Per Link Data-driven Network Replication Towards Self-Adaptive Digital Twins

MSWiM 2025

Universitat Politècnica de Catalunya (UPC)

Samir SI-MOHAMMED, Fabrice THÉOLEYRE

Associate Professor at University of Lorraine, France

CRAN/Polytech Nancy

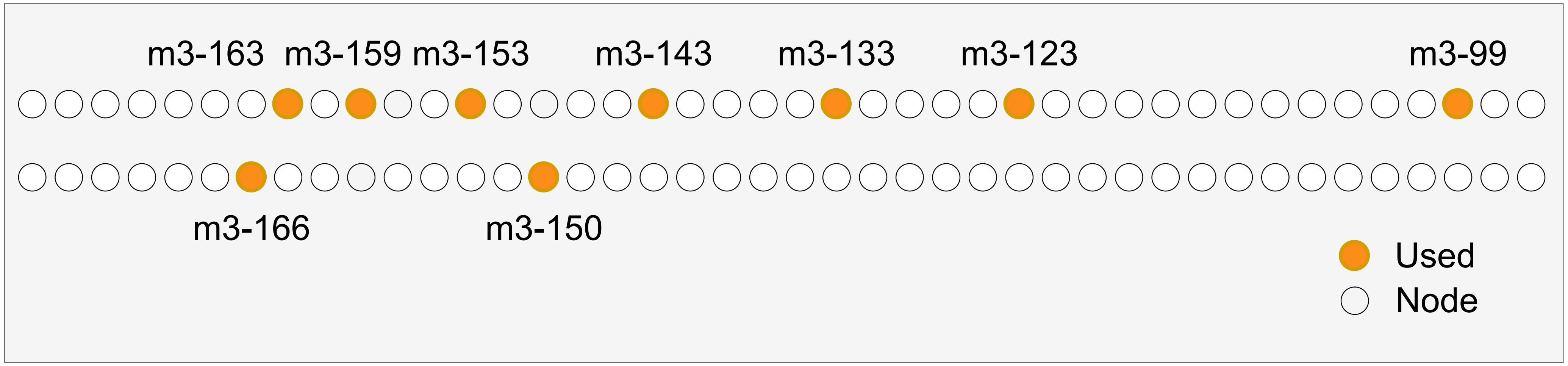
samir.si-mohammed@univ-lorraine.fr

28 October 2025

Problem Statement

❖ Case Study:

- Network architect deploying IoT solutions
- Industrial deployment on the experimental platform *FIT IoT-Lab* [1]
  - **Smart metering in an indoor environment**



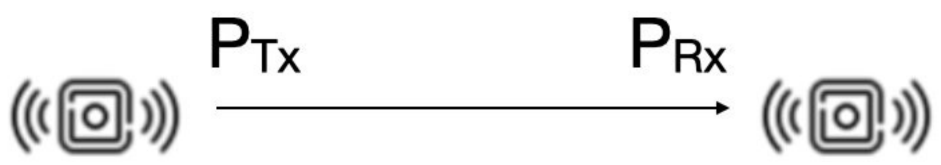
- 9 nodes on the Grenoble site, 24 hours deployment
- Communication technology: 802.15.4, on *Contiki-NG*
- Trafic: Broadcast, periodic (1 packet/s)

- **Objective: Reproduce and predict the evolution of the PDR (Packet Delivery Ratio)**

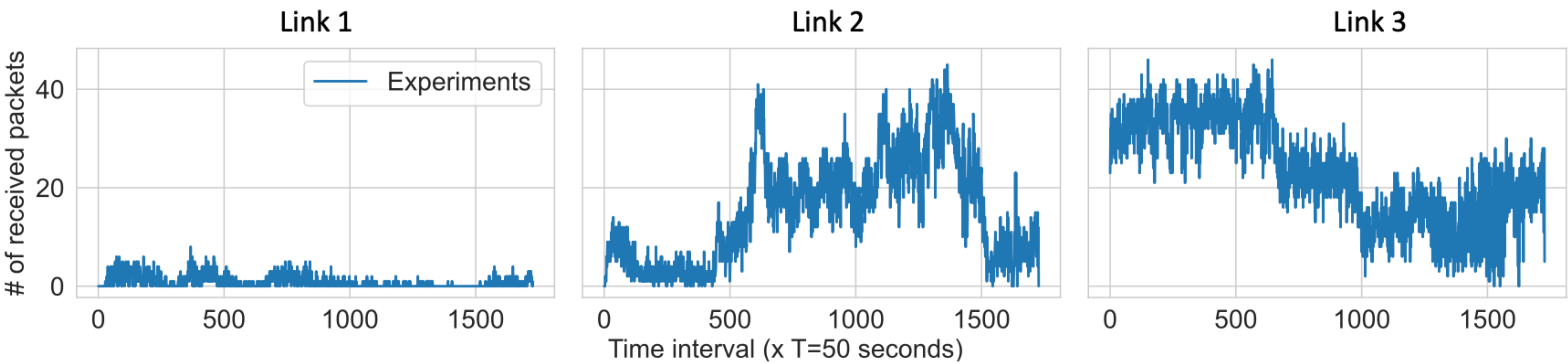
[1] Adjih, Cedric, et al. "FIT IoT-LAB: A large scale open experimental IoT testbed." 2015 IEEE 2nd World Forum on Internet of Things (WF-IoT). IEEE, 2015.

Problem Statement

❖ Compare between the PDR evolution:



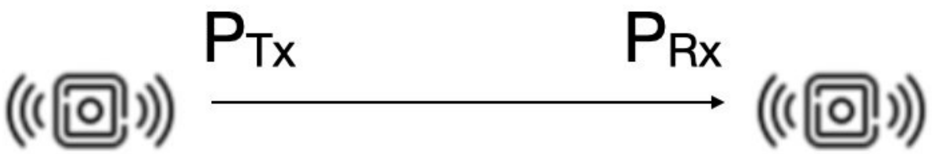
- From real measures
- From the Cooja simulator [2] :
  - Calibrated with an average PDR value per link (fixed)



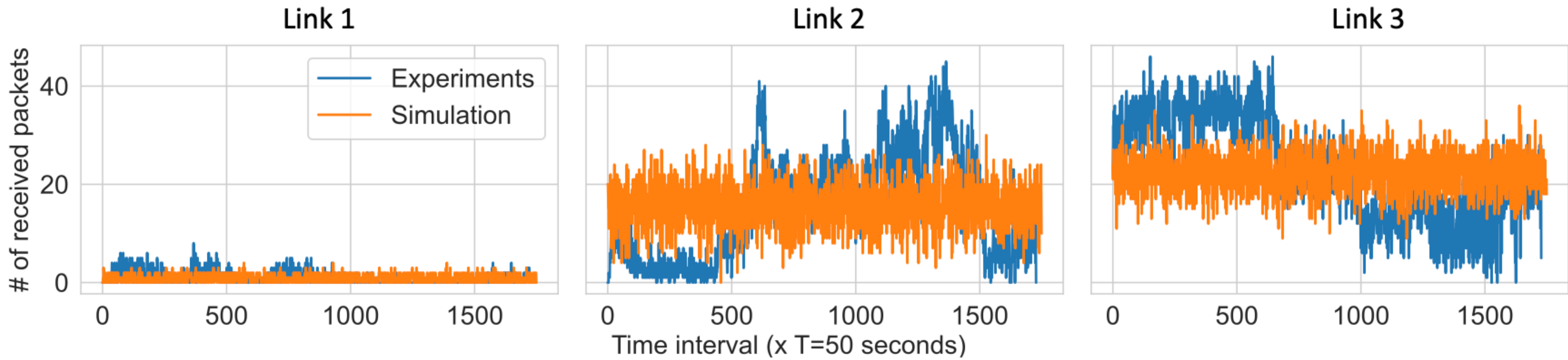
[2] Osterlind, Fredrik, et al. "Cross-level sensor network simulation with cooja." Proceedings. 2006 31st IEEE conference on local computer networks. IEEE, 2006.

Problem Statement

❖ Compare between the PDR evolution:



- From real measures
- From the Cooja simulator [2] :
  - Calibrated with an average PDR value per link (fixed)

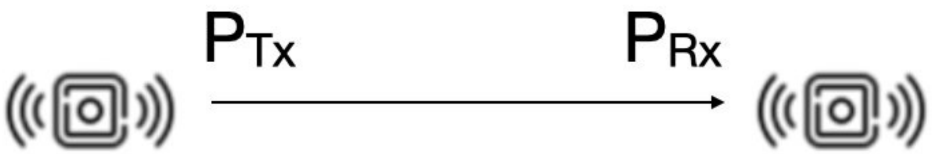


[2] Osterlind, Fredrik, et al. "Cross-level sensor network simulation with cooja." Proceedings. 2006 31st IEEE conference on local computer networks. IEEE, 2006.

Problem Statement

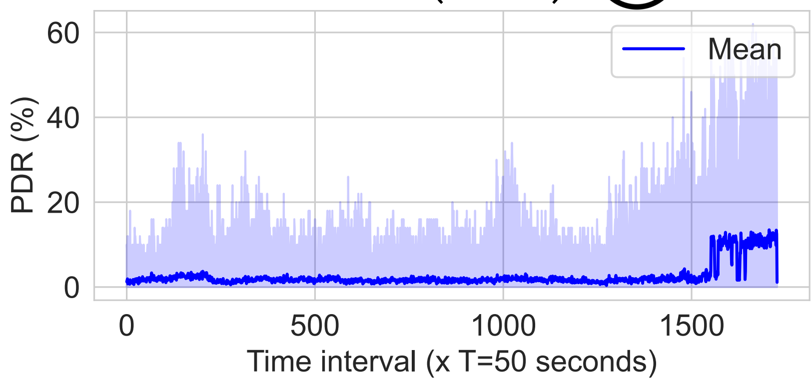
❖ Compare between the PDR evolution:

- From real measures
- From the Cooja simulator [2] :
  - Calibrated with an average PDR value per link (fixed)

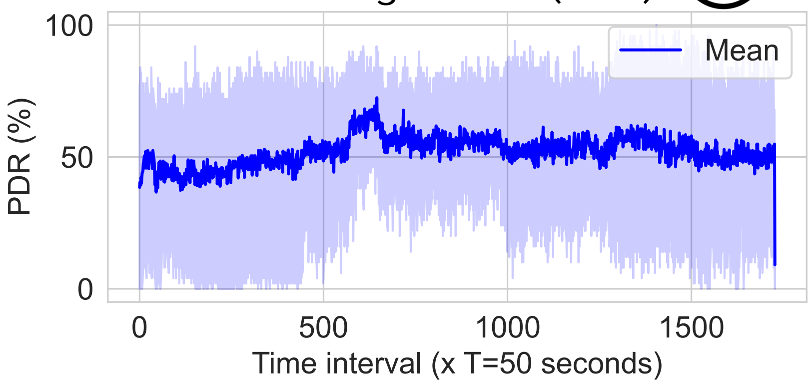


➢ **Classification according to the mean PDR: 0% - 20% - 60% - 75% - 100%**

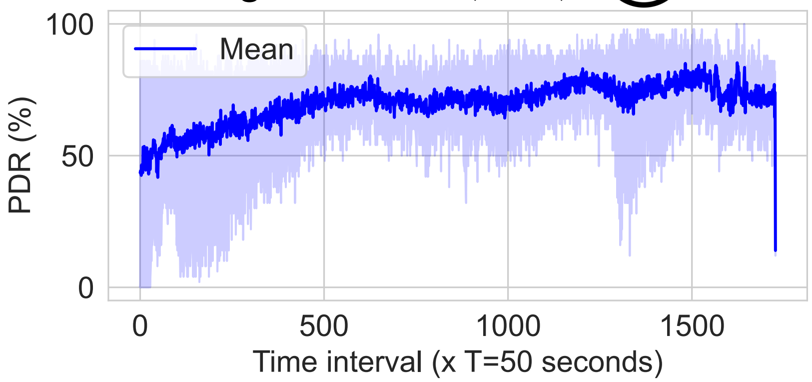
bad links (n=27)



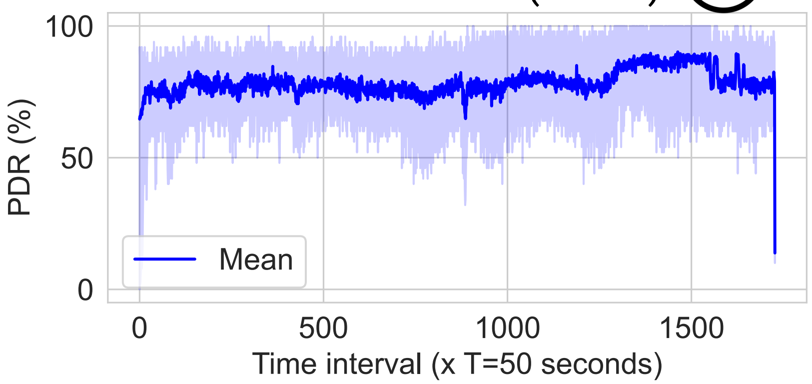
average links (n=6)



good links (n=8)



excellent links (n=12)

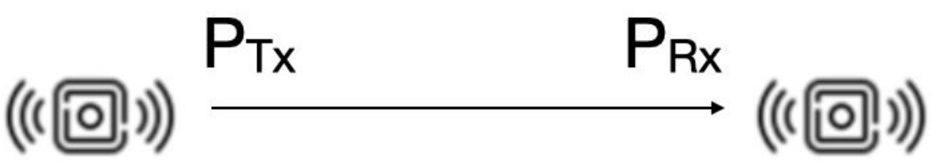


[2] Osterlind, Fredrik, et al. "Cross-level sensor network simulation with cooja." Proceedings. 2006 31st IEEE conference on local computer networks. IEEE, 2006.

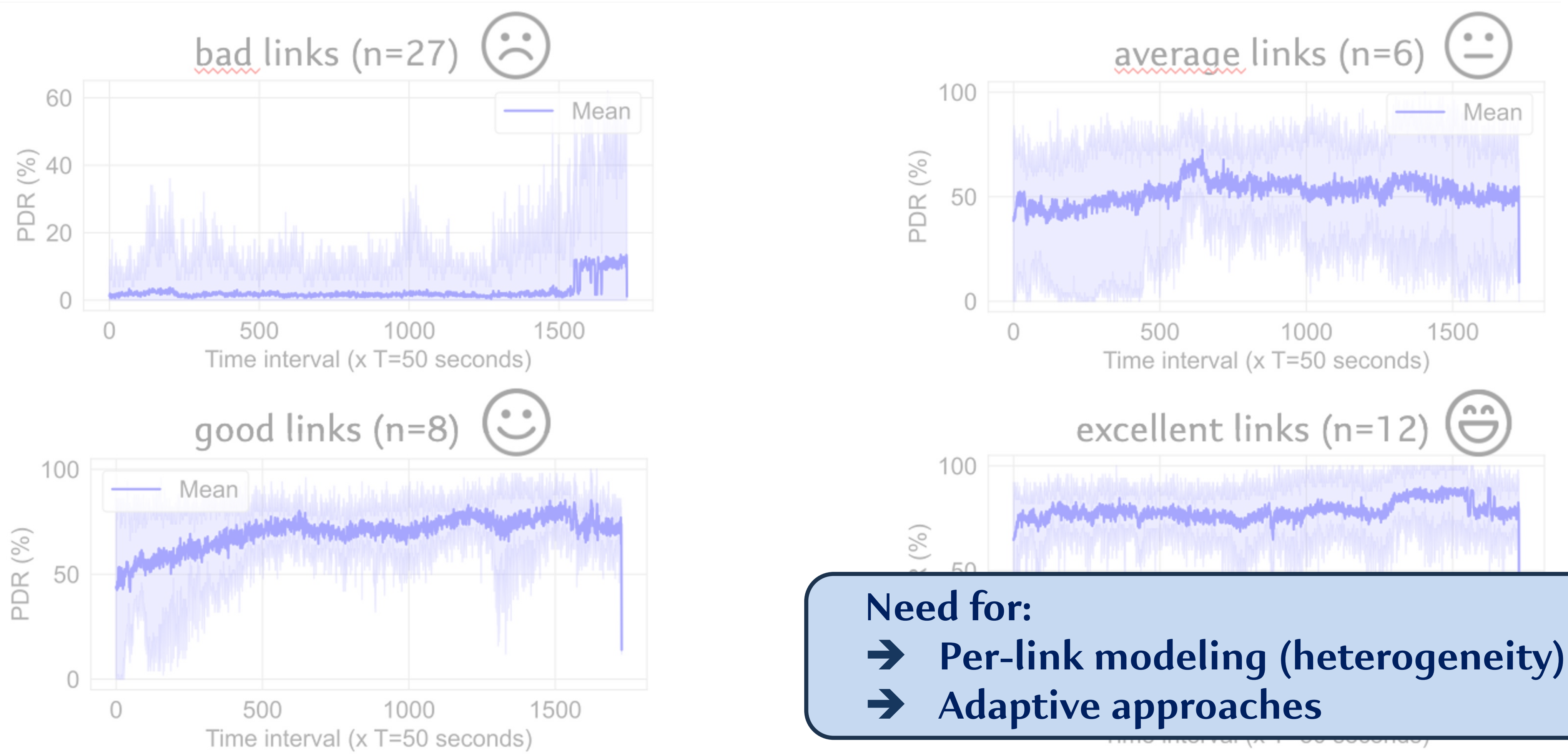
Problem Statement

❖ Compare between the PDR evolution:

- From real measures
- From the Cooja simulator [2] :
  - Calibrated with an average PDR value per link (fixed)



➢ **Classification according to the mean PDR: 0% - 20% - 60% - 75% - 100%**



Need for:
➔ Per-link modeling (heterogeneity)
➔ Adaptive approaches

[2] Osterlind, Fredrik, et al. "Cross-level sensor network simulation with cooja." Proceedings. 2006 31st IEEE conference on local computer networks. IEEE, 2006.

Problem Statement

❖ **How to accurately predict the evolution of a KPI (PDR) in a network?**

❖ **State of the Art:**

- **Ray-tracing** approaches [3];
- **Empirical studies** [4, 5] to generate propagation models;
- **ML-based heuristics** [6, 7] for predicting performances.

❖ **Limits:**

- Ignoring the links heterogeneity – **Modeling the network as a whole**
- Absence of adaptivity in the modeling.

[3] Ruah, C. et al., (2024). Calibrating wireless ray tracing for digital twinning using local phase error estimates. *IEEE Transactions on Machine Learning in Communications and Networking*.
[4] Brun-Laguna, K., Minet, P., Watteyne, T., & Gomes, P. H. (2019). Moving beyond testbeds? Lessons (we) learned about connectivity. IEEE Pervasive Computing, 17(4), 15-27.
[5] Baccour, N. et al., (2012). Radio link quality estimation in wireless sensor networks: A survey. ACM Transactions on Sensor Networks (TOSN), 8(4), 1-33.
[6] Sindjoung, M. L. F., & Minet, P. (2019, November). Wireless link quality prediction in iot networks. In 2019 PEMWN (pp. 1-6). IEEE.
[7] Benadji, H., Zitoune, L., & Vèque, V. (2023, December). Predictive modeling of loss ratio for congestion control in IoT networks using deep learning. In GLOBECOM 2023. IEEE.

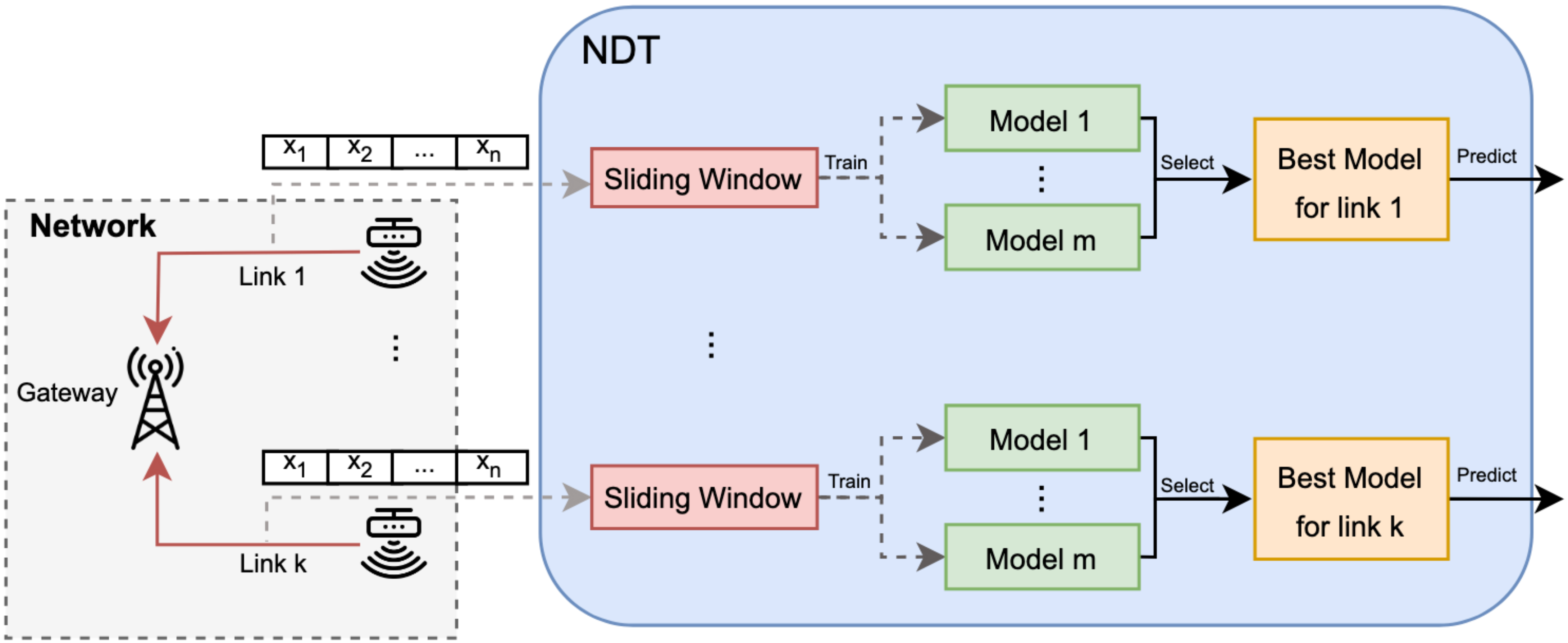
Proposed Solution

❖ **Contribution: Data-driven approach**

1. Replace the **physical layer** simulation by measurements
2. **Separately** model each radio link
3. **Continuously** select the best model for each link

❖ **Interest:**

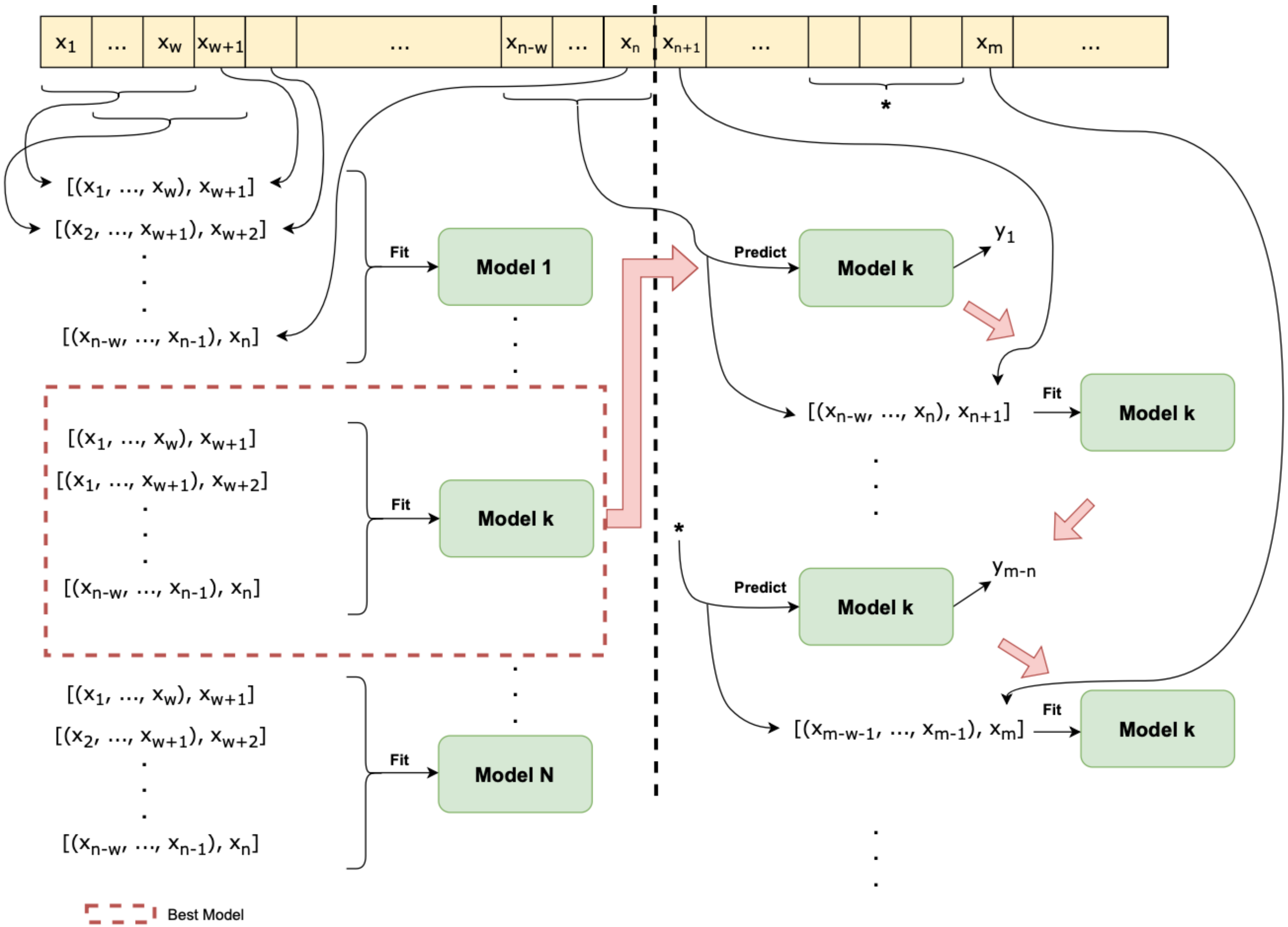
- Flexibility
- Temporality (anomaly Detection, etc.)



Proposed Solution – One Step

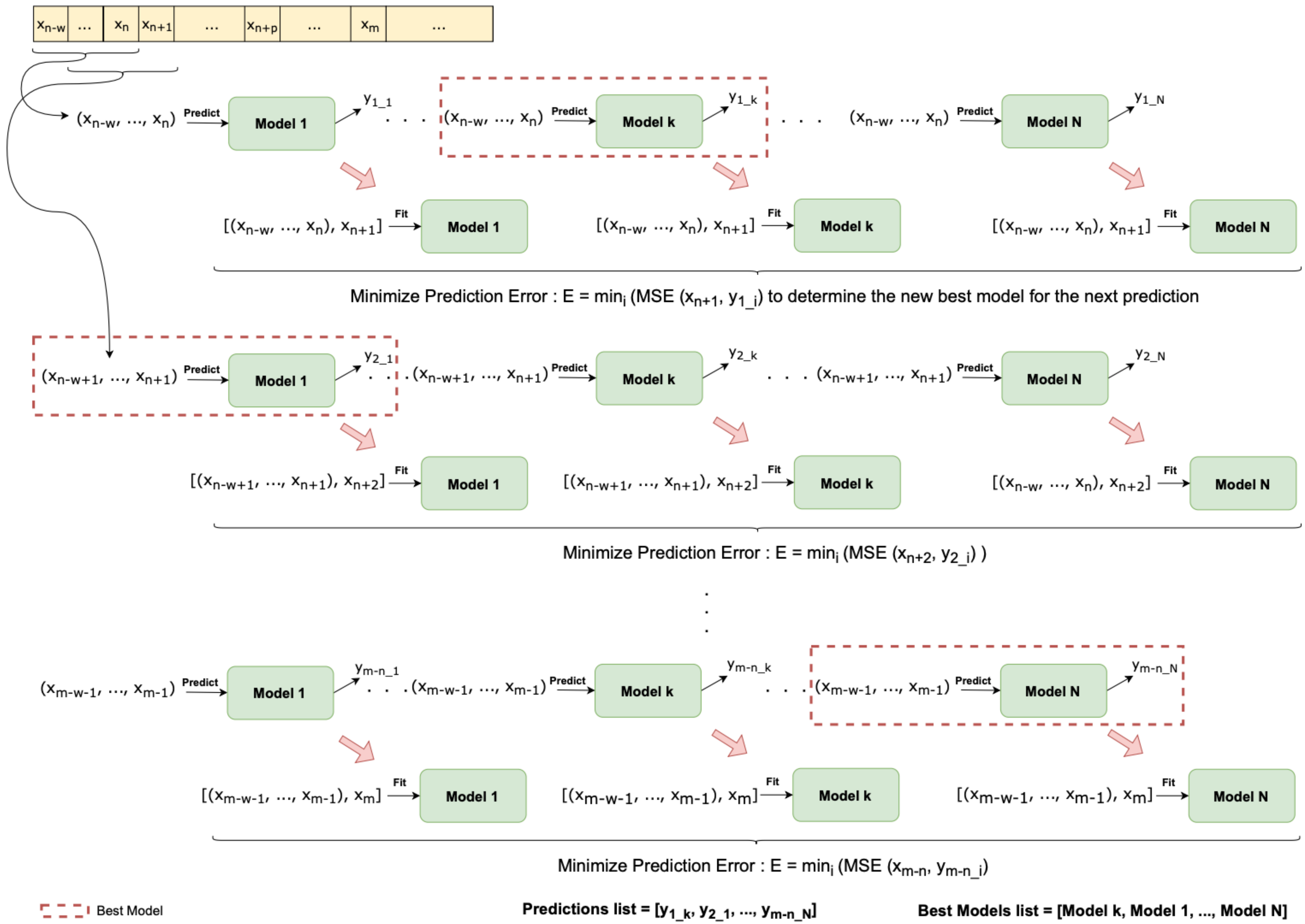
❖ **Fixed Model:** For each model:

- Split between Training and Testing data
- Create sliding windows for the regression
- Select the best model to use for predictions



Proposed Solution – One Step

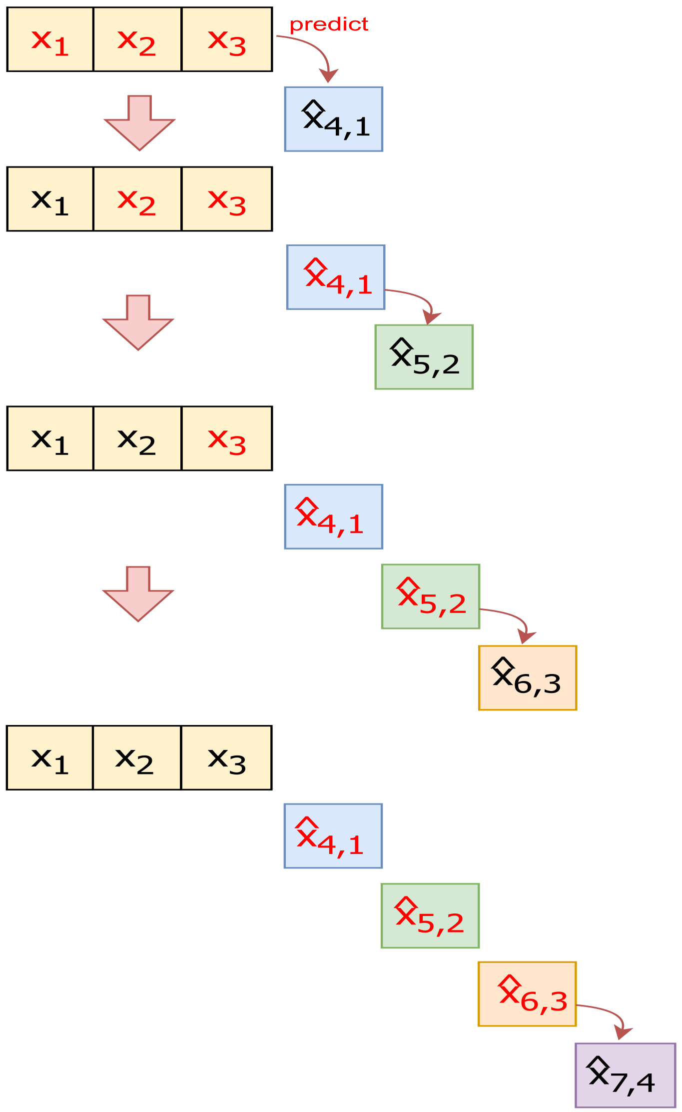
❖ Adaptive Model:



Proposed Solution – Multi-Step

❖ Multi-step prediction:

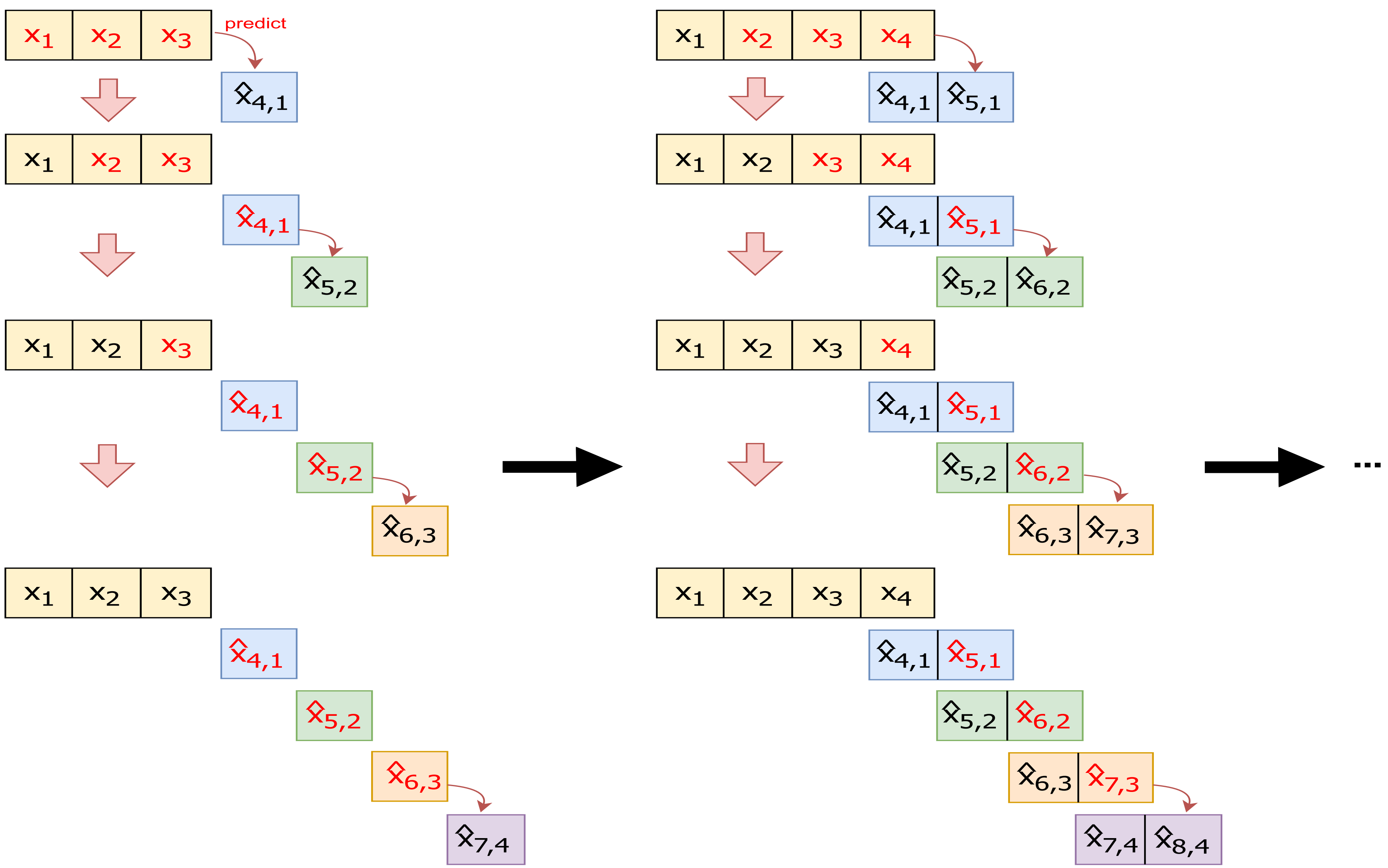
- Distinguish between the steps predictions for the evaluation



Proposed Solution – Multi-Step

❖ Multi-step prediction:

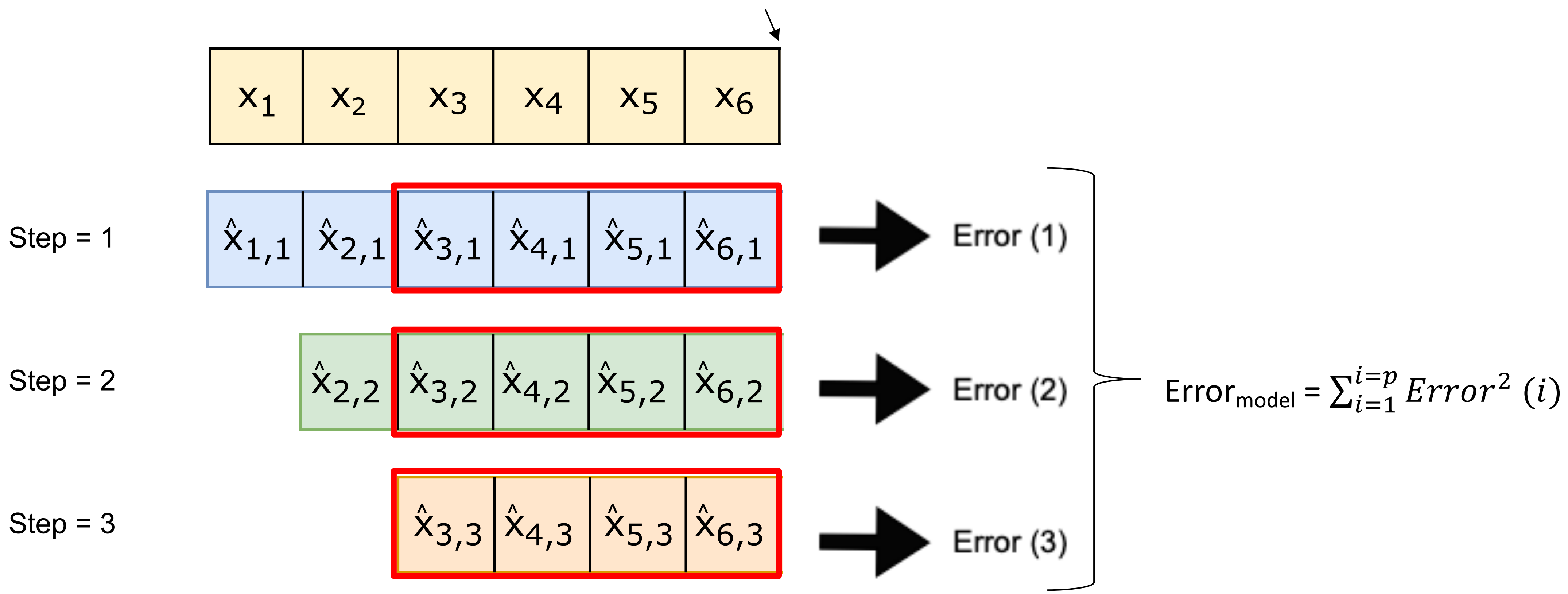
- Distinguish between the steps predictions for the evaluation



Proposed Solution – Multi-Step

❖ Error computing for the Adaptive approach:

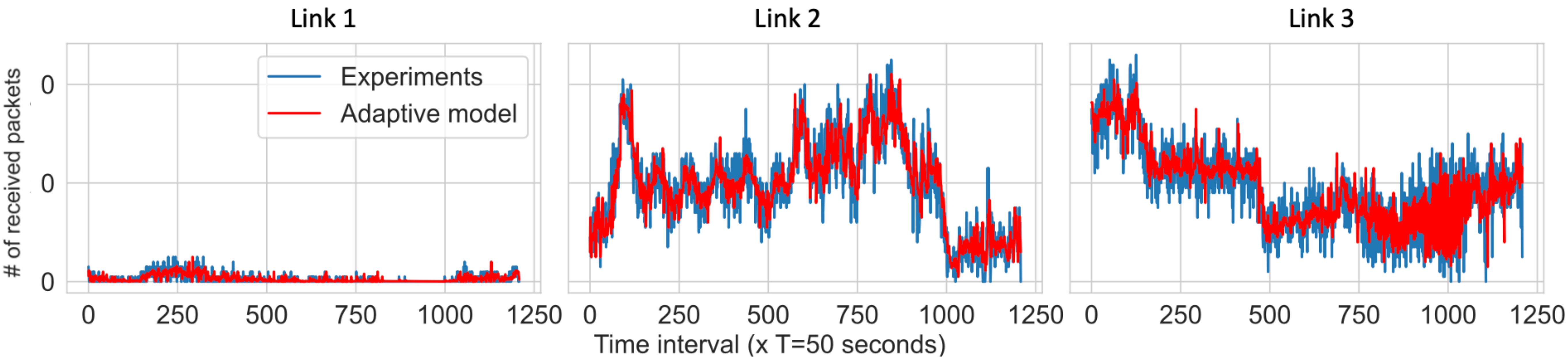
- At each interval, compute the error for the last *q (e.g., q*=4) predictions, according to the different steps errors of each model:



❖ **Choose Best Model** = $\boldsymbol{argmin}\ (\boldsymbol{Error_{model}})$ for the next prediction

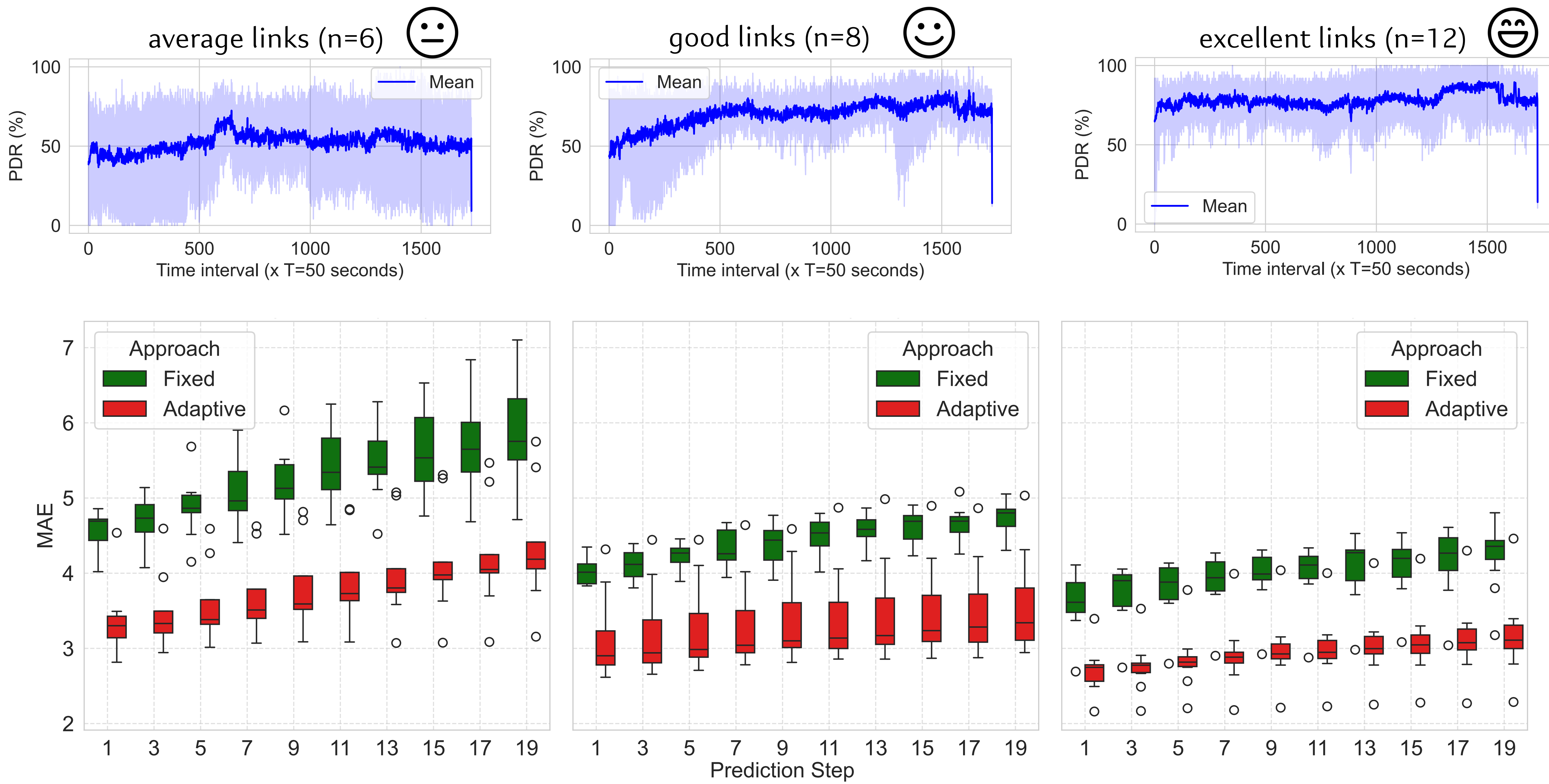
Results

- **Short-term predictions (p=1):**



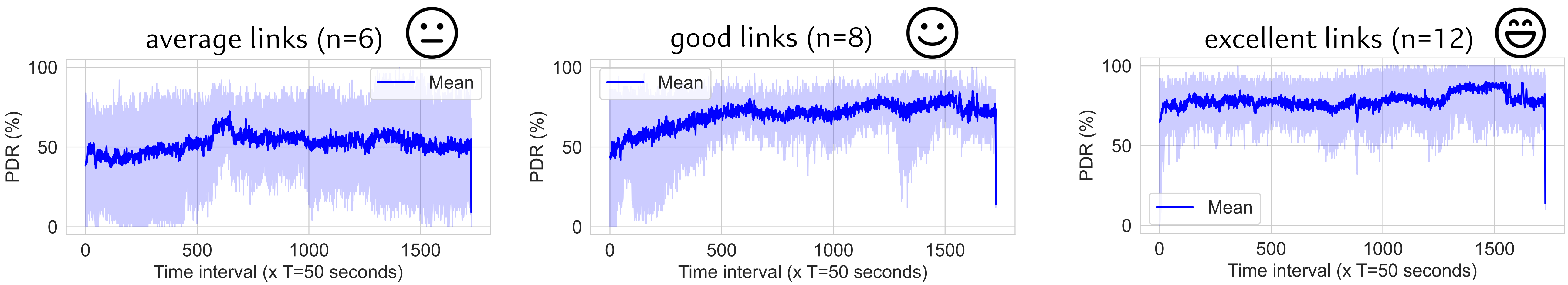
Results

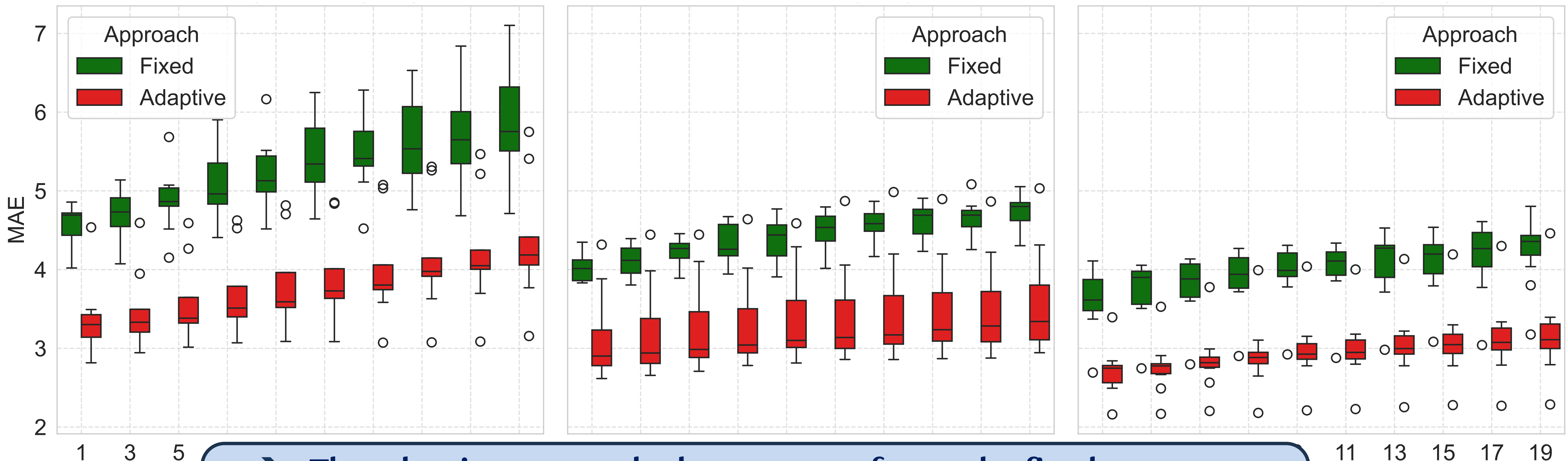
❖ Mid-term predictions (p>1):



Results

❖ Mid-term predictions (p>1):

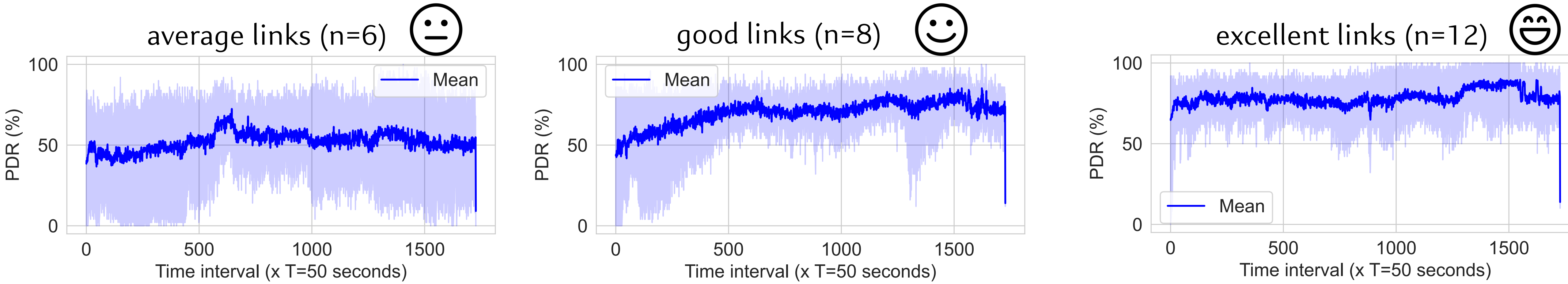


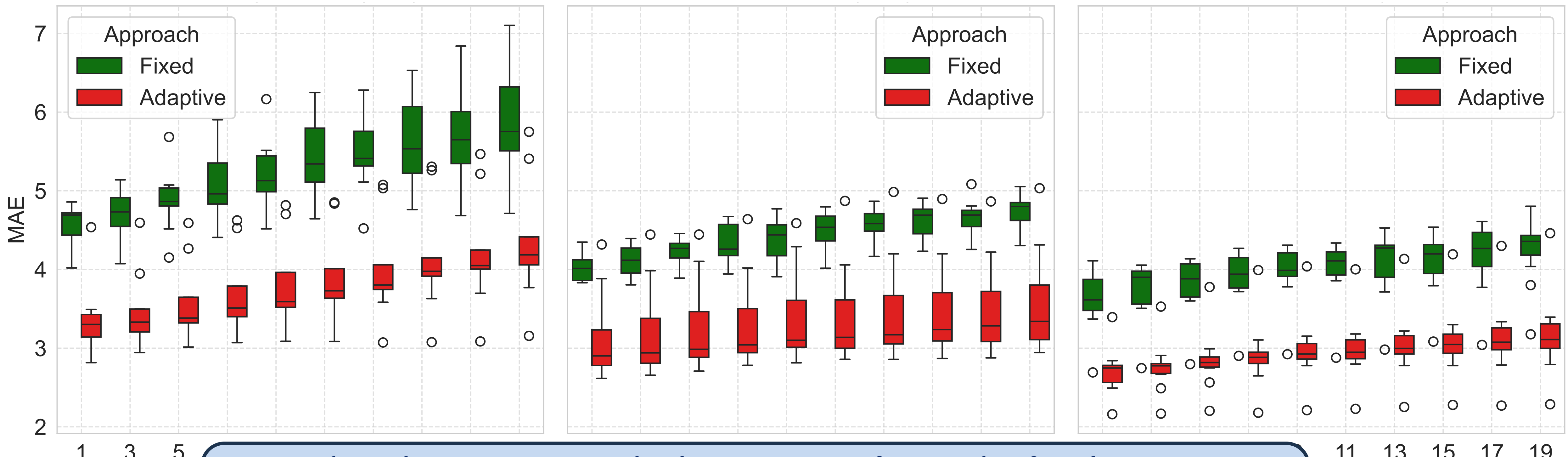


➔ **The adaptive approach always outperforms the fixed one**

Results

❖ Mid-term predictions (p>1):

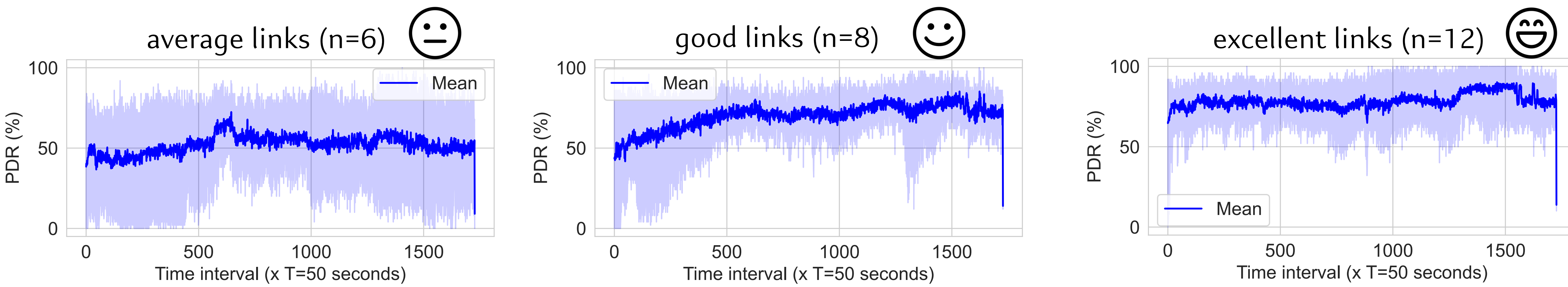


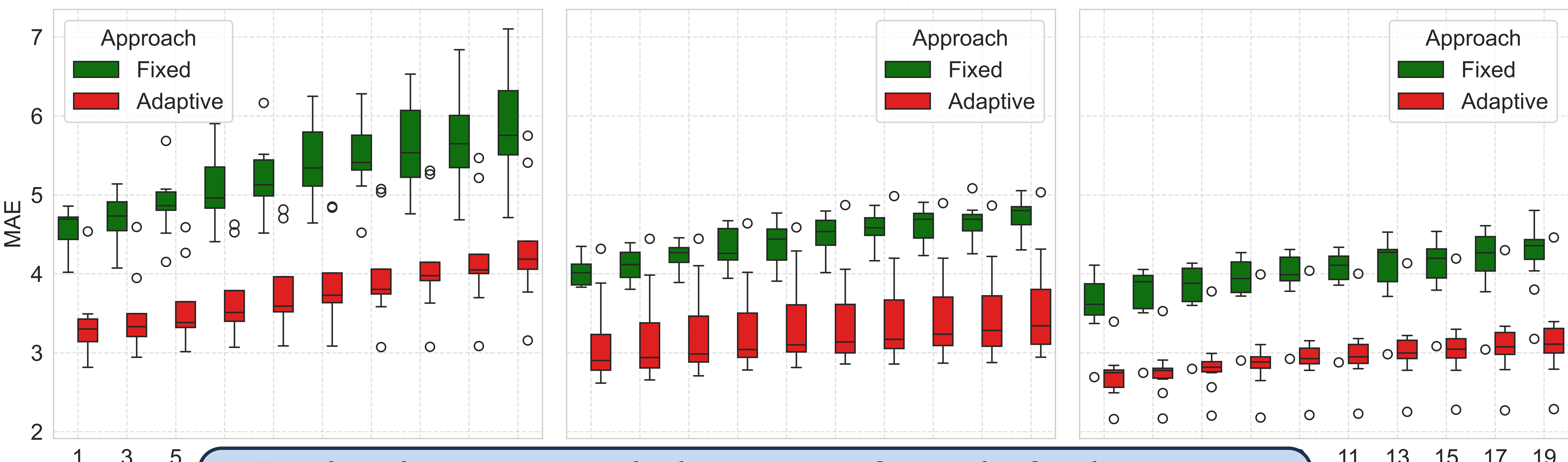


- ➔ **The adaptive approach always outperforms the fixed one**
- ➔ **Low prediction errors for short/mid term, with a linear increasing**

Results

❖ Mid-term predictions (p>1):

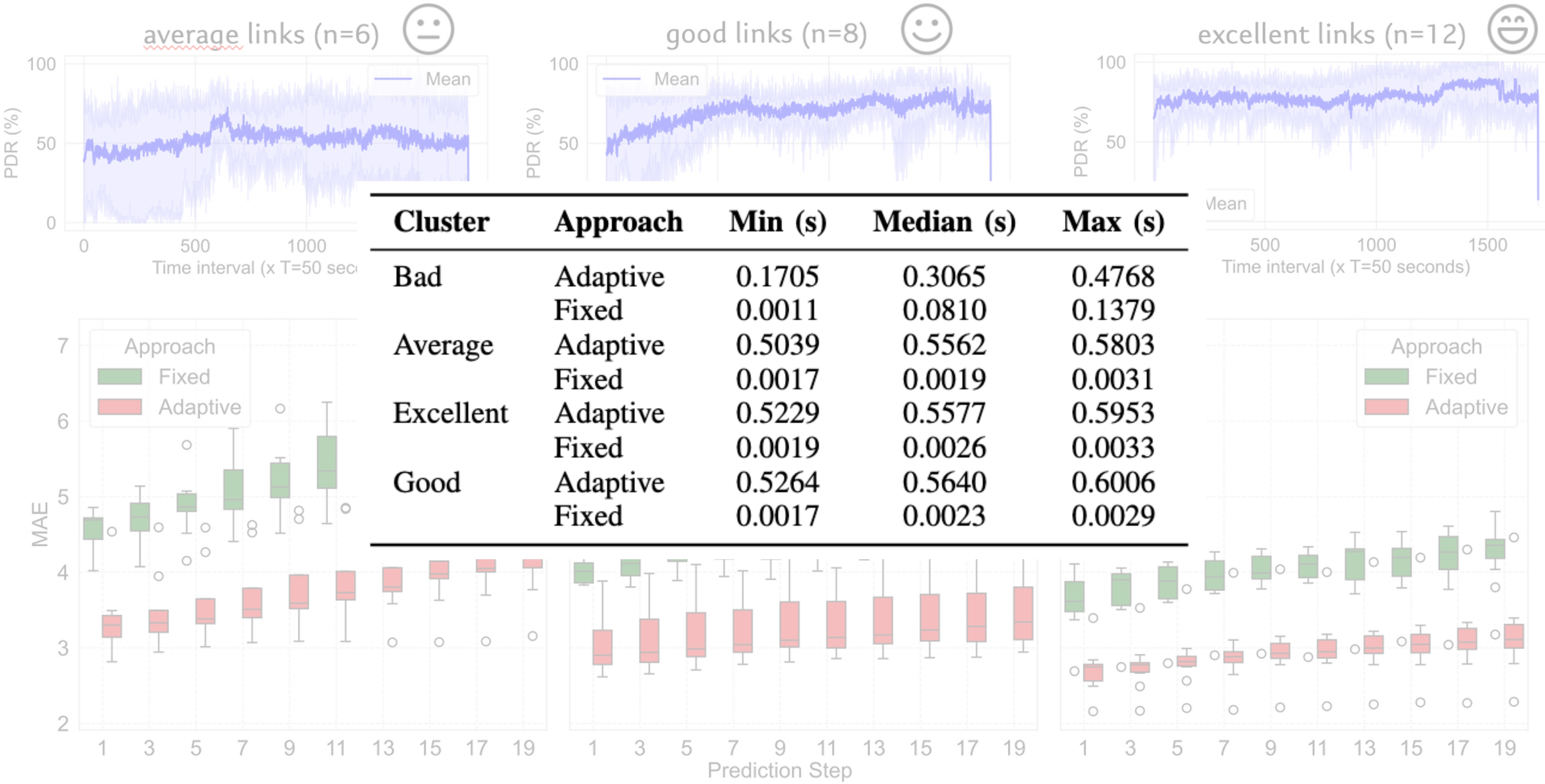




- ➔ **The adaptive approach always outperforms the fixed one**
- ➔ **Low prediction errors for short/mid term, with a linear increasing**
- ➔ **Better predictions for stable links**

Results

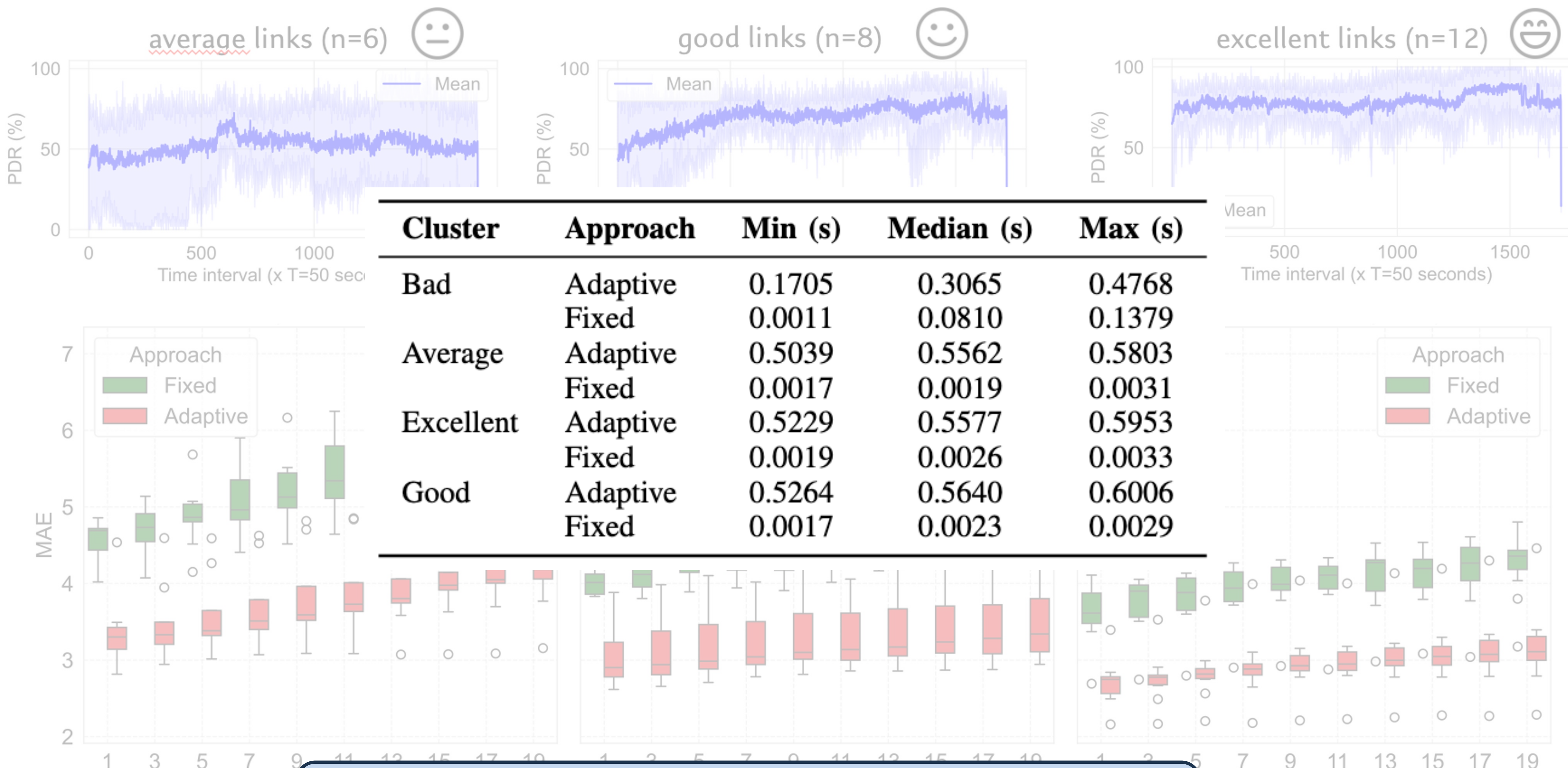
❖ Mid-term predictions (p>1):



Cluster	Approach	Min (s)	Median (s)	Max (s)
Bad	Adaptive	0.1705	0.3065	0.4768
	Fixed	0.0011	0.0810	0.1379
Average	Adaptive	0.5039	0.5562	0.5803
	Fixed	0.0017	0.0019	0.0031
Excellent	Adaptive	0.5229	0.5577	0.5953
	Fixed	0.0019	0.0026	0.0033
Good	Adaptive	0.5264	0.5640	0.6006
	Fixed	0.0017	0.0023	0.0029

Results

❖ Mid-term predictions (p>1):

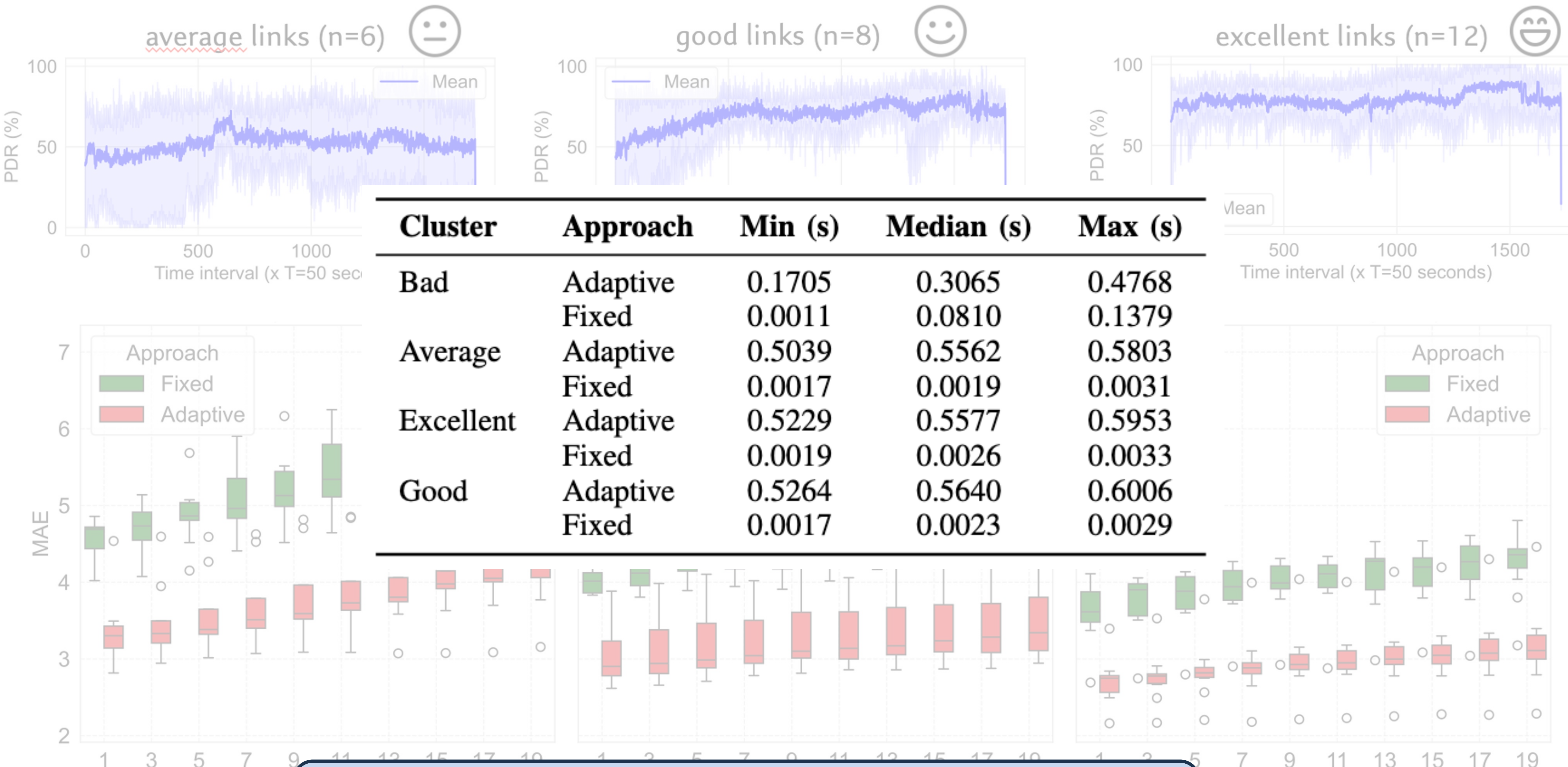


Cluster	Approach	Min (s)	Median (s)	Max (s)
Bad	Adaptive	0.1705	0.3065	0.4768
	Fixed	0.0011	0.0810	0.1379
Average	Adaptive	0.5039	0.5562	0.5803
	Fixed	0.0017	0.0019	0.0031
Excellent	Adaptive	0.5229	0.5577	0.5953
	Fixed	0.0019	0.0026	0.0033
Good	Adaptive	0.5264	0.5640	0.6006
	Fixed	0.0017	0.0023	0.0029

➔ Low prediction times for both approaches

Results

❖ Mid-term predictions (p>1):



Cluster	Approach	Min (s)	Median (s)	Max (s)
Bad	Adaptive	0.1705	0.3065	0.4768
	Fixed	0.0011	0.0810	0.1379
Average	Adaptive	0.5039	0.5562	0.5803
	Fixed	0.0017	0.0019	0.0031
Excellent	Adaptive	0.5229	0.5577	0.5953
	Fixed	0.0019	0.0026	0.0033
Good	Adaptive	0.5264	0.5640	0.6006
	Fixed	0.0017	0.0023	0.0029

➔ **Low prediction times for both approaches**
➔ **Accuracy comes at a cost...**

Contributions

- **Individual link modeling**
  - **Captures radio links heterogeneity**
  - **Uses lightweight models**

- **Adaptive algorithm for model selection**
  - **Captures radio links dynamics**
  - **Showcases tradeoff between accuracy and complexity**

- **Completely reproducible experiments and results:** https://github.com/SamirSim/Wireless-Link-Quality-Prediction

Takeaways & Perspectives

Takeaways:

- Importance of deployment conditions
  - **Need for reproducible experimental data**
- Network heterogeneity and dynamics
  - **Need for per-link modeling in network performance evaluation**
  - **Need for adaptive approaches**

Perspectives:

- Metrology with a dynamic traffic?
- Scalability?
- Generalization capabilities?

Takeaways & Perspectives

Takeaways:

- Importance of deployment conditions
  - **Need for reproducible experimental data**
- Network heterogeneity and dynamics
  - **Need for per-link modeling in network performance evaluation**
  - **Need for adaptive approaches**

Perspectives:

- Metrology with a dynamic traffic?
- Scalability?
- Generalization capabilities?

Thank you for your attention!