

Bring the Noise: Three Rules for Improving Thermospheric Density Retrieval From LEO POD Data

Charles Constant

University College London

Shaylah M. Mutschler

Space Environment Technologies

Santosh Bhattarai

University College London

Marcin Pilinski

University of Colorado Boulder, Laboratory for Atmospheric and Space Physics

ABSTRACT

Thermospheric drag represents one of the largest sources of operational uncertainty for Low Earth Orbit (LEO) operations today. This uncertainty translates into degraded operational products and ultimately a worse characterisation of the future risk of collision between space objects. This problem is particularly acute during geomagnetic storms, where density at a given altitude may rise by up to an order of magnitude in a matter of hours.

Accurate, timely knowledge of current and future thermospheric density in LEO is increasingly critical for the Space Situational Awareness (SSA) and satellite operations communities as the space object population continues to grow. Available methods of providing density estimates fall short of the community's needs on several counts (e.g. accuracy, latency, and licensing restrictions). As a result, interest has grown in leveraging Precise Orbit Determination (POD) data streams generated by the growing number of LEO spacecraft to enable a step change in the quality and quantity of data available to drive the next generation of assimilative thermospheric density models.

One frequently used type of method for using satellite tracking data to invert for thermospheric density along orbital trajectories is called the Energy Dissipation Rate (EDR) method. This approach relies on converting state vectors to orbital energy and translating the fluctuations in energy into estimates of thermospheric density.

We quantify how four factors influence the performance of density retrievals using the EDR method on POD data: POD noise level, drag acceleration the satellites are exposed to, time interval over which the estimation is carried ("fit-span"), and the number of spacecraft used simultaneously to generate observations.

We simulate the orbits of satellites along Starlink-like trajectories over the month of April 2023, using HASDM as the density model. To improve on the commonly used white-noise representation for POD noise, we propose and implement a higher-fidelity POD noise model that incorporates anisotropy, autocorrelation, and geometric dilution of precision (GDOP) modulation, reflecting more realistic performance bounds. We test this model at three noise levels, informed by values reported in the literature. We then modify the semi-major axis of the orbits to sample ten drag different levels, vary the fit-span over which the density is inverted from one to five orbits, and measure the effect of averaging simultaneous observations from up to 18 equally spaced co-orbital satellites.

Three clear rules of thumb emerge:

1. **Fit-span:** Doubling the observation window roughly halves the RMS percentage error.
2. **Satellite-count rule:** For satellites distributed along a single orbital track, adding satellites to the average reduces error following an inverse power law: $\sim N^{-0.6}$ for unassimilated retrievals and $\sim N^{-0.25}$ for averages of individually assimilated estimates.
3. **Drag rule:** Density retrieval performance follows a logistic-like response to drag: for weak signals ($< 2 \times 10^{-8} \text{ ms}^{-2}$) improvements are negligible; in a transition band ($2 \times 10^{-8} - 2 \times 10^{-7} \text{ ms}^{-2}$) each $\sim 3\times$ increase in drag yields $\sim 50\%$ error reduction; beyond 10^{-6} ms^{-2} returns diminish. This pattern is consistent across all three noise levels studied.

When the number of contributing satellites is sufficiently large or the POD data sufficiently low-noise, a simple mean of unassimilated retrievals can match or exceed the average of individually assimilated estimates. Because this result arises from a per-satellite assimilation scheme, it should be interpreted as a lower bound on assimilative performance; more advanced multi-satellite frameworks are expected to perform better. Assimilation remains essential for providing four-dimensional density estimates.

Overall, the results provide a first sampling of the density-retrieval error space and a compact analytical approximation of how errors scale with POD noise, drag regime, fit-span, and satellite count. These results aim to support the development of the next generation of assimilative thermospheric density models.

1. INTRODUCTION

The rapid expansion of Low Earth Orbit (LEO) satellite constellations over the past half-decade has fundamentally altered both the promise and the risk profile of the near-Earth environment. With over ten thousand spacecraft already in operation and certain projections suggesting a further order-of-magnitude increase by the early 2030s [2, 20], the risk of collisions and the attendant threat of a cascading Kessler-type event have escalated [27]. Effective risk mitigation depends heavily on a key operational component: accurate, high-cadence, nowcasting and forecasting specifications of thermospheric mass density [42, 29, 56, 38].

At the same time, the proliferation of hardware in space presents an unprecedented data-gathering opportunity, one which offers the potential to offset the very risks these constellations pose. Each satellite in LEO can in principle serve as a moving sensor, indirectly sampling the thermosphere along its orbital path [18, 47, 39]. At the current scale of deployment, and with even greater numbers on the horizon, the space-weather community stands poised to achieve a transformation in empirical density modelling akin to the shift in real-time traffic forecasting enabled by smartphone navigation apps. Just as millions of mobile devices now provide continuous, high-resolution data to map congestion and predict delays with remarkable accuracy, a dense network of LEO satellites could deliver the spatio-temporal coverage required to unlock the next generation of thermospheric density models.

Realising this opportunity, however, depends on the type of tracking data available. Operational satellites usually actively cooperate in the tracking process by releasing Precise Orbit Determination (POD) products, state vectors, or other telemetry; we refer to these as cooperatively tracked objects. Others provide no such data, and their orbits must be reconstructed from external observations such as radar or optical measurements, most widely disseminated in the form of TLEs; these are uncooperatively tracked objects. Both categories enable thermospheric density estimation, but with differences in accuracy, latency, and ultimately operational value.

To enable the step-change in modelling promised by large constellations, the community needs a clear understanding of how the increasing pool of potential cooperative tracking data can be exploited, their respective limitations, and strategies to address them. Research in this direction is emerging, and the present paper aims to contribute by generating results directly relevant to the primary stakeholders in model ideation and development—namely assimilative model developers and satellite operators. Providing empirically derived rules of thumb on factors such as required POD quality, number of satellites, and the time span of observations is essential to inform sound decision-making in this domain.

1.1 State of Thermospheric Density Modelling

At present, three broad thermospheric density model families serve the operations community:

1. **Data-driven climatology models (non-circulation / non-fluid-solving):** These are typically parametric in nature and attempt to represent the statistically most probable state of the thermosphere for a given set of inputs. They do not solve the governing fluid equations, and as such cannot reproduce the thermosphere's evolution as a coupled dynamical system. Models such as NRLMSISE-00 underpin many operational workflows, and remain widely used due to their low computational cost, favourable usage licenses, and thorough characterization within the literature [38, 6, 45].
2. **First-principles General Circulation Models (circulation / fluid-solving):** By solving the coupled physical equations of the upper atmosphere (e.g. Navier–Stokes), General Circulation Models are capable of reproducing physical features that emerge from system dynamics, across a variety of spatial and temporal scales. Examples include terminator waves, post-storm over-cooling, and the equatorial temperature and wind anomaly. However, the computational burden of these models is substantial, and they remain prohibitive for many stakeholders (e.g.

satellite operators). They are most often run by specialist space weather organisations (e.g. National Oceanic and Atmospheric Administration Space Weather Prediction Center (NOAA SWPC) and the Met Office Space Weather Operations Center (MOSWOC)). Machine-learning surrogates are in development [28] and may provide a valuable compromise between physical fidelity and efficiency, echoing recent successes in terrestrial weather modelling [26, 7].

3. **Assimilative models:** In terrestrial weather modelling this approach is the gold standard [21]. These models blend observations to tie a model (data-driven or first-principles) back to reality. The core premise is that, where observations are too sparse or erroneous, their combination with a high-quality model provides an answer closer to the truth. While this framework is the backbone of terrestrial numerical weather prediction, it is still in its infancy for the space weather community- HASDM (operational within the U.S. Government) is perhaps the most well-known model of this variant, with alternatives being made operational such as the recent deployment of the Advanced ENsemble electron density Assimilation System (AENeAS) [15] at the MOSWOC.

1.2 Observation Bottleneck

The skill of any assimilative model is bounded, on one hand by the quality of the data available and, on the other, by the quality of the underlying model it is assimilating these observations into. In practice, three attributes of the assimilated observations drive the operational performance of an assimilative scheme: volume, accuracy, and latency. State-of-the-art systems such as HASDM rely on calibration data from a network of 80+ uncooperatively tracked satellites. While this approach provides best-in-class performance, these observations are sparse in both space and time relative to the potential pool of cooperative data streams that could be harnessed from current and planned LEO constellations.

This sparsity introduces several shortcomings. Firstly, the temporal coverage may be insufficient to resolve the rapid density enhancements and relaxations that occur during geomagnetic storms, where increases of nearly an order of magnitude can unfold within only a few hours [57, 41]. Second, the set of calibration objects occupies orbits that are not necessarily distributed across regions where satellite traffic is most heavily concentrated. As a result, the regions of the thermosphere where accurate density knowledge is most operationally critical may not be adequately sampled. By contrast, POD data have the potential to address these limitations: they offer a source of high-volume, high-cadence measurements within the most densely populated orbital shells. Moreover, thanks to their higher accuracy and reduced noise, the density observations that cooperative data can offer enable faster processing and dissemination than uncooperative data.

1.3 Density Observation Sources

The oldest and most widely used density proxy stems from uncooperative tracking data. Perhaps the most ubiquitous example is the Two-Line Element (TLE) set, in continuous operational use since the late 1950s [23, 13, 44]. TLEs encode a satellite's ballistic coefficient along with changes in the mean motion that are sensitive to drag and neutral density, providing an orbit-averaged measure of density. While globally available, the data are noisy, low-cadence, and integrated over unspecified time periods (typically hours to days), limiting their value for timely, high-resolution modelling.

In contrast, high-resolution accelerometer data from missions such as CHAMP and GRACE-FO have underpinned thermosphere research for more than two decades [30, 52]. These dedicated missions furnish high-accuracy, high-temporal-resolution (often ≤ 30 s) measurements of non-gravitational forces, enabling precise local density retrievals that represent the gold standard for validation across the literature [52, 38, 31]. Operationally, however, the computationally intensive processing, the latency of data production (often days to weeks), and the very limited number of satellites restrict their utility for real-time model assimilation.

Between these extremes lies the use of precise orbit determination (POD) solutions derived from on-board GNSS receivers. Over the past 10–15 years, this approach has been explored as an intermediate solution, offering higher cadence than uncooperative measurements and lower latency than accelerometer data [10, 48, 17, 24]. A key advantage is that POD-derived densities are effectively a data of opportunity: they exploit GNSS hardware already carried on board spacecraft for navigation, avoiding the cost and complexity of dedicated accelerometers. This results in lower mission hardware costs, while still enabling density retrievals at useful spatio-temporal scales. Until recently, reported use of POD data was confined to scientific studies, though various organisations are now considering their integration into operational pipelines. Wider adoption is currently constrained by accessibility and dissemination— only a small

number of operators release near-real-time POD data for LEO satellites [50, 40]. Nevertheless, if these barriers can be overcome, POD data represent a scalable and cost-effective pathway toward the dense, low-latency observational network required for next-generation thermospheric density modelling.

1.4 GNSS-Derived Density: A New Opportunity

Mega-constellations carry the potential to deliver the very asset assimilative frameworks need for a paradigm-shifting change: abundant, high-resolution, high-cadence measurements. Because a POD solution embeds the cumulative drag experienced by a satellite along its trajectory, if it is of high enough quality, it has the potential to be inverted into a time series of densities that ought to outperform TLE-based density inversions in terms of spatial and temporal resolution. The fidelity of those densities, however, rests on the signal-to-noise ratio of the drag embedded in the POD data. This signal-to-noise ratio is itself dependent on several factors: the intrinsic accuracy of the POD solution (e.g. susceptible to GNSS receiver noise, processing strategy, attitude knowledge, force-mismodelling) and the strength of the drag signal. Satellites orbiting at lower altitudes or through geomagnetically disturbed conditions, where atmospheric drag presents a higher disturbing acceleration, will result in improved density retrieval quality, whereas high-altitude, quiet-time arcs tend to yield weaker, noisier estimates [48].

Development efforts towards any future models which aim to leverage this rich set of signals-of-opportunity must therefore account for these different sources of error if POD-derived observations are to realise their full potential.

1.5 Scope of This Study

Effective use of POD data streams for thermospheric density estimation requires a clear understanding of the variables that govern retrieval performance. Both satellite operators and model developers must anticipate, at least approximately, the accuracy attainable from different data sources to design systems that meet operational requirements. Although accuracy and uncertainty thresholds for orbit prediction and conjunction assessment are frequently discussed in the literature [1, 46, 5, 12], analyses of the specific factors driving density-retrieval error remain limited [48]. Establishing these relationships is essential for propagating performance requirements back to data quality, assimilation parameters, and satellite sampling strategies. This study therefore provides a quantitative foundation intended both as a blueprint for further analyses and as a set of practical rules of thumb to guide the design of next-generation density assimilation systems.

We address this objective through controlled simulations that quantify how orbit-averaged density-retrieval error, obtained via the EDR method applied to POD trajectories, depends on four principal variables: (1) the noise characteristics of the POD products, (2) the magnitude of the drag acceleration the satellites are experiencing, (3) the number of orbits over which a single density observation is estimated (also referred to as the “fit-span”), and (4) the number of satellites contributing to each observation. We further examine how assimilation modifies these dependencies when the retrieved densities are integrated into a background model.

To this end, we develop a more realistic representation of POD noise than the prevailing white-noise assumption, constructing three representative noise tiers (“low”/“medium”/“high”) informed by values reported in the literature. These noise levels are superimposed on a set of 18 reference ephemerides propagated along Starlink-like trajectories using the HASDM density model over the month of April 2023. The resulting synthetic datasets allow us to evaluate and quantify the impacts of POD noise, drag strength, fit-span, and satellite count on retrieval performance.

Specifically, we ask:

- How do density retrieval errors respond to a realistic POD noise model?
- How do density retrieval errors scale with fit-span ?
- How do density retrieval errors change as additional co-orbital satellites are incorporated into a single observation?
- How do these behaviours differ when densities are processed with and without an assimilative scheme?

2. EXPERIMENTAL SETUP

2.1 Satellite Selection and Reference Data Generation

We begin by selecting one platform around which to focus: the Starlink V1.0 platform. The Starlink V1.0 satellite currently represents one of the major large-scale commercial constellations in orbit for which some information is available in the scientific literature [18]. Orbital elements for this constellation vehicle were obtained from Two-Line Elements, accessed through `space-track.org`. Reference cross-sectional areas and masses were taken from literature sources [18, 47].

“Truth” ephemerides are then generated by propagating the Keplerian elements in Table 1 using the numerical integrator and force model summarized in Table 2. Unless otherwise specified, all analyses and results pertain to the first week of April 2023. The $F_{10.7}$ index at this time ranged from 125.1 solar flux units (sfu) to 137.6 sfu representing a medium solar activity period. The period is geomagnetically quiet with a peak Kp index of $K_p = 4^-$ being reached over the period.

Section 3.3 investigates the simultaneous use of multiple satellites spread along a common orbital track to resolve densities. Ephemerides for these additional satellites were generated synthetically by replicating the reference orbit and offsetting only the true anomaly. In practice, this means that each satellite follows an identical trajectory, shifted in phase along the orbit so that the group forms a uniformly spaced “string of pearls” configuration. For $N = 18$, for example, the satellites are spaced 20° apart in true anomaly, ensuring full coverage of the orbit with equal angular separation. The aim of this setup is to isolate the effect of satellite count on density retrievals by holding all other orbital parameters constant.

Geometric properties (projected area, drag coefficient, and mass) are held constant throughout each simulation. Although a time-varying attitude model would increase physical fidelity [18, 17], it was omitted to focus on the relationship between density retrieval errors and POD model errors.

Table 1: Initial Keplerian Elements of the Starlink-like Satellite Simulated

Satellite	a (km)	e	i ($^\circ$)	RAAN ($^\circ$)	ω ($^\circ$)	M ($^\circ$)
Starlink V1.0	6928.1	0.00001	53.0	125.0	90.0	40.0

Table 2: Force Model and Numerical Integrator

Component	Description	Reference
Numerical Integrator	Dormand–Prince (DOP853), variable step size	[14]
Earth Gravity	EGM2008 to degree and order 120	[43]
Third-Body Perturbations	Sun and Moon point-mass gravity	[19]
Solar Radiation Pressure	Cannonball model, Cone shadow model	[32]
Atmospheric Density Model	HASDM,	[53]
Geometric Model	0.05 (for both SRP and Drag), Panel Model	[47, 17]

The reference (or “truth”) neutral-density time series against which benchmarking of the different strategies is carried out was generated by sampling the HASDM densities along each propagated orbit and computing the mean density over the period corresponding to one full orbital revolution of the satellite.

2.2 Energy Dissipation Rate Based Density Retrieval

Whilst a number of approaches exist for estimating density from POD data, we focus on the EDR method. The EDR method can be applied to many different kinds of observations and is computationally more tractable than alternative methods such as POD accelerometry [48, 55]. The use of this kind of method underpins state-of-the-art models such the initial version of HASDM [53] and has been applied to POD data across the scientific literature [54, 39, 38, 17, 55] providing valuable sources of comparison.

We follow the method outlined below:

For an Earth-centred, Earth-fixed (ECEF) position $\mathbf{r} = (x, y, z)$ and velocity $\mathbf{v} = (\dot{x}, \dot{y}, \dot{z})$, the orbit energy is

$$\varepsilon(t) = \frac{V^2}{2} - \frac{\omega_{\oplus}^2}{2} (x^2 + y^2) - \frac{\mu}{r} - U_{\text{nsp}}(\mathbf{r}) \quad (1)$$

with $r = \|\mathbf{r}\|$, $V = \|\mathbf{v}\|$, $\mu = 3.986004418 \times 10^{14} \text{ m}^3 \text{ s}^{-2}$, and $\omega_{\oplus} = 7.2921159 \times 10^{-5} \text{ rad s}^{-1}$. The non-spherical potential U_{nsp} is evaluated from the EGM2008 gravity field (evaluated to degree and order 120).

Assuming the atmosphere co-rotates with Earth, the relative-wind velocity is given by

$$\mathbf{V}_{\text{rel}} = \mathbf{v} - \omega_{\oplus} \times \mathbf{r} \quad V_{\text{rel}} = \|\mathbf{V}_{\text{rel}}\|$$

Under the assumption that drag is the dominant source of dissipation of orbital energy over one complete orbit, the rate of energy decay is given by

$$-\frac{d\varepsilon}{dt} = \frac{1}{2m} C_D A_{\text{ref}}(t) \rho(t) V_{\text{rel}}^3 \quad (2)$$

where m is spacecraft mass, C_D the drag coefficient, and $A_{\text{ref}}(t)$ the ram profile of the spacecraft in the direction of the oncoming atmospheric flux: \mathbf{V}_{rel} .

Integrating between successive perigees from t_0 to t_1 ,

$$\Delta\varepsilon = \varepsilon(t_1) - \varepsilon(t_0) = -\frac{\rho_{\text{eff}}}{2m} \int_{t_0}^{t_1} C_D A_{\text{ref}} V_{\text{rel}}^3 dt \quad (3)$$

which yields the orbit-effective density:

$$\rho_{\text{eff}} = -\frac{2m\Delta\varepsilon}{\int_{t_0}^{t_1} C_D A_{\text{ref}} V_{\text{rel}}^3 dt} \quad (4)$$

Perigees are detected by locating local minima in the radial distance $r(t)$ where \dot{r} changes sign from negative to positive. To suppress false triggers in noisy POD data, a candidate is accepted only if r subsequently increases by at least 2 m over a minute-long window.

2.3 POD Noise Model

POD solution noise is one of the primary mechanisms through which the quality of density retrievals are degraded. In previous studies looking at modelling the effect of POD noise on density retrieval, POD noise errors were assumed to be isotropic white noise [47, 48]. We aim to refine that assumption to provide a noise model which is more representative of the reported decimetre- to metre-level performance typical of commercial small-sat platforms [4]. Unfortunately, the open literature on commercial GNSS-based POD solutions is sparse, arguably a key reason the white-noise modelling approach persists in the literature. [51] present an excellent meta-analysis of science-grade LEO-POD performance, yet no comparable survey exists for the commercial sector. To the best of the authors' knowledge, to date, detailed metrics of a commercial constellation's POD product exist in the public domain only for the Spire satellites [4]; we therefore combine those figures with the Sentinel "navigation" solution [25] as a plausible lower bound on what a well-resourced commercial operator might achieve, while recognising that many fleets are likely to deliver lower accuracy.

Sentinel-1/-2/-6 near-real-time (NRT) orbits, validated via International Laser Ranging Service (ILRS) data, exhibit $\sim 0.3\text{--}0.7$ m three-dimensional (3-D) RMS position error, whereas their science-grade orbits remain below 7 cm [25, 35]. Spire's commercial POD product lies in between 27–30 cm 3D RMS when compared against Bernese/NAPEOS references that agree at the 6–7 cm level [4]. For comparison with commercial uncooperative tracking data, at the time of writing (August 2025), LeoLabs list 41 m accuracy and 16 m precision for their uncooperatively derived data (as reported at https://platform.leolabs.space/system_metrics).

Across the reviewed literature characterizing LEO POD solutions [34, 25, 22, 4, 33, 35], along-track errors consistently dominate and tend to follow a $1/f$ ("pink-noise") spectrum; cross-track errors are then smaller and more isotropic, and radial errors tend to be the smallest and are also largely isotropic [25]. The errors in POD solutions display temporal autocorrelation. In the case of Sentinel-1 and -2's "nav" solutions this behaviour can be captured by a first-order autoregressive process (AR(1)) with 600–800 s decay. [25] also cite that the variance is heteroscedastic, oscillating

in tandem with geometric dilution of precision (GDOP) at 0.5–3 m peak-to-peak. Along-track bias is also commonly reported (e.g. -0.2 to -0.5 m for Sentinel-2A and B [25]).

The sources of these POD errors are numerous and may be attributed to effects such as: receiver noise and multipath [51, 35], ionospheric and/or tropospheric delay, and mismodelled accelerations [3].

Note that most science missions benefit from being able to routinely verify their POD solutions against a mm-level accuracy, independent measurement source. For example, Satellite Laser Ranging (SLR). However, the presence of retro-reflectors is exceedingly rare on commercial satellites thus requiring different and less robust validation approaches (e.g. orbit overlaps). For instance [4] validate POD solutions from the Spire constellation against two independent POD processing software packages (Bernese, NAPEOS). This provides a relative rather than fully absolute verification. The absence of SLR or equivalent truth data complicates both error characterisation and performance guarantees for commercial fleets.

To reproduce the dominant signatures reported in the LEO POD noise characterization literature, a synthetic noise model is overlaid onto each state vector within the reference ephemerides (section 2.1). The noise model is composed of the following components: anisotropic noise, pink ($1/f$) colouring in the along-track direction, short-range autocorrelation, bias, and heteroscedastic scaling of the noise as a function of Geometric Dilution of Precision (GDOP) [25, 36]. Figure 1 illustrates the effect of each step on the along-track residuals of a 24-hour arc of simulated POD data when compared to the “truth” simulated reference ephemeris. We outline the method to generate and apply this noise model below:

For each epoch t we draw a zero-mean position residual

$$\boldsymbol{\varepsilon}_w(t) = \begin{bmatrix} \varepsilon_H \\ \varepsilon_C \\ \varepsilon_L \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \text{diag}(\sigma_H^2, \sigma_C^2, \sigma_L^2))$$

where the standard deviations $\sigma_{H,C,L}$ (height, cross-track, along-track) of each noise model tier (high, medium, low) are defined based on values in the literature (values presented in table 3).

The along-track component is coloured by a frequency-domain filter (meaning the noise is adjusted so that slow oscillations become stronger and rapid oscillations weaker)

$$\hat{\varepsilon}_{\text{pink}}(f) = \hat{\varepsilon}_w(f) (f/f_0)^{-1/2}$$

where hats denote Fourier coefficients and f_0 is the fundamental frequency of the window. An inverse Fast Fourier Transform returns the time-domain series $\varepsilon_{\text{pink}}(t)$, which replaces ε_L .

Short-range correlation is imposed with a first-order autoregressive process,

$$x_t = \alpha x_{t-1} + e_t \quad e_t \sim \mathcal{N}(0, \sigma^2)$$

using a 10-s step size and a target folding time of $k_{\text{decay}} = 70$ samples ($= 700$ s), as noted in [25].

$$\alpha = 1 - \frac{1}{k_{\text{decay}}} = 0.9857$$

The AR(1) operator is applied to each positional axis (H/C/L).

A constant offset $b_L = -0.5$ m is added in the along-track of all solutions, aligning with the mean bias reported for Sentinel NRT solutions [25].

As per [25], POD errors tend to be modulated by GDOP, which is caused by variations in the geometry of the overhead GNSS satellites that the LEO orbiters use to generate their own POD solutions. We simulate reception from a dual GPS/Galileo constellation by propagating broadcast ephemerides [49, 16] over the period of interest. To ensure robust visibility and geometry, GNSS satellites with elevation angles above 5° are retained, following standard practice in space-borne GNSS-POD processing [11]. The receiver antenna is assumed to be zenith-facing and nadir-mounted, a common configuration on commercial LEO spacecraft [4].

With line-of-sight unit vectors $\mathbf{u}_i(t)$ to n visible satellites, the design matrix is $A = [\mathbf{u}_1^\top \dots \mathbf{u}_n^\top]^\top$ and the GDOP is

$$g(t) = \sqrt{\text{tr}[(A^\top A)^{-1}]}$$

Let $\tilde{g} = \text{median}\{g(t)\}$. We select an exponential heteroscedastic factor

$$m(t) = \exp[g(t) - \tilde{g}]$$

consistent with the multiplicative behaviour of GDOP observed by [25]. The final position error is

$$\eta(t) = m(t) [\varepsilon_w(t)_{(\text{coloured})}] + \begin{bmatrix} 0 \\ 0 \\ b_L \end{bmatrix}$$

For each simulated ephemeris in this study, three different POD noise solution tiers are created. The magnitude of the noise associated with these three noise tiers were selected in an attempt to reflect the range of solutions found in the literature and those that may be encountered by commercial POD data providers. The values selected are listed in Table 3.

Table 3: H/C/L error used for each noise model, corresponding 3D RMS error, and source for selected values.

Noise Model	Height (m)	Cross Track (m)	Along Track (m)	3D RMS (m)	Reference/Justification
Low	0.2	0.1	0.2	0.300	[25]
Medium	0.1	0.2	0.4	0.458	[4]
High	0.8	0.4	1.6	1.830	$\sim 4 \times \text{Medium case}$

2.4 Effect of Drag Regime on Retrieval Error

Ten drag-acceleration levels were sampled. These were selected to encompass the typical range of drag regimes encountered by LEO satellite constellations across a range of solar and geomagnetic activity levels. The drag-acceleration bounds were defined as ranging from 6.5×10^{-8} to $1.5 \times 10^{-5} \text{ m s}^{-2}$. Given the broadly exponential relationship between thermospheric density and altitude, the samples were spaced uniformly in logarithmic space. Since the satellite's reference cross-sectional area, mass, and drag coefficient were all kept constant, variation in the semimajor-axis was the only variable modified to sample the drag-acceleration space. Figure 2 shows the sampled altitudes and corresponding mean orbital drag accelerations over the first week of April 2023 for each sample.

2.5 Effect of Fit-Span on Retrieval Error

The fit-span parameter used in the EDR method effectively represents the amount of time over which the integral is carried out (equation 3). Increasing the fit-span reduces the noise by increasing the length of the observation and thus the signal of any given density estimate. This process will also result in a smoothing of the extracted density signal as the result of the integral is now attributed to a longer orbital arc. In the results presented, we varied this over integer values from 1 to 5 orbits. For each fit-span N , the change in orbital energy due to drag was measured by integrating over N consecutive orbits using the simulated trajectories over which the different POD noise levels were overlaid. The retrieved density time series for each span was then compared to the reference density data to measure the RMS% error. Note that when studying the impact of increasing fit-span we compared the mean HASDM density over the same N -orbit interval. This procedure ensures that comparisons across fit spans are performed on a consistent temporal averaging basis, however, this means that the loss of temporal resolution which occurs as a result of increasing the fit-span is not accounted for.

2.6 Assimilating POD-Derived Densities

POD data alone enable the provision of density estimates along the orbits of the satellites for which the POD data is available. However, most applications that satellite operators and SSA service providers are interested in (e.g. collision avoidance, orbit propagation) typically require the use of density values spanning much larger spatial scales (potentially across the entirety of LEO). Leveraging observations to this end is typically achieved through an assimilation scheme which makes use of the observations- in this case POD-derived EDRs- to calibrate a background model. The following section is provided under the expectation that future density models will rely heavily on POD-derived densities within such frameworks.

The process of assimilating density observations into a background model will warp the relationship between retrieval error and the underlying variables of fit-span, drag-strength, POD quality, and satellite count. This section aims to provide an assessment of the impact of assimilation on the density retrieval errors.

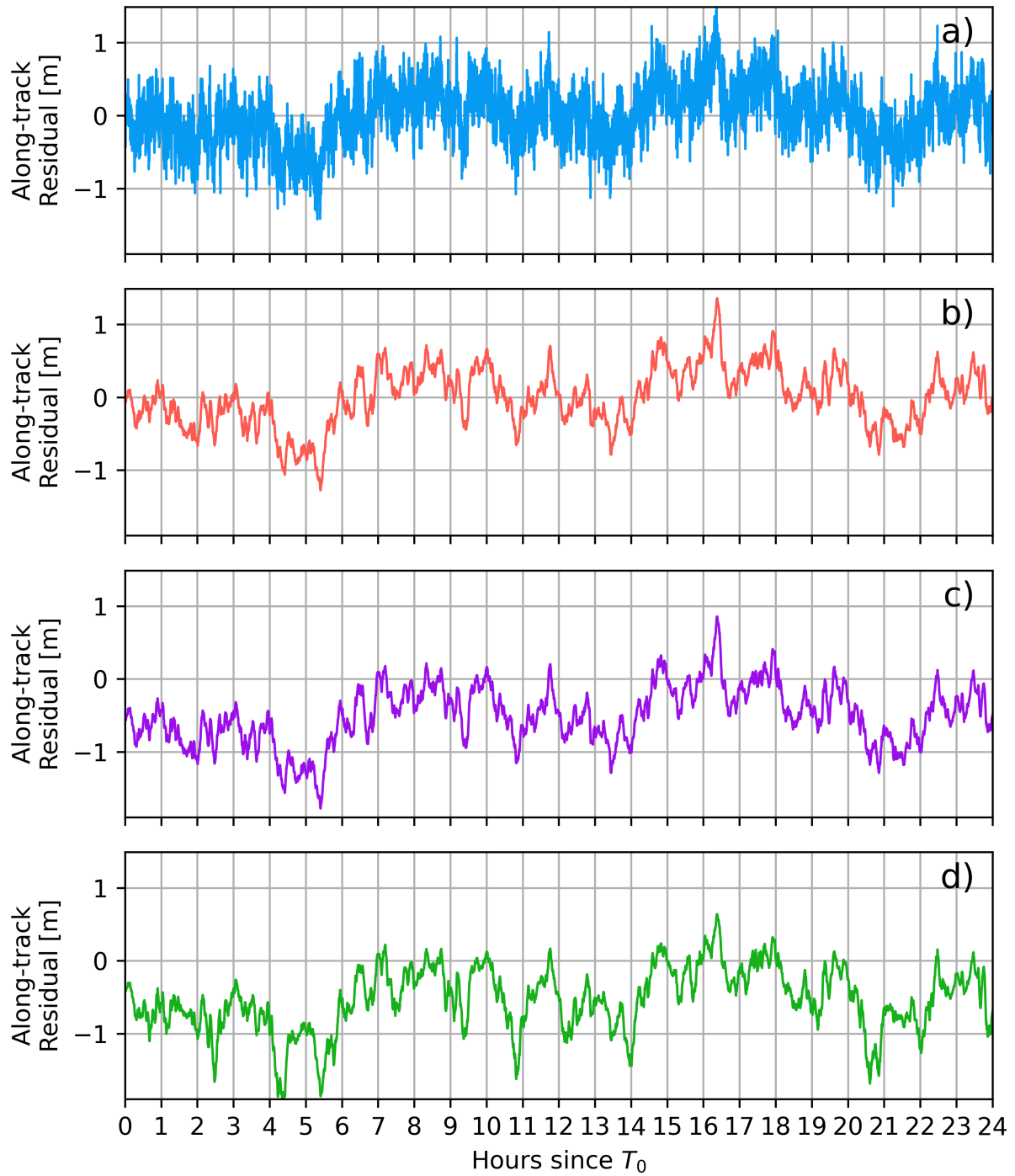


Fig. 1: Along-track residuals of a 24-hour trajectory arc with respect to the reference orbit. Going from top to bottom, each graph corresponds to the addition of a new component of noise. Panel a) shows the modulation of the white noise into pink noise. Panel b) shows the effect of the application of the autocorrelation model (AR(1)). Panel c) shows the effect of adding a -0.5 m bias. Panel d) shows the effect of GDOP-modulation of the signal.

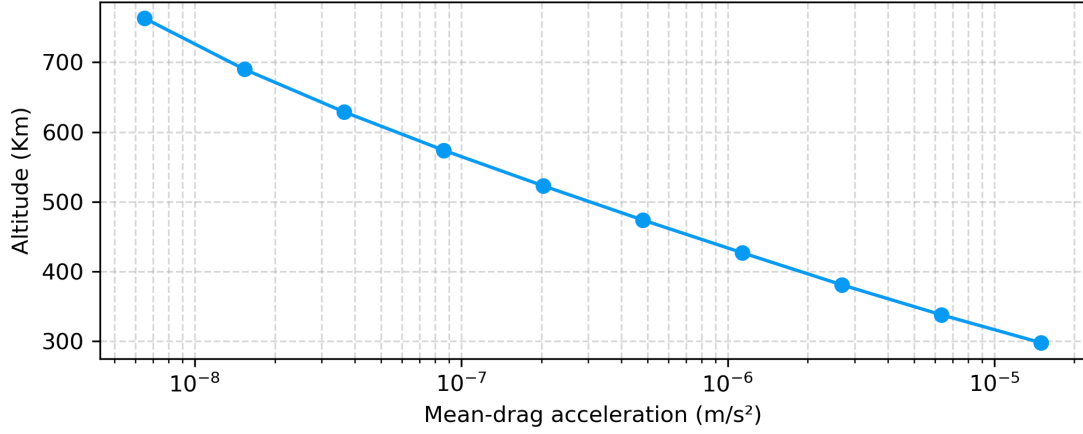


Fig. 2: Mean altitude and corresponding drag-acceleration of each of the 10 altitudes sampled over the first week of April 2023. Densities sampled from the HASDM density model.

During geomagnetically quiet periods and over relatively short time spans (\leq a few days), orbit errors can often be addressed through static bias adjustments (e.g. ballistic coefficient estimation), such that the assimilation process may become akin to a bias tuning exercise. In contrast, storm-time conditions introduce steep temporal gradients in density. These features stress-test the ability of assimilation schemes to resolve time-dependent corrections rather than simply absorb static biases, making the comparison more discriminative and operationally relevant. For this reason, we focus this section of the analysis on the week spanning 21st–28th April 2023. This week-long window includes a geomagnetically quiet phase and a G4-class geomagnetic storm (including main phase, and post-storm recovery).

2.6.1 Assimilation Scheme Description: Ensemble–Kalman Adjustment of NRLMSISE-00

The assimilation process used in this section of the work is outlined below. It follows the process outlined in [37]. Orbit-averaged EDRs are assimilated into NRLMSISE-00 using an Ensemble Kalman Filter (EnKF). The filter estimates a correction series $\Delta A_p(t)$ that perturbs the 3-hourly geomagnetic index A_p ; all other inputs remain fixed, apart from an empirically selected solar-flux correction. A value of $\Delta F_{10.7} = -40$ sfu is applied over the period of the experiment (April 2023) to prevent baseline over-prediction.

A 48-hour window is sampled every $\Delta t_{\text{step}} = 3$ h. At epoch t_k , the state vector collects the present and next three A_p -corrections,

$$\mathbf{x}_k = [\Delta A_p(t_k), \Delta A_p(t_{k+1}), \Delta A_p(t_{k+2}), \Delta A_p(t_{k+3})]^\top \in \mathbb{R}^4$$

so the sliding estimation window influencing any analysis spans 12 h. The sliding window moves forward by 3 hours after each estimation epoch. Because consecutive states overlap by three elements, every epoch is re-analysed as new data arrive and no temporal gaps form. Between analyses the correction field is assumed constant, $\mathbf{x}_{k+1}^f = \mathbf{x}_k^a$. To maintain ensemble spread, the newest element is redrawn at each step from $\mathcal{N}(0, \sigma_{A_p}^2)$ with $\sigma_{A_p} = 50$.

Each EDR measurement represents the density averaged over a complete orbital arc centred on time t_j . For ensemble member i , the corresponding model prediction is obtained by integrating density returned by NRLMSISE-00 along that arc while applying ΔA_p and $\Delta F_{10.7}$,

$$h_j(\mathbf{x}_i) = \bar{\rho}_{\text{NRLMSISE-00}}(t_j; A_p(t_j) + \Delta A_p(t_j), F_{10.7} + \Delta F_{10.7})$$

Let $\mathbf{y} \in \mathbb{R}^m$ stack the $m = 6\text{--}7$ orbit-mean densities that fall within the current window, and let $\mathbf{h}(\mathbf{x}_i)$ be the corresponding model outputs for ensemble member i . Although the experiment is simulated and the uncertainties can be accurately characterised, we deliberately emulate the operational setting where per-arc uncertainties are not known accurately a priori: we therefore fix a single relative observation uncertainty for all arcs, chosen as an intermediate

value between the two representative levels shown in Figure 3. We adopt a one-sigma fractional error of 30% for every orbit-mean density, yielding

$$\mathbf{R} = \text{diag}((\alpha y_1)^2, \dots, (\alpha y_m)^2) \quad \alpha = 0.30$$

We use normalised ensemble anomalies \mathbf{X}', \mathbf{Y}' (columns scaled by $1/\sqrt{N_e - 1}$) so that the sample covariances are $\mathbf{P}_{xy} = \mathbf{X}'\mathbf{Y}'^T$ and $\mathbf{P}_{yy} = \mathbf{Y}'\mathbf{Y}'^T$. The EnKF analysis is given by:

$$\mathbf{K} = \mathbf{P}_{xy}(\mathbf{P}_{yy} + \mathbf{R})^{-1}, \quad \mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}(\mathbf{y} - \mathbf{h}(\mathbf{x}_i^f))$$

After all windows are processed, an epoch may appear in up to four analyses; their arithmetic mean yields the final estimate $\widehat{\Delta A_p}(t_k)$. At the temporal edges, where fewer analyses contribute, a linear weight proportional to order of appearance compensates for reduced redundancy without bias. The resulting A_p correction time series is fed back into NRLMSISE-00 to generate the assimilated density time series.

All assimilated EDR time series generated were processed through this framework. Note that assimilation is performed on a per-satellite basis rather than jointly across all satellites. This was a practical choice motivated by the complexity of setting up a multi-satellite variant; nevertheless, this configuration was deemed sufficient to expose the strengths and limitations of using POD-derived orbit-mean densities within an assimilative pipeline.

2.7 Aggregating Simultaneous Density Observations From Multiple Satellites

Most commercial satellite constellations tend to follow clearly defined orbital configurations (e.g. Walker-Delta, Flower) to carry out their mission. These orbital configurations usually involve geometrically near-identical satellites orbiting along a nominal orbital path. This mode of operation opens up a further avenue of exploration for POD derived densities. At present, within existing studies, the main lever that is actioned to reduce noise and increase accuracy is to extend the measurement window (typically by increasing the fit-span). Constellation data offers a different mechanism- using a greater number of simultaneous measurements along the same orbital path. To investigate the effect of using data from more satellites on retrieving a density time series, we simulate combinations of equally spaced $N = \{1, 2, 3, 6, 9, 18\}$ satellites along the nominal Starlink-like trajectory in Table 1. All Keplerian elements are held constant while true anomaly is modified to produce even angular separation (e.g., 20° spacing for $N = 18$). Symmetry was ensured to minimize sampling bias. With the fit-span fixed at one orbit, the EDR method is then modified to provide neutral densities at predetermined epochs tied to the orbital period of the first satellite (rather than perigee-to-perigee as outlined in section 2.2). This modification is required so that samples can be averaged at a common cadence. The mean retrieved density across the N satellites was then compared to the “truth” background density model: HASDM. For this section of the analysis, we use the same storm time period outlined in section 2.6, and compare the effects of averaging density observations from multiple satellites both with and without use of the assimilation tool to investigate the effect of noise on the density retrieval exercise.

3. RESULTS

3.1 Impact of improved Noise model

Existing studies on the effect of POD noise on density retrieval performance tend to assume POD noise to be white [48, 47]. Thus, as a first point of comparison we consider the impact of modifying this commonly used approach to a higher-fidelity one (outlined in section 2.3). To do so, we apply a white noise model and the noise model outlined in section 2.3 to the reference ephemeris for the nominal-altitude Starlink-like satellite over the first week of April 2023. We then use both ephemerides to generate a time series of effective densities using the EDR method (see Section 2.2) to analyse the effect of changing the noise model on the end product we are interested in: thermospheric density time series. To ensure fair comparison, the white and higher-fidelity noise models are made equivalent in 3D RMS terms. For this experiment, the absolute noise level of both noise models is set to 0.458 m (3D RMS). For the higher-fidelity noise model, the 1-sigma anisotropic noise distribution is 0.2 m in the height direction, 0.1 m in the cross-track direction and 0.4 m in the along-track direction. For the white noise case this translates to a 1-sigma error of 0.265 m in all three directions. To generate an estimate of the RMS percentage error and uncertainty of the density time series retrieved using each noise model, we sample each noise distribution by repeating this process 50 times, each

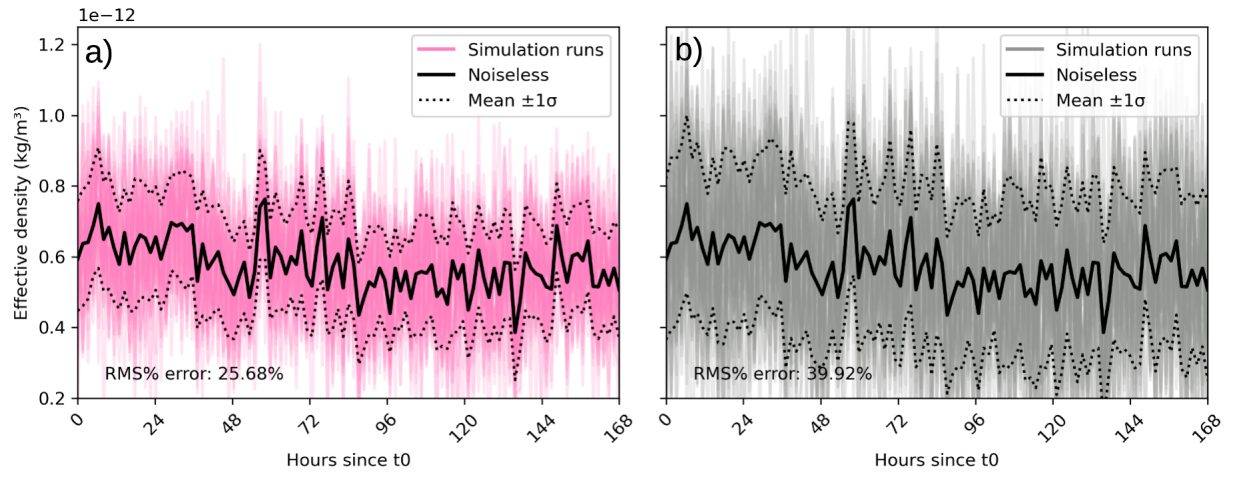


Fig. 3: Monte-Carlo Sampling of the error associated with different noise modelling approaches across the Starlink-like reference trajectory over the first week of April 2023. Panel a) shows the error for the higher-fidelity noise model outlined in section 2.3. Panel b) shows the error for a white noise model. Both models are equivalent in 3D RMS terms (0.458 m). Each panel contains 50 samples of each noise model.

time sampling at random from both the noise distributions. Figure 3 shows the 50 density time series generated for each noise model and the density time series returned using a noise-free ephemeris.

The RMS% error is calculated as follows: Let M be the number of Monte-Carlo simulations (50 in this case) and N the number of observations in the density time series (110 in this case). We denote by

$$\rho_{i,j}, \quad i = 1, \dots, M, \quad j = 1, \dots, N$$

the effective-density from the i th simulation at time index j , and by

$$\rho_{0,j}, \quad j = 1, \dots, N$$

the “noiseless” effective-density at time index j . Then the RMS-percentage error is computed as

$$\text{RMS\% error} = \frac{\sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (\rho_{i,j} - \rho_{0,j})^2}}{\frac{1}{N} \sum_{j=1}^N \rho_{0,j}} \times 100\% \quad (5)$$

We see that with both noise models being attributed identical 3D RMS noise values, the white noise model results in larger error. We attribute this in part to the decreased level of error in the height direction of the higher-fidelity noise model when compared to the white noise model. Moreover the autocorrelation present within the higher-fidelity solution likely has the effect of reducing the variance in the solution, resulting in a smaller spread around the mean.

3.2 Effect of Drag Strength and Fit-Span

When retrieving densities from POD data along the path of a single satellite, the three key variables driving the performance of the EDR method are: POD solution quality (noise/bias), strength of the drag signal (drag acceleration), and number of orbits over which the EDR integral (Eq.3) is carried out (fit-span). Increases in all three of these variables are expected to result in lower density retrieval error. In the following section we investigate whether this result holds and to what extent. From this point on, all noise is modelled using the higher-fidelity noise model outlined in section 2.3.

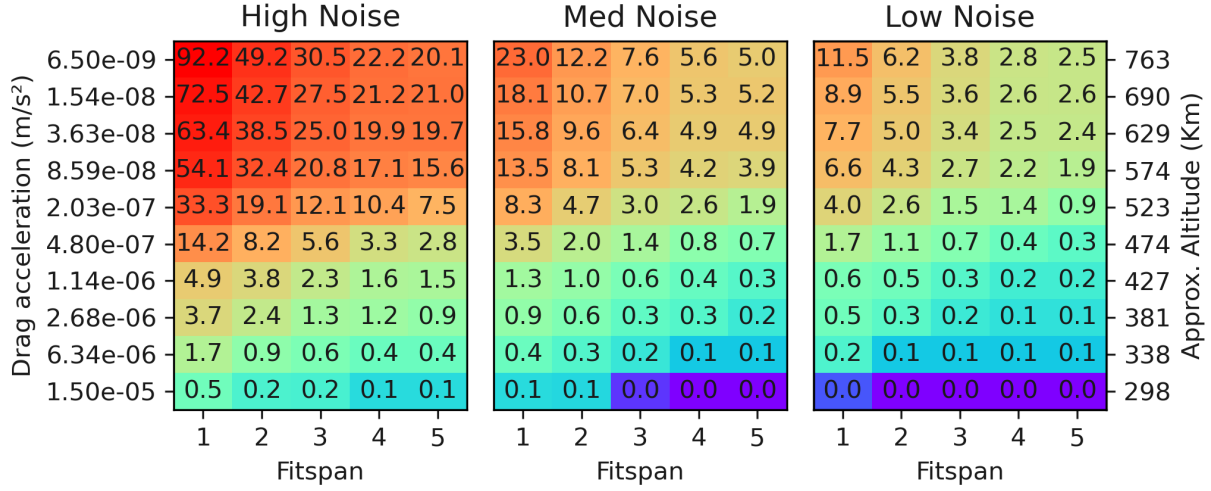


Fig. 4: Median RMS percentage errors of retrieved densities along the orbits queried over the first week of April 2023. Each matrix corresponds to a POD noise model specified in Table 3. Values are median absolute errors across 100 orbits. Over this period (near solar maximum) the $F_{10.7}$ index ranged from 125.1 sfu to 137.6 sfu and the maximal K_p value was $K_p = 4^-$. The secondary y-axis (right-hand side) provides approximate altitudes which correspond to the drag-accelerations on the left hand-side y-axis at the time of analysis (1-7th April 2023) for an object with an area-to-mass ratio of 0.05.

Figure 4 displays the absolute median error across 100 density retrievals sampling across the parameter space of interest: POD noise strength (low/medium/high as defined in table 3), fit-span length and drag acceleration. The density retrievals are carried out in the first week of April 2023 (geomagnetically calm). The absolute median error is used here instead of the mean to reduce the impact of outliers and to give a sense of the typical error (50th percentile). Going from left to right across figure 4, each matrix corresponds to a different noise grade (as defined in section 2.3). The rows of each matrix correspond to the drag accelerations outlined in figure 2, and the columns corresponds to increasing fit-spans (from 1 to 5). The top two rows of each matrix corresponds to a very low Earth orbit regime (≤ 350 km). In this case the strength of the drag signal is such that changes in POD quality and increases in fit-span make only marginal improvements to the quality of the retrieved densities- the signal is already far exceeding the noise of the POD solution. On the other hand, the bottom row in the matrices is representative of an altitude of roughly 750 km. Here, the drag signal is around four orders of magnitude weaker than for satellites in the top-most row, and each increment in both fit-span and POD quality, significant changes in the quality of the retrieved densities are made.

Comparing any one cell across the three matrices shows a clear ordering: the low-noise solution error exhibits roughly half the error of the medium-noise case, and the high-noise solution is typically about four times larger than the medium-noise case. This ordering and relative error magnitude is consistent with the noise levels of the POD products that these densities are derived from (see Table 3). Across all noise levels, errors decrease with increasing fit-span and with increasing drag acceleration.

We note that two of the currently operational mega-constellations (SpaceX Starlink and Amazon Kuiper- operating at around 550 and 650 km altitude respectively) operate in drag regimes where changes in fit-span, and POD noise quality stand to make a significant impact. We also note that this relationship is likely to become exacerbated over the next decade as the drag signal these constellations are exposed to will reduce in tandem with the solar cycle, which is currently exiting its maximum phase. In comparison, the frequently studied GRACE-FO platform provides ideal conditions for retrieving low error POD density estimates- especially given the well-characterized, high-quality nature of the POD data that is available for it [52].

Figure 4 shows how median RMS% errors vary with drag acceleration and fit-span for each POD-noise level. Having sampled the error landscape through simulation, we aim to capture these behaviours with simple functions. The aim of this exercise is to provide lightweight analytical expressions that can serve as first-order approximations of how

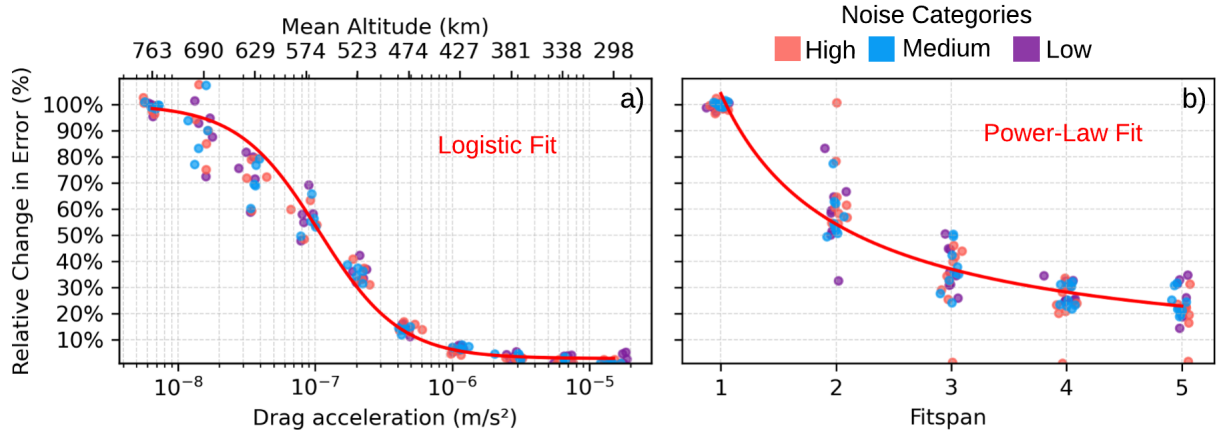


Fig. 5: Relative change in RMS-percentage error as a function of (a) orbit-averaged drag acceleration and (b) EDR fit-span. A logistic function and a power-law provide empirical approximations for panels (a) and (b), respectively.

density-retrieval error scales with drag acceleration and fit-span without the need for running a full simulation.

To construct these fits, we first normalise each row and column of the matrix by its maximum to remove absolute scale. We then analyse one axis at a time: rows for drag-dependence and columns for fit-span-dependence. These normalized values are displayed in figure 5. As in figure 4, error decreases as a function of both drag acceleration and fit-span. However there are differences in how the error decreases. Drag displays a sharper gradient in the middle of the curve, whereas fit-span sees a sharper decrease at the beginning of the curve which then tails off.

Panel a) demonstrates that the error-drag relationship is well described by a logistic curve. At very low drag accelerations ($\sim 5 \times 10^{-9}$ – $2 \times 10^{-8} \text{ m s}^{-2}$), increasing drag yields only modest error reductions (10%). Between $\sim 2 \times 10^{-8}$ and $2 \times 10^{-7} \text{ m s}^{-2}$, the curve steepens: each tripling of drag acceleration reduces the error by roughly half. Beyond $\sim 2 \times 10^{-7} \text{ m s}^{-2}$, improvements taper as the curve approaches a floor.

This logistic response and its transition region appear consistently across all three noise levels. The noise level scales the absolute error scale (see figure 4). However, the relative drop in error occurs at similar drag accelerations/altitudes. We attribute this common behaviour to the orbit-averaged nature of the EDR method: across the range of noise magnitudes tested, the work done by drag over the entire orbit is consistently sufficient to produce a comparable level of improvement relative to POD noise. This behaviour mirrors the findings of [48](their Figure 13b). Other density retrieval methods such as POD-accelerometry are reported to exhibit higher sensitivity to POD noise [48]. Sampling a broader range of noise values would clarify whether a noise dependence develops at greater noise values. Nevertheless, the current results suggest that, regardless of absolute POD noise magnitude (within the range sampled), there is a region in the drag-acceleration space where density retrieval error is disproportionately improved by a unit increase in drag acceleration.

Panel (b) shows that the fit-span dependence follows a simple power-law: the first increase in fit-span yields an approximate halving in error, and subsequent increments in fit-span have roughly half the impact of the previous increment.

Using these two mathematical fits, we can write the behaviour of the relative RMS percentage error as a separable function of drag acceleration d and fit-span f . The drag dependence in figure 5 a) is represented by a four-parameter logistic function:

$$E_{\text{drag}}(d) = E_{\text{min}} + \frac{E_{\text{max}} - E_{\text{min}}}{1 + (d/d_{50})^{\kappa}} \quad (6)$$

where E_{max} is the low-drag plateau (e.g., $d 2 \times 10^{-8} \text{ m s}^{-2}$), E_{min} is the modelling floor (here 5×10^{-6}), d_{50} is the half-transition drag, and κ controls the steepness near the inflection. Over the week-long experiment, a nonlinear least-squares fit yields $E_{\text{max}} = 97\% \pm 2\%$, $E_{\text{min}} = 1.8\% \pm 0.3\%$, $d_{50} = (2.1 \pm 0.2) \times 10^{-7} \text{ m s}^{-2}$, $\kappa = 1.12 \pm 0.05$, and $R^2 = 0.985$.

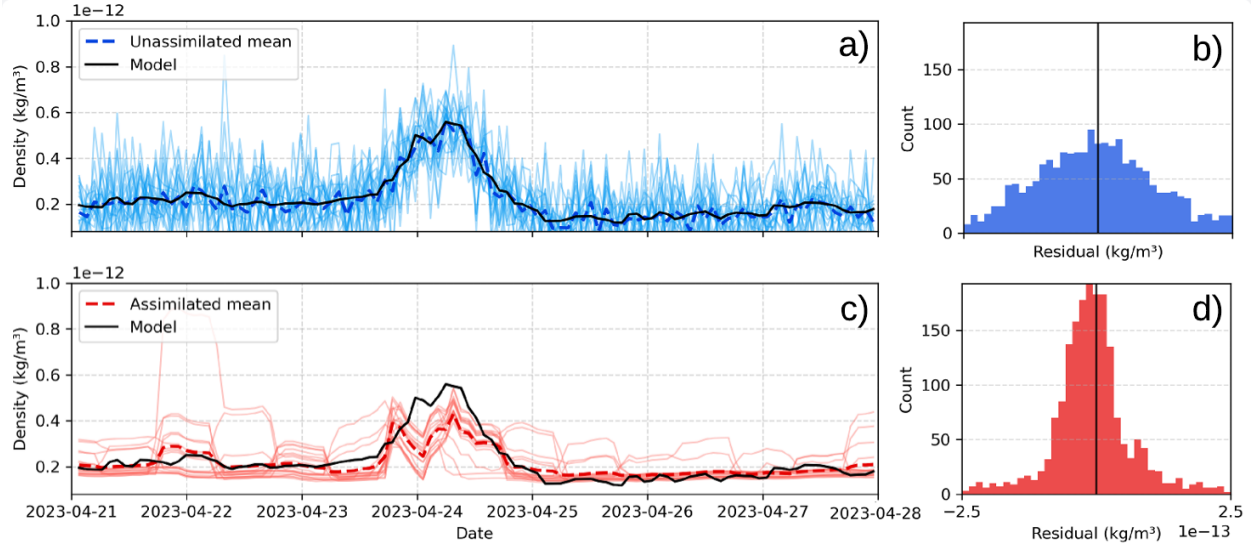


Fig. 6: Time series of densities retrieved by 18 satellites simultaneously under the high-noise model using a fit-span of 1. The thin blue lines in panel a) show the direct density retrievals and the thick blue dotted line represents the mean of these. The thin red lines in panel c) represent the density time series for the same satellites when processed through the assimilative pipeline. The red dotted line is the mean of those estimates. In both panels a) and c) the black line is the truth model (HASDM). Panels b) and d) provide the corresponding residuals of all the time series relative to the truth model.

The fit-span sensitivity follows

$$E_{\text{span}}(f) = 2^{-(f-1)}, \quad (7)$$

so that doubling the fit-span approximately halves the error.

Combining both effects gives

$$E(d, f) = \left[E_{\min} + \frac{E_{\max} - E_{\min}}{1 + (d/d_{50})^{\bar{\kappa}}} \right] \times 2^{-(f-1)} \quad (8)$$

which explains between 96% and 98% of the variance observed in the matrices for all three POD-noise levels (figure 4).

From equations (6) and (7), two practical rules emerge:

- **Fit-span rule:** Each doubling of the fit-span approximately halves the RMS-percentage error.
- **Drag rule:** For $d \in [3 \times 10^{-8}, 3 \times 10^{-7}] \text{ m s}^{-2}$ (roughly 630-450 km), each threefold increase in d approximately halves the RMS-percentage error. Below this range, improvements are relatively minor; above it, performance approaches a floor.

3.3 Effect of Increasing Satellite Numbers and Usage Within an Assimilative Context

A central question for operational use of POD-derived densities is how much benefit can be gained by increasing the number of satellites contributing independent measurements along a given orbit.

Panels a) and c) of figure 6 show the density time series provided by all 18 satellites along the same orbital trajectory and their mean in both processing modes: (i) direct, unassimilated density retrievals and (ii) retrievals processed through the assimilative EnKF pipeline described in section 2.6.1. In both cases, the fit-span is fixed at 1. Individually, the assimilated time series (thin red lines in panel c) exhibit lower variance than the direct density observations (thin blue lines in panel a), as reflected by narrower residual distributions in panel d) when compared to panel b). However, when averaged, the assimilated time series exhibit bias, with a skew and offset visible in the residuals (panel d)). In

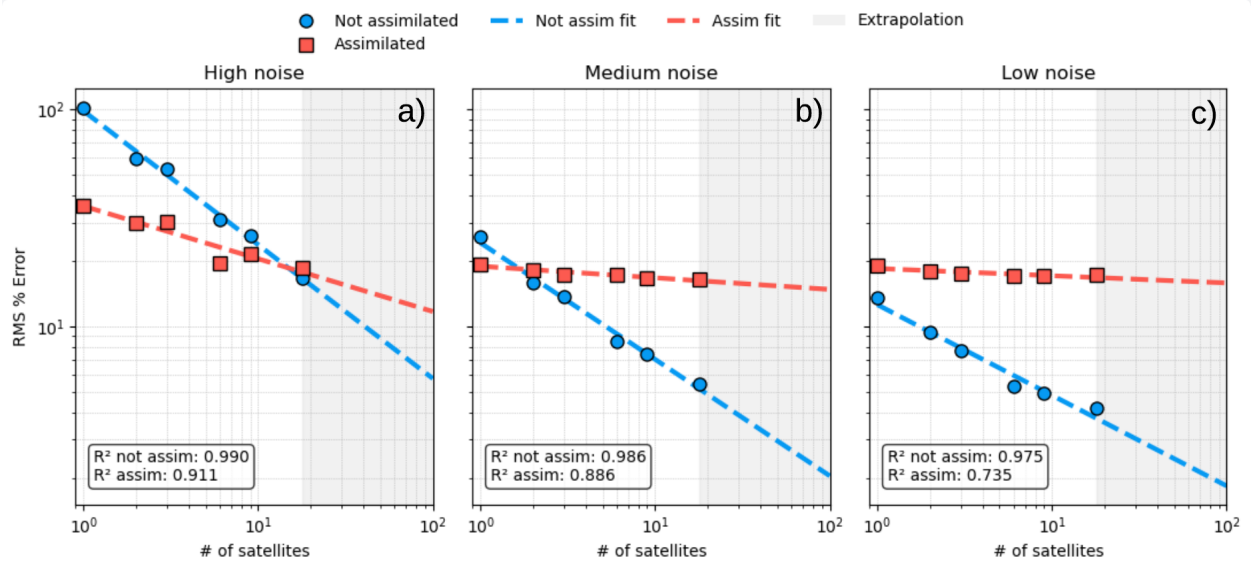


Fig. 7: Mean RMS% error of mean-orbital density estimated by increasing numbers of equally spaced satellites. Panels a), b), and c) correspond to high, medium and low noise levels respectively. Red line/data points are the assimilated densities. Blue line/data points are the direct density observations (unassimilated). The lines fit to the observations are inverse power-law fits. The fit lines extrapolated (grey area of graph) to 100 satellites.

contrast, individual time series in panel a) have larger per-arc errors but their mean aligns more faithfully with the density profile of the truth model (HASDM). A clear distinction emerges in the shape of the retrieved storm response by the assimilated and unassimilated models. The mean of the assimilated solutions tends toward a slightly different storm profile- characteristic of bias introduced by the underlying background model (NRLMSISE-00).

Furthering our effort to provide a mathematical characterisation of the drivers of POD retrieval error, we examine how error scales with the number of satellites contributing density observations. Figure 7 summarizes this relationship under the three different noise levels, both with and without assimilation. As expected, adding satellites consistently reduces error. In an ideal case of unbiased, independent, and identically distributed (i.i.d.) observations, the error should decrease according to an inverse square-root law ($\propto N^{-1/2}$). While this scaling provides a useful first-order approximation, our results are more accurately described by an inverse power law with a shallower slope. We believe this departure from the theoretical $\sim N^{-1/2}$ behaviour arises from the structure of the proposed noise model, which introduces autocorrelation, bias, and common error modulation across satellites (i.e. through GDOP effects), thereby breaking the i.i.d. assumption.

Both the absolute value of the error (intercept), and to a lesser extent the gradient of the slope, are modulated by POD noise level and whether assimilation is used. Under the high POD noise case, a given increase in the number of satellites used yields a stronger improvement for the unassimilated case. When densities are generated via the assimilative tool, the incremental gain from additional satellites is smaller.

An important feature to note in figure 7 is the point at which the blue lines fall below the red ones. Beyond this threshold, the mean of the unassimilated retrievals outperforms the assimilated estimates in resolving densities along the arc. It is clear from figure 6 that the assimilation step can introduce a persistent model-driven bias that degrades the quality of the density retrieval process.

We stress that this does not imply that assimilation lacks value. The RMS% errors reported in Fig. 7 quantify performance along the specific orbital arc. Assimilation provides a global density field, whereas the unassimilated mean is constrained to the sampled orbit. Hence, while the unassimilated mean can achieve lower error along this specific orbit given enough satellites, its utility is inherently more limited in scope.

We find a good approximation of the relationship between the error $E(N)$ and the number of satellites N is provided

by the power-law rule:

$$\frac{E(N)}{E(1)} \approx N^{-b} \implies E(N) \approx E(1)N^{-b} \quad (9)$$

Fitting across all POD tiers yields

$$b \approx \begin{cases} 0.6 & \text{(unassimilated)} \\ 0.25 & \text{(assimilated)} \end{cases}$$

The RMS% error of the direct (unassimilated) observations decays approximately as $E(N) \approx E(1)N^{-0.6}$, so doubling the satellite count reduces error by $2^{-0.6} \approx 0.6598$ (a 34.0% reduction). In the assimilative context, the decay is shallower, $E(N) \approx E(1)N^{-0.25}$, so doubling the satellite count reduces error by $2^{-0.25} \approx 0.8409$ (a 15.9% reduction). Goodness-of-fit for the full three-parameter approximation is reported below.

3.4 An analytical approximation of density inversion performance

To incorporate the effect of increasing the number of satellites N into our mathematical model, we combine equation 9 with equation 8. The full three-parameter error surface is given by

$$E(d, f, N) = E_{\min} + \frac{E_{\max} - E_{\min}}{1 + (d/d_{50})^\kappa} 2^{-(f-1)} N^{-b} \quad (10)$$

where E is the RMS percentage error, d is the orbit-averaged drag acceleration (m s^{-2}), f is the fit-span (in number of orbits), and N is the satellite count. The exponents κ and b are dimensionless; $b \approx 0.6$ for unassimilated retrievals and $b \approx 0.25$ for assimilated retrievals.

Relative to the cell-wise RMS% errors in figure 4, the full analytical model (equation 10) captures around 89% of the variation in the simulation data. Expressed as the coefficient of determination (r^2), which measures the fraction of variance explained by the model, the values are $r_{\text{high}}^2 = 0.894$, $r_{\text{medium}}^2 = 0.892$, and $r_{\text{low}}^2 = 0.887$, corresponding to the high-, medium-, and low-noise tiers respectively.

4. DISCUSSION

In the preceding section we found that retrieval error decreases approximately as $2^{-(f-1)}$ with fit-span, follows a logistic-like response to drag with the steepest section of the logistic curve between 2×10^{-8} – $2 \times 10^{-7} \text{ m s}^{-2}$, and declines with satellite count as $N^{-0.6}$ (unassimilated) or $N^{-0.25}$ (assimilated).

The halving in density-retrieval error with increasing fit-span (equation 7) is consistent with variance reduction by averaging and persists in the presence of coloured, anisotropic, and autocorrelated POD errors (section 2.3). This indicates that, across independent orbit-averaged density observations of one same spacecraft, the dominant error structures in the modelled LEO POD noise, exhibit sufficiently short temporal correlation that successive arcs contribute quasi-independent information. The benefit of increasing the fit-span must, however, be balanced against the temporal scale of the density signal one wishes to resolve: certain features of a storm can, in certain cases, materialise on time scales below that of a single orbit (100 minutes) [57, 41]. Under such conditions, using large fit-spans risks smoothing signals of operational interest. In a modelling context, the fit-span should therefore be selected to preserve the temporal scales of interest while reducing retrieval variance, rather than to minimise variance alone.

The relationship between retrieval error and drag acceleration magnitude is well described by a logistic curve (figure 5(a); equation 6), with a low-drag plateau, a transition region in which each $\sim 3 \times$ increase in drag acceleration reduces error by approximately one half, and a high-drag floor with diminishing returns. In our experiments, this steep region lies between 2×10^{-8} – $3 \times 10^{-7} \text{ m s}^{-2}$.

The high-drag floor primarily reflects effects intrinsic to the EDR method, and thus still persist even in a fully simulated, internally consistent setup. First, by construction ρ_{eff} is a weighted mean of $C_D A V_{\text{rel}}^3$ (see equation 4), whereas our reference “truth” is the simple mean of HASDM densities along the orbit. These coincide perfectly only if the weight $C_D A V_{\text{rel}}^3$ is constant or statistically independent of ρ ; in practice V_{rel} varies along the arc and ρ varies with latitude and local time, so a small bias between the two averages remains even as random noise vanishes. Second, finite perigee-to-perigee segmentation and discrete quadrature of $\epsilon(t)$ and the denominator in equation 4 introduce

small, drag-independent biases. As drag increases and random error becomes negligible, this methodological floor dominates, producing the observed saturation.

Because the trajectories and force model are internally consistent in our simulations, residual non-drag mismodelling should not contribute; however, in operational POD data such effects can matter [47]. While precise thresholds are expected to vary with satellite geometry, area-to-mass ratio, attitude law, and the chosen reference (“truth”) density model, the overall shape of the relation is consistent with prior studies on the sensitivity of EDR to POD noise [48]. Although spacecraft operators typically aim to minimise the effects of drag on their spacecraft, the results presented here indicate that in some cases it may be advantageous to deliberately increase the drag experienced by a subset of platforms within a constellation. This could be achieved, for example, by maintaining a slightly lower orbital altitude or by temporarily increasing the projected area into the atmospheric flow. A stronger drag signal enhances the accuracy of density retrievals, and this in turn improves the quality of the underlying assimilative model. Better density characterization will propagate into reduced state uncertainties and orbit propagation errors. In turn, these can reduce manoeuvre frequency, for instance by improving the reliability of conjunction probability assessments [1, 9, 8].

Constellations typically provide multiple satellites distributed along a common orbital track. Exploiting this geometry offers a third lever via platform count, an effect that has been less explored in the literature. For direct (unassimilated) means of simultaneous density observations, we find that the error drops according to the relation: $E(N) \sim N^{-0.6}$ (figure 7), which is slightly steeper than the $N^{-1/2}$ reduction expected of independent, identically distributed errors. This behaviour may be suggestive of partial phase cancellation: error components with systematic orbital-phase structure (i.e. driven by geometry or GDOP modulation) tend to have opposite signs at different phase offsets; equal spacing of satellites therefore cancels part of this structured component when averaging, in addition to reducing random variance (see section 2.3). In contrast, averaging individually *assimilated* estimates yields a much shallower $N^{-0.25}$ decay in error as satellites are added. In the results presented, this is likely indicative of background-model bias leakage. When assimilation is performed per-satellite (method outlined in section 2.6.1), common biases in the background density model are not removed and are re-introduced in each density time series, so averaging more assimilated series does not cancel the shared component. This behaviour may be influenced by the fact that each orbital plane provides a limited picture of the storm-time variations. It may be that the specific orbital configuration used (see table 1) was not placed in the most optimal way to resolve the storm behaviour. Expanding the scheme to include simultaneous assimilation of multiple satellite data streams of various orbital geometries are expected to recover a steeper N -scaling than the per-satellite baseline presented here. The exponent $b \approx 0.25$ (in equation 9) should therefore be regarded as a lower bound for assimilated pipelines that do not explicitly address common error sources.

A practical finding for current operators is that, along a well-sampled orbital path, the simple arithmetic mean of several unassimilated EDR-based density retrievals can provide highly accurate along-track density estimates (depending on POD noise and number of satellites used). This does not diminish the role of assimilation, which remains essential for producing a coherent four-dimensional density field and for propagating information to unsampled regions [53, 15, 54]. Rather, it identifies useful cases in which assimilation may not even be necessary. Firstly, if the objective is simply the specification of density along a single orbital track, direct averaging can provide high quality density information at relatively low computational cost and technical complexity. Secondly, this “string-of-pearls” configuration also enables effective anomaly detection and outlier removal by comparing each satellite to its in-plane peers. Anomalies could for example be the result of unanticipated thruster or attitude manoeuvres or a bad onboard GNSS solution.

Taken together, the fitted relationships (equation 10) provide a practical first-order mapping from performance requirements to design choices. For large constellations, even a modest, well-distributed subset of higher-quality POD solutions should be able to deliver high absolute accuracy density estimates (approaching $\leq 5\%$ RMS% under favourable conditions) when coupled appropriately with other processing parameters.

Several assumptions bound the applicability of these results. Using HASDM as the reference density model likely influences the error floor and the apparent logistic inflection (figure 5); repeating the analysis with alternative reference density models (e.g., NRLMSISE-00, JB2008, WAM-IPE) will shift thresholds but is unlikely to alter the qualitative relationships described. Increasing the physical fidelity of the drag model (time-varying area, realistic attitude law, time-varying C_D) will introduce additional variability, primarily through modulation of the drag signal [17]. While the noise model used in this study advances beyond a white-noise assumption, it does not include all operational non-stationarities (e.g. receiver duty cycles, clock events). Moreover, absolute noise levels were drawn from a limited set of literature sources because independent validation of commercial-class POD remains scarce.

Further advances in this line of work are likely to be limited by the lack of public characterisation of commercial-class LEO POD quality [4, 51]. Reliable assessments of density retrieval method performance depend on credible, independently validated error budgets. Science missions routinely employ SLR for this purpose [25, 35]; equipping even a small number of spacecraft within large constellations with SLR retro-reflectors, would materially strengthen the certainty and quality of density retrievals and accelerate improvements in processing and assimilation. Wider dissemination of low-latency LEO POD data (both commercial and scientific), where licensing permits, would also further advance work in this field.

5. CONCLUSIONS

This study establishes a quantitative, operational link between three controllable levers: fit-span, satellite count, and drag regime, and the accuracy of orbit-averaged density retrievals from POD using the EDR method. By replacing an isotropic white-noise surrogate with a literature-informed higher-fidelity POD error model (anisotropy, pink along-track colouring, short-range autocorrelation, GDOP-modulated heteroscedasticity, and a small along-track bias) and by assessing a per-satellite EnKF that adjusts A_p to a background model (NRLMSISE-00), we recover analytical scalings which aim to support the design of next-generation assimilative pipelines.

First, the fit-span lever behaves as expected from variance-reduction principles but remains effective even under realistic error models: every doubling of the EDR fit-span reduces RMS-percentage error by approximately one half. This scaling persisted across the noise levels tested, indicating that successive orbit-mean observations contain sufficiently independent information even when along-track POD errors are somewhat autocorrelated. The practical consequence for model designers is not to maximise fit-span indiscriminately, but to select the shortest span which still preserves storm-time variability of interest while still meeting accuracy targets; over-smoothing forfeits the temporal resolution of signals that matter operationally.

Second, the drag-signal lever follows a logistic-like response. At very low drag accelerations ($2 \times 10^{-8} \text{ ms}^{-2}$) increasing the drag signal yields only marginal improvement. Across a transition band of approximately 2×10^{-8} to $2 \times 10^{-7} \text{ ms}^{-2}$, each $\sim 3\times$ increase in drag delivers a $\sim 50\%$ reduction in error. Beyond 10^{-6} ms^{-2} the curve approaches a floor, reflecting diminishing returns.

Many operational and planned satellite constellations span this transition region (450–650 km); modest changes in operating altitude or in the effective projected area of the satellite (e.g. via changes in attitude) can be used to tune the strength of the drag acceleration and achieve density retrieval accuracy improvements. In practice, however, both levers are tightly constrained by primary mission requirements (e.g. lifetime, coverage, pointing for power/communications), so density retrieval is likely to remain a secondary objective. Nevertheless, awareness of these scalings enables opportunistic drag tuning: operators may exploit existing attitude modes or scheduled operations that transiently change projected area (e.g. downlink pointing, momentum dumps, sun-pointing/charging attitudes) or natural altitude dispersions within their broader constellation to harvest higher signal-to-noise EDR observations without the need for dedicated manoeuvres.

Thirdly, increasing the number of simultaneously used satellites provides a powerful lever. When density estimates from N equally spaced co-orbital satellites are combined, the direct (unassimilated) mean of those densities follows $E(N) \approx E(1)N^{-0.6}$: errors fall faster than the idealised $N^{-1/2}$ expectation associated with random noise. We suggest this may be due to partial cancellation of phase-structured components (e.g. GDOP-modulated signatures) in addition to random-variance reduction. When each satellite's time series is first assimilated independently and the results are then averaged, the improvement is shallower, $E(N) \approx E(1)N^{-0.25}$. In our configuration this behaviour is attributable to common background-model bias that persists across per-satellite runs and thus does not cancel under averaging. The implication is twofold. Along a well-sampled orbital path, a simple arithmetic mean of multiple unassimilated EDR retrievals can even exceed the average of individually assimilated series, provided the POD noise is low enough and N is sufficiently large. At the same time, assimilation remains essential for producing a coherent four-dimensional density field and propagating information away from sampled tracks; a joint, multi-satellite assimilation is expected to mitigate shared-bias effects and recover steeper N -scaling than the per-satellite baseline assessed here.

The fidelity with which POD errors are characterized and then modelled matters. For an equivalent absolute 3D RMS error level, the higher-fidelity noise model produced smaller density-retrieval errors than a white-noise approach. Simplistic noise models may channel error through pathways that are unrepresentative of operational POD products. This result motivates the need for further LEO POD data gathering and characterization efforts, especially of the kinds

that are used by commercial satellite operators.

The observed behaviours can be summarized in an analytical form that explains the bulk of the observed variance ($\geq 89\%$) and is provided in the aim to support design-to-requirement mapping:

$$E(d, f, N) = E_{\min} + \frac{E_{\max} - E_{\min}}{1 + (d/d_{50})^k} 2^{-(f-1)} N^{-b}$$

where d is orbit-averaged drag acceleration, f the fit-span (in orbits), and N the in-plane satellite count, with $b \approx 0.6$ for direct averaging and $b \approx 0.25$ for the per-satellite assimilation used here. This analytical approximation is able to capture $\sim 89\%$ of the variance in the simulated data.

6. FUTURE WORK

Further work would benefit from incorporating simultaneous multi-satellite assimilation, higher-fidelity force modelling (time-varying area and C_D), investigating the use of adaptive fit-spans, and a broader range of orbital geometries. We also plan to further validate the results presented herein against real POD data arcs and independent truth sources (e.g. accelerometer-derived densities), developing the results provided into a more detailed analysis which could be used as a document to support assimilative model development by enabling developers to quantitatively specify requirements (e.g. POD quality) and processing strategies for the next generation of assimilative thermospheric density models.

ACKNOWLEDGEMENTS

We thank Dr. Eric Sutton for insightful feedback on early results that shaped the design and interpretation of our experiments. We are also grateful to the participants of the 2025 Satellite Drag Meeting for constructive comments and discussion that greatly improved this work. Charles Constant is supported by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant EP/R513143/1. This project was also supported by the UCL Innovation and Enterprise Impact Acceleration Account 2022–27 (EPSRC grant EP/X525649/1). Dr. Marcin Pilinski serves as a consultant for Space Environment Technologies in matters related to satellite drag.

REFERENCES

- [1] Salvatore Alfano and Daniel Oltrogge. “Probability of Collision: Valuation, variability, visualization, and validity”. In: *Acta Astronautica* 148. July 2019 (2018), pp. 301–316. ISSN: 00945765. DOI: 10.1016/j.actaastro.2018.04.023. URL: <https://doi.org/10.1016/j.actaastro.2018.04.023>.
- [2] Andrea D Ambrosio et al. “Projected Orbital Demand and LEO Environmental Capacity Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA”. In: (2022).
- [3] Daniel Arnold et al. “Impact of Ionosphere on GPS-based Precise Orbit Determination of Low Earth Orbiters”. In: *AGU Fall Meeting Abstracts* Abstract G31B-1115 (2015). URL: <http://abstractsearch.agu.org/meetings/2015/FM/G31B-1115.html>.
- [4] Daniel Arnold et al. “Precise orbit determination of Spire nano satellites”. In: *Advances in Space Research* 72.11 (2023), pp. 5030–5046. ISSN: 18791948. DOI: 10.1016/j.asr.2023.10.012.
- [5] T. E. Berger et al. “Flying Through Uncertainty”. In: *Space Weather* 18.1 (Jan. 2020). ISSN: 15427390. DOI: 10.1029/2019SW002373.
- [6] T. E. Berger et al. “The Thermosphere Is a Drag: The 2022 Starlink Incident and the Threat of Geomagnetic Storms to Low Earth Orbit Space Operations”. In: *Space Weather* 21.3 (2023). ISSN: 15427390. DOI: 10.1029/2022SW003330.
- [7] Kaifeng Bi et al. “Accurate medium-range global weather forecasting with 3D neural networks”. In: *Nature* 619.7970 (2023), pp. 533–538. ISSN: 14764687. DOI: 10.1038/s41586-023-06185-3.
- [8] Romain Buchs and Marie-Valentine Florin. *Collision risk from space debris: Current status, challenges and response strategies*. Tech. rep. 2021. DOI: 10.5075/epfl-irgc-285976.

- [9]Charles D. Bussy-Virat, Aaron J. Ridley, and Joel W. Getchius. “Effects of Uncertainties in the Atmospheric Density on the Probability of Collision Between Space Objects”. In: *Space Weather* 16.5 (May 2018), pp. 519–537. ISSN: 15427390. DOI: 10.1029/2017SW001705.
- [10]Andrés Calabia, Shuanggen Jin, and Robert Tenzer. “A new GPS-based calibration of GRACE accelerometers using the arc-to-chord threshold uncovered sinusoidal disturbing signal”. In: *Aerospace Science and Technology* 45 (2015), pp. 265–271. ISSN: 12709638. DOI: 10.1016/j.ast.2015.05.013. URL: <http://dx.doi.org/10.1016/j.ast.2015.05.013>.
- [11]US Coast Guard Navigation Center. “NAVSTAR GPS User Equipment Introduction”. In: *Department of Defense Document MZ10298* September (1996), p. 215. URL: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:User+equipment+introduction#0>.
- [12]Charles Constant, Santosh Bhattarai, and Marek Ziebart. “Limitations of Current Practices in Uncooperative Space Surveillance : Analysis of Mega-Constellation Data Time-Series”. In: *AMOS Tech.* 2023.
- [13]GE Cook, Desmond King-Hele, and Doreen Walker. “The contraction of satellite orbits under the influence of air drag. II. With oblate atmosphere”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 264.1316 (1961), pp. 88–121. ISSN: 0080-4630. DOI: 10.1098/rspa.1961.0186.
- [14]J. R. Dormand and P. J. Prince. “A family of embedded Runge–Kutta formulae”. In: *Journal of Computational and Applied Mathematics* 6.1 (1980), pp. 19–26. DOI: 10.1016/0771-050X(80)90013-3. URL: <https://www.sciencedirect.com/science/article/pii/0771050X80900133>.
- [15]Sean Elvidge and Matthew J. Angling. “Using the local ensemble Transform Kalman Filter for upper atmospheric modelling”. In: *Journal of Space Weather and Space Climate* 9 (2019). ISSN: 21157251. DOI: 10.1051/swsc/2019018.
- [16]European Union. *Galileo Open Service Signal-In-Space interface control document (OS SIS ICD)*. Tech. rep. 2. 2023.
- [17]D J Fitzpatrick. “Applying Energy Dissipation Rate GNSS Accelerometry to a Non , Circular Orbiting Satellite”. In: (2024), pp. 1–22.
- [18]David J Fitzpatrick et al. “Harnessing Satellite Constellations as Signals of Opportunity for Atmospheric Forecasting and Enhanced Space Situational Awareness”. In: (2024). URL: www.amostech.com.
- [19]Boggs DH Folkner WM Williams JG. *The planetary and lunar ephemeris DE 421*. Jet Propulsion Laboratory, Pasadena, California, 2008.
- [20]Government Accountability Office. “Large Constellations of Satellites – Mitigating Environmental and Other Effects”. In: September (2022), p. 80.
- [21]Nils Gustafsson et al. “Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres”. In: *Quarterly Journal of the Royal Meteorological Society* 144.713 (2018), pp. 1218–1256. ISSN: 1477870X. DOI: 10.1002/qj.3179.
- [22]André Hauschild et al. “Orbit determination of sentinel-6A using the galileo high accuracy service test signal”. In: *GPS Solutions* 26.4 (2022). ISSN: 15211886. DOI: 10.1007/s10291-022-01312-5.
- [23]D.G. King-Hele and Doreen M C Walker. *The prediction of satellite lifetimes*. Tech. rep. Farnborough: Royal Aircraft Establishment, 1987, p. 69.
- [24]D. Kuang et al. “Measuring atmospheric density using GPS-LEO tracking data”. In: *Advances in Space Research* 53.2 (Jan. 2014), pp. 243–256. ISSN: 02731177. DOI: 10.1016/j.asr.2013.11.022.
- [25]P. Kuchynka et al. “Uncertainties in GPS-based operational orbit determination: A case study of the Sentinel-1 and Sentinel-2 satellites”. In: *Aeronautical Journal* 124.1276 (2020), pp. 888–901. ISSN: 00019240. DOI: 10.1017/aer.2020.8.
- [26]Thorsten Kurth et al. “FourCastNet: Accelerating Global High-Resolution Weather Forecasting Using Adaptive Fourier Neural Operators”. In: *Proceedings of the Platform for Advanced Scientific Computing Conference, PASC 2023* (2023). DOI: 10.1145/3592979.3593412.
- [27]Hugh G Lewis et al. “Critical Number of Spacecraft in Low Earth Orbit : A New Assessment of the Stability of The Orbital Environment”. In: April (2025).
- [28]Richard J. Licata and Piyush M. Mehta. “Reduced Order Probabilistic Emulation for Physics-Based Thermosphere Models”. In: *Space Weather* 21.5 (2023), pp. 1–19. ISSN: 15427390. DOI: 10.1029/2022SW003345.
- [29]Richard J. Licata, Piyush M. Mehta, and W. Kent Tobiska. “Impact of Space Weather Driver Forecast Uncertainty on Drag and Orbit Prediction”. In: *Advances in the Astronautical Sciences* 175.February (2021), pp. 1941–1959. ISSN: 00653438.

- [30]G. March et al. “CHAMP and GOCE thermospheric wind characterization with improved gas-surface interactions modelling”. In: *Advances in Space Research* 64.6 (2019), pp. 1225–1242. ISSN: 18791948. DOI: 10.1016/j.asr.2019.06.023. URL: <https://doi.org/10.1016/j.asr.2019.06.023>.
- [31]Piyush M. Mehta et al. “New density estimates derived using accelerometers on board the CHAMP and GRACE satellites”. In: *Space Weather* 15.4 (2017), pp. 558–576. ISSN: 15427390. DOI: 10.1002/2016SW001562.
- [32]Oliver Montenbruck and Eberhard Gill. *Satellite Orbits Models, Methods and Applications*. 3rd. Springer-Verlag Berlin, 2000. ISBN: 9783642635472.
- [33]Oliver Montenbruck, Florian Kunzi, and André Hauschild. “Performance assessment of GNSS-based real-time navigation for the Sentinel-6 spacecraft”. In: *GPS Solutions* 26.1 (2022), pp. 1–11. ISSN: 15211886. DOI: 10.1007/s10291-021-01198-9. URL: <https://doi.org/10.1007/s10291-021-01198-9>.
- [34]Oliver Montenbruck and Pere Ramos-Bosch. “Precision real-time navigation of LEO satellites using global positioning system measurements”. In: *GPS Solutions* 12.3 (2008), pp. 187–198. ISSN: 10805370. DOI: 10.1007/s10291-007-0080-x.
- [35]Oliver Montenbruck et al. “Sentinel-6A precise orbit determination using a combined GPS/Galileo receiver”. In: *Journal of Geodesy* 95.9 (2021). ISSN: 14321394. DOI: 10.1007/s00190-021-01563-z.
- [36]Ruben Morales-Ferre et al. “GDOP-based analysis of suitability of LEO constellations for future satellite-based positioning”. In: *WiSEE 2020 - 8th Annual IEEE International Conference on Wireless for Space and Extreme Environments, Proceedings* (2020), pp. 147–152. DOI: 10.1109/WiSEE44079.2020.9262624.
- [37]S. Mutschler et al. “First results of a new inversion tool for thermospheric neutral mass density computations during severe geomagnetic storms”. Submitted August 2025. 2025.
- [38]Shaylah Mutschler, Penina Axelrad, and Eric Sutton. “Application of SoleiTool for Density Estimation using CubeSat GPS Data”. In: (2021).
- [39]Shaylah M. Mutschler et al. “Physics-Based Approach to Thermospheric Density Estimation Using CubeSat GPS Data”. In: *Space Weather* 21.1 (2023). ISSN: 15427390. DOI: 10.1029/2021SW002997.
- [40]NASA. *Physical Oceanography Distributed Active Archive Center (podaac)*. 2025. URL: <https://podaac.jpl.nasa.gov>.
- [41]D. M. Oliveira et al. “Thermosphere Global Time Response to Geomagnetic Storms Caused by Coronal Mass Ejections”. In: *Journal of Geophysical Research: Space Physics* 122.10 (2017), pp. 762–10. ISSN: 21699402. DOI: 10.1002/2017JA024006.
- [42]William E. Parker et al. “Influences of Space Weather Forecasting Uncertainty on Satellite Conjunction Assessment”. In: *Space Weather* 22.7 (2024). ISSN: 15427390. DOI: 10.1029/2023SW003818.
- [43]Nikolaos K. Pavlis et al. “The development and evaluation of the Earth Gravitational Model 2008 (EGM2008)”. In: *Journal of Geophysical Research: Solid Earth* 117.B4 (2012), B04406. DOI: 10.1029/2011JB008916. URL: <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2011JB008916>.
- [44]J. M. Picone, J. T. Emmert, and J. L. Lean. “Thermospheric densities derived from spacecraft orbits: Accurate processing of two-line element sets”. In: *Journal of Geophysical Research: Space Physics* 110.A3 (2005), pp. 1–19. ISSN: 21699402. DOI: 10.1029/2004JA010585.
- [45]J. M. Picone et al. “NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues”. In: *Journal of Geophysical Research: Space Physics* 107.A12 (2002), pp. 1–16. ISSN: 21699402. DOI: 10.1029/2002JA009430.
- [46]Aubrey B Poore, Jeffrey M Aristoff, and Joshua T Horwood. “Covariance and Uncertainty Realism in Space Surveillance and Tracking”. In: *Working Group on Covariance Realism* (2016), p. 161.
- [47]Vishal Ray et al. “A long-term neutral density database using commercial satellite data for atmospheric model calibration Vishal Ray Kayhan Space Corp”. In: *Advanced Maui Optical and Space Surveillance Technologies (AMOS) Conference* (2023). URL: www.amostech.com.
- [48]Vishal Ray et al. “Error Assessment of Thermospheric Mass Density Retrieval With POD Products Using Different Strategies During Solar Minimum”. In: (2024), pp. 1–25. DOI: 10.1029/2023SW003585.
- [49]SAIC. *NAVSTAR GPS Space Segment/Navigation User Interfaces IS-GPS-200*. Tech. rep. 66. 2021, p. 16.
- [50]P. A. Schreiner, K. Neumayer, and R König. *GFZ Rapid Science and Near Real Time Orbit Products*. Tech. rep. 2022. DOI: <https://doi.org/10.48440/GFZ.B103-22067>.
- [51]Kannan Selvan et al. “Precise orbit determination of LEO satellites: a systematic review”. In: *GPS Solutions* 27.4 (2023), pp. 1–17. ISSN: 15211886. DOI: 10.1007/s10291-023-01520-7. URL: <https://doi.org/10.1007/s10291-023-01520-7>.

- [52]Christian Siemes et al. “New thermosphere neutral mass density and crosswind datasets from CHAMP, GRACE, and GRACE-FO”. In: *Journal of Space Weather and Space Climate* 13.November 2013 (2023), p. 16. ISSN: 21157251. DOI: 10.1051/swsc/2023014.
- [53]Mark F. Storz, Bruce R. Bowman, and Major James I. Branson. “High accuracy satellite drag model (HASDM)”. In: *AIAA/AAS Astrodynamics Specialist Conference and Exhibit* August (2002). DOI: 10.2514/6.2002-4886.
- [54]Eric K. Sutton. “A New Method of Physics-Based Data Assimilation for the Quiet and Disturbed Thermosphere”. In: *Space Weather* 16.6 (2018), pp. 736–753. ISSN: 15427390. DOI: 10.1002/2017SW001785.
- [55]Eric K. Sutton et al. “Toward Accurate Physics-Based Specifications of Neutral Density Using GNSS-Enabled Small Satellites”. In: *Space Weather* 19.6 (2021), pp. 1–15. ISSN: 15427390. DOI: 10.1029/2021SW002736.
- [56]W Kent Tobiska et al. “Understanding variability in HASDM to support space traffic management”. In: 2000 (2022).
- [57]Eftyhia Zesta and Denny M. Oliveira. “Thermospheric Heating and Cooling Times During Geomagnetic Storms, Including Extreme Events”. In: *Geophysical Research Letters* 46.22 (2019), pp. 12739–12746. ISSN: 19448007. DOI: 10.1029/2019GL085120.