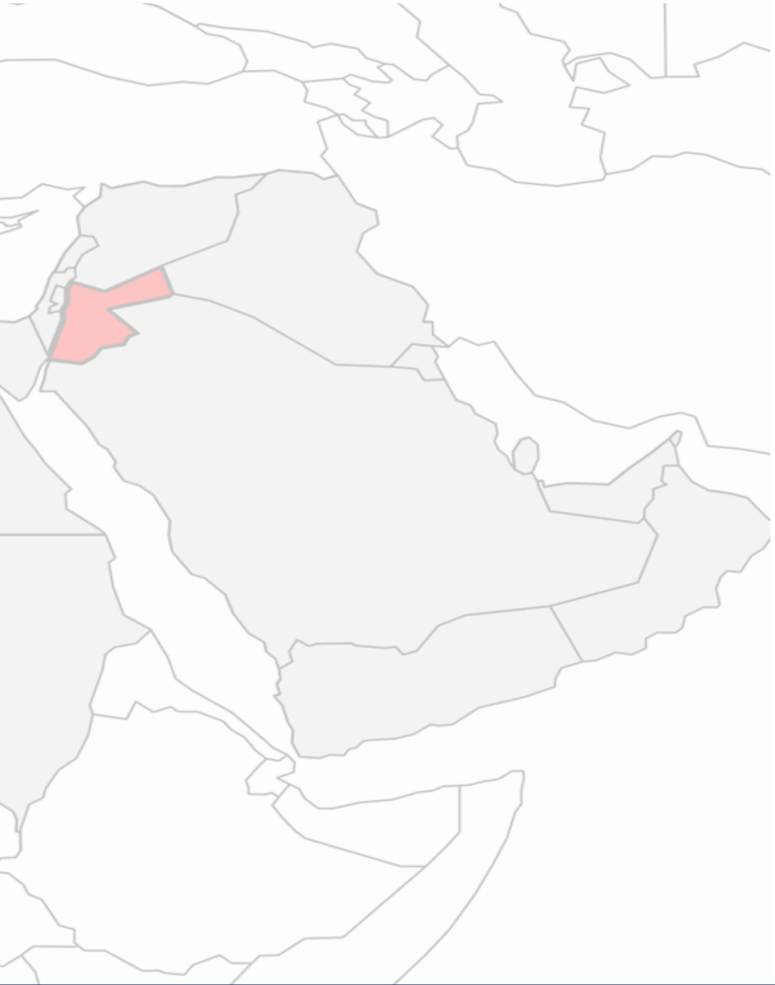
Innovation in Temperature Prediction Using ECMWF's Forecasts Event (UEF2026)

إدارة الأرصاد الجوية
المملكة الأردنية الهاشمية



According to the World Meteorological Organization (WMO) report on the State of the Climate in the Arab Region, 2024 was the hottest year on record, with accelerated warming and more frequent extremes (WMO, 2025).

Measurements of meteorological variables are conducted according to internationally recognized standards established by the World Meteorological Organization. Temperature observations collected at monitoring stations worldwide follow unified scientific protocols to ensure consistency, reliability, and comparability of recorded data. Yet despite these standardized practices, accurate temperature forecasting in the Middle East–North Africa (MENA) region remains challenged by increasing climate variability and extreme heat conditions. Addressing these challenges requires innovative approaches that combine advanced AI techniques with global forecast products and local observational truth

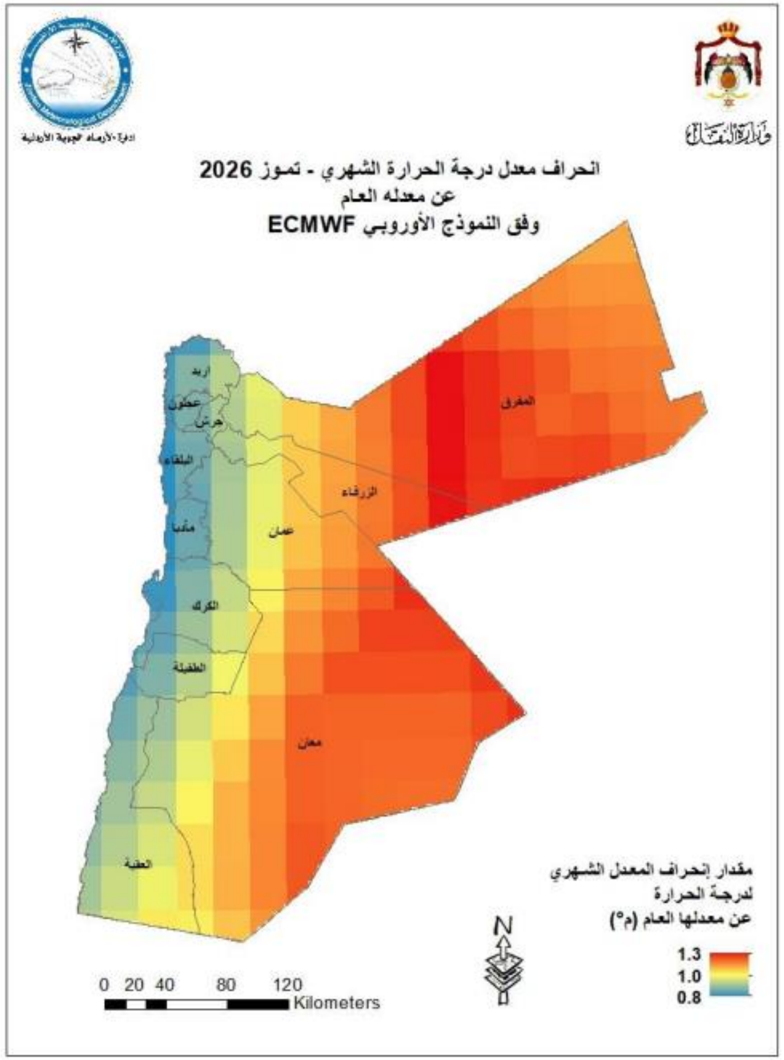
Ibrahim Hadyeh, MSc
Cybersecurity & AI Researcher | Data Analyst

Communications Systems Section
Jordan Meteorological Department

The Middle East and North Africa (MENA) region is characterized as one of the driest and hottest regions in the world. Rising temperatures and extreme heat pose severe, scientifically documented challenges to both aviation safety and infrastructure resilience on a global scale. Jordan’s climate is influenced by a transition between the Mediterranean basin and the arid desert

According to a recently issued seasonal outlook from the Jordan Meteorological Department (JMD), model guidance indicates that mean temperatures during July 2026 are likely to be approximately 0.8–1.3 °C above the long-term climatological average across most regions of the Kingdom, with probabilities reaching up to 70%, the shown figure illustrates the forecast monthly temperature anomaly for July 2026 relative to the climatological mean, based on the May 2026 ECMWF seasonal forecast cycle.

This type of climatological temperature anomaly outlook illustrates why interpretable anomaly detection and confidence-aware forecasting support are becoming increasingly important operationally.



Source: Jordan Meteorological Department (JMD)

Ibrahim Hadyeh, MSc
Cybersecurity & AI Researcher | Data Analyst

Communications Systems Section
Jordan Meteorological Department

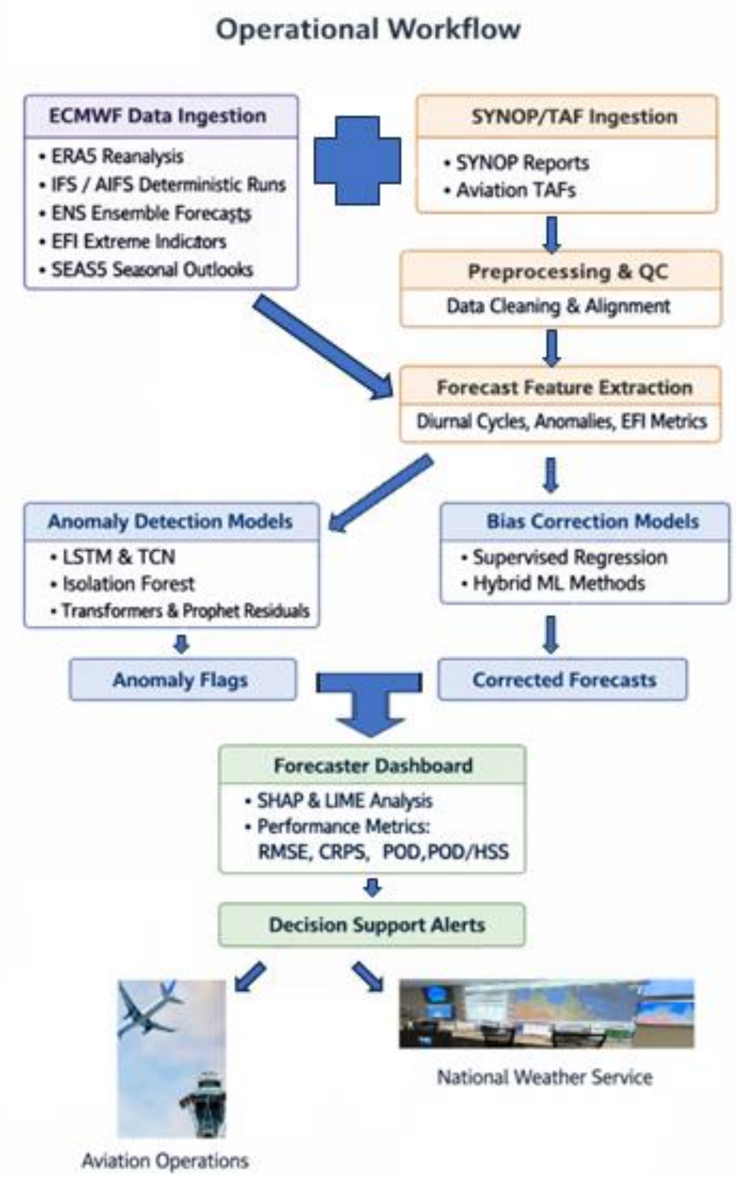
Innovation in Temperature Prediction

Aviation weather forecasting remains a critical challenge: pilots, dispatchers, airlines and ATC need accurate, timely, and trustworthy forecasts to ensure safety and efficiency. In aviation, even small errors in predicting wind shear, visibility, or convective storms can compromise safety. Together, While global models and satellite data underpin aviation meteorology, SYNOP and TAF remain indispensable at the operational scale, ensuring situational awareness and forecast guidance at airports., but they struggle with extreme events (heat spikes, dust storms, convective storms) and uncertainty quantification. SYNOP and TAF are often treated separately in datasets, but in aviation meteorology they are deeply interconnected. SYNOP provides the real-time observational backbone, while TAF translates those observations into actionable forecasts for pilots and controllers. Without SYNOP, TAF accuracy suffers, especially dangerous during extreme weather events. yet their inherent limitations, such as sparse spatial coverage, human bias, and short forecast horizons, creating critical vulnerabilities when confronting one of the above-mentioned extreme events.

Innovation in temperature prediction, through ECMWF AI products and machine learning models, fills the operational gap by enhancing accuracy, reliability, and resilience. This integration ensures that aviation meteorology can anticipate extremes with greater confidence. Jordan provides a compelling case study: its diverse climate, ranging from desert heatwaves to cold spells in highland regions, poses significant challenges for SYNOP and TAF forecasts. By applying ECMWF ensemble products and AI-driven anomaly detection to Jordanian SYNOP and TAF data, we demonstrate how advanced modelling can correct biases, strengthen reliability, and deliver actionable guidance for aviation operations in a region where extremes are both frequent and impactful.

AI-Driven Innovation in Temperature Prediction: Anomaly Detection in SYNOP and TAF Forecasts Using ECMWF Products: A Case Study from Jordan

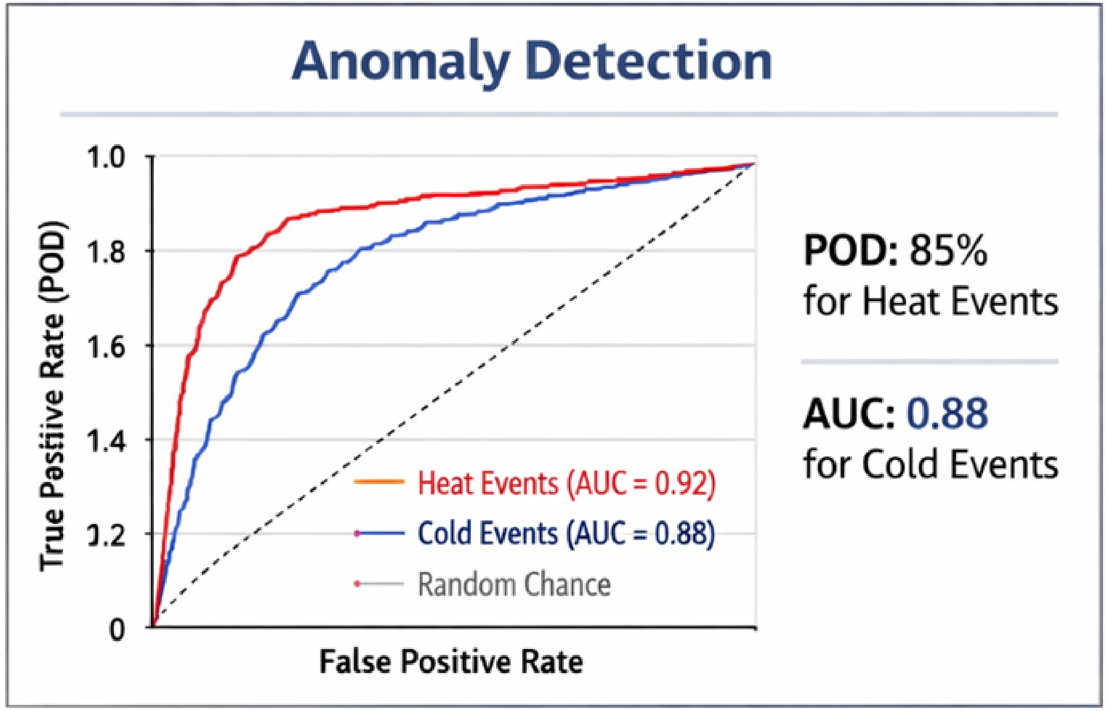
Building on these advances, this study employs AI-enhanced forecast products as the foundation of its methodology. High-resolution datasets such as reanalysis archives, deterministic runs, ensemble forecasts, extreme event indicators, and seasonal outlooks are integrated with SYNOP observations and aviation TAF reports. After preprocessing and feature extraction, the framework applies time-series machine-learning architectures (recurrent networks, convolutional models, isolation methods, and residual-based approaches) for anomaly detection, alongside supervised regression and hybrid techniques for bias correction. Model performance is evaluated using standard error metrics and classification measures (e.g., RMSE, CRPS, Brier score, ROC, POD, HSS), while explainable AI tools (such as SHAP and LIME) ensure interpretability through a forecaster dashboard. The resulting decision-support alerts enhance aviation operations and national weather services, delivering more reliable early warnings of temperature extremes in Jordan’s climate-vulnerable context.



Ibrahim
HADYEH

KEY RESULTS AND INSIGHTS

This ROC curve illustrates the performance of our anomaly detection models by plotting the trade-off between detection capability and false alarms across varying thresholds. The upward curvature of both event lines demonstrates strong separation between anomalous and normal states, well above the random-chance diagonal. The chart confirms that the models consistently achieve high detection skill, with heat events showing slightly stronger separability than cold events.

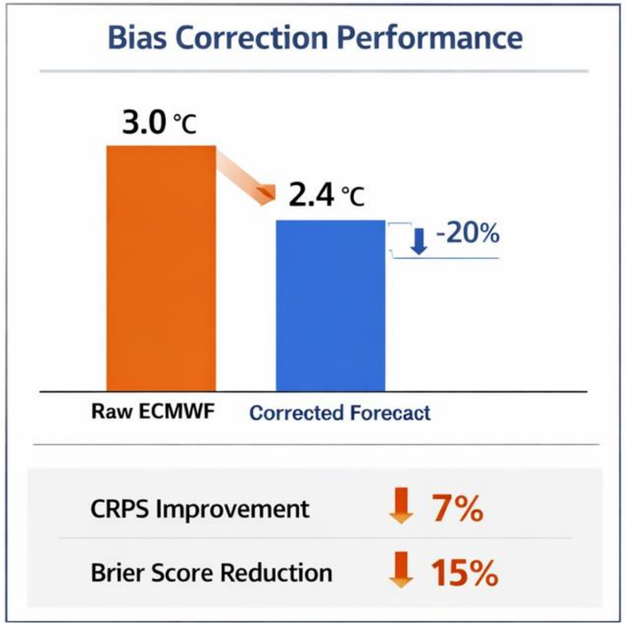


In anomaly detection, Probability of Detection (POD) reflects the sensitivity of the model to true events, ensuring anomalies are captured reliably. The Area Under the Curve (AUC), derived from the ROC curve, provides a threshold-independent measure of overall discrimination skill. Together, POD and AUC demonstrate both the operational effectiveness and statistical robustness of the proposed framework

Ibrahim HADYEH

AI-Driven Innovation in Temperature Prediction: Anomaly Detection in SYNOP and TAF Forecasts Using ECMWF Products: A Case Study from Jordan

KEY RESULTS AND INSIGHTS



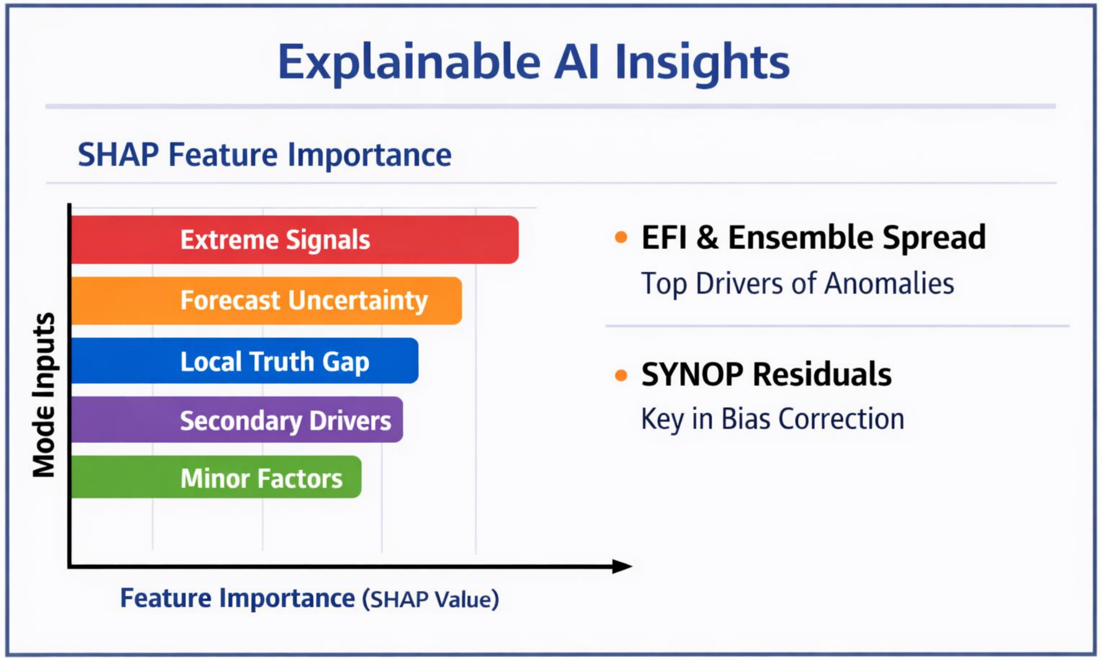
Bias correction improved ensemble forecast skill, reducing RMSE by 20%, CRPS by 7%, and Brier Score by 15%. These gains indicate sharper, better-calibrated forecasts that statistically closer to observed truth.

Other recent works have reported comparable RMSE reductions. For instance, Hui Zhang, Junming Chen, and colleagues (2025) applied Bayesian Model Averaging to ECMWF 2-meter temperature forecasts and achieved a 24.7% RMSE reduction, improving accuracy from ~73% to over 91%. Likewise, Lee and Zhang (2024) developed an AI framework using ConvLSTM with residual self-attention, which reduced RMSE by about 20% over 1–7 day forecasts compared to operational ECMWF outputs. These studies illustrate that RMSE reductions in the range of 20–25% are consistently achievable across different methodologies.

Ibrahim HADYEH

KEY RESULTS AND INSIGHTS

This SHAP feature importance chart provides a transparent view of how the model interprets meteorological inputs. The top bar, labeled Extreme Signals, represents the influence of the Extreme Forecast Index (EFI), highlighting how the model responds to unusual weather patterns and extremes. Forecast Uncertainty captures the ensemble spread, reflecting how variability among ensemble members drives anomaly detection. The Local Truth Gap quantifies the difference between forecasts and SYNOP observations, serving as the foundation for bias correction. Beneath these, Secondary Drivers contribute supportive but less dominant effects, while Minor Factors represent features with minimal influence. Together, these layers illustrate a balanced model architecture



The Ensemble Spread reflects forecast uncertainty, showing how variability among ensemble members signals potential anomalies. In contrast, SYNOP Residuals capture the local truth gap, quantifying differences between ECMWF forecasts and ground observations to correct systematic bias. Together, they balance global uncertainty with local observational truth, ensuring reliable anomaly detection.

CONCLUSION & FUTURE WORK

Conclusion

- Our framework demonstrated robust anomaly detection skill, effective bias correction, and transparent interpretability. ROC analysis confirmed strong detection capability.
- Bias correction reduced systematic errors and improved forecast reliability, and explainable AI insights revealed ensemble spread and EFI as key drivers of anomalies, with SYNOP residuals anchoring bias correction.
- By integrating ECMWF ensemble products with OGIMET observations, the system directly addresses the forecasting challenges outlined at the start, delivering both scientific rigor and operational transparency for the MENA region.

Future Work – Validation & Benchmarking

- Expand datasets (ERA5, SYNOP, TAF, ECMWF ENS/EFI/SEAS5).
- Validate anomaly detection skill (ROC, POD, HSS) across diverse events.
- Benchmark bias correction against published state-of-the-art (Bayesian averaging, ConvLSTM, hybrid ML).
- Operational testing with aviation partners for real-world reliability.

Thank you

Thesis Architecture