The WMO Radar Quality Control Quantitative Precipitaton Estimation Inter-comparison Project

RQQI[1]

V3.0
20111130
Paul Joe

Revision History

Date	Version	Author	Changes
2.0	20110401	PJ	As presented to IOC
3.0	20111114	PJ	• Major Revision • Revision to reflect decisions, discussion at Kick Off meeting. • Added dual-polarization to the project. • Added consideration of nowcasting, NWP and climate applications. • Added glossary of terms and included definitions of raw data for consistency with data protocol. • Merged first workshop implementation plan to reduce number of documents.

This is a living document and will be updated as necessary. It consolidates several documents into one for ease of tracking.

[1] Pronounced Rickey

Project Statement

Inter-comparison workshops will be conducted to identify, document and exchange the best techniques for quality control of ground based Doppler weather radar data primarily for quantitative precipitation estimation in a variety of radar scenarios and in different weather and environment regimes and to develop data quality metrics for global and regional applications. In addition, requirements for spatial radar data (including non-precipitating echoes, radial velocity fields, and precipitation echo identification) for nowcasting, NWP data assimilation, verification, and for climate diagnostic and trend analysis require harmonization of radar data biases and variances both regionally and globally. In addition, dual-polarization radars are now being deployed operationally and need to be taken into consideration.

Goals

1. Undertake systematic inter-comparison and validation of radar QC algorithms evaluated under a variety of environmental conditions for QPE, nowcasting, NWP and climate applications.
2. Provide guidance to WMO members on quality control processes employed in radar quality control algorithms.
3. Characterize and assess errors involved in radar quality control algorithms.
4. Report on algorithms employed in radar QC.

Objectives

1. To develop a framework for QC algorithms inter-comparison.
2. To collect, collate and create inter-comparison and validation test data sets that would consider a variety of radar types, scanning modes and environmental conditions
3. To develop quantitative radar data quality metrics.
4. To compare and evaluate radar QC algorithms in a series of focused inter-comparison workshops.
5. To develop a data quality framework (metadata)
6. To develop or promote existing data and product exchange formats that include data quality.
7. To conduct and report on inter-comparison workshops with recommendations approved by a International Organizing Committee (IOC) of experts

The Need

Advances in severe weather nowcasting, data assimilation, weather and climate model validation, satellite rainfall verification, and hydrological applications have led to new and enhanced requirements for high quality regional and global radar datasets and radar precipitation products. Recent advances in radar technology, signal and data processing have brought the field to the brink of operational readiness for the quantitative use of these products. In the past, radars were perceived to address only local and qualitative applications, such as for severe weather diagnosis and warning. This view is rapidly changing as scientific needs for assimilation of radar data in Numerical Weather

Prediction models and diagnostic and trend analysis for climate applications require harmonized radar data and quality information (unbiased, known variance). This has been made possible by high speed telecommunication networks that allow vast amounts of data to be transferred and by storage technology that permit data archiving in a central location.

While the progress in the radar QPE has been impressive, it is also recent and there are many differing approaches and solutions. It is therefore necessary to harmonize, consolidate, validate, verify, identify the best algorithms and under what conditions to specify the quality of the products. Non-precipitating radar echoes (due to insects and Bragg scattering) can reveal valuable Doppler wind fields for NWP and for the identification of low level convergence boundaries for nowcasting convective initiation. In the latter situation, reflectivity fields are also useful. Dual-polarization radar is an emerging operational technology that provides considerable data quality information. It is able to identify ground clutter, distinguish biological targets, rain-snow boundaries and the presence of hail. Therefore, a data quality framework that can distinguish or classifies the radar targets is needed.

Radar QC processing is a common problem for all NMHS' and a collaborative and sharing approach of the techniques and results will have mutual benefits. Processing differences include techniques or algorithms to mitigate ground clutter at the signal and data processing stages, to determine the appropriate vertical profile of reflectivity, to identify attenuation and partial blockage effects and to make bias corrections. Product differences include temporal and spatial scales of the data, accuracy and precise, data format exchange standards.

Stake Holders

This project addresses a core issue that contributes to other WMO mandates and initiatives. Within WMO, the project is lead by the CIMO ET on Upper and Remote Sensing. Other stakeholders include other CIMO ET groups (particularly precipitation and cryosphere teams), WWRP Nowcasting and Mesoscale Weather Forecasting Research groups, the THORPEX DAOS group and WCRP (WGPRN), amongst many others.

The best approaches need to be identified, documented and shared with all member countries. Commonality of approach, when possible, is needed to promote collaboration and efficiencies. Data quality metrics and standards need to be identified and developed so that these data and products can be integrated as part of the Global Observing System and as part of Global Earth Observation System of Systems(GEOSS)[2]. Processes for maintaining, monitoring and meta-data reporting radars need to be established to support a long term archive.

[2] The GEOSS goal is to establish a global, coordinated and sustained observing system to meet societal needs particularly in respect of severe weather warning and disaster management.

Many new radars are being established in developing nations where societal impact is high. An example is Africa and this proposed project will help them attain the Millennium Development Goals through improvements in the use of their radars for climate purposes. Other related WMO programs will also benefit from this project and include: WMO Information Systems, Quality Management Framework, Natural Disaster Prevention and Mitigation, Flash Flood Initiative, contribute to foster closer collaborations between NHS and NMS's and International Exchange of Data and Products and enhance the social and economic benefits of NHMS'.

The Data Assimilation Perspective: The assimilation of high resolution radar data is becoming an operational reality. For example, ECMWF has demonstrated the impact of assimilating the U.S. radar data on the forecasts in Europe (Lopes and Bauer, ????). Different assimilation systems process either the reflectivity, radial velocities or derived wind profiles (VAD or Velocity Azimuth Display winds) from Doppler radar networks. Information from a variety of radar networks will be needed to describe their quality.

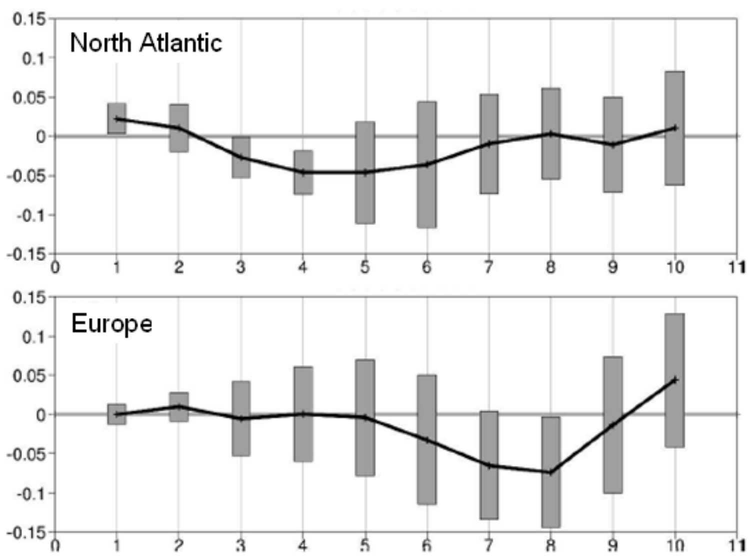


Figure 1: Impact of assimilating U.S. radar data on the forecasts in the North Atlantic and Europe from the ECMWF model. Lower RMSFE scores indicate the positive impact (Lopez and Bauer, 2006).

The Data Assimilation Perspective: The extremes in the hydrological cycle manifest in changing extremes of precipitation. Precipitation rate is so highly variable in time and space that it is one of the most difficult of the essential climate variables to measure and monitor precisely, particularly the more extreme (rarer) events. Current surface precipitation gauge arrays are much too sparse in spatial coverage and (usually) report too infrequently to resolve precipitation variations and can easily miss extreme events. Although indirect satellite precipitation products do exists at relatively high space and time resolution, the indirect nature of these methods may not be able to properly capture the extreme events that represent non-average conditions by their very nature. They also may have biases that need in-situ and ground based weather radars to remove. Spaceborne radars have good spatial resolution but have very low time sampling frequencies compared with the variations of precipitation.

Weather radars, specifically designed to observe precipitation at high temporal and spatial resolutions, are becoming ubiquitous in many nations around the world and are currently the only precipitation-measuring system with the requisite space-time sampling. However, until now these systems have also been generally sparsely distributed and operated and analyzed separately in a case-study approach. The recent significant growth in the number of operational radar networks, now makes possible high-resolution determinations of precipitation over extensive land areas and over long time periods. Such datasets, when merged with very stable spaceborne radar reflectivities would provide the observational basis for learning how the small-scale extreme events are connected with the large-scale atmospheric circulation and how this may be changing in time. To provide this type of data, requires systematic collection and analysis of data from these radar networks in as many different climate regimes as possible on a retrospective basis to produce appropriate "climate-scale" statistics. The potential of these radar networks for climate studies is high but has not been realized.

The detailed corrections and adjustments are a pre-requisite for the generation of appropriate products to study climate variations and extremes. A consistent global measure of radar data quality applicable for a variety of radar networks in different climate regimes and recording of quality meta-data are needed to make use of the radar data for these types of studies.

The Data Assimilation Perspective: Radar and radar networks are critical to the provision of severe weather warnings and nowcasting. The observation of clear air echoes are critical. These are created predominantly by insects but can also be created by Bragg scattering.(turbulent fluctuations at half the wavelength of the radar). These observations provide good wind data but also provide the identification of boundaries that are useful for nowcasting convective initiation.

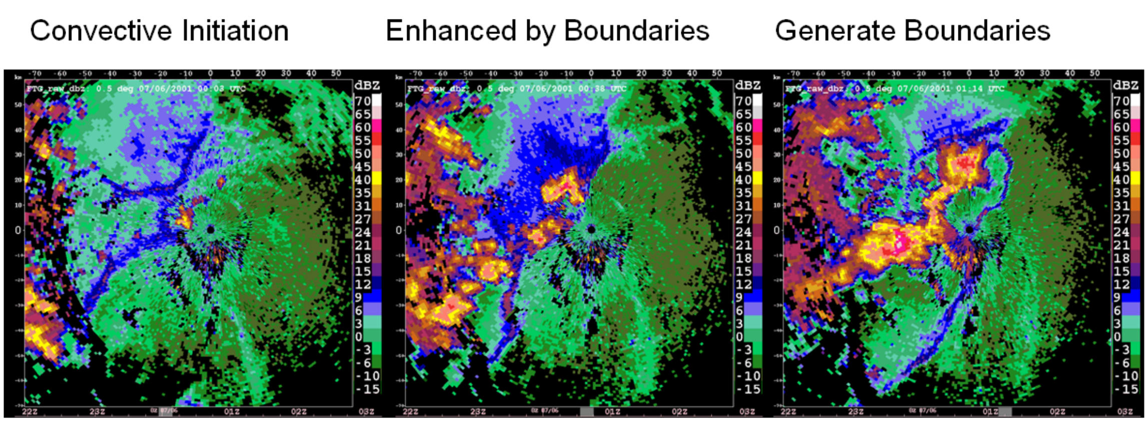


Figure 2: An example of non-precipitating echoes (insects, green and blue echoes) showing the presence of convergent boundaries (blue linear echoes). Thunderstorms initate and grow on the boundaries. They generate their own boundaries, leading to secondary convective initiation. The relationship between the boundary and thunderstorm movements are used to nowcast the evolution of the thunderstorms (Wilson, personal communication).

Ground echoes have utility to reveal humidity structures in the atmosphere through the retrieval of refractivity fields through changes in the propagation paths (Fabry, ????). Propagation paths can be measured by the phase shift from the Doppler signal. Changes in path length are due to primarily humidity in the atmosphere as pressure and temperature (other relevant parameters) fields tend to be smoother and show less variation.

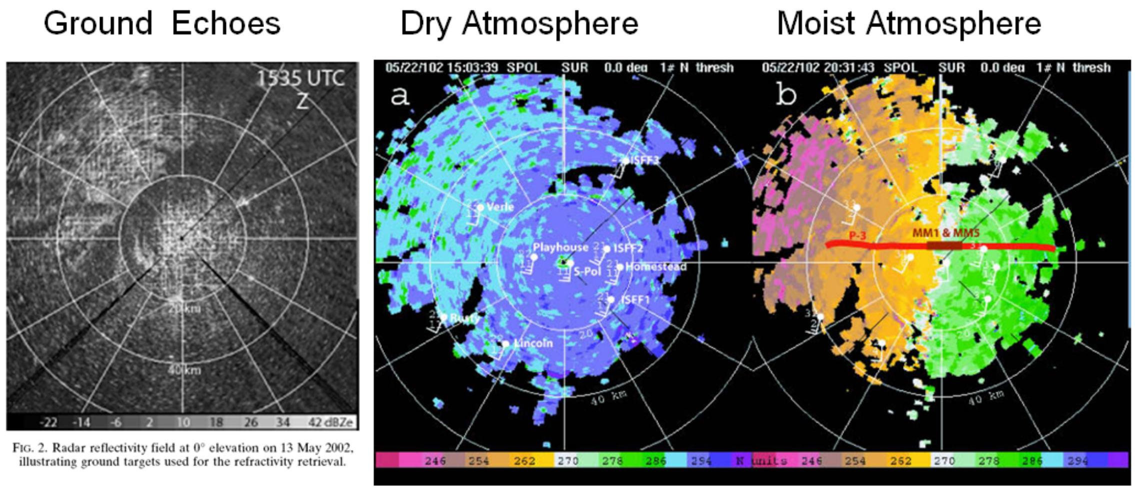


Figure 3: An example of humidity retreival from refractivity technique using ground echoes (Weckwerth et al, 2004).

Radar Technology

Radars are complex instruments and there can be many technical differences related to the radar hardware technology (conventional, Doppler, polarization, data resolution), equipment location (flat land, urban/rural environments, mountain valley or top) and its environment (dry prairie, moist coastal, desert, tropical, winter, summer, ground echoes) and purpose (cycle times and scan strategy).

Radar Signal and Data Processing: The radar technology and signal processing options have significant impact on the approaches available to address one of the primary issues for rainfall measurement – ground clutter and anomalous propagation removal. Conventional radars still form the bulk of the operational radar technology globally. Reflectivity statistical techniques, ground clutter maps, texture of the data and high altitude CAPPI maps are still used to mitigate ground clutter. Doppler radars afford the possibility of using velocity signal processing to remove stationary targets (Fig. 4a). This has been a significant improvement but there are some drawbacks (see Fig. 4b). Polarimetric radars are now being deployed operationally and offer significant potential through differential phase measurements to identify and separate clutter targets, mitigate the impact of drop size distribution effects, attenuation in precipitation, and a number of non-meteorological error sources.

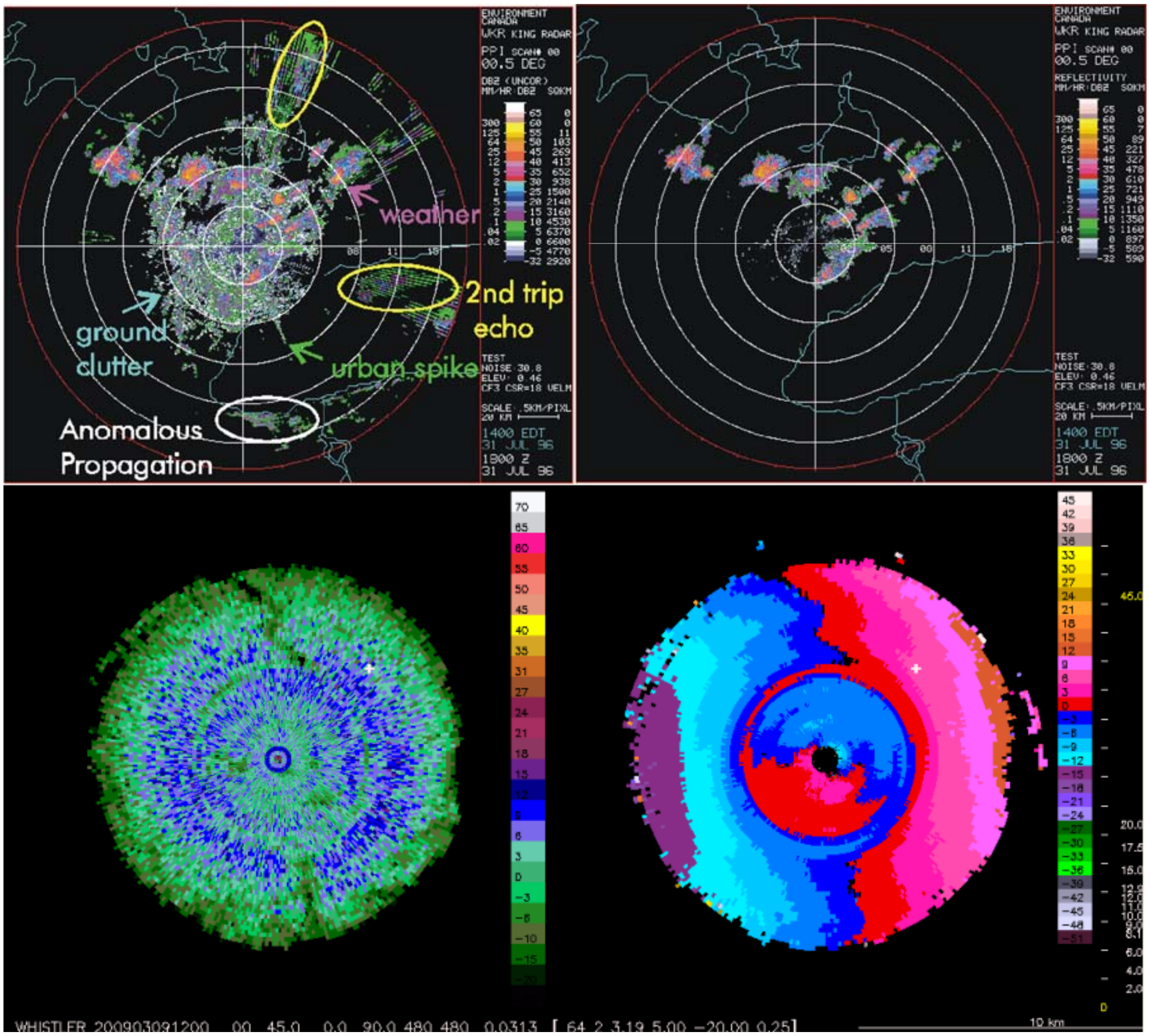


Figure 4: (a) Improvement of ground clutter mitigation using Doppler technology. (b) Depending on the weather situation zero velocity notching can remove too much echo.

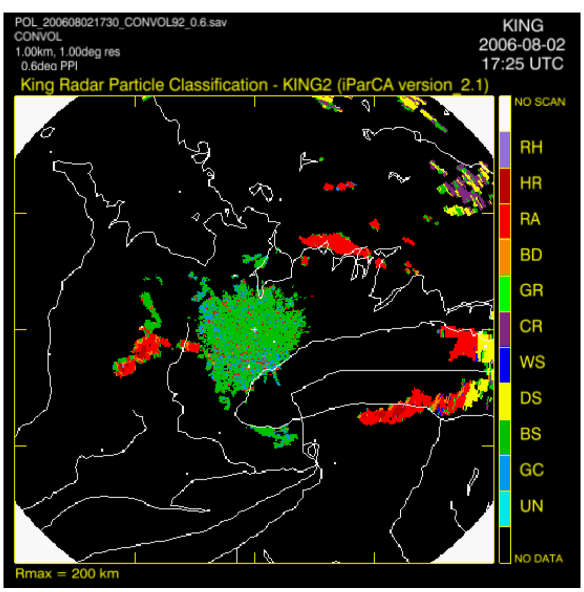


Figure 5: Dual-polarization classification of radar targets. The categories are: RH = rain/hail, HR = heavy rain, RA = rain, BD = big drops, GR = graupel, CR = crystals, WS = wet snow, DS = dry snow, BS = biological scatterers (includes birds, bugs, chaff), GC = ground clutter, UN = unclassified

Performance, Calibration and Maintenance: A number of engineering, system design and related sources impact the accuracy of measurement and may dominate without careful treatment . These are loosely lumped into "radar calibration" but can be separated to performance, electronic calibration and system maintenance. Performance refers to the sources of uncertainty related to the radar specifications such as the antenna pattern, gain and sidelobe structure, receiver and filter losses, pulse shape, frequency drifts, noise and even antenna tilt. For the radar engineer, calibration refers to the accuracy and precision of the received power only. In the past, this was a major source of uncertainty and some research radar receivers were calibrated on a daily basis. The receivers are much more stable now. Maintenance refers to monitoring for component failures and drift. Changes in these factors over time is requires a systematic approach to metadata definition and reporting. Considerable experience has been gained over a number of years with a range of radars and the application of various techniques for analysis and treatment of this problem.[3] For the purposes of RQQI, it is assumed that the radar is calibrated and maintained. Other WMO activities address this issue (Fig. 5).

Figure 6: A photo from the WMO radar technical training course conducted by the Turkish Meteorological Service on calibration and maintenance.

Weather and Radar Environment Variation: Fig. 6 shows some of the impacts of the environment on the weather radar data quality. These are factors need to be mitigated and many are the subject of the inter-comparisons.

[3] The issue of network calibration, important to many agencies employing different radar hardware has been investigated extensively within the Nordic Weather Radar Network and within the U.S. Nexrad radar network. Polarimetric techniques also offer ways forward for self-consistent reflectivity bias calibration.

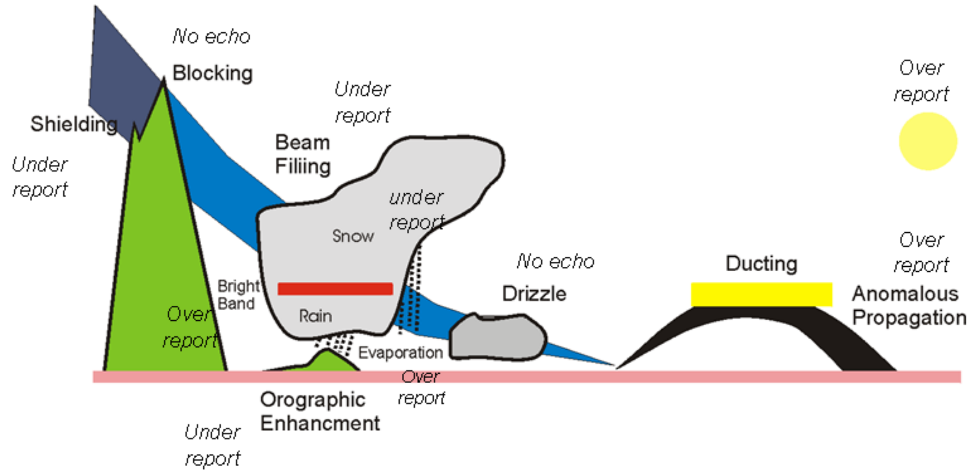


Figure 7: A cartoon depicting some of the physical effects that impact on radar QPE. Even if a radar is perfectly calibrated, these physical effects dominate the use of radar data for QPE. Substantial progress has been made in this area in the past few years and operational systems correcting and adjusting the radar data for these factors are now just emerging. Bold text indicates the physical effect and the italics indicate the impact on QPE.

Processing Radar for Precipitation Estimation

The following figure (Fig. 8) describes the steps required to process radar data to mitigate the environmental effects for quantitative precipitation estimation. The purpose is to illustrate the complexity of the data processing steps that are required to convert radar measurements to precipitation estimates. While the details and sequence may be debated, it serves to illustrate the problems that need to be solved. Historically, many of the steps are and were not performed. The objective of the inter-comparison is to quantify the impact of the various steps for radar QC and QPE. While, the QPE application is the prime focus, the nowcasting application where clear air echo and convergence line detection are used for the diagnosis of convective initiation is also addressed.

Radar hardware changes over time due to failure, under-performance and technology and in order to make use of radar data archives, calibration-related meta data on each radar is needed in order to perform re-analysis of the data.

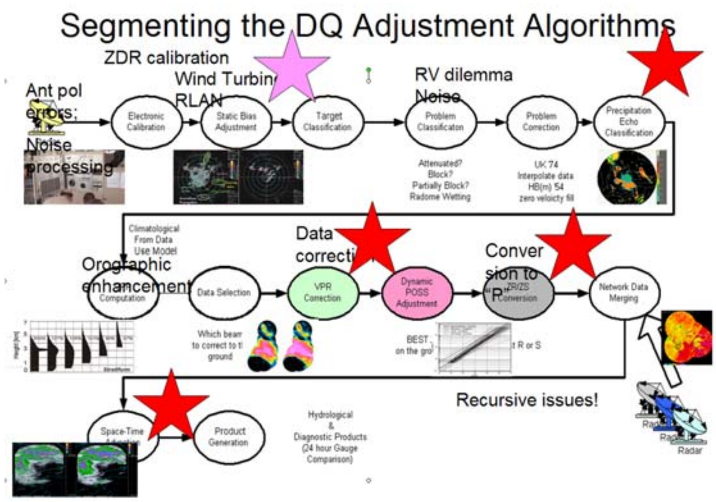


Figure 8: A representation of the processing chain for radar QPE. Currently, only electronic calibration and Z to R conversion is universally done. Impact of other adjustments can be substantially larger depending on the environment and application. The red stars indicate convenient breakpoints for inter-comparison workshops (see text).

1. *Electronic Calibration:* The processing chain begins with radar data (moment data). It is assumed that the radar has been electronically calibrated and antenna gain and polarization issues are resolved as well as noise is properly accounted for (electronic calibration bubble). Signal processing (various Doppler notch or reflectivity statistical techniques) for the removal of ground clutter may be
2. *Static bias adjustment:* This step (second bubble) refers to the potential use of static ground clutter maps that are created to identify ground clutter targets or techniques that adaptively identify ground clutter targets and subsequently used adjust radar data by subtracting the ground clutter pattern, by identifying locations where QPE estimates are likely to be poor, or to determine locations for applying the signal processing clutter filters. Wind turbines or radio-frequency interference are considered static clutter targets though they may be time varying. In practice, ground clutter echoes fluctuate due to fluctuations in the beam propagation paths due to atmospheric humidity, temperature and pressure fluctuations. These changes in propagation paths result in changes in reflectivity fluctuations. Dual-polarization (Zdr signatures) may be used to identify clutter targets. Speckle filters (in either the signal or data processing steps) may be used to remove point targets such as airplanes and ships.
3. *Target Classification:* The next step is to classify the character of the radar echo. At this stage, there will still be non-precipitating echoes. For QPE applications, precipitation or non-precipitating targets (ground, insects, airplanes, etc) are not desirable. The introduction of dual-polarization plays a significant role here. In the past, reflectivity thresholds were often used. However, there is overlap in reflectivity values from non-precipitating and precipitating echoes. This step may include the identification of second trip echoes.
4. *Problem Classification:* The next step is to further classify the quality the precipitation echo – whether it is a blocked, partially blocked or attenuated (by

precipitation or by a wet radome). Dual-polarization techniques are very useful here. For non dual-polarization radars, blocked or partially blocked or attenuated beams may be identified through accumulation analysis, use of digital elevation models, using neighbouring beams or other techniques.

5. *Problem Correction:* Removal and adjustment for the presence of ground echoes, blocked or partially blocked echoes is then envisaged to produce a clean data set of radar data. Data from vertical scans may be used to compensate for blocked or partially blocked echoes. *At this point, the best "cleaned volume" of radar data has been produce* - precipitation echoes are now identified; clutter echoes have been removed; non-precipitating echoes have been identified; the data has been adjusted for partially or blocked echoes. This is often sufficient for NWP assimilation where the NWP models will invert the data itself. For QPE applications, the next steps are to prepare to estimate surface precipitation.
6. *Precipitation Target Classification:* At this point, the three dimensional precipitating weather echo volume is classified as stratiform, convective, bright band, rain or snow or hail in preparation for the vertical profile correction and Z-P[4] relationship.
7. *Orographic Adjustement:* If it is appropriate, an orographic (enhancement) correction is applied. The correction is different if the situation is convective or stratiform. Statistical or climatological information may be used to account for orographic enhancements. However, external information such a Froude numbers may be used to determine the nature of the correction. This is really a orographic VPR correction.
8. *Base Data Selection:* In preparation for the VPR correction, a base data is selected for extrapolation to the surface. This is usually, but not necessarily, the lowest elevation angle. A higher elevation angle (may be a function of the azimuth and range and also on the algorithm) may be of higher quality (not censored by the clutter filters or not in a blocked sector).
9. *VPR Computation and Correction*: At this point, the base data is extrapolated to the ground using an appropriate VPR. The correction will depend on range and also the precipitation target classification (stratiform or convective, etc) determined from an earlier step. The VPR may be determined by different means – climatology, data near the radar, sounding, NWP model or other. The objective is to compute the reflectivity value at the surface. Note that this includes the snow VPR corrections where snow measured at 1 or 2 km above the surface may be advected many tens of kilometers due to strong low level winds.
10. *Dynamic DSD Adjustment:* If there is a DSD device that can used for calibration – where Z values from the presumably highly calibrated DSD device and the weather radar are compared and adjusted accordingly. This step is a conceptual step and it is not clear that it should be implemented but for completeness, it is included in the processing chain. The DSD device can also be used to dynamically determine the Z-P relationship. *At this point, the optimal or best reflectivity at the surface has been determined.*

[4] I use Z-P, P for precipitation, instead of Z-R to indicate that the relationship may be with rain, snow, hail or mixed conditions.

11. *Z-P:* Once the best Z is determined, a conversion from Z to precipitation is done using a Z-R, Z-R, Z-A relationship. Depending on the spatial and temporal scales of the QPE application, the relationship may be static (same all the time), quasi-static (seasonal) or dynamic (convective, stratiform, etc). *At this point, the optimal or best physical precipitation estimate has been generated.*
12. *Raingauge Adjustment*: This step is NOT shown in the processing chain. The objective of the processing was to produce a physical correction scheme so that errors would be identified and corrected at their source. If there are sufficient rain gauges available, believed and an appropriate adjustment technique (depends on the corrections already applice) is available, then the R or P values may be statistically adjusted. *At this point, the best statistical precipitation estimate has been generated.*
13. *Network Merging:* Then the radar data may be merged on a network basis. It may be done on a volume basis, or on a surface product basis. It may be interpolated to a 3-D grid dynamically as data arrives or it may be done on a fixed schedule to account for temporal scan cycle differences. Envisioned techniques include nearest radar, maximum value or best quality. The merging of rainfall estimates from overlapping and non-overlapping radars into one product can overcome: (i) the sudden attenuation due to a wet radome or heavy rainfall, (ii) issues relating to identifying and removing artefacts from data and (iii) merging rainfall estimates from radars with different operating and error characteristics (if data quality estimates can be provided from single radar products). An inter-radar adjustment at the boundaries (for example, by comparing neigbouring radars) could be included. *At this point, the best regional or network estimate of precipitation intensity has been generated.*
14. *Space-Time Merging:* Generation of accumulation products requires a space-time merging. Motion fields are computed and used to interpolate the precipitation estimates (and precipitation type) to finer time resolution and then accumulated. This accounts for the coarse temporal sampling for fast moving storms. *At this point, the best regional or network estimate of precipitation accumulation has been generated.*
15. *Product Generation:* Different applications require different types of products – for example, at specific verification points, areal basin estimates, daily, monthly or storm totals. This is strongly user-dependent.

Notes on the Processing Chain:

1. Final comment is the the diagram shows a linear process but recursion and repetition of steps could occur.
2. Underlying the entire chain are: (i) the scan strategy, data sampling and signal processing, (ii) the atmospheric and clutter environment, (iii) type of weather.
3. Inherent in the processing chain are concepts of a “raw” Data Model. Annex 2 describes a coarse framework or definition of “raw data”. Within this framework, the inter-comparison deals with different levels of “raw data” or “level 2” data (using the terminology of Annex 2). Each of the steps in the processing chain can

be conceptually used to define “raw data levels” 2a, 2b, 2c, etc. The concept is sufficient for now and each raw data level will be defined as the project progresses.

4. Related to data levels are data quality concepts. Each step represents an improvement in the quality or maturity level of the data. Quality representations are needed. These concepts are immature and will need to be addressed as the project progresses.
5. The processing chain is complex and there are natural breakpoints in the chain to make it more of a tractable problem. In the figure, these breakpoints are indicated by red stars and in the text by italicized font (steps 5, 10, 11, 12, 13 and 14). A series of inter-comparison workshops are envisaged. To reduce the number of workshops, topics will likely be combined as experience and tractability dictates.
6. Currently, perhaps only electronic calibration and Z-R are globally implemented (red stars). The effect of the other adjustment factors can be substantially greater, particularly at long ranges from the radar.
7. *Vertical Profile Effects:* Other physical processes can also dominate the estimation process in a practical sense including the contamination of rain by melting hydrometeors e.g., hail and the brightband effect. These effects are particularly important in stratiform rain for non-tropical winter time situations and have regional, seasonal and local variations that impact the estimation process. Fig. 9 (left) show theoretical vertical profiles of snow, stratiform and convective situations and the effect of beam smoothing and beam propagation as a function of range. At long ranges, considerable reduction in surface precipitation is observed by the radar due to the broadening of the beam and the increasing altitude of the beam with range. On the right, is correction of these effects in the Finnish radar network. At long range, correction of 20+ dB are not unreasonable. This is considerably more significant that 1-3 dB electronic calibration uncertainties.

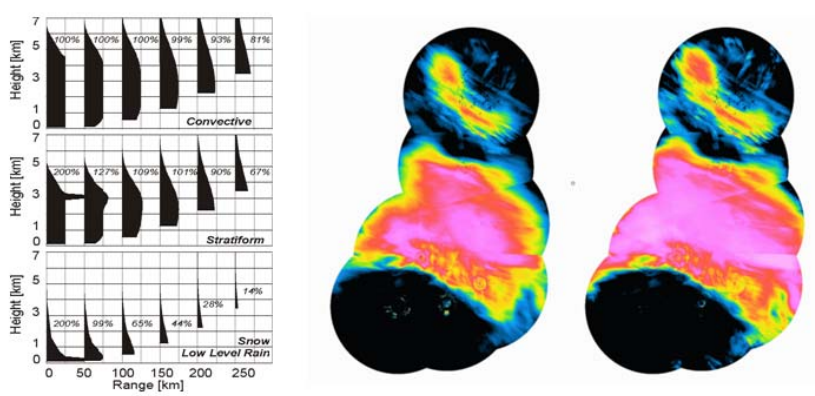


Figure 9: An example of a precipitation accumulation product from the Finnish radar network where the vertical profile of reflectivity corrections have been applied to a case of a quasi-stationary front where precipitation accumulations are expected to be uniform. On the left is an uncorrected image and it show a drop off of accumulation with range from the radar. This is due to beam filling (beam not filled and hence precipitation is under-reported) and vertical profile decrease of reflectivity (VPR) with height leading to underestimate of surface precipitation. The figure on the right shows an accumulation after VPR correction.

8. In the following figure (Fig. 10), the precipitation is accumulated over a winter season from Finland. The fall off with range in the accumulation on the left most figure is due to the VPR effect. The near circular and uniform fall off is indicative of a very high quality radar site free of artifacts. On the right of the figure, is a comparison with the range gauge data. The falloff is evident and the scatter is relatively small. *This is a key image upon which the concept of the radar quality metrics are based.*

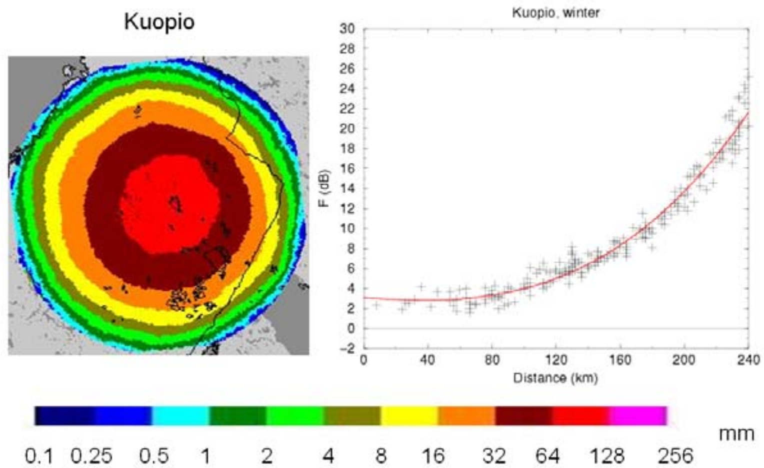


Figure 10: A accumulation of radar derived precipitation for Kuopio radar Finland. On the right, is the comparison of precipidation accumulation as a function of range. Figure courtesy of Daniel Michelson.

The following figure shows a similar plot but this site has terrain blockage and it is for summer. The increased scatter is due to the greater variability of the Z-R relationship in convective situations but also due to the data in the partially blocked area.

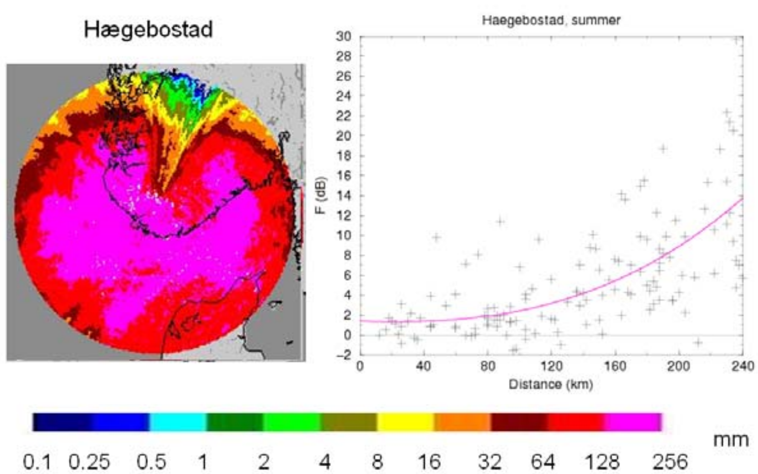


Figure 11: Similar to the previous plot except for a radar with partial blockage and for summer. The greater scatter in the figure on the right is due both to the area of blockage but also due to the greater scatter in the Z-R relationship for summer precipitation. Figure courtesy of Daniel Michelson.

9. *Radar Scanning Geometry:* The limitations imposed by the above physical process are further confounded by issues related to the scan strategy, geometry and data/signal of the radar viewing of precipitation. Potential non-uniform filling of the radar beam with increased range, the height of the radar beam above the ground, blockage of the beam by obstacles especially in mountainous terrain, ground clutter are examples of non-meteorological factors that can further limit the accuracy and bias precipitation measurements. Fig. 3 shows an example of a corrected precipitation map. These factors vary significantly from site to site and under various conditions.
10. *Drop Size Distribution Factor:* Precipitation is inferred most frequently through the reflectivity factor Z, which is well known to suffer significant limitations imposed by the microphysical changes impacting drop size distribution. With a climatological Z-R the limit of measurement accuracy is approximately 30-40%. With adaptive methods based on precipitation type this can be improved as it can be through the adoption of polarimetric approaches (combinations of Z_{DR} and K_{DP}). The physical limits to accuracy are then somewhere near 10-15% with systematic variations in dropsize characteristics. Such systematic variations occur in orographic situations with low concentrations of large drops, convective maritime with high concentrations of smaller size drops, convective subtropical/continental modes with lower concentrations of large drops and of course within different stratiform precipitation processes e.g., melting of large dry snow and cases with small rimed ice. The accuracy of QPE is strongly situation dependent and therefore it is necessary to sample a wide range of weather, locations, and radar configurations when characterizing QPE errors.

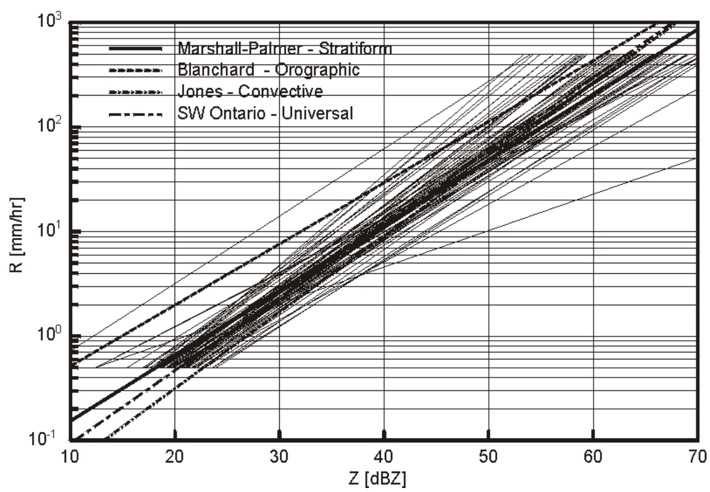


Figure 12: A survey of Z-R relationships found in the literature. The thick line in the middle is the original Marshall-Palmer relationship.

Validation

Absolute Validation: The validation of the radar estimates of rainfall, typically undertaken by intercomparison with gauges raises a number of serious issues. The critical issues are :

1. radars and gauges measure rainfall on different scales and both are affected by individual sampling and instrument errors.

2. Dense gauge networks are crucially important with attention to calibration. So, quality control of network gauge data is also an important consideration.
3. Utilization of independent, systematic and consistent validation processes and techniques are certainly desirable and validation techniques are still in development. This is also true for NWP and satellite validation and climate change applications. In the latter cases, three kinds of validation are envisaged:
   a. physical validation where the physical assumptions of the retrievals are validated,
   b. integrated validation where networks or hydrological basin outflows are used for validation and
   c. statistical validation where space-time statistics are used to compare results.
4. As absolute validation is application dependent and only needed at later stages (step 11+), these issues will be addressed at later stages of the project.

Relative Validation: A tractable concept for validation is "relative validation". Fig. 10 shows what a "perfect radar" (almost) should produce when radar data is accumulated over a winter season. The image (Fig. 10, left) show an annular pattern which indicates beamfilling issues (and perhaps some attenuation) of shallow precipitation. It is largely free of environmental artefacts (ground clutter, airplanes, EMI, turbines, etc). The comparison with gauges (Fig. 10, right) shows a smooth relationship with range (not range dependent artefacts) very tight scatter (low variance). There appears to be a 2-3 dB systematic bias at near ranges and the smoothness with range indicates that the bias is consistent throughout the radar range.

Variance Metric for Validation

Concept: Fig. 10 (right) shows the relationship between radar derived precipitation estimates versus gauges. The analysis was conducted over a winter season and it introduces a Z to P conversion that is not always consistent at different locations. In fact, in summer the scatter is often larger (Fig. 11). For the intercomparison, it is desirable to remove this factor since there may not be many gauges under the radar coverage and it can be application and environment dependent.

The curve can be generated from the radar data alone by a similar process:
1. accumulate reflectivity data[5]
2. compute the average and variance of reflectivity with range (in an annulus)
3. if the data is free from artefacts, then the curve should be smooth with range and the scatter about the mean should be small
4. An overall metric would be various statistics of the range dependent variance such as average or maximum variance value. The maximum value is sensitive to

[5] This condition is relaxed later. Only a "uniform" reflectivity field is needed. This could be produced by a long seasonal data set, a short data set if the accumulated reflectivity field is uniform (e.g., widespread band of precipitation that is uniform over a few hours or a day) or even a single data set (of uniformly widespread precipitation)

outliers whereas the average is much less sensitive but may be dominated by the sheer numbers of good comparisons and may not readily reveal the benefits of the adjustments. Other statistics of variance (e.g. RMS, pairwise difference comparisons, etc) may be more revealing as experience dictates but it is sufficient for now to use the overvall average and maximum variance as the validation metric.[6]

Restated, the inter-comparison validaton metric is a statistic of the variance of reflectivity with range. This statistic can be used to describe the relative improvement of the adjustment procedures.

Computationally: It follows from above that

$$\sigma^2(R) = \Sigma_\theta(Z(R)-\langle Z(R)\rangle)^2/N_\theta(R) \tag{1}$$

where R is range from radar
Z is linear reflectivity
Σ_θ is the summation in azimuth
$N_\theta(R)$ is the number of points in azimuth at the specific range
$\sigma^2(R)$ is the variance of reflectivity as a function or range

and

$$\sigma_{Ave}^2 = \langle\sigma^2\rangle = \Sigma_R\, \sigma^2(R) / N_r \tag{2}$$

$$\sigma_{Max}^2 = \max(\, \sigma^2(R)\,) \tag{3}$$

where σ_A^2 is the overall average variance
σ_{Max}^2 is the overall maximum variance

Inter-comparison Modality

Modalities: Two modalities are identified for the inter-comparisons:

1. Comparing algorithms against a common dataset
2. Comparing an algorithm on diverse radar datasets (in terms of hardware and technology configuration but also in different weather and beam propagation regimes - mountains, coastal, flat terrain, etc).

Process: The proposed intercomparison process is summarized as follows:

- Existing rainfall estimation applications developed by various groups are applied to a set of radar datasets.

[6] The premise for the RQQI data quality metric is therefore that perfectly corrected reflectivity accumulations should be smooth. Spatial correlations or variograms can therefore be defined and used to quantify the smoothness of the resulting ground clutter and AP corrected fields. Given the weather, the shape of the curve may be steeper or shallower.

- These datasets are collected from a set of representative environments and radar configurations (scan strategies, sampling, etc).
- Datasets in native format will be provided. However, they may need to be converted to a common radar data format. The OPERA ODIM_H5 format has been selected by the IOC as common radar data format at the Exeter IOC meeting of April 2011/
- Individual groups will run their estimation applications on these new independent datasets.
- Ancillary information may also be required (e.g. only geographical location and DEM information for the initial inter-comparisons are envisaged, NWP or other data may be needed later)
- The algorithms will provide output in agreed product form (that is, cleaned data in some simple form or in ODIM_H5 format)
- The output will be processed and qualitatively (expertise) and quantitatively (variance metric) analyzed by the Core Project Analysis and Writing team.
- A draft report will be prepared for review and approval by the IOC.

Characteristics of a Inter-comparison Dataset

Uniformity: It also follows from Fig. 10, that "uniform" radar fields are needed in order to compute the variance metric in order to remove the impact of the variability of the weather.

1. Seasonal (Fig. 10 for a winter example or Fig. 11 for a summer example) or even multi-year accumulation of low level reflectivity is one way of producing such a data set. This is particularly true in meteorological environments where widespread precipitation in not present or where data may only be archived sporadically.
2. A uniform weather pattern is often found as widespread winter stratiform precipitation can also be produced by a single radar data set (Fig. 4b). Even if true, it would be prudent to analyze several hours of such data.
3. Uniformity may be achieved with shorter data sets (say several hours to a day) if the weather is banded but persists for the extensive time period (see Fig. 12).
4. Uniform data sets may be generated through simulation or synthesis. Simulation refers to generating data from theoretical basis model. For example, ground echoes could be theoretically generated assuming a theoretical model of clutter (e.g. gaussian echo with a narrow velocity spectrum with a mean of zero velocity) and weather (e.g. gaussian echo with broader spectrum with a variety of mean velocities). Synthetic data refers to data sets that are a combination of measure data (e.g., I, Q data from ground targets without the presence of weather could be combined with I, Q data from real weather withour ground echoes – say from, high elevation angles – from which level 2 data could be generated). This has great appeal but requires a high level of expertise.

For a data set to be considered for the intercomparison, the accumulation of reflectivity (linear, log) or rainfall amount, are expected to be reasonably "uniform" after the *corrections have been made. They should contain artefacts before correction. This is somewhat subjective and it will be the responsibility of the IOC to use their expertise to select appropriate cases.*

Representative Data: Radar data are collected a vast variety of ways. There are:

- Data characteristics - range and azimuth resolution, samples
- Signal and data processing filters
- Scan strategy (number of sweeps, angles employed) and cycle time
- Wavelength and attenuation characteristics
- Range-Velocity effects
- etc

Data sets from a variety of radars and variety of radar configurations will be needed to assess the global radar data quality.

Representative Radar Environments: Radar data are collected in a wide variety of atmospheric environments with different prevailing artifacts. These include:

- Rural clutter
- Urban clutter
- Mountain clutter (mountain peak vs valley)
- Sea clutter
- Total and partial blockage
- Anomalous propagation (super and sub refraction)
- Multi-path
- Wind turbines
- Multi-trip echoes
- Airplanes, ships, cars, power lines
- etc

Data sets from a variety of atmospheric and radar environments will be needed to assess global radar data quality.

Sample Regimes: The nature of the weather (strength, convective, stratiform, Arctic, Lake Effect snow, Tropical, etc) is an important factor. While the project will be reliant on volunteer contributions, the following is a list of envisaged or potential contributors:

i. High Latitude Regime Including Mixed Phase Precipitation (Winter rain)–
    1. Nordic weather radar network (FMI)
    2. UK MetOffice
    3. Polarimetric Radar Montreal (McGill University)

ii. Tropical Maritime (Tropical convection)
    1. Darwin, (CPOL, Polarimetric), Australia (BMRC)
    2. Melbourne, Florida or Kwajalein (NASA TRMM)

iii. Sub Tropical Regime (Moist severe weather regime)
    1. Brisbane (Polarimetric -CP2), Australia (BMRC)

        2. Beijing, PRC (CMA/BMB, B08 FDP)
        3. Japan (JMA)
        4. Korea (KMA)
        5. Hong Kong (HKO)

    iv. Continental Regime (Arid regimes with severe weather)
        1. KOUN Oklahoma, USA (NNSL)
        2. CHILL, Colorado, USA (CSU)
        3. SAWS, South Africa

    v. Mountainous (Orographic enhancement process and blocking)
        1. Meteoswiss, Switzerland
        2. Catalunya, GRAHI-UPC, Spain
        3. Germany (DWD)
        4. France (Meteofrance)
        5. Canada (Vancouver 2010 project)

    vi. Signal processing QC
        1. NCAR
        2. NSSL

Project Deliverables

- Intercomparison data sets
- Standard data quality metrics.
- Intercomparison Workshops
- Workshop summary reports.
- Reports/publications on best techniques and best practices for radar QC and QPE.
- Recommendations for global implementations
- Training workshops

Annex 1: Glossary

Data Processing – processing of the moment data.

IOC – International Organizing Committee, responsible for reviewing and approving the report and making recommendations to WMO member.

Project Team – subset of IOC responsible for analysis, writing draft report and ensuring project schedule

Products – refers to highly processed raw data into relevant end-user units such as mm/h

Raw Data – generally refers to the moment data but is described in Annex 2.

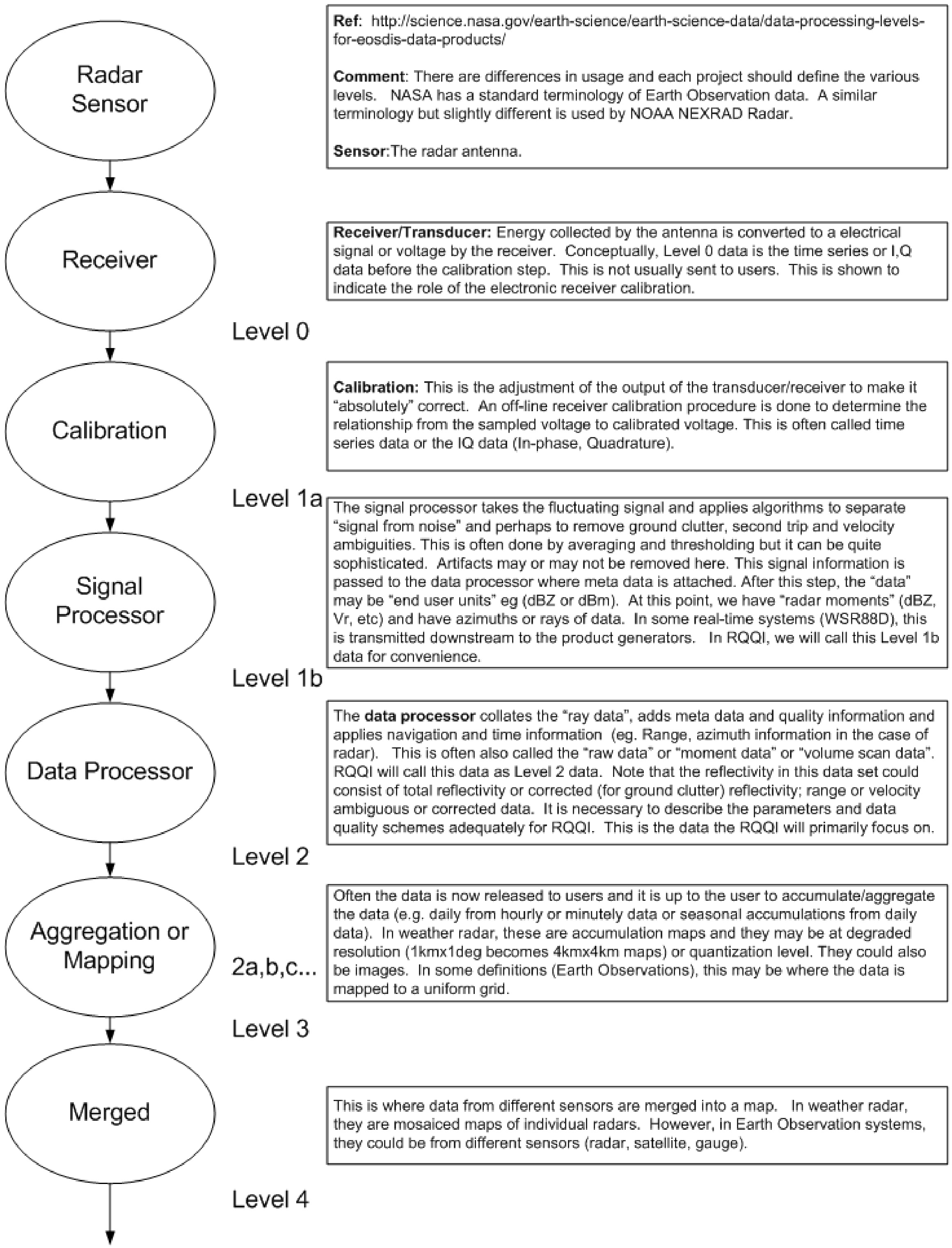
Signal Processing – processing of the time series, or IQ data

Simulated Data – Raw Data that is created based on theoretical concepts

Synthetic Data – Raw Data that is create based on the combining of actual measurements

Time Series Data – this is the voltages or the in-phase or quadrature data

Annex 2: Raw Data Definitions



Annex 3: IOC and Project Team Membership

Name (* indicate Project Team)	Organization
Paul Joe (chair)*	Canada, Environment Canada
Liping Liu*	China, Chinese Meterological Agency
Yoshihisa Kimata	Japan, Japan Meteorological Agency
Alan Seed	Australia, Bureau of Meteorology
Estelle de Coning	South Africa, South African Weather Service
Vincenzo Levizzani	Italy, WCRP/WGPRN
Daniel Michelson*	Sweden, Swedish Meteorological Hydrological Institute
Daniel Sempere-Torres	Spain/Expert, GRAHI
Nicholas Gaussiat	UK/UKMO-OPERA HUB
Tim Crum	USA/ NOAA-ROC
John Hubbert*	USA/Expert, NCAR
Roberto Calheiros	Brazil, IPMET

Other Experts

Norman Donaldson*	Canada/Expert, Environment Canada
Michael Dixon	USA/Expert, NCAR
Ken Howard	USA/Expert, NOAA
Jian Zhang	USA/Expert, NOAA
Mark Curtis	Australia/Expert, BOM
Jan Sturz	Poland
Pier Paolo Alberoni	Italy
Ronald Hannesson	Germany/Gematronix
Heikki Pohjola	Finland/Vaisala
Malcolm Kitchen	UK/UKMO
Hidde Lijense	NL
Sorin Burcea	Romania

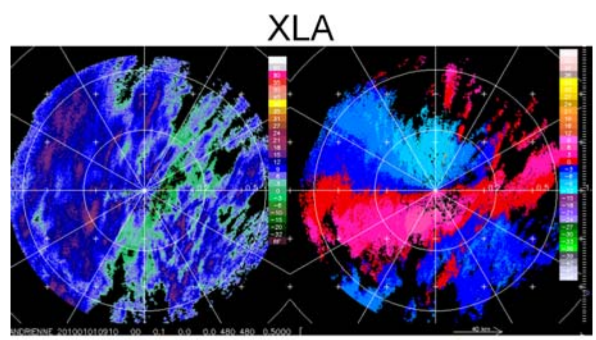
Annex 4: Exeter Kick-Off Meeting - Pilot Mini-Project

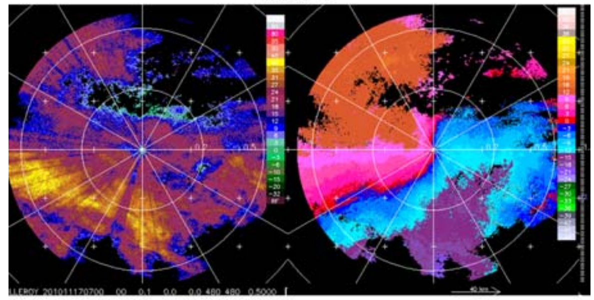
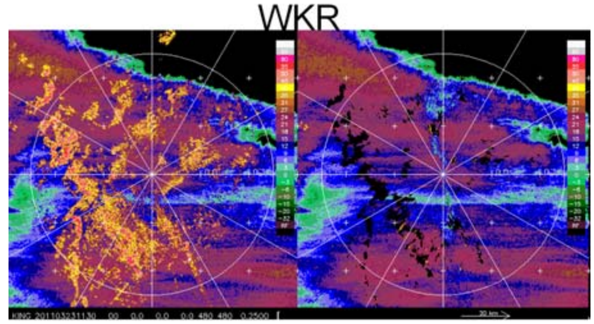
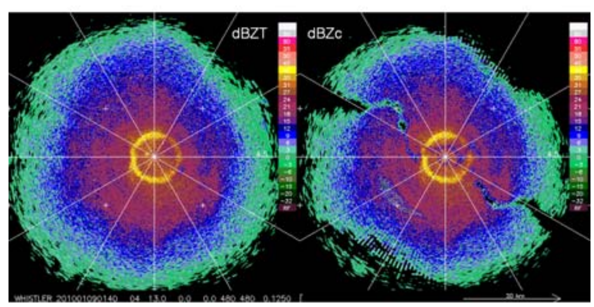
In preparation for the Exeter kick-off meeting of the IOC, a pilot inter-comparison project was conducted to test the feasibility of the inter-comparison modalities. Sample data sets were collated, processed and analyzed. The following briefly describes the mini-project and lessons learned are summarized.

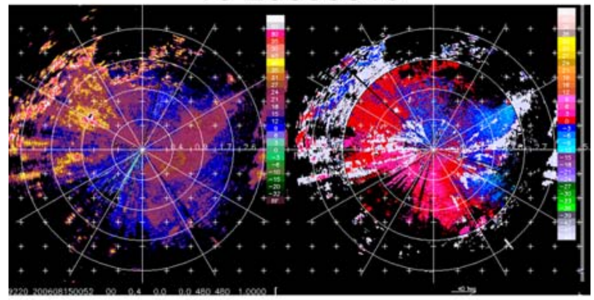
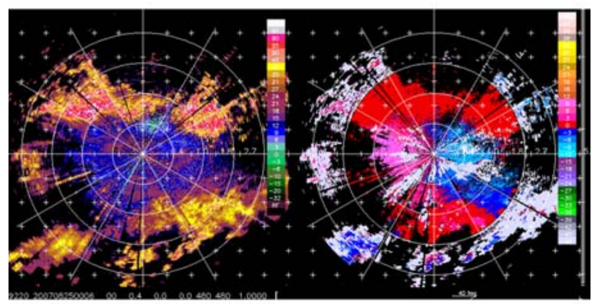
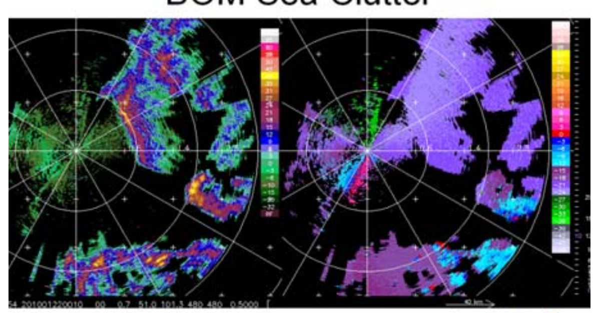
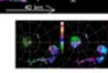
The objective of the pilot was:

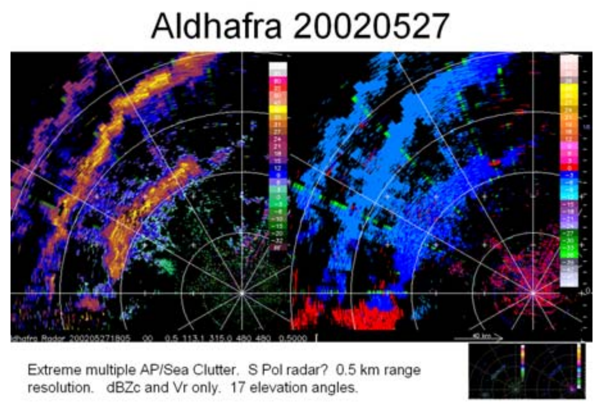
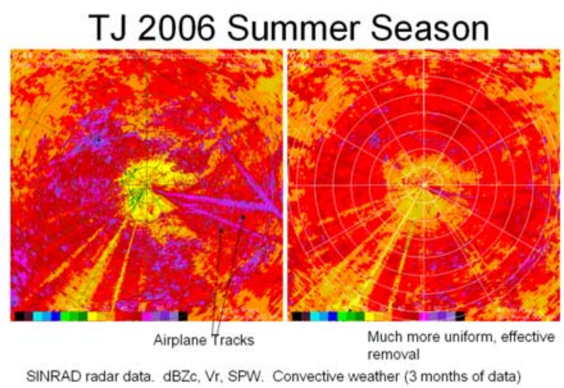
- to mechanically test the procedure,
- to identify gaps in the process and
- to determine if the variance metric makes sense.

The Pilot Case Data: The following table lists the cases selected for the pilot. The cases were selected to illustrate a variety of situations – clutter, weather and radar data difference. Many of the cases were already available to the author. Except the strong multi-path anomalous sea clutter case (Saudi Arabia) which was provided by Cathy Kessinger of NCAR.

	Case	Comment	Sample
1	Uniform Wx with Local Clutter (XLA)	EC volumes scans have 24 elevation angles of Zt data and the angles are optimized to make a CAPPI product. Zt and Zc PPI data are also available. So various techniques may be applied and evaluated. This is a “null” case in that the weather is uniform and clutter is minimal and therefore should produce the lowest variance.	 The data accumulates to uniform pattern. Widespread snow. A baseline case. IRIS formatted data. 24 elevation angles. Doppler (dBZT, dBZc, Vr, SPW) at low levels. Range res = 1km or 0.5 km. Az res = 1 or 0.5 degrees.

2	Uniform Wx with partial blocking (WVY)	EC data. In contrast with case 1, the variance should be higher.	WVY  The data accumulates to uniform pattern with an area of blockage. Widespread snow. A baseline case. IRIS formatted data. 24 elevation angles. Doppler (dBZT, dBZc, Vr, SPW) at low levels. Range res = 1km or 0.5 km. Az res = 1 or 0.5 degrees.
3	Urban Clutter/Niagara Escarpment (WKR)	EC data. In contrast with case 1, should have higher variance	WKR  The data accumulates to uniform pattern with an area of blockage. Widespread snow. Urban (skyscrapers) and small terrain clutter. IRIS formatted data. 24 elevation angles. Doppler (dBZT, dBZc, Vr, SPW) at low levels. Range res = 1km or 0.5 km. Az res = 1 or 0.5 degrees.
4	Zero Notch Filter removes too much echo (VVO 2009)	EC data. BB case with uniform weather	Zero Notch Filtering  Same as last image except dBZT vs dBZc

4	Strong Anomalous Propagation Echo (TJ 2006)	CMA data – coarse elevation scans, VCP21. Has zero notch filtering applied. Have QC and NonQC data sets. Very strong AP data; variance should show considerable difference. 2 month accumulations already available. Convective weather.	TJ 20060815  Very strong AP echo, clear air echo and 2nd trip. SINRAD format (1/0.25 km range bins, split scans) dBZc, Vr, SPW. VCP21 (9 elevations).
5	Strong AP with Weather (TJ 2007)	Similar to previous case except for the Beijing radar, with mountain clutter and application of sector GC filtering.	TJ AP with Weather 20070825  Very strong AP echo. clear air echo, 2nd trip and weather echo. SINRAD format (1/0.25 km range bins, split scans) dBZc, Vr, SPW. VCP21 (9 elevations).
6	Sea Clutter (Sydney AU, Kurnell)	Uniform sea clutter, different scan angles than EC data, no Zt data available.	BOM Sea Clutter  Kurnell doppler radar. Sea clutter in AP conditions with weather. 0.25 km dBZc and Vr data. 

7	Sea Clutter / Multi-path AP (Saudi 2002)	Multi-path AP mixed with Sea Clutter; very strong; a published test case for Fuzzy Logic SC.	
8	Convective Weather with Airplane Tracks - One season (TJ Radar 2007)	CMA data with many point targets.	

Pilot Algorithms: The author had access to various clutter removal algorithms or had before/after algorithm data sets. The following algorithms were available.

Algorithm	**Brief Description**
CAPPI	Constant altitude slice through the volume scan data were made at 1.0, 1.5 and 3.0 km AGL. EC uses a 24 elevation scan with Zt data.
Doppler Notch Filtering	Signal processing technique. EC collects PPI data with both Zt and Zc data as part of the archive. Can also compare with the CAPPI techniques.
Fuzzy Logic – AP	The author coded up a version of the NCAR Fuzzy Logic technique that removes Anomalous Propagation echoes. This was applied to the sea clutter case as well.
Fuzzy Logic - SC	The author coded up a version of the NCAR Fuzzy Logic technique that removes Sea Clutter echoes. This was applied to all the other cases as well. This algorithm differs from the previous one in that radial velocity is not available and the membership functions are tuned differently.
Fuzzy Logic – CMA	The author had both QC and NONQC data from CMA that was part of the B08 FDP project.
PRECIP-	Norman Donaldson processed the data using a prototype EC algorithm that

ET	uses echo top, vertical reflectivity gradient to identify clutter. It was only run on a limited number of cases.

Variance Results

These results have not been extensively validated and are presented to illustrate the inter-comparison process.

Case	Technique							
	RAW	CAPPI15	CAPPI30	Precip-ET	DOPPLER	REC_AP	REC_SC	CMA_AP
BOM_seaClutter		7.3	8.1		8.5	8.7	8.5	
Saudi_20020517		23.6	20.3		26	7.3	26.4	
Saudi_20020527		18.8	15.3		21.2	10	20	
TJ_bigAP_20060815_nonQC		8.2	7.2		8.7	7	9.4	
TJ_mix_20070825_nonQC		8.4	8.8		8.2	8.5	8.8	7.2
TJ_mix_20070825_QC		7.5	7.9		7.2	7.9	8.1	
WO_zeroNotch_snow	7.3	7.1	5.6		6.7	6.7	6.7	
WKR_cnTower_Snow_20110324_dop*	11.5	10.3	8.4	7.4	9	10.3	10.6	
WYR_uniform_blockage_20101117_dop	9.3	9.4	8.1	6	8.8	8.6	8.8	
XLA_uniform_20100101_dop	5.9	4.4	3.1	3.2	4.6	4.4	4.5	

Lessons Learned

1. Most systems only archive corrected reflectivity and only a few archive total reflectivity. A separate experiment for I, Q data should be considered to evaluate notch filters.
2. A reminder that this test evaluates the ability to produce smooth fields with low variance as expected and should not be interpreted as the best reflectivity or best estimate of precipitation at the surface.
3. The algorithms should be run at the native resolution of the data. That is, the data should not be interpolated to a common azimuth and range resolution. There are too many assumptions needed to do the interpolation.
4. Data sets must have uniform precipitation to properly evaluate over-correction of the adjustment techniques.
5. Not all techniques can be fairly applied to all data sets.
    a. This is fairly obvious but this should be kept in mind when processing case data in batch mode. For example, many techniques require a substantial number of elevation angles (greater than four; that is, a volume scan) to properly apply.
    b. Some techniques require radial velocity and some require spectral width. So these techniques can not be applied to those data sets.
6. The variance or variations of this metric are reasonable to quantitatively evaluate the techniques. Additional qualitative analysis is needed to determine if it makes logical sense.
7. It seems that some corrections (zero notch) introduce low reflectivity artifacts.
8. Conversion to ODIM_H5 may require information/metadata external to the data files themselves.

Annex 5: Milestones

See Exeter Meeting Record.