Advancing Heliostat Aiming Strategies in Solar Tower Plants

SolarPACES

J.A. Carballo*[1,2], J. Bonilla[1,2], N.C. Cruz[3], J. Fernández-Reche[1], J.D. Álvarez[2], A. Avila-Marin[1], M. Berenguel [2]

[1] CIEMAT - Plataforma Solar de Almería (PSA), Ctra. de Senés km 4.5, E-04200 Tabernas, Almería, Spain

[2] CIESOL, Solar Energy Research Centre, Joint Institute University of Almería - CIEMAT, Department of Informatics, E-04120, Almería, Spain

[3] Department of Computer Engineering, Automatics, and Robotics. University of Granada, Granada, Spain

*Correspondence: J.A. Carballo, jcarballo@psa.es

1. Introduction

Solar Tower (ST) plants play a crucial role in renewable energy systems, achieving high thermal efficiency and enabling large-scale energy storage. A critical challenge in ST plants is modelling and optimizing the heliostat aiming strategy to maximize energy capture while ensuring operational safety and receiver longevity [1,2,3]. In our previous work [4], we introduced a model-free deep Reinforcement Learning (RL) approach using the Soft Actor-Critic (SAC) algorithm. It achieved an 8.8% increase in yearly absorbed power compared to traditional fixed-point aiming strategies. However, this study identified key areas for improvement, including the optimization of training efficiency and the incorporation of realistic solar variability:

1. Training Efficiency and Scalability: The original SAC implementation required extensive computational resources and time. We integrate Ray RLlib, a distributed RL framework, to enhance training scalability and enable advanced hyperparameter tuning.

2. Real-World Solar Variability: Instead of a constant DNI assumption, we now train the model using real-world DNI data from the Plataforma Solar de Almería (PSA) year-type dataset, capturing transient effects such as cloud coverage and seasonal variations.

These advancements aim to produce a more robust and adaptable aiming strategy better suited for real-world deployment in existing and future ST plants.

2. Methodology

As commented before, Ray RLlib [5] has been used in this work. It provides a scalable and flexible framework for distributed RL training, significantly improving upon the original TensorFlow-based SAC implementation. The key enhancements include:

- Parallel Environment Sampling: Multiple environment instances run concurrently, accelerating data collection and policy updates.

- Hyperparameter Optimization: Automated tuning of learning rates, network architectures, and exploration parameters to maximize policy performance.

Besides, using the PSA year-type dataset provides minute-resolution DNI measurements, enabling the RL agent to learn under realistic solar conditions that challenge traditional control strategies.

3. Preliminary Results and Discussion

Adopting Ray RLlib has transformed the training paradigm, improving both efficiency and performance. While the original SAC implementation required 9 days to train a single agent configuration, the new distributed framework completes simultaneous training and hyperparameter optimization of 500 distinct SAC configurations in less than 4 days - representing a speedup in experimental throughput. This massive parallelization capability, which evaluated over 500 parameter combinations including network architectures, learning rates, and exploration strategies, not only accelerates development cycles but also enhances policy robustness through comprehensive configuration space exploration. Moreover, the top-performing agents achieve a consistent 9.1% increase in annual energy yield over conventional methods, surpassing our previous 8.8% benchmark. These results, achieved under realistic variable DNI conditions with cloud transients, demonstrate both the technical superiority of our approach and its readiness for industrial-scale deployment in solar tower plants.

4. Conclusions and Future Work

This study demonstrates how integrating Ray RLlib with real-world DNI data significantly advances RL-based heliostat aiming optimization, achieving both superior scalability and operational robustness. The distributed framework enables faster and more efficient training, reducing computation time and making the approach practical for industrial-scale deployment. By incorporating actual PSA irradiance data with its inherent variability, the developed strategies show remarkable resilience to environmental fluctuations, maintaining stable performance even during cloud transients. Looking ahead, we plan to validate these results through field testing at the CESA-1 plant while exploring transfer learning techniques to adapt pre-trained policies to new solar installations with minimal retraining, potentially improving how solar tower plants optimize their aiming strategies across different geographic locations and receiver configurations.

References

[1] X. Lin, X. Zhao, Z. Liu, W. Huang, Y. Zhao & J. Feng. Real-time and high-accuracy radiative flux distribution simulation based on analytical model for solar power tower system. Solar Energy, 287, 113208. 2025. doi: https://doi.org/10.1016/j.solener.2024.113208

[2] M. Kuhl, M. Pargmann, M. Cherti, J. Jitsev, D.Maldonado & R. Pitz-Paal. Flux density distribution forecasting in concentrated solar tower plants: A data-driven approach. Solar Energy, 282, 112894. 2024. doi: https://doi.org/10.1016/j.solener.2024.112894

[3] F. Xu, Y. Sun & M. Guo. Prediction of Solar Flux Density Distribution Concentrated by a Heliostat Using a Ray Tracing-Assisted Generative Adversarial Neural Network. Energies, 18, 1451. 2025. doi: https://doi.org/10.3390/en18061451

[4] J. A. Carballo, J. Bonilla, N. C. Cruz, J. Fernández-Reche, J.D. Álvarez, A. Avila-Marin, & M. Berenguel. Reinforcement learning for heliostat aiming: Improving the performance of Solar Tower plants. Applied Energy, *377*, 124574. 2025. doi: https://doi.org/10.1016/j.apenergy.2024.124574

[5] E. Liang, R. Liaw, R. Nishihara, P. Moritz, R. Fox & J. Gonzalez. Ray rllib: A composable and scalable reinforcement learning library. arXiv preprint arXiv:1712.09381, 85, 245. 2025.