Comments on

# ASPRS Accuracy Standards for Digital Geospatial Data

By Harold Schuch, 12/02/13

Introduction

First of all, it is great that ASPRS should again get involved in the accuracy debacle. Also, I’m fully aware of the difficulty of creating standards while technology continues to be in flux. As such, my comments are not just limited to the “map scale” angle, but more generally to a modern world. Hope you don’t mind all my comments. They are late in coming, because I wanted to give them sufficient time to develop and review. SO, here we go:

Page 1, line 8

Given that so many Public Works departments continue to work with contours as their main “vertical” product, one will still see ½ contour interval as the best-understood vertical accuracy standard in the long term. This, with ¼ ci for spots should remain (the ¼ ci accuracy can be converted to 2 sigma confidence interval). This would help strengthen the link with published maps. I chose 2 sigma arbitrarily since a 95.44% confidence interval is close to 95% that most lay people find desirable and easy to explain in court. I say that because 1 sigma or RMSE is not (68.26% CI “sounds” too lenient).

Page 1, line 19

What is “network accuracy”? Is this the result of an adjusted traverse? Or is it the baseline accuracy of a GPS network offered by NGS or a vendor? Or is it the RMSE of a least-squares-adjusted trilateration-triangulation network? I would either drop the reference to “network”, or be more specific.

How do you relate otho product accuracy to this? I suggest bringing into play some form of message based on error propagation, such as “control point RMSE shall be at least 1.4 times as accurate as required ortho product accuracy” [ taking the simplest case of RMSE(ortho)= ±sqrt(2xRMSE(control)) and setting RMSE(control)=1 ]. This is in line with what is presented in the appendices.

I like to arbitrarily ask for RMSE(control) to be 4 times better than RMSE(final product). Or, since block adjustment is normally involved, then: RMSE(control) is better than 2x RMSE(block adjustment), which is better than 2x RMSE(tested end product). But I would never consider less than an x1.4 factor (your 1.414 in the appendices).

Or, better yet, I would directly address error propagation as it is incurred: From original reference point to final mapped product, something like this (one leads to the next, and error propagates down the line):

Base station coordinate accuracy

> Base station GDOP

> AGPS GDOP

> AGPS/IMU accuracy (= a priori perspective center accuracy, which

includes coordinates and space angles)

> Block adjusted perspective center accuracy (if push-broom or LiDAR)

> Sensor accuracy (LiDAR, scanner, camera – as based on calibration or not)

> Block adjustment pass/tie point precision (if frame-based imagery)

> Block adjustment control point precision (if frame-based imagery)

> Map product accuracy

> Published map accuracy of measurements thereon.

Page 1, line 22

I feel that this paragraph does not go far enough. The reasons for this are the following:

* An original digital image file has a scale simply because the following are involved: (a) There is a flying height (or object distance), (b) there is a scan angle (for push-broom or even LiDAR) or frame size (for image sensors), plus a “look angle” (vertical, low-oblique, high-oblique, convergent). Laymen and the courts can easily be convinced that a scale is involved.
* Once data have been georeferenced, only then will there be no scale (or 1:1) because data are stored in ground coordinates. Only then will the argument of “no scale” in the usual sense hold. This suggests that one may have to separate “raw data scale” from final product “no scale”.
* Bringing pixel sizes into the scale discussion is unnecessary and dangerous. The reasons are that (a) raw pixels generally are not square, (b) GSD is only square when the ’look angle of the sensor” is perpendicular to the surface scanned (as in Nadir-located data). Otherwise, pixels will cover a non-square ground (object) surface (near the horizon they can even cover vast areas). An even pixel size and square format can only be specified for ortho maps after processing (after resampling), and not for “Digital Imagery” (title of this paragraph).
* As such, the term “GSD”, due to its general non-square nature for raw data, has to be defined (and identified) as a characteristic of either raw images (non-square) or final ortho product (square). My choice would be for the second.
* So, should the specs cover only a few square image pixels (e.g. Nadir)? I think that they should only cover ortho products. Any initial data collection should be considered “intermediate production data” to be left out of the specs. This would be in line with all the recent specs (USGS, FEMA, FGDC) that let the user define the accuracy of end products. It would then be the user’s specifications that have to set the “purposes” of a project and “what detail has to be seen” on the end product. Of course, the alternative would be to introduce “precision” (=resolution) as another measure of intermediate product “quality” in support of final product “accuracy”.
* Alternatively, if GSD is applied to raw images, should it not reflect an average of pixel sizes within a raw image/swath-covered area (after cookie-cutting to cull neighboring coverage where overlaps occur)?

Page 2, line 48

In truth, I find that the handling of vertical accuracy needs to be linked to horizontal accuracy. As much as it is true that TIN models “normally” don’t show breaklines, and denser point sampling can remediate that to a practical degree, the whole issue of horizontal positioning is not addressed. For example, specs have to call for any vertical data have to be “correctly” positioned in a horizontal sense (including being in the right datum!) prior to any processing. One of the existing specs (I forget which) even admits that horizontal accuracy of, say, LiDAR, is four or five times less accurate than vertical accuracy. This needs to be addressed and firmed up.

I understand that most large-area projects follow “historic” LiDAR specs that call for one point every meter or so (or two points within a 2m x 2m cell), which makes matching image detail with vertical DEM or TIN data essentially impossible. However, I also understand that technology now allows (a) much denser LiDAR clouds combined with (b) intensity and even color data, which makes image matching possible. This should be brought to the foreground and included in specifications.

Page 2, line 57

First of all, I would now move away from using “photos” in the specs language (like in the title).

I feel that the language of this whole section needs to be strengthened in various ways. However, the principal two aspect to be addressed are the following, to help the laymen and the courts. At the center of my comment is my belief that “RMSE” and “1 sigma” should be left out of all discussions for two important reasons:

* Any mention of an accuracy level related to “only” 68% of data is normally seen as suspect by laymen and the courts.
* Talking about RMSE and sigma on one side, and 95% confidence interval on the other is an unnecessary duplication.

It is clear that all this discussion should be related to “end-product sampling results”, and that this does not mean that this is the accuracy of each and every one of the points tested. For example, there are still statements out in the marketplace that rave about sub-pixel RMSEs for block adjustments (really only a measure of “precision”), giving the customer the impression that every pixel in the whole ortho product has that “accuracy”. Customers normally don’t know that it take at least three adjoining pixels in a row or 3x3 square to start being able to see anything. (Note that this is also something used for GPS measurements, when a resulting “sub-centimeter” position is the result of the mean of hundreds of positions that are definitely not in that range.)

Then, what constitutes a test sample? What is a minimum sample? Can one realistically test 10% of all the pixels in an ortho map? Is one to test “20” points in each ortho tile, or in each coverage, or each image/swath, etc.? Or should one ask for 20 points every square mile / kilometer? I have seen specs asking for “20 points” for projects covering very large areas.

In terms of calculation, does one sample each tile and then calculate a mean for all the tiles? Or not? I’m throwing things out that may be ludicrous, but I have seen it done.

When we are talking “RMSEx and RMSEy is Pixel size x 1.0”, is that raw image or end-product pixel size? (Also relate to non-square GSD discussion above.) This table leaves all recent projects in the dust, simply because they (a) specify 3” or 6” GSDs, but use LiDAR data calling for 1 point within every 2m x 2m square on the ground. Therefore, this table asks for the impossible, unless more LiDAR points are collected.

Page 3, line 81:

All RMSE values should have a “±” in front (as the result of the “R” in RMSE).

Apart from that, I notice that the 1.25% limit, taking a 1:1,200 scale map (or 1”=100’), corresponds to about 1/125th of an inch, or 20 microns. That’s a far cry from the already difficult-to-implement 1/30th or 1/50th inch. What tool will be used to measure at that precision? What paper map will require that accuracy? What paper material will support that accuracy (shrinkage)?

General Comments

As it turns out, the single most important contributor to digital data accuracy (whether printed or not) is the AGPS/IMU platform. Especially for push-broom cameras, the use of ground control is made difficult, and projects are mostly held in place by AGPS/IMU data (perspective center coordinates and space angles). This is especially true for LiDAR data collection.

For starters, and as I tell my students, you cannot stop the plane at a point in space and take 500 GPS readings over a period of four hours. Therefore, the kinematic use of GPS by airplanes introduces error (a type of “position smear”) that is far above the usual “cm accuracy” claimed by vendors. The actual adjusted AGPS/IMU error level, after adjustments, has been estimated variously at about ±1 foot.

Apart from the AGPS error, there are other major aspects that have to be considered:

* IMU Borehole calibration
* GPS antenna vs. sensor position calibration
* Correct timing between units (AGPS/IMU/sensor)
* Correct use of complex software to process data
* Use of a “good” vs. a “poor” AGPS/IMU unit (expensive vs. less so)

As such, it may become necessary to establish best-practice parameters for this technology, as a concern that may override any discussion about ground control (unless square sensor cameras are deployed, allowing the use of pass/tie points and ground control).

Also, in line with some of the above comments, there is a need to firm up a few things that are normally not being addressed (but should be):

* How closely does the DEM or TIN have to be positioned within the georeference that is held by imagery to be orthorectified (horizontally and vertically)?
* What are acceptable AGPS/IMU error levels to achieve the stated standards?
* To what an accuracy does time have to be tracked to achieve these accuracy levels? The reason for this is that when a GNSS unit is in motion, all the internal timing issues come to the foreground, such as each satellite is really observed at a slightly different time, processing speeds in hardware/firmware/software, time to excite the antenna, time to output and store data, etc.

Finally, I have noticed that there is less than a general awareness of basic statistics in the community. I would include in Appendix C a brief section on the Standard Distribution Curve that includes:

* An indication of the source for some of the numbers. They can be extracted from many sources, such as <http://www.mathsisfun.com/data/standard-normal-distribution-table.html> (especially pretty and practical implementation).
* A clear message that accuracy levels can be expressed in various ways, but only one should be used consistently to avoid confusion. I mean that the following really express the same accuracy level for sample testing:
* RMSEx = ±5 feet
* X error = ±5 feet at 1 σ
* X error = ±5 feet at 68.26% CI (CI=confidence interval)
* X error = ±8.23 feet at 90% CI (σ x 1.645 from Normal Distribution tables)
* X error = ±9.80 feet at 95% CI (σ x 1.960 from tables)
* X error = ±10 feet at 2 σ
* X error = ±10 feet at 95.44% CI (σ x 2.000 from tables)

My suggestions here are the following for any document explaining accuracy:

* It should use only one version consistently (I suggest the one using 95% CI as highlighted above) to avoid general confusion and shenanigans, and
* It should contain an explanation that accuracy can be expressed in various equivalent ways (giving samples like above).

Whew! I look forward to your comments.

Best regards,

Harold Schuch, PE, CP, LSI