Reinforcement Learning Applied in Energy Management in Wearable IoT with Energy Harvesting

Daniel Iglesias
Magíster en Ciencias de la Ingeniería
Universidad de La Frontera
Temuco, Chile

Carlos Muñoz
Departamento de Ingeniería Eléctrica
Universidad de La Frontera
Temuco, Chile

October 2024

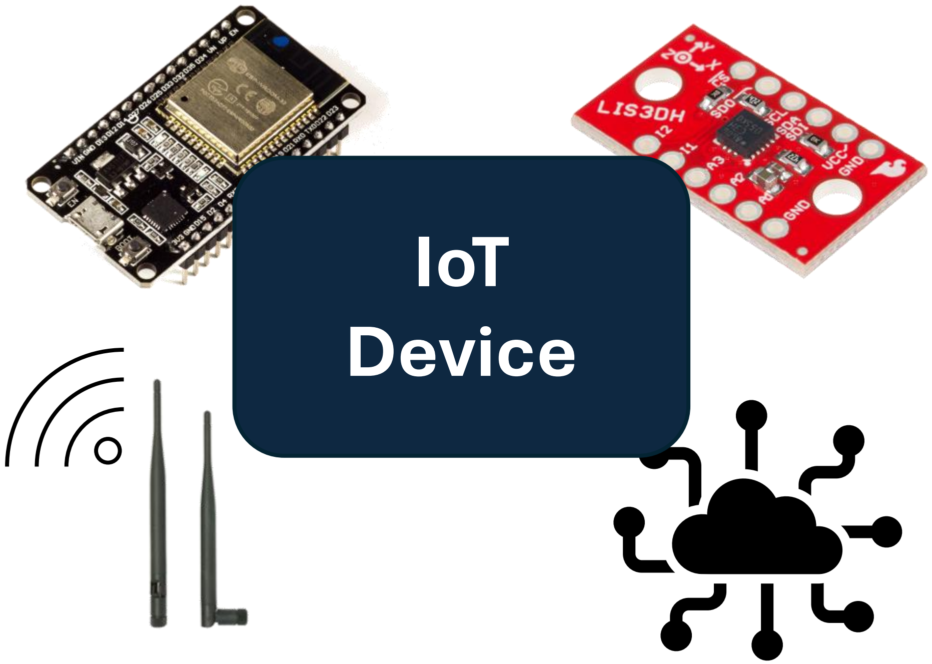
OUTLINE

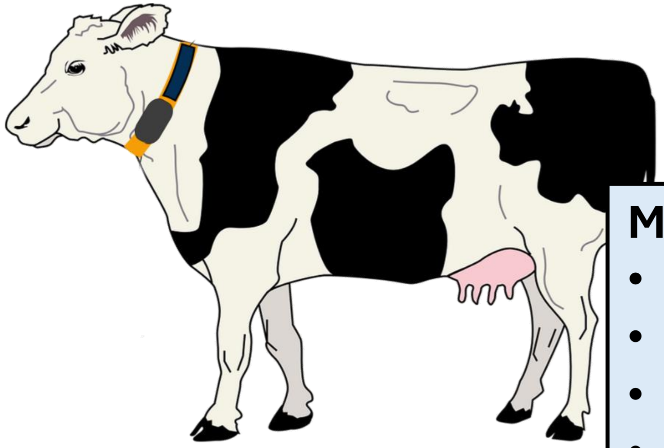
- Introduction
- Related work
- Methodology
  - Wearable IoT device
  - Energy Management System
  - Environment and agent
  - Environment simulation
  - Training
  - Agent evaluation
- Results and Discussion
- Conclusions, Limitations and Future Work

INTRODUCTION

Internet of Things (IoT) Devices

Precision Livestock Farming (PLF)





Monitoring:
- Health
- Fertility
- Stress
- Feeding
- Rumia
- Location
- Rest

INTRODUCTION

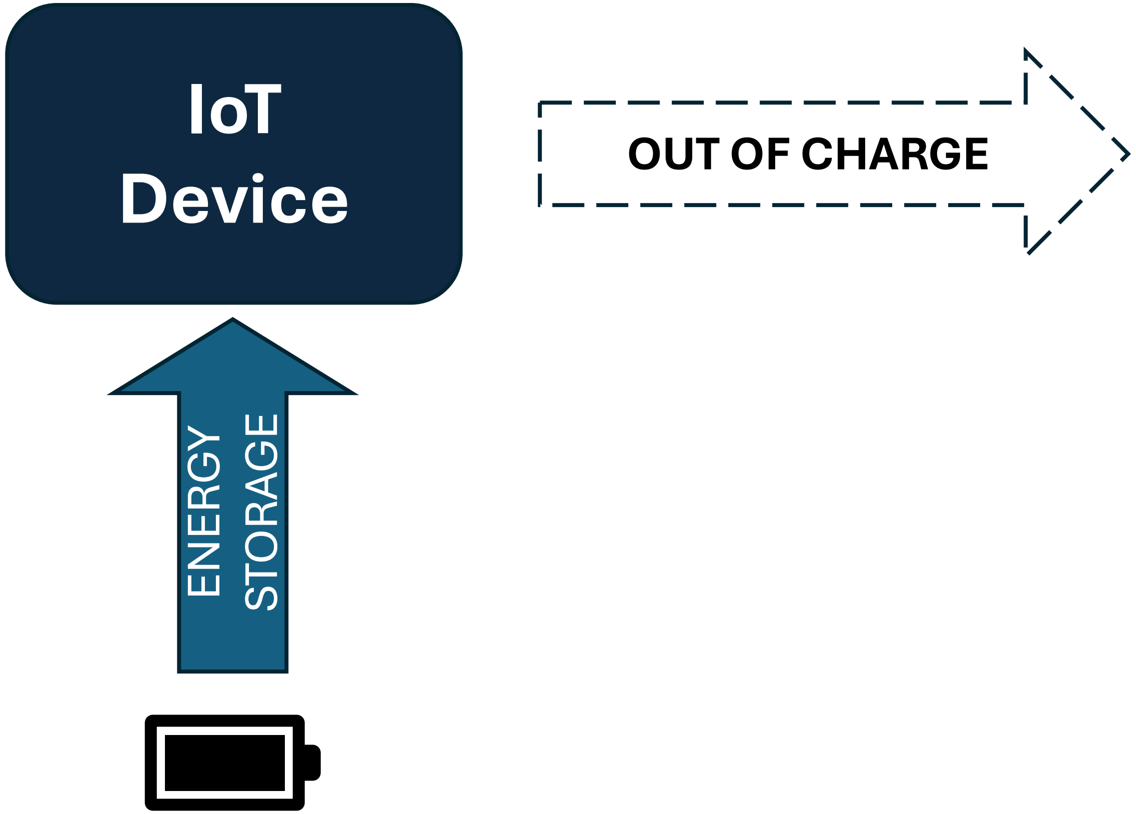
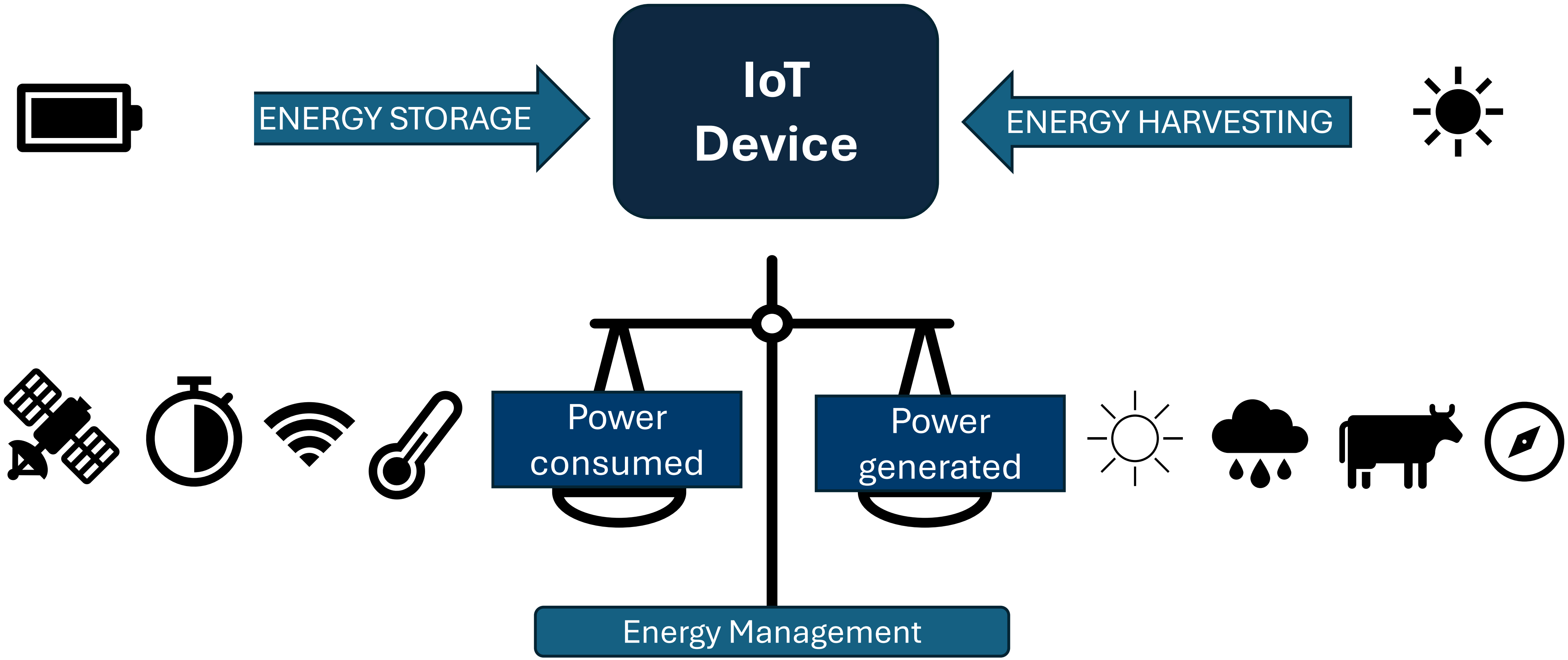
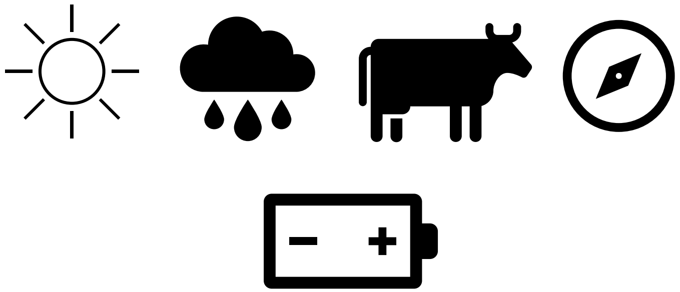


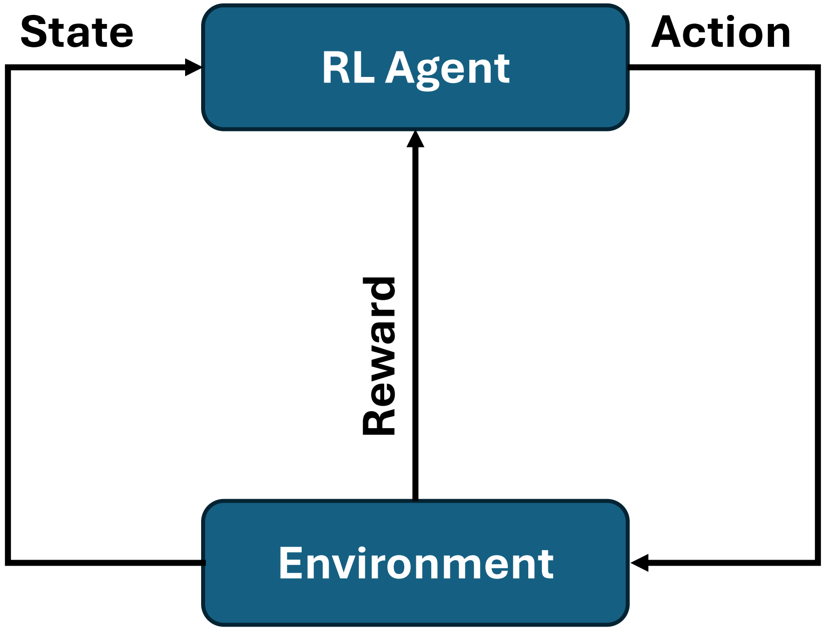
Figure 1: IoT collar replacement process.

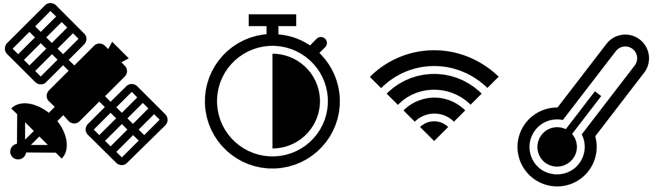
INTRODUCTION



INTRODUCTION







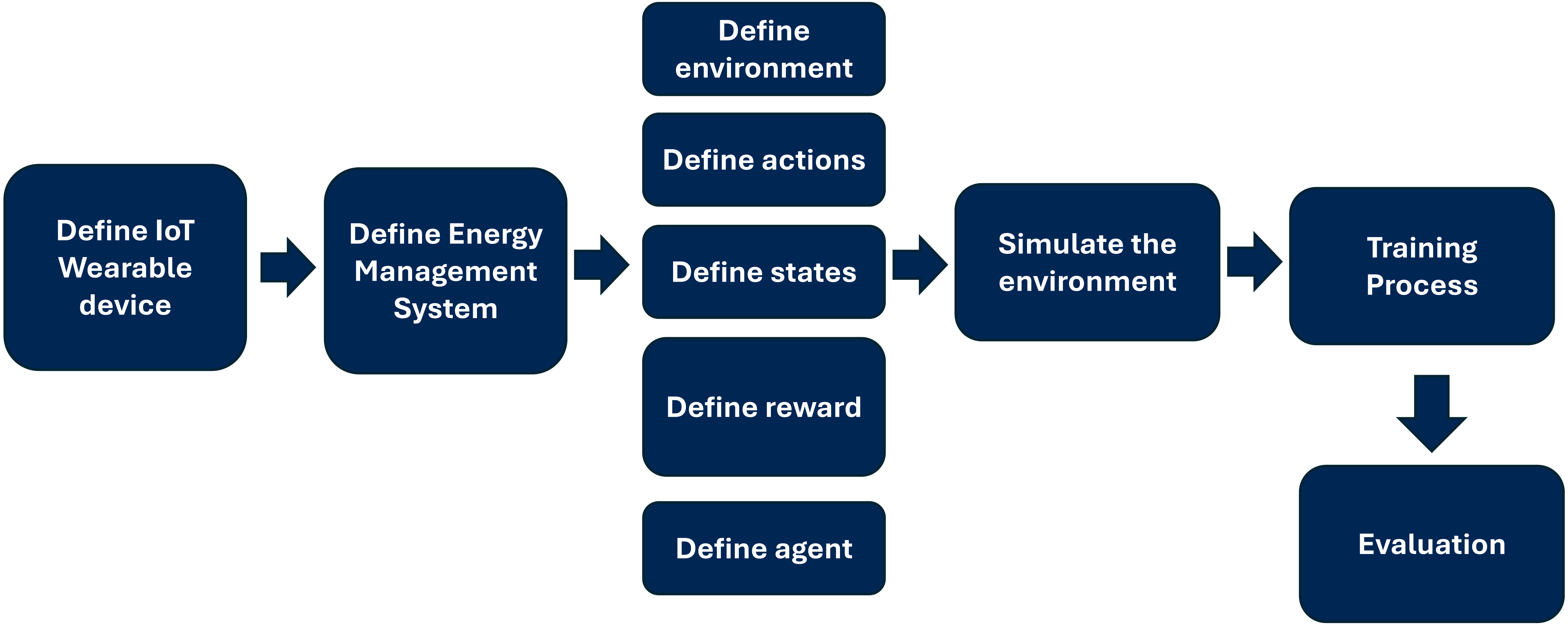
This work aims to demonstrate that an energy manager based on reinforced learning can be used for energy management in a wearable IoT application for livestock monitoring with solar energy harvesting.

RELATED WORK

Table 1: Application of RL in Energy Management for IoT devices.

RL algorithm	Year	Energy harvesting	Application	Ref.
Q-Learning	2021	General	General	[1]
Q-Learning	2021	Solar	General	[2]
Q-Learning	2021	Solar	Mobile	[3]
Q-Learning	2022	Solar	General	[4]
PPO	2022	General	Wearable	[5]
FQL	2022	Solar	General	[6]
Q-Learning	2023	Solar	General	[7]
DQL	2023	TEG	Ambiental	[8]

METHODOLOGY



Wearable IoT device

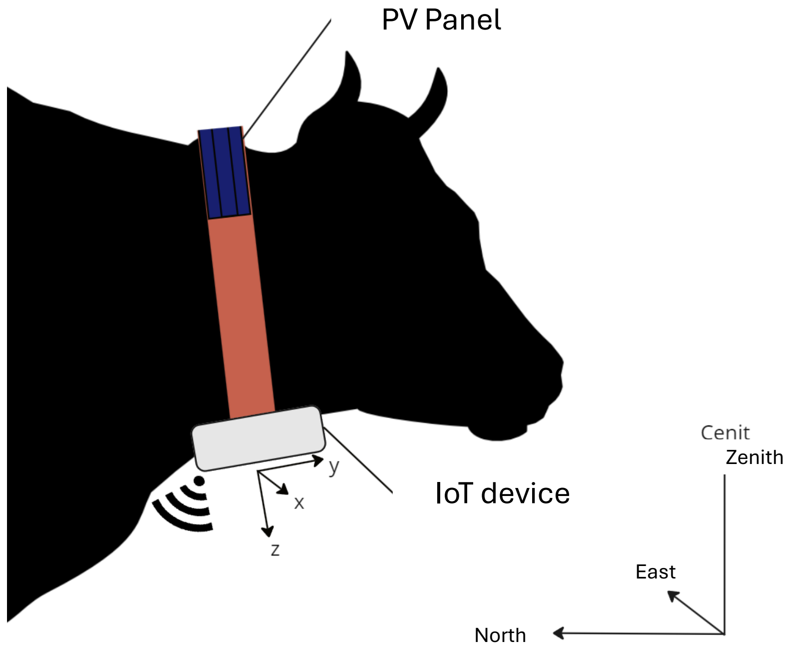


Figure 2: Wearable IoT device.

The device consists of a IoT collar used by dairy cows:

- Microcontroller with LoRaWAN Wireless communication.
- Magnetometer, accelerometer and gyroscope sensors. Sampling frequency 10 Hz.
- GPS
- Local data processing and sending of results.
- Li-ion batteries and flexible photovoltaic panel.

Energy Management System

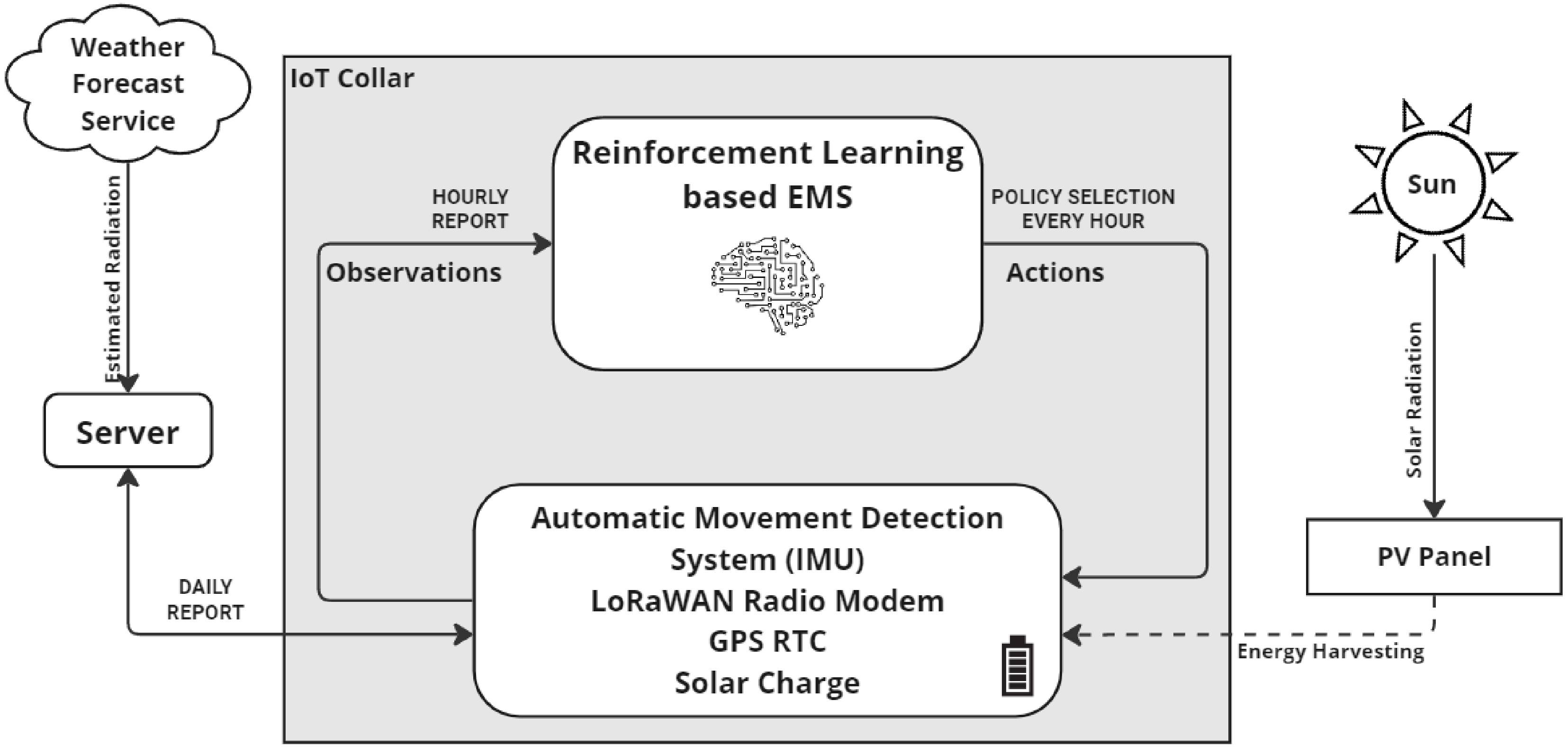


Figure 3: Energy Management System proposed.

Environment and agent

The **environment** is what the agent interacts with, in this case it consists of the energy dynamics present in the wearable device.

- **Energy harvested** from the environment through the solar panel (500-mW).
- **Energy stored** in a battery (Capacity: 10 Wh).
- **Energy consumption:**
  - IMU operation and data processing.
  - Data transmission.
  - GPS location.

Environment and agent

The reinforcement learning agent receives **observations** describing the device and the estimated behavior of the surrounding environment throughout the day.

Table 2: List of observations.

Observation	Range	Physical value
Current battery charge	[0,1]	[0, 100]%
Target battery charge	[0, 1]	[0, 100]%
PV Panel Orientation	[-1, 1]	
Current Radiation	[0, 1]	[0, 1200] W/m2
All-day radiation	[0, 1]	[0, 1200] W/m2
Time of day	[-1, 1]	[0, 24] h

Environment and agent

The reinforcement learning agent takes actions by modifying the duty cycle, data transmission frequency, and GPS execution frequency.

Table 3: List of actions.

Action	Range	Physical value
Duty Cycle	[0.1,1]	[10, 100]%
Transmission frequency	[0.1, 1]	[1, 10] trans./h
GPS frequency	[0.1, 1]	[1, 10] exec./h

A reward function was defined that rewards getting closer to the desired value and penalizes battery discharge.

$$R = 0{,}5 \cdot (|\Delta B| \leq 0{,}02) - 0{,}05 \cdot |\Delta B| - 10 \cdot (B_A \leq 0)$$

Environment and agent

Three agents were defined based on state-of-art algorithms: TD3, PPO and SAC:

- They were implemented in MATLAB R2024a.
- Default hyperparameters were used.
- The discount factor was set to 0.9.
- 32 hidden units were considered.

Environment simulation

A simulation of the environment was implemented in Simulink.

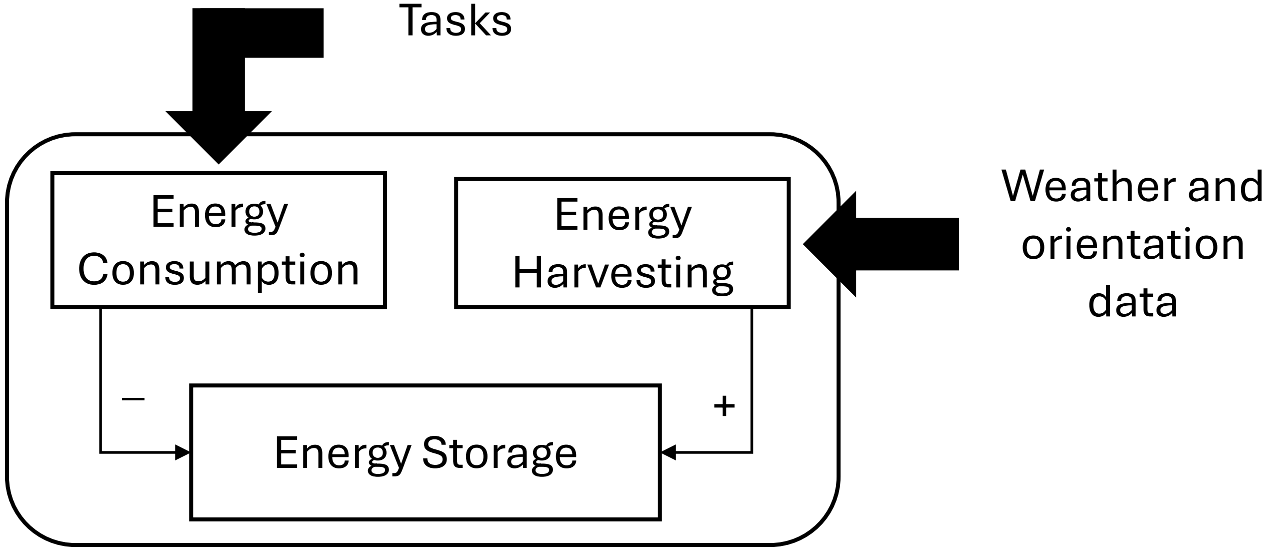
- To estimate consumption, values from and IoT collar capable of acquitting data from an IMU, transmitting wirelessly, and obtaining GPS location were used.

Table 4: Energy consumption of the collar.

Task	Consumption (mW)	Duration	Consumption per execution (mWh)
Active IMU	102	D	-
Inactive IMU	22	1 - D	-
Transmission	254	1.133 s	0.080
GPS exec.	437	48.520 s	5.887

Environment simulation

- To estimate the energy harvesting, a Photovoltaic Generation Model [9] was used, which utilizes solar panel parameters and environmental variables such as radiation, wind speed, ray direction, and the angle of the surface.



Environment simulation

Different data were used for training and evaluation:

Table 5: Data set for training and evaluation.

Datos	Entrenamiento	Evaluación
Datos climáticos	2009 - 2015	2016
Datos de orientación	Variable aleatoria (movimiento del collar)	Datos reales (192 horas de datos)

Figure 4: Location “Campo Experimental Maquehue.”

Training

Characteristics of the training:

- Training during 10 000 episodes (max. 24 steps).
- Initial battery charge: 20% ~ 100%.
- Target battery charge: $\Delta B \leq 10\%$.
- Random panel orientation (average normal vector pointing zenith)
- Solar orientation for a day of the year
- Weather data between 2009 and 2015.

Agent evaluation

Characteristics of the evaluation:

- Initial battery charge B_{ini}: {30%, 60%, 90%}
- Target battery charge B_D: B_{ini} + {-5%, 0, +5%}
- Weather data and solar position: Jan 1 2016 to Dec 25 2016
- Estimated panel orientation from a **real implementation**.

RESULTS AND DISCUSSION

Over the year, the agent based on the TD3 algorithm performed better, with a difference of 4.9% with respect to the target battery charge.

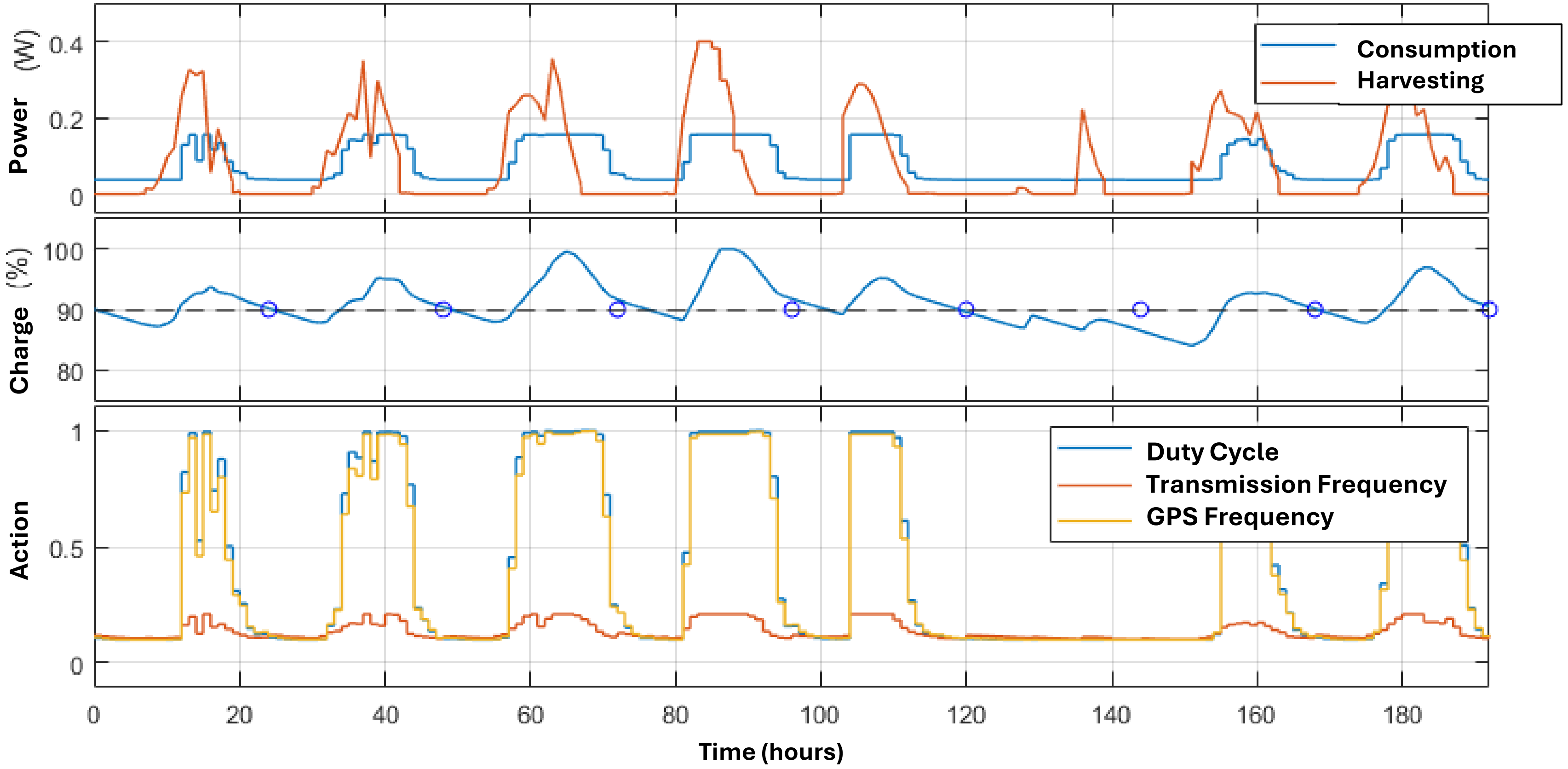
The worst performances were obtained in autumn and winter.

Table 5: Difference between target and current battery charge at the end of the day.

Algorithm	Summer	Autumn	Winter	Spring	Average
TD3	1.56%	7.34%	7.76%	2.58%	**4.91%**
PPO	6.70%	6.88%	7.25%	6.51%	6.84%
SAC	2.44%	9.68%	10.20 %	3.65%	6.62%

RESULTS AND DISCUSSION

Illustrative example of the behavior of the manager based on the RL TD3 agent.



CONCLUSIONS

- The obtained results demonstrate that an energy manager based on reinforced learning can be used for energy management in a PLF wearable IoT application.
- The manager can adapt to the varying conditions experienced by the simulated device, adjusting the power consumption to the environmental and operating variables.
- The applicability of the state-of-the-art algorithms TD3, PPO and SAC was shown, being TD3 the one that obtained a better performance.

CONCLUSIONS

- This work focused on the implementation of this kind of managers in a wearable application for animals, being a subject little treated in the literature.

LIMITATIONS

- This work considers a completely accurate weather forecast, which differs from a real implementation.

- The impact on the consumption of the manager execution was not considered.

FUTURE WORK

- Consider the difference between forecast and actual weather.
- Implement the manager on a real device and evaluate its performance.
- Search for optimization processes of the agent after its implementation, according to new data that will be obtained.
- Try new combinations for observations and actions.

ACKNOWLEDGE

This research was funded by ANID- Proyecto Fondecyt Regular 1220178. Daniel Iglesias' master studies are funded by ANID - Subdireccion de Capital Humano / Magíster Nacional / 2024-22241455.

Thank you!

REFERENCES

[1] Y. Kim and T. J. Lee, "Learning nodes: machine learning-based energy and data management strategy," Eurasip Journal on Wireless Communications and Networking, vol. 2021, no. 1, pp. 1–16, 12 2021. [Online]. Available: https://link.springer.com/articles/10.1186/s13638-021-02047-6 https://link.springer.com/article/10.1186/s13638-021-02047-6

[2] M. Prauzek and J. Konecny, "Optimizing of Q-LearningDay/Night Energy Strategy for Solar Harvesting Environmental Wireless Sensor Networks Nodes," Elektronika ir Elektrotechnika,vol. 27, no. 3, pp. 50–56, 6 2021. [Online]. Available: https://eejournal.ktu.lt/index.php/elt/article/view/28875

[3] M. Prauzek, T. Paterova, J. Konecny, and R. Marti-nek, "Data-Driven Self-Learning Controller for Power-Aware Mobile Monitoring IoT Devices," Computers, Materials &Continua, vol. 70, no. 2, pp. 2601–2618, 9 2021. [On-line]. Available: https://www.techscience.com/cmc/v70n2/44649/htmlhttps://www.techscience.com/cmc/v70n2/44649

[4] F. Namjoonia, M. Sheikhi, and V. Hakami, "Fast reinforcement learning algorithms for joint adaptive source coding and transmission control in IoT devices with renewable energy storage," Neural Computing and Applications, vol. 34, no. 5, pp. 3959–3979, 3 2022. [Online].Available: https://springerlink.ufro.elogim.com/article/10.1007/s00521-021-06656-6

[5] T. Basaklar, Y. Tuncel, and U. Y. Ogras, "tinyMAN: Lightweight Energy Manager using Reinforcement Learning for Energy Harvesting Wearable IoT Devices," 2 2022. [Online]. Available: http://arxiv.org/abs/2202.09297

[6] R. C. Hsu, T. H. Lin, and P. C. Su, "Dynamic Energy Management for Perpetual Operation of Energy Harvesting Wireless Sensor Node Using Fuzzy Q-Learning," Energies 2022, Vol. 15, Page 3117, vol. 15, no. 9,p. 3117, 4 2022. [Online]. Available: https://www.mdpi.com/1996-1073/15/9/3117/htm https://www.mdpi.com/1996-1073/15/9/3117

[7] H. Malekijou, V. Hakami, N. T. Javan, and A. Malekijoo, "Q-learning-based algorithms for dynamic transmission control in IoT equipment," Journal of Supercomputing, vol. 79, no. 1, pp. 75–108, 1 2023. [Online]. Available: https://springerlink.ufro.elogim.com/article/10.1007/s11227-022-04643-9

[8] M. Prauzek, J. Konecny, and T. Paterova, "An Analysis of Double Q-Learning-Based Energy Management Strategies for TEG-Powered IoT Devices," IEEE Internet of Things Journal, vol. 10, no. 21, pp. 18 919–18 929, 11 2023

[9] A. Molina and F. Martínez, "Modelo de Generación Fotovoltaica," 3 2017. [Online]. Available: https://solar.minenergia.cl/downloads/fotovoltaico.pdf