Application of Temporal Fusion Transformer to Trail Running Predictions

Master's Degree Final Project

Eric Aguayo

Universidad San Francisco de Quito

December 2025

🏔️ The Challenge

Trail Running Complexity

- Long Endurance activities (3-6 hours)
- Extreme **elevation changes**
- Complex **fatigue dynamics**
- **Nutrition & hydration** critical

Why Prediction Matters

- ⚠️ **30% under-prediction** → dehydration, bonking
- 📊 **Over-prediction** → suboptimal pacing
- 🎯 **Goal**: Accurate race time estimation

⚠️ The problem: cold-start predictions

> *"How can we predict race completion time **before the race begins**, without any data from the current session?"*

Traditional Methods

Method	Limitation
Average pace	Ignores terrain complexity
Naismith's rule	No fatigue modeling
Regression	Misses temporal dependencies

💡 The Solution

Temporal Fusion Transformer (TFT)

Architecture Components:

- 🚪 Gated Residual Networks (GRN)
- 📊 Variable Selection Networks
- 🔄 LSTM Encoder-Decoder
- 🎯 Multi-head Self-Attention

Key Advantages

- Handles **known** vs **unknown** futures
- Learns **long-range** dependencies
- Provides **interpretable** attention weights
- Supports **multi-target** forecasting

🔬 Contributions

1. **Novel Application**: First documented TFT for trail running
2. **Cold-Start Methodology**: Synthetic encoder approach
3. **Asymmetric Loss Function**: Corrects under-prediction bias
4. **Distance-Domain Resampling**: Pace-independent predictions
5. **Multi-Target Forecasting**: Duration, HR, temp, cadence
6. **Error Cancellation Analysis**: Robust evaluation guidance

📊 Data Pipeline

Distance-Domain Resampling

```
106 Polar sessions (79 train / 16 val / 11 test)
Time Domain (1 sec) → Distance Domain (5 meters)
7 Garmin sessions (5 train / 1 val / 1 test)
```

Input Features:

- Heart Rate
- Altitude
- Cadence
- Speed
- Temperature

Derived Features:

- Elevation diff/gain/loss
- Duration per interval
- Fatigue proxies
- Rolling averages

🧊 Cold-Start Solution

Synthetic Encoder Approach

$$x_{synthetic} = \frac{\sum_{s=1}^{S} w_s \cdot x_{s,0}}{\sum_{s=1}^{S} w_s}$$

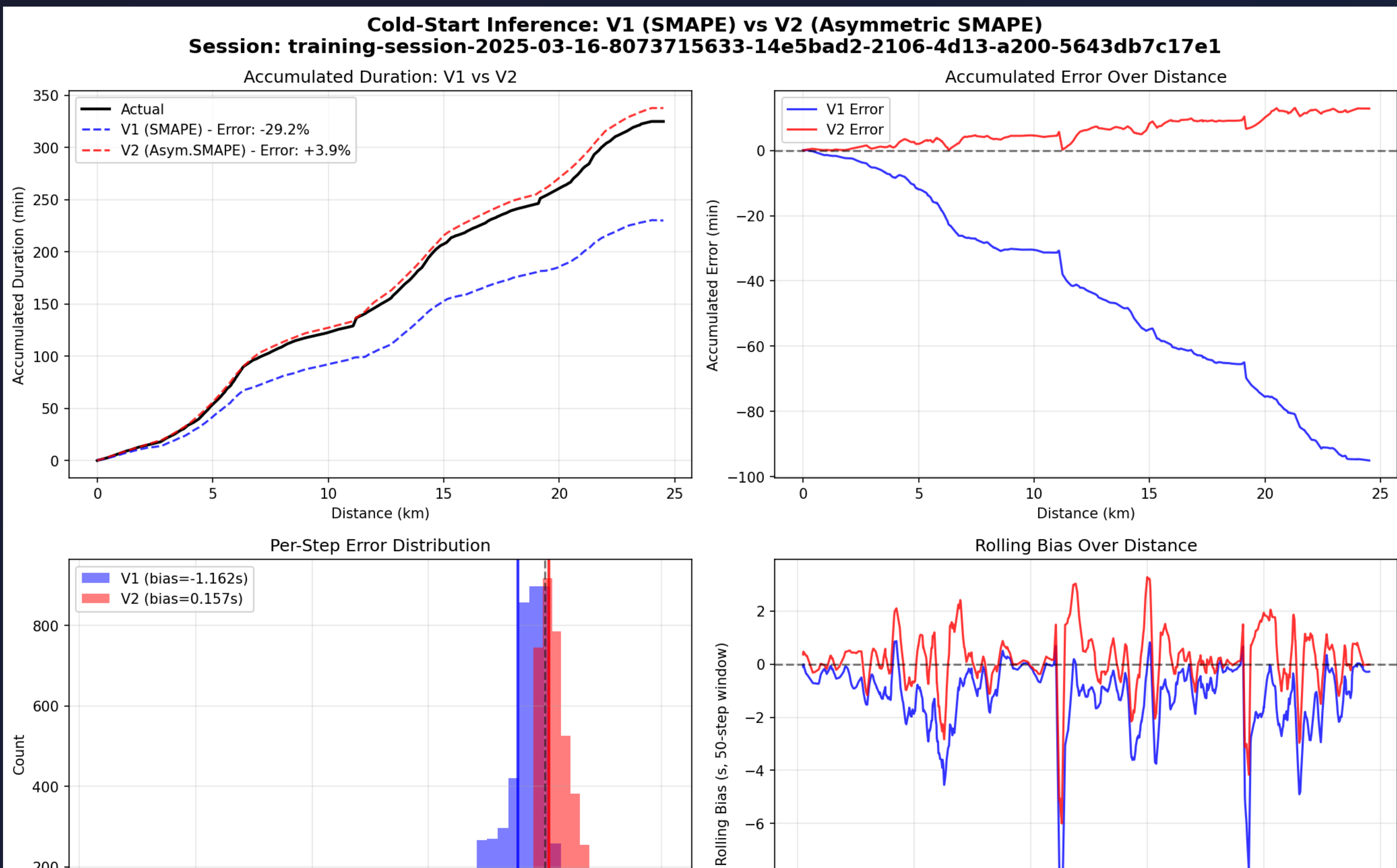
- Weight historical first samples
- Recent sessions weighted higher
- Use actual **terrain data** (known)
- Estimate physiological baseline

Synthetic Encoder Purpose

- Capture current **fitness level**
- Leverage **population patterns**
- GPS route **preview** available
- Add an initial input without any previous race data! 🎯

📊 Results

Accumulated Duration Prediction



📈 Results: V1 vs V2

Cold-Start on 24.5 km Session / 324.9 min

Metric	V1 (sMAPE)	V2 (Asym)	Change
MAE (s/5m)	1.347	1.066	**-20.9%**
Bias (s/5m)	-1.162	+0.157	**+113.5%**
Predicted	229.9 min	337.7 min	+46.9%
Accumulated Error	**-29.2%**	**+3.9%**	✅

Across all 11 test sessions: -30.4% → **+3.7%** 🎯

⚖️ Asymmetric Loss Function

Correcting Under-Prediction Bias

$$\mathcal{L}_{asym} = w \cdot \frac{|y - \hat{y}|}{|y| + |\hat{y}| + \epsilon} \cdot 2$$

$$w = \begin{cases} \alpha & \text{if } y > \hat{y} \quad \text{(under-prediction)} \\ 1 - \alpha & \text{if } y \leq \hat{y} \quad \text{(over-prediction)} \end{cases}$$

With α = 0.51: Slight penalty for under-prediction

⚠️ *Highly sensitive parameter:*
α>=0.55 caused high over-prediction

🔄 V3: Transfer Learning

Fine-tuning with Garmin + Nutrition Data

New Features (V3):

- Rate of Perceived Exertion (RPE)
- Water intake
- Electrolyte intake
- Food intake

Via NutritionLogger app

Key Finding:

Metric	V3 (Garmin)	V2 (Polar)
MAE (s/5m)	0.168	0.633
Final Error	+2.2%	+1.8%

⚠️ ***Error Cancellation Warning:***
Lower cumulative error can mask

⚠️ Limitations

Limitation	Impact	Mitigation and Needs
Single-Athlete Dataset	Multi-athlete generalization	Multi-athlete data needed
106 Sessions	Overfitting risk	Dropout=0.25 / More data needed
Geographic Specificity	Andes-trained mostly	More data needed
Missing Features	Weather, sleep, HRV absent	Future sensor integration
Fine Tuning Limitation	Restricted to base model size	Increase model complexity

V3 Lesson: 5 sessions insufficient for sparse

Practical Applications

🏃 Race Planning

Estimate finish time **before race start** based on route profile

🍎 Nutrition Planning

Predicted duration informs **caloric & fluid** needs

Practical Applications

📈 Training Optimization

Analyze **predicted vs actual** performance

⏱️ Pacing Strategy

Real-time predictions for **pacing adjustments**

🔮 Future Work

1. **Multi-Athlete Data** – Incorporate data from Strava and Garmin Connect to improve generalization.
2. **Few-Shot Analysis** – Evaluate predictions when a small amount of prior session data is available.
3. **Weather Integration** – Include external factors such as temperature, humidity, and wind.
4. **Uncertainty Quantification** – Provide confidence intervals for model predictions.
5. **Interpretability** – Analyze attention weights to better understand feature importance.

6. **Model Compression** – Enable on-device inference, e.g., from phone to watch.
7. **Multi-Metric Evaluation** – Assess both per-step and cumulative prediction accuracy.
8. **Session Type Classification** – Distinguish between training and race sessions.
9. **Race Planning Optimization** – Develop strategies to minimize race completion time.
10. **IMU Sensor Integration** – Incorporate preprocessed accelerometer/gyroscope data for terrain technicality analysis.

🎓 Conclusions

1. **TFT Successfully Applied** to trail running prediction
2. **Cold-Start Prediction Achieved** via synthetic encoder approach
3. **Asymmetric Loss** reduced error from **-30.4% to +3.7%**
4. **Distance-Domain Processing** enables pace-independent predictions
5. **Transfer Learning Works** - V2 model transferred to Garmin but more data is needed for true validation
6. **Critical Insight:** Evaluate both **per-step** and **cumulative** metrics

> 🎯 *"Conservative over-prediction is preferable for race planning"*

Thank You!

Questions?