Statistical Analysis of Military and Civilian Navigation Error Data Services

Ted Driver, *Analytical Graphics Incorporated*

BIOGRAPHY

Ted Driver is the Senior Navigation Engineer at Analytical Graphics Inc. Ted has worked on the Navigation Tool Kit for three years, having previously been the technical lead for the Navigation Tool Kit at Overlook Systems Technologies. He has led the engineering team in developing the navigation algorithms and data stream definitions and is currently working on statistical prediction models for GPS accuracy. He was previously the senior GPS Operations Center analyst within the 2nd Space Operations Squadron at Schriever Air Force Base. He has worked in the GPS field for 9 years, previously designing the environment and navigation models for the GPS High Fidelity Simulator currently in use at Schriever Air Force Base. Mr. Driver received his Bachelors Degree in Physics from the University of California at San Diego and his Masters Degree in Physics from the University of Colorado. He is a former President of the Rocky Mountain Section of the Institute of Navigation, currently holding the position of Vice-chair, and has recently helped the 2nd Space Operations Squadron re-establish the GPS Performance Analysis Working Group; a forum where Air Force personnel and GPS professionals discuss GPS performance status and issues.

ABSTRACT

Objective:
Determine a statistical difference in navigation accuracy, if any exists, between data provided by a Military Error Data Service and Civilian Error Data Service.

Methodology:
With the advent of both civilian and military real-time GPS satellite ephemeris and clock correction streams into the GPS user community, the question naturally arises: Which is better? The Jet Propulsion Laboratory and Navcom Technologies both have civilian data streams available to the public and the military is in the process of standing up the GPS Information Service (GPSIS) emanating from the GPS Operations Center (GPSOC). While civilian users only have civilian sources from which to choose, military users will have the capability to choose between the civilian data service and the military data service. To aid in making that choice, this paper will define a decision making criteria for choosing one service over another. The analysis will search for statistical significance in the differences in computed accuracy derived from the two services. The accuracy will be calculated as Signal-In-Space (SIS) using the Navigation Tool Kit as a Navigation Error modeling and analysis tool.

Using a Paired Data Analysis (PDA) technique, differences in accuracy for a specific location will be analyzed. An initial check is made of the normality of the differences between the accuracies derived from the two data sets and then hypothesis testing is used to test the value of the differences. The null hypothesis is that there is no difference between the accuracies derived from the Military Data Stream and the Civilian Data Stream at a specified threshold level. The alternative hypothesis is that the accuracy derived from the Military Data Stream is better than the accuracy derived from the Civilian Data Stream for the given threshold.

Anticipated Results:
I intend to show that either the null hypothesis is supported or evidence is strong to reject it, using a 95% confidence threshold.

Significance:
This comparison technique is not unique to military and civilian error correction data; it can be applied to other error data sources as well. The outcome of this particular analysis however will shed light on problems inherent in error data processing and their statistical analysis. Lastly, the results of the analysis made lead to a better informed decision by the military, as to which data service to choose, when the choice must be made.

INTRODUCTION

Navigation error data is becoming more prevalent lately – especially in the civilian domain. World-wide networks of monitoring stations send GPS measurements to processing facilities; those facilities determine ephemeris and clock errors to a high degree of accuracy in near real time and are made available to the public. A very important newcomer to this field is the military. The Air Force is in the process of creating a service to deliver navigation error data to military users – an unparalleled accomplishment – that will benefit military navigation planners across the globe. The military's data is currently

derived from 5 monitoring stations, but in the near future, will include 6 NGA stations as well.

This paper investigates the following questions – with both civilian and military navigation error data available, is there a statistically significant difference between them? Is one error data stream more accurate than another for navigation planning purposes? As a military planner, with a choice of available data, is there any reason to choose a civilian error data set over the military error data set?

This paper will outline a series of statistical tests for determining a quantitative answer to these questions. First, I'll provide a discussion of the datasets used for analysis and the data used for a truth comparison. Then, I'll outline the statistical tests performed on the data and the results of those tests. Finally, I'll interpret these results using different comparison criteria; to allow the user of these two data streams to know which stream is better to use given their specific mission criteria.

CIVILIAN NAVIGATION ERROR DATA

The data used to represent the civilian navigation errors are provided by Navcom Technologies Inc.; a business partner of Analytical Graphics Inc. (AGI). Navcom provides an error data delivery service to customers of AGI for use in the Navigation Tool Kit (NTK) – a tool to aid the mission planner in determining navigation accuracy under a variety of circumstances and environments. The provided data consists of calculated GPS satellite ephemeris and clock errors in two different types of delivery modes. In one mode, the data is posted on an FTP site daily in an XML format and contains ephemeris and clock errors for each healthy GPS satellite at 15 minute intervals. This is known as the 'archive' set of files. The other mode of delivery is via the real time service Navcom provides to AGI customers. This data consists of ephemeris and clock errors for healthy GPS satellites and GPS satellite subframe 1 health indications. Both are delivered at one minute intervals. The archive data is identical to the real time data, only the archive data is collected once every 15 minutes and placed in the archive file for posting on the FTP site once per day. For both modes of delivery, the data is calculated using a Kalman filtering process that takes in GPS measurements from 57 monitoring stations. The monitoring stations track L1 carrier and C/A code as well as L2 carrier. A proprietary method is used to obtain dual frequency error corrected measurements at these sites.

Here's an example of the data in an archive file:

```
<?xml version="1.0"?>
<GPSISFILE  FILEID="PAF"  SYSID="GPS"  SOURCE="STARFIRE"
VERSION="2">
<CREATION YEAR="2005" DOY="025" HR="23" MIN="45" SEC="00" />
<REFERENCE YEAR="2005" DOY="025" HR="23" MIN="45" SEC="00" />
<PAF_RECORD>
  <VALID YEAR="2005" DOY="25" HR="0" MIN="0" SEC="0" />
     <PAF_FIELD
        SVID="1"
        DATA_AVAILABLE="YES"
        POS_ERROR_X="-0.953"
        POS_ERROR_Y="2.492"
        POS_ERROR_Z="1.094"
        CLOCK_PHASE_ERROR="0.945"
        VEL_ERROR_X="0.0001220703"
        VEL_ERROR_Y="0.0004882812"
        VEL_ERROR_Z="-0.0002441406"
        CLOCK_FREQUENCY_ERROR="0.0000261937"
        AGE_OF_DATA="0.000"
  />
  <PAF_FIELD
        SVID="2"
        DATA_AVAILABLE="YES"
        POS_ERROR_X="0.719"
        POS_ERROR_Y="0.445"
        POS_ERROR_Z="-0.078"
        CLOCK_PHASE_ERROR="-0.486"
        VEL_ERROR_X="0.0000000000"
        VEL_ERROR_Y="0.0001220703"
        VEL_ERROR_Z="-0.0001220703"
        CLOCK_FREQUENCY_ERROR="0.0000065815"
        AGE_OF_DATA="0.000"
  />
```

Note that the source of the data is the STARFIRE network and that the position and velocity coordinates are in the Earth centered, Earth fixed coordinate system. The acronym 'PAF' means Performance Analysis File (as opposed to a Prediction Support File (PSF) file used to support predictions of GPS accuracy) and each PAF_RECORD has its own time of validity. A separate PAF_FIELD is defined for each SVID (equivalent to PRN) with position and velocity errors and clock phase and frequency errors for that SVID at that time. Note that all units are in meters, the time component being scaled by the speed of light to achieve the proper unit. The AGE_OF_DATA field in the STARFIRE produced files will always be 0, since the age of the navigation upload on a given satellite cannot be determined using the navigation data stream broadcast by the satellites. In the analysis described in this paper, the ephemeris values will be converted to the Radial, Along-Track, Cross-Track (RAC) coordinate system. The Civilian errors then will be referred to using the following notation:

Civilian Radial Error:	$dR^C_{sv_i}(t_j)$
Civilian Along-Track Error:	$dAT^C_{sv_i}(t_j)$
Civilian Cross-Track Error:	$dCT^C_{sv_i}(t_j)$
Civilian Clock Error:	$dC^C_{sv_i}(t_j)$

MILITARY NAVIGATION ERROR DATA

The data used to represent the military navigation errors are created at the 2SOPs Master Control Station (MCS). The military data consists of the same constituents as the civilian data – ephemeris and clock errors for each healthy GPS satellite for a given time. The difference here is that the military data is derived from measurements made at five monitoring stations – using true dual frequency tracking of carrier and code. The measurements are made using P(Y) code instead of C/A code and are passed to the MCS Kalman filtering process. For those familiar with the MCS processing, the

ephemeris and clock errors in the military data are the estimated range deviations (ERDs) produced by the MCS Kalman filter.

The military data is in the same PAF file format as the civilian data, as an example shows:

```
<? xml version="1.0" standalone="no"?>
<GPSISFILE    FILEID="PAF"    SYSID="GPS"    SOURCE="GOCGIS"
VERSION="2">
<CREATION YEAR="2005" DOY="052" HR="10" MIN="37" SEC="07"/>
<REFERENCE YEAR="2005" DOY="039" HR="00" MIN="00" SEC="00"/>
<PAF_RECORD>
        <VALID    YEAR="2005"   DOY="039"   HR="00"   MIN="00"
SEC="00"/>
        <PAF_FIELD
              SVID="1"
              DATA_AVAILABLE="YES"
              POS_ERROR_X="0.134764909744263"
              POS_ERROR_Y="-1.233568191528320"
              POS_ERROR_Z="-0.173474848270416"
              CLOCK_PHASE_ERROR="1.776714324951172"
              VEL_ERROR_X="-0.000000448327009"
              VEL_ERROR_Y="0.000000182344774"
              VEL_ERROR_Z="-0.000001318738688"
              CLOCK_FREQUENCY_ERROR="-0.0000243409178"
              AGE_OF_DATA="735"/>
        <PAF_FIELD
              SVID="2"
              DATA_AVAILABLE="YES"
              POS_ERROR_X="0.447188258171082"
              POS_ERROR_Y="0.355884671211243"
              POS_ERROR_Z="0.249038577079773"
              CLOCK_PHASE_ERROR="-0.255644679069519"
              VEL_ERROR_X="-0.000000119867917"
              VEL_ERROR_Y="-0.000000972098082"
              VEL_ERROR_Z="0.000000720030691"
              CLOCK_FREQUENCY_ERROR="0.00000293469481"
              AGE_OF_DATA="1290"/>
```

Note that the source of the data is the GOSGIS (GPS Operations Center GPS Information Service) and that there are additional digits in the military data. The AGE_OF_DATA field in the GOCGIS produced files are filled in, since the MCS has access to this information. In the analysis described in this paper, the ephemeris values will also be converted to the RAC coordinate system. The military errors will be referred to using the following notation:

Military Radial Error: $dR_{SV_i}^M(t_j)$

Military Along-Track Error: $dAT_{SV_i}^M(t_j)$

Military Cross-Track Error: $dCT_{SV_i}^M(t_j)$

Military Clock Error: $dC_{SV_i}^M(t_j)$

TRUTH DATA
To compare the Military and Civilian errors, a truth source is needed. Once found, both datasets can be differenced from the truth to obtain residuals. It is upon these residuals from truth that the statistical analysis will be performed.

Since both the civilian and military error data is composed of ephemeris and clock errors from the broadcast ephemeris and clock states, a good truth source would consist of the actual ephemeris and clock states differenced from the broadcast states. GPS satellite ephemeris and clock states are published by the National Geospatial-Intelligence Agency (NGA) in the form of 3-day post fit SP3 files. These files are available on their website [1]. The SP3 file format contains each satellite's ephemeris in Earth-centered, Earth-fixed coordinates as well as each satellite's clock state. Each of these states are provided at 15 minute intervals, for a 24 hour period, in a single file.

The broadcast ephemeris from each satellite is available in RINEX2 format from a variety of sites on the World Wide Web. For this analysis, I retrieved the *Global Navigation* file for each day of the analysis from the Continuously Operating Reference Station Network (CORS) [2] operated by the National Oceanic and Atmospheric Organization (NOAA). This file contains information in RINEX2 format specifically for subframes 1, 2 and 3 of the GPS Navigation Data, broadcast by each GPS satellite. This data consists of precise, broadcast ephemeris and clock parameters for each satellite for a given time span. Algorithms defined in IS-GPS-200D [3] detail the processing necessary to obtain the GPS satellite's position and clock state at a given time.

Once the position is known from both the NGA files and the RINEX2 files, the differencing can be done. The truth ephemeris and clock error states are then constructed thusly:

$$\Delta \vec{R}_{SV_i}^T(t_j) = \vec{R}_{SV_i}^{NGA}(t_j) - \vec{R}_{SV_i}^{Brdc}(t_j)$$

$$\Delta C_{SV_i}^T(t_j) = C_{SV_i}^{NGA}(t_j) - C_{SV_i}^{Brdc}(t_j)$$

The suffix *NGA* denotes the data from the NGA precise ephemeris file. The suffix *Brdc* denotes the data calculated by the algorithms [4] in IS-GPS-200D. The true ephemeris errors will also be converted to the RAC coordinate system and be referenced as:

True Radial Error: $dR_{SV_i}^T(t_j)$

True Along-Track Error: $dAT_{SV_i}^T(t_j)$

True Cross-Track Error: $dCT_{SV_i}^T(t_j)$

True Clock Error: $dC_{SV_i}^T(t_j)$

STATISTICAL TESTS
Now that the test datasets and the truth dataset are defined, we can proceed with the statistical tests on the data. First, however, some background is required on the types of statistical tests available for describing and classifying data.

Statistical Background
There are a variety of statistical tests that can be performed on a dataset – all with specific reasons for their use. In our case, we'd like to know whether one dataset is better than another – we'd like to know which dataset is closer to truth. The best statistical analysis for this type of situation is a hypothesis test [5]. Generally, in hypothesis testing one creates a null hypothesis H_0, which describes a

presumed given situation, and then and alternate hypothesis H_a, (also denoted as H_1) that describes the conjecture to test the dataset against. Hypothesis testing then dictates that one of several types of numerical algorithms be performed on the data with the numerical results of that test compared against standard table values to decide whether the null hypothesis should be rejected in lieu of the alternate hypothesis. Some types of numerical tests that can be performed during a hypothesis experiment include Z-testing, T-testing and Paired Data Analysis techniques. Z-testing can be performed when the data consists of a large number of samples; the central limit theorem suggests a minimum of 30 samples for this type of test to apply. The Z-test measures the Z-score of a statistical variable, identical to the number of standard deviations (in a normal distribution) from the mean value of that variable. The Z-test measures how well the differences in two datasets match the hypothesis criteria, assuming each dataset is independent of the other. A T-test is typically applied when then number of samples is small (< 30). In this case a different statistical distribution applies and the T-score is compared to a table of values derived from the T-distribution. The T-test also assumes that the datasets are independent of one another. The Paired Data Analysis technique is typically used when two samples of data are taken from the same source, but are treated differently. The PDA technique does not assume that the data samples are independent and it uses a modification of the T-test to derive a T-score. In our case, both military and civilian errors are derived from measuring the same GPS pseudoranges, thus the PDA technique is the best choice for this analysis.

The T-test used in the PDA technique derives a T-score using the following:

$$T_{score} = \frac{\overline{D} - \Delta_0}{S_D / \sqrt{N}}$$

With

$$\overline{D} = \overline{X} - \overline{Y}$$

$$\overline{X} = \text{Dataset 1 Mean}$$

$$\overline{Y} = \text{Dataset 2 Mean}$$

$$\Delta_0 = \text{Expected Mean Difference}$$

And

$$S_D = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(D_i - \overline{D})^2}$$

$N = $ Number of statistical observations in X and Y

Once the T-score is defined, it must be compared to standard table values at a prescribed confidence level to determine if the hypothesis should be accepted or rejected at that confidence level. The table values are chosen based on the number of degrees of freedom as well.

Because I am using 113 statistical observations, I should choose a table value based on 112 degrees of freedom, the number of degrees of freedom for the test is one less than the number of observations.

Another consideration for the tests is to determine if a two-tailed or single tailed test is needed. The radial and clock hypotheses will use a two-tailed test because their values can be positive or negative while the accuracy hypothesis will use a single-tailed test since navigation position error can only be positive.

Hypotheses
In measuring GPS accuracy, the radial ephemeris and clock measurements are arguably the most important metrics to study; therefore I'll concentrate my initial analysis on those two components only.

For the radial and clock tests, I'm stating the following null and alternate hypotheses:

$$H_0 : \overline{\Delta R}_{SV_i}^{M} - \overline{\Delta R}_{SV_i}^{C} = \Delta_0$$

$$H_a : \overline{\Delta R}_{SV_i}^{M} - \overline{\Delta R}_{SV_i}^{C} \neq \Delta_0$$

Where

$$\overline{\Delta R}_{SV_i}^{M} = \frac{1}{N}\sum_{j=1}^{N}\left[dR_{SV_i}^{M}(t_j) - dR_{SV_i}^{T}(t_j)\right]$$

$$\overline{\Delta R}_{SV_i}^{C} = \frac{1}{N}\sum_{j=1}^{N}\left[dR_{SV_i}^{C}(t_j) - dR_{SV_i}^{T}(t_j)\right]$$

The average is taken over a single day, with N being determined by the number of available samples for a given SV. N may change as a function of SV and time due to data losses or planned outages. The H_0 hypothesis states that there is no statistically significant difference between the military and civilian errors at level Δ_0.

The hypotheses are similar for the Clock Errors:

$$H_0 : \overline{\Delta C}_{SV_i}^{M} - \overline{\Delta C}_{SV_i}^{C} = \Delta_0$$

$$H_a : \overline{\Delta C}_{SV_i}^{M} - \overline{\Delta C}_{SV_i}^{C} \neq \Delta_0$$

Where

$$\overline{\Delta C}_{SV_i}^{M} = \frac{1}{N}\sum_{j=1}^{N}\left[dC_{SV_i}^{M}(t_j) - dC_{SV_i}^{T}(t_j)\right]$$

$$\overline{\Delta C}_{SV_i}^{C} = \frac{1}{N}\sum_{j=1}^{N}\left[dC_{SV_i}^{C}(t_j) - dC_{SV_i}^{T}(t_j)\right]$$

The dataset I'm using to analyze these hypotheses consists of 113 days of observations – each day consisting of a mean radial or clock error over 96 K-Points (15 minute segments) [6].

The more important quantity to analyze however is user accuracy. While radial and clock errors are important to the analyst, it is end-user accuracy that counts to a military planner. Therefore, I'm including the following hypotheses as well:

$$H_0 : \overline{Acc_{Diff}} = \Delta_0$$

$$H_a : \overline{Acc_{Diff}} < \Delta_0$$

Where

$$\overline{Acc_{Diff}} = \frac{1}{113} \sum_{i=1}^{113} \left[\overline{Pacc_i^M} - \overline{Pacc_i^C} \right]$$

And

$$\overline{Pacc_i^M} = \frac{1}{96} \sum_{j=1}^{96} \left[Pacc_i^M(t_j) \right]$$

$$\overline{Pacc_i^C} = \frac{1}{96} \sum_{j=1}^{96} \left[Pacc_i^C(t_j) \right]$$

and $Pacc_i^M(t_j)$ and $Pacc_i^C(t_j)$ are the position accuracy for the ith day at the jth K-point for the Military and Civilian datasets, respectively.

The Δ_0 level in the accuracy hypothesis will quantify the importance of Military Operational Impacts – the military planner can choose a threshold Δ_0 that corresponds to the particular mission plan and make a decision on which data service to use based on the results of the tests outlined in this paper.

RADIAL AND CLOCK STATISTICAL RESULTS

The radial and clock errors were tested against a Δ_0 value of 0. This choice will show how close the military and civilian errors are to the truth source, including any inherent biases. Figures 1 and 2 highlight the radial errors in two cases; Figure 1 shows the mean radial differences from truth for a 113 day span for PRN 22. Figure 2 shows the same but for PRN 10 instead. It's apparent that the civilian data has a bias in it's determination of the PRN 22 radial error component, whereas PRN 10 does not. Also note that the military differences vary more than the civilian differences, despite the bias. This indicates that the civilian data processing models the actual errors better than the military processing, once the biases are taken into account. Figures 3 and 4 show the standard deviations of the radial errors for both PRNs, highlighting the greater variability of the military service over the civilian service.

The statistical test results on the radial errors are shown in Table 1. At a threshold of 0 meters, only 5 satellites show no differences between military and civilian data services. This is mostly due to the fact that several satellites showed biases in the civilian data stream. None of the military radial errors showed biases from truth.

Table 1 - Initial Radial Statistical Results

SV PRNs	$\Delta_0 = 0$ Level	Count
1,4,7,27,30	Pass	5
2,3,5,6,8-11,13-16,18-26,28, 29,	Fail	23
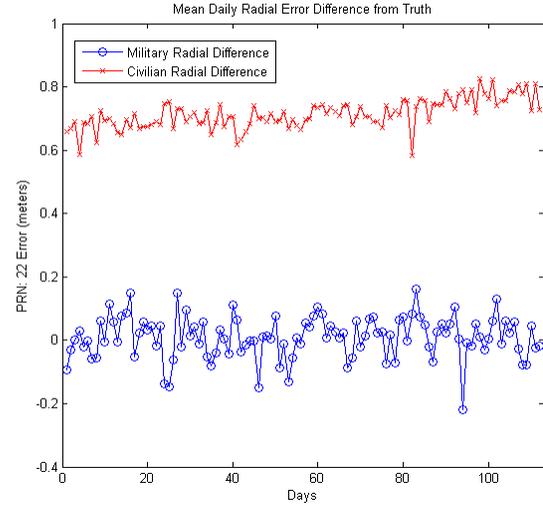
Figure 1 - PRN 22 Mean Radial Errors
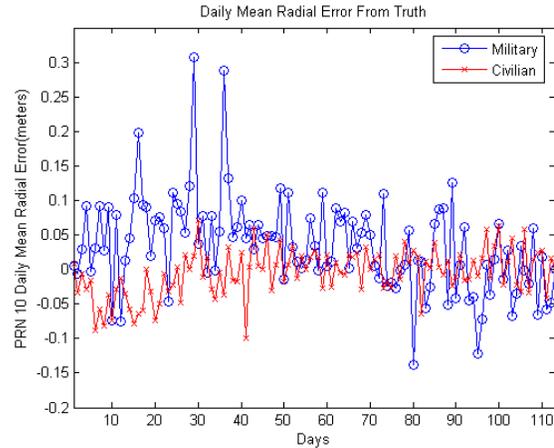
Figure 2 - PRN 10 Mean Radial Errors
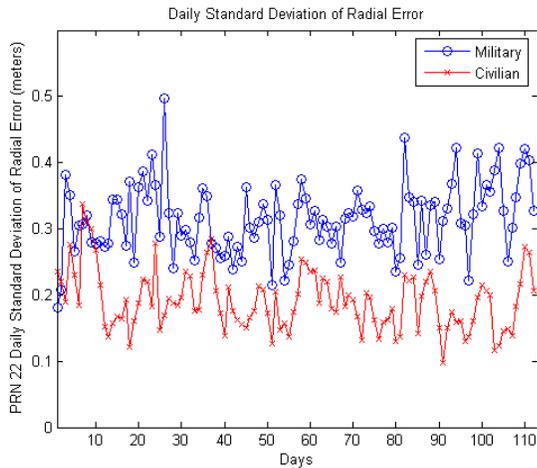
Figure 3 - PRN 22 Daily Radial Standard Deviation
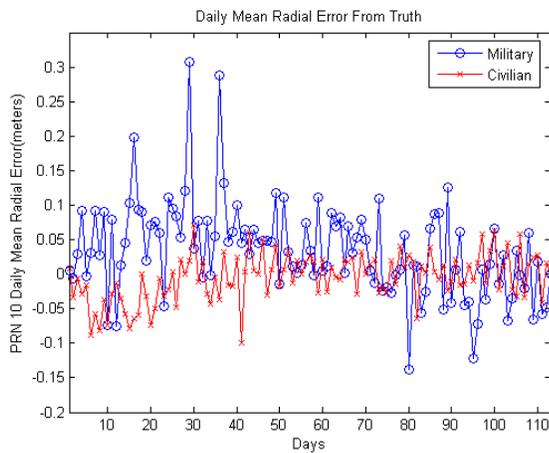
Figure 4 - PRN 10 Daily Radial Standard Deviation

The clock errors are shown for PRN 22 in Figure 5. Note that the military clock errors from truth show no apparent long term bias, while the civilian errors do appear to have a long term bias. The standard deviations of the clock errors are shown in Figure 6. The variances of the clock errors are almost identical between the two data services.

Table 2 shows the statistical test results for the clock errors. Here, at a threshold of 0 meters, all satellites show that the military clock errors are better than the civilian clock errors.

Table 2 - Initial Clock Statistical Results

SV PRNs	$\Delta_0 = 0$ Level	Count
	Pass	0
1-11,13-16,18-30	Fail	28
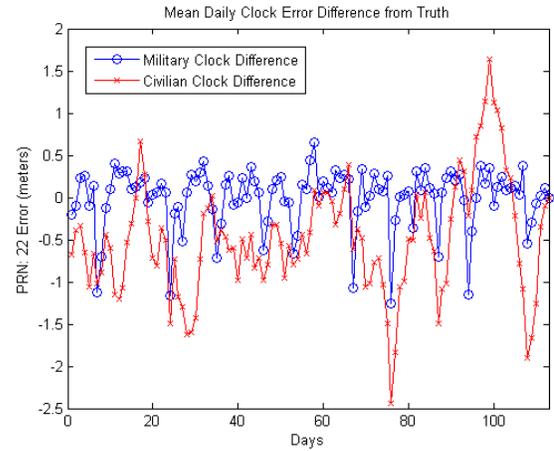
Figure 5 - PRN 22 Mean Clock Errors
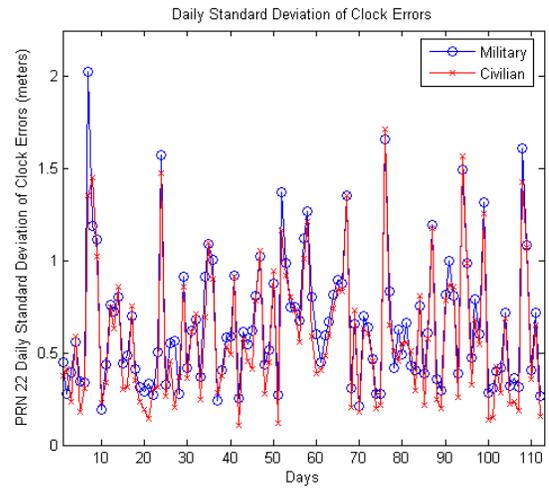
Figure 6 - Clock Daily Standard Deviations

Improvements in Civilian Data Service
After discovering these biases in the civilian data streams, efforts were undertaken to modify the civilian data stream to better match the truth data. After defining a solution, preliminary results were generated to show how much better the civilian data service had become. Figure 7 shows the T-score for each SV, when compared to the Military errors at a threshold level of 0 meters. This graph indicates an improvement in the modeling of the radial component. Bars in the graph falling within the blue lines (95% confidence) are considered a pass, those outside the blue lines are considered a fail. Summarized results for the modified radial errors are included in Table 3.
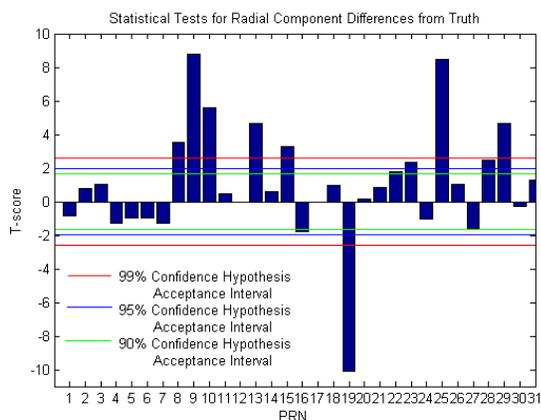
Figure 7 - Preliminary Radial Statistical T-scores

Table 3 - Preliminary Radial Statistical Results

SV PRNs	$\Delta_0 = 0$ Level	Count
1-7,11,14,16,18,20-22,24, 26, 27, 30	Pass	19
8-10,13,15,19,23, 25, 28, 29	Fail	9

The clock errors were also modified based on the initial results. Preliminary results show a significant improvement in the clock error modeling in the civilian data service. Figure 8 shows the T-scores for the clock error differences from the military errors. Here, all but 2 satellites now show that there is no difference between military and civilian data services. Table 4 shows a summary of the results.
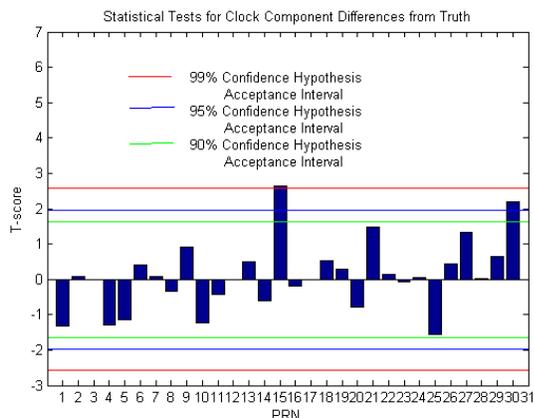
Figure 8 - Preliminary Clock Statistical Results

Table 4 - Preliminary Clock Statistical Results

SV PRNs	$\Delta_0 = 0$ Level	Count
1-11,13,14,16,18-29	Pass	26
15,30	Fail	2

It must be noted that these results are preliminary and further testing will be performed to determine the exact benefit experienced as a result of these changes. It is expected that these changes will be in effect by October 2005.

ACCURACY STATISTICAL RESULTS

Accuracy at a particular site is determined by combining the line-of-sight ephemeris and clock errors and the geometry of the satellites into a point position solution [7]. Navigation Tool Kit models this process, using PAF files as inputs, and then calculating the navigation errors at each site the user has specified. Many factors can affect the magnitude of the navigation errors, including visibility to the satellites (due to physical or electronic barriers), atmospheric conditions and receiver model characteristics. For the accuracy comparisons required by this study, I set up the Navigation Tool Kit to produce Signal-In-Space accuracy results: no other effects were modeled. A single scenario was created in NTK and then modified by changing the date and exchanging the civilian input data for the military input data. All other characteristics remained constant between the two scenarios. Figure 9 shows the military and civilian navigation errors for a typical day. The ordinate axis is in K-Points – 15 minute segments. The accuracy at the site was determined at 15 minute intervals, to coincide with the spacing of the input PAF data.
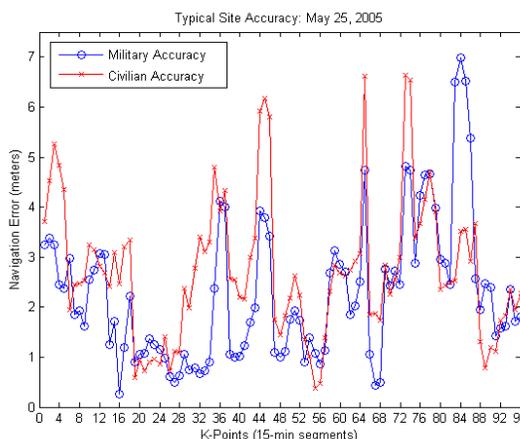
Figure 9 - Accuracy for a Typical Day

The military and civilian navigation errors for each day were then averaged, obtaining a daily mean error. Figure 10 shows the daily mean errors for the 113 day analysis period. Also shown in the figure is the difference between the military and civilian daily mean errors.
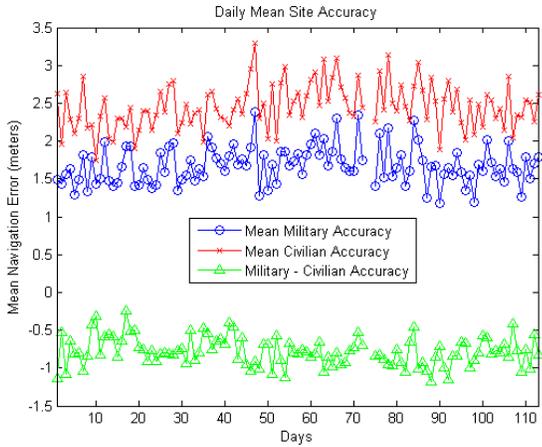
Figure 10 - Daily Mean Navigation Errors

To perform the PDA on this set of data, the differences in errors must be normally distributed. To test that the differences were normally distributed, I performed a Lilliefors normality test [8]. A normality plot is shown in Figure 11. If the data is normal, the normality plot will show that the data falls very nicely along a straight line. The normality plot shows notionally that the differenced data is normal and the results of the Lilliefors test confirm that this is indeed the case, at a 99% confidence level.
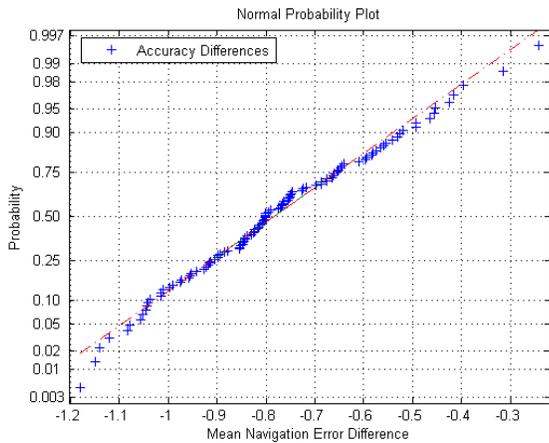
Figure 11 - Normal Probability Plot

The histogram in figure 12 shows how the daily mean accuracy values are distributed. Also indicated is the peak at approximately -80 centimeters.
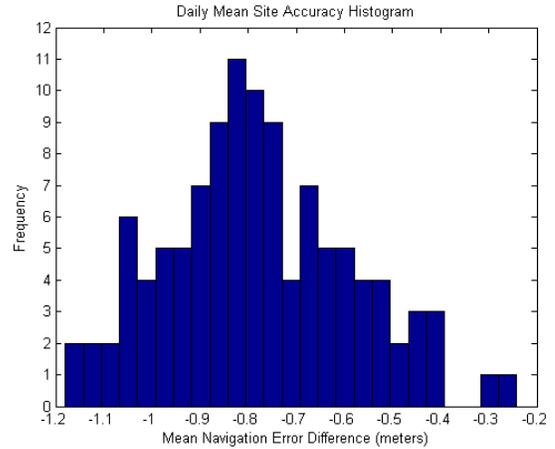
Figure 12 - Histogram of Daily Mean Accuracy

Now that we know the differenced accuracy data is normal, we can proceed with the statistical tests to see if our H_0 accuracy hypothesis is correct.

To make this analysis pertinent to a military planner, I'm defining several threshold (Δ_0) levels. The military planner can then look at his or her threshold of interest and then look at the results of the tests to know whether one service will be better for their particular mission. To complete the table, separate statistical tests were performed at each threshold level and a determination of pass or fail was made based on a 95% confidence level as prescribed by the T-test. Table 5 shows the results of those tests.

Table 5 – Military Operational Impact: Accuracy Statistical Results

Operational Threshold (Δ_0) Signal-In-Space	T-test Pass/Fail	T-score
20 Meters	Pass	1058.9
10 Meters	Pass	508.0
5 Meters	Pass	232.5
3 Meters	Pass	122.3
2 Meters	Pass	67.2
1 Meter	Pass	12.1
50 Centimeters	Fail	-15.4

In this table, if *Pass* is stated next to your threshold for operations, then there is no statistically significant reason to choose one data service over the other. There may well be other logistical or procedural reasons to choose one service over the other, but this analysis shows that either service will provide accuracy answers to meet your mission plan. If *Fail* appears next to your threshold for operations, then there is a statistically significant reason to choose the military service over the civilian service. In this case, the military service will provide better accuracy answers for your mission.

Civilian Service Derived Accuracy Improvements
As a result of the modifications made to the radial and clock errors on the civilian data service, new accuracy scenarios were executed – to get some preliminary evidence one how much the site accuracy had improved. Again, I created identical scenarios in NTK, changing only the input error source data – from Military to Civilian. This time however, I used the new preliminary PAF file results for the civilian scenarios. To make a statistical test, I created 29 days worth of scenarios, thus making 29 statistical observations to which I could apply the Paired Data T-test. Applying the T-test with 28 degrees of freedom, I ran the tests and came up with the results shown in Table 6. In this case, note that the Military Operational Threshold has been tested down to 5 centimeters. The corrections applied to the civilian data service have led to an average increase in accuracy of approximately 60 centimeters. It must be noted that these results are preliminary and further testing will be performed to determine the exact benefit experienced as a result of these changes.

Table 6 – Military Operational Impact: Preliminary Improved Accuracy Statistical Results

Operational Threshold (Δ_0) Signal-In-Space	T-test Pass/Fail	T-score
1 meter	Pass	50.0
50 centimeters	Pass	19.3
40 centimeters	Pass	13.2
30 centimeters	Pass	7.1
20 centimeters	Pass	0.9
10 centimeters	Fail	-5.2
5 centimeters	Fail	-8.3

Prior to the changes made on the civilian data service, accuracy for a typical day was as shown in Figure 13. The new graph of accuracy for the same day is shown in Figure 14. This graph shows a marked improvement in accuracy for that day. Figure 15 shows the mean accuracy by day for the new service as well as the old service – a clear distinction can be made, leading to the conclusion that the modifications to the civilian data service will indeed improve accuracy assessments.
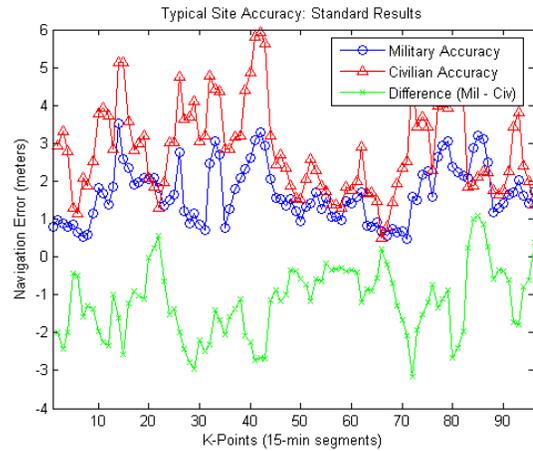
Figure 13 - Typical Accuracy for a Day
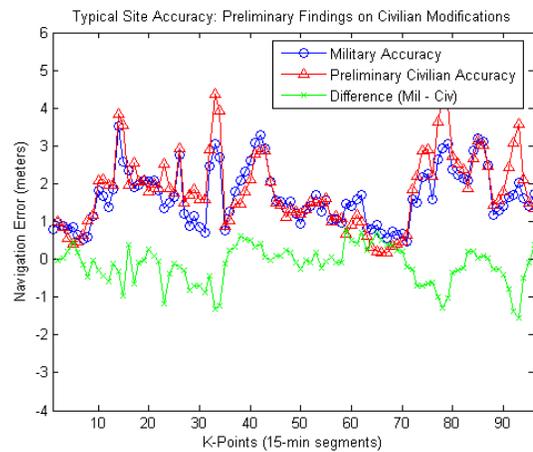
Figure 14 - Improved Accuracy for Same Day
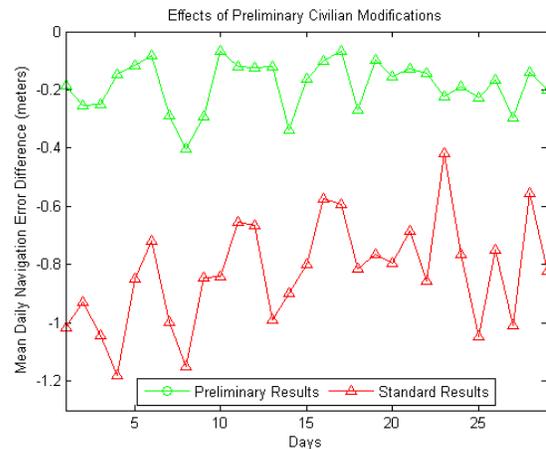
Figure 15 – Improved Daily Mean Accuracy

CONCLUSIONS
A full statistical comparison of military and civilian navigation error data services has not been performed before – though a clear need for this type of comparison exists from a military point of view. The military has the option of choosing a navigation data error source if

conditions warrant – but no way of determining what level of service the civilian data service provides. In this paper I have not only outlined and performed the tests necessary to judge the differences in the two services, I have also laid out tables of Military Operational Thresholds for quick reference to the results.

I compared first the radial and clock errors of the two data services to a truth source derived from the NGA precise ephemeredes and the broadcast ephemeredes for each satellite. I then compared the accuracy calculated by the Navigation Tool Kit using the data delivered by both services. My analysis showed that only 5 of the 28 GPS satellites analyzed had no statistically significant reason to choose the military over the civilian data service based on radial error alone. However, the initial clock errors showed that there was a statistically significant reason to choose the military data service for every satellite based on clock errors. Based on the initial results, I conferred with Navcom to determine possible modifications that could be made to the civilian data service. After implementing some of the planned changes I retested the errors against the truth data and presented the improved results.

My results show that for the military planner, an operational signal-in-space threshold of greater than 78 centimeters would lead them to choose either the military or civilian navigation error data service – there is no statistically significant reason to do otherwise. Below an operational threshold of 78 centimeters, the military planner would do better to choose the military data service. That is the case for only the very near future however. Within the next month or two, modifications will be put in place within the civilian navigation error data service that will bring the level of statistical significance to approximately 20 centimeters. That is, with an operational threshold of 20 centimeters or greater, there is no statistically significant reason to choose one service over the other.

An additional consideration, based off of the data in Figures 3 and 4, is the greater variability of the military errors over that of the civilian errors. While this increased variability was not enough the affect the outcome of the statistical tests, it is cause for further study.

ACKNOWLEDGEMENTS

REFERENCES

[1] NGA Geospatial Sciences Division website: earth-info.nga.mil/GandG/sathtml.

[2] CORS website: www.ngs.noaa.gov/CORS

[3] Interface Specification – Global Positioning System – 200D, available from the US Coast Guard Navigation Center: www.navcen.uscg.gov/gps/geninfo/IS-GPS-200D.pdf

[4] Ibid. See the calculations in Table 20-IV and Section 20.3.3.3.3.1

[5] Devore, Jay L.: Probability and Statistics for Engineering and the Sciences, Third Edition

[6] March 4, 2005 to June 24, 2005

[7] Parkinson, Bradford W., Spilker, James J., Axelrad, Penina: Global Positioning System: Theory and Applications Volume 1, Chapter 9

[8] Lilliefors, H. (1967), On the Kolmogorov-Smirnov test for normality with mean and variance unknown, Journal of the American Statistical Association, 62.