Implementing Reinforcement Learning via Markov Decision Process (MDP) for Wind Shelter Modelling: A Precursor to the Dynamic Line Rating (DLR) Technology Trial in Queensland's Transmission Network

Rudolf Espada (Australia)

Key words: Implementation of plans

SUMMARY

Embracing the purpose of connecting Queenslanders to a world-class energy future, Powerlink Queensland (PQ) is the state’s leading provider of high voltage electricity transmission services. To support the rapidly expanding renewable energy sector in the State, there is a need to improve thermal calculations for PQ’s overhead transmission lines using dynamic line rating (DLR) sensor technology. As a test, this case study explores a methodology for geographically locating critical spans that are heavily wind-sheltered along the 1,700 km of network/feeder from northern Cairns to the New South Wales (NSW) border. We examined the use of Reinforcement Learning, a type of Machine Learning, specifically through the framework of Markov Decision Process (MDP) for spatial modelling of wind-sheltered ground spans. PQ infrastructure datasets (i.e., built section, conductor height, & ground span) and resampled 10m pixel data from externally sourced datasets (i.e., network-wide vegetation height, wind(W), aspect(A), & relative topographic position(Rtp)) were used in the modelling. At the ground span level, the pixel centroids of the vegetation height were generated. To account for maximum conductor operating temperature, the conductor heights that are associated with the minimum vertical clearance were isolated from the entire conductor dataset. After computing the slope (%) between the vegetation and conductor heights, the wind shelter score (%) per ground span was calculated and then reclassified into five categories: low(0-25%), light(26-40%), moderate(41-75%), high(76-95%), and extreme(>95%). To find the optimal critical locations at the pixel level, we configured the MDP 4-tuple mathematical framework, i.e., [current/new states(si/si’), action/assigned criticality(ai), reward R(si,ai,si’), and transition probability T(si,ai,si’)]. The wind shelter score from above was introduced as the current state (si) variable. Using the ECMWF-ERA5 36 years(1985-2020) of eastward and northward horizontal wind speed components at 10m altitude, the mean monthly wind direction was calculated. The probability of the wind being deflected was then derived after applying the fuzzy Gaussian distribution function. To account for the influx of solar energy, the probability of pixel’s

compass direction was also generated from the aspect data. Furthermore, the probability of terrain evenness/ruggedness was calculated after applying the fuzzy near distribution function to local/landscape scale of Rtp data. Then the joint probability of wind score and WARtp probabilities was derived to generate T(si,ai,si'). An R(si,ai,si') of -10 with low-, -5 with light-, 0 with moderate-, +5 with high-, and +10 with extreme-criticality(si') were assigned to low, light, moderate, high, and extreme T(si,ai,si'), respectively. To find the expectimax value that maximises R(si,ai,si') per pixel, the Bellman equation and its iterative operations with a discounting factor of 0.9 were implemented in the R environment. Accuracy assessment showed 99.33% capture ratio of ground spans. Moreover, the MDP algorithm showed a 1:1 correspondence between the pre-assigned criticalities and resulting optimal criticalities, although this is subject to interpretation and decision-making appetite. Finally, we successfully identified extremely wind-sheltered and optimal critical areas in PQ's transmission networks, both at the ground span and pixel levels. These findings provide initial baseline information for DLR sensor site installation, number of sensors to install, and the operational costs involved.