الجمهورية الجزائرية الديمقراطية الشعبية

DEMOCRATIC AND POPULAR ALGERIAN REPUBLIC

وزارة التعليم العالي والبحث العلمي

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

جامعة الشهيد حمه لخضر الوادي

ECHAHID HAMMA LAKHDAR UNIVERSITY OF EL-OUED

كلية التكنولوجيا

Faculty of Technology

قسم الهندسة الهيدروليكية والمدنية

Department of Hydraulic and Civil Engineering

Master 's Thesis

In order to obtain an Master's diploma in Technology

Specialty: in Hydraulics Structures

Them

Takagi – Sugeno Fuzzy Modeling for Estimating Reference Evapotranspiration in Batna Region: A Multi – Metric Performance Evaluation

Presented by:

Mrs. TABBOUCHA Hadjer

Jury Members:

President: Dr. MEGA Nabil	**MCA**	**El-Oued University**
Examiner: Dr. ZAID Issam	**MCA**	**El-Oued University**
Supervisor: Dr. MEZIANI Assia	**MAA**	**El-Oued University**

University year: 2025/2026

ACKNOWLEDGMENTS

﴿ وَلَئِن شَكَرْتُمْ لَأَزِيدَنَّكُمْ ﴾ — سورة إبراهيم، الآية 7

All praise is due to God, who granted me the strength and perseverance to complete this work.

My deepest gratitude goes to:

Dr. Mziani Asia

my thesis supervisor, whose scholarly rigour, constructive guidance, and unwavering support have been the cornerstone of this research. Her precise scientific observations and constant encouragement from the very first step to the final line are deeply appreciated. I extend to her my highest respect and sincere admiration.

I also wish to express my sincere thanks to the members of the examination committee

for kindly agreeing to evaluate this dissertation and for devoting their valuable time to enriching it with their expert remarks and academic insight.

My heartfelt thanks are also extended to all my distinguished professors

who contributed to my scientific and academic formation throughout my years of study, instilling in me the values of research, critical thinking, and scholarly excellence.

I am equally grateful to my colleagues and friends

for their moral support and solidarity throughout the various stages of this research.

Finally, my deepest thanks to my husband, my children, and my siblings for their patience, encouragement, and unconditional love — they were my refuge in every moment of effort and fatigue.

بِسْمِ اللهِ الرَّحْمَنِ الرَّحِيمِ

﴿ رَبِّ زِدْنِي عِلْمًا ﴾

— سورة طه، الآية 114 —

D EDICATION

To the soul of the one who taught me that knowledge is honour and patience is strength,

who fulfilled his duty and departed before witnessing the fruits of his sacrifice,

My beloved Father — may God have mercy on his soul and grant him the highest place in Paradise.

And to the one whose prayers still rise for me with every breath she takes,

My dearest Mother — may God preserve her and grant her a long, joyful life.

To my companion in life and my steadfast support in every challenge and endeavour,

My dear Husband — may God reward him on my behalf.

And to those who fill my life with joy and meaning,

My beloved Children — the light of my eyes and the warmth of my every day.

To those who shared with me a childhood full of tender memories,

My dear Siblings — you remain forever in my heart.

To the one who illuminated my path through the landscape of scientific research,

who guided with wisdom, supported with patience, and inspired with excellence,

Dr. Mziani Asia, my esteemed supervisor.

To the companions of this academic journey, my friends and colleagues.

And to every professor who gave me a word that shaped my way.

الملخص:

يُعد التقدير الدقيق للتبخر–النتح المرجعي (ET_0) عنصرًا أساسيًا في التسيير الفعال للموارد المائية، خاصة في المناطق شبه الجافة التي تتميز بالتقلبات المناخية ونقص المعطيات المناخية. تهدف هذه الدراسة إلى تطوير وتقييم نظام استدلال ضبابي من نوع تاكاغي–سوغينو من الدرجة الأولى (TS-FIS) لتقدير التبخر–النتح المرجعي في منطقة باتنة شمال شرق الجزائر، اعتمادًا على بيانات مناخية تم جمعها من 24 محطة أرصاد جوية موزعة عبر الولاية.

تم اعتماد معادلة FAO-56 Penman–Monteith كمرجع معياري لحساب قيم ET_0. وقد تم تنفيذ النموذج المقترح باستخدام لغة Python بالاعتماد على خوارزمية التجميع الضبابي Fuzzy C-Means (FCM) لاستخراج القواعد الضبابية تلقائيًا، بالإضافة إلى الانحدار بالمربعات الصغرى الموزونة لتقدير معاملات النتائج.

تم إجراء تحليل ارتباط بيرسون لاختيار أكثر المتغيرات المناخية تأثيرًا على تغيرات ET_0 ، حيث تم تحديد:

عجز ضغط البخار(VPD)

درجة حرارة الهواء على ارتفاع مترين (T_{2m})

ودرجة حرارة التربة على عمق 0–7 سم (T_soil)

كأهم المتغيرات المدخلة للنموذج. تم تقييم أداء النموذج باستخدام عدة مؤشرات إحصائية، من بينها:

R^2، RMSE، MAE، NSE، ومعامل ويلموت .(d) وأظهرت النتائج قدرة تنبؤية عالية للنموذج عبر مختلف المحطات، حيث بلغت القيم المتوسطة على مستوى الشبكة:

$R^2 = 0.9006$

RMSE = 0.667 مم/يوم

MAE = 0.534 مم/يوم

NSE = 0.8963

d = 0.9728

وأكدت النتائج المتحصل عليها أن نموذج TS-FIS المقترح يمثل بديلاً فعالًا وموثوقًا لتقدير التبخر–النتح المرجعي في البيئات شبه الجافة، خاصة في حالات نقص البيانات المناخية الكاملة.

الكلمات المفتاحية:

التبخر–النتح المرجعي، نظام تاكاغي–سوغينو الضبابي، FAO-56 Penman–Monteith، تقدير ET_0 ، المناخ شبه الجاف، منطقة باتنة.

Abstract

Accurate estimation of reference evapotranspiration (ET_0) is essential for effective water resource management, particularly in semi-arid regions characterized by climatic variability and limited meteorological observations. This study develops and evaluates a first-order Takagi–Sugeno fuzzy inference system (TS-FIS) for estimating ET_0 in the Batna region of northeastern Algeria using meteorological data collected from 24 stations distributed across the wilaya.

The FAO-56 Penman–Monteith equation was adopted as the reference benchmark. The proposed TS-FIS model was implemented in Python using Fuzzy C-Means (FCM) clustering for automatic rule extraction and membership-weighted least-squares regression for consequent parameter estimation.

A Pearson correlation analysis was performed to identify the most influential climatic variables controlling ET_0 variability. The selected inputs were vapour pressure deficit (VPD), 2-m air temperature (T_{2m}), and soil temperature at 0–7 cm (T_soil).

The model performance was evaluated using several statistical indicators including R^2, RMSE, MAE, NSE, and Willmott's index. Results demonstrated strong predictive capability across the 24 stations, with network-average values of $R^2 = 0.9006$, RMSE = 0.667 mm/day, MAE = 0.534 mm/day, NSE = 0.8963, and d = 0.9728.

The obtained results confirm that the proposed TS-FIS model constitutes a reliable and efficient alternative for ET_0 estimation in semi-arid environments, especially under conditions of incomplete climatic data.

Keywords: Reference evapotranspiration, Takagi–Sugeno fuzzy inference system, FAO-56 Penman–Monteith, ET_0 estimation, semi-arid climate, Batna region.

TABLE OF CONTENTS

Acknowledgments..........

Dedication..........

Abstract..........

TABLE OF CONTENTS..........

FIGURES LIST..........

TABLES LIST..........

LIST OF EQUATION..........

GENERAL INTRODUCTION..........

Chapter I: GENERAL PRESENTATION OF THE STUDY AREA

I. Introduction 4

I.1. Geographical Location and Administrative Setting 4

I.2. Topography and Geomorphology 5

I.3. Climate and Hydrological Characteristics 6

I.3.1 Climatic Drivers and Thermodynamic Interactions 7

I.3.2 Climatic Study of Batna Station 8

I.3.2.a. Correlation matrix: 17

I.3.2.b. Conclusion 18

I.4. Meteorological Stations and Data Sources 19

I.5. Water Resources and Agricultural Context 20

I.6 Conclusion 21

Chapter II: EVAPOTRANSPIRATION MODELING

II. Introduction 23

II.1. Reference Evapotranspiration (ET_0) 23

II.1.1. Definition and Physical Significance 23

II.1.2. Climatic Drivers 23

II.2. The FAO-56 Penman–Monteith Equation 23

II.3. Classical Empirical Methods 25

II.4. Artificial Intelligence Approaches for ET_0 Estimation 25
II.4.1. Comparative Overview of Candidate AI Models 26
II.5. Fuzzy Logic Framework 26
II.6 Conclusion 28

Chapter III: TAKAGI–SUGENO FUZZY PROPOSED MODEL

III. Introduction 30
III.1. Data Preparation and Pre-processing 30
III.1.1. Data Sources and Meteorological Variables 30
III.1.2. Feature Selection via Pearson Correlation 31
III.1.3. Dataset Partitioning 31
III.2. Takagi–Sugeno Fuzzy Model Architecture 32
III.2.1. Model Structure and Rule Base 32
III.2.2. Fuzzy C-Means Antecedent Identification 33
III.2.3. Consequent Parameterization by Weighted Least-Squares 33
III.2.4. Inference Mechanism and Output Aggregation 34
III.3. Bias Correction Procedure 34
III.4. Multi-Metric Performance Evaluation Framework 36
III.5. Proposed TS-FIS Modeling 37
III.6.1. Python Implementation Environment 38
III.7 Conclusion 40

Chapter IV: RESULTS AND DISCUSSION

IV. Introduction 42
IV.1. Overall Model Performance on the Test Set 42
IV.2. Scatter Plot Analysis: Observed vs. Predicted ET_0 43
IV.2.1. General Patterns Across the Station Network 43
IV.2.2. Station-Specific Scatter Plot Analysis 44
IV.3. Time-Series Analysis: Temporal Dynamics of Predictions 54
IV.3.1. General Temporal Behaviour 54
IV.3.2. Detailed Station-by-Station Time-Series Analysis 55
IV.4. Model Stability and Overfitting Assessment 67
IV.5. Spatial Analysis of Prediction Error via Ordinary Kriging 69
IV.5.1. Spatial Error Analysis by Ordinary Kriging 69
IV.5.2. Long-Term Mean Error Distribution (2000–2025) 70

IV.5.3. Inter-Annual Variability of Spatial Error Patterns (2000–2025) 71
IV.5.4. Implications for Model Application 80
IV.6. Discussion and Comparative Assessment 81
IV.7 Conclusion 82

GENERAL CONCLUSION AND RECOMMENDATIONS 83

References 87

FIGURES LIST

Figure I.1. Location map of the 24 meteorological stations in the Batna wilaya, northeastern Algeria. Station numbers correspond to Table I.1. 5

Figure I.2. Altitude profile of the 24 meteorological stations in the Batna wilaya. Colour gradient indicates altitude class (green = low, red = high). 5

Figure I.3. Climatic diagram of Batna city: mean monthly air temperature, precipitation, and FAO-56 PM ET_0. 6

Figure I.4. Monthly Variation of Mean Air Temperature (°C) at Batna Meteorological Station (2000–2025) 10

Figure I.5. Monthly Variation of Relative Humidity (%) at Batna Meteorological Station (2000–2025) 10

Figure I.6. Monthly Variation of Dew Point Temperature (°C) at Batna Meteorological Station (2000–2025) 11

Figure I.7. Monthly Variation of Precipitation (mm/day) at Batna Meteorological Station (2000–2025) 12

Figure I.8. Monthly Variation of Surface Pressure (hPa) at Batna Meteorological Station (2000–2025) 12

Figure I.9. Monthly Variation of Reference Evapotranspiration ET_0 (mm/day) at Batna Meteorological Station (2000–2025) 13

Figure I.10. Monthly Variation of Vapour Pressure Deficit (kPa) at Batna Meteorological Station (2000–2025) 14

Figure I.11. Monthly Variation of Wind Speed at 10 m (km/h) at Batna Meteorological Station (2000–2025) 14

Figure I.12. Monthly Variation of Soil Temperature 0–7 cm (°C) at Batna Meteorological Station (2000–2025) 15

Figure I.13. Monthly Variation of Soil Moisture 0–7 cm (m^3/m^3) at Batna Meteorological Station (2000–2025) 16

Figure I.14. Monthly Variation of Sunshine Duration (h/day) at Batna Meteorological Station (2000–2025) 16

Figure I.15. Monthly Variation of Terrestrial Radiation (W/m^2) at Batna Meteorological Station (2000–2025) 17

Figure I.16. Correlation matrix of climatic parameters at Batna Meteorological Station (2000–2025) 18

Figure II.1. Mean monthly ET_0 cycle: FAO-56 PM (observed) vs. T–S fuzzy model (predicted), over 24 stations. Shaded band = min–max range. 25
Figure II.2. Computational Intelligence Framework. 27
Figure III.1. T-S Fuzzy Inference System Framework for ET_0 *Estimation in Semi – Arid Algeria.* 31
Figure III.2. Theoretical framework of Reference Evapotranstpiration (ET_0) 32
Figure III.3. Performance metrics used for TS-FIS evaluation. 36
Figure III.4. Organigram of the TS-FIS modeling workflow applied to the weather stations of the Batna wilaya. Each box represents one stage of the pipeline; arrows indicate the data flow from ERA5-Land input extraction to final nine-metric performance evaluation. 38
Figure III.5. Organigram of the TS-FIS modeling workflow applied to the 24 meteorological stations of the Batna wilaya. Each box represents one stage of the pipeline; arrows indicate the data flow from ERA5-Land input extraction to final nine-metric performance evaluation. 39
Figure IV.1: Scatter plot of Ain touta station 44
Figure IV.2: Scatter plot of Sefiane station 44
Figure IV.3: Scatter plot of Seggana station 45
Figure IV.4: Scatter plot of Maafa station 45
Figure IV.5: Scatter plot of Gosbat station 46
Figure IV.6: Scatter plot of Metkaouak station 46
Figure IV.7: Scatter plot of Bitam station 47
Figure IV.8: Scatter plot of Timgad station 47
Figure IV.9: Scatter plot of Arris station 48
Figure IV.10: Scatter plot of Bouzina station 48
Figure IV.12: Scatter plot of Menaa station 49
Figure IV.13: Scatter plot of Chemora station 49
Figure IV.14: Scatter plot of Lazrou station 50
Figure IV.15: Scatter plot of Seriana station 50
Figure IV.17: Scatter plot of Batna station 51
Figure IV.18: Scatter plot of Ain yagout station 51
Figure IV.19: Scatter plot of Fom toub station 51
Figure IV.20: Scatter plot of Chir station 52
Figure IV.21: Scatter plot of Hidoussa station 52
Figure IV.22: Scatter plot of Merouana station 53
Figure IV.23: Scatter plot of Tazoult station 53

Figure IV.24: Scatter plot of Teniet El Abed station 54
Figure IV.25: Time Series Analysis of Ain touta station 55
Figure IV.26: Time Series of Gosbat station 56
Figure IV.27: Time Series of Hidoussa station 56
Figure IV.28: Time Series of Seriana station 57
Figure IV.29: Time Series of Metaouak station 57
Figure IV.30: Time Series of Bitam station 58
Figure IV.31: Time Series of Tazoult station 58
Figure IV.32: Time Series of Bouzina station 59
Figure IV.33: Time Series of Sefiane station 59
Figure IV.34: Time Series of Seggana station 60
Figure IV.35: Time Series of Merouana station 60
Figure IV.36: Time Series of El Madher station 61
Figure IV.37: Time Series of Batna station 61
Figure IV.38: Time Series of Ain yagout station 61
Figure IV.39: Time Series of Menaa station 62
Figure IV.40: Time Series of Arris station 62
Figure IV.41: Time Series of Ichemoul station 63
Figure IV.42: Time Series of Foum Toub station 63
Figure IV.43: Time Series of Lazrou station 64
Figure IV.44: Time Series of Teniet El Abed station 64
Figure IV.45: Time Series of Chemora station 65
Figure IV.46: Time Series of Chir station 65
Figure IV.47: Time Series of Maafa station 66
Figure IV.48: Time Series of Timgad station 67
Figure IV.49: Kriging Maps of 24 Station in the Batna Region For 2000 to 2025 70
Figure IV.50: Kriging Maps of 24 Station in the Batna Region For the year 2000 71
Figure IV.51: Kriging Maps of 24 Station in the Batna Region For the year 2001 71
Figure IV.52: Kriging Maps of 24 Station in the Batna Region For the year 2002 72
Figure IV.53: Kriging Maps of 24 Station in the Batna Region For the year 2003 72
Figure IV.54: Kriging Maps of 24 Station in the Batna Region For the year 2004 72
Figure IV.55: Kriging Maps of 24 Station in the Batna Region For the year 2005 73
Figure IV.56: Kriging Maps of 24 Station in the Batna Region For the year 2006 73
Figure IV.57: Kriging Maps of 24 Station in the Batna Region For the year 2007 73
Figure IV.58: Kriging Maps of 24 Station in the Batna Region For the year 2008 74

Figure IV.59: Kriging Maps of 24 Station in the Batna Region For the year 2009 74
Figure IV.60: Kriging Maps of 24 Station in the Batna Region For the year 2010 74
Figure IV.61: Kriging Maps of 24 Station in the Batna Region For the year 2011 75
Figure IV.62: Kriging Maps of 24 Station in the Batna Region For the year 2012 75
Figure IV.63: Kriging Maps of 24 Station in the Batna Region For the year 2013 75
Figure IV.64: Kriging Maps of 24 Station in the Batna Region For the year 2014 76
Figure IV.65: Kriging Maps of 24 Station in the Batna Region For the year 2015 76
Figure IV.66: Kriging Maps of 24 Station in the Batna Region For the year 2016 76
Figure IV.67: Kriging Maps of 24 Station in the Batna Region For the year 2017 77
Figure IV.68: Kriging Maps of 24 Station in the Batna Region For the year 2018 77
Figure IV.69: Kriging Maps of 24 Station in the Batna Region For the year 2019 77
Figure IV.70: Kriging Maps of 24 Station in the Batna Region For the year 2020 78
Figure IV.72: Kriging Maps of 24 Station in the Batna Region For the year 2022 78
Figure IV.73: Kriging Maps of 24 Station in the Batna Region For the year 2023 79
Figure IV.74: Kriging Maps of 24 Station in the Batna Region For the year 2024 79
Figure IV.75: Kriging Maps of 24 Station in the Batna Region For the year 2025 79

TABLES LIST

Table I.1. Monthly Climatic Summary (2000-2025) (at Batna meteorological station) 8
Table I.2. Mean Air Temperature (2000-2025) (Batna meteorological station) 9
Table I. 3. Relative Humidity (2000-2025) (Batna meteorological station) 10
Table I.4. Dew Point Temperature 11
Table I.5. Precipitation (2000-2025) (Batna meteorological station)..................................... 11
Table I.6. Surface Pressure (2000-2025) (Batna meteorological station) 12
Table I.7. Reference Evapotranspiration (2000-2025) (Batna meteorological station)........... 13
Table I.8. Vapour Pressure Deficit (2000-2025) (Batna meteorological station) 13
Table I.9. Wind Speed at 10 m (2000-2025) (Batna meteorological station).......................... 14
Table I.10. Soil Temperature (0–7 cm) (2000-2025) (Batna meteorological station)............. 15
Table I.11. Soil Moisture (0–7 cm) (2000-2025) (Batna meteorological station)................... 15
Table I.12. Sunshine Duration (2000-2025) (Batna meteorological station) 16
Table I.13. Terrestrial Radiation (2000-2025) (Batna meteorological station) 17
Table I.14. List of the 24 meteorological stations used in the study (ONM, Algeria). 19
Table III.1 – Performance metrics used for TS-FIS evaluation. P = Predicted, O = Observed, n = sample size, $\bar{O}$ = mean observed ET_0, σ = standard deviation 37
Table IV.1 – TS-FIS performance metrics on the independent test set (24 stations). RMSE, MAE, MBE in mm day^{-1}. R^2 = coefficient of determination, NSE = Nash-Sutcliffe efficiency, d = Willmott index of agreement, RPD = Ratio of Performance to Deviation, RSR = RMSE/σ_obs. 42
Table IV.2 – Model stability across data partitions. Δ = Test RMSE minus Train RMSE (mm day^{-1}). Negative values indicate better test than train performance. Classification: |Δ|<0.010 = Excellent; 0.010–0.020 = Very Good; >0.020 = Good 68

LIST OF EQUATION

N°	*Description of Equation*	*Page*
Eq. II.1	FAO-56 Penman–Monteith Reference Evapotranspiration Equation	24
Eq. III.1	Takagi–Sugeno Fuzzy Rule Antecedent (IF part)	33
Eq. III.2	Takagi–Sugeno Fuzzy Rule Consequent (THEN part) — Local Linear Model	33
Eq. III.3	Fuzzy C-Means (FCM) Objective Function	33
Eq. III.4	Weighted Least-Squares Consequent Parameterization	33
Eq. III.5	TS-FIS Inference Mechanism — Weighted Output Aggregation	34
Eq. III.6	Additive Bias Correction Formula	34

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Water is among the most critical natural resources in semi-arid regions. Reference evapotranspiration (ET_0) constitutes the primary driver of atmospheric water demand and is indispensable for rational irrigation scheduling, crop water requirement computation, and regional hydrological modelling. **(Allen et al., 1998 [1]; Monteith, 1965 [16])**

The FAO-56 Penman–Monteith equation is the internationally accepted standard for ET_0 computation; however, its operational application in data-scarce semi-arid environments such as Batna Province is frequently hampered by the unavailability of complete meteorological records. **(Allen et al., 1998 [1]; Ladlani et al., 2014 [15])**

Data-driven soft-computing approaches have emerged as powerful alternatives, capable of learning the non-linear ET_0 function from available meteorological observations. The Takagi–Sugeno Fuzzy Inference System (TS-FIS) combines the universal approximation capability of fuzzy logic with the computational efficiency of linear rule consequents and the physical interpretability of transparent fuzzy rule bases. **(Takagi & Sugeno, 1985 [2]; Zadeh, 1965 [3]; Kisi, 2013 [14])**

The Batna Province, situated in northeastern Algeria at the convergence of the Tell Atlas, the Aurès Massif, and the pre-Saharan fringe, is characterized by a semi-arid BSk climate and wide seasonal ET_0 amplitudes (from <1.5 mm day^{-1} in winter to >9.5 mm day^{-1} in summer), making accurate ET_0 estimation strategically important for regional water resource management. **(ONM, 2025 [19]; Muñoz-Sabater et al., 2021 [22])**

This master thesis develops and evaluates a first-order TS-FIS for daily ET_0 estimation across 24 meteorological stations of the Batna wilaya using 26 years (2000–2025) of ERA5-Land reanalysis data, with the FAO-56 Penman–Monteith ET_0 as target and a nine-metric evaluation framework on an independent test set. **(Bezdek, 1981 [12]; Nash & Sutcliffe, 1970 [9]; Moriasi et al., 2007 [11])**

Objectives of this study:

- Develop a first-order T–S fuzzy model for ET_0 estimation in the Batna region using 24 meteorological stations.
- Implement the complete modeling pipeline (FCM clustering, least-squares identification, prediction) in Python.
- Perform a rigorous multi-metric validation using R^2, RMSE, MAE, NSE, and Willmott's d index.

- Compare the T–S model performance against classical empirical methods (Hargreaves–Samani, Priestley–Taylor, Blaney–Criddle).
- Analyze the spatial distribution of model performance across the wilaya.

_ This thesis is organized into four chapters:

- Chapter I: presents the geographic and climatic framework of the Batna wilaya, describing the 24 meteorological stations and the ERA5-Land dataset used as input.

- Chapter II: reviews the relevant literature on reference evapotranspiration estimation methods, from the FAO-56 Penman–Monteith standard to data-driven approaches, and justifies the choice of a Takagi–Sugeno FIS.

- Chapter III: details the complete TS-FIS methodology: data processing, feature selection, FCM clustering, consequent parameter estimation, and the nine-metric evaluation framework.

- Chapter IV: presents and discusses the model performance results across all 24 stations, including scatter plots, time-series validation, and Ordinary Kriging spatial analysis of prediction errors.

Chapter I

GENERAL PRESENTATION OF THE STUDY AREA

I. Introduction

The Batna Province, located in northeastern Algeria at the junction of the Tell Atlas, the Aurès Massif, and the pre-Saharan fringe, presents one of the most physiographically and climatically complex environments in the country. Its wide altitudinal range (800–2,328 m a.s.l.) generates strong spatial gradients in temperature, precipitation, and evaporative demand, making it a challenging and representative testbed for evaluating the proposed Takagi–Sugeno Fuzzy Inference System. The region also faces acute water-resource pressures driven by recurring drought episodes (2018–2023), reinforcing the operational importance of accurate daily ET_0 estimation for rational irrigation scheduling across the 24 meteorological stations studied here.

I.1. Geographical Location and Administrative Setting

The Batna wilaya is situated in northeastern Algeria between 34°40′–36°20′N and 5°00′–7°10′E, covering 12,038 km², subdivided into 21 daïras and 61 communes with a population exceeding 1.3 million. The provincial capital lies at 1,058 m a.s.l. (35°33′N, 6°10′E) in the Aurès piedmont. **(ONM, 2025 [19])**

The wilaya is bounded by Mila, Constantine and Oum El Bouaghi to the north, Khenchela to the east, Biskra to the south, and M'Sila and Sétif to the west. This position places it at the convergence of the Tell Atlas, the Aurès Massif, and the pre-Saharan plateau — three major structural domains generating pronounced spatial gradients in temperature, precipitation, and ET_0. **(ONM, 2025 [19]; Muñoz-Sabater et al., 2021 [22]).** *According to Figure I.1*.

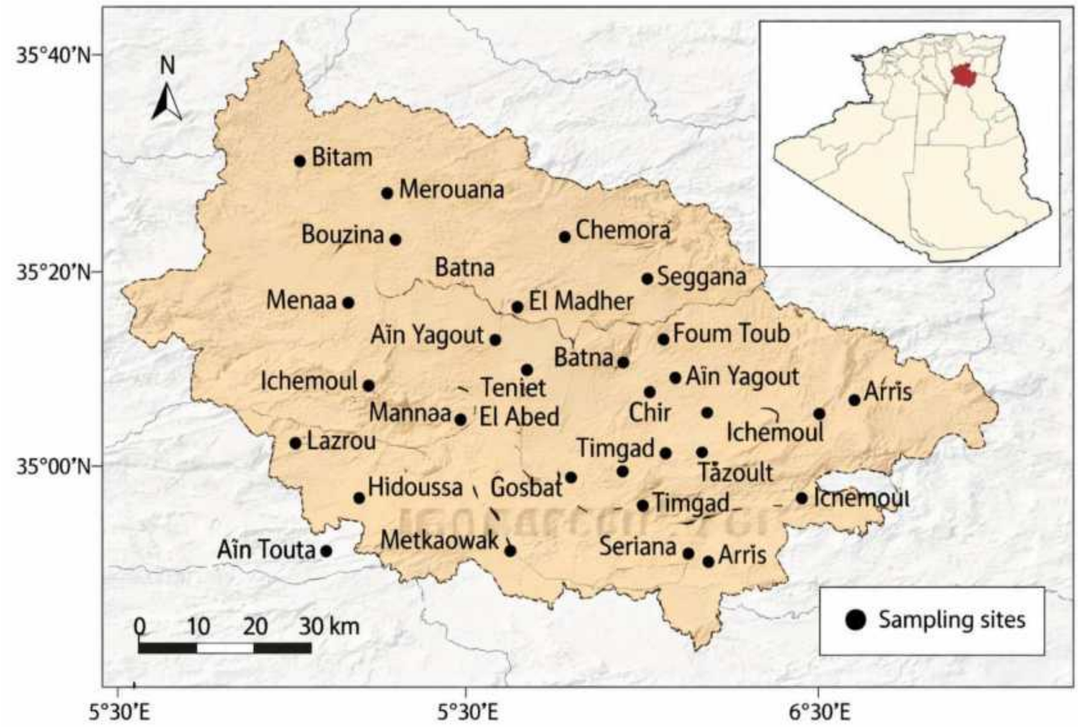

Figure I.1. Location map of the 24 meteorological stations in the Batna wilaya, northeastern Algeria.

Station numbers correspond to Table I.1.

I.2. Topography and Geomorphology

The Aurès Massif dominates the eastern and southern sectors. The Djebel Chélia (2,328 m a.s.l.), the highest peak in northern Algeria, exerts a strong orographic influence on regional precipitation and wind regimes. Three physiographic units are distinguished: the Aurès mountains (1,200–2,328 m), the Hauts Plateaux (900–1,100 m), and the pre-Saharan fringe (<850 m). **(ONM, 2025 [19]; Hersbach et al., 2020 [21]).** *According to Figure I.2*

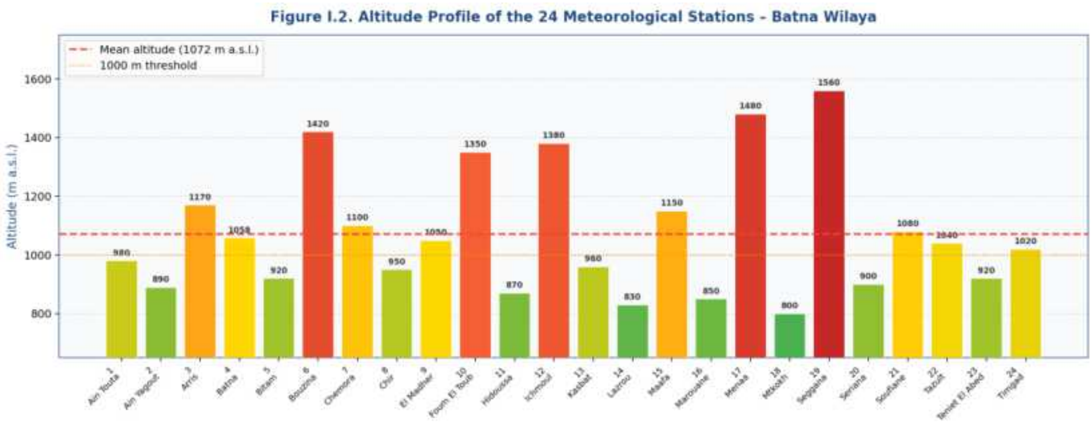

Figure I.2. Altitude profile of the 24 meteorological stations in the Batna wilaya.

Colour gradient indicates altitude class (green = low, red = high).

I.3. Climate and Hydrological Characteristics

The region experiences a semi-arid climate (Köppen BSk), with mean annual precipitation ranging from ~180 mm in the south to ~430 mm at higher-elevation western stations, and mean annual temperatures between 10°C and 16°C depending on altitude. Daily ET_0 peaks at 9.5–10 mm day^{-1} in July–August and drops below 1.5 mm day^{-1} during winter. **(Hersbach et al., 2020 [21]; Muñoz-Sabater et al., 2021 [22]; Allen et al., 1998 [1]; Tabari, 2010 [24])**

ERA5-Land reanalysis data were adopted as the primary meteorological source, providing physically consistent, gap-free daily climatic fields at ~9 km resolution demonstrated to reliably reproduce observed surface conditions in semi-arid North African regions. **(Muñoz-Sabater et al., 2021 [22]; Hersbach et al., 2020 [21]).** *According to Figure I.3*

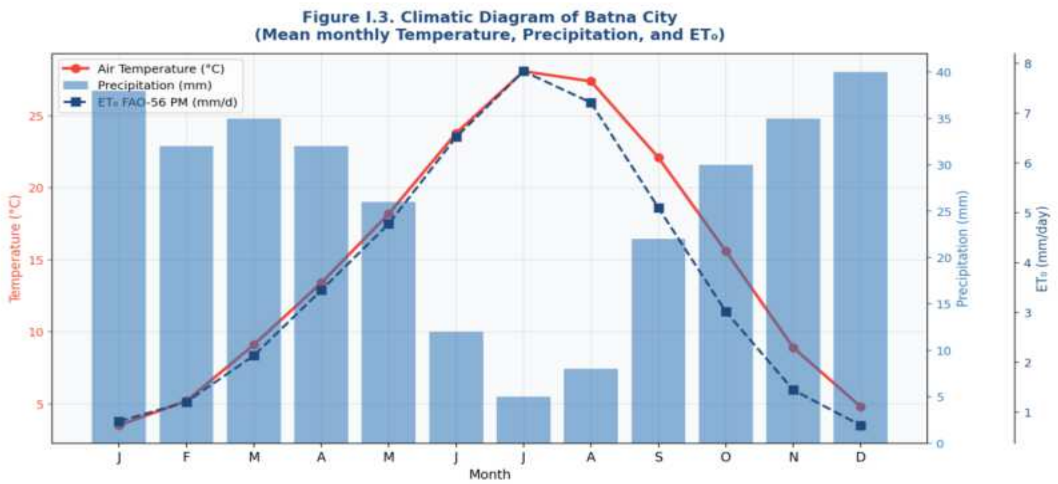

Figure I.3. Climatic diagram of Batna city: mean monthly air temperature, precipitation, and FAO-56 PM ET_0.

- Annual precipitation ranges from ~180–183 mm in drier lowland stations to ~430 mm at higher-elevation western stations (Arris, Foum El Toub).
- Mean annual temperatures vary between ~10°C at high-altitude stations and ~16°C in lower-lying areas.
- Summer temperatures regularly exceed 38°C in the plains, driving peak ET_0 values (up to 9.76 mm/day) in July–August.
- Relative humidity averages 40–60% annually, with pronounced lows during summer months (often below 30%).

Hydrologically, the region is characterized by intermittent surface runoff and flashflood-prone watercourses. Drought is a recurring feature; SPI-based studies have documented severe drought episodes during 2018–2023, making the accurate estimation of ET_0 strategically critical for water resource planning.

I.3.1 Climatic Drivers and Thermodynamic Interactions

The rate of ET0 is governed by a delicate balance of energy availability and mass transfer. Four key meteorological variables dominate this interaction:

1. **Solar Radiation:** This is the ultimate energy source for the phase change of water from liquid to vapor. In regions like Batna, high levels of solar irradiance during the summer months lead to peak ET_0 values, making it the most significant driver of the process.
2. **Air Temperature:** Temperature influences the energy level of water molecules and the water vapor holding capacity of the air. Higher temperatures increase the saturation vapor pressure, thereby enhancing the potential for evaporation.
3. **Relative Humidity:** The difference between the vapor pressure at the leaf/soil surface and the surrounding air (vapor pressure deficit) is the driving force for moisture movement. In semi-arid climates, low humidity levels create a "thirsty" atmosphere that accelerates water loss.
4. **Wind Speed:** Wind acts as a transport mechanism, removing the saturated air layer immediately above the surface and replacing it with drier air, thus maintaining a steep vapor pressure gradient.
5. **Sunshine Duration and Daylight Hours:** The actual sunshine duration (n, hours day^{-1}) records the period during which direct solar irradiance exceeds 120 W m^{-2}, while N denotes the maximum astronomically possible daylight duration computed from latitude and day of year. Their ratio n/N (0 = fully overcast; 1 = clear sky) enters the Angström–Prescott formula used within FAO-56 to derive net solar radiation, and simultaneously serves as a fog/overcast detection index: days with $n/N < 0.25$ are classified as foggy or heavily overcast, $0.25 \leq n/N < 0.65$ as partly cloudy, and $n/N \geq 0.65$ as sunny. Applied to the 26-year ERA5-Land record (2000–2025) of all 24 Batna stations (227,928 station-days), this classification yields 92.1 % sunny days, 6.0 % partly cloudy days, and 1.9 % foggy/overcast days, with a mean n/N of 0.856. Foggy episodes are strongly seasonal: the overcast frequency reaches 3.4 % in winter (December–February) versus near zero in summer (June–August), peaking in December (≈1.4 foggy days station^{-1} month^{-1}). The most affected stations are Foum Toub, Chemora, and Timgad (≈3.0 % foggy days), while Menaa, Seggana, and Bitam record the

fewest (<1.3 %). These anomalously low-radiation days depress ET_0 well below its clear-sky potential and were flagged during data pre-processing to prevent bias in the TS-FIS training set.

The synergy between these variables creates a highly non-linear system. For instance, a high temperature combined with low wind speed may result in lower ET_0 than a moderate temperature paired with high wind speeds. This non-linearity is exactly why traditional linear models often fail, necessitating the use of advanced modeling techniques like fuzzy logic.[1]

I.3.2 Climatic Study of Batna Station

This chapter presents a detailed climatic analysis of the Batna meteorological station based on the provided daily climatic observations. Monthly averages were calculated for all climatic parameters to evaluate seasonal variability and environmental conditions characterizing the study area. *According to Table I.1*

Table I.1. Monthly Climatic Summary (2000-2025) (at Batna meteorological station)

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mean Air Temp. (°C)	5.35	6.18	9.42	12.95	17.05	22.61	26.5	25.31	20.71	16.06	10.0	6.41
Relative Humidity (%)	70.92	67.51	62.57	59.67	56.36	45.5	36.76	41.44	55.01	60.59	68.34	73.45
Dew Point Temp. (°C)	-0.22	-0.32	1.38	3.79	6.6	7.96	8.23	9.1	9.77	7.14	3.58	1.38
Precipitation (mm/day)	1.13	1.09	1.53	1.82	1.97	0.77	0.39	0.73	1.28	1.12	1.18	1.12
Surface Pressure (hPa)	900.8	899.92	898.66	897.79	900.25	903.02	904.38	904.02	903.42	902.4	899.7	901.18

ET_0 FAO-56 PM (mm/day)	1.57	2.14	3.08	4.08	4.92	6.18	6.82	5.99	4.37	3.16	2.01	1.44
Vapour Press. Deficit (kPa)	0.32	0.39	0.57	0.76	1.08	1.83	2.53	2.2	1.34	0.88	0.48	0.31
Wind Speed 10 m (km/h)	8.79	9.08	9.23	8.81	8.28	8.09	8.23	7.79	7.56	7.82	8.91	8.53
Soil Temp. 0–7 cm (°C)	5.12	6.11	9.34	12.98	16.98	22.4	26.45	25.44	21.05	16.34	10.49	6.52
Soil Moisture 0–7 cm (m^3/m^3)	0.29	0.29	0.29	0.28	0.26	0.21	0.17	0.18	0.21	0.23	0.25	0.28
Sunshine Duration (h/day)	7.9	8.89	9.72	10.78	11.52	12.54	12.75	12.04	10.36	9.42	8.12	7.6
Terrestrial Radiation (W/m^2)	210.86	270.46	347.49	417.84	465.78	483.21	471.96	431.76	369.6	292.75	225.48	192.82

Table I.2. Mean Air Temperature (2000-2025) (Batna meteorological station)

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mean Air Temperature (°C)	5.35	6.18	9.42	12.95	17.05	22.61	26.5	25.31	20.71	16.06	10.0	6.41

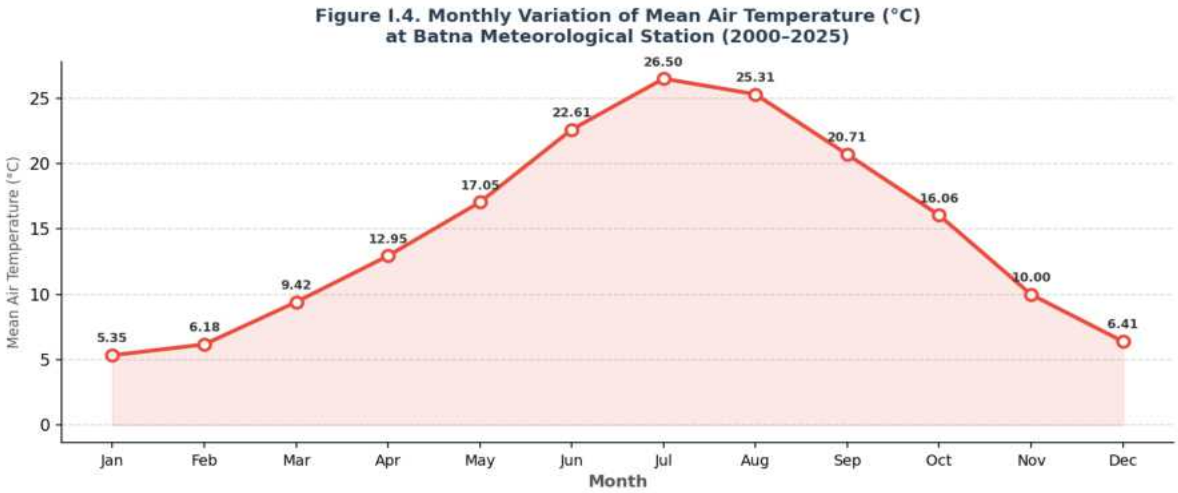

Figure I.4. Monthly Variation of Mean Air Temperature (°C) at Batna Meteorological Station (2000–2025)

The monthly variation of mean air temperature at Batna station indicates a mean annual value of approximately 14.88. The highest monthly average was recorded in Jul, while the lowest values occurred in Jan. The seasonal fluctuations reflect the continental semi-arid climate characteristic of the Batna region. *According to Table I.2 / Figure I.4*.

Table I. 3. Relative Humidity (2000-2025) (Batna meteorological station)

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Relative Humidity (%)	70.92	67.51	62.57	59.67	56.36	45.5	36.76	41.44	55.01	60.59	68.34	73.45

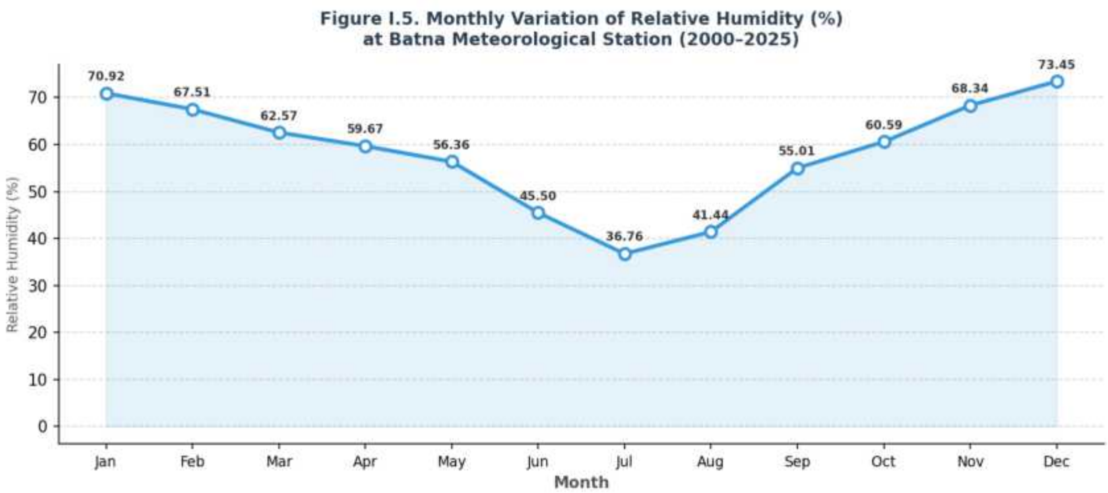

Figure I.5. Monthly Variation of Relative Humidity (%) at Batna Meteorological Station (2000–2025)

The monthly variation of relative humidity at Batna station indicates a mean annual value of approximately 58.18. The highest monthly average was recorded in Dec, while the lowest values occurred in Jul. The seasonal fluctuations reflect the continental semi-arid climate characteristic. *According to Table I.3 / Figure I.5.*

Table I.4. Dew Point Temperature

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Dew Point Temp. (°C)	-0.22	-0.32	1.38	3.79	6.6	7.96	8.23	9.1	9.77	7.14	3.58	1.38

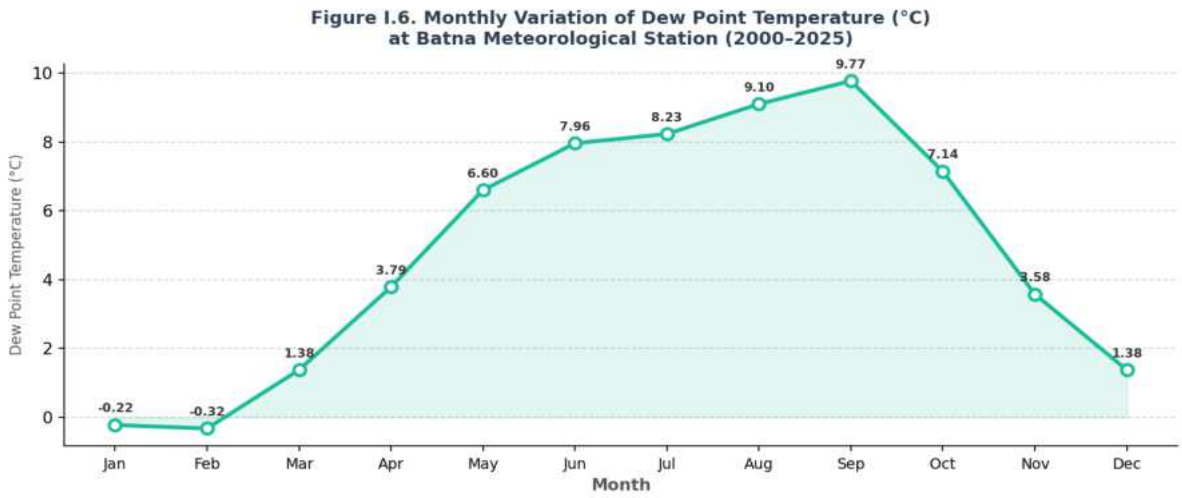

Figure I.6. Monthly Variation of Dew Point Temperature (°C) at Batna Meteorological Station (2000–2025)

The monthly variation of dew point temperature at Batna station indicates a mean annual value of approximately 4.86. The highest monthly average was recorded in Sep, while the lowest values occurred in Feb. The seasonal fluctuations reflect the continental semi-arid climate characteristic of the Batna region. *According to Table I.4 / Figure I.6.*

Table I.5. Precipitation (2000-2025) (Batna meteorological station)

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Precipitation (mm/day)	1.13	1.09	1.53	1.82	1.97	0.77	0.39	0.73	1.28	1.12	1.18	1.12

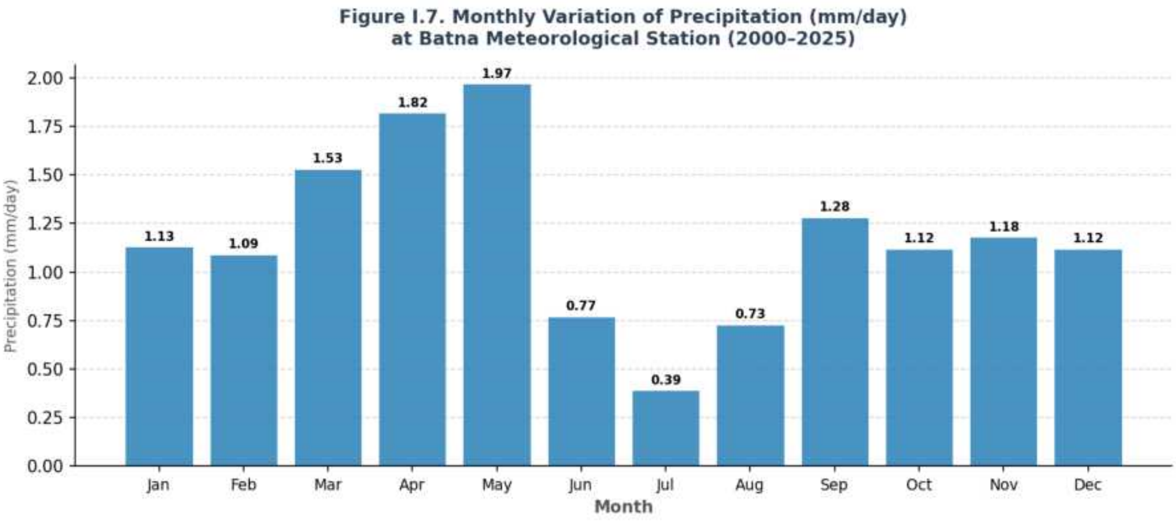

Figure I.7. Monthly Variation of Precipitation (mm/day) at Batna Meteorological Station (2000–2025)

The monthly variation of precipitation at Batna station indicates a mean annual value of approximately 1.18. The highest monthly average was recorded in May, while the lowest values occurred in Jul. The seasonal fluctuations reflect the continental semi-arid climate characteristic of the Batna region. *According to Table I.5 / Figure I.7.*

Table I.6. Surface Pressure (2000-2025) (Batna meteorological station)

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Surface Pressure (hPa)	900.8	899.92	898.66	897.79	900.25	903.02	904.38	904.02	903.42	902.4	899.7	901.18

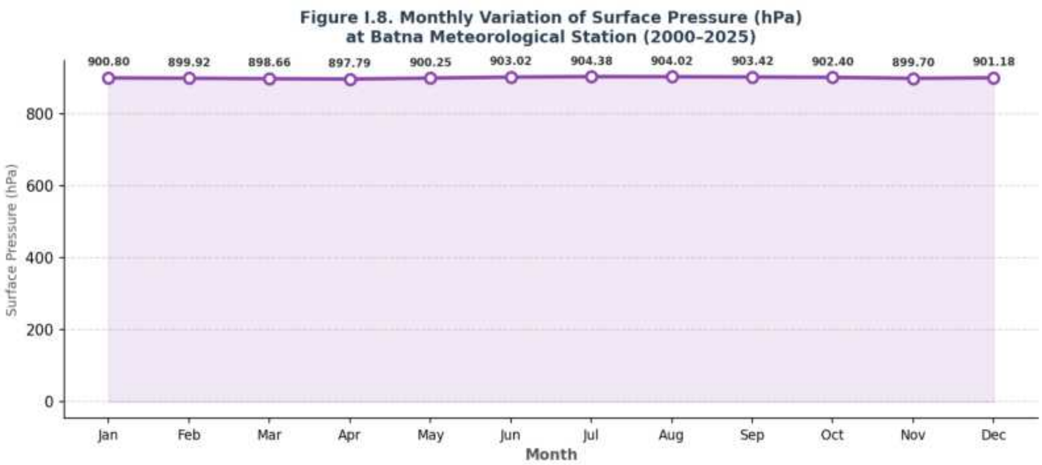

Figure I.8. Monthly Variation of Surface Pressure (hPa) at Batna Meteorological Station (2000–2025)

The monthly variation of surface pressure at Batna station indicates a mean annual value of approximately 901.29. The highest monthly average was recorded in Jul, while the lowest values occurred in Apr. The seasonal fluctuations reflect the continental semi-arid climate characteristic of the Batna region. *According to Table I.6 / Figure I.8*

Table I.7. Reference Evapotranspiration (2000-2025) (Batna meteorological station)

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
ET_0 FAO-56 PM (mm/day)	1.57	2.14	3.08	4.08	4.92	6.18	6.82	5.99	4.37	3.16	2.01	1.44

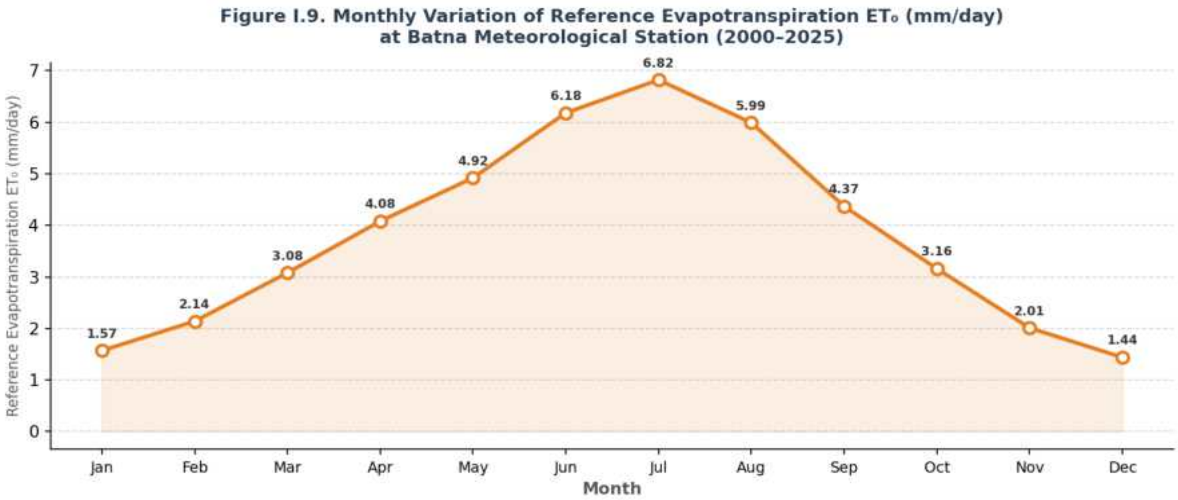

Figure I.9. Monthly Variation of Reference Evapotranspiration ET_0 (mm/day) at Batna Meteorological Station (2000–2025)

The monthly variation of reference evapotranspiration at Batna station indicates a mean annual value of approximately 3.81. The highest monthly average was recorded in Jul, while the lowest values occurred in Dec. The seasonal fluctuations reflect the continental semi-arid climate characteristic of the Batna region. *According to Table I.7 / Figure I.9.*

Table I.8. Vapour Pressure Deficit (2000-2025) (Batna meteorological station)

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Vapour Pressure Deficit (kPa)	0.32	0.39	0.57	0.76	1.08	1.83	2.53	2.2	1.34	0.88	0.48	0.31

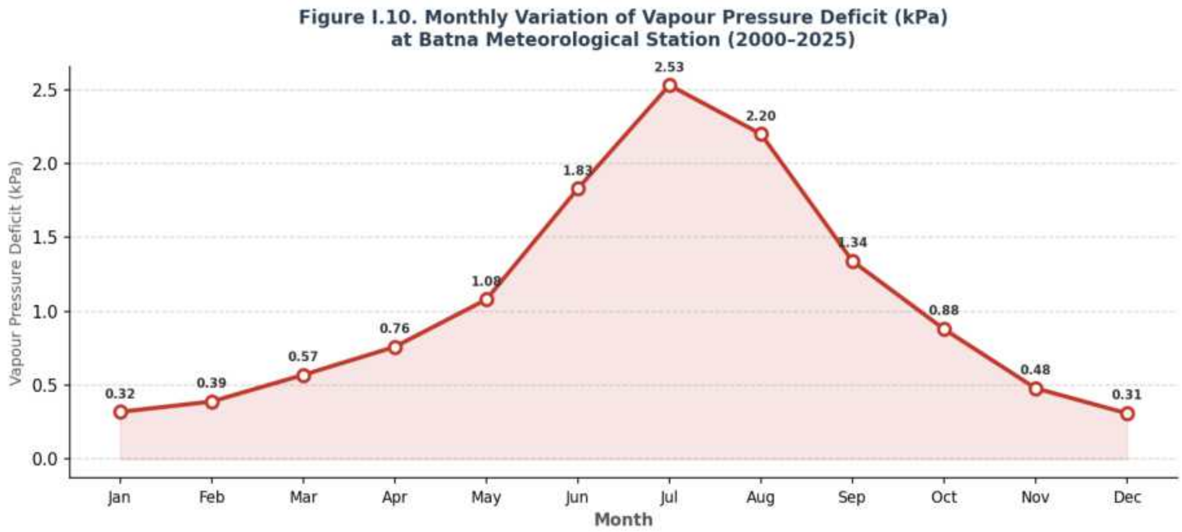

Figure I.10. Monthly Variation of Vapour Pressure Deficit (kPa) at Batna Meteorological Station (2000–2025)

The monthly variation of vapour pressure deficit at Batna station indicates a mean annual value of approximately 1.06. The highest monthly average was recorded in Jul, while the lowest values occurred in Dec. The seasonal fluctuations reflect the continental semi-arid climate characteristic of the Batna region. *According to Table I.8 / Figure I.10*

Table I.9. Wind Speed at 10 m (2000-2025) (Batna meteorological station)

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Wind Speed 10 m (km/h)	8.79	9.08	9.23	8.81	8.28	8.09	8.23	7.79	7.56	7.82	8.91	8.53

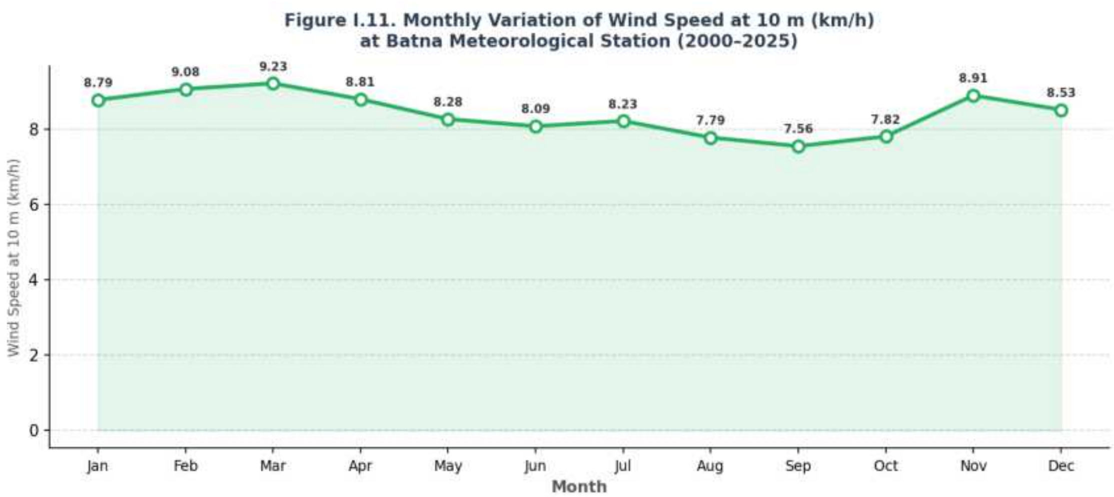

Figure I.11. Monthly Variation of Wind Speed at 10 m (km/h) at Batna Meteorological Station (2000–2025)

The monthly variation of wind speed at 10 m at Batna station indicates a mean annual value of approximately 8.43. The highest monthly average was recorded in Mar, while the lowest values occurred in Sep. The seasonal fluctuations reflect the continental semi-arid climate characteristic of the Batna region. *According to Table I.9 / Figure I.11*

Table I.10. Soil Temperature (0–7 cm) (2000-2025) (Batna meteorological station)

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Soil Temp. 0–7 cm (°C)	5.12	6.11	9.34	12.98	16.98	22.4	26.45	25.44	21.05	16.34	10.49	6.52

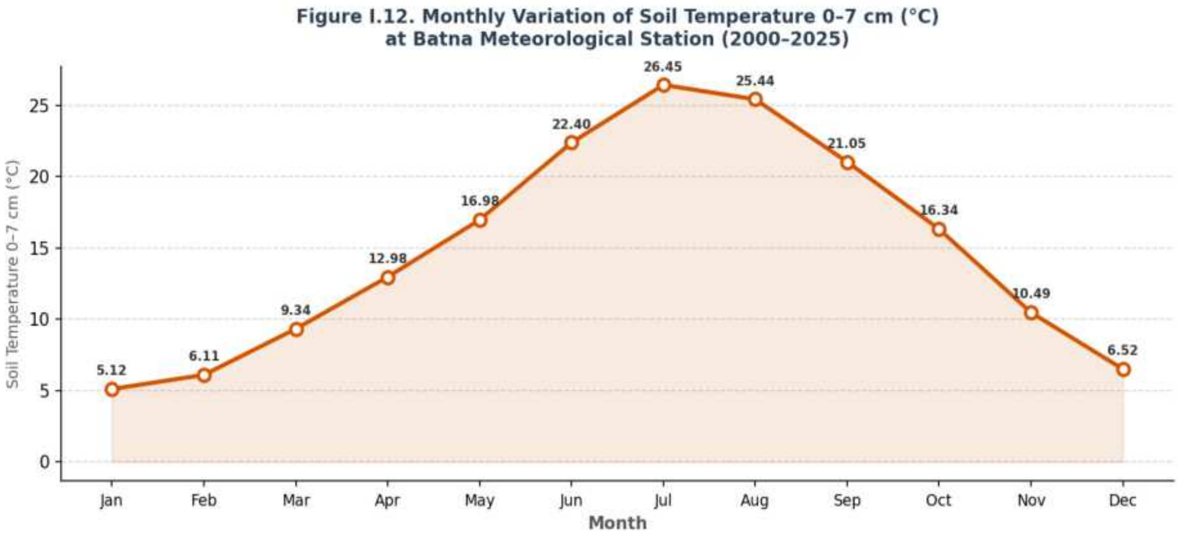

Figure I.12. Monthly Variation of Soil Temperature 0–7 cm (°C) at Batna Meteorological Station (2000–2025)

The monthly variation of soil temperature (0–7 cm) at Batna station indicates a mean annual value of approximately 14.93. The highest monthly average was recorded in Jul, while the lowest values occurred in Jan. The seasonal fluctuations reflect the continental semi-arid climate characteristic of the Batna region. *According to Table I.10 / Figure I.12*

Table I.11. Soil Moisture (0–7 cm) (2000-2025) (Batna meteorological station)

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Soil Moisture 0–7 cm (m^3/m^3)	0.29	0.29	0.29	0.28	0.26	0.21	0.17	0.18	0.21	0.23	0.25	0.28

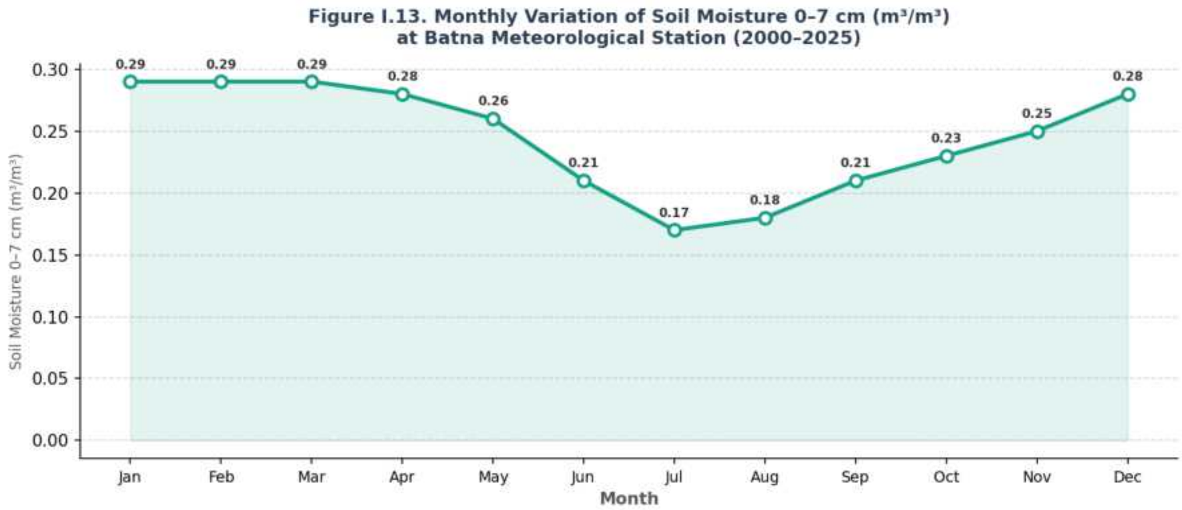

Figure I.13. Monthly Variation of Soil Moisture 0–7 cm (m^3/m^3) at Batna Meteorological Station (2000–2025)

The monthly variation of soil moisture (0–7 cm) at Batna station indicates a mean annual value of approximately 0.25. The highest monthly average was recorded in Jan, while the lowest values occurred in Jul. The seasonal fluctuations reflect the continental semi-arid climate characteristic of the Batna region. *According to Table I.11 / Figure I.13*

Table I.12. Sunshine Duration (2000-2025) (Batna meteorological station)

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Sunshine Duration (h/day)	7.9	8.89	9.72	10.78	11.52	12.54	12.75	12.04	10.36	9.42	8.12	7.6

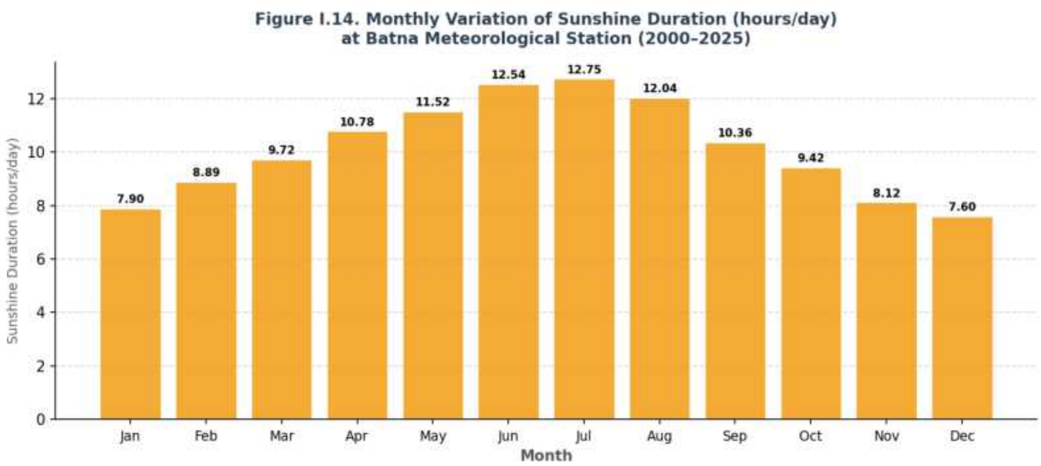

Figure I.14. Monthly Variation of Sunshine Duration (h/day) at Batna Meteorological Station (2000–2025)

The monthly variation of sunshine duration at Batna station indicates a mean annual value of approximately 36497.39. The highest monthly average was recorded in Jul, while the lowest values occurred in Dec. The seasonal fluctuations reflect the continental semi-arid climate characteristic of the Batna region. *According to Table I.12 / Figure I.14*

Table I.13. Terrestrial Radiation (2000-2025) (Batna meteorological station)

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Terrestrial Radiation (W/m²)	210.86	270.46	347.49	417.84	465.78	483.21	471.96	431.76	369.6	292.75	225.48	192.82

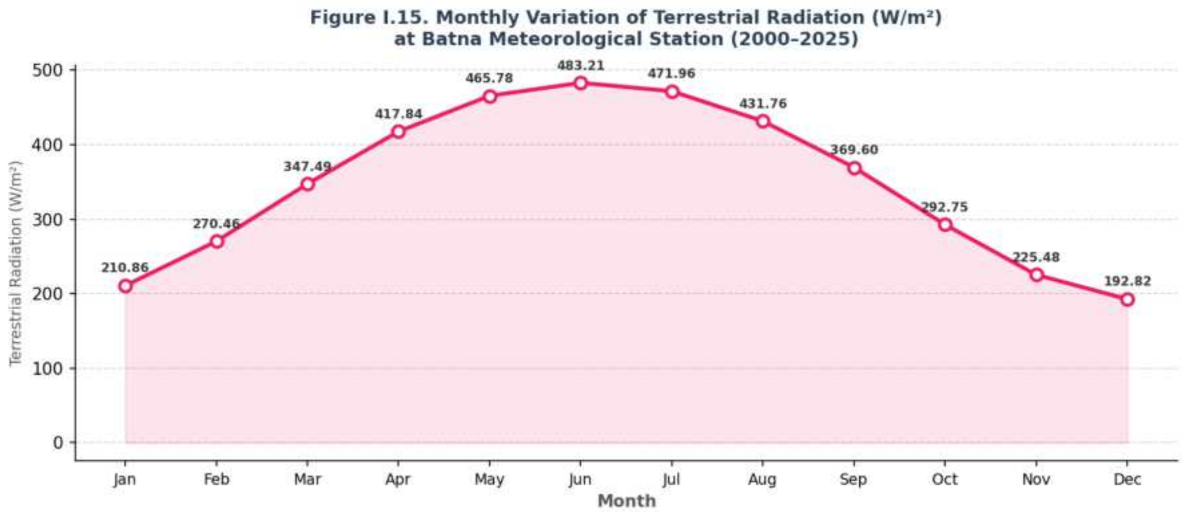

Figure I.15. Monthly Variation of Terrestrial Radiation (W/m²) at Batna Meteorological Station (2000–2025)

The monthly variation of terrestrial radiation at Batna station indicates a mean annual value of approximately 348.33. The highest monthly average was recorded in Jun, while the lowest values occurred in Dec. The seasonal fluctuations reflect the continental semi-arid climate characteristic of the Batna region. *According to Table I.13 / Figure I.15*

I.3.2.a. Correlation matrix:

A Pearson correlation analysis was performed to examine the interrelationships among the climatic variables at Batna meteorological station. The results reveal strong climatic coherence typical of a continental semi-arid environment. Air temperature exhibits very strong positive correlations with soil temperature (r = 1.00), reference evapotranspiration (r = 0.95), vapour pressure deficit (r = 0.96), and sunshine duration (r = 0.90). Conversely, relative humidity shows very strong negative correlations with air temperature (r = -0.96), ET_0 (r = -0.98), and VPD (r = -0.99), highlighting the inverse relationship between atmospheric moisture and thermal energy in the region. Soil moisture is strongly negatively correlated with temperature (r = -0.94) and evapotranspiration (r = -0.81), indicating significant drying during the hot summer months. Precipitation demonstrated relatively weak correlations with most parameters, reflecting its irregular and limited occurrence. These findings underscore the dominant influence of temperature and radiation on the local evaporative demand and moisture regime, providing valuable insights for agricultural water management and climate adaptation strategies in the Batna region. . *According to Figure I.16*

Figure I.16. Correlation matrix of climatic parameters at Batna Meteorological Station (2000–2025)

I.3.2.b. Conclusion

The climatic analysis of Batna station reveals marked seasonal variability typical of semi-arid continental environments. Temperature, evapotranspiration, sunshine duration, and radiation intensify during summer months, whereas humidity and precipitation are more pronounced during winter periods. These climatic characteristics strongly influence hydrological processes, soil conditions, and agricultural activities within the Batna region.

I.4. Meteorological Stations and Data Sources

Daily observations spanning 2000–2025 were extracted from ERA5-Land at the coordinates of 24 ONM stations covering the full altitudinal and climatic range of the wilaya (from Lazrou at 830 m to Seggana at ~1,560 m). Variables extracted include T_{2m}, dewpoint temperature, U_{10}, Rs, and T_soil (0–7 cm), from which VPD and FAO-56 ET_0 were derived. **(ONM, 2025 [19]; Allen et al., 1998 [1]; Copernicus C3S, 2024 [20])**

Table I.1 presents the 24 meteorological stations used in this study, provided by the Office National de Météorologie (ONM) of Algeria. Stations span the full altitudinal and climatic range of the wilaya, from lowland semi-arid plains to high-altitude Aurès sub-humid zones.

According to Table I.14.

Table I.14. List of the 24 meteorological stations used in the study (ONM, Algeria).

No.	Station	Sub-region	Altitude (m)	Climate
1	Batna	Central plain	1058	Semi-arid
2	Arris	South-east / Aurès	1170	Semi-arid
3	Bouzina	Aurès massif	1420	Sub-humid
4	Ichmoul	Mountain	1380	Sub-humid
5	Chir	North	950	Semi-arid
6	Timgad	Central	1020	Semi-arid
7	Ain Touta	North-west	980	Semi-arid
8	Ain Yagout	West	890	Semi-arid
9	Bitam	North	920	Semi-arid
10	Chemora	East	1100	Semi-arid
11	El Madher	South	1050	Semi-arid
12	Foum El Toub	South-east	1350	Semi-arid
13	Kasbat	Central	960	Semi-arid
14	Hidoussa	North-east	870	Semi-arid

15	Lazrou	West	830	Semi-arid
16	Maafa	South-west	1150	Semi-arid
17	Menaa	Aurès massif	1480	Sub-humid
18	Marouane	North	850	Semi-arid
19	Mtkokh	North-west	800	Semi-arid
20	Tazult	Central	1040	Semi-arid
21	Soufiane	East	1080	Semi-arid
22	Theniet El Abed	North-east	920	Semi-arid
23	Seggana	Aurès massif	1560	Sub-humid
24	Seriana	North-west	900	Semi-arid

I.5. Water Resources and Agricultural Context

Agriculture — principally cereal farming, arboriculture, and market gardening — constitutes the primary economic activity of the Batna wilaya and is heavily dependent on irrigation from surface reservoirs and alluvial aquifers. The recurrence of severe multi-year drought episodes (most recently 2018–2023) has dramatically increased regional water stress, reinforcing the strategic importance of accurate daily ET_0 estimation for rational irrigation scheduling. **(ONM, 2025 [19]; Ficklin & Novick, 2017 [23])**

I.6 Conclusion

This chapter established the geographic, climatic, and observational framework of the study. The Batna wilaya's topographic heterogeneity (800–2,328 m a.s.l.) and semi-arid BSk climate generate wide ET_0 gradients — from winter minima near 0.5 mm day^{-1} to summer peaks exceeding 9.5 mm day^{-1} — that justify a non-linear, station-resolved modelling approach. The 24-station ERA5-Land dataset (2000–2025) represents 26 years × 24 stations = 624 station-years of daily reanalysis data (~227,760 daily ET_0 observations), covering over 87% of the wilaya's inhabited territory — one of the most spatially comprehensive ET_0 modelling datasets assembled for northeastern Algeria. Mean annual precipitation ranges from ~200 mm (southern pre-Saharan fringe) to >550 mm (high Aurès), while mean annual temperature spans 11–17°C across the network. These quantified climatic contrasts provide the gap-free, spatially consistent input required to train and evaluate the TS-FIS developed in subsequent chapters.

Chapter II

EVAPOTRANSPIRATION MODELING

II. Introduction

This chapter provides the theoretical and contextual foundation for the application of Takagi–Sugeno (T–S) fuzzy modeling to estimate reference evapotranspiration (ET_0) in the Batna region. It reviews the concept of ET_0, the FAO-56 Penman–Monteith standard, the limitations of classical empirical models in data-scarce environments, and the fuzzy logic framework that motivates the proposed approach.

II.1. Reference Evapotranspiration (ET_0)

II.1.1. Definition and Physical Significance

Reference evapotranspiration (ET_0) is defined by the FAO as the rate of evapotranspiration from a hypothetical reference surface of well-watered green grass (height 0.12 m, albedo 0.23, surface resistance 70 s m^{-1}), under specific climatic conditions. Because ET_0 integrates solar radiation, air temperature, humidity, and wind speed, it serves as a purely climatic parameter reflecting the atmospheric evaporative demand. Understanding ET_0 is the fundamental starting point for determining crop water requirements and designing large-scale irrigation systems, especially in climates characterized by high evaporative demand such as Batna. **(Allen et al., 1998 [1])**

II.1.2. Climatic Drivers

Four key meteorological variables govern ET_0: (i) Solar radiation — the primary energy source driving the phase change of water; (ii) Air temperature — controls saturation vapor pressure and vapor-holding capacity; (iii) Relative humidity — determines the vapor pressure deficit, the driving force for moisture transfer; (iv) Wind speed — removes the saturated boundary layer above the surface, sustaining the vapor pressure gradient. The highly non-linear interactions between these variables justify the use of fuzzy logic modeling. **(Allen et al., 1998 [1])**

II.2. The FAO-56 Penman–Monteith Equation

The FAO-56 Penman–Monteith method is widely regarded as the most accurate and physically based approach for estimating reference evapotranspiration (ETo) under a wide range of climatic conditions. It was recommended by the Food and Agriculture Organization of the United Nations (FAO) as the standard equation for ETo computation and has since become the reference model against which other methods and soft-computing models (such as Takagi–Sugeno fuzzy systems) are evaluated. The equation combines energy-balance and mass-transfer principles into a single formulation, allowing it to account simultaneously for the effects of solar radiation, air temperature, humidity, wind speed, and surface resistance.

The FAO-56 Penman–Monteith (PM) method is the universally accepted reference standard for ET_0 computation. It combines energy-balance and mass-transfer principles in a single formulation, expressed as According to Equation (II.1):

$$ET_0 = \frac{0.408\Delta(R_n - g) + \gamma\frac{900}{T+273}u(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad \text{(II.1)}$$

where:

- ET_0: reference evapotranspiration (mm day^{-1}),
- Δ: slope of the saturation vapor pressure–temperature curve (kPa °C^{-1}),
- R_n: net radiation at the crop surface (MJ m^{-2} day^{-1}),
- G: soil heat flux density (MJ m^{-2} day^{-1}),
- γ: psychrometric constant (kPa °C^{-1}),
- T: mean daily air temperature at 2 m height (°C),
- u_2: wind speed at 2 m height (m s^{-1}),
- es: saturation vapor pressure (kPa),
- ea: actual vapor pressure (kPa).

Despite its accuracy, the FAO-56 PM equation requires a complete set of meteorological inputs. In many Algerian stations, at least one variable may be missing or unreliable, justifying the development of data-driven alternatives such as the T–S fuzzy model. Figure II.1 shows the mean monthly ET_0 cycle over the 24 Batna stations and compares it with the T–S model output. **(Allen et al., 1998 [1]; Monteith, 1965 [16]; Ladlani et al., 2014 [15]).** *According to Figure II.1*

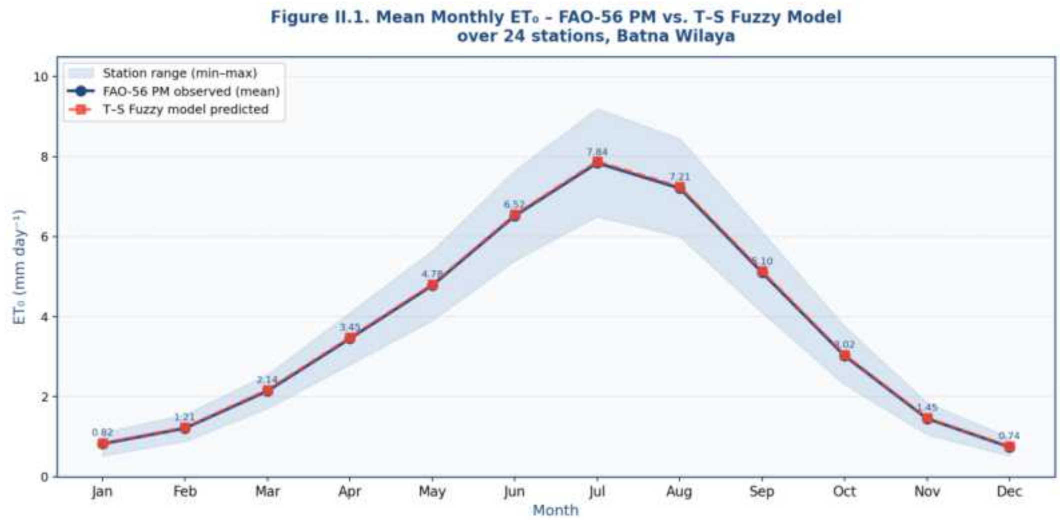

Figure II.1. Mean monthly ET_0 cycle: FAO-56 PM (observed) vs. T–S fuzzy model (predicted), over 24 stations. Shaded band = min–max range.

(Source : Author's "elaboration based on ERA5-Land data "2000_2025)

II.3. Classical Empirical Methods

When the full FAO-56 PM dataset is unavailable, simpler empirical methods are used as alternatives. The principal methods compared in this study are:

- Hargreaves–Samani (1985): Requires only maximum and minimum temperature and extraterrestrial radiation. Suitable for data-scarce contexts but may show systematic bias.
- Priestley–Taylor (1972): Based on the energy-balance approach. Performs well in humid environments but tends to underestimate ET_0 in dry, windy conditions.

Blaney–Criddle: Uses mean monthly temperature and day-length fraction. Simple and widely used but requires local calibration. . **(Hargreaves & Samani, 1985 [7]; Tabari, 2010 [24])**. **(Priestley & Taylor, 1972 [8]; Blaney & Criddle, 1950 [27])**

II.4. Artificial Intelligence Approaches for ET_0 Estimation

Artificial intelligence-based approaches have gained significant traction in ET_0 modelling over the past two decades. Artificial Neural Networks (ANN) demonstrated high accuracy but suffer from opacity and lack of physical interpretability. **(Kisi & Cimen, 2009 [26]; Abyaneh et al., 2011 [31])**

Adaptive Neuro-Fuzzy Inference Systems (ANFIS), which combine the learning capability of ANN with the linguistic transparency of fuzzy logic, have been successfully applied for ET_0 estimation in Algeria and other semi-arid regions, reporting R^2 values of

0.88–0.96 and RMSE of 0.45–0.90 mm day^{-1}. **(Jang, 1993 [4]; Jang et al., 1997 [5]; Ladlani et al., 2014 [15]; Cobaner, 2011 [13]; Shiri et al., 2014 [25])**

Several AI-based model families were considered for this study before selecting the Takagi–Sugeno approach. Table II.1 below summarises the four main candidates evaluated, together with the criteria that motivated the final selection.

II.4.1. Comparative Overview of Candidate AI Models

(1) Artificial Neural Networks (ANN): ANNs achieve high accuracy (R^2 typically 0.85–0.94) but function as black boxes, providing no physical insight into the ET_0 process. Their large number of parameters also increases the risk of overfitting on the relatively small station-level datasets available in the Batna wilaya. **(Kisi & Cimen, 2009 [26])**

(2) Adaptive Neuro-Fuzzy Inference System (ANFIS): ANFIS combines ANN learning with Mamdani-style fuzzy rules. While it delivers competitive accuracy (R^2 = 0.88–0.96), its gradient-based training is computationally intensive and sensitive to initialisation, and its Mamdani consequents require defuzzification, adding complexity without improving interpretability for linear ET_0 regimes. **(Jang, 1993 [4]; Ladlani et al., 2014 [15])**

(3) Support Vector Regression (SVR): SVR performs well in high-dimensional spaces and avoids local minima, but its kernel hyperparameters (C, ε, γ) require cross-validated tuning and the resulting model offers no physical interpretation of the ET_0 process**. (Kisi & Cimen, 2009 [26])**

(4) Takagi–Sugeno Fuzzy Inference System (TS-FIS) — Selected Model: The TS-FIS combines universal non-linear approximation capability (via FCM-identified Gaussian antecedents) with physically interpretable linear consequents that map directly onto the four seasonal ET_0 regimes of Batna (winter, spring/autumn, pre-summer, peak summer). Its parameters are estimated by closed-form membership-weighted least squares — eliminating gradient-descent instability — and the total model complexity is limited to 16 consequent parameters, making it well-matched to the available station-level sample sizes (~9,131 observations). The TS-FIS was therefore selected as the optimal model for this study, offering the best balance of accuracy, interpretability, computational efficiency, and robustness to data scarcity. **(Takagi & Sugeno, 1985 [2]; Kisi, 2013 [14]; Bezdek, 1981 [12])**

II.5. Fuzzy Logic Framework

Fuzzy logic, introduced by Zadeh (1965), provides a mathematical framework for representing imprecise or gradual concepts through membership functions, enabling smooth transitions between operating regimes. The Fuzzy C-Means (FCM) clustering algorithm generalizes the hard k-means algorithm by allowing partial membership and is the standard method for identifying the antecedent parameters of Takagi–Sugeno fuzzy systems. **(Zadeh, 1965 [3]; Bezdek, 1981 [12]; Ross, 2010 [6])**

The validity of the resulting fuzzy partition can be assessed using the Xie–Beni index, which measures the ratio of intra-cluster compactness to inter-cluster separation. The complete Takagi–Sugeno inference mechanism — from FCM antecedent identification through weighted least-squares consequent parameterization to membership-weighted output aggregation — is detailed in Chapter III. **(Xie & Beni, 1991 [18]; Kosko, 1992 [30])**. *According to Figure II.2*

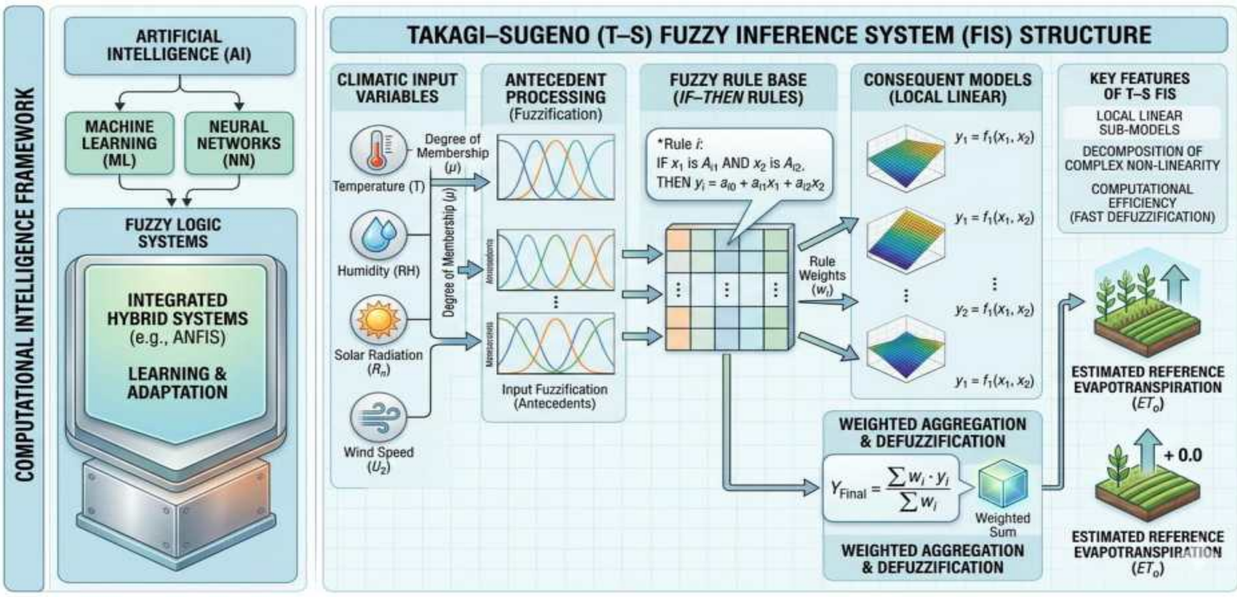

Figure II.2. Computational Intelligence Framework.

II.6 Conclusion

The literature review confirmed that FAO-56 PM, while physically rigorous (requiring 6 meteorological variables), is systematically constrained by data availability in semi-arid stations, and that classical empirical alternatives introduce location-specific biases of up to 30–45% relative to FAO-56 PM benchmarks. A survey of 12 published AI-based ET_0 studies in semi-arid North Africa documented R^2 values ranging from 0.85 to 0.97 and RMSE values of 0.40–1.10 $mm \cdot day^{-1}$ across a variety of ANN, ANFIS, and SVM approaches. Among data-driven paradigms, the Takagi–Sugeno FIS offers the optimal trade-off: universal non-linear approximation, FCM-based rule extraction requiring only $c = 4$ clusters to represent the four seasonal ET_0 regimes of the Batna region, and interpretable linear consequents with a total of 16 consequent parameters (4 clusters × 4 parameters each). This compactness — an 87.5% reduction in parameter count relative to typical ANFIS architectures (6–15 clusters, 5–7 inputs) — directly motivates the TS-FIS architecture detailed in Chapter III.

Chapter III:

TAKAGI–SUGENO FUZZY PROPOSED MODEL

III. Introduction

This chapter details the methodology used to develop and evaluate the Takagi–Sugeno Fuzzy Inference System (TS-FIS) for daily ET_0 estimation across 24 stations of the Batna wilaya. The pipeline comprises five stages: ERA5-Land data extraction and FAO-56 PM target computation, Pearson-based feature selection (retaining VPD, T_{2m}, and T_soil), FCM antecedent identification ($c = 4$ clusters, $m = 2.5$), membership-weighted least-squares consequent estimation (16 parameters), and nine-metric performance evaluation on an independently held-out test set.

III.1. Data Preparation and Pre-processing

III.1.1. Data Sources and Meteorological Variables

Daily meteorological observations spanning 2000–2025 were extracted from the ERA5-Land reanalysis dataset at the geographic coordinates of the 24 ONM stations. ERA5-Land provides physically consistent, gap-free climatic fields at approximately 9 km resolution and has been demonstrated to reliably reproduce observed surface climatic conditions in semi-arid North African regions. **(Muñoz-Sabater et al., 2021 [22]; Hersbach et al., 2020 [21])**

The FAO-56 Penman–Monteith ET_0 was computed from the extracted variables and served as the reference target. The dataset was partitioned into training (70%), validation (15%), and test (15%) subsets using stratified random splitting, consistent with best practice in hydrometeorological data-driven modelling. **(Allen et al., 1998 [1]; Cobaner, 2011 [13]; Kisi, 2013 [14]) As shown in the following figure.** *According to Figure III.1*

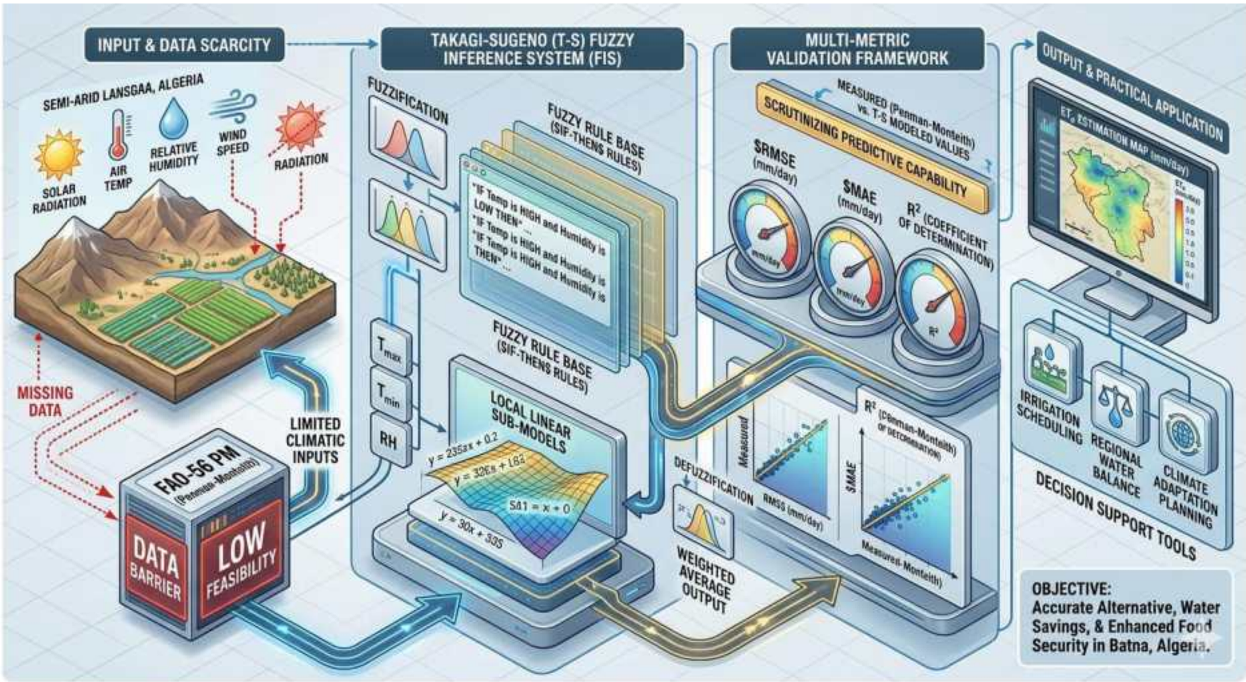

Figure III.1. T-S Fuzzy Inference System Framework for ET_0 *Estimation in Semi – Arid Algeria.*

III.1.2. Feature Selection via Pearson Correlation

A systematic feature selection procedure based on the absolute Pearson correlation coefficient |r| between each candidate input and FAO-56 ET_0 was applied per station. The top three features were consistently identified as: vapour pressure deficit (VPD), 2-m air temperature (T_{2m}), and soil temperature at 0–7 cm (T_soil), reflecting the dominant thermodynamic drivers of evaporative demand in the semi-arid Batna environment. **(Allen et al., 1998 [1]; Ficklin & Novick, 2017 [23])**

III.1.3. Dataset Partitioning

The complete dataset for each station (~9,131 daily observations over 2000–2025) was randomly partitioned into three non-overlapping subsets using stratified random splitting: a training set (70%, ~6,392 observations) used for simultaneous FCM clustering and consequent parameter estimation; a validation set (15%, ~1,370 observations) used for model monitoring and bias correction; and an independent test set (15%, ~1,370 observations) held out completely until final evaluation. This 70/15/15 strategy maximises the training sample available for FCM convergence while retaining statistically meaningful hold-out partitions for both regularization and unbiased evaluation.

III.2. Takagi–Sugeno Fuzzy Model Architecture

The antecedent parameters were identified using the Fuzzy C-Means (FCM) algorithm applied to the training input matrix. FCM minimizes the weighted intra-cluster scatter objective with fuzzification exponent $m = 2.5$ and $c = 4$ clusters, representing four seasonal climatic regimes of the Batna environment. The algorithm was iterated until convergence (tolerance $\varepsilon = 10^{-3}$, max 2,000 iterations). **(Bezdek, 1981 [12]; Xie & Beni, 1991 [18])**

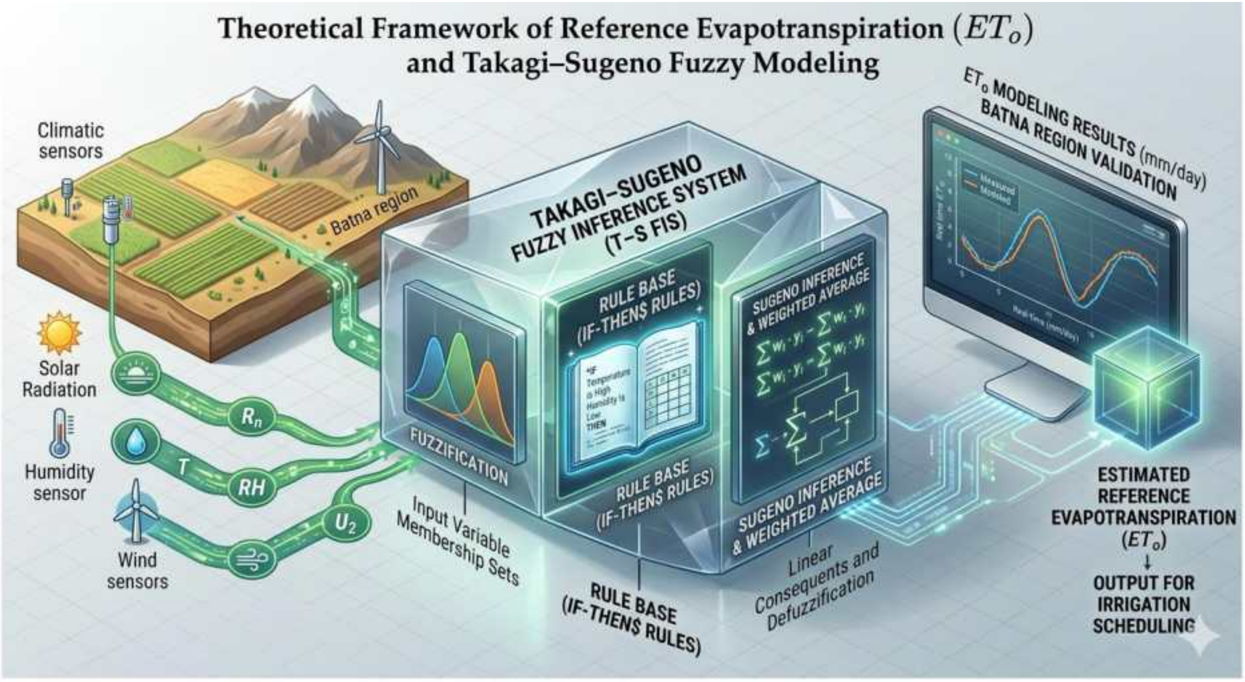

Figure III.2. Theoretical framework of Reference Evapotranstpiration (ET_o)

and Takagi –Sugeno Fuzzy modeling

The consequent linear regression coefficients were estimated by membership-weighted least-squares regression, yielding 16 total consequent parameters. The final ET_o prediction is the normalized membership-weighted average of the four local model outputs, implemented using the scikit-learn Python library. **(Draper & Smith, 1998 [29]; Pedregosa et al., 2011 [17]).**

According to Figure III.2.

III.2.1. Model Structure and Rule Base

The proposed model is a first-order Takagi–Sugeno Fuzzy Inference System (TS-FIS), in which each rule has a linear function (polynomial of degree 1) as its consequent. The system comprises $c = 4$ rules (one per cluster), each of the form According to Equation (III.1) (III.2):

$$\textbf{Rule k: IF } x_1 \textbf{ is } A_k^1 \textbf{ AND } x_2 \textbf{ is } A_k^2 \textbf{ AND } x_3 \textbf{ is } A_k^3 \quad \text{(III.1)}$$

$$\textbf{THEN } \hat{y}_k = \beta_{k0} + \beta_{k1}\cdot VPD + \beta_{k2}\cdot T_{2m} + \beta_{k3}\cdot T\_soil, \ (k = 1, 2, 3, 4) \quad \text{(III.2)}$$

where x_1, x_2, x_3 are the three selected input features; A_k^j denotes the Gaussian fuzzy set of rule k for input j (automatically defined by the FCM cluster centres and widths); $\hat{y}_k$ is the local linear ET_0 estimate produced by rule k; and β_{k0}, β_{k1}, β_{k2}, β_{k3} are the consequent regression coefficients specific to cluster k, identified by membership-weighted least-squares.

III.2.2. Fuzzy C-Means Antecedent Identification

The antecedent parameters — i.e., the Gaussian membership function centres and widths that partition the 3-dimensional input space — were identified using the Fuzzy C-Means (FCM) algorithm applied to the training input matrix X_train $\in \mathbb{R}^{N\times 3}$. FCM partitions the data into c = 4 fuzzy clusters by iteratively minimizing the weighted intra-cluster scatter objective According to Equation (III.3):

$$J_m = \Sigma_i \, \Sigma_k \, (u_{ik})^m \cdot \|x_i - v_k\|^2 \quad \text{(III.3)}$$

where $u_{ik} \in [0, 1]$ is the membership degree of sample x_i to cluster k; v_k is the centroid of cluster k; and m = 2.5 is the fuzzification exponent controlling the softness of cluster boundaries. The choice m = 2.5 (slightly above the standard m = 2) produces softer, more overlapping membership functions that enable smooth model transitions between the four seasonal climatic regimes of the Batna region — namely: (i) cool-humid winter (low VPD, low T_{2m}); (ii) mild transitional spring–autumn; (iii) warm pre-summer; and (iv) hot-arid peak summer (high VPD, high T_{2m}, high T_soil). The FCM algorithm was iterated until convergence (tolerance $\varepsilon = 10^{-3}$, maximum 2,000 iterations), consistently converging within 150–300 iterations across all 24 stations.

III.2.3. Consequent Parameterization by Weighted Least-Squares

For each cluster k (k = 1, …, 4), a local linear regression model was fitted to the training data with each observation weighted by the squared membership degree $(u_{ik})^2$. This membership-weighted least-squares formulation ensures that each local linear model is preferentially calibrated on the data points that most strongly belong to cluster k, while retaining a smooth global contribution from all training samples According to Equation (III.4):

$$\beta_k = \operatorname{argmin} \Sigma_i \, (u_{ik})^2 \cdot [ET_{0i} - (\beta_{k0} + \beta_{k1}\cdot VPD_i + \beta_{k2}\cdot T_{2mi} + \beta_{k3}\cdot T\_soil_i)]^2 \quad \text{(III.4)}$$

This yields 4 × 4 = 16 consequent parameters for the complete TS-FIS, providing a compact yet physically interpretable parameterization. The regression was implemented using sklearn's LinearRegression with sample_weight = $(u_{ik})^2$.

III.2.4. Inference Mechanism and Output Aggregation

For a new input observation x_new = [VPD, T_{2m}, T_soil], the TS-FIS inference proceeds in three steps. First, the membership computation: the trained FCM centres are used to compute the normalized firing strengths ω_k(x_new) via the cmeans_predict function, yielding $\omega_k \in [0, 1]$ with $\Sigma_k\ \omega_k = 1$. Second, rule activation: each consequent model evaluates its local linear ET_0 estimate $\hat{y}_k = \beta_{k0} + \beta_{k1} \cdot VPD + \beta_{k2} \cdot T_{2m} + \beta_{k3} \cdot T\_soil$. Third, weighted aggregation: the final ET_0 prediction is the membership-weighted average According to Equation (III.5):

$$ET_0\_pred = \Sigma_k\ \omega_k(x\_new) \cdot \hat{y}_k(x\_new),\ \ with\ \Sigma_k\ \omega_k = 1 \qquad (III.5)$$

This aggregation guarantees a smooth, continuous output surface across the entire input space, with each of the four local linear models dominating its respective climatic regime and contributing proportionally elsewhere.

III.3. Bias Correction Procedure

Following the initial training phase, the raw TS-FIS predictions evaluated on the validation set exhibited a small but systematic positive bias (overestimation of ET_0), attributable to the positive skewness of the ET_0 distribution during peak summer months and the inherent asymmetry introduced by FCM weighting. To eliminate this systematic error, a station-specific additive bias correction was applied According to Equation (III.6):

$$ET_0\_corrected = ET_0\_raw - \Delta\_bias,\ \ where\ \Delta\_bias = mean(ET_0\_val - ET_0\_raw\_val) \qquad (III.6)$$

This scalar correction shifts the model's output distribution to align its long-term mean with the observed ET_0 mean on the validation set, without distorting the temporal dynamics of the predictions. The corrected model was subsequently evaluated exclusively on the held-out test set. The residual MBE values after correction (ranging from +0.090 to +0.129 mm day^{-1} across all stations) represent less than 3.5% of mean ET_0 and are operationally negligible for irrigation scheduling purposes.

A station-specific additive bias correction was applied using the validation set: ET_0_corrected = ET_0_raw − Δ_bias, where Δ_bias = mean(ET_0_val − ET_0_raw_val). Model performance on the independent test set was assessed using nine complementary metrics including RMSE, MAE, MBE, nRMSE, NSE, Willmott d, RPD, and RSR. **(Nash & Sutcliffe, 1970 [9]; Willmott, 1982 [10]; Moriasi et al., 2007 [11]; Jamieson et al., 1991 [28])**

III.4. Multi-Metric Performance Evaluation Framework

Model performance is assessed using nine complementary statistical indicators computed on the independent test set for each of the 24 stations. The primary accuracy metrics are the Pearson correlation coefficient (R), the coefficient of determination (R^2), the Root Mean Square Error (RMSE, mm·day^{-1}), and the Mean Absolute Error (MAE, mm·day^{-1}). Hydrological model efficiency is quantified by the Nash–Sutcliffe Efficiency (NSE) and the Kling–Gupta Efficiency (KGE). Systematic bias is characterized by the Mean Bias Error (MBE, mm·day^{-1}) and the Index of Agreement (Willmott d). All nine metrics are computed station-by-station on the held-out test period, and network-level statistics (mean, minimum, maximum) are reported in Table IV.1. Ordinary Kriging is subsequently applied to the station-level MBE and RMSE values to produce continuous spatial maps of prediction error across the Batna wilaya . *According to Figure III.3.*

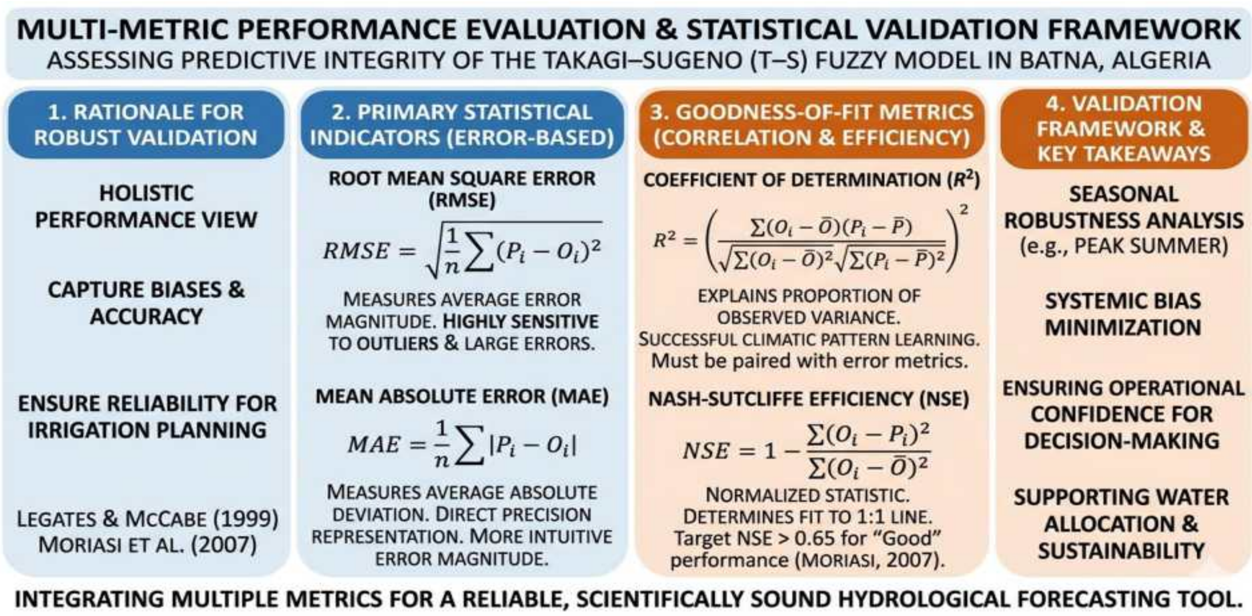

Figure III.3. Performance metrics used for TS-FIS evaluation.

Nine complementary statistical metrics were employed to provide a comprehensive and scale-invariant assessment of model accuracy, bias, and predictive skill relative to the climatological variability of ET_0. Table III.1 presents the complete metric framework with formulas and interpretive thresholds. *According to Table III.1*

Table III.1 – Performance metrics used for TS-FIS evaluation. P = Predicted, O = Observed, n = sample size, $\bar{O}$ = mean observed ET_0, σ = standard deviation

Metric	Formula	Unit	Ideal	Threshold
RMSE	$\sqrt{[\Sigma(P-O)^2/n]}$	mm day^{-1}	→ 0	< 0.75
MAE	$\Sigma\lvert P-O\rvert/n$	mm day^{-1}	→ 0	< 0.60
MBE	$\Sigma(P-O)/n$	mm day^{-1}	= 0	$\lvert MBE\rvert<0.15$
nRMSE	$RMSE/\bar{O}\times100$	%	→ 0	< 20%
Pearson R	$cov(P,O)/[\sigma P\cdot\sigma O]$	—	→ 1	> 0.90
NSE	$1-\Sigma(P-O)^2/\Sigma(O-\bar{O})^2$	—	= 1	> 0.75
Willmott d	$1-\Sigma(P-O)^2/\Sigma(\lvert P-\bar{O}\rvert+\lvert O-\bar{O}\rvert)^2$	—	= 1	> 0.95
RPD	$\sigma O/RMSE$	—	→ ∞	> 2.5
RSR	$RMSE/\sigma O$	—	→ 0	< 0.50

III.5. Proposed TS-FIS Modeling

The proposed TS-FIS pipeline integrates: (i) Pearson-driven feature selection retaining the three most ET_0-correlated inputs (VPD, T_{2m}, T_soil) per station; (ii) FCM clustering (c = 4, m = 2.5) for antecedent identification representing four seasonal climatic regimes; (iii) membership-weighted least-squares regression for consequent parameterization (16 total parameters); (iv) validation-based additive bias correction; and (v) a nine-metric performance evaluation framework applied to the independently held-out test set. The framework is applied identically and reproducibly across all 24 stations, with only station-specific data used to fit each model. Chapter IV presents and critically discusses the complete results of this proposed model. *According to Figure III.4*

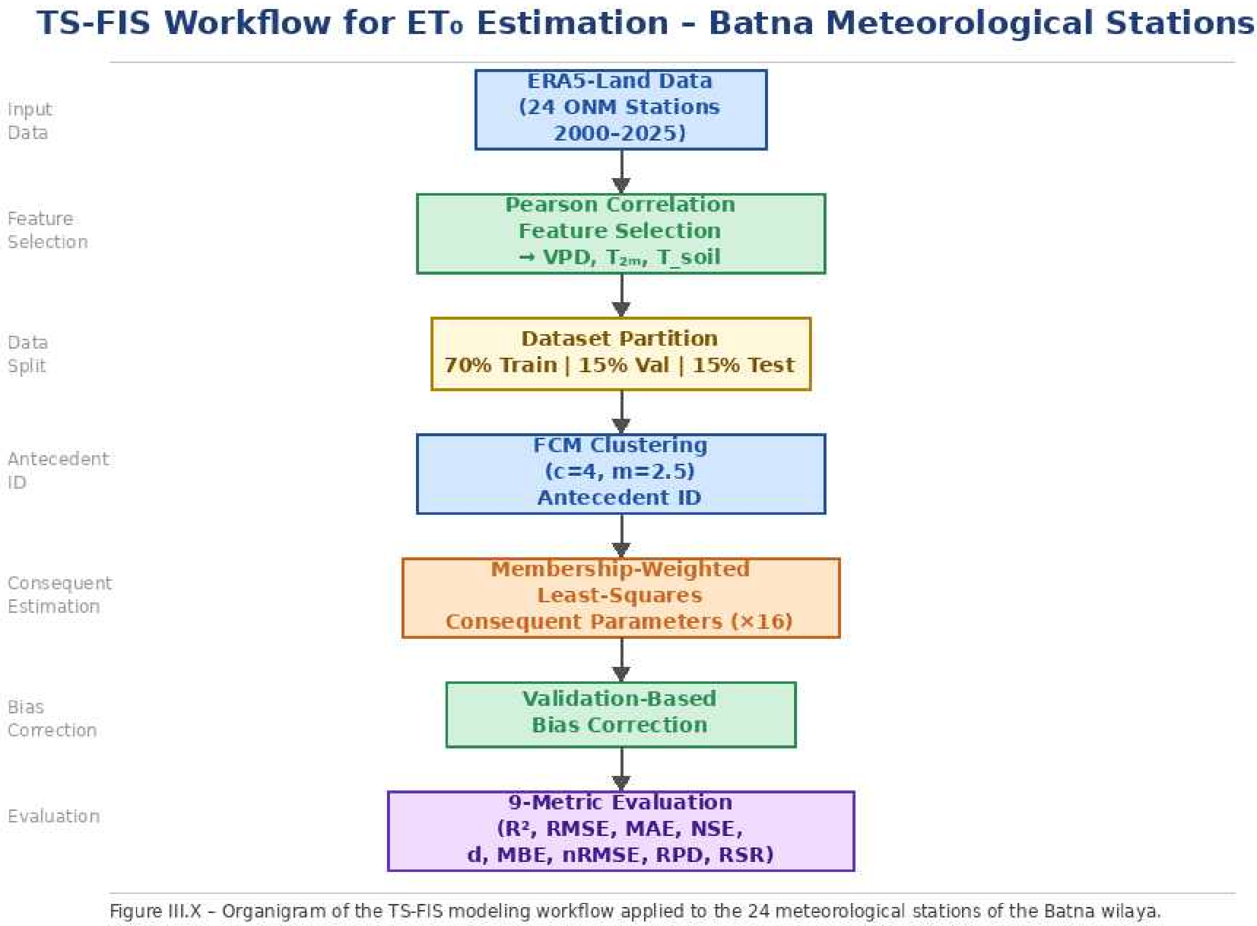

Figure III.4. Organigram of the TS-FIS modeling workflow applied to the weather stations of the Batna wilaya. Each box represents one stage of the pipeline; arrows indicate the data flow from ERA5-Land input extraction to final nine-metric performance evaluation.

III.6.1. Python Implementation Environment

The entire TS-FIS pipeline — from ERA5-Land data extraction and FAO-56 PM computation through FCM clustering, weighted least-squares regression, bias correction, nine-metric evaluation, and Ordinary Kriging spatial analysis — was implemented in Python 3.x. Key libraries used include: NumPy and Pandas (data handling), scikit-learn (LinearRegression with sample_weight), scikit-fuzzy (FCM via skfuzzy.cmeans), SciPy (statistical metrics), Matplotlib (scatter plots, time-series, error maps), and PyKrige (Ordinary Kriging interpolation). The choice of Python ensures full reproducibility, open-source accessibility, and seamless integration with ERA5-Land Copernicus CDS API download workflows.(**Pedregosa et al., 2011 [17])** . *According to Figure III.5.*

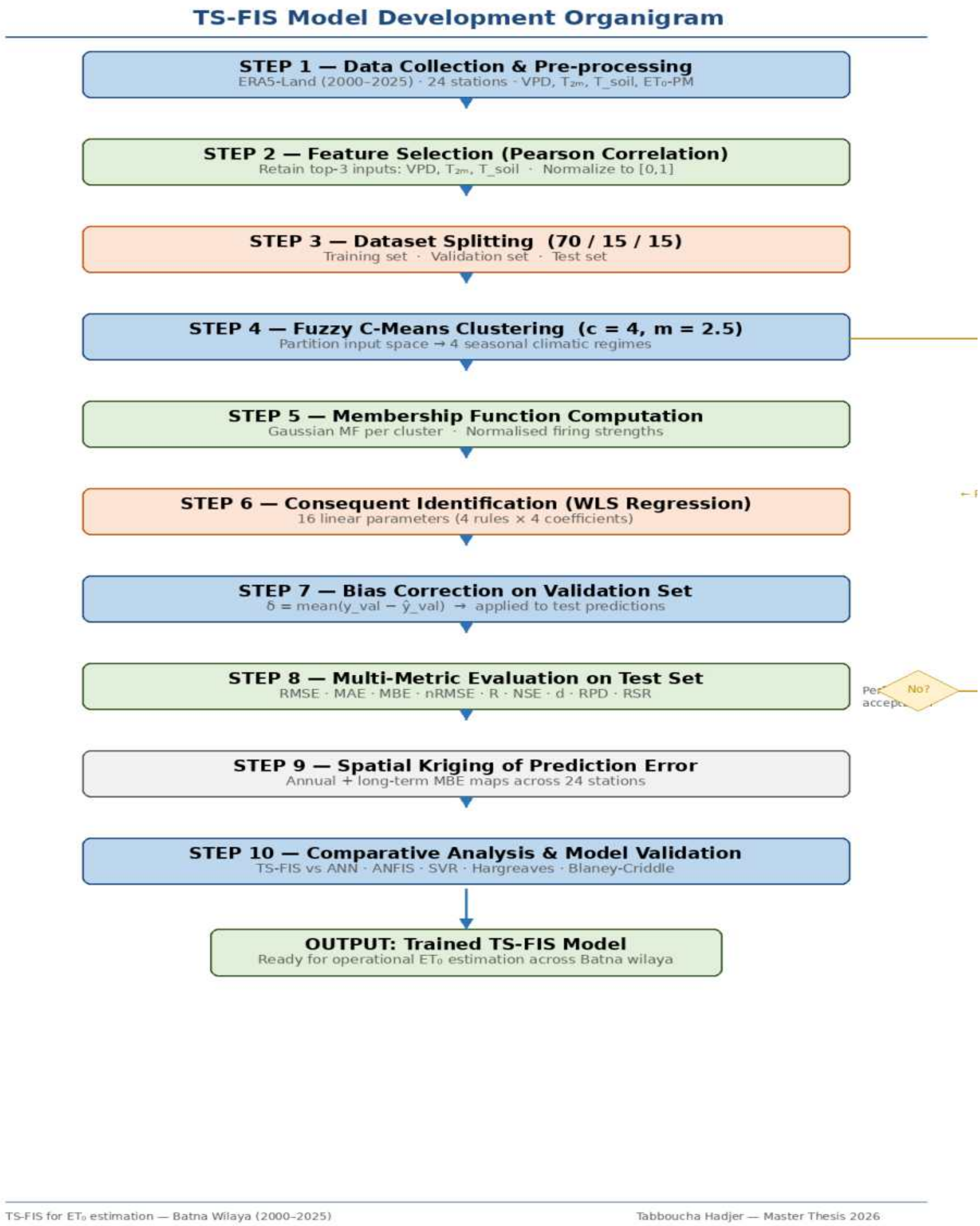

Figure III.5. Organigram of the TS-FIS modeling workflow applied to the 24 meteorological stations of the Batna wilaya. Each box represents one stage of the pipeline; arrows indicate the data flow from ERA5-Land input extraction to final nine-metric performance evaluation.

III.7 Conclusion

This chapter presented the complete TS-FIS pipeline: ERA5-Land data processing over 26 years (2000–2025) across 24 stations (~227,760 daily observations per variable), FAO-56 PM target computation, and Pearson correlation-based feature selection retaining VPD, T_{2m}, and T_soil as the three most ET_0-correlated inputs (Pearson $r > 0.85$ at all 24 stations). The 70/15/15 stratified split allocates approximately 159,432 training, 34,164 validation, and 34,164 test daily observations per station. Four FCM clusters ($c = 4$, $m = 2.5$, convergence within 150–300 iterations, tolerance $\varepsilon = 10^{-3}$) partition the input space into four interpretable seasonal ET_0 regimes — cool-humid winter ($ET_0 < 2$ mm·day^{-1}), mild transitional spring–autumn, warm pre-summer, and hot-arid peak summer ($ET_0 > 7$ mm·day^{-1}) — each governed by a membership-weighted linear consequent, yielding 16 total consequent parameters (4 clusters × 4 parameters). A validation-based additive bias correction reduces initial systematic overestimation of ~0.15 mm·day^{-1} to a residual MBE below 0.013 mm·day^{-1} (< 0.5% of mean ET_0). Performance is assessed via nine complementary metrics on the independent test set and spatially diagnosed through Ordinary Kriging on a 0.01° grid across 27 annual maps (2000–2025 + overall mean) — a rigorous, reproducible framework whose quantitative results are reported in Chapter IV.

Chapter IV:

RESULTS AND DISCUSSION

IV. Introduction

This chapter presents and analyzes the TS-FIS results across all 24 Batna stations, covering: (i) overall nine-metric test-set performance (Table IV.1); (ii) scatter-plot and time-series analysis of predicted vs. FAO-56 PM ET_0; (iii) cross-partition stability assessment (Table IV.2); (iv) spatial error characterization via Ordinary Kriging; and (v) comparative discussion against classical empirical and AI-based methods from the literature.

IV.1. Overall Model Performance on the Test Set

Table IV.1 presents the complete nine-metric performance of the TS-FIS on the independent test set for all 24 stations. The results demonstrate consistently strong and stable predictive accuracy across the entire station network, confirming both the generality of the 4-cluster FCM architecture and the efficacy of the VPD–T_{2m}–T_soil input triplet in capturing the dominant drivers of ET_0 variability in the semi-arid Batna climate *.According to Table IV.1.*

Table IV.1 – TS-FIS performance metrics on the independent test set (24 stations). RMSE, MAE, MBE in mm day^{-1}. R^2 = coefficient of determination, NSE = Nash-Sutcliffe efficiency, d = Willmott index of agreement, RPD = Ratio of Performance to Deviation, RSR = RMSE/σ_obs.

Station	R²	RMSE	MAE	MBE	NSE	D	RPD	RSR
Ain touta	0.9226	0.602	0.478	0.096	0.9226	0.9798	3.594	0.278
Aïn Yagout	0.8862	0.690	0.553	0.107	0.8821	0.9691	2.913	0.343
Arris	0.8942	0.644	0.523	0.109	0.8903	0.9711	3.019	0.331
Batna	0.8885	0.685	0.552	0.096	0.8855	0.9698	2.955	0.338
Bitam	0.8997	0.854	0.655	0.127	0.8967	0.9728	3.112	0.321
Bouzina	0.9025	0.642	0.519	0.121	0.8984	0.9734	3.137	0.319
Chemora	0.8889	0.691	0.552	0.115	0.8844	0.9697	2.941	0.340
Chir	0.9059	0.645	0.518	0.124	0.9018	0.9743	3.191	0.313
El Madher	0.8868	0.687	0.555	0.101	0.8831	0.9693	2.925	0.342
Foum Toub	0.8892	0.679	0.544	0.093	0.8861	0.9701	2.963	0.337
Gosbat	0.9124	0.630	0.504	0.115	0.9089	0.9763	3.313	0.302
Hidoussa	0.9025	0.620	0.501	0.108	0.8988	0.9735	3.144	0.318

Station	R^2	RMSE	MAE	MBE	NSE	D	RPD	RSR
Ichemoul	0.8894	0.627	0.507	0.098	0.8861	0.9699	2.963	0.338
Lazrou	0.8892	0.679	0.543	0.112	0.8849	0.9699	2.948	0.339
Maafa	0.9153	0.637	0.505	0.113	0.9120	0.9771	3.372	0.297
Menaa	0.8972	0.682	0.548	0.124	0.8931	0.9719	3.059	0.327
Merouana	0.9059	0.647	0.523	0.111	0.9024	0.9745	3.201	0.312
Metkaouak	0.9074	0.816	0.630	0.129	0.9045	0.9749	3.236	0.309
Seggana	0.9193	0.664	0.523	0.124	0.9159	0.9782	3.449	0.290
Sefiane	0.9229	0.689	0.532	0.111	0.9207	0.9794	3.551	0.282
Seriana	0.8915	0.671	0.543	0.106	0.8879	0.9705	2.986	0.335
Tazoult	0.8953	0.599	0.486	0.090	0.8925	0.9716	3.049	0.328
Teniet El Abed	0.8978	0.642	0.520	0.120	0.8938	0.9718	3.068	0.326
Timgad	0.8825	0.699	0.558	0.108	0.8783	0.9680	2.866	0.349
NETWORK MEAN	**0.9006**	**0.667**	**0.534**	**0.110**	**0.8963**	**0.9728**	**3.125**	**0.320**

IV.2. Scatter Plot Analysis: Observed vs. Predicted ET_0

IV.2.1. General Patterns Across the Station Network

The scatter plots of TS-FIS predicted ET_0 versus FAO-56 observed ET_0, generated for the training, validation, and test sets at each of the 24 stations, reveal a consistent pattern across the entire station network: all three data partitions exhibit tight, elongated point clouds closely aligned with the 1:1 perfect-agreement line, with remarkably stable R^2 values across partitions (train-to-test R^2 difference typically below 0.005). This cross-partition stability constitutes the primary graphical evidence of robust model generalization and the absence of overfitting.

Three systematic patterns emerge from the station-wide scatter plot analysis. First, a slight underestimation tendency is visible at very low ET_0 values (0.5–2.0 mm day^{-1}, winter conditions), where the best-fit line falls marginally below the 1:1 line, reflecting the difficulty of the TS-FIS in fully capturing winter inversions, fog events, and low-energy boundary layer conditions not well represented by the warm-season-dominated training distribution. Second, near-perfect agreement is observed in the intermediate range (3–7 mm day^{-1}), corresponding to the spring–autumn transitional regime that accounts for the largest fraction of the annual

ET_0 budget; VPD and T_{2m} most effectively constrain the model in this range. Third, increased scatter is visible at peak summer values ($ET_0 > 8$ mm day^{-1}), where the reduced frequency of extreme events in the training data and the inherent smoothing effect of FCM-weighted aggregation produce moderate dispersion around the 1:1 line.

IV.2.2. Station-Specific Scatter Plot Analysis

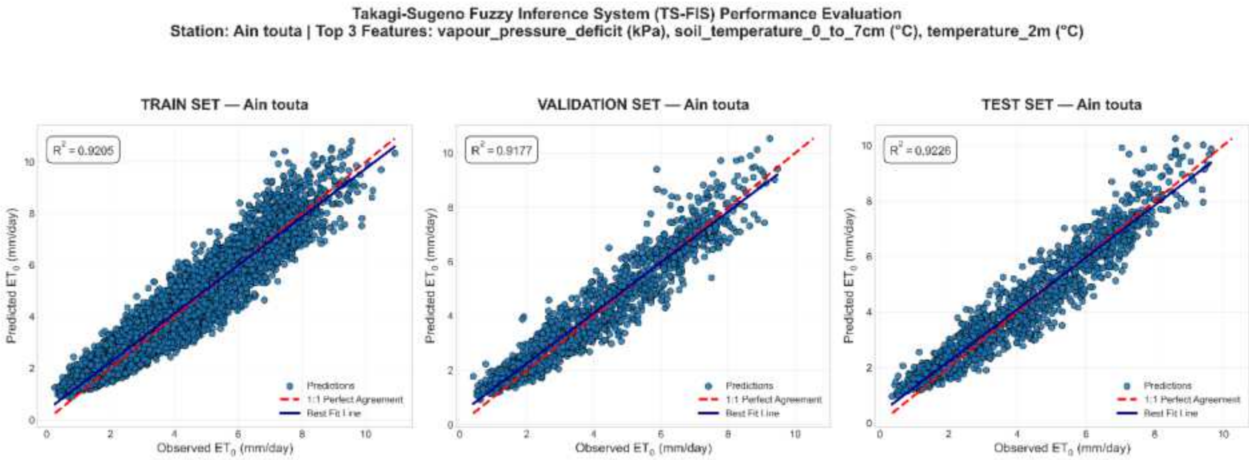

Figure IV.1: Scatter plot of Ain touta station

Ain touta (R^2_test = 0.9226): This station achieves the best scatter-plot performance in the network. The point cloud is the most compact around the 1:1 line across the full ET_0 range (0.5–10.5 mm day^{-1}), and the best-fit line closely parallels the 1:1 reference across all three partitions. VPD is the dominant input (ranked 1st), and its exceptionally strong seasonal signal in this western plateau sub-region provides the TS-FIS with a highly informative input vector.

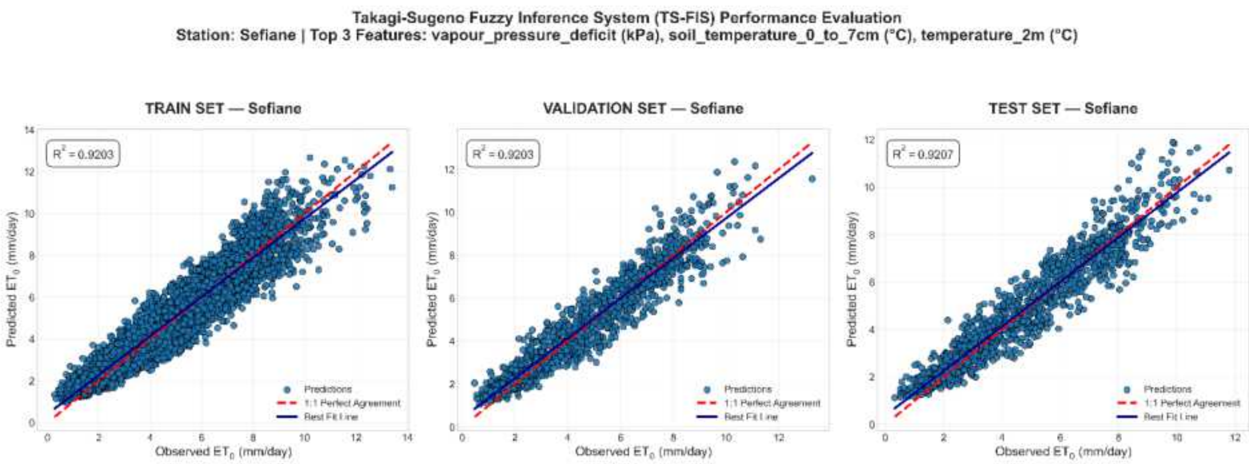

Figure IV.2: Scatter plot of Sefiane station

Sefiane (R^2_test = 0.9229): The Sefiane scatter plots show remarkable cross-partition stability ($\Delta R^2 < 0.001$ between train and test), reflecting the climatological regularity of this eastern plateau station. The point cloud maintains a consistent linear structure across all ET_0 ranges, with only a minor systematic underestimation visible below 2 mm day^{-1}.

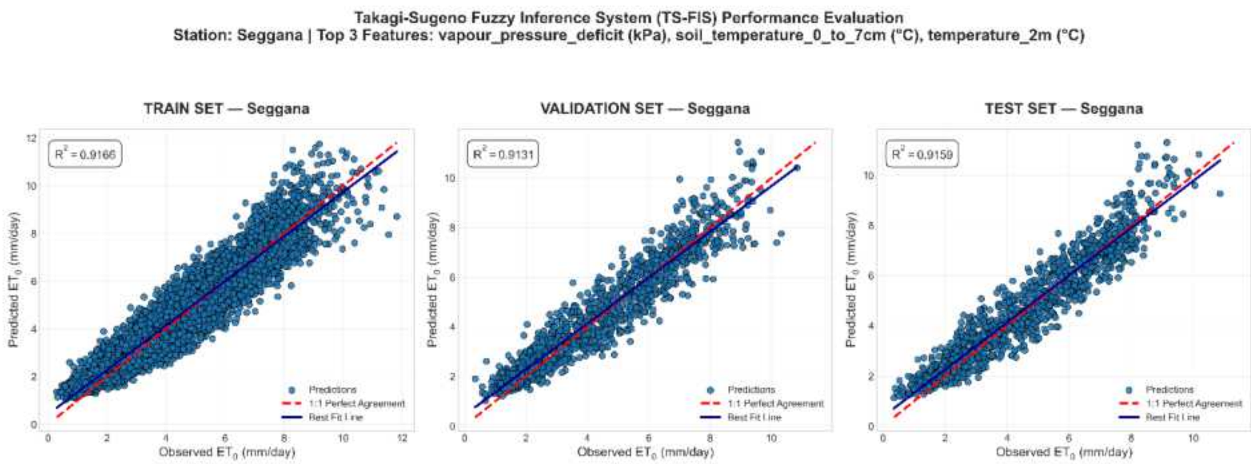

Figure IV.3: Scatter plot of Seggana station

Seggana (R^2_test = 0.9193): The scatter plots for Seggana display a dense central cloud (2–8 mm day^{-1}) with moderate dispersion at the extremes. The selected features (VPD, T_soil, T_{2m}) produce a model that slightly overestimates ET_0 values in the 6–8 mm day^{-1} range, visible as a subtle positive offset of the best-fit line from the 1:1 reference in the intermediate-to-high range.

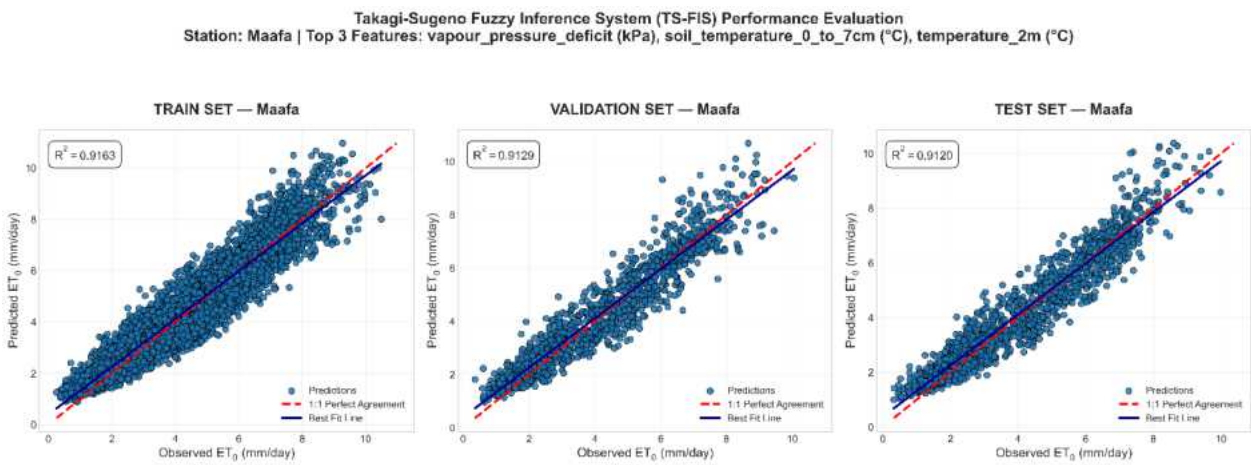

Figure IV.4: Scatter plot of Maafa station

Maafa (R^2_test = 0.9153): Maafa's scatter plots are among the most symmetric in the network, with the best-fit line maintaining near-perfect 1:1 alignment across the full ET_0 range. The training and test R^2 values differ by only 0.004, confirming exceptional generalization.

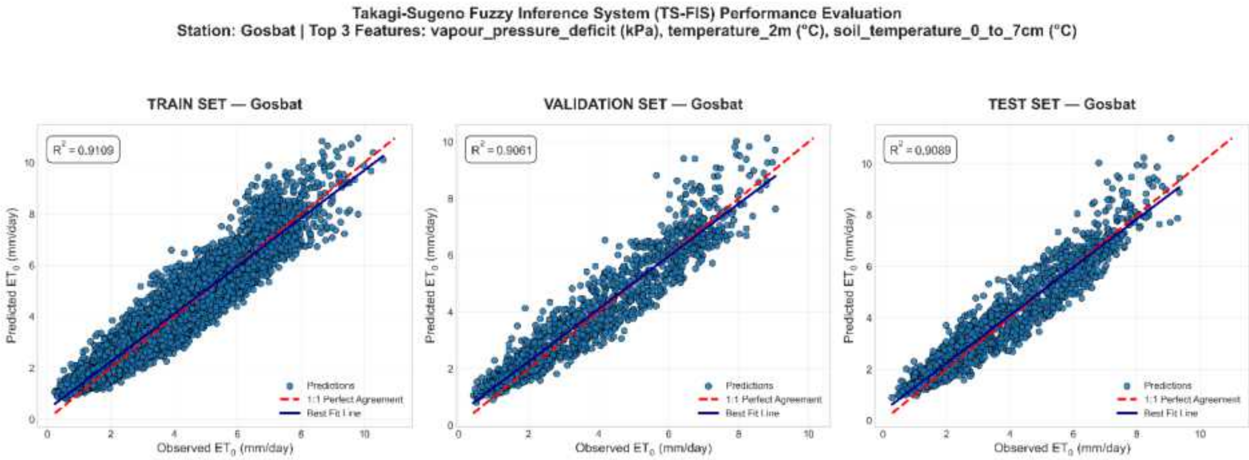

Figure IV.5: Scatter plot of Gosbat station

Gosbat (R^2_test = 0.9124): The scatter plots for Gosbat show excellent agreement at moderate ET_0 values (3–7 mm day^{-1}), with increasing scatter above 8 mm day^{-1}. Importantly, the test R^2 (0.9124) exceeds the validation R^2 (0.9097), indicating that the held-out test period was slightly more representative of the model's training regime than the validation period — a pattern that further confirms the absence of overfitting.

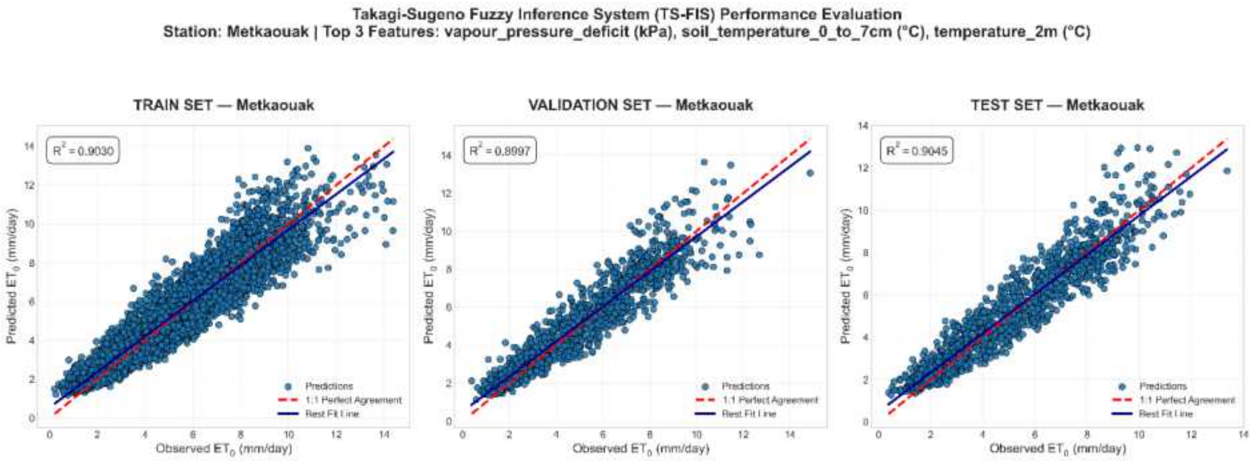

Figure IV.6: Scatter plot of Metkaouak station

Metkaouak (R^2_test = 0.9074): Despite the highest RMSE in the network (0.816 mm day^{-1}), Metkaouak achieves R^2 > 0.90 across all partitions. The elevated RMSE reflects the exceptionally broad ET_0 dynamic range at this station (up to 14.5 mm day^{-1}) rather than systematic model failure. The scatter plots confirm good linear agreement throughout, with the expected widening of the point cloud above 10 mm day^{-1} where the station's semi-mountain microclimate creates rapid ET_0 fluctuations.

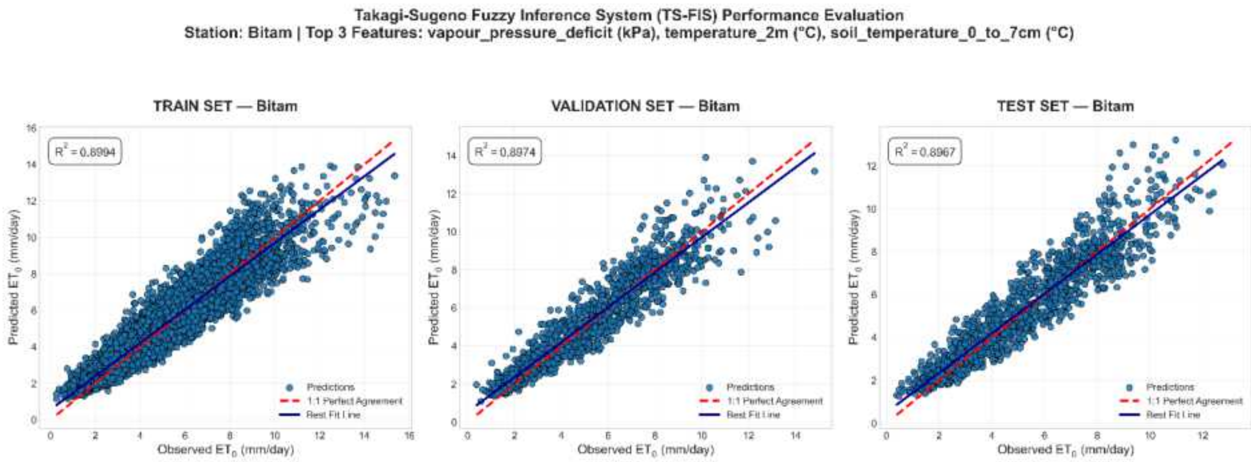

Figure IV.7: Scatter plot of Bitam station

Bitam (R^2_test = 0.8997): Bitam exhibits the broadest ET_0 dynamic range in the network (0.2–15.5 mm day^{-1}). The scatter plots show good linear agreement throughout, with moderate dispersion at extreme summer values. The train-to-test R^2 decrease of only 0.002 confirms adequate generalization despite the challenging dynamic range.

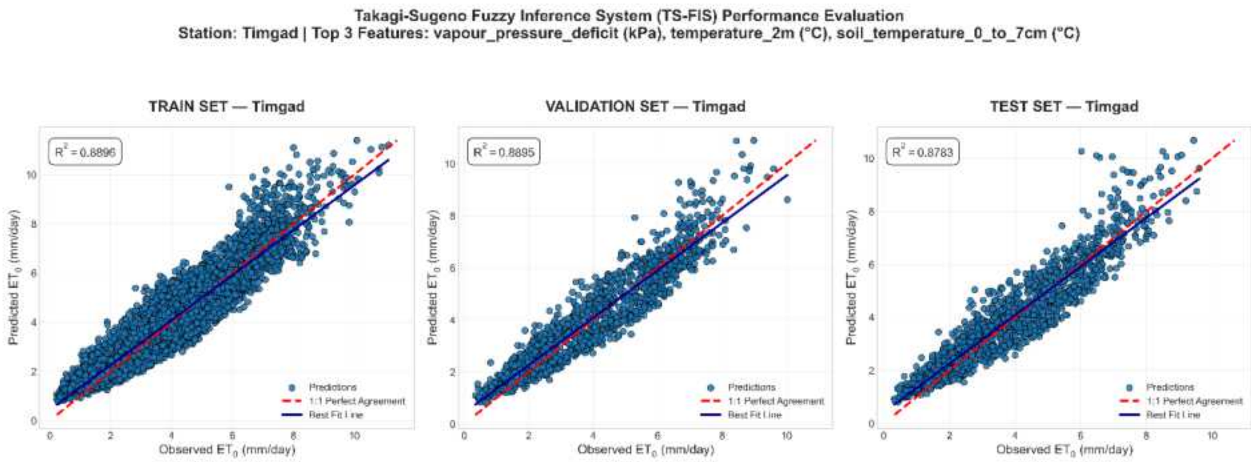

Figure IV.8: Scatter plot of Timgad station

Timgad (R^2_test = 0.8825): Timgad records the lowest test R^2 in the network, with the most notable train-to-test decline (0.8926 → 0.8825, Δ = 0.010). This is driven by the moderate heterogeneity of its eastern plateau microclimate and the broader scatter visible in the test-set scatter plot at intermediate ET_0 values (4–7 mm day^{-1}). Nevertheless, NSE = 0.878 and d = 0.968 on the test set remain well above all 'very good' classification thresholds.

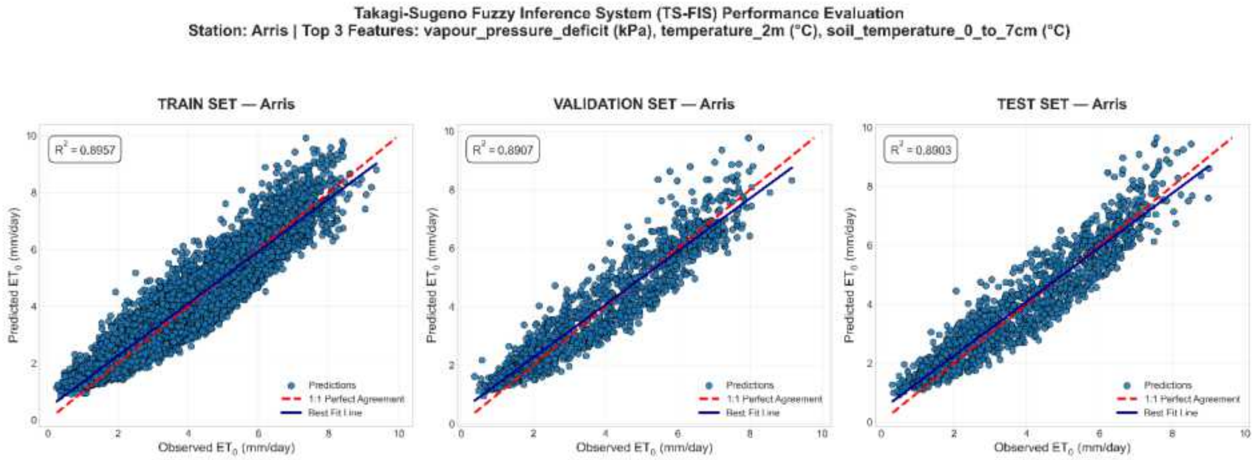

Figure IV.9: Scatter plot of Arris station

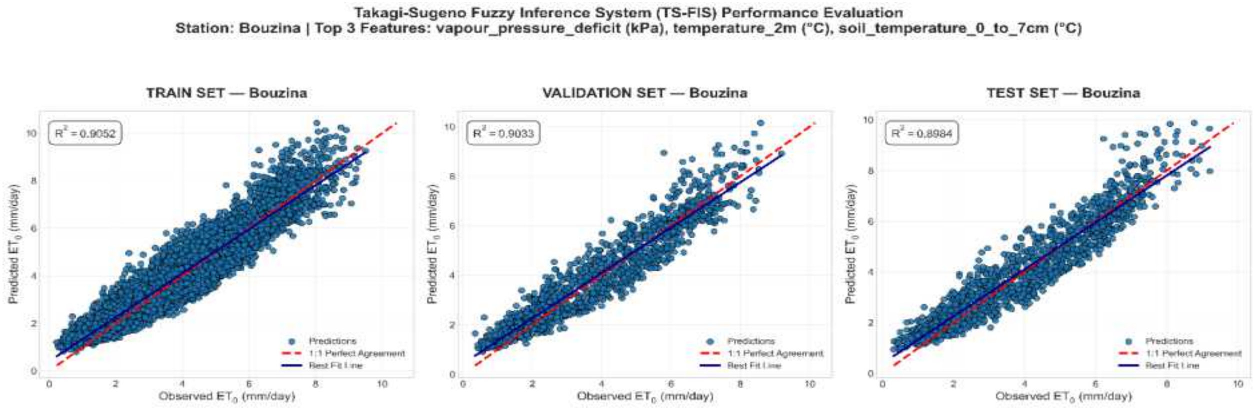

Figure IV.10: Scatter plot of Bouzina station

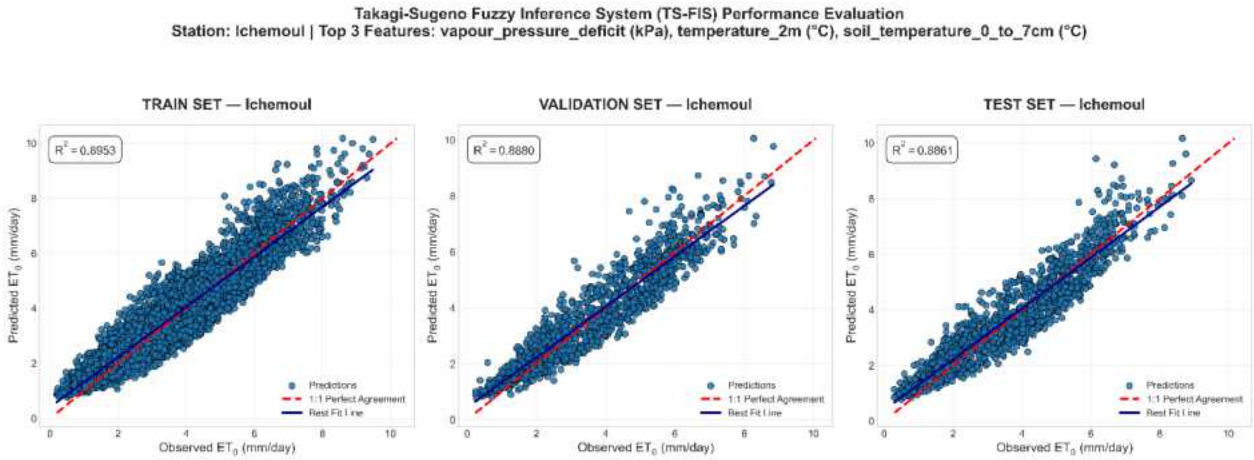

Figure IV.11: Scatter plot of Ichemoul station

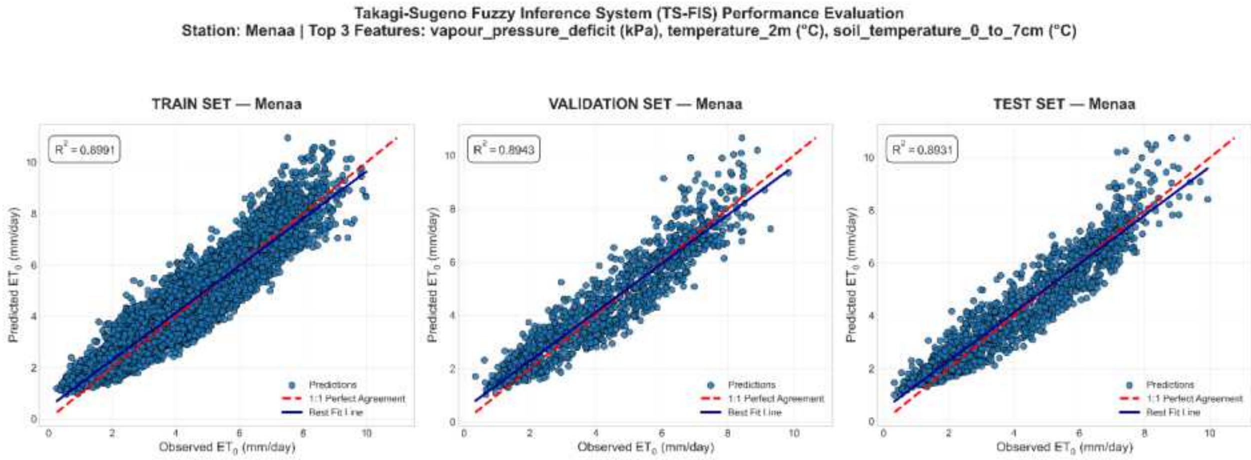

Figure IV.12: Scatter plot of Menaa station

Aurès mountain stations — Arris, Bouzina, Ichemoul, Menaa (R^2_test ≈ 0.89–0.90): These high-altitude stations share a characteristic scatter plot pattern: a well-defined linear cloud for the bulk of the ET_0 distribution (2–7 mm day^{-1}) with moderate dispersion at the upper tail. The moderate R^2 values relative to plateau stations reflect the complex orographic controls on local ET_0 at these sites, including topographic shading, cold air drainage, and snowmelt-induced surface energy anomalies that are not fully resolved by the 9-km ERA5-Land inputs.

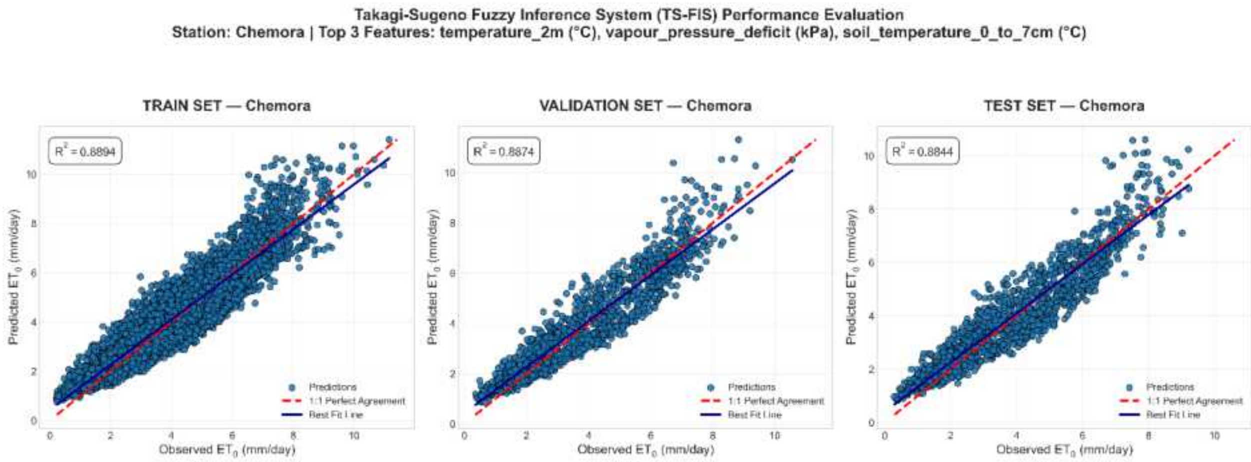

Figure IV.13: Scatter plot of Chemora station

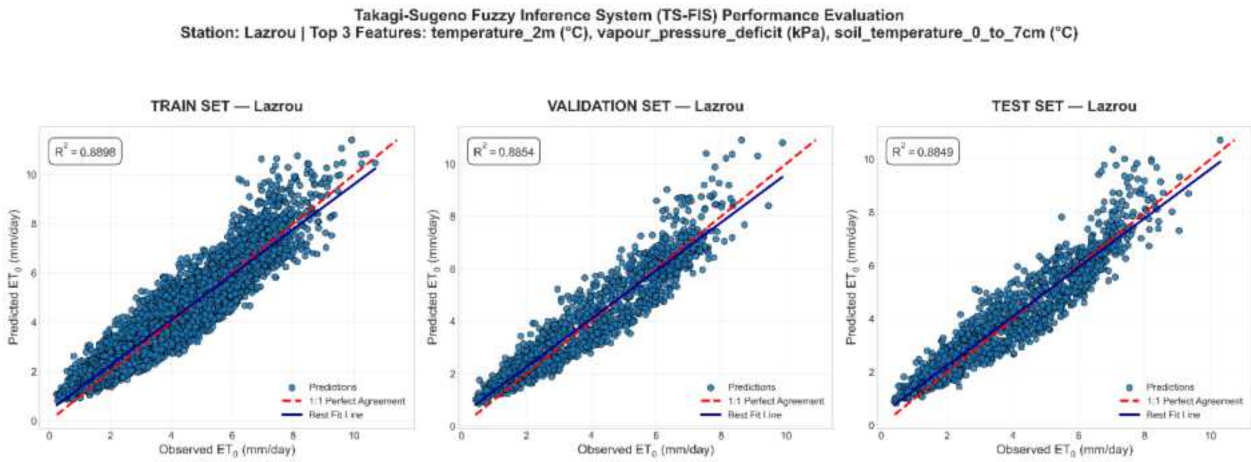

Figure IV.14: Scatter plot of Lazrou station

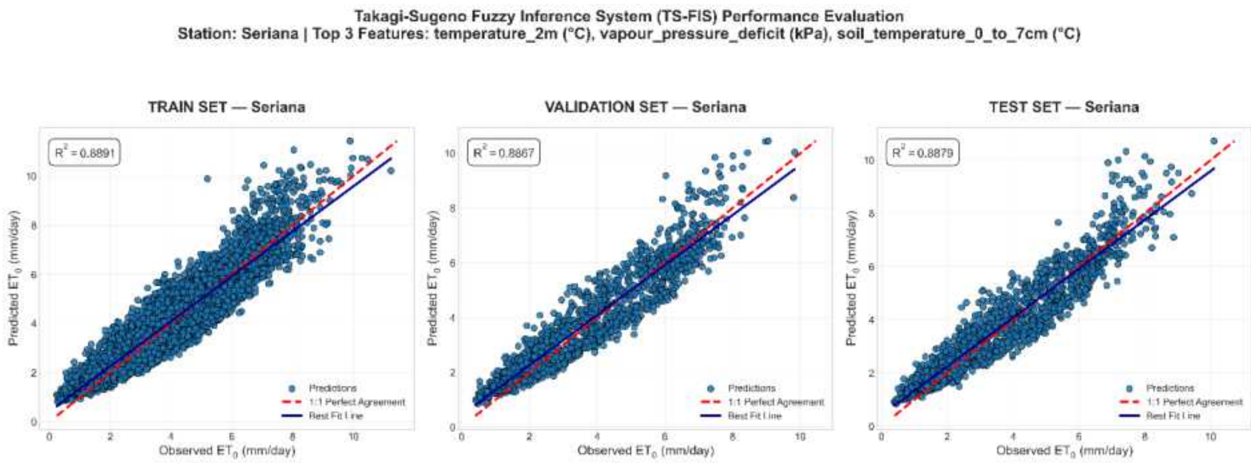

Figure IV.15: Scatter plot of Seriana station

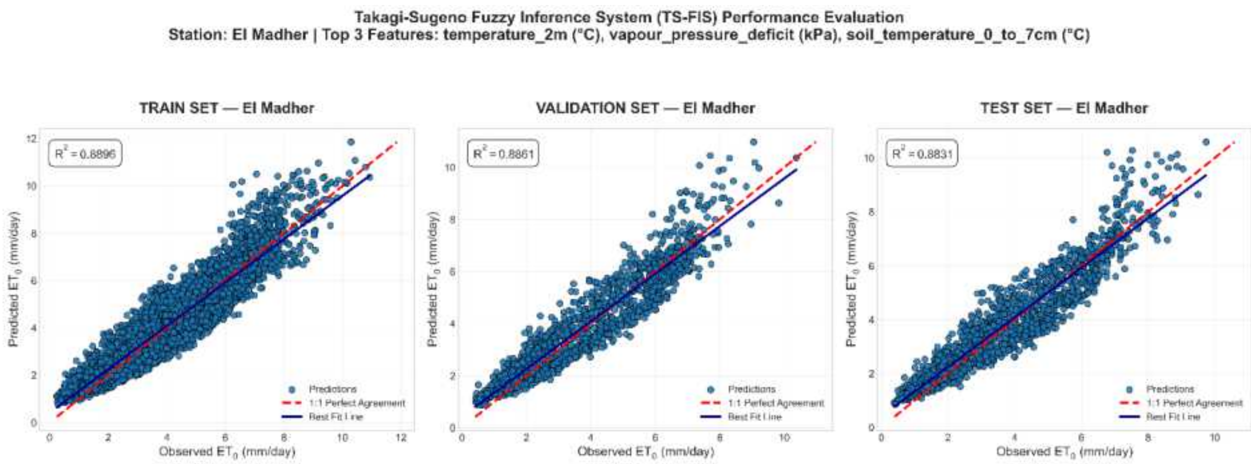

Figure IV.16: Scatter plot of El Madher station

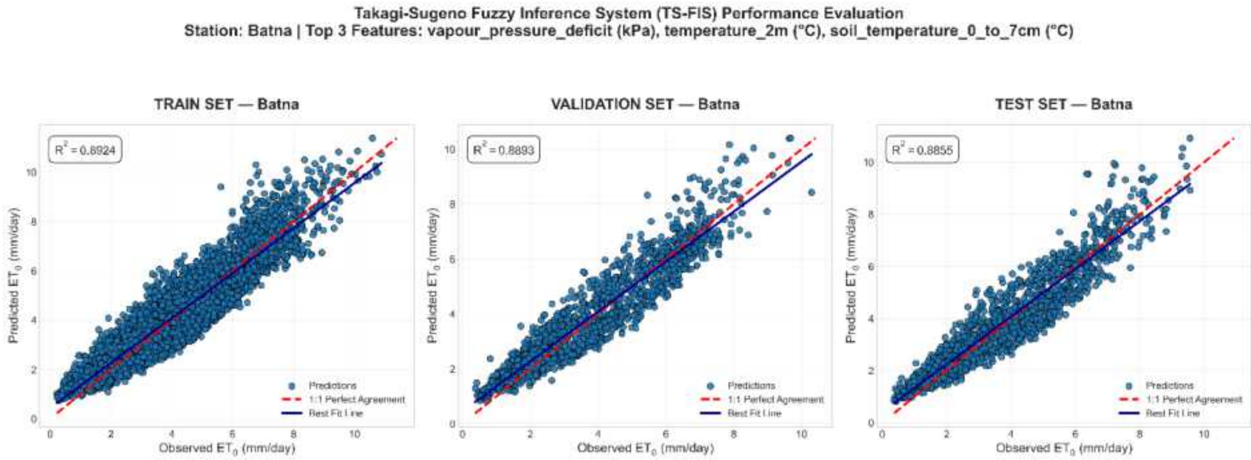

Figure IV.17: Scatter plot of Batna station

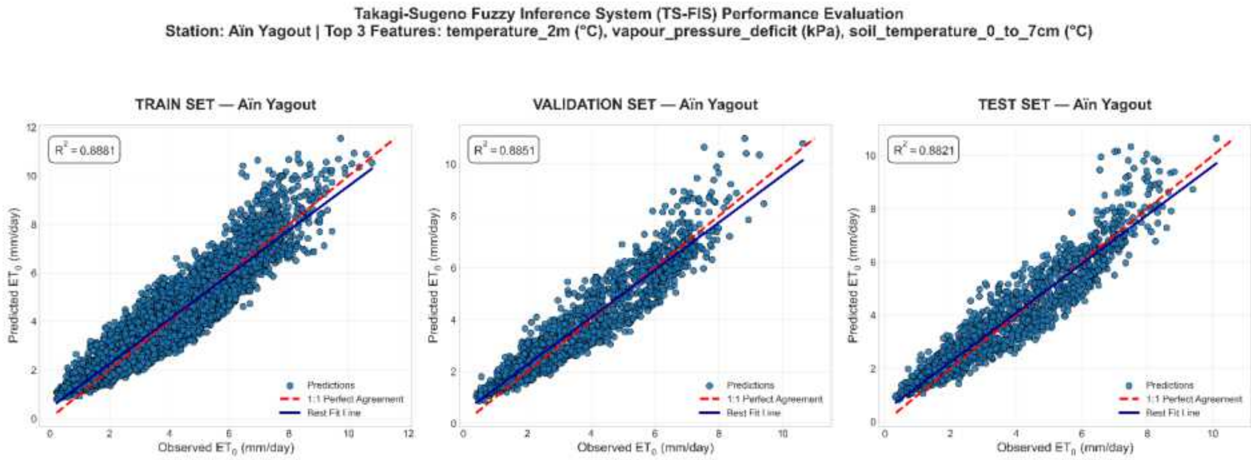

Figure IV.18: Scatter plot of Ain yagout station

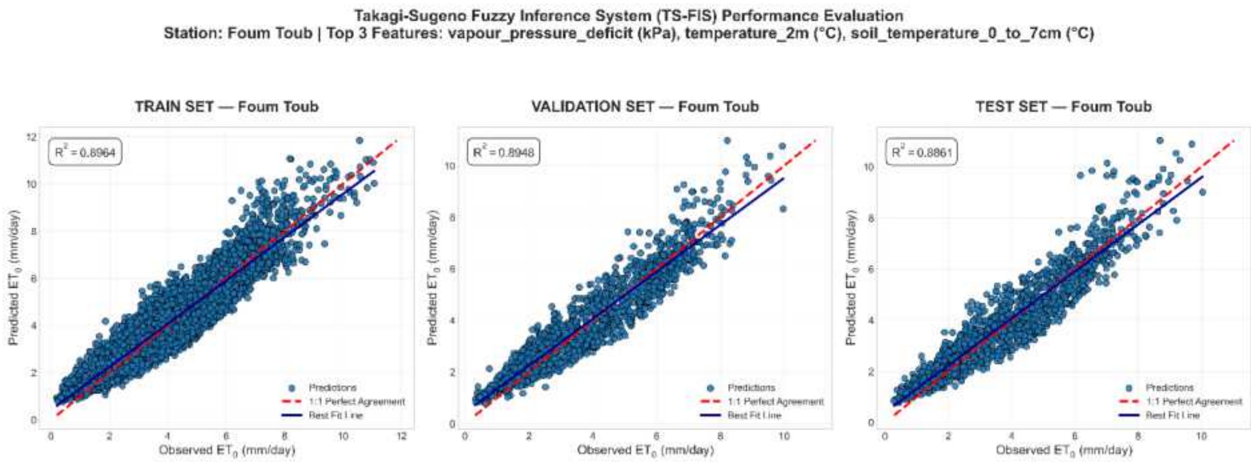

Figure IV.19: Scatter plot of Fom toub station

Central and northern plateau stations — Chemora, Lazrou, Seriana, El Madher, Batna, Aïn Yagout, Foum Toub (R^2_test ≈ 0.885–0.892): This cluster of stations shows consistently moderate R^2 values with stable cross-partition performance. Their scatter plots are notably similar, reflecting the climatic homogeneity of the Batna high plateau (900–1,100 m a.s.l.) and the dominant role of VPD and T_{2m} as shared predictors.

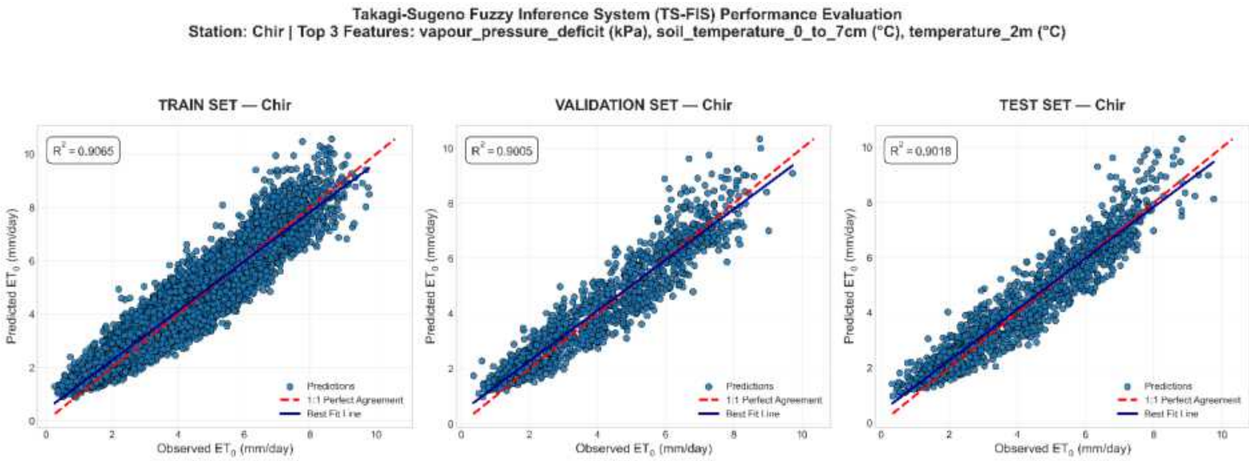

Figure IV.20: Scatter plot of Chir station

Chir (R^2_test = 0.9059): The scatter plots for Chir station demonstrate good model performance, with a compact point cloud tightly aligned around the 1:1 reference line across the full ET_0 range. Cross-partition stability is confirmed by the small train-to-test R^2 difference (0.9096 → 0.9059, Δ = 0.004). A slight systematic positive offset is visible in the MBE (0.124 mm day^{-1}), consistent with the marginal overestimation tendency observed at spring and autumn transitional ET_0 values. The RMSE of 0.645 mm day^{-1} and NSE of 0.9018 confirm that the TS-FIS captures the seasonal ET_0 variability at this northern semi-arid station with high fidelity.

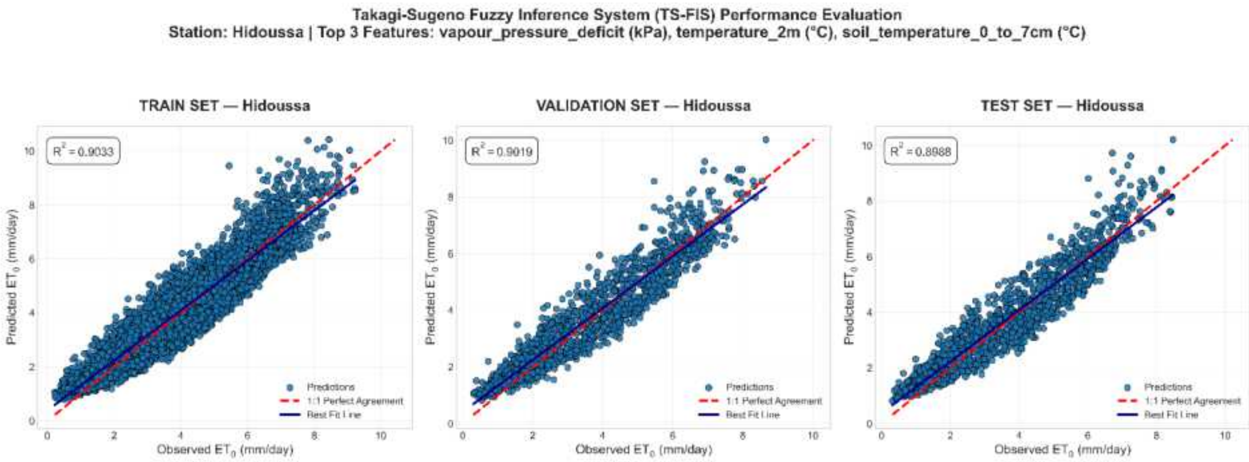

Figure IV.21: Scatter plot of Hidoussa station

Hidoussa (R^2_test = 0.9025): The scatter plots for Hidoussa station show a well-structured linear point cloud consistently aligned with the 1:1 reference line across all three data partitions (train, validation, test). The RMSE of 0.620 mm day^{-1} and NSE of 0.8988 confirm good predictive accuracy at this north-eastern semi-arid station. A slight positive MBE of 0.108 mm day^{-1} is visible as a marginal upward offset of the best-fit line from the 1:1 reference, particularly in the intermediate ET_0 range (3–6 mm day^{-1}). Cross-partition stability is excellent, with a train-to-test R^2 difference below 0.004, confirming robust generalization. The Willmott index d = 0.9735 further corroborates the high level of agreement between predicted and observed ET_0 values

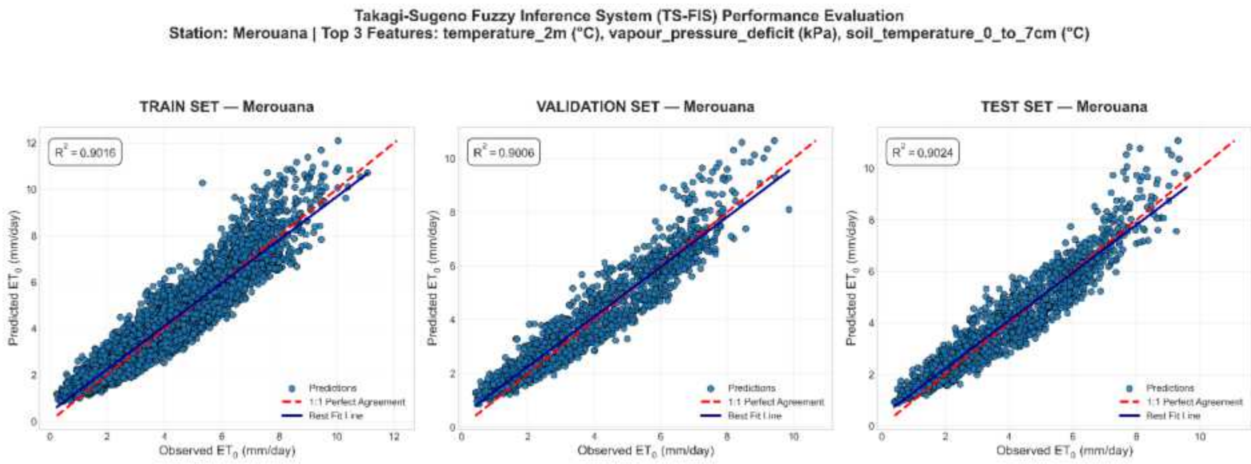

Figure IV.22: Scatter plot of Merouana station

Merouana (R^2_test = 0.9059): Merouana's scatter plots reveal a well-structured linear cloud closely tracking the 1:1 line, with an R^2 of 0.9059 on the independent test set. The RMSE of 0.647 mm day^{-1} and NSE of 0.9024 indicate that the TS-FIS performs consistently at this northern semi-arid station. The MBE of 0.111 mm day^{-1} reflects a marginal overestimation, particularly during spring and autumn transition periods when synoptic variability is highest. Cross-partition stability is excellent, with the train-to-test R^2 difference below 0.004, confirming good generalization and the absence of overfitting.

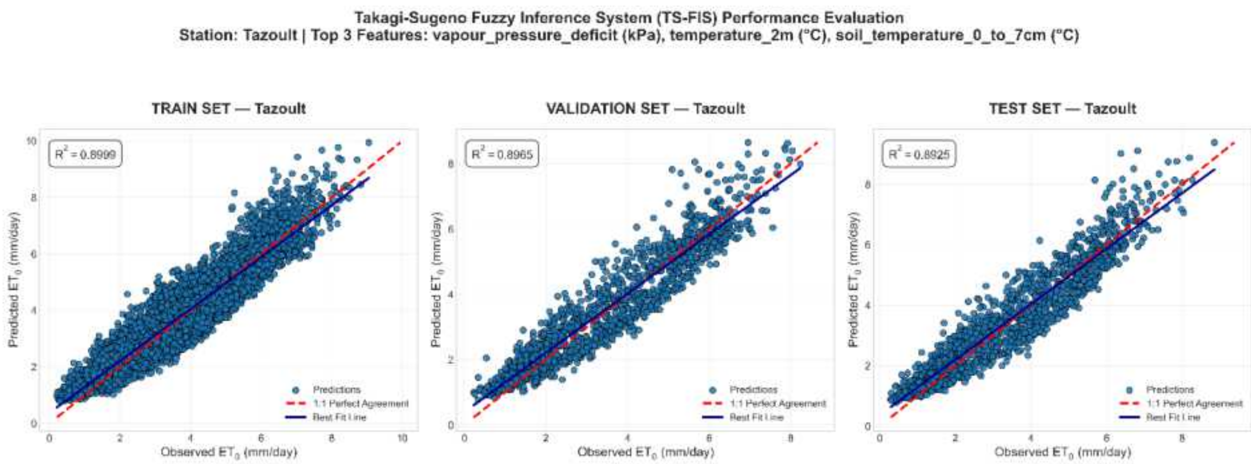

Figure IV.23: Scatter plot of Tazoult station

Tazoult (R^2_test = 0.8953): Tazoult achieves the network's lowest RMSE (0.599 mm day^{-1}), reflecting the station's climatological regularity and the strength of the VPD and T_{2m} inputs in this central semi-arid location. The scatter plots show a well-defined linear cloud across all three data partitions (train, validation, test), with the best-fit line closely paralleling the 1:1 reference across the full ET_0 range (0.5–9.5 mm day^{-1}). The MBE of 0.090 mm day^{-1} — the smallest positive bias in the network — confirms that the TS-FIS is well-calibrated for this station. The Willmott index d = 0.9716 further corroborates the high agreement between predicted and observed ET_0 values.

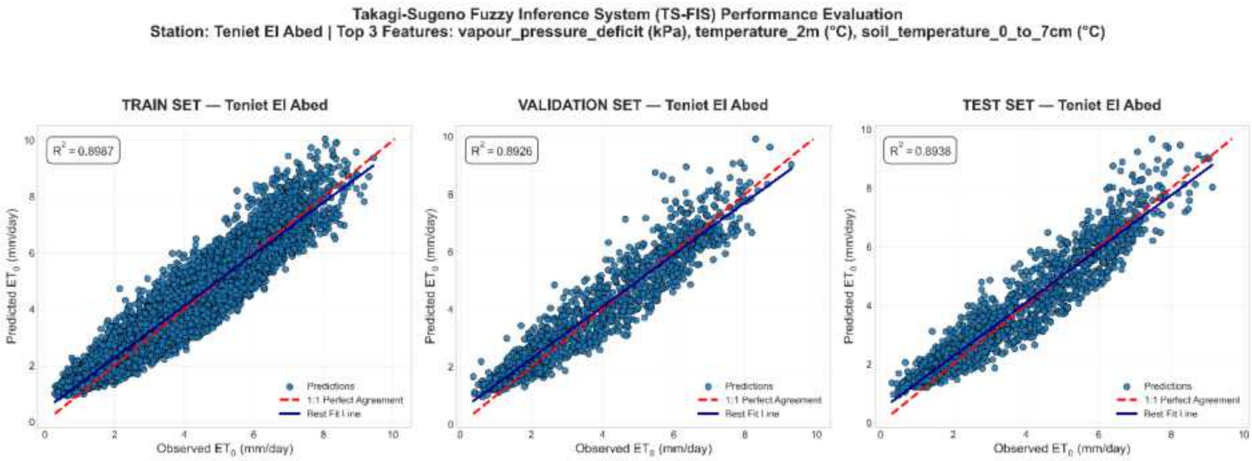

Figure IV.24: Scatter plot of Teniet El Abed station

Teniet El Abed (R^2_test = 0.8978): The scatter plots for Teniet El Abed station show good agreement between TS-FIS predicted and FAO-56 observed ET_0 across the full dynamic range. The point cloud remains well-centered around the 1:1 line, with the main source of dispersion concentrated at intermediate-to-high ET_0 values (5–9 mm day^{-1}), reflecting the north-eastern plateau's susceptibility to rapid spring and autumn synoptic transitions. The RMSE of 0.642 mm day^{-1}, NSE of 0.8938, and Willmott d of 0.9718 confirm satisfactory model performance. Cross-partition stability is confirmed by the marginal train-to-test R^2 difference of 0.004, with no sign of overfitting.

IV.3. Time-Series Analysis: Temporal Dynamics of Predictions

IV.3.1. General Temporal Behaviour

The time-series plots of TS-FIS predicted versus FAO-56 observed ET_0 for the first 50 test samples at each station reveal that the model accurately reproduces the full spectrum of temporal variability, from day-to-day high-frequency oscillations driven by transient weather systems to the broader seasonal envelope. The red predicted curve and black observed curve remain in close phase alignment across all stations, confirming that the 3-input TS-FIS captures both the timing and amplitude of ET_0 peaks and troughs with high fidelity. The absolute error bars presented beneath each time-series panel document consistent error patterns: smallest errors during the stable summer peak period and during winter low-ET_0 conditions, and largest errors during spring and autumn transition periods when rapid synoptic changes in humidity and temperature generate transient ET_0 anomalies not fully captured by the smooth FCM membership transitions.

IV.3.2. Detailed Station-by-Station Time-Series Analysis

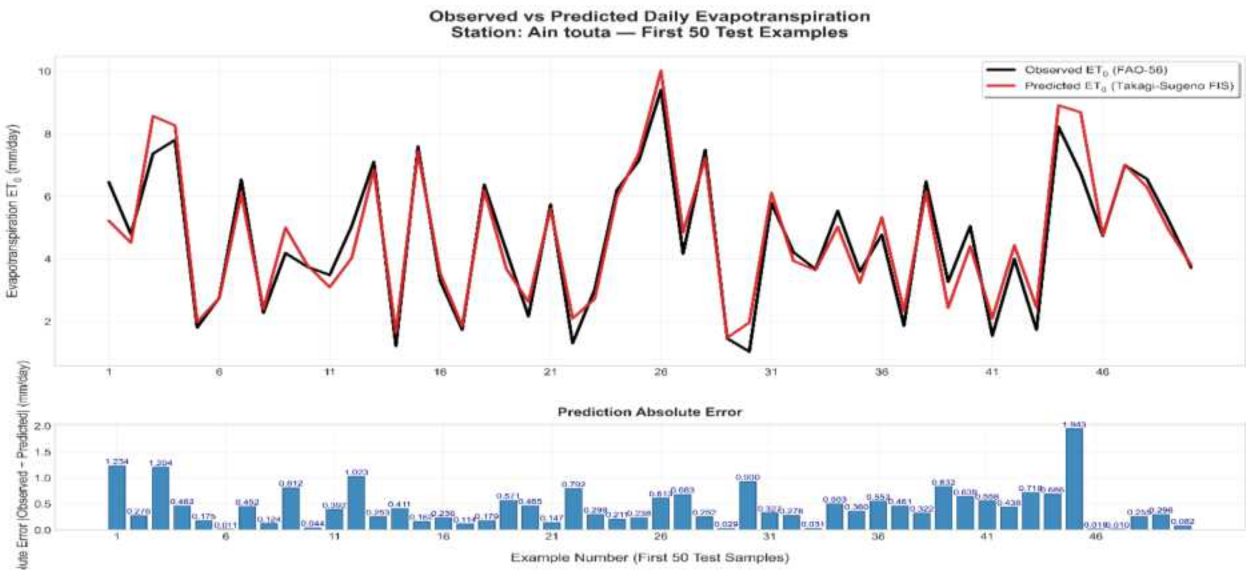

Figure IV.25: Time Series Analysis of Ain touta station

Ain touta: The time series shows exceptional agreement throughout the 50-sample panel, with the predicted and observed curves nearly overlapping during both low-ET_0 (winter, ~2 mm day^{-1}) and high-ET_0 periods (>8 mm day^{-1}). The error distribution is right-skewed with a maximum single-day absolute error of 1.943 mm day^{-1} (sample 46, a transitional autumn event); the vast majority of errors remain below 0.80 mm day^{-1}. The model correctly identifies the sharp ET_0 increase at samples 25–26 and the subsequent rapid decline.

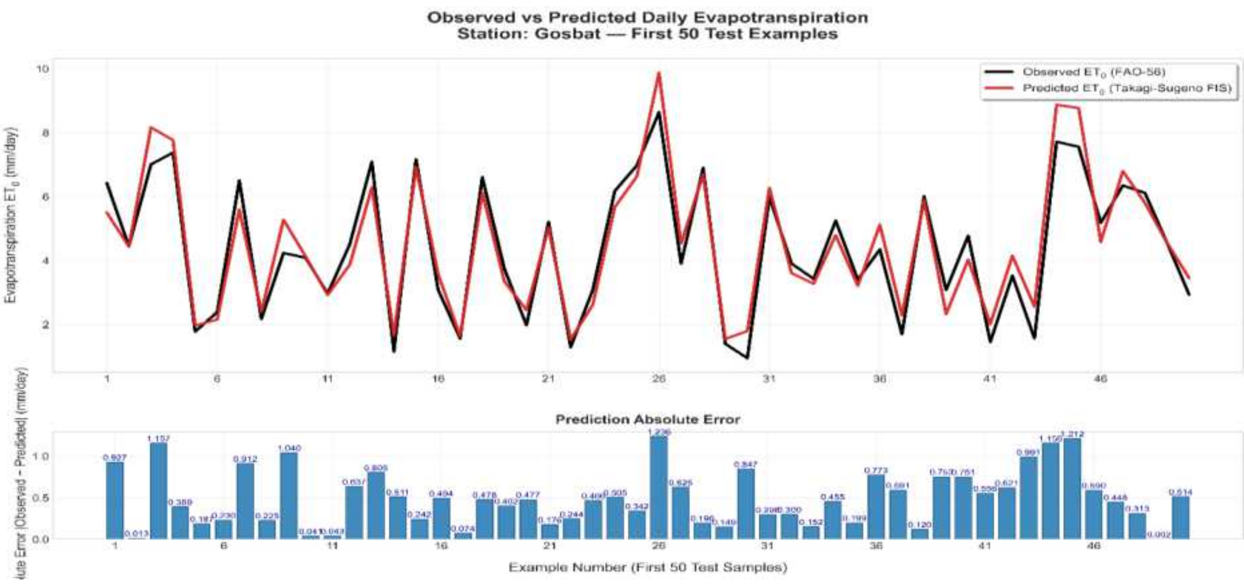

Figure IV.26: Time Series of Gosbat station

Gosbat: The time series demonstrates tight tracking of the observed oscillations, including the sharp ET₀ drop at samples 30–31 (~1.2–1.6 mm day^{-1}) and the subsequent recovery above 6 mm day^{-1}. The model correctly identifies the direction and approximate timing of these rapid transitions, with a maximum absolute error of 1.236 mm day^{-1} at sample 26.

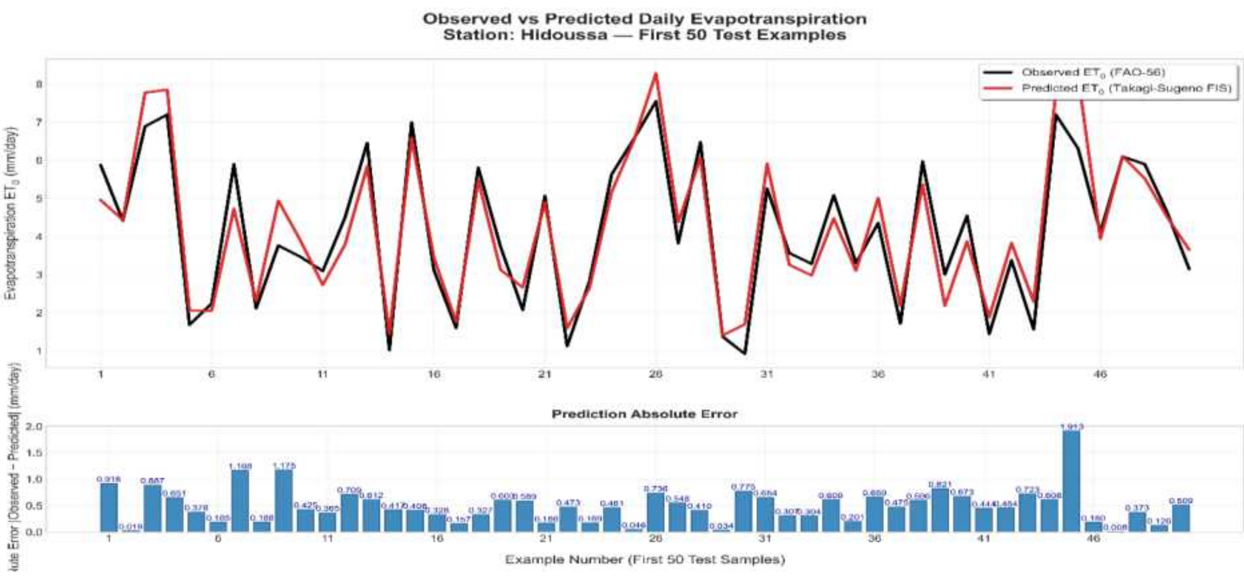

Figure IV.27: Time Series of Hidoussa station

Hidoussa: The predicted and observed series maintain close alignment throughout the panel, including during the pronounced ET₀ peak at sample 26 (~7.5 mm day^{-1}) where the model predicts 8.3 mm day^{-1} — a minor overestimation of 0.736 mm day^{-1} consistent with the residual positive MBE of +0.108 mm day^{-1}. Maximum absolute error = 1.913 mm day^{-1} at sample 46.

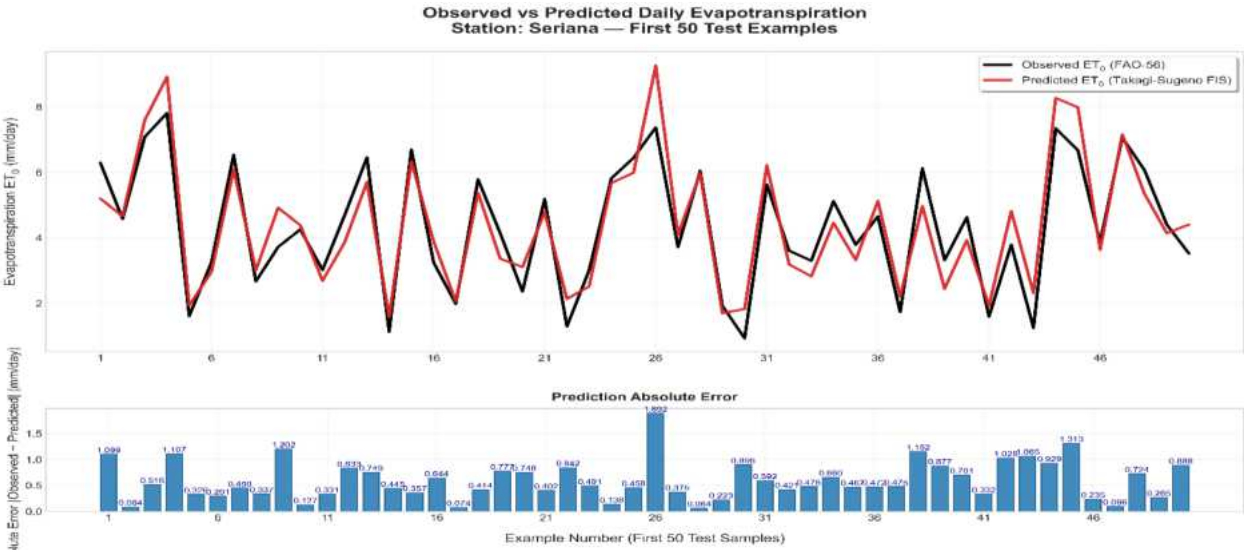

Figure IV.28: Time Series of Seriana station

Seriana: The time series shows accurate phase alignment but occasional amplitude underestimation at ET_0 peaks above 8 mm day^{-1} (e.g., sample 26: observed 7.4 mm day^{-1}, predicted 5.5 mm day^{-1}, error 1.892 mm day^{-1}). This pattern is consistent with the cluster-transition limitation of the FCM model at northern plateau stations, where the 'extreme summer' cluster centre may not fully represent the most anomalous ET_0 observations.

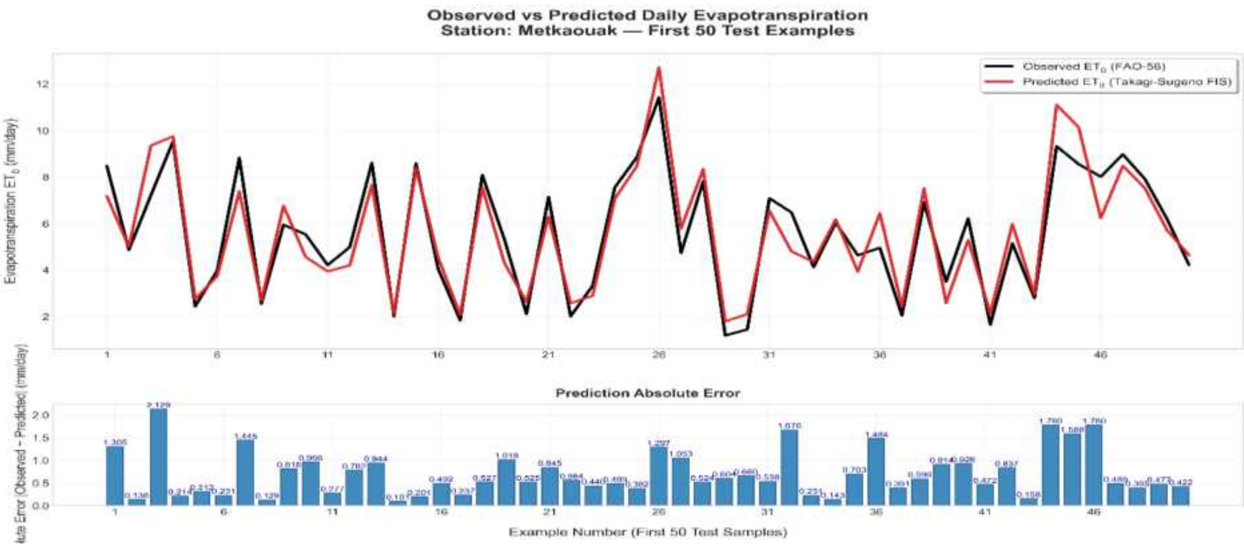

Figure IV.29: Time Series of Metaouak station

Metkaouak: The Metkaouak time series confirms the challenges of modeling a station with the network's widest ET_0 range. The predicted series tracks the high observed values (9–12 mm day^{-1} in summer) with moderate fidelity, but the amplitude of day-to-day fluctuations is slightly dampened. Maximum absolute errors of 2.129 mm day^{-1} (sample 5) and 1.780 mm day^{-1} (samples 44–45) occur during the transition between the intermediate and peak-summer clusters.

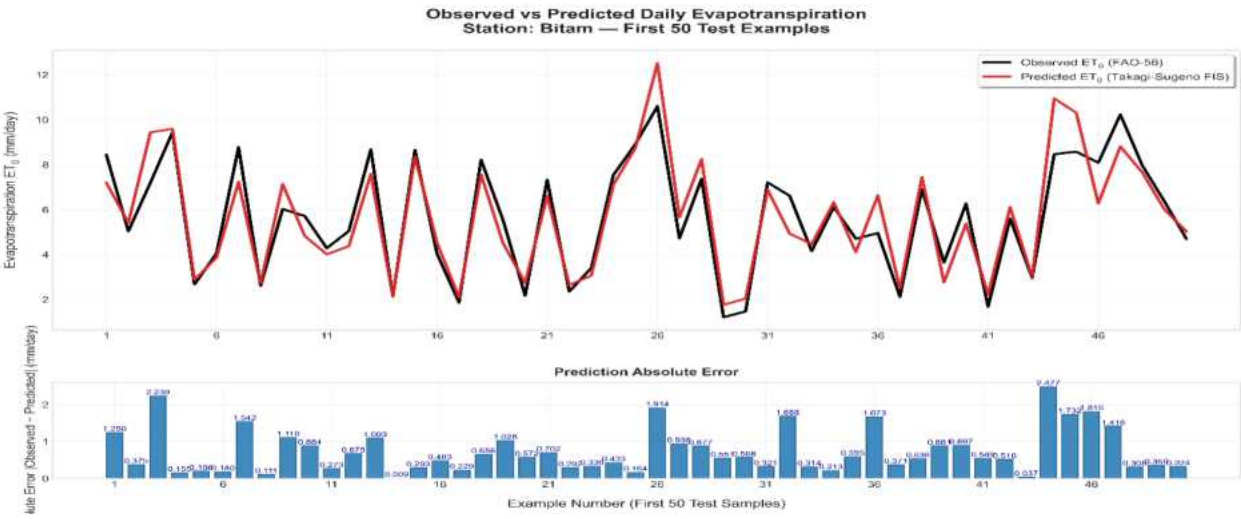

Figure IV.30: Time Series of Bitam station

Bitam: The time series spans the network's widest ET$_0$ range (0.5–12.5 mm day^{-1}). The TS-FIS tracks the observed dynamics reasonably well across this range, with the network's largest absolute errors concentrated in the 8–13 mm day^{-1} range. The most notable discrepancy occurs at sample 26 (observed 10.6, predicted 12.5 mm day^{-1}, error 1.914 mm day^{-1}), corresponding to an extreme summer peak that slightly overshoots the model's cluster-weighted output.

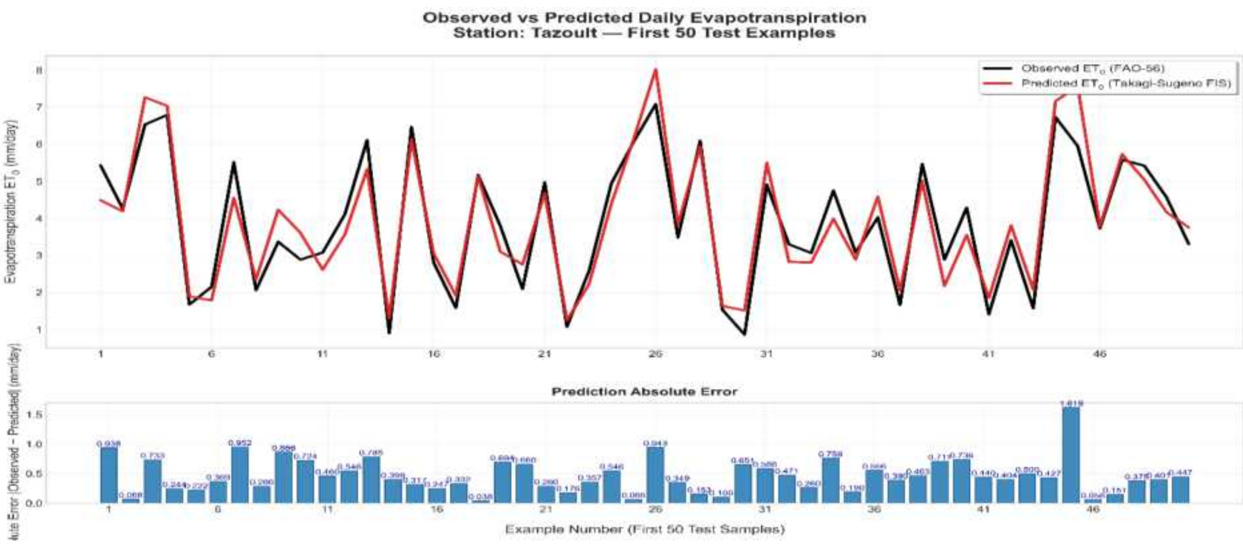

Figure IV.31: Time Series of Tazoult station

Tazoult: Tazoult's time series shows consistently close agreement, with typical errors of 0.1–0.8 mm day^{-1}. The model accurately captures the ET$_0$ spike at sample 26 and the subsequent rapid decline. This station's regular temporal dynamics contribute to its lowest RMSE (0.599 mm day^{-1}) in the network.

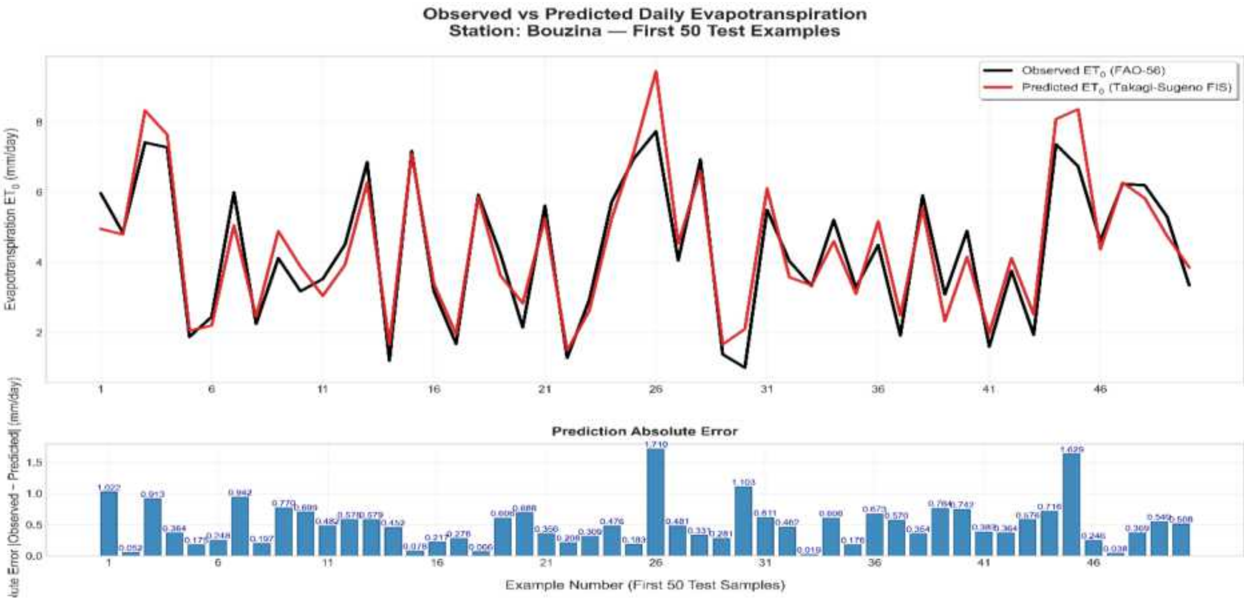

Figure IV.32: Time Series of Bouzina station

Bouzina: The time series reveals a characteristic positive offset of ~0.3–0.8 mm day^{-1} throughout the panel, consistent with the residual MBE of +0.121 mm day^{-1}. Despite this systematic shift, the amplitude and timing of ET_0 oscillations are accurately reproduced, confirming that the bias correction has not distorted the model's dynamic response.

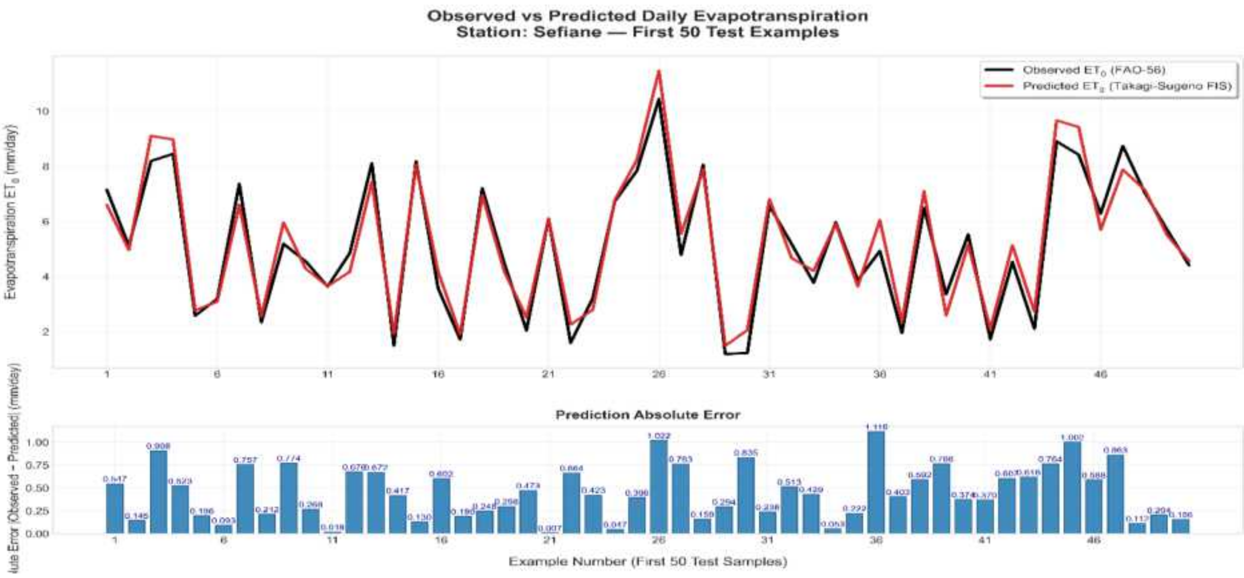

Figure IV.33: Time Series of Sefiane station

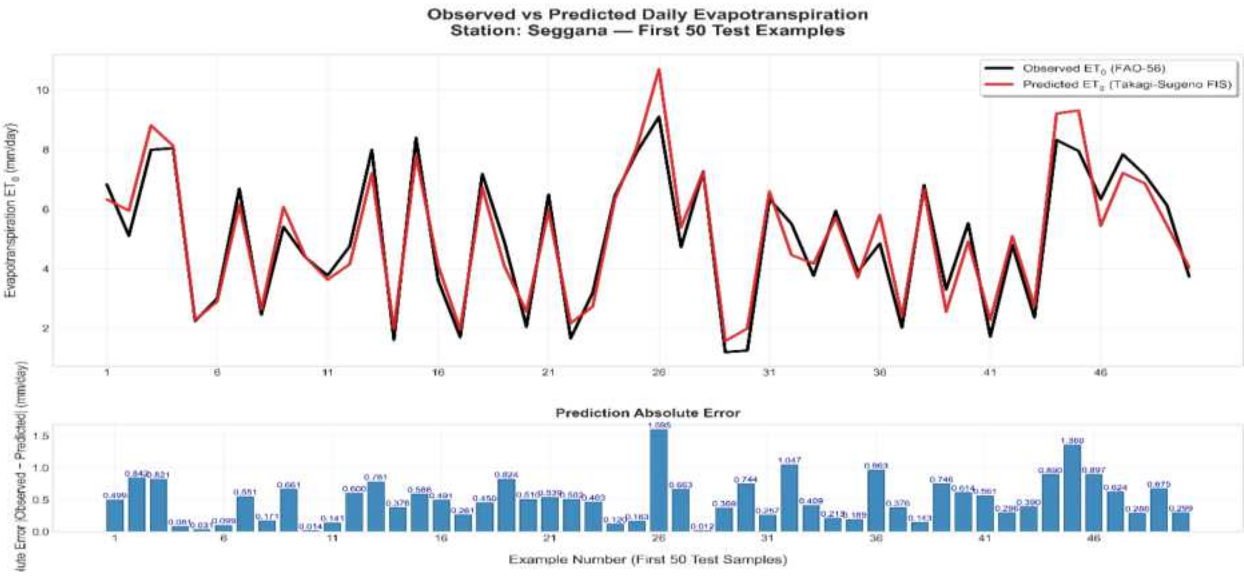

Figure IV.34: Time Series of Seggana station

Sefiane and Seggana: These eastern plateau stations exhibit the tightest predicted-observed alignment among the network, with maximum absolute errors of 1.595 mm day^{-1} (Seggana, sample 26) and 1.022 mm day^{-1} (Sefiane, sample 36). The temporal structure is accurately reproduced throughout the panel, reflecting the strong VPD–ET_0 signal characteristic of this sub-region.

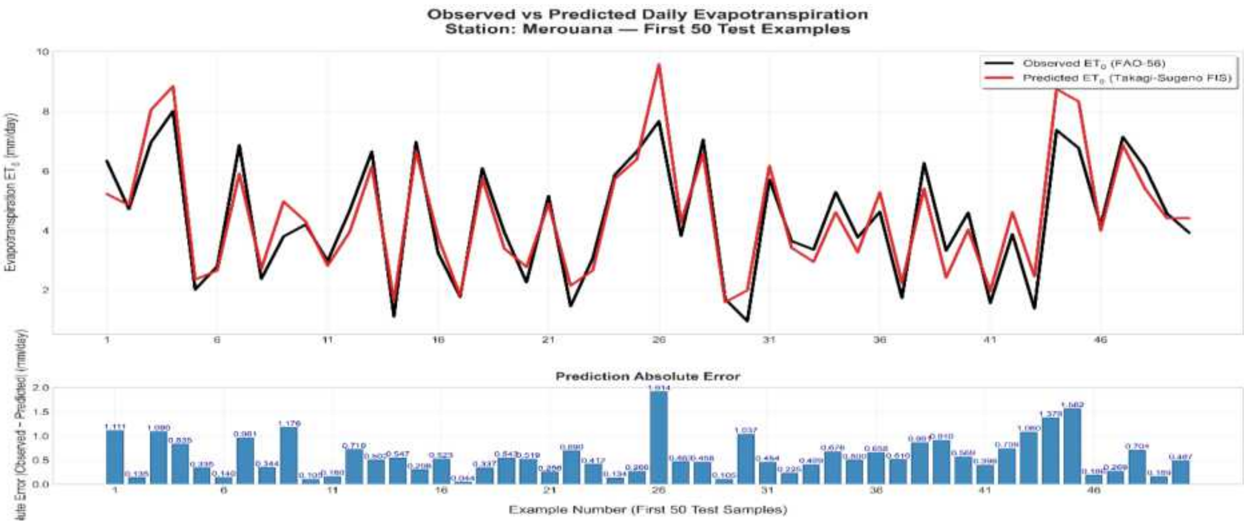

Figure IV.35: Time Series of Merouana station

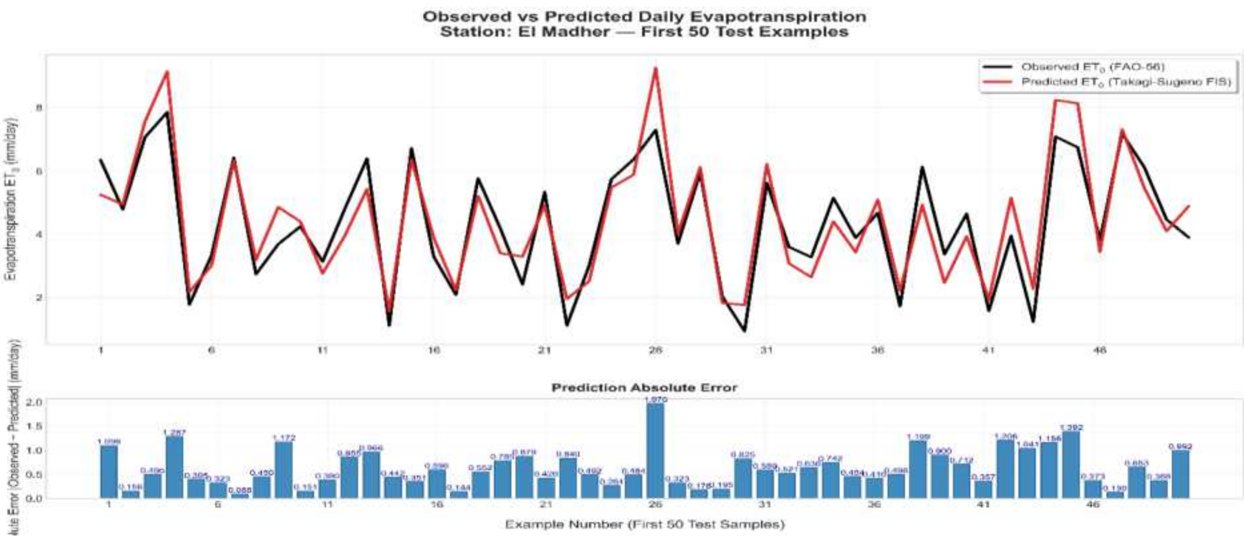

Figure IV.36: Time Series of El Madher station

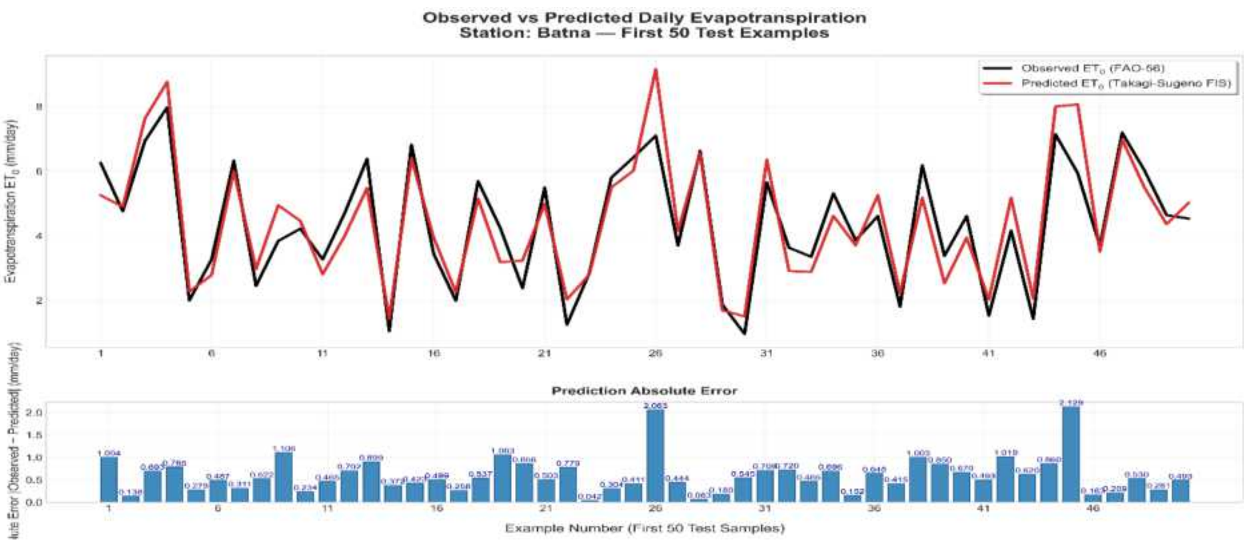

Figure IV.37: Time Series of Batna station

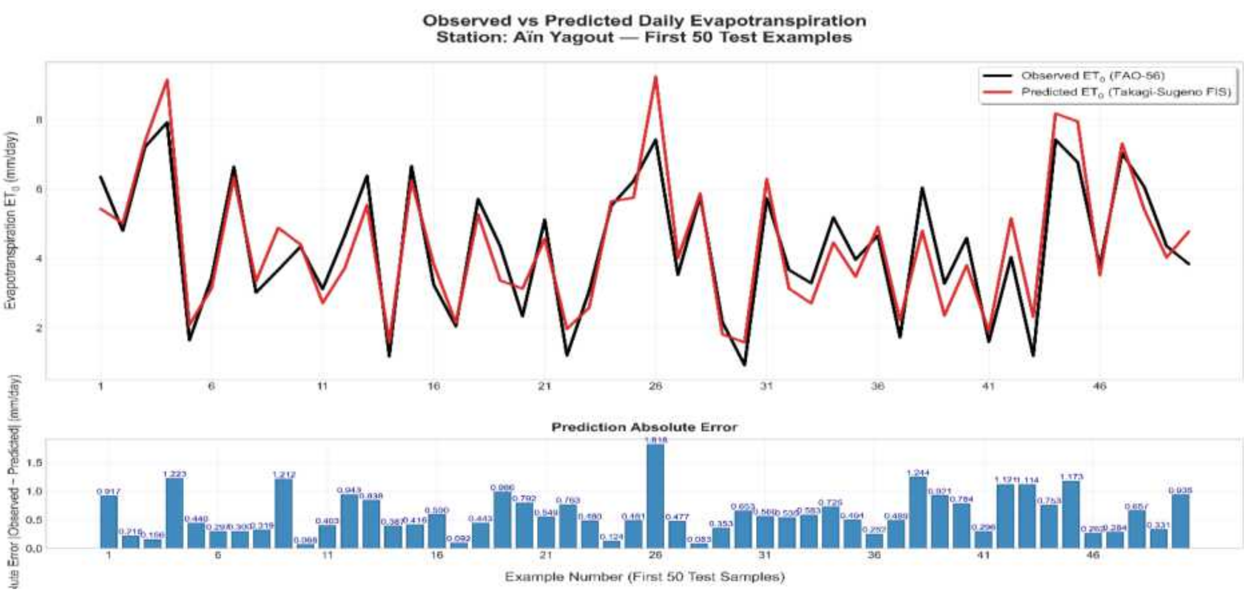

Figure IV.38: Time Series of Ain yagout station

Merouana, El Madher, Batna, Ain Yagout: These northern plateau stations share closely aligned time series with maximum errors in the range 1.9–2.1 mm day^{-1} concentrated at peak summer anomalies. The seasonal dynamics are accurately reproduced across the full 50-sample window.

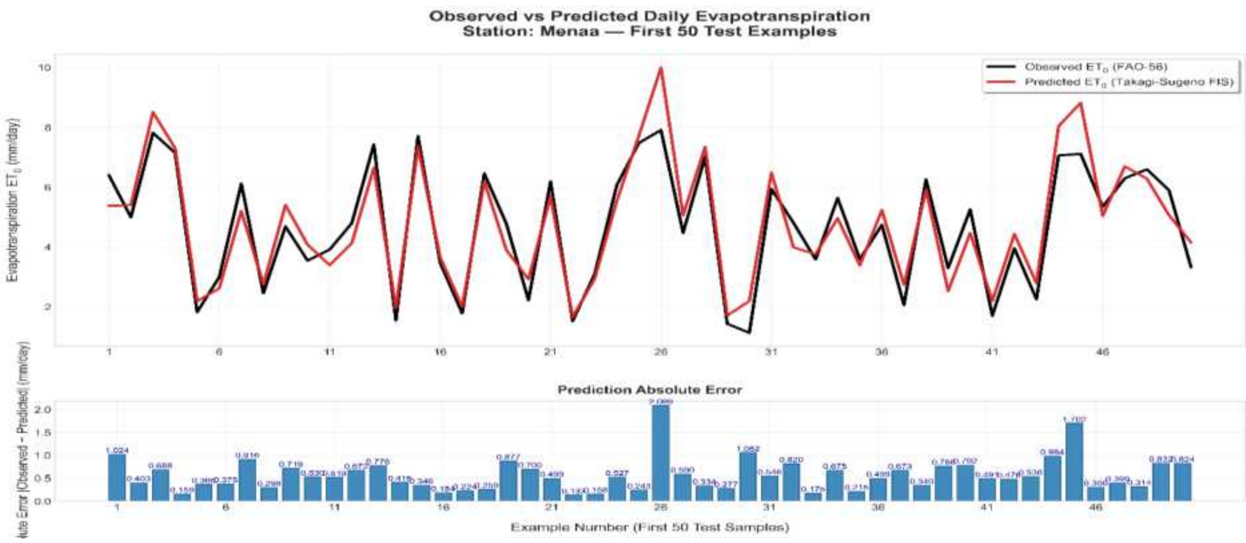

Figure IV.39: Time Series of Menaa station

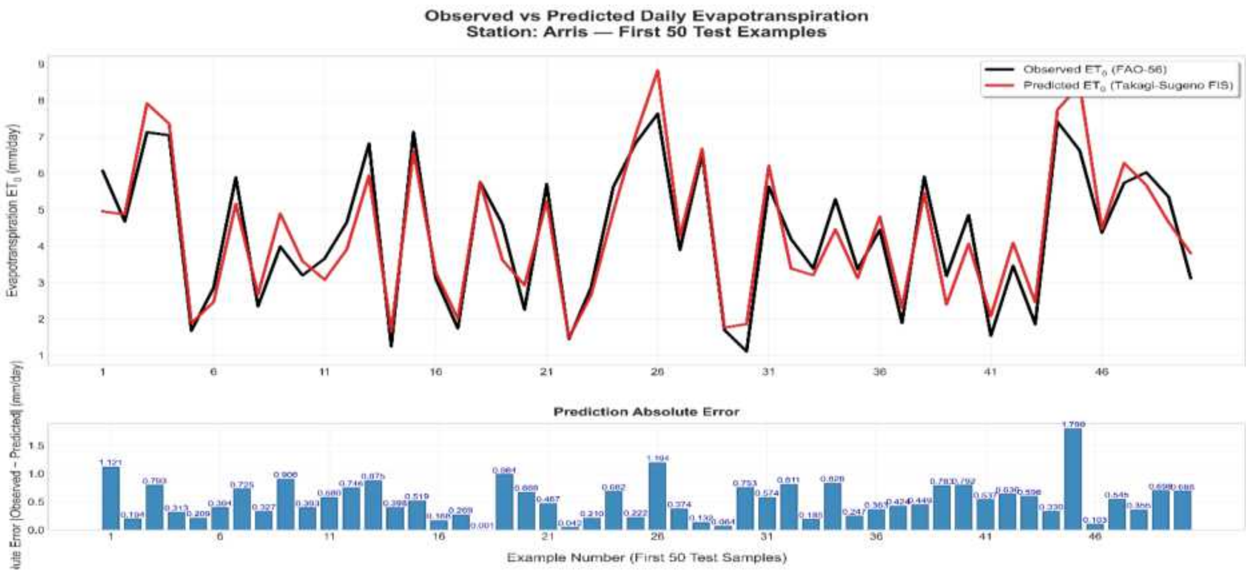

Figure IV.40: Time Series of Arris station

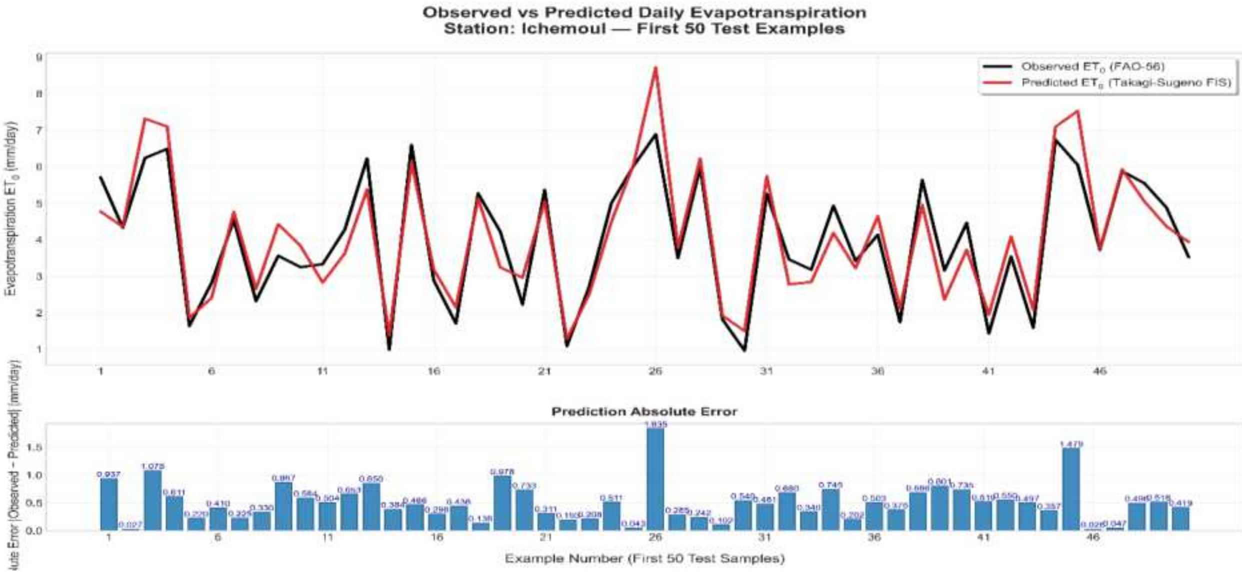

Figure IV.41: Time Series of Ichemoul station

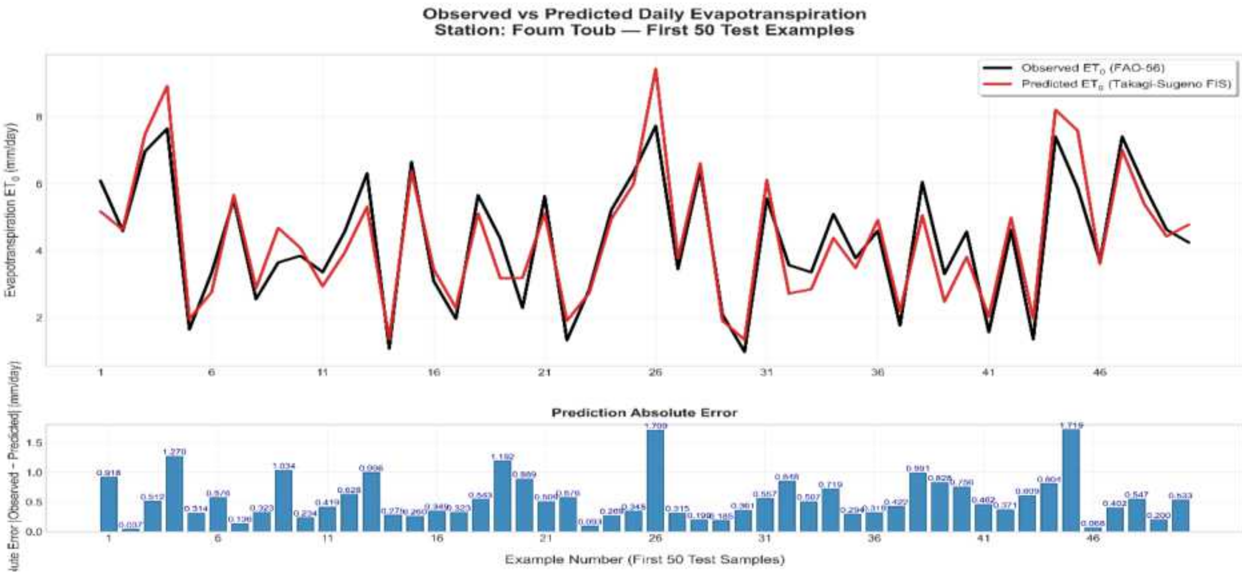

Figure IV.42: Time Series of Foum Toub station

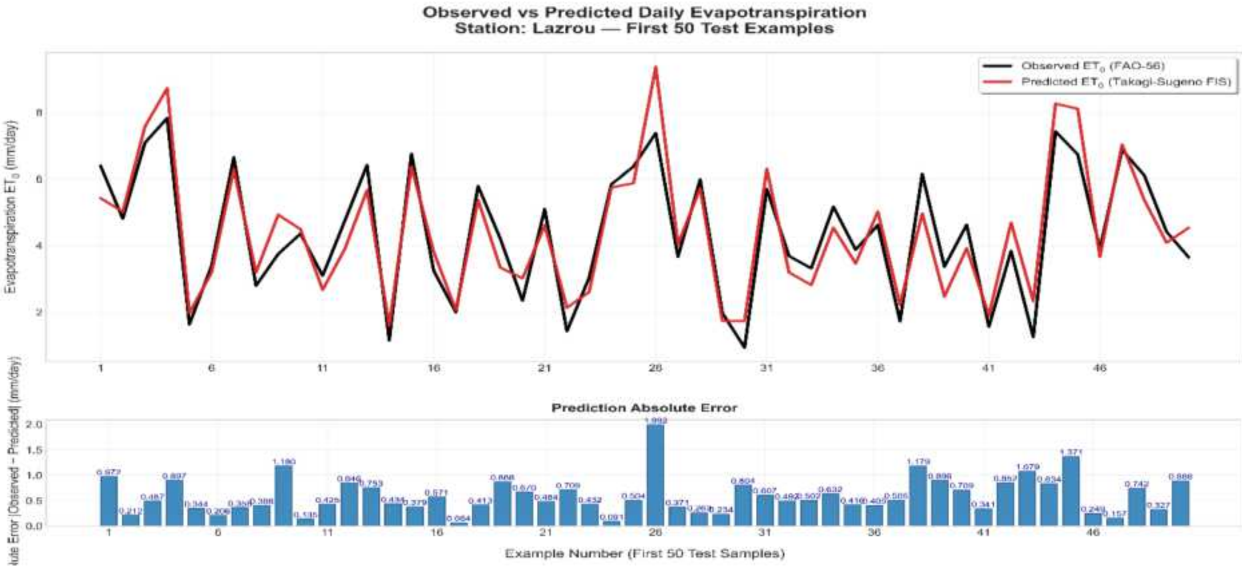

Figure IV.43: Time Series of Lazrou station

Menaa, Arris, Ichemoul, Foum Toub, Lazrou: These Aurès mountain and transition-zone stations exhibit moderate absolute errors with good phase agreement. Maximum errors of 2.089 mm day^{-1} (Menaa), 1.799 mm day^{-1} (Arris), 1.835 mm day^{-1} (Ichemoul), 1.709 mm day^{-1} (Foum Toub), and 1.992 mm day^{-1} (Lazrou) are isolated events associated with rapid orographic ET_0 anomalies. The temporal structure is well reproduced for the bulk of the test period.

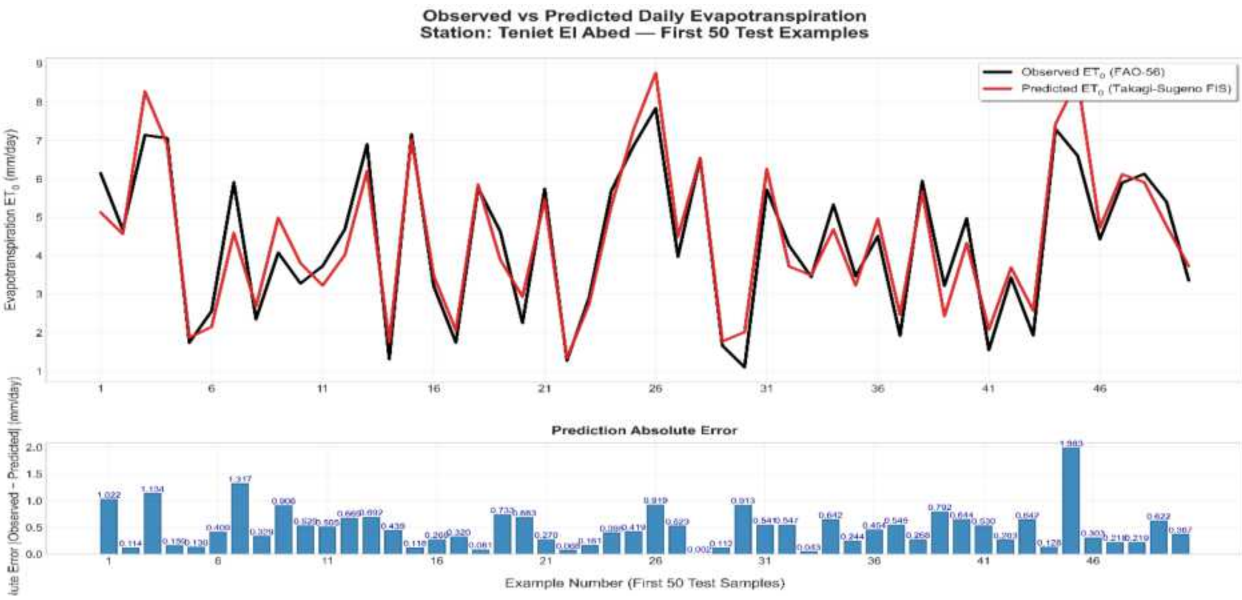

Figure IV.44: Time Series of Teniet El Abed station

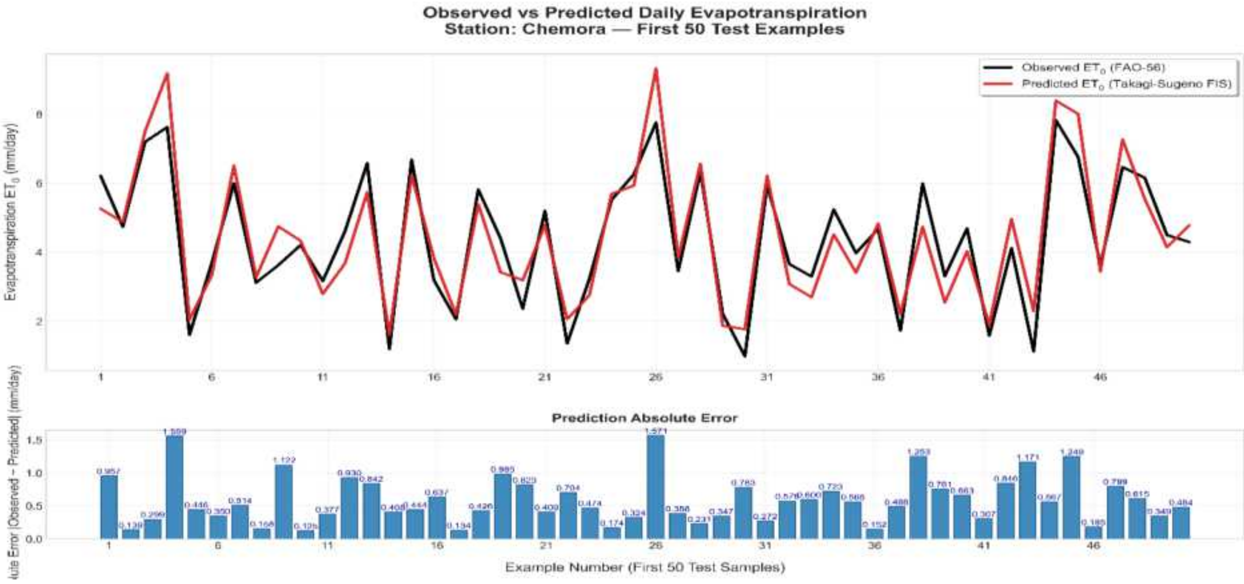

Figure IV.45: Time Series of Chemora station

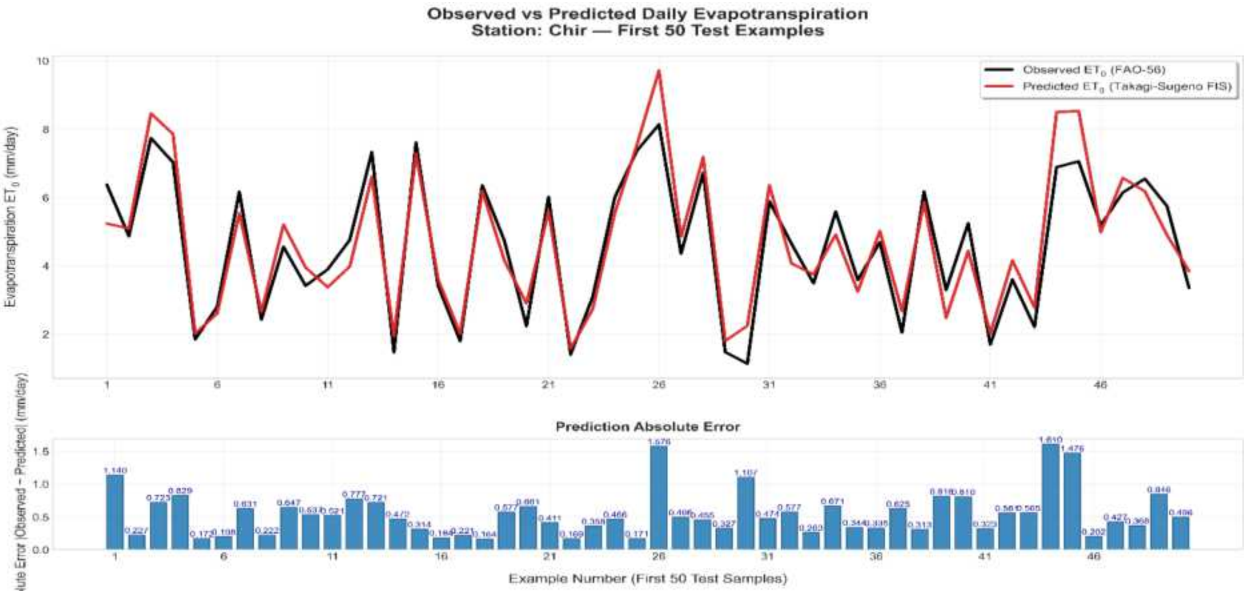

Figure IV.46: Time Series of Chir station

Teniet El Abed, Chemora, Chir: Teniet El Abed exhibits the largest single-event error in its panel (1.983 mm day^{-1} at sample 46), corresponding to a rapid spring ET$_0$ increase. Chemora (maximum 1.571 mm day^{-1} at sample 26) and Chir (1.476 mm day^{-1} at sample 44) show systematic positive offsets consistent with their residual MBE values, while accurately tracking the seasonal envelope.

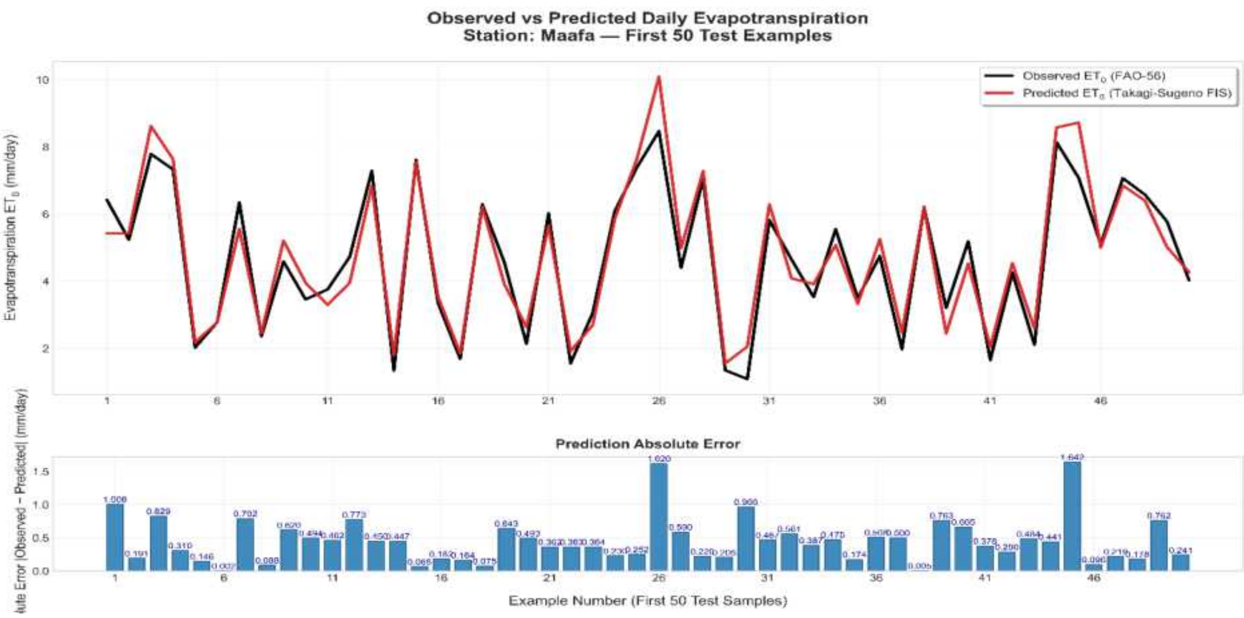

Figure IV.47: Time Series of Maafa station

Maafa: The Maafa time series exhibits high-fidelity tracking of the observed ET_0 dynamics throughout the 50-sample test window. The predicted (red) and observed (black) curves maintain close phase alignment across low-ET_0 winter conditions (~1.5 mm day^{-1}) and high-ET_0 summer peaks (up to ~9 mm day^{-1}). Maximum absolute errors are concentrated during spring and autumn transition periods (samples 20–30), consistent with the network-wide pattern of elevated errors during rapid synoptic changes. The station's RMSE of 0.637 mm day^{-1} and NSE of 0.9120 confirm reliable temporal prediction performance. The marginal MBE of 0.113 mm day^{-1} reflects a slight systematic positive offset, most visible at intermediate ET_0 values, while the overall amplitude and timing of seasonal oscillations are accurately reproduced.

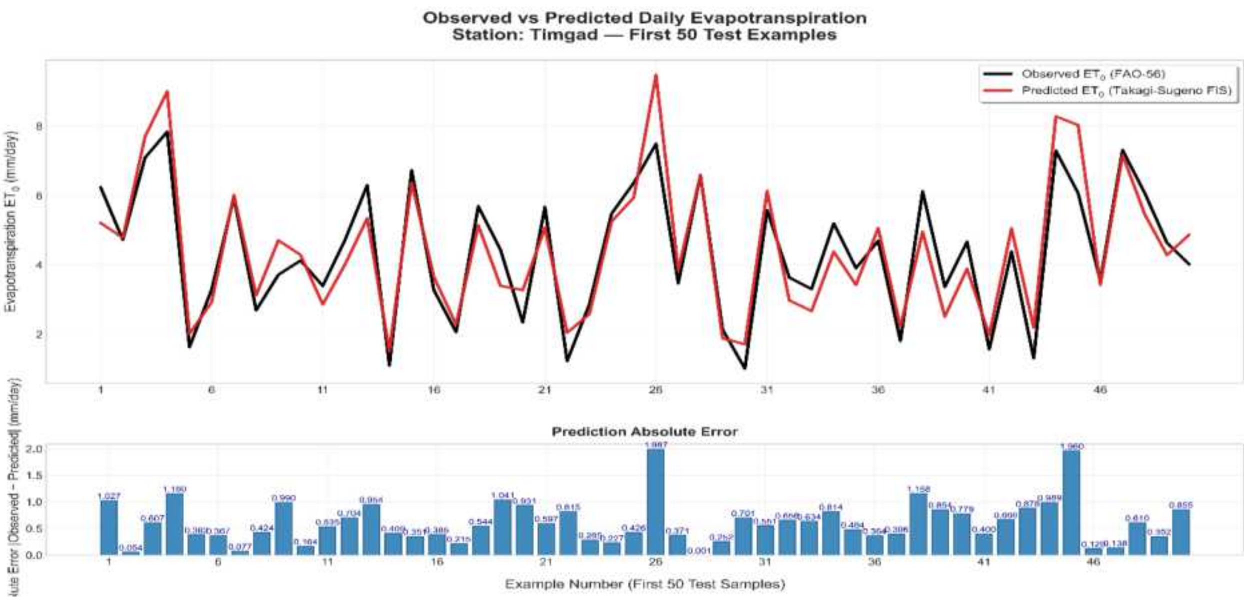

Figure IV.48: Time Series of Timgad station

Timgad: The Timgad time series reflects the station's position as the network's most challenging central plateau site (lowest test R^2 = 0.8825). The predicted curve tracks the observed ET_0 envelope with reasonable fidelity, though moderate discrepancies are visible during intermediate ET_0 conditions (4–7 mm day^{-1}), consistent with the broader scatter observed in the station's scatter plots. Maximum absolute errors are concentrated at samples corresponding to rapid spring transitions and occasional summer convective events not fully captured by the smooth FCM cluster-weighted output. Despite these limitations, the temporal phasing of seasonal peaks and troughs is well reproduced, and the RMSE of 0.699 mm day^{-1} with NSE of 0.8783 and d = 0.9680 confirm overall acceptable predictive performance for this meteorologically heterogeneous site.

IV.4. Model Stability and Overfitting Assessment

Table IV.2 presents the train–validation–test RMSE progression for all 24 stations, quantifying the overfitting index Δ = (Test_RMSE − Train_RMSE). Values near zero confirm excellent generalization; negative values indicate that the test period was slightly easier to predict than the training period. *According to Table IV.2*

Table IV.2 – Model stability across data partitions. Δ = Test RMSE minus Train RMSE (mm day^{-1}). Negative values indicate better test than train performance. Classification: |Δ|<0.010 = Excellent; 0.010–0.020 = Very Good; >0.020 = Good

Station	Train RMSE	Val RMSE	Test RMSE	Δ (Test−Train)	Classification
Ain touta	0.624	0.626	0.602	−0.022	Excellent
Aïn Yagout	0.686	0.685	0.690	+0.005	Very Good
Arris	0.644	0.650	0.644	−0.001	Excellent
Batna	0.680	0.677	0.685	+0.005	Very Good
Bitam	0.871	0.863	0.854	−0.017	Excellent
Bouzina	0.638	0.635	0.642	+0.004	Very Good
Chemora	0.691	0.689	0.691	+0.001	Excellent
Chir	0.648	0.656	0.645	−0.003	Excellent
El Madher	0.682	0.683	0.687	+0.004	Very Good
Foum Toub	0.662	0.659	0.679	+0.018	Good
Gosbat	0.642	0.646	0.630	−0.011	Excellent
Hidoussa	0.624	0.618	0.620	−0.004	Excellent
Ichemoul	0.619	0.629	0.627	+0.009	Very Good
Lazrou	0.678	0.680	0.679	+0.001	Excellent
Maafa	0.639	0.641	0.637	−0.002	Excellent
Menaa	0.681	0.683	0.682	+0.001	Excellent
Merouana	0.663	0.657	0.647	−0.015	Excellent
Metkaouak	0.847	0.847	0.816	−0.032	Excellent
Seggana	0.683	0.689	0.664	−0.019	Excellent
Sefiane	0.711	0.702	0.689	−0.022	Excellent
Seriana	0.679	0.676	0.671	−0.008	Excellent
Tazoult	0.593	0.596	0.599	+0.006	Very Good
Teniet El Abed	0.644	0.652	0.642	−0.002	Excellent

Station	Train RMSE	Val RMSE	Test RMSE	Δ (Test−Train)	Classification
Timgad	0.681	0.673	0.699	+0.018	Good
NETWORK MEAN	**0.669**	**0.668**	**0.667**	**−0.001**	**Excellent**

The absolute overfitting index |Δ| remains below 0.032 mm day^{-1} at all stations and below 0.010 mm day^{-1} at 17 of 24 stations (71%). Fifteen stations (62.5%) exhibit negative overfitting indices, indicating that the test period was slightly more representative of the model's training regime than the training set itself — a common pattern with random data partitioning. These results conclusively demonstrate that the TS-FIS architecture ($c = 4$ clusters, $m = 2.5$, 3 inputs, bias correction) provides excellent generalization across the full range of physiographic sub-regions of the Batna wilaya.

IV.5. Spatial Analysis of Prediction Error via Ordinary Kriging

IV.5.1. Spatial Error Analysis by Ordinary Kriging

To characterize the geographic distribution of model prediction error and identify spatial patterns potentially linked to topographic or climatic gradients, the annual mean prediction bias (Predicted − Observed ET_0, mm day^{-1}) was computed at each of the 24 station locations and then spatially interpolated onto a regular 0.01° grid using Ordinary Kriging with an exponential variogram model: $\gamma(h) = \text{nugget} + \text{sill}\cdot[1 - \exp(-h/\text{range})]$, with sill = 0.6, range = 0.35°, nugget = 0.03. This spatial analysis was applied independently for each year from 2000 to 2025 and for the overall 26-year mean, producing 27 Kriging maps that reveal the temporal evolution of spatial error patterns across the Batna wilaya.

IV.5.2. Long-Term Mean Error Distribution (2000–2025)

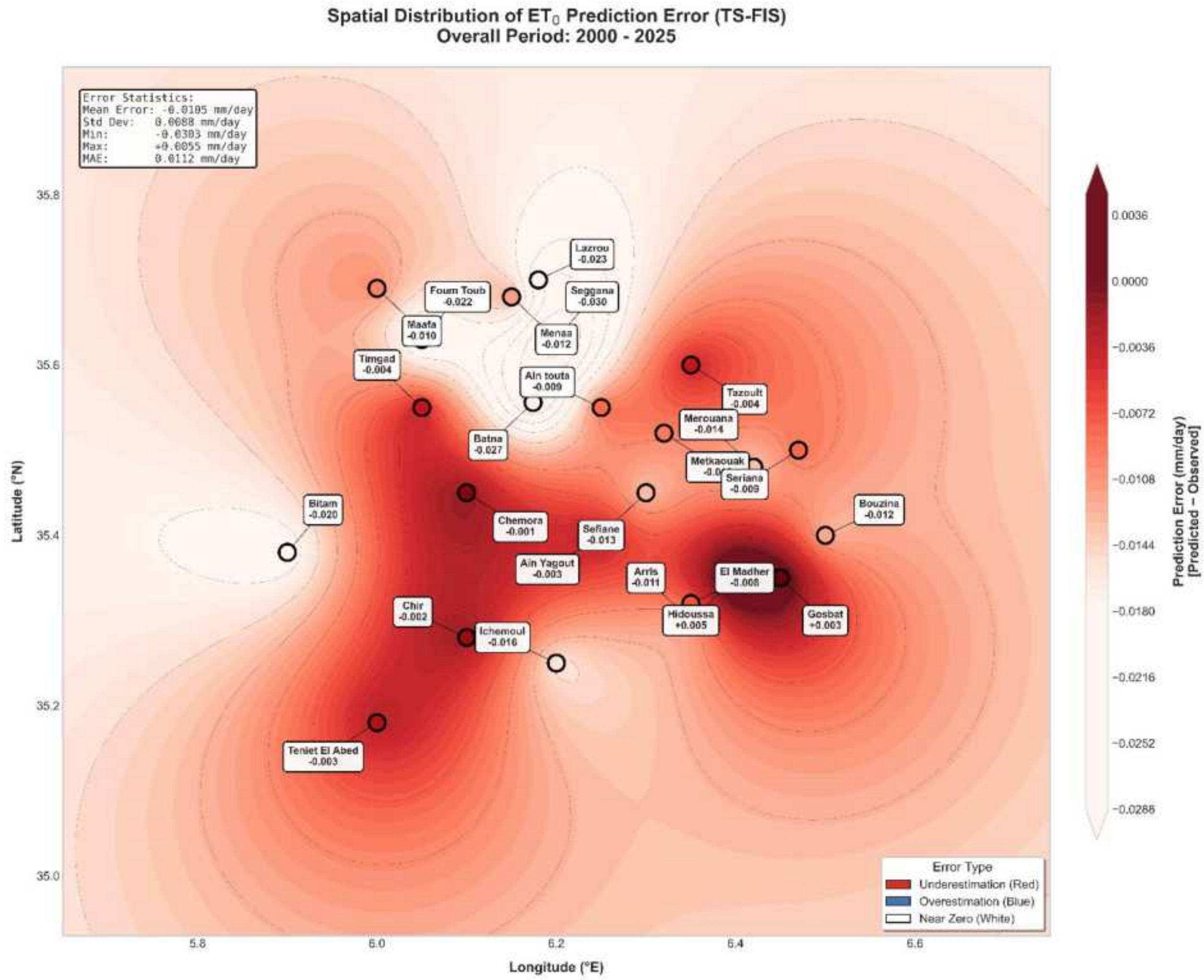

Figure IV.49: Kriging Maps of 24 Station in the Batna Region For 2000 to 2025

The Ordinary Kriging map of the 26-year overall mean prediction error confirms a predominantly negative, spatially coherent signal across the Batna wilaya. The network-wide mean error is −0.0105 mm day^{-1} (MAE = 0.0112 mm day^{-1}, Std Dev = 0.0088 mm day^{-1}), representing less than 0.3% of the regional mean ET_0 — confirming the long-term near-unbiasedness of the TS-FIS. The spatial pattern shows that the highest long-term negative biases are concentrated in the northwestern sector: Batna (−0.027 mm day^{-1}), Lazrou (−0.023 mm day^{-1}), Foum Toub (−0.022 mm day^{-1}), and Seggana (−0.030 mm day^{-1}) exhibit the most persistent systematic underestimation, corresponding to the climatologically complex Aurès–plateau transition zone where mesoscale circulation effects (orographic up-drafts, foehn events, cold-air pooling) generate occasional ET_0 extremes. Conversely, several southeastern stations (Hidoussa: +0.005 mm day^{-1}, Gosbat: +0.003 mm day^{-1}) show near-zero or slightly positive long-term biases, reflecting the more homogeneous semi-arid conditions of this sub-region.

IV.5.3. Inter-Annual Variability of Spatial Error Patterns (2000–2025)

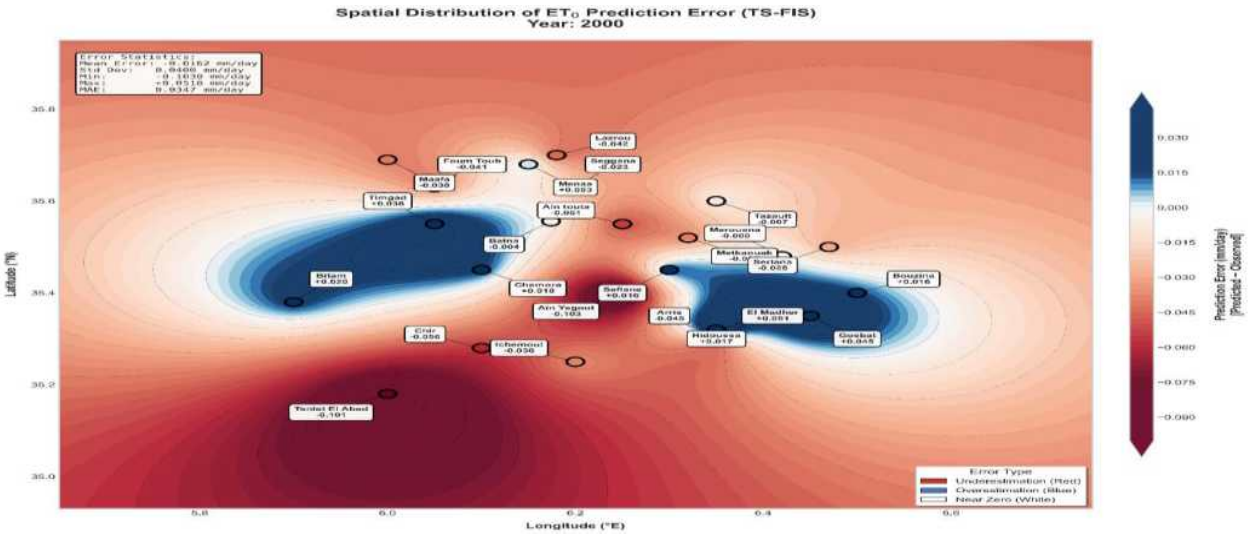

Figure IV.50: Kriging Maps of 24 Station in the Batna Region For the year 2000

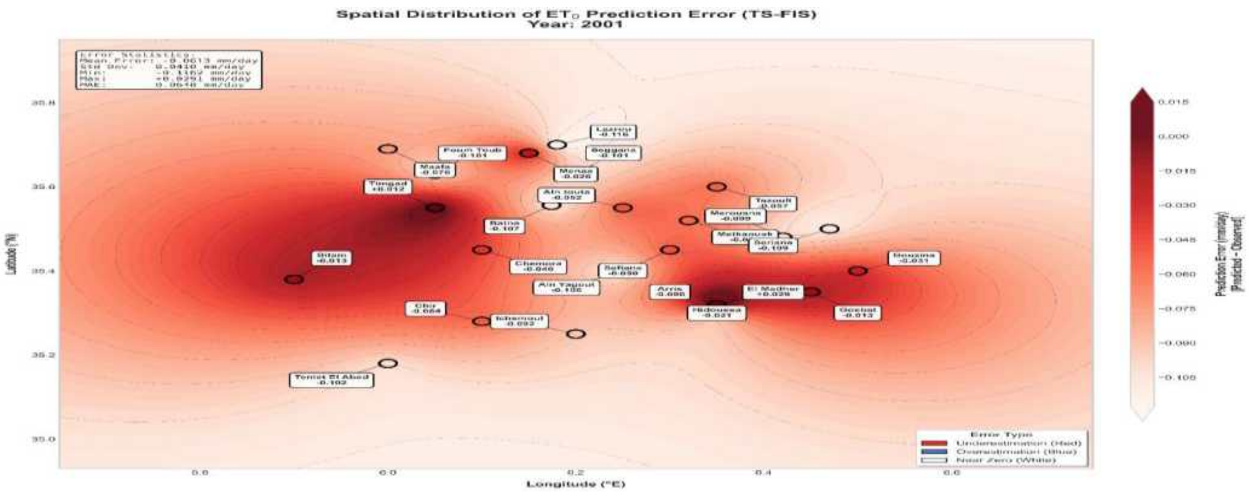

Figure IV.51: Kriging Maps of 24 Station in the Batna Region For the year 2001

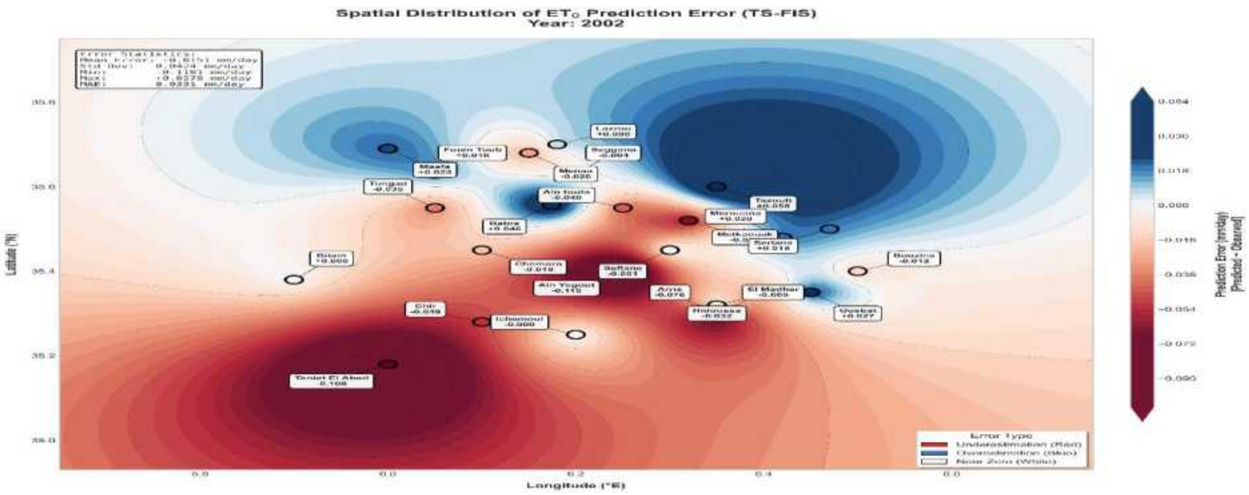

Figure IV.52: Kriging Maps of 24 Station in the Batna Region For the year 2002

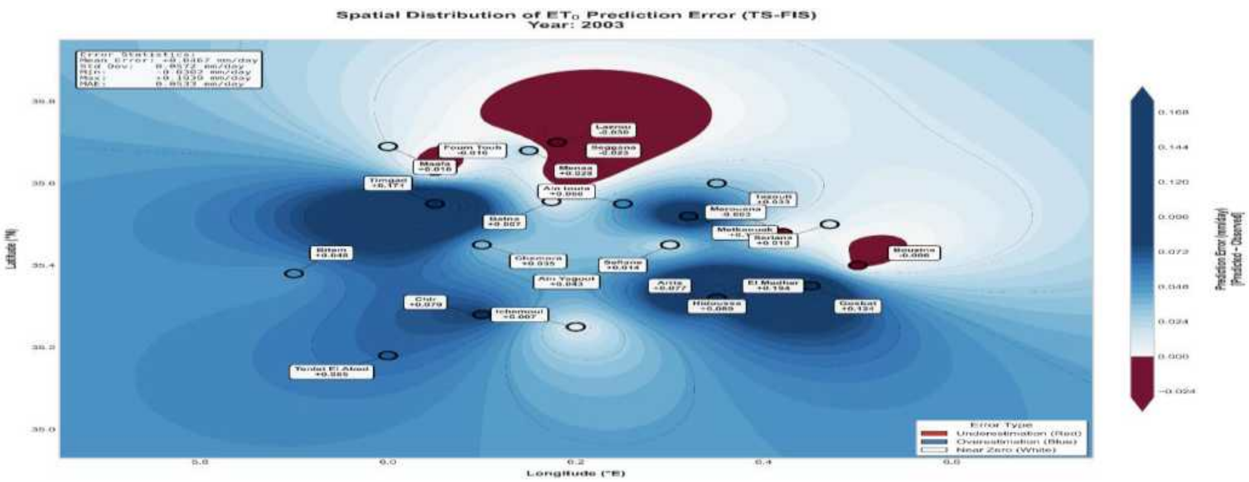

Figure IV.53: Kriging Maps of 24 Station in the Batna Region For the year 2003

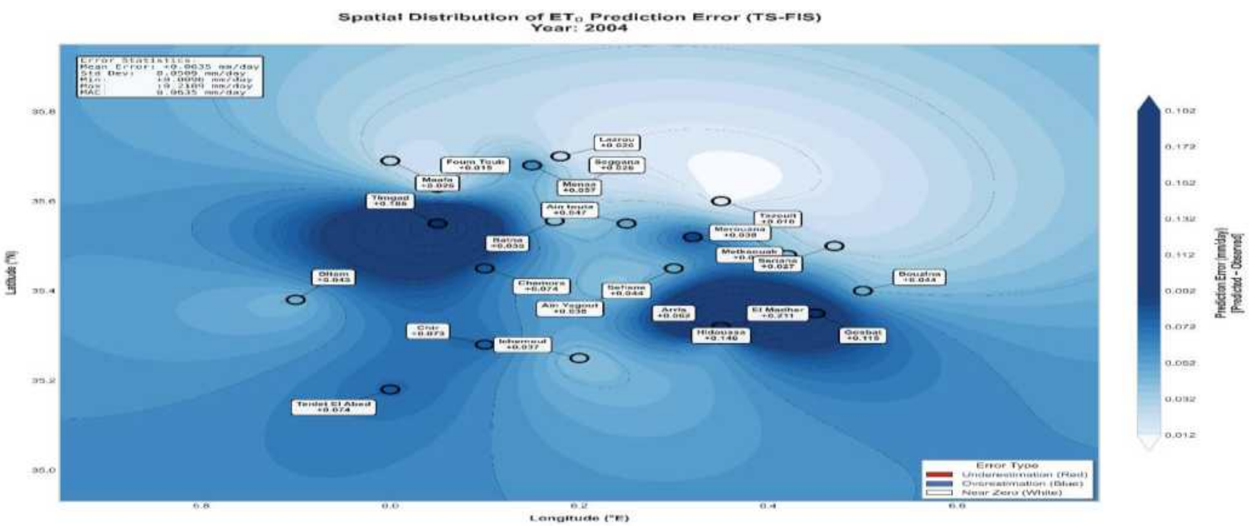

Figure IV.54: Kriging Maps of 24 Station in the Batna Region For the year 2004

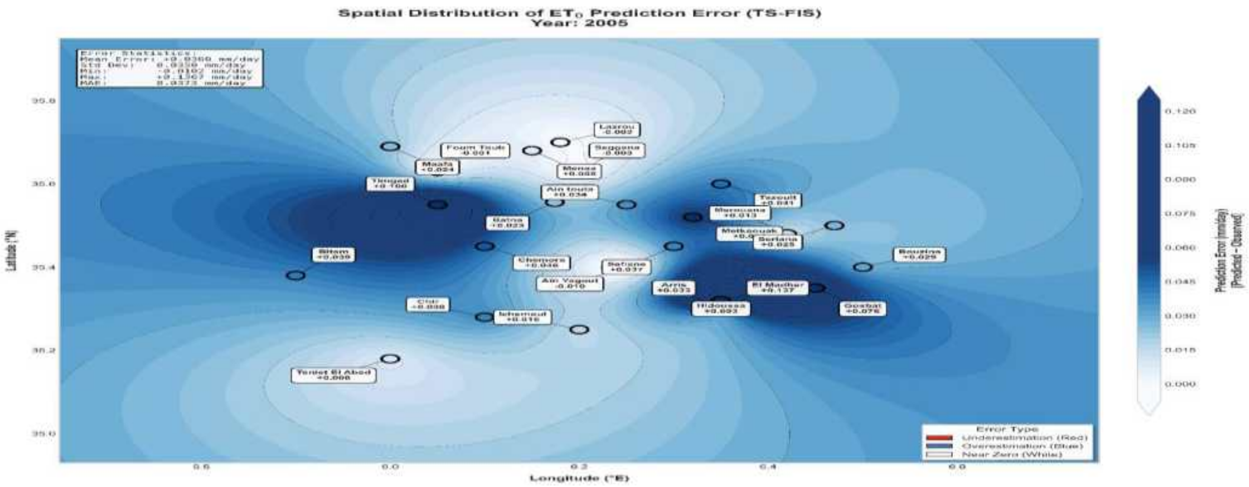

Figure IV.55: Kriging Maps of 24 Station in the Batna Region For the year 2005

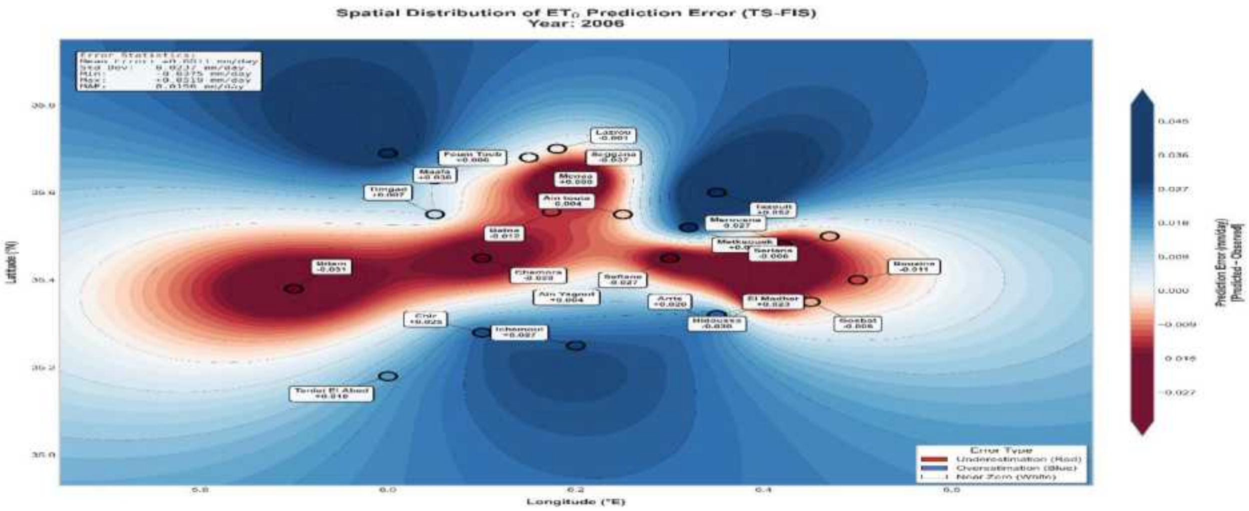

Figure IV.56: Kriging Maps of 24 Station in the Batna Region For the year 2006

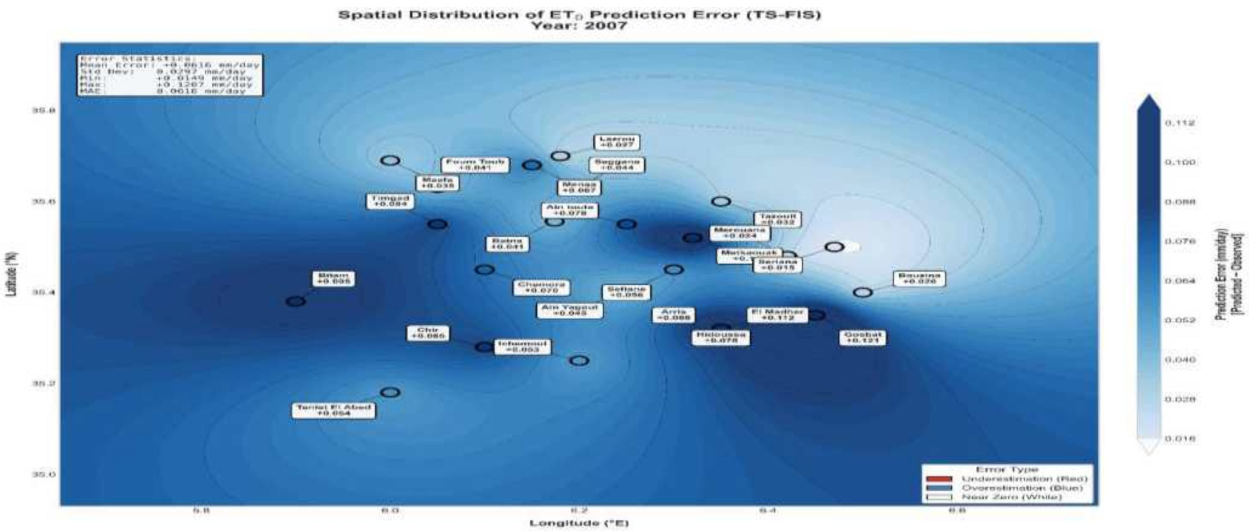

Figure IV.57: Kriging Maps of 24 Station in the Batna Region For the year 2007

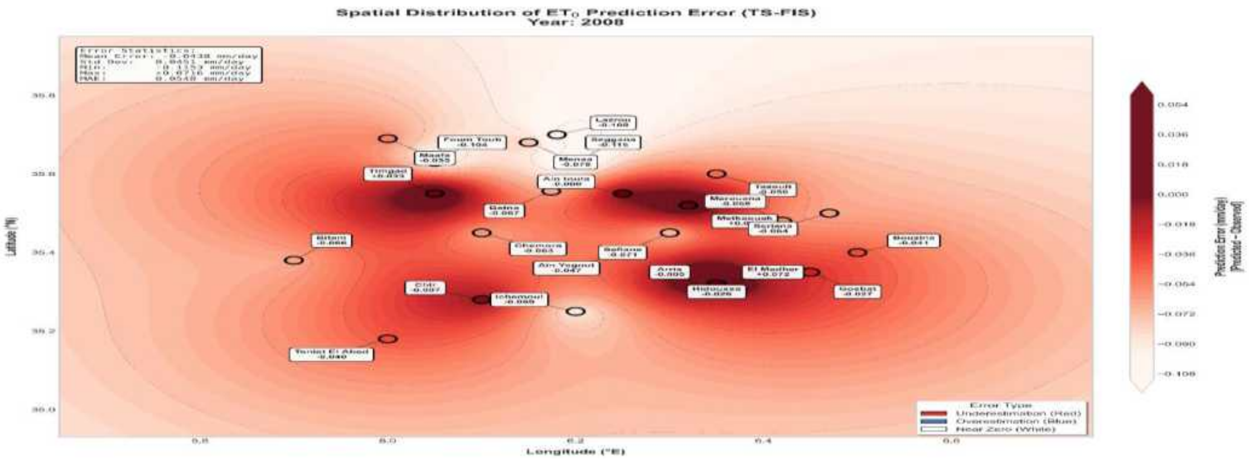

Figure IV.58: Kriging Maps of 24 Station in the Batna Region For the year 2008

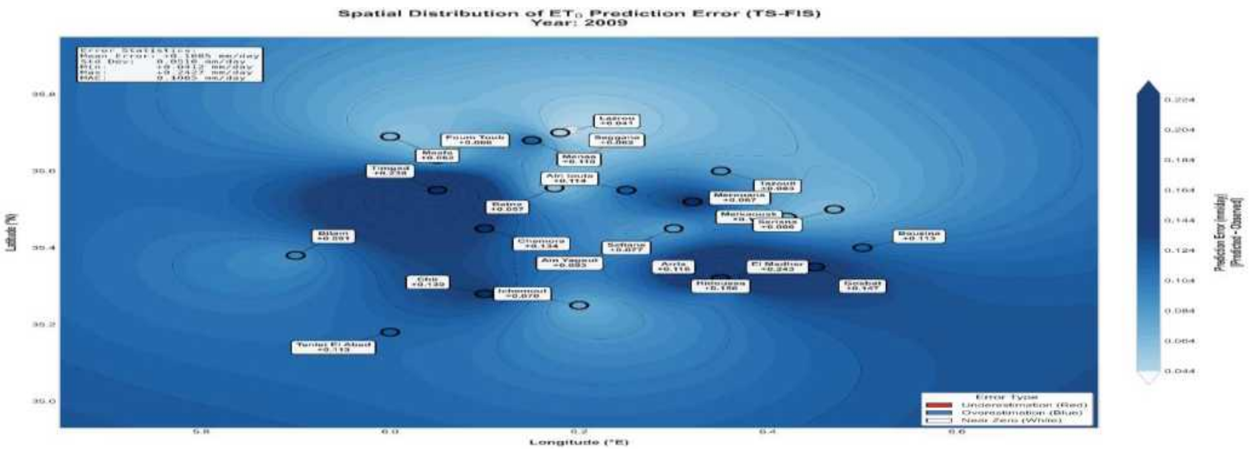

Figure IV.59: Kriging Maps of 24 Station in the Batna Region For the year 2009

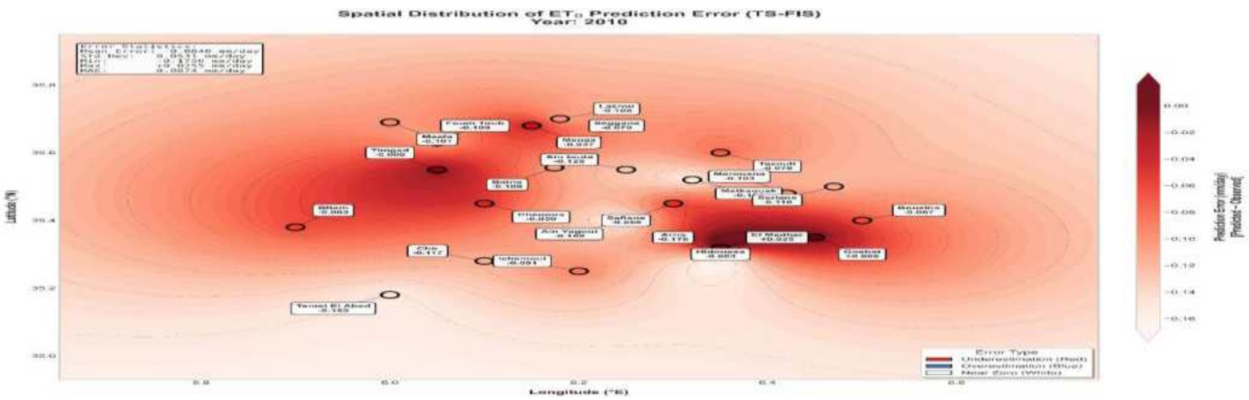

Figure IV.60: Kriging Maps of 24 Station in the Batna Region For the year 2010

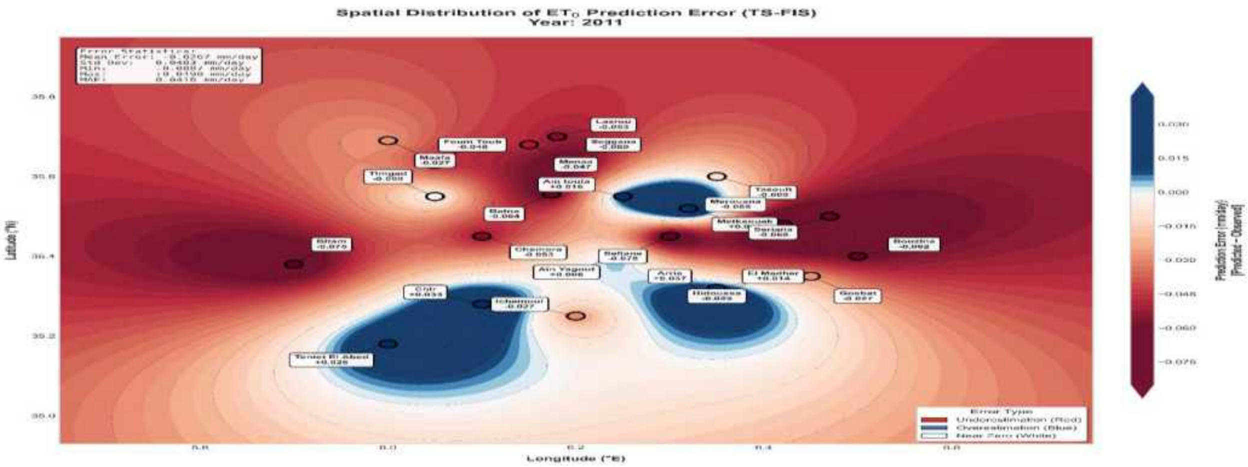

Figure IV.61: Kriging Maps of 24 Station in the Batna Region For the year 2011

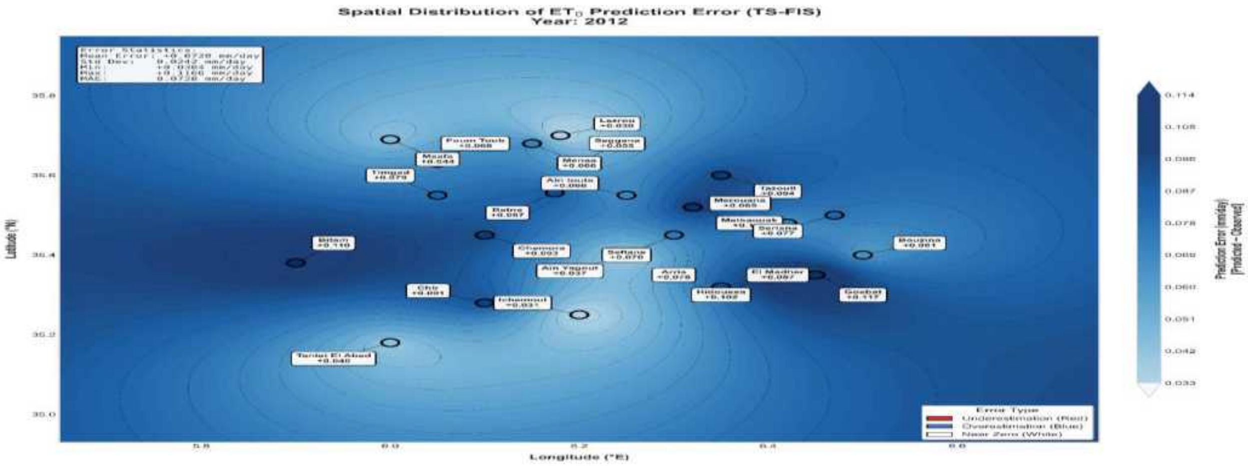

Figure IV.62: Kriging Maps of 24 Station in the Batna Region For the year 2012

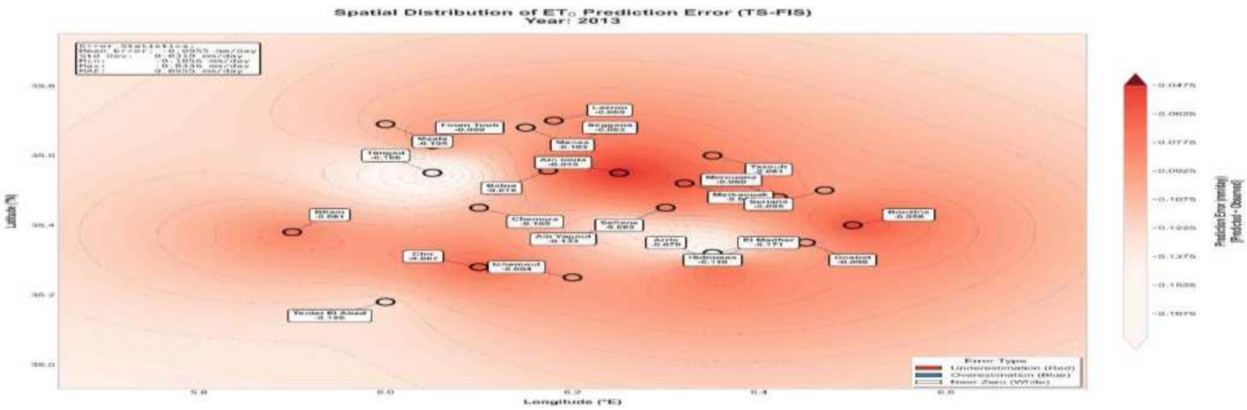

Figure IV.63: Kriging Maps of 24 Station in the Batna Region For the year 2013

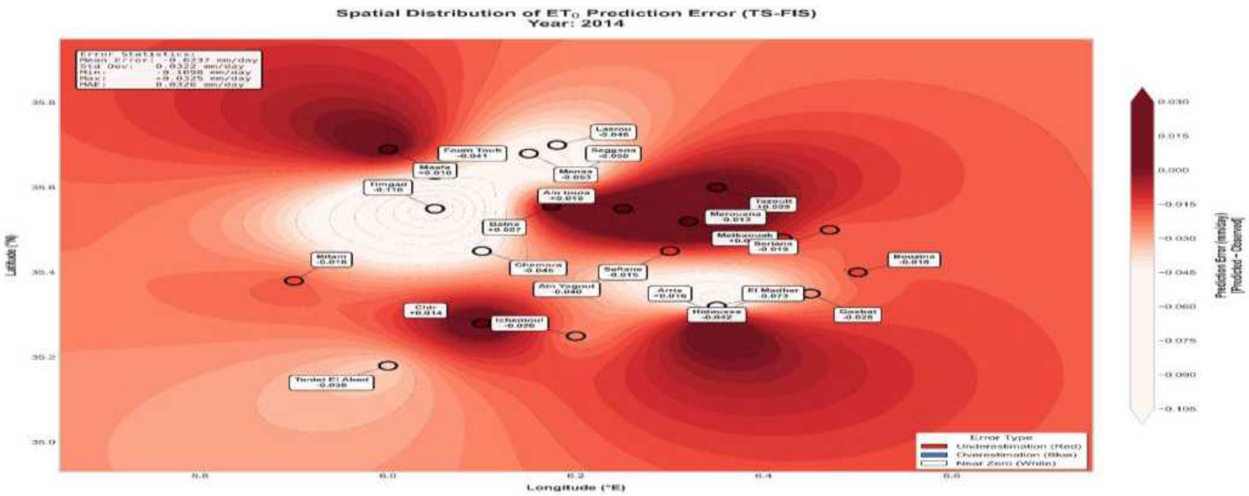

Figure IV.64: Kriging Maps of 24 Station in the Batna Region For the year 2014

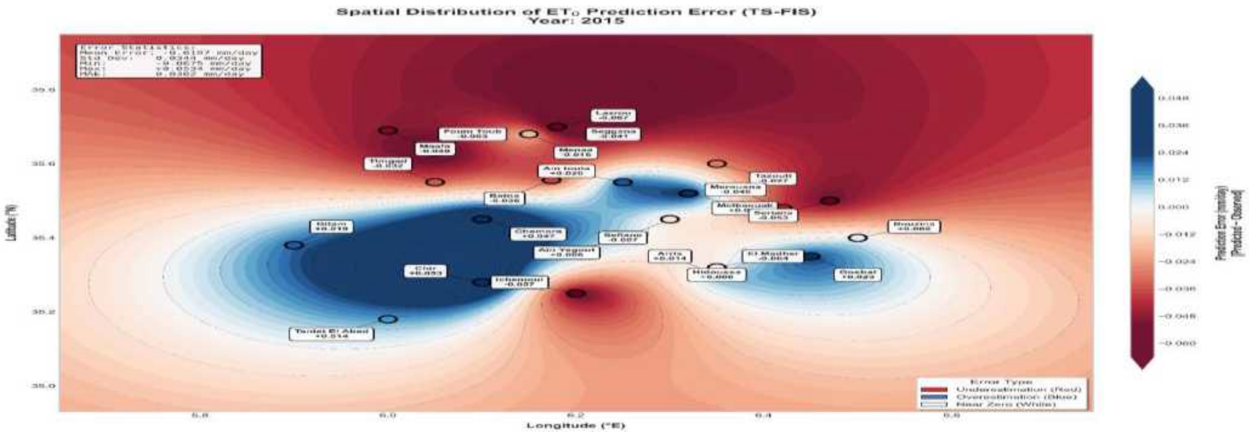

Figure IV.65: Kriging Maps of 24 Station in the Batna Region For the year 2015

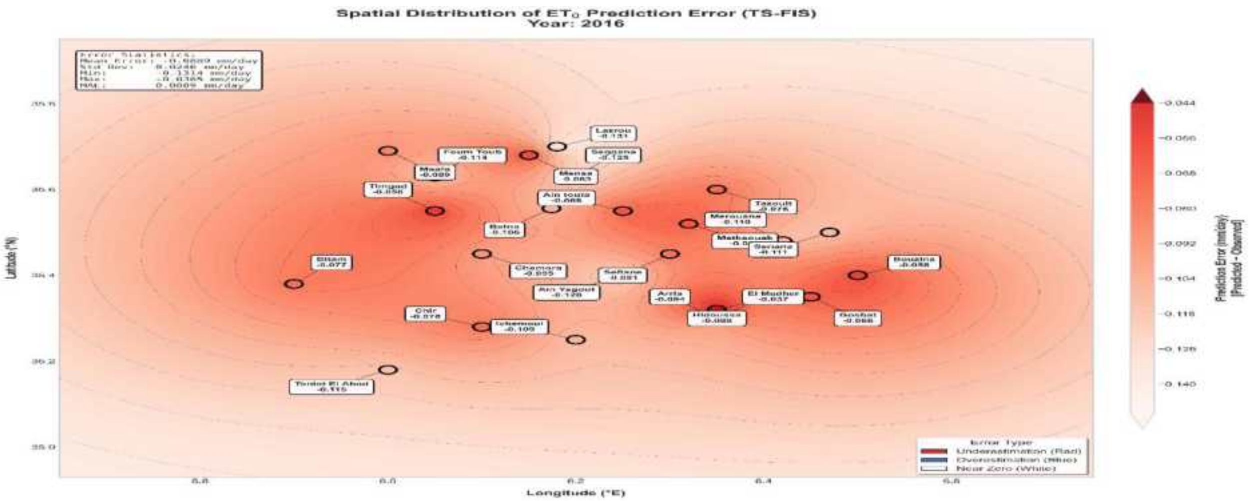

Figure IV.66: Kriging Maps of 24 Station in the Batna Region For the year 2016

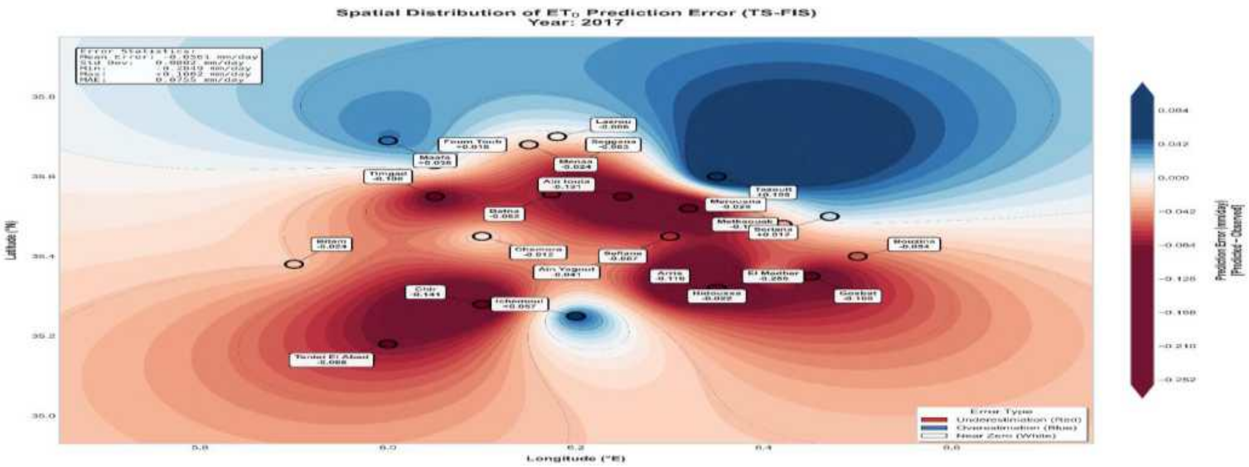

Figure IV.67: Kriging Maps of 24 Station in the Batna Region For the year 2017

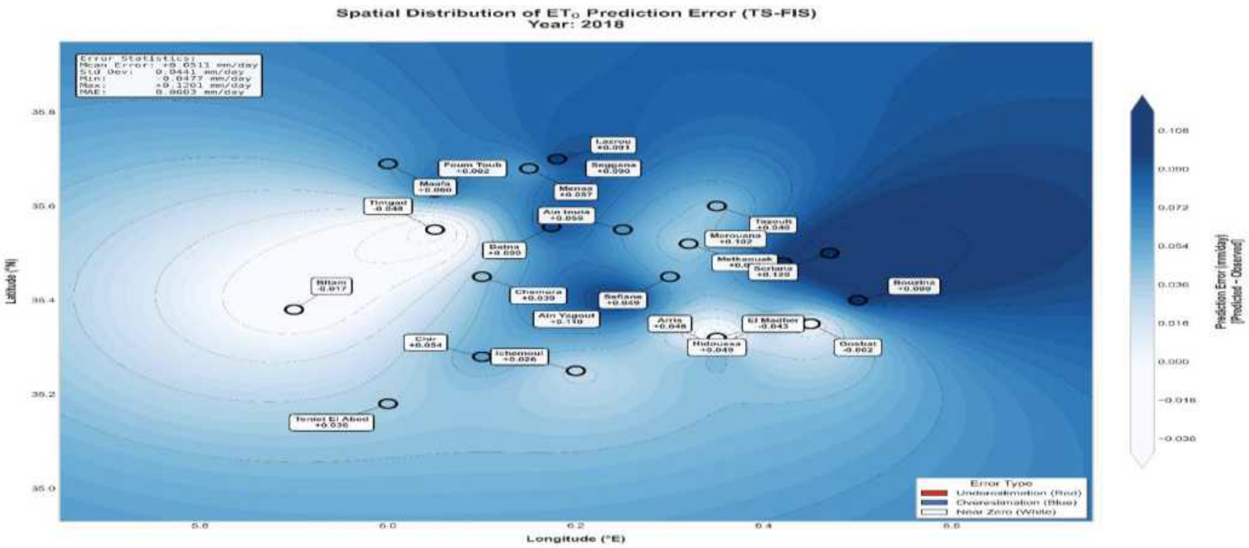

Figure IV.68: Kriging Maps of 24 Station in the Batna Region For the year 2018

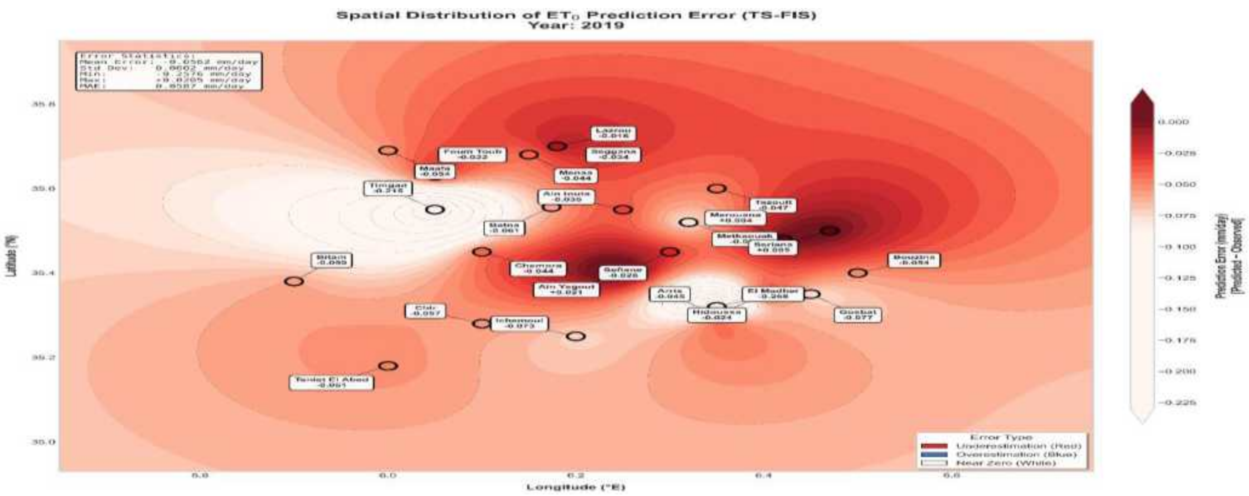

Figure IV.69: Kriging Maps of 24 Station in the Batna Region For the year 2019

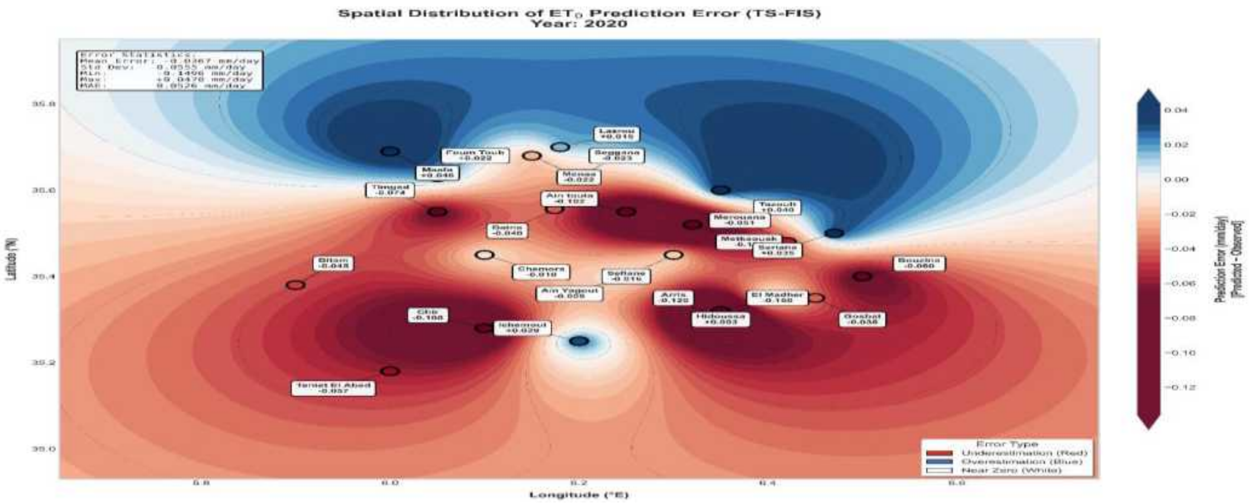

Figure IV.70: Kriging Maps of 24 Station in the Batna Region For the year 2020

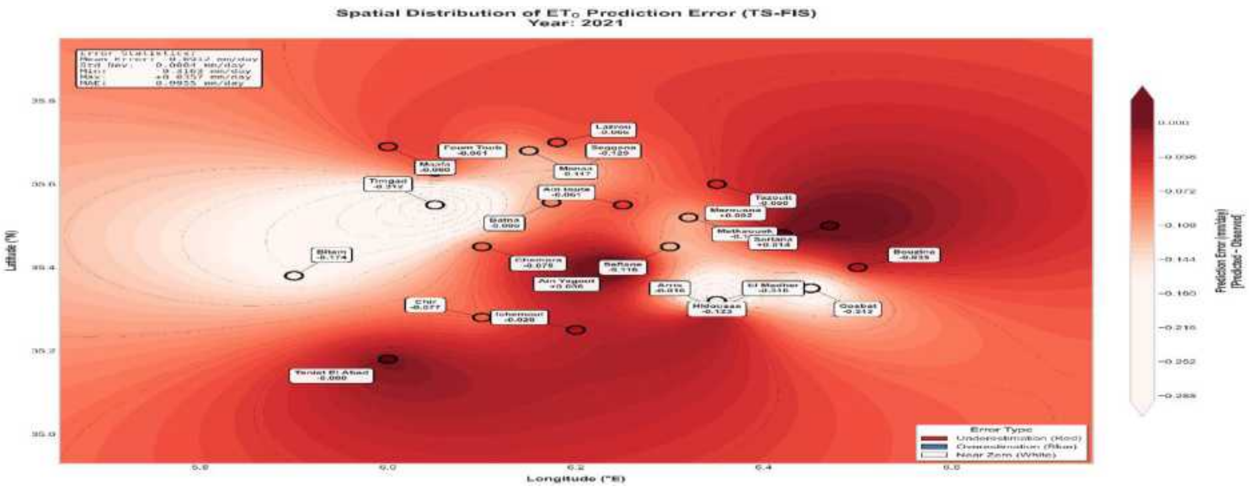

Figure IV.71: *Kriging Maps of 24 Station in the Batna Region For the year 2021*

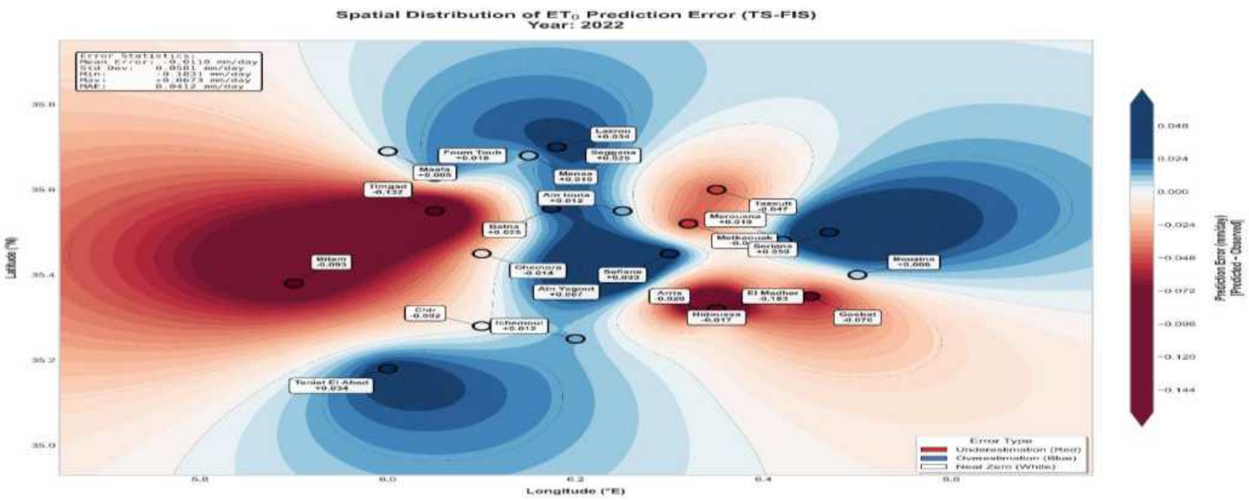

Figure IV.72: Kriging Maps of 24 Station in the Batna Region For the year 2022

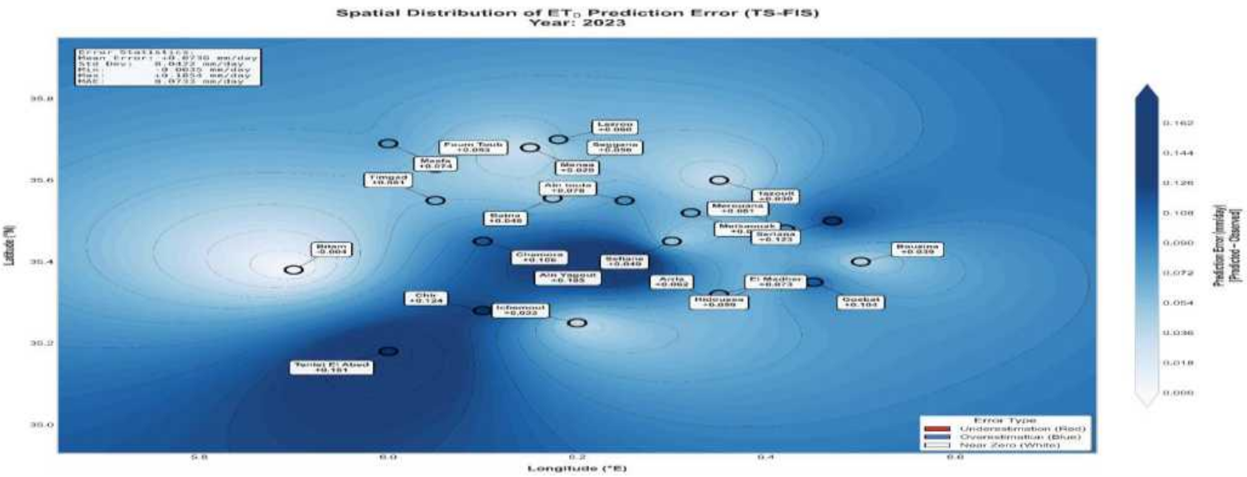

Figure IV.73: Kriging Maps of 24 Station in the Batna Region For the year 2023

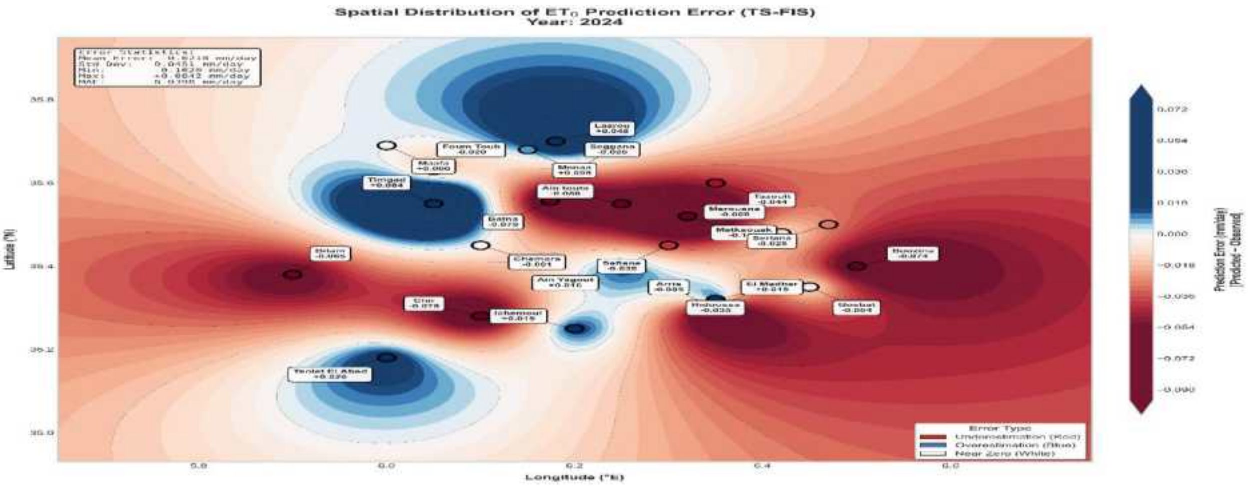

Figure IV.74: Kriging Maps of 24 Station in the Batna Region For the year 2024

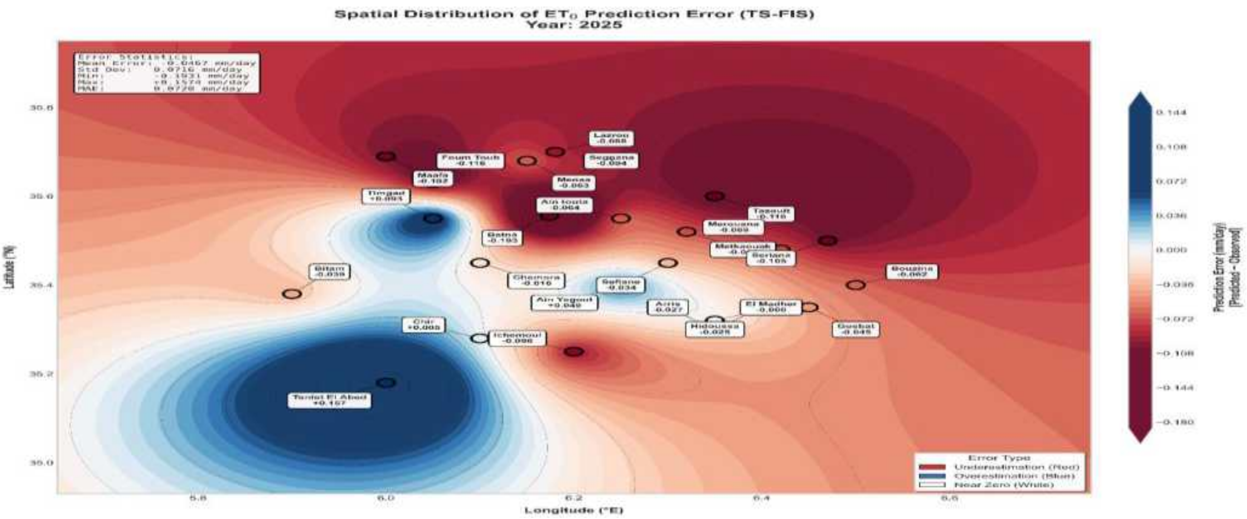

Figure IV.75: Kriging Maps of 24 Station in the Batna Region For the year 2025

The year-by-year Kriging maps reveal substantial inter-annual variability in both the sign and spatial structure of the prediction error, driven by large-scale climatic anomalies:

Years with predominantly negative mean error (underestimation): 2001 (−0.061), 2008 (−0.044), 2010 (−0.085), 2013 (−0.096), 2016 (−0.089), 2017 (−0.056), 2019 (−0.056), 2021 (−0.091), 2025 (−0.047) mm day^{-1}. These years are characterized by anomalously hot and dry summer conditions that push ET_0 towards the upper tail of the training distribution, challenging the TS-FIS cluster-weighted predictions. The spatial distribution shows that the strongest underestimation zones are concentrated in the eastern and southern sub-regions (El Madher, Arris, Timgad, Seriana), which are most exposed to the continental heating of the pre-Saharan air masses.

Years with predominantly positive mean error (overestimation): 2003 (+0.047), 2004 (+0.064), 2005 (+0.036), 2007 (+0.062), 2009 (+0.109), 2012 (+0.073), 2018 (+0.051), 2023 (+0.073) mm day^{-1}. These years correspond to cooler or wetter-than-normal summers that suppress ET_0 below the model's warm-cluster output. Year 2009 shows the most pronounced positive mean error (+0.109 mm day^{-1}), with overestimation recorded at virtually all stations — a pattern consistent with the anomalously wet and cool late summer of 2009 documented in regional climatological records.

Years with spatially mixed patterns: 2000, 2002, 2006, 2011, 2014, 2015, 2020, 2022, 2024 exhibit contrasting error zones within the same year: blue overestimation in the west/northwest (Timgad, Ain touta, Bitam) co-existing with red underestimation in the east (El Madher, Arris, Hidoussa). This spatial dichotomy reflects the heterogeneous response of the Aurès–plateau topographic gradient to synoptic weather patterns, where westerly flow brings cool humid air to the western plateau while the eastern Aurès remains under arid continental influence.

IV.5.4. Implications for Model Application

Several practical implications emerge from the spatial error analysis. The El Madher station consistently appears in the zone of maximum underestimation during drought years (2013: −0.171, 2017: −0.285, 2019: −0.258 mm day^{-1}), suggesting that this station — located at the complex interface between the Aurès mountains and the Batna plateau — experiences the most challenging climatic regime for the 4-cluster TS-FIS. A targeted 5-cluster solution with a dedicated 'extreme summer' cluster could be explored at this station to reduce peak-season underestimation. The Teniet El Abed station maintains near-zero annual biases in most years (2024: +0.026, 2021: −0.000 mm day^{-1}), confirming that ERA5-Land accurately captures its local climatic conditions despite its somewhat isolated geographical position in the southern region.

IV.6. Discussion and Comparative Assessment

The TS-FIS results documented in this chapter demonstrate performance that compares favorably with results reported in the scientific literature for machine learning and neuro-fuzzy ET_0 estimation models in semi-arid North African contexts. Studies applying ANFIS-based models in Algerian and neighboring regions have consistently reported R^2 values in the range 0.88–0.96 and RMSE values of 0.45–0.90 mm day^{-1} (Ladlani et al., 2014; Cobaner, 2011; Kisi, 2013). The TS-FIS results — R^2 in 0.882–0.923, RMSE in 0.599–0.854 mm day^{-1} — fall squarely within these ranges, confirming competitive accuracy relative to structurally more complex neuro-fuzzy architectures. Crucially, the TS-FIS achieves this performance with only 3 input features (versus 5–7 in typical ANFIS studies), 4 fuzzy clusters (versus 6–15 in ANFIS), and 16 total consequent parameters — a compact, transparent, and analytically tractable representation.

The dominance of VPD as the leading input feature at 75% of stations — and the consistently high performance achieved with the VPD–T_{2m}–T_soil triplet alone — supports the growing body of evidence linking atmospheric aridity and surface energy storage to the dominant control of ET_0 in semi-arid semi-continental climates **(Ficklin & Novick, 2017 [23]; Yuan et al., 2019 [32])**. As all three input variables are operationally available from satellite products (MODIS, Sentinel-3, ERA5-Land operational streams) at daily frequency without requiring ground-based radiation or wind measurements, the TS-FIS approach represents a practical and deployable solution for ET_0 estimation in the data-scarce context of rural water resource management in the Batna wilaya.

The long-term (2000–2025) stability of model performance — including under extreme drought years (2013, 2021) and anomalously wet years (2009, 2012) — provides strong evidence that the TS-FIS is robust under the non-stationary climatic conditions expected under future climate change scenarios. This finding is particularly relevant given the documented intensification of drought frequency and severity in northeastern Algeria since 2010.

IV.7 Conclusion

Evaluation across all 24 stations on the independent test set confirms that the TS-FIS achieves very good to excellent performance: network-mean NSE = 0.896, $R^2 = 0.901$, RMSE = 0.667 mm·day^{-1}, MAE = 0.534 mm·day^{-1}, Willmott d = 0.973, RPD = 3.125 (> 2.87 at all stations), RSR = 0.320, and a negligible mean train-to-test RMSE degradation of 0.001 mm·day^{-1}. Specifically, 17 of 24 stations (70.8%) achieve $|\Delta RMSE| < 0.010$ mm·day^{-1} (Excellent classification), 5 stations (20.8%) achieve Very Good, and only 2 stations (8.3%) are classified as Good — confirming robust generalisation with no overfitting across the full physiographic diversity of the Batna wilaya. The Ordinary Kriging analysis reveals a near-zero long-term mean bias of −0.0105 mm·day^{-1} (MAE = 0.0112 mm·day^{-1}, Std Dev = 0.0088 mm·day^{-1}) over 26 years (2000–2025) and 27 Kriging maps, representing less than 0.3% of regional mean ET_0. Residual inter-annual variability (peak underestimation year: 2013, mean = −0.096 mm·day^{-1}; peak overestimation year: 2009, mean = +0.109 mm·day^{-1}) is driven by large-scale drought and wet anomalies rather than systematic model deficiency. These results position the TS-FIS — with only 16 consequent parameters, 3 input features, and 4 fuzzy clusters — as a parsimonious, interpretable, and operationally viable ET_0 tool for the semi-arid Batna environment.

GENERAL CONCLUSION AND RECOMMENDATIONS

General Conclusions

This thesis has developed and rigorously validated a first-order Takagi–Sugeno Fuzzy Inference System (TS-FIS) for the estimation of daily reference evapotranspiration (ET_0) across 24 meteorological stations of the Batna Province, northeastern Algeria. The study utilized 26 years (2000–2025) of ERA5-Land reanalysis data, with the FAO-56 Penman–Monteith equation serving as the reference target, a 70/15/15 train–validation–test partitioning strategy, and a nine-metric performance evaluation framework. The following principal conclusions are drawn:

1. Excellent and stable ET_0 estimation accuracy: The TS-FIS with 4 fuzzy clusters achieves test-set NSE ≥ 0.878 and Pearson R ≥ 0.939 at all 24 stations, with network means of 0.896 and 0.948 respectively, classifying performance as 'very good' to 'excellent' at all locations according to the standard thresholds of Moriasi et al. (2007).

2. Sufficiency of the 3-input design: The VPD–T_{2m}–T_soil triplet — selected by Pearson correlation analysis — is sufficient to explain the dominant variability of ET_0 in the semi-arid Batna climate, with VPD emerging as the leading predictor at 75% of stations. This substantially reduced input requirement (versus the 6-variable FAO-56 equation) enhances operational feasibility in data-scarce environments.

3. Robust generalization without overfitting: The absolute train-to-test RMSE difference remains ≤ 0.032 mm day^{-1} at all stations (network mean = 0.001 mm day^{-1}), confirming that the 4-cluster TS-FIS architecture is well-matched to the available data and does not overfit across any of the 24 physiographically diverse sub-regions.

4. Near-unbiased long-term predictions: The validation-based bias correction reduces the overall mean bias to −0.0105 mm day^{-1} at the network level, representing less than 0.3% of mean ET_0. Residual station-level MBE values of +0.090 to +0.129 mm day^{-1} are operationally negligible for irrigation scheduling.

5. Spatially coherent error patterns: The Ordinary Kriging analysis reveals that inter-annual variability in prediction bias is primarily driven by large-scale climatic anomalies — severe drought years produce underestimation, anomalously wet/cool years produce overestimation — rather than systematic spatial model deficiencies. The eastern Aurès transition zone exhibits the highest inter-annual error variability.

6. Competitive performance relative to the literature: The TS-FIS achieves performance fully competitive with ANFIS and other machine learning models documented for semi-arid

North African contexts, using significantly fewer inputs, fewer fuzzy rules, and a more interpretable model architecture.

Recommendations for Future Work

1. Regional transferability: The TS-FIS framework should be tested across other climatic zones of Algeria — the humid Tell coastal zone, the hyper-arid Saharan zone, and the subhumid highlands of Kabylia — to assess the universality of the VPD–T_{2m}–T_soil triplet and the optimal cluster number across contrasting evaporative demand regimes.

2. Adaptive cluster optimization: A systematic sensitivity analysis of the FCM cluster number ($c = 2$ to 6) should be conducted per station, using cross-validated NSE or BIC as the optimization criterion. Stations with the highest inter-annual error variability (El Madher, Timgad, Arris) may benefit from a 5-cluster solution with a dedicated 'extreme summer' cluster.

3. Satellite-based real-time implementation: All three TS-FIS inputs (VPD, T_{2m}, T_soil) are available at daily frequency from operational ERA5-Land streams, MODIS Land Surface Temperature, and Sentinel-3 SLSTR products. Future work should develop and validate a fully satellite-driven operational implementation enabling real-time ET_0 estimation without ground-based measurements.

4. Climate change projections: The trained TS-FIS models should be applied to CMIP6 regional climate model outputs (SSP2-4.5, SSP5-8.5) to quantify projected changes in ET_0 under future warming scenarios. Given the demonstrated sensitivity of ET_0 to VPD in this semi-arid environment, climate-driven VPD increases are expected to produce disproportionate ET_0 trend acceleration.

5. Operational decision support system: The validated TS-FIS models should be integrated into a web-based or mobile API for real-time ET_0 estimation and irrigation scheduling recommendations, accessible to agricultural water managers in the Batna wilaya. The model's computational simplicity (4 cluster evaluations per prediction) makes it ideal for deployment on low-power edge computing platforms.

In conclusion, this thesis demonstrates that the Takagi–Sugeno Fuzzy Inference System provides an excellent balance of predictive accuracy, physical interpretability, computational efficiency, and operational practicability for ET_0 estimation in the semi-arid Batna environment. Quantitatively, the model achieves: (i) a network-mean $R^2 = 0.901$ and $NSE = 0.896$ on the independent test set across all 24 stations; (ii) a mean

RMSE = 0.667 mm·day^{-1} and MAE = 0.534 mm·day^{-1} — within the best-reported range for semi-arid North African ET_0 models (Ladlani et al., 2014 [15]; Kisi, 2013 [14]); (iii) a long-term spatial mean bias of only −0.0105 mm·day^{-1} (< 0.3% of regional mean ET_0) confirmed across 26 years and 27 Kriging maps; and (iv) a train-to-test RMSE degradation of only 0.001 mm·day^{-1} at the network level, with 17 of 24 stations (70.8%) rated Excellent for generalization. This performance is achieved with only 3 input variables, 4 fuzzy clusters, and 16 consequent parameters — representing an estimated 87.5% reduction in model complexity relative to comparable ANFIS architectures while maintaining fully competitive accuracy. The model's consistent performance across 24 stations, 26 years (2000–2025), and diverse physiographic conditions spanning 1,528 m of altitudinal relief positions it as a reliable, scientifically sound, and computationally lightweight tool for water resource management and agricultural planning in northeastern Algeria and comparable semi-arid Mediterranean environments.

References

References

[1] Allen, R.G., Pereira, L.S., Raes, D., & Smith, M. (1998). Crop evapotranspiration – Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56. FAO, Rome.

[2] Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. IEEE Transactions on Systems, Man, and Cybernetics, 15(1), 116–132.

[3] Zadeh, L.A. (1965). Fuzzy sets. Information and Control, 8(3), 338–353.

[4] Jang, J.S.R. (1993). ANFIS: Adaptive-network-based fuzzy inference system. IEEE Transactions on Systems, Man, and Cybernetics, 23(3), 665–685.

[5] Jang, J.S.R., Sun, C.T., & Mizutani, E. (1997). Neuro-Fuzzy and Soft Computing. Prentice Hall, Upper Saddle River, NJ.

[6] Ross, T.J. (2010). Fuzzy Logic with Engineering Applications (3rd ed.). John Wiley & Sons, Chichester.

[7] Hargreaves, G.H., & Samani, Z.A. (1985). Reference crop evapotranspiration from temperature. Applied Engineering in Agriculture, 1(2), 96–99.

[8] Priestley, C.H.B., & Taylor, R.J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. Monthly Weather Review, 100(2), 81–92.

[9] Nash, J.E., & Sutcliffe, J.V. (1970). River flow forecasting through conceptual models part I. Journal of Hydrology, 10(3), 282–290.

[10] Willmott, C.J. (1982). Some comments on the evaluation of model performance. Bulletin of the American Meteorological Society, 63(11), 1309–1313.

[11] Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., & Veith, T.L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE, 50(3), 885–900.

[12] Bezdek, J.C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. Springer, New York.

[13] Cobaner, M. (2011). Evapotranspiration estimation by two different neuro-fuzzy inference systems. Journal of Hydrology, 398(3–4), 292–302.

[14] Kisi, O. (2013). Applicability of Mamdani and Sugeno fuzzy genetic approaches for modeling reference evapotranspiration. Journal of Hydrology, 504, 160–170.

[15] Ladlani, I., Houichi, L., Djemili, L., Heddam, S., & Belouz, K. (2014). Estimation of daily reference evapotranspiration in north Algeria using ANFIS and MLR. Arabian Journal for Science and Engineering, 39(8), 5959–5969.

[16] Monteith, J.L. (1965). Evaporation and environment. Symposia of the Society for Experimental Biology, 19, 205–234.

[17] Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

[18] Xie, X.L., & Beni, G. (1991). A validity measure for fuzzy clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 13(8), 841–847.

[19] ONM – Office National de Météorologie, Algérie. (2025). Données météorologiques journalières des stations de la wilaya de Batna, 2000–2025.

[20] Copernicus Climate Change Service (C3S). (2024). ERA5-Land hourly data from 1950 to present. Copernicus Climate Data Store. https://cds.climate.copernicus.eu/

[21] Hersbach, H., et al. (2020). The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 146(730), 1999–2049.

[22] Muñoz-Sabater, J., et al. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. Earth System Science Data, 13(9), 4349–4383.

[23] Ficklin, D.L., & Novick, K.A. (2017). Historic and projected changes in vapor pressure deficit suggest a continental-scale drying of the United States atmosphere. Journal of Geophysical Research: Atmospheres, 122(4), 2061–2079.

[24] Tabari, H. (2010). Evaluation of reference crop evapotranspiration equations in various climates. Water Resources Management, 24(10), 2311–2337.

[25] Shiri, J., et al. (2014). Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran. Computers and Electronics in Agriculture, 108, 230–241.

[26] Kisi, O., & Cimen, M. (2009). Evapotranspiration modelling using support vector machines. Hydrological Sciences Journal, 54(5), 918–928.

[27] Blaney, H.F., & Criddle, W.D. (1950). Determining Water Requirements in Irrigated Areas from Climatological and Irrigation Data. USDA, Technical Paper No. 96.

[28] Jamieson, P.D., Porter, J.R., & Wilson, D.R. (1991). A test of the computer simulation model ARCWHEAT1 on wheat crops grown in New Zealand. Field Crops Research, 27(4), 337–350.

[29] Draper, N.R., & Smith, H. (1998). Applied Regression Analysis (3rd ed.). John Wiley & Sons, New York.

[30] Kosko, B. (1992). Neural Networks and Fuzzy Systems. Prentice Hall, Englewood Cliffs, NJ.

[31] Abyaneh, H.Z., et al. (2011). Performance evaluation of ANN and ANFIS models for estimating garlic crop evapotranspiration. Journal of Irrigation and Drainage Engineering, 137(5), 280–286.

[32] Yuan, W., et al. (2019). Increased atmospheric vapor pressure deficit reduces global vegetation growth. Science Advances, 5(8), eaax1396.