POLYTECHNIC UNIVERSITY OF TURIN

Master’s Thesis in Environmental and Land Engineering

Department of Environmental and Land Engineering

Evaluation of Low-Cost GNSS for Tropospheric Delay Estimation: Physical Consistency, Processing Sensitivity, and Spatial Mapping

Supervisors:

Prof. Paolo Dabove

Dr. Milad Bagheri

Candidate:

Zahra Chaam (s310048)

Academic year 2025-2026

To my homeland,
wounded, yet unbroken,
burdened, yet luminous with memory and promise.
May it find healing, freedom, and the glory that has
always belonged to it.

To my family,
the roots that kept me grounded,
who never allowed me to surrender to doubt.

And to the love of my life,
who never stopped believing in me
when the path felt uncertain and the nights felt
long.

Table of Contents

ABSTRACT 6

CHAPTER 1: INTRODUCTION 7

- 1.1 Background 7
- 1.2 Problem Statement 8
- 1.3 Objectives 9
- 1.4 Research Questions 9
- 1.5 Thesis Structure 10

CHAPTER 2: THEORETICAL BACKGROUND 11

- 2.1 GNSS Signal Propagation and Atmospheric Effects 11
- 2.2 GNSS Observation Equations 11
  - 2.2.1 Fundamental GNSS Observables 11
  - 2.2.2 Ionosphere-Free Carrier-Phase Observation 11
- 2.3 Slant Tropospheric Delay and Zenith Representation 12
  - 2.3.1 Slant Delay Concept 12
  - 2.3.2 Mapping from Zenith to Slant 12
- 2.4 Zenith Hydrostatic Delay (ZHD) 13
  - 2.4.1 Physical Origin 13
  - 2.4.2 Analytical Expression 13
- 2.5 Zenith Wet Delay (ZWD) 13
  - 2.5.1 Physical Origin 13
  - 2.5.2 Integral Formulation 13
- 2.6 Mapping Functions 14
  - 2.6.1 Purpose of Mapping Functions 14
  - 2.6.2 Common Mapping Function Models 14
  - 2.6.3 Implementation in This Thesis 14
- 2.7 Precise Point Positioning (PPP) 14
  - 2.7.1 Conceptual Framework 14
  - 2.7.2 PPP Observation Model 14
  - 2.7.3 Troposphere Estimation in PPP 15
- 2.8 Meteorological Quantities Relevant to Tropospheric Delay 15
  - 2.8.1 Temperature and Dew Point 15

2.8.2 Surface Pressure 15
2.8.3 Precipitation 15
2.9 Position Stability and Its Relation to Tropospheric Estimation 15
CHAPTER 3: STUDY AREA AND GNSS STATIONS 17
3.1 Rationale for Station and Area Selection 17
3.2 Overview of the GNSS Station Set 17
3.3 Environmental Classification of Stations 18
3.3.1 Coastal Stations 18
3.3.2 Inland Lowland Stations 19
3.3.3 Plateau Stations 19
3.3.4 Mountain and Complex Terrain Stations 19
3.4 The Florence Study Area: RUF01 and IGMI00ITA 20
3.4.1 Geographic Context 20
3.4.2 RUF01 20
3.4.3 IGMI00ITA (Reference Station) 20
3.4.4 Relative Geometry and Height Difference 20
3.5 GNSS Networks and Data Sources 21
3.5.1 EUREF Permanent GNSS Network 21
3.5.2 Centipede and RENAG/RESIF Networks 21
3.6 Representativeness of the Station Set 21
CHAPTER 4: DATA AND METHODOLOGY 22
4.1 Chapter purpose and methodological logic 22
4.2 Data sources and file types 22
4.2.1 GNSS stations and observation data 22
4.2.2 Precise products and external reference products 23
4.2.3 ERA5 meteorological data and units 23
4.3 Tropospheric parameter definitions and physical equations 23
4.4 Multi-station case study workflow (CSRS-PPP + ERA5) 24
4.4.1 CSRS-PPP tropospheric outputs and time handling 24
4.4.2 Hourly aggregation 24
4.4.3 Statistical comparisons used in the case study 24
4.5 RUF stability workflow (RTKLIB PPP + EUREF + ERA5) 25
4.5.1 RTKLIB PPP processing 25

4.5.2 Dual precise-product strategies: CDDIS vs GFZ....25
4.5.3 ZTD parsing from RTKLIB outputs....26
4.6 Quality control and hourly aggregation....26
4.6.1 Basic physical and statistical QC....26
4.6.2 Robust outlier handling: rolling median + MAD....26
4.6.3 Forced hourly aggregation at :30 UTC....27
4.7 Intercomparison metrics....27
4.8 ZTD–ERA5 relationship modelling....28
4.9 Positioning stability analysis....28
4.9.1 Robust (typical behaviour) positioning statistics....29
4.10 France spatial ZTD mapping workflow....30
4.10.1 Batch PPP and epoch-level ZTD extraction....30
4.10.2 QC and aggregation for France (C1–C4)....30
4.10.3 Station-level filtering and definition of the interpolation-ready dataset....30
4.10.4 Vertical trend modelling and detrending....31
4.10.5 Residual kriging in projected coordinates....31
4.10.6 DEM preprocessing and final ZTD reconstruction....32
4.10.7 Cross-validation framework and comparative interpolation methodology....32
CHAPTER 5: Multi-Station Tropospheric Case Study Results under Dry and Rainy Regimes....34
5.1 Chapter Scope and Analytical Context....34
5.2 Urban lowland basin: RUF01 (Florence region, Italy)....36
5.2.1 Dry summer regime (15 July 2024)....36
5.2.2 Rainy synoptic regime (3 March 2024)....37
5.3 High mountain environment: SCV1 (Alpine region, ~2329 m)....38
5.3.1 Dry summer regime (15 July 2024)....38
5.3.2 Rainy regime (3 March 2024)....38
5.4 Plateau environment: BTCHZ (Swiss plateau)....39
5.4.1 Dry summer regime (15 July 2024)....39
5.4.2 Rainy regime (3 March 2024)....39
5.5 Inland lowland environment: STAVE (Belgium)....39
5.5.1 Dry summer regime (15 July 2024)....40
5.5.2 Rainy synoptic regime (3 March 2024)....40
5.6 Coastal environments: SOPH (Mediterranean) and PYLA (Atlantic)....40

5.6.1 SOPH (Mediterranean coast) 40
5.6.2 PYLA (Atlantic coast) 41
5.7 Mountain and complex terrain: AIGL (Lozère mountains, France, ~1619 m) 42
5.7.1 Dry summer regime (15 July 2024) 42
5.7.2 Rainy synoptic regime (3 March 2024) 42
5.8 Cross-station summary table 43
5.9 Cross-station synthesis 44
5.10 Key findings from the case study 44
CHAPTER 6: Long-Term Stability Analysis of the Low-Cost GNSS Station RUF01 45
6.1 Chapter Scope and Analytical Objectives 45
6.2 Data Set and Processing Overview 45
6.3 ZTD Time Series Generation and Quality Control 46
6.4 Intercomparison of ZTD Time Series 47
6.4.1 Comparison Against the EUREF Reference 47
6.4.2 Internal Consistency and Processing-Strategy Sensitivity 48
6.5 Meteorological Consistency of Long-Term ZTD Series 51
6.6 PPP Positioning Stability Analysis 55
6.6.1 Workflow And Metrics 55
6.6.2 ENU Residual Behaviour 56
6.6.3 Implications for Tropospheric Estimation 57
6.7 Synthesis and answers to the research questions 57
CHAPTER 7: FRANCE-WIDE SPATIAL RECONSTRUCTION OF ZENITH TOTAL DELAY (MARCH 2024) 59
7.1 Introduction 59
7.2 Station Dataset and Filtering 59
7.3 Elevation Dependence of Monthly Mean ZTD 60
7.4 Station-Level Results and Residual Analysis 61
7.5 Residual Kriging 63
7.6 DEM Integration and Final ZTD Field 64
7.7 Interpretation and Implications 66
7.8 Cross-Validation and Comparative Interpolation Performance 66
7.8.1 IDW Validation 67
7.8.2 Universal Kriging Validation 67
7.8.3 Regression–Kriging Validation 67

7.8.4 Interpretation of Validation Results 68
7.9 Chapter Conclusions and Scientific Implications 68
CHAPTER 8: SYNTHESIS, LIMITATIONS, AND FUTURE RESEARCH 70
8.1 Integrated Synthesis of Findings 70
8.2 Scientific Implications 70
8.3 Methodological Limitations 71
8.4 Directions for Future Research 71
8.5 Concluding Perspective 72
REFERENCES 73

ABSTRACT

Tropospheric delay is a major error source in high-precision Global Navigation Satellite System (GNSS) positioning. When estimated within a Precise Point Positioning (PPP) framework, however, it constitutes a physically meaningful atmospheric observable directly linked to refractivity in the lower atmosphere. The zenith total delay (ZTD) integrates both hydrostatic and wet refractivity components and can therefore be exploited for meteorological monitoring and atmospheric research.

This thesis evaluates the physical consistency, temporal stability, and spatial representativeness of GNSS-derived tropospheric delays, with particular emphasis on low-cost GNSS receivers. The study progresses from short-term physical validation to long-term stability assessment and finally to national-scale spatial reconstruction.

First, GNSS-derived hydrostatic (TRODRY) and wet (TROWET) delay components are compared with ERA5 reanalysis variables at seven European GNSS stations spanning coastal, lowland, plateau, and mountainous environments. Two contrasting atmospheric regimes, a dry summer day and a rainy winter day, are analysed. TRODRY exhibits strong near-linear dependence on surface pressure (correlation coefficients frequently > 0.9; slopes $\approx 1\text{–}2.5$ mm hPa^{-1}), consistent with hydrostatic theory. TROWET shows strong correlation with dew point temperature in coastal and lowland regions, while mountainous stations exhibit more variable behaviour due to vertical humidity stratification and complex terrain effects.

Second, the stability of a low-cost GNSS station (RUF01, Florence region) is evaluated over a two-month period and compared against a nearby geodetic EUREF station (IGMI00ITA) under two independent precise product strategies. After rigorous quality control and hourly aggregation, RUF01 demonstrates a stable and physically coherent ZTD signal, with a mean bias of approximately -39 mm, RMS difference of ~48 mm, and correlation of ~0.62 relative to the EUREF reference. Internal consistency between processing strategies is high, indicating that processing choices can influence systematic behaviour as strongly as receiver class. Position residual analysis further confirms centimetre-level typical 3D stability after robust filtering, supporting the suitability of low-cost PPP solutions for atmospheric applications.

Third, a France-wide batch PPP campaign using the Centipede low-cost GNSS network was conducted for March 2024, generating a dense national ZTD dataset. Elevation explains 99.32% of the spatial variance in monthly mean ZTD (slope $\approx$ -284 mm km^{-1}), confirming dominant hydrostatic control. After detrending, regression–kriging yields a cross-validated RMSE of 3.66 mm ($R^2 = 0.998$), substantially outperforming distance-based interpolation methods. The resulting 1 km ZTD field over metropolitan France demonstrates strong physical coherence and highlights the capability of dense low-cost GNSS networks to resolve regional atmospheric structure.

Overall, the results demonstrate that PPP-derived tropospheric delays from low-cost GNSS receivers are physically consistent, temporally stable under controlled processing, and capable of supporting robust national-scale atmospheric mapping when deterministic elevation effects are explicitly modelled.

CHAPTER 1: INTRODUCTION

1.1 Background

Global Navigation Satellite Systems (GNSS), including GPS, Galileo, GLONASS, and BeiDou, have become indispensable for positioning, navigation, and timing across scientific, engineering, and societal applications. While GNSS was originally developed for navigation and geodesy, it has increasingly been recognised as a powerful remote sensing tool for monitoring the Earth system [1, 2]. One of the most significant contributions of GNSS to Earth observation is its capability to sense the neutral atmosphere through the propagation delay experienced by radio signals.

As GNSS signals travel from satellites to ground-based receivers, they are delayed by refraction in both the ionosphere and the neutral atmosphere [1, 3]. While ionospheric delay is frequency dependent and can be largely eliminated using dual-frequency combinations, the tropospheric delay is non-dispersive at GNSS frequencies and must therefore be explicitly modelled or estimated. In precise positioning, tropospheric delay is traditionally treated as a nuisance parameter because it introduces centimetre- to decimetre-level range errors if not properly accounted for. From an atmospheric perspective, however, this delay represents a direct manifestation of the refractive properties of the lower atmosphere and therefore contains physically meaningful information [4].

The tropospheric delay is commonly expressed in terms of the zenith total delay (ZTD), which represents the integrated delay along the vertical direction above the receiver. ZTD can be decomposed into two physically distinct components: a hydrostatic (dry) component, which depends primarily on the total mass of dry air and is closely linked to surface pressure [5], and a wet component, which arises from atmospheric water vapour and exhibits strong spatial and temporal variability [4, 5]. This decomposition makes GNSS-derived tropospheric delays particularly attractive for meteorological and climatological applications, as they provide continuous, all-weather observations that complement conventional instruments such as radiosondes, microwave radiometers, and surface meteorological stations [4].

In parallel with these scientific developments, the rapid proliferation of low-cost GNSS receivers has fundamentally changed the landscape of GNSS observations [6]. Low-cost hardware, often deployed in dense community-based or research networks, offers the possibility of dramatically increasing the spatial density of tropospheric observations [7]. If tropospheric products derived from such receivers can be demonstrated to be stable and physically meaningful, they could significantly enhance the monitoring of atmospheric water vapour at regional and national scales. For such applications, positioning stability is not a secondary consideration, unstable coordinates and intermittent PPP failures can propagate into tropospheric estimates. Consequently, any rigorous evaluation of low-cost GNSS tropospheric sensing must treat position residual behaviour as an integral quality indicator alongside ZTD validation [6, 8].

By integrating physical validation, processing sensitivity analysis, positioning stability assessment, and national-scale spatial reconstruction, this thesis aims to provide a comprehensive evaluation of low-cost GNSS as an atmospheric observing system.

1.2 Problem Statement

Despite the well-established theoretical foundations of GNSS tropospheric sensing, several practical and scientific challenges remain unresolved, particularly when low-cost receivers and large-scale applications are considered.

First, the quality and stability of tropospheric delays estimated from low-cost GNSS receivers are not guaranteed a priori. Compared to geodetic-grade stations, low-cost installations often rely on simpler antennas, less rigid monuments, and non-ideal site conditions. These factors can increase susceptibility to multipath, signal obstruction, and environmental noise, potentially degrading both positioning and tropospheric parameter estimation [6].

Second, the relationship between GNSS-derived tropospheric parameters and conventional meteorological variables is not uniform across different environments. While the hydrostatic component of the delay is expected to closely follow surface pressure under all conditions, the wet component is influenced by complex atmospheric processes, including vertical humidity stratification, boundary-layer dynamics, and orographic effects. Consequently, the strength and linearity of relationships between wet delay and near-surface humidity indicators, such as dew point temperature, may vary substantially between coastal, lowland, plateau, and mountainous regions [4, 5].

Third, PPP-derived tropospheric products are inseparable from PPP positioning behaviour. PPP solutions may contain convergence transients, cycle-slip-driven degradations, and rare but large failures driven by poor satellite geometry or data quality. These effects tend to be amplified for low-cost receivers and can contaminate tropospheric estimates unless they are explicitly detected and treated. A robust stability assessment must therefore include a position residual analysis in a local ENU frame, separating typical behaviour from failure modes using transparent quality filters and robust statistical measures [8, 9].

Fourth, tropospheric delay estimates derived from PPP are inherently dependent on the processing strategy, including the choice of precise orbit and clock products, modelling assumptions, and filtering approaches. Without careful control, differences attributed to receiver performance may in fact reflect differences in processing [9]. A rigorous assessment therefore requires parallel processing strategies applied consistently to both low-cost and geodetic stations. This sensitivity is particularly relevant for low-cost receivers, where modelling errors can be amplified [2, 9].

Finally, while many studies focus on single-station validation, the broader scientific value of low-cost GNSS networks lies in their ability to resolve spatial patterns of atmospheric water vapour. Constructing regional-scale tropospheric delay maps from dense low-cost networks introduces additional challenges related to data volume, quality control, station stability, and spatial interpolation, which must be addressed systematically.

1.3 Objectives

The objectives of this thesis are organised into three complementary themes. The first objective is the physical validation of GNSS tropospheric delays, with the aim of quantifying and interpreting the relationships between GNSS-derived tropospheric parameters (ZTD, TRODRY, and TROWET) and meteorological variables from the ERA5 reanalysis at multiple European stations under contrasting atmospheric conditions.

The second objective is the assessment of the stability and reliability of tropospheric products derived from a low-cost GNSS station. This involves the implementation of a complete batch PPP processing workflow for the low-cost station RUF01 using RTKLIB demo5, the evaluation of the stability, bias, and temporal consistency of its ZTD estimates through comparison with the nearby geodetic EUREF station IGMI00ITA using multiple precise product strategies, and the quantification of PPP positioning stability through ENU residual analysis as an integral indicator of solution reliability.

The third objective is the spatial representation of tropospheric delay at the regional scale. This objective extends the analysis from individual stations to a national scale by constructing spatial ZTD maps over France using the dense Centipede low-cost GNSS network, thereby assessing the feasibility and limitations of regional tropospheric monitoring based on low-cost GNSS observations.

1.4 Research Questions

The thesis addresses the following research questions:

1) How robust is the relationship between the hydrostatic component of tropospheric delay and surface pressure across different European environments and synoptic regimes?
2) To what extent does the wet component of tropospheric delay reflect near-surface humidity indicators such as dew point temperature, and how does this relationship depend on station environment and topography?
3) What are the bias, RMS difference, and correlation between ZTD estimates from a low-cost GNSS station and an established EUREF reference station when processed using consistent PPP strategies?
4) How sensitive are GNSS-derived tropospheric delays to the choice of precise products and processing strategy compared with the influence of receiver class?
5) What is the positioning stability of the low-cost station relative to a geodetic reference under identical PPP strategies, and how does processing strategy affect ENU residual dispersion and outlier behaviour?
6) To what extent can dense low-cost GNSS networks resolve spatially coherent ZTD fields at the national scale, given station-level stability constraints and the elevation dependence of tropospheric delay?

1.5 Thesis Structure

The thesis is structured as follows: Chapter 2 presents the theoretical framework, including GNSS observation models, tropospheric delay decomposition, mapping functions, PPP methodology, and the physical basis linking tropospheric delays to meteorological variables. Chapter 3 describes the study areas and GNSS stations, including environmental classification and station characteristics.

Chapter 4 details the data sets and methodologies used in this study, covering CSRS-PPP outputs, ERA5 reanalysis data, batch PPP processing with RTKLIB, quality control procedures for tropospheric and positioning parameters, and the statistical analysis techniques applied. Chapter 5 presents the results of the multi-station case study under dry and rainy conditions, focusing on the physical relationships between tropospheric components and meteorological variables.

Chapter 6 analyses the long-term stability of the low-cost station RUF01, including multi-strategy PPP ZTD validation against EUREF, processing sensitivity, meteorological consistency, and PPP positioning stability through ENU residual analysis. Chapter 7 introduces the spatial ZTD mapping over France, describing the large-scale batch processing workflow, station filtering, interpolation methods such as inverse distance weighting and kriging, and uncertainty assessment, and motivates extensions based on digital elevation models. Finally, Chapter 8 synthesises the results, discusses limitations, and outlines directions for future research.

CHAPTER 2: THEORETICAL BACKGROUND

2.1 GNSS Signal Propagation and Atmospheric Effects

GNSS positioning is fundamentally based on the measurement of signal travel time between satellites and ground-based receivers. GNSS satellites transmit electromagnetic signals in the microwave band, which propagate through the Earth's atmosphere before reaching the receiver antenna. During this propagation, the signals experience delays caused by the refractive properties of the ionosphere and the neutral atmosphere [1, 2].

The ionospheric delay is dispersive at GNSS frequencies, meaning that its magnitude depends on signal frequency. This property allows the first-order ionospheric effect to be largely eliminated through linear combinations of dual-frequency observations. In contrast, the tropospheric delay is essentially non-dispersive in the GNSS frequency range and therefore affects all frequencies in nearly the same way. As a result, it cannot be removed by frequency combinations and must be explicitly modelled or estimated.

The troposphere extends from the Earth's surface to approximately 8–10 km in mid-latitudes and up to 16–18 km in the tropics [10]. It contains nearly all atmospheric water vapour and is characterised by strong vertical gradients in temperature, pressure, and humidity. These properties make the tropospheric delay both a critical error source for precise positioning and a valuable vertically integrated observable of the atmospheric state [3, 8].

2.2 GNSS Observation Equations

2.2.1 Fundamental GNSS Observables

GNSS receivers provide two primary observation types: (i) pseudorange (code) measurements, which deliver absolute signal travel time but are comparatively noisy, and (ii) carrier-phase measurements, which provide much higher precision but include an unknown ambiguity term [1, 2]. In PPP practice, both code and carrier phase are used together: code observations stabilise the absolute range/clock behaviour, while carrier phase provides the precision required for centimetre-level positioning and reliable troposphere estimation [9].

2.2.2 Ionosphere-Free Carrier-Phase Observation

For dual-frequency GNSS observations, the ionosphere-free (IF) linear combination is formed to eliminate the first-order ionospheric delay [1, 2]. Expressed in metres, the IF carrier-phase observation equation can be written as [9, 11]:

$$\Phi_{IF} = \rho + c(\delta t_r - \delta t_s) + T + \lambda_{IF} N_{IF} + \varepsilon$$

where

Φ_{IF} is the ionosphere-free carrier-phase observable (m),

ρ is the geometric distance between satellite and receiver antenna phase centres (m),

c is the speed of light in vacuum (m s^{-1}),

δt_r and δt_s are the receiver and satellite clock offsets (s),

T is the tropospheric signal delay along the signal path (m),

$\lambda_{IF} N_{IF}$ is the ionosphere-free carrier-phase ambiguity term expressed in metres (it is generally not an integer after forming the IF linear combination) [11],

ε includes unmodelled effects and measurement noise (m).

In this formulation, the tropospheric delay appears explicitly as an additive term that must be either modelled or estimated.

2.3 Slant Tropospheric Delay and Zenith Representation

2.3.1 Slant Delay Concept

The tropospheric delay T in the observation equation represents the slant tropospheric delay, i.e., the integrated refractive delay along the actual satellite–receiver signal path. This delay depends strongly on satellite elevation angle and atmospheric structure [3, 4].

Direct estimation of individual slant delays for each satellite would be impractical and numerically unstable. Instead, the troposphere is parameterised using a zenith delay combined with elevation-dependent mapping functions [2, 3]. Since the tropospheric delay appears as a slant-path quantity in the observation equation, practical estimation requires parameterisation through zenith delays combined with elevation-dependent mapping functions.

2.3.2 Mapping from Zenith to Slant

The slant tropospheric delay at the elevation angle el is expressed as [12, 13] :

$$T(el) = mhyd(el) \cdot ZHD + mwet(el) \cdot ZWD + G(el)$$

Where

ZHD is the Zenith Hydrostatic Delay (m),

ZWD is the Zenith Wet Delay (m),

$m_{\mathrm{hyd}}(el)$ and $m_{\mathrm{wet}}(el)$ are the hydrostatic and wet mapping functions (dimensionless),

$G(el)$ represents optional horizontal gradient contributions.

The Zenith Total Delay (ZTD) is defined as [4, 5, 12]:

$$ZTD = ZHD + ZWD$$

This decomposition separates the physically distinct contributions of dry air and water vapour to the total tropospheric delay.

2.4 Zenith Hydrostatic Delay (ZHD)

2.4.1 Physical Origin

The hydrostatic component arises from the refractivity of dry atmospheric gases, primarily nitrogen and oxygen. Because dry air is well mixed in the troposphere and exhibits relatively smooth spatial variability, the hydrostatic delay is highly stable and predictable [5, 11].

2.4.2 Analytical Expression

A commonly used approximation for the zenith hydrostatic delay is the Saastamoinen-type formulation [5]:

$$ZHD \approx \frac{0.0022768 \cdot P}{1 - 0.00266\cos(2\phi) - 0.00028h}$$

Where

ZHD is expressed in metres,

P is the surface atmospheric pressure (hPa),

ϕ is the geodetic latitude of the station (radians),

h is the ellipsoidal height of the station (km).

To first order, the denominator introduces only small corrections, meaning that ZHD is nearly linearly proportional to surface pressure. This theoretical property underpins the expectation of strong linear relationships between GNSS-derived hydrostatic delay (TRODRY) and ERA5 surface pressure.

2.5 Zenith Wet Delay (ZWD)

2.5.1 Physical Origin

The wet component of the tropospheric delay arises from the refractivity of atmospheric water vapour. Unlike dry air, water vapour is highly variable in both space and time and is typically concentrated in the lower troposphere [4, 12].

2.5.2 Integral Formulation

The zenith wet delay can be expressed as [4, 12]:

$$ZWD = 10^{-6}\int_0^{\infty}\left(k_2'\frac{e}{T} + k_3\frac{e}{T^2}\right)dz$$

where

e is the partial pressure of water vapour (hPa),

T is the absolute temperature (K),

z is height above the station (m),

k_2' and k_3 are refractivity constants.

This formulation highlights two key properties. First, ZWD is a vertically integrated quantity that depends on the full atmospheric moisture profile. Second, near-surface humidity indicators such as dew point temperature act only as proxies for this integrated column. It should also be noted that ERA5 dew point represents a grid-cell average rather than point-scale station measurements, which can further reduce apparent coupling in complex terrain. As a result, ZWD exhibits stronger variability and weaker predictability than ZHD, particularly in complex terrain [14, 15].

2.6 Mapping Functions

2.6.1 Purpose of Mapping Functions

Mapping functions convert zenith delays into slant delays as a function of elevation angle and account for the increased atmospheric path length at low elevations [11, 13].

2.6.2 Common Mapping Function Models

Early GNSS applications relied on simple elevation-angle-based mapping functions. Modern precise GNSS processing, however, employs more advanced models that distinguish between hydrostatic and wet components and account for dependencies on latitude, season, and, in some cases, numerical weather model information. Commonly used models include the Niell Mapping Function (NMF) and the Vienna Mapping Functions (VMF1 and VMF3) [13].

2.6.3 Implementation in This Thesis

Both CSRS-PPP and RTKLIB demo5 internally apply mapping functions as implemented by their respective PPP engines to relate zenith delays to slant delays for each elevation angle [9, 11]. In this thesis, mapping-function coefficients are not manually specified; instead, the analysis relies on the standard internal separation into hydrostatic and wet components provided by the PPP solutions.

For the multi-station case study, the estimated zenith parameters are interpreted as TRODRY ≈ ZHD and TROWET ≈ ZWD for physical comparison with ERA5 pressure and humidity. For the stability and spatial-mapping analyses, the combined zenith total delay (ZTD = ZHD + ZWD) is used as the primary observable.

2.7 Precise Point Positioning (PPP)

2.7.1 Conceptual Framework

Precise Point Positioning is a GNSS processing technique in which a single receiver is processed using precise satellite orbit and clock products. Unlike relative positioning, PPP does not rely on nearby reference stations [9].

2.7.2 PPP Observation Model

In PPP, the ionosphere-free carrier-phase observation can be written as [2, 9, 11]:

$$\Phi_{IF} = \rho + c(\delta t_r - \delta t_s) + m_{\text{hyd}}(el)ZHD + m_{\text{wet}}(el)ZWD + \lambda_{IF}N_{IF} + \varepsilon$$

The unknown parameters typically estimated include station coordinates, receiver clock offsets, zenith tropospheric delays, and carrier-phase ambiguities. Estimation is performed using least-squares adjustment and/or sequential filtering (implementation-dependent).

2.7.3 Troposphere Estimation in PPP

In PPP, the zenith delay is usually modelled as a stochastic parameter that is allowed to vary slowly in time. This approach enables the extraction of ZTD time series that reflect real atmospheric variability rather than relying solely on empirical tropospheric models [8, 9].

In the CSRS-PPP outputs used in Chapters 4–5, the zenith components are provided directly as TRODRY and TROWET, which correspond to the zenith hydrostatic and zenith wet components (i.e., TRODRY $\approx ZHD$ and TROWET $\approx ZWD$), with $ZTD = ZHD + ZWD$. The CSRS product files may also include horizontal gradient terms (north/east) [16]. These gradients are not analysed further in the present thesis, which focuses on zenith behaviour and its relationship to surface meteorological variables. Because ZTD is estimated as a stochastic parameter within the PPP filter, its temporal behaviour reflects both atmospheric variability and the stability of the underlying positioning solution, making subsequent quality control and positioning assessment essential for physical interpretation.

2.8 Meteorological Quantities Relevant to Tropospheric Delay

2.8.1 Temperature and Dew Point

ERA5 provides 2 m air temperature $t2m$ and 2 m dew point temperature $d2m$, originally expressed in Kelvin and converted to degrees Celsius. Dew point temperature represents the temperature at which air becomes saturated with respect to water vapour and is therefore a direct indicator of near-surface moisture content [14, 15].

2.8.2 Surface Pressure

Surface pressure sp is provided by ERA5 in Pascals and converted to hectopascals (1 hPa = 100 Pa). It directly controls the hydrostatic component of the tropospheric delay [5].

2.8.3 Precipitation

Total precipitation tp is provided in metres of water equivalent per hour and converted to millimetres. Precipitation reflects the outcome of atmospheric processes rather than instantaneous water vapour content, and its relationship with ZTD is therefore indirect and potentially lagged [4].

2.9 Position Stability and Its Relation to Tropospheric Estimation

Although tropospheric delay estimation is often treated independently of positioning performance, the two are fundamentally linked in PPP. Because height and tropospheric delay parameters are strongly correlated in PPP, instability in the Up component can bias ZTD estimates [8].

To assess PPP solution reliability, station coordinates are analysed in a local East–North–Up (ENU) frame relative to fixed reference coordinates. Position residuals are defined as:

$$\Delta\mathbf{x} = \mathbf{x}_{PPP} - \mathbf{x}_{ref}$$

These residuals are transformed from ECEF to local ENU coordinates and summarised using both raw dispersion metrics and robust statistics after quality filtering. Particular attention is paid to the vertical (Up) component, which is typically more weakly constrained and more sensitive to tropospheric modelling.

In this thesis, positioning stability is treated as a necessary quality indicator for interpreting ZTD results, especially for low-cost receivers.

CHAPTER 3: STUDY AREA AND GNSS STATIONS

3.1 Rationale for Station and Area Selection

The objective of this thesis is not limited to evaluating tropospheric delay estimation under idealised or homogeneous conditions. Instead, it aims to assess how GNSS-derived tropospheric parameters behave across contrasting atmospheric regimes, topographic settings, and station configurations, including both geodetic-grade and low-cost installations [4, 6, 17].

To this end, the study deliberately combines stations spanning coastal, inland lowland, plateau, and mountainous environments, exposed to marine, continental, alpine, and Mediterranean climatic influences [17, 18]. The selected dataset further integrates stations belonging to official geodetic reference networks and dense low-cost community networks, and includes both short-term regime-focused case studies and longer-term stability analyses based on continuous PPP processing [6, 7, 9, 19].

This diversity enables the investigation of two complementary aspects: first, the physical consistency of GNSS-derived tropospheric delays with atmospheric variables under differing environmental conditions; second, the robustness and transferability of tropospheric products across receiver classes and deployment scenarios.

3.2 Overview of the GNSS Station Set

The multi-station tropospheric case study is based on seven Centipede low-cost GNSS stations; IGMI00ITA is included as a geodetic reference station for the Florence stability branch. The station set comprises [7, 19]:

RUF01 (RUF): Urban lowland low-cost GNSS station, Italy,

SCV1: Alpine low-cost GNSS station, France,

AIGL: Mountainous or foothill low-cost GNSS station, France,

SOPH: Mediterranean coastal low-cost GNSS station, southeastern France,

PYLA: Atlantic coastal low-cost GNSS station, western France,

BTCHZ: Plateau low-cost GNSS station, Switzerland,

STAVE: Inland lowland low-cost GNSS station, Belgium,

IGMI00ITA (IGMI): Geodetic EUREF GNSS station, Italy.

Together, these stations span a wide range of elevations, from near sea level to high alpine terrain, and represent diverse climatic and atmospheric conditions. This configuration provides a suitable basis for evaluating how tropospheric delay behaviour varies with environment, topography, and atmospheric structure.

Table 3.1 summarises the GNSS stations analysed in this thesis, including approximate coordinates, data network, and a simplified environmental classification used to interpret regime-dependent behaviour in Chapters 5–6.

Table 3.1. GNSS stations analysed in this thesis and the environmental classification [7, 19].

GNSS station	***Latitude (°)***	***Longitude (°)***	***Height (m)***	***Network***	***Station ID (RINEX/EPN)***	***Environment type***
IGMI00ITA (Tuscany, IT)	43.7956	11.2138	95.1	EUREF/EPN	IGMI00ITA	Urban lowland
RUF01 (Tuscany, IT)	43.6952	11.3262	248.7	CentipedeRTK	A33100ITA	Urban lowland
STAVE (Poperinge, BE)	50.9310	2.6540	57.6	CentipedeRTK	A38600BEL	Inland lowland
BTCHZ (Payerne, CH)	46.7572	6.7824	754.5	CentipedeRTK	A22000CHZ	Plateau
SCV1 (Briançon, FR)	44.9202	6.5496	2328.5	CentipedeRTK	A52300FRA	Mountainous
AIGL (Massif Central, FR)	44.1214	3.5813	1618.8	CentipedeRTK	A68000FRA	Mountainous
PYLA (Gironde, FR)	44.6274	-1.1450	57.2	CentipedeRTK	A45500FRA	Coastal (Atlantic)
SOPH (Sophia Antipolis, FR)	43.6114	7.0541	178.8	CentipedeRTK	A70500FRA	Coastal (Mediterranean)

Coordinates and heights are provided for descriptive and classification purposes; PPP processing and ERA5 extraction use the station identifiers and time-synchronised series as described in Chapters 4–5.

3.3 Environmental Classification of Stations

To support a physically grounded interpretation of GNSS–meteorology relationships, the selected stations are grouped into four environmental categories based on geographic setting, climatic influence, and topographic context. This environmental classification is not intended as a climatological taxonomy, but as a physically motivated framework that acknowledges atmospheric structure, boundary-layer behaviour, and vertical moisture distribution vary systematically across environments and directly affect both tropospheric delay estimation and its linkage to surface meteorological variables [17, 18].

3.3.1 Coastal Stations

Coastal stations are characterised by their direct exposure to large water bodies, which exert strong control on local atmospheric conditions. Humidity levels are typically high, diurnal temperature ranges are reduced by the thermal inertia of the sea, and the atmospheric boundary layer is frequently influenced by marine air masses. Sea–land breeze circulations are common, particularly during summer, introducing pronounced diurnal variability in temperature, humidity, and wind fields [20].

Within this study, the coastal category includes the SOPH station along the Mediterranean coast and the PYLA station on the Atlantic coast. These sites often exhibit strong coupling between near-surface moisture indicators, such as dew point temperature, and column-integrated water vapour represented by the GNSS wet delay. However, this coupling can become regime-dependent, particularly under stable summer conditions, when shallow marine boundary layers and local circulations may decouple surface humidity from the deeper moisture column [4, 17, 20].

3.3.2 Inland Lowland Stations

Inland lowland stations are typically associated with relatively homogeneous atmospheric conditions and limited topographic complexity. Under fair-weather conditions, the boundary layer is often well mixed, allowing near-surface meteorological variables to be broadly representative of the lower troposphere. These environments usually exhibit strong seasonal variability driven by the annual cycle of radiative forcing and synoptic-scale circulation [17].

The inland lowland category includes the RUF01 and STAVE stations. Owing to their atmospheric characteristics, such environments are particularly favourable for interpreting GNSS-derived wet tropospheric delay using surface meteorological proxies. The relatively direct linkage between surface conditions and column-integrated water vapour makes these stations valuable reference cases for evaluating the physical consistency of GNSS–meteorology relationships.

3.3.3 Plateau Stations

Plateau stations occupy intermediate elevations and are generally subject to more continental atmospheric regimes with reduced direct marine influence. These stations may exhibit a combination of lowland-like and mountain-like characteristics, depending on synoptic conditions and season. During winter, synoptic-scale forcing often dominates, whereas transitional behaviour may emerge during other periods [17, 18].

In this study, the BTCHZ station, located on the Swiss plateau, represents this category. Its moderate elevation allows investigation of how increasing height above sea level affects the representativeness of near-surface meteorological observations for tropospheric delay estimation, particularly in comparison with both lower-elevation inland stations and high-altitude mountainous sites.

3.3.4 Mountain and Complex Terrain Stations

Mountain stations are located in regions with pronounced topographic variability and strong vertical gradients in atmospheric properties. These environments are characterised by frequent vertical stratification of humidity, which can reduce the representativeness of surface meteorological conditions for the free troposphere. Orographic lifting, downslope flows, and valley circulations further contribute to complex and highly variable atmospheric dynamics [18].

The mountain and complex terrain category includes the SCV1 and AIGL stations [7]. In such settings, discrepancies between model orography and true station elevation in atmospheric reanalysis products may introduce additional uncertainties. Consequently, these stations constitute particularly stringent test cases for evaluating GNSS–meteorology relationships [14, 15].

3.4 The Florence Study Area: RUF01 and IGMI00ITA

3.4.1 Geographic Context

The Florence region in central Italy is situated within a broad inland basin surrounded by low mountain ranges. The regional climate is influenced by Mediterranean conditions, with warm and dry summers, mild to cool winters, and frequent synoptic-scale precipitation events during autumn and winter. This combination produces pronounced seasonal and synoptic variability while avoiding the extreme topographic complexity associated with high alpine terrain [21].

These characteristics make the Florence area well-suited for GNSS tropospheric studies, enabling investigation of tropospheric delay variability across a wide range of atmospheric conditions within a relatively controlled topographic setting.

3.4.2 RUF01

RUF01 is a low-cost GNSS station installed as part of the Centipede network in an urban or semi-urban environment at moderate elevation above sea level [7]. The station employs a non-geodetic antenna with limited calibration information and records observations at a 30 s sampling interval. While this temporal resolution is sufficient for PPP-based tropospheric estimation, it introduces challenges related to increased noise levels, multipath susceptibility, and long-term stability [6].

RUF01 is representative of a rapidly growing class of low-cost GNSS installations that are inexpensive and easily deployable. Demonstrating the stability and physical reliability of tropospheric products derived from such stations is therefore central to assessing their potential for dense atmospheric monitoring networks.

3.4.3 IGMI00ITA (Reference Station)

IGMI00ITA is a geodetic GNSS station belonging to the EUREF Permanent GNSS Network [19]. It is equipped with a geodetic-grade GNSS receiver, a calibrated antenna, and stable monumentation, and operates according to established geodetic standards. The station provides high-quality tropospheric products, including hourly zenith total delay estimates distributed in SINEX-TRO format [22], which are widely used as reference data in European GNSS meteorology.

3.4.4 Relative Geometry and Height Difference

Although RUF01 and IGMI00ITA are geographically close, they differ in elevation, with RUF01 located approximately 150 m higher than IGMI00ITA. This height difference introduces a systematic offset in zenith tropospheric delay, as the integrated atmospheric mass above the station decreases with elevation. The co-location of these stations within the same regional atmospheric regime, combined with their contrasting hardware characteristics and elevation difference, provides a well-controlled configuration for evaluating temporal coherence, interstation bias stability, and the sensitivity of low-cost PPP-derived tropospheric estimates to station height and instrumentation.

3.5 GNSS Networks and Data Sources

3.5.1 EUREF Permanent GNSS Network

The EUREF Permanent GNSS Network constitutes a core reference infrastructure for geodetic and atmospheric research in Europe. It provides harmonised, high-quality GNSS observations and derived products processed according to rigorous standards. Among its principal outputs are tropospheric parameters distributed in SINEX-TRO format, which serve as benchmarks for assessing the quality and physical reliability of GNSS-derived tropospheric delays.

3.5.2 Centipede and RENAG/RESIF Networks

The Centipede network, supported by RENAG/RESIF, provides a dense deployment of low-cost GNSS stations with fully open-access observation data. The network has expanded rapidly, offering high spatial resolution at substantially lower cost than traditional geodetic infrastructures. Although individual stations may not meet the strict performance standards of geodetic-grade sites, the combination of network density, spatial coverage, and accessibility makes Centipede particularly attractive for atmospheric applications, where spatial representativeness can partially compensate for reduced single-station performance.

3.6 Representativeness of the Station Set

The combined station set enables this thesis to address several scientific dimensions simultaneously. It allows evaluation of the physical consistency of GNSS-derived tropospheric delays across climates, the environmental sensitivity of ZTD–meteorology relationships, the robustness of low-cost GNSS installations, and the scalability of GNSS tropospheric sensing from individual stations to dense regional networks [4, 6, 7].

By intentionally including both favourable and challenging environments, the study avoids overly optimistic conclusions and instead provides a realistic and balanced assessment of GNSS tropospheric capabilities.

CHAPTER 4: DATA AND METHODOLOGY

4.1 Chapter purpose and methodological logic

This thesis combines (i) short, controlled multi-station experiments that isolate physical relationships between GNSS tropospheric parameters and meteorological variables, (ii) a long-term low-cost station validation designed to separate processing-strategy effects from receiver-class effects, and (iii) a national-scale spatial reconstruction of ZTD over France to assess whether dense low-cost GNSS networks can support atmospheric mapping [6].

These three analysis branches are organised within a single methodological framework that includes data acquisition and format harmonisation (GNSS observations, precise products, ERA5 reanalysis data, and geodetic reference products), processing steps based on CSRS-PPP outputs for the two case-study days and batch PPP processing using RTKLIB demo5 for RUF01 and IGMI00ITA under two precise-product strategies (CDDIS and GFZ), standardised and station-agnostic quality control and aggregation procedures, and consistent temporal alignment strategies. Temporal alignment involves hourly aggregation for the case-study days, a forced hourly grid at :30 UTC for the stability study, and a +30 min time shift applied to ERA5 data for the Florence stability branch. Statistical analysis and diagnostic evaluation are then performed using bias, RMS difference, and correlation metrics, regression sensitivity analyses, precipitation lag tests, and ENU positioning stability indicators.

The purpose of this structured approach is not only to compute results, but to ensure that each processing step and data transformation remains physically meaningful, methodologically consistent, and fully reproducible.

4.2 Data sources and file types

4.2.1 GNSS stations and observation data

Two GNSS usage regimes are considered.

(A) Multi-station case study (two dates; seven stations).

Tropospheric products are obtained from CSRS-PPP outputs (PPP processing performed externally). Raw GNSS observation files exist upstream but are not part of the internal processing chain for this branch; the internal workflow begins from the CSRS-PPP tropospheric output files [16].

(B) RUF stability and sensitivity (two months; two stations; two product strategies).

Raw GNSS observations are processed internally using RTKLIB demo5. The observation set includes RUF01 (low-cost station, Centipede/RENAG origin) and IGMI00ITA (geodetic EUREF station) [7, 19]. IGMI00ITA is processed internally to enable controlled comparisons under an identical estimator and configuration. Observations are sampled at 30 s and cover 2024-03-01 to 2024-04-30 (61 days). Data are typically distributed as Hatanaka-compressed RINEX (.crx.gz) and converted to RINEX (.rnx) prior to processing [24, 25].

4.2.2 Precise products and external reference products

Two independent precise-product strategies are used to isolate product sensitivity.

CDDIS strategy (NASA CDDIS distribution; IGS/MGEX products).

Precise orbits are provided in SP3 format, precise clocks in CLK format, and Earth rotation parameters in ERP format [26]. Navigation information is taken from BRDM (multi-GNSS broadcast navigation). Bias-related auxiliary files (e.g., BIA/OSB) are used when available and compatible with the implemented processing chain.

GFZ strategy (GFZ final products).

The same core product classes are used (SP3/CLK/ERP) [27]. No separate GFZ bias file is applied in the sensitivity branch, consistent with the implemented workflow.

Independent geodetic ZTD reference (anchor).

Hourly ZTD for IGMI00ITA is extracted from EUREF SINEX-TRO products and treated as the external anchor in the stability assessment [22].

4.2.3 ERA5 meteorological data and units

ERA5 is used as an independent atmospheric reference. The analysis relies on four hourly UTC ERA5 variables: 2 m temperature, 2 m dew point temperature, surface pressure, and total precipitation, each converted to units suitable for direct statistical comparison with GNSS-derived time series [14, 15]. For the Florence stability branch, ERA5 timestamps are shifted by +30 minutes to align ERA5(t) with the forced GNSS ZTD grid at :30 (Section 4.6).

Table 4.1. ERA5 variables and conversions

ERA5 variable	*ERA5 name*	*Native unit*	*Unit used in thesis*	*Conversion applied*
2 m temperature	t2m	K	°C	$T_{°C} = T_K - 273.15$
2 m dew point	d2m	K	°C	$T_{d,°C} = T_{d,K} - 273.15$
Surface pressure	sp	Pa	hPa	$P_{hPa} = P_{Pa}/100$
Total precipitation	tp	m	mm	$tp_{mm} = tp_m \times 1000$

ERA5 variables represent grid-cell averages rather than point-scale observations at the GNSS antenna, and are therefore treated as physically consistent but imperfect proxies for local conditions, particularly in complex terrain.

4.3 Tropospheric parameter definitions and physical equations

The physical definition of ZTD, its hydrostatic/wet decomposition, mapping functions, and the hydrostatic (Saastamoinen-type) and wet refractivity formulations are developed in Chapter 2.

In this chapter, these relationships are used as interpretive constraints (e.g., expected near-linearity of hydrostatic delay with surface pressure), while the methodological focus is on how tropospheric time series are produced, quality-controlled, aggregated, synchronised, and evaluated. The Niell Mapping Function (NMF) was used to project slant delays to the zenith direction, while the Saastamoinen model provided the a priori hydrostatic delay component, with the residual ZTD estimated as a stochastic parameter within the PPP adjustment.

4.4 Multi-station case study workflow (CSRS-PPP + ERA5)

4.4.1 CSRS-PPP tropospheric outputs and time handling

For each station and each analysed day, CSRS-PPP provides tropospheric estimates at 30 s resolution. Epoch timestamps are encoded as:

$$yy{:}\,DOY{:}\,sssss$$

where $sssss$ denotes seconds-of-day. These timestamps are converted to UTC datetimes, after which tropospheric parameters are aggregated to hourly values for consistent comparison with hourly ERA5 variables [16].

4.4.2 Hourly aggregation

For any variable $x(t)$ sampled within an hour H with n_H samples, the hourly mean is:

$$\bar{x}_H = \frac{1}{n_H}\sum_{i=1}^{n_H} x_i$$

This aggregation is applied to TRODRY and TROWET (and to ZTD when required). Hourly aggregation is adopted as a compromise between temporal resolution and estimator noise reduction, under the assumption that sub-hourly PPP variability is dominated by stochastic positioning effects rather than physically meaningful atmospheric signals.

4.4.3 Statistical comparisons used in the case study

The statistical framework is intentionally simple and transparent. Pearson correlation is computed following standard statistical definitions [28]:

$$r_{XY} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$$

Linear relationships are assessed using simple linear regression:

$$y = ax + b$$

where a is the slope (e.g., mm hPa^{-1} for TRODRY vs pressure; mm $^\circ\mathrm{C}^{-1}$ for TROWET vs dew point) and b is the intercept. Model error is quantified by RMSE:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Precipitation coupling is examined using lagged correlations between TROWET and total precipitation. For lag k hours:

$$r(k) = corr\big(TROWET(t), tp(t+k)\big)$$

with $k \in \{-1,0,+1\}$, Positive lag values ($k > 0$) indicate precipitation occurring after increases in wet delay, allowing assessment of lead-lag behaviour.

4.5 RUF stability workflow (RTKLIB PPP + EUREF + ERA5)

4.5.1 RTKLIB PPP processing

The stability period spans 61 daily files and high-rate epochs, motivating a batch daily processing design. For each day, RTKLIB demo5 produces a position time series (.pos) and a statistics file (.pos.stat) containing, among other outputs, troposphere records [29]. The PPP configuration is held constant across stations (RUF01 and IGMI00ITA) and across product strategies (CDDIS and GFZ), so that differences in ZTD behaviour can be attributed to product strategy or station class rather than configuration drift.

4.5.2 Dual precise-product strategies: CDDIS vs GFZ

The sensitivity analysis uses a controlled parallel-branch design in which estimator settings and modelling choices are fixed and only the correction source is changed [9].

Table 4.2. Controlled PPP parallel-branch design

Component	***Held constant***	***Changed between branches***
RTKLIB mode	static PPP	no
Constellations	GPS + Galileo	no
Frequency handling	ionosphere-free dual-frequency	no
Troposphere model	Saastamoinen a priori model + ZTD estimated as a stochastic parameter	no
Mapping functions	Niell Mapping Function (NMF)	no
Elevation mask	~10°	no
RTKLIB config file	same stochastic model	no
Precise products	—	CDDIS vs GFZ (SP3/CLK/ERP)

4.5.3 ZTD parsing from RTKLIB outputs

From the .pos.stat files, $TROP records are parsed to obtain ZTD [m], $\mathrm{ZTD}_{\mathrm{STD,i}}$ [m] (formal uncertainty), and timestamps [29]. A unified QC and aggregation pipeline is then applied in a station-agnostic manner to RUF_CDDIS, RUF_GFZ, IGMI_CDDIS, and IGMI_GFZ, and the resulting hourly series are compared against IGMI_EUREF as an external anchor.

4.6 Quality control and hourly aggregation

This section defines the transformation from raw PPP-derived tropospheric estimates into a defensible scientific time series.

4.6.1 Basic physical and statistical QC

First-stage filters applied to each epoch i,

Physical plausibility:

$$0.5 \le ZTD_i \le 4.5 \quad [m]$$

Formal uncertainty threshold:

$$0 < ZTD_{STD_i} \le 0.25 \quad [m]$$

These bounds are conservative for mid-latitude zenith delays and are designed to remove gross failures without over-filtering [11].

4.6.2 Robust outlier handling: rolling median + MAD

To reduce the influence of transient PPP disturbances while preserving sampling regularity, outliers are handled using a rolling median and MAD approach applied per station time series, using a centred window of ±15 minutes around each epoch t_i. MAD-based filtering is preferred over standard deviation-based methods because it is robust to non-Gaussian tails and isolated PPP failures, which are common in low-cost GNSS processing. Because replacement can locally reduce variance, the number of replaced epochs is reported explicitly and remains small relative to the total epoch count; conclusions are based on synoptic-scale coherence and inter-series consistency rather than on high-frequency variance at the 30 s level. If W_i denotes the set of ZTD values within that window, the rolling median is [30]:

$$m_i = median(W_i)$$

The median absolute deviation (MAD) is:

$$MAD_i = median(\mid w - m_i \mid) \text{ for } w \in W_i$$

Using the Gaussian-consistent scaling:

$$\sigma_i \approx 1.4826 \cdot MAD_i$$

An epoch is flagged as an outlier if:

$$\mid ZTD_i - m_i \mid > 3\sigma_i$$

Flagged epochs are not deleted; they are replaced to preserve continuity:

$$ZTD_i \leftarrow m_i$$

This choice is motivated by the need to preserve temporal continuity for hourly aggregation; all replacement counts are explicitly reported to ensure transparency and to avoid masking systematic errors.

4.6.3 Forced hourly aggregation at :30 UTC

For the RUF stability study, an hourly grid is imposed at :30 UTC. For each hour bin H containing n_H valid epochs ZTD_j, the hourly mean is computed as:

$$\bar{ZTD}_H = \frac{1}{n_H}\sum_{j=1}^{n_H} ZTD_j$$

A minimum of 30 samples per hour (corresponding to approximately 15 minutes of valid data at 30 s sampling) is required for an hourly estimate to be retained. ERA5 timestamps are shifted by +30 minutes so that ERA5(t) aligns with $\bar{ZTD}$ at :30 UTC.

4.7 Intercomparison metrics

Intercomparison is performed on synchronised hourly series. For the two ZTD series A_i and B_i in mm, the difference series is:

$$d_i = A_i - B_i$$

Bias (mean difference):

$$Bias_{A-B} = \frac{1}{N}\sum_{i=1}^{N} d_i$$

RMS difference:

$$RMS_{A-B} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} d_i^2}$$

Standard deviation of differences:

$$STD_{A-B} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(d_i - \bar{d})^2}$$

Mean absolute difference:

$$MAE_{A-B} = \frac{1}{N}\sum_{i=1}^{N} | d_i |$$

Pearson correlation r_{AB} follows the definition in Section 4.4.3. These metrics are computed (i) for each series versus IGMI_EUREF (anchor) and (ii) for internal consistency checks (RUF_CDDIS vs RUF_GFZ, IGMI_CDDIS vs IGMI_GFZ).

4.8 ZTD–ERA5 relationship modelling

For each ZTD series (mm) and each ERA5 predictor X, with $X \in \{T_{^\circ C}, T_{d,^\circ C}, P_{hPa}, tp_{mm}\}$, the relationship is modelled as:

$$ZTD = aX + b$$

The slope a has units mm $^\circ\mathrm{C}^{-1}$ for temperature and dew point, mm hPa^{-1} for pressure, and mm mm^{-1} for precipitation. For each fit, the reported diagnostics are r, slope a, $R^2 = r^2$ under simple linear regression, and RMSE (Section 4.4.3). This modelling is used as a consistency check rather than a full physical decomposition: slopes and signs must align with known atmospheric behaviour, and low-cost versus geodetic series should yield comparable sensitivities if the GNSS products are physically meaningful [4, 5, 12].

4.9 Positioning stability analysis

Positioning stability is treated as a methodological pillar of the RUF stability chapter because coordinate instability can contaminate ZTD estimation. The analysis uses RTKLIB .pos outputs and is restricted to PPP solutions by retaining epochs with solution quality Q = 6, avoiding mixing nominal PPP epochs with degraded solutions [29].

Reference coordinates are defined per station. For RUF01, Centipede network coordinates are used and treated as fixed over the two-month period [7]. For IGMI00ITA, published EUREF coordinates with velocities are propagated from reference epoch 2010.0 to approximately 2024.25 in ECEF [19].

Positioning stability is assessed by forming coordinate residuals between the PPP-derived receiver position and an external reference position, then expressing these residuals in a local East–North–Up (ENU) frame centred at the station. For each epoch, PPP geodetic coordinates are converted to ECEF, differenced against the reference ECEF coordinates, and rotated to ENU using the station latitude and longitude. For stations with published velocities, the reference coordinates are linearly propagated from the reference epoch to the analysis period. Full details are summarised in *Table 4.3*.

Table 4.3. Station reference coordinates used for PPP ENU residuals

Station	***Reference type***	***Geodetic reference (lat, lon, h)***	***ECEF reference at t0=2010.0 (m)***	***Velocities (m/yr)***	***Propagated ECEF at t=2024.25 (m)***	***ENU rotation origin (lat, lon)***
RUF01	Fixed (Centipede)	43.69519427° 11.32616259° 248.735 m	—	—	—	43.69519427° 11.32616259°
IGMI00ITA	EUREF + linear velocity	—	X0=4,523,251.190 Y0=896,760.115 Z0=4,391,796.404	VX= -0.0163 VY= +0.0186 VZ= +0.0126	Xref=4,523,250.9577 Yref=896,760.3801 Zref=4,391,796.5836	43.795650° 11.213803°

Let $X_{PPP} = [X, Y, Z]^T$ denote PPP ECEF coordinates [m] and X_{REF} the reference ECEF coordinates [m]. The ECEF residual is:

$$\Delta X = X_{PPP} - X_{REF}$$

Residuals are rotated into the local ENU frame at the reference latitude ϕ and longitude λ:

$$\begin{bmatrix} dE \\ dN \\ dU \end{bmatrix} = R(\phi, \lambda)\, \Delta X$$

where $R(\phi, \lambda)$ is the standard ECEF→ENU rotation matrix. Horizontal and 3D magnitudes are defined as:

$$dH = \sqrt{dE^2 + dN^2}, \qquad d3D = \sqrt{dE^2 + dN^2 + dU^2}$$

Raw metrics are reported to retain diagnostic information, including spikes, and include means and standard deviations of dE, dN, dU, along with $std(d3D)$, $median(\mid d3D \mid)$, $p95(\mid d3D \mid)$, $RMS(d3D)$, and maxima.

4.9.1 Robust (typical behaviour) positioning statistics

In addition to raw dispersion metrics, a robust filtering stage is applied to characterise typical PPP behaviour after excluding transient failures. First, only converged PPP epochs are retained using the RTKLIB solution-quality flag (Q = 6). Second, epochs with insufficient tracked satellites are removed using a minimum satellite-count threshold (ns ≥ 7). Third, component-wise outliers in the ENU residuals are rejected using a MAD-based rule applied separately to dE, dN, and dU: an epoch is flagged if any component exceeds median ± 6·(1.4826·MAD) [30]. Finally, to suppress the influence of extreme tails without imposing an arbitrary hard bound, the upper 0.5% of |dU| and |d3D| values are trimmed. The outlier rate is reported as the percentage of PPP-valid epochs removed by this robust filtering pipeline. Robust statistics (median(|d3D|), p95(|d3D|), and related metrics) are computed on the retained 30 s epochs.

4.10 France spatial ZTD mapping workflow

The France-wide spatial reconstruction of ZTD for March 2024 is based on a complete end-to-end processing chain that includes batch PPP processing, epoch-level quality control, station-level filtering, vertical trend modelling, residual kriging, and DEM-based reconstruction. The objective is to generate a physically consistent national ZTD field and to evaluate alternative interpolation strategies through cross-validation in order to assess methodological robustness.

4.10.1 Batch PPP and epoch-level ZTD extraction

The batch processing workflow encompasses 10,713 daily RINEX observation files covering metropolitan France for March 2024. Static PPP processing is performed using RTKLIB demo5 under the configuration described in Section 4.5, producing 10,712 corresponding .pos.stat files.

From these outputs, $TROP records are parsed to extract ZTD [m], $\mathrm{ZTD}_{\mathrm{STD},i}$ [m] (formal uncertainty), and epoch timestamps. Across the campaign, $TROP records were successfully parsed for 10,617 station-day files and passed to the quality-control stage. The resulting dataset includes 360 stations with usable ZTD time series prior to station-level filtering.

4.10.2 QC and aggregation for France (C1–C4)

The national-scale quality-control implementation follows exactly the station-agnostic C1–C4 pipeline defined in Section 4.6. The same physical plausibility bounds and formal-uncertainty thresholds are applied without modification ($0.5 \leq \mathrm{ZTD} \leq 4.5$ m; $0 < \mathrm{ZTD\_STD} \leq 0.25$ m), ensuring consistency with the mid-latitude tropospheric constraints discussed previously.

Robust outlier handling is implemented using the rolling-median and MAD-based procedure described in Section 4.6.2, employing a centred ±15 min window and a 3σ detection threshold with Gaussian-consistent scaling [30]. Flagged epochs are replaced by the local rolling median rather than deleted, preserving temporal continuity for subsequent aggregation while maintaining transparency through explicit reporting of replacement counts.

After quality control, epoch-level ZTD estimates are aggregated to hourly means using the formulation provided in Section 4.6.3. This ensures that the France-wide spatial reconstruction is based on temporally consistent, variance-reduced station-hour values derived under an identical QC logic to that used in the stability and case-study branches. The resulting hourly station dataset constitutes the baseline input for spatial interpolation.

4.10.3 Station-level filtering and definition of the interpolation-ready dataset

After hourly aggregation, station-level filters are applied sequentially to ensure spatial representativeness and positional stability. Stations are first required to provide at least 25 valid days during March 2024, and to exhibit a station-level ZTD standard deviation not exceeding 30 mm. This coverage and variability criterion reduces the dataset from 360 to 332 stations.

The spatial domain is then restricted to metropolitan France using a geographic bounding box, yielding 322 stations. To prevent coordinate instability from contaminating the spatial tropospheric field, a stability criterion based on Rad95 ≤ 0.10 m is applied using daily position solutions derived from the PPP .pos outputs (Section 4.9). The application of this stability filter produces the final interpolation-ready set of 274 stations.

These steps ensure that the spatial reconstruction is based exclusively on stations exhibiting sufficient temporal coverage, stable positioning behaviour, and geographic consistency.

4.10.4 Vertical trend modelling and detrending

Zenith Total Delay exhibits a strong and physically expected dependence on station elevation, primarily driven by the hydrostatic component of atmospheric delay as described in Chapter 2. To isolate horizontal spatial variability, the dominant vertical trend is explicitly modelled prior to interpolation.

A least-squares linear regression is fitted to station-mean ZTD values (mm) as a function of station elevation h(m) for the 274 filtered stations. The resulting relationship is:

$$ZTD_{hat}(h) = -0.283838\,h + 2395.059384$$

with $R^2 = 0.9932$, RMSE = 6.721 mm, and $N = 274$.

Although atmospheric pressure decreases exponentially with height, over the elevation range of metropolitan France this dependence is well approximated by a linear model, as indicated by the high coefficient of determination. This regression captures the dominant vertical structure of the tropospheric field. This linear formulation is adopted as a first-order approximation suitable for the elevation range and monthly-mean scale considered in this study, rather than as a strict physical representation of atmospheric stratification.

Residuals are defined as:

$$Resid_i = ZTD_{obs,i} - ZTD_{hat}(h_i)$$

This detrending procedure removes the primary topographic signal, ensuring that subsequent interpolation targets horizontal atmospheric variability rather than elevation-driven bias.

4.10.5 Residual kriging in projected coordinates

The spatial interpolation is performed on the residual field rather than on raw ZTD values. Interpolation is carried out using Universal Kriging in the EPSG:2154 (Lambert-93) projected coordinate system to ensure metric consistency across the metropolitan France domain.

The interpolation grid is defined at 1 km spatial resolution over metropolitan France. Variogram parameters are obtained through automatic fitting procedures, and a local neighbourhood search is adopted with an approximate radius of 200 km, using between 8 and 20 neighbouring stations. Automatic model selection within the kriging implementation determines the optimal theoretical variogram form (e.g., spherical or exponential) based on empirical semivariogram fitting, while preserving consistency across validation runs. By interpolating only the residual component, elevation-related artefacts are avoided and the kriging model focuses exclusively on horizontally structured atmospheric variability.

Cross-validation metrics, including RMSE, NRMSE, and R^2, are used to evaluate interpolation performance.

4.10.6 DEM preprocessing and final ZTD reconstruction

Elevation information at each grid node is required to reconstruct the full ZTD field. A Digital Elevation Model derived from SRTM data is obtained via OpenTopography and preprocessed through mosaicking of relevant tiles, reprojection to EPSG:2154, resampling to 1 km resolution, and clipping to the metropolitan France extent.

Let $h(x, y)$denote the DEM elevation at grid node $(x|y)$. The final ZTD field is reconstructed using a regression–kriging formulation:

$$ZTD_{final}(x, y) = [-0.283838 \cdot h(x, y) + 2395.059384] + Residual_{kriged}(x, y)$$

This formulation ensures that the physically dominant vertical dependence is explicitly represented through the regression component, while the kriged residuals capture horizontal atmospheric variability. The resulting national ZTD grid is therefore consistent with station observations, topographic structure, and geostatistical interpolation principles.

4.10.7 Cross-validation framework and comparative interpolation methodology

To assess the methodological robustness of the France-wide ZTD reconstruction and to quantify the influence of modelling assumptions on spatial prediction, a structured cross-validation framework is implemented. This framework is designed to compare alternative interpolation strategies under identical data constraints, ensuring that performance differences arise from modelling philosophy rather than data treatment.

The comparison is conducted under the simplifying assumptions adopted for the national reconstruction. Zenith Total Delay is treated as a single integrated quantity without explicit separation into hydrostatic and wet components, although the wet component does not exhibit strictly linear height dependence [12]. Furthermore, the relationship between ZTD and elevation is assumed spatially constant and globally linear across metropolitan France. Finally, interpolation is performed on monthly mean station ZTD values rather than on daily reconstructions, reducing temporal variability and enabling controlled methodological evaluation.

Within this framework, three spatial interpolation strategies are examined: Inverse Distance Weighting (IDW), Universal Kriging (UK) applied directly to station-mean $ZTD_{obs}$with elevation included as a predictor, and Regression–Kriging (RK), in which a deterministic elevation-dependent trend is first estimated and removed prior to kriging of the residual field.

Inverse Distance Weighting is implemented as a deterministic estimator in which predictions at location x_0 are computed as a weighted average of neighbouring observations, with weights inversely proportional to horizontal distance raised to a power parameter p [31]. This formulation assumes that spatial similarity decreases solely as a function of Euclidean distance and does not incorporate physical predictors such as elevation.

Universal Kriging extends the classical kriging framework by decomposing the spatial field into deterministic and stochastic components [32]. The field is expressed as:

$$Z(x) = m(x) + \varepsilon(x),$$

where $m(x)$ represents a deterministic trend incorporating elevation and projected spatial coordinates, and $\varepsilon(x)$ is a spatially correlated zero-mean residual. Variogram parameters are estimated from the empirical semivariogram and used to solve the kriging system of equations within the projected EPSG:2154 coordinate system.

Regression–Kriging adopts a two-stage formulation in which the deterministic elevation-dependent trend is estimated globally and removed prior to spatial interpolation [32, 33]. A linear regression between station-mean ZTD and elevation is fitted to obtain:

$$ZTD_{hat}(h) = ah + b,$$

After which residuals are computed and interpolated using kriging. The final reconstructed field is obtained by the superposition of the deterministic trend and the interpolated residual surface. This formulation explicitly separates vertical structure from horizontal atmospheric variability.

Predictive performance for each method is evaluated using leave-one-out cross-validation (LOOCV). Each station is removed sequentially, predicted from the remaining stations, and the prediction error is computed. Performance metrics, including RMSE, MAE, Bias, and R^2, are calculated according to the definitions provided in Section 4.7. The use of a unified LOOCV framework ensures that all interpolation strategies are evaluated under identical statistical definitions and data conditions, providing a coherent methodological basis for the comparative analysis presented in Chapter 7.

CHAPTER 5: Multi-Station Tropospheric Case Study Results under Dry and Rainy Regimes

5.1 Chapter Scope and Analytical Context

This chapter presents the results of a controlled multi-station case study designed to evaluate the physical consistency of GNSS-derived tropospheric delays under contrasting meteorological regimes. Building on the methodology described in Chapter 4, zenith tropospheric delay components obtained from Precise Point Positioning (PPP) are compared with collocated ERA5 meteorological variables at several European GNSS stations.

Two 24-hour periods are selected to represent contrasting lower-tropospheric regimes: a dry summer day (15 July 2024), characterised by relatively stable boundary-layer conditions [17], and a rainy winter day (3 March 2024), dominated by synoptic-scale forcing and widespread precipitation [10]. The objective of this chapter is to assess the consistency between TRODRY (hydrostatic delay) and surface pressure, and to investigate the sensitivity of TROWET (wet delay) to near-surface humidity and precipitation under different atmospheric regimes.

Because the analysis is restricted to 24-hour windows ($N \approx 24$), correlations and regression slopes are interpreted as regime-specific diagnostics rather than climatological statistics.

The study considers seven European GNSS stations spanning different climatic and geographic settings, including coastal lowlands, inland plains, and mountainous environments (Chapter 3). This diversity allows identical diagnostics to be evaluated under markedly different atmospheric structures, including well-mixed boundary layers, shallow marine layers, vertically stratified moisture fields, and orographically forced precipitation [17, 18, 20].

Tropospheric delay is analysed through its two standard zenith components. The hydrostatic component (TRODRY), governed primarily by the vertically integrated mass of dry air, is expected to remain tightly linked to surface pressure across environments and regimes. The wet component (TROWET), which reflects the refractivity contribution of atmospheric water vapour, is a column-integrated quantity and is therefore expected to exhibit stronger dependence on atmospheric regime and site-specific representativeness of near-surface humidity.

For each station and day, Zenith tropospheric delays were obtained using the CSRS-PPP web application [16] and decomposed into hydrostatic (TRODRY) and wet (TROWET) components at a 30-second sampling interval. CSRS-PPP is treated here as a black-box precise positioning service providing ZTD estimates consistent with IGS standards. The GNSS estimates were converted from the CSRS epoch format (yy:DOY:sssss) to UTC and aggregated to hourly means to match ERA5 temporal resolution. CSRS-PPP zenith-delay components are provided in metres and are converted to millimetres (×1000) prior to reporting and regression, so that slopes are expressed in mm hPa^{-1} and mm $^{\circ}\text{C}^{-1}$. ERA5 variables were extracted at each station location and include 2 m temperature, 2 m dew point temperature, surface pressure, and hourly total precipitation [14, 15]. Simple linear models of the form: $TROWET = a \cdot T\_d + b$ and $TRODRY = a \cdot P\_s + b$ are used as diagnostic tools to characterise sensitivity and coherence over the selected 24-hour period. Reported statistics include Pearson correlation (r), regression slope (a), coefficient of determination (R^2), and RMSE of the residuals. These diagnostics are not intended to define universal station properties, but rather to evaluate physical consistency and regime dependence over the analysed day.

Precipitation is treated explicitly as an event-timing indicator rather than a direct proxy for instantaneous column water vapour. Where relevant, its relationship with TROWET is examined using short temporal lags (±1 h) to accommodate timing offsets between precipitation accumulation and GNSS-sensed moisture evolution (Chapter 4, Section 4.4.3).

Interpretation follows two guiding principles. First, TRODRY should remain pressure-controlled across all regimes, including disturbed synoptic conditions, because dry-air mass is well represented by surface pressure even during storms [5, 8]. Second, TROWET–dew point coupling should vary with the representativeness of the near-surface layer for the full moisture column, which may be reduced by vertical stratification, shallow marine boundary layers, sea–land breeze dynamics, or complex orography [12, 17, 18, 20]. Although ZWD represents the vertically integrated water vapour content of the entire atmospheric column, ERA5 2 m dew point temperature (d2m) is used here as a near-surface moisture proxy. The physical justification lies in the fact that, under well-mixed boundary layer conditions, surface humidity variations are strongly coupled with the lower tropospheric moisture profile, which typically dominates the total precipitable water content. Therefore, d2m can reasonably track short-term ZWD variability. However, this coupling weakens under stable stratification, complex orography, or decoupled humidity layers, which explains the reduced correlation observed at mountainous stations and during dry or convectively shallow regimes. Data sources and processing steps for TRODRY and TROWET are summarised in Chapter 4 (Sections 4.4–4.8).

5.2 Urban lowland basin: RUF01 (Florence region, Italy)

5.2.1 Dry summer regime (15 July 2024)

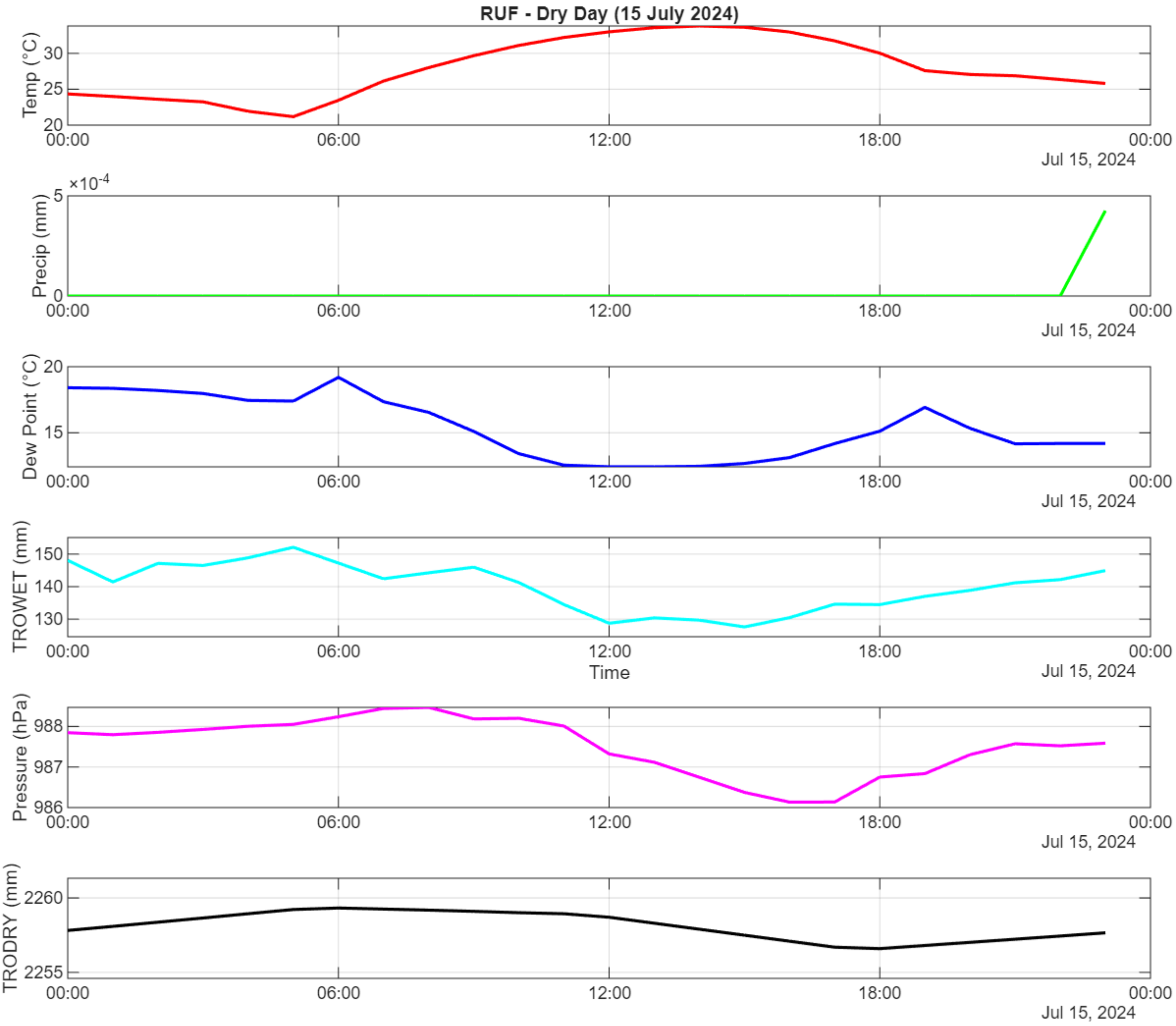

Figure 5.1. Hourly evolution of TRODRY, TROWET, ERA5 surface pressure, ERA5 dew point temperature, and precipitation at RUF01 for 15 July 2024.

Under dry summer conditions, RUF01 exhibits clear separation between hydrostatic and wet behaviour (Figure 5.1). TROWET shows strong coupling with 2 m dew point temperature ($r = 0.79$; $R^2 = 0.62$), indicating that near-surface humidity explains a substantial fraction of wet-delay variability over this 24-hour window. The regression sensitivity is moderate (2.49 mm $°C^{-1}$), lower than values observed at the coastal and plateau stations, which is consistent with an inland-basin setting in which surface humidity variations may reflect changes confined to a relatively shallow layer rather than proportional changes across the full moisture column [12, 17].

The hydrostatic component behaves as expected for a pressure-controlled delay: TRODRY is strongly associated with surface pressure ($r = 0.81$) with a near-linear response (slope = 1.05 mm hPa^{-1}), confirming physically consistent hydrostatic estimation under stable conditions (*Table 5.1*) [5].

Table 5.1. RUF01 results, dry day (15 July 2024)

Relationship	***r***	***Slope***	***R^2***	***RMSE***
TROWET vs dew point	0.79	2.49 mm °C^{-1}	0.62	4.37 mm
TRODRY vs pressure	0.81	1.05 mm hPa^{-1}	0.66	0.52 mm

5.2.2 Rainy synoptic regime (3 March 2024)

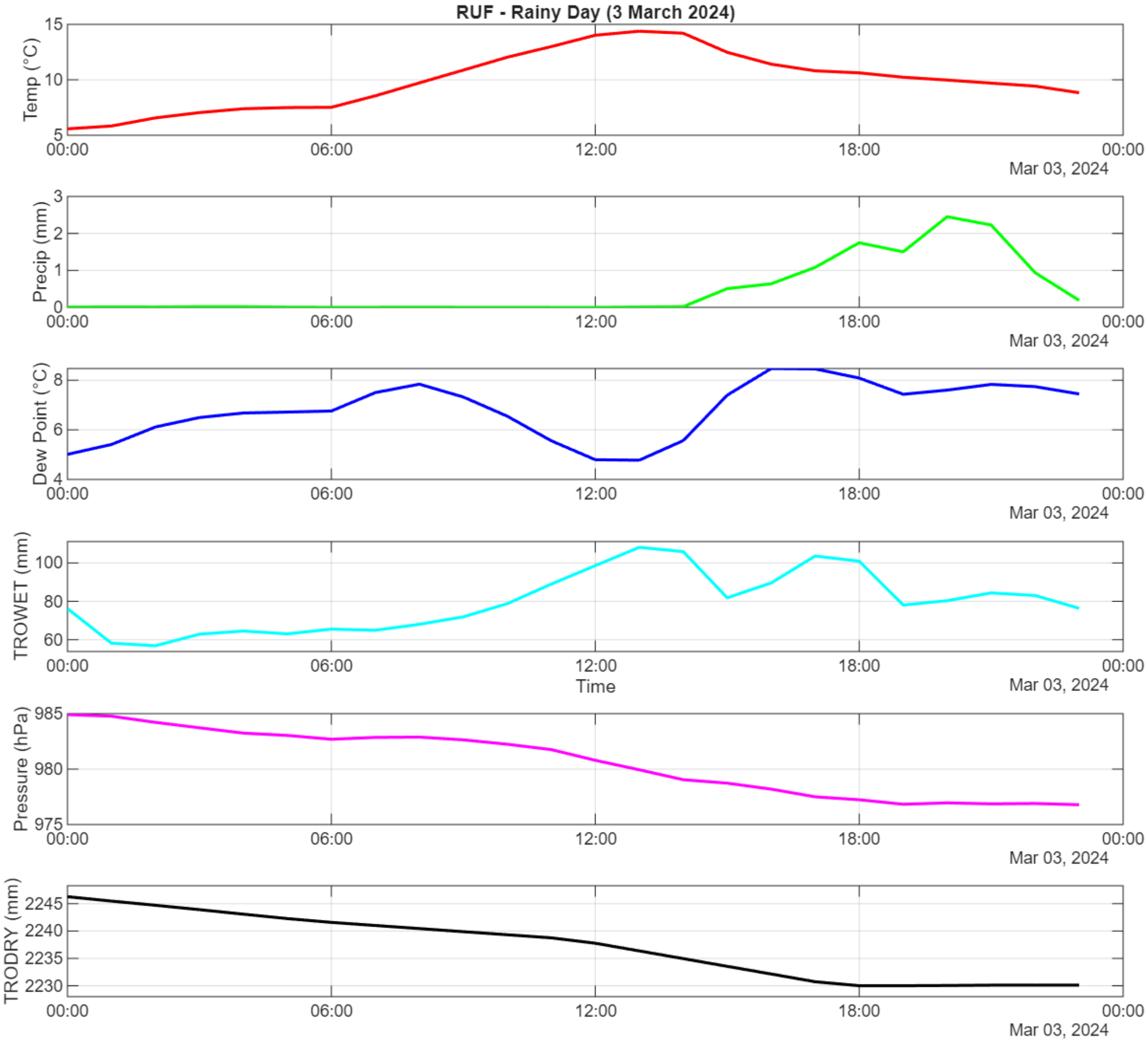

Figure 5.2. Hourly evolution of TRODRY, TROWET, ERA5 surface pressure, ERA5 dew point temperature, and precipitation at RUF01 for 3 March 2024.

During the rainy regime, the TRODRY–pressure relationship remains extremely strong (r = 0.99; slope = 1.96 mm hPa^{-1}), confirming that hydrostatic delay behaviour is robust under synoptic disturbance (*Table 5.2*) [5, 10]. In contrast, TROWET shows no meaningful linear relationship with dew point over this 24-hour window (r = -0.04; slope = -0.50 mm °C^{-1}; R^2 ≈

0.00). This indicates that, for this event at RUF01, near-surface dew point is not a reliable proxy for the column-integrated moisture variability sensed by GNSS, consistent with regime-dependent vertical structure and/or timing effects between low-level humidity and precipitation processes [10, 12, 17].

Table 5.2. RUF01 results, rainy day (3 March 2024)

Relationship	***r***	***Slope***	***R²***	***RMSE***
TROWET vs dew point	-0.04	-0.50 mm $°C^{-1}$	≈ 0.00	15.16 mm
TRODRY vs pressure	0.99	1.96 mm hPa^{-1}	0.99	0.60 mm

5.3 High mountain environment: SCV1 (Alpine region, ~2329 m)

5.3.1 Dry summer regime (15 July 2024)

At this high-altitude site, TRODRY shows weaker but still coherent dependence on surface pressure ($r = 0.58$; $R^2 = 0.34$). The reduced correlation relative to lowland stations is consistent with limited daily pressure range and enhanced representativeness errors in reanalysis pressure fields over complex terrain [14, 15, 18].

TROWET remains positively related to dew point ($r = 0.68$; $R^2 = 0.46$) with a steep sensitivity (9.57 mm $°C^{-1}$), indicating that even modest surface humidity variability at high elevation can correspond to substantial changes in the integrated wet delay under fair-weather conditions [12].

Table 5.3: SCV1 results, dry day (15 July 2024)

Relationship	***r***	***Slope***	***R²***	***RMSE***
TROWET vs dew point	0.68	9.57 mm $°C^{-1}$	0.46	7.89 mm
TRODRY vs pressure	0.58	1.48 mm hPa^{-1}	0.34	0.44 mm

5.3.2 Rainy regime (3 March 2024)

During the rainy day, SCV1 exhibits strong regime dependence: TROWET becomes negatively associated with dew point ($r = -0.35$; slope = -1.43 mm $°C^{-1}$), consistent with conditions where column moisture, condensation, and vertical redistribution evolve under orographic forcing while the near-surface humidity proxy varies independently [18]. TRODRY remains coherently linked to pressure ($r = 0.72$; $R^2 = 0.52$).

Precipitation timing is assessed using the ±1 h lag correlations reported in the cross-station summary (*Tables 5.12–5.13*). At SCV1 on the rainy day, precipitation correlations are negative and reach their largest magnitude at +1 h ($r = -0.83$), consistent with event timing and moisture-phase changes under orographic forcing.

Table 5.4. SCV1 results, rainy day (3 March 2024).

Relationship	*r*	*Slope*	*R²*	*RMSE*
TROWET vs dew point	-0.35	-1.43 mm °C⁻¹	0.13	3.57 mm
TRODRY vs pressure	0.72	1.12 mm hPa⁻¹	0.52	1.44 mm

5.4 Plateau environment: BTCHZ (Swiss plateau)

5.4.1 Dry summer regime (15 July 2024)

BTCHZ provides one of the clearest examples of strong surface representativeness under stable conditions. TROWET correlates extremely strongly with dew point ($r = 0.93$; $R^2 = 0.86$) and shows a steep sensitivity (15.77 mm $°C^{-1}$), consistent with a well-mixed boundary layer over relatively homogeneous terrain. TRODRY shows coherent pressure control ($r = 0.73$; slope = 0.95 mm hPa^{-1}).

Table 5.5. BTCHZ results, dry day (15 July 2024)

Relationship	*r*	*Slope*	*R²*	*RMSE*
TROWET vs dew point	0.93	15.77 mm °C⁻¹	0.86	6.65 mm
TRODRY vs pressure	0.73	0.95 mm hPa⁻¹	0.53	1.06 mm

5.4.2 Rainy regime (3 March 2024)

Under synoptic rainfall, dew point loses explanatory power for TROWET ($r = 0.08$; $R^2 = 0.01$), while precipitation timing becomes informative with short lags ($r = 0.64$–0.70, peaking at +1 h). This pattern is consistent with synoptic forcing, where moisture evolution and rainfall accumulation are linked but not necessarily aligned at lag 0. TRODRY remains strongly pressure-driven ($r = 0.96$; $R^2 = 0.91$).

Table 5.6. BTCHZ results, rainy day (3 March 2024).

Relationship	*r*	*Slope*	*R²*	*RMSE*
TROWET vs dew point	0.08	0.98 mm °C⁻¹	0.01	8.47 mm
TRODRY vs pressure	0.96	1.33 mm hPa⁻¹	0.91	1.11 mm

5.5 Inland lowland environment: STAVE (Belgium)

STAVE provides a lowland reference case where hydrostatic control is exceptionally stable across regimes, while wet-delay behaviour remains only moderately predictable from surface moisture.

5.5.1 Dry summer regime (15 July 2024)

At STAVE, hydrostatic control is exceptionally stable: TRODRY correlates nearly perfectly with pressure (r = 0.98; slope = 1.79 mm hPa^{-1}). TROWET shows only moderate coupling with dew point (r = 0.50), and the large RMSE (26.54 mm) indicates substantial residual variability around the linear fit despite a moderate correlation, consistent with the fact that dew point is a near-surface proxy while TROWET integrates the full moisture column.

Table 5.7. STAVE results, Dry day (15 July 2024)

Relationship	***r***	***Slope***	***R²***	***RMSE***
TROWET vs dew point	0.50	6.99 mm °C^{-1}	0.25	26.54 mm
TRODRY vs pressure	0.98	1.79 mm hPa^{-1}	0.97	0.84 mm

5.5.2 Rainy synoptic regime (3 March 2024)

On the rainy day, hydrostatic control remains extremely strong (r = 0.99; slope = 2.21 mm hPa^{-1}). TROWET again shows only moderate coupling with dew point (r = 0.44), with lower RMSE than on the dry day, suggesting more coherent moisture evolution under synoptic forcing but still limited predictability from a surface proxy alone.

Table 5.8. STAVE results, rainy day (3 March 2024).

Relationship	***r***	***Slope***	***R²***	***RMSE***
TROWET vs dew point	0.44	8.86 mm °C^{-1}	0.19	10.18 mm
TRODRY vs pressure	0.99	2.21 mm hPa^{-1}	0.99	0.87 mm

5.6 Coastal environments: SOPH (Mediterranean) and PYLA (Atlantic)

5.6.1 SOPH (Mediterranean coast)

The SOPH station represents a Mediterranean coastal environment, where tropospheric variability is influenced by the proximity of the sea, frequent marine air masses, and pronounced boundary-layer regime changes between dry summer conditions and rainy synoptic events. These characteristics make SOPH a particularly informative site for assessing regime-dependent behaviour of GNSS-derived tropospheric delays.

On the dry summer day (15 July 2024), TRODRY at SOPH shows only small absolute variability because ERA5 surface pressure also varies weakly over this 24-hour window. Under such a limited predictor dynamic range (N = 24), Pearson correlation becomes sensitive to noise and representativeness differences, and the slightly negative TRODRY–pressure correlation should be treated as non-diagnostic rather than physically meaningful. The hydrostatic behaviour is therefore interpreted primarily through station-days with sufficient pressure variability, where the expected near-linear dependence is consistently recovered (e.g., *Tables 5.1–5.8* and *Table 5.13*).

The wet component (TROWET) at SOPH during the dry day shows a notably weaker and less coherent relationship with near-surface dew point temperature compared to other coastal and lowland stations. This reduced sensitivity is consistent with the presence of a shallow and stratified marine boundary layer under summer conditions, in which near-surface humidity does not fully represent the vertically integrated water vapour content. Local sea-breeze circulations and diurnal boundary-layer decoupling likely contribute to this behaviour by introducing vertical moisture gradients that weaken the coupling between surface indicators and column-integrated wet delay [20].

In contrast, during the rainy winter day (3 March 2024), SOPH exhibits a markedly different response. TRODRY displays a strong positive correlation with surface pressure, consistent with hydrostatic theory and enhanced synoptic-scale pressure variability during winter. At the same time, TROWET shows a clear and physically consistent positive relationship with ERA5 dew point temperature. This strengthened coupling reflects a deeper, more vertically homogeneous boundary layer during synoptic precipitation events, under which near-surface humidity becomes a more reliable proxy for column-integrated water vapour.

The SOPH results demonstrate that, at coastal Mediterranean sites, the interpretability of GNSS-derived wet delay is strongly regime-dependent. Under dry summer conditions, shallow boundary layers and local circulations can decouple surface meteorology from the integrated atmospheric column, whereas under rainy synoptic regimes, the physical linkage between TROWET and humidity is restored. This behaviour highlights the necessity of regime-aware interpretation when using GNSS tropospheric products in coastal environments.

Table 5.9. SOPH results (15 July 2024; 3 March 2024).

Relationship	***r***	***Slope***	***R²***	***RMSE***
TROWET vs dew point	-0.06	-1.44 mm °C⁻¹	≈ 0.00	14.50 mm
TRODRY vs pressure	-0.23	-0.28 mm hPa⁻¹	0.05	0.48 mm

Relationship	***r***	***Slope***	***R²***	***RMSE***
TROWET vs dew point	0.87	11.87 mm °C⁻¹	0.77	8.42 mm
TRODRY vs pressure	0.95	1.39 mm hPa⁻¹	0.90	1.09 mm

5.6.2 PYLA (Atlantic coast)

PYLA shows consistently strong hydrostatic behaviour and strong wet coupling during dry conditions. On the dry day, TROWET correlates strongly with dew point ($r = 0.85$; $R^2 = 0.72$) and exhibits a steep sensitivity (15.38 mm °C^{-1}), consistent with an ocean-influenced boundary layer in which surface dew point variability reflects substantial changes in column moisture.

On the rainy day, the dew point relationship weakens (r = 0.27; R² = 0.07; slope = 1.40 mm °C⁻¹), consistent with synoptic regime complexity and vertical humidity redistribution. TRODRY remains nearly perfectly pressure-controlled on both days (r = 0.99 dry and rainy), confirming the stability of hydrostatic control across regimes at this coastal site.

Table 5.10. PYLA results (15 July 2024; 3 March 2024).

Relationship	***r***	***Slope***	***R²***	***RMSE***
TROWET vs dew point	0.85	15.38 mm °C⁻¹	0.72	14.22 mm
TRODRY vs pressure	0.99	1.77 mm hPa⁻¹	0.99	0.68 mm

Relationship	***r***	***Slope***	***R²***	***RMSE***
TROWET vs dew point	0.27	1.40 mm °C⁻¹	0.07	5.03 mm
TRODRY vs pressure	0.99	2.11 mm hPa⁻¹	0.99	0.98 mm

5.7 Mountain and complex terrain: AIGL (Lozère mountains, France, ~1619 m)

AIGL provides a second complex-terrain test case that differs from SCV1 in setting but shows similarly strong regime dependence in wet-delay interpretability.

5.7.1 Dry summer regime (15 July 2024)

On the dry day, TROWET is only weakly related to dew point (r = 0.12; R² = 0.02), indicating limited surface representativeness for the integrated moisture column under summer conditions in complex terrain. TRODRY remains strongly pressure-controlled (r = 0.89; slope = 2.56 mm hPa⁻¹).

5.7.2 Rainy synoptic regime (3 March 2024)

On the rainy day, TROWET–dew point coupling strengthens substantially (r = 0.84; R² = 0.71) with a moderate sensitivity (3.27 mm °C⁻¹), consistent with reduced column depth at elevation and saturation/condensation processes limiting vapour growth. TRODRY remains very strongly pressure-driven (r = 0.99).

Table 5.11. AIGL results for dry and rainy regimes (15 July 2024; 3 March 2024).

Relationship	***r***	***Slope***	***R²***	***RMSE***
TROWET vs dew point	0.12	0.98 mm °C⁻¹	0.02	9.25 mm
TRODRY vs pressure	0.89	2.56 mm hPa⁻¹	0.79	0.92 mm

Relationship	*r*	*Slope*	*R²*	*RMSE*
TROWET vs dew point	0.84	3.27 mm °C⁻¹	0.71	3.33 mm
TRODRY vs pressure	0.99	1.91 mm hPa⁻¹	0.98	0.95 mm

5.8 Cross-station summary table

Tables 5.12 and *5.13* consolidate the station-level diagnostics into a single cross-station view for the dry and rainy regimes, respectively. For each station-day, the tables report: (i) the correlation between TROWET and ERA5 2 m dew point temperature (d2m) together with the fitted sensitivity (mm °C⁻¹); (ii) the correlation between TRODRY and ERA5 surface pressure (sp) together with the fitted sensitivity (mm hPa⁻¹); and (iii) correlations between TROWET and ERA5 hourly total precipitation (tp) evaluated at short lags (-1 h, 0 h, +1 h) where precipitation variability supports a meaningful lag test.

These summaries support the synthesis in Sections 5.9–5.10 by separating the behaviour expected to be robust across regimes (hydrostatic pressure control) from behaviour that is inherently regime- and environment-dependent (wet-delay coupling and precipitation timing). Where precipitation is negligible, nearly constant, or dominated by isolated single-hour spikes within the 24-hour window, the lag-correlation test is not statistically meaningful; in such cases, all precipitation-lag entries for that station-day are reported as "—".

Table 5.12. Cross-station summary of hourly diagnostics (dry day)

Station	*r_TROWET d2m*	*r_tp lag0*	*r_tp Lag+1h*	*r_tp Lag-1h*	*r_TRODRY sp*	*Slope_TROWET mm °C⁻¹*	*Slope_TRODRY mm hPa⁻¹*
AIGL	0.12	0.04	0.02	-0.04	0.89	0.98	2.56
BTCHZ	0.93	0.69	0.59	0.77	0.73	15.77	0.95
PYLA	0.85	0.49	0.19	0.72	0.99	15.38	1.77
RUF	0.79	—	—	—	0.81	2.49	1.05
SCV1	0.68	0.31	0.39	0.12	0.58	9.57	1.48
SOPH	-0.06	0.42	0.55	0.15	-0.23	-1.44	-0.28
STAVE	0.50	0.29	0.14	0.49	0.98	6.99	1.79

Table 5.13. Cross-station summary of hourly diagnostics (rainy day)

Station	*r_TROWET d2m*	*r_tp lag0*	*r_tp Lag+1h*	*r_tp Lag-1h*	*r_TRODRY sp*	*Slope_TROWET mm °C⁻¹*	*Slope_TRODRY mm hPa⁻¹*
AIGL	0.84	0.36	0.28	0.49	0.99	3.27	1.91
BTCHZ	0.08	0.69	0.70	0.64	0.96	0.98	1.33
PYLA	0.27	0.30	-0.20	0.45	0.99	1.40	2.11
RUF	-0.04	0.30	0.20	0.37	0.99	-0.50	1.96
SCV1	-0.35	-0.69	-0.83	-0.57	0.72	-1.43	1.12
SOPH	0.87	0.52	0.33	0.65	0.95	11.87	1.39
STAVE	0.44	0.46	0.45	0.47	0.99	8.86	2.21

5.9 Cross-station synthesis

Across the full station set, TRODRY behaves as a robust, pressure-dominated observable across environments and regimes [5, 8]. Strongest correlations generally occur on the rainy synoptic day, when pressure variability is typically larger; weaker correlations occur primarily in complex terrain or short windows with limited pressure range, reflecting representativeness and sampling limitations rather than loss of physical meaning [14, 15].

In contrast, TROWET exhibits pronounced dependence on both atmospheric regime and site characteristics [12, 17, 18, 20]. Strong coupling with dew point emerges when the near-surface layer is broadly representative of the moisture column (e.g., plateau and Atlantic-coastal dry regime), whereas weak or negative coupling occurs when representativeness is degraded by vertical stratification, shallow marine boundary layers, local coastal circulations, or orographic forcing. Negative or weak correlations between TROWET and near-surface humidity do not imply physically negative moisture sensitivity, but reflect decoupling between surface indicators and the vertically integrated moisture column under specific regimes. Precipitation-based diagnostics show that wet delay is better interpreted as a process-sensitive indicator than as a universal linear predictor; where lagged relationships appear, they tend to be secondary compared with humidity control and are sensitive to hourly accumulation timing.

5.10 Key findings from the case study

Across all stations and regimes, the hydrostatic component (TRODRY) exhibits behaviour consistent with its theoretical dependence on surface pressure (Chapter 2, Section 2.4). When pressure variability within the 24-hour window is sufficient, correlations are strong and regression slopes fall within the expected mm hPa^{-1} range. Apparent anomalies, such as weak or negative correlations over short windows, are attributable to limited predictor range and short-sample effects rather than a physical inversion.

The wet component (TROWET) shows strong spatial and regime dependence. A strong positive coupling with dew point temperature is observed in settings and regimes where near-surface humidity is more representative of the atmospheric column, such as during dry conditions at BTCHZ and PYLA. In contrast, decoupling occurs under coastal Mediterranean summer conditions and in complex terrain subject to stratified or orographically forced regimes.

Hourly precipitation is not a direct proxy for column water vapour, and its zero-lag correlation with TROWET is generally weak. Where lag effects are present, they indicate timing relationships between moisture evolution and rainfall accumulation; however, these effects remain subordinate to the primary controls exerted by humidity and pressure and are sensitive to event structure and the timing of hourly accumulation.

Overall, the multi-station case study demonstrates that PPP-derived tropospheric delays behave in a manner that is physically consistent with atmospheric theory while remaining sensitive to environmental setting and synoptic regime. The results confirm that GNSS tropospheric products provide robust and interpretable atmospheric information when analysed with appropriate awareness of boundary-layer structure, topography, and temporal context.

CHAPTER 6: Long-Term Stability Analysis of the Low-Cost GNSS Station RUF01

6.1 Chapter Scope and Analytical Objectives

The objective of this chapter is to evaluate the long-term stability, reliability, and physical interpretability of tropospheric delays estimated from a low-cost GNSS receiver using Precise Point Positioning (PPP). While the multi-station case study in Chapter 5 demonstrated that PPP-derived tropospheric components behave consistently with atmospheric physics under controlled short-term conditions, long-term monitoring applications impose more stringent requirements. These include temporal stability, robustness against processing artefacts, and resistance to occasional positioning failures.

The analysis focuses on the low-cost station RUF01, located in the Florence region, and spans a continuous two-month period from 1 March to 30 April 2024. To separate receiver-related effects from processing-related effects, RUF01 is processed using two independent precise-product strategies. In parallel, a nearby geodetic EUREF station, IGMI00ITA, is processed using identical PPP settings and serves two complementary roles: (i) as a processing control that isolates hardware effects under a common estimator, and (ii) as an external reference through published EUREF tropospheric products (hereafter IGMI_EUREF).

The chapter addresses four interlinked questions.

1. How stable and continuous is the PPP-derived ZTD time series at the low-cost station RUF01 over the 61-day period?
2. To what extent do different precise products (CDDIS vs GFZ) affect the magnitude and variability of ZTD at both RUF01 and IGMI00ITA?
3. Are the GNSS-derived ZTD series physically consistent with the independent meteorological information from ERA5?
4. Does PPP positioning stability support the interpretation that the tropospheric estimates are reliable?

The analysis builds directly on the PPP processing strategy, quality-control pipeline and positioning-residual methodology detailed in Chapter 4, and applies them in a focused way to the RUF01–IGMI pair.

6.2 Data Set and Processing Overview

RUF01 is a low-cost dual-frequency GNSS receiver located in the Florence area at an ellipsoidal height of approximately 249 m; IGMI00ITA is a collocated high-grade EUREF station at about 95 m ellipsoidal height. Both stations are within a few tens of kilometres of each other and sample a very similar atmospheric column, which is favourable for differential tropospheric comparison.

The stability analysis is based on continuous GNSS observations collected at RUF01 and IGMI00ITA at a 30-second sampling interval over 61 consecutive days. RUF01 data (Centipede network) are provided as compact RINEX (CRX), converted to standard RINEX, and IGMI00ITA observations were downloaded from the EUREF archive. All raw observations are processed internally using the demo5 version of RTKLIB in static PPP mode, with identical configuration settings that were adopted in Chapter 4 (Section 4.4), ionosphere-free linear combinations, an elevation mask of 10°, the Niell Mapping Function (NMF), precise orbits, clocks and Earth rotation parameters. This controlled setup ensures that observed differences in ZTD behaviour can be attributed to receiver characteristics or precise-product strategy rather than estimator configuration.

Two parallel PPP processing branches are implemented [26, 27]. The first branch is based on PPP driven by NASA CDDIS multi-GNSS precise products, including BRDM navigation files, SP3 orbits, CLK clocks, and associated ERP and bias files. The second branch is driven by GFZ final multi-GNSS precise products, using SP3 orbits, CLK clocks, and ERP, while applying the same RTKLIB configuration.

In addition to the PPP-derived ZTD time series, IGMI_EUREF is taken from the EUREF SINEX-TRO solution (analysis centre ASI0EPNFIN), providing hourly ZTD for IGMI00ITA over the study period and serving as an external reference for PPP ZTD. ERA5 hourly reanalysis data for the Florence grid cell are used as independent meteorological information to assess the meteorological consistency of all ZTD series.

Tropospheric parameters are extracted from the PPP .pos.stat files following the $TROP parsing procedure described in Section 4.6.1. For the CDDIS processing branch, 151,541 30-s ZTD epochs at RUF01 pass the basic physical and formal-error screening, with 702 epochs flagged as outliers and replaced by the rolling-median filter. At IGMI00ITA, 165,628 epochs pass the same basic screening, and 655 epochs are replaced by the filter.

For the GFZ processing branch, the same quality-control thresholds and filtering strategy are applied. At RUF01, 149,142 epochs are parsed, of which 692 are replaced, while at IGMI00ITA, 165,503 epochs are parsed and 391 epochs are replaced. These quality-control statistics ensure that the subsequent stability analysis is based on internally consistent ZTD time series across both stations and processing strategies.

6.3 ZTD Time Series Generation and Quality Control

The transformation from QC'd 30-s ZTD estimates to analysis-ready time series follows the generic pipeline defined in Sections 4.6.1–4.6.3, ensuring that differences between stations or precise-product strategies are not contaminated by inconsistent processing.

Epoch-level screening removes physically implausible ZTD values, using conservative mid-latitude bounds consistent with Chapter 4, together with epochs whose formal ZTD standard deviation exceeds the adopted threshold. To suppress the impact of transient PPP disturbances (cycle slips, poor geometry, temporary loss of convergence), each series is then filtered with a rolling-median and median absolute deviation (MAD) scheme. Outliers identified by this robust filter are replaced by the local rolling median rather than discarded [30]. This preserves regular 30-second sampling and prevents gaps or sampling biases in the subsequent hourly means.

Hourly ZTD values are computed on a forced grid centred at :30 UTC, requiring at least 30 valid 30-second epochs per hour. All five series are mapped onto the same 1464-epoch grid spanning 1 March–30 April 2024, and ERA5 timestamps are shifted accordingly so that meteorological variables refer to the centre of each ZTD aggregation window. After quality control and aggregation, five ZTD time series are available for analysis on this common grid, RUF_CDDIS, RUF_GFZ, IGMI_CDDIS, IGMI_GFZ and IGMI_EUREF. Their availability on the 1464-epoch grid is 1420, 1316, 1402, 1388 and 1416 hourly values, respectively.

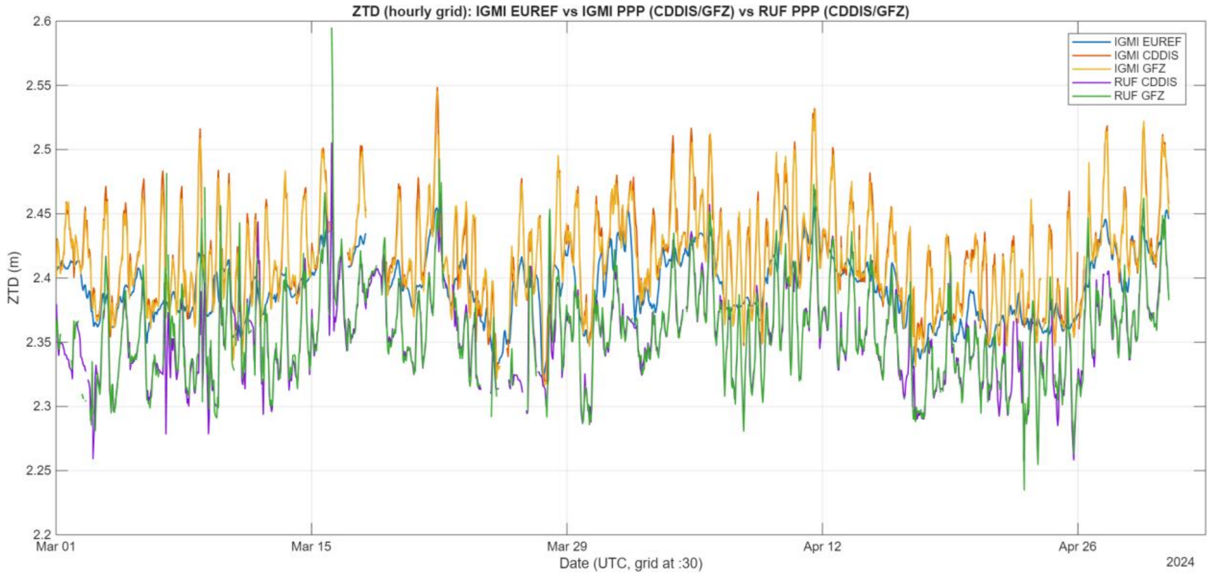

Figure 6.1. Hourly ZTD time series on the fixed :30 UTC grid for all five solutions (IGMI_EUREF, IGMI_CDDIS, IGMI_GFZ, RUF_CDDIS, RUF_GFZ) over 1 March–30 April 2024.

Figure 6.1 shows the resulting hourly ZTD series. Despite the small differences in data availability, the time series exhibit a high degree of temporal coherence across stations and processing strategies, seasonal evolution, synoptic-scale fluctuations and diurnal oscillations are shared by all five solutions. Absolute offsets between stations and between CDDIS and GFZ branches are clearly visible but remain remarkably stable over time, indicating systematic rather than random discrepancies.

6.4 Intercomparison of ZTD Time Series

6.4.1 Comparison Against the EUREF Reference

The primary validation step compares the PPP-derived ZTD series against the EUREF reference for IGMI00ITA. *Table 6.1* summarises the pairwise statistics between each PPP-derived series and the IGMI_EUREF reference on the common hourly grid, the number of paired hourly epochs available after synchronising the two series being compared (N), mean bias, RMS, standard deviation and Pearson correlation.

For RUF01 processed with CDDIS products, the ZTD time series exhibits a mean bias of -39.2 mm relative to IGMI_EUREF, with an RMS difference of 48.1 mm and a correlation coefficient r = 0.63. The GFZ-based RUF01 series shows a very similar behaviour, with a bias of -37.1 mm, RMS 46.9 mm and r = 0.63. This indicates that, for the low-cost station, both the magnitude and temporal structure of the bias are largely insensitive to the precise-product strategy.

For IGMI00ITA processed in PPP mode, positive biases of +24.1 mm (CDDIS) and +22.4 mm (GFZ) are observed relative to the EUREF product, with RMS differences of 35.9 mm and 33.9 mm and correlations r = 0.71 and r = 0.72, respectively. These values are typical of differences between single-station PPP estimates and combined network solutions and confirm that PPP-derived ZTD for a geodetic station closely follows the reference both in magnitude and temporal behaviour.

The sign and magnitude of the RUF01 bias are physically consistent with its higher elevation relative to IGMI00ITA, which reduces the atmospheric column above the antenna. Importantly, the bias remains stable over the entire two-month period, and the correlation values indicate that temporal variability is largely preserved despite the systematic offset.

Table 6.1: ZTD comparison against IGMI_EUREF (hourly, March–April 2024)

Series vs IGMI_EUREF	***N (hourly)***	***Bias (mm)***	***RMS (mm)***	***Corr r***
RUF_CDDIS	1,372	-39.2	48.1	0.63
RUF_GFZ	1,276	-37.1	46.9	0.63
IGMI_CDDIS	1,400	+24.1	35.9	0.71
IGMI_GFZ	1,388	+22.4	33.9	0.72

The statistics confirm that the low-cost station RUF01 produces a stable, physically interpretable ZTD signal whose systematic offset is mainly controlled by station height, while the geodetic station IGMI00ITA shows tighter agreement with the EUREF reference as expected.

6.4.2 Internal Consistency and Processing-Strategy Sensitivity

Internal consistency between processing strategies is assessed by comparing, at each station separately, the CDDIS- and GFZ-based PPP ZTD series. *Table 6.2* lists matching-epoch counts, mean bias, RMS and standard deviation of the differences, robust MAD and correlation.

For RUF01, the difference between CDDIS- and GFZ-based series has a mean bias of -1.54 mm, an RMS difference of 10.47 mm and a correlation r = 0.962. For IGMI00ITA, the corresponding bias is +1.72 mm, RMS 7.01 mm and r = 0.984. In both cases, the correlation is extremely high and the RMS difference is an order of magnitude smaller than the absolute ZTD values, indicating that the temporal structure of the atmospheric signal is essentially unchanged when switching precise products.

Table 6.2. Internal consistency between PPP processing strategies (CDDIS vs GFZ)

Station	***Series A***	***Series B***	***N (hourly)***	***Bias A−B (mm)***	***RMS (mm)***	***Corr r***
RUF01	RUF_CDDIS	RUF_GFZ	1,312	-1.54	10.47	0.962
IGMI00ITA	IGMI_CDDIS	IGMI_GFZ	1,387	+1.72	7.01	0.984

Figures 6.2 and *6.3* visualise these results.

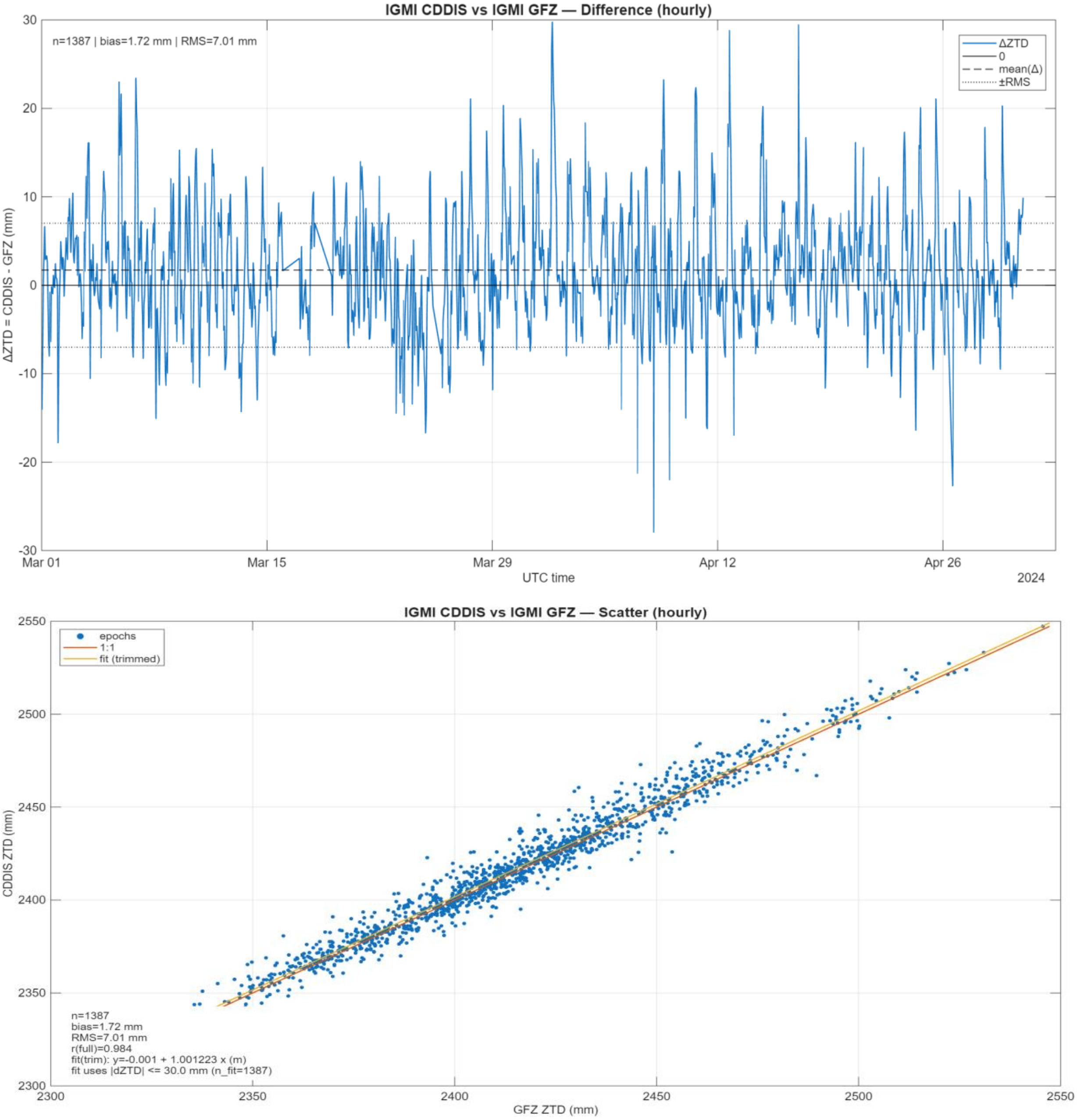

Figure 6.2. IGMI00ITA, internal comparison of CDDIS and GFZ hourly ZTD: (a) time series of CDDIS–GFZ differences with mean and ±RMS; (b) scatter plot of CDDIS vs GFZ ZTD with 1:1 line and trimmed linear fit.

For IGMI00ITA (*Figure 6.2*), the hourly differences CDDIS–GFZ remain mostly within ±10 mm, with only a small, nearly constant positive bias. The scatter plot of CDDIS vs GFZ ZTD is tightly clustered around the 1:1 line, with a trimmed regression slope very close to unity (≈ 1.00). This behaviour confirms that, for a geodetic-grade station, different precise-product strategies yield virtually indistinguishable tropospheric signals apart from a small quasi-stationary offset.

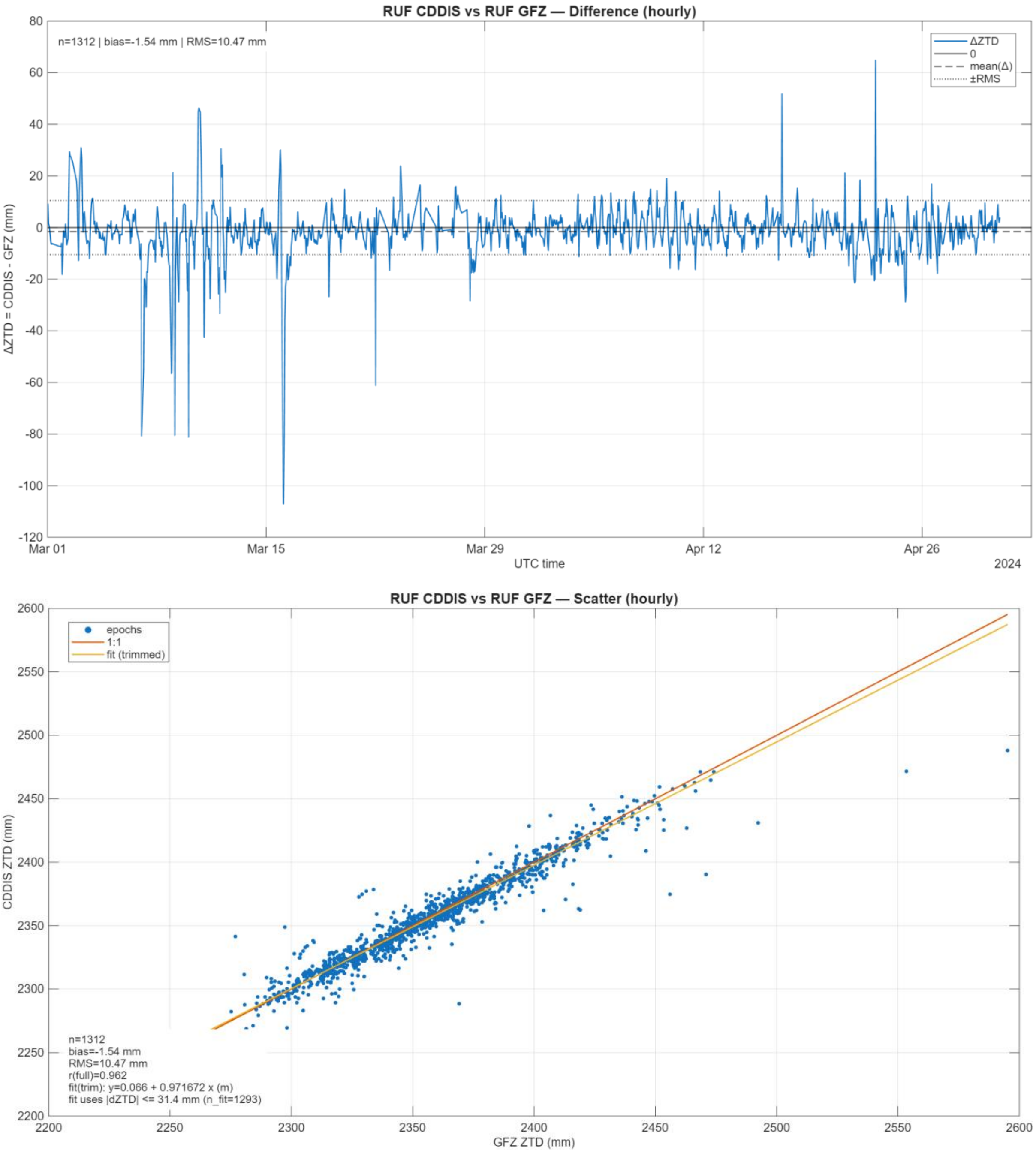

Figure 6.3. RUF01, internal comparison of CDDIS and GFZ hourly ZTD: (a) time series of CDDIS–GFZ differences with mean and ±RMS; (b) scatter plot of CDDIS vs GFZ ZTD with 1:1 line and trimmed linear fit.

For RUF01 (*Figure 6.3*), CDDIS–GFZ differences exhibit somewhat larger dispersion, with occasional excursions up to a few centimetres. Nevertheless, the mean bias remains small (-1.54 mm), the RMS difference is still limited (10.47 mm), and the correlation remains very high (r = 0.962). The scatter plot again shows points closely aligned with the 1:1 line, with a trimmed regression slope slightly below unity (≈ 0.97), indicating a very minor compression of the GFZ range relative to CDDIS.

Taken together, *Table 6.2* and *Figures 6.2–6.3* demonstrate that the choice between CDDIS and GFZ precise products introduces small, slowly varying differences relative to absolute ZTD magnitude and does not materially alter the temporal structure of the atmospheric signal; the dominant persistent offsets in the RUF–IGMI comparisons remain consistent with station geometry and systematic modelling differences.

6.5 Meteorological Consistency of Long-Term ZTD Series

The physical plausibility of the long-term ZTD series is assessed through comparison with ERA5 near-surface meteorological variables for the Florence region: 2-m temperature, 2-m dew point, surface pressure and hourly total precipitation. The analysis follows the general framework of Section 4.8 but is specialised here to the five ZTD series introduced above.

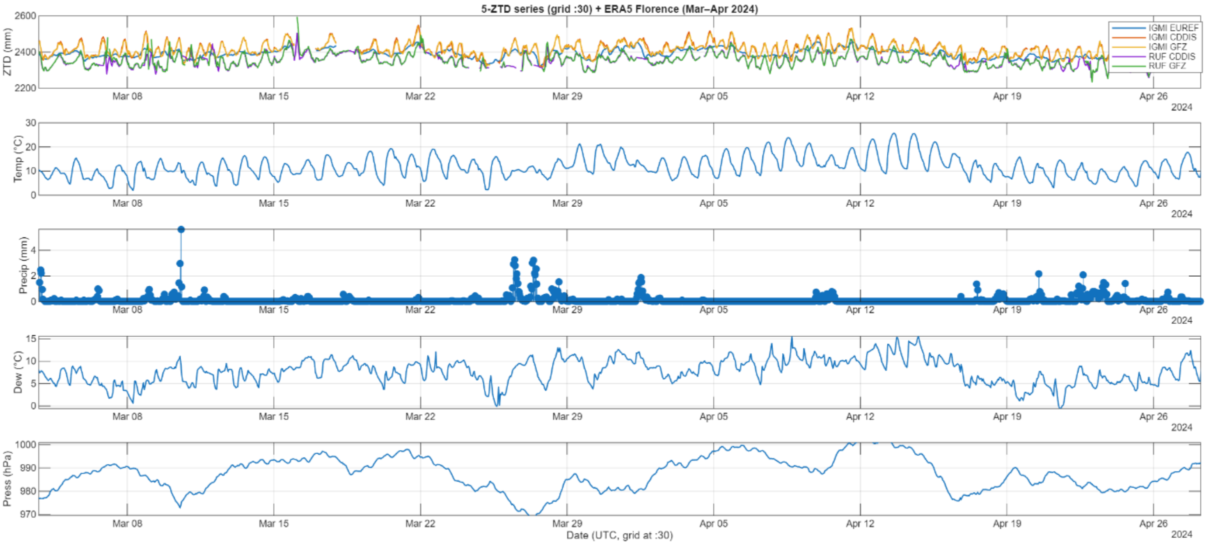

Figure 6.4. Time series of the five hourly ZTD series (IGMI_EUREF, IGMI_CDDIS, IGMI_GFZ, RUF_CDDIS, RUF_GFZ) together with ERA5 temperature, dew point, precipitation and surface pressure over March–April 2024.

Figure 6.4 shows the joint evolution of the five hourly ZTD series with ERA5 temperature, precipitation, dew point and surface pressure over March–April 2024. ZTD variations at both stations and in both PPP strategies track the combined influence of synoptic systems, diurnal temperature cycles and slower seasonal trends. Periods of enhanced ZTD coincide with warm, moist conditions and lower surface pressure, while marked ZTD minima occur during cooler, drier and higher-pressure episodes. Precipitation events tend to occur near, but not always exactly at, local ZTD maxima, reflecting the complex timing of convective and frontal processes relative to the integrated water vapour field.

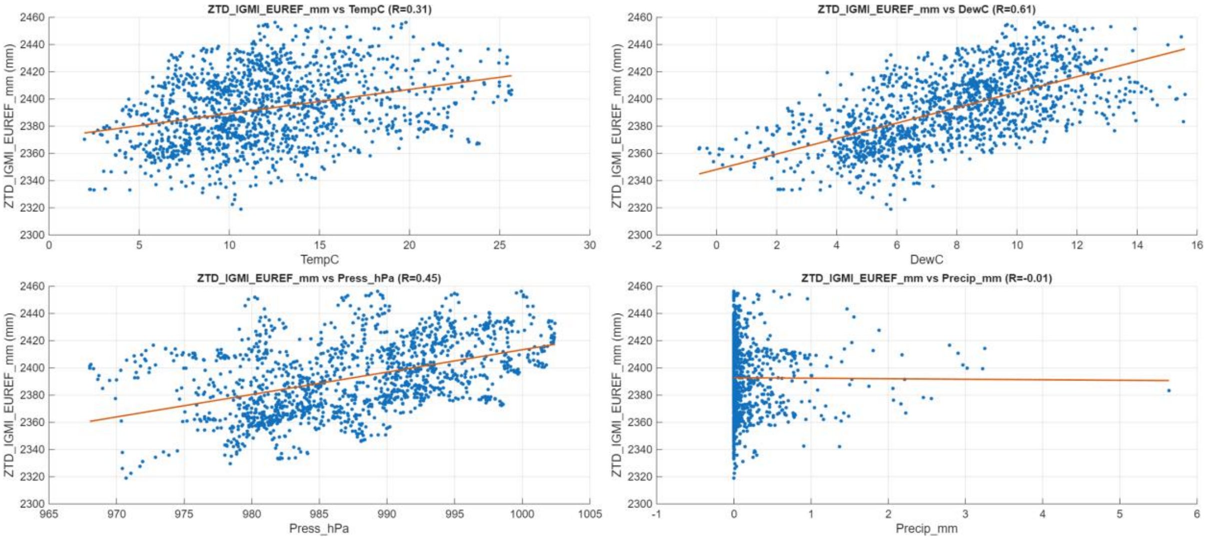

Figure 6.5. Scatter plots of IGMI_EUREF ZTD versus ERA5 near-surface temperature, dew point, surface pressure and hourly precipitation, with linear regression fits and correlation coefficients.

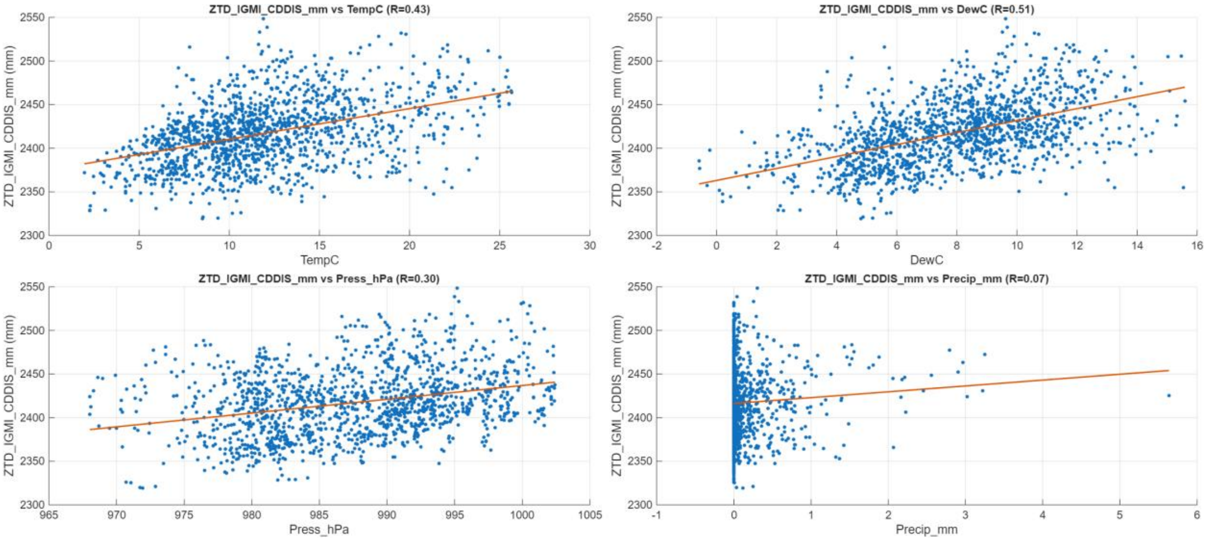

Figure 6.6. As Figure 6.5, but for IGMI_CDDIS ZTD.

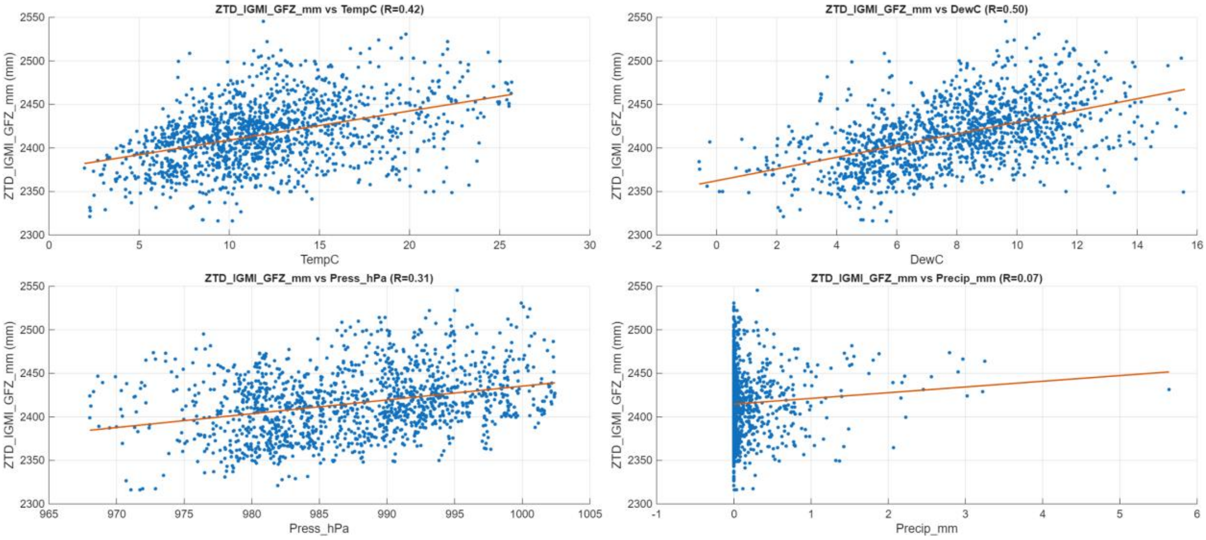

Figure 6.7. As Figure 6.5, but for IGMI_GFZ ZTD.

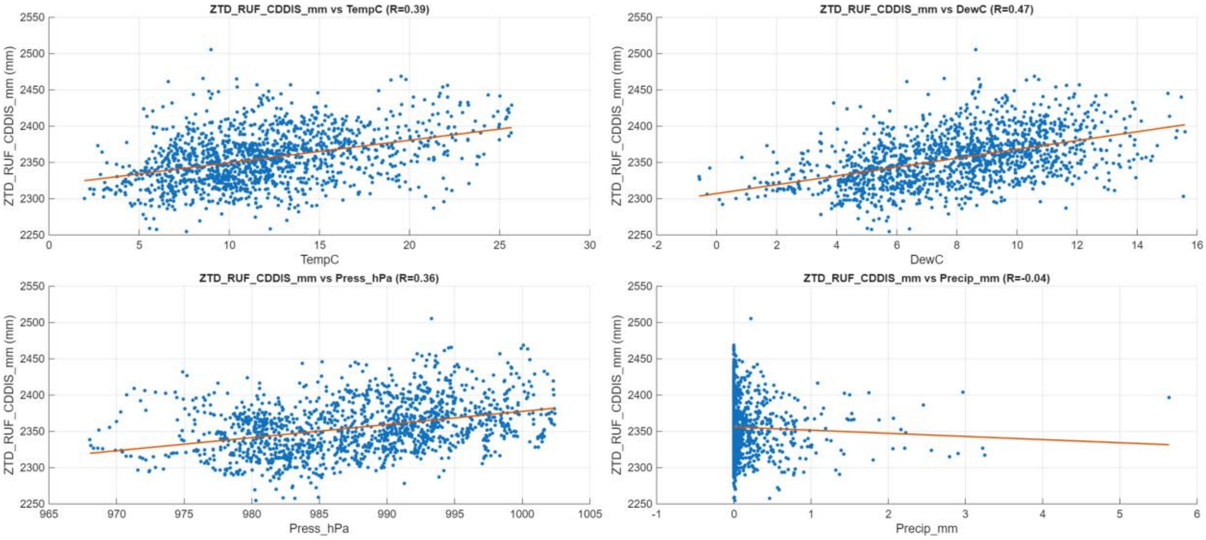

Figure 6.8. As Figure 6.5, but for RUF_CDDIS ZTD.

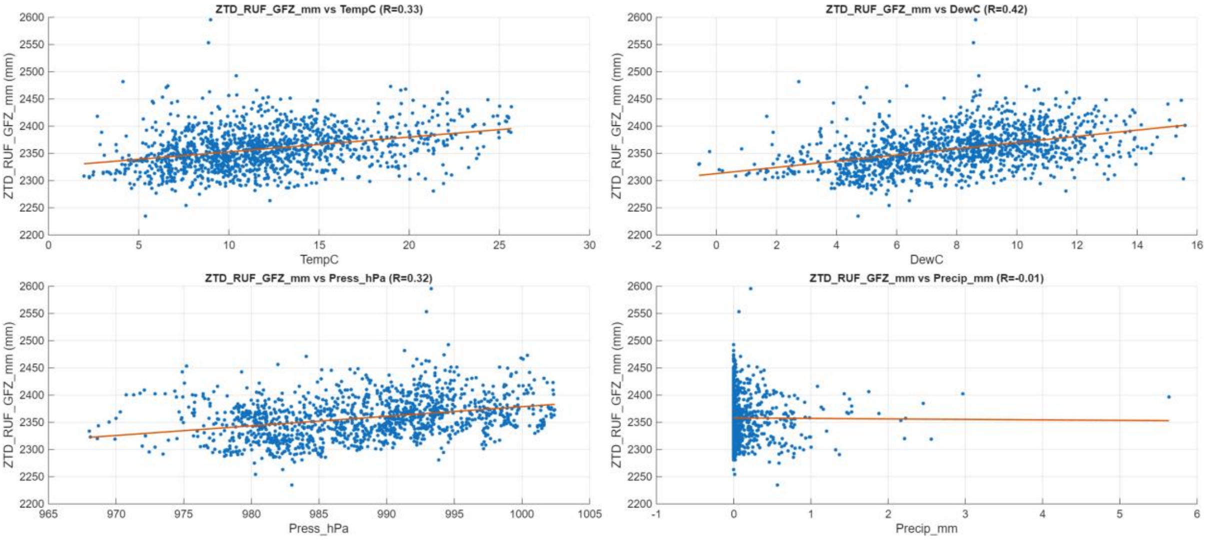

Figure 6.9. As Figure 6.5, but for RUF_GFZ ZTD.

Figures 6.5–6.9 present scatter plots and linear regressions between hourly ZTD and ERA5 variables for each of the five ZTD series: IGMI_EUREF, IGMI_GFZ, IGMI_CDDIS, RUF_GFZ and RUF_CDDIS. Corresponding correlation coefficients are summarised in *Table 6.3.* In all cases, ZTD increases with temperature, dew point and pressure, while the relationship with hourly precipitation is weak and noisy. This is consistent with the interpretation of ZTD as a proxy for integrated column water vapour and hydrostatic mass loading, whereas precipitation represents discrete outcomes of moist-process dynamics.

For IGMI_EUREF *(Figure 6.5)*, correlations with temperature, dew point and pressure are r = 0.31, 0.61 and 0.45, respectively, while precipitation shows virtually no instantaneous correlation (r ≈ -0.01). Dew point clearly emerges as the strongest predictor of ZTD variability, with a well-defined positive slope reflecting the close link between column water vapour and near-surface humidity. Temperature and pressure exhibit moderate positive relationships, reflecting their combined control on air density and hydrostatic mass loading.

The IGMI_GFZ and IGMI_CDDIS PPP series (*Figures 6.6* and *6.7*) display very similar patterns. Dew point retains the highest correlations (r ≈ 0.50–0.51), followed by temperature (r ≈ 0.42–0.43) and surface pressure (r ≈ 0.30–0.31), with only weak correlations with precipitation (r ≈ 0.07). The regression slopes are close to those obtained for IGMI_EUREF, indicating that both PPP branches reproduce the same underlying meteorological relationships as the network-combined reference. This confirms that the geodetic PPP solutions reproduce the physical relationships present in the EUREF product rather than introducing spurious variability, and that the choice of precise products (CDDIS vs GFZ) mainly affects small-scale noise and bias rather than the large-scale meteorological signal.

For the low-cost station RUF01, both PPP strategies (*Figures 6.8* and *6.9*) again show dew point as the dominant predictor (r ≈ 0.42–0.47), with temperature and pressure providing moderate correlations (r ≈ 0.33–0.39 and 0.32–0.36, respectively). Precipitation is essentially uncorrelated at zero lag. Regression slopes with respect to dew point and temperature are very close to those at IGMI00ITA, confirming that the low-cost receiver senses the same meteorological forcing, despite its different hardware and local environment.

Table 6.3 synthesises these findings by listing, for each ZTD series, the correlation coefficients with ERA5 temperature, dew point, pressure and precipitation. Across all five series, dew point consistently yields the highest correlations (typically 0.4–0.6), followed by temperature and pressure (≈0.3–0.4). Hourly precipitation exhibits only weak instantaneous correlation, reinforcing the interpretation that rainfall is an outcome of moist dynamics rather than a direct proxy for column water vapour.

Table 6.3. Correlation (r) between ZTD and ERA5 variables (hourly, March–April 2024)

ZTD series	*r(t2m)*	*r(d2m)*	*r(sp)*	*r(tp)*
IGMI_EUREF	0.31	0.61	0.45	- 0.01
IGMI_CDDIS	0.43	0.51	0.30	0.07
IGMI_GFZ	0.42	0.50	0.31	0.07
RUF_CDDIS	0.39	0.47	0.36	- 0.04
RUF_GFZ	0.33	0.42	0.32	- 0.01

The close similarity of correlation patterns and regression slopes between EUREF, geodetic PPP and low-cost PPP series demonstrates that the long-term ZTD estimates used in this chapter are physically consistent with the local meteorological environment. In particular, the low-cost RUF01 station captures the same humidity-driven ZTD variability as the neighbouring geodetic station.

6.6 PPP Positioning Stability Analysis

6.6.1 Workflow And Metrics

Positioning stability provides an independent check on PPP quality and, by extension, on the reliability of the derived ZTD. The analysis follows the workflow introduced in Section 4.9.

For each station and PPP branch, daily PPP solutions from RTKLIB are compared against fixed reference coordinates: Centipede network coordinates for RUF01 and EUREF coordinates (propagated to epoch 2024.25) for IGMI00ITA. Cartesian differences are rotated into the local East–North–Up (ENU) frame, and scalar metrics are derived for horizontal (dH) and three-dimensional (d3D) residuals.

Both raw statistics and robust, outlier-filtered summaries are computed. The latter employs thresholds based on the MAD to focus on the “typical” behaviour while excluding large excursions associated with convergence, data gaps or unmodelled disturbances.

Over the two-month analysis window, using a single mid-period propagated reference instead of a fully time-dependent reference produces differences that are negligible compared with the centimetre-to-decimetre residual statistics reported, because with velocities on the order of 1–2 cm yr^{-1}, the additional displacement over ±1 month would be at the millimetre level.

6.6.2 ENU Residual Behaviour

The positioning stability of RUF01 and IGMI00ITA was assessed using PPP-derived ENU residuals computed relative to fixed reference coordinates, retaining only converged static solutions (Q = 6). Residual statistics were derived from the full 30 s time series over March–April 2024 and evaluated using the robust filtering procedure introduced in Section 4.9 (ns ≥ 7, MAD-based filtering and tail trimming).

At both stations, the horizontal components (dE, dN) remain tightly centred around zero for the majority of epochs, reflecting stable horizontal geometry under static PPP conditions. As expected from GNSS observation geometry, the vertical component (dU) exhibits larger dispersion than the horizontal components, consistent with the weaker constraint in the Up direction.

The geodetic reference station IGMI00ITA displays lower overall instability, with an outlier rate below 12% for both processing strategies. The vertical dispersion, expressed as Std(dU), is approximately 3.7–3.8 cm (*Table 6.4*), consistent with expectations for a well-installed geodetic monument. Median absolute three-dimensional residuals |d3D| are approximately 8.0–8.1 cm, with 95th-percentile values near 15 cm. These values reflect typical static PPP behaviour and include small systematic effects associated with reference-coordinate propagation and epoch alignment.

The low-cost station RUF01 exhibits the same overall residual structure but with more frequent transient disturbances. The outlier rate reaches approximately 27% for CDDIS processing and approximately 25% for GFZ processing. Vertical dispersion is larger than at IGMI00ITA, with Std(dU) ≈ 5.3 cm under CDDIS and ≈ 5.0 cm under GFZ. Importantly, once transient degradations are removed, the median absolute three-dimensional residuals remain comparable to IGMI00ITA (≈ 7.4 cm), and 95th-percentile values remain below approximately 15 cm.

A comparison of precise-product strategies indicates slightly reduced dispersion under GFZ-based processing at both stations, particularly in the vertical component at RUF01, where Std(dU) decreases by approximately 0.3 cm relative to CDDIS. This modest improvement is consistent with the ZTD internal-consistency results discussed in Section 6.4.2 and confirms that precise orbit/clock strategy exerts a secondary but measurable influence on positioning stability.

Overall, the robust statistics demonstrate that, despite a higher rate of transient disturbances, the typical positioning behaviour of RUF01 remains stable at the centimetre scale once episodic outliers are excluded. This level of stability is sufficient to support reliable tropospheric delay estimation and confirms that low-cost installations, when processed under a controlled PPP framework and evaluated using robust filtering, can achieve performance consistent with atmospheric monitoring requirements.

Table 6.4. Robust PPP positioning statistics (ENU residuals, March–April 2024)

Station	***Processing Strategy***	***Median |d3D| (cm)***	***P95(|d3D|) (cm)***	***Std(dU) (cm)***	***Outlier Rate (%)***
RUF01	CDDIS	7.4	14.4	5.3	26.8
	GFZ	7.4	14.2	5.0	24.8
IGMI00ITA	CDDIS	8.1	15.2	3.7	11.9
	GFZ	8.0	15.1	3.8	11.1

6.6.3 Implications for Tropospheric Estimation

The positioning analysis demonstrates that, despite occasional spikes, the typical PPP positioning behaviour of RUF01 is sufficiently stable to support reliable tropospheric delay estimation when appropriate quality control is applied. The coherence between ZTD variability and meteorological forcing, combined with centimetre-level typical position residuals, confirms that the tropospheric signals extracted from RUF01 are not artefacts of unstable positioning.

6.7 Synthesis and answers to the research questions

This chapter addressed the four questions posed in Section 6.1.

(i) Temporal stability of low-cost ZTD (RUF01)

The RUF01 ZTD series exhibits stable behaviour over March–April 2024. After QC, the hourly time series (*Figure 6.1*) shows no evidence of drift, only a nearly constant offset relative to IGMI, attributable to altitude and local effects. Positioning residuals remain at the centimetre scale after robust filtering (*Table 6.4*), consistent with a mechanically stable monument.

(ii) Sensitivity to precise products (CDDIS vs GFZ)

Internal comparisons at each station (*Table 6.2; Figures 6.2–6.3*) reveal millimetric mean biases and centimetric RMS differences between CDDIS- and GFZ-based PPP solutions, with correlation coefficients exceeding 0.95. For IGMI00ITA, both strategies reproduce the EUREF ZTD reference within a few centimetres (*Table 6.1*), confirming that the overall tropospheric signal is largely insensitive to the choice of precise products.

These results indicate that differences between precise-product strategies are small relative to the intrinsic temporal variability of the tropospheric delay and to the systematic offsets associated with station height and hardware characteristics. Importantly, this does not imply that the receiver class is negligible. Rather, the analysis demonstrates that, when identical PPP settings and rigorous quality control are applied, the correction strategy itself can introduce biases of comparable magnitude to hardware-related effects. Consequently, both receiver characteristics and processing strategy contribute to systematic offsets, but processing consistency emerges as a primary determinant of inter-series bias stability.

(iii) Meteorological response and physical consistency

Correlations and regressions with ERA5 (*Figure 6.4; Figures 6.5–6.9; Table 6.3*) demonstrate that all five ZTD series respond coherently to near-surface temperature, dew point and pressure. Dew point is the dominant predictor of ZTD variability, followed by temperature and pressure, while precipitation shows negligible instantaneous correlation. The low-cost RUF01 station mirrors the geodetic behaviour, confirming that its ZTD estimates are physically meaningful.

(iv) Positioning stability and tropospheric reliability

PPP positioning residuals at RUF01 and IGMI00ITA are typically at the few-centimetre level, with occasional spikes that coincide with ZTD outliers handled by the rolling-median filter (*Table 6.4*). This behaviour indicates that the PPP solutions are geometrically stable and that the tropospheric estimates are not dominated by positioning errors.

In summary, a carefully processed low-cost GNSS station such as RUF01 can provide stable and physically meaningful ZTD estimates over multi-week periods, provided that robust QC and validation are applied and that processing strategies are controlled. The demonstrated temporal coherence, processing robustness, and physically consistent meteorological response of RUF01 support the assumption that, after appropriate station-level filtering, similar low-cost stations within the Centipede network can contribute meaningfully to regional-scale ZTD mapping, as explored in the following chapter.

CHAPTER 7: FRANCE-WIDE SPATIAL RECONSTRUCTION OF ZENITH TOTAL DELAY (MARCH 2024)

7.1 Introduction

The preceding chapters established the methodological and physical foundations required for spatial tropospheric reconstruction. Chapter 5 demonstrated the physical consistency of GNSS-derived tropospheric components under controlled multi-station experiments, confirming expected relationships between hydrostatic delay and surface pressure, and between wet delay and moisture indicators. Chapter 6 extended this validation temporally, showing that low-cost GNSS receivers, when processed under a controlled PPP framework and subjected to strict stability filtering, produce ZTD time series consistent with geodetic reference products.

Building upon these validated pillars, the present chapter addresses a fundamentally different scale of analysis, national spatial reconstruction. Rather than evaluating isolated stations or short temporal windows, the objective here is to determine whether a dense network of low-cost GNSS stations can support a physically coherent and statistically robust Zenith Total Delay (ZTD) field over metropolitan France at 1 km spatial resolution.

Zenith Total Delay integrates the entire atmospheric refractivity column and is therefore dominated by hydrostatic pressure effects, which are strongly elevation-dependent [4,5]. However, horizontal variability driven by mesoscale moisture processes, coastal gradients, and synoptic transport remains superimposed upon this deterministic vertical structure [12,17]. A national reconstruction must therefore separate these two components to avoid conflating topographic control with atmospheric variability.

The methodological framework adopted follows the regression–kriging strategy defined in Chapter 4. The elevation-dependent deterministic structure is first modelled explicitly using least-squares regression, after which the residual field, representing horizontally structured atmospheric variability, is interpolated geostatistically in projected coordinates. This separation respects both atmospheric physics and geostatistical assumptions of second-order stationarity.

The central research question of this chapter is therefore twofold: whether the monthly mean ZTD over France is overwhelmingly controlled by elevation, as predicted by hydrostatic theory, and whether the residual atmospheric variability can be reconstructed reliably using a dense low-cost GNSS network.

7.2 Station Dataset and Filtering

The France-wide campaign began with 360 stations providing usable ZTD time series after epoch-level quality control and hourly aggregation. As described in Section 4.10, station-level filtering was then applied to ensure temporal coverage, statistical stability, and positional robustness.

Stations were required to provide at least 25 valid days during March 2024 and to exhibit a monthly ZTD standard deviation not exceeding 30 mm. This reduced the dataset to 332 stations. Restricting the domain to metropolitan France yielded 322 stations. Finally, a positioning stability constraint of Rad95 ≤ 0.10 m was imposed using PPP-derived coordinate solutions (Section 4.9), resulting in a final interpolation-ready dataset of 274 stations.

The Rad95 threshold ensures that coordinate instability does not propagate into tropospheric estimates, consistent with the sensitivity of ZTD to station height errors discussed in Chapter 6 and in previous literature [8]. The final 274-station network provides dense spatial coverage across lowlands, coastal regions, central plateaus, and high mountain environments including the Alps and Pyrenees.

7.3 Elevation Dependence of Monthly Mean ZTD

Zenith Total Delay is dominated by the hydrostatic component, which scales primarily with surface pressure and therefore with elevation [4,5]. In order to isolate horizontal atmospheric variability, the vertical trend was explicitly modelled using a least-squares regression between station mean ZTD (mm) and elevation h(m).

The resulting relationship is:

$$ZTD_{hat}(h) = -0.283838\,h + 2395.059384$$

The regression statistics are:

$$R^2 = 0.9932, \qquad RMSE = 6.721\,\text{mm}, \qquad N = 274$$

The coefficient of determination indicates that 99.32% of the spatial variance in monthly mean ZTD across metropolitan France is statistically explained by elevation alone within the adopted linear model. This extremely high explanatory power confirms that hydrostatic pressure dominates the large-scale spatial structure of ZTD at monthly timescales, fully consistent with classical GNSS meteorology theory [4,12].

The fitted slope corresponds to approximately -284 mm per kilometre of elevation gain. Although atmospheric pressure decreases exponentially with height, over the limited elevation range encountered in metropolitan France (0–2300 m), the hydrostatic contribution to ZTD is well approximated by a linear model, as evidenced by an RMSE of 6.721 mm, which is small relative to the national ZTD range exceeding 600 mm.

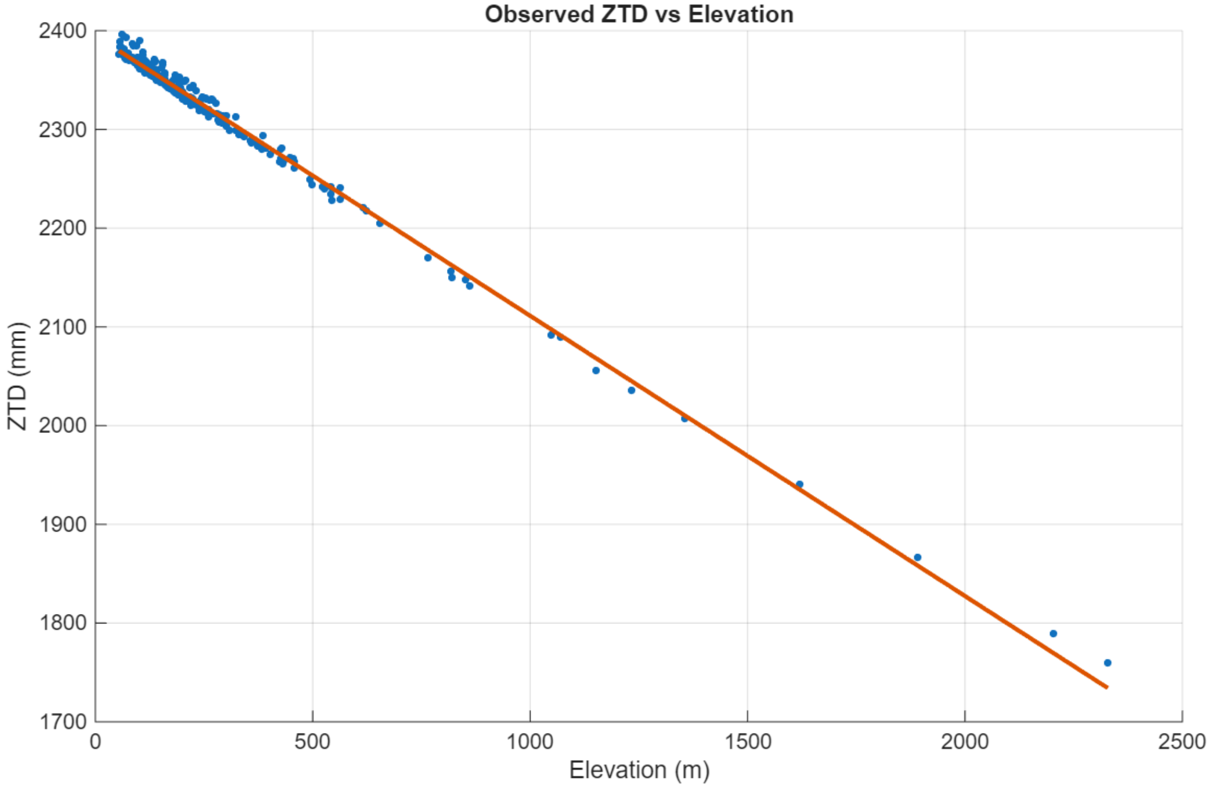

Figure 7.1. Observed monthly-mean ZTD as a function of station elevation (March 2024)

The plot reveals a highly coherent linear structure with minimal dispersion, indicating that the regression effectively captures the dominant vertical atmospheric control.

7.4 Station-Level Results and Residual Analysis

A condensed selection of the regression dataset for the 274 stations is provided in *Table 7.1*, reporting representative stations and the most extreme residual cases (largest positive and negative values). For each listed station, the table includes latitude, longitude, elevation, Rad95 observed monthly mean ZTD, predicted ZTD from the regression model and residual value.

Residuals are defined as:

$$Resid_i = ZTD_{obs,i} - ZTD_{hat,i}$$

The observed monthly mean ZTD values range from 1760.11 mm at SCV100FRA (2328 m elevation) to 2396.73 mm at PALI00FRA (60 m elevation). After removal of the elevation trend, residuals are generally small, with most stations falling within approximately ±10 mm. This magnitude is small relative to both the national ZTD range (>600 mm) and the regression RMSE (6.7 mm), indicating that residual variability represents secondary atmospheric structure rather than unmodelled topographic bias.

The residual range across the network is:

$$Resid_{min} = -12.93 \text{ mm}, \qquad Resid_{max} = +25.98 \text{ mm}$$

The largest positive residual occurs at SCV100FRA (+25.98 mm), indicating that at very high elevation, the linear model slightly underestimates the observed ZTD. The largest negative residual (-12.93 mm at WLBH00FRA) remains small relative to the full ZTD magnitude.

Table 7.1. Stations with Largest Positive and Negative Residuals

Largest Positive Residuals (Model Underestimation)

Station	*Lat (deg)*	*Lon (deg)*	*Hgt (m)*	*Rad95 (m)*	*ZTD_obs (mm)*	*ZTD_hat (mm)*	*Resid (mm)*
SCV100FRA	44.92	6.54	2328.52	0.087	1760.11	1734.14	+25.98
CITR00FRA	42.28	9.52	102.79	0.083	2389.85	2365.88	+23.97
CHA200FRA	45.12	5.89	2202.87	0.069	1789.45	1769.80	+19.66
PALI00FRA	43.37	4.81	60.38	0.077	2396.73	2377.92	+18.81
AGDE00FRA	43.29	3.46	65.80	0.077	2394.71	2376.38	+18.33
MONT00FRA	43.30	3.09	71.33	0.071	2393.08	2374.81	+18.27
CETT00FRA	43.39	3.65	60.27	0.077	2395.95	2377.95	+18.00
LLEN00FRA	43.39	5.17	65.54	0.084	2394.19	2376.46	+17.73
PAS600FRA	43.51	-0.92	94.25	0.073	2384.59	2368.31	+16.28
HYE100FRA	43.12	6.12	154.02	0.075	2367.32	2351.34	+15.98

Largest Negative Residuals (Model Overestimation)

Station	*Lat (deg)*	*Lon (deg)*	*Hgt (m)*	*Rad95 (m)*	*ZTD_obs (mm)*	*ZTD_hat (mm)*	*Resid (mm)*
WLBH00FRA	48.41	7.35	819.14	0.066	2149.62	2162.56	-12.93
CBER00FRA	50.01	5.74	543.63	0.079	2227.94	2240.76	-12.81
AUBU00FRA	48.21	7.19	1151.89	0.096	2055.79	2068.11	-12.32
FR5200FRA	48.09	5.48	499.10	0.083	2243.89	2253.40	-9.50
FE0800FRA	49.68	4.52	309.42	0.051	2298.74	2307.23	-8.50
LAJ300FRA	46.37	5.97	1234.31	0.067	2036.23	2044.72	-8.48
MAGC00FRA	45.79	2.88	861.98	0.092	2141.80	2150.40	-8.60
MNE100FRA	45.25	6.41	764.39	0.082	2169.72	2178.10	-8.37
SLIN00FRA	50.72	5.57	200.99	0.096	2330.63	2338.01	-7.39
ALEX00FRA	50.60	4.37	207.90	0.084	2328.66	2336.05	-7.39

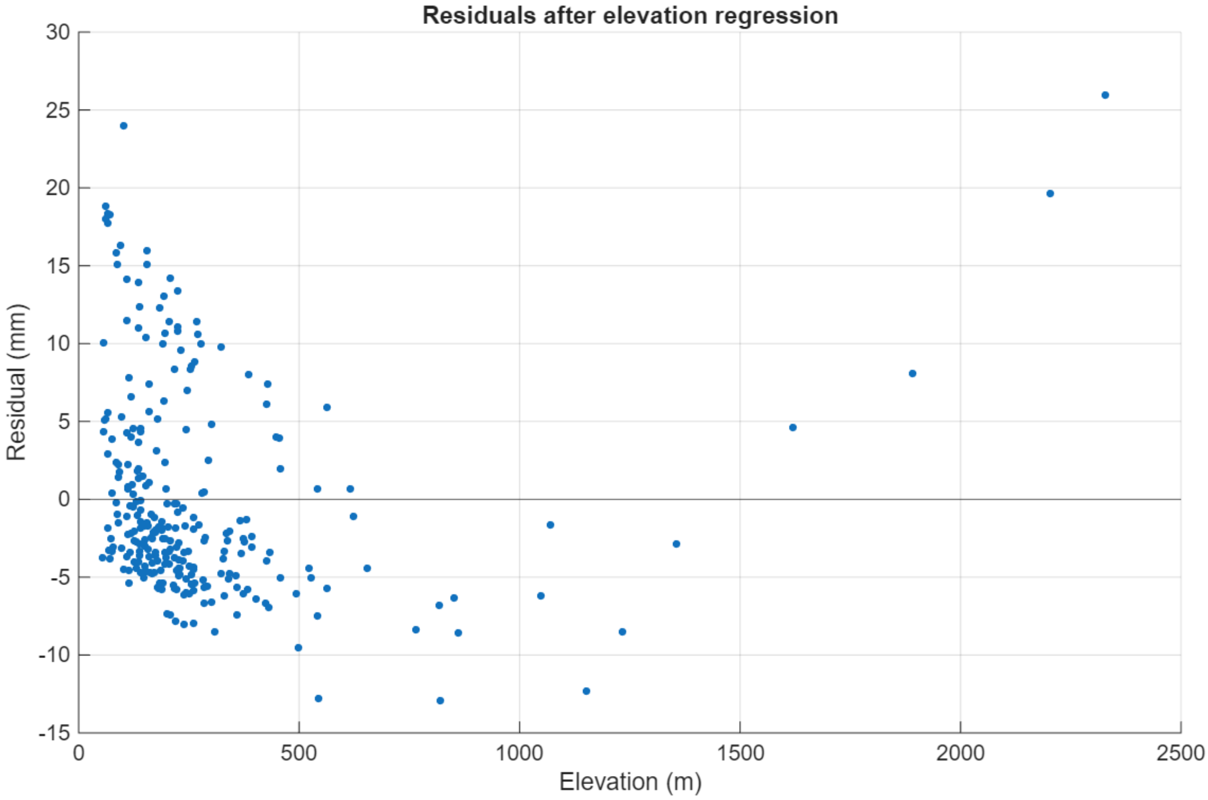

Figure 7.2. ZTD residuals vs station elevation after linear detrending (March 2024)

Figure 7.2 shows the residuals as a function of elevation. No remaining linear trend is visible, confirming that the dominant altimetric component has been successfully removed. Residual dispersion does not exhibit a systematic linear dependence on elevation, although isolated high-elevation stations show larger deviations, supporting the adequacy of a linear elevation model over the elevation range covered by the French network. The remaining residual structure likely reflects secondary meteorological controls such as regional moisture transport, coastal humidity gradients, and mountain boundary-layer processes [17,18]. These signals correspond to the horizontal atmospheric variability targeted by the subsequent geostatistical interpolation stage.

7.5 Residual Kriging

Spatial interpolation was performed on the residual field rather than on raw ZTD values, ensuring that deterministic elevation structure does not bias the kriging model. Universal Kriging was applied in the EPSG:2154 (Lambert-93) projected coordinate system to maintain metric consistency across the national domain.

The interpolation grid was defined at 1 km resolution. Variogram parameters were determined automatically, and a local neighbourhood search radius of approximately 200 km was employed, using between 8 and 20 neighbouring stations for each prediction node. The empirical semi variogram of the residual field exhibited near-isotropic behaviour, and an automatically fitted spherical model with a small nugget component provided stable predictions. These settings balance spatial smoothness and local representativity in accordance with geostatistical best practices [32].

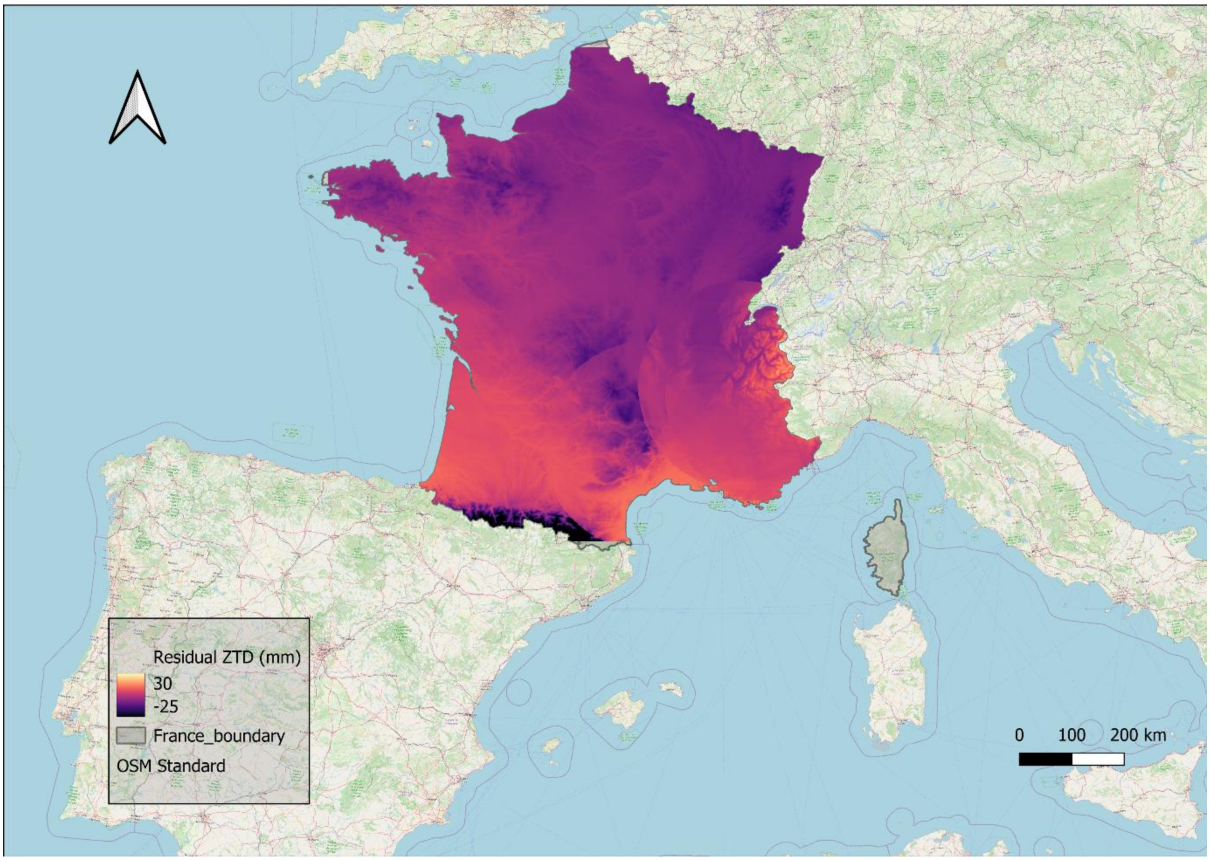

Figure 7.3. Universal-kriged residual ZTD field (March 2024, metropolitan France, EPSG:2154, 1 km)

Figure 7.3 presents the Universal Kriging residual surface. The interpolated residuals exhibit coherent spatial structure. Positive residuals are evident along Mediterranean coastal regions, while moderate negative residuals appear in parts of northeastern France and selected Alpine areas. No systematic bullseye artefacts are observed, indicating stable variogram behaviour and sufficient station density across most of the domain.

The residual field is displayed within a constrained visualization range (-25 to +30 mm) to improve interpretability. Extreme extrapolation values reaching approximately -60 mm occur locally at the southwestern boundary. These extrapolated values are artefacts of kriging near domain boundaries where local station density decreases and should not be interpreted as physically meaningful atmospheric anomalies.

7.6 DEM Integration and Final ZTD Field

Elevation at each grid node was obtained from a 1 km Digital Elevation Model derived from SRTM data. The DEM was mosaicked, reprojected to EPSG:2154 (Lambert-93), resampled to 1 km resolution, and clipped to metropolitan France to ensure spatial consistency with the kriging grid.

The final ZTD field was reconstructed using the regression–kriging formulation:

$$ZTD_{final}(x, y) = -0.283838 \cdot h(x, y) + 2395.059384 + Residual_{kriged}(x, y)$$

This formulation explicitly preserves the physically dominant vertical structure quantified by the elevation regression while superimposing horizontally interpolated atmospheric variability represented by the kriged residual component.

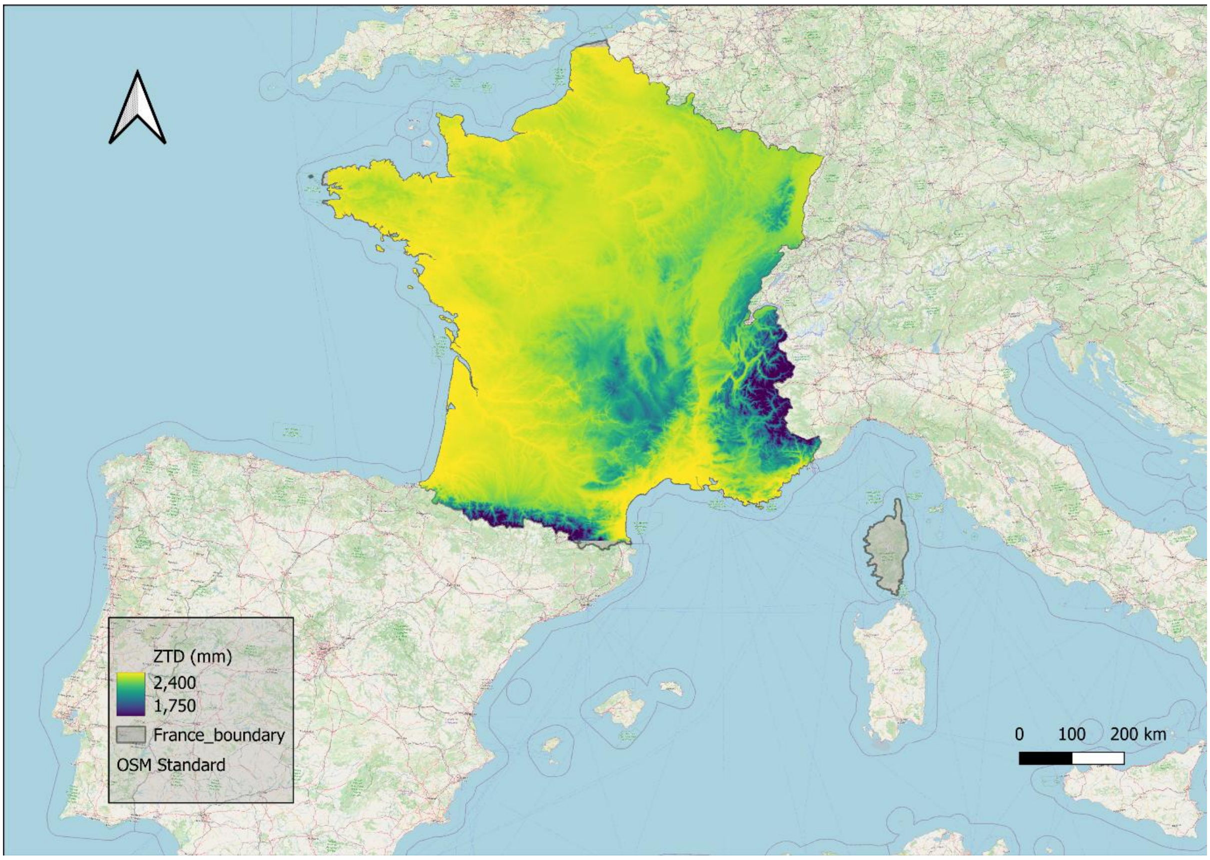

Figure 7.4. Final reconstructed ZTD map of metropolitan France at 1 km resolution for March 2024

Figure 7.4 presents the final reconstructed ZTD field for March 2024. The national-scale pattern exhibits strong physical coherence. Lowland regions display ZTD values near 2390 mm, while high-altitude regions of the Alps and Pyrenees exhibit pronounced reductions below 1800 mm. Some Mediterranean coastal sectors show slightly elevated ZTD values relative to adjacent inland areas at comparable elevation, consistent with enhanced low-level moisture availability. No discontinuities or projection artefacts are visible.

The large-scale spatial gradient reflects the dominant hydrostatic control quantified by the elevation regression (Section 7.3). The pronounced reductions over the Alps and Pyrenees correspond directly to increased elevation, while extensive lowland regions exhibit comparatively homogeneous ZTD values. This spatial structure confirms that the regression–kriging framework successfully separates deterministic elevation-driven variability from the weaker horizontally correlated atmospheric component.

7.7 Interpretation and Implications

The France-wide reconstruction confirms that the monthly mean ZTD over metropolitan France is governed primarily by elevation-driven hydrostatic structure. The elevation regression explains 99.32% of the spatial variance ($R^2 = 0.9932$), and the fitted slope of approximately -284 mm km^{-1} quantifies the dominant reduction of atmospheric delay with height in a manner consistent with refractivity theory [4,5,12].

After removal of this deterministic component, station residuals are generally confined to the ±10 mm range and rarely exceed ±20 mm, indicating that horizontally structured atmospheric variability exists but remains small relative to the national altimetric gradient. The residual field nevertheless exhibits spatial coherence, with patterns consistent with mesoscale moisture influences and coastal effects superimposed on the hydrostatic baseline [10,17,18].

This scale separation is critical for geostatistical modelling. Detrending suppresses topography-driven non-stationarity and allows kriging to target the weaker, horizontally correlated component under assumptions closer to second-order stationarity [32,33]. The resulting residual surface is smooth and physically plausible, and kriging variance increases primarily in peripheral mountainous regions where station density decreases, consistent with geostatistical expectation.

7.8 Cross-Validation and Comparative Interpolation Performance

To assess the predictive reliability of the France-wide ZTD reconstruction and to quantify the methodological benefit of the regression–kriging framework, a structured leave-one-out cross-validation (LOOCV) analysis was conducted. The validation was performed on monthly mean station ZTD values in order to maintain full consistency with the deterministic regression framework adopted for spatial reconstruction. In this procedure, each station was sequentially removed from the dataset, its ZTD value was predicted using the remaining stations, and the resulting prediction error was recorded. The same validation procedure was applied consistently across all interpolation strategies to ensure strict comparability under identical statistical definitions [28,32].

Three alternative interpolation approaches were examined within this framework. First, Inverse Distance Weighting (IDW) was applied directly to station-mean ZTD without explicit modelling of elevation dependence. Second, Universal Kriging (UK) was applied directly to station-mean ZTD with elevation incorporated into the deterministic trend component. Third, a regression–kriging (RK) approach was implemented, in which the deterministic elevation-dependent trend was first removed through global regression and the residual field was subsequently interpolated prior to final reconstruction. All validation metrics were computed according to the formal definitions introduced in Section 4.7.

For strict comparability, IDW was also recomputed on the 263-station subset used for UK and RK validation. Performance metrics are therefore computed over identical station sets, ensuring that differences between IDW, UK, and RK reflect methodological performance rather than sample-size effects.

7.8.1 IDW Validation

Under the harmonised leave-one-out cross-validation (LOOCV) framework applied to the 263-station subset used consistently for Universal Kriging and Regression–Kriging evaluation, direct IDW interpolation of absolute ZTD produced a bias of +1.615 mm, a mean absolute error (MAE) of 26.041 mm, and a root mean square error (RMSE) of 56.165 mm, with a coefficient of determination of $R^2 = 0.540$.

The large RMSE confirms that IDW fails to reproduce the dominant vertical structure of ZTD. Because IDW relies exclusively on horizontal distance weighting, it cannot represent the strong altimetric dependence demonstrated in Section 7.3. Prediction errors were particularly pronounced in high-elevation regions, where differences in station elevation exceed horizontal proximity effects. This behaviour reflects the inability of purely distance-based estimators to represent vertically stratified atmospheric structure.

These results demonstrate that deterministic elevation effects must be explicitly modelled when interpolating ZTD at the national scale.

7.8.2 Universal Kriging Validation

Universal Kriging was then applied directly to station-mean ZTD with elevation incorporated into the deterministic component. Under the LOOCV framework, 263 stations yielded valid predictions, as a small number of peripheral stations did not meet the minimum-neighbour constraint imposed by the local search configuration. The resulting cross-validation performance indicated an RMSE of 7.93 mm and a coefficient of determination of $R^2 = 0.991$.

The substantial improvement relative to IDW reflects the explicit inclusion of elevation within the kriging trend model. By embedding deterministic structure into the trend function, Universal Kriging captures the majority of large-scale spatial variance. Although overall RMSE remains low, isolated prediction errors exceeding 20–30 mm were observed in complex terrain, reflecting sensitivity to local station geometry and variogram behaviour. These residual discrepancies likely arise from sensitivity to local geometry, variogram fitting behaviour, and irregularities in station distribution [32].

7.8.3 Regression–Kriging Validation

The regression–kriging framework yielded the strongest predictive performance among the evaluated methods. After restricting the comparison to the 263 stations used in the Universal Kriging validation, the leave-one-out cross-validation results indicated a bias of -0.076 mm, a mean absolute error (MAE) of 1.849 mm, and a root mean square error (RMSE) of 3.658 mm, with a coefficient of determination of $R^2 = 0.998$. The near-zero bias confirms the absence of systematic distortion in the reconstructed field. The reduction of RMSE by more than a factor of two relative to Universal Kriging highlights the statistical advantage of explicitly separating deterministic elevation-dependent structure from stochastic horizontal variability prior to interpolation.

Table 7.2 summarises the comparative validation results.

Method	*RMSE (mm)*	*R^2*	*Principal Limitation*
IDW	56.17	0.54	No explicit elevation modelling
Universal Kriging	7.93	0.991	Sensitive to trend specification
Regression–Kriging	3.66	0.998	Requires global trend estimation

7.8.4 Interpretation of Validation Results

The validation results provide quantitative confirmation of the physical reasoning developed earlier in this chapter. Given that 99.32% of spatial variance is statistically explained by elevation within the adopted linear model, interpolation methods that do not explicitly represent this deterministic structure are expected to incur large predictive errors.

Universal Kriging improves performance by embedding elevation into the trend model, but still estimates deterministic and stochastic components simultaneously. Regression–Kriging, by contrast, first isolates the dominant altimetric control through global regression and then applies kriging only to the residual field. This two-stage separation yields a residual process closer to second-order stationarity, which is a fundamental assumption of kriging theory [32,33].

The reduction of RMSE to 3.658 mm indicates that the residual spatial variability at the monthly timescale is weak relative to the deterministic altimetric structure but remains sufficiently structured to be predicted under second-order stationarity assumptions. Importantly, the magnitude of cross-validated prediction error remains well below the intrinsic spatial ZTD range exceeding 600 mm, demonstrating that the regression–kriging framework produces a stable and physically coherent national product.

From a methodological standpoint, these results validate the modelling philosophy adopted in Chapter 4. The deterministic elevation-dependent structure must be treated explicitly before applying geostatistical interpolation. When this separation is respected, dense low-cost GNSS networks can support national-scale atmospheric mapping with predictive performance comparable to traditional geostatistical applications.

7.9 Chapter Conclusions and Scientific Implications

This chapter demonstrates that a dense low-cost GNSS network is capable of supporting national-scale reconstruction of monthly mean ZTD over metropolitan France when processed under a controlled PPP workflow and subjected to rigorous coverage and positional stability filtering. Elevation dominates the spatial structure of monthly mean ZTD ($R^2 = 0.9932$; slope ≈ -284 mm km^{-1}), confirming the expected hydrostatic control on large-scale delay. After detrending, the residual field is small but spatially coherent, and regression–kriging provides the most reliable prediction performance in cross-validation (RMSE = 3.658 mm), substantially outperforming distance-only interpolation and improving upon Universal Kriging applied to raw ZTD.

The resulting 1 km field should be interpreted as a climatological-scale product derived from temporal averaging, in which high-frequency PPP noise is strongly suppressed and physically meaningful spatial gradients become observable. The results demonstrate that modelling assumptions represent a primary contributor to reconstruction uncertainty at the monthly timescale, often exceeding the influence of residual measurement noise. Collectively, these findings confirm that physically constrained regression–kriging provides a robust and scalable framework for transforming dense GNSS observations into coherent national atmospheric products.

CHAPTER 8: SYNTHESIS, LIMITATIONS, AND FUTURE RESEARCH

8.1 Integrated Synthesis of Findings

The results presented in Chapters 5 through 7 collectively establish a coherent validation chain across spatial and temporal scales. Chapter 5 confirmed the short-term physical consistency of GNSS-derived tropospheric components across contrasting meteorological regimes. Hydrostatic delay (TRODRY/ZHD) exhibited strong linear dependence on surface pressure, while wet delay (TROWET/ZWD) demonstrated sensitivity to moisture indicators, with site-dependent variability reflecting boundary-layer structure and topographic complexity. These findings validated the physical interpretability of PPP-derived zenith delays under controlled multi-station conditions.

Chapter 6 extended this validation temporally and methodologically. The low-cost RUF01 station produced a ZTD time series statistically consistent with geodetic EUREF reference products when processed under identical PPP configurations. Processing strategy (CDDIS vs GFZ precise products) introduced systematic differences of comparable magnitude to receiver-class differences, demonstrating that correction modelling plays a critical role in bias control. Positioning stability analysis further confirmed that tropospheric reliability is linked to coordinate solution robustness, particularly in the vertical component.

Chapter 7 addressed the largest spatial scale of analysis. Monthly mean ZTD over metropolitan France was shown to be overwhelmingly controlled by elevation-driven hydrostatic structure ($R^2 = 0.9932$; slope ≈ -284 mm km^{-1}). After explicit removal of this deterministic component, the residual field was small but spatially coherent. Regression–kriging yielded cross-validated prediction errors (RMSE = 3.658 mm) substantially lower than both IDW and Universal Kriging applied directly to raw ZTD. This confirms that physically informed separation of deterministic and stochastic components is essential for stable national-scale atmospheric reconstruction.

Taken together, the thesis demonstrates that GNSS-derived tropospheric delays from both geodetic and low-cost receivers are physically consistent, temporally stable under controlled processing, and spatially reconstructable at the national scale when deterministic structure is explicitly modelled.

8.2 Scientific Implications

Several broader implications emerge from this multi-scale analysis.

First, hydrostatic control dominates spatial ZTD structure at the monthly timescale. The elevation regression explains more than 99% of spatial variance across France, indicating that national-scale tropospheric mapping must explicitly treat altimetric structure prior to geostatistical interpolation. Failure to do so produces large predictive errors, as demonstrated by the IDW validation results.

Second, the processing strategy constitutes a primary source of systematic bias. Differences between precise product providers can produce offsets of tens of millimetres, comparable to receiver-class differences. This highlights the importance of consistent correction models and reinforces the necessity of standardised PPP workflows in comparative tropospheric studies.

Third, low-cost GNSS networks, when filtered for positional stability and processed under rigorous PPP frameworks, are capable of supporting physically coherent atmospheric products. The results therefore support the potential expansion of dense low-cost networks for climatological-scale tropospheric monitoring.

Finally, the regression–kriging framework demonstrates that physically constrained spatial decomposition improves geostatistical performance. By isolating deterministic hydrostatic structure prior to interpolation, the residual process more closely satisfies second-order stationarity assumptions, resulting in superior predictive accuracy.

8.3 Methodological Limitations

Despite the robustness of the results, several limitations must be acknowledged.

First, the France-wide reconstruction was performed on monthly mean ZTD values. Temporal averaging suppresses high-frequency variability and PPP noise, but it also limits the interpretation of synoptic or convective-scale processes. The reconstructed field therefore represents a climatological-scale product rather than a real-time atmospheric analysis.

Second, ERA5 surface variables were used as meteorological reference proxies in the multistation validation experiments. While physically justified under well-mixed boundary-layer conditions, near-surface dew point does not directly represent vertically integrated water vapour. In complex terrain, representativeness errors between station conditions and grid-cell means may reduce linear coupling.

Third, the elevation dependence was modelled using a first-order linear regression. Although statistically justified by the extremely high R^2 and absence of residual trend, atmospheric pressure decreases exponentially with height. The linear approximation is valid over the elevation range considered (0–2300 m), but may require refinement for higher topographic domains.

Fourth, cross-validation was conducted using leave-one-out methodology on the same temporal window. While this ensures spatial predictive reliability, independent temporal validation across multiple months or seasons would further strengthen generalisation.

Finally, station distribution remains uneven in mountainous and peripheral regions, leading to increased kriging variance in these areas. Although regression–kriging mitigates non-stationarity, interpolation reliability remains partly dependent on network geometry.

8.4 Directions for Future Research

Several avenues for future investigation arise naturally from this work.

Seasonal extension of the France-wide reconstruction would allow assessment of how elevation dominance and residual moisture variability evolve under differing atmospheric regimes. Inclusion of additional covariates, such as temperature gradients or integrated water vapour proxies, could refine residual modelling beyond purely spatial kriging.

Real-time or near-real-time regression–kriging frameworks could be explored to evaluate operational feasibility for atmospheric monitoring or data assimilation applications.

Further sensitivity experiments using additional precise product providers would clarify the extent to which the correction strategy influences long-term bias structure. Harmonisation of PPP processing standards across networks could reduce systematic discrepancies.

Finally, integration with independent atmospheric datasets, such as radiosonde or numerical weather prediction outputs, would enable deeper investigation into vertical moisture structure and improve understanding of ZWD variability beyond near-surface proxies.

8.5 Concluding Perspective

This thesis demonstrates that GNSS tropospheric delay estimation from low-cost receivers is not merely technically feasible but scientifically robust when embedded within physically informed processing and modelling frameworks. Across scales ranging from individual stations to national domains, deterministic hydrostatic structure and stochastic moisture variability can be separated and reconstructed with high statistical reliability.

The results highlight a central principle: accurate atmospheric mapping with GNSS requires explicit treatment of physical structure before statistical interpolation. When this principle is respected, dense GNSS networks, regardless of receiver cost class, can serve as valuable components of large-scale atmospheric monitoring systems.

REFERENCES

1. Hofmann-Wellenhof, B., Lichtenegger, H., & Wasle, E. (2008). *GNSS – Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and more*. Springer, Vienna. https://doi.org/10.1007/978-3-211-73017-1.
2. Teunissen, P. J. G., & Montenbruck, O. (Eds.). (2017). *Springer Handbook of Global Navigation Satellite Systems*. Springer, Cham. https://doi.org/10.1007/978-3-319-42928-1.
3. Davis, J. L., Herring, T. A., Shapiro, I. I., Rogers, A. E. E., & Elgered, G. (1985). Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimates of baseline length. *Radio Science*, 20(6), 1593–1607. https://doi.org/10.1029/RS020i006p01593.
4. Bevis, M., Businger, S., Herring, T. A., Rocken, C., Anthes, R. A., & Ware, R. H. (1992). GPS meteorology: Remote sensing of atmospheric water vapor using the Global Positioning System. *Journal of Geophysical Research: Atmospheres*, 97(D14), 15787–15801. https://doi.org/10.1029/92JD01517.
5. Saastamoinen, J. (1972). Contributions to the theory of atmospheric refraction. *Bulletin Géodésique*, 46(1), 279–298. https://doi.org/10.1007/BF02521844.
6. Hamza, V., Stopar, B., Sterle, O., & Pavlovčič-Prešeren, P. (2024). Observations and positioning quality of low-cost GNSS receivers: A review. *GPS Solutions*, 28, 149. https://doi.org/10.1007/s10291-024-01686-8.
7. CentipedeRTK. (n.d.). *Centipede low-cost GNSS network documentation and station archive*. https://docs.centipede.fr/.
8. Tregoning, P., & Herring, T. A. (2006). Impact of a priori zenith hydrostatic delay errors on GPS estimates of station heights and zenith total delays. *Geophysical Research Letters*, 33(23), L23303. https://doi.org/10.1029/2006GL027706.
9. Kouba, J., & Héroux, P. (2001). Precise Point Positioning using IGS orbit and clock products. *GPS Solutions*, 5(2), 12–28. https://doi.org/10.1007/PL00012883.
10. Holton, J. R., & Hakim, G. J. (2013). *An Introduction to Dynamic Meteorology* (5th ed.). Academic Press. https://doi.org/10.1016/C2009-0-63394-8.
11. European Space Agency (ESA). (2023). *ESA GNSS Handbook* (TM-23/1). European Space Agency.
12. Bevis, M., Businger, S., Chiswell, S., Herring, T. A., Anthes, R. A., Rocken, C., & Ware, R. H. (1994). GPS meteorology: Mapping zenith wet delays onto precipitable water. *Journal of Applied Meteorology*, 33(3), 379–386. https://doi.org/10.1175/1520-0450(1994)033<0379:GMMZWD>2.0.CO;2.

13. Niell, A. E. (1996). Global mapping functions for the atmosphere delay at radio wavelengths. *Journal of Geophysical Research: Solid Earth*, 101(B2), 3227–3246. https://doi.org/10.1029/95JB03048.
14. Hersbach, H., Bell, B., Berrisford, P., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. https://doi.org/10.1002/qj.3803.
15. Copernicus Climate Change Service (C3S). (2017). *ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate*. Copernicus Climate Data Store. https://cds.climate.copernicus.eu.
16. Natural Resources Canada (NRCan). (n.d.). *Canadian Spatial Reference System – Precise Point Positioning (CSRS-PPP).* https://webapp.csrs-scrs.nrcan-rncan.gc.ca/geod/tools-outils/ppp.php.
17. Stull, R. B. (1988). *An Introduction to Boundary Layer Meteorology*. Springer. https://doi.org/10.1007/978-94-009-3027-8.
18. Whiteman, C. D. (2000). *Mountain Meteorology: Fundamentals and Applications*. Oxford University Press. https://doi.org/10.1093/oso/9780195132717.001.0001.
19. EUREF Permanent Network (EPN). (n.d.). *Station metadata and tropospheric products*. https://epncb.oma.be/.
20. Simpson, J. E. (1994). *Sea Breeze and Local Winds*. Cambridge University Press.
21. Consorzio LaMMA. (n.d.). *Climatologia di Firenze 1991–2020*. https://www.lamma.toscana.it/clima-e-energia/climatologia/clima-firenze.
22. International GNSS Service (IGS). (2018). *SINEX_TRO Format, Version 2.00: Tropospheric Parameters Exchange Format*. https://files.igs.org/pub/data/format/sinex_tro_v2.00.pdf.
23. Paziewski, J. (2020). Recent advances and perspectives for positioning and applications with smartphone GNSS observations. *Measurement Science and Technology*, 31(9), 091001. https://doi.org/10.1088/1361-6501/ab8a7d.
24. Hatanaka, Y. (2008). A compression format and tools for GNSS observation data. *Bulletin of the Geographical Survey Institute*, 55, 21–30. https://www.gsi.go.jp/common/000045517.pdf.
25. Gurtner, W., & Estey, L. (2013). *RINEX: The Receiver Independent Exchange Format, Version 3.02*. International GNSS Service (IGS). https://files.igs.org/pub/data/format/rinex302.pdf.

26. NASA Crustal Dynamics Data Information System (CDDIS). (n.d.). *GNSS products archive (SP3, CLK, ERP)*. https://cddis.nasa.gov/archive/gnss/products/ (accessed 29 Dec 2025).

27. GFZ German Research Centre for Geosciences. (n.d.). *GNSS products (final products access)*. https://gnss.gfz.de/products.

28. Taylor, J. R. (1997). *An Introduction to Error Analysis* (2nd ed.). University Science Books.

29. Takasu, T. (n.d.). *RTKLIB: An Open Source Program Package for GNSS Positioning*. https://github.com/tomojitakasu/RTKLIB.

30. Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association*, 88(424), 1273–1283. https://doi.org/10.1080/01621459.1993.10476408.

31. Shepard, D. (1968). A two-dimensional interpolation function for irregularly spaced data. In *Proceedings of the 1968 ACM National Conference*, 517–524. https://doi.org/10.1145/800186.810616.

32. Cressie, N. (1993). *Statistics for Spatial Data* (Revised ed.). Wiley.

33. EPSG. (n.d.). *RGF93 v1 / Lambert-93 (EPSG:2154): definition and scope.* https://epsg.io/2154.