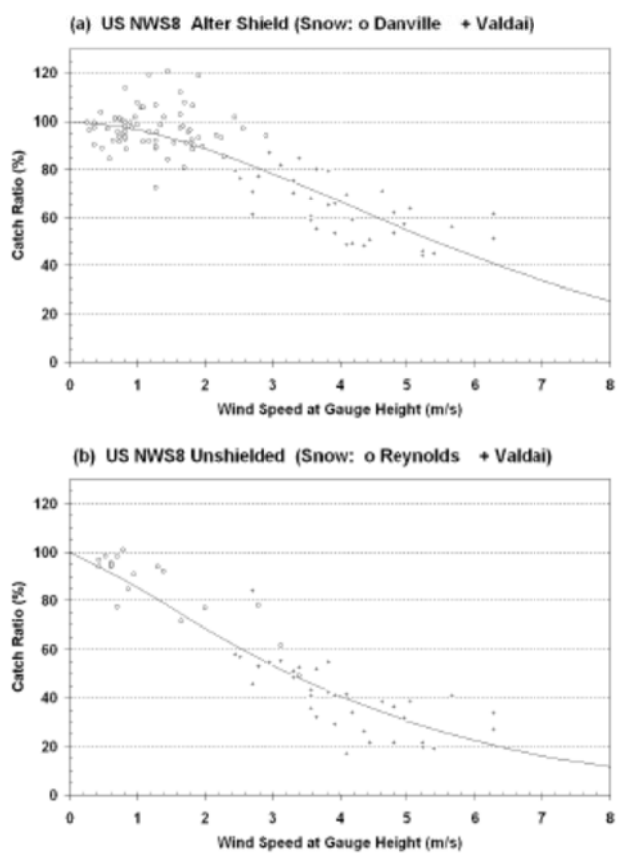


Figure 4.6.2 Regression analyses of Catch Ratio (Nipher/DFIR) versus Wind Speed at Gauge Height

Figure 4.7.1 Daily catch ratio (%) of the NWS 8" non-recording gauge to the DFIR as a function of daily wind speed (m/s) at the gauge height for (a) Alter-shielded, snow; (b) unshielded, snow



Goodison, et al., 1998, chapter 4.8 (summary):

"Regression analysis was used to develop relationships of catch ratio versus wind and temperature. Multiple linear regression (MLR) was applied for the initial screening of significant variables.

[..]

Using the MLR results as a guide, non- linear regression analysis was applied to obtain improved fits where appropriate. In some cases, using a second order polynomial or exponential form of the wind variable improved the regression coefficient. The final regression equations […] are given in Table 4.8.1:

Table 4.8.1 Regression equations for catch ratio versus wind and temperature for Nipher, Tretyakov, US NWS8" and Hellmann gauges

Gauge	Catch Ratio versus Wind and Temperature	n	r^2	SE
	Snow			
Nipher	$CR_{NIPHER} = 100.00 - 0.44*W_s^2 - 1.98*W_s$	241	0.40	11.05
Tretyakov	$CR_{Tretyakov} = 103.11 - 8.67 * W_s + 0.30 * T_{max}$	381	0.66	10.84
US NWS 8" Sh.	$CR_{NWS\ 8\text{-}Alter\ Shield} = \exp(4.61 - 0.04*W_s^{1.75})$	107	0.72	9.77
US NWS 8" Unsh.	$CR_{NWS8\text{-}unshield} = \exp(4.61 - 0.16*W_s^{1.28})$	55	0.77	9.41
Hellmann	$CR_{Hellmann,\ unsh.} = 100.00 + 1.13*W_s^2 - 19.45*W_s$	172	0.75	11.97
	Mixed			
Nipher	$CR_{NIPHER} = 97.29 - 3.18*W_s + 0.58* T_{max} - 0.67*T_{min}$	177	0.38	8.02
Tretyakov	$CR_{Tretyakov} = 96.99 - 4.46 *W_s + 0.88 * T_{max} + 0.22*T_{min}$	433	0.46	9.15
US NWS 8" Sh.	$CR_{Alter\ Shield} = 101.04 - 5.62*W_s$	75	0.59	7.56
US NWS 8" Unsh.	$CR_{Unshield} = 100.77 - 8.34*W_s$	59	0.37	13.66
Hellmann	$CR_{Hellmann,\ unsh.} = 96.63 + 0.41*W_s^2 - 9.84*W_s + 5.95 * T_{mean}$	285	0.48	15.14

Discussion
(Goodison et al., 1998, chapter 4.9):

Applying a DFIR > 3 mm minimum threshold to data sets

This threshold was set in the protocol for analyzing the data, in order to minimize the large variations in catch ratio caused by small absolute differences between the gauges. By applying this threshold, many observations were not included in the analysis. Empirical investigations at Jokioinen showed that this limit is not necessary. The observed catch ratios showed variances of the same magnitude for small precipitation values as for precipitation values satisfying the constraint. Some of the site-analyses were carried out without the threshold (or a much lower one, i.e. 0.2 mm for Jokioinen)

Bush/DFIR adjustments

Most of the analyses for the study were performed with an included transfer to the BUSH-gauge, as this provides the most 'true' precipitation. The quality of this step does depend on the quality and validity of the DFIR-bush transformation. It is questionable if it can be applied on derived data (with a transfer equation from another gauge) as well as on raw data (directly measured with DFIR). The suggested transformation DFIR to bush is complex in the sense that it is not possible to evaluate statistically, the consequences of implementing the DFIR-Bush transformation prior to analysis of the catch ratios versus e.g. wind speed and temperature. Some of the site-analysis were carried out without applying the bush-transformation and adjusted only to DFIR data.

Catch ratio vs correction factor

For the purpose of intercomparing national gauges against the reference DFIR, it is preferable to use catch ratios ($CR=P_{gauge}/P_{DFIR}$) since it provides a comparison of the relative catch efficiency of the national gauge against the reference. For the purpose of developing adjusting procedures, it is preferable to use the correction factor ($CF=P_{DFIR}/P_{gauge}$) which should be derived independently.

Log transformation of ratios

A logarithmic transformation of the CR or CF's is likely to produce homogeneous variances (homoscedasticity) which is preferable for most of the statistical analysis undertaken. In fact, direct modeling of the CR's in relation to wind speed and temperature runs the risk of biased results and inaccurate confidence limits. A range of regression analyses has been used in the country reports, some of which have applied the logarithmic transformation to the ratios before further analysis.

The effect of how averages of wind speed and temperature are estimated

With data from Jokioinen the effect of different averaging periods were assessed. Wind speed and temperature averages for snow events were estimated (based on 10 min temperature and wind speed measurements) as:

- Storm averages
- 12-h averages
- Averages using 3-hourly observations only
- Average of the period maximum and minimum temperature

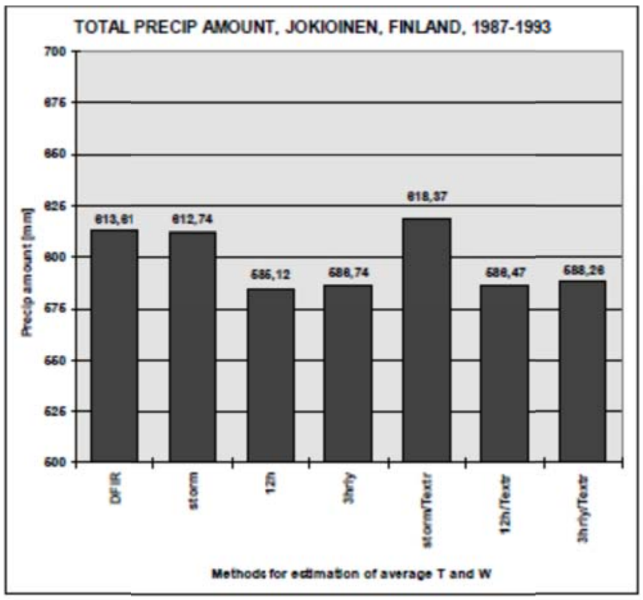


Figure 4.9.2 Adjusted snowfall amount measured by the Danish Hellmann gauge at Jokioinen, Finland, 1987-1993, using various methods of estimating average wind speed and temperature. DFIR=Tretyakov reference gauge, storm=adjustment using storm average of temperature T and wind speed W, 12h=using 12-hour averages, 3hrly=using average of 3 hourly measurements, Textr=average temperature from mean of maximum and minimum temperature. The figures above the columns are the total precipitation amount.

The figure shows that the storm averages perform best, but this is not surprising because the chosen adjustment model is based on storm data. 12-hour and 3-hourly data result in a slight underestimation of the total sum of snowfall (in accordance with that they underestimate the storm-temperatures), but the differences are in fact small over the 7-year period

Different statistical models

The intercomparison study extends over very different geographical and climatological fields, which gives rise to 'errors' of very different nature. Furthermore, a series of statistical models have been used which cannot be compared directly even within a common statistical framework.

2. Yang, D., 2014: DFIR vs Bush Gauge

The statistical tools used, such as regression and correlation analyses of gauge catch ratios as a function of wind speed, have been recommended and tested in the previous WMO gauge intercomparisons.

The regression analysis derives a best-fit curve by means of the least square estimation, i.e. the relationship between catch ratio and wind speed, which is statistically 95% significant.

3. Førland, E., 1996

Using adaptions of earlier used equations for the correction factors of different gauges and calculated the best fitting coefficients based on the data set from Jokioinen, Finland. Correction factors are calculated separately for liquid and solid precipitation. The mixed phase correction factors are a combination of liquid and solid correction factors.

For stations without explicit windspeed measurements, a simple correction method is introduced, using a correction factor based on the exposure of the site.

The study includes very concrete recipes and calculation examples. Often cited.

4. Methodology for correction of solid precipitation data at col de Port by Y. Lejeune, S. Morin, (Météo-France – CNRS, CNRM-GAME, Centre d'Etudes de la Neige, Grenoble, France)

Correcting raw solid precipitation data for gauge undercatch is critical for a proper evaluation of mass fluxes at the snowpack/atmosphere boundary. At Col de Porte, France, a GEONOR T200-1B – 600 mm with a single alter shield has been operating since the summer 1993 and provides hourly measurements of the mass of the collection bucket. Additional precipitation gauges include two so-called PG2000 gauges developed and operated by EDF (tipping bucket, 2000 cm catch surface area, no wind shield – see below for details).

A description of the environmental setting of the site can be found in Morin et al. (2012), which provide an overview of meteorological and snow conditions at the site for the period from 1993 to 2011. The information in this document remains appropriate for the following snow seasons.

The description of snow precipitation correction provided in this document is reprinted below :

« The master precipitation gauge at CDP is the GEONOR gauge. Complementary precipitation data are provided by the two PG2000, only one of both is heated as soon as the collector temperature drops below 5°C. Note that the heat rate is adjusted so that the temperature of the precipitation collector remains lower than 5°C to avoid evaporation as much as possible.

Precipitation data are manually partitioned between rain and snow using all possible ancillary information, primarily air temperature but also the information from the heated/nonheated rain gauge, snow depth and albedo measurements. Relative humidity data are used to rule out spurious precipitation events, i.e. small but non-zero hourly recordings of the GEONOR gauge occurring while RH is lower than 70 %. The GEONOR gauge is corrected for windspeed and temperature following Forland et al. (1996), using a heated cup anemometer placed a short distance from the gauge (1 m horizontally, same height above ground), since the 1999–2000 snow season. For completeness, we provide here the equations used for the correction factor (multiplying the raw precipitation rate). In the case of solid precipitation, the following equation is used as long as the windspeed is between 1.0 and 7.0 ms^{-1}:

$$k = \exp(\beta_0 + \beta_1 ug + \beta_2 T + \beta_3 u_g T) \quad (1)$$

where $\beta_0 = -0.12159$, $\beta_1 = 0.18546$, $\beta_2 = 0.006918$, $\beta_3 = -0.005254$, u_g is the windspeed at gauge height (in ms^{-1}) and T is air temperature (in °C). For windspeed values below 1.0 ms^{-1}, no correction is applied, and above 7.0 ms^{-1} the correction for a windspeed of 7.0 ms^{-1} is used. Similarly, the equation is used only when the temperature is above −12°C ; below this value, the correction factor at T = −12°C is used.

In the case of liquid precipitation, the following equation is used:

$k = \exp(\alpha_0 + c + \alpha_1 ug + \alpha_2 \ln(I) + \alpha_3 u_g \ln(I))$ (2)

where $\alpha_0 = 0.007697$, $\alpha_1 = 0.034331$, $\alpha_2 = -0.00101$, $\alpha_3 = -0.012177$, $c = -0.05$, u_g is the windspeed at gauge height (in ms^{-1}) and I is the precipitation rate (in $kgm^{-2}h^{-1}$). In the case of mixed-phase precipitation, a mixed correction factor is obtained by averaging the two correction factors with a weighting coefficient according to the relative snow- and rainfall rate. Before the 1999–2000 snow season, the precipitation data were multiplied using a scaling factor adjusted for each year by minimizing the difference between the precipitation record and the observed amount of fresh snow recorded using a snow board; this factor remained on the order of 10 %, with year-to-year variations. »

Below we provide additional information on the preliminary spurious data removal and how the partitioning between solid and liquid precipitation is carried out. What is dealt with in the following is hour-to-hour variations of bucket weight, corresponding to raw precipitation rate. First of all, bucket purges and antifreeze refills are removed from the record. An additional treatment is required to handle hourly precipitation data consistently. Indeed, weight accumulation data (reported with a precision of 0.1 $kgm^{-2}h^{-1}$) from the Geonor frequently exhibit spurious slightly negative/positive values (from -0.1 to -0.5, large values mostly related to wind-speed), which is related to uncertainties in the measurement technique (vibrating wires perturbed by structure vibrations induced by the wind) and the data acquisition precision. While the cumulative sum of weight differences would remain correct, it is not appropriate to report negative precipitation values, as well as potentially erroneous (even small) positive precipitation values. A dedicated algorithm is applied on hourly weight increments to remove spurious negative values. Starting from the first record of the series and moving forward in time, once a negative precipitation value is identified, a search (forward and backward in time) is carried out on neighboring time steps to identify positive values which are reduced correspondingly. Note that this algorithm can thus only be applied a posteriori, given that the number of hours forward and backward in time to mitigate a negative precipitation data can range from 1 to ca. 100 hours and more (in the case of the 2013-2014 season). Large numbers correspond to long (several days) periods without precipitation with a negative precipitation record around its middle. As also stated by Morin et al. (2012), “relative humidity data are used to rule out potentially spurious precipitation events, i.e. small isolated non-zero hourly recordings of the GEONOR gauge occurring while RH is lower than 70 %” (actually rather 90% in practice). In case they are spurious, the corresponding precipitation value at the time step is set to 0, but the value is affected to the nearest precipitation event.

Once this step is carried out, what remains to be achieved is to partition the precipitation rate between solid and liquid precipitation, and apply gauge undercatch corrections.

As indicated in Morin et al. (2012), the partitioning between solid and liquid precipitation is carried out “using all possible ancillary information, primarily air temperature but also the information from the heated/nonheated rain gauges, snow depth and albedo measurements.” In short, if air temperature lies below 0°C, precipitation is considered snow. If air temperature exceeds 2°C, it is considered rain. This simple set of threshold is applied to the whole time series, and intermediate cases are then handled on a case-by-case basis. In such cases, not only temperature data but their time evolution (increasing or decreasing during the event) are taken into account. Variations of snow surface temperature are considered (a decrease of snow surface temperature may indicate fresh snow precipitation, while an increase or stagnation at the melt point may indicate rain precipitation). During the day, the measured snow albedo is used: albedo increase indicate fresh snow precipitation, while albedo stagnation or decrease can be linked to rain precipitation. Snow height variations, and measurements of snowboard measurements are also indicative of whether solid or liquid precipitation occurred, but they need to be considered together with the measurements referred to above, as settling for instance, renders the interpretation of snow height measurements difficult in terms of snow accumulation. All of the above leads to inform decisions regarding precipitation phase during intermediate air temperature cases.

However, practical experience indicates that assigning a 50% partitioning between solid and liquid precipitation regarding such intermediate cases is better than crudely attributing the phase to either liquid or solid (this needs to be quantified, however – to be done). One good way to test such impacts is to use a detailed snowpack model (Crocus in our case) to compare model output, driven by precipitation data explicitly separating liquid and solid precipitation but also other drivers of the surface energy and mass balance, with measurements of snow height and snow water equivalent. This check is regularly carried out as a consistency check of the whole correction procedure (i.e., not only precipitation). Compensations may occur, but precipitation remains the main driver of positive snowpack mass variations.

5. Norway, Haukeliseter

The main difference here is the use of Bayesian statistic to analyse the data instead of a more traditional regression analysis. In Wolff et al. (2014) the used method is explained in more details.

Three winters with precipitation data were collected at Haukeliseter test site and analysed during the study. Precipitation events were identified and afterwards filtered in order to pick only those events which were not disturbed by not-controllable parameters, as for example affected wind measurements. The classification of the dataset after key parameters which possibly influence precipitation loss gave a good idea of the shape of possible adjustment functions**.**

The following attributes for an adjustment function, for a given temperature are proposed:

a. The ratio between true and observed precipitation is a function of only wind speed.
b. The ratio is monotonically decreasing from unity when V = 0 to a limit greater or equal to zero when the wind speed V approaches infinity.
c. The ratio decreases exponentially as a function of wind speed.
d. The rate of change of ratio varies significantly as a function of wind speed, and can take the value of zero in parts of the domain.

Bayesian statistics were then used to (objectively) choose the model describing the data set best. Only wind speed and air temperature are input variables for the derived adjustment. It calculates the catchment efficiency of a Geonor with Alter windshield compared to Geonor inside a double fence construction. Two equations were derived, one for the use of 10-m wind measurements, the other for the use of gauge-height wind measurements.

How well the derived function and associated covariates describes the actual catch ratio was evaluated by analysing the residual plots.

No signs of model misspecification could be seen for the covariates wind speed and temperature. Plotting the residuals against the true precipitation, measured in the DFIR, yields a trumpet shape, which may indicate that the noise variance is dependent of the amount of true precipitation. Furthermore, the residuals seem to have a heavier tail than a normal distribution, which also indicates a non-sufficient description of the noise or uncertainty of the adjusted values.

6. Thériault, J.M. et al., 2012: Dependence of Snow Gauge Collection Efficiency on Snowflake Characteristics

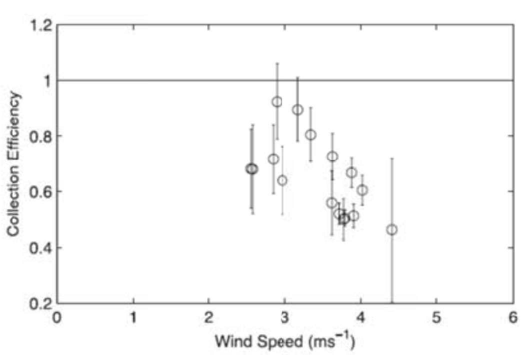


FIG. 12. The collection efficiency of a Geonor placed in a single Alter shield as a function of wind speed for the 20 Feb 2010 snowstorm. The error bar shows the standard deviation of the collection efficiency during the 10-min average.

Intercomparison of the theoretical collection efficiency retrieved from the results of fluid dynamic modelling (Lagrangian model for trajectories) and measured collection efficiency.

Collection efficiency is defined as the ratio of the precipitation rate averaged over 10 min measured by the Alter-shielded gauge to that of the Geonor in the DFIR. The wind speed value used to correlate with the collection efficiency is also averaged over the 10-min period (at gauge height). The following figure shows measurements from one snow storm, 20th February 2010. The mean diameter of the snowflakes falling in the gauge has been calculated and the collection efficiency is also plotted versus the mean diameter of each sample. In general, the snowflake size increases with decreasing collection efficiency. This is consistent with the theory that larger snowflakes will fall in the gauge cause larger snowflakes are less affected by the wind than smaller ones. It also depends on the type of snowflake, however. For example, a large dendrite will still fall slower than small graupel.

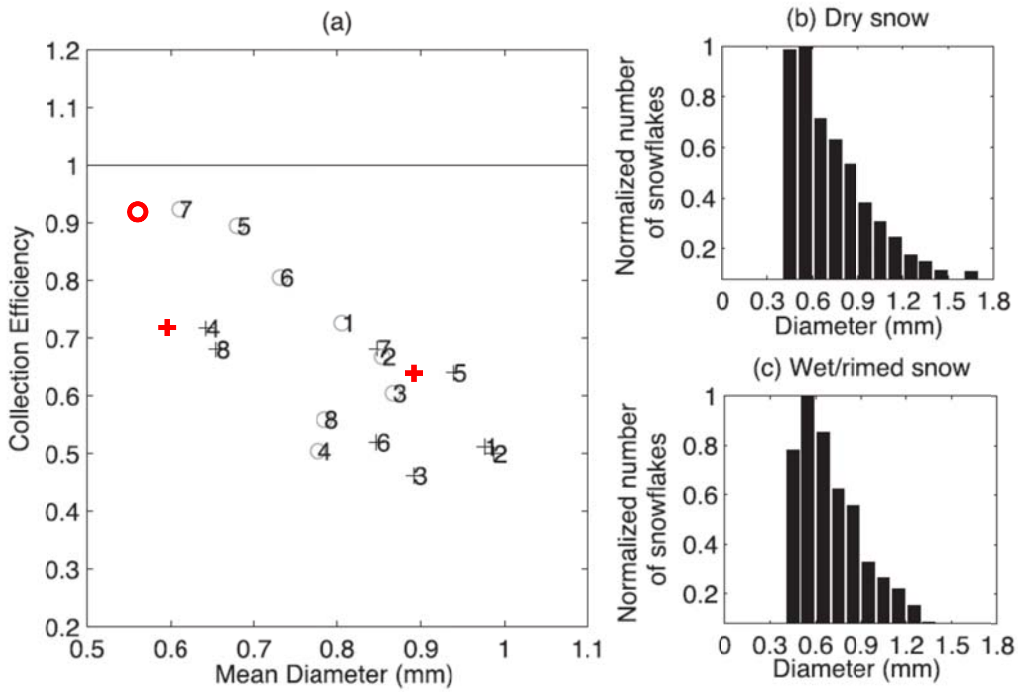


FIG. 13. (a) The collection efficiency of a Geonor in a single Alter shield vs the mean snowflake diameter falling inside the gauge. The gray open circles are for wet/rimed snow, and black plus signs indicate dry snow. There are eight samples of wet/rimed snow and dry snow, numbered from 1 to 8 for each category. Also shown is the normalized number of snowflakes per diameter for (b) dry snow and (c) wet/rimed snow.

From observations, each sample has been numbered and classified as “wet/rimed” or “dry” snow, there are eight samples each. The dry snowflakes are generally larger than the wet/rimed snowflakes. Dry-snow samples 4 and 8 are anomalies. Sample 4 exhibits lightly rimed dendrites, and sample 8 exhibits radiating assemblage of plates, whereas the other samples depict moderately to heavily rimed dendrites.

Using the mean diameter of the snowflakes falling inside the gauge, the mean terminal velocity of each sample was computed using equation and parameters from Rasmussen et al. (1999). They were divided into dry and wet/rimed snow categories. Results are shown in the following figure. The dry snowflakes fall at a terminal velocity that is close to 0.65 m/s, whereas the wet/rimed snowflakes have a terminal velocity that is close to 1.3 m/s, independent of diameter. As expected, the faster-falling snow crystals are associated with a higher catch efficiency. The slower-falling snowflakes (dry) are associated with a lower catch efficiency. For example, the larger dry-snow samples 4 and 5 have a lower collection efficiencies than does wet-snow sample 7 for the same wind speed

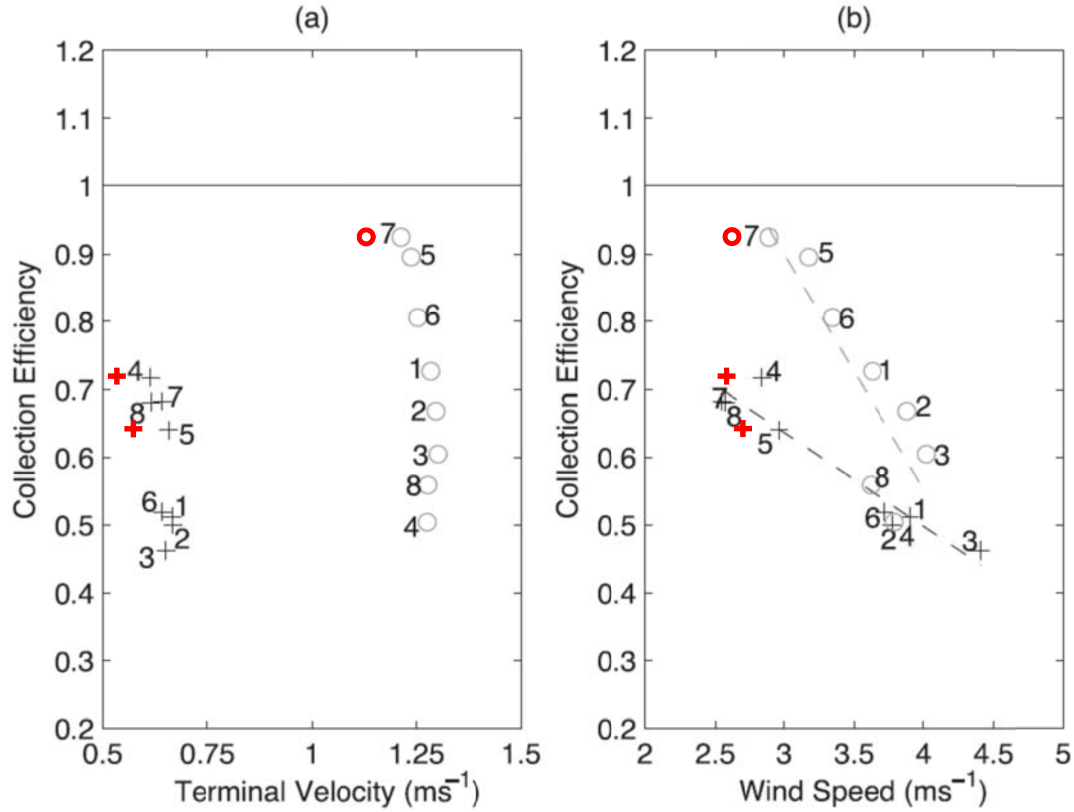


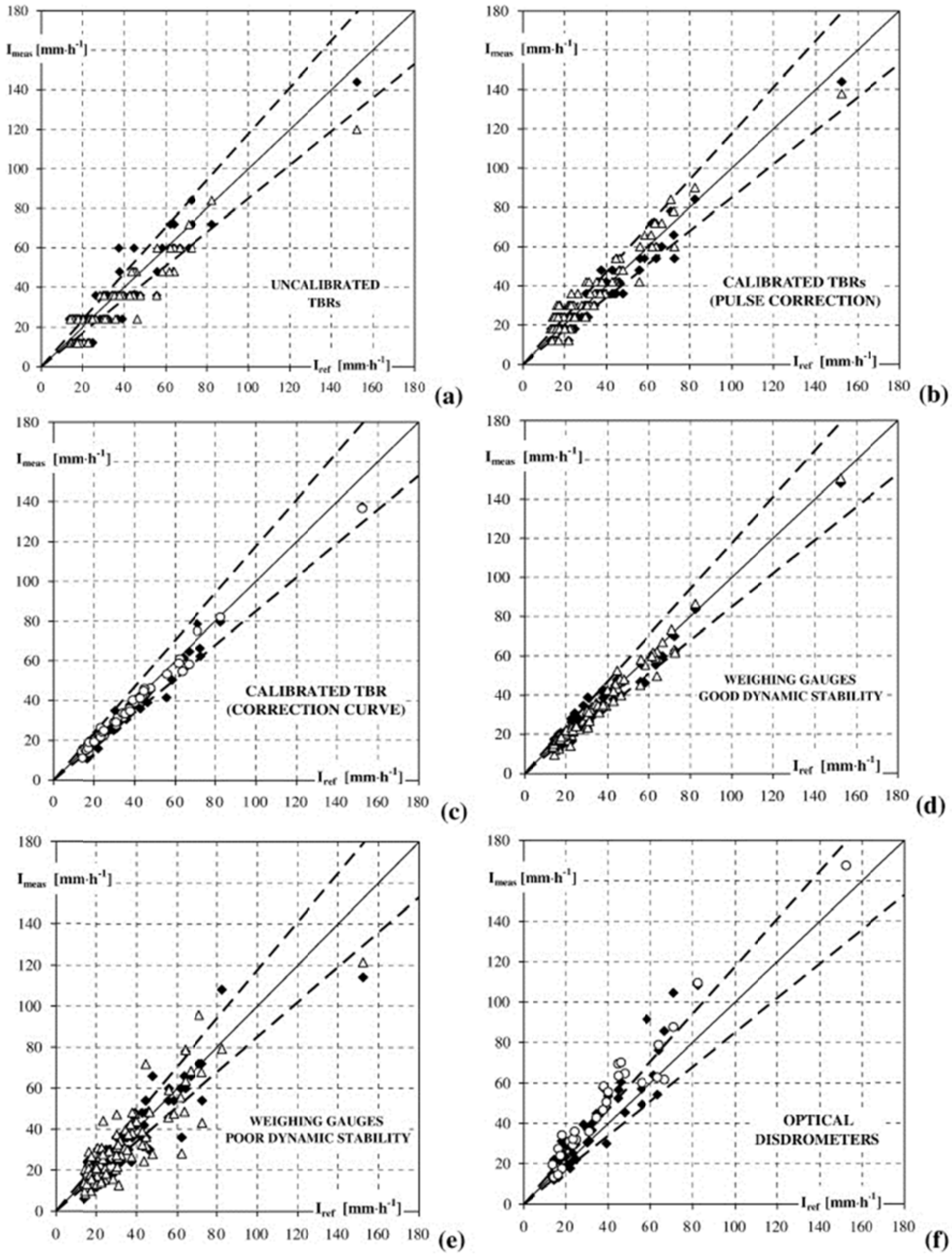
FIG. 14. The collection efficiency of a Geonor in a single Alter shield vs (a) the computed mean terminal velocity of snowflakes falling inside the gauge and (b) the wind speed. The black (gray) dashed lines are the linear-fit curve of dry (wet/rimed) snow samples. The gray open circles are for wet/rimed snow, and black plus signs are for dry snow. There are eight samples of wet/rimed snow and dry snow, numbered from 1 to 8 for each category. These correspond to the samples in Fig. 13.

7. WMO Field Intercomparison of Rain Intensity Gauges (Italy)

Interesting retrieval of the field reference data set, not presented here.

Comparison of measuring performance

In order to compare the measurement results of several installed gauges with respect to the calculated reference a series of scatter plots is used. The one-minute rainfall intensity measured by each rain gauge was plotted with linear scales versus the one-minute reference intensity, for the three best events available. Rain gauges were categorized in groups according to the measurement principle or data output averaging time. Dashed lines indicate the ±15% interval around the one-to-one fit in all graphs, see figure below. Some tests were done (not reported), if the observed differences are depending on wind speed or temperature. No dependency could be seen. That was expected, as only high intensity events with comparable large dropsizes were studied.



8 A few remarks/explanations on inverse model theory (something very related to the Bayesian Model likelihood) from a non-professional, ;-)

I think the book from Tarantola, 2005, gives a good (but unfortunately not simple for a non-mathematician) theoretical background on inverse model theory or how to estimate models and model parameters. I personally like figure 1.12 from the book:

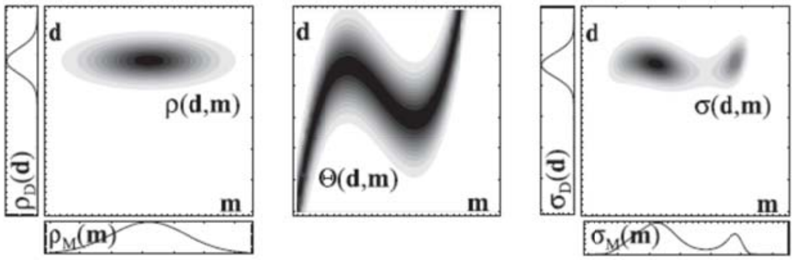


Figure 1.12. *Left: The probability densities* $\rho_D(\mathbf{d})$ *and* $\rho_M(\mathbf{m})$ *respectively represent the information on observable parameters (data) and the prior information on model parameters. As the prior information on model parameters is, by definition, independent of the information on observable parameters (measurements), the joint probability density in the space* $\mathfrak{D} \times \mathfrak{M}$ *representing both states of information is* $\rho(\mathbf{d},\mathbf{m}) = \rho_D(\mathbf{d})\,\rho_M(\mathbf{m})$. *Center:* $\Theta(\mathbf{d},\mathbf{m})$ *represents the information on the physical correlations between* $\mathbf{d}$ *and* $\mathbf{m}$, *as predicted by a (nonexact) physical theory. Right: Given the two states of information represented by* $\rho(\mathbf{d},\mathbf{m})$ *and* $\Theta(\mathbf{d},\mathbf{m})$, *their conjunction is* $\sigma(\mathbf{d},\mathbf{m}) = k\,\rho(\mathbf{d},\mathbf{m})\,\Theta(\mathbf{d},\mathbf{m})\,/\,\mu(\mathbf{d},\mathbf{m})$ *and represents the combination of the two states of information. From* $\sigma(\mathbf{d},\mathbf{m})$ *it is possible to obtain the marginal probability densities* $\sigma_M(\mathbf{m}) = \int d\mathbf{d}\,\sigma(\mathbf{d},\mathbf{m})$ *and* $\sigma_D(\mathbf{d}) = \int d\mathbf{m}\,\sigma(\mathbf{d},\mathbf{m})$. *By comparison of the posterior probability density* $\sigma_M(\mathbf{m})$ *with the prior one,* $\rho_M(\mathbf{m})$, *we see that some information has been gained on the model parameters thanks to the data* $\rho_D(\mathbf{d})$ *and the theoretical information* $\Theta(\mathbf{d},\mathbf{m})$.

I think inverse model theory offers an (objective) tool to choose the best-fitting model from a large amount of possible curves including a comparable evaluation of the uncertainty and likelihood of the fit.

Classic regression analysis is often limited to evaluate only a few possible models. Least-square-methods assume Gaussian uncertainty distributions, which performs badly if outliers are present.

Another book, explaining explicitly Bayesian statistics is Gelman, A., et al. , 2013.

References

Goodison, B.E., P.Y.T. Louie, and D. Yang,, 1998: WMO solid precipitation measurement intercomparison, Final report. Instruments and Observing Methods Rep. 67, WMO/TD-No. 872, World Meteorological Organization, Geneva, Switzerland.

Forland, E., Allerup, P., Dahlström, B., Elomaa, E. T., Perälä, J., Rissanen, P., Vedin, H., and Vejen, F.: Manual for operational correction of Nordic precipitation data, Tech. rep., Det Norske Meteorologiske Institutt., 1996.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B.: Bayesian data analysis, third ed., Chapman & Hall, United States, 2013.

Lanza, L. G., E. Vuerich, 2009: The WMO Field Intercomparison of Rain Intensity Gauges. Atmospheric Research, 94, 534-543.

Morin, S., Lejeune, Y., Lesaffre, B., Panel, J.-M., Poncet, D., David, P., and Sudul, M. : An 18-yr long (1993–2011) snow and meteorological dataset from a mid-altitude mountain site (Col de Porte, France, 1325 m alt.) for driving and evaluating snowpack models, Earth Syst. Sci. Data, 4, 13-21, doi :10.5194/essd-4-13-2012, 2012.

Rasmussen, R., J. Vivikanandan, J. Cole, B. Meyers, and C. Masters, 1999: The estimation of snowfall rate using visibility. J. Appl. Meteor., 38, 1542-1563.

Tarantola, A., 2005: Inverse Problem Theory and Methods for Model Parameter Estimation. Society for Industrial and Applied Mathematics, Philadelphia, 358 p.

Thériault, J. M., R. Roy, I. Kyoko, and S. Landolt, 2012: Dependence of Snow Gauge Collection Efficiency on Snowflake Characteristics, J. Appl. Meteor. Climatol., pp. 745–762.

Yang D., 2014: Double Fence Intercomparison Reference (DFIR) vs. Bush Gauge for “true” snowfall measurement. Journal of Hydrology, 509, http://dx.doi.org/10.1016/j.jhydrol.2013.08.052.

Wolff, M. A, K. Isaksen, A. Petersen-Øverleir, K. Ødemark, T. Reitan, and R. Brækkan, 2014: Measurements of wind-induced loss of solid precipitation: results of a Norwegian field study. Hydrol Earth Syst. Sci., in preparation.

WMO-CIMO, 1992: Final report of the sixth session of the IOC for the WMO Solid Precipitation Measurement Intercomparison, WMO, Geneva.

ANNEX VIII

OVERVIEW OF ANTIFREEZE AND OIL ASSESSMENT

Jeff Hoover, Stephnie Watson, Samuel Morin

Automatic precipitation gauges used to measure solid precipitation in cold climates typically use an initial ‘charge’ of antifreeze and oil. The antifreeze solution is used to prevent freezing of the bucket mixture as precipitation accumulates within the bucket. The oil reduces evaporation of the bucket contents (antifreeze and water mixture), and is particularly important for volatile antifreeze mixtures. Freezing of the bucket contents can lead to damage of the bucket, inaccurate measurements, and difficulty draining the gauge. As well, slush and ice formation above the antifreeze mixture can lead to unwanted evaporation and/or sublimation loss.

In general, the desirable properties for both oil and antifreeze are low density (density lower than that of water is a necessity) and low viscosity. For oil, low density and viscosity allow solid precipitation to pass through the oil layer more easily. For antifreeze, low density encourages the self-mixing of precipitation within the antifreeze mixture. Low viscosity of antifreeze also promotes self-mixing, and limits stratification and slush/ice accumulation above the antifreeze mixture.

A series of laboratory tests were conducted to assess the density and dynamic viscosity of a number of oil and antifreeze samples at 0 °C, -20 °C, and -40 °C. The types of oils tested included various silicone oils, isoparaffinic hydrocarbons, hydraulic fluid, electrical insulating oil, mineral oil, and raw linseed oil. The antifreeze mixtures tested included propylene glycol/methanol, diluted (with distilled water) propylene glycol/methanol, and diluted propylene glycol. The oil tests indicated that isoparaffinic hydrocarbons (e.g. ExxonMobil, Isopar M) display the best combination of low density and low viscosity at all temperatures tested. The antifreeze tests demonstrated that mixtures of methanol and propylene glycol have sufficiently low density at the temperatures tested; however, their high volatility necessitates mixing with oil to prevent the evaporation of methanol.

A separate study investigated the freezing of 40% propylene glycol / 60% methanol mixtures for different initial charge and accumulated precipitation volumes (the latter effectively serving to dilute the antifreeze mixture). The results indicate that as the volume of the gauge charge is decreased, the ability of the antifreeze to prevent freezing at a given accumulation is reduced. For example, considering a 600 mm capacity gauge with an orifice area of 200 cm^2, a 6 L antifreeze charge will prevent freezing at temperatures as low as -40°C if 6 L of precipitation (corresponding to 300 mm) are accumulated. A 4 L antifreeze charge, on the other hand, will allow for freezing at warmer temperatures (between -20 °C and -40 °C) for the same precipitation accumulation.

ANNEX IX

WMO SPICE SNOW ON GROUND ANALYSIS OUTLINE

1. Experiment Objectives and Anticipated Outcomes

The study objectives for the WMO SPICE Snow on Ground (SoG) study, as defined in the experiment plan, are as follows:

Primary Objectives

1) Characterize and recommend automated methods, appropriate for the measurement of total snow depth in a range of cold climate conditions. Different measurement strategies may be recommended for different climatic regimes and for various measurement purposes. The final report is expected to include recommendations regarding measurement thresholds, siting of instruments, availability of ancillary measurements for correcting the measurements of the snow depth sensors (e.g.location and siting of temperature for sonic corrections), instrument sensitivity and accuracy.
2) Investigate and report on the measurement and reporting performance of snow depth sensors measuring total snow depth, over various time periods (minutes, hours, days), linking these measurements to the reference gauge measurements of total precipitation where possible.

Secondary Objectives

3) Assessment of the feasibility to derive reliable spatial representativeness based on using point measurements of snow depth and recommendations for future initiatives focusing on this issue.
4) Assessment of the capability of automated sensors to determine the Snow Water Equivalent (SWE) of accumulated or freshly fallen snow, linking these measurements to reference gauge precipitation measurements and snow depth measurements (where possible).

Anticipated outcomes include:

1) Integrate the results from multiple sites to characterize the instruments under test measuring snow on ground, and identify their limitations. This should include assessments of thresholds, location and siting of snow depth instruments, and the corrections applied (such as temperature for sonic corrections), instrument sensitivity and accuracy.
2) Assess and characterize measurement errors and biases, and their correlation with the operational environment.
3) Assess the ability to estimate the SWE as a result of the correlation between the measurement of snow depth sensors (total or differential over a given period) and the total precipitation amount (R2, where available). Based on this analysis, recommendations can be made for making these linkages in national operation networks to estimate SWE where one set of this instrumentation are not available.
4) Assess the capabilities of automated non-conventional SWE measurement instrumentation to capture the changes in snowpack SWE over both the accumulation and melt periods

commenting where possible on the sensor footprint, resolution, and other capabilities and limitations.

5) Report on the performance of various materials used as targets for the measurements based on the SPICE results, if available, and the review of the literature available on this topic.

2. Analysis Plan

1. Instrument Performance

In order to assess automated methods of measuring snow on the ground (snow depth AND SWE), a general metric of instrument performance is required. This can be categorized three ways: a) Instrument reliability, b) Instrument repeatability, and c) Instrument accuracy.

a) Instrument reliability is to be assessed by looking at the failure rate of the instrument that can be attributed to instrument specific factors (e.g. electronic failures) or environmental factors (e.g. instrument fails because of temperature) when using the instrument according to manufacturer design. Period of instrument failure need to be cross-referenced with site meta and ancillary data. Analysis will also document the frequency and reason for human intervention during the intercomparison period.

b) Instrument repeatability is the ability of the instrument to measure the static snow pack with minimal variability in the measurement. The instrument will be assessed by choosing relatively long periods when the snow pack (or snow free surface) is not changing or changing very slowly. During these conditions, variability in the instrument measurement can be considered noise and evaluated as such. Sources of instrument noise will be examined. Although this could be more easily done with no snow on the ground, it would be preferable to examine periods with snow as a snowpack could introduce complicating factors (such as a non-solid surface) that need to be factored into instrument performance.

c) Instrument accuracy will be assessed by comparing the instrument measurement to the manual reference measurement (frequency being dependent on instrument type and reference method) and assessing biases related to instrument errors or environmental factors.

2. Instrument Thresholds

Instrument threshold analysis can be categorized by a) physical and b) temporal thresholds.

a) The physical threshold is the threshold of the snow pack at which the sensor is able to make a reliable and accurate measurement. The various instruments under test will have different minimum and/or maximum physical thresholds. Minimum thresholds will be evaluated by assessing how soon and accurately the instrument registers a snow pack measurement as snow begins to collect on the target surface. This will be accomplished through manual reference measurements (and other manual observations), web camera observations (where available) and observations made with other SoG instrumentation at the site. Metrics for Instrument Performance also apply. The surface target needs to be considered in this assessment.

b) Temporal threshold is the minimum or maximum measurement frequency of the instrument at which it can make a reliable and accurate (and meaningful) measurement. The metrics for instrument performance will apply but this will also be dependent on the nature of the snowpack (how rapidly it changes) which will be environmentally dependent.

3. Linking SoG to Gauge Measurements

Analysis and procedures will be developed to link the snow depth as measured by SoG depth instrumentation to the liquid water equivalent of snow events as measured using the R2 or R3 reference. This will be done for case studies where we have a high degree of confidence in the accuracy of both the snow depth and the precipitation gauge measurements. These case studies will be used to develop and assess the capability of estimating snow depth from gauge measurements or vice versa, examining the density of event based snowfall, and testing the commonly used 10:1 snow depth to precipitation conversion factor. Ancillary data will be used to determine when and why procedures break down. This link will be examined at various time scales including hourly, 6-hourly, daily, and event scales. This analysis will be repeated for automated SWE measurements, linking these measurements to the R2 or R3 reference.

Further to this, SoG instrumentation will be assessed for the capability of determining start and end times of snowfall events, determining transition from rain to snow (or the reverse), and precipitation typing. Other ancillary meteorological data will be factored into this analysis.

4. Assessing spatial representation of point measurements

Where possible, point or multiple point measurements will be assessed, using established techniques, on their capability to represent spatial averages of snow depth or SWE. This is accomplished by comparing the spatial average of snow depth or SWE obtained using a multi-point snow course. Although this is more of a siting issue than an instrument issue, reporting should indicate recommendations as an outcome of SPICE.

Table 1: Site/Instrument/Analysis Matrix. Numbers indicate planned analysis (i.e. 1=Instrument Performance,…)

	Site						
Instrument	**Sodankyla**	**Hala Gasienicowa**	**Col de Porte**	**CARE**	**Caribou Creek**	**EVK2-CNR**	**Weissfluhjoch**
SR50ATH	1,2,3,4	1,2	1,2,3,4	1,2,3,4		1,2,3,4	1,2,3
SR50	1,2,3,4		1,2,3,4		2,3,4(?)		
SHM30	1,2,3,4		1,2,3,4	1,2,3,4			1,2,3
SL300	1,2,3,4			1,2,3,4			
USH-8	1,2,3,4			1,2,3,4			
CS725	1,2,3,4				1,2,3,4		
Lysimeter							1,2,3
SP3 Snow Pillow							1,2,3
SSG Snowscale							1,2,3

____

ANNEX X

SPICE WORKPLAN FOR 2014-15

CATEGORY	TASK
REGISTRY OF PROBLEMS AND SOLUTIONS:	Create google doc with list of problems of instruments under test Accessible to all site managers (that have instruments under tests and of those who have completed their commissioning report?) Ensure that there is follow up on raised issues
	Site managers to input in the REGISTER information on problems encountered and (later) the way they were resolved. Focus on the relevance to data analysis and result interpretation
RECURRING PROBLEMS WITH INSTRUMENTS FROM INSTRUMENT PROVIDERS	Task one person to contact manufacturer on reported Recurring PROBLEMS: For problems with gauges/sensors that are relevant to many site (f.ex. Pluvio2 phantom accumulation)
	In each case, the designated individual to Report on recommended solution to concerned site managers and IOC
	Request IOC decision in case of need to change a configuration/software version.
INSTRUMENTS FROM PROVIDERS	Cross-check that all instruments provided by Instrument Providers are included in commissioning reports of the sites where the instruments have been allocated.
	Verify that all data from Instruments that have been provided by Instrument Providers is available at NCAR and that data is meaningful (e.g. it's reasonably precip, etc.)
Relationship with the Instrument Providers	Send Letter informing them of : Prolongation of experiment Possibility to obtain precipitation data Invite them to liaise with sites (look at data, visit sites) Motivate them to liaise (in your interest as SPICE will report on your instruments) Report to Rodica in case of problem. Who: letter from Secretariat
Letter to Site Managers	Encourage to liaise with instrument providers Recall publication guidelines (inform provider in case of reporting results from instruments under test) Appropriate time for feedback: at least 3 weeks in case of “negative/problematic” results. Warn that information is scattered through various SPICE IOC meeting reports Circulate a summary of the Data Protocol General guidance: Keep as is to have 2 winters of consistent data: guidance to site managers

phantom Accumulation Issues (Pluvio)	No modification to instruments from provider (at least if no formal information from provider to project leader); INCLUDE DECISION IN THE REPORT Update of Pluvio firmware related to phantom accumulation:
phantom Accumulation Issues (Pluvio)	For references, **Yves-Alain** will invite Nemeth to Payerne to discuss and develop way forward to be proposed to the IOC.
Configuration Issues (Sodankylä)	Put additional wind sensor at height of R3 and gauges under test to have indication of wind speed difference between R2 and R3
	Install disdrometer in DFIR
	Install an additional Pluvio with slats in right direction before season 2014/2015
	Orientation of Alter shield slats (for R3 and gauge under test): maintain as available in May 2014.
	The IOC noted the height difference between the gauge in R2 vs. that of the R3 gauges. While recognizing the impact, the IOC decided to maintain the current configuration, to allow for a continuous consistent 2-year dataset to be collected
Data quality control	Flag system for bad quality data Possibility to flag data for both site-managers and data-analysts
	Alternative to implementation at NCAR: sites could create a flag-file similar to an extra instrument file with to defined details (disadvantage: difficult to add flags for data-analysts) Concept in Brussels IOC-3 meeting (2012, annex nr. 2)
	Document flexibility/possibility for changes of **QC and Event** selection (thresholds, filter widths, time interval lengths, …) as implemented at NCAR
	Include the description of QC procedures in the reference report
	Information supporting the qc assessment (manual and automatic): Develop and implement qc for sensor under test data, for automatic qc: define fields/thresholds;
Data analysis of SUT	Document the implemented data aggregation for all instruments under test, related to the derivation of the accumulation during events.
	Data aggregation for the sensors under test, corresponding to events will be implemented at NCAR (in process);
	Include in the data analysis the ability to detect and assess false reports from the sensors under test.

Data Analysis: derivation of transfer functions	Transfer functions for gauges under test: start simple with only one (continuous) transfer function towards DFIR for catch ratio dependent on wind and temp for selected gauge configurations including data from all sites. John is using his Matlab tool, GyuWon the Bayesian statistics [C,] applied to: unshielded Geonor, unshielded Pluvio single alter Geonor, single Alter Pluvio double alter Geonor, double Alter Pluvio tipping buckets non-catchment type instruments Transfer functions solely based on precipitation type (and not temperature) might be more accurate, but will be working for just a few stations as necessary data are not available
	Compare various methods known to be used for the derivation of transfer functions, using the same data set and report on the differences/similarities
	check for other dependencies (i.e. intensity) with residuals, evt. introduce another parameter
	apply the same transfer functions on data sets from different sites for validation and look for site-specific biases; Can sites be grouped according to these differences? Compare these differences to differences from R3-analysis[C,]
	Look into continuous equation vs. type-specific equation
	Document all methods explored for future reference and refinement of methods in later data analysis iterations
Data statistics	describe statistics from the event files: how many events, total snow amount, wind, per winter per site: average catch ratio per gauge per site per winter Note: season catch ratios might include bad-quality data (phantom accumulation Pluvio, bucket emptying, …) add quality control and redo – what are the differences? That will become an important issue when applying corrections – what kind of QC needs to be applied to data before adjustment? Gauge accumulation from event file vs. gauge accumulation from beginning and end of the season. How much do we get
Data analysis for Non-Catchment Type Instruments	Performed as separate topic, informing/connecting to transfer equation development: it requires a consolidated analysis and results plan. While recognizing that this is a very complex topic, and transfer functions are not feasible at this stage, use the data from the sites where manual observations are available to assess the ability of these instruments to report the precipitation type, when snow or mixed, estimate amount and intensity
	Deliverable: Assess the possibility of using combinations of sensors (disdrometers, snowfall, radiation,….) for deriving the precipitation type

Snow on the Ground	A complete analysis plan needs to be developed to identify specifics steps/tasks and linkages with the falling precipitation project. Finalise the event selection methodology for snow on ground events. Focus on building on the derivation of events for falling precipitation; possible at sites where the R2, R3 references are available As manual reference for SoG measurements are generally less frequent, the correlation between these and accumulation events can still be used to mark longer periods with/without precipitation, and assess their contribution to accumulation over various intervals.
	Note, that the SPICE accumulation event identification does not catch trace amounts, the setting of snow or snowmelt periods, which might be interesting to compare. Sites without a suitable precipitation gauge (Italian Forni site and Nepal, maybe more?) the site teams need to identify events based on their experience. Method needs to be documented.
	Identify and provide requirements for parameters for SoG events, if additional parameters are needed, to complement current set.
Snowfall	Compare snow-fall (relative distance change) instead of the whole snow pack. Fresh-snow-density observations are useful, use simple situations with cold temperatures (for preventing melting). Any site which does this fresh-snow measurements?
	All SoG-sites with density analysis need to record their fresh layer separately!!
	Wherever possible should snow-density measurements be included
	Yuri looks for a possibility at Valday
	Leena wants to make grain size measurements and from the other snow-site in Sodankylä available
	Leena will try organize snow density measurements for Feb-April 2015 Sodankylä
	Assess how to expand the use of the approach currently used at Col de Porte: relative snowfall measurements are compared with models and are related to the accumulation
R0-R1 assessment	Continue assessing R0-R1 based on new data (Valdai, Caribou Creek) and past projects; utilize various methods currently available: descriptive statistics, Bayesian method (GyuWon), John's approach Look into continuous equation vs. type-specific equation
	Daqing to assess how the two gauges inside the Canadian Bush gauge compare to each other, evt. apply a similar method as in Russia (combination

	of max 3 gauges in the bush)
	Continue assessing R1-R2 based on new data, other SPICE data (e.g. include Marshall manual obs) and past projects (Bratt's Lake, Jokioinen, SPICE-1; Look into continuous equation vs. type-specific equation
	Roy/Bruce apply more data to their unshielded/shielded relation (if possible with and without wind-binning) Analyse adjustment functions for DFIR-single alter and DFIR-unshielded (see above) further Check if site-categories of these two methods are related Include temperature impacts, for temperatures where mixed precip would occur.
Uncertainty	Calibration procedure provides a bias for the gauges (and hopefully only that) à useful for gauge characterization Additional we want to derive a random error range for each gauge The uncertainty is related to the transfer function we chose – residual analyses of the results. Uncertainty analysis with comparing multiple gauges of same configuration at one site (basically references);
	John will also take GyuWon matrix method; using also old Marshall data
	Leena will do the Pluvio Uncertainty analysis following John's approach
	John is giving a talk on uncertainty and afterwards takes a lead on a paper on that, cooperation with GyuWon.
	Daqing gives the Tretyakov-stuff to John from the Russians
Derivation of Reference Data Set Using the Pluvio^2 gauge output	Decision at Boulder meeting to use 'rawest possible' data from Pluvio^2 – Bucket RT, confirmed with OTT; operationally, OTT recommends to use Accumulated NRT output; Need to better understand the differences between Bucket RT/Accumulated NRT during precip events and non-precip.
	Employ Accumulated NRT data as an additional 'quality controlled' data stream for input to event selection algorithm; compare output, quantify differences. Assess how to use all Pluvio2 data products to improve the quality of the data used for the reference data set

Derivation of Reference Data Set Using the Geonor gauge output	Several methods are available for averaging data from three transducers, with varying levels of complexity/ease of implementation: Simple average – quick and easy Weighted average – reduce contribution of noisy wires to average; can create ‘cleaner’ dataset, but risk biasing average output too-high or too-low Detailed performance assessment – look at wire offsets, differences in magnitude of noise, different response to precipitation; establish thresholds for excluding ‘compromised’ wires when computing average output for reference datasets Noted that spectral analysis of data could be useful for identifying and removing features that would impact three-wire average; focus on non precipitating conditions Analysis of precipitation data from CARE by GyuWon’s group has indicated that selection of simple average or weighted average does not lead to significant differences; not necessarily the case for all data sets
	Further refine the algorithm by using simple averaging to generate first version of quality-controlled reference dataset for event selection and further analysis; pursue more detailed performance assessment, with potential to generate an improved reference dataset in the future
	Investigate spectral analysis as a means of quality-controlling Geonor data prior to averaging
Uncertainty/noise in Geonor and Pluvio data used in derivation of reference datasets;	**Proposal 1:** test additional 6 s datasets of both gauges in identical configurations and quantify residual noise following filtering with increasing filter widths; identify smallest possible* filter width that reduces noise in datasets from both gauges to a comparable level (*to avoid ‘smoothing over’ any small precip features) Noise in reference datasets for Geonor and Pluvio2 gauges should be of similar magnitude; no ‘preferential’ noise reductions for one gauge type or the other using recommended QC methodology; establish ‘common ground’ for subsequent event selection and analysis. Recent analysis of 6 s CARE data for pairs of Geonor and Pluvio2 gauges in identical configurations has indicated that extending filter width from 2 minutes (current recommendation) to 5 minutes and longer periods can reduce noise in both gauge types to similar magnitude
	Proposal 2: conduct similar analysis as proposed above using 1 min datasets in which single Bucket RT value output every 1 min [Mike] Related analysis at CARE indicated that selection of filter width was not critical for 1 min data from both gauge types; BUT this was for 1 min data generated by averaging 10 x 6 s data points (averaging = inherent filtering)

Time resolution of reference datasets	Proposal: produce reference datasets with 1 min time resolution to serve as the basis for all applications – can extend aggregation/averaging interval to longer periods to match specific applications (e.g. hourly operational reports); if higher uncertainty in Pluvio data at 1 min relative to Geonor data at same time interval persists following filter width investigation, consider generating 5 min reference dataset for further analysis. Analysis by GyuWon's group has indicated that for data at smaller timescales (less than about 5-10 minutes), there is greater uncertainty in Pluvio gauges relative to Geonor gauges; uncertainty becomes similar as average to longer timescales
Data Transfer to NCAR	All site managers have to transfer all their data to NCAR
	Some sites needs assistance for Data transfer to NCAR (A)
	Russia
	Italy
	Korea (Rodica)
	Nepal
	Review/update/distribute procedure for Manual measurements reporting to NCAR (Precip Amount, SoG), including data format, procedure. Build on the Switzerland guidelines
	Identify gap in the data available at NCAR (planned to be transmitted vs. what is at NCAR today: conversation with Site Managers and assessment of NCAR site)
	NCAR to create webpages for all sites that don't have one yet
	Site managers to provide to NCAR/Andy the info for establishing the site webpages (description, photos, etc.)
	Data repository at NCAR: Try to identify concrete problems with sending data, what are the reasons – help individually (addressed in other reports)
Site Configuration management	Major changes to site configuration need to be reviewed/acknowledged/approved by IOC, if have significant impact on data or represent a major departure from the original configuration, as described in the commissioning report
	Maintenance to be carried out strictly according to use manual and any additional guidance provided by manufacturer [A,] In case of doubt consult manufacturer.
	Finalise Commissioning Reports
Site Metadata made available to DAT(s)	Site Managers to validate the data files that are for SPICE, in case files other than those intended for SPICE have been transferred to NCAR.
	Site managers to indicate which precipitation detector/disdrometer is to be used for the derivation of the reference data (specify the file name)
	Site managers to indicate which wind sensor is to be used for gauge height wind/10 m wind/other (related to SPICE!)
	The structure of files on the webpages of various sites is not always the same: awareness of it is needed.

	Concern regarding the linking of metadata with data, and having the appropriate metadata available: to verify with NCAR whether all the metadata provided is used in the database and is easy for access.
	Develop an evergreen table of the instruments under test, including their data fields, eventually by consolidating existing tables. It is proposed that alternative for storage of such table are explored, e.g. using Google drive. Use this as a template for other on-line collaboration document development and maintenance (site configuration, etc.)
SITE Logs	All maintenance & calibration actions have to be included in site tracking sheet (available on Rodica´s ftp server: A, Rodica and Audrey relocate file to google drive SPICE account or alternatives) Including calibration performed by manufacturers!
	If calibration is conducted by the Instrument provider, the results must be documented in the site log
	Each site needs to store and share the information on site events that could affect the data and would trigger flags. These could be used for data analysis and manual qc. [A,)
	Assess the possibility of storing the data/site logs (site changes tracking) at NCAR, to be uploaded by Site Manager at least one per season, at the end of the season.
	Expand the site data logs (site changing tracking) to include events that would affect the data
	upload site changing tracking log to NCAR.
Publications	Issue paper
	Conference(s)
	Presentations made at conferences to be shared within the team, and stored at a place for future reference
	Write short version of SPICE-disclaimer

____

ANNEX XI

SNOW MEASUREMENTS AND DATA ANALYSIS IN SODANKYLÄ

Finnish Meteorological Institute's Arctic Research Centre (FMI ARC) in Sodankylä has large snow measurement sites for monitoring the development of natural seasonal snowpack. Measurements are mainly related to satellite data validation, instrument deployment and development, and interpretation algorithm development. Intensive Observation Area (IOA) is the main snow measurement site in Sodankylä. The site hosts several microwave instruments, and numerous automatic and manual reference measurements. Another important measurement site is wetland site. Pine forest and wetland are the most common land cover types in the wider Sodankylä area. Those sites are not operational and data is used only for research applications. Measurements are made since 2006.

Microwave radiometers, scatterometer and optical spectroradiometer are reference and validation instruments for satellites in Sodankylä. Passive microwave instruments have frequencies 1.4, 10, 19, 21, 36, 90 and 150 GHz with H and V polarizations. Optical ASD spectroradiometer for radiance and reflectance measurements is mounted on top of the 30 m mast.

Sodankylä has extensive automatic in-situ measurements of snow. Automatic Weather Station (AWS) system consists of air temperature and snow depth sensors. Disdrometer measures precipitation intensity and detects different types of precipitation. Gamma Water Instrument (GWI) is an experimental prototype of automatic sensor for snow water equivalent (SWE) measurements. The measurement is based on attenuation of gamma radiation from cesium source. GWI is calibrated manually by using manual SWE measurements. Snow temperature profile measures temperature every 10 cm up to 120 cm.

Regular manual snow measurements include snowpit measurements in several sites, snow course measurements and organic elemental carbon measurements. Snowpit measurements include measurements of snow height, SWE, temperature profile, density profile, snow layers, grain size and type, hardness, wetness, density and moisture profile (Toikka Snowfork) and specific surface area (A2 Photonic Sensors IceCube). Snowpit measurements are made weekly in both main measurement sites. FMI ARC has two short snow depth variability courses and a four kilometer long SWE and depth variability course.

In addition to regular measurements, snow measurement campaigns are organized in Sodankylä. Nordic Snow Radar Experiment (NoSREx) campaigns for the satellite instrument development and deployment were organized during 2009-2013. During those campaigns lots of in-situ measurements and microwave measurements were made together with cooperation partners. Artic Snow Measurement Experiment (ASMEx) was organized first time in 2014. Campaign includes radiometer measurements and manual reference measurements from homogenous snow slabs of natural snow. Smaller local campaigns, such as optical snow laboratory measurements, are also organized.

Sodankylä has large amount of different data to analyze. Satellite reference instrument and manual measurement quality checks are made along with research. Data of automatic instruments are quality checked semi-automatically after every winter. Spikes are removed, too small and large values are removed, data level during summer is checked and data gaps are analyzed. Final data analysis of FMI ARC data in all cases is made along with research.

____

ANNEX XII

BRIEF OVERVIEW OF THE SODANKYLA SPICE SITE

The Sodankylä SPICE site is situated in Sodankylä Arctic Research Centre, near the new satellite antenna area. It has been operational from summer 2013. There is almost full winter season of measurements from the first year 2012-2013. In the field there are 21 SPICE instruments, 6 FMI's own instruments and also 2 manual measurements are made. Precipitation, snow depth, present weather and ancillary weather data is measured.

Precipitation is measured with 3 different bucket instruments: OTT Pluvio^2, Vaisala rain gauge (VRG) and Meteoservis MRH3. There are six Pluvio^2 instruments, four with the opening of 200 cm^2 and two with the opening of 400 cm^2. The Pluvio^2 instruments have different wind shielding (DFIR, Alter and no shielding). There are also two VRG instruments with different wind shielding (Alter and Tretyakov). Precipitation can also be measured with OTT Parsivel disdrometer which is an optical present weather sensor.

Snow depth is measured with four different kind of optical instruments (USH8, SR50ATH, SL300 and SHM-30) and also four manual measurements are made once every day. There are two USH8 and SR50ATH instruments with different setups. One is installed on horizontal bar and the other diagonal bar. The idea behind the two setups is that the instrument installed on a diagonal bar should measure less snow, since there is no accumulation of snow on the bar itself where it could fall into the measurement area. The manual snow measurements are made remotely using a webcam and snow sticks.

Snow water equivalent is measured with GMON3 (CS725) and SGG1000 instruments. SGG1000 is a snow scale that weighs the snowpack above. GMON3 measures the attenuation of gamma rays coming from the radiation source on the ground. Manual measurements are made once every two weeks.

Present weather is measured with three different Vaisala present weather detectors (FS11P, EPI33 and PWD22) and the OTT Parsivel disdrometer. Ancillary data includes measurements of wind speed and direction, temperature, relative humidity and air pressure. Also a rain detector (DRD11A) can be found in the field.
