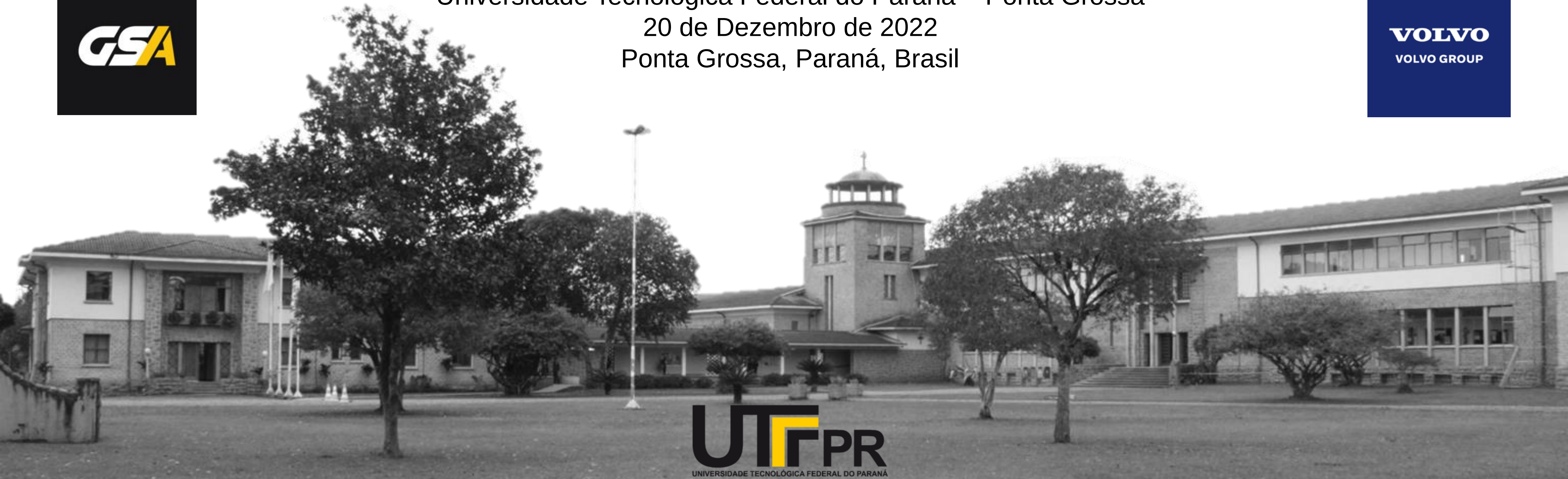
DESIGN AND DEVELOPMENT OF A MODEL PREDICTIVE CONTROL SYSTEM FOR THE THERMAL REGULATION OF THE REFRIGERANT FLUID

Master's Student: João Eduardo Hoffmann
Advisor: Max Mauro Dias Santos

Programa de Pós-Graduação em Engenharia Elétrica
Universidade Tecnológica Federal do Paraná – Ponta Grossa

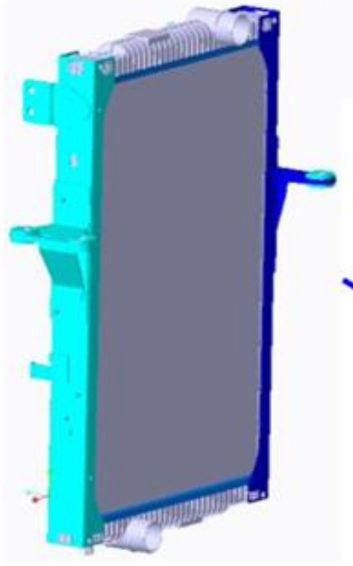
20 de Dezembro de 2022
Ponta Grossa, Paraná, Brasil

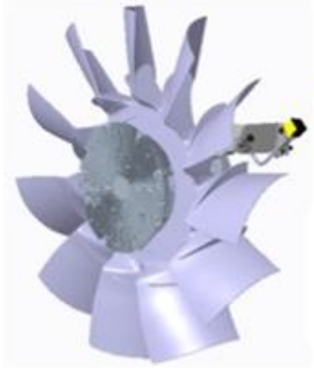


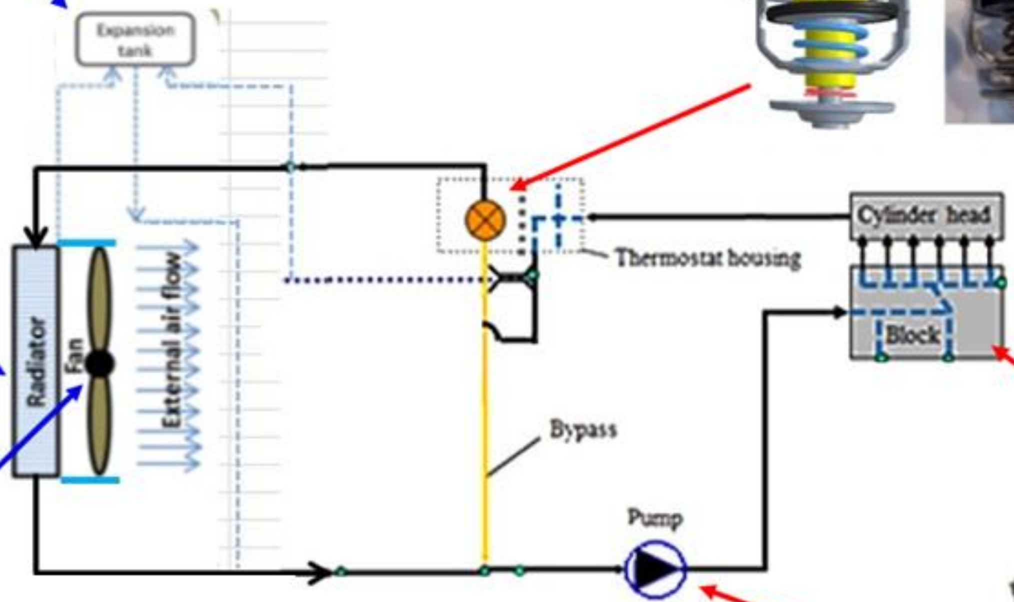
UTFPR
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

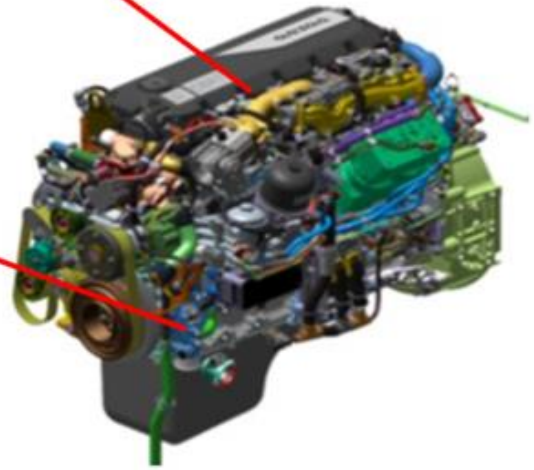
1. Introduction

Cooling system circuit (coolant side only)





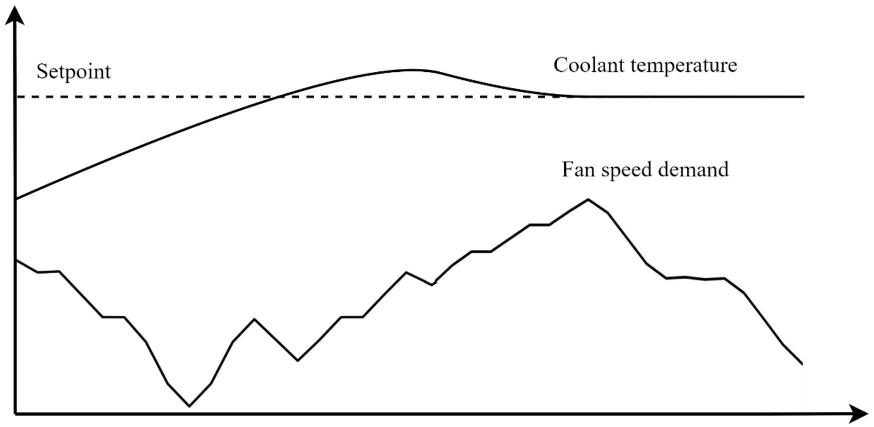




HDE11 Euro6 / US10 Engine output	~90kW
Oil pump absorbed power	3kW
Coolant pump absorbed power	1kW
Fan absorbed power	20W *(out of 15kW if fully engaged)*

1.2. Coolant temperature regulation

- **How:** Performed by controlling the coolant flow and temperature to achieve a high driveline overall efficiency.
  - Fan controller regulation aims to maintain the coolant temperature at a defined setpoint.

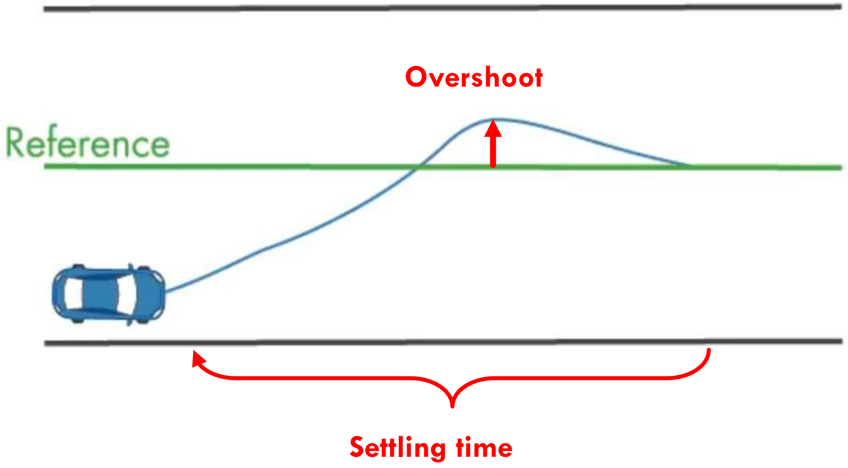


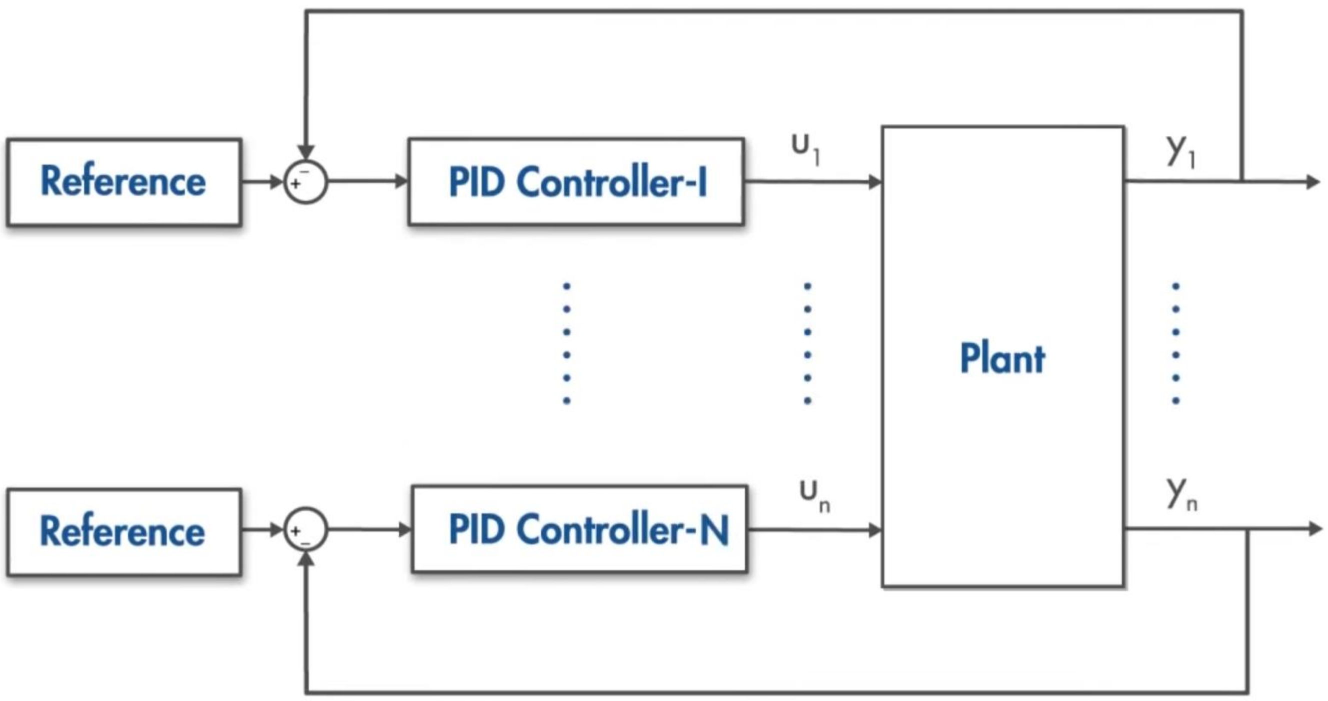
Fuel consumption reduction

Power/torque increase

Emissions reduction

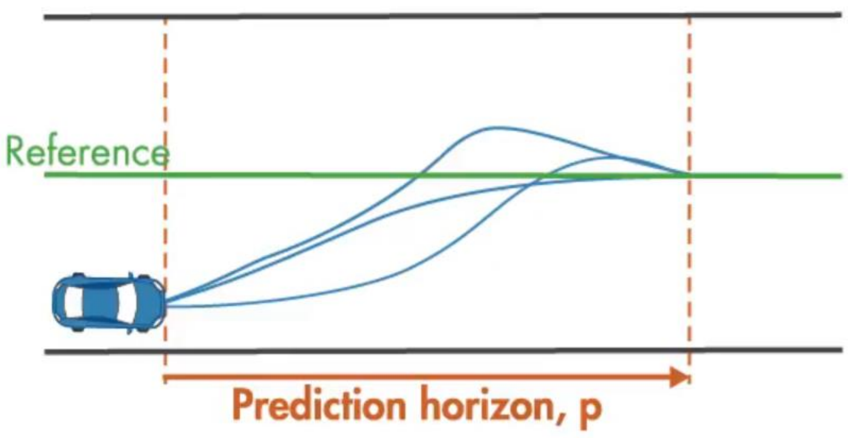
1.3. Control and optimization systems

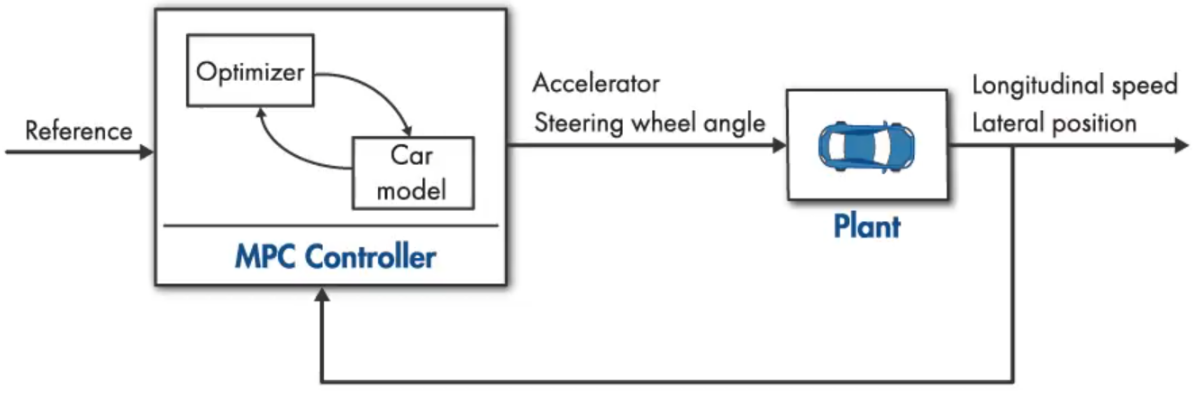


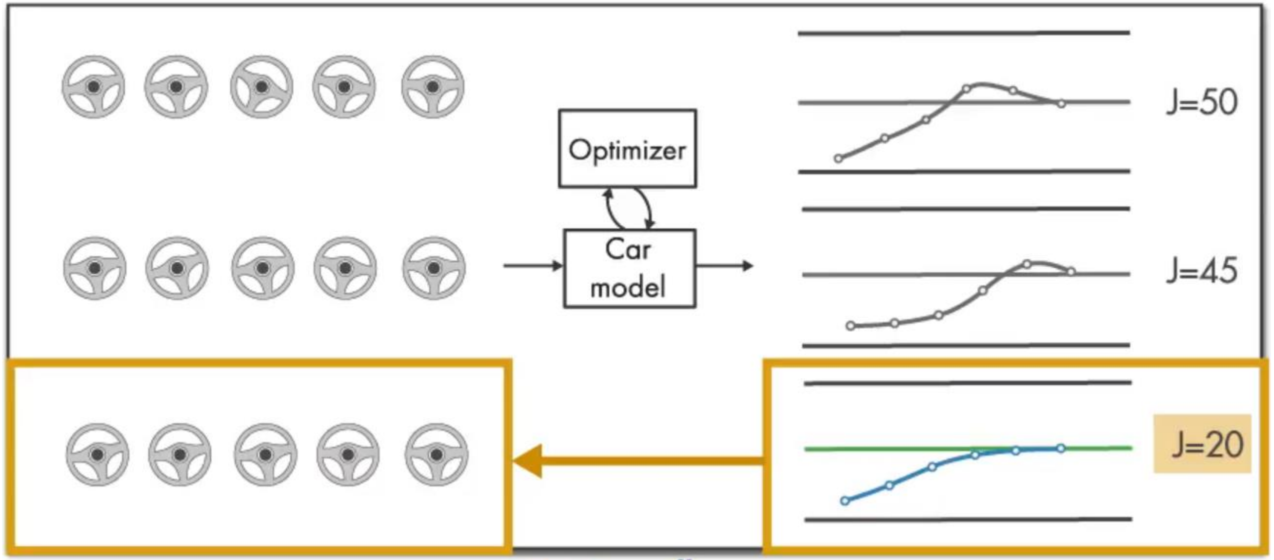


- **Reactive**
- **Constraints**
- **MIMO**

1.3. Control and optimization systems



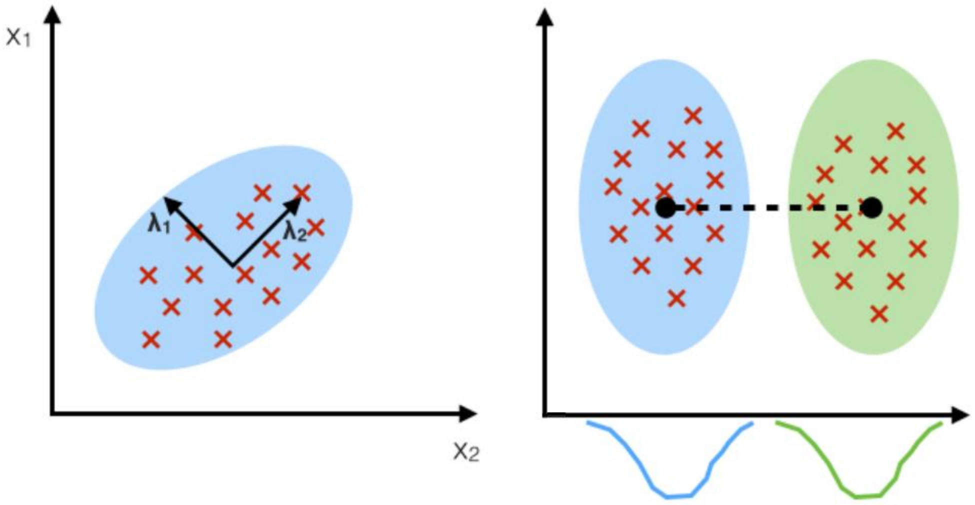


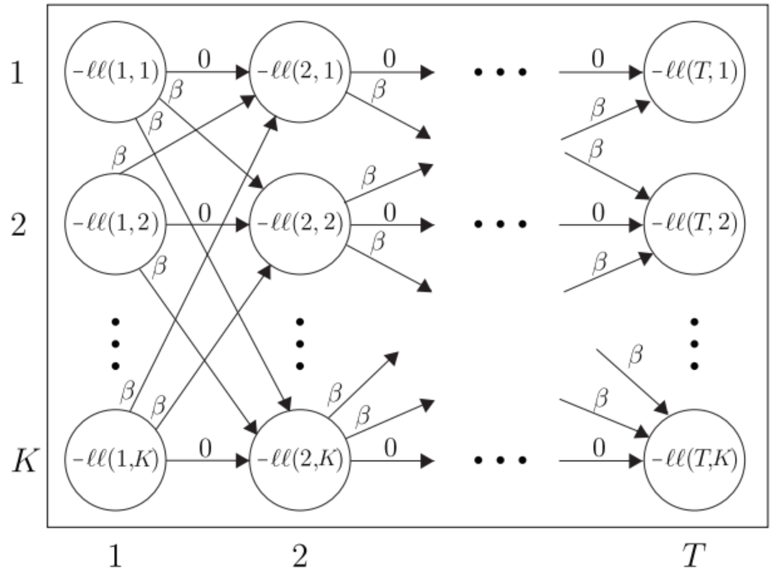


- Improved accuracy
- Constraints
- Heavy process

1.4. Feature extraction

- Extraction of relevant patterns;
- Reduction to the appropriate dimension:
  - PCA, LDA, clustering.
- Maintenance of sequential relevance:
  - Graph-based clustering.
- Dynamic and visually supervised optimization (Human-interpretability):
  - TICC.





2. Justification

- Although classic control systems present light and fast control processes, their regulation accuracy and complexity are negatively affected as proportional to the dimension of a given dynamic system.

- Such is the case of the vehicle and thermal dynamics systems, impacted by multiple internal and external factors that present a high-dimensional correlation towards different thermal rejection states for the vehicle's engine.

- Advances in ML and intelligent systems have shown that system identification, dimensionality reduction and optimizations can be achieved for highly non-linear systems while establishing high-level settings for a regulation objective.

3. Objectives

- The main objective of this work is to assess an MPC system proposition for the thermal regulation of the cooling system.

- The project targets 3 specific objectives to accomplish this goal:
  - Define the methods and strategies for the control and optimization systems;
  - Design and develop the ML models and architectures;
  - Evaluate the MPC system.

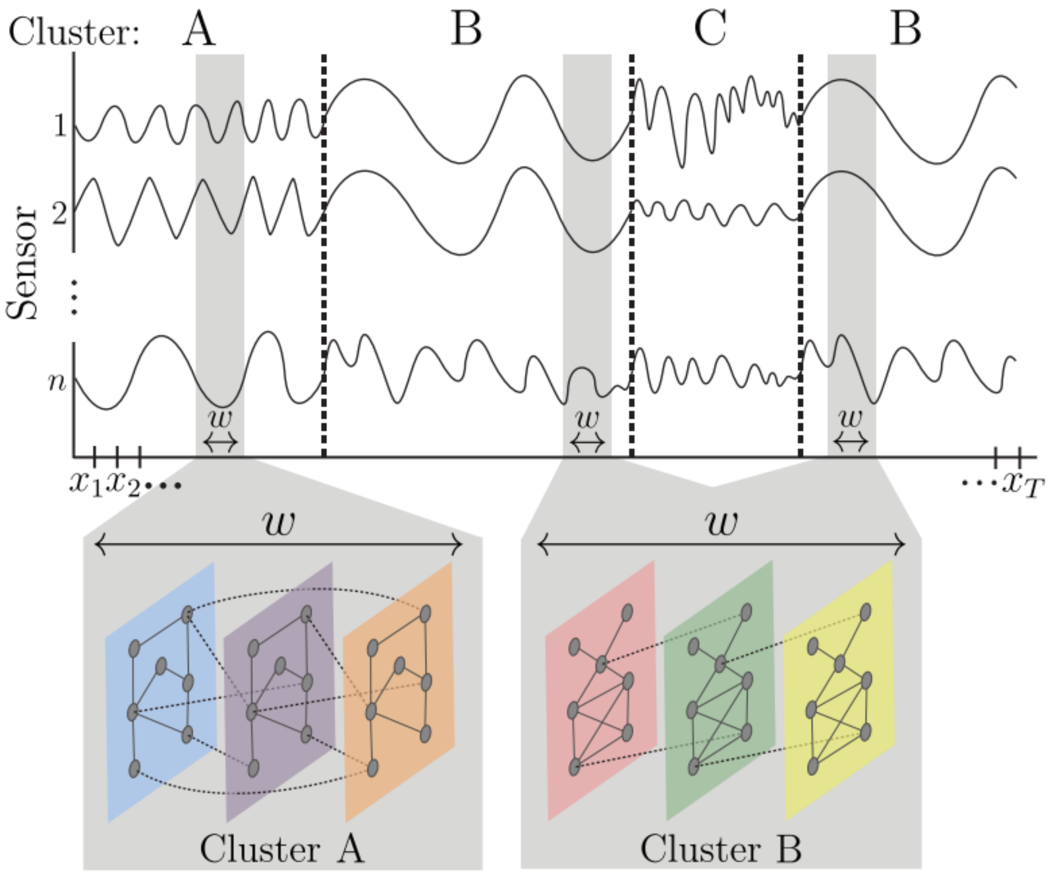
4. Feature extraction

4.1. Toeplitz Inverse Covariance-based Clustering (TICC)

- Clustering of high dimensional and temporal thermal impact data
- Simple and robust configuration:
  - (k , w , λ , β)

$$\underset{\Theta \in \mathcal{T}, \mathrm{P}}{\operatorname{argmin}} \sum_{i=1}^{K} \left[\overbrace{\|\lambda \circ \Theta_i\|_1}^{\text{sparsity}} + \sum_{X_t \in P_i} \left(\overbrace{-\ell\ell(X_t, \Theta_i)}^{\text{log likelihood}} + \overbrace{\beta \mathbb{1}\{X_{t-1} \notin P_i\}}^{\text{temporal consistency}} \right) \right] \quad (1)$$

- Human interpretable cluster definition
- Extraction of outputs as nonlinear "distances" to each defined cluster



D. Hallac, S. Vare, S. P. Boyd, and J. Leskovec, "Toeplitz inverse covariance-based clustering of multivariate time series data," CoRR, vol. abs/1706.03161, 2017. [Online]. Available: http://arxiv.org/abs/1706.03161

4.2. Human-Interpretable Feature Extraction

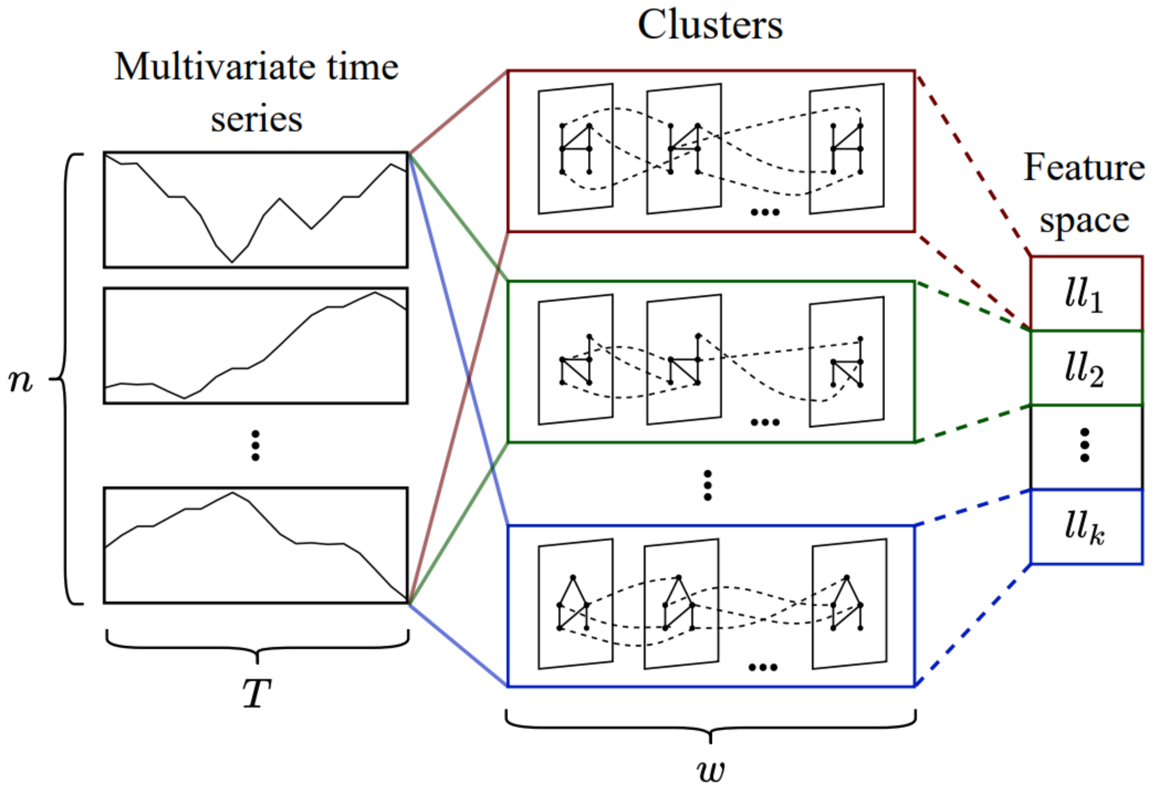
Algorithm 1 Extract the Feature Space

---

given X = time series of T sequential observations, μ_i = i-th cluster's empirical mean, Θ_i = i-th cluster's inverse covariance matrix.
initialize LL = list of K zeros.
for $i = 1, \ldots, K$ **do**
  x = $X - \mu_i$
  LD = $\log(\det(\Theta_i))$
  LL[i] = $\mathrm{x}^T \odot (\Theta_i \odot \mathrm{x})$ + LD
return LL

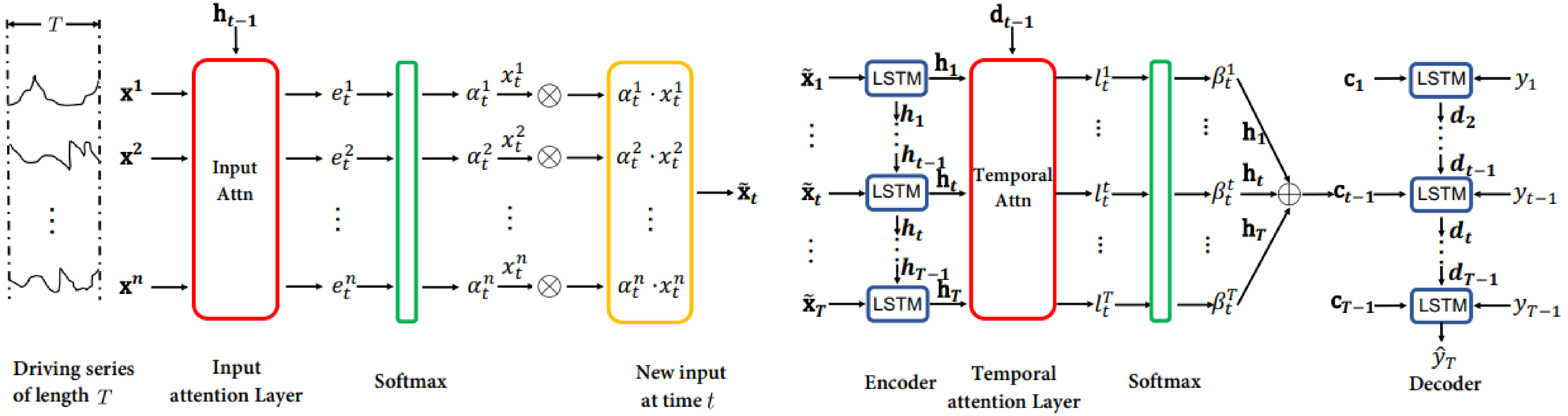
Source: Own Authorship.

Figure 6 – Human-interpretable feature extraction method.



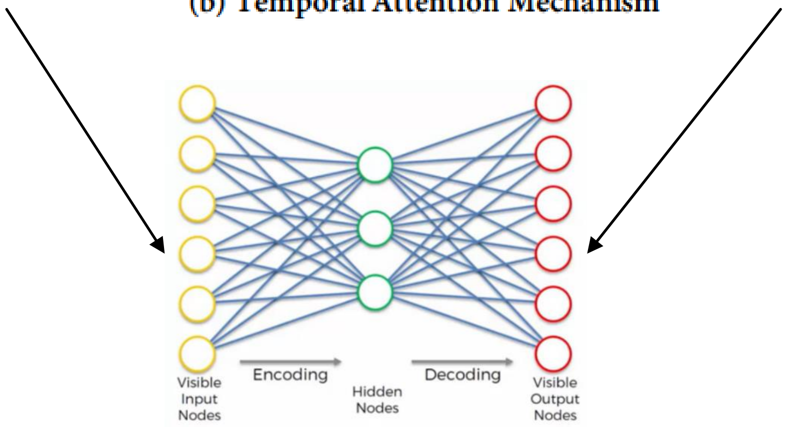
Source: Own Authorship.

4.3. NARX and DA-RNN



(a) Input Attention Mechanism (b) Temporal Attention Mechanism

Y. Qin, D. Song, H. Chen, W. Cheng, G. Jiang, and G. W. Cottrell, “A dual-stage attention-based recurrent neural network for time series prediction,” CoRR, vol. abs/1704.02971, 2017. [Online]. Available: http://arxiv.org/abs/1704.02971



4.4. Evaluation

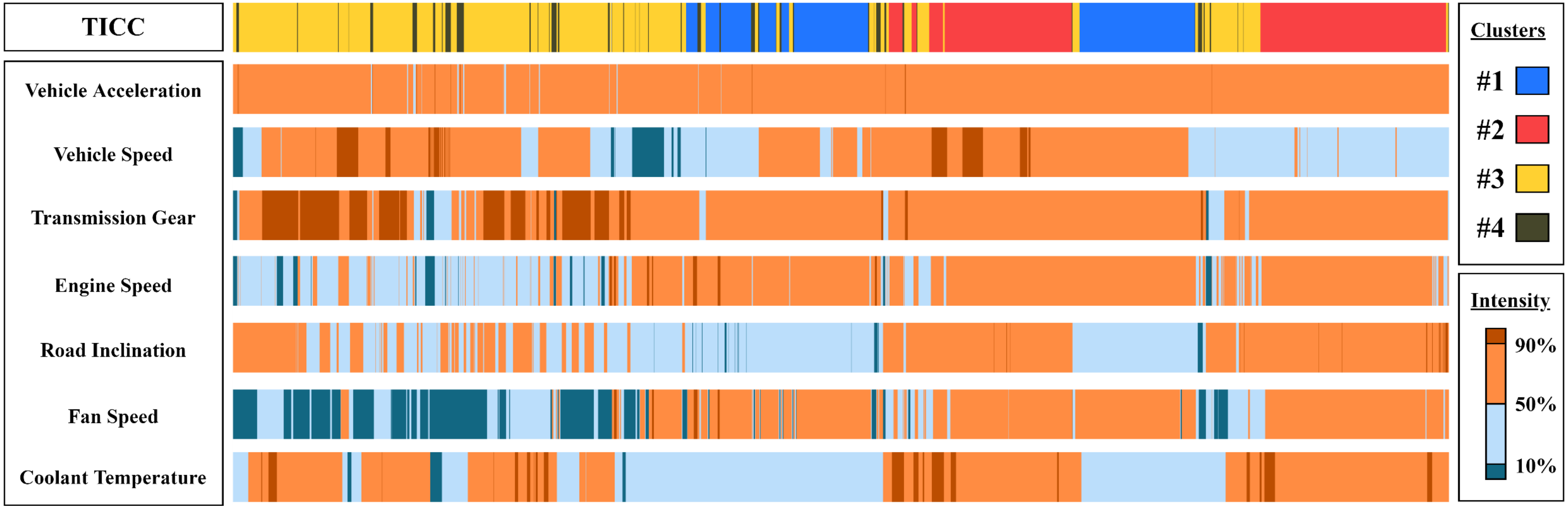
- Comparison between accuracy and number of parameters of:
  - Derived: ML model with the proposed feature space as input
  - Custom: ML model with the original multivariate time series as input

- Accuracy metrics:

$$\mathbf{MAE} = \frac{1}{N} \sum_{i=1}^{N} \left| y_t^i - \hat{y}_t^i \right| \qquad \mathbf{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(y_t^i - \hat{y}_t^i \right)^2}$$

4.4.

Figure 12 – Outputs of the TICC visualization algorithm.



Source: Own Authorship.

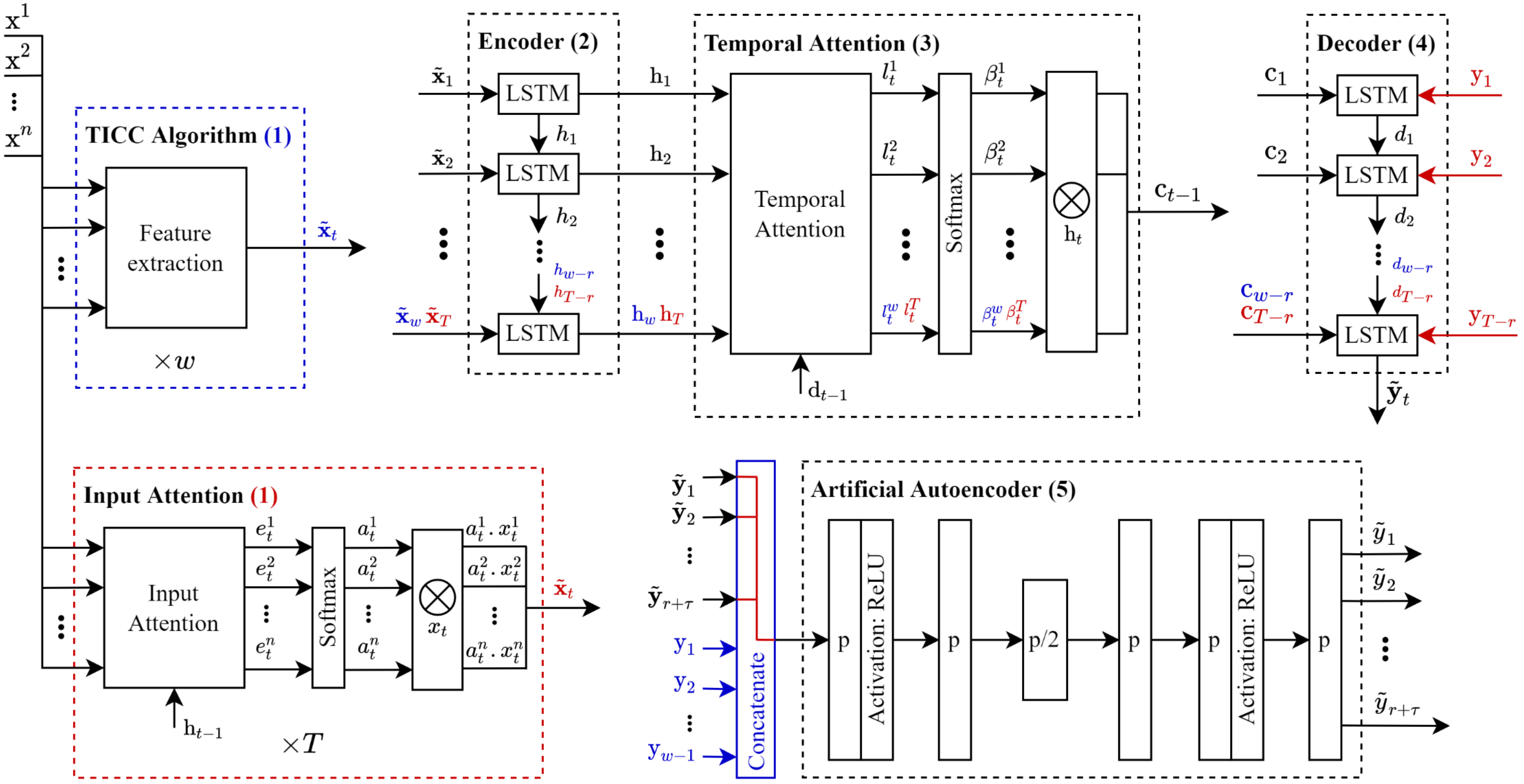
Table 3 – Importance matrix.

Clusters	Acc.	Spd.	Gear	Eng. Spd.	Road Incl.	Fan Spd.	Cool. Temp.
#1	84.033	0	0	255.55	41.516	137.506	39.396
#2	42.671	0	0	256.497	28.174	47.543	19.115
#3	50.503	67.539	39.996	179.106	53.944	65.774	43.138
#4	41.107	32.725	27.268	130.192	30.645	46.96	20.102

Source: Own Authorship.

4.4. Evaluation

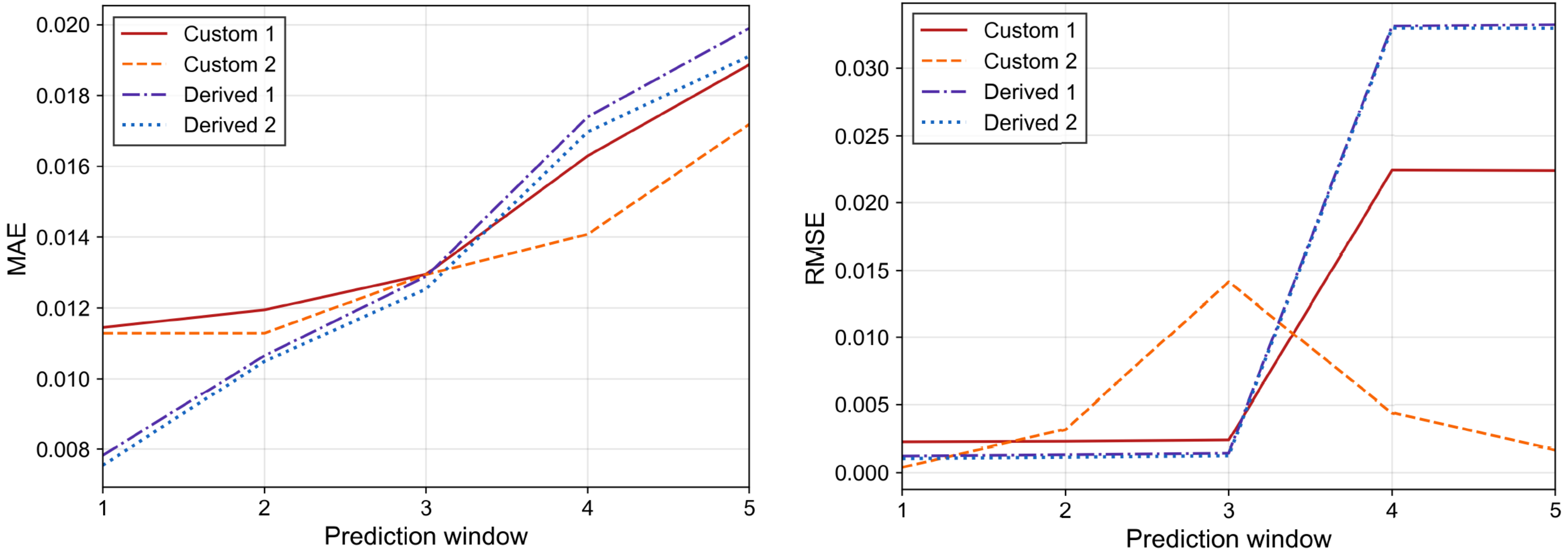
Figure 7 – Custom and derived DA-RNN architectures.



Source: Own Authorship.

4.4.

Figure 13 – Error × prediction window curve of all evaluation models for the (a) MAE metric and (b) RMSE metric.



(a) MAE **(b) RMSE**

Table 4 – Evaluation results.

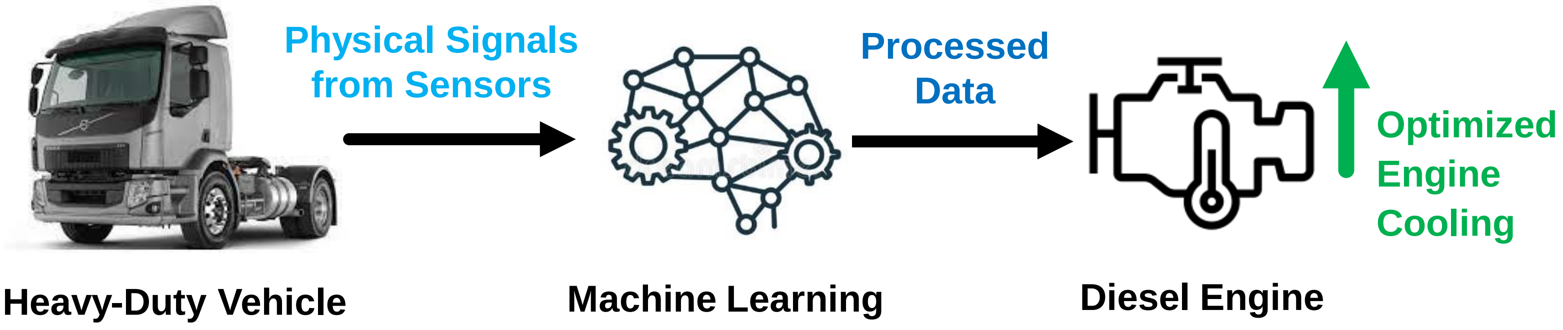
Metrics		Models			
		Custom 1	Derived 1	Custom 2	Derived 2
Number of Param.		3427	3169	7118	7106
Accuracy	MAE	0.0143	0.0137	0.0135	**0.0133**
	RMSE	0.0103	0.0141	**0.00475**	0.0139

Source: Own Authorship.

5. Model Predictive Controller (MPC)

- **How:** An ML model maps past coolant temperature and thermal impact variables to a future horizon of fan speeds.
  - Data-driven model design, model-based clustering and RL labeling strategies.

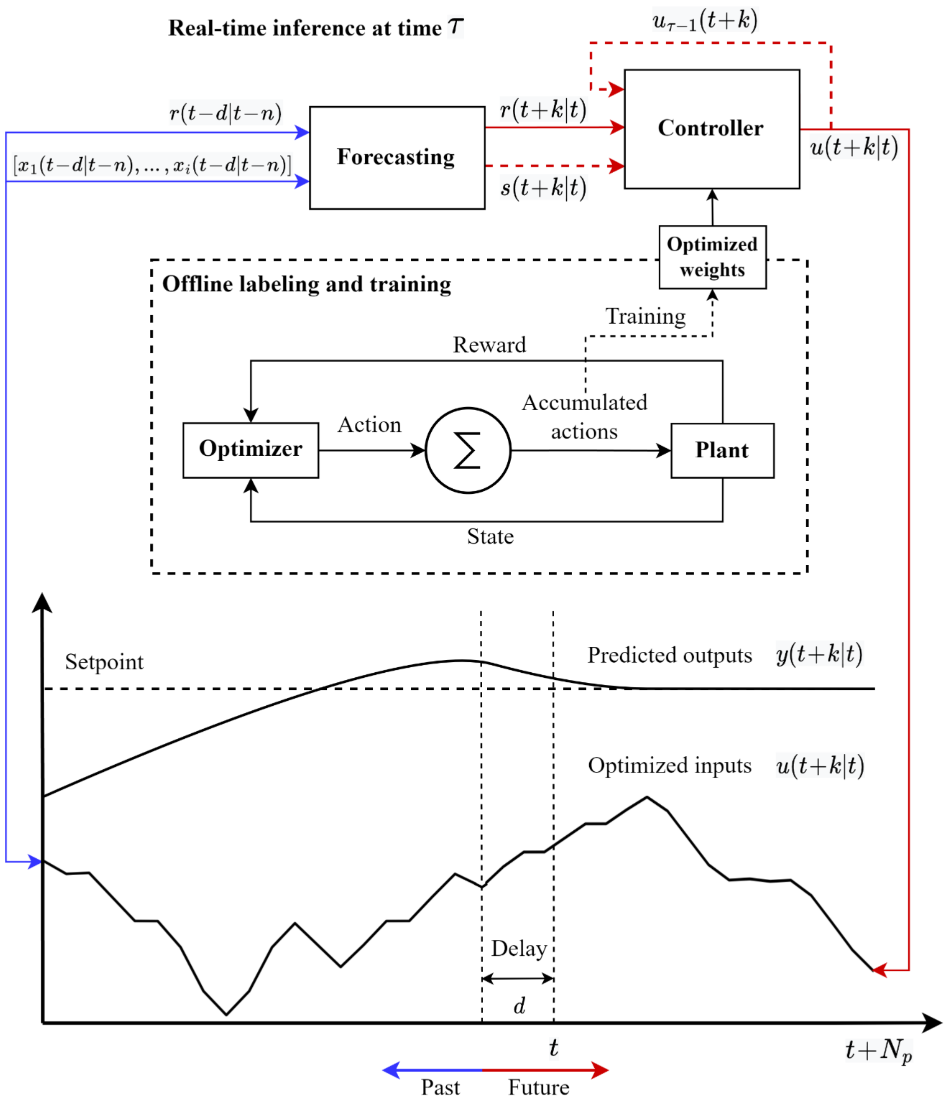
- **Achievements:** Regulation accuracy improvements due to a better knowledge of the cooling temperature response system:
  - Potential fuel economy improvements



5.1. Method and Strategies

- **Main strategies:**
  - **Forecasting:** Predicts coolant temperature and thermal impact states.
  - **Controller:** Predicts a horizon of fan speed demands.
  - **Offline Labeling and training:** Provides optimized fan speed demands for the controller training

Figure 8 – Proposed MPC method.



Source: Own Authorship.

5.2. Offline Reinforcement Learning (RL) Labeling

Figure 10 – Visual representations of the action space.

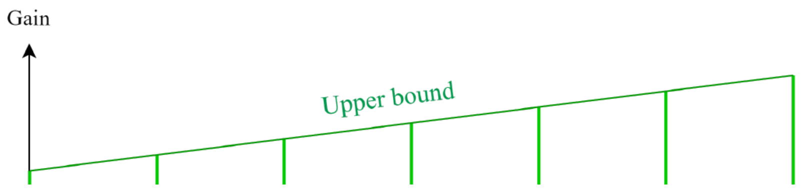
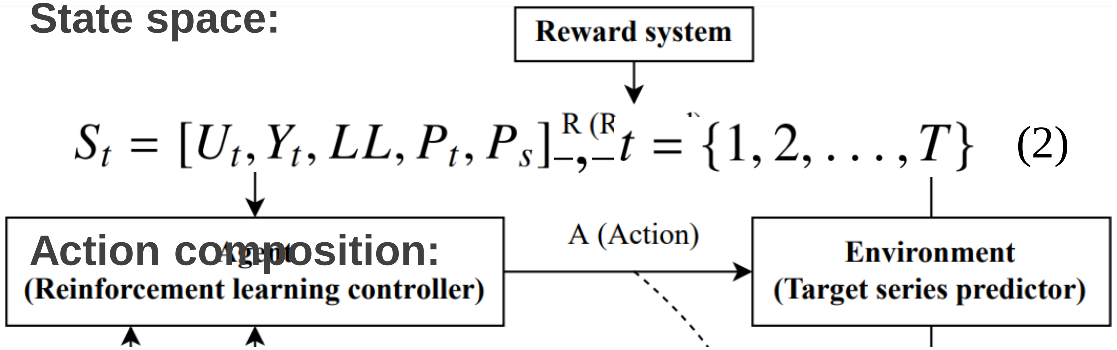


Figure 9 – RL labeling model.

- **State space:**



$$S_t = [U_t, Y_t, LL, P_t, P_s], t = \{1, 2, \dots, T\} \quad (2)$$

- **Action composition:**

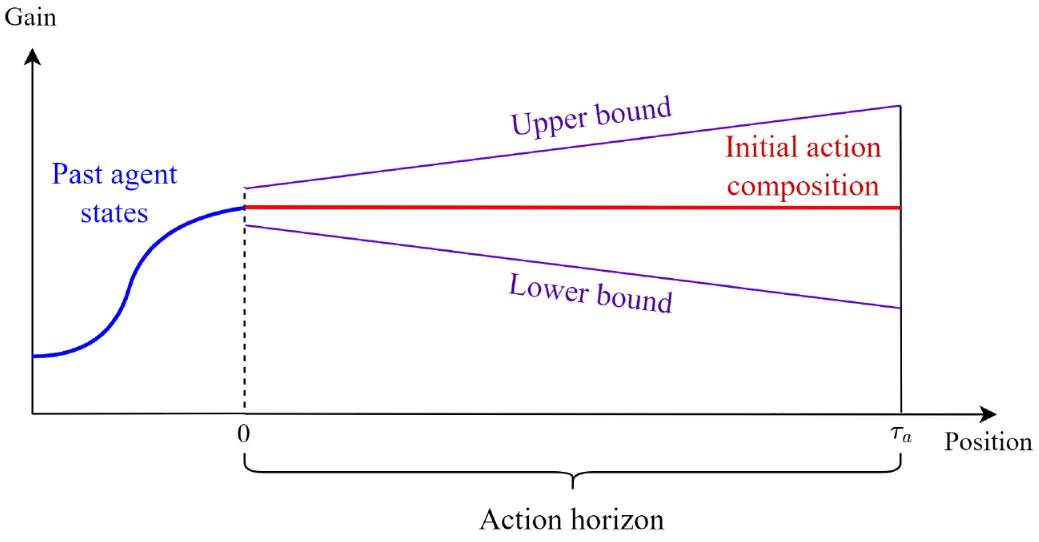
Table 2 – Action setting.

ID	Description	Summary
#1	Increase the fan speed	$+k$
#2	Decrease the fan speed	$-k$
#3	Move forward 2 positions	$+2p$
#4	Move back 2 positions	$-2p$
#5	Move forward 1 position and increase the fan speed	$+p+k$
#6	Move forward 1 position and decrease the fan speed	$+p-k$
#7	Move back 1 position and increase the fan speed	$-p+k$
#8	Move back 1 position and decrease the fan speed	$-p-k$

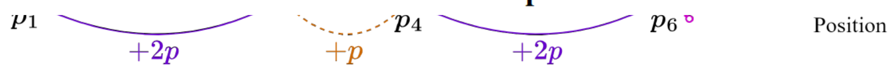
Source: Own Authorship.

Source: Own Authorship.

Figure 15 – Initial action composition.



Source: Own Authorship.



(b) Strategy example

Source: Own authorship.

5.2. Offline Reinforcement Learning (RL) Labeling

- Reward:
  - Function as a regression problem (MSE)
  - Cumulative distribution of reward and penalties

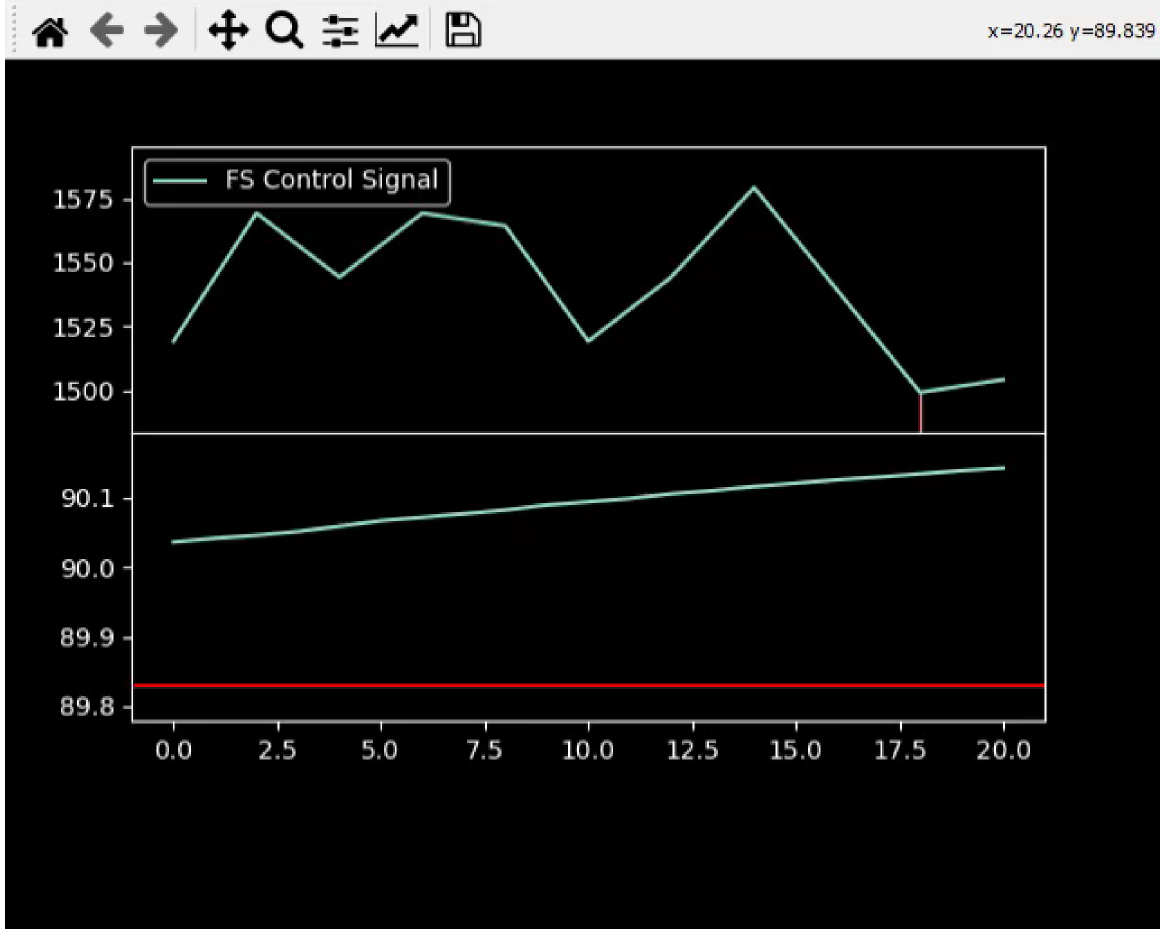
- Labeling:
  - Learned through exploration
  - Optimally achieved with exploitation

Algorithm 2 – Optimize the fan actuation and extract the labels

$baseError = MSE(Y_i[P_s,\ldots,\tau_e], \hat{Y}[P_s,\ldots,\tau_e])$
for $t = 1,\ldots,T$ **do**
  $U_t, P_t = DQNAgentInference(S_t, reward)$
  $Y_t = DNNEnvironmentInference(X_i, U_t)$
  $S_t = [U_t, Y_t, LL, P_t, P_s]$
  $currentError = MSE(Y_t[P_s,\ldots,\tau_e], \hat{Y}[P_s,\ldots,\tau_e])$
  $proximityScore = \frac{(currentError - (2 * baseError)) * (2 * r_{max})}{(goalError - (2 * baseError))} - r_{max}$
  $reward = 0$
  while $-s < milestone < s$ **do**
    if $milestone \geq 0$ **then**
      if $proximityScore \geq (r\_list[milestone] + d)$ **then**
        $reward \mathrel{+}= r\_list[1]$
        $milestone{+}{+}$
      else if $proximityScore \leq (r\_list[milestone] - d)$ **then**
        $reward \mathrel{-}= r\_list[1]$
        $milestone{-}{-}$
      else
        break
      end if
    else
      if $proximityScore \geq (-r\_list[-milestone] + d)$ **then**
        $reward \mathrel{+}= r\_list[1]$
        $milestone{+}{+}$
      else if $proximityScore \leq (-r\_list[-milestone] - d)$ **then**
        $reward \mathrel{-}= r\_list[1]$
        $milestone{-}{-}$
      else
        break
      end if
    end if
  end while
  if $milestone > oldMilestone$ **then**
    $oldMilestone = milestone$
    $L_i = GetLabel(X_i, U_t)$
  end if
end for
return L_i

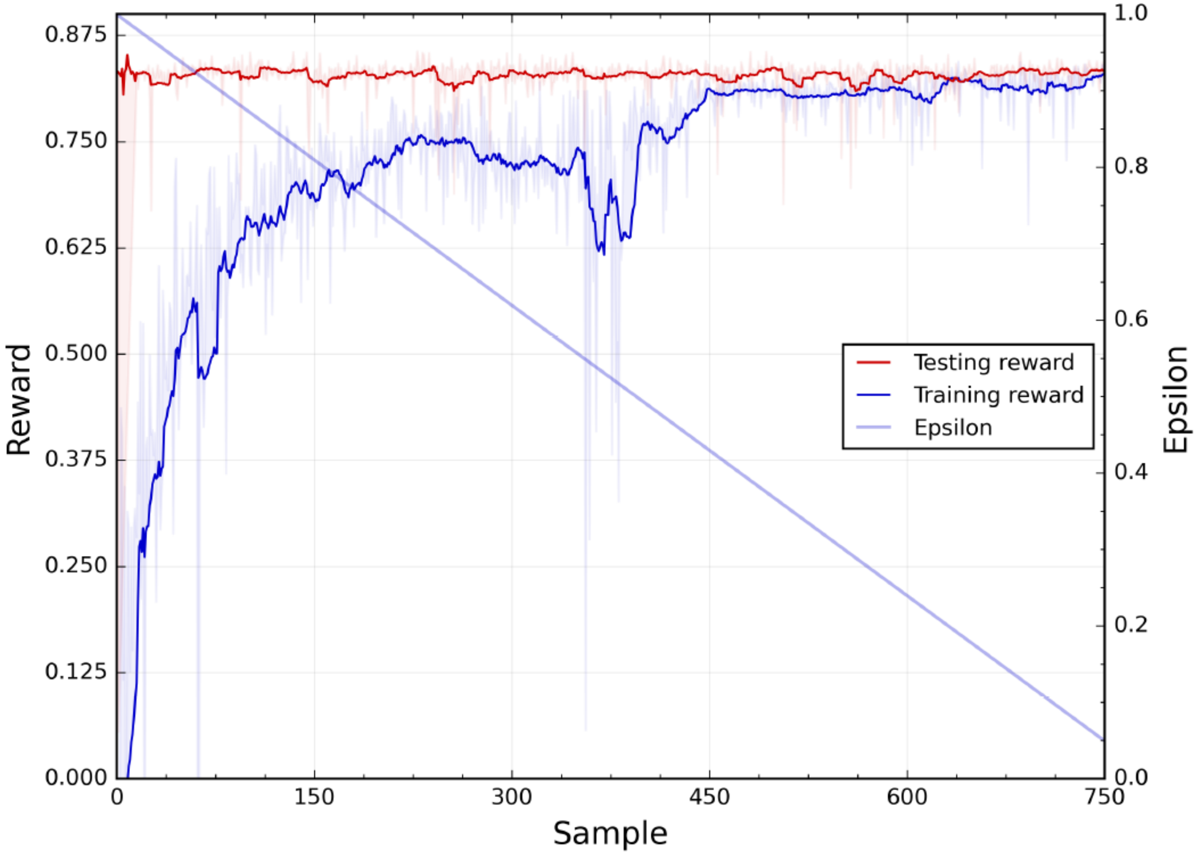
Source: Own Authorship.

5.2. Offline Reinforcement Learning (RL) Labeling



Exploitation

Figure 21 – Reward evolution of the labeling model.



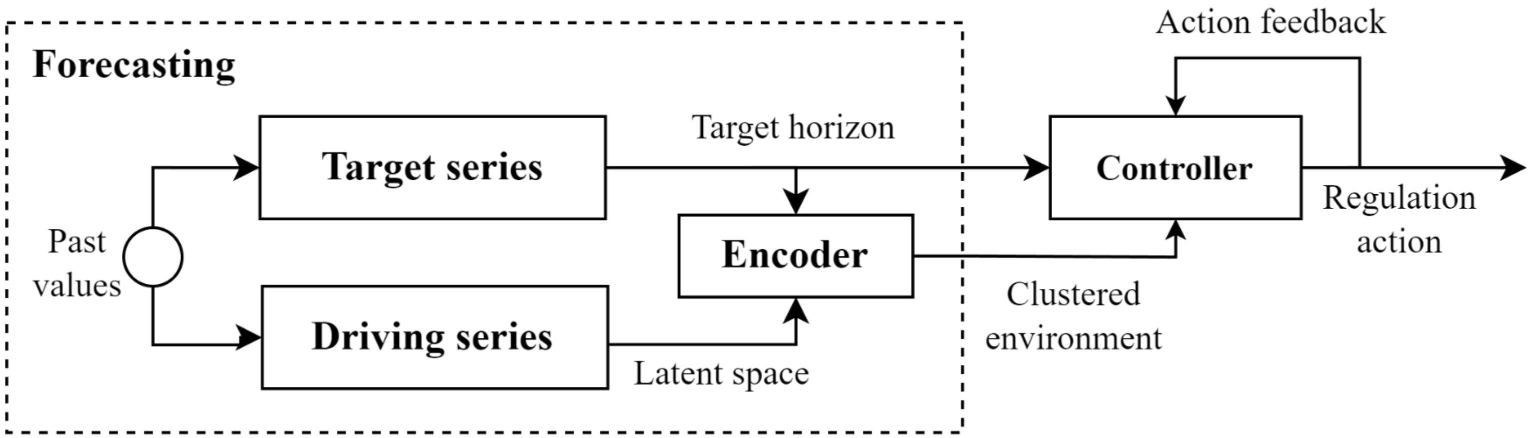
Source: Own Authorship.

5.3. Models and Architectures

- Forecasting:
  - Parellel prediction of the target and thermal impact series
  - Custom DA-RNN (Sequential attention)

- Controller:
  - Inspired by the RL agent
  - Translation model

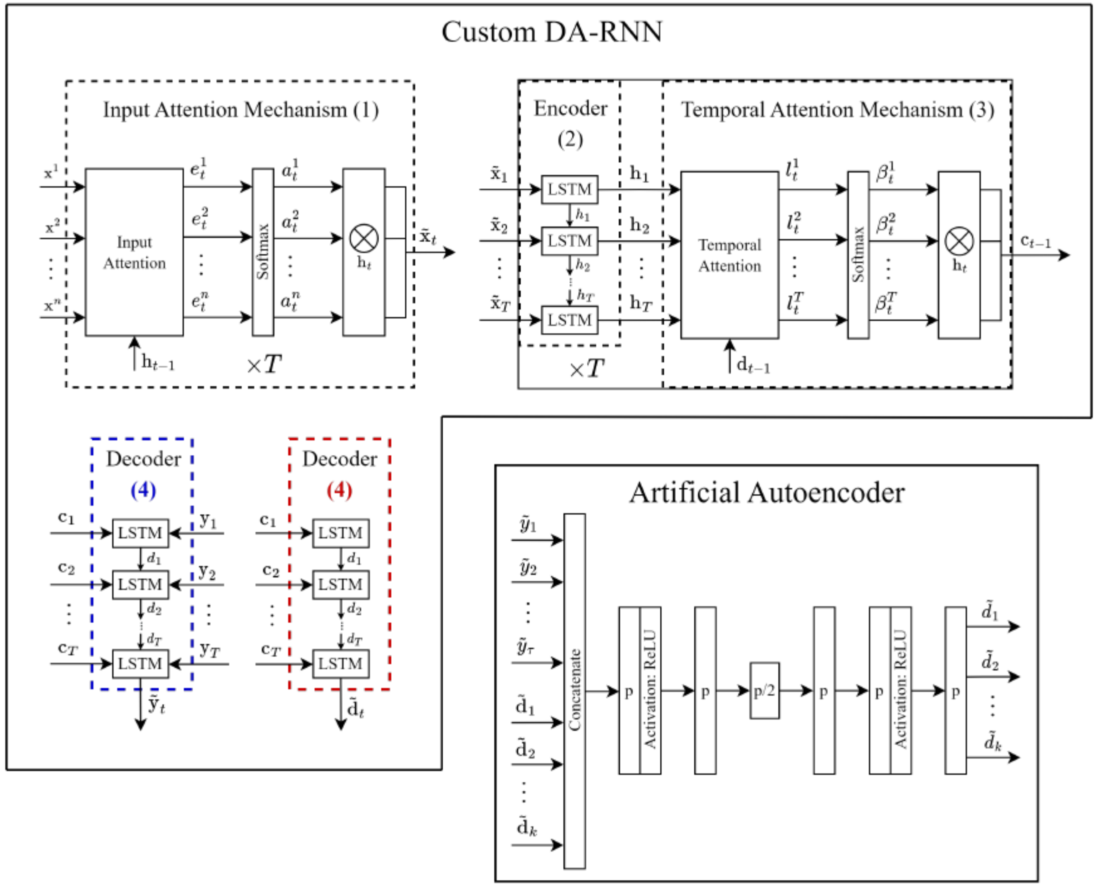
Figure 11 – Internal structure.



Source: Own Authorship.

5.3. Models and Architectures

Figure 12 – Forecasting architecture.



Source: Own Authorship.

Figure 13 – Controller architecture.

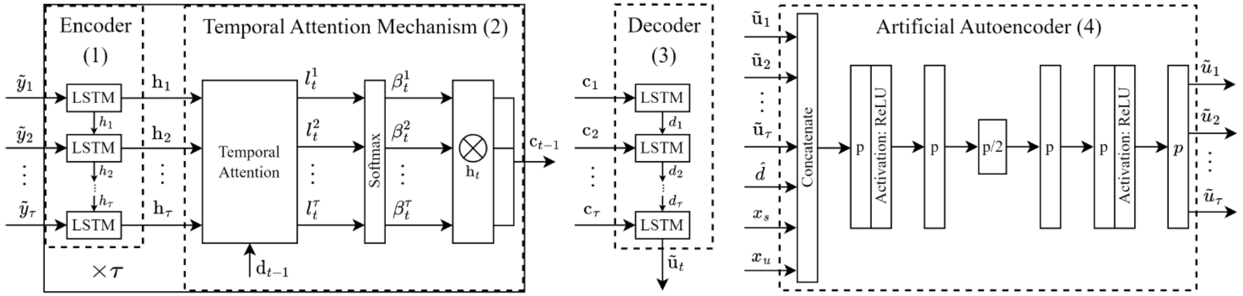
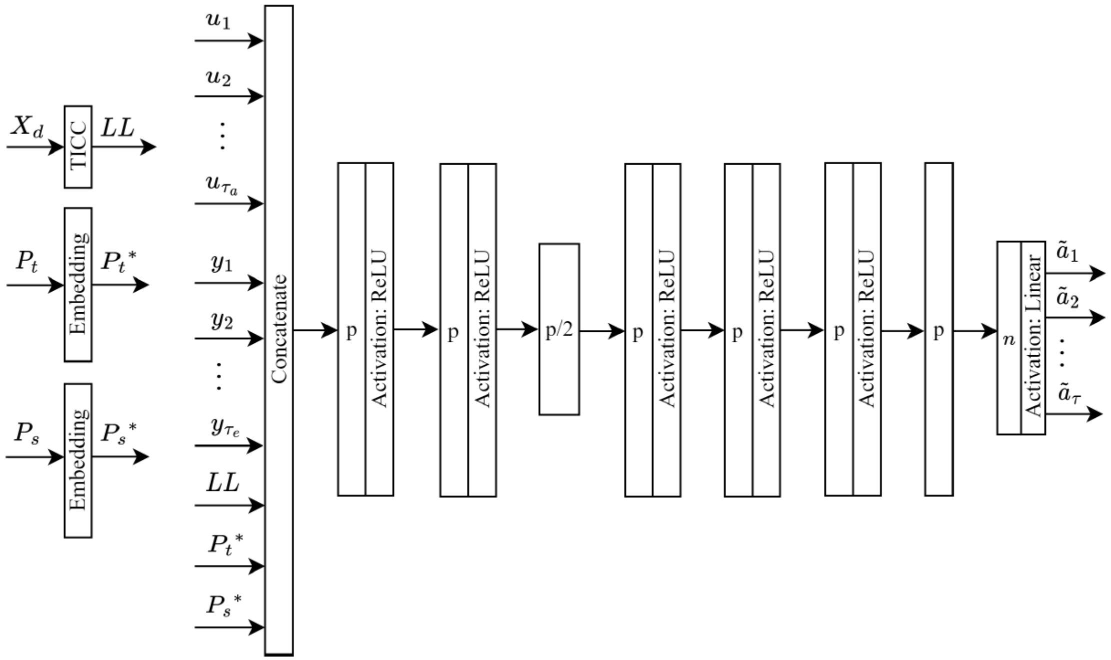


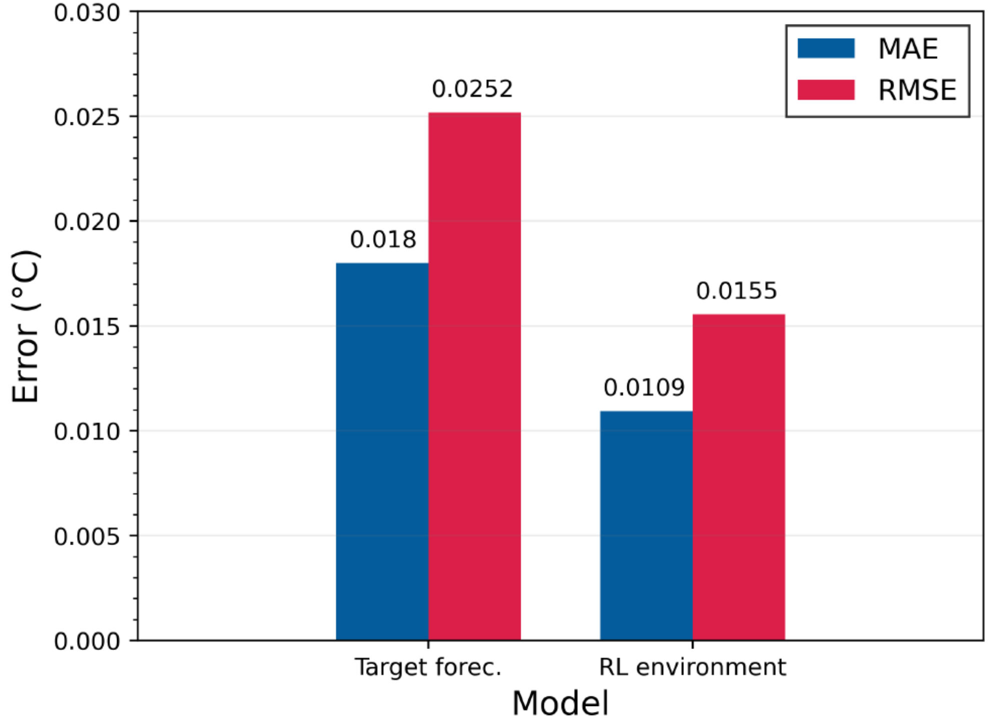
Figure 14 – Agent architecture.



Source: Own Authorship.

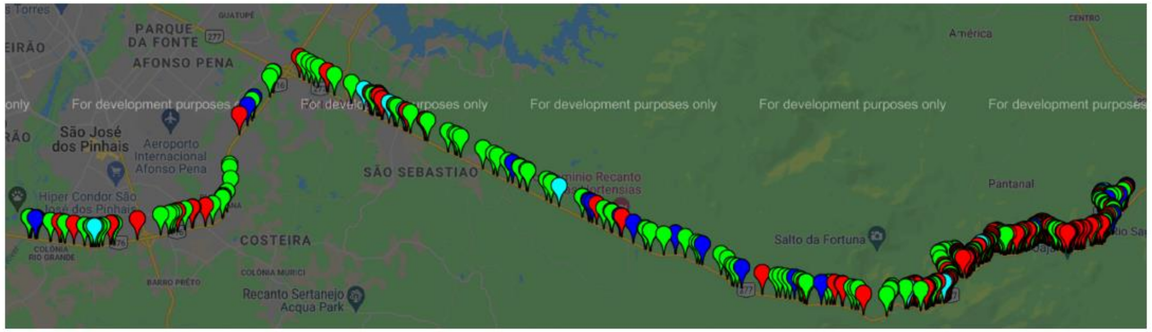
5.4. Training and Testing

Figure 19 – Accuracy of the target series forecasting and environment models.

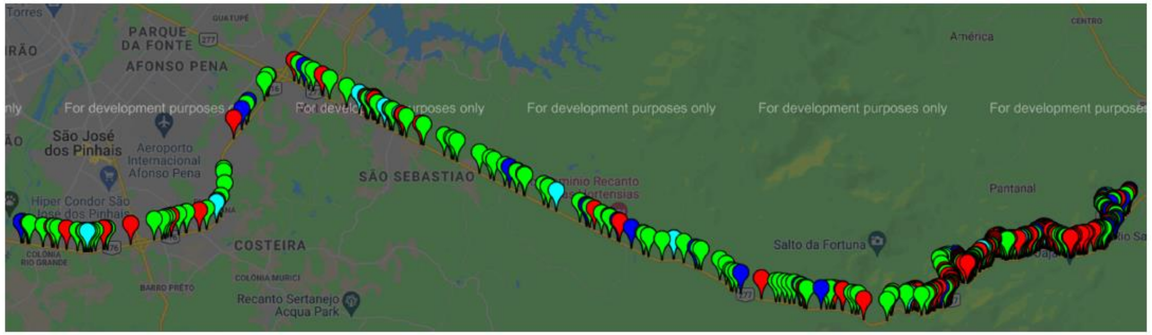


Source: Own Authorship.

Figure 20 – Single-class performance of the driving series forecasting model.



(a) Ground truth



(b) Prediction

Source: Own authorship.

Table 6 – Accuracy of the driving series forecasting model.

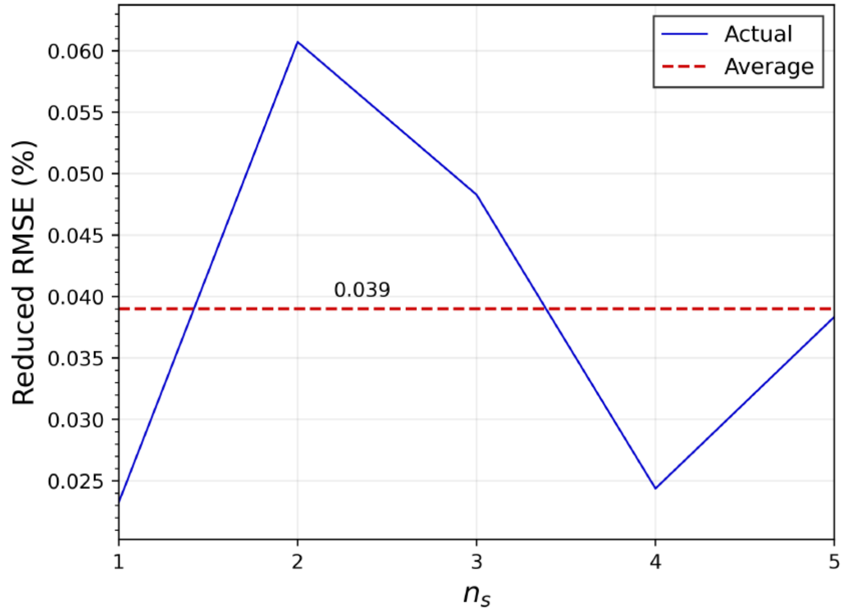
Multi-class		Single-class
MAE	RMSE	PCP (%)
0.2391	0.3892	93.9

Source: Own Authorship.

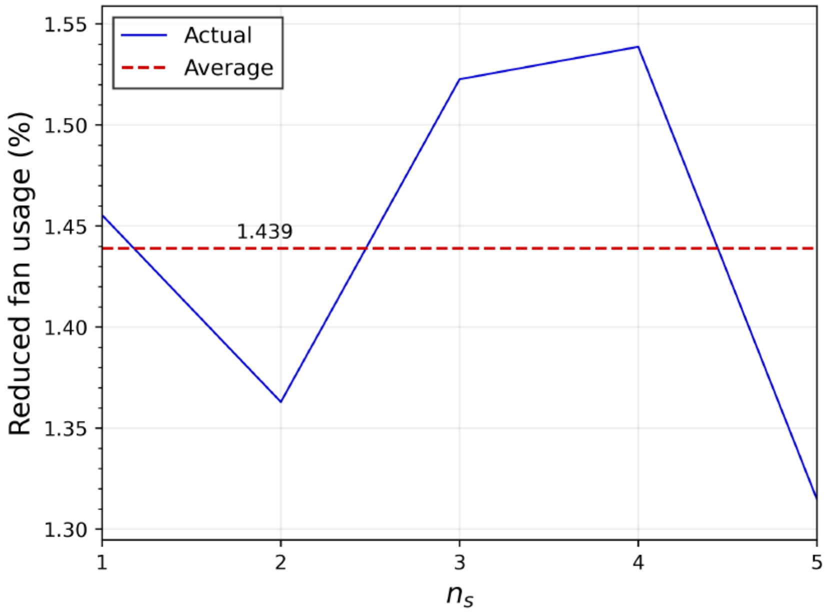
5.4. Training and Testing

- Complete controller solution:
  - Multiple runs with a progressive time evolution;
  - Performance metrics:
    - Fan usage;
    - RMSE.

Figure 22 – Results of the MPC solution.



(a) RMSE



(b) Fan usage

Source: Own authorship.

6. Physical Evaluation

- **Evaluation method:**
  - Comparison between truck runs with intercallated controller actuations;
  - Route and conditions similar to the training and testing samples with a fixed driver control behavior (35 km/h | avoidance of gear changes).
  - Positional comparisons;

- **Metrics:** Fuel consumption and fan usage.

- **How:** Python/ATI VISION integration for the acquisition, controller execution and impostion of the values of interest.

- **Evaluation samples selection:**
  - Selection over similar external conditions (6 samples);
  - Removal of inconsistent samples with TICC by defining a clustered environment that associates samples with similar internal and external impacts in the coolant temperature.

6. Physical Evaluation

- Evaluation algorithm:
  - Multiprocessing and threading;
  - Acquisition;
  - Controller execution;
  - Imposition;
  - Data logging;

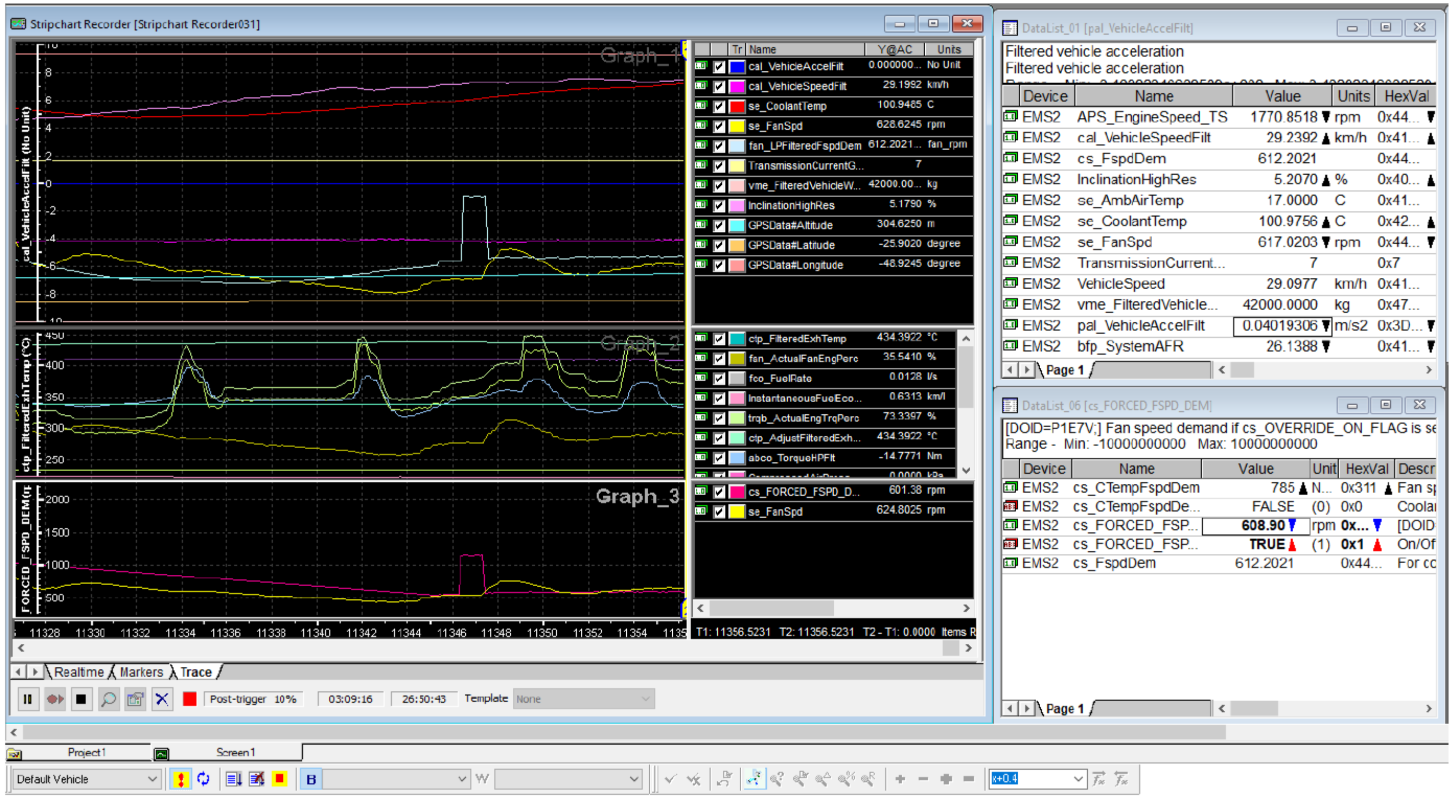
- ATI Vision visualization tool:

Algorithm 4 – Execute the controller inference

Given S_{list} = sample from the data acquisition process, t_s = sampling time.

while 1 **do**

Figure 16 – Configured ATI VISION screen for visualization.



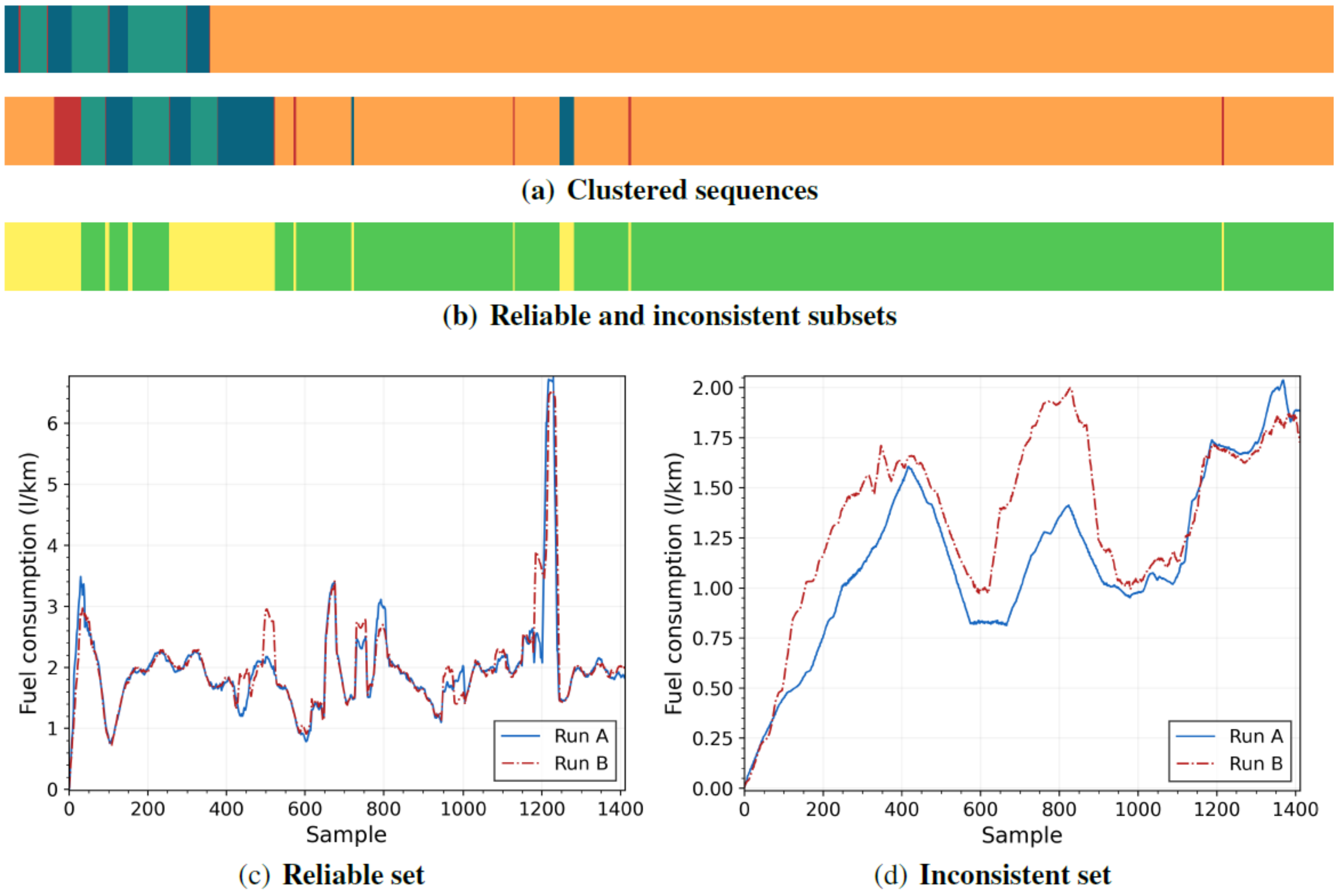
Source: Own Authorship.

end while

Source: Own Authorship.

6. Physical Evaluation

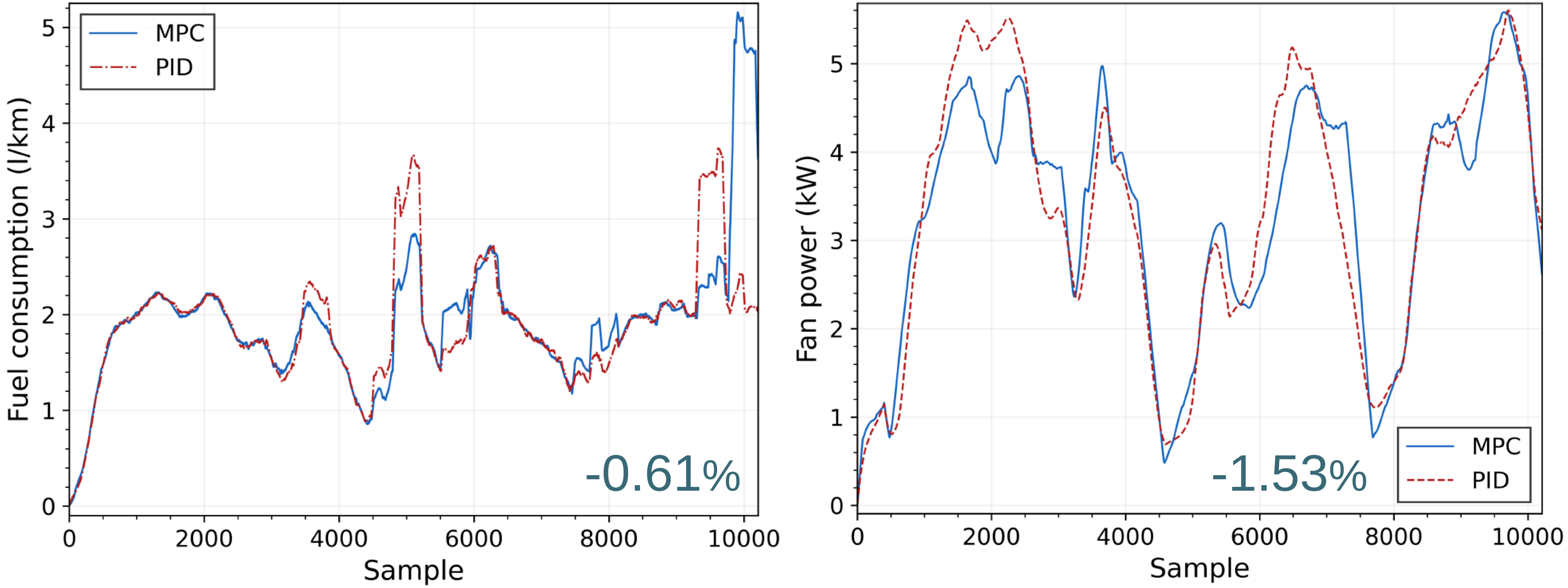
Figure 23 – Results of the extraction of reliable samples.



Source: Own authorship.

6. Physical Evaluation

Figure 24 – Regulation performance of the MPC and PID solutions.



(a) Fuel consumption (b) Fan power

Source: Own authorship.

7. Conclusion

- Application of supervised and semi-supervised ML strategies:
  - Improvements in the processed knowledge over the plant, reference and controlled signals;
  - Addressing of processing time concerns with the generation of low-dimensional feature spaces.
  - Multi-agent solution, where different cooling agents will be controlled while performing an individual and collective exploration of a new simulation environment;

- Controller evaluation results:
  - Average fan usage reduction of 1.439%, with remote plant simulations, and 1.53%, as resulted of the physical evaluation, when compared to the PID solution;
  - Average fuel consumption reduction of 0.63%;
  - Consequence of the predictive actuation of the proposed controller, learning without the specification of a range for actuation and generating a non-linear demand of fan speeds as proportional to a foreseen thermal rejection states.

- Additionally, considerations for different metrics of interest, in the conception of an optimization system, would generate improvements and a more consistent thermal regulation towards fuel efficiency.

8. Next Steps

- Design and development of a new controller version:
  - New input treatment model for the controller and RL agent, containing the input of different vehicle characteristics relative to the thermal objective (scalable to multiple configurations);
  - Developed into an embedded system, involving production scale concerns;
  - Multi-agent solution, where different cooling agents will be controlled while performing an individual and collective exploration of a new simulation environment;
  - Inner multi-agent setting, for each cooling agent, with a new decision-making system.

- The remote and physical tests and evaluations will be enhanced for multiple vehicle configurations and engine load conditions, according to the new achievements of a new version.

References

- ADETOLA, V.; DEHAAN, D.; GUAY, M. Adaptive model predictive control for constrained nonlinear systems. **Systems & Control Letters**, v. 58, n. 5, p. 320–326, 2009. ISSN 0167-6911. DOI: https://doi.org/10.1016/j.sysconle.2008.12.002. Available from: https://www.sciencedirect.com/science/article/pii/S0167691108002120.
- ARULKUMARAN, K. et al. Deep Reinforcement Learning: A Brief Survey. **IEEE Signal Processing Magazine**, v. 34, n. 6, p. 26–38, 2017.
- AUER, P. Using Confidence Bounds for Exploitation-Exploration Trade-Offs. **J. Mach. Learn. Res.**, JMLR.org, v. 3, null, p. 397–422, Mar. 2003. ISSN 1532-4435.
- AUER, P.; CESA-BIANCHI, N.; FISCHER, P. Finite-Time Analysis of the Multiarmed Bandit Problem. **Mach. Learn.**, Kluwer Academic Publishers, USA, v. 47, 2–3, p. 235–256, May 2002. ISSN 0885-6125. Available from: https://doi.org/10.1023/A:1013689704352.
- BEGUM, N. et al. Accelerating Dynamic Time Warping Clustering with a Novel Admissible Pruning Strategy. In: PROCEEDINGS of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Sydney, NSW, Australia: Association for Computing Machinery, 2015. (KDD ’15), p. 49–58. ISBN 9781450336642.
- BEMPORAD, A.; MORARI, M., et al. The Explicit Linear Quadratic Regulator for Constrained Systems. **Automatica**, Pergamon Press, Inc., USA, v. 38, n. 1, p. 3–20, Jan. 2002. ISSN 0005-1098. Available from: https://doi.org/10.1016/S0005-1098(01)00174-1.
- BEMPORAD, A.; RICKER, L. N.; MORARI, M. **Get Started with Model Predictive Control Toolbox — mathworks.com**. [S.l.: s.n.], 2020. https://www.mathworks.com/support/search.html?fq[]=asset_type_name:video&fq[]=category:mpc/getting-startedwith-model-predictive-control-toolbox&page=1. [Accessed 20-Aug-2022].
- BEMPORAD, P. Explicit Model Predictive Control. In: [s.l.: s.n.], Jan. 2013. P. 1–9. ISBN 978-1-4471-5102-9.
- BENGIO, Y.; SIMARD, P.; FRASCONI, P. Learning long-term dependencies with gradient descent is difficult. **IEEE Transactions on Neural Networks**, v. 5, n. 2, p. 157–166, 1994.
- BERNDT, D. J.; CLIFFORD, J. Using Dynamic Time Warping to Find Patterns in Time Series. In: KDD Workshop. [S.l.: s.n.], 1994.
- BOYD, S. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. **Foundations and Trends® in Machine Learning**, Now Publishers, v. 3, n. 1, p. 1–122, 2010.
- BRANDES, U. A Faster Algorithm for Betweenness Centrality. **Journal of Mathematical Sociology**, v. 25, p. 163–177, 2001.

References

- CHISCI, L.; FALUGI, P.; ZAPPA, G. Gain-scheduling MPC of nonlinear systems. **International Journal of Robust and Nonlinear Control**, v. 13, n. 3-4, p. 295–308, 2003. DOI: https://doi.org/10.1002/rnc.819. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/rnc.819.
- CHO, K.; MERRIENBOER, B. van, et al. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches. **CoRR**, abs/1409.1259, 2014. arXiv: 1409.1259. Available from: http://arxiv.org/abs/1409.1259.
- CHO, K.; MERRIËNBOER, B. van, et al. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. In: PROCEEDINGS of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar: Association for Computational Linguistics, Oct. 2014. P. 1724–1734. DOI: 10.3115/v1/D14-1179. Available from: https://aclanthology.org/D14-1179. DIACONESCU, E. The use of NARX neural networks to predict chaotic time series.**WSEAS Transactions on Computers archive**, v. 3, p. 182–191, 2008.
- ESTER, M. et al. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In: PROCEEDINGS of the Second International Conference on Knowledge Discovery and Data Mining. Portland, Oregon: AAAI Press, 1996. (KDD'96), p. 226–231.
- FENG, G. A Survey on Analysis and Design of Model-Based Fuzzy Control Systems. **IEEE Transactions on Fuzzy Systems**, v. 14, n. 5, p. 676–697, 2006.
- FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. Sparse inverse covariance estimation with the graphical lasso. en. **Biostatistics**, v. 9, n. 3, p. 432–441, July 2008. ISSN 1465-4644, 1468-4357. DOI: 10.1093/biostatistics/kxm045. Available from: http://biostatistics.oxfordjournals.org/content/9/3/432. Visited on: 15 Mar. 2013.
- GAO, Y.; ER, M. J. NARMAX time series model prediction: feedforward and recurrent fuzzy neural network approaches. **Fuzzy Sets and Systems**, Elsevier BV, v. 150, n. 2, p. 331–350, Mar. 2005.
- GRUSLYS, A. et al. The Reactor: A fast and sample-efficient Actor-Critic agent for Reinforcement Learning. In: INTERNATIONAL Conference on Learning Representations. [S.l.: s.n.], 2018. Available from: https://openreview.net/forum?id=rkHVZWZAZ.
- GU, S.; LILLICRAP, T.; GHAHRAMANI, Z., et al. **Q-Prop: Sample-Efficient Policy Gradient with An Off-Policy Critic**. [S.l.]: arXiv, 2016. Available from: https://arxiv.org/abs/1611.02247.
- GU, S.; LILLICRAP, T.; SUTSKEVER, I., et al. **Continuous Deep Q-Learning with Model-based Acceleration**. [S.l.]: arXiv, 2016. Available from: https://arxiv.org/abs/1603.00748.

References

- GU, S. (et al. Interpolated Policy Gradient: Merging On-Policy and Off-Policy Gradient Estimation for Deep Reinforcement Learning. In: GUYON, I. et al. (Eds.). **Advances in Neural Information Processing Systems**. [S.l.]: Curran Associates, Inc., 2017. v. 30. Available from: https://proceedings.neurips.cc/paper/2017/file/a1d7311f2a312426d710e1c617fcbc8c-Paper.pdf.
- HALLAC, D.; NYSTRUP, P.; BOYD, S. **Greedy Gaussian Segmentation of Multivariate Time Series**. [S.l.]: arXiv, 2016. Available from: https://arxiv.org/abs/1610.07435.
- HALLAC, D.; VARE, S., et al. Toeplitz Inverse Covariance-Based Clustering of Multivariate Time Series Data. In: PROCEEDINGS of the 27th International Joint Conference on Artificial Intelligence. Stockholm, Sweden: AAAI Press, 2018. (IJCAI'18), p. 5254–5258. ISBN 9780999241127.
- HAUSKNECHT, M.; STONE, P. **Deep Recurrent Q-Learning for Partially Observable MDPs**. [S.l.]: arXiv, 2015. Available from: https://arxiv.org/abs/1507.06527.
- HAYKIN, S. **Adaptive filter theory**. 4th. Upper Saddle River, NJ: Prentice Hall, 2002.
- HEESS, N. et al. **Memory-based control with recurrent neural networks**. [S.l.]: arXiv, 2015. Available from: https://arxiv.org/abs/1512.04455.
- HIMBERG, J. et al. Time series segmentation for context recognition in mobile devices. In: PROCEEDINGS 2001 IEEE International Conference on Data Mining. [S.l.: s.n.], 2001. P. 203–210.
- HOCHREITER, S.; SCHMIDHUBER, J. Long Short-term Memory. **Neural computation**, v. 9, p. 1735–80, Dec. 1997.
- KIM, J.-S. Recent advances in adaptive MPC. In: ICCAS 2010. [S.l.: s.n.], 2010. P. 218–222.
- KIM, J.-S.; YOON, T.-W.; SHIM, H. Switching Adaptive Output Feedback MPC for Input-constrained Neutrally Stable Linear Plants. In: PROCEEDINGS of the 44th IEEE Conference on Decision and Control. [S.l.: s.n.], 2005. P. 777–782.
- KIM, T.-H.; SUGIE, T. Adaptive receding horizon predictive control for constrained discrete-time linear systems with parameter uncertainties. **International Journal of Control**, Taylor & Francis, v. 81, n. 1, p. 62–73, 2008. DOI: 10.1080/00207170701266779. Available from: https://doi.org/10.1080/00207170701266779.
- KULKARNI, T. D. et al. Deep Successor Reinforcement Learning. **ArXiv**, abs/1606.02396, 2016.
- LEE, J.; CHANG, H.-J. An application of explicit model predictive control to electric power assisted steering systems. In: 2017 11th Asian Control Conference (ASCC). [S.l.: s.n.], 2017. P. 2119–2124.

References

- LILLICRAP, T. P. et al. Continuous control with deep reinforcement learning. In: BENGIO, Y.; LECUN, Y. (Eds.). **4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings**. [S.l.: s.n.], 2016. Available from: http://arxiv.org/abs/1509.02971.
- LIN, T. et al. Learning long-term dependencies in NARX recurrent neural networks. **IEEE Transactions on Neural Networks**, v. 7, n. 6, p. 1329–1338, 1996.
- MAATEN, L. van der; HINTON, G. Visualizing Data using t-SNE. **Journal of Machine Learning Research**, v. 9, n. 86, p. 2579–2605, 2008.
- MAKRIDAKIS, S.; HIBON, M. ARMA Models and the Box-Jenkins Methodology. **Journal of Forecasting**, Wiley, v. 16, n. 3, p. 147–163, May 1997.
- MIURA, K. An Introduction to Maximum Likelihood Estimation and Information Geometry. **Interdisciplinary Information Sciences (IIS)**, v. 17, Nov. 2011.
- MNIH, V.; BADIA, A. P.; MIRZA, M.; GRAVES, A.; LILLICRAP, T., et al. Asynchronous Methods for Deep Reinforcement Learning. In: BALCAN, M. F.; WEINBERGER, K. Q. (Eds.). **Proceedings of The 33rd International Conference on Machine Learning**. New York, New York, USA: PMLR, June 2016. v. 48. (Proceedings of Machine Learning Research), p. 1928–1937. Available from: https://proceedings.mlr.press/v48/mniha16.html.
- MNIH, V.; BADIA, A. P.; MIRZA, M.; GRAVES, A.; LILLICRAP, T. P., et al. Asynchronous Methods for Deep Reinforcement Learning. arXiv, 2016. Available from: https://arxiv.org/abs/1602.01783.
- MNIH, V.; KAVUKCUOGLU, K., et al. Human-level control through deep reinforcement learning. **Nature**, Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., v. 518, n. 7540, p. 529–533, Feb. 2015. ISSN 00280836. Available from: http://dx.doi.org/10.1038/nature14236.
- MÜLLNER, D. **Modern hierarchical, agglomerative clustering algorithms**. [S.l.]: arXiv, 2011. Available from: https://arxiv.org/abs/1109.2378.
- NA, S.; XUMIN, L.; YONG, G. Research on k-means Clustering Algorithm: An Improved k-means Clustering Algorithm. In: 2010 Third International Symposium on Intelligent Information Technology and Security Informatics. [S.l.: s.n.], 2010. P. 63–67.
- NG, A. Y.; JORDAN, M. I.; WEISS, Y. On Spectral Clustering: Analysis and an Algorithm. In: PROCEEDINGS of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic. Vancouver, British Columbia, Canada: MIT Press, 2001. (NIPS'01), p. 849–856.

References

- O'DONOGHUE, B. et al. Combining policy gradient and Q-learning. In: INTERNATIONAL Conference on Learning Representations. [S.l.: s.n.], 2017. Available from: https://openreview.net/forum?id=B1kJ6H9ex.
- OH, J. et al. **Control of Memory, Active Perception, and Action in Minecraft**. [S.l.]: arXiv, 2016. Available from: https://arxiv.org/abs/1605.09128.
- PARIKH, N. Proximal Algorithms. **Foundations and Trends® in Optimization**, Now Publishers, v. 1, n. 3, p. 127–239, 2014.
- QIN, Y. et al. A dual-stage attention-based recurrent neural network for time series prediction. In: AAAI PRESS. PROCEEDINGS of the 26th International Joint Conference on Artificial Intelligence. [S.l.: s.n.], 2017. P. 2627–2633. Available from: https://www.ijcai.org/proceedings/2017/0366.pdf.
- RUE, H.; HELD, L. **Gaussian Markov Random Fields: Theory And Applications (Monographs on Statistics and Applied Probability)**. [S.l.]: Chapman & Hall/CRC, 2005. ISBN 1584884320.
- RYBAKOV, S. et al. Learning interpretable latent autoencoder representations with annotations of feature sets. Cold Spring Harbor Laboratory, Dec. 2020. Available from:https://doi.org/10.1101/2020.12.02.401182.
- SALEM, F.; MOSAAD, M. I. A comparison between MPC and optimal PID controllers: Case studies. In: MICHAEL Faraday IET International Summit 2015. [S.l.: s.n.], 2015. P. 59–65.
- SCHULMAN, J. et al. High-Dimensional Continuous Control Using Generalized Advantage Estimation. In: PROCEEDINGS of the International Conference on Learning Representations (ICLR). [S.l.: s.n.], 2016.
- SCOTT, S. L. A modern Bayesian look at the multi-armed bandit. **Applied Stochastic Models in Business and Industry**, v. 26, n. 6, p. 639–658, 2010. DOI: https://doi.org/10.1002/asmb.874. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/asmb.874.
- SERFLING, R. J. Probability Inequalities for the Sum in Sampling without Replacement. **The Annals of Statistics**, Institute of Mathematical Statistics, v. 2, n. 1, p. 39–48, 1974. Available from: https://doi.org/10.1214/aos/1176342611.
- SHOUCHE, M. et al. Simultaneous Constrained Model Predictive Control and Identification of DARX Processes. **Automatica**, v. 34, n. 12, p. 1521–1530, 1998. ISSN 0005-1098. DOI: https://doi.org/10.1016/S0005-1098(98)80005-8. Available from: https://www.sciencedirect.com/science/article/pii/S0005109898800058.
- SILVER, D. et al. Deterministic Policy Gradient Algorithms. In: XING, E. P.; JEBARA, T. (Eds.), 1. **Proceedings of the 31st International Conference on Machine Learning**. Bejing, China: PMLR, June 2014. v. 32. (Proceedings of Machine Learning Research, 1), p. 387–395. Available from: https://proceedings.mlr.press/v32/silver14.html.

References

- SMYTH, P. Clustering Sequences with Hidden Markov Models. In: PROCEEDINGS of the 9th International Conference on Neural Information Processing Systems. Denver, Colorado: MIT Press, 1996. (NIPS'96), p. 648–654.
- SUTSKEVER, I.; VINYALS, O.; LE, Q. V. Sequence to Sequence Learning with Neural Networks. In: PROCEEDINGS of the 27th International Conference on Neural Information Processing Systems - Volume 2. Montreal, Canada: MIT Press, 2014. (NIPS'14), p. 3104–3112.
- SUTTON, R. S.; BARTO, A. G. **Reinforcement Learning: An Introduction**. Second. [S.l.]: The MIT Press, 2018. Available from: http://incompleteideas.net/book/the-book-2nd.html.
- SWIEF, A.; EL-ZAWAWI, A.; EL-HABROUK, M. A Survey of Model Predictive Control Development in Automotive Industries. In: 2019 International Conference on Applied Automation and Industrial Diagnostics (ICAAID). [S.l.: s.n.], 2019. v. 1, p. 1–7.
- SZITA, I.; LÖRINCZ, A. Optimistic Initialization and Greediness Lead to Polynomial Time Learning in Factored MDPs - Extended Version. **CoRR**, abs/0904.3352, Apr. 2009.
- VITERBI, A. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. **IEEE Transactions on Information Theory**, v. 13, n. 2, p. 260–269, 1967.
- WANG, Z.; BAPST, V., et al. Sample Efficient Actor-Critic with Experience Replay. In: INTERNATIONAL Conference on Learning Representations. [S.l.: s.n.], 2017. Available from: https://openreview.net/forum?id=HyM25Mqel.
- WANG, Z.; SCHAUL, T., et al. **Dueling Network Architectures for Deep Reinforcement Learning**. [S.l.]: arXiv, 2015. Available from: https://arxiv.org/abs/1511.06581.
- WIERSTRA, D. et al. Recurrent policy gradients. **Logic Journal of the IGPL**, v. 18, n. 5, p. 620–634, 2010.
- XU, C.; SU, Z. Identification of cell types from single-cell transcriptomes using a novel clustering method. **Bioinformatics**, v. 31, n. 12, p. 1974–1980, Feb. 2015. ISSN 1367-4803. eprint: https://academic.oup.com/bioinformatics/article-pdf/31/12/1974/17100675/btv088.pdf.
- YAN, L.; ELGAMAL, A.; COTTRELL, G. W. Substructure Vibration NARX Neural Network Approach for Statistical Damage Inference. **Journal of Engineering Mechanics**, American Society of Civil Engineers (ASCE), v. 139, n. 6, p. 737–747, June 2013.