THE SPECIFICATION AND VALIDATION OF PREDICTED ACCURACY CAPABILITIES FOR COMMERCIAL SATELLITE IMAGERY

John T. Dolloff and **Henry J. Theiss**
InnoVision Basic and Applied Research Office
Sensor Geopositioning Center
National Geospatial-Intelligence Agency (contractors)
7500 GEOINT Dr, Springfield, VA 22150
{john.t.dolloff, henry.j.theiss}.ctr@nga.mil; PA case # 14-199

ABSTRACT

A primary application for commercial satellite imagery is Multi-image Geopositioning (MIG). It utilizes a weighted least squares (WLS) algorithm in order to solve for a best estimate of 3D ground location typically using a stereo pair of imagery. It also simultaneously provides a corresponding 3x3 error covariance matrix characterizing the predicted accuracy of the particular solution. Previously, specifications for commercial imagery only addressed the accuracy of the solution, typically in terms of 90% horizontal and vertical error, or CE90 and LE90. The CE90 and LE90 values are typically fixed, and as such, intended to be conservative, i.e., interpreted as at least 90% of horizontal errors and 90% of vertical errors are less than the specified CE and LE, respectively, applicable to all image sets within the specified range of operational imaging geometry. No specification for the reliability of the MIG solution's predicted accuracy had been included. Along with imaging geometry, predicted accuracy is influenced primarily by the sensor support data (sensor position, attitude, etc.) error covariance matrices and temporal correlation models, typically provided in the metadata. It varies for each particular set of images used in a MIG. The lack of a specification for accuracy prediction capabilities is a significant liability for various organizations that rely on commercial imagery for MIG-type activities. This paper presents a proposed method for the inter-related specification of both accuracy and accuracy prediction capability requirements for commercial satellite imagery and includes various examples of its validation. In addition, both absolute and relative accuracy are addressed.

KEYWORDS: accuracy prediction, specification, validation, imagery, multi-image geopositioning

INTRODUCTION

This paper presents recommended methods for the specification of both accuracy and accuracy prediction capability requirements for commercial satellite Electro-Optical (EO) imagery. There are specific specifications corresponding to stereo extraction absolute accuracy, both horizontal and vertical, and mono extraction relative accuracy, horizontal only. The method for the specification of accuracy prediction is new. In fact, the authors know of no such previous specification, regardless of form. The performance of accuracy prediction capabilities is of vital importance to many geopositioning applications and to the National Geospatial-Intelligence Agency and its customers in particular. The recommended methods are also practical and directly related to sample-based validation of the requirements relative to ground-truth points. These methods are also readily extendable to non-commercial and air-borne imagery as well.

The specifications are in terms of solution accuracy assuming Multi-image Geopositioning (MIG) based extractions – an optimal Weighted Least Squares (WLS) solution technique, where measurement weights are based on sensor support data (sensor position, attitude, etc.) uncertainty projected to image space plus mensuration uncertainty. A MIG not only provides an estimate of the target's 3D ground location, but a corresponding 3x3 error covariance matrix that represents the solution's "predicted accuracy". This predicted accuracy is tailored (applicable) to the specific imagery with its particular image geometry and support data uncertainty used in the solution – it is not simply an overall prediction applicable to all MIGs and all imagery for a specific type of sensor. For stereo extraction, one target is extracted and two images are used; for mono extraction, multiple targets are extracted simultaneously from one image in conjunction with an external elevation source.

For a specific MIG (target) solution, the corresponding 3x3 error covariance matrix can be used by an

application to generate a 90% 3D confidence ellipsoid such that when centered at the MIG's 3D coordinate solution, there is a 90% probability that the true target location is within the 3D ellipsoid. If the solution is to support validation of accuracy and accuracy prediction, ground truth is also available and the ellipsoid when centered at zero is termed an error ellipsoid. There is a 90% probability that the solution error (target minus ground truth) is within this ellipsoid. Of course, this assumes that the ellipsoid (error covariance matrix) is correct, which is to be validated.

Figure 1 presents an example of a 90% confidence ellipsoid centered at the MIG solution for stereo extraction. The two lines represent the image rays. Note that as discussed later in this paper, a 90% 2D confidence ellipse for horizontal coordinates can also be generated from the upper left 2x2 portion of the error covariance matrix (relative to local tangent plane coordinates) and a 90% confidence line for vertical coordinates can also be generated from the lower right 1x1portion of the error covariance matrix. In addition, from these, CE90 and LE90 can be generated as well. The former is defined as the radius of a circle such that there is a 90% probability that the true target horizontal location residues within the circle; the later the length of a line such that there is a 90% probability that the true target vertical solution resides within +/- the line. Figure 2 presents the CE-LE confidence cylinder corresponding to the same error covariance used to generate the confidence ellipsoid of Figure 1; CE90 forms the radius of the cylinder. Figure 2 also includes the ellipsoid along with the cylinder for comparison.

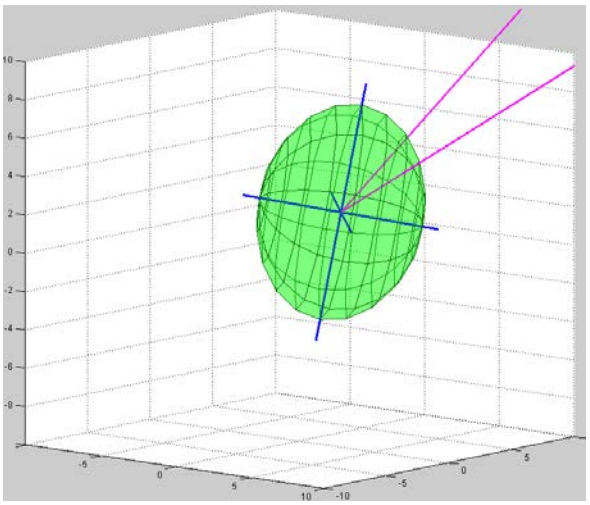

Figure 1: MIG stereo extraction 3D 90% confidence ellipsoid

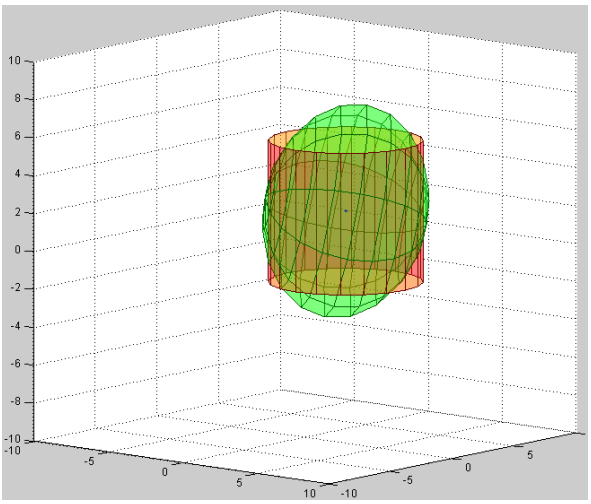

Figure 2: Corresponding CE-LE cylinder added

For the purposes of both specification and validation of the requirements, one pixel (one-sigma) mensuration error is assumed. For purposes of specification associated with monoscopic extraction, the one-sigma error of the external elevation source is assumed negligible. For corresponding validation, the external elevation is typically set equal to the elevation of the ground truth point which the MIG solution is to be compared to. Also, as mentioned earlier, during validation of the requirements, all MIG solutions are compared to corresponding ground truth points in order to compute corresponding errors for analyses. Ground truth uncertainty is assumed negligible (one sigma errors per 3D component less than 0.1 meters). If not, the appropriate uncertainties must be root-summed-squared into the accuracy requirements during validation. (When a MIG solution is compared to (subtracted from) ground truth, this is termed "measured error".) Also, for monoscopic extraction horizontal relative accuracy, there are actually 4 groups of similar specifications, each associated with a different distance bin for an arbitrary point pair within the image. The accuracy requirements are in terms of meters, but the distance bin range is in terms of pixels.

The following accuracy and accuracy prediction requirements apply to the physical sensor model and corresponding image metadata (sensor support data), which includes: (1) the sensor position and attitude data streams, (2) the sensor position and attitude error covariance matrix data streams and any other applicable error covariance matrices associated with sensor adjustable parameters (systematic sensor errors), (3) their corresponding strictly positive definite correlation functions (spdcf) for the temporal correlation of the corresponding sensor errors between images, (4) the "unmodelled" intra-image error covariance matrix of sensor "high frequency" errors which are not associated with sensor adjustable parameters, and (5) their corresponding spdcf for intra-image correlation of unmodelled errors. The above metadata may be contained in a combination of explicit image-specific metadata and vendor published data-base parameters. The requirements for both accuracy and accuracy prediction levy requirements on the quality of this metadata, including that associated with the statistical description of sensor uncertainty.

Although not addressed explicitly in this paper, the specifications can be extended in a straightforward manner for the RPC sensor model. In general, the accuracy prediction requirements for RPC are somewhat less stringent than for the physical model due to inherent limitations in the RPC uncertainty model. The latter is assumed to be the updated RPC uncertainty model as described in [1]. Without the updated model, predicted accuracy for RPC would be unreliable and virtually unspecifiable. Other recommended references associated with this paper include those for image-based extraction processes which make use of the Community Sensor Model (CSM) interface [2], the MIG in general [3], the MIG process applied to commercial satellite imagery and its sensor support data uncertainty [4], stereo extraction accuracy of WorldView-1 imagery [5], and the sensor and MIG error covariance matrices, spdcf, and error ellipsoids [6]. Unmodelled error is covered in [2] and [4] as well.

Finally, the accuracy specifications presented in this paper do not include specifications for relative accuracy for stereo imagery or absolute accuracy for mono-imagery. The added complexity and "expense" of twice as many specifications and corresponding validation activities is felt unwarranted. It is felt that if the sensor models meet the more limited set of requirements, the others requirement would most likely be met as well. This is because they all involve similar error sources, including the same corresponding sensor error covariance matrices, "unmodelled error" covariance matrices, temporal correlation functions, and intra-image correlation functions.

The remaining sections of this paper now go on to present the recommended specification for accuracy and accuracy prediction capabilities in detail. Following this, a simulation-based example of the validation of these requirements is presented. It provides insight into the need for each of the specific requirements detailed previously as well as the recommended validation process itself, including sample size. For completeness, at the end of this paper alternate metrics associated with the validation of accuracy prediction are presented which better represent MIG-generated "elongated" 3D and 2D error ellipsoids. These are typically not applicable for commercial satellite EO imagery with its tightly vendor-controlled and favorable imaging geometry, and hence, the more familiar CE and LE are used for the metrics. However, this is not the case for many other types of imagery and the alternate metrics are better suited. Finally, insight into the actual values of the "tolerances" used in the specifications applicable to both types of metrics are presented.

SPECIFICATION METHOD OR FORM

Stereo Extraction

The following two subsections present specifications associated with MIG-based stereo extraction.

Stereo Extraction Absolute Accuracy

prob{(dH<=CE90_spec)}>=.90 (1)
prob{(dV<=LE90_spec)}>=.90 (2)

dH is defined as horizontal radial error, in meters, relative to "ground truth", i.e., $dH = (\varepsilon x^2 + \varepsilon y^2)^{1/2}$. Similarly, dV is defined as absolute vertical error, in meters, relative to "ground truth", i.e., $dV = (\varepsilon z^2)^{1/2}$. More specifically, dH and dV are assumed relative to a stereo MIG solution and as compared to ground truth. "CE90_spec" and "LE90_spec" are specification "inputs", in meters, that depend on the sensor type and possibly the procuring agency/customer for the imagery. Their specific values are not important relative to the specification's method and are not presented here. Finally, of course, "prob" is an abbreviation for probability.

Specifications (1-2) are the primary absolute accuracy specifications. They are expressed in terms of specified values for CE90 and LE90, i.e., CE90_spec and LE90_spec, which is typical for many satellite commercial imaging systems. An interpretation of (1) is as follows: it is required the horizontal error associated with an arbitrary stereo extraction is less than the fixed value CE90_spec at least 90% of the time. The following specifications (3-4) now address accuracy outliers:

prob{(dH<=1.8*CE90_spec)}>=0.99 (3)
prob{(dV<=1.9*LE90_spec)}>=0.99 (4)

Specification (3) addresses horizontal "outliers". Theory dictates the use of approximately 1.5*CE90_spec in (3) above to limit the number of outliers at the 0.99p probability level, assuming that the CE_90 specification is "just" met ; the 1.8 multiplier includes a pad to account for possible non-Gaussian errors and the effects of only a reasonable number of validation samples. A "pad" is not included in specification (1) as it does not reference

"outliers" and the system design should have plenty of margin to begin with; as such, a "pad" is felt unwarranted. Specification (4) is the vertical counterpart to (3).

Stereo Extraction Absolute Accuracy Prediction

prob{(dH/CE99_pred)<=1}>=0.97 (5)
prob{(dH/CE90_pred)<=1}>=0.86 (6)
prob{(dH/CE50_pred)>1}>=0.42 (7)

prob{(dV/LE99_pred)<=1}>=0.97 (8)
prob{(dV/LE90_pred)<=1}>=0.86 (9)
prob{(dV/CE50_pred)>1}>=0.42 (10)

All "CEXX_pred" are computed from a MIG solution error covariance matrix corresponding to dH. This error covariance matrix is the upper left 2x2 submatrix of the MIG solution's 3x3 error covariance matrix relative to a local tangent plane coordinate system centered at the ground point's location. As an example, CE50_pred corresponds to 50% circular error probable as computed per Appendix A. Similarly, all "LEXX_pred" are computed from the lower right 1x1 submatrix of the MIG solution's error covariance matrix corresponding to dV. Note that for a given MIG solution, the corresponding dH and CEXX form a specific pair from which the ratios in (5-7) are computed; similarly, the corresponding dV and LEXX form a specific pair from which the ratios in (8-10) are computed .

Specifications (5-7) and (8-10) are the primary absolute horizontal and vertical accuracy prediction specifications, respectively. As an example of their interpretation, consider (6): it is required that horizontal error associated with an arbitrary stereo extraction be no larger than its corresponding predicted CE90 computed from the MIG solution's error covariance matrix at least 86% of the time.

For both horizontal and vertical predicted accuracies, there are actually three specifications to better control the distribution of error-to-predicted accuracy ratios. Also note that the inequality sign inside the probability brackets is "greater than" for Specifications (7) and (10), while "less than or equal to" for the others. Both the need for three specification and the difference in inequality signs are addressed later on in this document. The following specifications (11-12) now address predicted accuracy outliers in a manner similar to (3-4) for accuracy outliers:

prob{(CE90_pred<=1.6*CE90_spec)}>=0.99 (11)
prob{(LE90_pred<=1.7*LE90_spec)}>=0.99 (12)

All of the above (5-12) include reasonable size "pads" to account for a vendor's non-perfect sensor error models, possible non-Gaussian errors, and the effects of only a reasonable number of validation samples. For example, for requirement (6), the value 0.86 is used instead of 0.90 – their difference (and one-sided inequality) is the "pad".

Mono Extraction

The following are counterparts to the above accuracy and accuracy prediction specifications (1-12), but applicable to the relative extraction of a pair of points using monoscopic MIG. Only horizontal extraction is addressed as elevation is an external input for mono extraction.

The following horizontal relative accuracy and accuracy prediction requirements shall be met for each of four distance bins (*i=1,..,4*) with corresponding range intervals in pixels of: (0,200], (200,1000], (1000,5000], (>5000), respectively. Note that these distance bins are in terms of pixels and a different set of values may be applicable to imagery other than commercial satellite imagery.

Mono Extraction Relative Accuracy

prob{(dH_rel<=CE90_rel_spec$_i$)}>=0.90 (13)
prob{(dH_rel<=1.8*CE90_rel_spec$_i$)}>=0.99 (14)

dH_rel is relative horizontal radial error, in meters, relative to "ground truth", and corresponding to an arbitrary target point pair within an image-based distance bin *i (i=1,..4).* $dH\_rel = (\varepsilon x\_rel^2 + \varepsilon y\_rel^2)^{1/2}$, where $\varepsilon x\_rel =$

$(\varepsilon x_1 - \varepsilon x_2)$ and $\varepsilon y\_rel = (\varepsilon y_1 - \varepsilon y_2)$. More specifically, the above dH_rel is assumed from a multi-target MIG solution as compared to ground truth. "CE90_rel_spec$_i$" is a specification "input", in meters.

Mono Extraction Relative Accuracy Prediction

prob{(dH_rel/CE99_rel_pred)<=1}>=0.97 (15)
prob{(dH_rel/CE90_rel_pred)<=1}>=0.86 (16)
prob{(dH_rel/CE50_rel_pred)>1}>=0.42 (17)
prob{(CE90_rel_pred<=1.6*CE90_rel_spec$_i$)}>=0.99 (18)

Predicted accuracy requirements (15-18) apply individually to each distance bin 1-4. All "CEXX_rel_pred" are computed from the 2x2 relative horizontal error covariance matrix and correspond to dH_rel. This relative horizontal error covariance matrix is the upper left 2x2 submatrix of the 3x3 relative error covariance matrix relative to a local tangent plane coordinate system centered at the average of the two ground points' location, and corresponding to the target point pair within distance bin *i*. The relative error covariance matrix is, in turn, computed from the appropriate 3x3 diagonal blocks and 3x3 cross-covariance block from the multi-target MIG error covariance matrix for the two points. As an example, CE50_rel_pred corresponds to 50% circular relative error probable.

Note, as a final thought regarding this section of the paper, the primary accuracy prediction requirements (5-10) and (15-17) are actually independent of the accuracy requirements and can be specified/validated independently, if so desired.

EXAMPLES OF STEREO EXTRACTION SPECIFICATION AND ITS VALIDATION

Let us assume hypothetical values CE90_spec=6 meters and LE90_spec=6 meters for specificity. The following presents simulated MIG solution errors (sample realizations) and related MIG accuracy and predicted accuracy metrics. These are based on assumed values for the MIG solution's 3D error covariance matrix output, which includes the effects of minimal predicted mensuration error uncertainty. The magnitude and the "shape" of the MIG error covariance matrix are reasonable for generic commercial satellite stereo imagery.

Two basic MIG solution error covariance matrices were used in an alternate fashion and with minor perturbations (a few %) to each corresponding to each sample realization. (Two were used for convenience; a larger "collection" applicable operationally.) The error covariance matrices were "full", i.e., there were non-zero cross-covariance elements and different valued diagonal elements. Three hundred independent samples (realizations) of MIG solution error were generated. For each realization, a random 3D error was generated consistent with the appropriate error covariance matrix. Figure 3 presents the corresponding basic 90% 3D error ellipsoids.

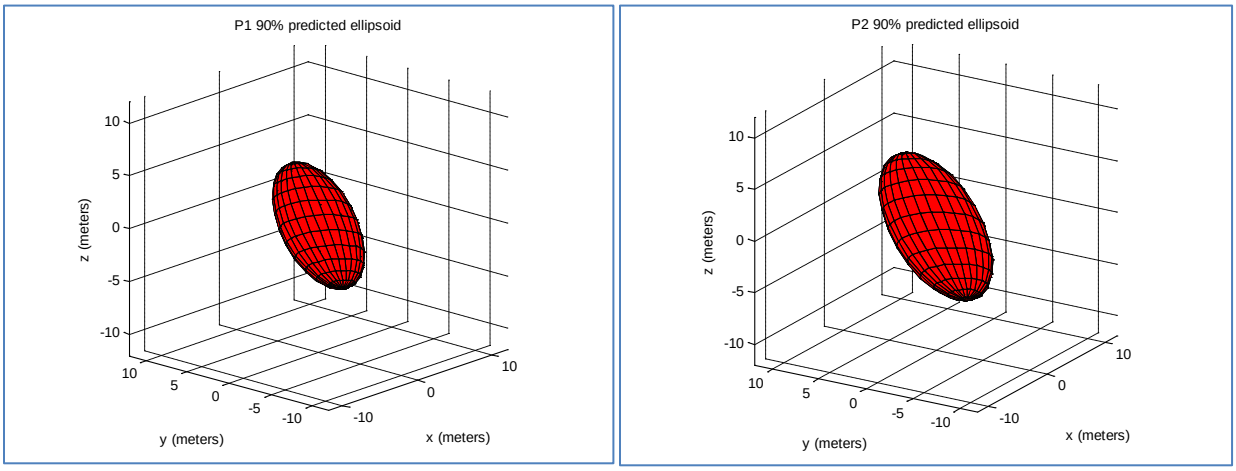

Figure 3: 90% 3D Error Ellipsoids: P1 (left) and P2 (right)

The two basic error covariance matrices are termed P1 and P2. The second (P2) is approximately a larger scaled version of the first (P1). The 2x2 error covariance matrices corresponding to horizontal errors are termed PH1 and PH2 and correspond to the upper left 2x2 sub-matrices of P1 and P2, respectively. Their corresponding 90% predicted 2D error ellipses are presented in Figure 4.

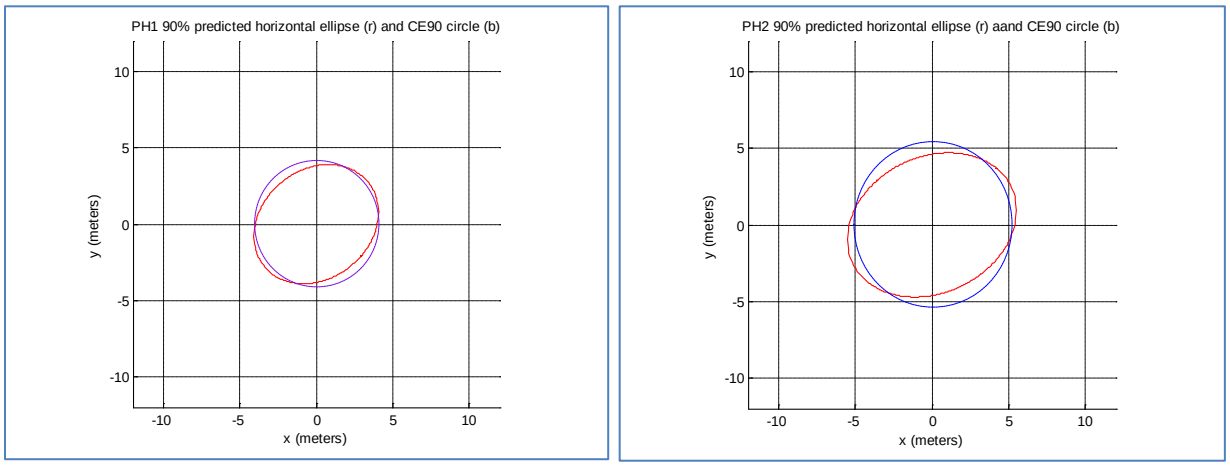

Figure 4: 90% 2D Error Ellipses and CE90 approximations: PH1 (left) and PH2 (right)

$$P1 = \begin{bmatrix} 3.60 & 0.69 & 0.37 \\ 0.69 & 3.30 & 2.87 \\ 0.37 & 2.87 & 3.90 \end{bmatrix} \quad P2 = \begin{bmatrix} 6.60 & 1.13 & 0.60 \\ 1.13 & 4.80 & 4.07 \\ 0.60 & 4.07 & 5.40 \end{bmatrix} \tag{19}$$

Note that the two horizontal ellipses presented in Figure 4 also correspond to associated CE90s, the circles included in the figure, with plenty of margin relative to the stereo extraction absolute accuracy specification's CE90_spec=6 meters, as typically the case. Also, because the ellipses are not appreciably elongated, they are approximated well by their corresponding CE circles.

Figure 5 corresponds to the validation of the horizontal absolute accuracy requirements (1) and (3). Figures 6-8 corresponds to the validation of the horizontal accuracy prediction requirements (6) and (11), (5), and (7), respectively, corresponding to CE90, CE99, and CE50, respectively. All figures reference the same 300 realizations or samples of horizontal error (blue circles) and their x-axis predicted CEXX values computed from the appropriate MIG covariance matrix. All requirements pass. (Note that all of the figures include the "slope 1" line, although it is not needed in Figure (5). Also, for example, "CE90_pred" is referred to as "predicted CE90" in figures (5) and (6).)

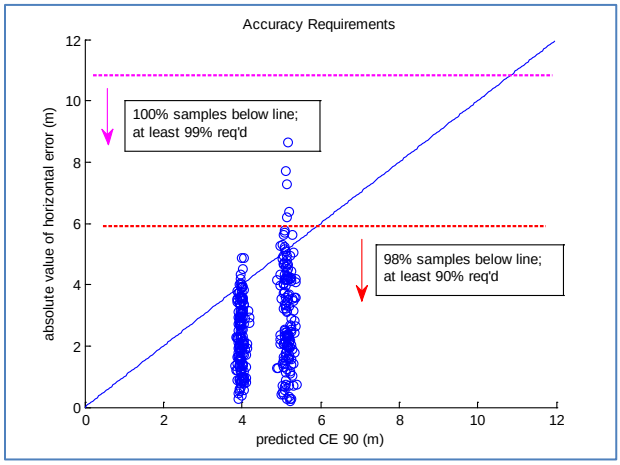

Figure 5: Validation of Accuracy Req's (1) and (3)

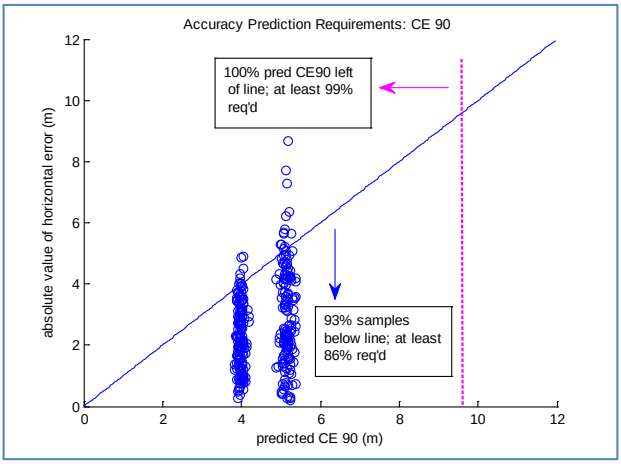

Figure 6: Validation of Accuracy Pred Req's (6) and (11)

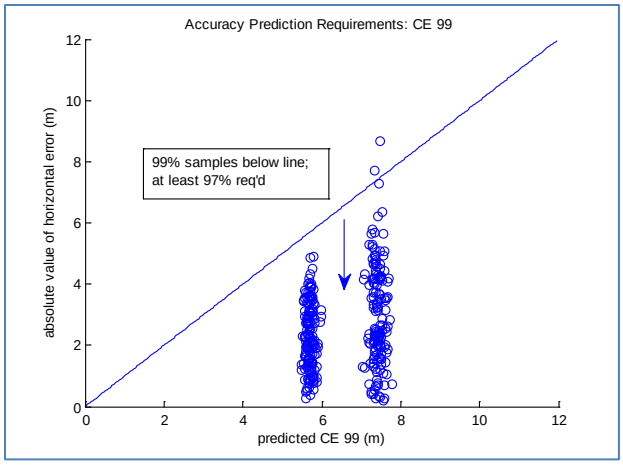

Figure 7: Validation of Accuracy Prediction Req't (5)

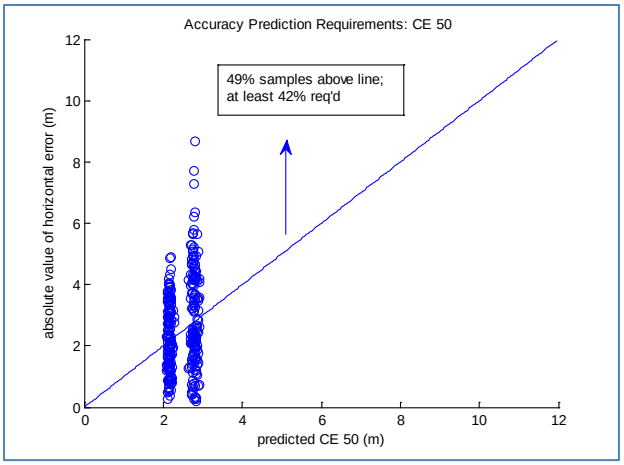

Figure 8: Validation of Accuracy Prediction Req't (7)

Recall that requirements (3) and (11) address (extreme) accuracy and accuracy prediction outliers, respectively. Also, the specification of the primary accuracy prediction requirements (5)-(7) are a practical way to ensure a reasonable relationship between horizontal accuracy and corresponding predicted accuracy. This is illustrated in the following hypothetical plot (Figure 9) that references predicted CE90 for specificity:

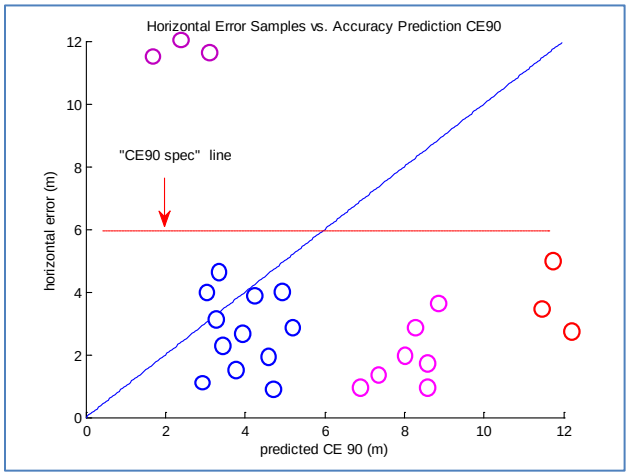

Figure 9: Hypothetical (horizontal error, predicted CE90) pairs plot

The blue circles represent a reasonable distribution of sample pairs that is "covered" (ensured) by the combination of specifications (5), (6), and (7). The other colored circles represent predicted CE90 value extreme outliers (red), horizontal error value extreme outliers (dark red), and systematic overly pessimistic predicted CE90 values (magenta). They are "covered" (virtually prohibited) by specifications (11), (3), and (7), respectively. (Note: the samples in the plot above were limited to a reasonable number since they were "hand drawn"; in reality, and primarily for the blue and magenta samples, each sample (circle) represents an approximate grouping of tens of samples. Also, the same principles apply if either predicted CE50 or CE99 are used instead for the x-axis of the plot.)

SAMPLE SIZE

As stated earlier, 300 samples were selected for the tests in the above examples. It is a reasonable number and comes close to eliminating sample size as an issue. Of course, even more than 300 samples are better, and fewer can

be tolerated if need be. The samples in the examples were independent. They should be independent in any actual real-world validation process as well. For stereo MIG, each sample should correspond to a single ground point in a different stereo pair; for example, 300 samples correspond to 300 different stereo pairs. Few, if any, of the stereo pairs should correspond to the same (orbital) pass. The various ground points should correspond to surveyed locations across various test sites around the world. (For monoscopic MIG relative accuracy validation, not discussed explicitly in this section of the paper, a single set of four distance bin pairs of points should be used per image. Few, if any, of the images should correspond to the same (orbital) pass.)

The effect of sample size was also explored systematically. Recall that there were three tests involving the ratio of sample error to predicted CEXX and predicted LEXX, corresponding to XX=50, 90, and 99, respectively. These tests correspond to the validation of horizontal accuracy prediction requirements (5)-(7) and vertical accuracy prediction requirements (8)-(10). These tests were performed for different sample sizes. Ideally, if sample size was not an issue, all three tests should pass since the (error) samples were generated consistently with the alternating P1 and P2, from which the CEXX and LEXX were computed. For a given test and sample size, the test was performed numerous times over independent groups of samples and the percent of time the test passed tabulated. The sample sizes selected were 10, 50, 100, 200, 300, 400, 600, and 1200 samples. The number of tests corresponding to 10 samples per test was 24000 (total samples = 10*24000=240000), corresponding to 50 samples per test was 4800, …, and corresponding to 1200 samples per test was 200.

Figure 10 (horizontal and vertical) presents % tests passed – we want 100% to pass. 300-400 samples looks good, 600 samples better, and 50 or less poor. Thus, for example, if 50 samples had to be used for validation of requirements, a non-trivial "pad" to requirements/tests would need to be added to compensate for insufficient samples, thus "watering down" the effectiveness of the validation testing.

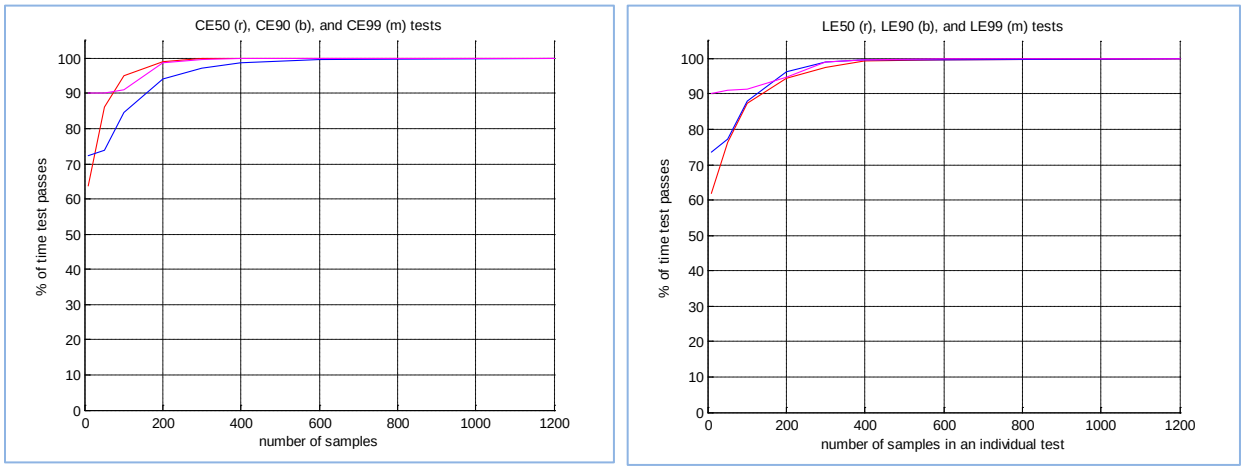

Figure 10: Percentage tests passed versus sample size per test; horizontal error/predicted CEXX (left), vertical error/predicted LEXX (right)

(Note that a specific number of samples adequate for the validation of stereo extraction absolute accuracy prediction requirements should also be adequate for the validation of stereo extraction absolute accuracy, monoscopic relative accuracy, and monoscopic relative accuracy prediction requirements as well.)

ALTERNATE METRICS AND METHODS

Ellipsoidal-Based Metrics

Another approach to the metrics used for specification and validation of accuracy and accuracy prediction are 3D and 2D error ellipsoids computed from the MIG solution's 3x3 error covariance matrix output, as opposed to CEXX and LEXX computed from the same error covariance matrix output. The ellipsoidal-based metrics are preferred when the ellipsoids (error covariance) are elongated. When 3D ellipsoids are used, they also afford the

opportunity to account for any non-trivial correlations between horizontal and vertical errors in the 3x3 error covariance matrix which yield a non-vertical 3D ellipsoid, for which CEXX and LEXX together cannot account for.

There can be 3D error ellipsoids for 3D accuracy and 2D error ellipsoids or ellipses for horizontal accuracy, and corresponding to 50%, 90%, and 99% probabilities. (1D vertical accuracy can also be included as done previously using LEXX.) The corresponding error samples are the absolute value of 3D error and 2D horizontal error, respectively. They are compared to the corresponding predicted "radial" error within the appropriate ellipsoid. The direction is dictated by the (signed) error sample. Figure11 below presents a hypothetical example corresponding to 2D horizontal error and a 2D 90% predicted error ellipse. Specifically, absolute value of 2D horizontal error = 4.3 m (blue + green) and corresponding underlying predicted 90% radial = 3.3 m (green). For this particular case, the actual error sample is larger than the corresponding radial from the 90% predicted error ellipse, so this particular sample fails the "90% level" test.

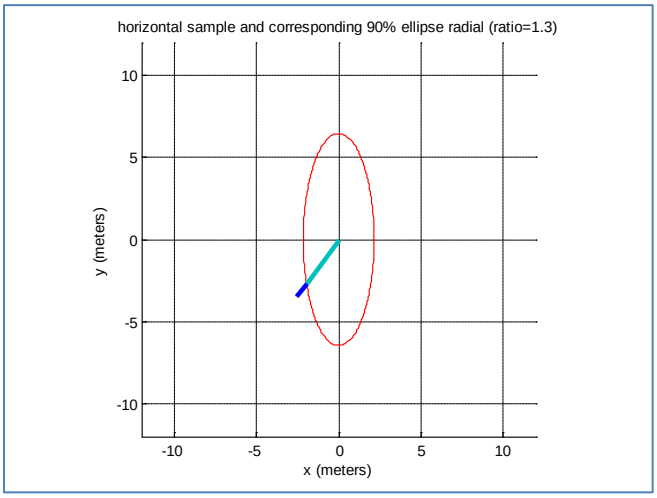

Figure 11: Horizontal error sample and corresponding 90% radial

In general, the formula for an nD error ellipsoid at the pp% probability level and corresponding to an nxn error covariance matrix P is as follows:

$$\varepsilon X^T P^{-1} \varepsilon X = d_{pp}^2, \tag{20}$$

where εX is nD error (meters) and the value of d_{pp} is dependent on the desired probability level: n=3, $d_{50} = 1.538$, $d_{90} = 2.500$, $d_{99} = 3.368$; n=2 (horizontal error, and upper left 2x2 of P), $d_{50} = 1.177$, $d_{90} = 2.146$, $d_{99} = 3.035$; n=1 (vertical error, and lower right 1x1 of P), $d_{50} = 0.674$, $d_{90} = 1.645$, and $d_{99} = 2.576$.

The sample-based tests corresponding to ellipsoids and probability level XX consist of testing if the ratio of absolute error to corresponding predicted ellipsoidal radial (predicted ellipsoid at the XX probability level) is less than or equal to 1.0 for the 99% and 90% level tests, and greater than or equal to 1.0 for the 50% level test.

For a given sample k, we designate the absolute error as $|\varepsilon X_k|$ and its corresponding radial as $\varepsilon X radial_k$, where

$$\varepsilon X radial_k = d_{XX} |\varepsilon X_k| (\varepsilon X_k^T (P\_pred_k)^{-1} \varepsilon X_k)^{-1/2}. \tag{21}$$

For a given sample k, we can also define the predicted normalized error as:

$$\varepsilon X norm_k \equiv (\varepsilon X_k^T (P\_pred_k)^{-1} \varepsilon X_k)^{1/2} / d_{XX}. \tag{22}$$

For a specific sample k, the corresponding XX accuracy prediction tests also correspond to whether $\varepsilon X norm_k \leq 1.0$. Therefore, for example, assuming XX=90 and 2 dof, the corresponding accuracy prediction specification can

be expressed as $prob\{\varepsilon Xnorm \leq 1\} \geq 0.86$. This specifcation is "equivalent" to the CE-based requirement (6). For actual validation purposes, however, the usual samples versus corresponding predicted accuracy plots are performed, where samples correspond to the usual absolute error but predicted accuracy corresponds to the radial.

In addition, with the use of ellipsoids instead of CE and LE, multiple requirement plots may be combined for convenience into one plot, since 50%, 90%, and 99% ellipsoids have the same shape. Therefore, for example, a 2D predicted radial at the 50% probability level is a fixed factor times the corresponding 2D predicted radial at the 90% probability level. (This is not true for CE50 relative to CE90.) In particular, only a "2D 90% error ellipsoid samples" plot needs to be performed, but with three lines: (1) the usual "slope 1 line" for which 2D sample absolute errors and the corresponding predicted 90% radials correspond to, (2) a smaller sloped line (d_{50}/d_{90}) for which 2D sample absolute errors and the corresponding 50% predicted radials correspond to, and (3) a large sloped line (d_{99}/d_{90}) for which 2D sample absolute errors and the corresponding 99% predicted radials correspond to. The following example (Figure 12) is based on the same simulation discussed earlier; all tests pass. The same basic form and probability levels for the requirements presented earlier using CEXX and LEXX also apply. Note the larger "spread" of samples across the plots as compared to the earlier plots. This is due to the variable direction of the (signed) error samples.

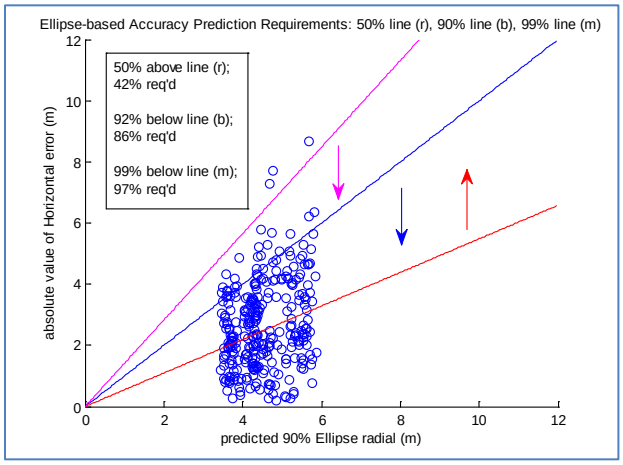

Figure 12: Ellipse-based Accuracy Prediction tests

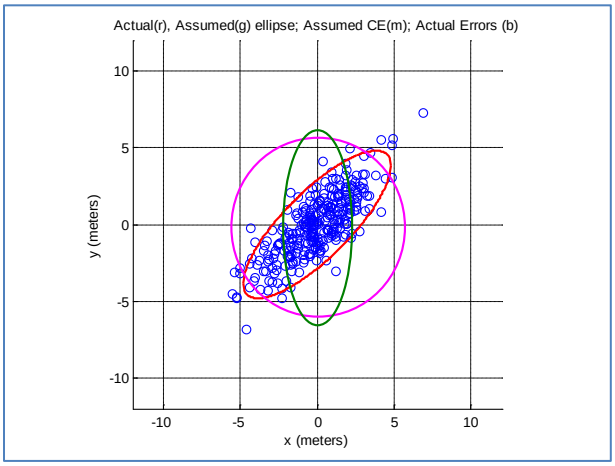

Figure 13: Example of Ellipse preferred

Finally, Figure 13 presents an example of why ellipse-based accuracy tests are preferred when the ellipse (or ellipsoid) is elongated. In this example, the red 90% ellipse is the correct ellipse corresponding to the correct MIG solution error covariance matrix. It is "known" or "truth" for this simulation-based experiment but unknown to the validation process. Instead, the uncertainty model associated with the sensor and actually used by the MIG as part of the validation process is incorrect and consequently the green 90% ellipse corresponds to the error covariance matrix actually calculated by the MIG – it is incorrect. If accuracy prediction tests are ellipse-based, they will fail using the green ellipse since a significant number of error samples are outside of this ellipse, which is what we want since it and its underlying error covariance matrix are incorrect. On the other hand, if the accuracy prediction tests use the baseline CE tests instead that are tailored to commercial satellite EO imagery with non-elongated ellipses, they will incorrectly pass since the vast majority of error samples are within the CE circle. Consequently, the calculated MIG error covariance matrix will be incorrectly assumed "correct" by the validation process.

Actual versus Assumed Error Covariance and the Chi-Square Distribution

The baseline accuracy prediction requirements can be met if the assumed error covariance matrix (from which the predicted accuracies are generated) is equal to (.925)^2 to (1.075)^2 times the actual error covariance matrix (from which the error samples are generated). That is, when the assumed one-sigma values are between +/- 7.5% of the actual (but generally unknown) values. In fact, the actual probability levels (theory + " pad") or thresholds used in requirements (5-7) and (8-10) (or their ellipsoidal counterparts) assumed this range of differences between assumed and actual in conjunction with a Gaussian distribution of errors. However, the "pad" is actually to be

interpreted to include the combined effects of assumed vs. actual error covariance matrix, errors that are only approximately Gaussian distributed, and some residual sampling effects.

Note that although this range may seem limited, it should still be reasonable in that samples for validation will actually be taken over many images taken over many months with the "averaging" of somewhat optimistic and pessimistic error propagation (MIG error covariance generated from sensor support data error covariance and temporal correlation model). Furthermore, the commercial vendor will have ample time to "tune" their accuracy prediction or uncertainty model.

The probability levels used in the requirements corresponding to an assumed error covariance matrix with one-sigma's +/- 7.5% the actual one-sigma's were derived based on ellipsoidal metrics as follows.

The (sample) metric $\varepsilon X^T P^{-1} \varepsilon X$ can be considered a random variable with a chi-square probability distribution with n degrees of freedom (dof), where n=3 if 3D errors, n=2 if 2D errors, and n=1 if 1D errors. This assumes that εX is a random variable with a multi-variate (joint), mean zero, Gaussian probability distribution (not a specific realization or sample). See reference [7] for further details regarding chi-square.

Figures 14 plots the chi-square cumulative probability distribution vs. normalized distance d for 3 dof (left) and 2 dof (right), and where $\varepsilon X^T P^{-1} \varepsilon X = d^2$. Multiple plots correspond to the "difference" between actual and assumed covariance (one-sigma's). (The three additional horizontal lines in the left and right plots correspond to 0.99, 0.9, and 0.5 cumulative probability levels; the three additional vertical lines correspond to their threshold distance d counterparts. Note that these d values are identical to those used in the definition of the corresponding error ellipsoids.)

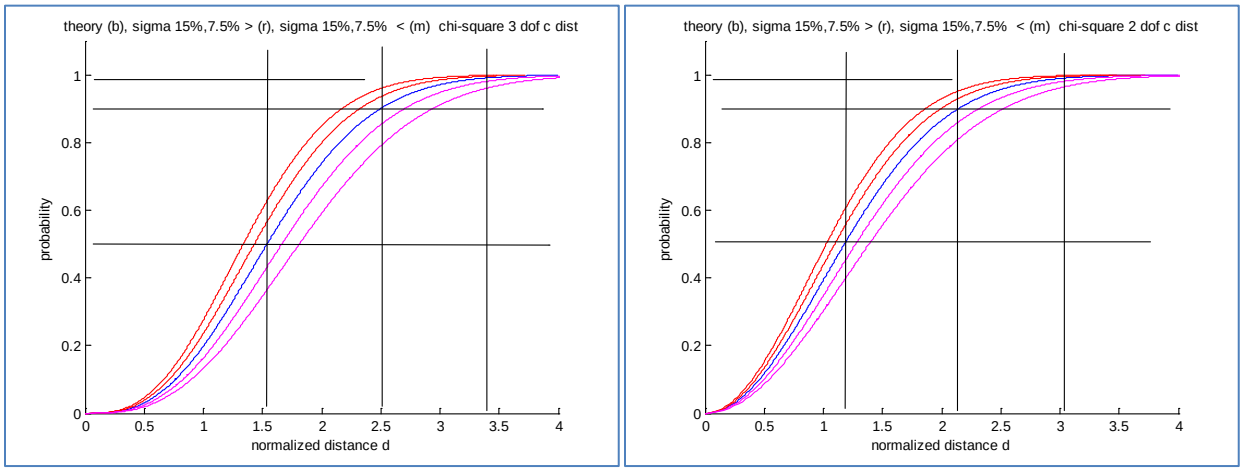

Figure 14: Chi-square cumulative probability distributions – 3 dof (left), 2 dof (right)

Thus, for example, based on the above plots, if we assume +7.5%/-7.5% sigma and a little "extra pad" for sampling and only approximate Gaussian errors (slightly more pad is selected for the 50% spec since this specification favors conservative error prop, i.e., the test is based on normalized error > 1, not < 1), we have the following specified horizontal error accuracy prediction requirements:

$$prob\{\varepsilon Xnorm_{99} \leq 1\} \geq 0.97 \qquad (23)$$
$$prob\{\varepsilon Xnorm_{90} \leq 1\} \geq 0.86$$
$$prob\{\varepsilon Xnorm_{50} > 1\} \geq 0.42,$$

The value 1.0 minus the value of the upper (red) 7.5% lines on the plots were used to compute the probability values corresponding to the XX=50 probability requirement, whereas the value of the lower 7.5% (magenta) lines on the plots were used to compute the probability values for the other requirements. Also, with regard to CE-based metrics, $prob\{\varepsilon Xnorm_{90} \leq 1\}$ and 2 dof is "equivalent" to prob{(dH/CE90_pred)<=1}. (Note: For a non-

commercial satellite imaging sensor with less controlled imaging geometry, etc., the 15% or similar lines may be more applicable than the 7.5% lines with a corresponding decrease in the values on the right side of Equation (23), i.e., a "loosening" of the specification for accuracy prediction capabilities.)

As an aside, a natural question arises. Instead of specification of accuracy prediction capabilities based on the three separate but specific tests above (per dof), why not simply specify that the "sample" chi-square cumulative probability distribution be within a tolerance of its theoretical value? Two reasons why the baseline approach is preferred instead: (1) the tolerance would need to be a function of the normalized distance d (the x-axis in the above plots) in order to be effective since we don't expect the assumed and actual covariance to be the same, and (2) trends cannot be observed that are a function of error covariance (CE/LE or ellipsoid radial) magnitude.

SUMMARY

Recommended methods for the specification and subsequent validation of accuracy and accuracy prediction capability requirements were presented for commercial satellite EO imagery. Accuracy prediction capabilities are necessary for many applications but methods for their specification and validation have not been presented previously per the authors' knowledge. The paper also presented a simulation-based example of validation of these requirements for insight into both the requirements and their validation process. The paper then extended the specification and validation from CE and LE-based metrics to ellipsoidal-based metrics for imaging sensors that yield MIG-generated elongated error covariance matrices or ellipsoids. These imaging sensors are typically not as controlled with favorable imaging geometry as are, for example, the WorldView and GeoEye commercial satellite imaging sensors. Finally, the paper related the various accuracy prediction specification and validation test thresholds to the chi-square distribution.

APPENDIX A: Evaluation of CEXX and LEXX

The following equation (A-1) defines CEXX and LEXX, respectively. A mean-zero distribution of multi-variate Gaussian errors is assumed and relative to a local tangent plane coordinate system. P is the 3x3 error covariance of 3D position, P_H the upper left 2x2 of P, and $P_z = \sigma_z^2$ the lower right 1x1 of P. Given a desired probability level XX expressed in integer percent, i.e., from 1 to 99, the equation(s) can be solved for numerically for the corresponding CEXX and LEXX in meters. There are also numerous approximations and tables for their solutions. For example, reference [1] presents a high fidelity approximation of CE90 based on the eigenvalues of P_H and LE90 based on σ_z.

$$0.XX = (2\pi\det(P_H))^{-0.5} \iint_{\sqrt{x^2+y^2}\le CEXX} e^{-0.5[x \quad y]P_H^{-1}\begin{bmatrix}x\\y\end{bmatrix}} dxdy;\ \ 0.XX = (2\pi\sigma_z^2)^{-0.5} \int_{\sqrt{z^2}\le LEXX} e^{-0.5z^2/\sigma_z^2} dz \quad \text{(A-1)}$$

REFERENCES

[1] Dolloff, J. T., 2012, "RPC Uncertainty Parameters: Generation, Application, and Effects", Proc.of ASPRS Annual Convention, Sacramento, California, March 19-23, 2012.

[2] Doucette, Peter, et al, 2013, "Measurement and Automation Practices in Photogrammetry", Chapter 11 in Manual of Photogrammetry, Book, Sixth Edition, Chris McGlone, editor, ASPRS.

[3] Dolloff, John, 2013, "Geopositioning", Chapter 10.3.4.1.1 in Manual of Photogrammetry, Book, Sixth Edition, Chris McGlone, editor, ASPRS.

[4] Dolloff, John and Hank Theiss, 2012, "Temporal Correlation of Metadata Errors for Commercial Satellite Images: Representation and Effects on Stereo Extraction Accuracy", International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXIX-B1.

[5] Dolloff, J., Settergren, R., 2010. "An Assessment of WorldView-1 Positional Accuracy based on 50 Contiguous Pairs of Stereo Imagery", PE&RS Journal, 76(8), pp. 935-943.

[6] Dolloff, John, 2013, "The full multi-state vector error covariance matrix: Why needed and its practical representation", Proc of SPIE, Vol 8747.

[7] Bar-Shalom, Yaakov and Thomas Fortmann, 1988, Tracking and Data Association, Academic Press.