Advancing Heliostat Aiming Strategies in Solar Tower Plants

Jose A. Carballo[a], Javier Bonilla[a], N.C. Cruz[b], Jesús Fernández-Reche[a], J.D. Álvarez[c], Antonio Avila-Marín[a], M. Berenguel[c]

a CIEMAT - Plataforma Solar de Almería (PSA), Ctra. de Senés km 4.5, E-04200 Tabernas, Almería, Spain
b Dept. of Computer Engineering, Automation, and Robotics. Faculty of Education, Economics, and Technology of Ceuta, University of Granada
c CIESOL, Solar Energy Research Centre, Joint Institute University of Almería - CIEMAT, Department of Informatics, E-04120, Almería, Spain

Aiming strategy – Motivation

A major challenge in solar tower (ST) plants is optimizing heliostat aiming to maximize energy capture while protecting receiver integrity and lifetime.

We previously proposed a model-free deep Reinforcement Learning (RL) approach [1] using the Soft Actor–Critic (SAC) algorithm. The task is framed as continuous control, where an SAC agent interacts with a SolarPilot/CoPylot model of the CESA-1 field, learning through simulated episodes. Rewards promote absorbed power while penalizing spillage, excessive motion, and unsafe flux peaks, ensuring policies converge to safe and efficient solutions.

The agent evaluates plant states and provides optimized aiming in real time, enabling fully automatic, adaptive strategies that outperform static fixed-point methods.

The present work focuses on enhancing training efficiency and incorporating realistic solar variability to ensure robust, transferable strategies for current and future ST plants.

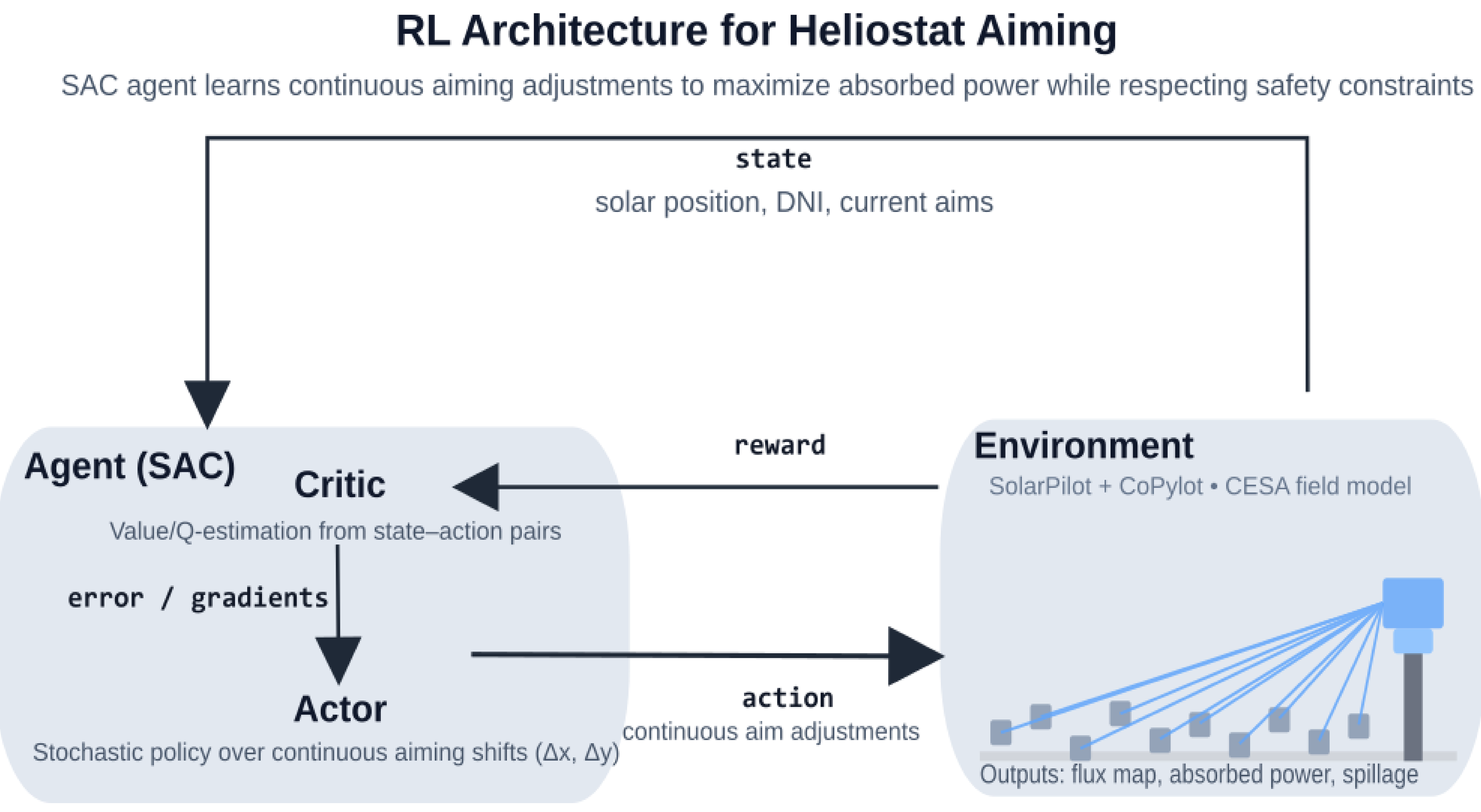

Fig 1.- RL in heliostat aiming strategy

MOTHODOLOGY

Building on previous results, we adopted the Ray framework to overcome the main limitations of our initial SAC approach. The implementation relies on Ray Rllib [2] for scalable distributed training, Ray Tune for large-scale hyperparameter optimization, and Ray Train for efficient execution.

The optimization process (Fig. 2) launches multiple parallel trainings with different configurations. Each training consists of episodes made of simulation steps, where the agent interacts with the environment, collects rewards, and updates actor–critic networks. Performance is evaluated through the mean episode return (R), and promising configurations are refined iteratively to converge towards robust policies.

The custom environment developed in the previous work in Python and integrating SolarPilot/CoPylot [3], simulates 114 heliostats of the CESA-I field has been modified to incorporate real PSA irradiance data, exposing the agent to variability and cloud transients.

All experiments were executed on a 225-core workstation at PSA, enabling large-scale parallel sampling and efficient policy evaluation. Best-performing policies were validated with annual simulations, assessing energy yield, reward trajectories, and flux distribution. This workflow ensures robustness and transferability of the learned strategies to real ST plant operation.

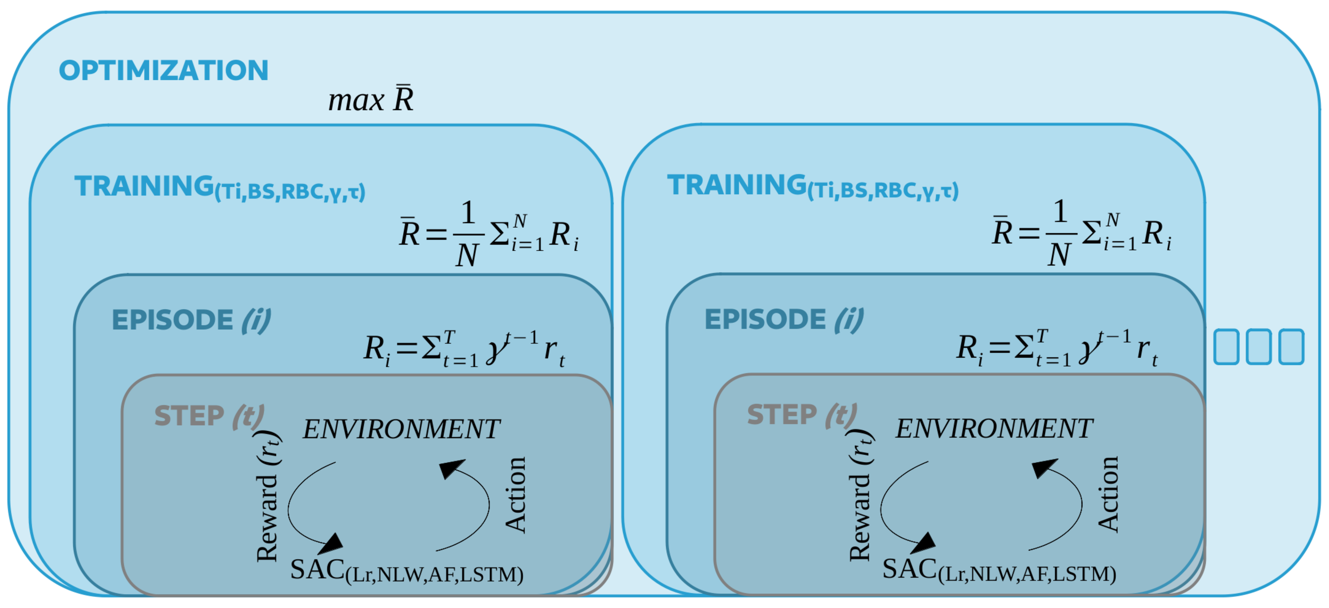

Fig 2.- Optimizatin and training process

Keras · PyTorch · XGBoost · python · RAY · SolarPILOT · RLlib · Ray Train · tune

Fig 3.- Software

RESULTS

The adoption of Ray RLlib greatly improved both efficiency and performance. Training time was reduced from ~9 days for a single SAC agent to <4 days for 500 parallel configurations, enabling systematic exploration of network architectures, learning rates, and exploration strategies.

This large-scale search enhanced policy robustness, ensuring generalization under realistic DNI variability, including cloud transients. Optimized agents achieved a 9.1% annual gain in absorbed power, surpassing the 8.8% improvement of our earlier study.

These results confirm the technical superiority and industrial readiness of distributed RL for heliostat aiming optimization in solar tower plants.

Fig 4.- Optimization/training results

Conclusions and future works

Integrating Ray RLlib with real DNI data advances RL-based heliostat aiming, achieving scalability, robustness, and industrial feasibility. The distributed framework reduces training time while maintaining stable performance under solar variability and cloud transients.

Future work will focus on field validation at the CESA-1 plant and the use of transfer learning to adapt pre-trained policies to new sites and receiver designs, enabling broader deployment of adaptive aiming strategies across solar tower plants.

References

[1] 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

[2] 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.

[3] Wagner, M.J., Wendelin, T. (2018). SolarPILOT: A power tower solar field layout and characterization tool, Solar Energy, Vol. 171, pp. 185-196, ISSN 0038-092X, https://doi.org/10.1016/j.solener.2018.06.063.

ASTERIX CAESAR

This work has been funded by ASTERIx-CAESar project which is funded by the European Union under the Grant Agreement Nº 101122231. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.