Machine Learning for Streaming Data

IJCAI Tutorial 2024

Heitor Murilo Gomes[1*], Nuwan Gunasekara[2]
Albert Bifet[2,3], Bernhard Pfahringer[2]

* Corresponding author: heitor.gomes@vuw.ac.nz

https://nuwangunasekara.github.io/ijcai2024/

[1] Victoria University of Wellington, New Zealand, [2] University of Waikato, New Zealand,
[3] TELECOM Paris, LCTI, France.

Prediction Intervals

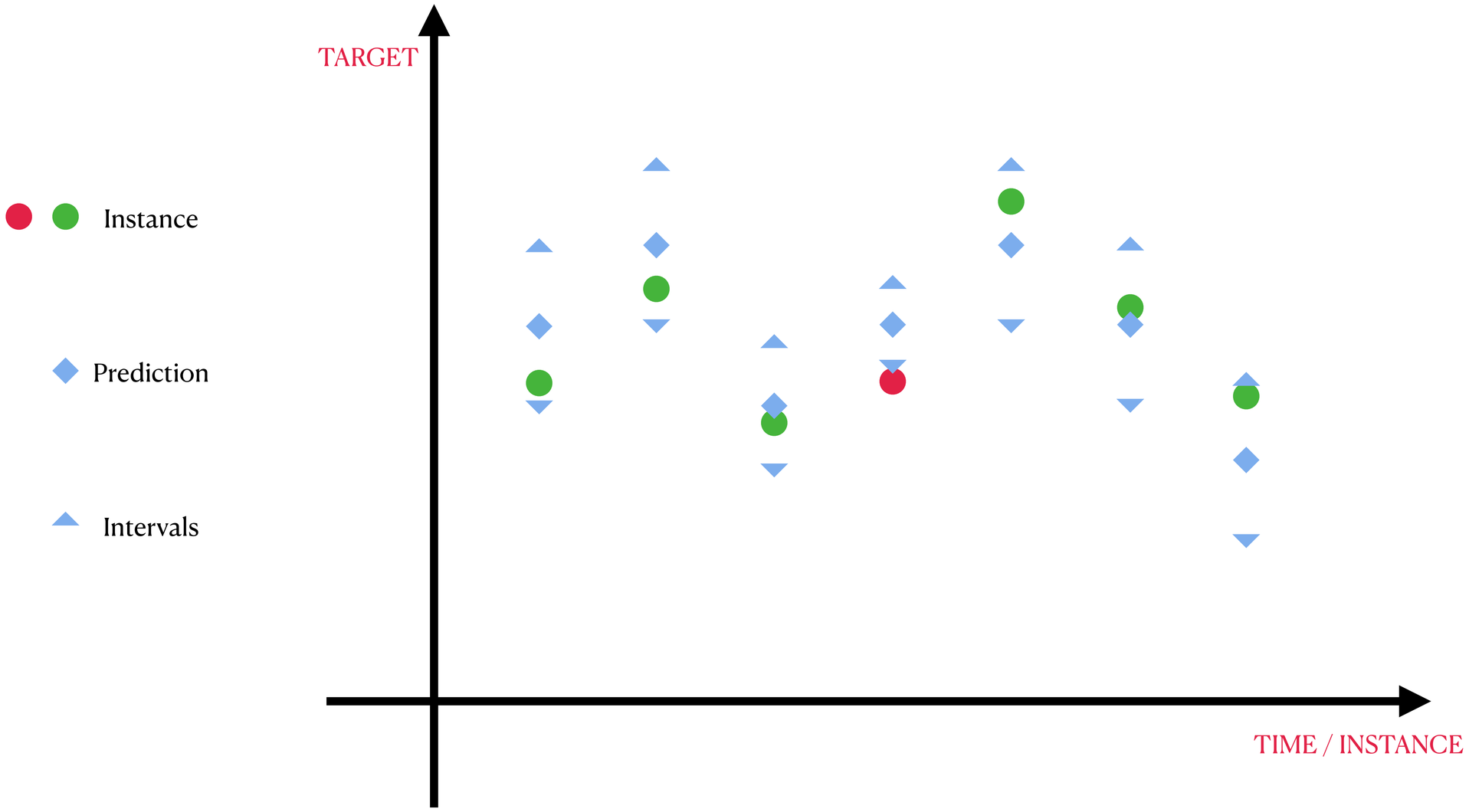
Prediction Intervals

Prediction Intervals (PIs) are very useful improve our confidence in predictions yield in regression tasks

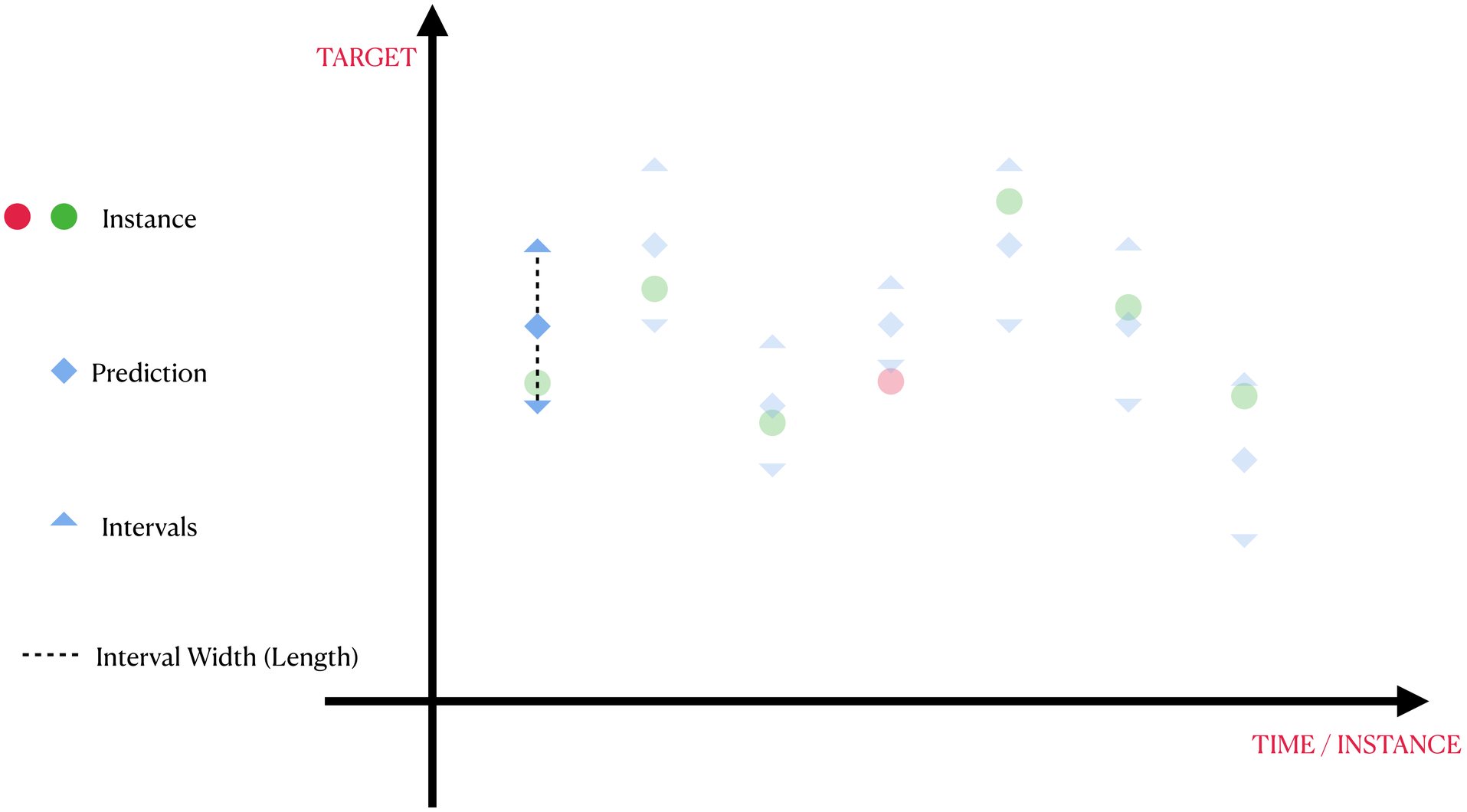
Challenge

Traditional PI methods were not designed to adapt to evolving streams (i.e. those with concept drift)

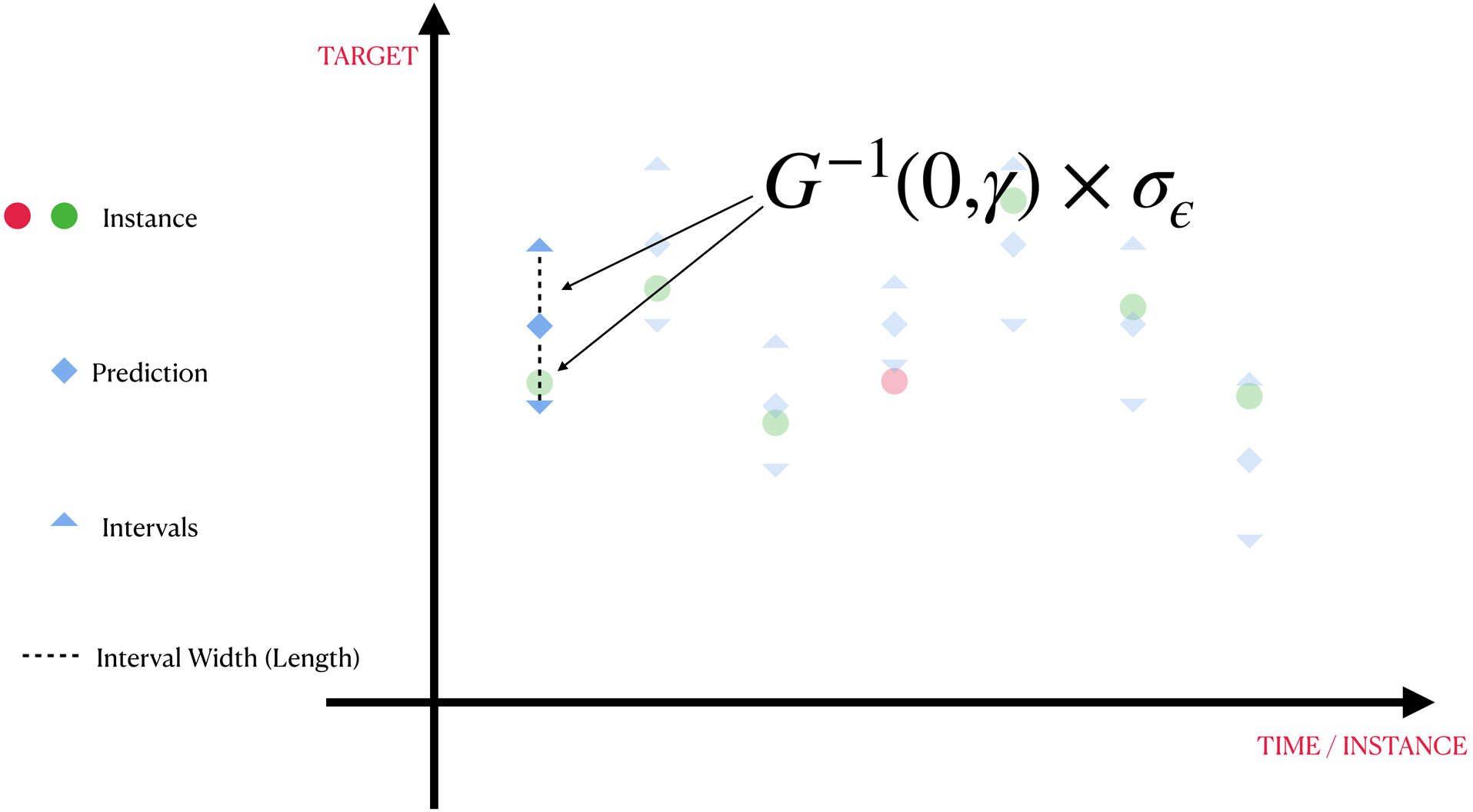
Prediction Interval



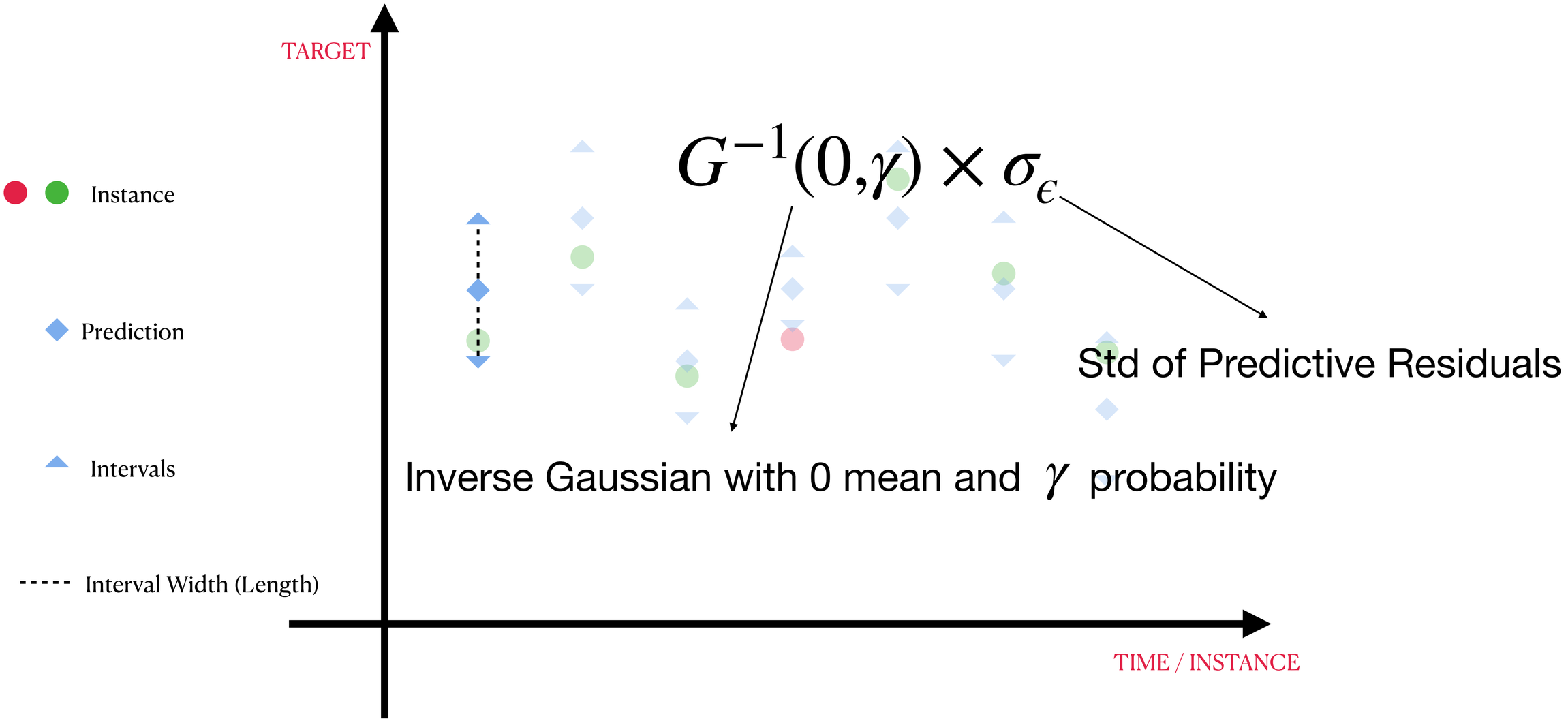
Prediction Interval



Mean and Variance Estimation (MVE)



Mean and Variance Estimation (MVE)



Mean and Variance Estimation (MVE)

$$\Pr\left(y \in \left[\hat{y} - G^{-1}(0,\gamma) \times \sigma_\epsilon, \hat{y} + G^{-1}(0,\gamma) \times \sigma_\epsilon\right]\right) \approx \gamma$$

$$\text{PI}_{\text{MVE}} \in \left(\hat{y} - G^{-1}(0,\gamma) \times \sigma_\epsilon \, , \hat{y} + G^{-1}(0,\gamma) \times \sigma_\epsilon \right)$$

γ : Confidence Level / Significance Level

Adaptive Prediction Interval (AdaPI)

$$\Pr\Big(y \in \big[\hat{y} - \mathcal{S} \times G^{-1}(0,\gamma) \times \sigma_\epsilon, \hat{y} + \mathcal{S} \times G^{-1}(0,\gamma) \times \sigma_\epsilon\big] \Big) \approx \gamma$$

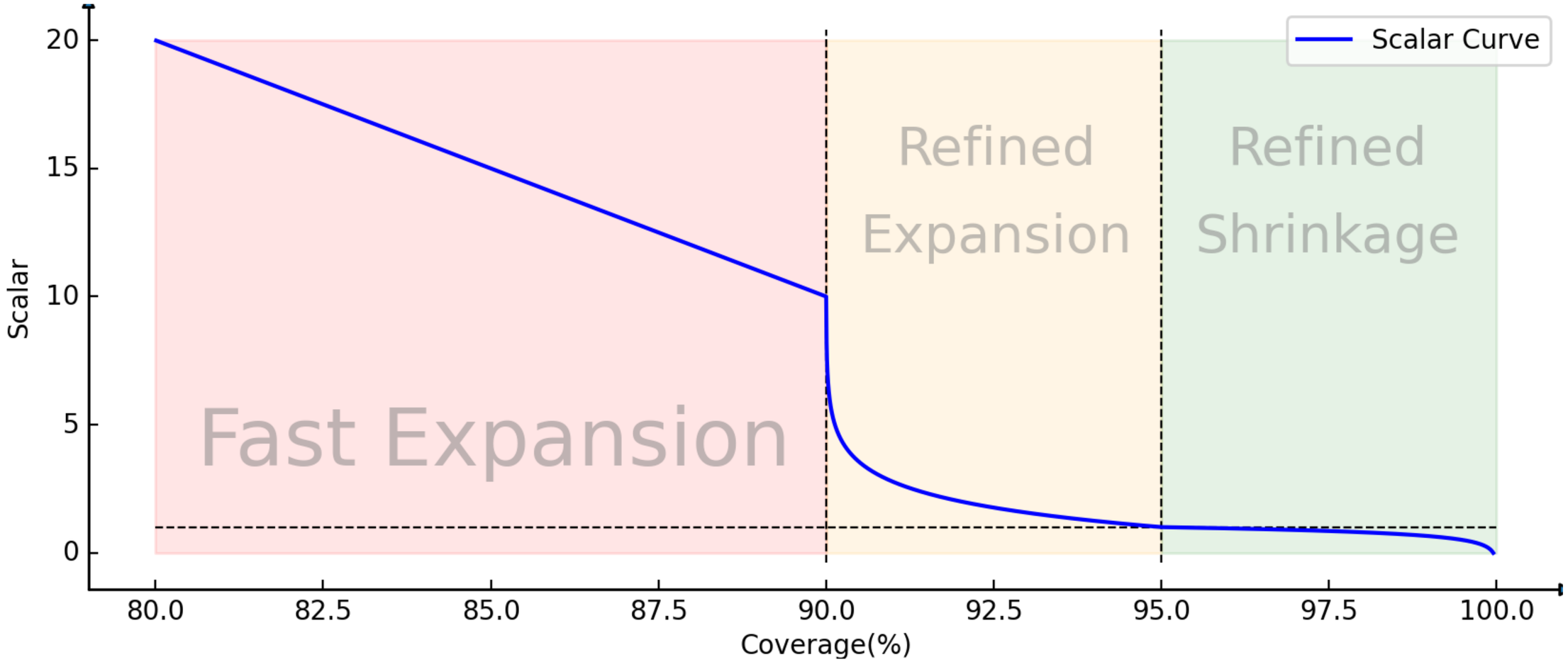
$$\mathrm{PI}_{\mathrm{AdaPI}} \in \big(\hat{y} - \mathcal{S} \times G^{-1}(0,\gamma) \times \sigma_\epsilon \ , \hat{y} + \mathcal{S} \times G^{-1}(0,\gamma) \times \sigma_\epsilon \ \big)$$

γ : Confidence Level / Significance Level

$\mathcal{S}$: Scalar

Yibin Sun, B Pfahringer, H M Gomes, and A Bifet. "Adaptive Prediction Interval for Data Stream Regression." PAKDD (2024)

Scalar for AdaPI



Yibin Sun, B Pfahringer, H M Gomes, and A Bifet. "Adaptive Prediction Interval for Data Stream Regression." PAKDD (2024)

Evaluation Metrics for PI

$$\text{Coverage} = \frac{1}{N}\sum_{i=1}^{N} I_i$$

$$NMPIW = \frac{\frac{1}{N}\sum_{i=1}^{N}(P_{u_i} - P_{l_i})}{R}$$

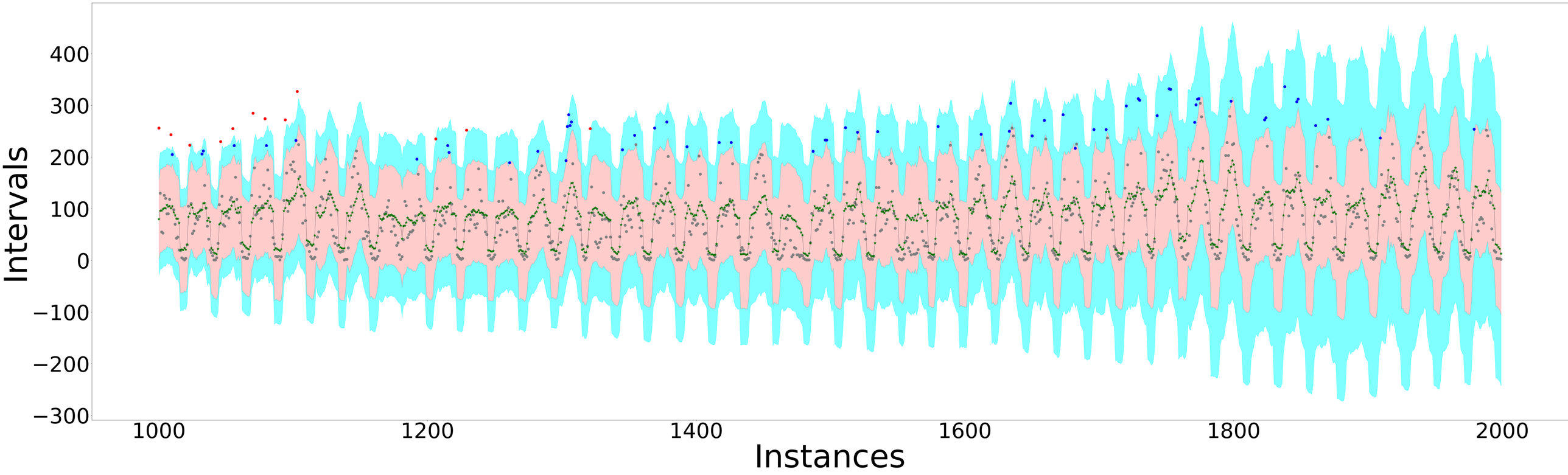
NMPIW : Normalized Mean Prediction Interval Width

R : Range of Target Values

P_u, P_l : Upper and Lower Bounds of Prediction Intervals

Expansion Case

Datasets: Bike

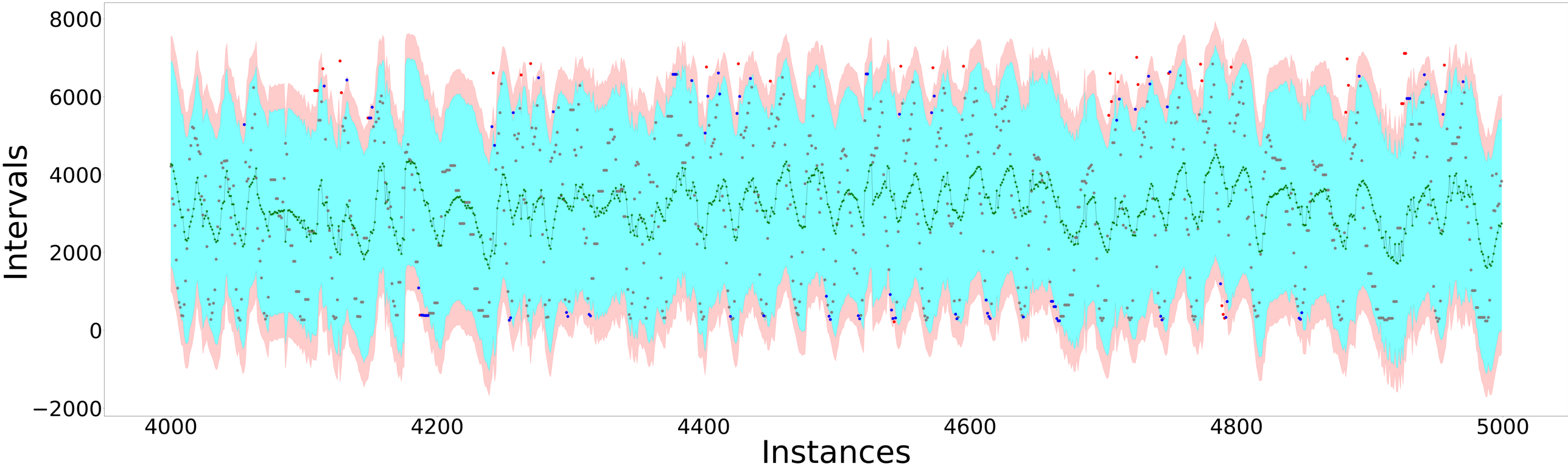


Blue area: AdaPI Red area: MVE

Yibin Sun, B Pfahringer, H M Gomes, and A Bifet. "Adaptive Prediction Interval for Data Stream Regression." PAKDD (2024)

Shrinkage Case

Datasets: MetroTraffic

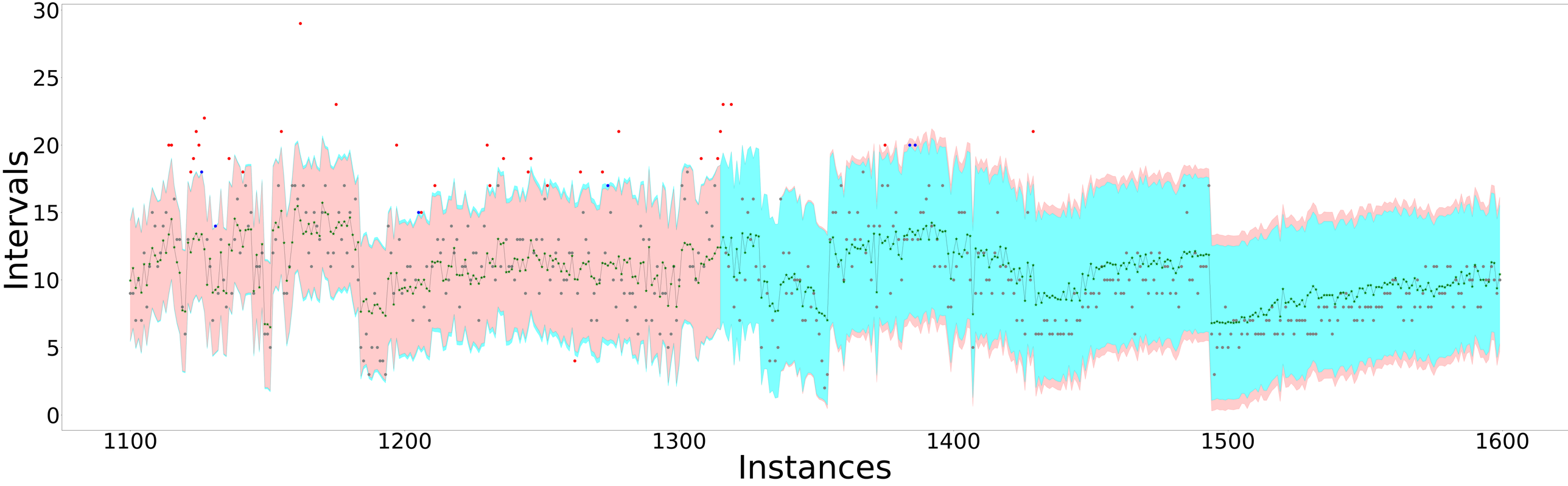


Blue area: AdaPI Red area: MVE

Yibin Sun, B Pfahringer, H M Gomes, and A Bifet. "Adaptive Prediction Interval for Data Stream Regression." PAKDD (2024)

Switching Case

Datasets: Abalone

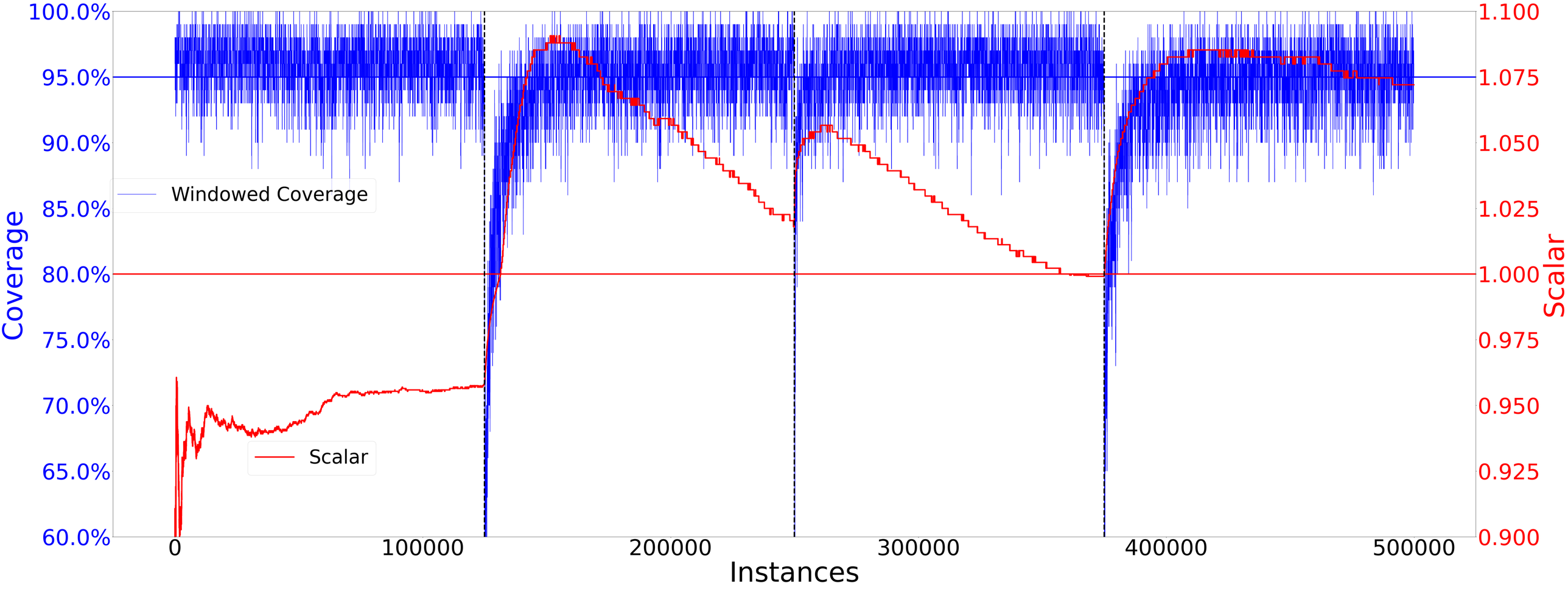


Blue area: AdaPI Red area: MVE

Yibin Sun, B Pfahringer, H M Gomes, and A Bifet. "Adaptive Prediction Interval for Data Stream Regression." PAKDD (2024)

Dealing with Drifts

Datasets: HyperA



Yibin Sun, B Pfahringer, H M Gomes, and A Bifet. "Adaptive Prediction Interval for Data Stream Regression." PAKDD (2024)

Prediction Intervals Summary

Prediction Interval (PI) is essential for uncertainty quantification in regression tasks.

Challenges

Traditional PI methods are not suitable for dynamic data streams.

Solution

Mean and Variance Estimation (MVE); ADAPI*

Evaluation

Coverage

Normalised Mean Prediction Interval Width (NMPIW)

Yibin Sun, B Pfahringer, H M Gomes, and A Bifet. "Adaptive Prediction Interval for Data Stream Regression.” PAKDD (2024)

Practical examples

IJCAI_2024_prediction_intervals.ipynb