GIMA

Geographical Information Management and Applications

Quantifying per-pixel radiometric uncertainties due to geolocation error: a CHIME case study

Mareike de Lanversin

Quantifying per-pixel radiometric uncertainties due to geolocation error

a CHIME case study

by

Mareike de Lanversin

Academic supervisor:	J. Timmermans
ESA supervisors:	I. Fernández Núñez, A. Gabriele
Responsible Professor:	P. van Oosterom
Project Duration:	September, 2025 - March, 2026

Acknowledgements

I would like to thank the CHIME team (Jens, Nacho, Laurent, Krisztina, Marco, Claudia, Antonio, Manfred, among others) for the opportunity to work on this topic and for welcoming me so openly into the team. Being involved in this research and learning about CHIME has been a really valuable experience.

A special thank you Ignacio (Nacho) Fernández Núñez for the weekly supervision, the many research ideas, and the enthusiasm he brought to our discussions. I am also grateful to Manfred Gawlas, my officemate and co-intern, for helping me think through the operational implications of this work, and for his contagious curiosity.

Finally, I would like to thank Dr.ir. Joris Timmermans for closely supervising this thesis; for the many discussions, coffees and cookies, and for helping shape this work into a strong MSc thesis. Thank you for your support throughout the highs and lows, and for showing me how engaging and rewarding research can be.

Abstract

Satellite remote sensing plays a critical role in biodiversity monitoring, environmental policy, and sustainable resource management by providing consistent, repeated and large-scale observations of the Earth's surface. As applications increasingly rely on quantitative, model-based analyses, users require transparent and reliable measures of uncertainty to ensure their confidence in derived products. Providing per-pixel uncertainty information is therefore essential to evaluate data reliability at the scale at which decisions are made and to enable traceable, reproducible use of satellite observations.

The Copernicus Hyperspectral Imaging Mission for the Environment (CHIME) is designed to meet these needs by delivering high-quality hyperspectral data together with detailed uncertainty products. A key component of this effort is the characterization of per-pixel uncertainties across the processing chain. However, current uncertainty models overlook the satellite geolocation error's contribution to radiometric uncertainty, which can lead to satellite data misinterpretations especially in areas of high radiometric contrast. This thesis addresses this gap by quantifying geolocation error induced radiometric uncertainties at a per-pixel level for the CHIME mission.

GenAI Statement

During the thesis process, I used generative AI (GenAI) tools to assist with code debugging, language editing, and LaTeX formatting. Specifically, I used GitHub Copilot within Visual Studio Code (based on OpenAI's GPT-4) for coding support and Google's NotebookLM for help in the writing process. These tools were used solely as support tools; all conceptual development, analysis, implementation, and final writing decisions remain my own responsibility.

Contents

1

Introduction

1.1. Food security and biodiversity loss

Despite current geopolitical tensions, environmental risks are projected to have the greatest societal impact over the coming decade. The Global Risks Report 2026 identifies biodiversity loss, ecosystem collapse, and food system disruptions among the most severe long-term threats to global stability (World Economic Forum, 2026).

Food systems and natural ecosystems are under increasing pressure from climate change, land degradation, and unsustainable resource use. Global progress toward ending hunger has stalled, and agricultural productivity is becoming increasingly vulnerable to climate variability and extreme weather events (Food and Agriculture Organization of the United Nations, 2023). At the same time, biodiversity decline is accelerating: up to one million species are estimated to be at risk of extinction, largely driven by land-use change and ecosystem degradation (Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, 2019). These challenges are interlinked, as agricultural expansion and intensification both respond to rising food demand and further contribute to environmental degradation, reinforcing a feedback loop between ecosystem decline and food system vulnerability.

Addressing food security while halting biodiversity loss therefore requires more sustainable, evidence-based land management. International frameworks such as the 2030 Agenda for Sustainable Development emphasize the importance of reliable environmental monitoring to guide policy and track progress (United Nations, 2015). In this context, Earth observation (EO) has emerged as a critical tool, providing spatially consistent and temporally repeated measurements of land surface conditions at scales relevant to both agricultural management and ecosystem monitoring (Drusch et al., 2012; Timmermans & Kissling, 2023). As monitoring objectives become more quantitative and decision-oriented, EO systems are increasingly required to deliver higher spatial resolution, improved radiometric accuracy, and robust uncertainty characterization.

1.2. EO Data Applications and User Requirements

Innovation in Earth observation for food security and biodiversity is fundamentally driven by user needs. Different users depend on reliable, and transparent EO data to guide sustainable management of our limited resources (Drusch et al., 2012; Gorroño et al., 2024; Timmermans & Kissling, 2023). Their requirements shape the design of missions such as the Copernicus Hyperspectral Imaging Mission for the Environment (CHIME).

To advance food security, farmers require repeated, high-resolution observations throughout the growing season to monitor crop development (Food and Agriculture Organization of the United Nations, 2023) and detect early signs of stress caused by drought, nutrient deficiencies, or pests. They rely on this information to make decisions that improve yield and reduce inputs, and this requires data with sufficient spatial and temporal resolution to retrieve accurate soil and vegetation variables. Likewise, to protect biodiversity, environmental managers also require high fidelity observations to monitor the

ecological state and impacts of ecological drivers and pressures and track progress toward sustainability (Knowledge Centre on Earth Observation, 2025). As such, their specific needs are spectral richness and the ability to quantify plant functional traits and soil parameters like chlorophyll content and soil moisture.

From these two fields of application, a set of common user requirements has been extracted (Nieke et al., 2023; Timmermans & Kissling, 2023; Verrelst et al., 2021):

- High spatial (10-30 m), temporal resolution (15-day revisit) and accuracy, as well as a global coverage, to capture environmental dynamics with precision and detail anywhere on earth.
- Transparency about data uncertainty, so that each measurement can be interpreted with an understanding of its confidence limits and reliability.

As user expectations evolve and applications demand finer discrimination of biophysical and biochemical variables, traditional EO systems face growing limitations. The next section traces how this led to a development from multispectral to hyperspectral imaging and why this transition represents a major leap in EO.

1.3. From Multispectral to Hyperspectral

Multispectral imaging has long been the foundation of EO, with missions such as Sentinel-2 providing data for land cover mapping, vegetation monitoring, and environmental assessment. However, multispectral sensors capture only a limited number of broad, non-contiguous spectral bands. This restricts their ability to detect subtle biochemical and structural differences between vegetation types or soil compositions. Because of this, many variables relevant for biodiversity, agriculture, and ecosystem monitoring are still poorly estimated (Qian, 2021).

In contrast, hyperspectral imaging (HSI) captures light in hundreds of narrow, contiguous bands (typically 5 to 10 nm wide) across the visible to shortwave infrared range (400-2500 nm). Each pixel therefore contains a continuous spectrum that represents a unique "fingerprint" of the observed surface (Figure 1.1). This spectral richness enables the precise retrieval of variables such as chlorophyll content or lignin concentration. HSI bridges the gap between remote sensing reflectance and detailed biophysical properties, enabling new measures to be made and which are key components for biodiversity and food security monitoring (Timmermans & Kissling, 2023; Vines & Zhang, 2022).

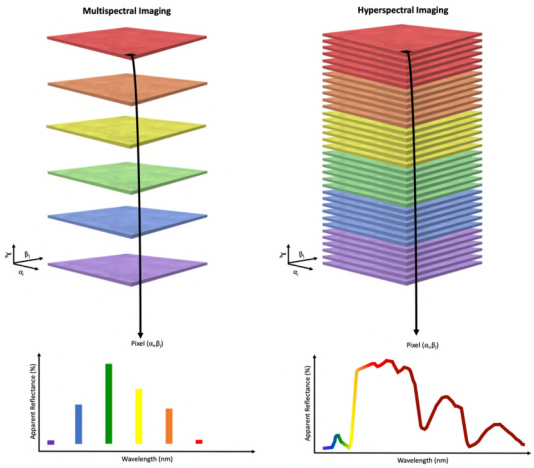


Figure 1.1: Illustration of the difference between multispectral and hyperspectral imaging (Figure retrieved from Vines and Zhang (2022))

While HSI offers unprecedented analytical potential, transforming spectral richness into operational, traceable products requires progress in instrumentation, calibration, and data management. Since the 1980s, hyperspectral imaging has evolved from early airborne sensors to pioneering spaceborne instruments such as Hyperion and CHRIS. These HSIs demonstrated the value of detailed spectral data but also revealed practical limitations: large data volumes, calibration complexity, and limited

uncertainty reporting. Therefore, the following missions made efforts to advance detector technology, radiometric calibration, and processing pipelines in order to improve data quality.

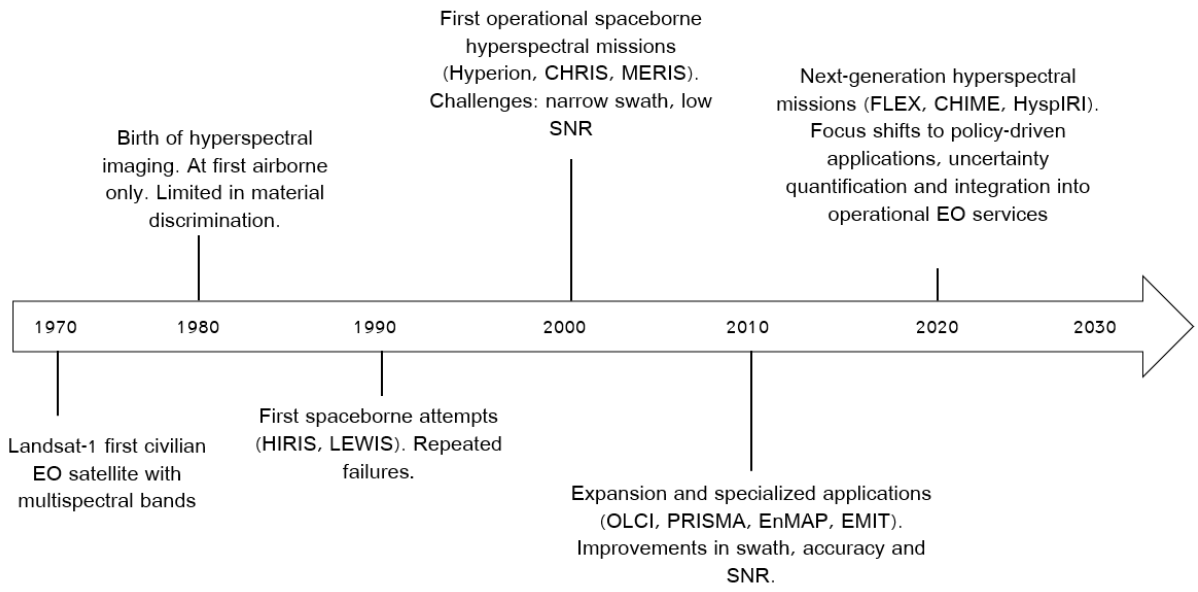


Figure 1.2: Timeline of hyperspectral imaging development

From these developments, three lessons were learnt:

a. HSI's spectral richness enables more accurate retrieval of biochemical and structural traits than multispectral systems.

b. Technological progress in sensors, calibration, and data handling improves performance but is challenged for large data volumes.

c. Transparent and per-pixel uncertainty quantification is essential to ensure data traceability and user confidence to map biodiversity (Timmermans & Kissling, 2023) and improve crop yields (Zheng et al., 2025).

These lessons paved the way for a new generation of hyperspectral missions designed to deliver operational, traceable products at global scales and culminating in CHIME.

1.4. The CHIME Mission

The lessons highlighted above paved the way for a new generation of hyperspectral missions designed to deliver operational, traceable products at global scales and culminating in CHIME.

CHIME is a next-generation hyperspectral satellite mission designed to deliver systematic, traceable observations that respond to emerging needs in ecological monitoring, natural resource management, and agricultural sustainability. It is also one of the first missions that aims to deliver per-pixel radiometric uncertainty data products along with its measurements, after Sentinel-2 and Sentinel-3 (Nieke et al., 2023).

To satisfy requires of both food security and biodiversity monitoring applications, CHIME is designed to acquire global hyperspectral data with 30 m spatial resolution, a 130 km swath, and a 12.5 day revisit, covering 400nm to 2500 nm with 256 contiguous bands at less than 10 nm intervals (Celesti et al., 2022; Rast et al., 2021).

CHIME's mission requirements (see Appendix A for full specifications) define its observation domain and minimum performance targets (Nieke et al., 2023):

- Observe all land and inland waters between 56° S - 84° N, including islands >100 km², coastal zones within 50 km of land, and shallow open waters (<50 m).

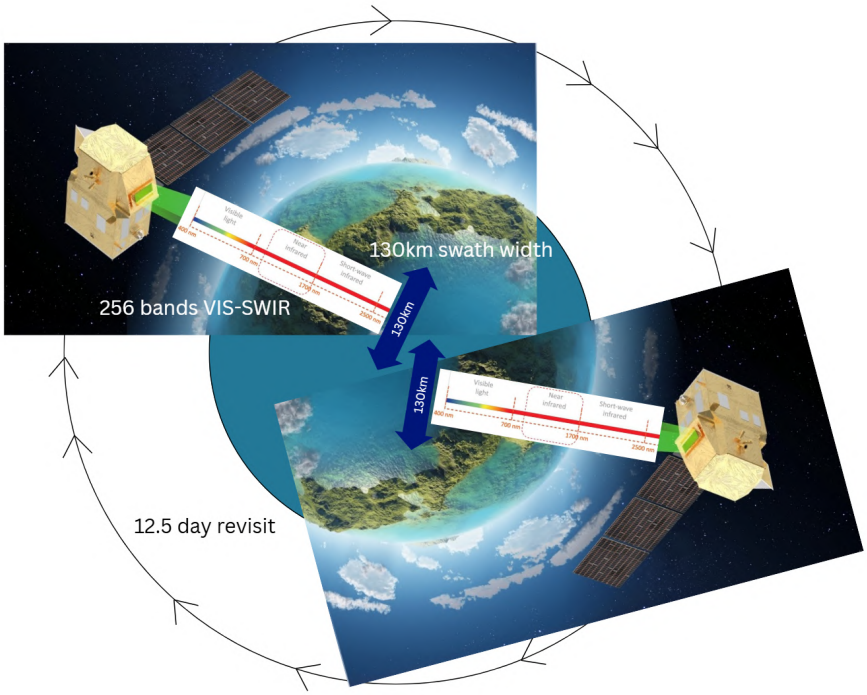


Figure 1.3: Visualization of CHIME Mission Main Characteristics

- Revisit every 12.5 days.
- Spectral coverage: 400–2500 nm at ≤ 10 nm bandwidth.
- Spatial resolution: 30 m with a > 0.25 MTF at Nyquist
- Radiometric accuracy: <5%.
- Spatial and spectral co-registration: <0.1 SSD/SSI.

Compared to multispectral sensors such as Sentinel-2 (which has 13 broad bands), CHIME's 256 narrow bands enable retrieval of previously inaccessible biophysical traits, including foliar nitrogen, phosphorus, and potassium, which are critical indicators of plant physiology and primary productivity. These mission requirements were explicitly defined to meet evolving user needs of greater spectral richness, calibration accuracy, and uncertainty transparency. Because of this, the CHIME products will support a wide range of applications, for example (Table 1.1):

Table 1.1: *CHIME Application domains*

Application Domain	Services Supported by CHIME	Quantifiable Variables (that CHIME can contribute to)
Sustainable agriculture & food security	Better services for food security and agriculture management (Rast et al., 2021)	Leaf-scale / canopy-scale vegetation variables: • Leaf Area Index (LAI) • Canopy chlorophyll content / Leaf chlorophyll content (LCC) • Leaf nitrogen concentration (LNC) or canopy nitrogen content (CNC) • Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) • Non-photosynthetic vegetation (NPV) fraction / dry biomass of litter or crop residue • Soil-vegetation interaction variables: soil organic carbon (SOC), topsoil moisture, soil texture/clay fraction in fields where vegetation is sparse or after harvest.
Raw materials exploration & management	Detailed mapping of mineral compositions and support for sustainable mining practices (Nieke et al., 2023)	• Mineralogical composition mappings: abundance of kaolinite, iron oxides, carbonates, sulfates at surface • Soil/regolith top-layer variables: clay/sand fraction, organic carbon, moisture, contaminants, soil texture, chemical state.
Biodiversity & management	State-and-change monitoring (Qian, 2021; Timmermans & Kissling, 2023)	• Vegetation structure/functional variables: LAI, canopy chlorophyll concentration, canopy nitrogen content, canopy dry mass content (CDMC) • Leaf-level variables: leaf chlorophyll content (LCC), leaf water content (LWC), leaf dry mass content (LDMC) • Fractional Vegetation Cover (FVC), non-photosynthetic vegetation (NPV) cover • Vegetation trait variables including lignin & cellulose (for NPV), functional plant traits (e.g., nutrient content)
Inland & coastal waters / aquatic ecosystem health	Monitoring of water-quality and aquatic ecosystem health (Qian, 2021)	• Water quality parameters: chlorophyll concentration in water, phytoplankton pigments • Optical properties of surface water: CDOM (coloured dissolved organic matter), suspended sediment concentration, diffuse attenuation coefficient (DAC) in water column. • Snow/ice optical properties in shallow waters or coastal zones: light absorbing impurities in snow/ice, snow grain size.
Environmental monitoring & hazards	Assessment of degradation, pollution and hazard impacts (Rast et al., 2021)	• Soil/surface degradation variables: soil organic carbon change, erosion indicators, contamination, salinity, bare soil fraction • Vegetation stress variables: chlorophyll loss, leaf water content decline, non-photosynthetic vegetation increase • Hazard-monitoring variables: burnt area cover, fire fuel dryness, mining tailings/waste detection, surface mineral change due to pollution.

Table 1.1: *CHIME Application domains*

Application Domain	Services Supported by CHIME	Quantifiable Variables (that CHIME can contribute to)
Forestry & cryosphere	CHIME is set to provide products related to snow and ice properties (Qian, 2021)	• Land cover/forest classification • Vegetation trait variables in forest canopy (chlorophyll, nitrogen, LAI) • Snow/ice cover and optical spectral indicators of snow/ice condition, reflectance properties, impurities in snow/ice • Atmospheric compounds (methane, CO_2)

1.5. Uncertainty Estimation

Satellite missions like CHIME provide a wealth of information about the Earth's surface, but the usefulness of these data depends not only on their radiometric accuracy, but also on how their associated uncertainties are quantified and communicated. Uncertainty estimation is the process of quantifying how confident we are in a measurement; in the context of satellite images, uncertainty estimation expresses how well a pixel's radiance value is recorded from space and later processed to represent a geophysical magnitude (e.g. reflectance, radiance, brightness, temperature) (Lamquin et al., 2013). As emphasized by the Quality Assurance Framework for Earth Observation (QA4EO), all satellite data products should be accompanied by traceable, quantified uncertainty estimates to allow users to judge whether the data are "fit for purpose" in their specific applications (Fox, 2010; Gorroño et al., 2024).

End users depend on reliable uncertainty information to make well-founded decisions. Without it, they risk misunderstanding the quality of the data and propagating hidden errors through models and analyses. For example, a crop growth model which uses satellite-derived leaf area index (LAI) data depends on local variations in illumination, viewing geometry, or atmospheric correction. If those variations are not known (or if they are incorrect), the outputs of the crop growth model such as estimated yields might be misinterpreted and lead to poor management decisions (Fernández et al., 2025; Hu, 2025). Thus, uncertainty is not a mere statistical abstraction - it directly influences scientific conclusions and operational reliability.

Traditionally however, most satellite missions have reported uncertainty at the scene or global level. This simplifies product validation but hides spatial variability, leaving users unaware of where measurements are more or less reliable. In heterogeneous landscapes such as patchy agricultural fields, mountainous terrain, or coastal zones, this can distort analyses and bias model results. As illustrated in Figure 1.4, assigning a single uncertainty value to an entire scene masks the local patterns of variability that are revealed when uncertainty is quantified for each pixel individually.

1.5.1. Per-pixel Uncertainty Estimation

Per-pixel radiometric uncertainty estimation represents a major step forward in Earth-observation missions data quality. Instead of assigning a single uncertainty value to an entire scene, per-pixel radiometric uncertainty estimation provides the uncertainty of each individual pixel by capturing local differences in illumination, viewing geometry, or atmospheric conditions and projecting these errors into a single spatial-metric (Fernández et al., 2025) (see Figure 1.4).

Per-pixel radiometric uncertainty maps are important for many applications. In sea-surface-temperature products, for instance, per-pixel uncertainty within a single acquisition can range from 0.1 K to 1.5 K depending on environmental and observational conditions (Gorroño et al., 2024). Using a single, scene-wide uncertainty value would obscure this variability and mislead downstream analyses. Similarly, in terrestrial monitoring, local uncertainty varies across field edges, heterogeneous canopies, and cloud-affected areas. This means that radiometric values at field edges are less reliable than those in the centres of fields - an insight critical for scientists developing crop-growth models that rely on these pixel-level radiance measurements.

What is more, at the instrument level radiometric uncertainties can be dependent on the dynamic range of hyperspectral instruments as both the signal to noise ratio (SNR) and calibration accuracy vary

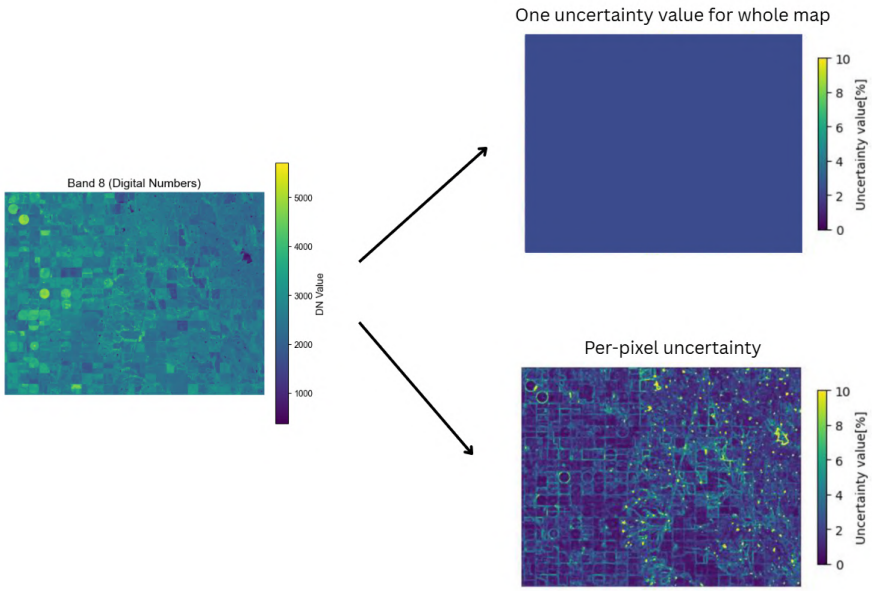


Figure 1.4: Example illustrating the difference between scene-wide and per-pixel uncertainty reporting. When a single uncertainty value is applied to an entire image (top right), spatial variability is lost. Per-pixel uncertainty (bottom right) reveals fine-scale differences in measurement reliability, allowing users to identify which parts of the scene are more uncertain.

depending on the incoming flux. Per-pixel uncertainty maps therefore enable more transparent, accurate and informed science by giving the tools for scientists to conduct traceable uncertainty propagation through their models and retrieval algorithms.

The current state-of-the-art approach to characterising per-pixel uncertainty has been adopted to the Sentinel-2 mission. Its Radiometric Uncertainty Tool (S2-RUT) applies the Guide to the Expression of Uncertainty in Measurement (GUM) to compute per-pixel radiometric uncertainties for both Level-1C (L1c, Top-of-Atmosphere) and Level-2A (L2a, surface-reflectance) products. While this approach does not provide operational uncertainty products, per-pixel uncertainty characterisation is openly accessible via the RUT extension in the ESA SNAP software via Github. At Level 1, S2-RUT accounts for major contributors such as instrument noise, ADC quantisation (±0.5 LSB), diffuser calibration, relative gain accuracy (≈ 0.4 %), and systematic effects diffuser ageing. At L2a, these terms are propagated through atmospheric correction, incorporating uncertainties from aerosol optical thickness, water vapour, adjacency effects, and surface anisotropy assumptions. Sentinel-2 also provides a spectral error-correlation matrix, enabling downstream algorithms such as LAI retrievals to correctly propagate uncertainties.

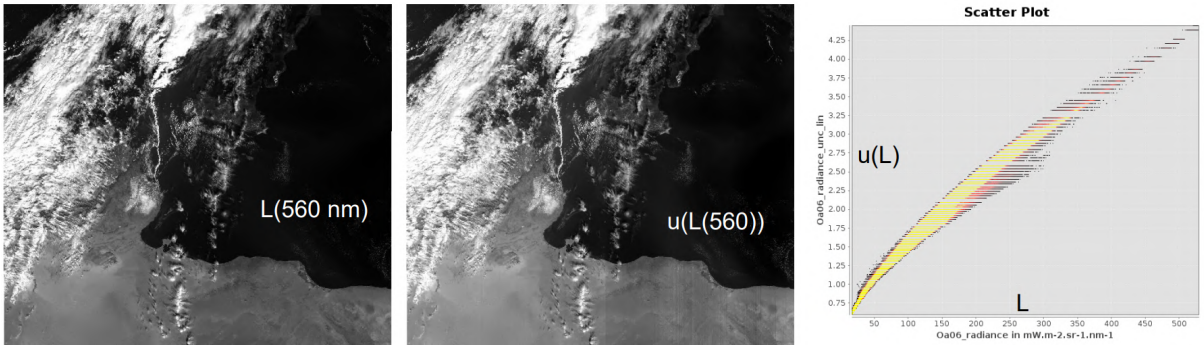


Figure 1.5: Example of per-pixel radiometric uncertainties for Sentinel-3 OLCI instrument at 560nm. Log-scaled to preserve relative accuracy (retrieved from Bourg and Laboratory (2023))

Since August 2022, OLCI Level-1B products also include per-pixel radiometric uncertainty estimates.

The uncertainty budget accounts for instrument noise and calibration-related effects, producing a per-pixel uncertainty map expressed in radiance units. However, this information is operationally available only at L1. The propagation of uncertainties to L2 ocean-colour products is not yet operational or fully validated, and users are advised to interpret L2 uncertainties with caution (Bourg & Laboratory, 2023; EUMETSAT & ESA, 2024).

By embedding per-pixel uncertainty into its products, Sentinel-2 and Sentinel-3 OLCI have set a benchmark for transparent and traceable radiometric quality. CHIME will build upon this foundation by extending the approach to hyperspectral data and including additional factors such as geolocation-induced radiometric uncertainty to increase user confidence and data reliability.

However, while Sentinel-2 and Sentinel-3 have paved the way for per-pixel uncertainty reporting, certain contributors to overall uncertainty remain uncharacterized. In particular, inaccurate satellite pointing (and correction) can lead to geolocation errors that are presently being overlooked (Roessel, 2024). Given CHIME's efforts to report uncertainties in a transparent and complete way, it is important that geolocation uncertainty is included in the overall radiometric uncertainty budget.

1.5.2. Research Objective

This thesis aims to improve the per-pixel estimation of the propagation of geolocation error induced radiometric uncertainty into radiometric uncertainty for CHIME L1 products. Building on Fernández et al. (2025)'s work, the research will improve the current geolocation uncertainty estimation efficiency and accuracy so that it can meet its performance targets.

2

Theory

To investigate how current geolocation uncertainty characterisations can be advanced, one needs to understand the theoretical basis that 1) explains the difference between errors and uncertainty, 2) identifies the different uncertainty components in hyperspectral monitoring, 3) delineates the propagation of uncertainty for CHIME. These are explained in the following paragraphs.

2.1. Error vs. Uncertainty

Error and uncertainty are related but distinct concepts. Error is defined as the difference between a measured value and the true value of a measurand, whereas uncertainty characterizes the expected magnitude and dispersion of that error. Uncertainty is typically expressed as a standard deviation and has to be estimated without direct knowledge of the true value (Povey & Grainger, 2015) (Figure 2.1).

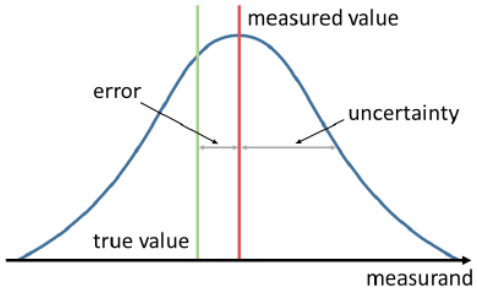


Figure 2.1: Simple illustration of error and uncertainty assuming a Gaussian uncertainty distribution (inspired from Povey and Grainger (2015)).

In controlled or simulated environments, such as during instrument design, calibration, and pre-launch performance assessment, the true value is either exactly known (for example through numerical simulations) or known with negligible uncertainty (like in laboratory reference standards). In these contexts, system performance can be evaluated directly in terms of errors, and satellite development therefore relies on performance error budgets in addition to uncertainty budgets.

In contrast, once a satellite is in orbit, the true value of the observed geophysical quantity is generally unknown. As a result, errors cannot be directly computed for routine satellite observations. Instead, satellite remote sensing products must report measurement uncertainty, which provides a quantitative, probabilistic description of the expected error based on instrument characterization, calibration knowledge, and processing assumptions.

Several multispectral missions, including Sentinel-2 and Sentinel-3 OLCI, have implemented per-pixel uncertainty estimation to support this requirement. However, for high-resolution hyperspectral missions such as CHIME, uncertainty estimation remains challenging due to the complexity of the measurement

process, the large data volumes involved, and the presence of multiple, strongly correlated sources of error.

2.2. Uncertainty Components in Hyperspectral Satellite Missions

The uncertainties described above arise at multiple stages along the hyperspectral measurement and processing chain, from photon collection and detection at the instrument level to calibration, geolocation, and higher-level processing. Figure 2.2 schematically illustrates how these uncertainties originate and propagate through the system.

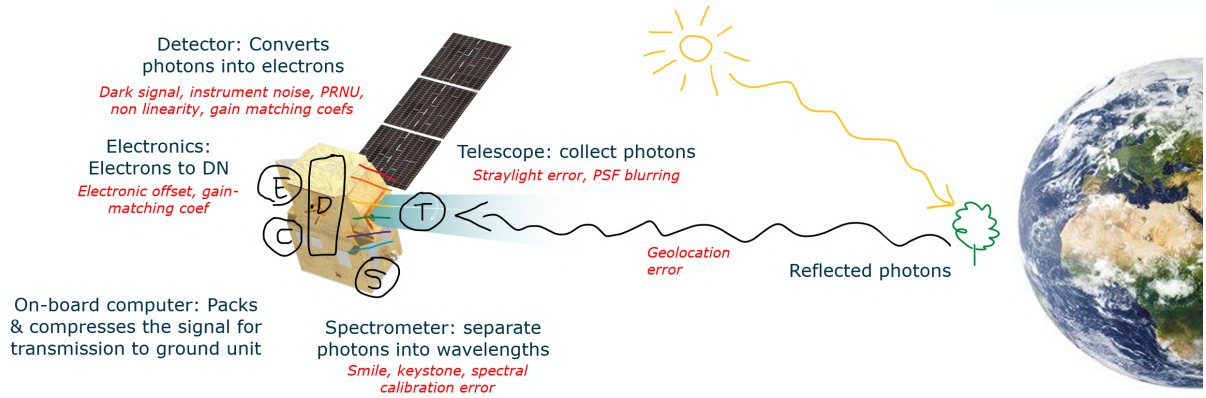


Figure 2.2: Schematic representation of uncertainty sources in CHIME and their propagation through the data acquisition and processing chain, from photon collection to calibrated radiometric products.

The different sources of uncertainty can be broadly categorised as follows:

- Sensor and instrumental uncertainties, originating at the detector, spectrometer, and telescope levels. These include photon noise, dark signal, readout and electronic noise, detector non-linearity, pixel response non-uniformity (PRNU), stray light, point spread function (PSF) blurring, and spectral effects such as smile and keystone.
- Calibration uncertainties, linked to imperfect knowledge of radiometric and spectral calibration parameters. These include uncertainties in gain and offset coefficients, reference diffuser characterisation, spectral calibration, and their temporal evolution due to instrument ageing and degradation.
- Digital number (DN) conversion uncertainties, affecting the transformation of raw digital counts into physical radiance or reflectance units. These uncertainties arise from the propagation of instrumental and calibration terms through the radiometric processing chain.
- Geometric and collocation uncertainties, related to knowledge of the spatial correspondence between measurements and the Earth's surface. Importantly, in EO it is not the absolute pointing knowledge error of the instrument that dominates uncertainty, but rather the pointing knowledge error, that is, uncertainty in where the instrument is believed to be pointing. Imperfect pointing knowledge, or pixel misregistration can lead to spectral mixing between neighbouring ground elements, particularly in heterogeneous scenes.

Each of these components contributes to the total radiometric uncertainty of hyperspectral measurements, with relative importance that may vary spatially, spectrally, and temporally. In hyperspectral systems, the large number of narrow, contiguous spectral bands further increases the complexity of uncertainty propagation, as many error sources exhibit strong spectral and spatial correlations.

2.3. Propagation of Uncertainty

In uncertainty analysis, each contributor is represented as a random variable with its own probability distribution. Instrument-related noise sources are commonly modelled as Gaussian, whereas other effects

like quantisation, digitisation or some calibration terms follow rectangular or uniform distributions. Systematic effects are handled via explicit correction terms in the measurement model; uncertainty associated with imperfect knowledge of those corrections is represented as standard uncertainty and propagated under GUM.

When several contributors act simultaneously, their combined effect is obtained through uncertainty propagation. Under the GUM framework, if the contributors are uncorrelated and the measurement model can be linearised, the combined standard uncertainty is given by the root-sum-square (RSS) of their individual uncertainties. When contributors are correlated, covariance terms must be included, as correlations can either amplify or reduce the total uncertainty: an important aspect for hyperspectral systems, where neighbouring spectral bands are often not fully independent.

Under the GUM framework, for a linearized measurement model, the expanded uncertainty is

$$U = k u_c(y), \quad u_c^2(y) = \sum_i u_i^2 + 2 \sum_{i<j} r_{ij} u_i u_j,$$

where u_i denotes the standard uncertainty contribution of term i expressed in output units and r_{ij} are correlation coefficients.

Building on these principles, Roessel (2024) developed an end-to-end approach for quantifying radiometric uncertainties in CHIME's L1 data products. The work focused on designing a Radiometric Uncertainty Tool (RUT) to model and compute per-pixel uncertainties in CHIME's Top-of-Atmosphere radiance (L1b) and reflectance (L1c) data. The primary uncertainty sources are introduced during the L1b and L1c processing steps, where the data are corrected for instrumental effects and calibration parameters to convert raw detector counts into meaningful radiance and reflectance values.

The methodology followed three main steps:

- Linking the radiometric model with the instrument and calibration parameters that influence radiance retrieval.
- Identifying and quantifying uncertainty contributors across the processing chain using physical models and metrological data.
- Combining these contributors following GUM principles to estimate the total per-pixel uncertainty.

The L1b measurement function expresses how raw digital counts are transformed into calibrated radiance values, with each correction step introducing its own uncertainty component (also to be seen in Figure 2.2):

$$L_{i,k}^{1b} = K_k^{rad} \cdot [K_{i,k}^{prnu} \cdot K_{i,k}^{gain} \cdot (NL_{i,k}(X_{i,k} - OS_{i,k}) - DS_{i,k} - SL_{i,k})] + 0$$

Each parameter in this function corresponds to a potential uncertainty contributor, summarised in Table 2.1 (See Appendix A for the complete uncertainty tree).

Table 2.1: Uncertainty contributors at L1b level included in the CHIME RUT.

L1b Contributor	Affected parameter
Gain Calibration Radiometric Knowledge	$K^{rad}{}_k$
Gain Calibration Temporal Knowledge	$K^{rad}{}_k$
Gain Calibration Spectral Knowledge	$K^{rad}{}_k$
Gain Calibration Geometric Knowledge	$K^{rad}{}_k$
Pixel Response Non-Uniformity Correction	$K^{prnu}{}_{i,k}$
Gain-matching coefficients	$K^{gain}{}_{i,k}$
Non-Linearity Correction	$NL_{i,k}$
Instrument Noise	*Raw counts* ($X_{i,k}$)
Crosstalk	*Raw counts* ($X_{i,k}$)
Polarisation	*Raw counts* ($X_{i,k}$)
Analog–Digital Conversion (ADC)	*Raw counts* ($X_{i,k}$)

Table 2.1: Uncertainty contributors at L1b level included in the CHIME RUT.

L1b Contributor	Affected parameter
Electronic Offset	$OS_{i,k}$
Dark Signal Correction	$DS_{i,k}$
Stray Light Correction	$SL_{i,k}$

2.4. Knowledge Gap

To date, advances in quantitative remote sensing have prioritized radiometric uncertainty linked to instrument performance, such as detector noise, calibration, and dark signal stability. Because these factors are easier to quantify using standard metrological frameworks like the GUM, they have been the primary focus of tools like the Sentinel-2 RUT and Sentinel-3 OLCI pipeline. Because of this, the impact of geolocation error on the final product uncertainty has remained comparatively under-explored (Fernández et al., 2025; Gorroño et al., 2024; Roessel, 2024).

The radiance recorded for a pixel is determined by where that pixel actually samples the Earth's surface (Figure 2.3). When the satellite's pointing knowledge is imperfect, the radiance value of a pixel may vary between acquisitions (temporally) due to slightly displaced ground locations, in other words geolocation error propagates to radiometric uncertainty (Fernández et al., 2025).

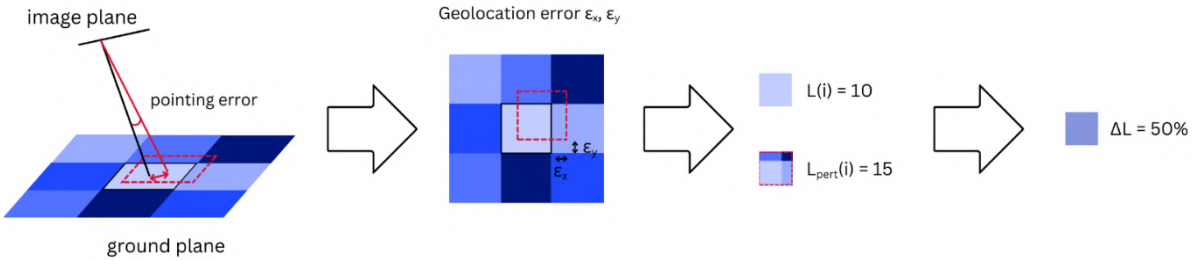


Figure 2.3: Schematic example of the propagation of geolocation error into radiometric uncertainty.

Because this effect scales with spatial contrast, geolocation uncertainty can become a dominant source of uncertainty in heterogeneous scenes with strong radiometric gradients (e.g. field edges, coastlines, small bright objects). In homogeneous areas, the effect is much smaller. Geolocation uncertainty is therefore fundamentally scene-dependent rather than instrument-dependent.

2.4.1. Deterministic formulation

Let $L(i, j)$ be the (spectral) radiance sampled on the image grid at pixel indices (i, j). The true sampling location is displaced by a random sub-pixel geolocation error $(\varepsilon_x, \varepsilon_y)$. The quantity of interest is the radiometric error random variable

$$\delta L(i, j) \equiv L_{\text{pert}}(i, j) - L(i, j),$$

where L_{pert} is the perturbed pixel value, the radiance after application of a geolocation shift.

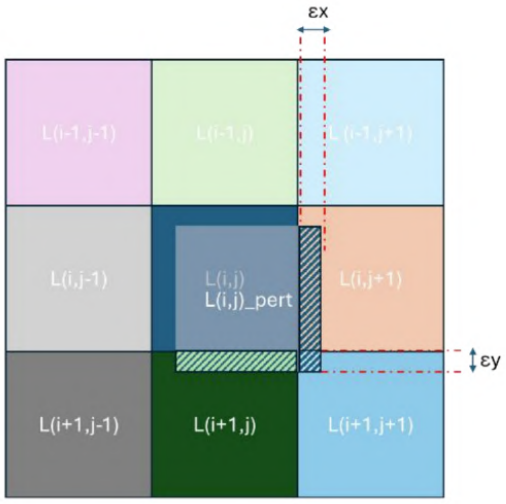


Figure 2.4: Depiction of bi-dimensional algorithm definition. Example where both components of geolocation error are positive (figure retrieved from Fernández et al. (2025)).

For a given displacement $(\varepsilon_x, \varepsilon_y)$ within a pixel, the perturbed value is modelled by bilinear interpolation of the surrounding radiance field. For the quadrant $\varepsilon_x \geq 0, \quad \varepsilon_y \geq 0$, the bilinear form yields

$$\delta L(i,j) = \varepsilon_x[L(i,j+1) - L(i,j)] + \varepsilon_y[L(i+1,j) - L(i,j)] + \varepsilon_x \varepsilon_y \tau,$$

with the corner term

$$\tau \;=\; L(i+1,j+1) - L(i+1,j) - L(i,j+1) + L(i,j).$$

This decomposition already reveals the key physical driver: geolocation uncertainty scales with local radiometric contrast, i.e. the spatial gradients of L and the magnitude of the pointing error.

2.4.2. Probabilistic framework

Building on this formulation, Fernández et al. (2025) derived a first-order probabilistic model of geolocation uncertainty (detailed in Appendix B). Assuming Gaussian geolocation knowledge errors and neglecting the second-order cross-term $\varepsilon_x \varepsilon_y \tau$, the resulting radiometric error distribution can be expressed as a finite mixture of signed half-normal contributions.

The shape of this distribution depends entirely on the local radiometric configuration (Figure 2.5). When neighbouring pixels exhibit mixed radiometry relative to the central pixel, the distribution is asymmetric but two-sided. When all adjacent pixels are brighter or darker, the distribution becomes strongly skewed and effectively one-sided under analytical approximations. This configuration dependence explains why geolocation uncertainty cannot, in general, be represented by a single symmetric variance term.

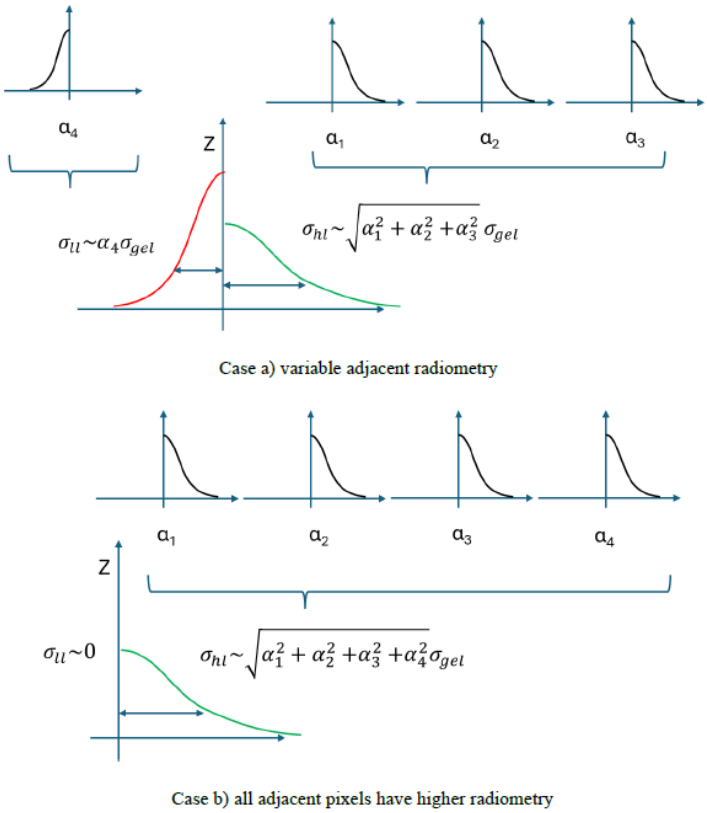


Figure 2.5: Conceptual illustration of geolocation error distributions under mixed and unidirectional adjacent radiometric configurations, highlighting the asymmetric and configuration-dependent nature of the error distribution.

2.4.3. Operational solvers

Direct Monte Carlo simulation of geolocation perturbations provides a numerical reference solution, but is computationally infeasible for operational hyperspectral processing. So, to enable operationally available per-pixel estimation, two solver strategies were introduced (see Appendix B for mathematical details):

- RSS solver: a fast analytical approximation based on root-sum-square aggregation of directional radiometric contrasts.
- Exact solver: a numerical approach that evaluates quantiles of the first-order mixture distribution via convolution and inversion.

The RSS solver is computationally efficient but structurally approximate, as it collapses the asymmetric mixture into Gaussian-equivalent one-sided scale parameters. The Exact solver is numerically consistent with the truncated first-order model but significantly more computationally demanding.

Although the abovementioned framework represents the state of the art, several limitations and open questions remained.

- First-order truncation and corner effects: The probabilistic formulation neglects the cross-term $\varepsilon_x \varepsilon_y \tau$ – however, in high-curvature scenes, τ may become large. In such cases, even when one geolocation component is small, the product $\varepsilon_x \varepsilon_y \tau$ may not be negligible relative to the linear terms. The practical validity of the first-order truncation therefore becomes scene-dependent.
- Use of product radiances: The derivation assumes access to the ideal radiance field L, whereas operational processing relies on perturbed product radiances. Although contrast equivalence holds under small-error conditions, its robustness across heterogeneous real scenes required empirical verification.
- Subpixel variability and PSF effects: The model assumes locally linear radiance behaviour at the

pixel scale. In practice, CHIME measurements are subject to PSF convolution and subpixel spatial variability. It was therefore necessary to assess whether these effects materially challenge the first-order formulation.

- Interpretation of solver discrepancies: Originally reported high relative errors were largely based on relative uncertainty width metrics. Relative metrics amplify deviations when the Monte Carlo reference uncertainty is small, potentially exaggerating discrepancies in low-contrast regions. This raised the question of whether solver discrepancies reflected structural modelling limitations or metric instability.
- Computational feasibility: The Exact solver improves consistency with the probabilistic model but remains computationally demanding, raising practical concerns for large hyperspectral data volumes.

Taken together, the state-of-the-art established a physically grounded first-order geolocation uncertainty estimation framework and operational solver strategies (Appendix B), but left unresolved questions regarding:

- The validity of the first-order truncation under realistic scene conditions,
- The impact of subpixel variability and PSF interactions,
- The reliability of product-based contrast estimation,
- The interpretation of solver discrepancies under robust validation metrics, and
- The balance between computational efficiency and numerical accuracy.

These open questions define the knowledge gap addressed in this thesis and motivate the research questions formulated below.

3

Research Questions

Building on the framework developed by Fernández et al. (2025), this thesis aims to improve the per-pixel estimation of geolocation uncertainty in CHIME L1 products. To achieve this, the following core research questions are defined:

- **RQ1** (quantification): What is the contribution of geolocation error to per-pixel radiometric uncertainty in CHIME L1 products?
- **RQ2** (distribution): What is the spatial distribution of per-pixel radiometric uncertainty attributable to geolocation uncertainty in CHIME Level-1 products?
- **RQ3** (algorithm development): How can the current geolocation uncertainty estimation algorithm be improved in terms of efficiency, accuracy, and robustness?
- **RQ4** (validation): How well do the improved analytical and numeric estimates agree with Monte Carlo simulations and proxy reference data across different scene types?

4

Methods

To address the research questions defined in the previous section, a structured methodology was developed and is summarised in Figure 4.1. The approach consists of three main components:

- Workflow development (RQ3): design and implementation of a modular geolocation uncertainty estimation framework allowing integration of baseline and newly proposed solvers.
- Workflow testing (RQ1-RQ3): evaluation of solver performance on synthetic grids and real scenes to analyse accuracy, spatial behaviour, and computational efficiency.
- Workflow validation (RQ2-RQ4): quantitative comparison of solver outputs against Monte Carlo reference simulations to assess bias, spread, and robustness across scene types.

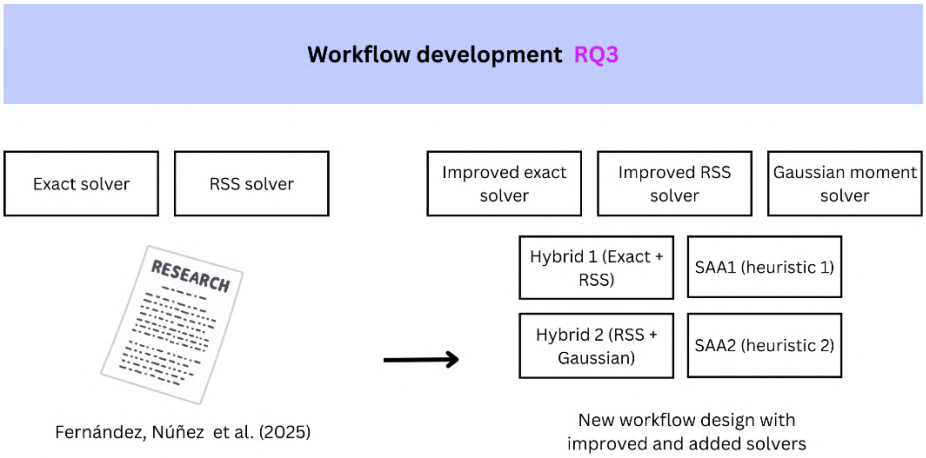
Each of these stages is described in detail in the following sections.

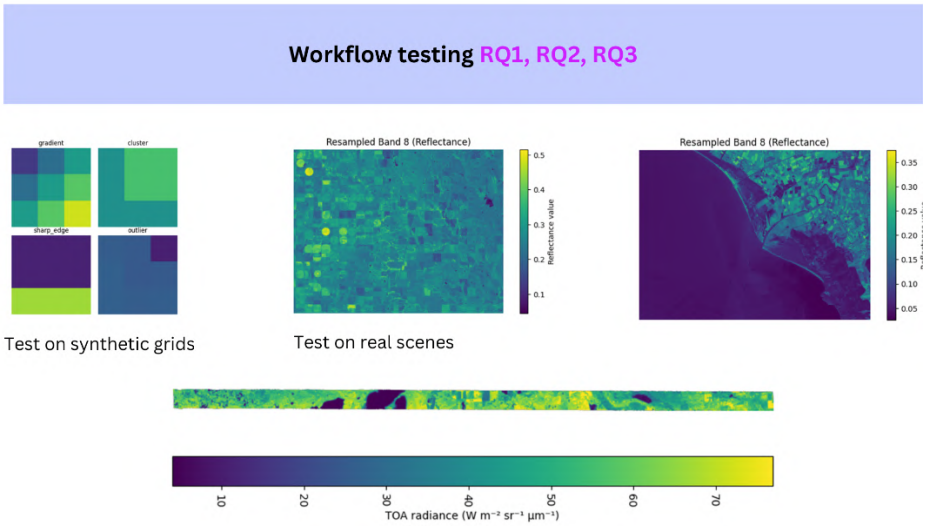
For reproducibility and methodological transparency, all implementation parameters are fixed and explicitly specified. Geolocation shifts are applied in pixel units and sampled from a zero-mean Gaussian distribution with standard deviation $\sigma_{\text{gel}} = 0.15$ pixels (corresponding to 0.30 px at 2σ). This value is selected to be consistent with typical post-refinement geolocation performance for Earth observation civil satellites (0.3–0.5 px at 2σ), as reported in Fernández et al. (2025).

An oversampling factor of 10 is applied throughout the workflow, consistent with the CHIME instrument design and its native oversampling strategy.

The instrument point spread function (PSF) is modelled as a Gaussian with standard deviation 7.3 pixels in the oversampled (3 m) geometry, corresponding to an effective spatial scale of approximately 22 m in the nominal grid. The Gaussian PSF assumption is consistent with the CHIME mission specification. The PSF kernel size is fixed at 41×41 pixels, ensuring adequate spatial support for the Gaussian model while avoiding truncation artefacts.

Pixels lacking all four direct neighbours are excluded from the analysis to maintain consistency with the first-order formulation described in Appendix B.





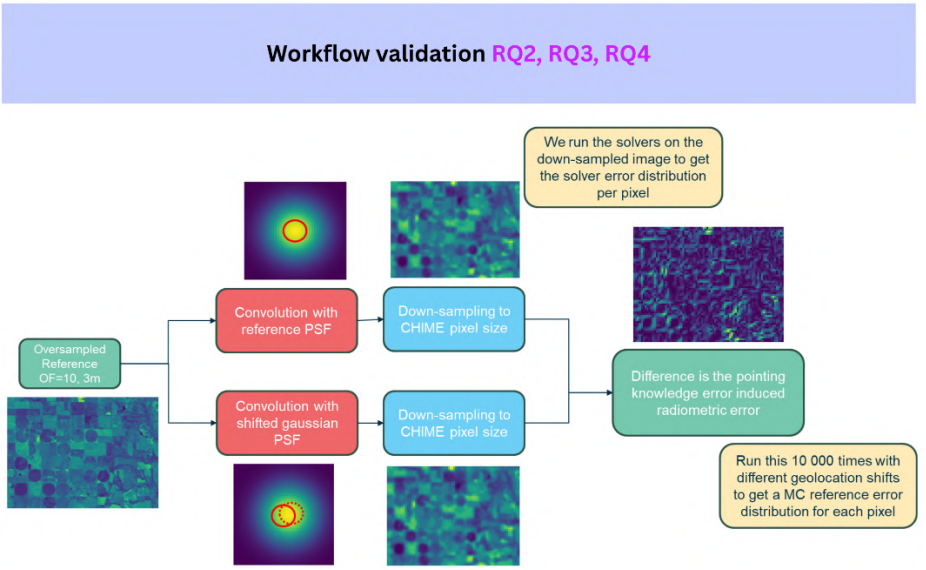


Figure 4.1: Overview of methodology

4.1. Workflow development

This section describes the solvers implemented in the workflow, their limitations, and the motivations for introducing additional methods. The focus is on algorithmic design and implementation choices; detailed mathematical formulations are provided in Chapter 5.

4.1.1. New solvers introduced

Five additional solvers are introduced to address the limitations of the baseline methods.

A Gaussian-moment solver performs a local Monte Carlo sampling of gaussian geolocation offsets and estimates the first two moments of the resulting radiance perturbation using bilinear resampling of the 3×3 patch. The error distribution is then approximated as a Gaussian using the estimated first two moments (mean and variance) to compute the 16^{th} and 84^{th} percentiles analytically. This approach is computationally cheaper than the MC per iteration and can be more stable than the exact and RSS solvers when linear gradient assumptions break down. Therefore, the Gaussian-moment solver is expected to perform well in scenes where the true error distribution is approximately symmetric. What is more, this solver indirectly accounts for second-order effects because it estimates the error distribution from a Monte Carlo distribution.

Two hybrid solvers are designed to combine complementary strengths of existing methods. Hybrid-1 applies the RSS solver by default and selectively upgrades pixels to the Exact solver based on contrast-derived criteria. Hybrid-2 similarly combines RSS with the Gaussian-moment solver. In both cases, solver switching thresholds are determined through sensitivity analysis to balance accuracy and runtime (the geolocation uncertainties estimated to be <2% are modelled with a simpler model like the RSS solver without introducing significant error in the total uncertainty estimate).

Two heuristic solvers, referred to as SAA1 and SAA2 (simple analytical solvers 1 and 2), are developed to address systematic biases observed in RSS estimates. SAA1 introduces scale factors and percentile adjustments to better match MC-derived bounds. SAA2 extends this formulation by adding coefficient-weighted linear terms intended to capture indirectly the second-order or curvature effects not represented in the first-order mixture model. Coefficients are determined empirically during development using the chosen scene and band.

Table 4.1: Overview of solvers developed.

Solver name	Solver type	Core idea	Uses MC information	Expected strengths	Main limitations
Gaussian-moment	Numerical	Approximates geolocation error distribution as Gaussian by matching first two moments of MC	Yes (during development, for moment estimation)	Accurate for near-symmetric error distributions; very fast	Cannot represent skewed or heavy-tailed distributions
Hybrid-1 (RSS + Exact)	Partially numerical	Applies RSS by default and for high uncertainty pixels switches to Exact solver	No (operationally)	Balances accuracy and runtime; preserves asymmetric bounds where needed	Requires threshold selection; still costly in highly textured scenes
Hybrid-2 (RSS + Gaussian)	Partially numerical	Applies RSS by default and for high uncertainty pixels switches to Gaussian moment solver	Yes (for Gaussian parameterization)	Faster than Hybrid-1; improves over RSS in moderately symmetric regimes	Limited accuracy in strongly skewed distributions
SAA1	Analytical	Applies scale factors and percentile corrections to RSS bounds	Yes (ad hoc for coefficient fitting)	Corrects systematic RSS bias with minimal added cost	Empirical; limited physical interpretability
SAA2	Analytical	Extends SAA1 with coefficient-weighted linear terms to capture second-order effects	Yes (ad hoc for coefficient fitting)	Best empirical match to MC across diverse contrast regimes	Most complex; coefficients scene-dependent

4.1.2. Engineering optimizations

Several engineering optimizations are implemented to enable large-scale evaluation of the Exact and Monte Carlo solvers within available computational resources.

The original Exact solver relied on SciPy's adaptive quadrature, which cannot be JIT-compiled and is prohibitively slow when executed in per-pixel loops. This was replaced with a Numba-compiled composite Simpson integrator, enabling vectorized execution and significant runtime reduction while preserving numerical accuracy.

The Monte Carlo implementation was optimized by replacing explicit Python loops over realizations with NumPy vectorized operations, reducing execution time from minutes to milliseconds for a fixed number of perturbations.

Quantile inversion in the Exact solver was made more robust by adapting the bracketing strategy to the asymmetric nature of the underlying distributions. Instead of using a fixed symmetric bracket, separate positive and negative scales are computed based on the sum of contributing standard deviations, and the root-finding interval is scaled accordingly. This prevents convergence failures in strongly asymmetric cases.

All implementations are benchmarked using single-core Python execution on CPU hardware. While operational implementations are expected to rely on multi-core C++ or GPU acceleration, these constraints ensure consistent algorithmic comparisons across solvers.

4.2. Workflow testing

Workflow testing is conducted to explore solver behaviour under controlled conditions and to identify numerical or structural failure modes prior to validation on real scenes.

4.2.1. Synthetic data: random grids

A dataset of 100 000 random 3×3 radiance grids is generated, with radiance values independently sampled from a uniform distribution between 0 and 100. A geolocation error standard deviation of 0.15 pixels is used, and 10 000 Monte Carlo realizations are generated per grid to form a reference distribution. These random grids represent texture-less, noise-dominated scenes and are used to stress-test solver robustness and numerical stability. Example grids are shown in Figure 4.2.

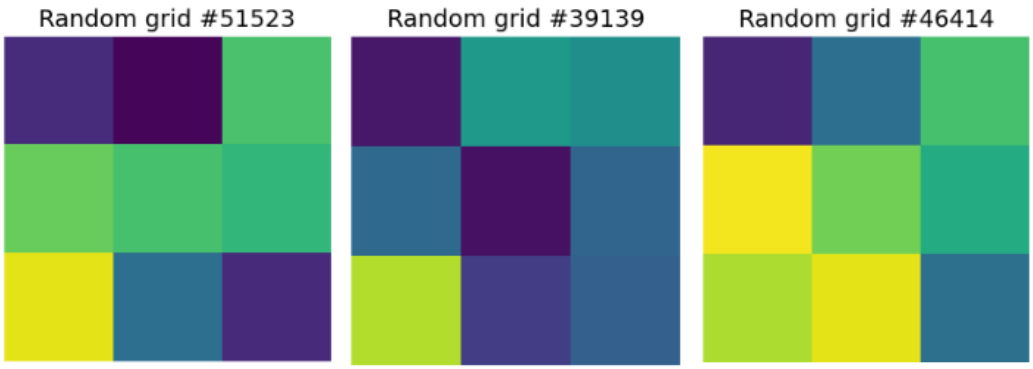


Figure 4.2: Subset of 3 randomly generated grids.

4.2.2. Synthetic data: patterned grids

To isolate the effect of spatial radiance patterns on solver performance, a second synthetic dataset is constructed using predefined spatial patterns. This step gives insight on algorithm limitations and performances under different conditions through in-vitro testing before the solvers are run on a real scene.

Four pattern classes are defined: smooth gradients, sharp edges, isolated outliers, and clustered regions. For each pattern type, radiance values are randomly sampled within predefined bounds, yielding a

dataset of 1 000 patterned grids. Representative examples are shown in Figure 4.3. Patterned scenes enable attribution of solver behaviour to identifiable contrast patterns which helps with interpretation of solver performance during validation.

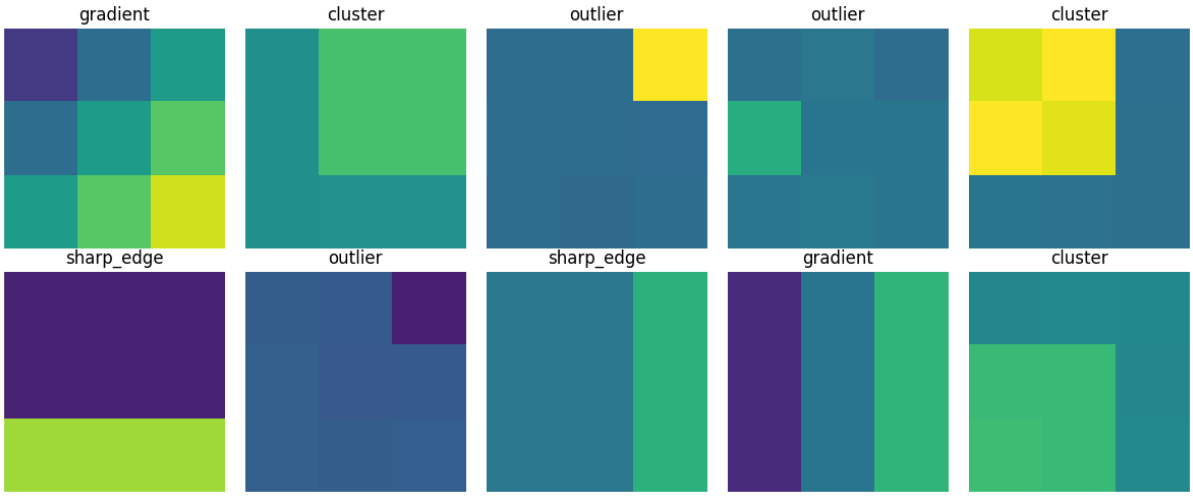


Figure 4.3: Subset of grids forming different characteristic patterns.

4.3. Workflow validation

4.3.1. CHIME-like forward model

To evaluate how the different algorithms estimated geolocation uncertainty, a framework was developed to output a MC geolocation uncertainty solver which outputs the "true" uncertainty for any scene. To achieve this goal, the validation framework developed retains the key physical aspects impacting the acquisition (i.e. instrument PSF, instrument down-sampling and geolocation error modelling) whilst having an imaging pipeline agile enough that could be run in an MC fashion. Validation is performed on real radiance scenes using a forward model that models the instrument PSF and then down samples the image to the CHIME 30m spatial resolution. The forward model is used to generate the MC reference image which is used for validation of solvers.

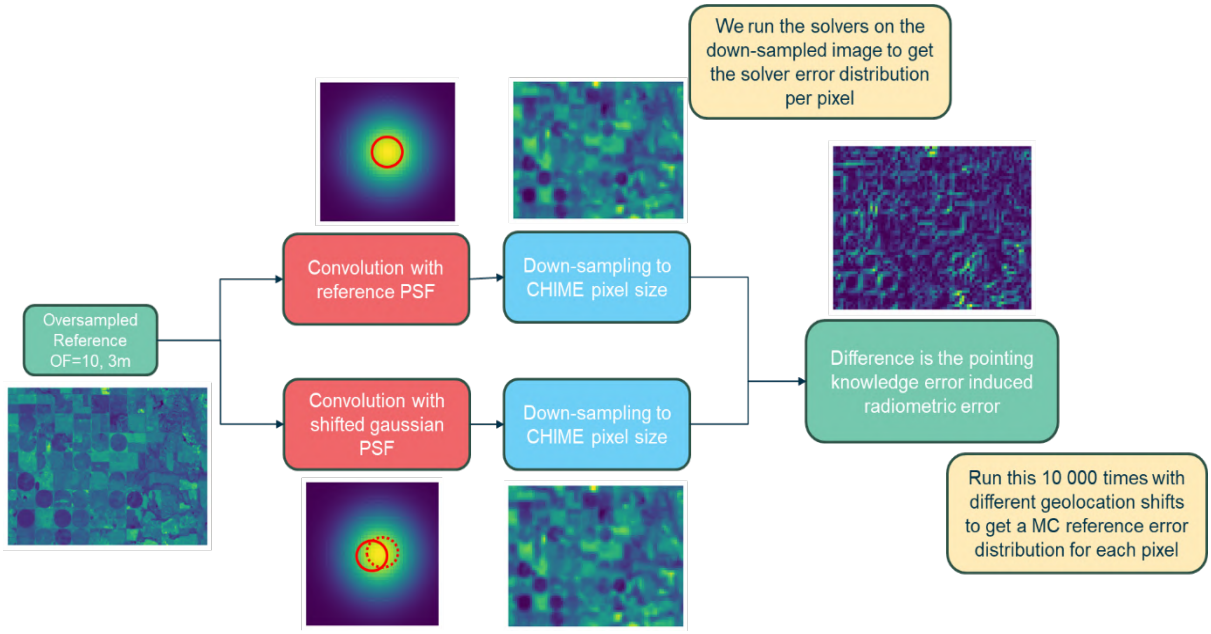


Figure 4.4: Overview of validation pipeline for generating the Monte Carlo baseline uncertainty map. The solvers are run on the down sampled scene with centred PSF.

Starting from an oversampled reference scene (OF = 10, 3 m), a Gaussian PSF convolution is applied to simulate the instrument response. A reference realization is generated using a centred PSF (upper part of Figure 4.4). For the Monte Carlo realizations, the PSF is spatially shifted (lower part of Figure 4.4), where the shift magnitude $(\varepsilon_x, \varepsilon_y)$ is sampled from a geolocation error distribution based on a given standard deviation based on known performances (σ_{gel}). All convolved images are then down sampled to the CHIME pixel size (30 m) via spatial averaging to emulate detector integration. The geolocation uncertainty is defined per pixel as the difference between each shifted down sampled realization and the reference down sampled image. This process is repeated for 10 000 realizations to estimate empirical per-pixel quantiles (mc_run = 10 000).

Initially, the framework was designed to include a pre-processing step in which the oversampled input scene was convolved with a top-hat kernel. The purpose of this was to suppress subpixel-scale spatial variability, following the considerations on subpixel effects discussed in Fernández Núñez (2025). However, after evaluating the behaviour of geolocation uncertainty with all modelled effects included, the solver performance was considered robust to subpixel variability for representative scenes. The additional top-hat convolution step was therefore omitted, and all subsequent validation experiments were conducted using the original, unmodified scenes.

4.3.2. Datasets

Validation is performed on two types of real scenes, selected to represent both previously studied cases and realistic CHIME-like observation conditions. In all cases, spectral bands with high radiometric contrast (bands in the near infrared, NIR) are intentionally chosen so that solver behaviour is tested under conditions where geolocation-induced radiometric uncertainty is also very high.

Sentinel-2 Band 8 in the NIR (10 m resolution) Two Sentinel-2 scenes over Grosseto (Italy) and Oklahoma (USA) are included to enable direct comparison with the study of Fernández et al. (2025). These scenes were provided by ESA and selected because they exhibit strong spatial and radiometric contrasts and because they represent in agricultural environments, which is representative of the anticipated CHIME application domain.

Band 8 (near-infrared, 10 m spatial resolution) is used for both scenes due to its sensitivity to vegetation structure, land–water boundaries, and crop heterogeneity, resulting in pronounced local radiance gradients. These characteristics make Band 8 particularly suitable for stressing geolocation uncertainty

solvers in regimes where uncertainty bounds are expected to be asymmetric and strongly scene dependent.

While Sentinel-2 provides high-quality, well-characterized imagery, its native spatial resolution is relatively close to the target CHIME resolution. Consequently, oversampling and subsequent PSF convolution and down sampling modify the original spatial statistics, limiting the fidelity of these scenes as strict CHIME analogues. For this reason, Sentinel-2 results are used primarily for contextual comparison and methodological continuity with prior work.

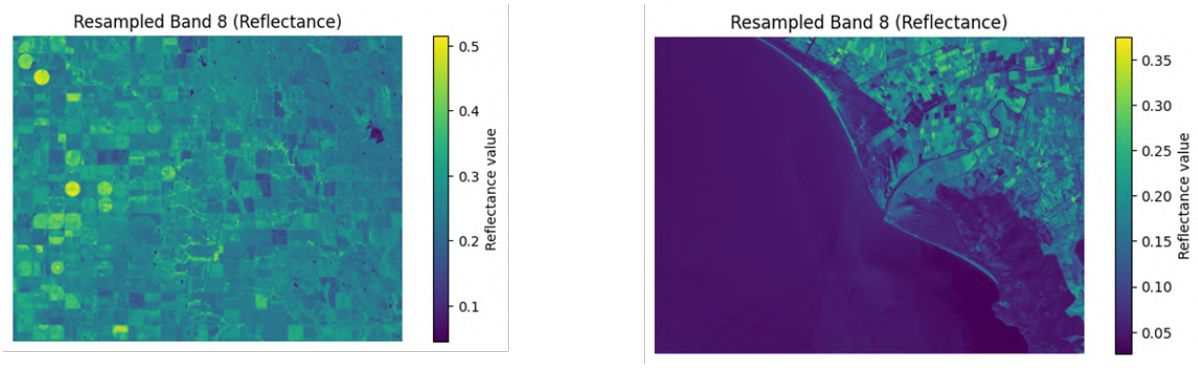


Figure 4.5: Left: Oklahoma scene (USA), right: Grossetto scene (Italy). Band 8 is used from the Sentinel-2 acquisition.

AVIRIS band 90 in the NIR (3 m resolution) As mentioned previously, the AVIRIS imagery at 3 m spatial resolution forms the primary validation dataset, as it allows realistic simulation of CHIME-like observations through down sampling by a factor of 10, which gives the final image's spatial resolution of approximately 30 m after PSF convolution.

Seven AVIRIS scenes were selected to span a wide range of biomes and radiometric contrasts relevant to CHIME science applications. They were downloaded via the AVIRIS data portal at MMGIS - AVIRIS and include scenes of alpine meadows, a California coastal region, high-biodiversity forest, a large agricultural area, mangroves in India, an agricultural region in Saskatchewan, and rainforest in India.

The selected AVIRIS scenes in NIR and RGB are shown in Figure 4.6 and in the Appendix C.

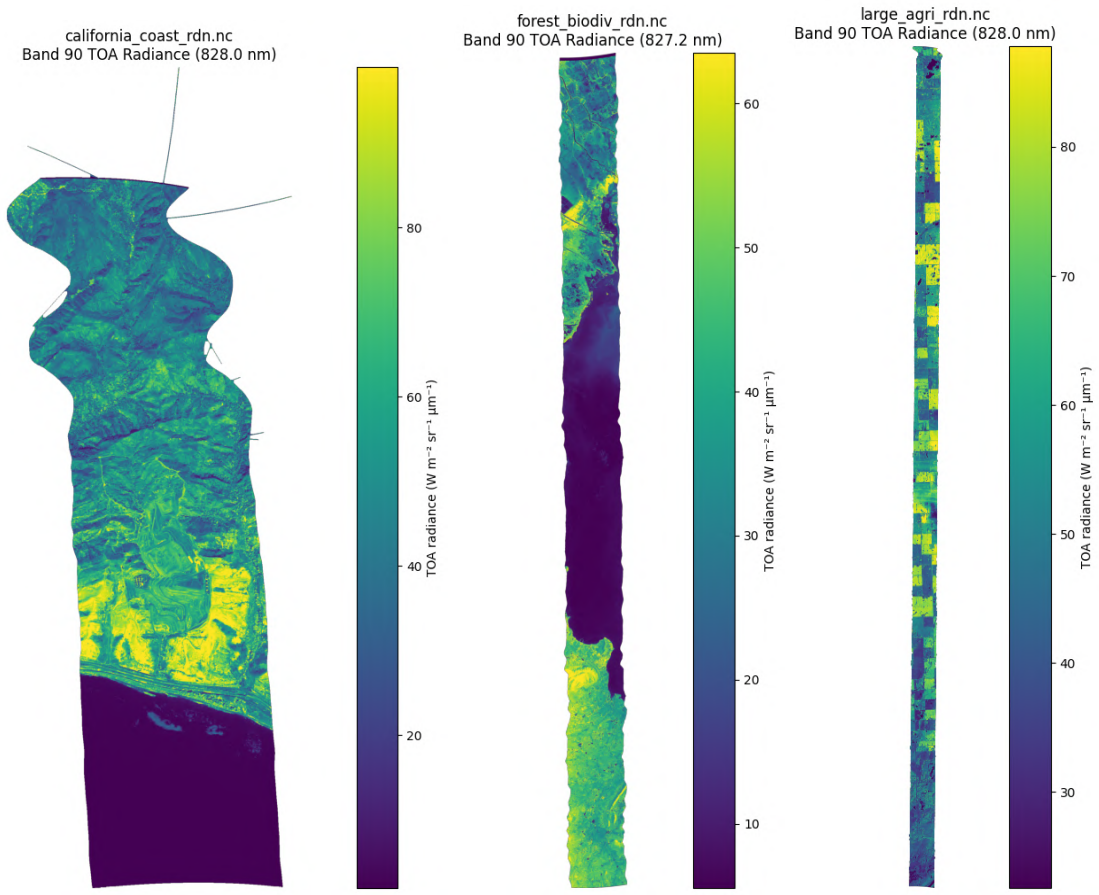


Figure 4.6: Left: Coastal scene in California (USA), middle: Forest with high biodiversity scene (India), right: large agricultural scene in Saskatchewan (Canada). AVIRIS band 90 is used and top-of-atmosphere radiance is depicted.

4.3.3. Validation metrics

Validation metrics are designed to quantify how accurately each solver estimates pointing-knowledge-induced radiometric uncertainty relative to a Monte Carlo reference, as well as to characterize the nature of any deviations. Because geolocation uncertainty distributions are generally asymmetric, evaluation focuses not only on uncertainty width but also on the individual lower and upper bounds.

Per-pixel error quantities For each pixel p, the Monte Carlo reference defines an empirical uncertainty interval through its 16th and 84th percentiles (high limit and low limit),

$$ll_p^{MC} = q_{16,p}^{MC}, \qquad hl_p^{MC} = q_{84,p}^{MC},$$

with corresponding width

$$w_p^{MC} = hl_p^{MC} - ll_p^{MC}.$$

For a given solver S, the estimated bounds and width are

$$ll_p^S, hl_p^S, w_p^S = hl_p^S - ll_p^S.$$

Validation is performed on the error of the uncertainty estimate, defined as the difference between solver-derived and MC-derived quantities:

$$\Delta ll_p = ll_p^S - ll_p^{MC}, \qquad \Delta hl = hl_p^S - hl_p^{MC}, \;\; \Delta w_p = w_p^S - w_p^{MC}.$$

These quantities measure, respectively, how well a solver reproduces the lower tail, upper tail, and total width of the geolocation uncertainty distribution.

All deviations are expressed in absolute percentage points, not as relative (normalized) errors. Relative errors were intentionally avoided because normalization by small reference values (so when ll_p^{MC}, hl_p^{MC}, or w_p^{MC} are close to zero) would artificially amplify deviations and lead to unstable or misleading performance indicators. Expressing errors in absolute percentage points ensures numerical stability and consistent interpretability across pixels with varying uncertainty magnitudes.

Figure 4.7 illustrates these error-of-the-error quantities schematically for a representative pixel, highlighting differences between solver-estimated bounds and the MC reference distribution.

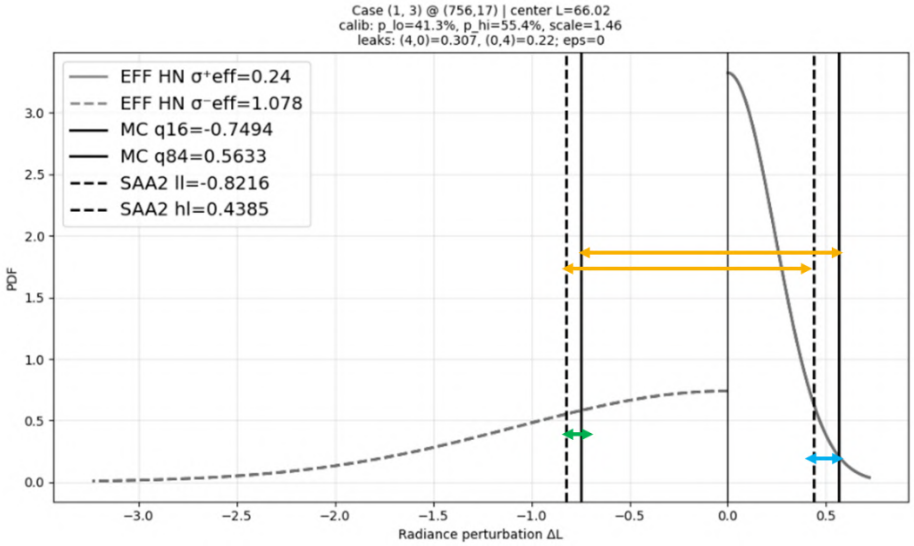


Figure 4.7: Schematic representation of validation metrics chosen to describe solver performance. Yellow: uncertainty width difference. Green: lower bound estimation difference. Blue: upper bound estimation difference.

Aggregate error metrics To summarize solver performance across all pixels in a scene, the distributions of Δll_p, Δhl_p, and Δw_p are characterized using the following statistics:

Mean Absolute Error (MAE):

$$MAE(\Delta x) = \frac{1}{N} \sum_{p=1}^{N} | \, \Delta x_p \, |,$$

where $\Delta x \in \{\Delta ll, \Delta hl, \Delta w\}$.

MAE measures the typical magnitude of solver deviation from the MC reference, independent of sign, and provides a good indicator of overall accuracy.

Root Mean Square Error (RMSE):

$$RMSE(\Delta x) = \sqrt{\frac{1}{N} \sum_{p=1}^{N} (\Delta x_p)^2}.$$

RMSE penalizes large deviations more strongly than MAE and is sensitive to outliers, making it useful for detecting rare but severe solver failures.

Mean (bias):

$$\mu(\Delta x) = \frac{1}{N} \sum_{p=1}^{N} \Delta x_p .$$

The mean quantifies systematic over- or under-estimation. For example, a positive mean in Δhl indicates a solver that systematically overestimates the upper uncertainty bound.

Median:

The median of Δx provides a robust measure of central tendency that is less sensitive to extreme values, particularly relevant for heavy-tailed error distributions.

Standard deviation:

$$\sigma(\Delta x) = \sqrt{\frac{1}{N} \sum_{p=1}^{N} (\Delta x_p - \mu(\Delta x))^2}.$$

The standard deviation characterizes the variability of solver errors across pixels and indicates scene-dependent sensitivity.

All metrics are computed separately for Δll, Δhl, and Δw, allowing asymmetric solver behavior to be explicitly quantified.

Rationale for metric selection The state-of-the-art validation approach focused only on uncertainty width as a validation metric. However, width-based metrics alone do not show the full picture of solver performance as geolocation error distributions are frequently skewed. By evaluating lower and upper bounds separately, the validation framework captures tail-specific biases that are directly relevant for downstream applications. For example, overestimation of the upper bound may lead to overly conservative uncertainty reporting, whereas underestimation of the lower bound may result in unrecognized radiometric risk. The combined use of width, bound-specific errors, and multiple summary statistics provides a comprehensive and physically interpretable assessment of solver performance.

Runtime metric In addition to accuracy, computational performance is evaluated through runtime measurements. Runtime in this study is defined as the time required to process a full hyperspectral data cube of dimensions $3661 \times 3661 \times 256$ pixels using single-core Python execution on local CPU hardware. This metric reflects algorithmic efficiency rather than operational performance; it enables consistent comparison between solvers under identical implementation constraints.

To assess scalability and hardware acceleration potential, additional benchmark tests were conducted in-house by an ESA CHIME team member on a reduced spectral subset (one quarter of the 256 bands) using both CPU and GPU implementations. These tests provide an indication of parallelisation benefits and spectral scaling behaviour, while maintaining comparability of solver structure. Operational runtimes are therefore expected to differ under optimized multi-core or GPU implementations, but the relative solver efficiency trends remain informative.

5

Results

This chapter presents the results of the modular geolocation uncertainty estimation workflow. The outcomes are structured according to the methodological stages introduced in Chapter 4: workflow development, workflow testing, and workflow validation. Results are reported descriptively, while their broader implications are addressed in Chapter 6.

5.1. Workflow development

5.1.1. Monte Carlo simulation structure

The Monte Carlo (MC) implementation serves as the numerical reference used for solver development and validation. Its purpose is to approximate, for each coarse pixel $p \equiv (i, j)$, the distribution of the geolocation errors as well as the error distribution's 16^{th} and 84^{th} quantiles, under an explicit forward imaging model (PSF blur, subpixel shift, and down sampling).

Forward model and sampling Let L^{nat} denote the radiance field at native (high) spatial sampling (AVIRIS 3 m GSD). A CHIME-like coarse pixel grid is obtained via (i) optical blur modelled by a Gaussian PSF, (ii) random geolocation perturbation, and (iii) spatial down sampling.

For Monte Carlo draw $k \in \{1, \ldots, M\}$, many 2-D shifts are sampled

$$\Delta x_k \sim \mathcal{N}(0, \sigma^2_{\text{nat}}), \qquad \Delta y_k \sim \mathcal{N}(0, \sigma^2_{\text{nat}}),$$

where the native-grid standard deviation of the geolocation shift is obtained by converting the coarse-grid standard deviation through the down sampling factor $DF = 10$:

$$\sigma_{\text{nat}} = DF \cdot \sigma_{\text{gel}}.$$

Correlation between Δx and Δy is neglected, consistent with the assumption of independent error components in the first-order geolocation uncertainty formulation (see Appendix B). While horizontal and vertical geolocation errors may in principle differ in magnitude and exhibit anisotropy, here an isotropic Gaussian model is adopted ($\sigma_x = \sigma_y$) consistent with the specified CHIME pointing knowledge performance. No privileged spatial direction is assumed. For typical residual geolocation errors after geometric refinement, this approximation is considered valid, as systematic directional biases are expected to be largely removed after geometric correction.

A shifted PSF kernel h_k is then generated (Gaussian with standard deviation σ_{psf} in native pixels, shifted by $(\Delta x_k, \Delta y_k)$), and applied through convolution:

$$\widetilde{L}^{\text{nat}}_k = L^{\text{nat}} * h_k.$$

Finally, coarse radiance is produced by block averaging over $DF \times DF$ native pixels. This step models

the detector sampling:

$$L_k^{\text{coarse}}(i,j) = \frac{1}{DF^2} \sum_{m=0}^{DF-1} \sum_{n=0}^{DF-1} \widetilde{L}_k^{\text{nat}}(DF \cdot i + m,\ DF \cdot j + n).$$

A reference (centred PSF) coarse radiance $L_{\text{ref}}^{\text{coarse}}$ is computed once using the same blur/down sampling pipeline but without geolocation perturbation. The purpose of this is to obtain both the non-perturbed image and the perturbed image, so that the difference can be computed as the error. For each coarse pixel, MC bounds are defined as empirical percentiles:

$$q_{16}^{MC}(i,j) = P_{16}\{\delta_k(i,j)\}_{k=1}^{M}, \qquad q_{84}^{MC}(i,j) = P_{84}\{\delta_k(i,j)\}_{k=1}^{M}.$$

5.1.2. Detailed mathematical description of solvers

All solvers operate on a coarse image L^{coarse} (the CHIME-like grid), as operationally the solvers will have as input only the product data and oversampled data is never accessible. For each valid pixel (i,j) with centre radiance $c = L^{\text{coarse}}(i,j)$, the four direct neighbours are:

$$n = \begin{bmatrix} L(i,j-1) \\ L(i,j+1) \\ L(i-1,j) \\ L(i+1,j) \end{bmatrix}, \qquad \alpha = n - c.$$

Solvers estimate percentile bounds (q_{16}, q_{84}) expressed in percent of centre radiance. Absolute radiance bounds $(\ell\ell, h\ell)$ are related by

$$(\ell\ell, h\ell) = \left(\frac{q_{16}}{100}c,\ \frac{q_{84}}{100}c\right).$$

Exact solver improvements Quantiles are obtained by solving

$$F_Z(z) = p, \qquad p \in \{0.16, 0.84\},$$

via bisection.

To ensure robust convergence for asymmetric mixtures, asymmetric support-aware bracketing is used. Let each signed half-normal contribution have scale σ_k and sign $s_k \in \{+1, -1\}$. Define the cumulative positive and negative scales:

$$S_+ = \sum_{k:s_k=+1} \sigma_k, \qquad S_- = \sum_{k:s_k=-1} \sigma_k.$$

The bracketing interval is

$$z_{\min} = -KS_-, \qquad z_{\max} = +KS_+,$$

with conservative tail factor K. This bracketing strategy avoids root-finding failures that were observed with symmetric brackets.

The resulting absolute bounds $(\ell\ell, h\ell)$ are then mapped to percent bounds by

$$q_{16} = 100\,\ell\ell/c, \qquad q_{84} = 100\,h\ell/c.$$

Gaussian-moment solver (local MC + Gaussian approximation) This solver approximates the local perturbation distribution by a Gaussian whose mean and variance are estimated from Monte Carlo resampling of a 3×3 patch using bilinear interpolation (without PSF and down sampling). This approximation allows the number of iterations to be considerably reduced compared to the full Monte Carlo method.

For pixel (i,j), extract the local patch $P \in \mathbb{R}^{3\times3}$ with centre $c = P_{1,1}$. Sample M shifts:

$$\varepsilon_x^{(k)}, \varepsilon_y^{(k)} \sim \mathcal{N}(0, \sigma_{\text{gel}}^2),$$

and compute bilinear-resampled values

$$\widehat{L}^{(k)} = \mathrm{Bilin}(P, \varepsilon_x^{(k)}, \varepsilon_y^{(k)}), \qquad \delta^{(k)} = \widehat{L}^{(k)} - c.$$

Estimate the first two moments (mean and standard deviation)

$$\widehat{\mu} = \frac{1}{M}\sum_{k=1}^{M} \delta^{(k)}, \qquad \widehat{\sigma} = \sqrt{\frac{1}{M}\sum_{k=1}^{M} (\delta^{(k)} - \widehat{\mu})^2}.$$

Then approximate quantiles using the normal inverse CDF:

$$\ell\ell = \widehat{\mu} + \widehat{\sigma}\,\Phi^{-1}(0.16), \qquad h\ell = \widehat{\mu} + \widehat{\sigma}\,\Phi^{-1}(0.84),$$

and map to percent.

Hybrid solvers (conditional refinement) Both hybrid solvers implement a two-pass strategy. In the first pass, the RSS solver is applied to all valid pixels to obtain preliminary percentile estimates $\left(q_{16}^{\mathrm{RSS}}, q_{84}^{\mathrm{RSS}}\right)$ and the corresponding uncertainty width

$$w^{\mathrm{RSS}} = q_{84}^{\mathrm{RSS}} - q_{16}^{\mathrm{RSS}}.$$

Pixels are flagged for refinement according to the criterion

$$w^{\mathrm{RSS}} > \tau,$$

where τ is a threshold in percentage points.

In the second pass, flagged pixels are refined using a higher-fidelity method:

- Hybrid-1: refine flagged pixels using the Exact solver.
- Hybrid-2: refine flagged pixels using the Gaussian-moment solver.

Both hybrids output (i) bound maps and (ii) a flagged-pixel mask, enabling reporting of refinement fraction.

The threshold parameter τ was selected by systematically evaluating candidate values

$$\tau \in \{0.5, 1.0, 1.5, 2.0, 3.0, 4.0\}\ \%.$$

(Figure 5.1) on the Oklahoma Sentinel-2 Band 8 scene. Rather than selecting τ solely based on refinement mask appearance, the choice was based on the relative magnitude of geolocation-induced uncertainty within the overall uncertainty budget.

Specifically, τ was chosen to be of the same order as the dominant contributors to the total width uncertainty. For the analysed scene, typical non-geolocation uncertainty contributions were on the order of 1–2%, with total uncertainty excluding geolocation effects around 4%. Selecting $\tau = 2\%$ therefore ensures that refinement is applied only where geolocation uncertainty is comparable in magnitude to other major contributors, thereby avoiding unnecessary computational overhead in regions where its contribution is negligible relative to the full error budget.

Visual inspection of the resulting refinement masks confirmed that $\tau = 2\%$ achieves an effective balance between physical relevance and computational efficiency in this high-contrast scenario. Future work could replace this fixed threshold with an adaptive strategy informed by local contrast or texture statistics of the scene.

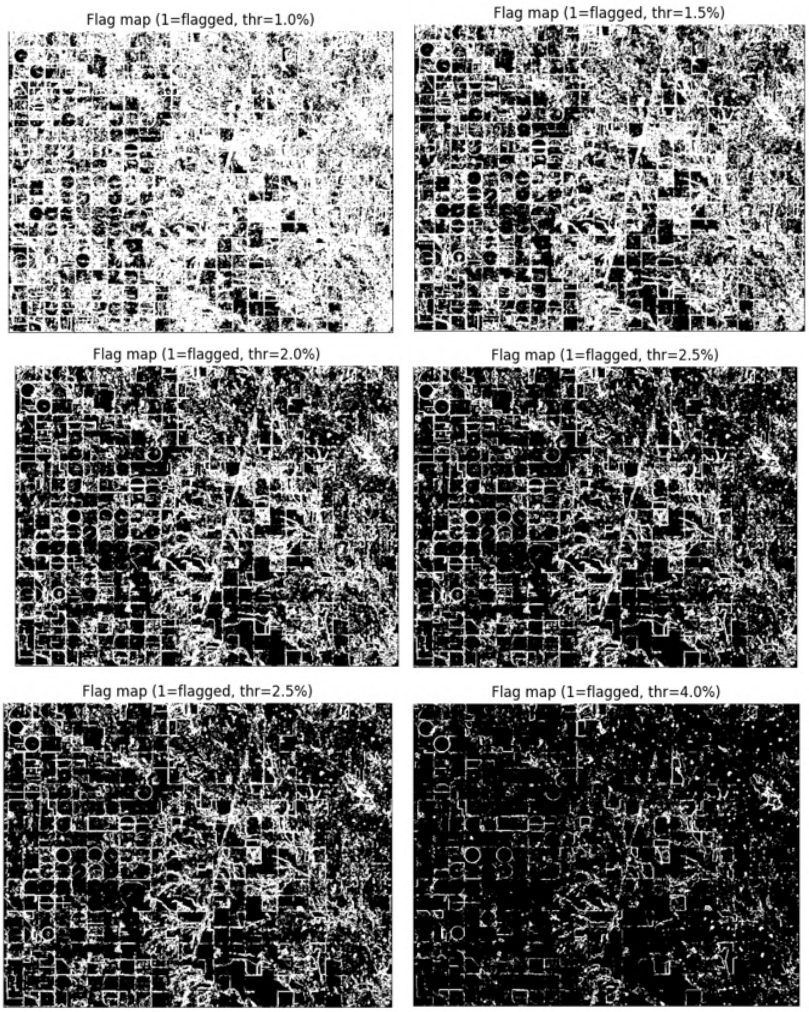


Figure 5.1: Sensitivity of number of flagged pixels to threshold value (in %). Black: flagged pixels.

SAA1: case-based half-normal approximation with empirical calibration Unlike RSS and the Exact solver, SAA1 and SAA2 are empirically calibrated heuristics rather than direct evaluations of the first-order probabilistic model.

The simple analytical solver 1 (SAA1) is a heuristic extension of the RSS approach, originally proposed by in-house ESA research, and further refined in this work. Its objective is to improve the estimation of asymmetric uncertainty bounds while retaining the computational simplicity of RSS.

SAA1 starts from the same first-order assumptions as RSS: pointing-induced radiometric error is approximated as a linear combination of independent signed half-normal contributions associated with local radiance gradients. As in RSS, local contrasts are defined as

$$\alpha_k = L_k - L_c, \qquad k \in \{1, \ldots, 4\},$$

and are separated into positive and negative contributions. The combined one-sided scales are computed as

$$\sigma_+ = \sigma_{\text{gel}} \sqrt{\sum_{\alpha_k > 0} \alpha_k^2}, \qquad \sigma_- = \sigma_{\text{gel}} \sqrt{\sum_{\alpha_k < 0} \alpha_k^2}.$$

Unlike RSS, SAA1 does not directly map these scales to bounds at zero or one standard deviation. Instead, it introduces case-dependent percentile estimation, motivated by the observation that the true first-order mixture distribution is skewed and that the 16th and 84th percentiles are not symmetric nor fixed multiples of $\sigma_\pm$.

Each pixel is classified according to the number of positive and negative gradients,

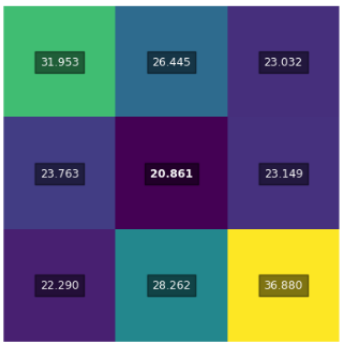
$$(N_+, N_-), \qquad N_+ + N_- \leq 4,$$

yielding five relevant cases: $(4,0), (3,1), (2,2), (1,3), (0,4)$.

Table 5.1: Description and illustration of the 5 cases used for the SAA solvers.

Case	Description	Example
(0,4)	All direct neighbouring pixels are darker than the central pixel	
(1,3)	1 direct neighbouring pixel is brighter and 3 are darker than the central pixel	
(2,2)	2 direct neighbouring pixels are brighter and 2 are darker than the central pixel	

Table 5.1: Description and illustration of the 5 cases used for the SAA solvers.

Case	Description	Example
(3,1)	3 direct neighbouring pixels are brighter and 1 is darker than the central pixel	
(4,0)	All direct neighbouring pixels are brighter than the central pixel	

For each case, SAA1 defines lower and upper bounds as

$$\ell\ell = -s\, Q_{p_-}(HN(\sigma_-)), \qquad h\ell = \; s\, Q_{p_+}(HN(\sigma_+)),$$

where $Q_p(HN(\sigma))$ denotes the p-th percentile of a half-normal distribution with scale σ, p_- and p_+ are case-dependent percentiles, and s is a multiplicative scale factor.

For one-sided cases $(4, 0)$ and $(0, 4)$, SAA1 additionally introduces a leakage term, allowing reconstruction of the nominally missing tail by defining

$$\sigma_-^{\text{eff}} = \lambda\, \sigma_+, \qquad \text{or} \qquad \sigma_+^{\text{eff}} = \lambda\, \sigma_-,$$

with leakage coefficient $\lambda \in (0, 1)$. This directly addresses the RSS degeneracy where the empty-side bound collapses to zero.

All SAA1 parameters, including the percentile locations $p_\pm$, scale factors s, and leakage coefficients λ, are estimated empirically by least-squares fitting to Monte Carlo–derived 16th and 84th percentiles. The fitting is performed independently for each (N_+, N_-) configuration using the full spatial extent of each AVIRIS scene. As a consequence, the current implementation calibrates the heuristic solvers on a per-scene and per-band basis. This scene-specific calibration strategy is inherently ad hoc and limits the generality of the resulting coefficients.

A more robust alternative would be to get calibration data across multiple scenes and bands, so that the solver's fitted parameters can be globally applicable. However, enforcing shared coefficients across heterogeneous scenes would inevitably reduce per-scene optimality and is therefore expected to degrade performance relative to scene-specific tuning.

The results reported in this thesis should thus be interpreted as a best-case performance scenario, providing an upper bound on the accuracy achievable by the SAA1 heuristic when its parameters are optimally adapted to each scene.

SAA2: linear correction of SAA1 using second-order contrast information While SAA1 aims to improve geolocation uncertainty estimation compared to RSS, systematic residual biases remain, particularly in scenes with high contrast and in corner cases where the second-order effects impact L_{pert} significantly. The SAA2 solver is based on and extends SAA1 by introducing a simple linear correction term intended to capture second-order effects not represented in the first-order mixture model.

Let (ℓ_1, u_1) denote the SAA1 bounds expressed in absolute radiance units. Define the signed contrast aggregate

$$T = \sigma_{\text{gel}} \sum_{k=1}^{4} \alpha_k,$$

and its magnitude $A = |T|$. SAA2 modifies the SAA1 bounds according to

$$\ell\ell_2 = \ell\ell_1 + c_{\text{norm}}\, T - c_{\text{abs}}\, A, \qquad h\ell_2 = h\ell_1 + c_{\text{norm}}\, T + c_{\text{abs}}\, A,$$

where c_{norm} and c_{abs} are also case-dependent coefficients like in SAA1.

This formulation can be interpreted as:

- a signed correction proportional to the net gradient imbalance (T), which shifts the interval midpoint,
- and a symmetric widening or narrowing proportional to the total gradient magnitude (A), which adjusts interval width.

As in SAA1, the coefficients $(c_{\text{norm}}, c_{\text{abs}})$ are estimated separately for each (N_+, N_-) case using least-squares regression against Monte Carlo residuals. Specifically, midpoint residuals are regressed onto T, and width residuals onto $2|T|$, yielding closed-form ridge-regularized solutions. As with SAA1, this fitting procedure is empirical, ad hoc, per band and per scene, therefore presenting the same limitations associated.

5.2. Workflow testing

5.2.1. Synthetic data: random grids

To evaluate each solver's behaviour under unconstrained local radiometric variability, the methods were first tested on randomly generated 3×3 synthetic grids. For each configuration, Monte Carlo simulations were used to derive the reference error distribution, and solver-derived 16th and 84th percentiles were overlaid for comparison (Figure 5.2).

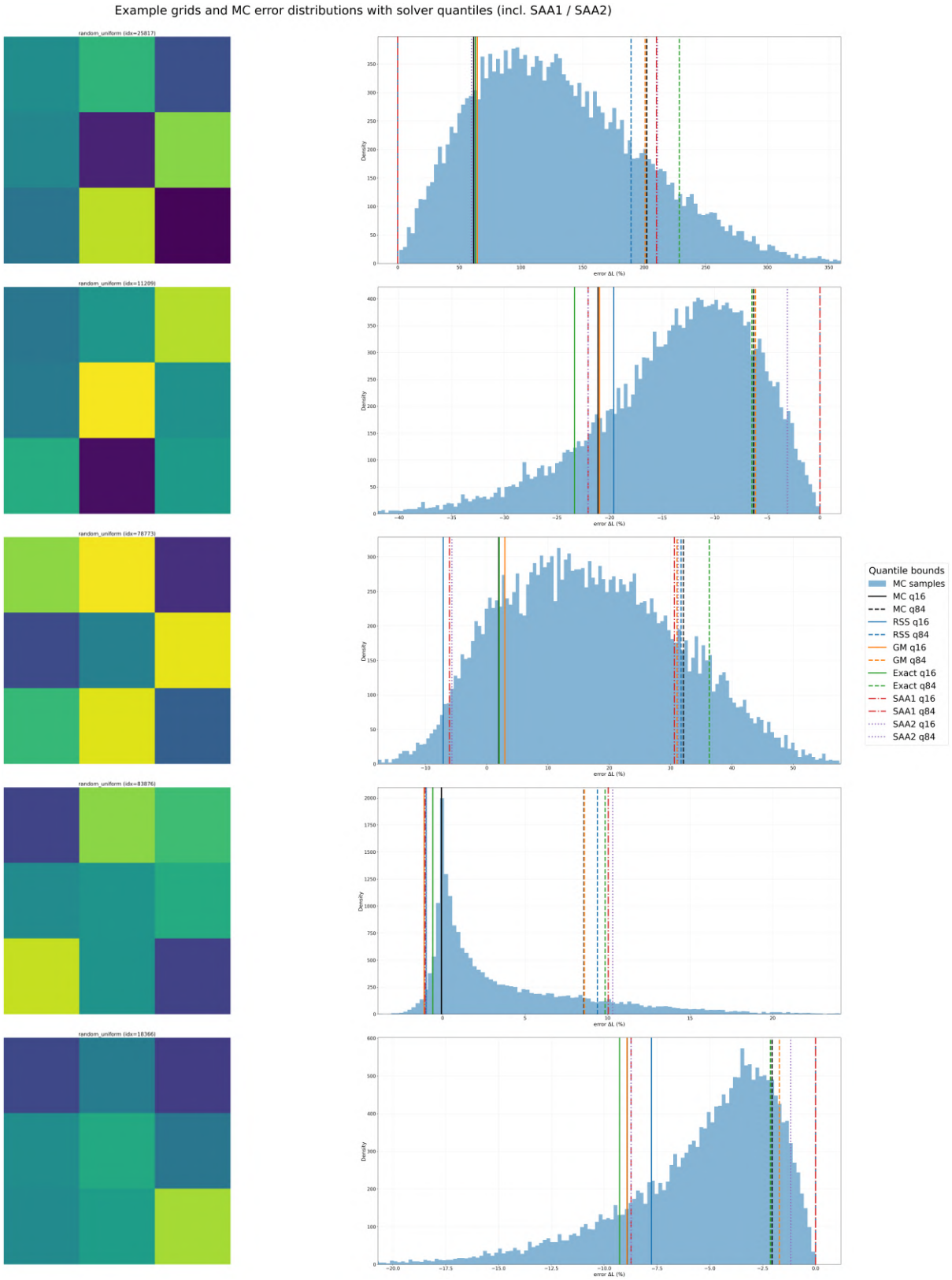


Figure 5.2: Subset of randomly generated grids (left) with their corresponding MC error distribution (blue histogram) and solver bound estimations using synthetic random grid dataset ($\sigma_{\text{gel}} = 0.15$ px) (right).

Across random grids, the Monte Carlo distributions' asymmetry and skew also depend on the contrast

configuration. From visual inspection, the RSS solver seems to capture the dominant tail direction but may misestimate the opposite bound in strongly unidirectional cases. The Exact solver closely matches the Monte Carlo quantiles, while the Gaussian-moment and hybrid solvers provide intermediate performance depending on distribution symmetry.

Overall, the synthetically generated grids show that solver performance is governed not only by contrast magnitude but also by the directionality and balance of local gradients. Even in the absence of deliberately structured patterns, the Monte Carlo error distributions show skewness. Across these grids, the Exact solver consistently aligns most closely with the MC quantiles, indicating stable behaviour under mixed-sign contributions. The RSS solver generally captures the dominant tail direction but tends to over- or underestimate the opposite bound when the error distribution is strongly one-sided. The Gaussian-moment and hybrid approaches perform well when the distribution is symmetric, but their deviations increase as skewness grows. The SAA solvers track the MC quantiles well in asymmetric cases although small systematic shifts appear in some contrast configurations.

5.2.2. Synthetic data: patterned grids

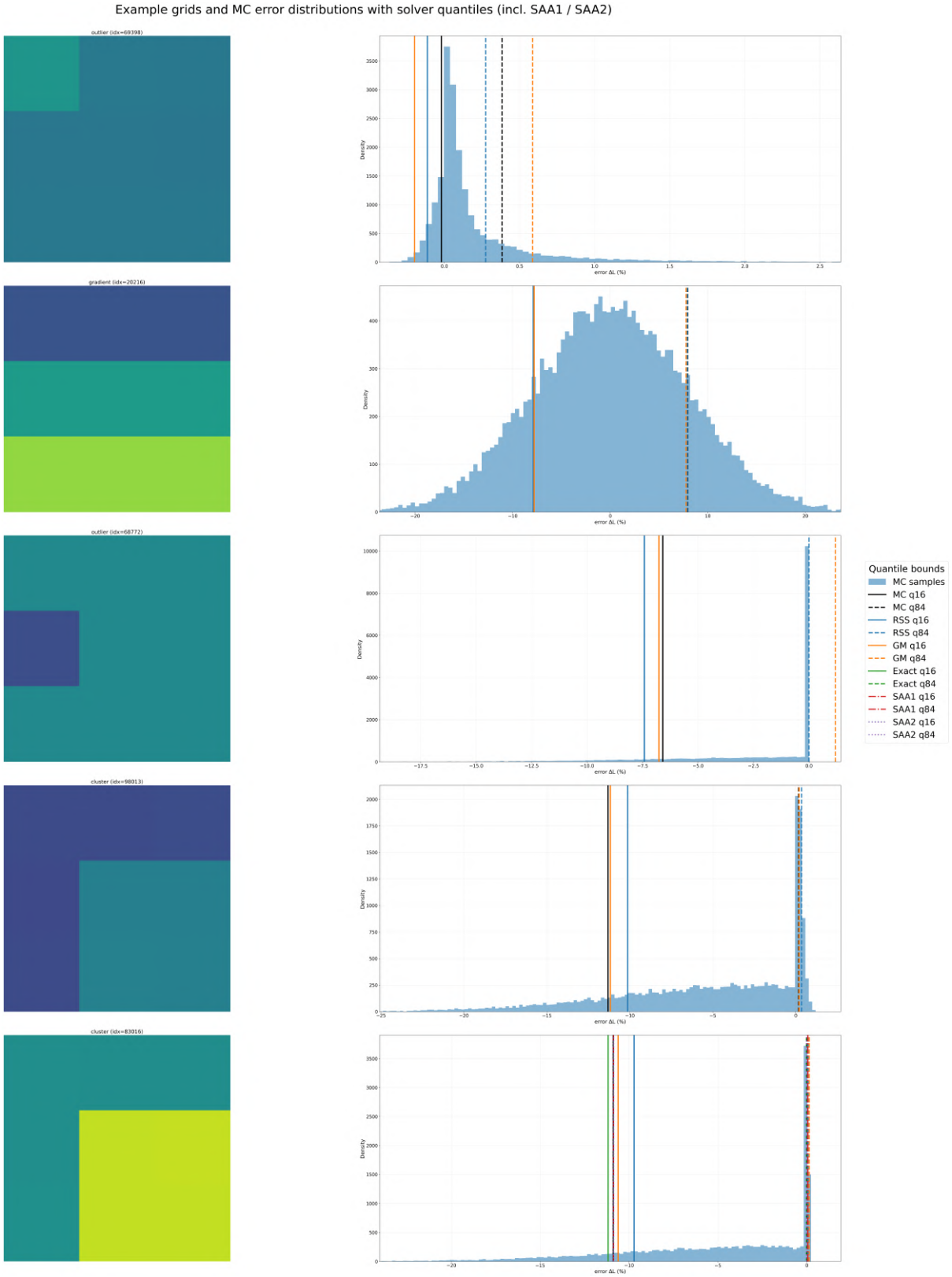


Figure 5.3: Subset of generated patterned grids (left) with their corresponding MC error distribution (blue histogram) and solver bound estimations using synthetic random grid dataset ($\sigma_{\text{gel}} = 0.15$ px) (right).

To isolate characteristic spatial structures, patterned synthetic grids (gradient, outlier, cluster, and sharp-edge configurations) were analysed (Figure 5.3). These controlled cases allow systematic comparison between Monte Carlo (MC) error distributions and solver-predicted quantiles.

The results show that solver behaviour is strongly pattern dependent. For smooth gradients, the MC error distributions are approximately symmetric, and all solvers reproduce the 16th and 84th percentiles closely. In contrast, outlier and sharp-edge configurations produce strongly skewed, often one-sided distributions. In these asymmetric regimes, structural differences between solvers become visible. RSS shows noticeable deviations due to tail degeneracy in unidirectional gradients, whereas the Exact solver remains consistent with the MC reference. The SAA variants generally track the Exact formulation more closely than RSS.

Table 5.2: Difference between solver uncertainty estimations and the MC "true" uncertainty. Width error is reported per pattern, per solver ($\sigma_{\text{gel}} = 0.15$ px).

Pattern	Solver	MAE	RMSE	Mean	Std
cluster	RSS	1.0633	1.2511	-1.0554	0.6719
	GM	0.3004	0.4012	-0.2852	0.2821
	Exact	0.3207	3.5230	0.3200	3.5085
	Hyb1	0.3179	3.4799	0.3077	3.4662
	Hyb2	0.3047	0.4018	-0.2846	0.2836
	SAA1	0.2860	0.4608	-0.1347	0.4407
	SAA2	0.5317	0.8787	-0.4967	0.7249
gradient	RSS	0.1211	0.1576	0.0956	0.1253
	GM	0.2260	0.3009	-0.0020	0.3009
	Exact	0.0707	0.0932	0.0015	0.0931
	Hyb1	0.0707	0.0932	0.0015	0.0931
	Hyb2	0.2257	0.3006	-0.0005	0.3006
	SAA1	1.0261	1.0811	1.0261	0.3407
	SAA2	1.0267	1.0818	1.0267	0.3407
outlier	RSS	2.3297	19.4729	2.1847	19.3500
	GM	1.2401	3.0454	0.8638	2.9204
	Exact	2.4558	13.6472	2.3303	13.4467
	Hyb1	2.0522	12.3421	1.8892	12.1966
	Hyb2	0.9204	2.6579	0.4194	2.6245
	SAA1	4.1089	33.8072	4.0049	33.5691
	SAA2	0.8514	1.8400	0.7162	1.6949
sharp_edge	RSS	1.8056	10.5839	1.1135	10.5251
	GM	30.5363	134.7429	30.5363	131.2371
	Hyb2	30.4742	133.3492	30.4742	129.8204

As such, the solvers' performance is dependent on the pattern of the grid. This is confirmed in Table 5.2, where the summary statistics of the width error distributions are reported for 10 000 generated grids and ordered per pattern type. The results show that for smooth gradient patterns, where the MC width distributions are approximately symmetric, the Exact and Hybrid1 solvers achieve the lowest errors, closely followed by RSS, while the SAA variants introduce a systematic positive bias.

For clustered patterns, moderate asymmetry leads to increased dispersion, with RSS showing a stronger negative bias and the SAA1 and hybrid solvers the lowest MAE. Where the error distributions are the most skewed (outlier pattern), large tail contributions cause larger RMSE inflation for RSS, Exact and SAA1 solvers, whereas the Hybrid2 and the SAA2 solvers demonstrate better robustness. Finally, in sharp-edge cases, the Gaussian-moment solver (and its hybrid counterpart) breaks down due to the Gaussian assumption being violated, while RSS remains comparatively stable.

Overall, Figure 5.3 confirms that solver performance is governed by the underlying spatial contrast

geometry: symmetric configurations favour classical analytical solutions, whereas asymmetric or high-contrast patterns require formulations that explicitly account for skewness in the error distribution.

5.2.3. Sentinel-2 scenes

Solver performance was further evaluated on real Sentinel-2 Band 8 scenes representing heterogeneous agricultural and coastal landscapes (Figures 5.4 and 5.5). Uncertainty maps derived from Monte Carlo simulation serve as the reference baseline.

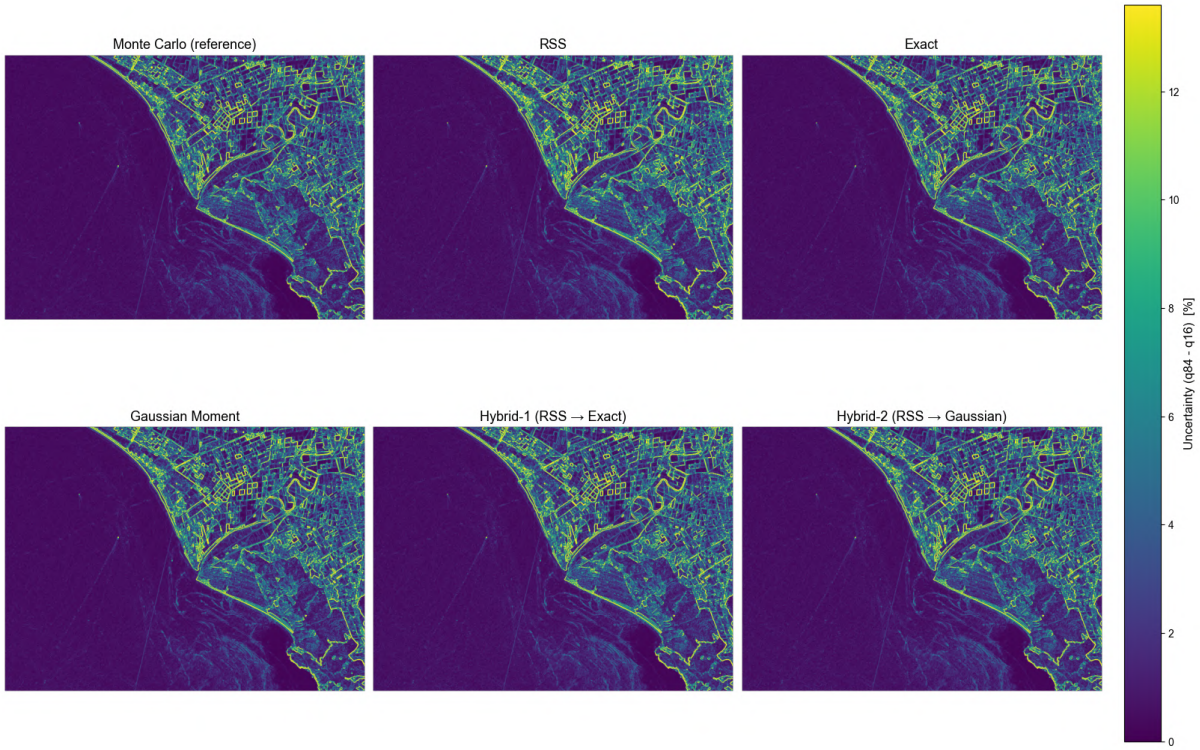


Figure 5.4: Uncertainty map Grossetto Italy via Sentinel-2 Band 8. Monte Carlo baseline "truth" and 5 solvers. $\sigma_{\text{gel}} = 0.15$ px.

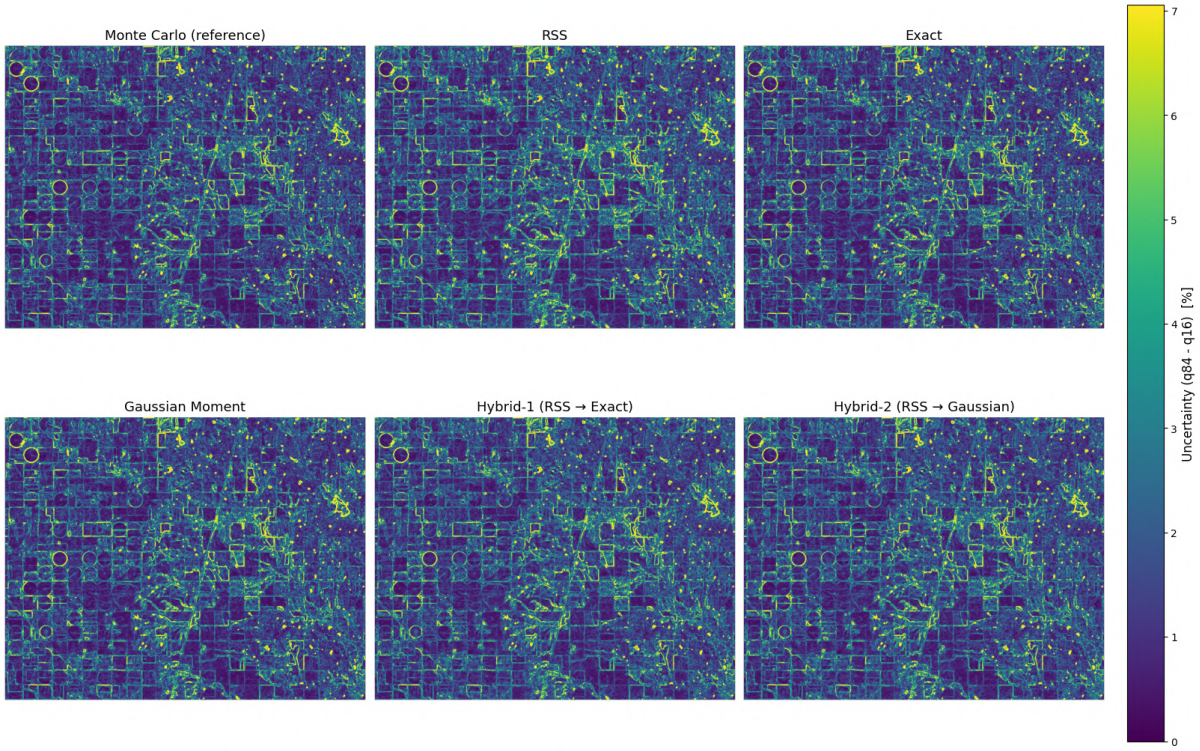


Figure 5.5: Uncertainty map Oklahoma USA via Sentinel-2 Band 8. Monte Carlo baseline "truth" and 5 solvers. $\sigma_{\text{gel}} = 0.15$ px.

As was observed, in both the Grosseto (Italy) and Oklahoma (USA) scenes, spatial uncertainty patterns are high in areas of high radiometric contrast, especially along field boundaries, coastlines, and high-contrast features. The RSS, Hybrid and Gaussian-moment approaches achieve spatial patterns visually comparable to the reference while reducing runtime relative to the Exact solver.

5.2.4. AVIRIS agricultural scene in Saskatchewan

In this section the band 90 (NIR) of an agricultural scene in Saskatchewan, Canada is chosen as a representative sample for the solver testing. The results for other scenes and biomes are depicted in Appendix C, all for the band 90 in the NIR and sampling from the same geolocation error distribution where $\sigma_{\text{gel}} = 0.15$ px.

Least squares parameter fitting for chosen scene Heuristic solver parameters (SAA1 and SAA2) were calibrated using least-squares fitting to Monte Carlo-derived 16th and 84th percentiles computed over large agricultural fields in Saskatchewan (AVIRIS NIR band 90, $\sigma_{\text{gel}} = 0.15$ px, 10 000 realizations). The fitted percentiles, scale factors, leakage coefficients, and second-order correction terms are summarised in Tables 5.3–5.7.

Table 5.3: Fitted low and high percentiles for each half-normal contribution across radiometric combination cases (n_+, n_-).

Case (n_+, n_-)	p_{lo} (q_{16})	p_{hi} (q_{84})
(4, 0)	72.1583	42.1951
(3, 1)	82.5222	73.3653
(2, 2)	68.3992	68.3992
(1, 3)	46.2330	57.8224
(0, 4)	40.0447	94.3786

The percentiles were fitted for negative and positive sides separately. For the balanced case (2, 2), a single percentile is used and stored symmetrically.

Table 5.4: Fitted scale parameters (s) for each half-normal contribution across radiometric combination cases (n_+, n_-).

Case (n_+, n_-)	Scale
(4, 0)	0.953
(3, 1)	0.771
(2, 2)	1.008
(1, 3)	1.340
(0, 4)	0.978

Table 5.5: Fitted leakage coefficients for radiometric combination cases in which all half-normal contributions have the same sign $(4, 0)$ and $(0, 4)$.

Case	Leak
(4, 0)	0.2626
(0, 4)	0.1369

Table 5.6: SAA2 signed (normal) correction coefficients (c_{norm}) for each radiometric combination case (n_+, n_-).

Case (n_+, n_-)	c_{norm}
(4, 0)	-0.01211
(3, 1)	-0.15896
(2, 2)	-0.31788
(1, 3)	-0.14893
(0, 4)	-0.02904

The signed linear correction term in the SAA2 model primarily controls bias compensation of the predicted confidence interval centre. Negative values indicate a systematic overestimation of the baseline interval that is corrected by shifting the interval inward.

Table 5.7: SAA2 absolute-value (width inflation) coefficients (c_{abs}) for each radiometric combination case (n_+, n_-).

Case (n_+, n_-)	c_{abs}
(4, 0)	0.01215
(3, 1)	0.06823
(2, 2)	0.05354
(1, 3)	0.05691
(0, 4)	0.02605

These coefficients scale the absolute-value correction term in the SAA2 model and regulate interval width inflation or contraction independently of sign, allowing additional control of interval coverage without introducing directional bias.

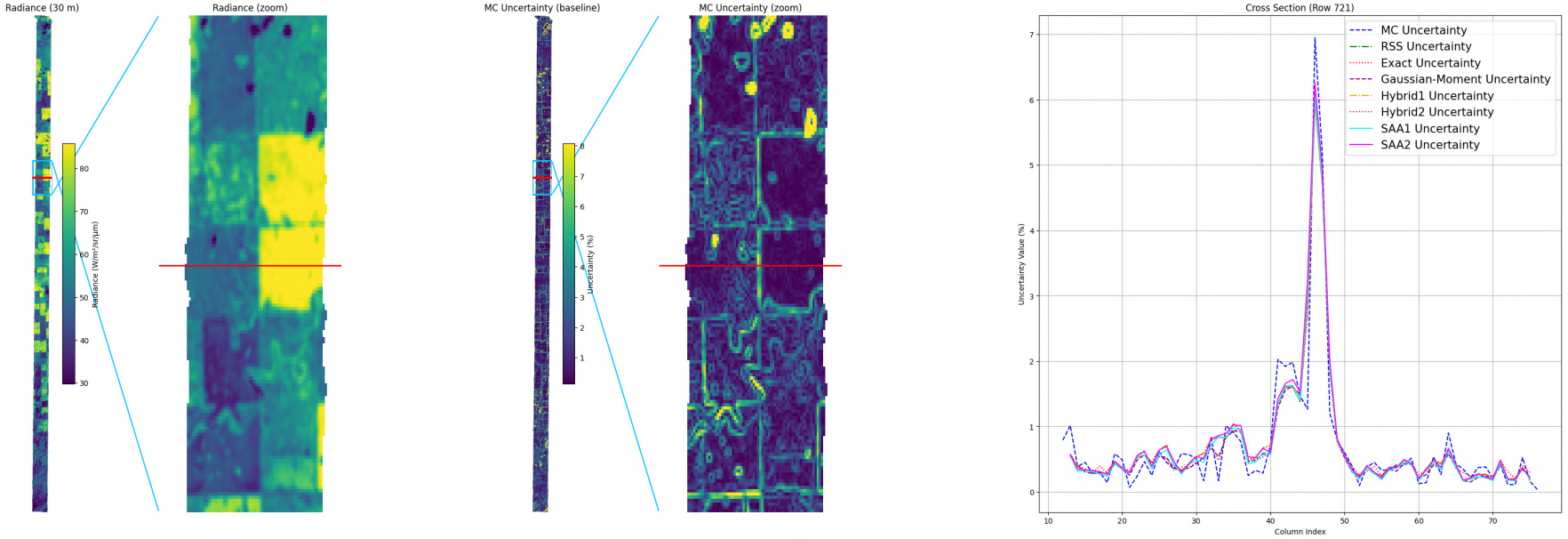


Figure 5.6: Left: radiance map of agricultural scene with closeout on area of interest (AOI) and cross section localization in red. Middle: Monte Carlo uncertainty map (run_mc = 10 000, $\sigma_{\text{gel}} = 0.15\,\text{px}$) with AOI and cross section. Right: Cross section with estimated geolocation uncertainty for each solver and for each pixel along the cross section.

Figure 5.6 shows the CHIME resolution radiance map, the Monte Carlo reference uncertainty map, and a solver cross-section for the agricultural scene. A spatial relationship is visible: elevated uncertainty coincides with sharp radiometric transitions, like along field boundaries as seen in the scene. In homogeneous field interiors, where radiometric contrast is low, geolocation-induced uncertainty remains small (below 1%). This behaviour is physically consistent, as uncertainty scales with local spatial gradients: small shifts only matter where contrast is high.

The cross-section confirms this interpretation quantitatively. Across most of the profile, uncertainty is low and stable; at the location corresponding to a field edge, a pronounced peak appears, reaching several percent. This demonstrates that although average uncertainty over the scene is small, locally it can become significant at high-contrast transitions.

Visually, all solvers follow the MC reference closely, including the peak structure. Differences are minor in low-contrast regions and slightly more apparent near asymmetric transitions. However, quantitative validation metrics are required to assess solver performance rigorously beyond this qualitative agreement.

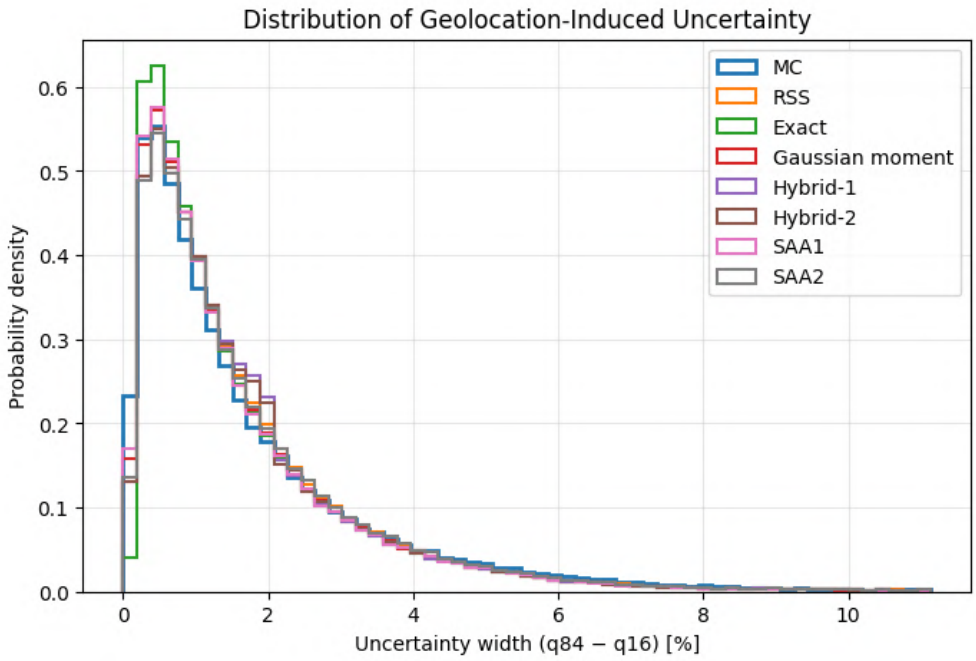


Figure 5.7: Geolocation uncertainty distribution for the Saskatchewan scene in NIR, per solver, $\sigma_{\text{gel}} = 0.15\,\text{px}$.

Figure 5.7 shows the full-scene distribution of uncertainty width ($q_{84} - q_{16}$) for each solver. The majority of pixels cluster at low uncertainty values, reflecting the predominance of small local contrasts in the tested scene. However, the distribution exhibits a long upper tail extending beyond 10% in localized some high-contrast areas.

Across the distribution, visual inspection indicated that the Exact solver and SAA solvers closely match the Monte Carlo reference, while RSS shows slight compression of the upper tail. Hybrid and Gaussian-moment approaches maintain similar distributional shapes while reducing computational cost. Quantitative metrics for this are reported in section 5.3.

5.3. Workflow validation

5.3.1. Monte Carlo reference geolocation uncertainty for one scene

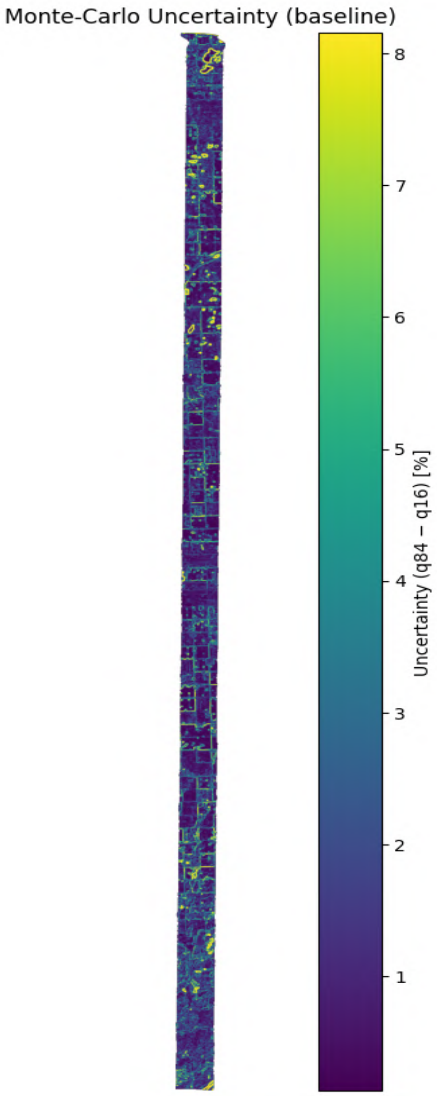


Figure 5.8: Geolocation uncertainty estimation map, complete version. Results are for the baseline "truth" MC method, using 10 000 PSF geolocation shifts and $\sigma_{\text{gel}} = 0.15$ px in the under sampled geometry.

The complete uncertainty map is shown in Figure 5.8. Furthermore, for visualization purposes the image was cut into four equal parts that were set side-by-side (Figure 5.9), allowing for better visual understanding of the uncertainties estimated by the Monte Carlo reference and their spatial distribution.

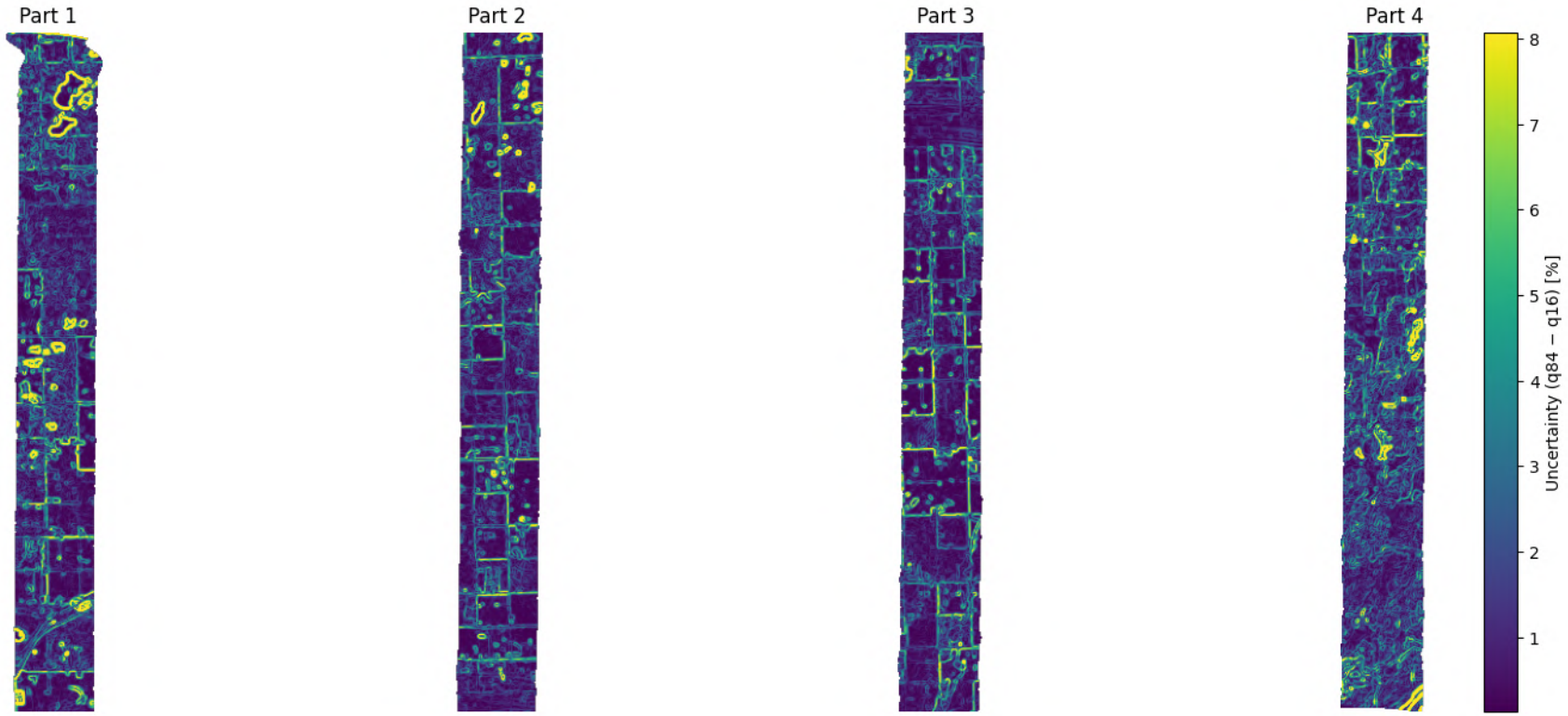


Figure 5.9: Geolocation uncertainty estimation map cut into four sections for visualisation, 10 000 realisations, $\sigma_{\text{gel}} = 0.15$ px (left-to-right sections correspond to top-to-bottom regions of the scene).

Across the majority of the scene, uncertainty remains low, typically below 1–2%. Elevated values are primarily concentrated along high-contrast features such as field boundaries, and sharp transitions between vegetation types. This spatial pattern confirms that geolocation-induced radiometric uncertainty is strongly contrast-driven and remains localized in structurally complex regions.

Although the scene-average uncertainty is modest, localized maxima exceed 8%, which shows that subpixel displacement can become significant in high-contrast areas despite small global geolocation error magnitude.

5.3.2. Validation metrics for one scene

On Saskatchewan band 90, all methods reproduce MC interval widths within sub-percentage-point accuracy, with SAA2 offering the best overall centring and the strongest improvement in lower-tail (q_{16}) placement, especially in unidirectional radiometric cases.

Tables 5.8 and 5.9 summarize solver performance against the Monte Carlo reference on a single test scene (Saskatchewan, band 90). Metrics are reported for the predicted uncertainty interval width $w = q_{84} - q_{16}$ and for the individual bounds q_{16} and q_{84}. Errors are expressed in absolute percentage points relative to the MC-derived percentiles.

Table 5.8: Absolute errors between solver-predicted and Monte Carlo-derived uncertainty interval widths (geolocation uncertainty) for a single scene, in percentage points.

Method	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.3704	0.5771	-0.0816	-0.0433	0.5713
EXACT	0.3655	0.5728	-0.1373	-0.0656	0.5561
GAUSSIAN_MOMENT	0.3823	0.6047	-0.1232	-0.0695	0.5920
HYBRID1	0.3719	0.5782	-0.1098	-0.0524	0.5676
HYBRID2	0.3826	0.6050	-0.0990	-0.0483	0.5969
SAA1	0.3688	0.5752	-0.1394	-0.0762	0.5580
SAA2	0.3652	0.5782	-0.0393	-0.0226	0.5768

Across all solvers, interval-width errors are small: MAE is below 0.40 percentage points for every method, indicating that all approaches reproduce MC width to within sub-percentage-point accuracy on this scene (Table 5.8). Among the tested solvers, SAA2 (MAE = 0.3652) and Exact (MAE = 0.3655) achieve the lowest MAE, with RSS and Hybrid-1 close behind. The Gaussian-moment and Hybrid-2 solvers show slightly larger dispersion (RMSE $\approx$ 0.60), consistent with increased variance rather than a large shift in central tendency.

Signed error (mean error) highlights systematic tendencies. Most solvers exhibit a negative mean width error which means they show a mild tendency to underestimate MC width. This bias is largest for the Exact and SAA1 methods and smallest for SAA2, which reduces the mean width bias substantially (MEAN = -0.0393). Importantly, the magnitudes of these signed errors remain small compared to typical uncertainty ranges, implying that differences are best interpreted as fine-scale calibration effects rather than gross model failures.

Table 5.9: Absolute errors of predicted uncertainty interval widths (geolocation uncertainty) relative to Monte Carlo reference values, stratified by radiometric combination case (n_+, n_-). Results are reported in percentage points.

Case	Method	MAE	RMSE	MEAN
(4,0)	RSS	0.425	0.764	0.265
	EXACT	0.368	0.653	0.183
	GAUSSIAN_MOMENT	0.391	0.700	0.238
	HYBRID1	0.396	0.674	0.231
	HYBRID2	0.407	0.711	0.255
	SAA1	0.348	0.612	0.029
	SAA2	0.358	0.631	0.077

Table 5.9: Absolute errors of predicted uncertainty interval widths (geolocation uncertainty) relative to Monte Carlo reference values, stratified by radiometric combination case (n_+, n_-). Results are reported in percentage points.

Case	Method	MAE	RMSE	MEAN
(3,1)	RSS	0.370	0.602	0.009
	EXACT	0.358	0.583	-0.088
	GAUSSIAN_MOMENT	0.373	0.619	-0.063
	HYBRID1	0.366	0.590	-0.045
	HYBRID2	0.377	0.621	-0.027
	SAA1	0.384	0.623	-0.128
	SAA2	0.387	0.634	0.029
(2,2)	RSS	0.382	0.583	-0.158
	EXACT	0.385	0.594	-0.185
	GAUSSIAN_MOMENT	0.406	0.632	-0.182
	HYBRID1	0.388	0.597	-0.175
	HYBRID2	0.402	0.630	-0.168
	SAA1	0.374	0.572	-0.135
	SAA2	0.373	0.586	-0.052
(1,3)	RSS	0.310	0.455	-0.014
	EXACT	0.301	0.444	-0.093
	GAUSSIAN_MOMENT	0.312	0.460	-0.075
	HYBRID1	0.308	0.451	-0.048
	HYBRID2	0.314	0.461	-0.036
	SAA1	0.326	0.488	-0.130
	SAA2	0.319	0.475	-0.031
(0,4)	RSS	0.296	0.420	0.168
	EXACT	0.258	0.367	0.108
	GAUSSIAN_MOMENT	0.273	0.390	0.147
	HYBRID1	0.289	0.403	0.160
	HYBRID2	0.291	0.411	0.166
	SAA1	0.243	0.354	-0.029
	SAA2	0.255	0.364	0.045

Table 5.9 shows that performance differences concentrate in highly unidirectional cases $(4, 0)$ and $(0, 4)$, where the induced perturbation distribution is strongly one-sided and tail behavior matters most. In these configurations, RSS shows the largest mismatch, which is consistent with degeneracy in the "missing tail" scenario: RSS bounds are driven by the amount of positive vs. negative contributions and can become biased when these contributions are not balanced.

By contrast, SAA1 and SAA2 reduce error in the unidirectional regimes, consistent with their explicit handling of one-sided behaviour through calibrated percentiles and leakage terms. In the most balanced case $(2, 2)$, solver performance converges: all methods show similar MAE/RMSE and all retain negative mean width error, consistent with a more symmetric MC reference distribution and reduced sensitivity to tail modeling details.

Case stratification suggests the solvers performance is dependent on the pixel's directional gradients. The clearest benefit of SAA1/SAA2 appears where the distribution departs most from symmetry and where first-order symmetric approximations (or RSS-style aggregation) are least appropriate.

To separate "width correctness" from "placement correctness," Tables 5.10 and 5.11 report errors for the lower and upper bounds individually.

Table 5.10: Absolute errors between solver-predicted and Monte Carlo–derived lower confidence bounds (q_{16}) for a single scene, in percentage points.

Method	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.3114	0.5003	0.0749	0.0416	0.4947
EXACT	0.3794	0.6091	0.1152	0.0634	0.5981
GAUSSIAN_MOMENT	0.3488	0.5596	0.1043	0.0654	0.5498
HYBRID1	0.3478	0.5901	0.1035	0.0449	0.5809
HYBRID2	0.3276	0.5454	0.0935	0.0455	0.5373
SAA1	0.2860	0.4721	0.0997	0.0495	0.4615
SAA2	0.2010	0.3178	0.0231	0.0133	0.3169

For the lower bound q_{16}, SAA2 achieves the lowest MAE (0.2010 percentage points), substantially outperforming the other solvers on this scene. SAA1 is the next best (MAE = 0.2860), while RSS/Exact/Gaussian-moment/Hybrid methods cluster at higher error levels. This indicates that SAA2 is particularly effective at placing the lower tail in alignment with MC.

Table 5.11: Absolute errors between solver-predicted and Monte Carlo–derived upper confidence bounds (q_{84}) for a single scene, in percentage points.

Method	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.3197	0.5551	-0.0067	-0.0392	0.5550
EXACT	0.3821	0.6536	-0.0220	-0.0603	0.6532
GAUSSIAN_MOMENT	0.3551	0.6214	-0.0189	-0.0629	0.6211
HYBRID1	0.3495	0.6354	-0.0064	-0.0403	0.6354
HYBRID2	0.3333	0.6080	-0.0055	-0.0417	0.6080
SAA1	0.2875	0.4961	-0.0398	-0.0418	0.4945
SAA2	0.2049	0.3390	-0.0162	-0.0096	0.3386

For the upper bound q_{84}, errors are more tightly grouped: all solvers remain within roughly 0.3–0.4 percentage points MAE, with SAA1/SAA2 again performing consistently well. The smaller separation on q_{84} relative to q_{16} suggests that the main difficulty on this scene lies in correctly capturing the negative (lower) tail, which is also consistent with the stronger influence of one-sided cases observed in the stratified results.

The bound-level results explain why width MAE differences are modest: some methods can achieve similar widths but with different allocations of error between the lower and upper bounds. The strong q_{16} performance of SAA2 indicates that its additional correction term is not just "nudging width," but improving tail placement, particularly for configurations where the lower tail is sensitive to missing/weak negative contributions.

Taken together, the one-scene validation supports three conclusions:

- All solvers are close to MC in width on this scene (sub-0.4 pp MAE), so differences are second-order relative to typical uncertainty magnitudes.
- The main separability comes from directional/tail-dominated cases $(4, 0)$ and $(0, 4)$, where RSS and purely first-order symmetric assumptions are most stressed.
- SAA2 is the most consistently "well-centred" (smallest signed width error) and provides the clearest improvement in lower-tail placement q_{16}, which appears to be the harder tail to match in this scene. However, as discussed, this solver is still a prototype and needs to be fit more globally to be used on all scenes. This might result in a degradation of solver performance.

5.3.3. Error histograms for one scene

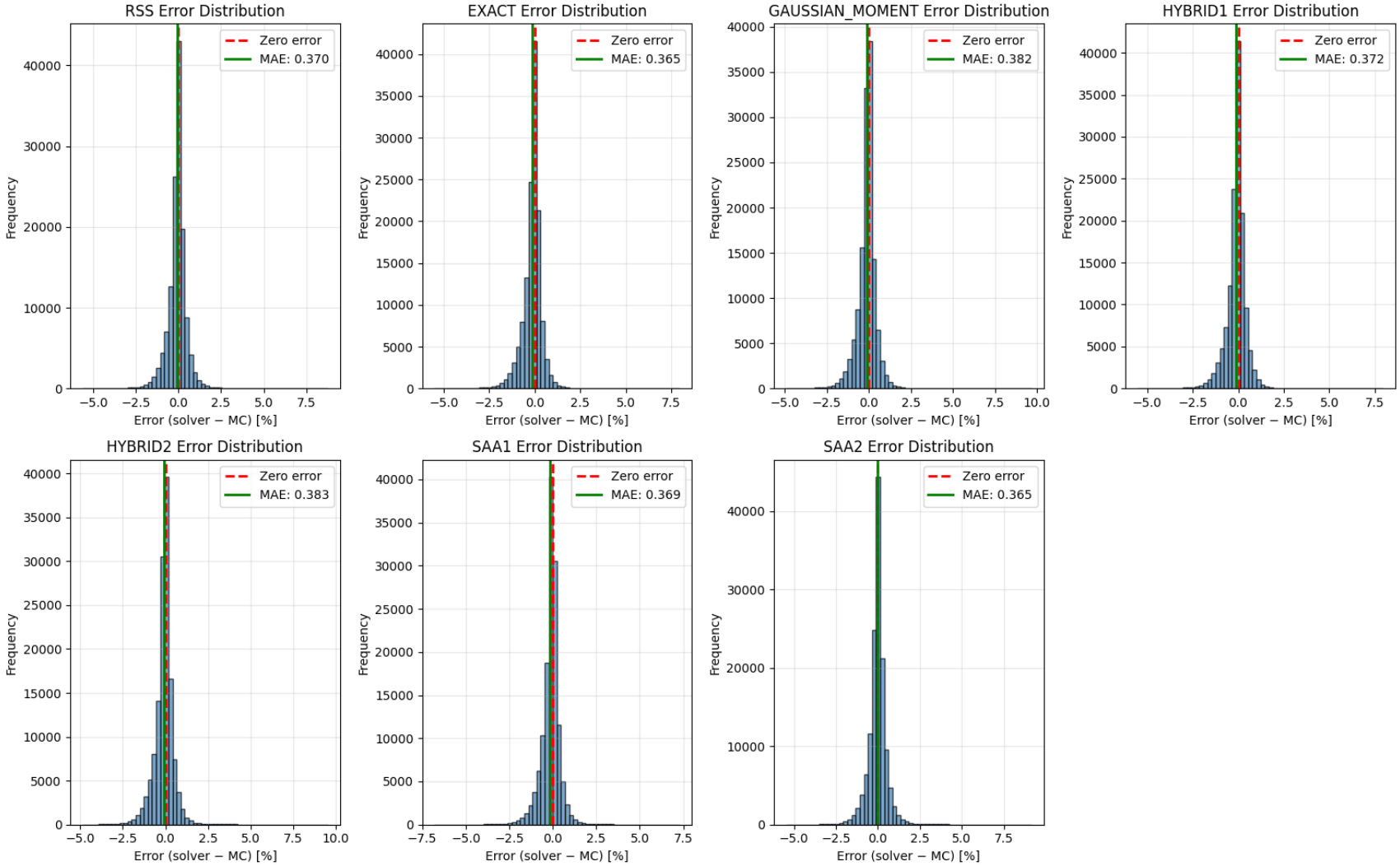


Figure 5.10: Error distributions for each solver and the Monte Carlo baseline, with MAE indicated in green.

Figure 5.10 shows full-scene error histograms (solver - MC) for predicted interval width. All distributions are centred close to zero, with narrow spreads relative to the total uncertainty range. While pixel-wise uncertainty values span from near-zero to >5%, solver prediction errors remain tightly concentrated around zero. This indicates that model discrepancies are small relative to the dynamic range of the uncertainty itself.

5.4. Runtime tests

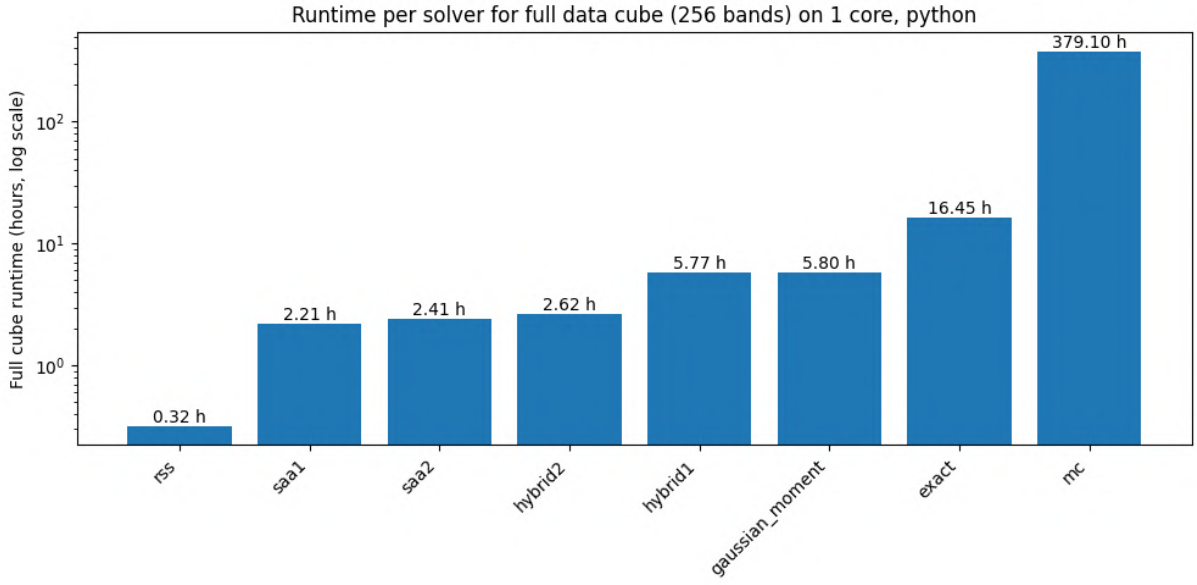


Figure 5.11: Single-core runtime per solver for full 256-band data cube (Python implementation)

Figure 5.11 shows the total runtime required to process a full 256-band data cube using each solver on a single CPU core in Python. Runtimes are presented on a logarithmic scale to highlight the large performance differences between methods. The RSS solver provides the fastest execution (0.32 h), while the Monte Carlo reference requires 379.10 h, representing the computational upper bound. The Exact solver (16.45 h) remains substantially more expensive than analytical approximations. The SAA1 and SAA2 solvers were executed without additional optimization, and their reported runtimes therefore reflect prototype implementations. Hybrid solvers exhibit intermediate performance, however their runtime depends on the scene, as only pixels exceeding the 2% uncertainty threshold are refined with higher-fidelity methods.

Overall, the results illustrate the strong trade-off between algorithm accuracy and computational efficiency, with the Monte Carlo reference being operationally infeasible for full-scene processing.

5.4.1. Optimized RSS execution

To assess potential performance gains under optimized execution, an in-house benchmark was conducted in parallel with this thesis research, using the RSS solver on a quarter-band subset of a CHIME-like data cube on M2 architecture.

Table 5.12: M2 CPU performance statistics for geolocation uncertainty RSS solver execution for a quarter-band CHIME data subset (in seconds).

Operation	Mean (s)	Min (s)	Max (s)
Processing (1 Core)	48.35	43.90	54.07
Processing (4 Cores)	15.54	14.31	16.88
Processing (8 Cores)	9.80	8.88	10.83
Processing (16 Cores)	8.02	7.33	8.93

Table 5.12: M2 CPU performance statistics for geolocation uncertainty RSS solver execution for a quarter-band CHIME data subset (in seconds).

Operation	Mean (s)	Min (s)	Max (s)
Load (All Cores)	19.21	16.89	22.49
Save (All Cores)	17.15	14.24	22.94

The CPU results (Table 5.12) demonstrate strong parallel scaling: mean processing time decreases from 48.35 s (single core) to 8.02 s (16 cores), corresponding to an approximate 6× speed-up. Load and save operations remain largely constant across configurations and constitute a significant fraction of total runtime, indicating I/O constraints beyond pure computational scaling.

Table 5.13: M2 GPU operation timing breakdown for RSS execution for a quarter-band CHIME data subset (in milliseconds).

Operation	Mean (ms)	Min (ms)	Max (ms)
Processing	6820.60	6459.00	8975.00
Host to Device	590.62	573.10	733.78
Kernel	6.1145	5.6673	7.6984
Device to Host	2911.57	2726.05	4320.13
Load	20172.72	18803.00	23555.00
Save	15861.12	13887.00	17105.00

The GPU timing breakdown (Table 5.13) further clarifies performance limitations. Kernel execution time is comparatively negligible (6 ms), while host–device transfer and especially file I/O dominate total runtime. Data loading (20 s) and saving (16 s) exceed kernel execution by several orders of magnitude, demonstrating that workflow performance is primarily constrained by data movement rather than arithmetic computation. These results indicate that further runtime improvements will depend more on optimized data handling and memory management than on kernel-level acceleration alone.

6

Discussion

6.1. Highlights

This thesis developed and validated a set of algorithms for estimating per-pixel radiometric uncertainty due to acquisition geolocation error. The primary contribution is a structured evaluation framework that enables comparison of analytical, semi-analytical, and stochastic approaches against a Monte Carlo reference. The resulting framework is physically consistent (rooted in local radiometric gradients and geolocation perturbation models) and operationally realistic for potential deployment within satellite ground processing pipelines.

Within this framework, we find that (apart from the operationally infeasible Monte Carlo reference) solver ranking depends on the chosen validation metric. Under global width-based metrics $(q_{84} - q_{16})$, the Exact solver performs marginally best. However, when lower and upper bounds are evaluated separately, SAA2 provides the most consistent accuracy, particularly in asymmetric high-contrast configurations where tail behaviour becomes important. Its correction terms reduce bias in individual bounds and improve centring relative to MC.

However, SAA2 remains a prototype: coefficients were fitted per scene and per band, and computational optimisation has not yet been finalised. Its reported performance therefore represents a best-case validation scenario rather than an immediately deployable operational solution.

If the anticipated global parameter stability of SAA2 cannot be achieved, the results indicate that RSS constitutes the most robust operational candidate, meaning that it provides stable, first-order-accurate uncertainty estimates with minimal bias at orders-of-magnitude lower computational cost. Earlier assessments based on synthetic contrast grids and relative error metrics suggested insufficient performance. However, validation on realistic scenes using absolute and percentile-based metrics reveals that RSS captures the dominant physics of geolocation-induced uncertainty with sub-percentage-point deviations from MC. This finding is important as it corrects the original assessment performed in Fernández et al. (2025) where the RSS performance error was overestimated.

Beyond identifying a numerically strong solver, the framework reveals three key behaviours:

- Geolocation-induced uncertainty is spatially heterogeneous and concentrated in high-contrast regions;
- The studied solvers' performance differs mainly in cases where the neighbouring pixel contrast results in error distributions that are very asymmetric. Typically, the difference is bigger in the estimation of the high and low confidence level rather than in the width estimation where performance is more consistent;
- Operational feasibility is governed more by throughput and data movement constraints than by marginal gains in numerical fidelity.

These findings are developed in detail in the following sections on spatial behaviour, solver accuracy, and runtime-accuracy trade-offs.

6.1.1. Spatial behaviour of geolocation uncertainty (RQ1 & RQ2)

Across all evaluated scenes, a clear spatial structure was observed: geolocation uncertainty is strongly scene-dependent and spatially heterogeneous, as shown with the synthetic tests. High uncertainty values systematically occur along sharp radiometric transitions such as field boundaries, coastline intersections, and small high-contrast objects. In homogeneous agricultural parcels, uncertainty remains low and often negligible relative to other contributors in the total uncertainty budget.

This dual behaviour is important to highlight because while the majority of pixels exhibit small geolocation-induced uncertainty, isolated high-contrast regions can experience substantially larger values. The distribution plots confirm that the uncertainty range spans from near-zero values to locally high outliers. Therefore, geolocation uncertainty must be included in uncertainty budgets not because it dominates globally, but because it can become locally significant in precisely those pixels where users often extract information (like crop boundaries where there is often higher biodiversity, stressed patches, heterogeneous parcels).

6.1.2. Solver accuracy relative to Monte Carlo reference (RQ3 & RQ4)

Solver performance was evaluated against a Monte Carlo reference that captures the relevant image acquisition physics while remaining computationally feasible. Overall, all solvers demonstrate good agreement with the MC-derived reference bounds. Mean absolute deviations remain small (sub-percentage-point), and inter-solver differences are moderate in global metrics.

The Exact solver generally provides the lowest deviations in strongly asymmetric radiometric configurations, particularly in cases such as (4,0) and (0,4). This behaviour is physically consistent: the Exact formulation directly solves the governing equations under first-order assumptions and therefore remains closest to the true solution as long as higher-order effects remain limited. In such cases, analytical approximations based on symmetric assumptions like the RSS will deviate due to tail degeneracy, while the Exact formulation preserves the asymmetry of the underlying half-normal contributions.

However, when second-order effects become more prominent (particularly in high-contrast or extreme-value regimes) the strictly first-order Exact solution may no longer be optimal. Gaussian-moment approaches attempt to incorporate second-order structure, but their reliance on Gaussian symmetry limits their performance in skewed distributions. In contrast, the SAA2 formulation introduces correction coefficients that partially account for higher-order behaviour without enforcing Gaussian symmetry. This likely explains its improved bound-level performance in certain asymmetric cases.

It is important to note that validation outcomes depend on the chosen metric. Width-based validation alone can obscure bound-level asymmetries. When q_{16} and q_{84} bounds are evaluated separately, solver differentiation becomes more pronounced. In balanced (2,2) cases, most solvers perform similarly due to near-symmetric error distributions. In strongly asymmetric regimes, however, structural differences between solver assumptions (first-order exactness, Gaussian approximation, or asymmetry-corrected formulations) become evident.

Overall, the results suggest a hierarchy of solver suitability:

- In predominantly first-order regimes with strong asymmetry, the Exact solver is most physically consistent.
- When higher-order contributions emerge but skewness remains important, SAA2 provides a pragmatic improvement (although it is still at the prototype level and must be calibrated and tested more globally).
- Gaussian-moment formulations capture some second-order structure but are limited by their symmetry assumption.

6.1.3. Runtime and operational feasibility (RQ3)

From an operational perspective, runtime differences are substantial. Monte Carlo simulation is computationally prohibitive for full-scene hyperspectral processing. Analytical solvers provide orders-of-magnitude acceleration. Although the Exact solver achieves slightly improved numerical fidelity, its runtime remains significantly higher than RSS or SAA variants.

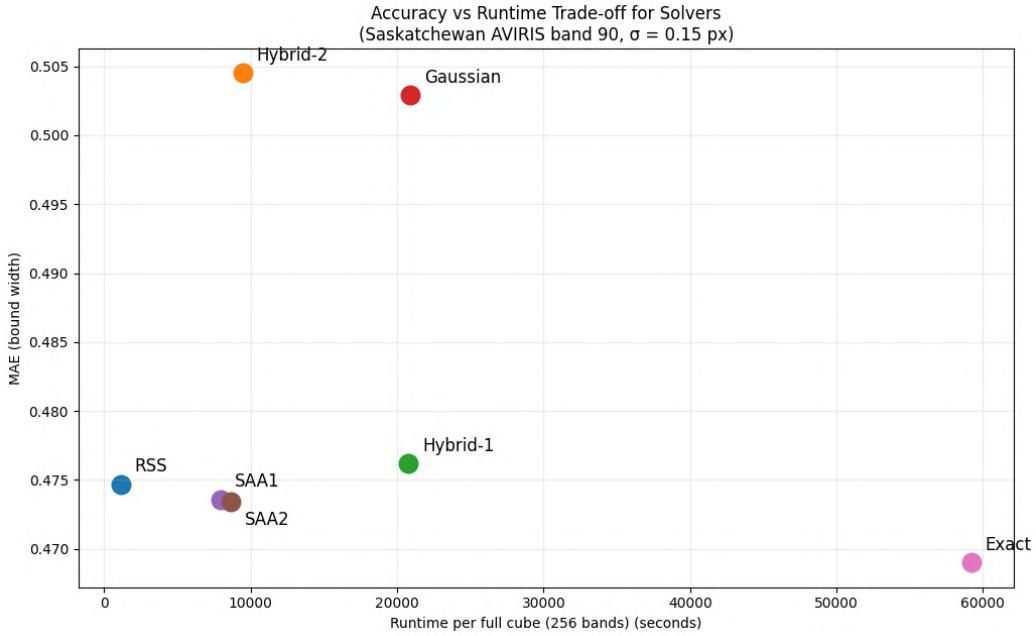


Figure 6.1: Runtime-accuracy trade-off for uncertainty width estimation (Saskatchewan AVIRIS Band 90, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

The runtime–accuracy trade-off analysis shows that:

- RSS offers the fastest execution with acceptable accuracy.
- Exact provides marginally better accuracy at substantially higher cost.
- Hybrid approaches balance these extremes.
- SAA variants show promising accuracy but remain prototype implementations with optimization potential.

Interestingly, RSS performed better than initially expected. Earlier assessments, based largely on synthetic high-contrast grids and relative error validation metrics, suggested that RSS was not performing well enough to be considered implementable. However, evaluation on realistic agricultural scenes using absolute validation metrics demonstrates that its performance is better than previously assumed. Similarly, optimization efforts indicate that Exact solvers may be more operationally viable than originally thought, especially when applied selectively (hybrid solvers).

The runtime–accuracy trade-off plots highlight that no solver dominates simultaneously in computational efficiency and numerical fidelity. However, the interpretation of "accuracy" depends on the selected validation metric. When uncertainty width $(q_{84} - q_{16})$ (or high limit – low limit) is used as the performance indicator, differences between solvers are relatively small.

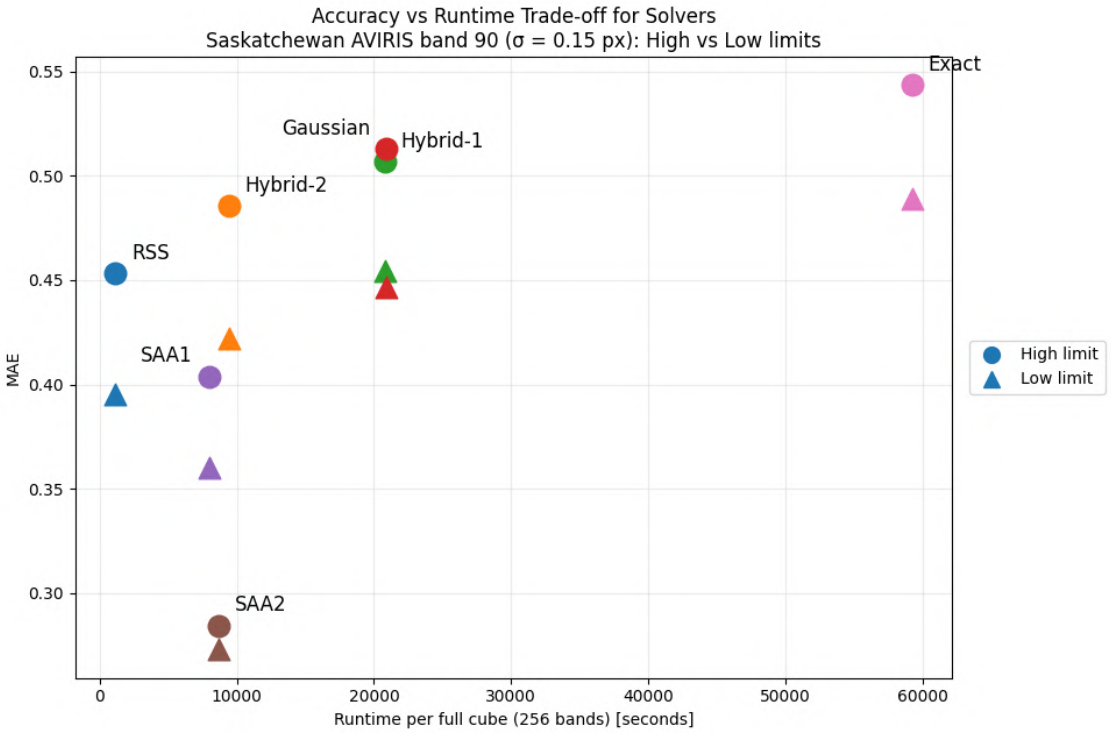


Figure 6.2: Runtime-accuracy trade-off for solver bound estimation (Saskatchewan AVIRIS Band 90, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

In contrast, when lower (q_{16}) and upper (q_{84}) bounds are evaluated separately, solver differences become more pronounced. Bound-level validation reveals asymmetry effects that are not visible in width-only metrics. In particular, SAA2 shows clear improvement in reducing bias in individual bounds, indicating better handling of skewed error distributions. The Exact solver also performs consistently well in bound-level metrics, although its advantage over analytical methods remains moderate.

This distinction is important:

- Width-based metrics assess whether the solver captures the overall magnitude of uncertainty.
- Bound-level metrics assess whether the solver correctly represents distribution asymmetry and bias.

Accordingly, solver preference depends on mission priorities and on which aspect of uncertainty representation is considered most critical:

- If accurate representation of asymmetric error bounds in high-contrast edge pixels is required (under the current validation metrics), Exact or SAA2 may be preferred.
- If operational throughput and large-scale processing are the primary constraints, RSS offers a good compromise, as its width-level accuracy remains competitive.
- Hybrid strategies show adaptive behaviour, concentrating computational effort in pixels where asymmetric effects are expected to be most pronounced.

6.2. Impact on food security

Given the capability of the framework to estimate geolocation error-induced per-pixel radiometric uncertainty, the question arises how this impacts applications such as food security and biodiversity monitoring. For this, an uncertainty propagation analysis was performed to evaluate how NDVI (as an agricultural metric) was affected.

NDVI is a spectral index derived from red and near-infrared (NIR) reflectance and is commonly used to assess vegetation vigour and photosynthetic activity. Healthy vegetation absorbs strongly in the red due

to chlorophyll and reflects strongly in the NIR because of leaf internal structure, resulting in high NDVI values. In contrast, stressed or diseased vegetation exhibits reduced NIR reflectance and/or increased red reflectance, leading to lower NDVI values (Judith et al., 2025).

To emulate the target mission geometry, the AVIRIS reflectance scene (native GSD $\approx$ 3 m) was first degraded to 30 m spatial resolution by applying a centred Gaussian PSF convolution followed by spatial down sampling (factor 10×10). The PSF standard deviation ($\sigma = 5$ native pixels $\approx$ 15 m) approximates the effective spatial response at coarse resolution. NDVI was then computed from the down sampled red ($\approx$ 668 nm) and NIR ($\approx$ 828 nm) bands as

$$\mathrm{NDVI} = \frac{N - R}{N + R}.$$

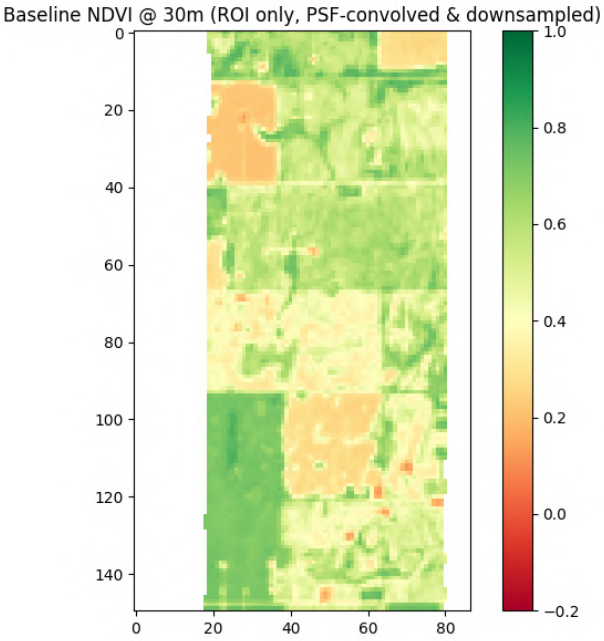


Figure 6.3: Baseline NDVI at 30 m (ROI) after PSF convolution and down sampling.

Figure 6.3 shows the spatial distribution of vegetation vigour within the selected region of interest (ROI). Field interiors appear relatively homogeneous, while strong gradients occur at field boundaries and in heterogeneous patches.

To represent a simplified operational decision layer, a binary crop-condition classification was derived from the baseline NDVI using a threshold of NDVI $\geq$ 0.60 to denote "healthy" vegetation with strong photosynthetic activity. This threshold is commonly used in agricultural monitoring literature as an indicator of a healthy crop (Judith et al., 2025).

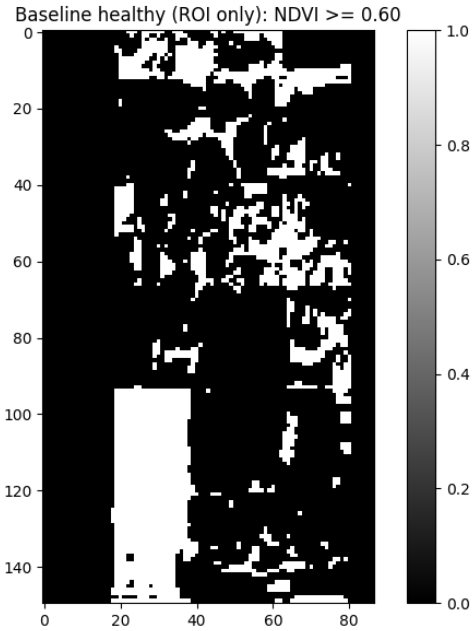


Figure 6.4: Baseline crop health mask (NDVI ≥ 0.60).

Figure 6.4 displays the resulting classification map. Healthy areas cluster within high-NDVI field interiors, while transitions between healthy and non-healthy categories occur predominantly along field edges and heterogeneous zones. These transition regions are particularly sensitive to small perturbations in NDVI and therefore provide a meaningful test for uncertainty propagation.

6.2.1. Propagation of red and NIR band TOA uncertainty to NDVI

Band-level geolocation-induced uncertainty bounds were estimated separately for the red and NIR bands over the ROI using the RSS solver ($\sigma_{\text{gel}} = 0.15\,\text{px}$). RSS produced the per-pixel low and high bounds (ll, hl) expressed in percent relative to the local radiance. These were converted into absolute uncertainties,

$$\Delta L = L \cdot (q/100),$$

yielding $(\Delta R_{ll}, \Delta R_{hl})$ and $(\Delta N_{ll}, \Delta N_{hl})$.

NDVI bounds were then computed under a strongly correlated perturbation assumption, applying the low perturbations simultaneously and the high perturbations simultaneously:

$$\text{NDVI}_{ll} = \text{NDVI}(N + \Delta N_{ll},\ R + \Delta R_{ll})\,, \qquad \text{NDVI}_{hl} = \text{NDVI}(N + \Delta N_{hl},\ R + \Delta R_{hl})\,.$$

The NDVI uncertainty interval was then formed as

$$\text{NDVI}_{ll} = \min(\text{NDVI}_{ll},\ \text{NDVI}_{hl})\,, \qquad \text{NDVI}_{hl} = \max(\text{NDVI}_{ll},\ \text{NDVI}_{hl})\,.$$

The plotted quantity is the NDVI uncertainty width $\text{NDVI}_{\text{max}} - \text{NDVI}_{\text{min}}$.

This step links the per-band geolocation-induced radiometric uncertainty estimate to an end-user index. The correlated perturbation model is deliberately simplified and reflects the fact that the same geolocation shift will affect both bands at the same pixel location (as perturbations are not independent between bands).

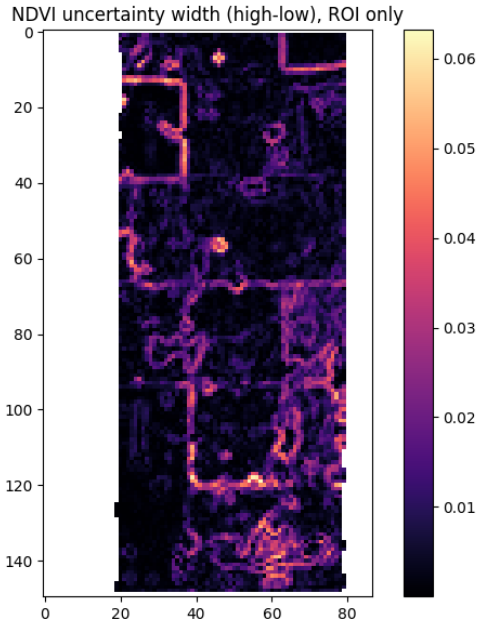


Figure 6.5: Propagated NDVI uncertainty width induced by geolocation error (ROI).

The NDVI uncertainty width is spatially structured and edge-dominated: values are lowest over homogeneous field interiors and increase along field boundaries and heterogeneous patches. In absolute magnitude, most NDVI uncertainty values remain small across the ROI, but they become locally elevated where contrast is strongest, consistent with the physical mechanism established earlier (uncertainty scales with local radiometric contrast).

6.2.2. Impact on decision-making

To assess the decision-level impact, the NDVI uncertainty interval was compared to the health threshold (NDVI = 0.60). Each pixel was categorized as:

- Definite healthy, if $\text{NDVI}_{\text{min}} \geq 0.60$
- Definite not healthy, if $\text{NDVI}_{\text{max}} < 0.60$
- Uncertain, if the interval spans the threshold

Figure 6.6: Crop health classification sensitivity to geolocation-induced uncertainty.

Figure 6.6 presents the resulting sensitivity map overlaid on a true-colour composite. Pixels classified as uncertain are primarily located along field boundaries and within heterogeneous transition zones. This spatial pattern mirrors the NDVI uncertainty width map and confirms that uncertainty effects are concentrated where radiometric contrast is highest.

The majority of pixels still remain robustly classified, demonstrating that geolocation-induced uncertainty is generally small in homogeneous agricultural areas. However, a non-negligible number of pixels becomes classification-ambiguous once uncertainty is considered.

This demonstration highlights a central conclusion of the thesis: geolocation-induced radiometric uncertainty does not primarily influence scene-average metrics, but rather affects localized, high-contrast regions that often drive decision-making. Even when the absolute magnitude of uncertainty is small, its spatial structure can affect the confidence in threshold-based interpretations such as crop health maps.

In food security monitoring, this ambiguity can influence yield estimation inputs, stress detection algorithms, or resource allocation decisions. Therefore, geolocation uncertainty should not be neglected on the basis of its small average magnitude. Instead, its spatially structured character should be explicitly communicated to users, particularly in heterogeneous agricultural landscapes where classification thresholds are applied.

The analysis presented here serves as a proof-of-concept demonstration of this propagation pathway. A comprehensive operational assessment would require multi-scene validation and integration with biophysical retrieval models, which constitutes a logical extension of the present work.

6.3. Impact on other applications

While the previous section demonstrated the implications of geolocation-induced radiometric uncertainty for NDVI-based crop health assessment, the relevance of PIRU extends beyond this application. In Earth observation, radiance measurements constitute the foundational input to a range of biophysical retrieval algorithms and ecosystem monitoring frameworks. So, any radiometric uncertainty component (and particularly one that is spatially structured) propagates into higher-level products.

Within biodiversity monitoring, satellite-derived products contribute to the quantification of Essential Biodiversity Variables (EBVs), particularly those related to species traits like morphology, physiology,

phenology, movement, or reproduction. Hyperspectral remote sensing is especially powerful in retrieving these variables; for example, it has long been the basis of plant physiological trait retrieval such as chlorophyll content, leaf area index (LAI), and canopy water content, as well as phenological indicators. These variables are typically derived using radiative transfer model (RTM) inversion frameworks or hybrid machine-learning approaches (Timmermans & Kissling, 2023; Verrelst et al., 2021). Because geolocation uncertainty scales with local radiometric contrast, it introduces spatially structured uncertainty that may bias trait retrievals in precisely those areas where biodiversity gradients are most pronounced.

6.4. Place within CHIME uncertainty budget

The following section discusses how the results of this thesis research could be placed within the broader CHIME uncertainty budget. As mentioned in Chapter 2, CHIME radiometric uncertainty is formalised through a measurement function as described below:

$$L_{i,k}^{\mathrm{L1B}} = K_k^{\mathrm{rad}} \left[K_{i,k}^{\mathrm{prnu}} K_{i,k}^{\mathrm{gain}} \left(NL_{i,k} \left(X_{i,k} - OS_{i,k} \right) - DS_{i,k} - SL_{i,k} \right) \right] + 0.$$

In its current form, this framework is made for uncertainty components that can be described by globally applicable sensor parameters (such as noise, offsets, gain, PRNU, calibration factors). By contrast, geolocation-induced radiometric uncertainty is fundamentally scene and pixel dependent, because it arises from the interaction between residual geolocation displacement and local radiance gradients. This spatial dependence complicates inclusion as a single global term within a conventional GUM-style budget.

One way to include geolocation uncertainty in the measurement function would be to absorb the geolocation contribution into the effective pixel location argument $X_{i,k}$, since geolocation error is a perturbation of sampling geometry rather than radiometric calibration. However, doing so explicitly would require extending the measurement function to incorporate spatial resampling and local scene structure, which goes beyond the current instrument-only formulation.

This also raises uncertainty product design questions that remain open. One option would be to provide geolocation uncertainty as an ad-hoc per-pixel uncertainty layer downloadable separately from the other uncertainty products, allowing users to propagate it into indices and higher-level retrievals as appropriate. A related question is whether geolocation uncertainty should be computed independently at L1 and L2 or computed at L1 and propagated through L2 processors. The latter is interesting for consistency, but propagation through non-linear retrieval algorithms may not preserve confidence-interval interpretation without explicit uncertainty propagation steps. In all cases, the present work supports the conclusion that geolocation-induced uncertainty should be represented explicitly within the CHIME uncertainty framework; the remaining challenge is to determine the most transparent and operationally robust communication strategy for end users.

7

Limitations and future work

7.1. Limitations

Although the proposed framework demonstrates strong agreement with MC reference simulations under CHIME-like conditions, several limitations constrain its generality and operational interpretation.

7.1.1. First-order truncation of the geolocation model

The analytical formulation is based on a first-order expansion of the bilinear interpolation model, neglecting the second-order cross-term τ. While τ is formally second order in displacement, its magnitude scales with local radiometric curvature. In scenes exhibiting sharp spatial transitions (field edges, point-like features, or high-contrast man-made structures), curvature terms may therefore become non-negligible and can introduce systematic bias in solvers that rely purely on the first-order mixture approximation (like the RSS solver for example).

Among the solvers considered here, only methods that either sample the local perturbation nonlinearly or absorb curvature effects through calibration can partially account for τ. The Gaussian-moment solver implicitly captures second-order behaviour because it estimates the perturbation distribution via a sample MC; however, it subsequently enforces a Gaussian approximation, which can be inaccurate when the true distribution is skewed or heavy-tailed. The SAA2 heuristic can also compensate for second-order effects indirectly: its fitted correction terms (learned against Monte Carlo bounds) can absorb residual biases caused by curvature, including those attributable to τ, albeit without explicitly modelling the cross-term.

7.1.2. Scene-specific calibration of heuristic solvers

The SAA1 and SAA2 solvers rely on empirically fitted coefficients estimated independently for each scene and spectral band. This calibration strategy represents a best-case scenario and does not reflect operational deployment.

Consequently:

- Reported performance of SAA solvers constitutes an upper bound.
- Coefficient transferability across biomes and spectral domains has not been demonstrated.
- The fitted parameters lack direct physical interpretability.

Without globally stable parameterisation, these heuristic solvers cannot yet be considered mission ready.

7.1.3. Simplified geolocation error statistics

Geolocation knowledge error was modelled as isotropic, Gaussian, and spatially uncorrelated. In practice, residual pointing errors after geometric correction may exhibit anisotropy, spatial correlation, or temporal persistence. The impact of non-Gaussian or correlated displacement statistics on percentile estimation and mixture behaviour has not been assessed.

7.1.4. Computational assessment under prototype implementation

Runtime comparisons were conducted using single-core Python implementations to ensure consistent solver benchmarking. Absolute runtimes therefore do not directly represent operational performance under optimized C++ or parallel implementations. The relative ranking of solver efficiency remains informative, but operational feasibility requires further optimisation analysis.

7.2. Future work

For SAA-type solvers to become operationally viable, the current scene-specific coefficient fitting must be replaced by a stable and transferable parameterisation strategy. Future work should therefore investigate:

- multi-scene and multi-biome calibration,
- spectral-band-independent parameterisation schemes,
- contrast-statistics-based adaptive coefficients

While relative solver efficiency has been established, operational deployment within the CHIME processing chain requires further optimisation. Future work should evaluate:

- vectorised or compiled implementations (e.g. C++ integration),
- multi-core and GPU parallelisation.

This step is necessary to confirm that per-pixel geolocation uncertainty estimation remains computationally tractable for full hyperspectral data cubes (256 bands) within operational processing timelines.

8

Conclusion

This thesis quantified how geolocation error contributes to per-pixel radiometric uncertainty in CHIME Level-1 products and developed computationally feasible algorithms to estimate this effect. A complete processing workflow was implemented and validated against Monte Carlo simulations using a CHIME-like forward model that included PSF convolution and spatial downsampling.

The results show that geolocation-induced radiometric uncertainty strongly depends on the scene and increases with local radiometric contrast. Even small, sub-pixel geolocation errors can lead to uncertainty contributions of several percentage points in areas with strong spatial gradients. Across all test scenes, the analytical and hybrid solvers reproduced the Monte Carlo–derived percentile bounds with good agreement. The RSS solver was the fastest method, while the Exact and SAA2 solvers achieved the highest accuracy.

Overall, the developed framework provides a physically consistent and operationally viable method to include geolocation error in the CHIME radiometric uncertainty budget. This supports transparent and traceable per-pixel uncertainty reporting and strengthens the reliability of hyperspectral Earth observation products for quantitative applications.

References

Bourg, L., & Laboratory, O. L. E. S. (2023, November 15). Olci level 1 processor & products: Recent and coming evolutions.

Celesti, M., Nieke, J., Adams, J., Isola, C., & Gascon, F. (2022). The copernicus hyperspectral imaging mission for the environment (chime): Status and planning. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 5011–5014. https://doi.org/10.1109/IGARSS47720.2022.9883664

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernández, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: Esa's optical high-resolution mission for gmes operational services. *Remote Sensing of Environment*, *120*, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

EUMETSAT & ESA. (2024, February 22). *Sentinel-3 product notice – olci level-1b processing baseline* [Product Notice ID EUM/OPS-SEN3/DOC/24/1401310, Version 1.0]. Retrieved February 18, 2026, from https://user.eumetsat.int/s3/eup-strapi-media/Sentinel_3_Product_Notice_OLCI_L1_PB_OL_L1_002_22_00_3d0884a890.pdf

Fernández, N. I., Bernard, F., Garcia, A., Alonso, K., Manolis, I., Celesti, M., Dransfeld, S., Gabriele, A., Barat Sanjuán, I., & Bolea Alamanac, A. (2025). Geolocation error impact on retrieval uncertainty: A new algorithm for high-resolution earth observation missions. *Proceedings of SPIE — Sensors, Systems, and Next-Generation Satellites XXIX*, *13667*, Article 136670L. https://doi.org/10.1117/12.3071924

Food and Agriculture Organization of the United Nations. (2023). *The state of food security and nutrition in the world 2023: Urbanization, agrifood systems transformation and healthy diets across the rural–urban continuum*. https://doi.org/10.4060/cc3017en

Fox, N. (2010, March). *A guide to expression of uncertainty of measurements* (tech. rep.) (QA4EO Guide Ref. QA4EO-QAEO-GEN-DQK-006). GEO.

Gorroño, J., Guanter, L., Valentin Graf, L., & Gascon, F. (2024). A framework for the estimation of uncertainties and spectral error correlation in sentinel-2 level-2a data products. *IEEE Transactions on Geoscience and Remote Sensing*, *62*, 1–13. https://doi.org/10.1109/TGRS.2024.3435021

Hu, S. (2025). *Satellite-assisted wofost yield prediction for green maize and winter barley in the netherlands* [Master's thesis, Delft University of Technology].

Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. (2019). *Global assessment report on biodiversity and ecosystem services*. https://doi.org/10.5281/zenodo.3831673

Judith, J., Tamilselvi, R., Beham, M. P., Lakshmi, S. S. P., Panthakkan, A., Al Mansoori, S., & Al Ahmad, H. (2025). Remote sensing based crop health classification using ndvi and fully connected neural networks. *arXiv*. https://doi.org/10.48550/arXiv.2504.10522

Knowledge Centre on Earth Observation. (2025). *Earth observation — knowledge for policy*. European Commission. Retrieved February 9, 2026, from https://knowledge4policy.ec.europa.eu/earthobservation_en

Lamquin, N., et al. (2013). *Sentinel-3 optical products and algorithm definition: Pixel-by-pixel uncertainty propagation in olci clear water branch* (tech. rep.) (Ref. S3-L2-SD-01-C01-ACR-TN).

Nieke, J., Despoisse, L., Gabriele, A., Weber, H., Strese, H., Ghasemi, N., Gascon, F., Alonso, K., Boccia, V., Tsonevska, B., Choukroun, P., Ottavianelli, G., & Celesti, M. (2023). The copernicus hyperspectral imaging mission for the environment (chime): An overview of its mission, system and planning status. *Sensors, Systems, and Next-Generation Satellites XXVII*. https://doi.org/10.1117/12.2679977

Povey, A. C., & Grainger, R. G. (2015). Known and unknown unknowns: Uncertainty estimation in satellite remote sensing. *Atmospheric Measurement Techniques*, *8*(11), 4699–4718. https://doi.org/10.5194/amt-8-4699-2015

Qian, S.-E. (2021). Hyperspectral satellites, evolution, and development history. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. https://doi.org/10.1109/JSTARS.2021.3090256

Rast, M., et al. (2021). Copernicus hyperspectral imaging mission for the environment (chime). *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 108–111. https://doi.org/10.1109/IGARSS47720.2021.9553319

Roessel, A. (2024). *Uncertainty review and analysis of chime's level 1 data products* [Master's thesis] [Master's Thesis ESA – ETH Zurich].

Timmermans, J., & Kissling, W. D. (2023). Advancing terrestrial biodiversity monitoring with satellite remote sensing in the context of the kunming-montreal global biodiversity framework. *Ecological Indicators*, *154*, 110773. https://doi.org/10.1016/j.ecolind.2023.110773

United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development*. Retrieved February 18, 2026, from https://sdgs.un.org/2030agenda

Verrelst, J., et al. (2021). Prototyping vegetation traits models in the context of the hyperspectral chime mission preparation. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 7678–7681. https://doi.org/10.1109/IGARSS47720.2021.9554407

Vines, P. L., & Zhang, J. (2022). High-throughput plant phenotyping for improved turfgrass breeding applications. *Grass Research*, *2*, 1. https://doi.org/10.48130/GR-2022-0001

World Economic Forum. (2026). *The global risks report 2026*. Retrieved February 18, 2026, from https://www.weforum.org/publications/global-risks-report-2026/in-full/

Zheng, J., Yu, L., Du, Z., & Xiao, L. (2025). Improving wheat phenology and yield forecasting with a deep learning-enhanced wofost model under extreme weather conditions. *EGUsphere*, *2025*, 1–29. https://doi.org/10.5194/egusphere-2024-4010

A

CHIME

CHIME Mission and System Summary	
Satellite heritage	• Platform: MILA (under development) • Instrument: HyperSpectral Instrument (HSI)
Payload	• On-ground swath: ~130km (at equator), composite of 3 identical spectrometers with 3 overlapping staggered sub swaths) • Spatial Sampling Distance (SSD): 30m • MTF: >0.25 ALT (along track); >0.4 ACT (across track) • > 200 spectr. channels within 400-2500nm, Spectral Sampling Interval (SSI) <10nm • Single channel solution offering excellent spatial co-registration across FoV • Spatial/Spectral co-registration at 0.1 SSD/SSI • Compliant with NEDL requirements • Radiometric accuracy < 4%; stability < 0.5% at Level 1
Satellite mass / power	• ~1.6 ton (Payload: 526kg) • ~1.4 kW (Payload: ~360 W)
AOCS	• Star Tracker optical heads interface directly to optical bench • Gyroless nominal operating mode • Safe Mode (w. magnetometer, sun sensors, reaction wheels and magnetic actuators) • Thrusters for orbit control • Autonomous LEOP sequence and simple NOM recovery
Payload data Tx	• Ka-Bands single polarisation: from 2 x 1.85 Gbps
Ground Segment	• Flight Operations Segment (FOS) • Core Ground Segment (GS) • High latitude stations (e.g., Svalbard) for Payload data downlink
Launcher	• VEGA-C (Vampire 1194 adaptor) in single launch configuration from CSG. • Compatibility with back-up launcher (e.g., Ariane 62) required until PDR

Figure A.1: CHIME mission and system summary (table extracted from Rast et al., 2021).

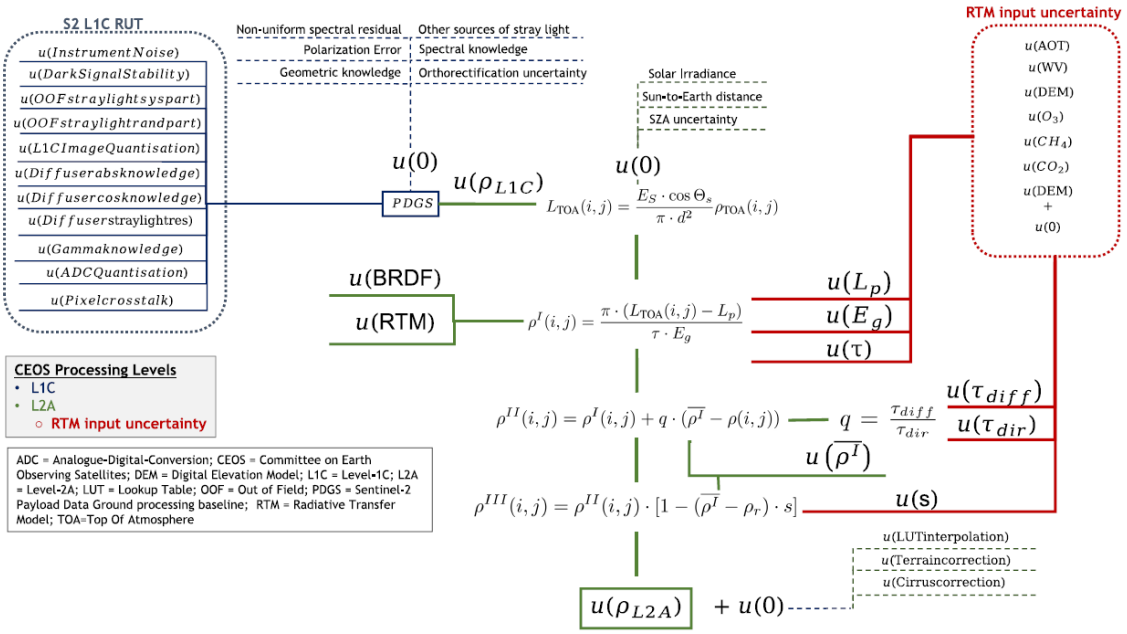


Figure A.2: Uncertainty tree diagram for the Sentinel-2 L2A products (flat terrain case). Figure retrieved from Gorrono et al. (2017).

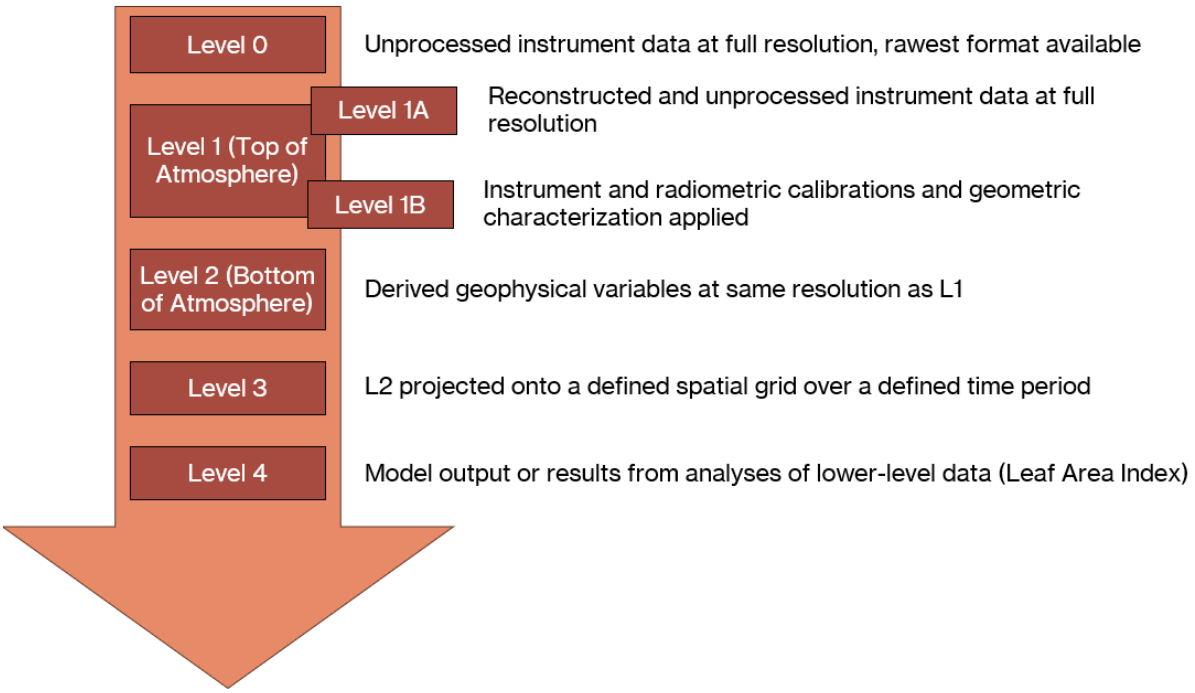


Figure A.3: Satellite data product level classification.

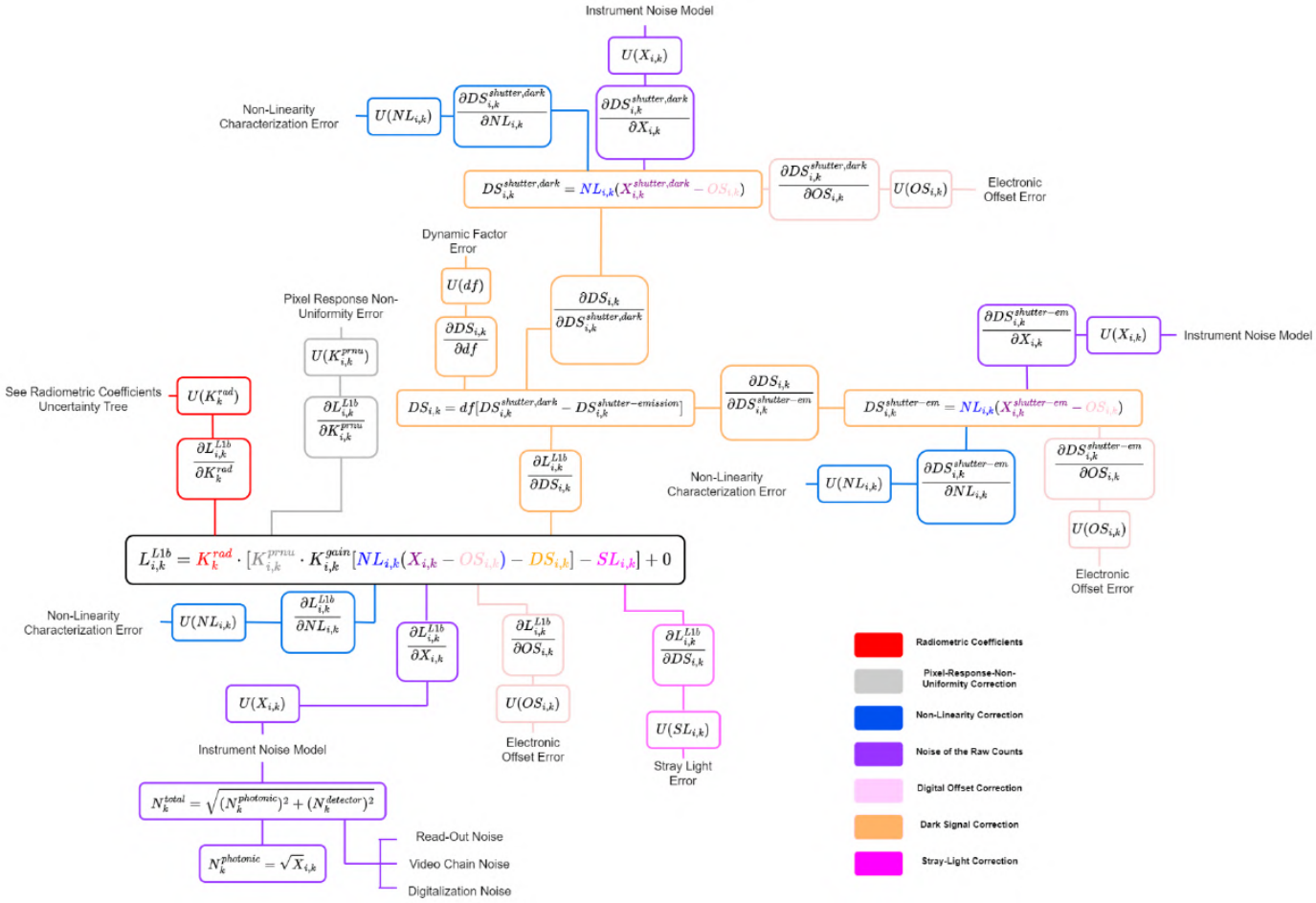


Figure A.4: Uncertainty tree diagram for L1B data.

Table A.1: Summary of optimisations implemented for the Exact and RSS solvers.

Optimisation	Description	Reason / Impact
Replaced `mpmath.quad` with `scipy.integrate.quad`	Used SciPy's wrapper of the QUADPACK Fortran library instead of Python-level arbitrary-precision integration.	`mpmath.quad` is slow due to arbitrary-precision arithmetic and Python overhead. QUADPACK runs in compiled Fortran/C, performs adaptive quadrature efficiently, and significantly reduces evaluation overhead.
Pre-built grids and samples	Pixel grids and geolocation-shift samples are constructed *once* and indexed directly during computation.	Avoids repeatedly rebuilding the 3×3 neighbourhood for every iteration, reducing repeated allocations and Python overhead.
Vectorised pixel storage	All pixel values stored in arrays rather than individual scalar variables.	Enables efficient NumPy-level operations, reduces Python object handling, and avoids repetitive scalar manipulation.
Vectorised neighbour comparisons	Replaced multiple `if/elif` blocks for neighbour–centre comparisons with array operations.	Removes slow Python branching, enables SIMD-friendly array operations, and simplifies code structure.
Generalised bound-calculation function (Exact solver)	Exact solver rewritten using a general function that handles all quadrants and sign combinations.	Eliminates long chains of conditional statements, reduces redundancy, and speeds up quantile computation.

B

Current geolocation uncertainty characterization

This appendix summarises the current state-of-the-art framework for quantifying geolocation-error-induced radiometric uncertainty, as introduced by Fernández et al. (2025). The objective is to clearly define the mathematical starting point of this thesis: the deterministic model, the probabilistic interpretation, the baseline solvers, and their known limitations.

This appendix therefore establishes the theoretical and algorithmic baseline from which the improvements presented in Chapters 4–6 were developed.

B.1. Deterministic formulation of geolocation uncertainty

Let $L(i, j)$ denote the spectral radiance at pixel (i, j) in the ideal geometry. Due to imperfect knowledge of the pixel barycentre, the true sampling location is displaced by subpixel geolocation errors ε_x and ε_y.

Under bilinear interpolation, the perturbed radiance can be written as a weighted combination of neighbouring pixels. Retaining only first-order terms and neglecting the second-order cross-term $\varepsilon_x \varepsilon_y$, the radiometric perturbation becomes

$$\delta L(i, j) = L_{\text{pert}}(i, j) - L(i, j) \approx \varepsilon_x \Delta_x L(i, j) + \varepsilon_y \Delta_y L(i, j).$$

This corresponds to a first-order linearisation of the bilinear resampling model and is equivalent to a discrete gradient formulation:

$$\delta L = \frac{\partial L}{\partial x} \varepsilon_x + \frac{\partial L}{\partial y} \varepsilon_y.$$

The local radiance contrasts are defined as

$$\begin{aligned} \alpha_1 &= L(i, j-1) - L(i, j), \\ \alpha_2 &= L(i, j+1) - L(i, j), \\ \alpha_3 &= L(i-1, j) - L(i, j), \\ \alpha_4 &= L(i+1, j) - L(i, j). \end{aligned}$$

Which neighbour contributes depends on the sign of the corresponding geolocation error component.

This deterministic first-order formulation reveals the physical driver of geolocation uncertainty: local radiometric contrast combined with geolocation error magnitude.

B.2. Probabilistic interpretation: mixture of signed half-normals

B.2.1. Geolocation error model

Geolocation errors are assumed to follow independent Gaussian distributions:

$$\varepsilon_x \sim \mathcal{N}(0, \sigma_x^2), \qquad \varepsilon_y \sim \mathcal{N}(0, \sigma_y^2), \qquad \varepsilon_x \perp \varepsilon_y.$$

In the isotropic case,

$$\sigma_x = \sigma_y = \sigma_{\text{gel}}.$$

B.2.2. Half-normal contributions

Because neighbour selection depends on the sign of ε_x and ε_y, each directional contribution is conditionally a scaled half-normal distribution.

If $Z \sim \mathcal{N}(0, \sigma^2)$, then $|Z|$ follows a half-normal distribution.

The horizontal contribution can be written as

$$X = \begin{cases} \operatorname{sgn}(\alpha_1)\,|\alpha_1|\,|\varepsilon_x|, & \varepsilon_x < 0, \\ \operatorname{sgn}(\alpha_2)\,|\alpha_2|\,|\varepsilon_x|, & \varepsilon_x \geq 0. \end{cases}$$

Similarly, the vertical contribution Y is defined analogously using α_3 and α_4.

B.2.3. Mixture structure

Since the sign of each geolocation component is equally likely, the horizontal term is a two-component mixture of signed half-normals, and the vertical term is another two-component mixture.

The total first-order error

$$Z = X + Y$$

is therefore a four-component mixture distribution.

Each component corresponds to the convolution of two independent signed half-normal distributions. No closed-form expression exists for the cumulative distribution function of this mixture in elementary functions. Consequently, quantiles must be computed numerically.

This mixture structure explains:

- The inherent asymmetry of geolocation error distributions,
- The dependence on local contrast configuration,
- The absence of a simple Gaussian description.

B.3. Baseline solver strategies

In practice, the goal is not to reconstruct the full probability density function of Z, but to estimate its 16th and 84th percentiles (equal-tailed 68% interval).

Two baseline approaches were proposed.

B.3.1. RSS solver (analytical approximation)

The RSS solver approximates the asymmetric mixture distribution by two one-sided scale measures.

Define the rectified linear unit:

$$r(x) = \max(0, x).$$

The upper scale parameter is

$$\sigma_{\text{hl}} = \sigma_{\text{gel}} \sqrt{r(\alpha_1)^2 + r(\alpha_2)^2 + r(\alpha_3)^2 + r(\alpha_4)^2}.$$

The lower scale parameter is

$$\sigma_{\text{ll}} = \sigma_{\text{gel}} \sqrt{r(-\alpha_1)^2 + r(-\alpha_2)^2 + r(-\alpha_3)^2 + r(-\alpha_4)^2}.$$

Bounds are approximated as

$$\text{upper} \approx +\sigma_{\text{hl}}, \qquad \text{lower} \approx -\sigma_{\text{ll}}.$$

This approach is computationally efficient and suitable for operational pipelines. However, it collapses the mixture distribution into Gaussian-equivalent one-sided scales and may:

- Overestimate uncertainty in many cases,
- Degenerate to zero on one side in strongly unidirectional contrast cases.

B.3.2. Exact first-order solver

The "Exact" solver evaluates the mixture CDF numerically via convolution and inverts it to obtain

$$l = F_Z^{-1}(0.16), \qquad u = F_Z^{-1}(0.84).$$

This solver is exact with respect to the first-order probabilistic model (neglecting $\varepsilon_x \varepsilon_y$ terms). It significantly reduces the overestimation observed in RSS but at the cost of:

- Per-pixel numerical integration,
- Per-pixel root finding,
- High computational burden.

B.4. Known limitations of the baseline framework

B.4.1. First-order truncation and neglected corner effects

The probabilistic formulation underlying both RSS and the Exact solver is derived from a first-order truncation of the bilinear resampling model. The second-order cross-term $\varepsilon_x \varepsilon_y \tau$ is neglected, where

$$\tau = L(i+1, j+1) - L(i+1, j) - L(i, j+1) + L(i, j).$$

The cross-term captures the interaction between horizontal and vertical gradients. It scales with the product $\varepsilon_x \varepsilon_y$. For small subpixel displacements, this product is assumed negligible.

However, its magnitude also depends on the curvature term τ. In scenes with strong radiometric curvature (e.g. coastlines, sharp field boundaries, small high-contrast objects), τ may become large. Under such configurations,

$$\varepsilon_x \varepsilon_y \tau$$

may no longer be negligible relative to first-order terms.

Consequently:

- The Exact solver is exact only with respect to the truncated first-order model.
- Even perfect numerical integration cannot eliminate modelling bias.
- Residual deviations relative to Monte Carlo simulations are theoretically expected in high-curvature scenes.

B.4.2. Degeneracy and asymmetry handling in RSS solver

The RSS solver collapses the four-component mixture into two one-sided scale parameters. If all gradients share the same sign (e.g. $\alpha_k > 0$ for all neighbours), one side of the interval collapses.

The true distribution remains skewed but non-degenerate. RSS therefore misrepresents tail behaviour in unidirectional contrast cases.

B.4.3. Dependence on ideal geometry

The derivation assumes access to the ideal radiance field $L(i, j)$. Operationally, only the perturbed radiance L_{pert} is available.

Under small-error conditions, contrasts may be approximated by product radiances, but this assumption may break down in high-contrast scenes.

B.4.4. Subpixel variability and PSF interactions

The first-order derivation assumes local linearity. Real scenes exhibit subpixel variability, and the instrument PSF couples spatial structure with geolocation error.

If the first-order model fails under realistic PSF conditions, solver development based on the truncated formulation would lack physical validity.

B.4.5. Validation metric sensitivity

Original validation emphasised relative width error:

$$\text{relative\_error} = \frac{\text{width}_{\text{solver}} - \text{width}_{\text{MC}}}{\text{width}_{\text{MC}}}.$$

Relative metrics become unstable when MC widths are small. Width-only validation also ignores asymmetry in lower and upper bounds.

Operational interpretation requires stable, physically meaningful metrics.

B.4.6. Computational scalability

The Exact solver requires per-pixel numerical convolution and root finding. While accurate under the first-order model, it is computationally expensive for hyperspectral image cubes.

A trade-off between numerical fidelity and runtime must therefore be quantified.

C

Scenes

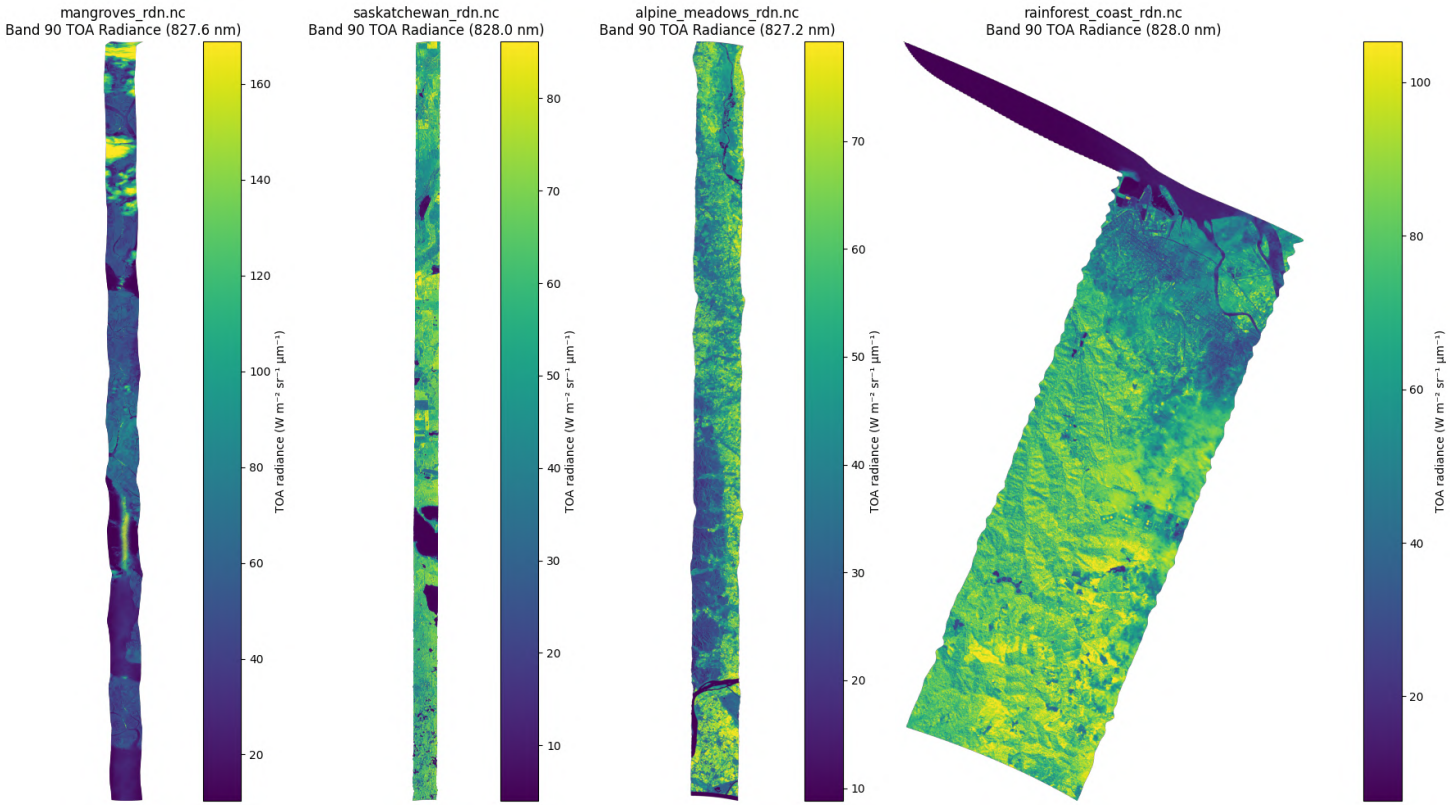


Figure C.1: Top-of-atmosphere NIR radiance for representative scenes: mangroves (India), agricultural field with water patches (Canada), alpine meadows (India), and rainforest (India), shown from left to right.

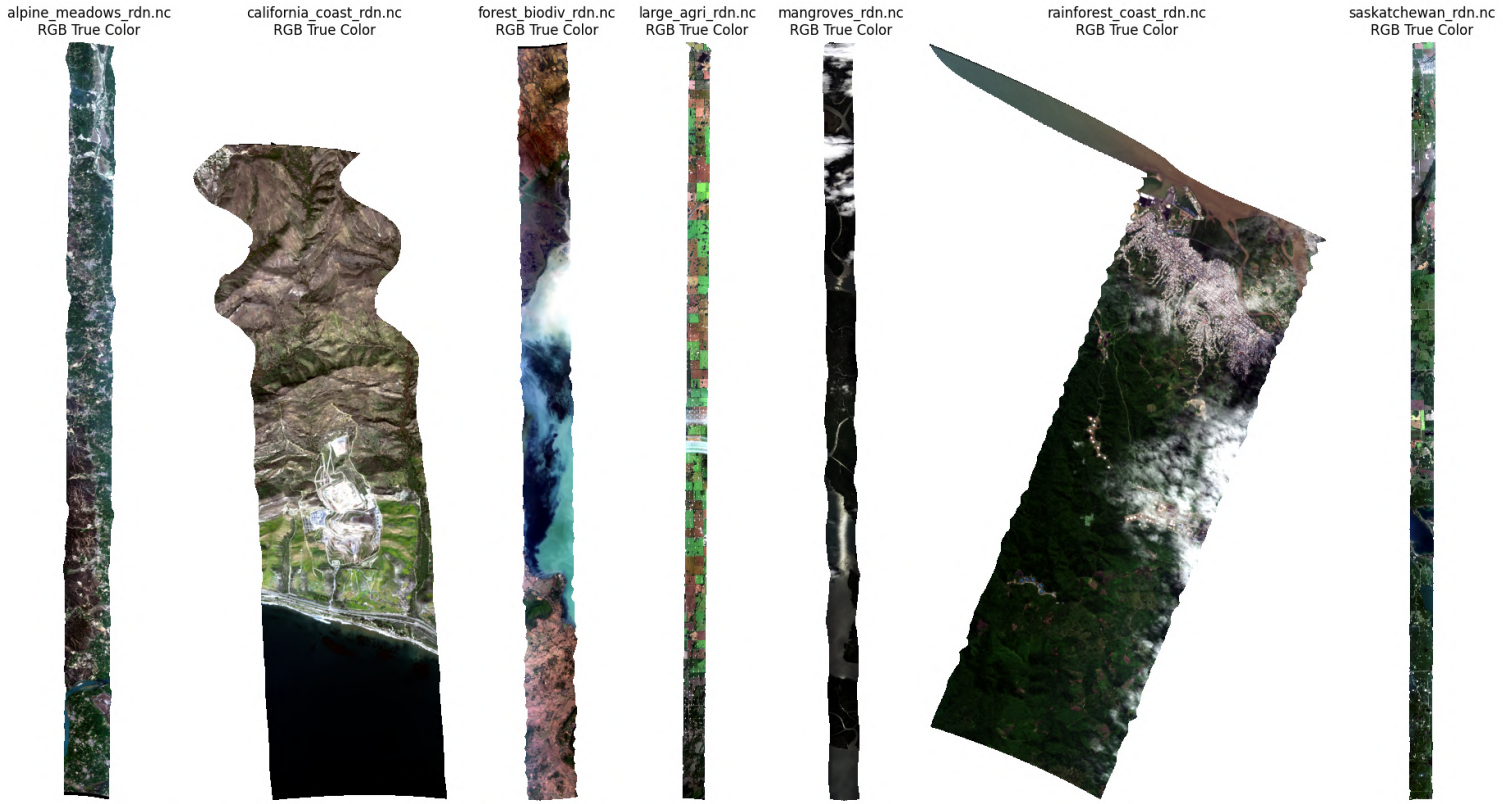
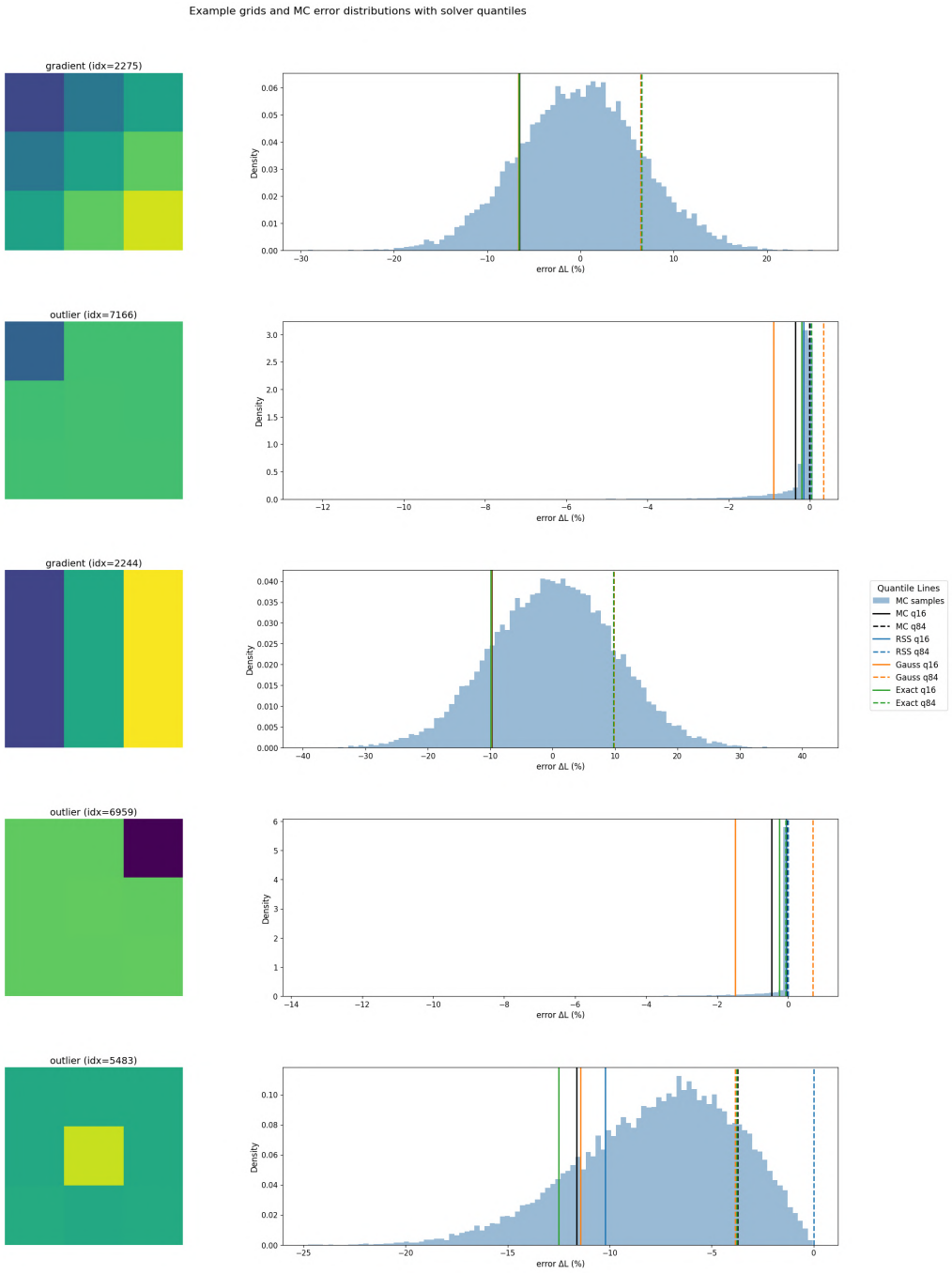


Figure C.2: Top-of-atmosphere RGB radiance for representative scenes: alpine meadows (India), California coast (USA), forest with high biodiversity (India), agricultural scene (Canada), mangroves (India), rainforest (India), and agricultural field with water patches (Canada), shown from left to right.

D

Results

D.1. Synthetic grids



captionMonte Carlo error distribution (blue histogram) and solver bound estimates for a synthetic random grid dataset ($\sigma_{\text{gel}} = 0.15\,\text{px}$).

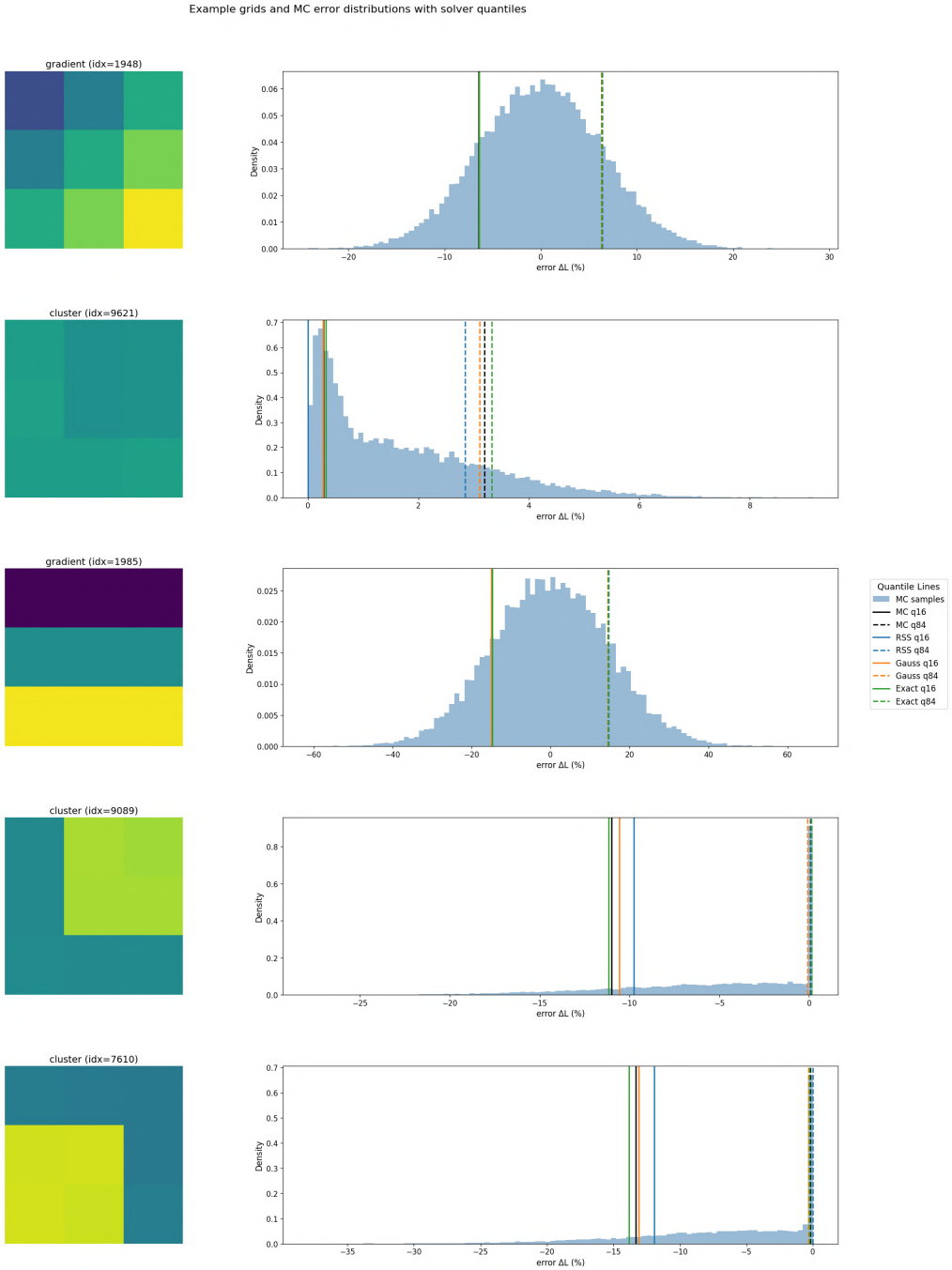


Figure D.1: Monte Carlo error distribution (blue histogram) and solver bound estimates for a synthetic random grid dataset ($\sigma_{\text{gel}} = 0.15\,\text{px}$).

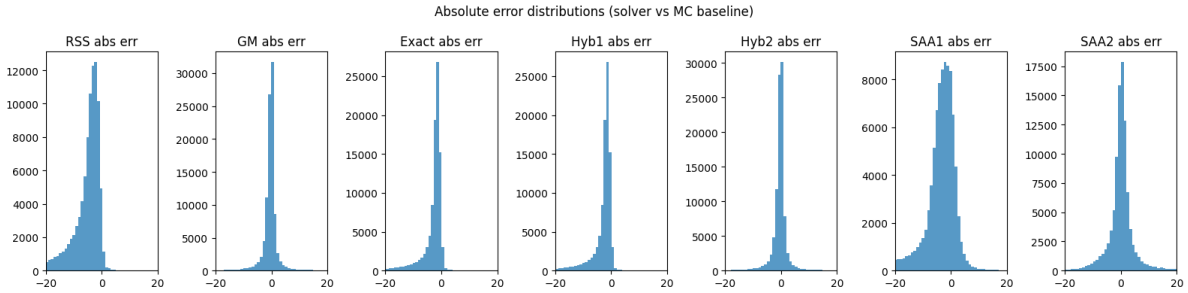


Figure D.2: Error distributions of solver prediction error (solver minus MC) for synthetically generated random grids ($\sigma_{\text{gel}} = 0.15\,\text{px}$).

D.2. Real scenes

D.2.1. Water and agricultural fields in Saskatchewan (NIR, AVIRIS)

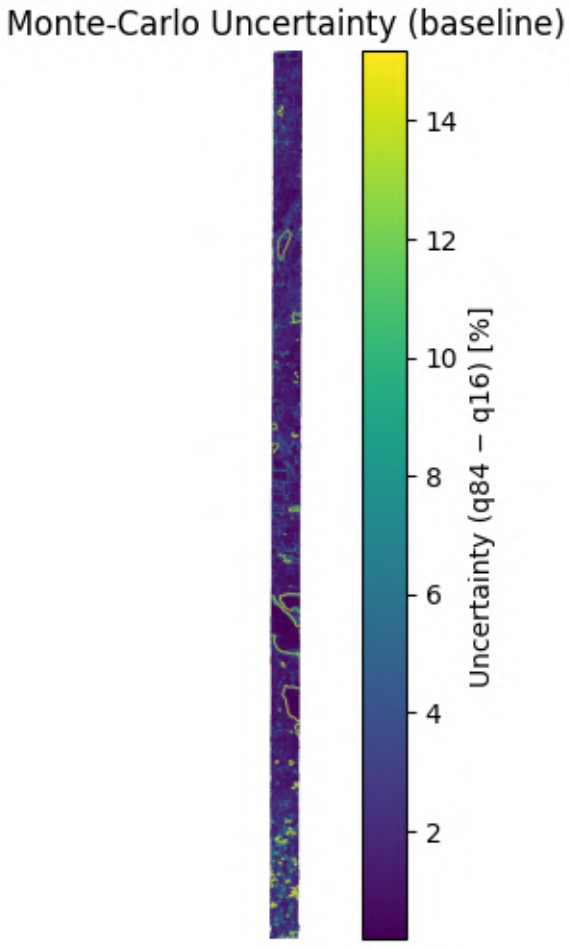


Figure D.3: Monte Carlo reference uncertainty map for the water and agricultural fields scene (Canada) (run_mc = 10,000, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

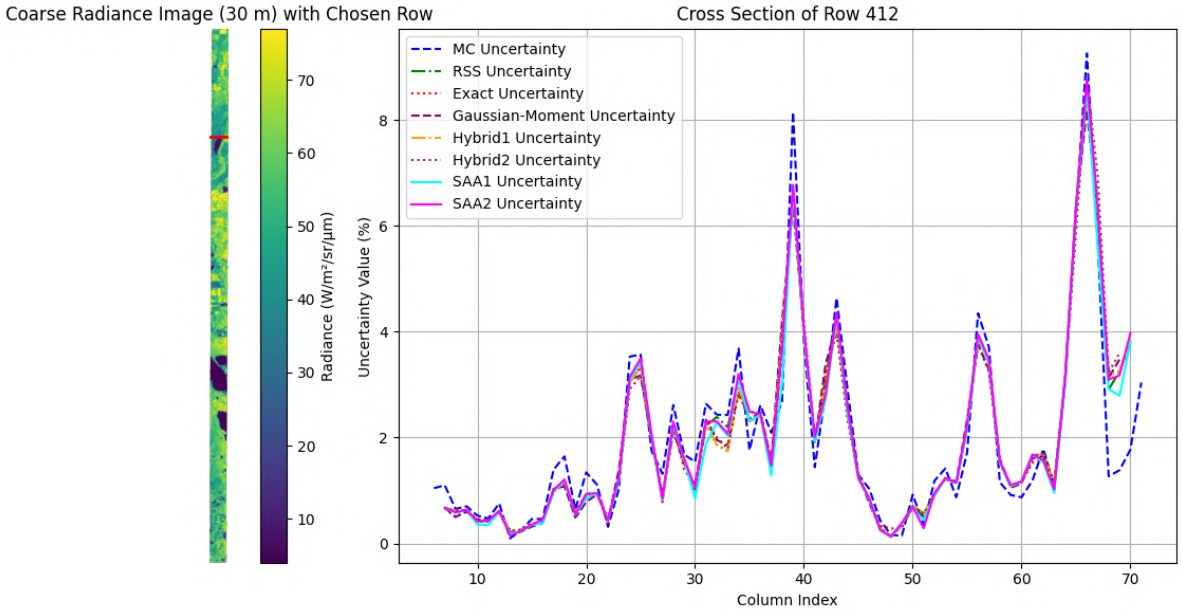


Figure D.4: Cross section of estimated geolocation uncertainty along the profile for each solver (water and agricultural fields; run_mc = 10,000, $\sigma_{\text{gel}} = 0.15$ px).

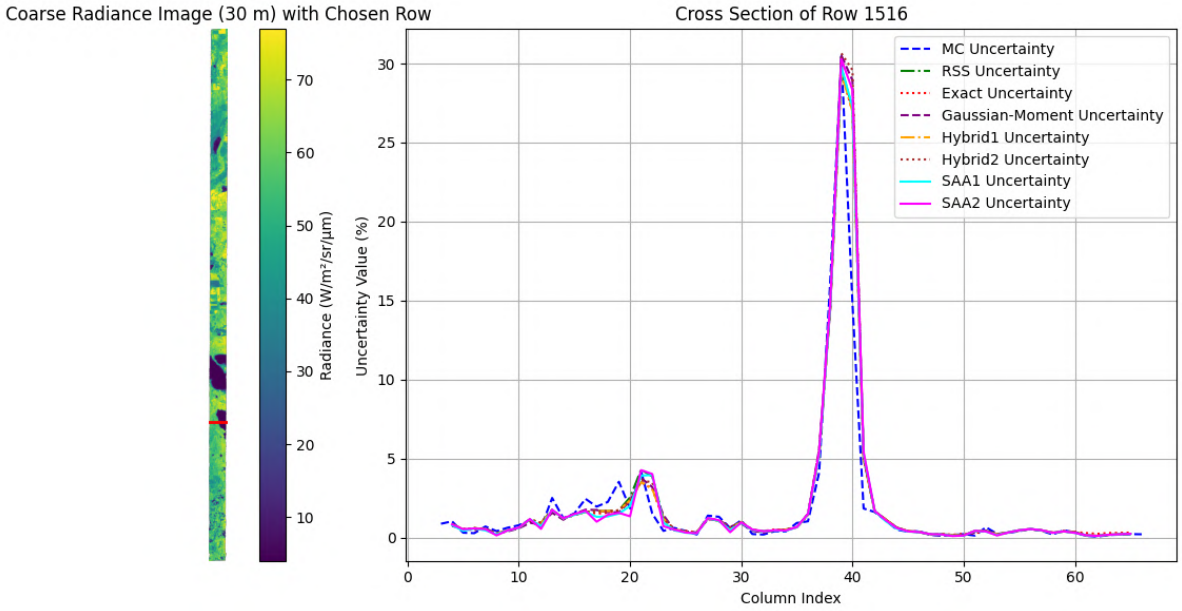


Figure D.5: Cross section of estimated geolocation uncertainty along the profile for each solver (water and agricultural fields; run_mc = 10,000, $\sigma_{\text{gel}} = 0.15$ px).

Table D.1: Absolute errors between solver-predicted and Monte Carlo-derived uncertainty interval widths (geolocation uncertainty) for the water and agricultural fields scene, in percentage points (pp).

Solver	MAE (pp)	RMSE (pp)	MEAN (pp)	MEDIAN (pp)	STD (pp)
RSS	0.4747	0.8971	-0.0302	-0.0367	0.8966
EXACT	0.4690	0.8953	-0.1015	-0.0791	0.8895
GAUSSIAN_MOMENT	0.5029	1.0063	-0.0804	-0.0757	1.0031
HYBRID1	0.4762	0.8995	-0.0677	-0.0501	0.8970

Solver	MAE (pp)	RMSE (pp)	MEAN (pp)	MEDIAN (pp)	STD (pp)
HYBRID2	0.5045	1.0042	-0.0512	-0.0470	1.0029
SAA1	0.4736	0.8735	-0.1272	-0.0951	0.8642
SAA2	0.4734	0.9269	-0.0656	-0.0663	0.9245

Table D.2: Absolute errors between solver-predicted and Monte Carlo-derived lower bound (q16) for the water and agricultural fields scene, in percentage points (pp).

Solver	MAE (pp)	RMSE (pp)	MEAN (pp)	MEDIAN (pp)	STD (pp)
RSS	0.3951	0.7069	0.1079	0.0402	0.6986
EXACT	0.4892	0.8655	0.1669	0.0733	0.8492
GAUSSIAN_MOMENT	0.4468	0.7747	0.1450	0.0703	0.7610
HYBRID1	0.4543	0.8493	0.1529	0.0456	0.8355
HYBRID2	0.4222	0.7612	0.1326	0.0455	0.7495
SAA1	0.3599	0.6757	0.1488	0.0577	0.6591
SAA2	0.2727	0.4874	0.0791	0.0549	0.4810

Table D.3: Absolute errors between solver-predicted and Monte Carlo-derived upper bound (q84) for the water and agricultural fields scene, in percentage points (pp).

Solver	MAE (pp)	RMSE (pp)	MEAN (pp)	MEDIAN (pp)	STD (pp)
RSS	0.4533	1.0682	0.0777	-0.0324	1.0654
EXACT	0.5436	1.2132	0.0654	-0.0640	1.2114
GAUSSIAN_MOMENT	0.5131	1.1770	0.0646	-0.0654	1.1752
HYBRID1	0.5068	1.2014	0.0852	-0.0349	1.1984
HYBRID2	0.4856	1.1633	0.0814	-0.0368	1.1605
SAA1	0.4039	0.9768	0.0216	-0.0414	0.9765
SAA2	0.2843	0.6106	0.0136	-0.0086	0.6105

Table D.4: Absolute errors of predicted uncertainty interval widths relative to Monte Carlo reference values, stratified by radiometric combination case (n_+, n_-), for the water and agricultural fields scene (percentage points).

Case	Solver	MAE	RMSE	MEAN
(4,0)	RSS	0.579	1.199	0.379
	EXACT	0.501	1.019	0.265
	SAA1	0.464	0.878	0.035
	SAA2	0.498	0.883	-0.355
(3,1)	RSS	0.475	0.914	0.040
	EXACT	0.460	0.869	-0.086
	SAA1	0.492	0.857	-0.184
	SAA2	0.493	0.947	0.032
(2,2)	RSS	0.519	0.967	-0.149
	EXACT	0.530	1.003	-0.183
	SAA1	0.517	0.967	-0.141
	SAA2	0.518	1.019	-0.080
(1,3)	RSS	0.367	0.577	-0.015
	EXACT	0.358	0.568	-0.120
	SAA1	0.395	0.638	-0.174
	SAA2	0.385	0.629	-0.130
(0,4)	RSS	0.376	0.530	0.220

Case	Solver	MAE	RMSE	MEAN
	EXACT	0.329	0.462	0.142
	SAA1	0.310	0.440	-0.031
	SAA2	0.331	0.483	-0.198

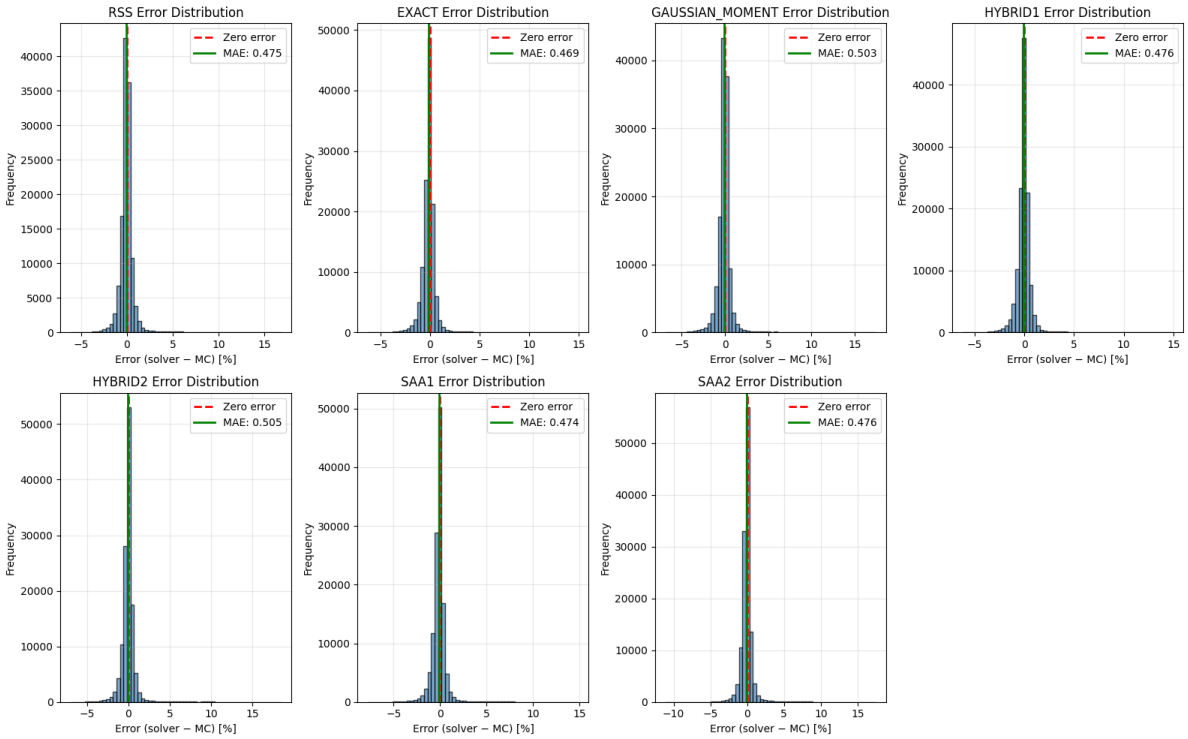


Figure D.6: Error distributions for each solver and the MC baseline for the water and agricultural fields scene, with MAE highlighted in green.

D.2.2. Alpine meadows (NIR, AVIRIS)

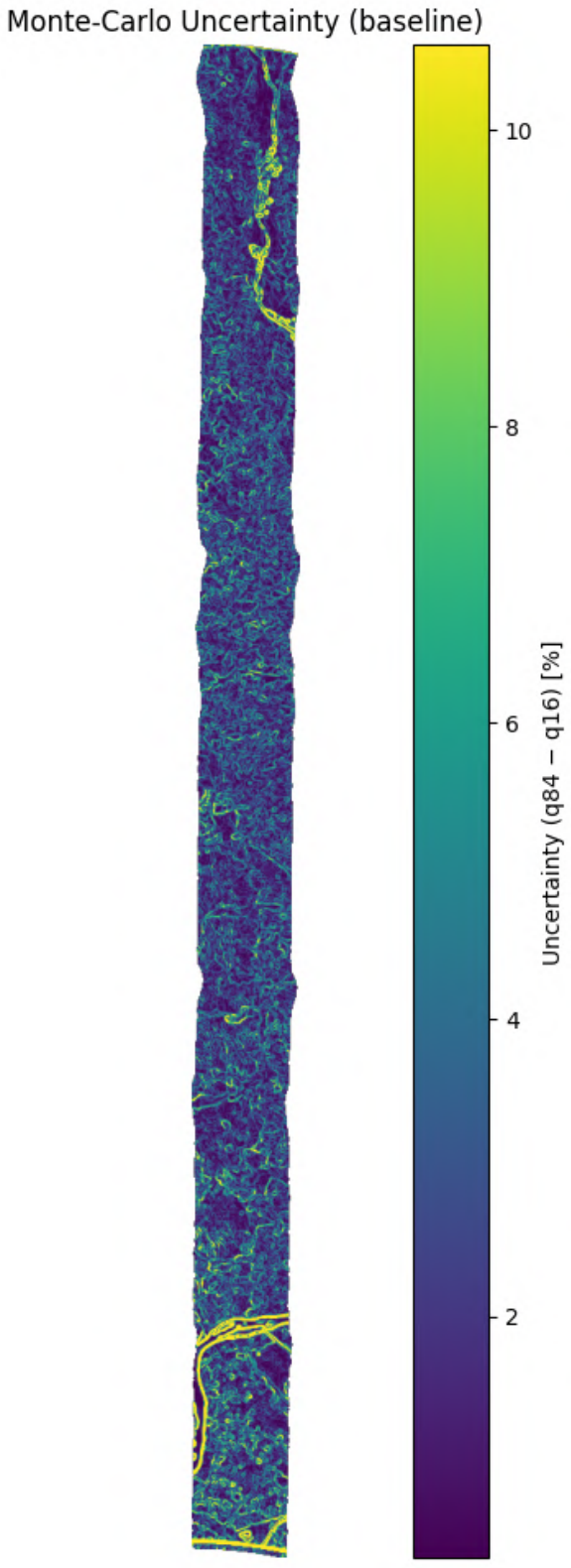


Figure D.7: Monte Carlo reference uncertainty map for the alpine meadows scene (run_mc = 10,000, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

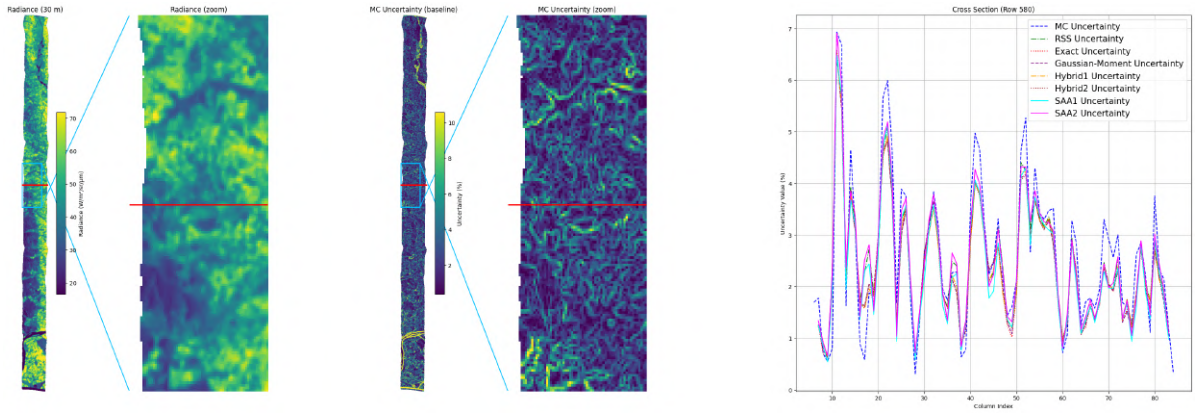


Figure D.8: Cross section of estimated geolocation uncertainty along the profile for each solver (alpine meadows; run_mc = 10,000, $\sigma_{\text{gel}} = 0.15$ px).

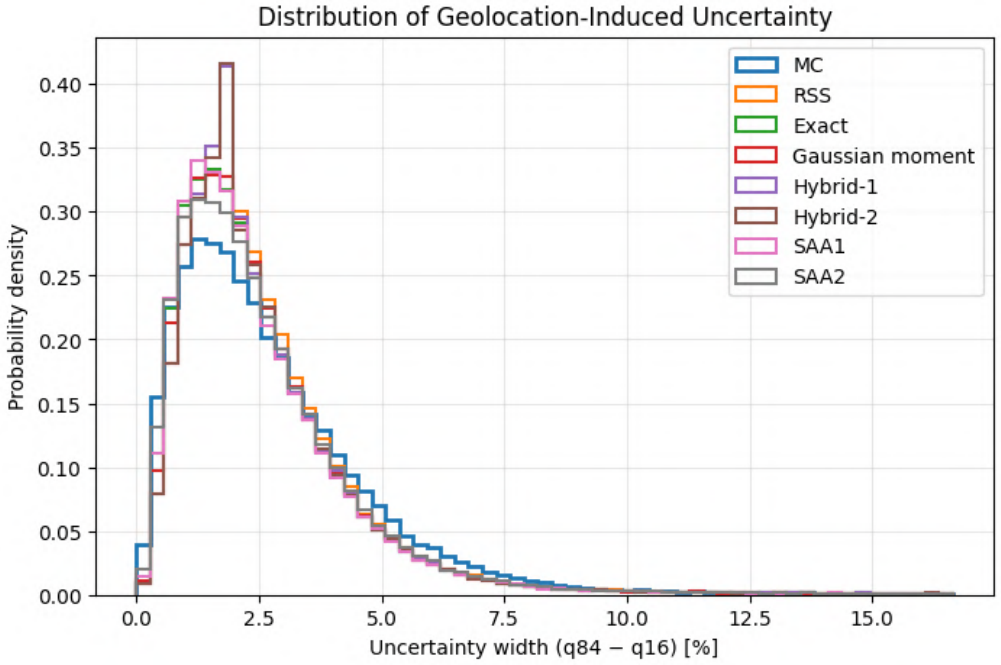


Figure D.9: Geolocation uncertainty distribution per solver for the alpine meadows scene (run_mc = 10,000, $\sigma_{\text{gel}} = 0.15$ px).

Table D.5: Absolute errors between solver-predicted and Monte Carlo-derived uncertainty interval widths for the alpine meadows scene, in percentage points (pp).

Solver	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.8649	7.3640	0.0593	-0.1686	7.3637
EXACT	0.8574	7.2322	-0.0556	-0.2496	7.2320
GAUSSIAN_MOMENT	0.9497	8.8718	0.0266	-0.2367	8.8718
HYBRID1	0.8646	7.2327	-0.0205	-0.2197	7.2327
HYBRID2	0.9508	8.8319	0.0502	-0.2128	8.8317
SAA1	0.8676	7.4013	-0.0457	-0.2514	7.4012
SAA2	0.8769	8.3912	0.1625	-0.1186	8.3896

Table D.6: Absolute errors between solver-predicted and Monte Carlo-derived lower bound (q16) for the alpine meadows scene, in percentage points (pp).

Solver	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.5191	1.0777	0.2144	0.1319	1.0562
EXACT	0.6493	1.2012	0.2829	0.2104	1.1674
GAUSSIAN_MOMENT	0.6089	1.3554	0.1995	0.1974	1.3406
HYBRID1	0.6152	1.1868	0.2688	0.1592	1.1559
HYBRID2	0.5827	1.3389	0.1897	0.1594	1.3254
SAA1	0.4724	1.0494	0.2630	0.1499	1.0159
SAA2	0.4142	2.7834	-0.0055	0.0652	2.7834

Table D.7: Absolute errors between solver-predicted and Monte Carlo-derived upper bound (q84) for the alpine meadows scene, in percentage points (pp).

Solver	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.8101	7.9326	0.2738	-0.0884	7.9279
EXACT	0.9303	7.8333	0.2273	-0.1588	7.8300
GAUSSIAN_MOMENT	0.8974	7.8674	0.2261	-0.1577	7.8642
HYBRID1	0.8947	7.8311	0.2483	-0.1074	7.8271
HYBRID2	0.8686	7.8329	0.2398	-0.1160	7.8292
SAA1	0.7452	7.9623	0.2173	-0.0971	7.9593
SAA2	0.5257	5.6681	0.1570	-0.0474	5.6659

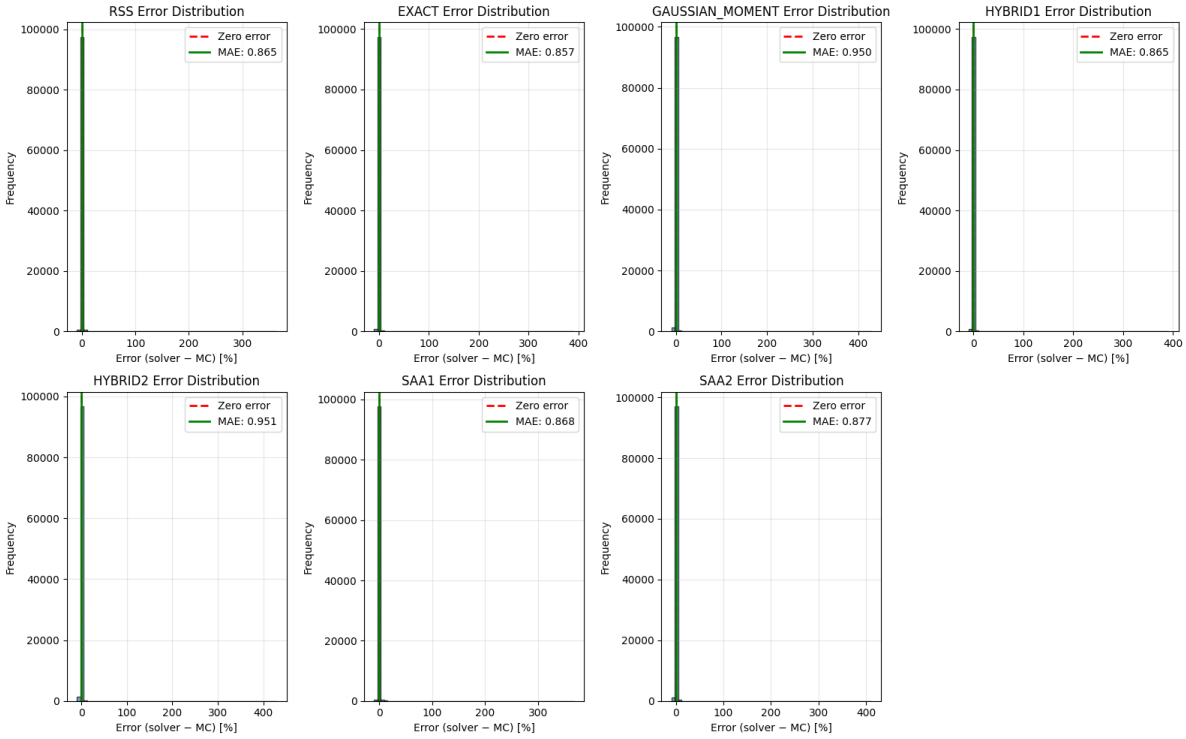


Figure D.10: Error distributions for each solver and the MC baseline for the alpine meadows scene, with MAE highlighted in green.

D.2.3. California coast (NIR, AVIRIS)

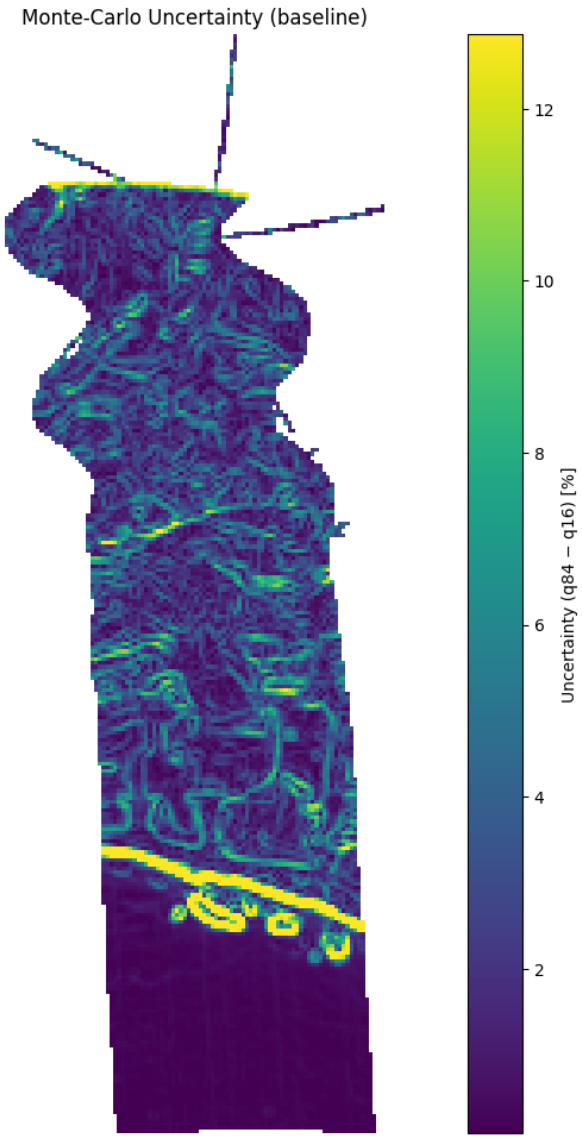


Figure D.11: Monte Carlo reference uncertainty map of the California coastal scene (USA) (run_mc = 10,000, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

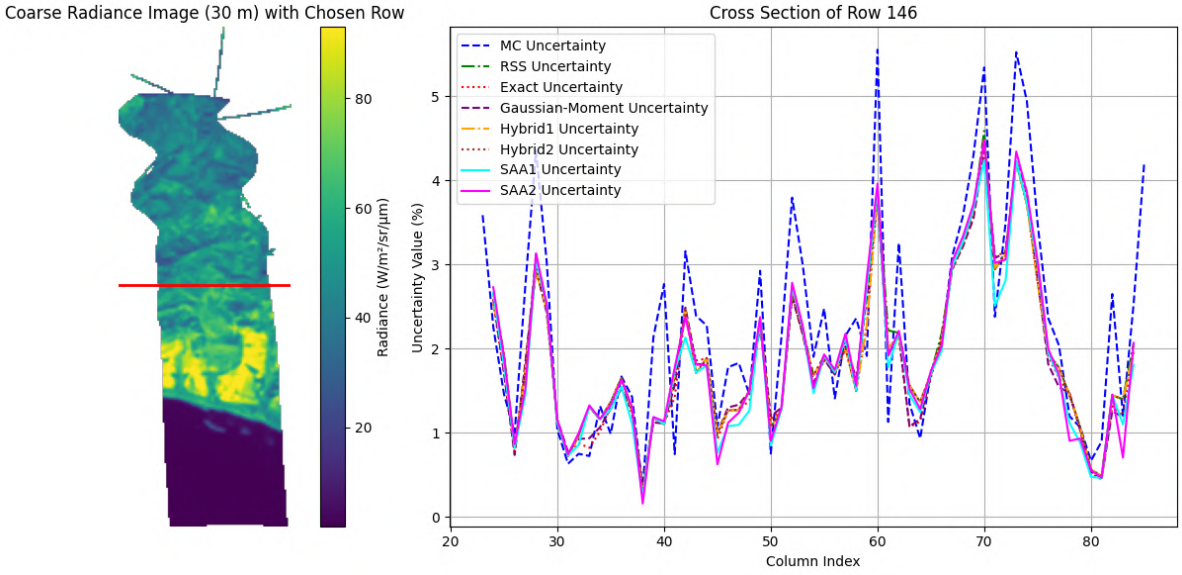


Figure D.12: Cross section of estimated geolocation uncertainty along the profile for each solver (California coastal scene; run_mc = 10,000, $\sigma_{\text{gel}} = 0.15$ px).

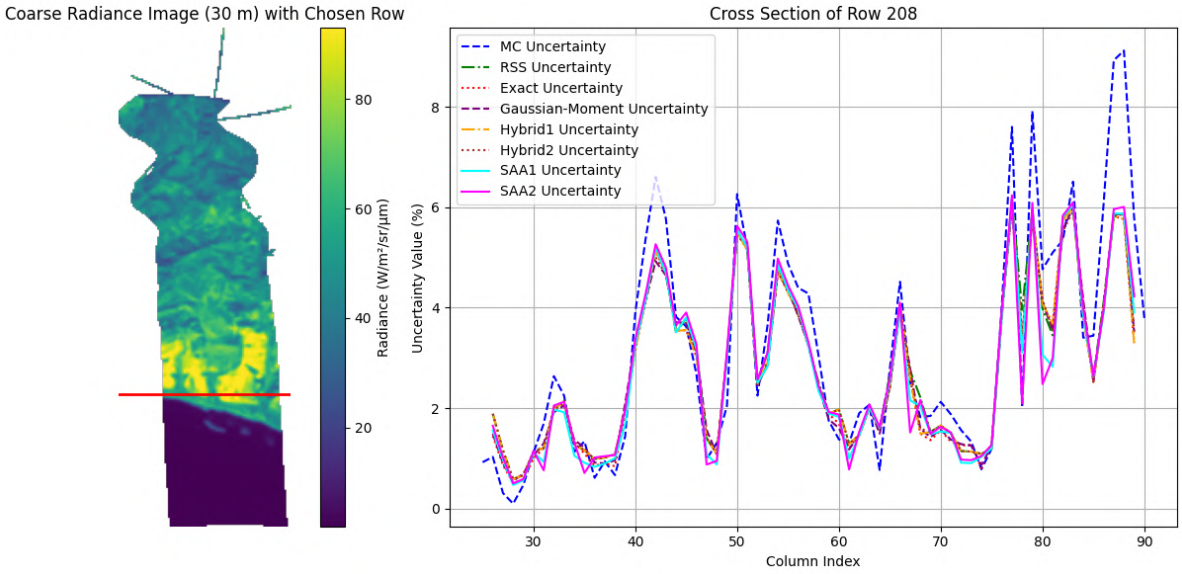


Figure D.13: Cross section of estimated geolocation uncertainty along the profile for each solver (California coastal scene; run_mc = 10,000, $\sigma_{\text{gel}} = 0.15$ px).

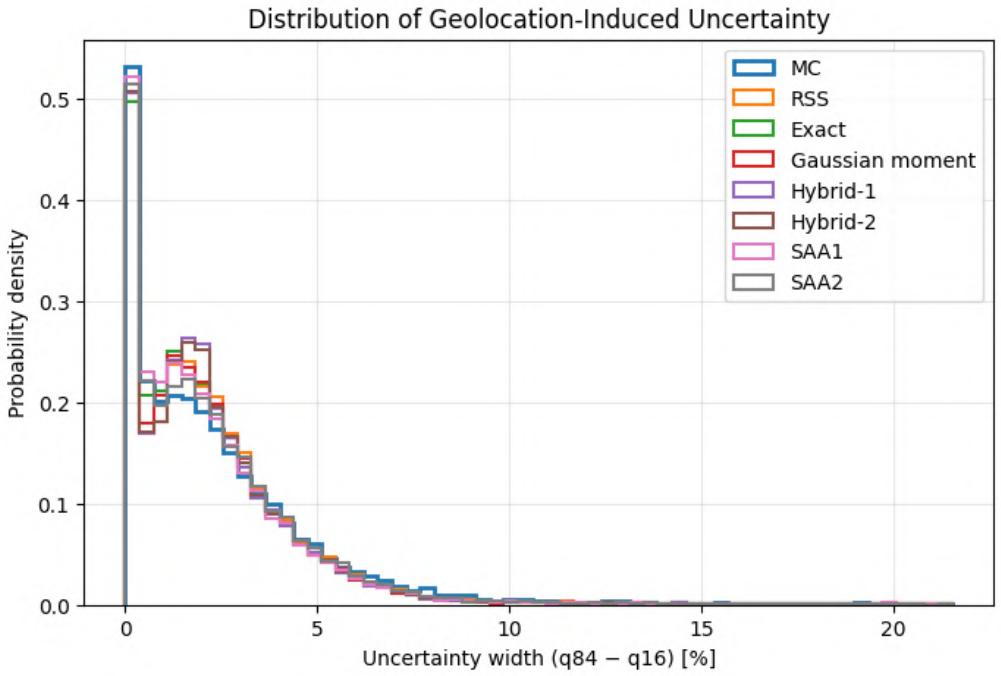


Figure D.14: Geolocation uncertainty distribution per solver for the California coastal scene (run_mc = 10,000, $\sigma_{\text{gel}} = 0.15$ px).

Table D.8: Absolute errors between solver-predicted and Monte Carlo-derived uncertainty interval widths (geolocation uncertainty) for the California coastal scene, in percentage points (pp).

Solver	MAE (pp)	RMSE (pp)	MEAN (pp)	MEDIAN (pp)	STD (pp)
RSS	0.5451	1.1407	-0.0272	-0.0148	1.1404
EXACT	0.5470	1.1411	-0.0957	-0.0136	1.1371
GAUSSIAN_MOMENT	0.5714	1.2937	-0.0715	-0.0300	1.2917
HYBRID1	0.5416	1.1425	-0.0851	-0.0250	1.1393
HYBRID2	0.5740	1.2964	-0.0522	-0.0218	1.2953
SAA1	0.5439	1.1556	-0.1422	-0.0493	1.1468
SAA2	0.5500	1.3186	-0.0008	-0.0104	1.3186

Table D.9: Absolute errors between solver-predicted and Monte Carlo-derived lower bound (q16) for the California coastal scene, in percentage points (pp).

Solver	MAE (pp)	RMSE (pp)	MEAN (pp)	MEDIAN (pp)	STD (pp)
RSS	0.4299	0.7728	0.1491	0.0227	0.7583
EXACT	0.5337	0.8915	0.1924	0.0301	0.8705
GAUSSIAN_MOMENT	0.4702	0.7528	0.1597	0.0361	0.7357
HYBRID1	0.5016	0.8765	0.1895	0.0258	0.8558
HYBRID2	0.4493	0.7389	0.1517	0.0257	0.7232
SAA1	0.3949	0.7403	0.1990	0.0342	0.7130
SAA2	0.2959	0.5799	0.0299	0.0079	0.5791

Table D.10: Absolute errors between solver-predicted and Monte Carlo-derived upper bound (q84) for the California coastal scene, in percentage points (pp).

Solver	MAE (pp)	RMSE (pp)	MEAN (pp)	MEDIAN (pp)	STD (pp)
RSS	0.5280	1.3665	0.1219	-0.0046	1.3610
EXACT	0.6289	1.4618	0.0966	-0.0014	1.4586
GAUSSIAN_MOMENT	0.5885	1.4008	0.0883	-0.0192	1.3981
HYBRID1	0.5953	1.4527	0.1044	-0.0074	1.4490
HYBRID2	0.5668	1.3961	0.0995	-0.0090	1.3925
SAA1	0.4806	1.3532	0.0569	-0.0150	1.3520
SAA2	0.3326	0.8745	0.0291	-0.0053	0.8740

Table D.11: Absolute errors of predicted uncertainty interval widths (geolocation uncertainty) relative to Monte Carlo reference values, stratified by radiometric combination case (n_+, n_-), for the California coastal scene (percentage points).

Method	Case	MAE	RMSE	MEAN
	(4,0)	0.733	1.090	0.503
	(3,1)	0.487	0.874	0.073
RSS	(2,2)	0.591	1.334	-0.118
	(1,3)	0.436	0.768	-0.020
	(0,4)	0.547	0.752	0.360
	(4,0)	0.629	0.967	0.361
	(3,1)	0.467	0.821	-0.064
EXACT	(2,2)	0.612	1.356	-0.141
	(1,3)	0.438	0.765	-0.132
	(0,4)	0.466	0.633	0.246
	(4,0)	0.564	0.840	0.081
	(3,1)	0.500	0.827	-0.161
SAA1	(2,2)	0.589	1.338	-0.110
	(1,3)	0.475	0.870	-0.205
	(0,4)	0.436	0.617	-0.002
	(4,0)	0.599	0.875	-0.390
	(3,1)	0.503	0.895	0.063
SAA2	(2,2)	0.591	1.429	-0.044
	(1,3)	0.463	0.866	-0.152
	(0,4)	0.451	0.673	-0.243

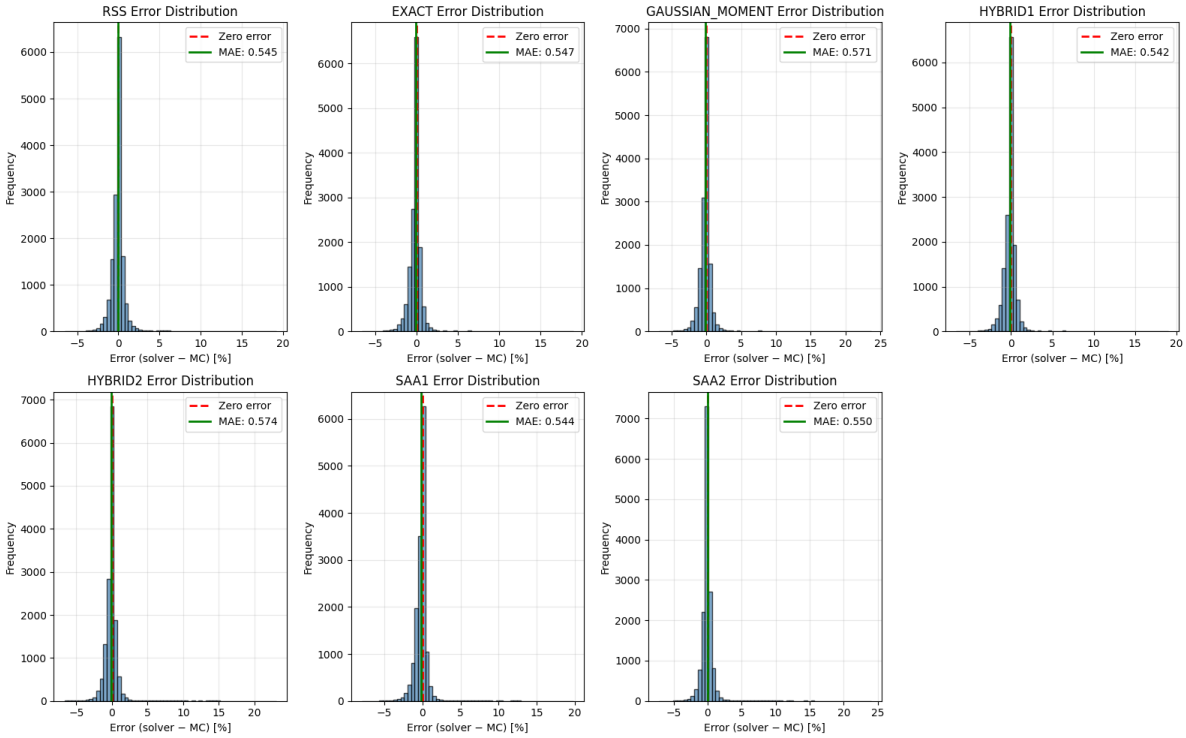


Figure D.15: Error distributions for each solver and the MC baseline for the California coastal scene, with MAE highlighted in green.

D.2.4. Forest with biodiversity in India (NIR, AVIRIS)

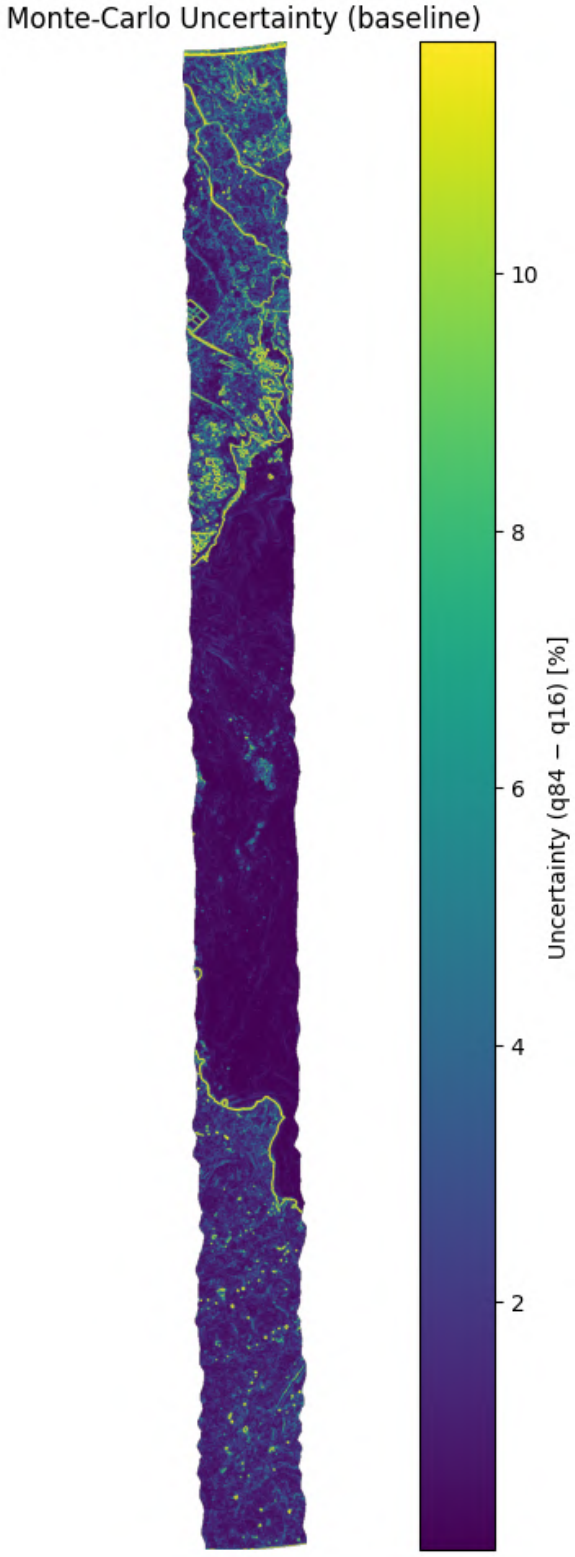


Figure D.16: Monte Carlo reference uncertainty map of the high-biodiversity forest scene (India) (run_mc = 10,000, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

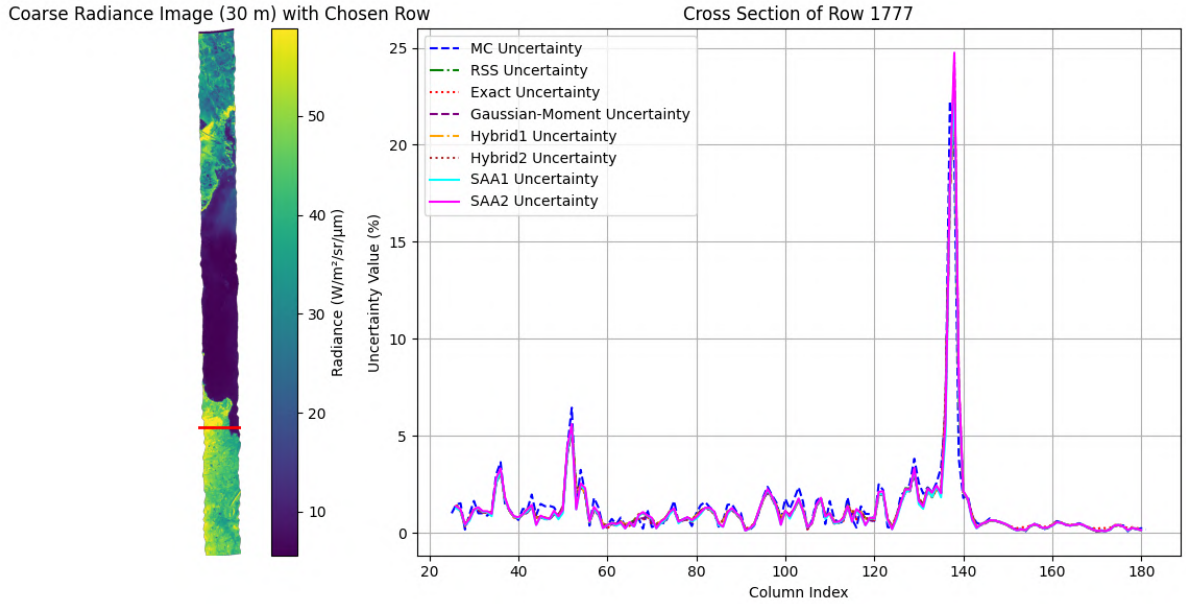


Figure D.17: Cross section of estimated geolocation uncertainty along the profile for each solver (high-biodiversity forest; run_mc = 10,000, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

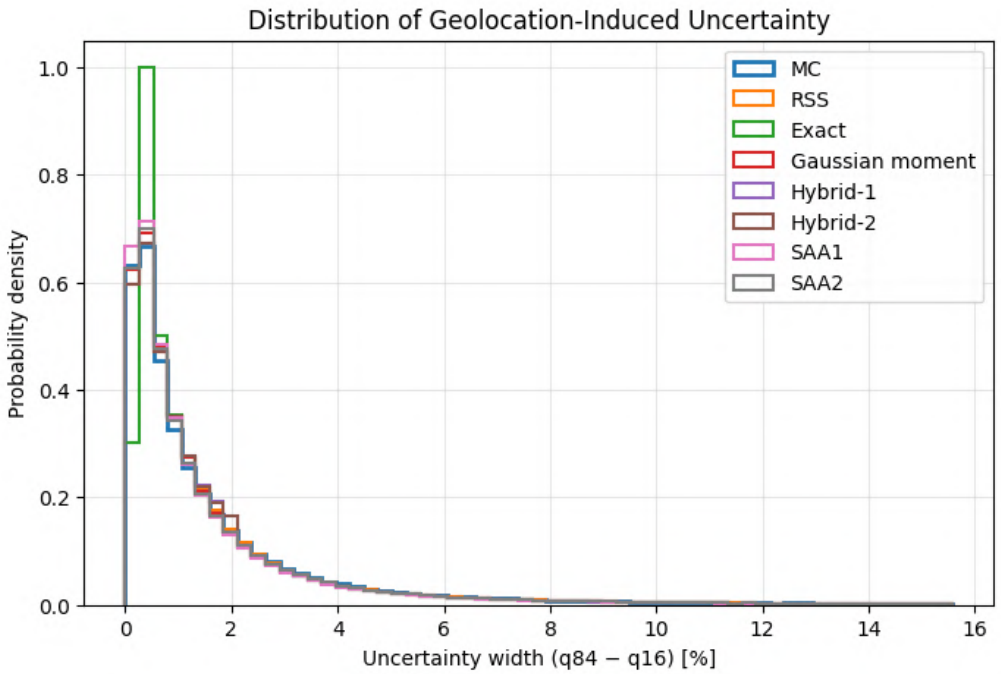


Figure D.18: Geolocation uncertainty distribution per solver for the high-biodiversity forest scene (run_mc = 10,000, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

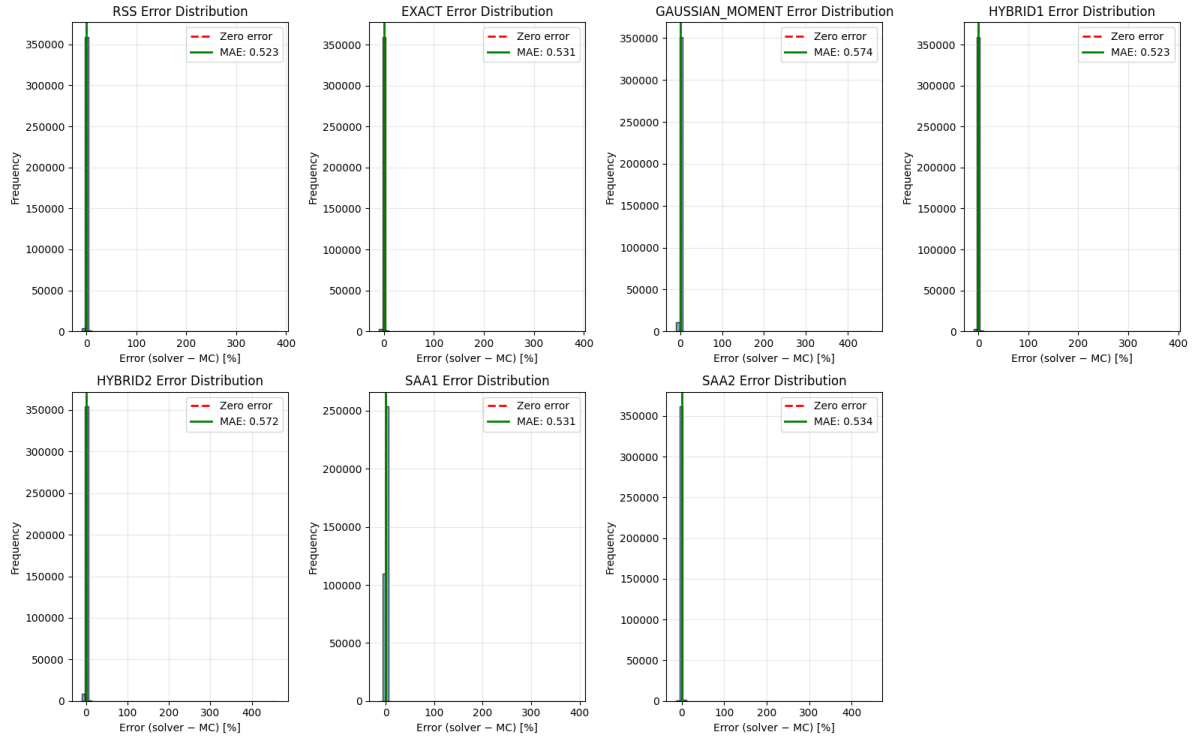


Figure D.19: Error distributions for each solver and the MC baseline for the high-biodiversity forest scene, with MAE highlighted in green.

Table D.12: Absolute errors between solver-predicted and Monte Carlo-derived uncertainty interval widths (geolocation uncertainty) for the high-biodiversity forest scene, in percentage points (pp).

Solver	n	MAE	RMSE	MEAN	MEDIAN	STD
RSS	363870	0.5234	5.6008	0.0142	-0.0313	5.6008
EXACT	363870	0.5308	5.5904	-0.0269	-0.0266	5.5903
GAUSSIAN_MOMENT	363040	0.5738	6.7068	-0.0074	-0.0457	6.7068
HYBRID1	363868	0.5228	5.5686	-0.0232	-0.0352	5.5686
HYBRID2	363661	0.5717	6.7177	0.0138	-0.0342	6.7177
SAA1	363870	0.5066	5.9218	0.0308	-0.0242	5.9217
SAA2	363870	0.5492	7.3421	0.2254	0.0162	7.3386

Table D.13: Absolute errors between solver-predicted and Monte Carlo-derived lower bound (q16) for the high-biodiversity forest scene, in percentage points (pp).

Solver	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.3207	0.8675	0.1319	0.0271	0.8574
EXACT	0.4044	0.9853	0.1648	0.0268	0.9714
GAUSSIAN_MOMENT	0.3764	1.0825	0.1296	0.0389	1.0747
HYBRID1	0.3718	0.9722	0.1645	0.0279	0.9582
HYBRID2	0.3563	1.0766	0.1213	0.0283	1.0697
SAA1	0.2852	0.8379	0.1235	0.0217	0.8288
SAA2	0.2487	2.2434	-0.0535	-0.0056	2.2428

Table D.14: Absolute errors between solver-predicted and Monte Carlo-derived upper bound (q84) for the high-biodiversity forest scene, in percentage points (pp).

Solver	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.4807	6.0463	0.1461	-0.0162	6.0445
EXACT	0.5607	6.0644	0.1379	-0.0116	6.0628
GAUSSIAN_MOMENT	0.5303	5.9400	0.1222	-0.0288	5.9387
HYBRID1	0.5247	6.0363	0.1412	-0.0166	6.0347
HYBRID2	0.5083	5.9478	0.1351	-0.0176	5.9463
SAA1	0.4521	6.3467	0.1543	-0.0078	6.3449
SAA2	0.3464	5.1793	0.1719	0.0110	5.1765

Table D.15: Absolute errors of predicted uncertainty interval widths relative to Monte Carlo reference values, stratified by radiometric combination case (n_+, n_-), for the coastal rainforest scene (percentage points).

Case	Solver	MAE	RMSE	MEAN
(4,0)	RSS	0.634	4.244	0.433
	EXACT	0.570	4.917	0.338
	SAA1	0.511	3.272	0.081
	SAA2	0.502	3.139	0.001
	GAUSSIAN_MOMENT	0.587	4.905	0.382
	HYBRID1	0.585	4.920	0.369
	HYBRID2	0.603	4.963	0.401
(3,1)	RSS	0.463	4.452	0.044
	EXACT	0.471	4.344	-0.044
	SAA1	0.458	3.729	-0.071
	SAA2	0.489	4.768	0.202
	GAUSSIAN_MOMENT	0.501	5.296	-0.017
	HYBRID1	0.461	4.344	-0.027
	HYBRID2	0.499	5.274	0.012
(2,2)	RSS	0.601	6.817	-0.029
	EXACT	0.614	6.765	-0.037
	SAA1	0.571	7.387	0.114
	SAA2	0.644	9.214	0.328
	GAUSSIAN_MOMENT	0.680	8.213	-0.015
	HYBRID1	0.605	6.765	-0.048
	HYBRID2	0.675	8.234	-0.002
(1,3)	RSS	0.377	3.138	-0.017
	EXACT	0.387	3.102	-0.092
	SAA1	0.412	3.734	-0.095
	SAA2	0.402	4.312	0.068
	GAUSSIAN_MOMENT	0.397	3.545	-0.088
	HYBRID1	0.379	3.102	-0.068
	HYBRID2	0.394	3.540	-0.052
(0,4)	RSS	0.365	0.579	0.179
	EXACT	0.323	0.503	0.108
	SAA1	0.318	0.519	-0.043
	SAA2	0.318	0.523	-0.072
	GAUSSIAN_MOMENT	0.330	0.525	0.139
	HYBRID1	0.344	0.527	0.150
	HYBRID2	0.351	0.544	0.165

D.2.5. Rainforest coast (NIR, AVIRIS)

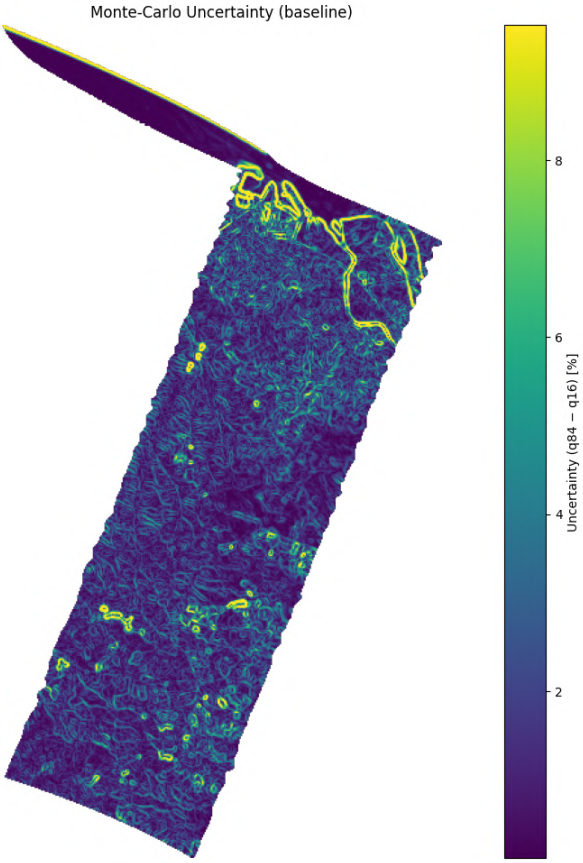


Figure D.20: Monte Carlo reference uncertainty map of the coastal rainforest scene (India) (run_mc = 10,000, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

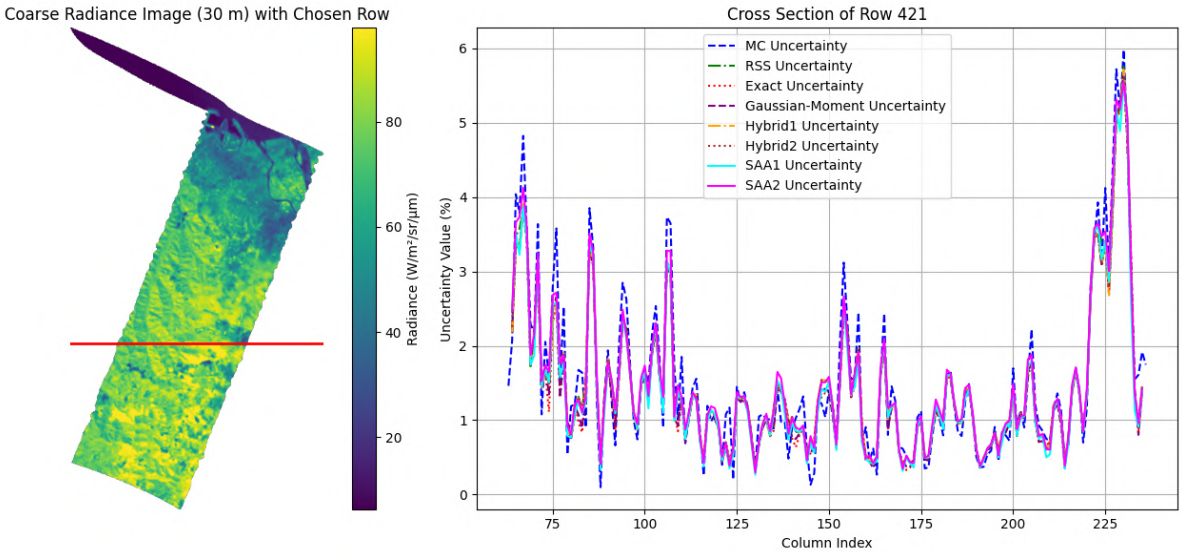


Figure D.21: Cross section of estimated geolocation uncertainty along the profile for each solver (coastal rainforest; run_mc = 10,000, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

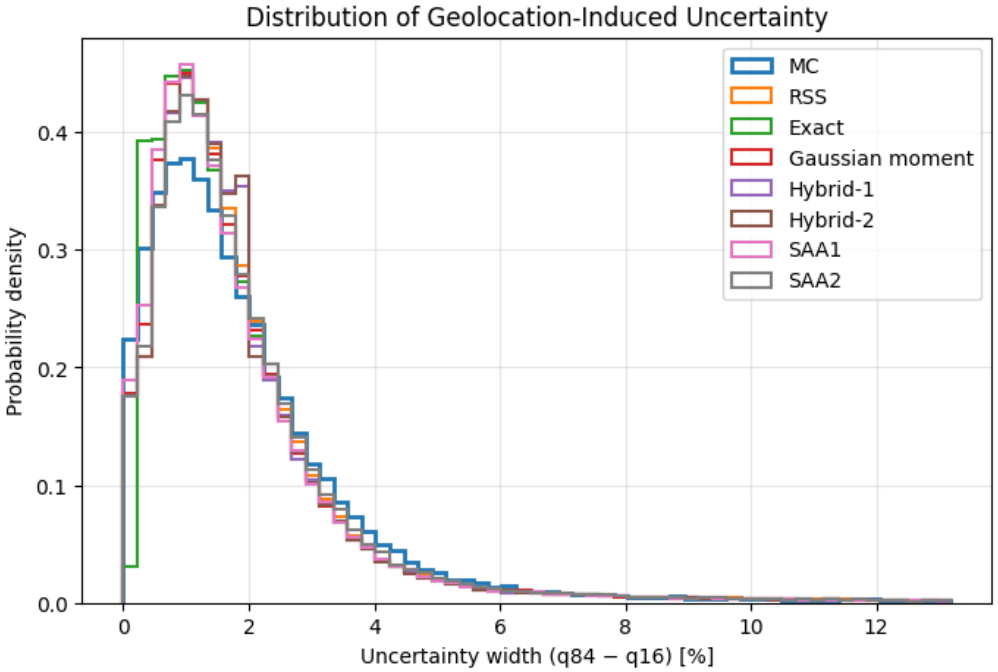


Figure D.22: Geolocation uncertainty distribution per solver for the coastal rainforest scene (run_mc = 10,000, $\sigma_{\text{gel}} = 0.15$ px).

Table D.16: Absolute errors between solver-predicted and Monte Carlo-derived uncertainty interval widths (geolocation uncertainty) for the coastal rainforest scene, in percentage points (pp).

Solver	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.3462	0.5538	-0.1062	-0.0708	0.5435
EXACT	0.3442	0.5444	-0.1652	-0.1155	0.5187
GAUSSIAN_MOMENT	0.3559	0.5663	-0.1561	-0.1092	0.5444
HYBRID1	0.3485	0.5506	-0.1344	-0.0850	0.5339
HYBRID2	0.3577	0.5679	-0.1270	-0.0821	0.5535
SAA1	0.3256	0.5190	-0.1254	-0.0836	0.5037
SAA2	0.3208	0.5419	-0.0119	-0.0102	0.5418

Table D.17: Absolute errors between solver-predicted and Monte Carlo-derived lower bound (q16) for the coastal rainforest scene, in percentage points (pp).

Solver	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.3036	0.4898	0.0900	0.0391	0.4814
EXACT	0.3865	0.5980	0.1297	0.0713	0.5838
GAUSSIAN_MOMENT	0.3538	0.5329	0.1193	0.0738	0.5194
HYBRID1	0.3454	0.5719	0.1158	0.0423	0.5601
HYBRID2	0.3243	0.5112	0.1052	0.0453	0.5002
SAA1	0.2703	0.4512	0.0921	0.0345	0.4417
SAA2	0.1869	0.3033	0.0116	0.0039	0.3031

Table D.18: Absolute errors between solver-predicted and Monte Carlo-derived upper bound (q84) for the coastal rainforest scene, in percentage points (pp).

Solver	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.3091	0.5380	-0.0162	-0.0436	0.5377
EXACT	0.3877	0.6390	-0.0355	-0.0776	0.6380
GAUSSIAN_MOMENT	0.3635	0.6107	-0.0368	-0.0795	0.6096
HYBRID1	0.3466	0.6147	-0.0186	-0.0455	0.6144
HYBRID2	0.3337	0.5913	-0.0218	-0.0491	0.5909
SAA1	0.2697	0.4908	-0.0334	-0.0393	0.4896
SAA2	0.1862	0.3315	-0.0003	-0.0047	0.3315

Table D.19: Absolute errors of predicted uncertainty interval widths relative to Monte Carlo reference values, stratified by radiometric combination case (n_+, n_-), for the coastal rainforest scene (percentage points).

Case	Method	MAE	RMSE	MEAN
(4,0)	RSS	0.433	0.725	0.317
	EXACT	0.360	0.583	0.222
	SAA1	0.333	0.532	0.056
	SAA2	0.346	0.562	0.110
(3,1)	RSS	0.328	0.522	-0.000
	EXACT	0.318	0.486	-0.105
	SAA1	0.353	0.537	-0.147
	SAA2	0.343	0.553	0.022
(2,2)	RSS	0.373	0.599	-0.226
	EXACT	0.383	0.610	-0.254
	SAA1	0.360	0.582	-0.201
	SAA2	0.341	0.575	-0.115
(1,3)	RSS	0.283	0.427	-0.033
	EXACT	0.281	0.421	-0.121
	SAA1	0.316	0.481	-0.165
	SAA2	0.292	0.446	-0.053
(0,4)	RSS	0.306	0.433	0.192
	EXACT	0.261	0.362	0.125
	SAA1	0.250	0.362	-0.032
	SAA2	0.261	0.373	0.053

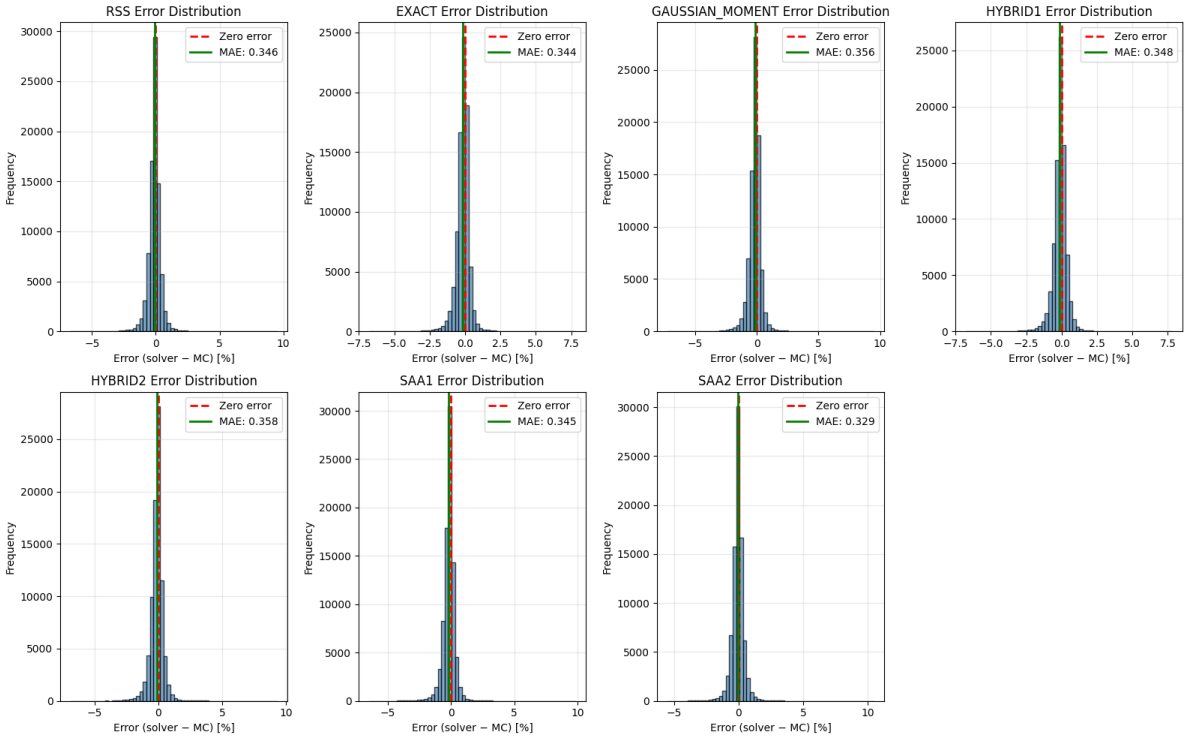


Figure D.23: Error distributions for each solver and the MC baseline for the coastal rainforest scene, with MAE highlighted in green.

D.2.6. Mangroves in India (NIR, AVIRIS)

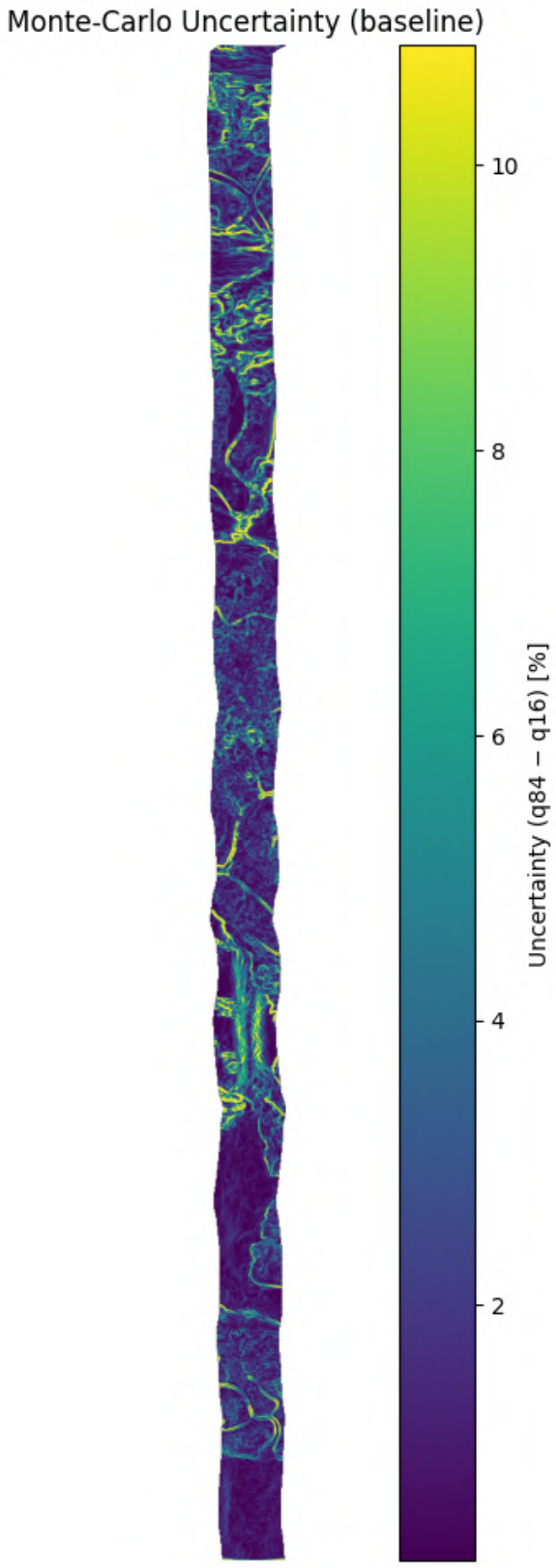


Figure D.24: Monte Carlo reference uncertainty map of the mangroves scene (India) (run_mc = 10,000, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

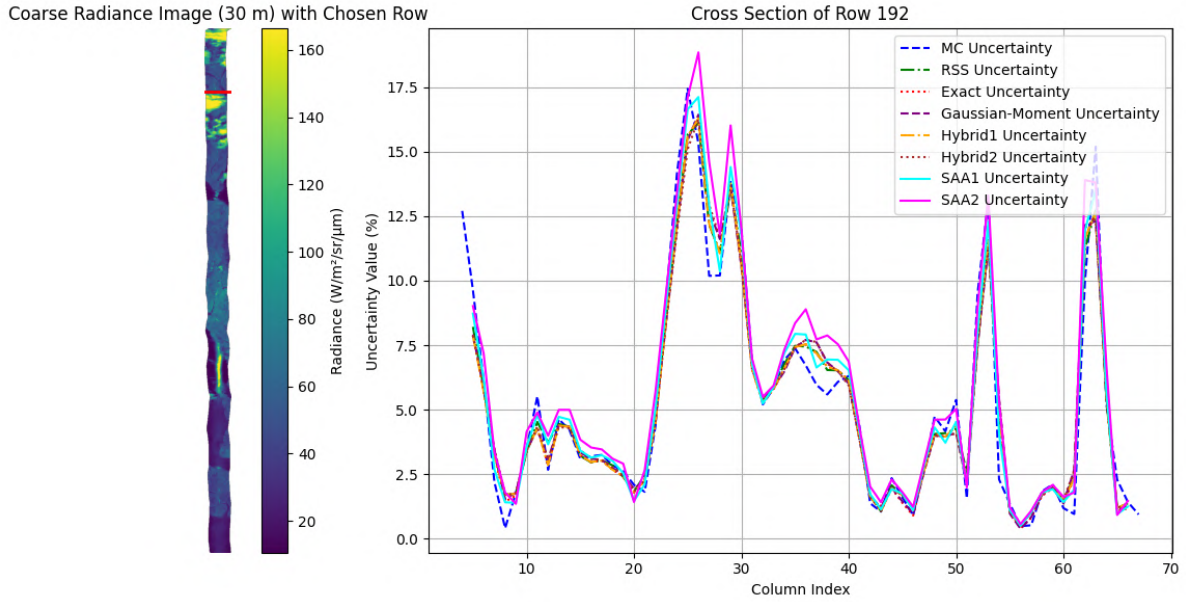


Figure D.25: Cross section of estimated geolocation uncertainty along the profile for each solver (mangroves; run_mc = 10,000, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

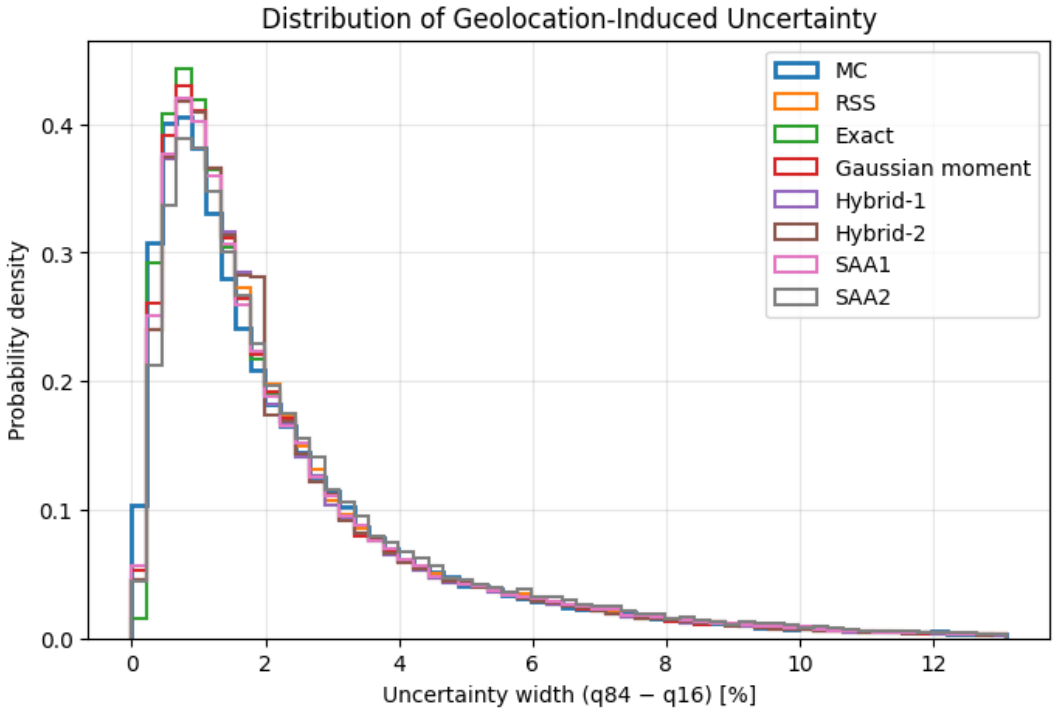


Figure D.26: Geolocation uncertainty distribution per solver for the mangroves scene (run_mc = 10,000, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

Table D.20: Absolute errors between solver-predicted and Monte Carlo-derived uncertainty interval widths (geolocation uncertainty) for the mangroves scene, in percentage points (pp).

Solver	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.3770	0.6216	-0.0332	-0.0067	0.6207
EXACT	0.3663	0.6003	-0.0943	-0.0327	0.5929

Solver	MAE	RMSE	MEAN	MEDIAN	STD
GAUSSIAN_MOMENT	0.3894	0.6432	-0.0779	-0.0308	0.6385
HYBRID1	0.3727	0.6054	-0.0711	-0.0187	0.6012
HYBRID2	0.3911	0.6442	-0.0600	-0.0160	0.6414
SAA1	0.3797	0.6202	-0.1392	-0.0625	0.6044
SAA2	0.3902	0.6364	-0.1698	-0.0794	0.6133

Table D.21: Absolute errors between solver-predicted and Monte Carlo-derived lower bound (q16) for the mangroves scene, in percentage points (pp).

Solver	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.3355	0.5401	0.0857	0.0251	0.5332
EXACT	0.4171	0.6528	0.1264	0.0481	0.6405
GAUSSIAN_MOMENT	0.3804	0.5912	0.1086	0.0474	0.5811
HYBRID1	0.3888	0.6364	0.1170	0.0296	0.6255
HYBRID2	0.3603	0.5774	0.1009	0.0302	0.5686
SAA1	0.3110	0.5176	0.1221	0.0401	0.5030
SAA2	0.2251	0.3689	0.0794	0.0365	0.3602

Table D.22: Absolute errors between solver-predicted and Monte Carlo-derived upper bound (q84) for the mangroves scene, in percentage points (pp).

Solver	MAE	RMSE	MEAN	MEDIAN	STD
RSS	0.3608	0.6350	0.0525	-0.0149	0.6328
EXACT	0.4364	0.7332	0.0321	-0.0373	0.7325
GAUSSIAN_MOMENT	0.4093	0.7000	0.0307	-0.0420	0.6994
HYBRID1	0.4075	0.7191	0.0459	-0.0189	0.7176
HYBRID2	0.3889	0.6890	0.0408	-0.0216	0.6878
SAA1	0.3178	0.5505	-0.0171	-0.0363	0.5503
SAA2	0.2238	0.3683	-0.0903	-0.0474	0.3570

Table D.23: Absolute errors of predicted uncertainty interval widths relative to Monte Carlo reference values, stratified by radiometric combination case (n_+, n_-), for the mangroves scene (percentage points).

Case	Method	MAE	RMSE	MEAN
(4,0)	RSS	0.609	0.974	0.410
	EXACT	0.513	0.807	0.283
	SAA1	0.474	0.738	-0.005
	SAA2	0.537	0.849	0.238
	GAUSSIAN_MOMENT	0.548	0.872	0.363
	HYBRID1	0.544	0.829	0.336
	HYBRID2	0.569	0.885	0.381
(3,1)	RSS	0.413	0.696	0.094
	EXACT	0.376	0.618	-0.033
	SAA1	0.417	0.673	-0.149
	SAA2	0.417	0.673	-0.148
	GAUSSIAN_MOMENT	0.398	0.665	0.009
	HYBRID1	0.387	0.625	0.012
	HYBRID2	0.405	0.668	0.043
(2,2)	RSS	0.371	0.611	-0.120

Case	Method	MAE	RMSE	MEAN
	EXACT	0.376	0.618	-0.149
	SAA1	0.380	0.624	-0.150
	SAA2	0.396	0.647	-0.220
	GAUSSIAN_MOMENT	0.401	0.663	-0.146
	HYBRID1	0.377	0.619	-0.141
	HYBRID2	0.399	0.662	-0.137
(1,3)	RSS	0.327	0.514	0.027
	EXACT	0.309	0.498	-0.068
	SAA1	0.343	0.556	-0.131
	SAA2	0.337	0.546	-0.091
	GAUSSIAN_MOMENT	0.324	0.521	-0.047
	HYBRID1	0.321	0.507	-0.023
	HYBRID2	0.331	0.525	-0.013
(0,4)	RSS	0.363	0.513	0.240
	EXACT	0.311	0.436	0.164
	SAA1	0.288	0.412	-0.027
	SAA2	0.296	0.436	-0.134
	GAUSSIAN_MOMENT	0.337	0.473	0.219
	HYBRID1	0.350	0.478	0.225
	HYBRID2	0.355	0.490	0.234

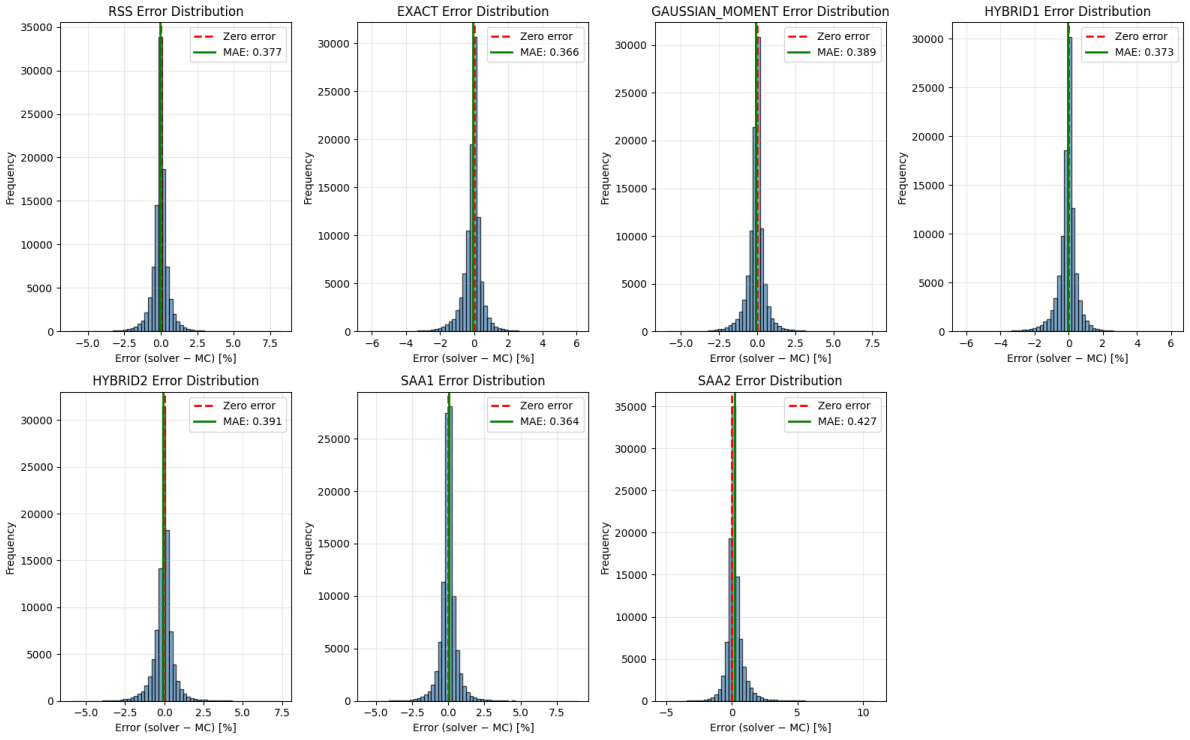


Figure D.27: Error distributions for each solver and the MC baseline for the mangroves scene, with MAE highlighted in green.

D.2.7. Oklahoma agricultural fields (NIR, Sentinel-2)

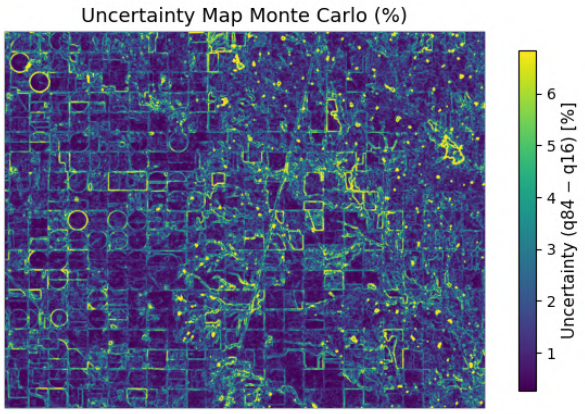


Figure D.28: Monte Carlo reference uncertainty map of the Oklahoma scene (USA) (run_mc = 10,000, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

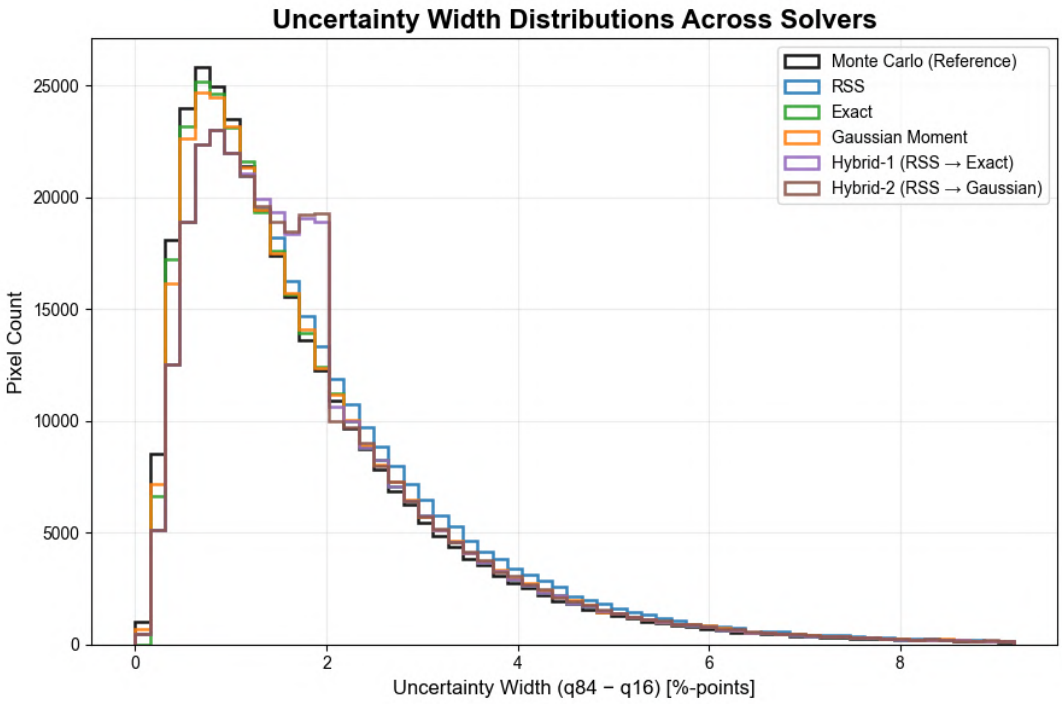


Figure D.29: Geolocation uncertainty distribution per solver for the Oklahoma scene (run_mc = 10,000, $\sigma_{\text{gel}} = 0.15\,\text{px}$).

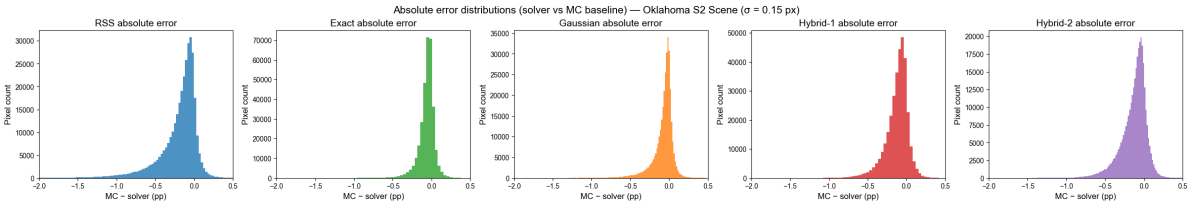


Figure D.30: Absolute error distributions for each solver and the MC baseline for the Oklahoma scene ($\sigma_{\text{gel}} = 0.15\,\text{px}$).

Table D.24: Validation metrics for each solver vs. the Monte Carlo reference for the Oklahoma scene (USA) ($\sigma_{gel} = 0.15$ px), in percentage points (pp).

Solver	r16 (pp)	Median (pp)	r84 (pp)	width$_{68}$ (pp)	MAE (pp)	RMSE (pp)
RSS	-0.3954	-0.1278	-0.0128	0.3826	0.2284	0.3897
EXACT	-0.1366	-0.0474	0.0129	0.1495	0.0863	0.1477
GAUSSIAN_MOMENT	-0.1680	-0.0400	0.0185	0.1865	0.1028	0.1881
HYBRID1	-0.2440	-0.0937	-0.0024	0.2415	0.1402	0.2047
HYBRID2	-0.2767	-0.0935	0.0063	0.2830	0.1555	0.2330

D.2.8. Grossetto coast (NIR, Sentinel-2)

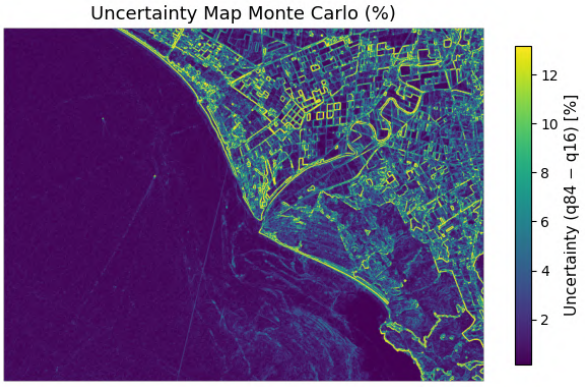


Figure D.31: Monte Carlo reference uncertainty map of the Grossetto scene (Italy) (run_mc = 10,000, $\sigma_{gel} = 0.15$ px).

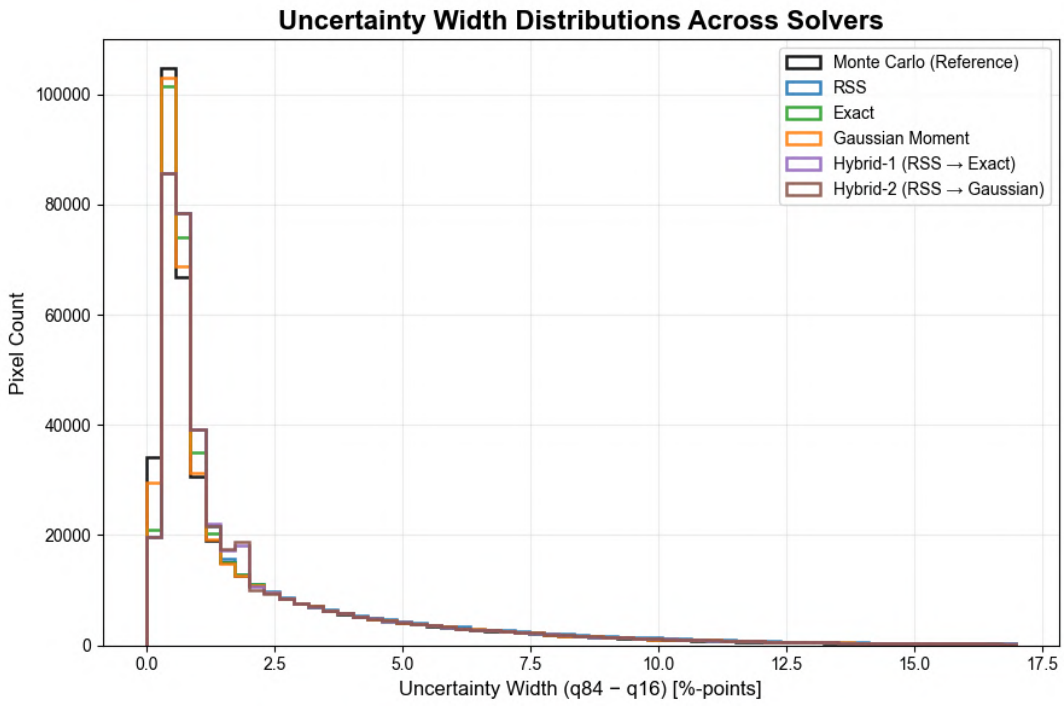


Figure D.32: Geolocation uncertainty distribution per solver for the Grossetto scene (run_mc = 10,000, $\sigma_{gel} = 0.15$ px).

Table D.25: Validation metrics for each solver vs. the Monte Carlo reference for the Grossetto scene (Italy) ($\sigma_{\text{gel}} = 0.15\,\text{px}$).

Solver	q16	|Median|	q84	Width_{68}	MAE	RMSE
RSS	0.4064	0.1101	0.0201	0.3863	0.2681	0.5231
EXACT	0.1757	0.0598	0.0061	0.1696	0.1254	0.3091
GAUSSIAN_MOMENT	0.1575	0.0226	0.0138	0.1713	0.1155	0.2652
HYBRID1	0.2578	0.0965	0.0156	0.2422	0.1606	0.3267
HYBRID2	0.2517	0.0891	0.0066	0.2451	0.1666	0.2882

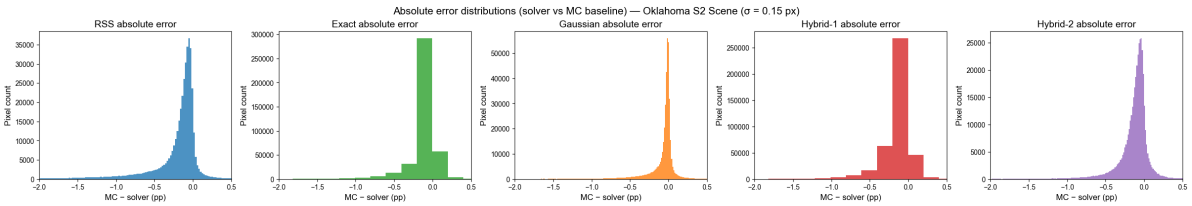


Figure D.33: Absolute error distributions for each solver and the MC baseline for the Grossetto scene ($\sigma_{\text{gel}} = 0.15\,\text{px}$).

E

Propagation to NDVI

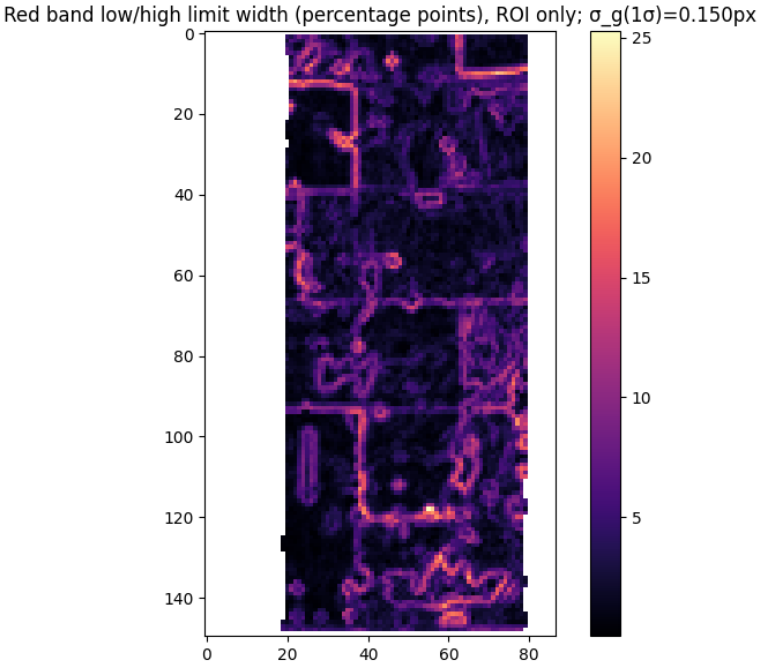


Figure E.1: Uncertainty interval width (geolocation uncertainty) using the RSS solver for band 59 (red, 673 nm; $\sigma_{\text{gel}} = 0.15$ px).

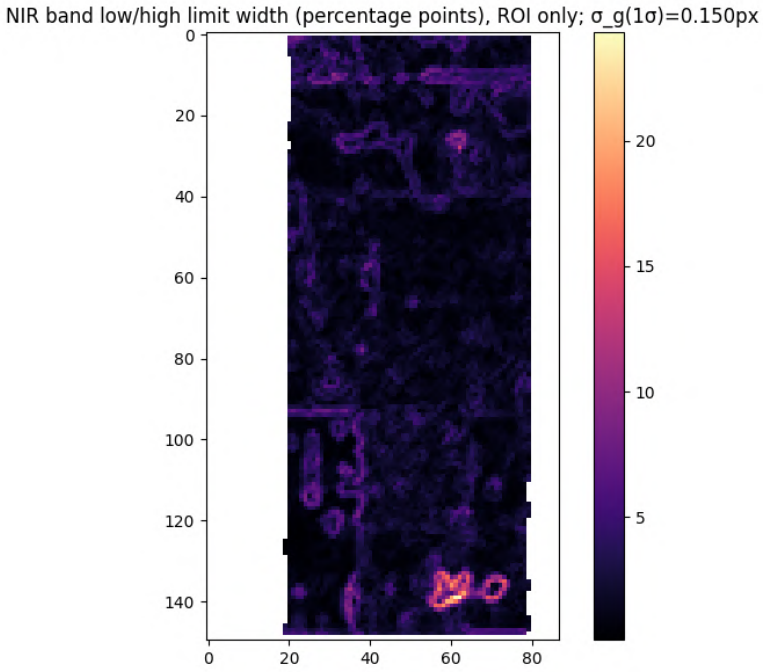


Figure E.2: Uncertainty interval width (geolocation uncertainty) using the RSS solver for band 90 (near-infrared, 827 nm; $\sigma_{\text{gel}} = 0.15$ px).

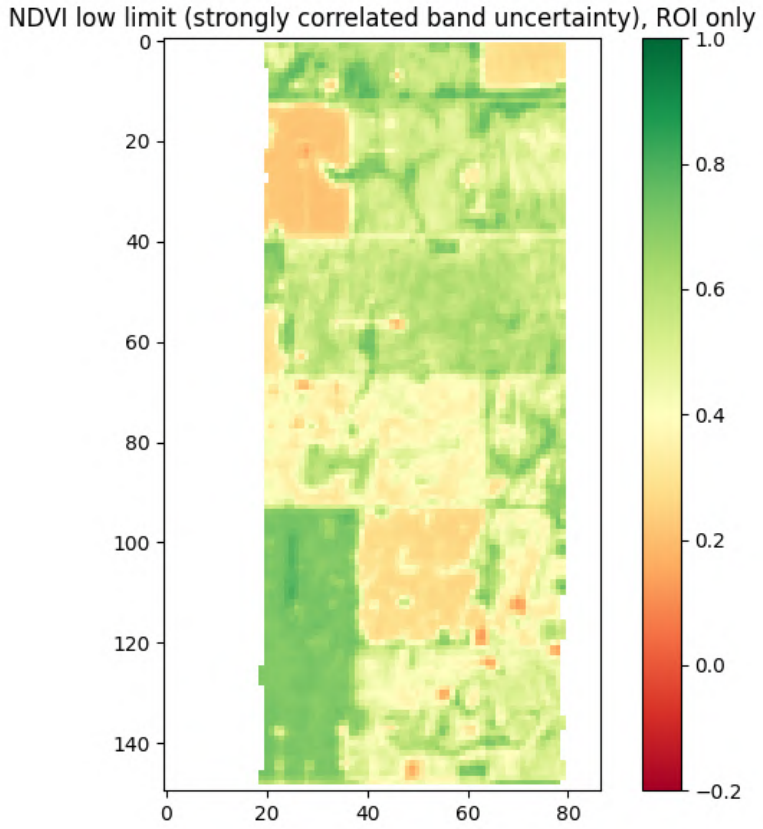


Figure E.3: Propagated NDVI uncertainty lower bound using RSS and the red and NIR band uncertainties ($\sigma_{\text{gel}} = 0.15$ px).

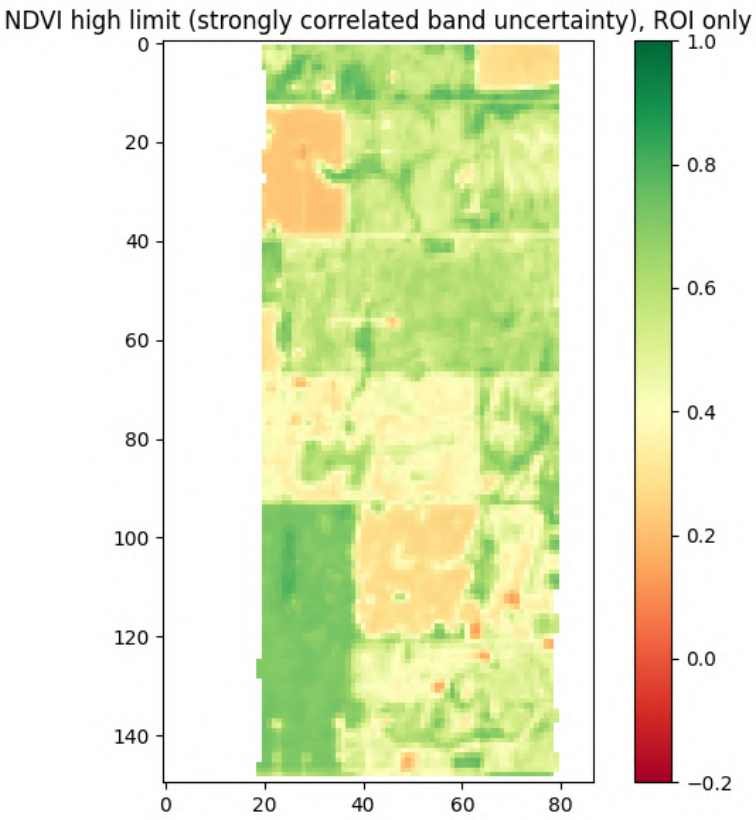


Figure E.4: Propagated NDVI uncertainty upper bound using RSS and the red and NIR band uncertainties ($\sigma_{\text{gel}} = 0.15\,\text{px}$).

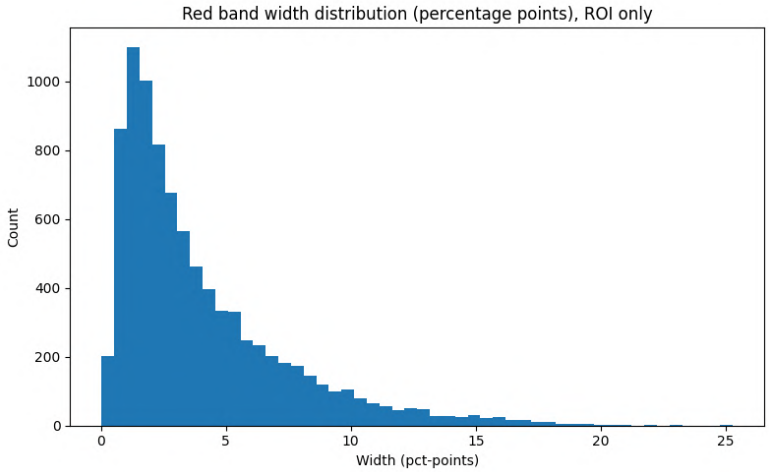


Figure E.5: Red-band uncertainty distribution for the ROI using the RSS solver ($\sigma_{\text{gel}} = 0.15\,\text{px}$).

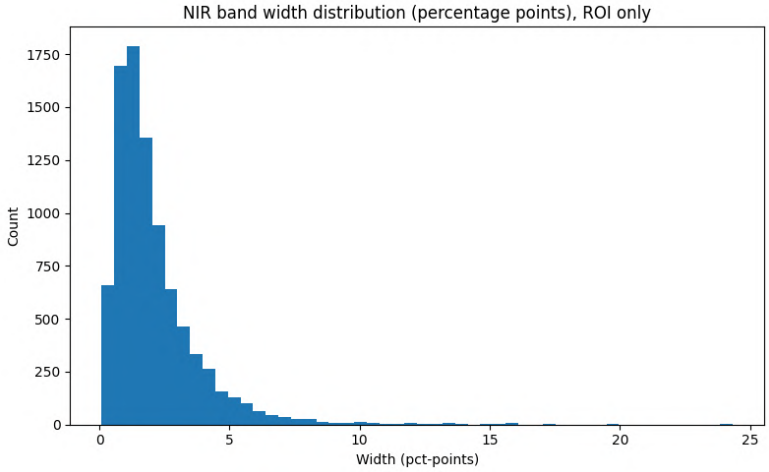


Figure E.6: NIR-band uncertainty distribution for the ROI using the RSS solver ($\sigma_{\text{gel}} = 0.15\,\text{px}$).

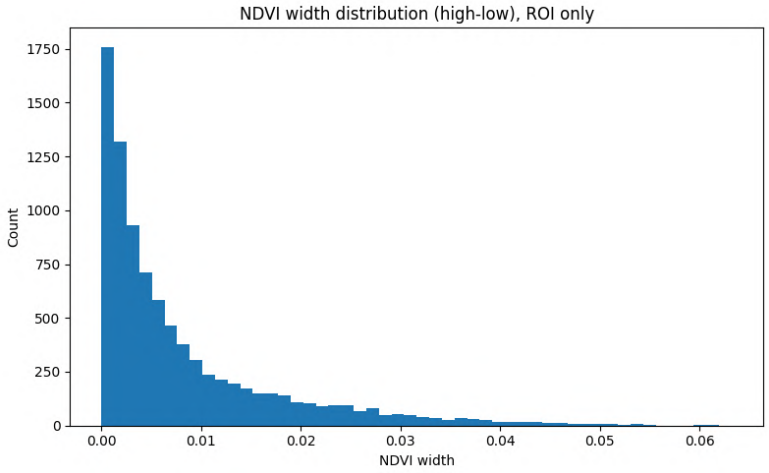


Figure E.7: Propagated NDVI uncertainty distribution for the ROI using the RSS solver ($\sigma_{\text{gel}} = 0.15\,\text{px}$).