Global Optimization of GPS Augmentation Architectures using Genetic Algorithms

Samuel P. Pullen, Per K. Enge, and Bradford W. Parkinson

Department of Aeronautics and Astronautics
Stanford University

BIOGRAPHIES

Sam Pullen is finishing his Ph.D. thesis as a Research Assistant in the Gravity Probe-B project at Stanford University. An S.B. Graduate of MIT, he conducts research on spacecraft design for reliability and robust control design along with studies of DGPS performance.

Per Enge is a Research Professor of Aeronautics and Astronautics at Stanford University. A Ph.D. graduate of the University of Illinois, his research focuses on DGPS aircraft landing applications. He previously taught at WPI and is an ION Satellite Division ex-Chairman.

Brad Parkinson is a Professor of Aeronautics and Astronautics at Stanford University and is the program manager for the Gravity Probe B spacecraft. He served as the first Program Director of the GPS Joint Program Office and was instrumental in the system's development.

ABSTRACT

The prediction of augmented GPS performance for spread-out user locations requires analyses of both accuracy under normal conditions and integrity in the case of system failures. Methods that combine covariance propagation and Monte Carlo simulation for the Wide Area Augmentation System (WAAS) have been developed, allowing system designers to study performance, risk, and cost tradeoffs. This process can be automated into computer search techniques that make WAAS network optimization possible

Revisions to our previously published accuracy and integrity algorithms have been made, including more detailed WAAS accuracy models and probability models for spacecraft, ionosphere, and ground errors. Updated results are given for the FAA testbed (NSTB) network and for an example WAAS system for Europe.

Next, a framework for network optimization is constructed from two bases. A top-level *user value model* expresses the relative quality of the combined accuracy and integrity evaluations for a given network. Global optimization is carried out using a *genetic algorithm* which maintains a population of possible network designs and "evolves" the next generation using operators derived from the Theory of Natural Selection. The optimization process is computer-intensive but has the potential to converge to the best possible network for a given application. A complete model for European WAAS network optimization is presented, and the prospects for improved computer speed using parallelized code are discussed.

1.0 Introduction

Networks of ground stations and geosynchronous satellites designed to augment civilian GPS navigation performance have been shown to provide corrected pseudorange accuracies of 1-2 meters, making aircraft precision approach using augmented GPS possible. Local Area Augmentation Systems (LAAS) broadcast corrections from a single site to nearby users. Wide Area Augmentation Systems (WAAS) instead use a network of spread-out reference stations (WRS's) which transmit their observations to a master station (WMS). This master site computes coordinated corrections for all GPS satellites in view of any WRS and uplinks them to communications satellites for downlink to any user within a very large geographic region [1].

While the augmented-GPS performance demonstrated to date in flight tests is very promising, the prediction of overall system performance for the entire user population is difficult. Previous studies conducted at Stanford have described new methods to predict normal-condition WAAS navigation *accuracy* over a large user area [2]. In addition, Monte Carlo simulation of specific WAAS failure modes allows a prediction of post-RAIM *integrity* risk, or the risk of being placed in a dangerous situation due to not being warned of a GPS/WAAS system failure [4]. Combining these two separate evaluations gives a comprehensive picture of the overall performance and acceptability of a given WAAS network.

This paper expands and extends our previous work in several respects. In Section 2.0, the WAAS coverage prediction model is summarized, and new results for the FAA WAAS testbed are shown which incorporate improved ranging error models. In addition, new results for a proposed European WAAS network are presented. Section 3.0 describes the integrity evaluation model, which uses Monte Carlo sampling to optimize RAIM thresholds in the presence of various rare-event errors. Integrity results of this evaluation for the WAAS networks studied in Section 2.0 are given there.

Section 4.0 summarizes the use of *genetic algorithms* (GA's) to provide a very flexible global optimization capability. The evolutionary search operators used in standard GA's are explained, and the encoding of a GPS augmentation network into a GA-compatible form is demonstrated. Section 5.0 outlines our proposed WAAS network objective function or value model which is used to combine DGPS performance analyses with a user cost/benefit assessment into a single top-level measure. This function goes a step beyond current requirements to suggest, at a policy level, what the basic goals of augmented-GPS navigation systems should be.

Section 6.0 demonstrates the potential of top-level GA optimization by showing the improvement obtained for the European WAAS network in the first few GA generations. Finally, Section 7.0 discusses the software improvements needed to allow optimization of GPS networks on a national and international scale. A set of conclusions from the latest augmented-GPS performance studies and optimization runs are then given. In particular, WAAS is shown to be a cost-effective approach to wide-area precision navigation, and further developments should provide the capacity to tailor future systems to specific user needs in a very efficient manner.

2.0 WAAS Coverage Prediction Methodology

While it is already apparent that WAAS has the potential to provide Category I accuracy for aircraft landing and that baselines of hundreds of kilometers are possible, a coverage prediction model is needed to predict accuracy across the entire geographic spread of users and to help determine just how many wide-area reference stations (WRS's) are needed to meet the RTCA MOPS accuracy requirements [1]. Our method is summarized here, and revisions to the WAAS error models are also explained along with new WAAS network results.

2.1 Summary of Coverage Prediction Method

The coverage prediction approach used here is based on the solution of least-squares covariance equations for given GPS and WRS geometries. This is a summary of the detailed explanation of the method contained in [2].

Accuracy predictions for large geographic areas are generated by a computer program which simulates a large number (between 1440 and 10,000) GPS and geosynchronous satellite geometries using a GPS orbit model. For each geometry, the matrix of direction cosines to each visible satellite $\mathbf{G}_w^i$ is computed for each WRS location *i* (using a 5^o mask angle). At this stage, the ranging observation errors for each satellite visible at each WRS are computed from the RMTSA model given in Table 1. The large WRS ionosphere covariance matrix $\mathbf{P}^p$ can then be computed element-by-element.

Noise Source	WRS Error (m)	User Error (m)
receiver noise	0.33	0.50
SA latency	not applicable	0.20
multipath	$0.20 / \tan(\varepsilon)$	$0.30 / \tan(\varepsilon)$
troposphere*	$0.07 / \sin(\varepsilon)$	$0.20 / \sin(\varepsilon)$

*changed since publication in [2]

Table 1: One-Sigma RMTSA Errors

The program then cycles through a grid of user locations separated by 1-4 degrees in latitude and longitude. For each user, the geometry matrix $\mathbf{G}_u$ is computed, and two separate processes of covariance propagation are carried out in parallel. The first is for *clock/ephemeris* error for satellites in view of the user (using a 7.5^o mask angle) based on the WRS's that can also see the satellite in question and can provide clock/ephemeris corrections. The second is for *ionospheric* spatial decorrelation projected from the pierce points observed by each WRS to the WMS, which fits a set of predictions to a grid, and finally to each user.

Covariance projections from these two error sources are brought together into a single pseudorange error covariance matrix $\mathbf{P}_v^*$ for each user which includes user-specific RMTSA errors. The weighted least-squares position error covariance $\hat{\mathbf{P}}_x$ is then computed, and the (Gaussian) vertical position error variance is given by the [3,3] entry of this final matrix. The vertical error for each geometry is stored in a histogram for that user, as is the Vertical DOP for the satellite geometry visible to that user [2]. "Availability" in this case is defined as the percentage of geometries for which a given user's vertical one-sigma error (given by $\sqrt{\hat{\mathbf{P}}_x[3,3]}$) is within either the ILS or RNP Category I one-sigma requirements of 2.05 and 3.6 meters respectively. Geometries for which this requirement is exceeded are deemed "non-available", and if this state persists over time, an "outage period" for Category I landings results.

Figure 1 gives a conceptual flow chart for this covariance propagation method. Note that propagation of error covariance follows two separate, parallel paths. The *clock/ephemeris* process results in a 4 x 4 matrix $\mathbf{P}_k^{SV}$ of error covariances for each satellite *k* (this is equal to the SPS covariance if no WRS's can see this satellite). For each user, the covariances for all the satellites he or she can see are arranged into $\mathbf{P}^{SV}$, which acts as the "plant" matrix for computation of the user position error covariance as follows:

$$\mathbf{P}_{\hat{x}} = \mathbf{G}_u^* \tilde{\mathbf{G}}_u \mathbf{P}^{SV} \tilde{\mathbf{G}}_u^T \left(\mathbf{G}_u^*\right)^T + \mathbf{G}_u^* \mathbf{P}_v^* \left(\mathbf{G}_u^*\right)^T \qquad (1)$$

Note that the User Differential Range Error (UDRE) for each satellite in view is given by the diagonal elements of the matrix $\tilde{\mathbf{G}}_u \mathbf{P}^{SV} \tilde{\mathbf{G}}_u^T$, which maps the clock/ephemeris error from the WMS correction through the user's satellite geometry (expressed by $\tilde{\mathbf{G}}_u$) [2].

The process of *ionosphere* covariance propagation instead fits the vector of combined WRS ionosphere pierce-point measurements to the WMS ionosphere grid of 5-15 degrees in latitude and longitude. A model for estimating the decorrelation between ionospheric delays at points that are far apart has been fitted to data in [3] and is detailed in [2]. The covariance of this fit, $\mathbf{P}^G$, is then propagated to each pierce point of each user in the user grid. The resulting ionosphere fit error covariance, $\mathbf{P}_\varepsilon^U$, is propagated to the overall position error as part of the "noise" term on the far right-hand-side of (1). The diagonal elements of $\mathbf{P}_\varepsilon^U$ also give the User Ionosphere Vertical Error (UIVE) variances.

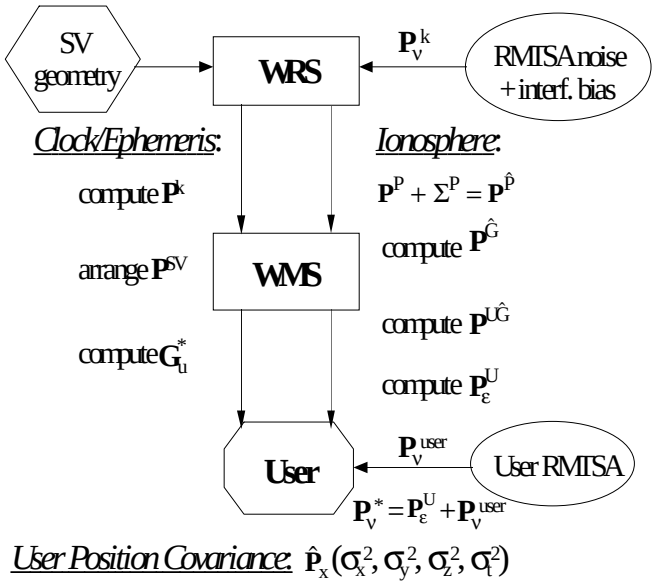


Figure 1: WAAS Covariance Overview

2.2 Revised Error Models

An ongoing effort is being made to update the error modes used for WRS and user observations. These revisions are based on the latest research on real-time algorithms at Stanford and the results from the experimental Stanford WAAS, which has three WRS's at Arcata and San Diego, California, and Elko, Nevada [9].

The Stanford WAAS implements *carrier smoothing* to reduce WRS observation errors. A Hatch/Eshenbach filter is used to average code psuedorange observations with much more precise carrier information (which has only 1-2 mm of noise) [11]. When a WRS first sees a given GPS satellite, the averaging process begins, leading to a reduction in the magnitude of receiver and multipath noise as a function of the time that satellite has been observed (without a cycle slip). Receiver noise has a short correlation time, but multipath takes much longer to average out. We now use an abstract exponential-decay model which gives a combined noise reduction factor *NRF* defined as follows:

$$NRF = \exp\left(-\frac{t_{obs}}{\tau_{cs}}\right) \qquad (2)$$

where the generic carrier smoothing time constant τ_{cs} is conservatively estimated to be 60 minutes. In the code, the cumulative time t_{obs} is tallied as the satellite geometry is updated. The receiver and (elevation-dependent) multipath standard deviations (from the RMTSA) are then reduced by multiplying by *NRF* from (2).

In addition, the assumption of a *bivariate Normal* distribution among ionospheric pierce-point observations has been relaxed. In [2], the variance ($\sigma^2(d)$) of the true ionosphere delay relative to an observed point a distance *d* away is given by a linear/exponential function of *d*. This variance is the converted to a covariance entry $\bar{\sigma}^2$ between two points using the bivariate Normal equation:

$$\bar{\sigma}_{i,j}^2 = \sigma_b^2 \left[1 - \left(\sigma_{i,j}(d)/\sigma_b\right)^2\right]^{0.5} \qquad (3)$$

where σ_b is a base deviation at a given point, assumed to be about 2.8 meters. The exponent 0.5 in the bivariate formulation results in closely correlated ionosphere measurements, even when separated by hundreds of kilometers. As a result, this exponent has been increased to 1.0 for our current studies, introducing more spatial decorrelation into the ionosphere correction process.

2.3 Ionosphere Observation Model Variants

Research on improving the calibration of satellite and receiver *interfrequency bias* suggests that this prevailing

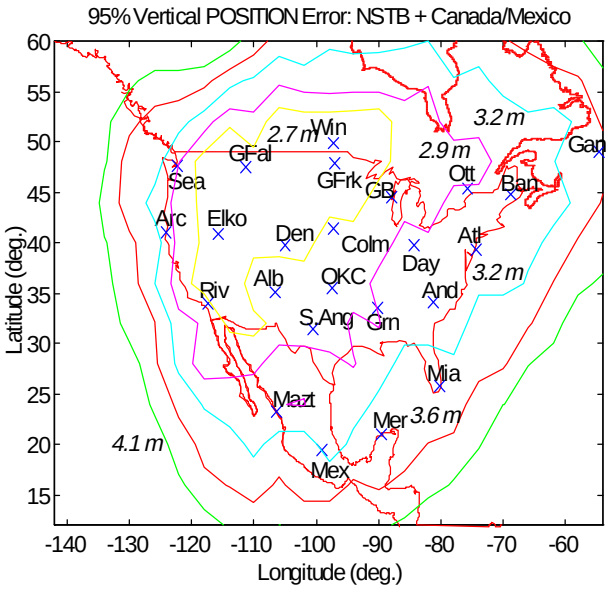


Figure 2: FAA NSTB 95% Vertical Accuracy

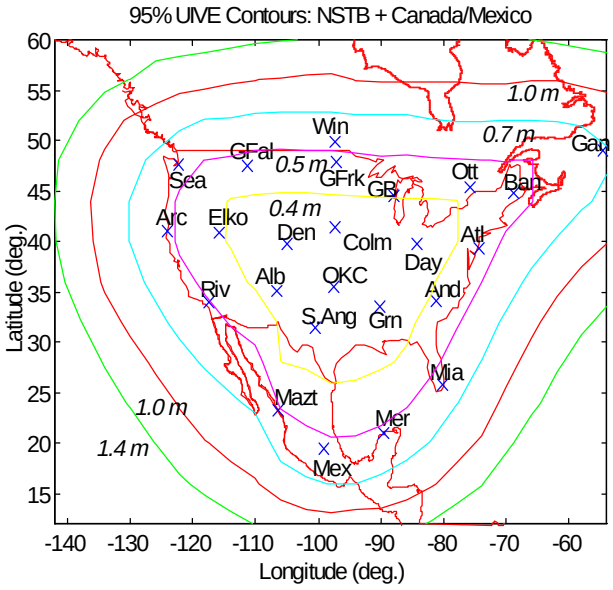


Figure 3: FAA NSTB 95% UIVE

bias error (which affects WRS ionospheric delay observations) can be cut almost in half from what is observed today. Currently, the ionosphere model in [2] assumes a 0.75-meter additional noise (1σ) term to model uncorrected interfrequency bias. Estimates of the slowly-changing bias parameters are possible over several hours of data-taking, making it possible to improve this to about 0.4 meters [12]. This has not yet been demonstrated in an end-to-end sense; thus this adjustment is considered a provisional improvement.

A more radical change to the WAAS network would be to assume only *single-frequency* WRS ionosphere observations. It is possible to extract a measurement of ionospheric delay from the code-carrier divergence on a single broadcast frequency, as described in [8,13]. Because dual-frequency measurements by definition require the use of L2, which is not part of the SPS service guaranteed to civilian users, WAAS networks deployed by non-U.S. agencies may choose to restrict themselves to the use of L1 measurements only [14].

From comparisons of single-frequency measurements to more accurate dual-frequency ones, our best current estimate is that the use of single-frequency observations would add a one-sigma *vertical* error of around 0.8 meters to the WRS ionosphere delay measurement. However, since single-frequency receivers cannot directly separate ionosphere from other error sources, the covariance propagation method used here must be re-worked to combine the clock/ephemeris and ionosphere into one larger estimator for this case. We now expect this change to reduce 95% accuracy by about 15-20%, but the effect on integrity could be much worse.

2.4 Results of Improved Models

All of the following results incoporate carrier smoothing (2) and revised ionosphere decorrelations (3). The effects of further changes are cited where applicable.

Figures 2-4 show 95% vertical error, 95% UIVE, and 95% UDRE, respectively, for the revised FAA WAAS testbed, or NSTB, which will precede the operational WAAS into service on an experimental basis. It now includes six additional WRS's, three in Canada and three in Mexico, giving a total of 24. From Figure 2, the best accuracy is obtained in the Western plains states and southern Manitoba, where users can see both the 180° and 55° W geosynchronous satellites. Overall accuracy, compared to the 18-WRS all-CONUS NSTB [2], has improved to the point that all of CONUS is within $2\sigma_v$ = 3.5 m under normal conditions. Improvement is most noticeable over the Eastern seaboard. The three WRS's in Mexico provide improved coverage to the Southwestern U.S. and provide better than 3.6 m accuracy over almost all of Mexico.

In Figures 3 and 4, both UIVE and UDRE contours tend to follow the outline of the outermost WRS's in the network. Ranging errors from both sources are comparable, and both are very small near the center of the network. It is a little surprising that UDRE degrades more rapidly as one moves away from the center, as fewer WRS's provide data to correct pseudoranges of satellites visible to these more distant users.

One problem with the proposed WRS layout is that, while the Gander WRS is in a very useful location, the Canadian WRS's in Ottawa and Winnipeg are too far

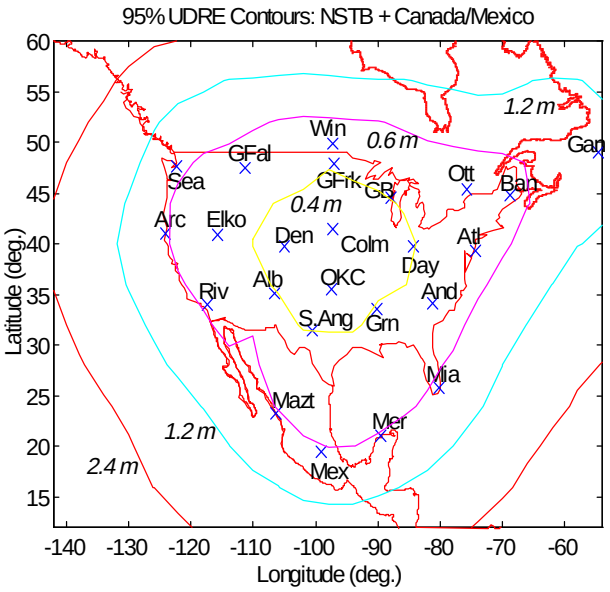


Figure 4: FAA NSTB 95% UDRE

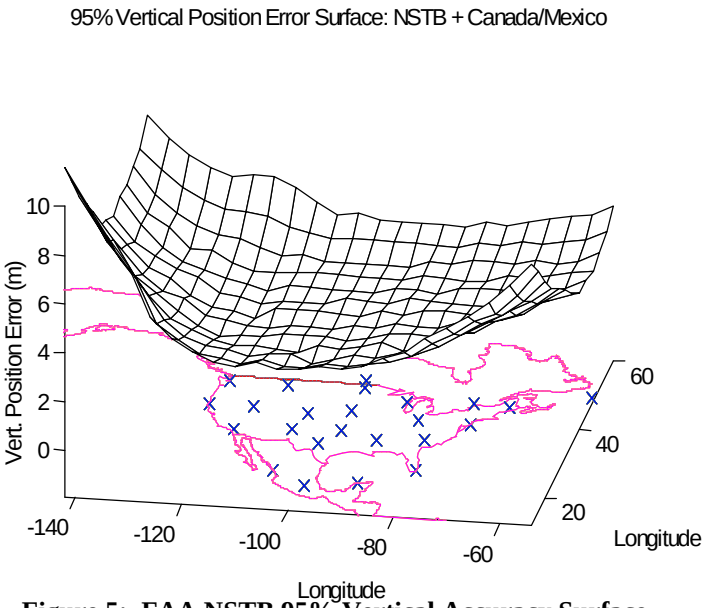


Figure 5: FAA NSTB 95% Vertical Accuracy Surface

East and are too close to the northern border of the U.S. to add significantly to coverage over most of Canada. Figure 5, which plots 95% vertical accuracy in 3-D, shows that accuracy degrades most rapidly in the Northwest corner of the user grid, over British Columbia. This may not be a significant drawback, since the network will be used as an experimental testbed only, but it is possible to achieve significantly better coverage of the Canadian provinces by moving the Ottawa and Winnipeg stations westward and northward.

Figure 6 shows 95% vertical position error results for a WAAS that could provide precision GPS corrections to Europe. It has a "four corner" arrangement of WRS's in Scotland, Spain, Northern Russia, and Turkey. Note that this minimal WRS arrangement provides sufficient accuracy to exceed the 4.1-meter vertical 95% ILS Category I precision approach requirement over almost all of Europe. This shows the potential of WAAS to provide high accuracy in a very cost-effective way, but the next question is whether such an arrangement also provides sufficient integrity, or user safety.

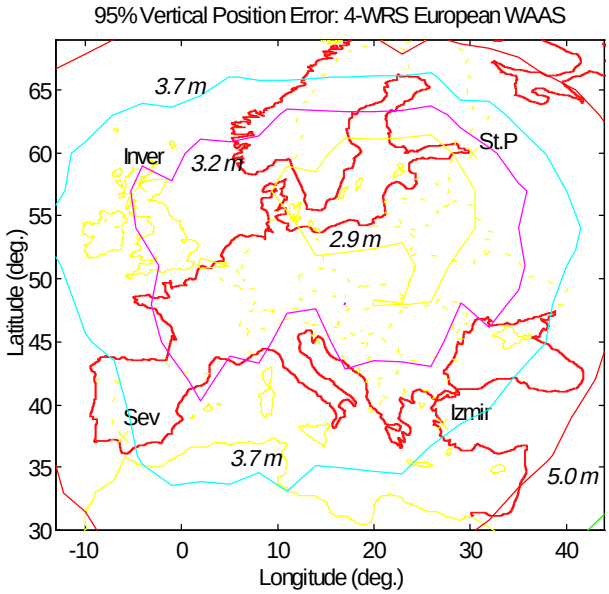


Figure 6: European WAAS 95% Vertical Accuracy

3.0 WAAS Integrity Simulations

3.1 Background and User Cost Model

Unlike the "normal conditions" assumed by the WAAS coverage prediction model, integrity threats are hazardous events that are presumed to occur rarely but have the potential to put the user in serious danger if he or she is not promptly warned. Since WAAS includes both ground-station and user elements, detecting these events is a shared responsibility of the augmentation network and of each user. Individual users can use Receiver Autonomous Integrity Monitoring (RAIM) to provide a warning from his overdetermined set of ranging measurements if he can see more than four satellites. This process uses the magnitude of a residual vector as the decision statistic; when it exceeds a pre-set threshold, the user is warned that conditions are unsafe. The ground stations can also monitor statistics that compare primary and redundant system measurements to each other. Warnings of unreliable satellites or corrections can then be included in the WAAS message to warn all affected users [5,6].

In [4], a method of setting WAAS RAIM residual thresholds based on a *user cost model* was developed and demonstrated. The cost model measures the likelihood of a fatal aircraft accident if certain outcomes, such as

missed detections and false alarms, occur. Table 2 gives this model expressed in terms of consequent user risks from various RAIM outcomes (also see Section 5.4). This relative measure of RAIM event costs allows us to optimize the detection thresholds rather than relying on the chi-square probability assumption used in [6].

RAIM Result	Base Cost	Variable Cost
good position	0	0
detected error	1	0
missed detect.	200	5
false alarm	1.012	0
non-available	0.012	0

Table 2: User RAIM Cost Parameters

3.2 WAAS Rare-Event Probability Models

The RAIM analyses in [4] focus on worse-than-Gaussian ionospheric spatial decorrelations, which are assumed to be possible in the worst 10% of cases based on previously published experimental data [3]. More severe decorrelations are possible in the worst 2% of cases. The effects of these "non-Normal" decorrelations must be analyzed by Monte Carlo simulation.

The approach taken here is based on [4] but uses the coverage prediction UIVE results. It samples ionospheric delay using the Normal distribution implied by the UIVE for a given site. If a sampled standard normal variable $|z| \le 1.645$, the vertical ionosphere ranging error is simply the sample z times one-half of the assumed 2σ 95% UIVE result. If the sample exceeds 1.645 (the 90% cutoff), the UIVE-based standard deviation is *multiplied* by an expansion factor *TM* which accounts for the abnormal spatial decorrelation possibility:

$$\begin{aligned} 1.645 \le |z| \le 2.33 \quad &\Rightarrow \quad TM = 1.13 \pm 0.11 \\ |z| \ge 2.33 \quad &\Rightarrow \quad TM = 1.40 \pm 0.25 \\ \text{otherwise} \quad &\Rightarrow \quad TM = 1.0 \end{aligned} \tag{4}$$

This inflated deviation is multiplied by the already-large sample z to give the vertical delay in this case. Finally, the vertical error is multiplied by the appropriate obliquity factor (a function of the satellite elevation angle ε) to provide the slant pseudorange error [4].

For dual-frequency WRS receivers (the base case), the actual *TM* is periodically sampled from the Normal distribution in (3), where $TM = 1.13 \pm 0.11$ becomes Normal (μ=1.13, σ=0.11). In the case of single-frequency WRS ionosphere observations (Section 2.3), rare-event spatial decorrelation uncertainty is much worse; thus the sampled TM from (3) is at least doubled.

In [4], N_J satellite geometries are updated in the same way as in Section 2.0, but each satellite is sampled to determine if it is "out of service" and is thus not usable. These failure probabilities, derived from [4,10], are:

$$\begin{aligned} \text{Pr (GPS satellite unhealthy)} &= 0.014 \\ \text{Pr (GEO satellite unhealthy)} &= 0.010 \end{aligned}$$

In addition, it is assumed that a failure of the WAAS ground network could lead to errors in the broadcast clock/ephemeris corrections that are much larger than UDRE. This could result from database or computation errors that are not caught by the WRS and WMS monitor systems. Because we are interested in sampling failure cases, the probability of having an increased UDRE is taken to be 0.001, which is at least 10 times higher than expected of an operational WAAS [5]. If a given satellite experiences this type of failure, its effective UDRE is increased by one plus a factor sampled from an *exponential* distribution with a mean $\mu = 2.0$.

3.3 Integrity Simulation Procedure

For each of N_J satellite geometries, N_K failure states are sampled. In each failure trial, each satellite is sampled to see if it is functioning normally, and unhealthy ones are removed from the user geometries. Each satellite next has the state of its ground correction sampled. The overall ionosphere decorrelation state, which applies to all users, is then sampled. A failure *bias* $|z| = 1.645$ is applied to this sample to insure that all cases at least have this amount of non-standard conditions.

Rather than attempting to run simulations for each user in the coverage prediction grid of Section 2.0, the user population is melded into a much smaller number N_U of user locations. For each sampled failure state, a weighted position fix is carried out for each user location. This position fix is based on the normal UIVE and UDRE for that location given by the coverage prediction method. Each user constructs a diagonal weighting matrix $\mathbf{W}$ from the RSS of his RMTSA, UIVE, and UDRE variances (from Section 2.0), but a 25% random factor is applied to prevent the use of perfect weighting information. The true psuedorange error vector *dx* and the weighted residual statistic *D* are given by [4,6]:

$$dx = \mathbf{G}^* z = \left(\mathbf{G}^{\mathrm{T}} \mathbf{W}^{-1} \mathbf{G}\right)^{-1} \mathbf{G}^{\mathrm{T}} \mathbf{W}^{-1} z \tag{5}$$

$$D^2 = z^{\mathrm{T}} \mathbf{W}^{-1} \left(\mathbf{I}_{\mathrm{m}} - \mathbf{G}\mathbf{G}^*\right) z \tag{6}$$

where $\mathbf{I}_{\mathrm{m}}$ is a m x m identity matrix (m is the number of functioning satellites in view), $\mathbf{G}$ is the (m x 4) user geometry matrix, and z is the (m x 1) vector of psuedorange measurements.

After all failure simulations are completed, the 90% of cases that would exist under normal DGPS conditions are added to the *dx* vs. *D* matrix stored for each discrete VDOP bin. Under normal conditions, *dx* and *D* are independent, with (vector) *dx* being Normally distributed with a covariance given by the vertical position error result from coverage prediction and D^2 having a chi-square distribution with variance parameter D^2/σ_z^2 and *m*-4 degrees of freedom [7]. The revised matrix is then searched to find the RAIM detection threshold *T* that *minimizes* the overall weighted RAIM cost (see Table 2) over all N_U user locations (each with a separate *dx* vs. *D* matrix). Note that only one set of thresholds is chosen for all users. If this optimal threshold *T** gives a cost that is lower than the non-availability cost in Table 2, RAIM is *available* for that VDOP and the system is usable. Otherwise, RAIM is *unavailable*, and the system incurs the non-availability inconvenience cost for those trials [4].

3.4 Integrity Results for 4-WRS European WAAS

Integrity evaluation for the European WAAS is done at 11 user locations shown as 'o' in Figure 7 (discussed in Section 4.1). These locations receive different weights as explained in Section 5.0. For these locations and the basic 4-WRS network of Figure 6, 1 million failure states and position fixes were simulated from 1000 satellite geometries for each of the 11 user locations. Table 3 contains the results in terms of the overall weighted RAIM cost and also probabilities of various hazardous events. For WAAS Category I approaches, the RPE or "required (vertical) protected error" is 19 meters at the 200-foot decision height.

Overall RAIM user cost	0.0019
Prob(RAIM available)	0.983
VDOP limit for availability	2.9
Prob(position error > RPE)	3.3×10^{-5}
Prob(missed detect. | error > RPE)	0.109
Prob(false alarm)	0.0007
Fatal Accident Prob. per approach	6.8×10^{-7}

Table 3: 4-WRS European WAAS User Integrity

Although it is clear from Figure 6 (Section 2.4) that this 4-WRS network meets the WAAS Cat. I accuracy requirements at all 11 user locations, it is equally clear here that this network does not provide adequate integrity given the failure uncertainty models from Section 3.2. Availability is not bad at 98.3%, but the probability of exceeding the RPE is too high. RAIM catches 90% of these events, but the remaining 10% that become "missed detections" translate into an unacceptably high fatal accident risk, which is computed by dividing the part of the RAIM cost due to integrity risk (0.0017) by the value of a single fatal accident in this cost model (2500). Note from this value that it implies that 10% of all missed detections (average cost of about 250) lead to fatal crashes. While current Cat. I requirements do not specify a maximum acceptable fatal accident risk, the implied requirement from the RNP and the Cat. III requirements is 10^{-9}. This network has considerably higher risk; so additional augmentations will be required.

4.0 Genetic Algorithm Optimization Model

By combining the coverage prediction model and integrity simulations, it is possible to generate overall evaluations for any GPS augmentation architecture. If the user population (or government agency) can derive a function that computes a top-level "figure of merit" based on the predicted geographic spread of accuracy and integrity performance, optimization of entire networks becomes possible. However, the use of complex covariance and simulation models to generate these evaluations requires a flexible optimal-search approach that does not require well-defined, deterministic problem formulations.

Evolutionary algorithms, a recent development of research in Artificial Intelligence (AI), now provide this capability. Several specific methods, including Simulated Annealing and Genetic Algorithms, have been used to solve a wide variety of problems. In general, they attempt to "evolve" better solutions over time by perturbing the best solutions found up to that point using semi-random operators that can avoid being "trapped" by local maxima or minima. They can also tolerate the noisy evaluations given by complex simulation models.

4.1 WAAS Network Design Encoding

Much of the work in designing an evolutionary search method for a specific application lies in tailoring the search to fit a natural encoding of the design space. For WAAS network optimization, the design variables can be expressed in a vector of binary (0/1) elements, or *genes*, which makes it possible to apply a standard *genetic algorithm* (GA) to evolve toward the optimal solution.

A computer search alone cannot design an optimal network -- the input of design engineers is crucial. In this case, we rely on human designers to provide a list of possible WAAS augmentation elements for the GA to consider. This list may include reference station sites, provision for independent monitor sites, and additional geosynchronous spacecraft to broadcast corrections and add redundant ranging measurements. Essentially, it can include any option that can be modeled in the GPS/WAAS accuracy and safety prediction algorithms.

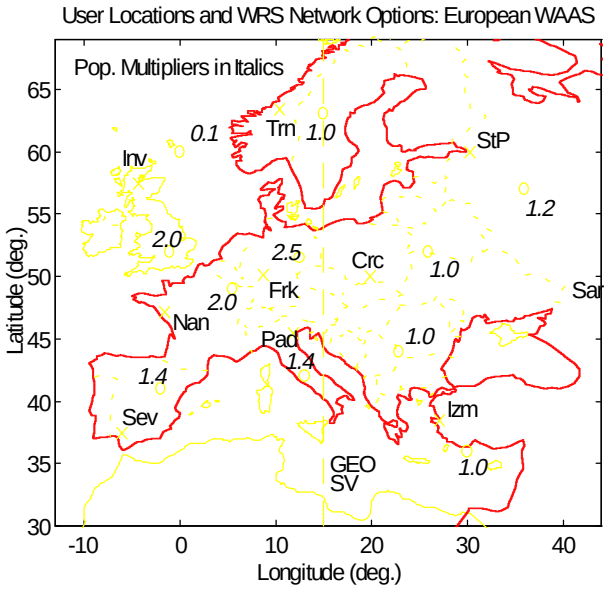


Figure 7: European WAAS Locations and Options

Table 4 below gives a list of 12 augmentation options for the European WAAS application that we shall consider in this paper. Ten of these are potential WRS locations, which are shown as 'x' in Figure 7. Next is the use of single-frequency ionosphere corrections (see Section 2.3) instead of dual-frequency ones, and the last is a third geosynchronous satellite placed over Central Europe at 15° E longitude (shown by a dashed line in Figure 7). Given this list of options, a design solution is simply a vector of N_d = 12 0-1 entries, where a 1 represents the presence of the relevant option and a 0 represents its absence. Note that the addition of new options thought up by the human designers can be handled simply by increasing the length of the design vector. This flexibility is important, as the results of early evaluations and optimization runs may motivate the designers to think of new augmentation options.

No.	Augmentation	Inc. Cost ($ K)
1	Inverness WRS (UK)	2000
2	Seville WRS (SP)	2000
3	St. Petersburg WRS (RU)	2000
4	Izmir WRS (TU)	2000
5	Padua WRS (IT)	2000
6	Trondheim WRS (NO)	2000
7	Saratov WRS (RU)	2000
8	Nantes WRS (FR)	2000
9	Frankfurt WRS (GE)	2000
10	Kracow WRS (PO)	2000
11	Single-Freq. WRS RCR's.	- 90 N_{WRS}
12	add'l. GEO SV at 15° E	25,000

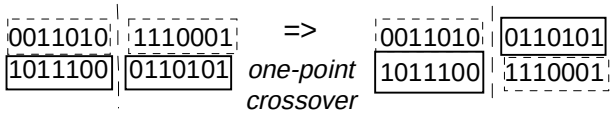
Table 4: European WAAS Augmentations

4.2 GA Population Evolution Operators

A genetic algorithm evolves a population of N_p design solutions (we use 10) from one generation to the next. The population is initialized by combining baseline solutions chosen by the designers (such as the 4-WRS network shown in Figure 6) and randomly generated solutions in a 50-50 ratio. This becomes the "zeroth" generation of the GA search. An accuracy and integrity evaluation of each is then conducted to provide an objective-function value for each initial network design.

The canonical GA used here evolves the next-generation design solutions based on the current-generation members and objective values. Three operators are used. The first is *reproduction*, in which a percentage P_R = 60% of the current solution members are chosen as parents of the next generation according to their objective value, or *fitness*. This is done by a variant of *roulette-wheel selection*, in which the parents are randomly chosen with probabilities that are proportional to their linearly normalized fitnesses. In addition, the best solution is automatically copied directly into the next generation (*elitism*) [15].

Parents selected by this process are "mated" together two at a time in the *crossover* operator, which simulates sexual reproduction. In the basic *one-point* crossover operator, a location between 1 and N_d is sampled from a Uniform distribution, and the two parents swap their genes before and after that point to make up two child solutions for the next generation:



Another possible choice is *uniform* crossover, in which two parents combine to produce one child. In this case, for each gene location, if the two parents have the same gene, the child gets that gene as well. If the two parents disagree, the child's gene is chosen by a 50-50 random sample [15]. A judicious combination of these different operators can yield a faster and more robust search.

Mutation is the final canonical GA operator. Once the N_n solutions that make up the next generation have been chosen by reproduction and crossover, a Uniform random sample is made for each gene in each solution. If this sample is lower than a chosen mutation probability P_M, that gene (bit) is flipped to its binary complement (e.g., 0 -> 1, 1 -> 0). This process, akin to rare genetic mutations in biological organisms, helps maintain the genetic diversity of the solution population, preventing a small set of apparently good solutions from achieving

premature dominance (i.e., a local optimum). Normally, P_M is chosen to be ≤ 0.01 (we have chosen 0.01), but higher mutation rates (inducing more diversity) have been successful for other problems [15].

001100111001 => 001101111001
one-bit mutation

4.3 GA Optimization Procedure

Figure 8 gives a flow chart of the procedure by which the GA "breeds" new generations of solutions and evaluates their fitnesses. Generation 0 is initialized as mentioned in Section 4.1, then a loop of generations begins. Given a generation *n*, the fitnesses of each of its N_p solution members are evaluated using both coverage (Section 2.0) and integrity (Section 3.0) analyses fed into the cost model of Section 5.0. Reproduction, crossover, and mutation are then applied to generate the new generation *n*+1. The GA evolution can be stopped when the population (or the value of its best solution) stops improving, or it can be ended after a set number of generations. Each re-evaluation of a given network is added to those conducted previously; thus statistical significance increases with each new evaluation. Once a given network evaluation converges to within an uncertainty tolerance, no further accuracy/integrity evaluations are needed. Therefore, later GA generations will run faster on the computer than earlier ones.

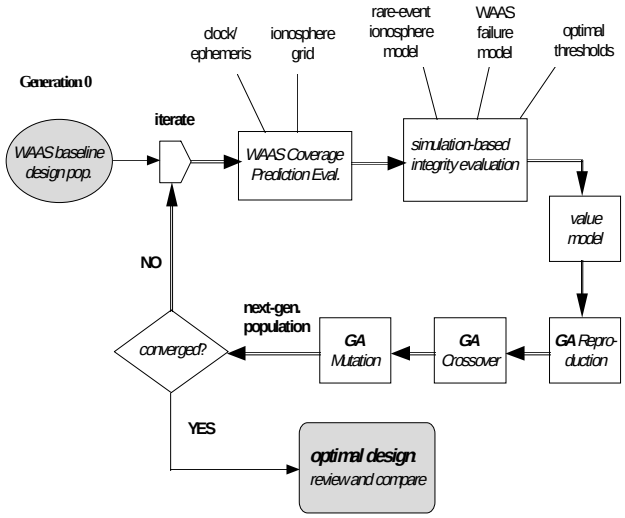


Figure 8: GA Optimization Procedure

5.0 WAAS Network Objective Function

Each of the possible solutions generated by GA evolution needs a *fitness* evaluation, or a measure of its relative "goodness". Because GA optimization is very flexible, there are no mathematical constraints on the form of this system objective function. We can thus construct a "value model" that attempts to express the system's top-level utility for the total user population. This is a key driver of the optimization process, as the GA evolution will tend to exploit any inconsistency or "hole" in the fitness model. For this reason, the elements of the objective function should be carefully considered, and the results of early GA runs may motivate changes in the value model.

The value model developed here is a provisional attempt to weigh user benefits and system costs in as wide a framework as possible, knowing that substantial revisions may be necessary as more designer and user input is received. The overall objective function $F(n)$ to be maximized is given by:

$$F_{\text{WAAS}}(n) \;=\; \sum_{u=1}^{11} PM_u \left[B_{\text{air}}^{u}\, f_{\text{acc}}^{u} - f_{\text{integ}}^{u} \right] - LCost_n \quad (7)$$

where $f_{\text{acc}}^{\ u}$ and $f_{\text{integ}}^{\ u}$ represent evaluations of coverage and integrity performance respectively for user location *u*, $B_{\text{air}}^{\ u}$ is the Cat. I user benefit for a given user location, PM_u is a "population multiplier" which measures the size of the user population near that location, and $LCost_n$ is the acquisition cost of a given WAAS network solution *n*, which includes the procurement cost and four years of OEM (operations and maintenance).

5.1 Population Multiplier

The basic definition for the population multiplier is:

$$PM_i \;=\; \begin{cases} \sqrt{p_i / p_c} & \text{where } p_i > p_c \\ 1 & \text{otherwise} \end{cases} \quad (8)$$

where p_i is the user population (which could be total population, number of air passengers, etc.) and p_c is a "critical value" which insures that all areas covered by WAAS get a minimum base priority. Locations which exceed this critical value do get a higher priority, but it does not scale linearly. The values of *PM* for the 11 user locations selected for the European WAAS is shown in Figure 7. Note that the location over the North Sea is valued at 10% of the overland site values since precision approaches cannot be done there. The maximum value of 2.5 given to the Leipzig, Ger. user location implies a critical value for overall population of about 8 million.

5.2 Network Acquisition Costs

The system acquisition cost for all WAAS networks assumes a well-equipped triply-redundant hardware setup at all ground stations. It includes a WMS procurement cost estimated at $6 million and four years of OEM at $2 million/year, giving a WMS acquisition cost of around $14 million. The incremental WRS cost

estimated to be $1.1 million, includes a $0.5 million procurement cost and $150 thousand per-year OEM cost. The cost savings obtained by using single-frequency receivers in the WRS's is estimated at 75% of the cost of a dual-frequency receiver set mutiplied by the number of WRS's in a given solution. For all ground augmentations, an 80% administrative and indirect cost factor is added, giving conservative final life cycle costs of $25 million for a WMS and $2 million for each WRS (as shown in Table 4). This is based on the high overall cost estimates for the FAA WAAS given in [16]. Finally, the cost of providing an additional geosynchronous satellite is assumed to be $25 million, the estimated cost of an inexpensive satellite designed for just this purpose. As in the Inmarsat case, leasing a GEO transponder may be an option, but the high value of the 15° E location suggests even a lease cost will be much higher than the $2 million/year paid by the FAA. The sensitivity of the optimal result to this cost should be examined further.

5.3 User Benefit Estimates

The calculation of benefits provided by Category I to precision approach users requires making significant assumptions. According to [17], WAAS is expected to increase the number of Category I approaches in the U.S. from 765 (in 1994) to over 5,000. It also suggests an overall user benefit for WAAS Cat. I to be $992 million, or about $200,000 per approach. In Europe, we estimate that this life-cycle per-approach benefit will be doubled due to the poorer weather there. In [18], Europe is estimated to have 326 Cat. I ILS facilities (1994), and we conservatively assume that WAAS will allow this to grow to 1200, giving a total user benefit of $480 million.

An estimate of the per-approach benefit of having Cat. I available is estimated by [18] as saving 2 minutes. Converted to aircraft per-hour fuel and direct operating costs of a weighted mix of passenger aircraft (about $4800), the benefit (conservatively) becomes an average of $160 per approach. Given 1200 Cat. I approaches each providing benefits of $400,000 on average over a four-year life cycle, approximately 3 million Cat. I approaches in Europe are expected in during this time.

A second user benefit to WAAS is removing the need to support and maintain the 326 current ILS facilities that now provide Cat. I capability. This cost is estimated by [18] to be $400,000 per ILS facility (per life cycle), which, multiplied by 326, gives an added benefit to WAAS of $130 million. While it can be argued [14] that the current ILS network has been recently upgraded and represents a "sunk cost," the continual maintenance of it would no longer be necessary after WAAS becomes operational. Under this model, the total life-cycle benefit of WAAS Cat. I is $610 million.

5.4 Accuracy and Integrity Evaluation

The WAAS accuracy evaluation f_{acc}^{u} is simply a percentage of the benefit for each user location, which is broken down from the $610 million total based on the population multiplier for that site. Perfect navigation gets 100% credit, a 2σ vertical error of 2.1 m gets 99%, 4.1 m (the ILS requirement) gets 90%, and 7.6 m (the WAAS RNP requirement) gets only 20% (since it is at the outermost limit of acceptability). A cubic polynomial fit gives, for a resulting 2σ vertical accuracy a (m):

$$f_{\text{acc}}^{u} = 1 - 0.005\,a + 0.0052\,a^2 - 0.0024\,a^3 \qquad (9)$$

where f_{acc}^{u} is in decimal terms (i.e. from 0 to 1).

Converting the RAIM user cost of Section 3.1 to this value framework requires two further assumptions. In Probabilistic Risk Assessment (PRA), it is considered valid to assign cost values to fatalities if the underlying risk is sufficiently small (below 10^{-4}) [19], which it is for GPS integrity. Assuming an average (based on the breakdown of aircraft sizes for Cat. I approaches) of 100 fatalities per fatal incident and a conservative "value per life" of $10 million, each fatal accident incurs a loss of about $1 billion. Since 3 million approaches are foreseen over the 4-year life cycle, and a fatal accident implies a cost of 2500 in the RAIM cost model, we can convert from RAIM cost (R_c) to overall value (f_{integ}^{tot}):

$$f_{\text{integ}}^{tot} = \frac{\left(\$1\times10^{9}\right)\left(3\times10^{6}\right)}{2500} R_{\text{c}} = 1.2\times10^{12}\, R_{\text{c}} \qquad (10)$$

Note that this calculation is also broken down by user location and population multiplier within the RAIM user cost optimization (Section 3.3). Also note that the non-availability cost per approach (0.012) from Table 2, which is included in the integrity evaluation, implies a nuisance cost equivalent to an average of an hour of added aircraft cost, including all consequent delays.

5.5 Value of 4-WRS Baseline European WAAS

The accuracy of the baseline 4-WRS European WAAS network (shown in Figure 6) translates into a accuracy multiplier (weighted by *PM*) of 0.958, giving an overall user benefit of $584 million. However, the RAIM user cost of 0.0019 from Table 3 translates (using (10)) into an integrity cost of $1.73 billion, or 3 times the user benefit. Clearly, this network is insufficient. Note that the acquisition cost of $33 million is dwarfed by the benefits and costs that result, indicating that additional augmentations would be very cheap relative to the possible performance improvement. Also, the fact that a 4-WRS WAAS network cannot provide sufficient integrity suggests that proposed augmented-GPS systems

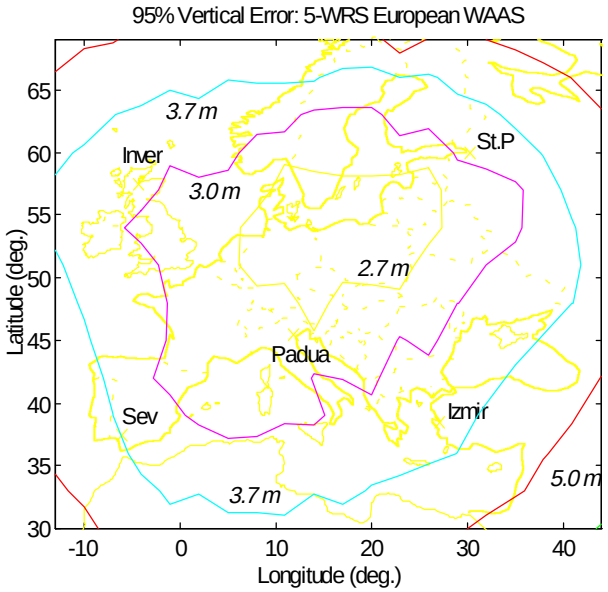


Figure 9: 95% Position Error for 5-WRS Network

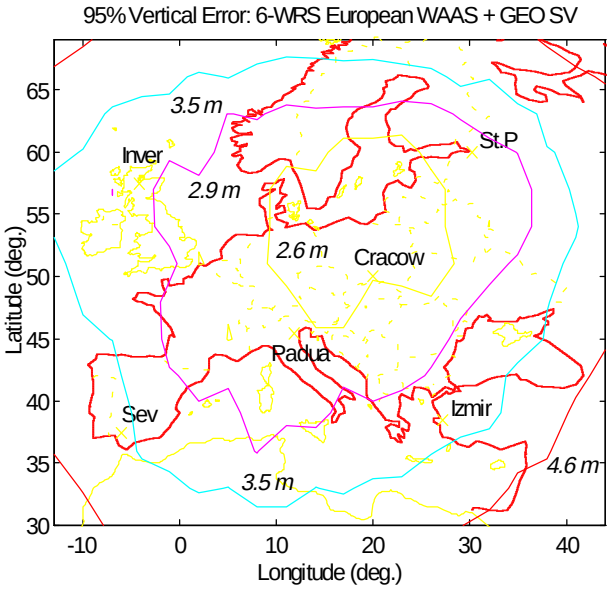


Figure 10: Pos. Error for 6-WRS + GEO Network

for large regions of Europe that are based on one or two DGPS sites would be insufficient as well, even though they may meet the Cat. I accuracy requirements [20].

6.0 "First-Generation" WAAS Results

In our efforts to run the GA optimization code on the European WAAS problem, we have discovered that the software needs to be re-written for parallel processing and that a computer with sufficient available processors will be needed to evolve a population of networks toward optimal convergence. However, we can conduct a first-generation evolution using the GA operators and manually investigate some of the networks that result. Results for two of these variants are shown here.

Figure 9 shows 95% vertical accuracy contours for a network coded [111110000000], which is simply the base 4-WRS network plus a 5th WRS in Padua, Italy, in south-central Europe. Compared to Figure 6, accuracy over highly-populated central Europe is significantly better, resulting in an accuracy benefit of $590.5 million. More importantly, integrity risk has decreased by a factor of 5.6 to give a total cost of $403.6 million. The acquisition cost is still only $35 million, giving a final value of about $152 million. The addition of a single WRS in a beneficial location thus has resulted in a feasible design.

Figure 10 shows vertical accuracy for a network coded [111110000101]. This adds a 6th WRS in Cracow, Poland, and it also uses the additional GEO at 15° E, giving an acquistion cost of $62 million. Although the accuracy contours continue to improve, the benefit has only slightly increased to $593 million. However, the addition of the GEO satellite has made a further substantial improvement to integrity. All failure trials were available, and the integrity cost has fallen to just $73 million. The total value for this network is $458 million, demonstrating that the addition of the 15° E GEO satellite is desirable even if its acquisition cost is much higher than $25 million.

Table 5 gives the relevant integrity parameters for both of these networks. We are continuing to run evaluations of the first and second-generation GA designs as well as manually-designed alternatives, but we have not yet found the point of "diminishing returns" beyond which further augmentations are not cost-effective.

Category	5-WRS	6-WRS+GEO
RAIM user cost	0.000336	0.000061
Prob(available)	0.9998	1.0
VDOP avail. limit	4.6	N/A
Pr(error > RPE)	1.4×10^{-5}	6×10^{-6}
Pr(MD|error > RPE)	0.055	0.020
Pr(false alarm)	0.0002	3.2×10^{-5}
Fatal Acc. Prob/app.	1.3×10^{-7}	2.4×10^{-8}

Table 5: Integrity for European WAAS Varaints

7.0 Conclusions and Further Work

Given the current state of information about augmented DGPS systems (WAAS in particular), it is difficult to make predictions regarding WAAS system-level performance from which network design decisions can be made. We have succeeded in doing so by

developing algorithms that combine covariance propagation to determine position accuracy for large areas of potential users with failure-case simulations that incorporate the best available current knowledge. Further improvements in these prediction methods are possible, including fitting more detailed error models to the rapidly-growing Stanford WAAS database. Better models of ground integrity can also be developed, allowing us to add detailed ground integrity monitor optimization to our current optimal-RAIM algorithm. Finally, the wealth of data to be collected by the FAA's NSTB starting in 1997 should dramatically reduce our uncertainty about potential failure sources, most notably including ionospheric spatial decorrelation.

The augmented-GPS network optimization results we have achieved to date are impressive. We have demonstrated the policy-level feasibility and desirability of using WAAS to provide Category I precision approach capability to Europe with the network designs of Section 6.0, and we are continuing to search for the best possible combination of WRS's and geosynchronous satellites to accomplish this. The 6-WRS + GEO SV combination looks very promising, as it meets all implied Cat. I requirements and provides a value benefit of over $450 million, depending on the cost of the GEO. We plan to expand the applicability of our optimization approach by revising the assumptions of European value model for networks in North America and the rest of the world.

As noted before, our ability to make this vision of augmented-GPS evolutionary optimization a reality requires implementing the coverage prediction and integrity simulation software on a multi-processor computer. This is intuitively easy because the evaluation of accuracy or integrity for each user location is a similar process that can be done simultaneously for as many locations as there are available processors. Stanford's GPS research groups plan to acquire a workstation with at least 16 fast processors by early next year. This computer will be used for extensive simulations of both LAAS and WAAS architectures, as Stanford is contracted by the FAA to evaluate the cost-benefit performance and certifiability of various competing LAAS systems. This work will utilize and further develop the GPS evaluation and optimization techniques reported in this paper.

ACKNOWLEDGEMENTS

The authors would like to thank the following people for their help with this research and the software on which it is based: Y.C. Chao, Dave Lawrence, Dr. Changdon Kee, Boris Pervan, Y.J. Tsai, and Dr. Todd Walter. The advice and interest of many other people in the Stanford GPS research group is appreciated, as is funding support from NASA, the FAA, and Boeing Commercial Airplane Group.

REFERENCES

[1] P. Enge and A.J. Van Dierendonck, "The Wide Area Augmentation System", *Proc. 8th Int'l. Flight Inspection Symposium*, Denver, CO., June 1994.

[2] S. Pullen, P. Enge, B. Parkinson, "A New Method for Coverage Analysis for the Wide Area Augmentation System (WAAS)", *Proc. of ION 51st Annual Meeting*, Colorado Springs, CO., June 5-7, 1995, pp. 501-513.

[3] J.A. Klobuchar, P.H. Doherty, and M.B. El-Arini, "Potential Ionospheric Limitations to Wide-Area Differential GPS", *Proceedings of ION GPS-93*, Salt Lake City, UT., Sept. 22-24, 1993, pp. 1245-1254.

[4] S. Pullen, P. Enge, B. Parkinson, "Simulation-Based Evaluation of WAAS Performance: Risk and Integrity Factors", *Proceedings of ION GPS-94*. Salt Lake City, UT., Sept. 20-23, 1994, pp. 975-983.

[5] R. Loh and J.P. Fernow, "Integrity Concepts for a GPS Wide-Area Augmentation System (WAAS)", *Proceedings of ION NTM-94*, San Diego, CA., Jan. 24-26, 1994, pp. 127-134.

[6] T. Walter, P. Enge, F. Van Graas, "Integrity for the Wide Area Augmentation System", *Proceedings of DSNS-95*. Bergen, Norway, April 24-28, 1995, No. 38.

[7] S. Pullen and B. Parkinson, "A New Approach to GPS Integrity Monitoring using Prior Probabilities and Optimal Threshold Search", *Proceedings of IEEE PLANS '94*. Las Vegas, NV., April 11-15, 1994, pp. 739-746.

[8] V. Ashkenazi, C.J. Hill, J. Nagel, "Wide Area Differential GPS: A Performance Study", *Proc. of ION GPS-92*, Albuquerque, NM., Sept. 16-18, 1992, pp. 589-598.

[9] T. Walter, C. Kee, *et.al.*, "Flight Trials of the Wide Area Augmentation System (WAAS)", *Proc. of ION GPS-94*. Salt Lake City, UT., Sept. 20-23, 1994, pp. 1537-1546.

[10] W. Phlong, B. Elrod, "Availability Characteristics of GPS and Augmentation Alternatives", *Navigation*, Vol. 40, No. 4, Winter 1993-94, pp. 409-428.

[11] R. Hatch, "The Synergism of GPS Code and Carrier Measurements", *Proc. 3rd Int'l. Geodetic Symposium on*

Satellite Doppler Positioning, Las Cruces, NM., Feb. 1982, pp. 1213-1232.

[12] Y.C. Chao, Y.J. Tsai, *et.al.*, "An Algorithm for Inter-Frequency Bias Calibration and Application to WAAS Ionosphere Modeling", *Proc. of ION GPS-95*, Palm Springs, CA., Sept. 12-15, 1995.

[13] C. Cohen, B. Pervan, B. Parkinson, "Estimation of Absolute Ionospheric Delay Exclusively through Single-Frequency GPS Measurements", *Proc. of ION GPS-92*, Albuquerque, NM., Sept. 16-18, 1992, pp. 325-330.

[14] W. Lechner, private conversation, Stanford GPS Research presentation, August 11, 1995.

[15] L. Davis, Ed., *Handbook of Genetic Algorithms*. New York: Van Nostrand Reinhold, 1991.

[16] *A Technical Report to the Secretary of Transportation on a National Approach to Augmented GPS Services*. U.S. Department of Commerce, NTIA Special Publication 94-30, December 1994.

[17] FAA HDQTRS - APO, "Projected GNSS Cat I/II/III Precision Landing Operations", *Proc. of ICAO Comm. /Ops. Meeting*. Montreal: Mar. 27-Apr. 7, 1995. Item 1.

[18] ICAO Secretariat, "Economic Evaluations of the Main Options for Precision Approach and Landing Systems", *Proc. of ICAO Comm./Ops. Meeting.* Montreal: Mar. 27-Apr. 7, 1995, Item 2-3, WP/23.

[19] J.D. Graham and J.W. Vaupel, "Value of a Life: What Difference Does It Make?", *Risk Analysis*, Vol. 1, No. 1, 1981, pp. 89-95.

[20] G. Schanzer, "Satellite Navigation for Precision Approach: Technological and Political Benefits and Risks", *Proeedings of ISPA '95*. Braunschweig, Ger., Feb. 21-24, 1995, pp. 25-30.