INTRODUCTION

Satellites belonging to the Global Navigation Satellite System (GNSS) emit radio signals that are used routinely to determine position. Over the past two decades, satellite-based augmentation systems (SBAS) have been developed to render GNSS position estimates safe and reliable for aircraft navigation. As a radio signal propagates through the ionosphere, it experiences delay due to the presence of charged particles along the signal ray path. Disturbances of the ionosphere can cause this delay to increase dramatically. Indeed, the ionosphere currently remains the largest source of positioning error for single-frequency GNSS users. From a GNSS perspective, the threat to positioning accuracy posed by an ionospheric disturbance depends upon its magnitude. This paper examines how the Wide Area Augmentation System (WAAS), the augmentation of the United States’
Global Positioning System (GPS), monitors the level of ionospheric perturbation over North America, and, in particular, how the algorithms that govern WAAS's response to the occurrence of ionospheric storms have been modified since WAAS's commissioning on July 10, 2003.

A subsequent companion paper, designated *WAAS and the ionosphere – a historical perspective: Mitigating mesoscale irregularities* (Sparks et al., n.d.) will address the evolution of the WAAS methodology designed to protect each user from the impact of mesoscale ionospheric irregularities that are almost, but not quite, severe enough to trigger a local irregularity detector or that are located in regions not sampled adequately by measurements; it focuses on the development of WAAS's undersampled ionospheric irregularity threat model. Note: throughout both the companion paper and the present paper, the terms local and irregularity are defined with reference to mesoscales (i.e., a local ionospheric irregularity is characterized by mesoscale gradients [~100 – 1000 km] in electron density). Our two papers may be regarded as supplementing the broad review of the improvements in WAAS service and integrity since 2003 (Walter et al., 2018). These improvements in WAAS have been achieved by expanding and upgrading its hardware and by updating its system software. The changes to algorithms associated with each new release of the system software are summarized in Appendix A. The history of all WAAS upgrades is discussed more fully in Walter et al. (2018).

WAAS receives GPS data at widely dispersed wide-area reference stations distributed across North America. Each reference station includes three dual-frequency receivers. In its initial operating capability (IOC), the WAAS receiver network consisted of 25 reference stations: 20 spread across the conterminous United States (CONUS) with additional sites in Alaska (Anchorage, Juneau, Cold Bay), Honolulu, Hawaii, and San Juan, Puerto Rico. In the update administered in June of 2006 (Localizer Performance with Vertical

![WAAS receiver sites in North America since March of 2008](image)
Guidance [LPV Release 4], the network was expanded to include four more stations in Alaska (Barrow, Bethel, Fairbanks, and Kotzebue). Three Mexican stations (Merida, Mexico City, and Puerto Vallarta) and two Canadian stations (Gander and Goose Bay) were incorporated into the network in September of 2007 (LPV Release 6/7). The current configuration of 38 stations (see Figure 1) was completed in March of 2008 (LPV Release 8/9.1) with the inclusion of two more Mexican stations (San Jose del Cabo, Tapachula) and two more Canadian stations (Iqaluit, Winnipeg).

WAAS transmits measurements collected by its receiver network to wide-area master stations (WMS), where these data are processed to determine differential satellite corrections, geosynchronous Earth orbit (GEO) orbital parameters, ionospheric delay corrections, and integrity confidence bounds. From the GPS measurement data, a master station derives estimates of the delay that would be encountered by signals propagating vertically through points on an ionospheric grid located at a height typical of the daytime peak of the electron density profile, chosen in WAAS to be 350 km. These ionospheric grid points (IGPs) are specified at regularly spaced intervals of latitude and longitude in a regional mask of the global IGP grid stipulated by the WAAS Minimum Operational Performance Standards (RTCA, 2016).

The vertical delay at an IGP and the safety-critical confidence bound of the error in that delay are designated, respectively, the ionospheric grid delay (IGD) and its grid ionospheric vertical error (GIVE). The IGDs and GIVEs at all IGPs are sent to ground Earth stations and uplinked to the GEO satellites where they are broadcast to users with WAAS-enabled GPS receivers. The broadcast GIVEs allow the user to bound the actual error of the IGDs with a high degree of certainty.

Space weather exerts substantial influence over the state of the ionosphere. Ionospheric storms are disturbances of the upper atmosphere that generate extended regions of enhanced charged-particle density for time periods that may last up to several hours. The presence of an ionospheric storm can cause a GNSS signal propagating toward an Earth-based receiver to experience an unusually large delay, thereby becoming a prospective source of significant positioning error. Of particular concern are compact mesoscale irregularities that have been observed to arise following extreme storms (i.e., large-scale disturbances that occur only rarely over the course of a solar cycle). When irregularities are compact, they can interfere with the propagation of signals detected by a user receiver while remaining poorly sampled by measurements of receivers in the system network. Thus, the impact of space weather represents a potential threat to the safety and reliability of GNSS aircraft navigation.

Under extreme storm conditions, it becomes necessary to deny the use of the precision approach (PA) service (en route [ER] service remains available during all ionospheric events). At all other times, satellite-based augmentation systems must monitor the state of the ionosphere continuously and provide accurate calibration of ionospheric delays whenever the level of ionospheric perturbation does not exceed the critical threshold that serves to render the system locally unavailable.

In addition to computing IGDs and GIVEs at IGPs, the WAAS GIVE Monitor is responsible for detecting ionospheric disturbances that may generate potential hazards to WAAS users. It uses two distinct metrics in real time to characterize the state of the ionosphere, and a third metric has proven convenient as a means of ranking the magnitudes of storms on a post-processing basis. The standard $\chi^2$ goodness-of-fit parameter associated with the vertical delay estimate at a given IGP has been found to be a reliable indicator of the local level of ionospheric perturbation in the vicinity of this IGP (Walter et al., 2001).
At each IGP, WAAS employs a local irregularity detector that trips when the WAAS irregularity metric, proportional to the local goodness-of-fit parameter, exceeds a specified threshold. To characterize the state of the ionosphere as a function of time on large spatial scales (i.e., over North America), WAAS defines the ionospheric perturbation metric (IPM) to be the maximum value of the irregularity metric over the entire WAAS IGP grid. This metric provides the basis for both the WAAS Extreme Storm Detector (ESD) and the WAAS Moderate Storm Detector (MSD). The ionospheric disturbance index (IDI), an index derived from the IPM and scaling with both the instantaneous magnitude of a disturbance and its duration, has proven useful for comparing the magnitudes of various ionospheric storms. Figure 2 provides a process flow diagram that summarizes the manner in which WAAS monitors the level of ionospheric disturbance, placing in larger context several of the equations presented in this paper.

The primary objectives of this paper are: a) to describe the various disturbance detectors that have been implemented in WAAS to monitor the state of the ionosphere; b) to highlight the availability and integrity concerns that have motivated their development; c) to report how the algorithms defining these detectors have evolved over the first 18 years of WAAS’s existence; and d) to rank the largest ionospheric storms occurring over the past two solar cycles according to the magnitude of their impact on WAAS.

Prior publications covering aspects of the technology discussed here have, for the most part, been limited to conference proceedings. Thus, a secondary objective is to provide a comprehensive bibliography for accessing these publications.
Section 2, we review the algorithmic formalism that has been used throughout the life of WAAS to perform an estimation of vertical delay at IGPs. Next, in Section 3, we discuss how measurements are selected for inclusion in the fits at these IGPs. In Section 4, we explain how the fit at each IGP is used to construct a local ionospheric irregularity detector. Sections 5 and 6 concern the WAAS Extreme Storm Detector (ESD), focusing on why it is needed and how it is defined. Section 7 addresses the implementation of the WAAS Moderate Storm Detector (MSD) and the benefits it provides. Section 8 shows how WAAS data may be used to rank the magnitudes of ionospheric storms. An appendix presents a table summarizing the changes that have been implemented over the past 18 years in the values of the WAAS operational system parameters related to monitoring the level of ionospheric perturbation.

2 | ESTIMATION OF IONOSPHERIC DELAY

The manner in which WAAS monitors ionospheric storms is based upon the methodology used to estimate the vertical delay at each IGP. WAAS computes vertical delay by fitting to a plane a set of estimates of the vertical delay at locations near the IGP using a geo-statistical technique known as kriging (Blanch, 2002, 2003; Blanch & Walter, 2002; Sparks et al., 2010, 2011). Kriging is a type of minimum mean square estimator adapted to spatial data that originated in the mining industry in the 1950s. It yields a smoothed representation of a spatially distributed variable that has been sampled by measurements at irregularly spaced locations. In general, kriging incorporates into a fit the degree to which neighboring measurements of ionospheric delay are correlated and generates a linear unbiased delay estimate.

From its commissioning in July of 2003 until WAAS Follow-On (WFO) Release 3 in 2011, however, WAAS calculated the vertical delay estimate and its integrity bound at each IGP from a planar fit of vertical delay estimates assumed to be uncorrelated. Deeming measurements to be completely uncorrelated may be regarded as a limiting case of the general kriging algorithm. Compared to the planar fit model, the kriging model can be tuned to achieve a better match to the observed random structure of the vertical delays. This section summarizes the equations that determine the kriging estimate of vertical delay at each WAAS grid point. The derivation of this algorithm in greater detail is presented in Sparks et al. (2011).

The delay that a radio signal experiences as it travels through the ionosphere may be expressed (in meters) as:

$$J_{rs} = \frac{40.30}{f^2} STEC_{rs}$$  

(1)

where $f$ is the signal frequency (in second$^{-1}$), and $STEC_{rs}$ is the slant total electron content (in electrons per meter$^2$) associated with the signal (i.e., the path integral of electron density along the line-of-sight from the receiver $r$ to the satellite $s$):

$$STEC_{rs} = \int_{l_r}^{l_s} d\ell \, n_e(\mathbf{x}_{rs}(\ell)),$$  

(2)

Here $n_e$ is the ionospheric electron density, and $\mathbf{x}_{rs}(\ell)$ identifies a set of points along the ray path parameterized by the path length $\ell$. In the local Cartesian coordinate system, this path integral may be rewritten as an integral over the altitude $h$:

$$STEC_{rs} = \int_{h_r}^{h_s} dh \, F(\alpha, h)n_e(x_{rs}(h))$$  

(3)
where:

\[
F(\alpha, h) = \left[ 1 - \left( \frac{R_e \cos(\alpha)}{R_e + h} \right)^2 \right]^{-1/2}
\]  

(4)

\(\alpha\) is the elevation angle of the satellite as viewed from the receiver, \(R_e\) is the Earth radius, and the path vector \(x_{rs}(h)\) is now considered to be a function of the altitude \(h\).

Under quiet or nominal ionospheric conditions at mid-latitude over North America, the spatial variation of vertical ionospheric delay has been found to be well-modeled locally by a plane (Walter et al., 2001). Consequently, WAAS performs local vertical delay estimation by adopting the standard thin-shell model of the ionosphere (i.e., all the electrons in the ionosphere are assumed to reside in an infinitesimally thin layer at a reference height \(h_i\)). This simplification serves two useful purposes: a) it allows each slant delay measurement to be associated with a unique ionospheric pierce point (IPP) – the point at which the satellite-to-station ray path penetrates the model ionospheric shell; and b) it provides a means of estimating the vertical delay at that IPP by multiplying the measured slant delay by a simple geometric factor. Under the thin-shell approximation, the slant delay becomes:

\[
I_{rs} \approx \frac{40.30}{f^2} F(\alpha, h_i) \int_{h_i}^{h} dh n_e(x_{rs}(h))
\]

\[= F(\alpha, h_i) I_{v}
\]

(5)

where \(I_v\) denotes the equivalent vertical delay at the IPP, and \(F(\alpha, h_i)\) is the obliquity factor used to convert between slant and vertical delay.

To estimate the vertical delay at an IGP, WAAS assumes that the local spatial variation of the true vertical delay \(I_{true}\) near the IGP can be modeled by a plane in Earth-centered, Earth-fixed Cartesian coordinates centered on the IGP:

\[
I_{true}(\Delta x) = a_0 + a_{east} \Delta x \cdot \hat{e}_{east} + a_{north} \Delta x \cdot \hat{e}_{north} + r(\Delta x)
\]

(6)

where \(\Delta x = x - x_{IGP}\) is the Euclidean vector describing the distance separating the IPP from the IGP, the coefficients \(a_0\), \(a_{east}\), and \(a_{north}\) define the planar trend, and \(r(\Delta x)\) is a scalar field describing small irregularities that are superimposed on the planar trend. The random field \(r(\Delta x)\) is assumed to be zero mean and wide sense stationary (in particular, the correlation between two points is only dependent on the distance between them). The vertical ionospheric delay measured at a location \(\Delta x\) is modeled as:

\[
I_{meas}(\Delta x) = I_{true}(\Delta x) + \varepsilon
\]

(7)

in which \(\varepsilon\) represents the measurement error that is independent of the ionosphere.

At each IGP, WAAS determines the vertical ionospheric delay by constructing a planar fit of a set of slant delay measurements whose IPPs reside in the vicinity of the IGP. The estimate for the vertical delay near the IGP may be expressed as a linear combination of \(N\) equivalent vertical delays as follows:

\[
\bar{I}(\Delta x) = \sum_{k=1}^{N} w_k I_{meas}(\Delta x_k)
\]

\[= w^T \bar{I}_{IPP}
\]

(8)
where:

\[
\mathbf{I}_{IPP} = \begin{bmatrix} \bar{T}_1 & \bar{T}_2 & \cdots & \bar{T}_N \end{bmatrix}^T
\]  

(9)

is the vector of equivalent vertical delays, each of which is specified by 
\(\bar{T}_\kappa = T_{\kappa} - I_{\kappa}\text{IPP}\) with \(\Delta x_\kappa = x_\kappa - x_{IGP}\) denoting the Euclidean displacement vector of the \(\kappa\)-th IPP from the IGP, and:

\[
\mathbf{w} = \begin{bmatrix} w_1 & w_2 & \cdots & w_N \end{bmatrix}^T
\]

(10)

is a vector of weights whose elements are determined by minimizing the estimation variance:

\[
\hat{\sigma}^2 \equiv \left( \bar{I}(\Delta x) - \bar{I}(\Delta x) \right)^2
\]

subject to the constraint that the estimator be unbiased:

\[
\left\langle \bar{I}(\Delta x) \right\rangle = \left\langle I_{true}(\Delta x) \right\rangle
\]

(12)

Substituting for the expressions on each side of the constraint equation gives:

\[
\left\{ \sum_{\kappa=1}^{N} w_\kappa \left[ a_0 + a_{\text{east}} \Delta x_\kappa \cdot \hat{e}_{\text{east}} + a_{\text{north}} \Delta x_\kappa \cdot \hat{e}_{\text{north}} + r\left(\Delta x_\kappa\right) + \varepsilon_\kappa\left(\Delta x_\kappa\right) \right] \right\} = \left\{ a_0 + a_{\text{east}} \Delta x \cdot \hat{e}_{\text{east}} + a_{\text{north}} \Delta x \cdot \hat{e}_{\text{north}} + r(\Delta x) \right\}
\]

(13)

Assuming that the expectation values of the scalar field and the measurement noise vanish (i.e., that \(\left\langle r(\Delta x) \right\rangle = \left\langle r(\Delta x_\kappa) \right\rangle = \left\langle \varepsilon_\kappa(\Delta x_\kappa) \right\rangle = 0\)), Equation (13) implies that the weighting vector \(\mathbf{w}\) must satisfy three conditions independent of the values of \(a_0\), \(a_{\text{east}}\), and \(a_{\text{north}}\). By considering separately the coefficients of \(a_0\), \(a_{\text{east}}\), and \(a_{\text{north}}\) in Equation (13), these three conditions may be expressed as the three components the following vector equation:

\[
\mathbf{G}^T \mathbf{w} = \mathbf{s}
\]

(14)

where:

\[
\mathbf{G} = \begin{bmatrix}
1 & \Delta x_1 \cdot \hat{e}_{\text{east}} & \Delta x_1 \cdot \hat{e}_{\text{north}} \\
1 & \Delta x_2 \cdot \hat{e}_{\text{east}} & \Delta x_2 \cdot \hat{e}_{\text{north}} \\
\vdots & \vdots & \vdots \\
1 & \Delta x_N \cdot \hat{e}_{\text{east}} & \Delta x_N \cdot \hat{e}_{\text{north}}
\end{bmatrix}
\]

(15)

and:

\[
\mathbf{s} = \begin{bmatrix} 1 & \Delta x_{\text{IPP}} \cdot \hat{e}_{\text{east}} & \Delta x_{\text{IPP}} \cdot \hat{e}_{\text{north}} \end{bmatrix}^T
\]

(16)

By constraining the estimator to be unbiased, the estimation variance simplifies to:

\[
\hat{\sigma}^2(\mathbf{w}) \equiv \left( r(\Delta x) - \mathbf{w}^T \mathbf{r}_{\text{IPP}} + \varepsilon_{\text{IPP}} \right)^2
\]

(17)
where:

\[
\mathbf{r}_{\text{IPP}} = \begin{bmatrix}
\mathbf{r}(\Delta \mathbf{x}_1) & \mathbf{r}(\Delta \mathbf{x}_2) & \cdots & \mathbf{r}(\Delta \mathbf{x}_N)
\end{bmatrix}^T
\]  

(18)

and:

\[
\mathbf{e}_{\text{IPP}} = \begin{bmatrix}
e_1 & e_2 & \cdots & e_N
\end{bmatrix}^T
\]  

(19)

Assuming that small deviations from the planar trend are uncorrelated with the measurement error (Sparks et al., 2011), this estimation variance can be expressed as the sum of two terms representing the variances of, respectively, process noise due to the ionosphere and measurement noise originating in the observing instruments:

\[
\hat{\sigma}^2(\mathbf{w}) = \hat{\sigma}^2_{\text{process}}(\mathbf{w}) + \hat{\sigma}^2_{\text{measurement}}(\mathbf{w})
\]  

(20)

where:

\[
\hat{\sigma}^2_{\text{process}}(\mathbf{w}) = \mathbf{w}^T \cdot \mathbf{C} \cdot \mathbf{w} - 2\mathbf{w}^T \cdot \mathbf{c} + c_0
\]  

(21)

\[
\hat{\sigma}^2_{\text{measurement}}(\mathbf{w}) = \mathbf{w}^T \cdot \mathbf{M} \cdot \mathbf{w}
\]

\[
\mathbf{C} = \langle \mathbf{r}_{\text{IPP}} \mathbf{r}_{\text{IPP}}^T \rangle
\]

\[
\mathbf{c} = \langle \mathbf{r}(\Delta \mathbf{x}) \mathbf{r}_{\text{IPP}} \rangle
\]

\[
c_0 = \langle \mathbf{r}^2(\Delta \mathbf{x}) \rangle
\]

\[
\mathbf{M} = \langle \mathbf{e}_{\text{IPP}} \mathbf{e}_{\text{IPP}}^T \rangle
\]

Here \( \mathbf{M} \) is the \( N \times N \) matrix describing the covariance of the measurement noise between measurement locations, \( \mathbf{C} \) is the \( N \times N \) matrix describing the covariance of the detrended ionospheric delay between measurement locations, \( \mathbf{c} \) is an \( N \)-vector whose elements specify the covariance between the scalar field at a position \( \Delta \mathbf{x} \) near the IGP and the detrended delays at the measurement locations, and \( c_0 \) is the variance of the scalar field, assumed to be a constant independent of \( \Delta \mathbf{x} \).

The weighting matrix \( \mathbf{W} \) is defined as:

\[
\mathbf{W} = (\mathbf{C} + \mathbf{M})^{-1}
\]  

(23)

Sparks et al. (2011) show how the general solution for the weighting vector becomes:

\[
\mathbf{w} = \left[ \mathbf{W} - \mathbf{W} \mathbf{G}^T \mathbf{W} \left( \mathbf{G}^T \mathbf{W} \right)^{-1} \mathbf{G}^T \right]^{-1} \mathbf{c} + \mathbf{W} \mathbf{G} \left( \mathbf{G}^T \mathbf{W} \right)^{-1} \mathbf{s}
\]  

(24)

Thus, to compute an estimate of the vertical delay \( \widetilde{I}(\Delta \mathbf{x}) \) at or near an IGP, a set of equivalent vertical delays \( \widetilde{I}_i \) must be calculated at IPPs near the IGP and then combined according to weights \( \mathbf{w} \) that depend upon \( \mathbf{M} \), \( \mathbf{C} \), and \( \mathbf{c} \) as specified in Equation (24) using Equation (23).

The \( L1 \) frequency (1,575.42 MHz) is used to convert measurements of total electron content to delay in units of meters of \( L1 \). The IGD that WAAS actually broadcasts at an IGP is quantized by rounding the estimated vertical delay \( \widetilde{I}(0) \) upwards to the nearest increment with 0.125-meter resolution (i.e., the broadcast IGDs range from zero to a maximum valid value of 63.750 meters in increments of
0.125 meters). When the estimated vertical delay exceeds this maximum value, the IGD is set to 63.875 meters to indicate that it is not to be used, i.e., the IGD is set to the minimum of 63.875 and \( \lceil (8 \, I_\kappa) / 8 \rceil \).

Each equivalent vertical delay \( \bar{T}_\kappa \) is computed from dual-frequency measurements with the satellite and receiver interfrequency biases removed as follows:

\[
\bar{T}_\kappa = \frac{f_{L2}^2 / f_{L1}^2}{f_{L1}^2 - f_{L2}^2} \left( PR_{L2,\kappa} - PR_{L1,\kappa} \right) - B_{L1-L2,r} - B_{L1-L2,s} \frac{F(\alpha_\kappa, h_1)}{F(\alpha_\kappa, h_1)} \]  

(25)

where \( f_{L1} \) and \( f_{L2} \) are, respectively, the \( L1 (1,575.42 \text{ MHz}) \) and \( L2 (1,227.60 \text{ MHz}) \) frequencies used by the Global Positioning System; \( PR_{L1,\kappa} \) and \( PR_{L2,\kappa} \) are the smoothed, multipath-corrected pseudorange measurements (in meters) associated with the \( \kappa \)-th IPP at \( L1 \) and \( L2 \), respectively; \( B_{L1-L2,r} \) and \( B_{L1-L2,s} \) are the \( L1-L2 \) interfrequency bias estimates (in meters) associated with the measurement receiver and satellite, respectively; and \( F(\alpha_\kappa, h_1) \) is the obliquity factor associated with the \( \kappa \)-th IPP. The estimation error for each equivalent vertical delay is characterized by the variance:

\[
\bar{\sigma}_\kappa^2 = \frac{\left( f_{L2}^2 / f_{L1}^2 \right)^2 \left( \sigma_{L1,\kappa}^2 + \sigma_{L2,\kappa}^2 \right) + \sigma_{L1-L2,r}^2 + \sigma_{L1-L2,s}^2}{F^2(\alpha_\kappa, h_1)} \]  

(26)

where \( \sigma_{L1,\kappa}^2 \) and \( \sigma_{L2,\kappa}^2 \) are the variances of the code noise and multipath error on, respectively, the \( L1 \) and \( L2 \) signals of the \( \kappa \)-th IPP, and \( \sigma_{L1-L2,r}^2 \) and \( \sigma_{L1-L2,s}^2 \) are the variances of the \( L1-L2 \) interfrequency bias estimates associated with the measurement receiver and satellite, respectively. Note: the conversion of delays to units of meters in Equation (25) and the conversion of variances to units of meters\(^2\) in Equation (26) are conducted using the \( L1 \) frequency. This equation is used to construct the diagonal elements of the measurement noise matrix:

\[
M = \begin{bmatrix}
\bar{\sigma}_1^2 & \sigma_{\text{bias}, 1, 2}^2 & \ldots & \sigma_{\text{bias}, 1, N}^2 \\
\sigma_{\text{bias}, 2, 1}^2 & \bar{\sigma}_2^2 & \ldots & \sigma_{\text{bias}, 2, N}^2 \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{\text{bias}, N, 1}^2 & \sigma_{\text{bias}, N, 2}^2 & \ldots & \bar{\sigma}_N^2 
\end{bmatrix}
\]  

(27)

where:

\[
\sigma_{\text{bias}, \kappa, \zeta}^2 = \begin{cases}
0 & \text{if } \kappa \text{ and } \zeta \text{ share no receiver or satellite;} \\
\frac{\sigma_{L1-L2,r}^2}{F(\alpha_\kappa, h_1) F(\alpha_\zeta, h_1)} & \text{if } \kappa \text{ and } \zeta \text{ share a common receiver;} \\
\frac{\sigma_{L1-L2,s}^2}{F(\alpha_\kappa, h_1) F(\alpha_\zeta, h_1)} & \text{if } \kappa \text{ and } \zeta \text{ share a common satellite.}
\end{cases}
\]

The off-diagonal elements here quantify the correlation of bias errors between measurements made with common receivers or common satellites.
To model $C$ and $c$ accurately requires a knowledge of how deviations of equivalent vertical delay from planarity, evaluated at two points, become decorrelated as a function of the distance separating the two points. The justification for the form assumed for $C$ and $c$ is discussed extensively in Sparks et al. (2011). Here we simply quote the conclusion, assuming that there is no deterministic underlying trend and that the function describing the deviations is stationary:

$$C_{\kappa, \zeta} = \left( \sigma_{\text{total decorr}}^2 \right)^2 - \left( \sigma_{\text{nominal decorr}}^2 \right)^2 \exp \left( -\frac{D_{\kappa, \zeta}}{d_{\text{decorr}}} \right)$$

and:

$$c_x(\Delta x) = \left( \sigma_{\text{total decorr}}^2 \right)^2 - \left( \sigma_{\text{nominal decorr}}^2 \right)^2 \exp \left( -\frac{d_x(\Delta x)}{d_{\text{decorr}}} \right)$$

where $(\sigma_{\text{nominal decorr}}^2)^2$ is the delay covariance associated with nearly coincident IPPs, $(\sigma_{\text{total decorr}}^2)^2$ is the delay covariance associated with widely separated (uncorrelated) IPPs, $d_{\text{decorr}}$ is the characteristic decorrelation distance and $D_{\kappa, \zeta}$ and $d_x(\Delta x)$ are defined as follows:

$$D_{\kappa, \zeta} = \sqrt{\left( \Delta x_{\kappa} - \Delta x_{\zeta} \right)^T \left( \Delta x_{\kappa} - \Delta x_{\zeta} \right)}$$

and:

$$d_x(\Delta x) = \sqrt{\left( \Delta x_{\kappa} - \Delta x \right)^T \left( \Delta x_{\kappa} - \Delta x \right)}$$

In the limit that the deviations of vertical delay from planarity are completely uncorrelated, i.e., that $(\sigma_{\text{nominal decorr}}^2)^2$ matches $(\sigma_{\text{total decorr}}^2)^2$, the covariance matrix $C$ becomes diagonal, the vector $c$ vanishes, and the weighting vector $w$ simplifies to:

$$w = W G \left( G^T W G \right)^{-1} s$$

which reproduces Equation (14) when each side of the equation is pre-multiplied by $G^T$. In addition to deriving from observations a set of values for the vertical delay $T_{\kappa}$, its variance $\sigma_{\kappa}^2$, and the off-diagonal elements $\sigma_{\text{bias,} \kappa, \zeta}^2$ of the matrix in Equation (27), exercising this methodology for performing vertical delay estimation requires specification of the following model parameters: $R_e$, $h_i$, $\sigma_{\text{nominal decorr}}$, $\sigma_{\text{total decorr}}$, and $d_{\text{decorr}}$ (see Table A1 in the Appendix).

In WAAS, the Earth radius $R_e$ is set to 6,378.1363 km, the value of the Earth’s semi-major axis in the EGM96 model (Lemoine et al., 1998). The reference height $h_i$ is set to the height of the ionospheric grid (i.e., 350 km), representative of the daytime peak of the F2 electron density layer. The remaining three parameters are determined from empirical studies of the dependence of vertical delay decorrelation with distance.

A study of zeroth-order decorrelation (Hansen et al., 2000b) confirmed that the difference between two equivalent vertical delay values is well modeled locally in the CONUS mid-latitude region by a planar trend with some uncertainty about that plane. In IOC WAAS and in all subsequent releases until WFO Release 3A, the deviations of equivalent vertical delay from planarity were treated as uncorrelated. Under this assumption, a decorrelation parameter was determined from a study of first-order decorrelation (Hansen et al., 2000a), similar to the study of zeroth-order
decorrelation except that the planar trend was subtracted from each equivalent vertical delay value and the correlation of the differences was assessed. This study found that a reasonable model of decorrelation required a value of $\sigma_{\text{decorr}}^{\text{nominal}} = \sigma_{\text{decorr}}^{\text{total}}$ in the range 35–50 cm. The lower bound of 35 cm was selected for WAAS releases prior to WFO Release 3A.

A full implementation of kriging was first fielded in 2011 (WFO Release 3A), in which the values selected for the kriging parameters were, respectively, $\sigma_{\text{decorr}}^{\text{nominal}} = 0.3$ m, $\sigma_{\text{decorr}}^{\text{total}} = 1$ m, and $d_{\text{decorr}} = 8,000$ km (Pandya et al., 2012). These values remain in force in the current system. Kriging’s use of spatial decorrelation to optimize the estimation of vertical delay exhibits three advantageous properties. First, observations in the fit that are situated near the estimation point of interest are given more weight than those more distant. Second, measurements close together are given less weight individually than those at the same distance from the estimation point but isolated—a declustering effect. The estimation variance provided by kriging is a function of the geometric location: points more distant from the measurements have a larger kriging variance in accord with intuition. The third useful property of kriging is the nugget effect, whereby the correlation between two measurements is modeled as discontinuous as the distance between them vanishes. This is a heritage from kriging’s origins in the mining industry. Gold grade values can change abruptly on a very small spatial scale when samples contain gold nuggets (i.e., where the presence of a nugget prevents two observations sampled arbitrarily closely from necessarily having the same value). The nugget effect is an appropriate means of dealing with the fact that two slant delay measurements having nearly the same IPP but traversing very different paths through the ionosphere can, when converted to vertical delay using the thin-shell obliquity factor, give rise to divergent estimates for the vertical delay at that IPP.

3  |  SELECTION OF MEASUREMENTS TO BE FIT

Figure 3 shows how the mask of ionospheric grid points at which fits are performed and at which IGDs and GIVEs are broadcast has evolved over time. At each IGP, a GIVE floor value is assigned representing the minimum permissible value of the GIVE that may be broadcast. In each of the plots in Figure 3, colored squares identify the values (in meters) of the GIVE floors at the IGPs in the IGP Working Set (i.e., where IGDs and GIVEs are generally available to the user). Figure 3(a) shows the WAAS IGP mask for IOC WAAS. In LPV Release 6/7, the WAAS IGP mask was expanded around Alaska, Mexico, and the East Coast, and IGPs above 55˚N (see Figure 3[b]) were replaced with the much denser IGPs of Band 9 as specified in the WAAS Minimum Operational Performance Standards (RTCA, 2016). In LPV Release 8/9.2, IGPs were added along the coasts of Alaska and the Northeast, and others were removed along the coast of Mexico, with one more IGP being added on the East Coast in WFO Release 3A (see Figure 3[c]). Finally, 12 low-latitude IGPs were removed in WFO Release 3A1 (see Figure 3[d]). Figures 3(b), 3(c), and 3(d) show the location of the pierce point filter line (PPFL) used to exclude low-latitude measurements from fits, as discussed below.

In each of the plots in Figure 3, black squares identify IGPs near Hawaii. Nearly all of the observations that are included in fits of vertical delay at these IGPs are derived from GPS measurements recorded at the WAAS reference station (WRS) in Honolulu, Hawaii (i.e., at a station very distant from all other stations in the WAAS network). The planar estimates of vertical delay at all the other IGPs are generated using mixtures of observations recorded at multiple WAAS stations.
When a problem that serves as a source of estimation error arises at a WRS (such as a hardware bias jump or ramp), the presence of observations recorded at multiple sites allows the error to be readily detected. Such error detection is not possible, however, when estimation data are restricted to measurements at an isolated site. Thus, while WAAS broadcasts ionospheric vertical delay estimates at the IGPs marked by black squares, the corresponding GIVEs are always set to “not monitored,” and consequently no vertical delay error bars can be computed for the vertical delay at these IGPs.

To estimate the vertical delay at a given IGP from a set of measurements belonging to a given epoch, WAAS selects a measurement subset whose IPPs reside closest to the IGP in question (see Figure 4). The fit domain is defined to be the region on the ionospheric shell enclosed by a circle whose points lie a Euclidean (straight-line) distance $R_{fit}$ from the IGP. All eligible IPPs that reside within a minimum distance $R_{fit} = R_{min}$ of the IGP are included in the fit. If the number of IPPs within this radius is less than $N_{target}$, the fit radius $R_{fit}$ is extended until it encompasses $N_{target}$ points. If $R_{fit}$ attains a maximum value of $R_{max}$ without encompassing $N_{target}$ points, it is held fixed at $R_{max}$ and the fit is performed using fewer than $N_{target}$ points. However, if
the number of IPPs within $R_{\text{max}}$ is less than a specified minimum $N_{\text{min}}$, no kriging estimate is computed, and the GIVE for that IGP is set to “not monitored.”

The WAAS IGP mask is designed to identify IGPs at which broadcast IGDs and GIVEs can generally be made available to the user (i.e., IGPs where there is a high probability that, in any given epoch, the number of IPPs within a radius $R_{\text{max}}$ will be greater than or equal to $N_{\text{min}}$). The spatial distribution of fit IPPs around an IGP is characterized by two metrics. In addition to $R_{\text{fit}}$, which serves as a proxy for the IPP density and becomes smaller as $R_{\text{fit}}$ increases, the ratio of $R_{\text{centroid}}$ to $R_{\text{fit}}$, where $R_{\text{centroid}}$ is the distance from the IGP to the centroid of this distribution, reflects the degree to which the distribution of fit IPPs is uniform about the fit center IGP. This ratio, designated the relative centroid metric (RCM), grows as the spread of IPPs within the fit domain becomes increasingly non-uniform. RCM values vary from zero to one, where zero represents a very well-balanced distribution and one represents the most unbalanced distribution possible.

Other metrics for characterizing the spatial distribution of fit IPPs have been considered (see, for example, Sparks et al. [2003a, 2003b]), but $R_{\text{fit}}$ and the RCM have proven to be the most useful: $R_{\text{fit}}$ and the RCM are simple to implement, and studies of alternative metrics have shown no significant benefits that would compensate for their added complexity. As noted in Table A1 of the Appendix, the values of the IPP selection parameters have not changed since WAAS’s commissioning in 2003: $R_{\text{min}}$, $R_{\text{max}}$, $N_{\text{target}}$, and $N_{\text{min}}$ have remained set
at, respectively, 800 km, 2,100 km, 30, and 10. When kriging was implemented in WAAS, a new study (Pandya et al., 2012) concluded that none of these values needed to be modified.

Not all observations are eligible to be included in a fit. Upstream monitors reject measurements when problems with signals are detected. Also excluded are measurements corresponding to elevation angles less than 5° and all measurements recorded by receivers in Hawaii. In 2007 (LPV Release 6/7) Mexican stations were incorporated into the WAAS network. This created a need to exclude from fits another set of observations, namely, those observations recorded by Mexican stations viewing GPS satellites at low elevation in the southern sky. Ray paths associated with such observations, in particular, those ray paths that pierce the equatorial ionization anomaly, frequently sample regions of the ionosphere characterized by large mesoscale electron density gradients. Vertical delay estimation in WAAS relies on the assumption that the true vertical delay can be represented by small deviations about a planar trend. This assumption, however, cannot be expected to hold in low-latitude regions (Rajagopal et al., 2004).

Figure 5 shows the median value of the 100 largest fit residuals for vertical delay estimates at IGPs tabulated from a given set of solar cycle 24 measurements. The data used to generate this figure consist of measurements recorded during the 16 storms used to generate the solar cycle 24 component of the calendar year 2018 (Release 51-CY18) threat model (see Sparks et al. [n.d.]), none of which may be considered extreme. The brown curves in Figure 5 represent contours of constant geomagnetic latitude. Note that the IGPs showing the largest median values, the three western-most points at 15°N latitude, are the three IGPs nearest the geomagnetic equator. Some of the disparity between the residual values plotted at mid-latitude and those at low-latitude is, no doubt, due to the influence of storms.
However, if we limited our analysis to quiet-time data only, the same trend would appear due to the influence of the equatorial ionization anomaly. The vertical delay is said to exhibit curvature at low latitudes, and it becomes necessary to eliminate from all fits those measurements that sample low-latitude ionospheric regions of enhanced delay. Otherwise, WAAS would experience a loss of availability at IGPs in the southern portion of the IOC IGP mask due to increased tripping of the irregularity detectors in this region.

To identify the low-latitude measurements to be excluded from fits, a two-step process is invoked. First, stations likely to sample low-latitude volumes of enhanced delay are specified in the low-latitude WRS mask. Second, a geographic region of higher ionospheric variability is identified on the ionospheric shell: measurements associated with stations in the low-latitude WRS mask are excluded from all fits when their IPPs lie within the area directly south of two connected line segments, designated the pierce point filter line (PPFL). These line segments roughly follow contours of geomagnetic latitude and are defined by three points, whose latitude ($\lambda$) and longitude ($\phi$) coordinates are distinguished, respectively, by the subscripts $W$ (west), $M$ (middle), and $E$ (east): the first segment connects the points [$\lambda_W$, $\phi_W$] and [$\lambda_M$, $\phi_M$] and the second connects the points [$\lambda_M$, $\phi_M$] and [$\lambda_E$, $\phi_E$]. All measurements with IPPs from WRSs not listed in the low-latitude WRS mask are eligible for inclusion in a fit. A measurement from a station in the low-latitude mask will be excluded from all fits if its IPP lies south of the PPFL and the IPP longitude is between $\lambda_W$ and $\lambda_E$.

In LPV Release 6/7, the values of the coordinates $\lambda_W$, $\lambda_M$, $\lambda_E$, $\phi_W$, $\phi_M$, and $\phi_E$ were, respectively, 130˚W, 80˚W, 60˚W, 26.75˚N, 18˚N, and 18˚N (see Figure 3[b] and Appendix Table A1). When four new Mexican stations were added to the WAAS network in 2008 (LPV Release 8/9.2), a trade study (Paredes et al., 2008) to determine the optimal position of the PPFL indicated that it should be moved 5˚ southward (see Figure 3[c]). In WFO Release 3A1 (January 20, 2012), the WRS at San Juan was added to the low-latitude WRS mask, and this necessitated one more modification to the PPFL: namely, extending it 20˚ eastward (see Figure 3[d]). Thus, the position of the PPFL is currently specified by $\lambda_W$, $\lambda_M$, $\lambda_E$, $\phi_W$, $\phi_M$, and $\phi_E$ having the values, respectively, of 130˚W, 80˚W, 40˚W, 21.75˚N, 13˚N, and 13˚N.

4 DETECTION OF LOCAL MESOSCALE IRREGULARITIES

The accuracy of the estimated IGD at each IGP and the reliability of its calculated formal uncertainty in bounding the estimation error depend upon the degree to which local ionospheric behavior conforms to the assumed planar model. To ensure integrity and improve performance, the GIVE Monitor allocates a portion of the burden of protection against deviations from planarity to an irregularity detector at each IGP. For WAAS purposes, an ionospheric irregularity is defined to be any ionospheric condition that cannot be accurately represented by the planar model. Thus, a disturbance that generates a steep gradient in total electron content constitutes an irregularity only if the planar model fails to reflect accurately its behavior. Any event that renders the computed integrity confidence bounds invalid, however, must be recognized and mitigated.

The standard chi-square goodness-of-fit statistic is a well-established means for determining whether a set of measurements is consistent with a specified model.
The chi-square statistic associated with the estimation of vertical delay measurements at a given IGP (Sparks et al., 2011) is given by:

\[
\chi^2 = I_{pp}^{\text{T}} \left( W - WG \left( G^T W G \right)^{-1} G^T W \right) I_{pp} \\
\]  

(32)

If fit residuals are independent and normally distributed, the chi-square statistic evaluated at an IGP can be expected to obey a chi-square distribution (Abramowitz & Stegun, 1972). This distribution is parameterized by the number of degrees of freedom \(N\) in each fit, which, in this case, is the number of fit measurements \(N\) minus the number of model parameters evaluated (i.e., 3). Given \(N = N - 3\) and an allowable false alarm rate \(P_{fa}\), the inverse \(\chi^2\) cumulative probability distribution (Abramowitz & Stegun, 1972) may be used to calculate a threshold value \(\chi^2_{\text{threshold}}\) such that, if \(\chi^2\) exceeds this value, it is highly likely that the input measurement variances or the ionospheric model (or both) are incorrect. A value of \(\chi^2\) below this threshold, however, does not guarantee that the model is valid: if fit residuals are larger than expected—perhaps due to an assumed decorrelation that does not accurately characterize the current ionospheric state—but only by a modest amount, there remains some probability that the chi-square statistic will not exceed the threshold.

For each estimate of vertical delay at an IGP, IOC WAAS defined an irregularity metric as:

\[
\chi^2_{\text{irreg}} = \frac{\chi^2}{\chi^2_{\text{threshold}}} \\
\]

(33)

where \(\chi^2_{\text{threshold}}\) was evaluated for a cumulative probability of 99.9\%, i.e., 1 minus a false alarm rate of \(P_{fa} = 10^{-3}\). (Throughout the life of WAAS, \(\chi^2_{\text{threshold}}\) has continued to be defined using \(P_{fa} = 10^{-3}\).) This metric, however, did not take into account the possibility that measurement noise could impair the ability of the metric to identify disturbed ionospheric conditions; while measurement noise for each satellite is well-characterized simply as a function of elevation and tracking time, process noise, which is influenced by the presence of ionospheric irregularities, can be quantified only through observations contaminated with measurement noise. To be confident that the irregularity metric does not underestimate the true level of ionospheric perturbation, it is necessary to inflate the original metric by a factor depending upon the measurement noise (Blanch et al., 2003). Accordingly, in LPV Release 6/7, the \(\chi^2_{\text{irreg}}\) metric was multiplied by a factor \(R_{\text{noise}}\) greater than or equal to one (Blanch, 2003):

\[
\chi^2_{\text{irreg}} = \frac{R_{\text{noise}} \chi^2}{\chi^2_{\text{threshold}}} \\
\]

(34)

where:

\[
R_{\text{noise}} = \prod_{\kappa=1}^{N} \left( 1 + \frac{M_{\kappa\kappa}}{C_{\kappa\kappa}} \right)^{\frac{1}{N}} \\
\]

(35)

and:

\[
\check{C}_{\kappa\kappa} = C_{\kappa\kappa} - \min_{\kappa \neq \zeta} \left( C_{\kappa\zeta} \right) \\
\]

(36)
where the $M_{xx}$ are given by Equation (27) and the $\tilde{C}_{xx}$ are given by Equation (28). A full derivation of the $\chi_{\text{irreg}}^2$ metric exceeds the scope of this paper. However, one can see how the $R_{\text{noise}}$ factor accounts for the masking effect of the measurement noise: as the measurement noise increases, the ratios $M_{xx}/\tilde{C}_{xx}$ become larger and the bound on the process noise increases accordingly. This $R_{\text{noise}}$ factor is key, because the process noise $r$ and the measurement noise $\varepsilon$ behave very differently in the user error bound. While the measurement noise can be averaged down, the process noise cannot.

Figure 6 shows the distribution of irregularity metric values across the current IGP grid for three storms of increasing magnitude. All plots have been generated for the same daytime epoch, 23:00 UTC, to minimize differences due to diurnal variations. The observations from which these results are derived, designated supertruth, consist of post-processed data merged from three co-located receivers at each WAAS station. These data serve to provide a truth standard for ionospheric delay measurements. Supertruth processing effectively eliminates measurement noise, permitting the covariance $M$ to be set to zero. The WAAS undersampled ionospheric irregularity threat model is constructed from supertruth data; we present a full description of how supertruth is generated in WAAS and the Ionosphere—A Historical Perspective: Mitigating Mesoscale Irregularities (Sparks et al., n.d.).

**FIGURE 6** The value of the irregularity metric of each IGP at 23:00 UTC for (a) a weak storm on July 26, 2004; (b) a moderate storm on October 24, 2011; and (c) an extreme storm on April 6, 2000. The peak value is indicated by the magenta rectangle in each plot.
WAAS assigns to each IGP in the working set a local irregularity detector based upon $\chi^2_{irreg}$. In IOC WAAS, a detector at an IGP was considered to have tripped when its $\chi^2_{irreg}$ metric, as defined in Equation (33), rose above an irregularity metric threshold of one. The tripping of the irregularity detector causes the GIVE Monitor to set the GIVE at that IGP to a conservative value, $GIVE_{max}$ that safely bounds the maximum WAAS ionospheric estimation error ever observed. When the GIVE at an IGP is $GIVE_{max}$, vertical guidance is no longer supported by WAAS at that location.

During an actual period of ionospheric disturbance, random statistical fluctuations can cause $\chi^2_{irreg}$ to drop below the threshold. To avoid resetting the GIVE to a lower value prior to the full recovery of the ionosphere, the GIVE remains at $GIVE_{max}$ for a duration $t_{irreg, hyst}$ designated the hysteresis time, after $\chi^2_{irreg}$ falls below the detection threshold. In IOC WAAS, the hysteresis time was set to 15 minutes; from LPV Release 6/7 onward, it has been set to 30 minutes. $GIVE_{max}$ has been set to 45 meters in all WAAS releases.

LPV Release 6/7 also introduced inflation of the formal estimation variance proportional to the chi-square goodness-of-fit statistic (see WAAS and the Ionosphere—A Historical Perspective: Mitigating Mesoscale Irregularities [Sparks et al., n.d.]). This diminished the necessity of having the local irregularity detector at each IGP provide protection against estimation error arising from well-sampled irregularities. Subsequently, the principal purpose served by local irregularity detectors has been to share responsibility for shielding users from the estimation error associated with undersampled irregularities.

The overbounding error variance upon which the GIVE at an IGP is based consists primarily of the sum of two terms: the inflated formal estimation variance of the vertical delay estimated at that IGP and a term, designated $\sigma^2_{undersampled}$ representing an augmentation of this variance to protect users from the adverse influence of undersampled irregularities (see Sparks et al. [n.d.]). For each fit, $\sigma^2_{undersampled}$ is retrieved from a lookup table in the undersampled ionospheric irregularity threat model. The lookup table is constructed by tabulating the maximum fit residuals that have occurred in past ionospheric storms as a function of the two IPP distribution metrics discussed above, namely, the fit radius $R_{fit}$ and the RCM of the IPPs associated with the measurements included in the fit. Fit residuals that occur in fits where the local irregularity detector has tripped are not included in the threat model.

Determining precisely how best to allocate, between the threat model and the irregularity detectors, the burden of protection against undersampling establishes the optimal value of the irregularity metric threshold $T_{irreg, trip}$ (see Sparks et al. [n.d.]). For LPV Release 6/7, a trade study determined the optimal value of this threshold to be 2.5 (Pandya et al., 2007). When kriging was fully implemented in WFO Release 3A, the study was revisited, and the trip threshold moved to 3.0 (Pandya et al., 2012). Currently the threshold value that the $\chi^2_{irreg}$ metric in Equation (34) must exceed to trip the irregularity detector at each IGP remains $T_{irreg, trip} = 3.0$.

5 | COMPACT MESOSCALE IRREGULARITIES ARISING FROM EXTREME STORMS

To ensure the integrity of delay corrections generated by WAAS, particular attention must be paid to the impact of ionospheric storms (Datta-Barua, 2004). Such storms can dramatically increase the overall level of ionization, generating
large absolute values of total electron content (TEC) along GPS ray paths. In addition, a storm can foster the development of large spatial and temporal gradients of TEC. Perhaps most important for WAAS, extreme ionospheric storms have been observed to produce compact regions of large plasma density that persist into night. As large spatial and temporal TEC gradients become more localized, the probability that the system will detect them decreases. A compact mesoscale irregularity that escapes detection potentially becomes a major source of positioning error should it be sampled by a signal traveling toward a user.

The IOC undersampled ionospheric irregularity threat model was constructed from observations recorded during the largest storms of solar cycle 23 that occurred prior to WAAS’s commissioning on July 10, 2003. Less than four months later, however, the system confronted an even larger geomagnetic disruption, the Halloween Storms of October 29–31, 2003 (Balch et al., 2004). These storms were triggered by two intense, X-class solar flares, both of which were correlated with the release of coronal mass ejections (CMEs). CMEs are large aggregations of plasma and magnetic fields that travel at supersonic speeds away from the Sun. If the magnetic fields within the interplanetary CMEs (ICMEs) are directed southward when the ICMEs impact the Earth’s magnetosphere, solar wind energy is transferred to the magnetosphere in the form of a magnetic storm.

The first solar flare erupted at 11:10 UTC on October 28, 2003. The associated ICME traveled at over 2,200 km/s, quickly arriving at Earth in roughly 19 hours. The onset of the resulting intense geomagnetic storm occurred at 06:13 UTC on October 29, 2003, with the $D_s$ index attaining a maximum depression of -353 nT between 00:00 and 02:00 UTC on the next day (see Figure 7) and $K_p$ reaching nine (corresponding to G5 [extreme] on the geomagnetic storm space weather scale of the National Oceanic and Atmospheric Administration [NOAA]).

The recovery of the $D_s$ index that began immediately thereafter was interrupted by the arrival of the second ICME at 16:20 UTC, following the second solar flare’s eruption at 20:49 UTC October 29, 2003, and another quick transit to the Earth in approximately 19 hours. This time the $D_s$ index dropped even further to a maximum depression of -383 nT between 22:00 UTC and 24:00 UTC on October 30, 2003, and again the $K_p$ index rose to nine. At the time of these events, the extreme values of the $D_s$ index attained over these two days represented, respectively, the ninth and seventh most severe drops ever recorded.

---

**FIGURE 7** The magnitude of the $D_s$ index (nT) from October 27, 2003 through November 2, 2003. [Available at: World Data Center for Geomagnetism, Kyoto, Final $D_s$ index (1957 – 2014), http://wdc.kugi.kyoto-u.ac.jp/dstdir/]
During local daytime, the Halloween Storms generated unusually large values of absolute TEC and large TEC gradients (Komjathy, et al., 2005), but it was what occurred at night that presented WAAS with a major problem. Under nominal conditions following sunset, the recombination of free electrons with positive ions reduces the background ionization to low levels. After the sunsets of October 30 and October 31, however, plasma density structures exhibiting large TEC gradients persisted for several hours into the night. The irregularity occurring on the morning of October 31 was particularly compact, exhibiting TEC values more than 60 TECU above the background (where 1 TECU is $10^{16}$ electrons /m$^2$). This dense plasma structure occurred over northern Florida and was subsequently designated the Florida Event. A compact mesoscale irregularity of such magnitude had not been observed in nighttime prior to this event. Such a tightly packed structure created for WAAS the prospect of a consequential irregularity going unsampled by the system.

Figure 8 uses the Florida Event to illustrate how a threat to integrity due to undersampling can arise. Figure 8(a) shows the equivalent vertical delay as sampled by a set of measurements recorded at 03:10:30 UTC on October 31, 2003, by receivers that comprised the Continuously Operating Reference Stations (CORS), a network consisting at the time of over 1,100 receivers operated by the National Geodetic Survey under the auspices of NOAA’s National Ocean Service. Note the orange and red dots over northern Florida and southern Georgia indicating a compact, dense irregularity persisting into the night (the local time of the data plotted is 22:10:30 EST). Figure 8(b) plots the equivalent vertical delay at the same time for the much more sparsely-distributed 25 receivers that comprised the WAAS network in 2003. Due to the lower density of receivers, the Florida Event nearly escaped detection.

The WAAS irregularity detectors at the grid points covering Florida on this occasion did, in fact, sample the irregularity displayed in Figure 8, activating the integrity algorithms designed to protect users in this area. No user position estimate was found to lie outside the error limits derived from the integrity confidence bounds of the vertical delay estimates broadcast at each IGP (i.e., no hazardously misleading information was broadcast). The Florida Event, however, highlighted the need to revise these algorithms. Had the irregularity in question been located slightly southwest, it could have gone undetected for a significant
period of time. To establish how best to modify the integrity algorithms has required understanding the mechanisms that give rise to such threats.

The spatial and temporal extent of the Florida Event was subsequently investigated to determine, in particular, the vertical distribution of electrons within the irregularity (Datta-Barua et al., 2005, 2008). TEC measurements from Jason satellite data constrained the bulk of the electrons to altitudes below 1,300 km. Dual-frequency GPS measurements from the receiver onboard the Satelite de Aplicaciones Científicas-C (SAC-C) implied that a non-negligible fraction of the charged-particle density resided above the spacecraft’s 700-km orbit altitude, and further analysis of ground and space measurements established the altitude of the nighttime peak of the density profile to be in the vicinity of 500 km. This is consistent with analysis of the dayside TEC enhancement of this storm which demonstrated plasma uplift, specifically, a 900% increase in the electron content, above the orbit of the Challenging Minisatellite Payload (CHAMP) satellite at middle latitudes (Mannucci et al., 2005a, 2005b).

Whether the Florida Event was truly isolated over Florida or represented an enhancement of a larger structure that extended northward from the equatorial region could not be readily determined with certainty due to the sparsity of measurements to the south. (In the WAAS network, only the station at San Juan, Puerto Rico, lay to the south of Florida at the time of this phenomenon.)

The physical mechanism that causes daytime plasma uplift in the equatorial region is known as the **Fountain Effect** (Kelly, 2009). At low latitude, the zonal component of the electric field is due primarily to winds in the E region driven by tidal oscillations of the atmosphere. During the day, this electric field is directed eastward. In the equatorial region, plasma rises due to \( E \times B \) drift until pressure forces, assisted by gravity, are large enough to cause the charged particles to slide poleward along magnetic field lines and drift to higher latitudes. This movement generates, on either side of the geomagnetic equator, regions of higher plasma density known as the **equatorial ionization anomaly**. Generally, the crests of this type of anomaly occur approximately ±10° away from the geomagnetic equator.

Plasma uplift can increase dramatically when an ICME strikes the Earth’s magnetosphere and causes a geomagnetic storm. Such a storm originates when there is a southward turning of the interplanetary magnetic field perpendicular to the ecliptic. The ensuing reconnection with the Earth’s magnetic field leads to particle and energy input into the near-Earth environment at high latitudes. Electric fields of solar wind origin then quickly penetrate into the magnetosphere and the dayside equatorial ionosphere to middle and low latitudes. These **prompt penetrating electric fields** (PPEFs) point from dawn to dusk on the dayside, reinforcing the diurnal zonal electric field generated by daytime neutral winds (Tsurutani et al., 2004). In contrast to this diurnal field which attains its maximum eastward magnitude near the dawn meridian, PPEFs are thought to exhibit their eastward maxima near local noon (Tsurutani et al., 2008). Under extreme storm conditions, these PPEFs can become substantially larger than the diurnal field, significantly magnifying the fountain effect. This serves to raise plasmas to much higher altitudes than would occur in the absence of the storm, from which charged particles subsequently diffuse downward along the magnetic field toward latitudes more distant from the geomagnetic equator.

For the geomagnetic storm of October 30, 2003, the crests of the equatorial anomaly moved poleward to geomagnetic latitudes of ±28°, and the plasma uplift may well have exceeded 800 km (Mannucci et al., 2005a, 2005b). At the same time, the TEC magnitude increased sharply at middle latitudes while decreasing
near the magnetic equator. This phenomenon of augmented plasma uplift and its subsequent redistribution from low to middle latitudes under the influence of PPEFs has been termed the Daytime Superfountain Effect (Tsurutani et al., 2004, 2008).

Ionospheric plasma density at any given altitude is determined by a competition between ionization and recombination. At F-region altitudes, the dominant mechanism for removing charged particles from the ionosphere is dissociative recombination, a process in which highly energetic, unstable neutral molecules are formed when electrons attach to positively charged molecular ions that have been produced from charge exchange reactions (Kelly, 2009). The amount of positively charged molecular ions available for recombination is limited by the background density of the neutral molecules from which they are created. At higher altitudes, the exponential fall-off in the density of these background neutral molecules reduces the overall rate of charged-particle recombination.

Compact nighttime irregularities such as the Florida Event are thought to be remnants of the dramatic increases in TEC observed to occur in daytime following the onset of extreme storms (Mannucci & Tsurutani, 2017). In the Daytime Superfountain Effect, two factors operate simultaneously to produce this highly enhanced daytime TEC. First, uplift of the equatorial plasma to altitudes exceeding those realized in the normal fountain effect reduces the charged-particle recombination rate below the already diminished values experienced by this plasma at the high altitudes encountered under nominal conditions. Second, photoionization at lower altitudes vacated by this plasma continues unabated, generating throughout daytime increasing amounts of charged particles. These TEC enhancements can be expected to form preferentially in the western Atlantic/21UT sector as a result of the influence exerted on the Atlantic sector magnetic field by the offset of the geomagnetic poles from the geographic poles and by magnetic declination effects near the South Atlantic Anomaly (Foster & Erickson, 2013). Once these enhancements pass into nighttime, the recombination rate at the lower altitudes remains sufficient to eliminate, shortly after dusk, most of the ionization that has occurred there. At higher altitudes, however, the recombination rate is presumed to be so low that the plasma can survive, in some cases, for several hours even in the absence of significant photoionization.

The epoch chosen for Figure 8, 3:10:30 UTC, occurred approximately 4½ hours after sunset in Miami, Florida, on the second day of the Halloween Storms. It is instructive to examine other extreme storms for evidence of localized electron density structures extending this far into nighttime. Figure 9 shows comparable plots of CORS data for each of the five largest storms in solar cycle 23 (including a plot 4½ hours after the Miami sunset on the first day of the Halloween Storms). As noted below, the storm of March 31, 2001, was much weaker than the others; Figure 9(c) shows a low level of background ionization not much different than if no storm had occurred. In contrast, Figures 9(a) and 9(e) demonstrate that the July 15–16, 2000, and the November 20–21, 2003 storms generated persistent electron density structures, respectively, to the west and to the east of Florida, as localized and having nearly the same magnitude as the Florida Event. In Figures 9(b) and 9(d), the April 6–7, 2000 storm and the first phase of the October 29–31, 2003 storms are characterized by nighttime electron density structures that produce TEC values even larger than those of the Florida Event; however, these structures are not as localized. In short, each of the truly extreme storms of solar cycle 23 corroborate the ability of the Daytime Superfountain Effect to generate regions of moderate to high electron density that persist hours into the night.
When defining the GIVEs to be broadcast, IOC WAAS assumed that the worst threats ever to go unrecognized by the local irregularity detectors are always present. Protecting against extreme conditions that, in fact, rarely occur was an overly conservative strategy that limited availability. Furthermore, this strategy

![Figure 9](image)

**Figure 9** The equivalent vertical delay (in meters) for observations of the CORS network over the conterminous United States (CONUS): (a) April 7, 2000, 04:09:00 UTC; (b) July 16, 2000, 04:43:30; (c) April 1, 2001, 04:07:00; (d) October 30, 2003, 03:10:30 UTC; and (e) November 21, 2003, 03:00:00 UTC. The epoch chosen for each figure is roughly 4½ hours after sunset in Miami, Florida.

6 | DETECTION OF EXTREME STORMS

When defining the GIVEs to be broadcast, IOC WAAS assumed that the worst threats ever to go unrecognized by the local irregularity detectors are always present. Protecting against extreme conditions that, in fact, rarely occur was an overly conservative strategy that limited availability. Furthermore, this strategy
did not even ensure safety, as it neglected the possibility that future storms could exhibit worse threats than the worst that had been observed historically. Indeed the Halloween Storms represented just such a case. Had the Florida Event gone undetected, the protection against the effects of undersampling provided by the IOC ionospheric threat model would have been insufficient to bound all positioning errors arising from GPS signals that intersected this irregularity on their way to the user.

To defend each user from the adverse influence of compact mesoscale irregularities such as the Florida Event, LPV Release 6/7 introduced into WAAS an Extreme Storm Detector (ESD) which distinguishes extreme storms from less severe disturbances (Sparks et al., 2005). The dense structures at high altitude produced by the Daytime Superfountain Effect typically arise one to two hours or more following extreme storm onset and can persist hours into the night. The ESD identifies the onset of an extreme storm and determines the duration following storm onset over which GIVEs are set to $GIVE_{\text{max}}$ at all WAAS IGPs. By extending this duration from storm onset to well beyond the recovery period of the storm, users are protected from errors arising from a failure to sample these superfountain-generated irregularities.

The operation of the original ESD and all subsequent versions is defined in terms of the behavior of an instantaneous ionospheric perturbation metric (IPM) that reflects the level of ionospheric perturbation over the WAAS service volume as a function of time. For the moment, we will defer specifying how the IPM is defined and focus instead on how the ESD is implemented.

At any particular moment, the ESD exists in one of four states: the Nominal State, the Onset Confirmation State, the Extreme Storm State, or the Recovery Confirmation State (see Figure 10). Under quiet ionospheric conditions, the ESD resides in the Nominal State. When the level of ionospheric disturbance increases to the point where the IPM crosses the ESD onset threshold $T_{\text{ESD,trip}}$ from below, the ESD enters the Onset Confirmation State. After the IPM has remained continuously above the onset threshold for a specified ESD onset confirmation interval $t_{\text{ESD,confirm}}$, the ESD confirms the existence of an extreme storm and enters the Extreme Storm State. The ESD then remains in this state until the IPM drops below the ESD recovery threshold $T_{\text{ESD,recovery}}$, at which point the ESD enters the Recovery Confirmation State.

As a conservative measure, the recovery threshold is assigned a value less than that of the onset threshold: this defers the beginning the ESD recovery period until a downward trend for the direction of the IPM has been clearly established. When the IPM has remained below the recovery threshold continuously for a recovery confirm interval $t_{\text{ESD,recovery}}$, the ESD declares a storm clear and returns to the Nominal State. On the other hand, if the IPM moves back above the onset threshold before the end of the recovery confirmation interval, the ESD returns to the Extreme Storm State. The history of the values of the parameters used to define the irregularity detectors, the ESD, and the MSD is presented in Table A1 of the Appendix.

Prior to fielding LPV Release 6/7, two different variables and their respective logarithms, computed routinely at every IGP in the WAAS grid, were each considered as a possible basis for defining the IPM: $\chi_{\text{irreg}}^2$, $\log \chi_{\text{irreg}}^2$, IGD, and $\log \text{IGD}$. For each of these variables, the simplest candidate IPM was its maximum value over the entire grid. In addition, candidate IPMs were considered that reflected the maximum level of ionospheric perturbation over more restricted geographic regions, namely, the maximum sum of the values of each variable over IGPs within a fixed distance from each IGP in the grid. The fixed distances examined were 800 km,
1,400 km, 2,000 km, and the entire grid. Finally, IPMs based upon observed delays were also proposed (e.g., the maximum mean vertical delay at IGPs within the same fixed distances and the maximum mean vertical delay of the top 25% vertical delays at IGPs within the same fixed distances). It was found that metrics based directly upon ionospheric delay (either estimated or observed) do not distinguish extreme storms from more moderate storms as readily or reliably as the ones based upon $\chi^2_{irreg}$.

While many of the proposed IPMs derived from sums of $\chi^2_{irreg}$ performed well, little appeared to be gained by adding complexity to the simplest IPM examined. Thus, the IPM at a given epoch was defined to be the maximum value that $\chi^2_{irreg}$ attains over the entire WAAS ionospheric grid:

$$IPM = \max_D (\chi^2_{irreg})$$

where $D = \{IGP\}$ is the set of all IGPs. In each plot displayed in Figure 6, a magenta-colored box identifies the IGP at which $\chi^2_{irreg}$ reaches its maximum value in the given epoch (i.e., the IGP that determines the value of the IPM). Since $\chi^2_{irreg}$ is dimensionless, both the IPM and $T_{ESD,trip}$ are also dimensionless.
The onset threshold $T_{ESD,trip}$ must be set low enough to protect the user from events that generate plasma structures like the Florida Event, but not so low that it causes the ESD to trip often. The tripping of the ESD removes the precision approach capability for several hours. Thus, it is imperative that false alarms be held to a minimum to avoid significantly impairing system availability. For LPV Release 6/7, the value assigned to $T_{ESD,trip}$ was 50. When kriging estimation was introduced in WFO Release 3A, $\chi^2_{irreg}$ values fell by roughly 1/3, and consequently $T_{ESD,trip}$ was reduced to 32. Such values for $T_{ESD,trip}$ vastly exceed the maximum values ever observed under nominal ionospheric conditions. For both releases (as well as all subsequent ones), the ESD onset confirmation interval $t_{ESD,confirm}$ has been set to one hour. This is consistent with the observation that extreme storms are associated with intense geomagnetic storms that last several hours. Adopting such values for $T_{ESD,trip}$ and $t_{ESD,confirm}$ ensures that the ESD trips only for large, rare storms. Had the ESD been operative throughout solar cycle 23, the use of these thresholds would have caused it to trip on six occasions: April 6, 2000; July 15, 2000; March 31, 2001; October 29, 2003; October 30, 2003; and November 20, 2003. Note: the March 31, 2001 storm was considerably weaker than the other five storms listed here, and, according to the metric used to rank storm magnitudes (as discussed below), its magnitude was only slightly larger than that of several other solar cycle 23 storms that would not have tripped the ESD; the fact that it would have done so reflects how conservatively the ESD thresholds have been set. Since the implementation of the ESD in 2007, the magnitude of ionospheric disturbances has never risen to a level sufficient to trip the ESD.

Observations of past storms demonstrate that once the IPM rises above $T_{ESD,trip}$, it generally remains there for an extended period of time. The time duration over which WAAS must provide protection against the impact of an extreme ionospheric storm is defined in terms of the recovery threshold $T_{ESD,recovery}$ and recovery confirm interval $t_{ESD,recovery}$. The recovery confirm interval must be several hours: $t_{ESD,recovery}$ is set according to the need to protect the user from events such as the Florida Event. For LPV Release 6/7, the values assigned to $T_{ESD,recovery}$ and $t_{ESD,recovery}$ were 45 and 8 hours, respectively. When kriging estimation was introduced in WFO Release 3A, $T_{ESD,recovery}$ was reduced to 29. Note: the eight-hour recovery confirmation interval is highly conservative—for all storms observed to date, an eight-hour duration following the drop of the IPM below the recovery threshold has found the ionosphere in a state of greatly diminished perturbation, with no compact plasma density structures present. Furthermore, in practice, one would not expect a disturbance like the Florida Event to remain unobserved by the local irregularity detectors for anywhere close to eight hours.

In principle, a sequence of fits at a single IGP can trip the ESD. It is found in practice, however, that whenever $\chi^2_{irreg}$ crosses the onset threshold at one IGP, its value at many neighboring IGPs is near the onset threshold as well (see Figure 6). This is undoubtedly due, not only to the wide impact of the extreme storm, but also to the fact that fits at neighboring IGPs are based on many of the same observations (i.e., the domain of each fit covers many IGPs).

The black line plotted in Figure 11 shows fairly typical behavior for the IPM as a function of time, as occurred during the moderate storm over the WAAS service region on October 24, 2011. As in Figure 10, the red dashed lines at 32 and 29 represent, respectively, the extreme storm onset threshold and the extreme storm recovery threshold in effect at the time of the storm—this storm did not trip the ESD.

Figures 12(a), 12(c), 12(e), 12(g), and 12(i) display similar plots for storms that would have tripped the ESD had it been in operation when the storms occurred.
FIGURE 11 Ionospheric perturbation metric (IPM) as a function of time for the moderate storm of October 24, 2011. The area of the region in red defines the ionospheric disturbance index (IDI) used to rank the magnitude of the storm. The dashed lines represent detector thresholds as they have been defined since WFO Release 3A: the ESD onset and recovery thresholds ($T_{ESD,trip}$ and $T_{ESD,recovery}$) are in red; the moderate storm detector onset threshold ($T_{MSD,trip}$) is in magenta; and the irregularity detector trip threshold ($T_{irreg,trip}$) is in blue.

during solar cycle 23. Figures 12(b), 12(d), 12(f), 12(h), and 12(j) show the corresponding maps of the vertical delay in the epoch of each storm when the IPM attained its peak value. These maps have been generated by interpolating the IGDs computed at each IGP.

One benefit of the ESD is that it improves system availability. Like the irregularity detectors, the ESD shifts another portion of the responsibility for protection from undersampling away from the ionospheric threat model. By relying on this detector to notify the system whenever an extreme storm is present, the threat model no longer needs to cover events that occur only in the aftermath of extreme storms. The threat model is now dedicated to addressing only those threats posed by less intense storms. This, in turn, results in a reduction of the amount by which the broadcast confidence bounds must be augmented to protect against positioning error due to undersampling. Consequently, WAAS availability improves significantly relative to that achieved without an ESD.

Implementing the ESD in WAAS has implications for system maintenance and safety as well. The strategy for compiling the IOC threat model required re-computation of the threat model whenever any subsequent storm produced worse undetected threats than those already incorporated into it. The Halloween Storms of 2003 were such events. A threat model constructed from observations that include data from this storm is significantly more conservative than one that excludes data from extreme storms. The ESD obviates the necessity of regenerating the threat model whenever a new extreme storm occurs and automatically provides protection from the impact of any future extreme storms that engender plasma structures denser and more localized than those observed previously.
FIGURE 12 Ionospheric perturbation metric (IPM) as a function of time and interpolated map of vertical delay (in m) at the peak of the IPM for the five extreme storms of solar cycle 23: (a) IPM for April 6–7, 2000; (b) vertical delay map for April 6–7, 2000; (c) IPM for July 15–16, 2000; (d) vertical delay map for July 15–16, 2000; (e) IPM for March 31, 2001; (f) vertical delay map for March 31, 2001; (g) IPM for October 29–31, 2003; (h) vertical delay map for October 29–31, 2003; (i) IPM for November 20–21, 2003; and (j) vertical delay map for November 20–21, 2003. Note that the length of each x-axis in the plots of IPM vs. time varies from storm to storm.
7 DETECTION OF MODERATE STORMS

Loss of WAAS availability occurs locally when a WAAS user computes a horizontal protection level (HPL) or a vertical protection level (VPL) that exceeds, respectively, the horizontal alert limit (HAL) or the vertical alert limit (VAL) specified for a given navigation mode and level of service (RTCA, 2016). Both the HPL and the VPL are dependent upon the broadcast GIVE, the VPL having a very strong dependence. On a nominal day, WAAS availability for the LPV level of service is generally 100% throughout most of the conterminous United States (CONUS). Prior to WAAS Release 46–CY16, however, availability was routinely found to be less than 100% along the coast of California. For example, in the first quarter of 2013, the average availability for the LPV level of service (with a 50-meter VAL) dropped to 99.6% just off the California coast. For the LPV 200 level of service (with a 35-meter VAL), the drop was even more dramatic, falling as low as 97% a short distance from the coast.

To improve availability, especially along the California coast, WAAS Release 46–CY16 introduced into the ionospheric threat model a dependence upon the regional level of ionospheric perturbation (Sparks & Altshuler, 2014). Prior to this release, the contribution of the threat model to the GIVE at a given IGP depended only upon the spatial distribution of the measurements used to estimate the IGD at that IGP. WAAS Release 46–CY16 separated the ionospheric threat model into two branches: the disturbed-time branch and the quiet-time branch (see Sparks et al. [n.d.]). The former, identical to the WFO Release 3A threat model, was used to define GIVEs whenever the ionosphere was determined to be sufficiently disturbed; the latter provided enhanced system availability by reducing the magnitudes of the GIVEs broadcast under nominal conditions.

To identify a degree of regional ionospheric perturbation that is less than extreme, WAAS Release 46–CY16 implemented a Moderate Storm Detector (MSD) based upon the same ionospheric perturbation metric and detection algorithm that are employed by the ESD. The MSD differs from the ESD only in using a lower disturbance detection threshold and a shorter confirmation interval. When constructing the quiet-time branch of the threat model, threats that cause the MSD to trip are eliminated from the tabulation of this branch in a manner analogous to how the tripping of the ESD serves to eliminate threats arising from extreme storms. Without these threats, the quiet-time branch of the threat model is much less conservative than the disturbed-time branch—the magnitudes of the tabulated values of $\sigma^2_{\text{undersampled}}$ are significantly reduced. In WAAS operations, the consequences of an MSD trip are very different from those of an ESD trip. The tripping of the ESD effectively shuts down vertical guidance by WAAS for several hours (a minimum of $t_{\text{ESD, recovery}}$ hours). When the MSD trips, the system simply inflates the values of the GIVEs broadcast using the more conservative values of $\sigma^2_{\text{undersampled}}$ retrieved from the disturbed-time branch of the ionospheric threat model.

To determine how often the MSD trips for reasonable values of the onset threshold $T_{\text{MSD, trip}}$ and the confirmation interval $t_{\text{MSD, confirm}}$, a study investigated MSDs defined by a range of values for each of these parameters, namely, thresholds of 5, 10, and 15, and confirmation intervals of 10, 20, and 30 minutes (Sparks & Altshuler, 2014). In this study $T_{\text{ESD, trip}}$ and $t_{\text{ESD, confirm}}$ were 32 and one hour, respectively, and $T_{\text{MSD, recovery}}$ and $t_{\text{MSD, recovery}}$ were assigned values that matched $T_{\text{MSD, trip}}$ and $t_{\text{MSD, confirm}}$, respectively. The study analyzed WAAS observations recorded over the period January 6, 2011, to November 13, 2013. For various combinations of parameters, the study computed the total amount of time that the MSD would have been in a tripped state (i.e., in a Storm State or in a Recovery Confirmation State)
had the MSD been operational in WAAS over the period in question. The computations were performed using Ionospheric Slant TEC Analysis using GNSS-based Estimation (*IonoSTAGE*), a software package developed at the Jet Propulsion Laboratory (Sparks, 2018).

Figure 13 reveals that, in all cases, the MSD would have rarely been in a tripped state. Even for an MSD with the highest trip rate (onset threshold = 5, confirmation interval = 10 minutes), the fraction of time it exists in a tripped state over the duration of the data set is less than 0.0006. Given that the current solar cycle has been unusually quiet, these results may be somewhat optimistic with regard to future solar cycles. Nevertheless, the implementation of the MSD can be expected to enable use of the quiet-time branch of the threat model to compute GIVEs under most circumstances.

For *Release 46–CY16*, the MSD onset threshold \( T_{MSD,trip} \) and confirmation interval \( t_{MSD,confirm} \) were set to 10 and 10 minutes, respectively. \( T_{MSD,recovery} \) and \( t_{MSD,recovery} \) were assigned matching values, respectively. Figure 11 displays the behavior of the IPM for a storm that would trip an MSD so defined but not the ESD. Had the ESD and the MSD been operational during the occurrence of the storms displayed in Figure 6, storms (b) and (c) would have both tripped this MSD, but only the latter would have also tripped the ESD.

![Figure 13](image.png)

**FIGURE 13** The number of minutes that the MSD is in a tripped state for the data interval January 6, 2011, to November 13, 2013

### 8 RANKING STORM MAGNITUDES

Prior to each upgrade of WAAS, it is necessary to conduct a thorough verification of the safety and integrity of the algorithms for aircraft navigation that have been modified. Included in such analysis is an assessment of the relationship of the magnitude of a given ionospheric storm and its impact on the ionospheric threat model. To perform this analysis, it has proven useful to devise an *ionospheric disturbance index* (IDI), a scalar metric that characterizes the level of ionospheric perturbation over a given duration.
Since the magnitude of a storm tends to increase with the magnitude of the IPM, an obvious IDI candidate is the maximum instantaneous value of the IPM over the course of the storm. This choice, however, fails to account for the temporal duration of the storm. A more useful IDI is the area that is both under the IPM curve and above the irregularity threshold over the duration in question:

\[
IDI = \int_{t_i}^{t_f} \left( IPM(t) - T_{irreg,trip} \right) dt
\]

(38)

where \([t_i, t_f]\) is the range of time over which the IPM exceeds the irregularity threshold (see the red-colored area in Figure 11).

Such an index exhibits several useful properties. It clearly scales with both the maximum instantaneous IPM magnitude and the storm duration. It vanishes over intervals where the irregularity metric at each IGP remains below the irregularity trip threshold. The time interval over which it is defined is arbitrary. As long as the time interval associated with a given storm includes all subintervals over which the IPM is above the irregularity threshold, the value of the IDI does not depend upon precisely how the duration of the storm is specified. Finally, the IDI of a time duration composed of a succession of adjacent subintervals can be computed by summing over the IDIs associated with the individual subintervals (appending a quiet subinterval to a given interval does not alter the value of the IDI).

Ionospheric storms can be ranked according to their IDIs. Figure 14 shows the magnitude of the IDI for each storm whose data have contributed to the generation of a component of the WAAS Release 51-CY18 ionospheric threat model (see Sparks et al. [n.d.]): (a) storms of solar cycle 23, and (b) storms of solar cycle 24. The largest storms of solar cycle 23, ranked according to IDI, match those that trip the ESD. Note that all the storms of solar cycle 24 rank among the weakest storms displayed for solar cycle 23 (i.e., note the different scales used to define the x-axis in each plot). Not one of these storms has been sufficiently intense to trip the ESD (the impact of the storms from 2011 through mid-2013 on WAAS navigation has been discussed elsewhere [Datta-Barua et. al., 2014]). When ranked by magnitude, Figure 14(a) shows that the storm order remains essentially unchanged after converting the method of delay estimation from the original planar fit approach to

FIGURE 14 The ionospheric disturbance index (IDI) for various storms occurring in (a) solar cycle 23 (defining the IDI using kriging estimation [WFO Release 3A] and planar fit estimation [Release 8/9.2]) and (b) solar cycle 24 (defining the IDI using kriging estimation). Note the different scales in each plot used for the x-axis.
kriging. This indicates the robustness of the definition of the IDI based upon the IPM; when other metrics were used to define the IDI, switching methods of estimation caused the storm ranking to vary considerably.

9 | SUMMARY

Ionospheric storms threaten the accuracy and integrity of ionospheric delay correction data broadcast by WAAS. This paper has described how WAAS monitors the level of ionospheric perturbation over North America and how its approach to mitigating threats posed by ionospheric storms has evolved since its commissioning in 2003. In particular, this paper has focused on the implementation and benefits of two detectors that were absent from the original incarnation of WAAS: the Extreme Storm Detector and the Moderate Storm Detector, both of which are based upon a system-wide ionospheric perturbation metric derived from the local mesoscale irregularity metrics evaluated at the ionospheric grid points that comprise the WAAS grid. An additional use of the ionospheric perturbation metric has been to provide a basis for defining an ionospheric disturbance index that serves as a convenient means of ranking the magnitudes of storms on a post-processing basis. From a WAAS perspective, there have been no extreme storms in solar cycle 24; since its implementation in 2007, the WAAS Extreme Storm Detector has never tripped.

In conclusion, it should be noted that the methodology described here has successfully contributed to the fact that, to date, no WAAS user position estimate has ever been found to lie outside the integrity error bounds derived from the data broadcast by WAAS. WAAS continues to avoid broadcasting what is technically designated hazardously misleading information.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. Bruce Tsurutani for reviewing the discussion of the superfountain effect and providing suggestions as to how to it might be improved. The authors also wish to thank Dr. Anthea Coster and Dr. William Rideout for providing data from the Madrigal Database of the MIT Haystack Observatory which have been used to plot Figures 8(a) and 9(a)-(e). GPS TEC data products and access through the Madrigal-distributed data system are provided to the community by the Massachusetts Institute of Technology under support from US National Science Foundation grant AGS-1952737. Data for the TEC processing is provided from the following organizations: UNAVCO, Scripps Orbit and Permanent Array Center, Institut Geographique National, France, International GNSS Service, The Crustal Dynamics Data Information System (CDDIS), National Geodetic Survey, Instituto Brasileiro de Geografia e Estatística, RAMSAC CORS of Instituto Geográfico Nacional de la República Argentina, Arecibo Observatory, Low-Latitude Ionospheric Sensor Network (LISN), Topcon Positioning Systems, Inc., Canadian High Arctic Ionospheric Network, Institute of Geology and Geophysics, Chinese Academy of Sciences, China Meteorology Administration, Centro di Ricerche Sismologiche, Système d’Observation du Niveau des Eaux Littorales (SONEL), RENAG : REseau NAtional GPS permanent, GeoNet - the official source of geological hazard information for New Zealand, GNSS Reference Networks, Finnish Meteorological Institute, SWEPOS - Sweden, Hartebeesthoek Radio Astronomy Observatory, TrigNet Web Application, South Africa, Australian
The research of Lawrence Sparks was performed at the Jet Propulsion Laboratory/California Institute of Technology under contract to the National Aeronautics and Space Administration and the Federal Aviation Administration. The research of Eric Altshuler was performed at the Sequoia Research Corporation under contract to Zeta Associates Incorporated and the Federal Aviation Administration. The research of Nitin Pandya was performed at Raytheon Company under contract with the Federal Aviation Administration. The research of Juan Blanch and Todd Walter was performed at Stanford University under cooperative agreements with the Federal Aviation Administration.

REFERENCES


APPENDIX A: RELEASE HISTORY OF STORM-RELATED OPERATIONAL SYSTEM PARAMETERS

Table A1 shows how the values of storm-related operational system parameters have evolved since the commissioning of the WAAS on July 10, 2003. For each WAAS release in which such parameters were modified, the first five rows show, respectively, the date of the release, the number of receiver stations in the network, the name of the IGP mask, the number of IGPs in the IGP mask, and whether the estimation technique assumed the fit measurements to be correlated. The parameters listed in the remaining rows are defined as follows:

- \( h \): the ionospheric shell height;
- \( \sigma_{\text{decorr}}^{\text{nominal}} \): the ionospheric delay decorrelation constant at the origin (0 km);
- \( \sigma_{\text{decorr}}^{\text{total}} \): the ionospheric delay decorrelation constant at large distances;
- \( d_{\text{decorr}} \): the characteristic ionospheric delay decorrelation distance;
- \( R_{\text{max}}, R_{\text{min}} \): the maximum and minimum fit radii;
- \( N_{\text{target}}, N_{\text{min}} \): the targeted number and minimum number of fit points;
- \( \lambda_W, \phi_W, \lambda_M, \phi_M, \lambda_E, \phi_E \): the coordinates of the PPFL;
- \( P_{\text{fa}} \): the allowed false alarm rate;
- \( R_{\text{noise}} \): the inflation factor to account for measurement noise;
- \( R_{\text{irreg}} \): the inflation factor to account for ionospheric and statistical uncertainty in the \( \chi^2 \) statistic (see Sparks et al. [n.d.])
- \( T_{\text{irreg,trip}} \): the irregularity detector trip threshold;
- \( t_{\text{irreg,hyst}} \): the irregularity detector hysteresis interval;
- \( GIVE_{\text{max}} \): the maximum broadcast GIVE;
- \( T_{\text{ESD,trip}} \): the ESD trip threshold;
- \( t_{\text{ESD,confirm}} \): the ESD onset confirm interval;
- \( T_{\text{ESD,recov}} \): the ESD recovery threshold;
- \( t_{\text{ESD,recov}} \): the ESD recovery confirm interval;
- \( T_{\text{MSD,trip}} \): the MSD trip threshold;
- \( t_{\text{MSD,confirm}} \): the MSD onset confirm interval;
- \( T_{\text{MSD,recov}} \): the MSD recovery threshold; and
- \( t_{\text{MSD,recov}} \): the MSD recovery confirm interval
TABLE A1
A history of updates to storm-related operational system parameters.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of stations</td>
<td>25</td>
<td>34</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>WAAS IGP mask</td>
<td>002</td>
<td>003</td>
<td>004</td>
<td>005</td>
<td>006</td>
<td>007</td>
</tr>
<tr>
<td>Number of IGP</td>
<td>190</td>
<td>321</td>
<td>317</td>
<td>318</td>
<td>306</td>
<td>306</td>
</tr>
<tr>
<td>Fit measurements</td>
<td>uncorrelated</td>
<td>uncorrelated</td>
<td>uncorrelated</td>
<td>correlated</td>
<td>correlated</td>
<td>correlated</td>
</tr>
<tr>
<td>$h_i$</td>
<td>350 km</td>
<td>350 km</td>
<td>350 km</td>
<td>350 km</td>
<td>350 km</td>
<td>350 km</td>
</tr>
<tr>
<td>$\sigma_{\text{nominal decorr}}$</td>
<td>35 cm</td>
<td>35 cm</td>
<td>35 cm</td>
<td>30 cm</td>
<td>30 cm</td>
<td>30 cm</td>
</tr>
<tr>
<td>$\sigma_{\text{total decorr}}$</td>
<td>35 cm</td>
<td>35 cm</td>
<td>35 cm</td>
<td>1 m</td>
<td>1 m</td>
<td>1 m</td>
</tr>
<tr>
<td>$d_{\text{decorr}}$</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>8000 km</td>
<td>8000 km</td>
<td>8000 km</td>
</tr>
<tr>
<td>$R_{\text{max}}, R_{\text{min}}$</td>
<td>2100, 800 km</td>
<td>2100, 800 km</td>
<td>2100, 800 km</td>
<td>2100, 800 km</td>
<td>2100, 800 km</td>
<td>2100, 800 km</td>
</tr>
<tr>
<td>$N_{\text{target}}, N_{\text{min}}$</td>
<td>30, 10</td>
<td>30, 10</td>
<td>30, 10</td>
<td>30, 10</td>
<td>30, 10</td>
<td>30, 10</td>
</tr>
<tr>
<td>$\lambda_{\text{EW}}, \phi_{\text{EW}}$</td>
<td>NA</td>
<td>130°W, 26.75°N</td>
<td>130°W, 21.75°N</td>
<td>130°W, 21.75°N</td>
<td>130°W, 21.75°N</td>
<td>130°W, 21.75°N</td>
</tr>
<tr>
<td>$\lambda_{\text{GM}}, \phi_{\text{GM}}$</td>
<td>NA</td>
<td>80°W, 18°N</td>
<td>80°W, 13°N</td>
<td>80°W, 13°N</td>
<td>80°W, 13°N</td>
<td>80°W, 13°N</td>
</tr>
<tr>
<td>$\lambda_{\text{EP}}, \phi_{\text{EP}}$</td>
<td>NA</td>
<td>60°W, 18°N</td>
<td>60°W, 13°N</td>
<td>40°W, 13°N</td>
<td>40°W, 13°N</td>
<td>40°W, 13°N</td>
</tr>
<tr>
<td>$P_{\text{fa}}$</td>
<td>$10^{-3}$</td>
<td>$10^{-3}$</td>
<td>$10^{-3}$</td>
<td>$10^{-3}$</td>
<td>$10^{-3}$</td>
<td>$10^{-3}$</td>
</tr>
<tr>
<td>$R_{\text{noise}}$</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$R_{\text{reg}}$</td>
<td>static</td>
<td>dynamic</td>
<td>dynamic</td>
<td>dynamic</td>
<td>dynamic</td>
<td>dynamic</td>
</tr>
<tr>
<td>$T_{\text{reg, trip}}$</td>
<td>1</td>
<td>2.5</td>
<td>2.5</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>$t_{\text{reg, hyst}}$</td>
<td>15 min</td>
<td>30 min</td>
<td>30 min</td>
<td>30 min</td>
<td>30 min</td>
<td>30 min</td>
</tr>
<tr>
<td>$GIVE_{\text{max}}$</td>
<td>45 m</td>
<td>45 m</td>
<td>45 m</td>
<td>45 m</td>
<td>45 m</td>
<td>45 m</td>
</tr>
<tr>
<td>$T_{\text{ESD, trip}}$</td>
<td>NA</td>
<td>50</td>
<td>50</td>
<td>32</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>$t_{\text{ESD, confirm}}$</td>
<td>NA</td>
<td>1 hr</td>
<td>1 hr</td>
<td>1 hr</td>
<td>1 hr</td>
<td>1 hr</td>
</tr>
<tr>
<td>$T_{\text{ESD, recovery}}$</td>
<td>NA</td>
<td>45</td>
<td>45</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>$t_{\text{ESD, recovery}}$</td>
<td>NA</td>
<td>8 hr</td>
<td>8 hr</td>
<td>8 hr</td>
<td>8 hr</td>
<td>8 hr</td>
</tr>
<tr>
<td>$T_{\text{MSD, trip}}$</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>$t_{\text{MSD, confirm}}$</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>10 min</td>
<td>10 min</td>
</tr>
<tr>
<td>$T_{\text{MSD, recovery}}$</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>$t_{\text{MSD, recovery}}$</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>10 min</td>
<td>10 min</td>
</tr>
</tbody>
</table>

How to cite this article: Sparks, L., Altshuler, E., Pandya, N., Blanch, J., & Walter, T. (2022) WAAS and the ionosphere – a historical perspective: Monitoring storms. NAVIGATION, 69(1). https://doi.org/10.33012/navi.503