PowerTech Madrid 2021
Paper No: 101

Imperial College London

Self-learning Control for Active Network Management

Julio Pérez Olvera, Timothy C. Green,
Adrià Junyent-Ferré

Imperial College London

Power & Energy Society®

This work was supported by the Mexican Energy Ministry (SENER) and the National Council for Science and Technology (CONACYT) through the PhD Scholarship number: 441647, awarded to J.P.O.

Outline

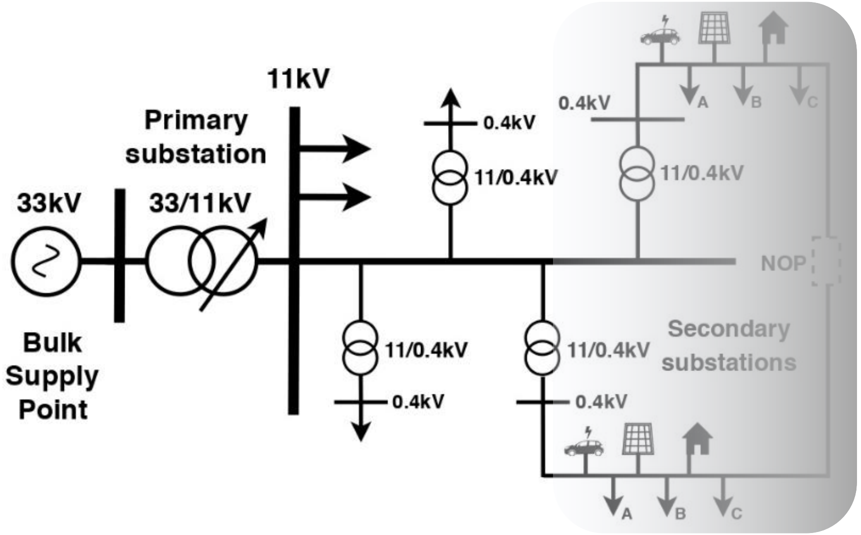
1. Challenges in distribution networks
2. Active Network Management (ANM) for distribution networks
3. Reinforcement learning for Active Network Management (ANM)
4. Case study
5. Conclusions

Challenges in distribution networks

❑ Large number of substations and customers in LV networks compared to transmission

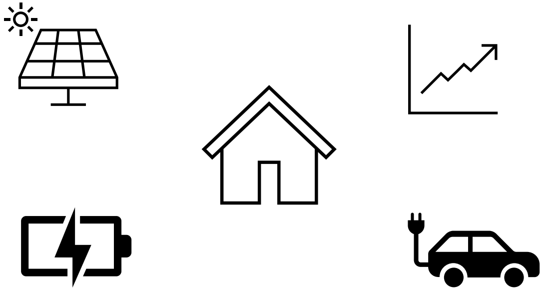
❑ Lack of SCADA monitoring and low visibility in LV networks

❑ LV connectivity information not always available to distribution network operators (DNOs)



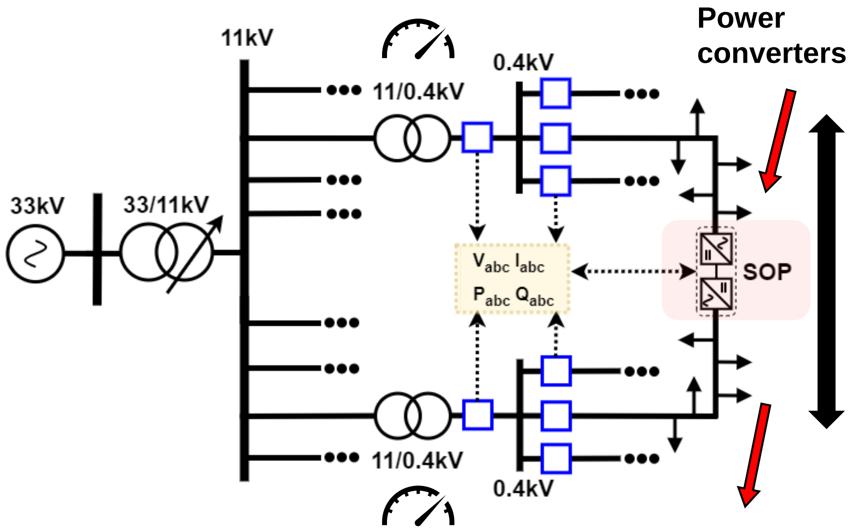
❑ Integration of low-carbon technologies (LCTs):

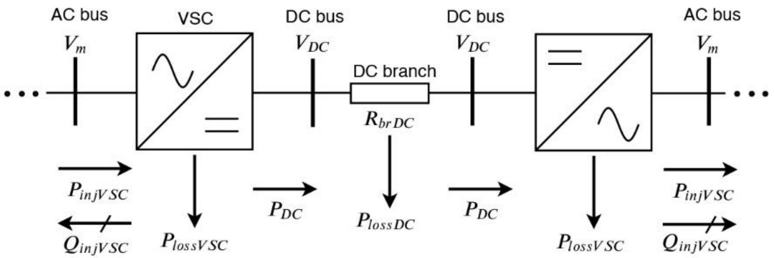
- **Solar PV** generation is highly variable, it can mask underlying load and cause voltage problems
- **Electric vehicles (EVs)** can cause a significant increase in demand and large power fluctuations



Active Network Management (ANM) for distribution networks

- **Power-electronics converters** can be used to mesh distribution networks
- Converters connected back-to-back through their DC bus, as a **soft-open-point (SOP)**
- The power flow can be regulated at all times
- Balance power flows between phases
- ANM can be used to distribute the power from highly loaded points to underutilised parts of the network
- More efficient use of the existing infrastructure
- **A control scheme is required to calculate the optimal set points for the converters**





IEEE PES Power & Energy Society® IEEE

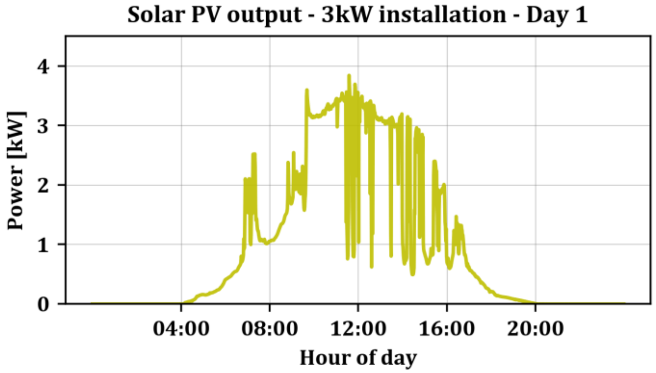
Active Network Management (ANM) for distribution networks

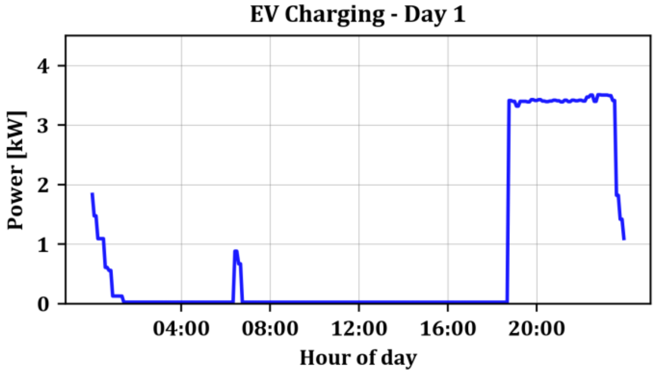
Designing a control that act optimally is challenging:

- ❑ Large number of substations and customers
- ❑ Lack of SCADA monitoring and low visibility
- ❑ High variability of low-carbon technologies (LCTs)

One option is using the Optimal Power Flow (OPF):

- ❑ Nonlinear optimisation approach that requires complete visibility of the network
- ❑ Can be computationally intensive for large networks
- ❑ **Optimizing the distribution of the power for every possible scenario using an OPF is not viable for real-time operation, especially when including LCTs**

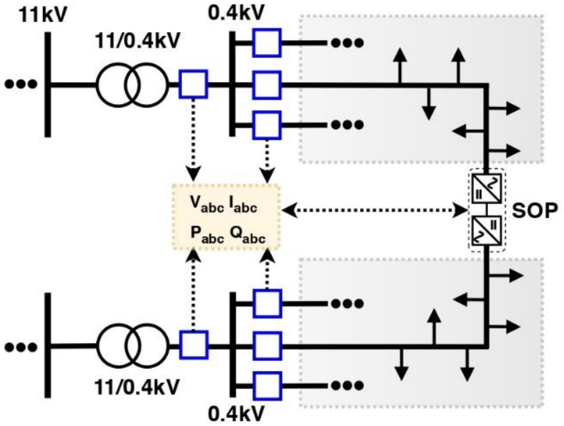


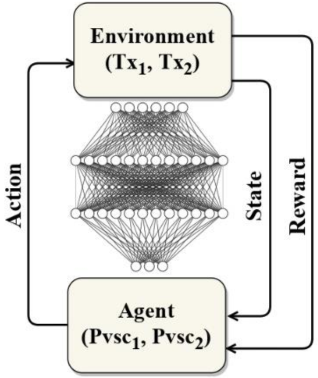


Reinforcement learning for Active Network Management

Using **reinforcement learning** we can:

- ❑ Solve problems involving a sequence of optimal decisions
- ❑ Train an agent to act optimally from repeated interactions with a model of the network
- ❑ **Deal with partial observability, only using available measurements**



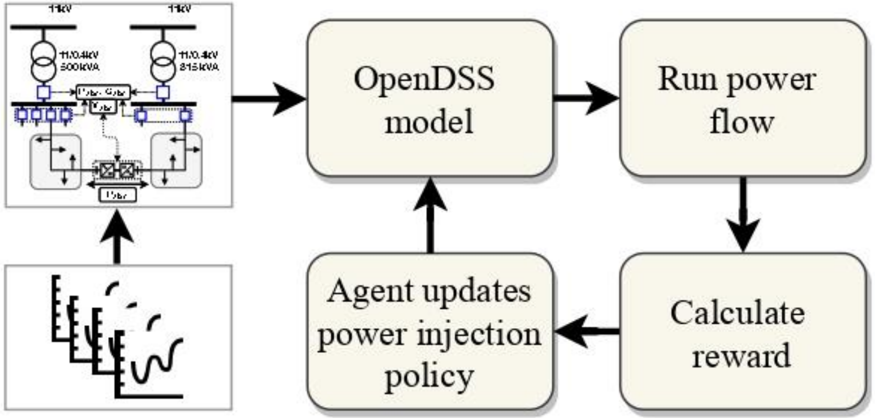


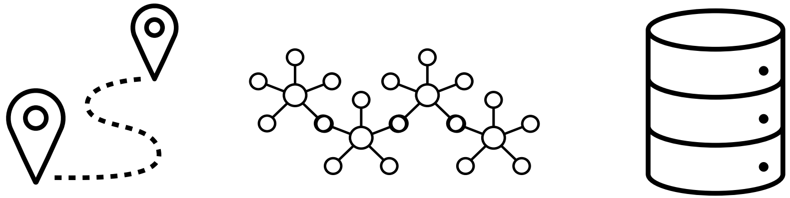
- ❑ The **state** comprises a series of network measurements, those available to the DNO:
  - Power, loading and losses of each **transformer**
  - Voltage at the terminals of each **converter**
- ❑ The **action** represents the three-phase power injection of the converters
- ❑ The **reward** function represents the **goal** of the agent controlling the converters: **to equalise the loading of two LV substations while minimizing the power losses**

Reinforcement learning for Active Network Management (ANM)

Model-free reinforcement learning:

- **Q-learning with neural networks** can be used to achieve near-optimal policies by interacting with a black-box model, using a **reward function**
- Interactions via power flow simulations**, which are computationally cheap**
- The optimal policy is obtained by following an **exploration policy** (ϵ-greedy)
- Deep Q Networks (DQN) in a two-stage format:
  - **Offline training** of neural network using on historical load profiles and network model
  - **Online recall** with trained neural network

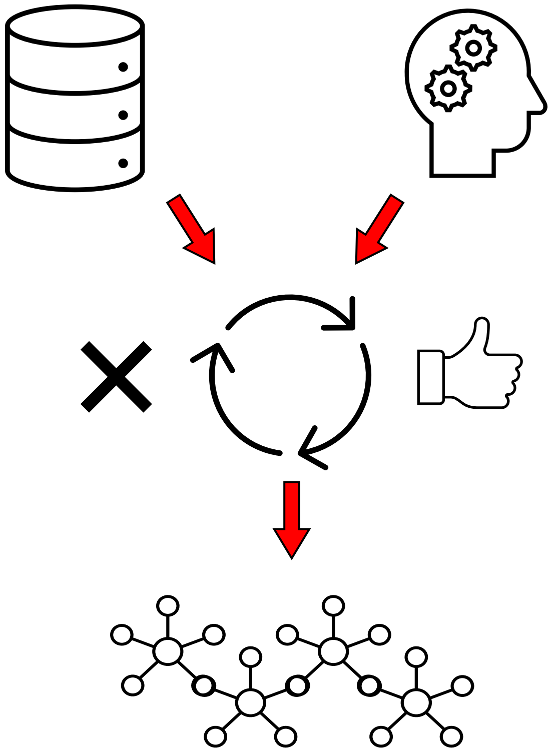




Reinforcement learning for Active Network Management (ANM)

Episodes and training:

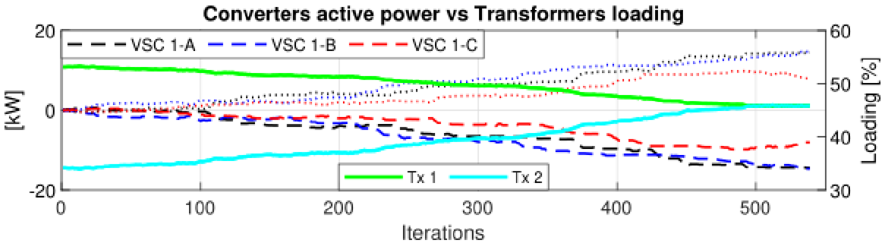
- ❑ During training, the agent stores the observed transitions in a **memory buffer**
- ❑ The **neural network** is trained using a batch of past observations from the memory buffer making an **efficient use of the data generated**
- ❑ A **reward** is given for actions resulting in improved loading equalisation or reduced power losses, otherwise a **penalty** is given
- ❑ The agent interacts with the environment for as many iterations as needed, until reaching a **termination state**
- ❑ Each minute is treated as a separate training episode, **allowing the agent to explore**

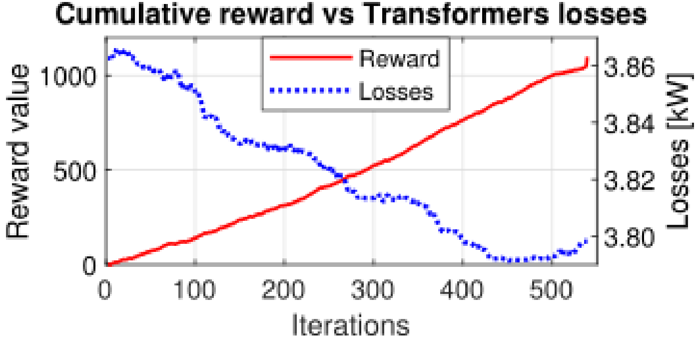


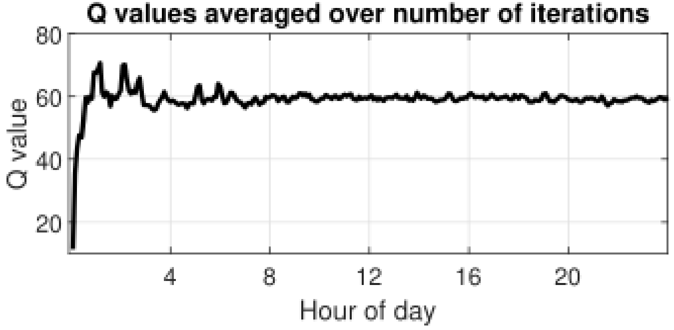
Case study

Results from the case study:

- ❑ The agent iteratively updates the power injections until loading equalisation is achieved
- ❑ The agent needs to balance the goals of
  - Achieving loading equalisation and
  - Lowering the power losses,
  - Loading equalization takes priority
- ❑ Evolution of the Q values:
  - Low value at start
  - Increasing as the agent learns
  - Settling around 60, the average cumulative reward the agent receives when meeting the goal





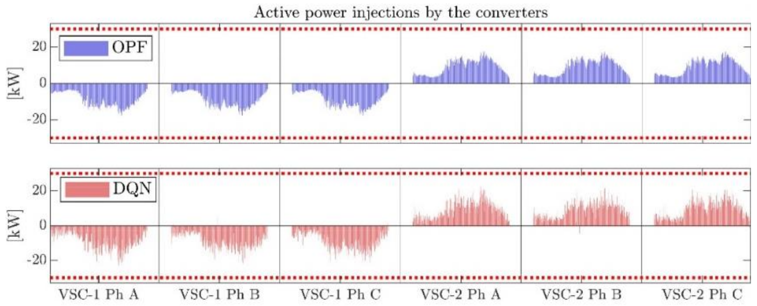


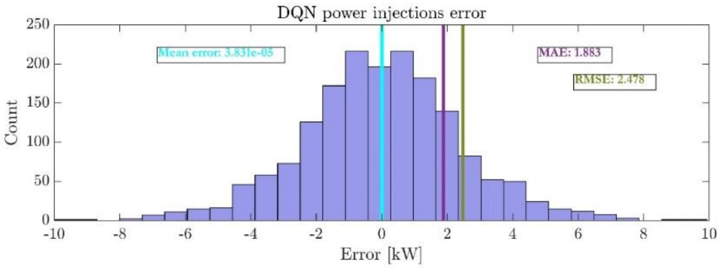
Case study

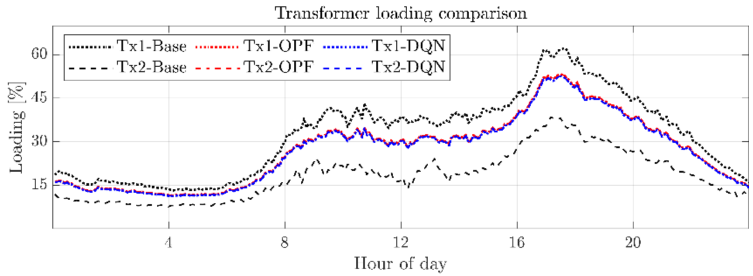
Results from the case study:

- ❑ Comparison of resulting power injections across all 288 time-steps, for the OPF and DQN
  - **The DQN algorithm agrees with the OPF**
  - Difference between the two (error) follows a normal distribution with a mean around zero.
  - Overall, the differences are less than ±4 kW

- ❑ The proposed control equalises the loading just as effectively as an OPF



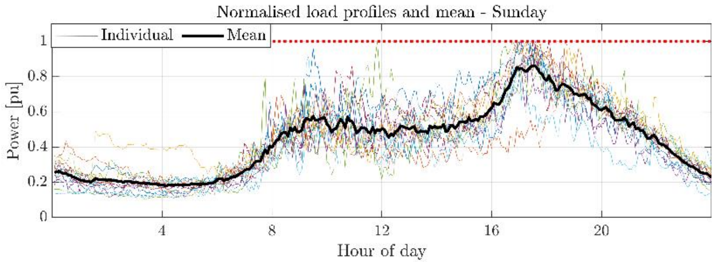


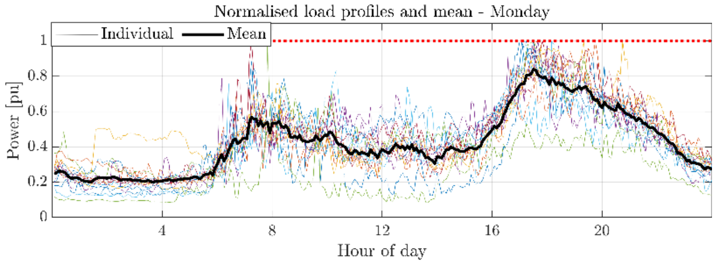


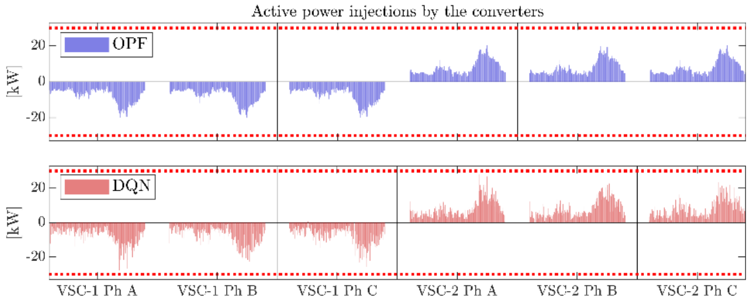
Case study

Results from the case study:

- ❑ The ability of the trained agent to generalise is validated on a load profile of a different day
  - The agent is trained using load profiles of **Sunday**
  - The trained agent is validated on load profiles of **Monday**
- ❑ The trained agent produces similar power injections as the OPF
- ❑ Unlike the OPF, it does not require a new optimization for the different set of loading conditions







IEEE PES Power & Energy Society® IEEE

Conclusions

- DNOs will need to adopt ANM schemes to cope with the integration of LCTs
- ANM using power electronic converters requires a suitable control strategy for real-time operation that can deal with partial visibility
- The OPF is computationally demanding and requires complete visibility, thus not a suitable option
- Reinforcement learning can be used to train an agent to make optimal decisions based on interactions with a network model, requiring only partial visibility
- The training uses power flow calculations, as it is computationally efficient, compared to the OPF
- The neural network enables the agent to operate with partial observability and to generalise
- After training the agent can be used online with only the available measurements as input
- Model-free algorithms show a performance comparable to traditional optimisation and they can generalise to different loading conditions

Thank you for your attention!

Questions?

Contact:

j.perez-olvera@imperial.ac.uk
juliioolvera1@gmail.com

The authors are grateful to Western Power Distribution (WPD) and EA Technology Ltd for their support and for granting access to the OpenLV project data to undertake this study. More information: OpenLV.net