P14.2

Representing Validity for Canadian Weather Radars

Norman Donaldson*

Cloud Physics and Severe Weather Research Section,
Environment Canada, Toronto, Ontario, M3H 5T4, Canada

1. INTRODUCTION

Environment Canada (EC) is receiving increasing numbers of requests to provide information on the quality of gridded radar data products. Such requests apply to both the entire datasets and individual measurements. This is largely driven by applications that use automated analysis to combine radar datasets with other datasets. Two common combinations are radar composite images and merged radar/gauge precipitation maps. In these applications the validity and relative quality of each data value is required in order to select or weight each source.

Three issues hinder responses: i) Users usually conceive quality in terms of only one of the factors that impact radar data quality. ii) Assessing the various quality factors is difficult. iii) EC has no general schema for representing quality in gridded remote sensing datasets. All three issues need to be addressed. Many groups have been working on assessing and quantifying quality, so this work is focused on the other two issues. The ultimate objective is a schema to handle quality that is suited to our real time systems. A preliminary step is to lay out taxonomy of quality related issues both to guide design and as an aid for response to users. An intermediate objective is to have adequate representation of validity. It is these two initial objectives that are the primary topic of this discussion.

In this discussion a distinction will be made between "validity" and "quality". Validity addresses questions of *whether* data is good for a specific application, while quality addresses questions about *how good* valid data is. Both depend on the context of the intended application.

* Corresponding Author: Norman Donaldson, Environment Canada, 14780 Jane St, King City, ON, Canada. E--mail Norman.donaldson@ec.gc.ca

The initial focus is on validity because it is the most basic property of a remote sensed gridded dataset. "Validity" includes factors like whether a data cell is empty because it was not scanned, or if data has been rejected for some reason. Even empty grid cells require validity information, although there is no measurement or other quality information.

Figure 1 illustrates the current situation. This image shows no radar targets, and leaves the user wondering. "Is there is nothing?", or simply "Do we know nothing?"

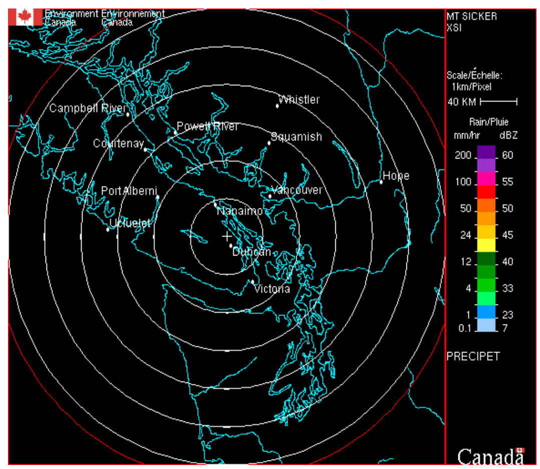

Figure 1 A sample radar image with no depiction of validity

The discussion will present an analysis of properties of data, aspects of quality and validity categories. Then two proposed minimal schema for handling these in our operational setting will be outlined.

2. BACKGROUND

To put EC's existing data representation in context, the weather radar program was originally intended for severe summer weather but has evolved into a multi-purpose program. The original requirement to display areas of intense precipitation has evolved to include severe weather techniques using weak clear air echoes, and winter weather. Newer requirements are being added by numerical weather prediction and hydrology applications.

The original severe weather context essentially had two categories of data: valid moderate-to-strong echoes and everything else ("0"). This crude representation of quality has been more or less inherited and is no longer adequate.

The existing radar processing system uses a distributed model of computing, both in terms of operating on computer clusters and in terms of intermediate products from one computing centre being forwarded to others. This places some limitations on file sizes. See Joe and Lapczak (2002) for an overview of the Canadian weather radar system.

Although there had been significant consideration of quality in terms of accuracy, the advent of dual-polarization on research radars has explicitly introduced other concepts, such as relevancy. For example, one can have accurately measured echoes that are assessed to be non-meteorological signals to some level of confidence.

In parallel to the work with the radar network, EC has a large on-going project called the Data Management Framework that will revise the handling and documentation of quality in more traditional measurement systems.

Our involvement with the NINJO project (Koppert, 2004) strongly suggests that handling of validity and quality issues is an urgent problem beyond just our own radar system. It is needed both for display purposes and for intelligent combination of data from multiple radars and potentially from other observational systems. Articles by Henley (2005, 2006) suggest such problems are common across the geosciences.

3. ANALYSIS OF DATA AND QUALITY

3.1 DATA PROPERTIES

There are seen to be four essential properties of a gridded dataset:

The measurement
The validity
The quality
The history of the data and processing

These are defined as follows.

- *The measurement* is what people usually think of when they refer to "the data". In a gridded dataset measurements might not be available at all locations. For example a point on the grid might be out of range of radars. Measurement in the current context is intended to include not just raw measurements. It also includes data that have been adjusted by post processing (e.g. VRP correction) or derived products, such as CAPPI or echo top maps. Further extension would include synthetic data, such as that obtained by interpolating over gaps.
- *Validity* answers the question "Is there relevant information here that addresses the problem at hand"? A grid location will always have validity, even where measurements are missing.
- *Qualit*y answers the question "how good is the measurement for the user's application. Obviously this only applies where there is a measurement. This is context dependent. For example, returns from insects in the boundary layer are irrelevant to hydrology applications, but can be exceedingly valuable to severe weather forecasters looking for convergence lines.
- *History* records where the data originated and what transformations were performed. Ideally this information would be available for each pixel. In a sophisticated system it is not unusual for dozens of tests and modifications to be undertaken so a complete history might contain two orders of magnitude more information than the measurements themselves. This seems completely impractical in our operational context.

The representation of history at every pixel is not addressed here, except as a suggestion that pixel history could exist as an ancillary dataset, rather than embedded in the main dataset. The existing EC data format includes a section of "meta data" that includes partial history of algorithms applied. A recommendation here is to ensure this is more complete.

Note that "validity" is highly intertwined with "quality". Invalid data can be thought of as having no quality and vice versa. Validity is however more basic for a gridded dataset, since it applies even in the absence of measurements.

3.2 QUALTITY ASPECTS

In a different context, Smith (2004) laid out a list of aspects of quality:

Accuracy and precision
Relevance
Timeliness
Coherence
Comparability
Acceptability and Clarity

All of these are relevant to weather radar to some degree. The first three are probably the most critical to weather radar.

- *Accuracy* of the measurement would include both the technical meaning of precision (e.g. number of decimal places) and the broader issue of variability of the measurement estimate, including uncertainty added by applying algorithms.
- *Relevance* would include such things as whether or not the targets are meteorologically interesting (ground clutter versus insects versus precipitation). This is tightly related to validity.
- *Timeliness* is whether data arrives quickly enough to be useful. This is primarily a measure of the quality of the data processing and communications systems rather than the data itself but could enter processing. For when delayed data arrives too late to be used with real time data, it might trigger reprocessing of earlier times. When combining datasets that are not synchronized, timeliness considerations will determine whether an older dataset should be combined with a new one. The concept can be extended in the radar context to include distance.
- *Comparability* in Smith's context means whether a dataset can be compared over time and to similar measurements from other sources. In the radar context it could reflect the size of the volume being sampled, so for example a large radar sample volume (several cubic kilometres) centred on a point at long range from one radar may not be directly comparable to a smaller sample volume around the same point from a closer radar, nor to a rain gauge with an aperture of 100 cm^2 underneath the radar point. In fact, since the size of radar sample volume increases with range, while the radar sensitivity decreases, the individual pieces of data along a single radar ray are not completely comparable, despite coming from the same sensor.
- *Coherence* is the uniformity of datasets that come from a combination of different sources. Three examples illustrate the variety of factors that can reduce coherence. i) A combination of Doppler corrected reflectivity measurements with conventional reflectivity measurements, or a combination of radar with NWP, would be less coherent than the original datasets. ii) If a radar composite is made from radars that scanned the atmosphere at different times then there is reduced temporal coherence. iii) a dataset where gaps are interpolated becomes less coherent because it will be a mix of primary data and secondary inferences.
- *Acceptability/clarity* refers to the usefulness/completeness of accompanying metadata. For example with a grid of radar data,
- "Clarity" reflects whether the grid's parameters are completely specified and the schema for missing cells is clearly indicated within the data system.

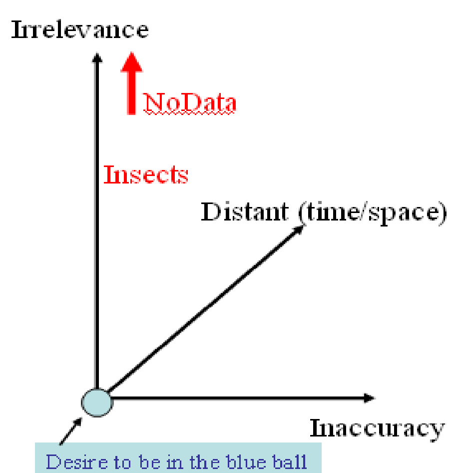

Figure 2: Quality of measurements for a particular purpose is multi-dimensional.

Figure 2 symbolically shows how quality for a specific application is multi-dimensional. Good quality means minimizing a number of problems.

3.3 CENSORING

If a measurement is deemed to have intolerably low quality it should be hidden in some way from subsequent processing. This is referred to as “censoring” and can take two forms. The first method, “hard censoring”, is to overwrite the measurement with some flag value such as the infamous “-999”. The second method is to create a parallel dataset holding information that data should not be used (and preferably the reason for rejection).

3.4 VALIDITY CATEGORIES

There are seen to be 7 validity categories:

A) Valid measurement
B) Below minimum detectable signal
C) Not sampled Type 1
D) Not sampled Type 2
E) Censored
F) Corrupted data
G) Substituted *

- A *valid measurement* is essentially “good” data that does not fall into one of the other validity categories.
- *Below minimum detectable signal* means that a measurement was attempted, but no significant signal was found. This is an example of partial validity. It provides an upper limit on the possible value but does not determine it to the precision of the instrument.
- *Not sampled Type 1* means that the location was not sampled at all, such as beyond maximum range of a radar. This is known from within the radar system.
- *Not sampled Type 2* means that it is known from some external information that no measurement is possible at the range and direction indicted. The prototype example of “not sampled Type 2” situation is a sample location that is hidden from the radar by an intervening mountain. In this example the sample location does not receive any radio energy from the radar, but the radar itself does not know that and it is even possible for artifacts to introduce measurement values into such hidden locations. It can be argued that Category D is a variation on either Category C or E, but to avoid pedantic debates about where it belongs, it is included as a separate category. In this situation there is no information and quality concepts do not apply.
- *Censored* means that the measurement was rejected for some quality reason. The prototype of this for weather radar would be rejection of ground clutter echoes.
- *Corrupted data* is data that has been invalidated during collection or transmission. This is useful to know, since it implies the data might be recovered by returning to the source.
- *Substituted indicates* the presence of some secondary information. In some contexts it is required to fill “holes” in data (original data would be in Categories C-F) with other data, such as by interpolation or from another source. It should also be noted that any data in Category G should also be linked back to the original status is Categories C-F. It can also be argued that substitution is a “history” issue and this is not a basic validity category at all. Once substituted the data “cell” moves to Category A.

In some contexts it is desirable to apply a software threshold on low reflectivities, say when weak insect echoes dominate over weak precipitation echoes. It is not envisioned that Category B would cover this situation. If a Category A signal is converted to a “zero” measurement, one would have a valid “zero” measurement. Depending on the representation scheme this situation may or may not be easy to logically distinguish. Let us label this situation as Category A, Type 2.

When the validity information is present one can convert the situation represented in Figure 1 to that in Figure 3, where grey areas around the outside show where no data is present (out of range) and the magenta areas show areas where data is absent. In the case of this mountain radar, there are sectors that are blocked by other mountains, and there are zones where ground echoes are so bad that one could not tell if weak precipitation were present.

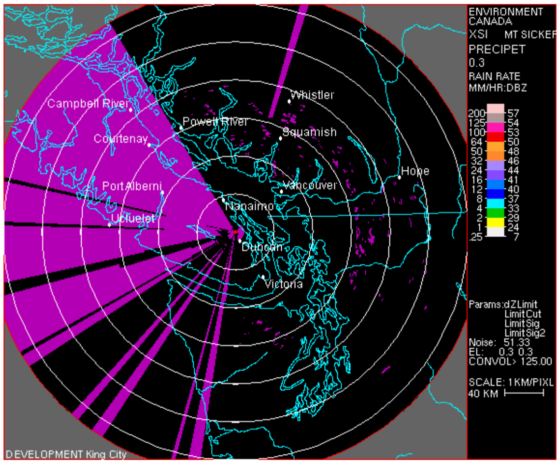

Figure 3 As with Figure 1, but now with validity information displayed.

4. REPRESENTATION OF VALIDITY

4.1 EXISTING DATA REPRESENTATION

Within the existing radar processing system data is stored intermediate files referred to as “meta files”. This has proved to be robust data containers. These files start with a plain-text header section, which includes metadata such as time, location, product type, and critical information like grid dimensions and geographic projection information. The actual data section follows and is most often a polar coordinate grid holding data. The data itself can be either ASCII representation of floating point numbers or one-byte indexed data. In the case of indexed data, the header section will include a conversion table for converting the indices to their true values.

As mentioned above there are essentially two types of situation represented, which are validity category A and everything else, which is represented as “0” in indexed files.

Despite its existing crude quality handling, the system was designed to be expandable in a number of ways. For example, although the data section is typically one single valued grid, it is extendable by either allowing the gridded data to be couplets or for there to be multiple gridded fields.

4.2 CONSIDERATIONS

To design a quality representation a number of factors need to be considered and balanced.

A practical real-time schema cannot place undue stress on computing resources such as CPU, disk IO, disk storage or telecommunications. Backward compatibility is desirable, and preferably existing software would ignore future extensions that it does not know.

Against these restrictions it is desired that the schema be as flexible and complete as practical.

Hard censoring is considered undesirable because it effectively destroys the original measurements. That means that if subsequent processing were to reassess the censoring criteria the original data could not be recovered.

It also needs to be considered that the various internal measures of quality must be aligned with quality criteria of multiple users. Holleman et al (2005) has a useful discussion of this issue.

It is desirable that that scheme be able to carry the same concepts between different processes, while still allowing different forms of storage (indexed data, floating point, local to software, etc).

One limitation of the entire quality representation is that the science for a merger of all quality aspects into a single number does not currently exist. For the foreseeable future the best that can realistically be hoped for is a quality index in the range (0, 1) that presents relative merit in a particular context.

Finally it is desired to bring the real time system as close to possible to accepted best practice. See for example Latini and Passerini (2003).

4.3 RECOMMENDATIONS

The preliminary recommendation is to move forward with two possible representation schemes. The first is to use a hard censoring schema that is more general than the existing one. The second scheme would be to use a soft-censorship scheme to augment the existing grid of measurements with a parallel grid containing a combined validity and quality field (VQ). The second scheme is the better solution but computer resource issues may require an intermediate step using the first scheme.

In both schemas, the mnemonic keywords in Table 1 provide a common link between different datasets and software. The numerical values representing each validity category do not need to be uniform between datasets, so each dataset needs to include header information specifying the values assigned to each key. Application software can remap validity categories into its internal representation using the keywords.

Table 1 Proposed validity flags and keys

Flag Value	Mnemonic Key	Category
0	VALID_ECHO	A
1	NO_ECHO	B
2	NOT_SAMPLED	C
3	NO_DATA	D
4	CENSORED	E
5	CLUTTER	E’
6	Censored - spare	E’
7	SUBSTITUTED	A’

For both schemes the header sections need to be made inclusive to hold information such as the ranges of valid data and specify the values of all special flags being used. A cumulative history of processing techniques applied to the entire dataset needs to be given.

The first proposed scheme would be to extend number of validity flags used. In the existing indexed scheme the single index “0” would be supplemented with other values. (See Table 1.) These additional flags would need to come from the existing set of indices, so there is a reduction in the number of measurements that can be represented. For floating point representations the issue of removing possible values is unimportant. The selection of flag values is flexible and needs only to be outside the valid range of the data. . The mnemonic keywords in Table 1 need to be linked to the values representing them, which might change between datasets. For the most part much backward compatibility in the EC systems can be achieved through the use of index tables in the headers. Category A, Type 2 situations are not easily handled in the indexed case.

The second proposed schema has the original gridded measurements left more or less intact. Another, parallel field would hold a merged quality and validity measure. Some validity issues still arise in the measurement section, such as distinguishing echoes that fall below measurable minimal detectible signal, but these are hoped to be more or less trivial. One storage possibility is that 1 byte data could be used, for the combined data, with 3 bits for validity flags and 5 bytes to hold a quality index, with the original range 0 to 1 being mapped to 32 levels. Preferably, if data storage permits, the validity and quality (VQ) data could be stored in a second and third grid.

Two approaches can be considered for handling the supplementary VQ data grid in data files. Either it can be embedded directly in the existing data files with the measurements, or it could be packaged in separate files. With embedded VQ data it is trivial to maintain a link to the related measurement fields. With VQ in separate files careful bookkeeping is required to keep the measurements and VQ linked together. On the other hand, having a separate VQ “product” file makes it easier to maintain a single base measurement field and then produce multiple VQ products aimed at different applications.

In terms of backward compatibility, for the second scheme it appears that much existing EC software would simply ignore the supplementary field if it were embedded, whereas new software could exploit the supplementary quality data. If data is stored as an independent product then there is an even higher degree of backward compatibility since existing software would be not know about the supplementary files.

5. CONCLUSIONS

A conceptual framework as been raised around the validity and quality issues associated with weather radars, with the aim of better representing these in Environment Canada’s real time processing systems.

Two recommendations regarding possible schema for representing validity within our systems is suggested. These are deliberately minimal, but it is clear that a complete representation is impractically cumbersome for real time systems. The proposals seem to be a reasonable compromise.

The next stages are discussions about operational implementation with the computer engineering teams and the development of some simple quality indices for real time users.

6. ACKNOWLEDGEMENTS

This project was undertaken with the support of Environment Canada's Atmospheric Science and Technology Directorate The author wishes to acknowledge many useful discussions with people inside E.C. and internationally. Bob Paterson also helped prototype concepts. Paul Joe has been helpful both of himself and by pointing me to other people. .

7. REFERENCES

Henley, Stephen, 2005: The man who wasn't there: The problem of partially missing data. Computers & Geosciences, Volume 31, Issue 6, Pages 780-785

Henley, Stephen, 2006: The problem of missing data in geoscience databases. Computers & Geosciences Volume 32, Issue 9, Nov 2006, Pages 1368-1377

Holleman, I, D. Michelson, G.Galli, U. Germann, 2005: Quality information for radars and radar data. OPERA project. Available from www.knmi.nl/publications

Koppert, H, T. S. Pedersen, B. Zürcher, and P. Joe, 2004: How to make an international Workstation Project successful. 20th International Conference on Interactive Information and Processing Systems (IIPS), Seattle, American Meteor Soc. http://ams.confex.com/ams/84Annual/techprogram/paper_71789.htm

Joe, P., and S. Lapczak, 2002: Evolution of the Canadian operational radar network. ERAD Conference, Delft, ERAD Publication Series, Vol. 1, ISBN 3-936586-04-7, 370-382.

Latini G, and Passerini G, 2003: Handling Missing Data: Applications to Environmental Analysis, Wit Pr/Computational Mechanics 200 pp (Book with CD).

Smith, Lisa 2004: RADAR Data Quality Framework. U.K. Dept Environment Food and Rural Affairs, (Defra), http://www.defra.gov.uk/foodfarm/farmanimal/diseases/vetsurveillance/documents/radar-dataqualityframework.pdf Last accessed 2009-09-28