Using stacked deep learning models to produce accurate historical precipitation records over Aotearoa New Zealand

William Naylor[1], Emily O'Riordan[1], Greg Bodeker[1], Greg Pearson[2], and Lexi Xu[1]

[1]Bodeker Scientific, Alexandra, [2]MetService, Wellington

Contact: emily@bodekerscientific.com

The Problem

- Climate change increases the frequency and intensity of extreme precipitation events, driving **demand for high-resolution, long-term precipitation data** for risk assessment and planning.
- While weather stations provide precipitation records, they're limited in coverage and consistency; reanalyses fill this gap by assimilating observations from various sources into a model run.
- ECMWF's ERA5 (0.25° resolution) and ERA5-land (0.1° resolution) reanalyses offer global precipitation data, but known biases affect their accuracy in certain regions, including New Zealand[1].
- MetService's **Quantitative Precipitation Estimation (QPE)** product offers a **higher-resolution, more accurate precipitation record (1 km)** for Aotearoa by combining data from a range of sources such as radar, stations and forecast models. However, **QPE data are available only from 2021 onward** and therefore do not provide the desired long-term, high-resolution record.

QPEnet

- **QPEnet is a deep learning model that transforms ERA5 atmospheric variables into high-resolution, QPE-like precipitation fields, matching both the intensity and location of precipitation events while downscaling from 9-30km to 1km resolution for Aotearoa New Zealand, covering 1950 to 2023.**
- Two networks are stacked together: a **ResUNet++ predicts precipitation location and intensity** at 9km resolution, and a Super Resolution Generative Adversarial Network **(SRGAN) generates the downscaling to 1km.**
- QPEnet addresses the limitations in both QPE and ERA5 for Aotearoa New Zealand's precipitation records by **enhancing the accuracy of ERA5** and **extending the coverage of QPE** back to 1950.
- The QPEnet model generates an historical, QPE-like data set over Aotearoa New Zealand, providing **a valuable resource for long-term precipitation analysis and data-driven forecasting**.

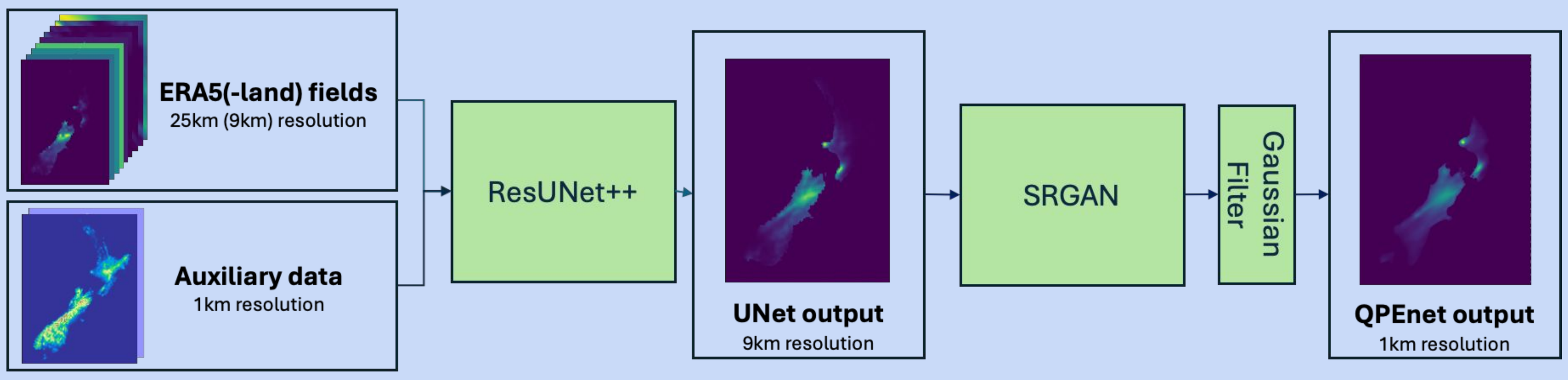

Figure 1: Schematic of the QPEnet model architecture including its inputs (ERA5 and auxiliary variables) which are fed into the ResUNet++ (a deep residual UNet convolutional network) to produce a coarse QPE-like field, which in turn is processed by a Super Resolution Generative Adversarial Network (SRGAN), followed by a Gaussian filter. The QPEnet model output is a high resolution precipitation field. A stacked two model approach was adopted, with ResUNet++ handling location/intensity prediction and SRGAN downscaling output to 1 km. This architecture performed better than a single convolutional model.

QPEnet Performance

- QPEnet's performance was compared to ERA5, ERA5-Land, and BARRA2 (a 12 km resolution data set over Oceania) by downscaling each to 1 km for consistency. The RainFARM downscaling algorithm[2] was used to create finer spatial resolutions by introducing variability at smaller scales through random Fourier phases.
- The fractional skill score (FSS) was used to evaluate **prediction accuracy within a neighbourhood area** rather than at single pixels. This more robust against location errors than pixelwise metrics (e.g., mean squared error), which are prone to the "double penalty" problem.
- **QPEnet demonstrated higher accuracy than ERA5-Land^D and approached the skill level of BARRA2^D**, especially for hourly precipitation data (Figure 2). BARRA2 assimilates a lot more data than ERA5 and so is expected to outperform the global reanalysis.
- When evaluated on 24-hour accumulations (Figure 3), QPEnet continued to outperform ERA5-Land^D and matched BARRA2^D's accuracy, especially for higher rainfall events.

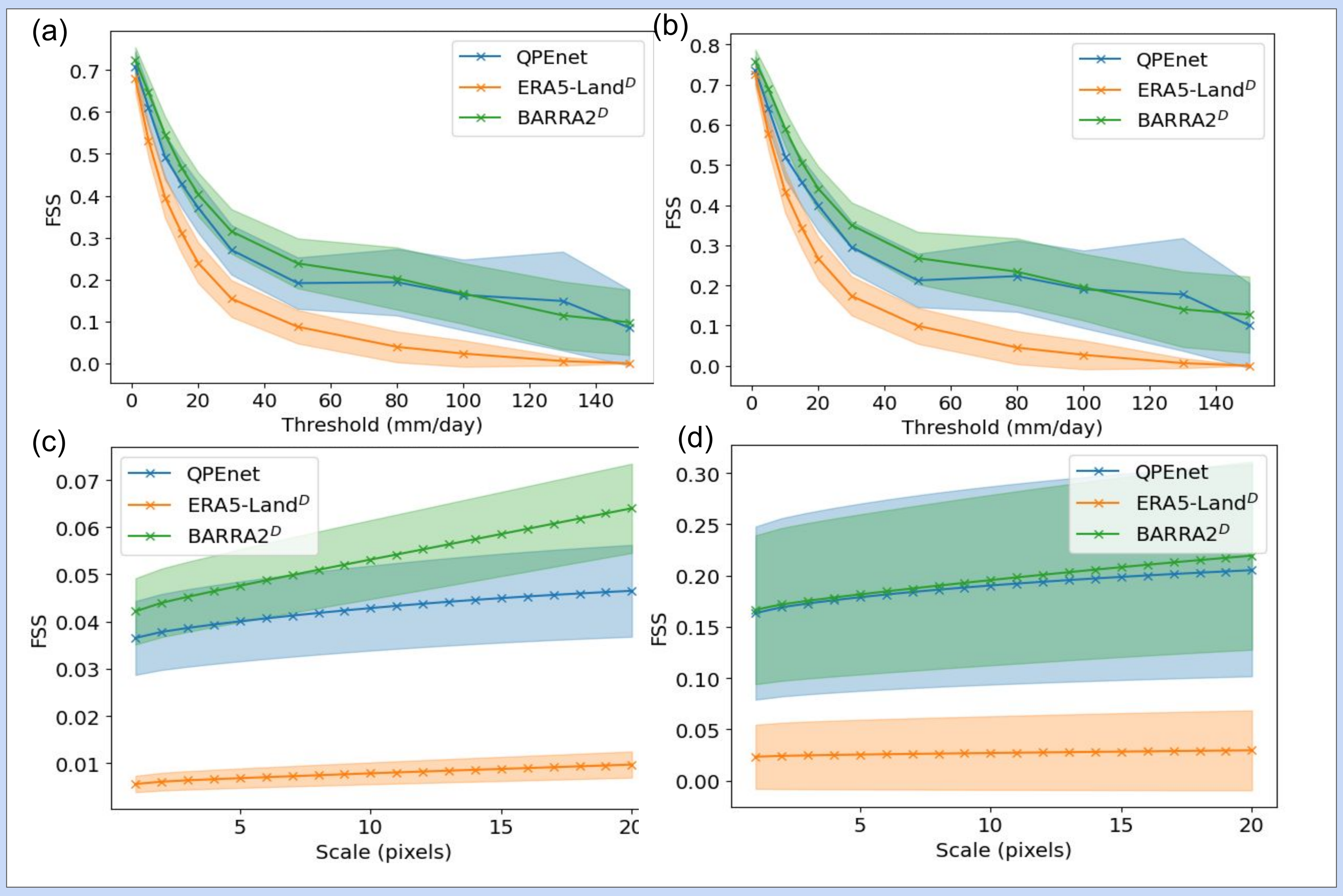

Figure 3: Fractional skill scores (FSS) for 24-hour precipitation accumulations for QPEnet, ERA5-Land^D and BARRA2^D, using the QPE test data as the truth. ^{D} denotes downscaled datasets. (a) with scale 1 pixel, (b) with scale 10 pixel, (c) with threshold 10 mm/hr and (d) with threshold 100 mm/hr.

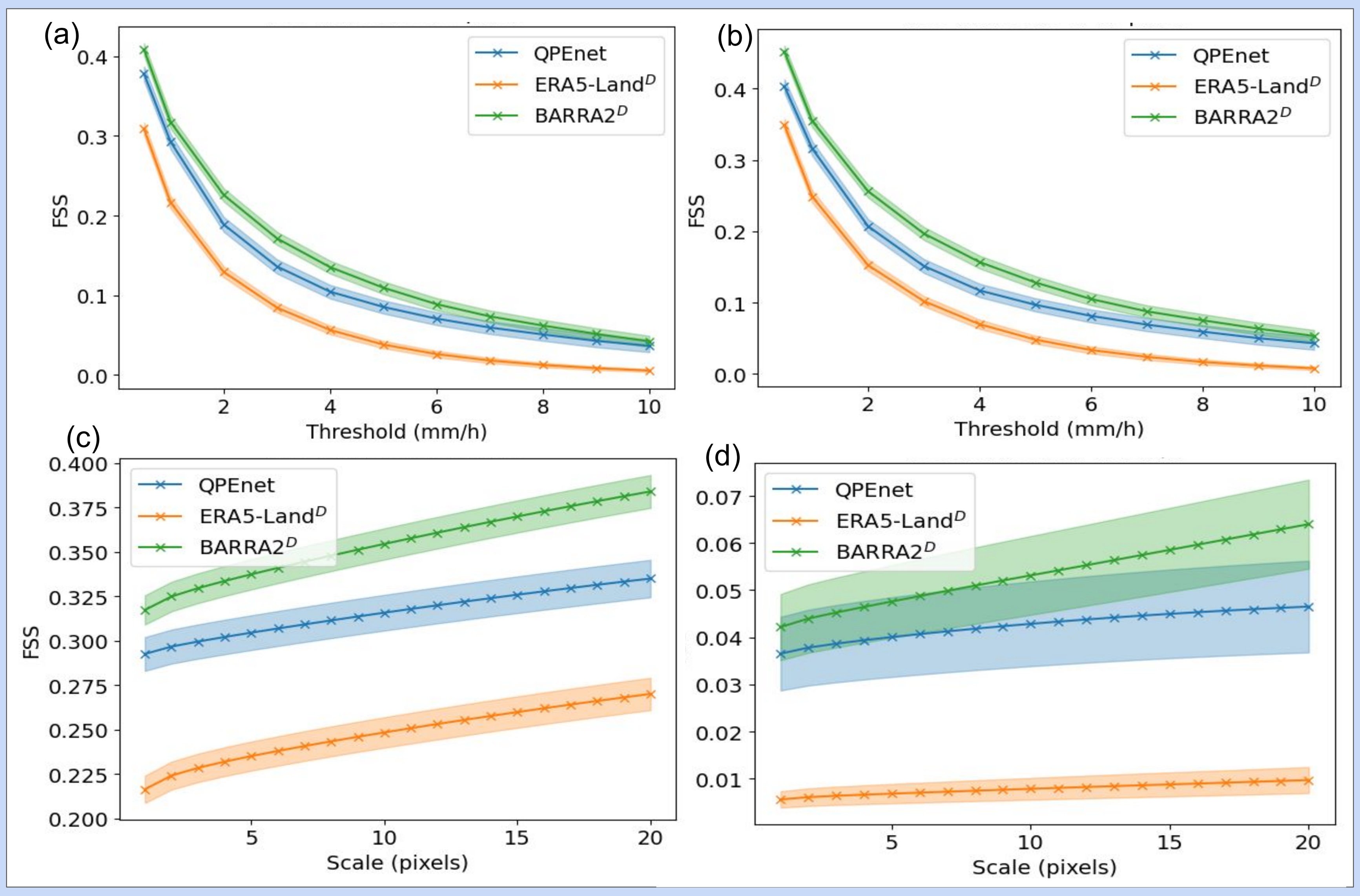

Figure 2: Fractional skill scores (FSS) of hourly QPEnet, ERA5-Land^D and BARRA2^D data, using the QPE test data as the truth data set. ^{D} denotes downscaled datasets. (a) with scale 1 pixel, (b) with scale 10 pixel, (c) with threshold 1 mm/hr and (d) with threshold 10 mm/hr.

Comparison of extreme precipitation depths

- **24-hour extreme precipitation depths** for a **100-year annual recurrence interval** (ARI) are compared between original ERA5 and QPEnet data. These are also compared to NIWA's High Intensity Rainfall Design System (HIRDS) field - see Figure 4.
- Using 1-year block maxima (annual maximum daily precipitation) derived from 24-hour rolling totals for each grid cell, Generalized Extreme Value (GEV) distributions were fitted to describe precipitation extremes. The cumulative distribution function resulting from the GEV fit describes precipitation depth relative to ARI, with linear interpolation used to determine 100-year precipitation depths for each cell.
- **QPEnet better represents the extreme precipitation values** and aligns more closely with HIRDS data than ERA5, **capturing spatial features in areas such as the Southern Alps, Mount Taranaki, North East Cape, Gisborne, and Coromandel Peninsula**.
- These findings indicate that QPEnet effectively enhances the fidelity of ERA5 data, improving its ability to capture precipitation extremes.

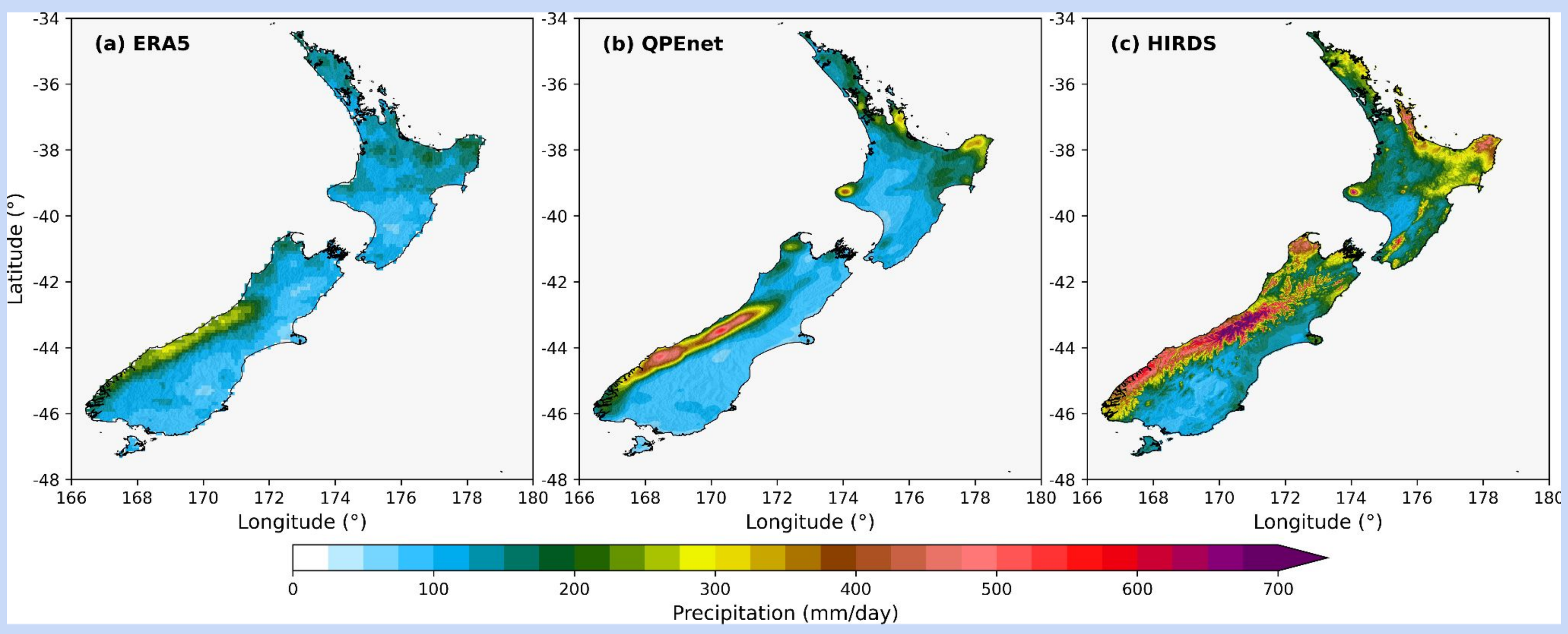

Figure 4: A comparison of 1-in-100 year precipitation depths derived from: (a) original ERA5 precipitation fields, (b) QPEnet precipitation fields, and (c) the precipitation field extracted from the High Intensity Rainfall Design System (HIRDS) data set.

Technical Notes

- ResUNet++ and SRGAN implemented in Pytorch.
- Loss function for all models: land-masked mean squared error.
- Adam optimizer, StepLR scheduler, learning rate of 10^{-5}.

Conclusion and Outlook

- **Deep learning can be used to emulate high fidelity meteorological products at low computational cost,** even with only a small volume of high quality meteorological data available to train with.
- **QPEnet approaches the accuracy of BARRA2** (a dynamical reanalysis that assimilates far more observational data for Aotearoa New Zealand than ERA5) at both the daily and the hourly scale, and significantly **improves the fidelity of ERA5**.
- For high rainfall events, QPEnet creates **more realistic extreme precipitation depths** than those produced by ERA5.
- Further investigation into deep learning architectures and loss functions may provide improved results in the future.

Acknowledgements

Ministry of Business, Innovation & Employment
HĪKINA WHAKATUTUKI

We acknowledge the New Zealand Ministry of Business, Innovation and Employment (MBIE) for funding the Deep Weather project under contract number BSCIF2201.

References

1. Pirooz, A. A. S., et al. "Evaluation of global and regional reanalyses performance over New Zealand." Weather and Climate 41.1 (2021): 52-71.
2. Pulkkinen, Seppo, et al. "Pysteps: An open-source Python library for probabilistic precipitation nowcasting (v1. 0)." Geoscientific Model Development 12.10 (2019): 4185-4219.