Understanding Precipitation Estimates from Commercial Microwave Links

Heron Ng, Hydro- JULES intern 2025

Department of Meteorology, University of Reading

DAMTP, University of Cambridge

Supervisors: John Wallbank (UKCEH), Steven Cole (UKCEH), Robert Moore (UKCEH), David Dufton (NCAS)

(Vodafone, 2025)

Project Objectives

- Understand the **current research work** on Commercial Microwave Links (CMLs) for precipitation estimation in literature
- Systematically process, clean, evaluate, and plot data
- **Pilot data** received from Vodafone for a **6-month period** (Jul 2024 – Jan 2025)
- Establish communication with Vodafone, NCAS and UKMO

Hydro- JULES project Overview: Understanding Precipitation Estimates from CMLs

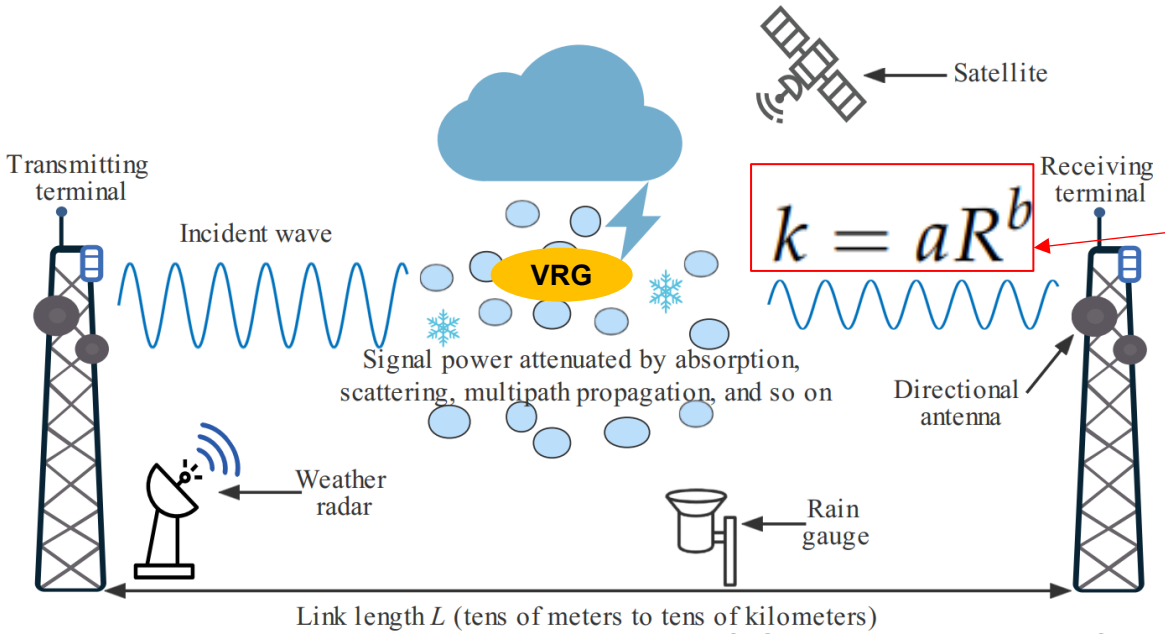


Figure: Basic operating principle of CML-based rainfall measurement (Lian et al., 2022)

VRG Virtual Rain Gauge = Midpoint of each CML path

Advantages

- High spatial density in urban areas
- Path close to ground surface

- CMLs are used around the world as part of the mobile phone network
- Signal attenuation increases with precipitation intensity
- Assessed using rain gauge and radar

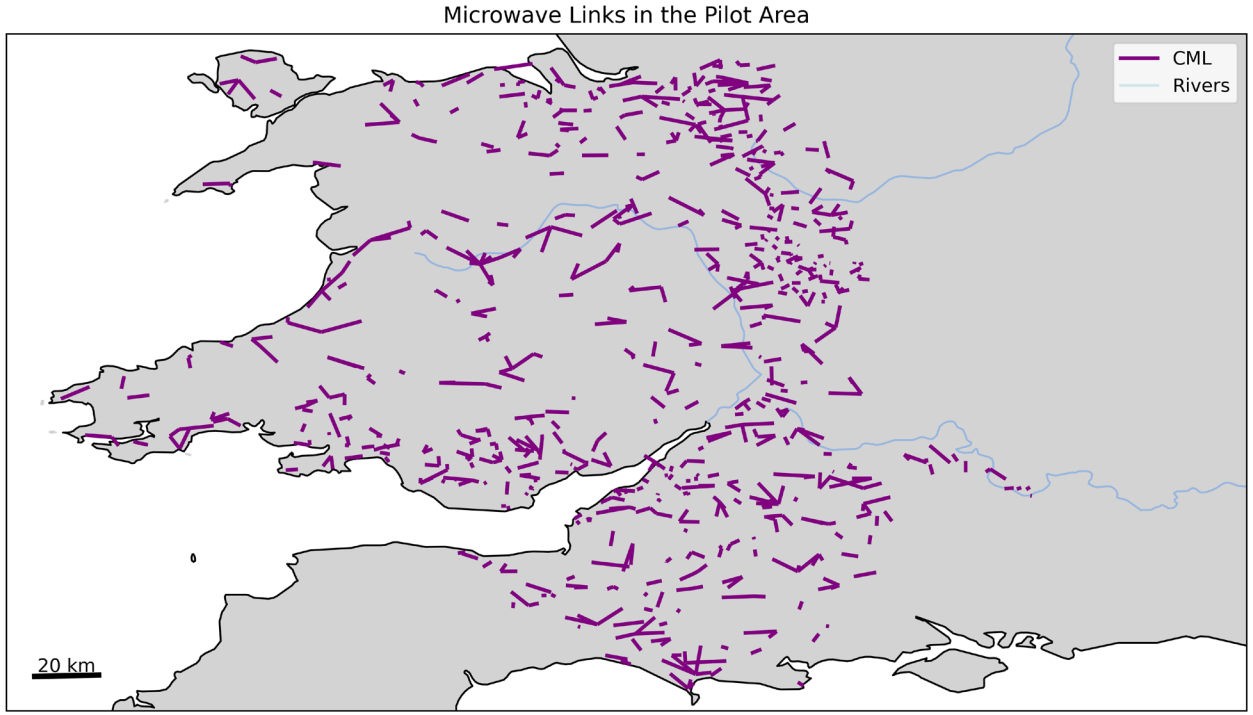
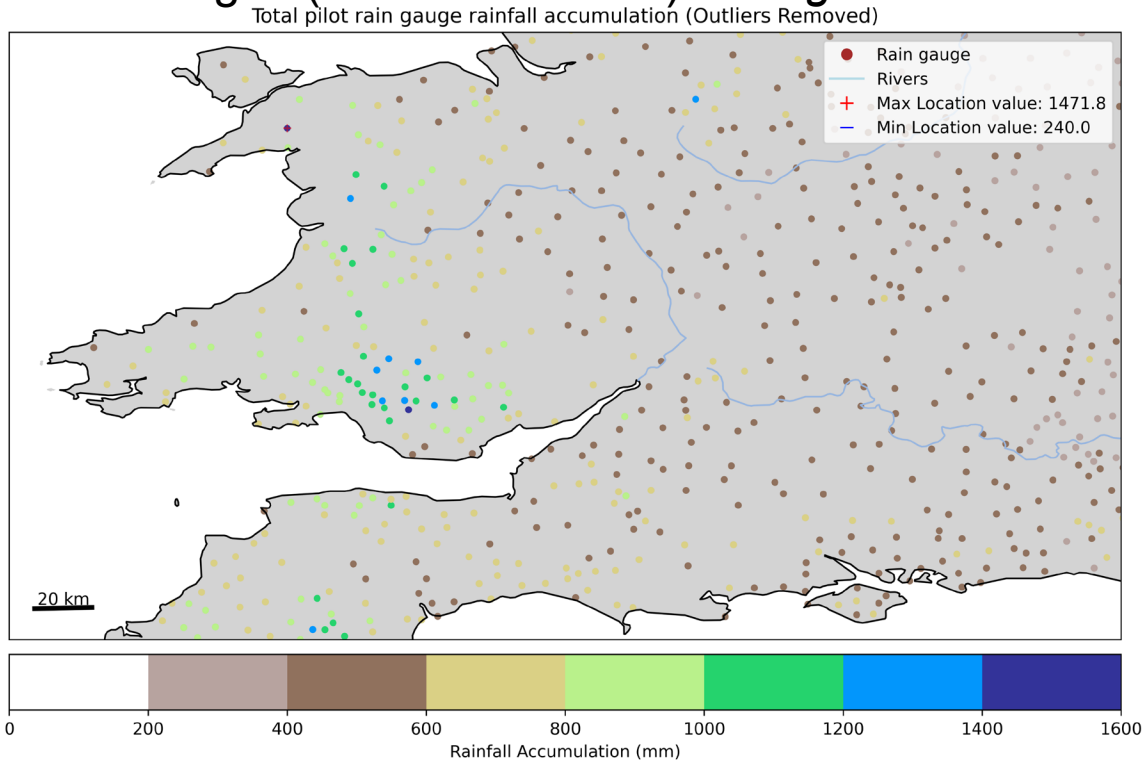


Figure: Locations of the CML paths provided in the pilot study

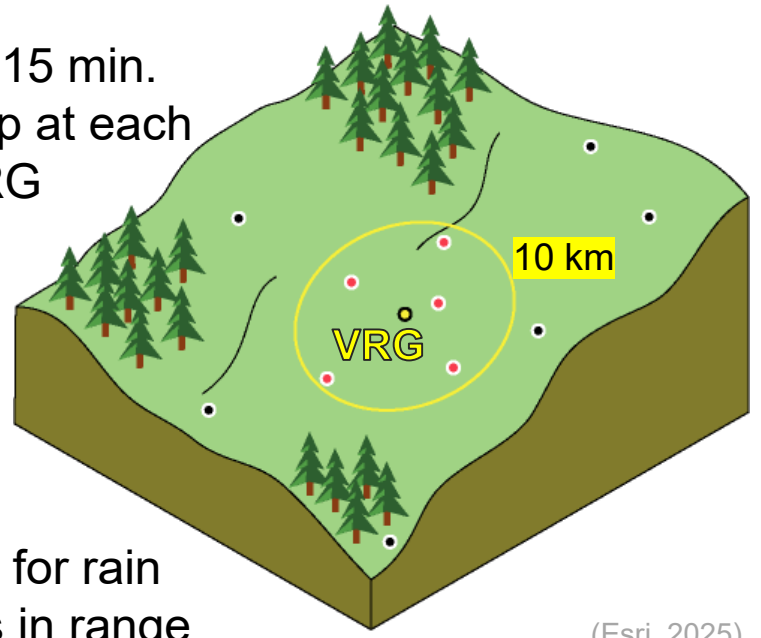
Rain Gauge Data processing for Validation

Rain Gauges (EA & NPW APIs) *"The ground truth"*



Inverse Distance Weighting (IDW) interpolation

At each 15 min. time-step at each CML VRG



Search for rain gauges in range

(Esri, 2025)

Data needs cleaning

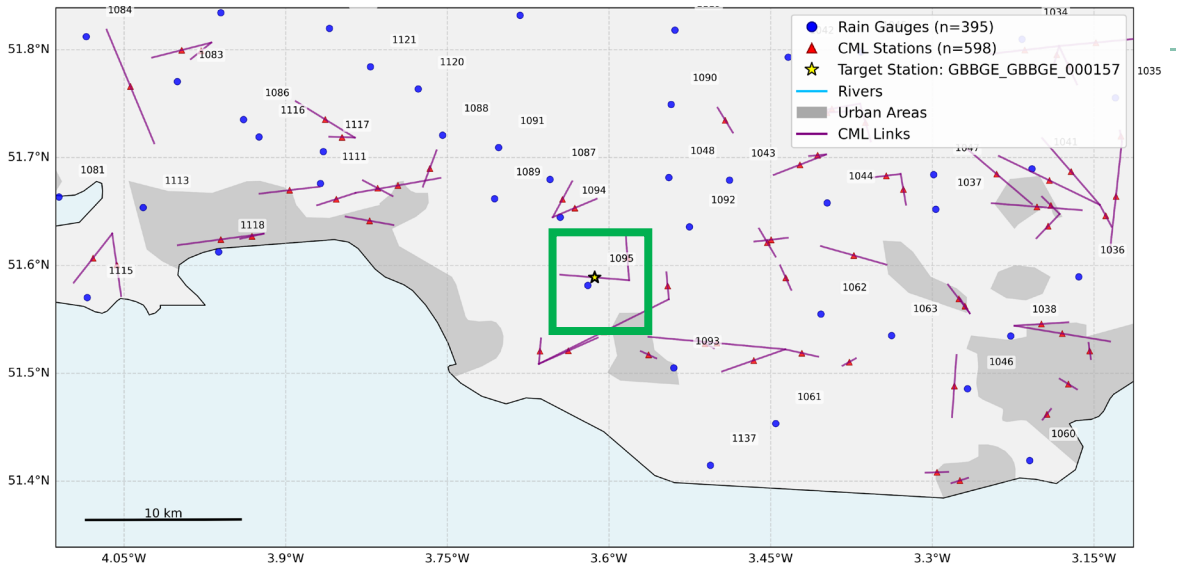
- Based on known physical constraints (mm/15min)
- Outlier (percentile) approach on total accumulation

Handling time step gaps

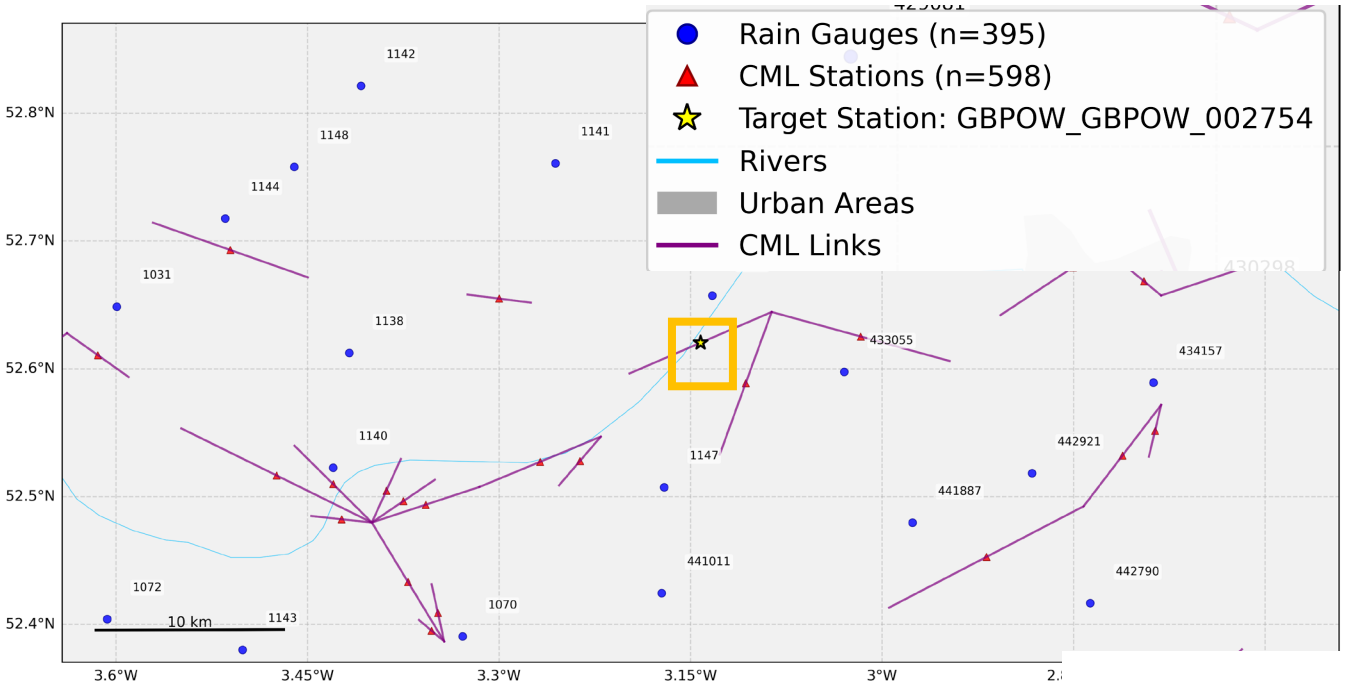
- Special care with NaN values, excluded from evaluation
- Only gauges with completeness > 90% are used

Selected Case Study CML VRGs

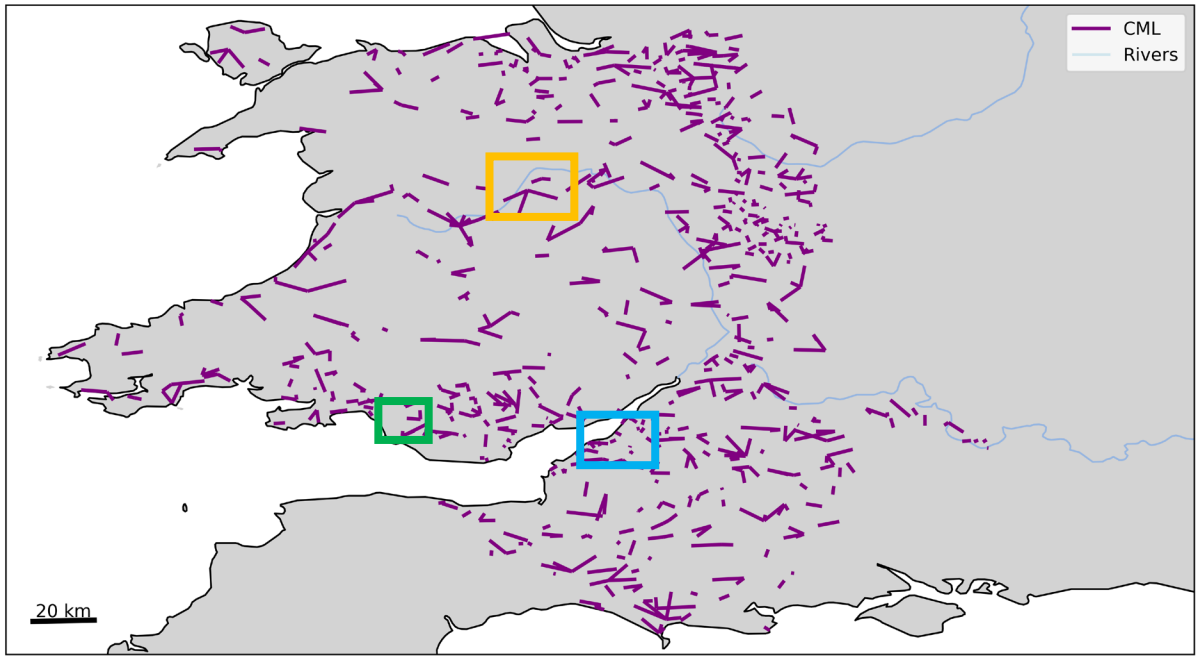
Southwest Wales – Good Correlation with interpolated RG



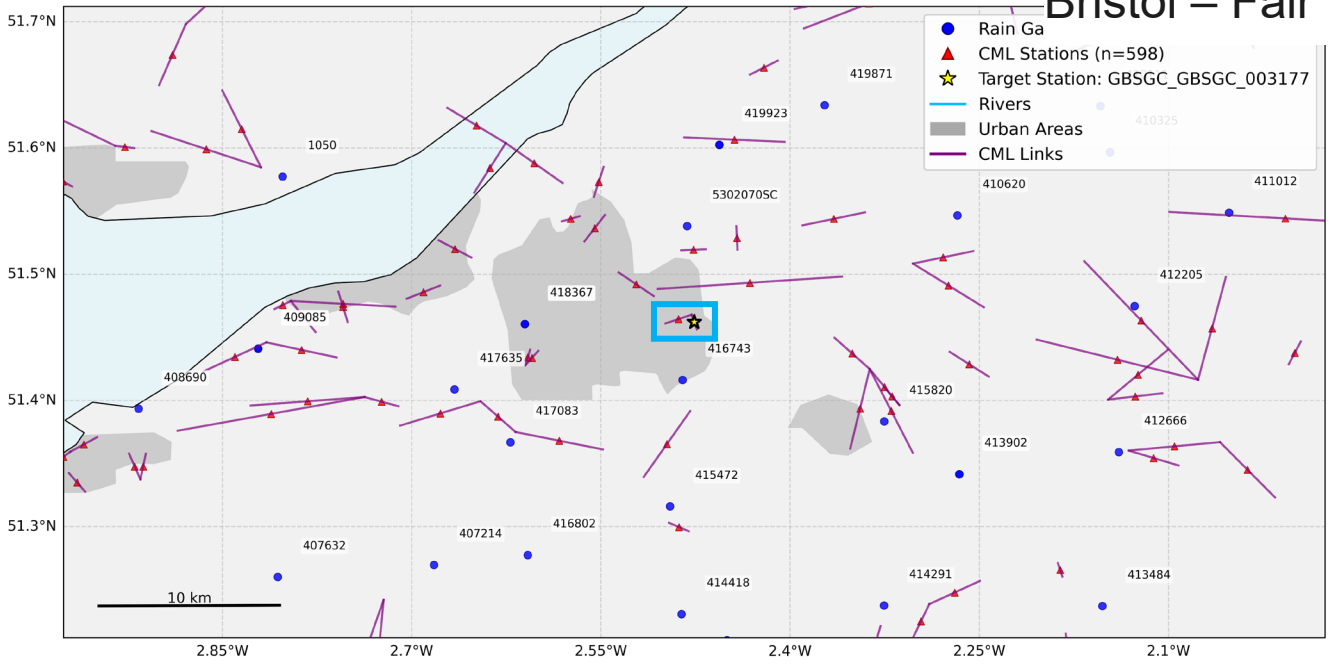
Upper River Severn - Poor



Microwave Links in the Pilot Area



Bristol – Fair



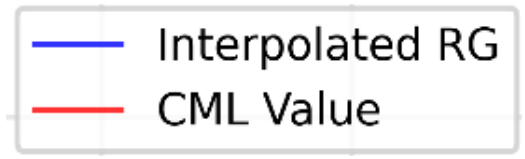
Southwest Wales – Good Correlation (Hourly accumulated rainfall time series)

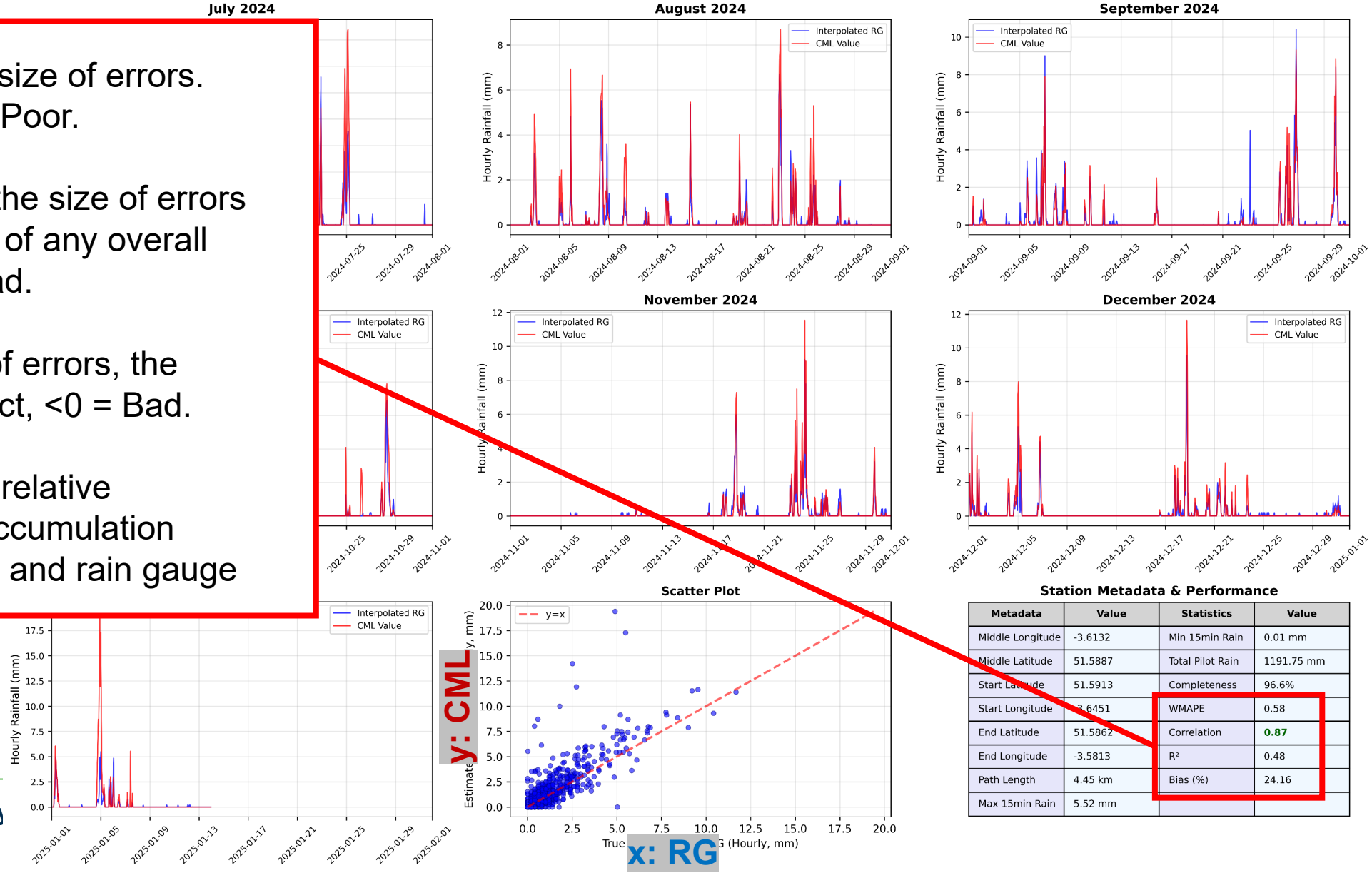
WMAPE: assesses the size of errors. 0% = Perfect, >100% = Poor.

Correlation: assesses the size of errors in the data independent of any overall bias. 1 = Perfect, 0 = Bad.

R^2 : assesses the size of errors, the tightness of fit. 1= Perfect, <0 = Bad.

Bias (%): assesses the relative difference in total rain accumulation between CML (or radar) and rain gauge

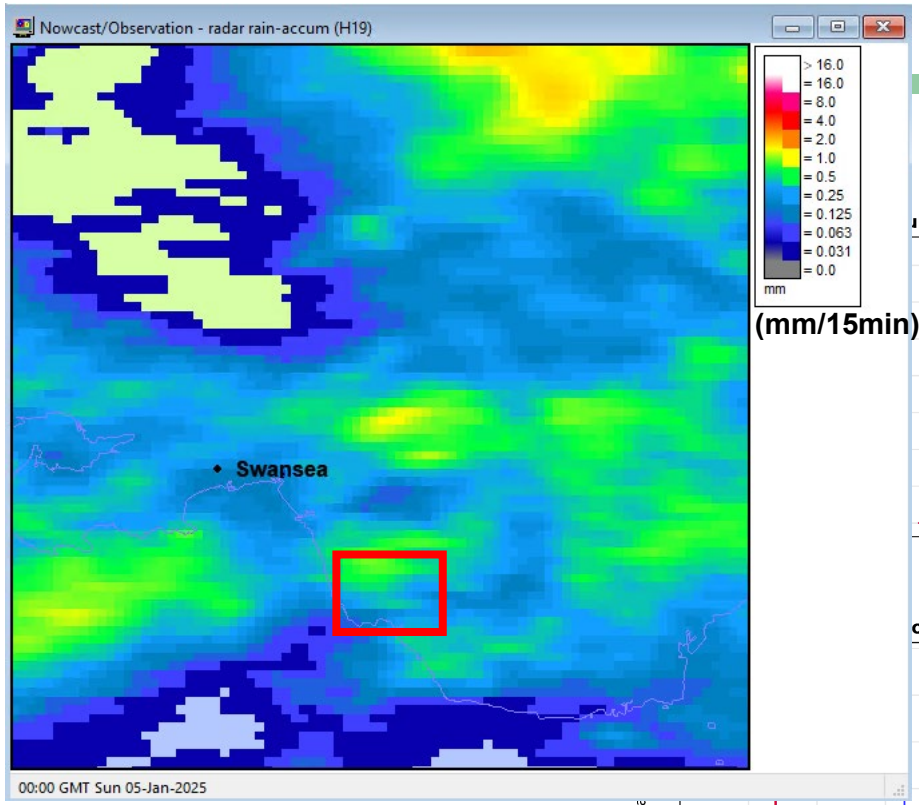




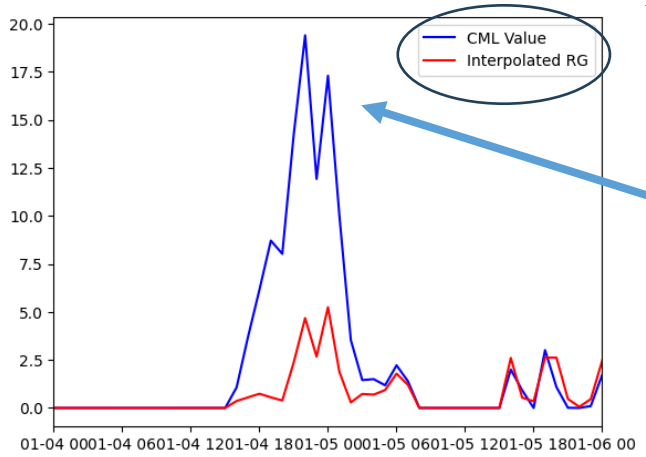
Station Metadata & Performance

Metadata	Value	Statistics	Value
Middle Longitude	-3.6132	Min 15min Rain	0.01 mm
Middle Latitude	51.5887	Total Pilot Rain	1191.75 mm
Start Latitude	51.5913	Completeness	96.6%
Start Longitude	-3.6451	WMAPE	0.58
End Latitude	51.5862	Correlation	0.87
End Longitude	-3.5813	R^2	0.48
Path Length	4.45 km	Bias (%)	24.16
Max 15min Rain	5.52 mm		

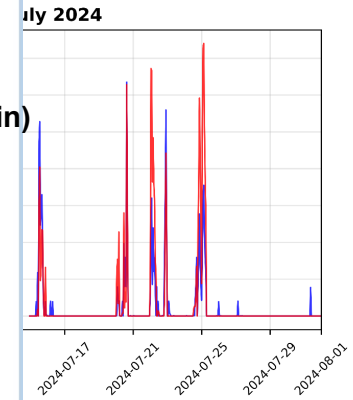
Southwest Wales – Good

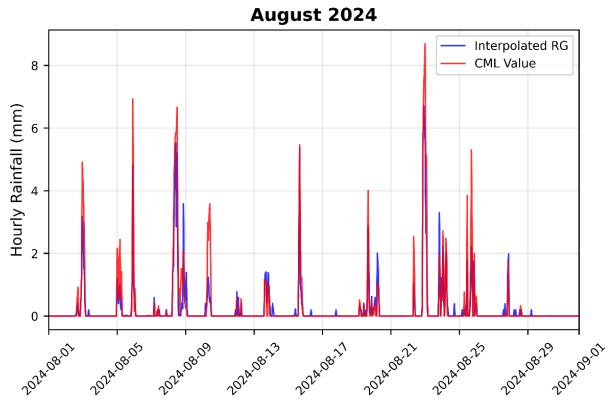


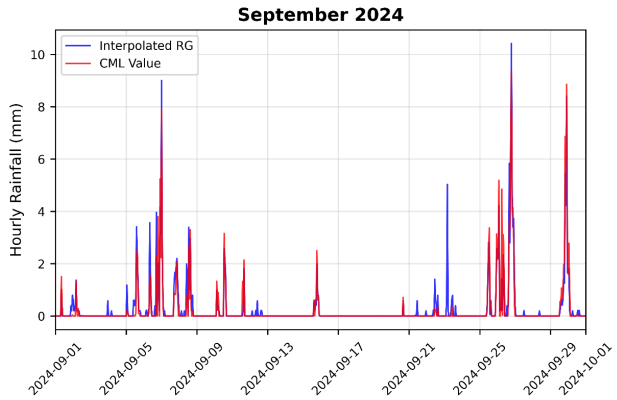
(mm/15min)

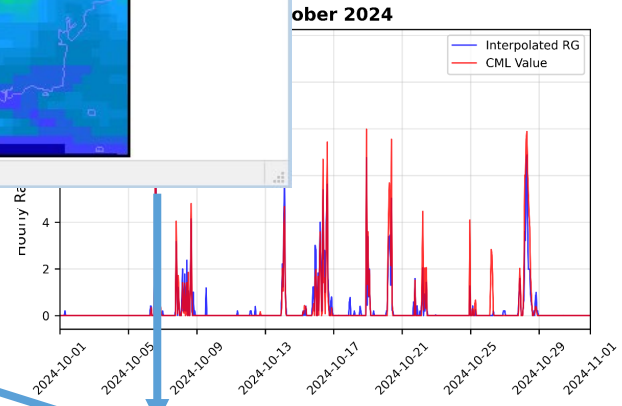


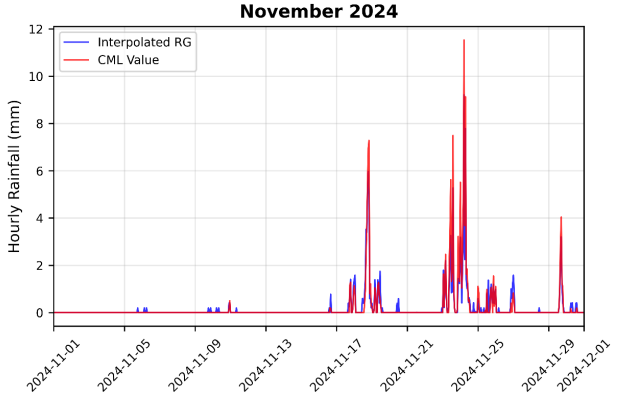
Hourly Rainfall Time Series - Station GBBGE_GBBGE_000157

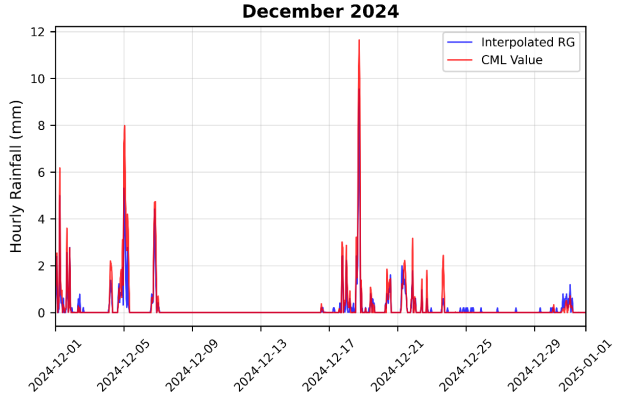


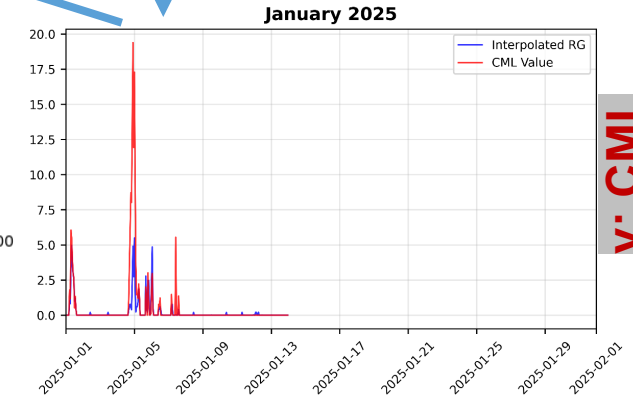


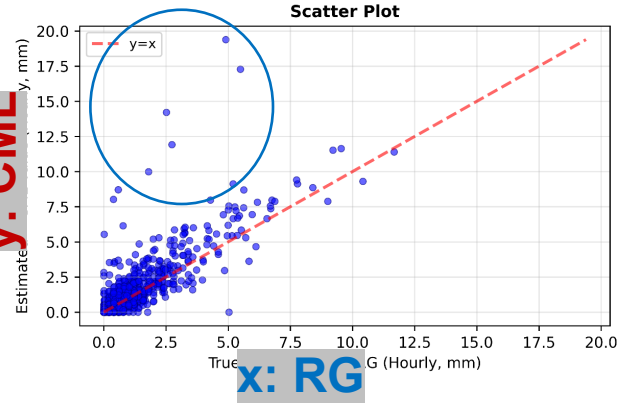










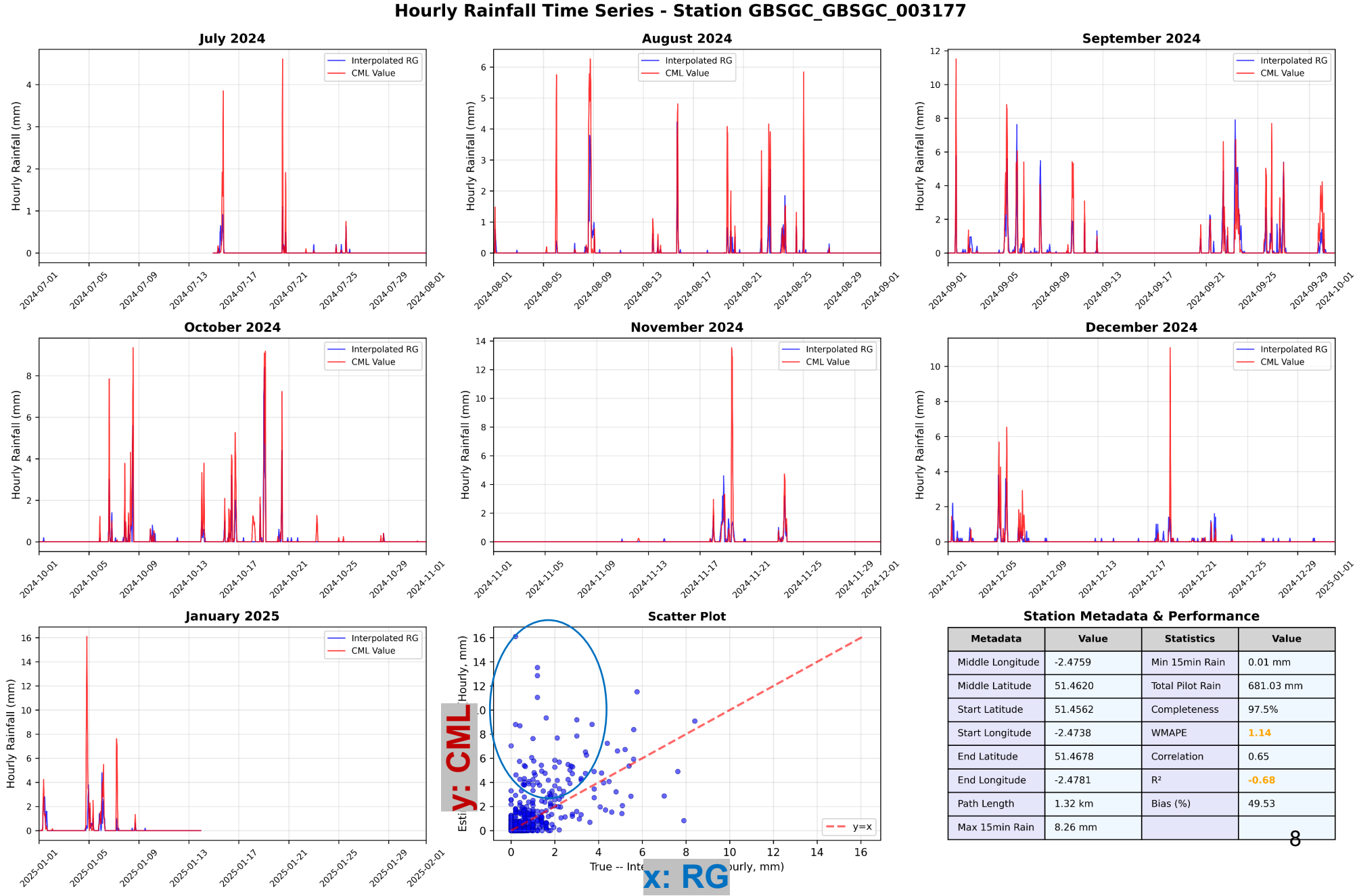


Station Metadata & Performance

Metadata	Value	Statistics	Value
Middle Longitude	-3.6132	Min 15min Rain	0.01 mm
Middle Latitude	51.5887	Total Pilot Rain	1191.75 mm
Start Latitude	51.5913	Completeness	96.6%
Start Longitude	-3.6451	WMAPE	0.58
End Latitude	51.5862	Correlation	**0.87**
End Longitude	-3.5813	R^2	0.48
Path Length	4.45 km	Bias (%)	24.16
Max 15min Rain	5.52 mm		

Bristol – Fair

- Correlation still high but high positive bias



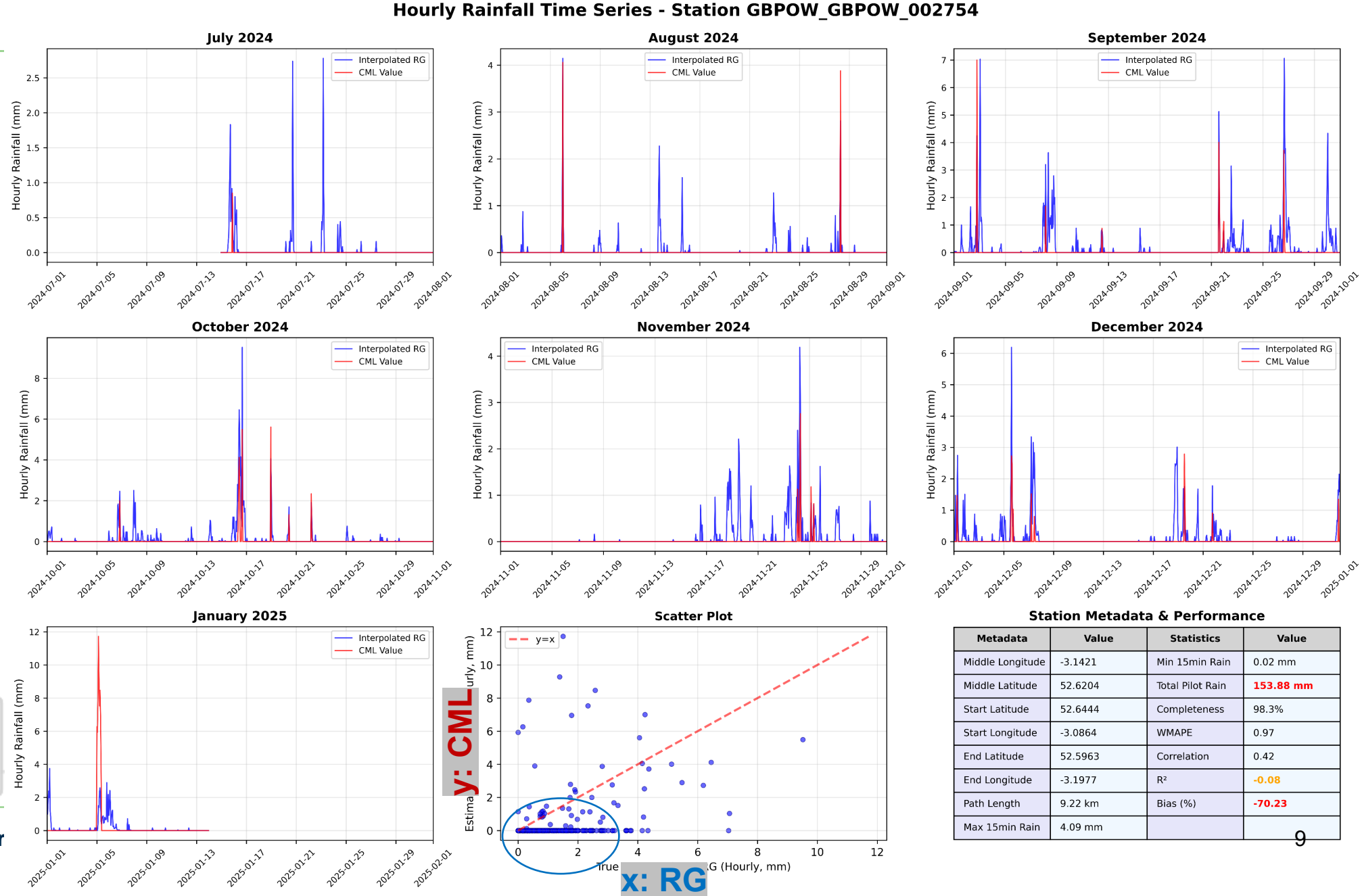
Station Metadata & Performance

Metadata	Value	Statistics	Value
Middle Longitude	-2.4759	Min 15min Rain	0.01 mm
Middle Latitude	51.4620	Total Pilot Rain	681.03 mm
Start Latitude	51.4562	Completeness	97.5%
Start Longitude	-2.4738	WMAPE	1.14
End Latitude	51.4678	Correlation	0.65
End Longitude	-2.4781	R^2	-0.68
Path Length	1.32 km	Bias (%)	49.53
Max 15min Rain	8.26 mm		

Interpolated RG
CML Value

Upper River Severn – Poor

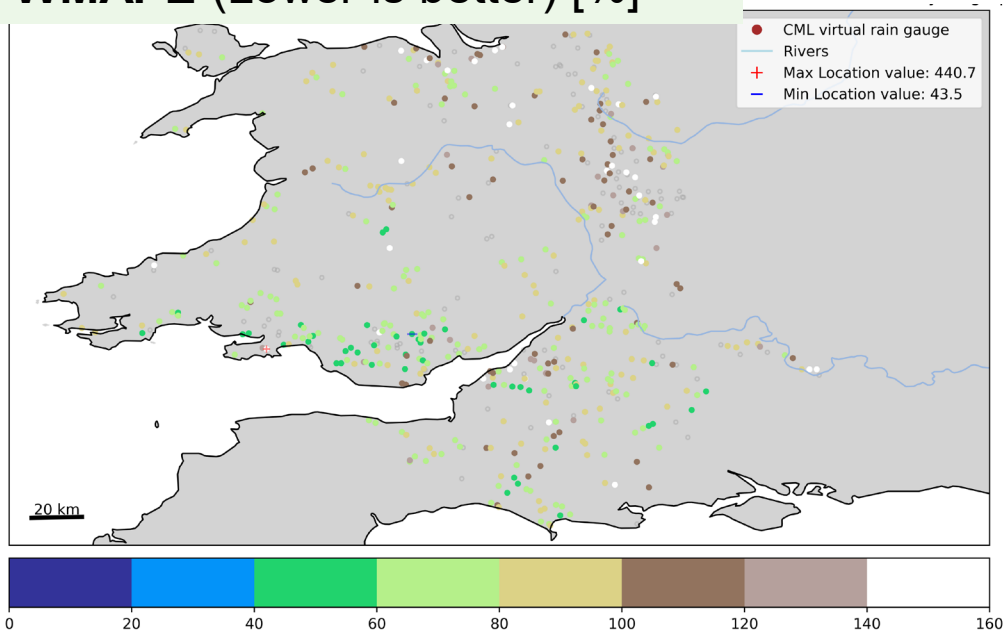
- Large negative bias



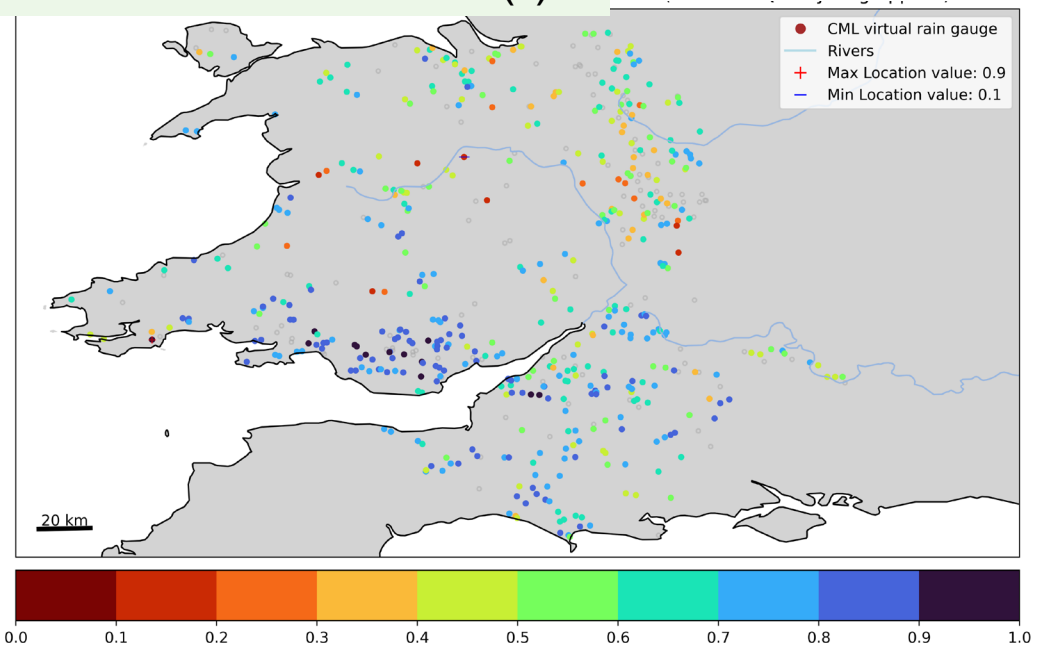
Metadata	Value	Statistics	Value
Middle Longitude	-3.1421	Min 15min Rain	0.02 mm
Middle Latitude	52.6204	Total Pilot Rain	**153.88 mm**
Start Latitude	52.6444	Completeness	98.3%
Start Longitude	-3.0864	WMAPE	0.97
End Latitude	52.5963	Correlation	0.42
End Longitude	-3.1977	R^2	**-0.08**
Path Length	9.22 km	Bias (%)	**-70.23**
Max 15min Rain	4.09 mm		

Evaluation of CML VRG performance with RG for hourly accumulation

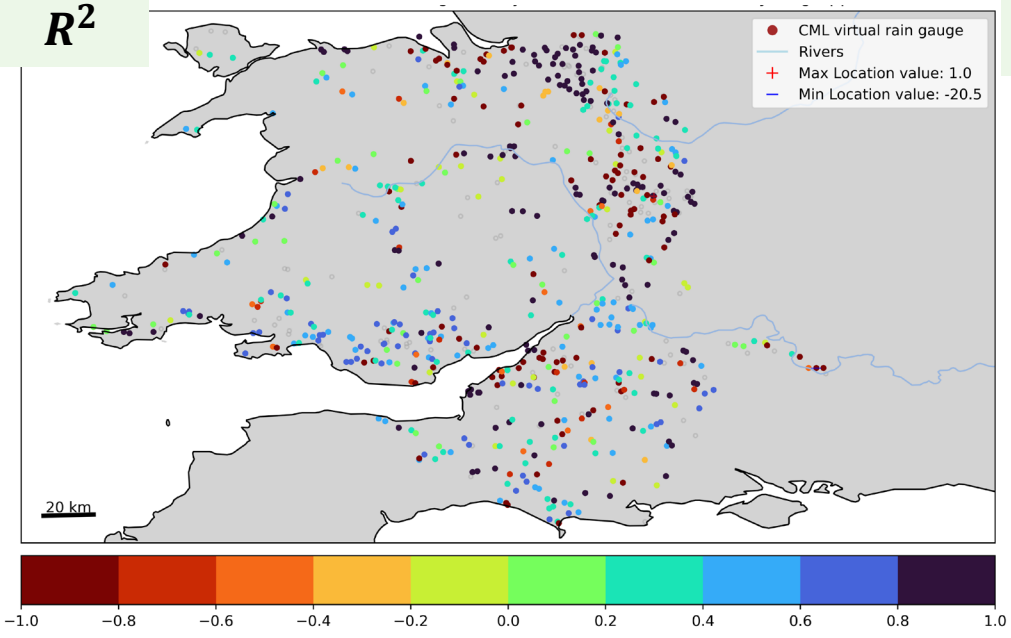
WMAPE (Lower is better) [%]



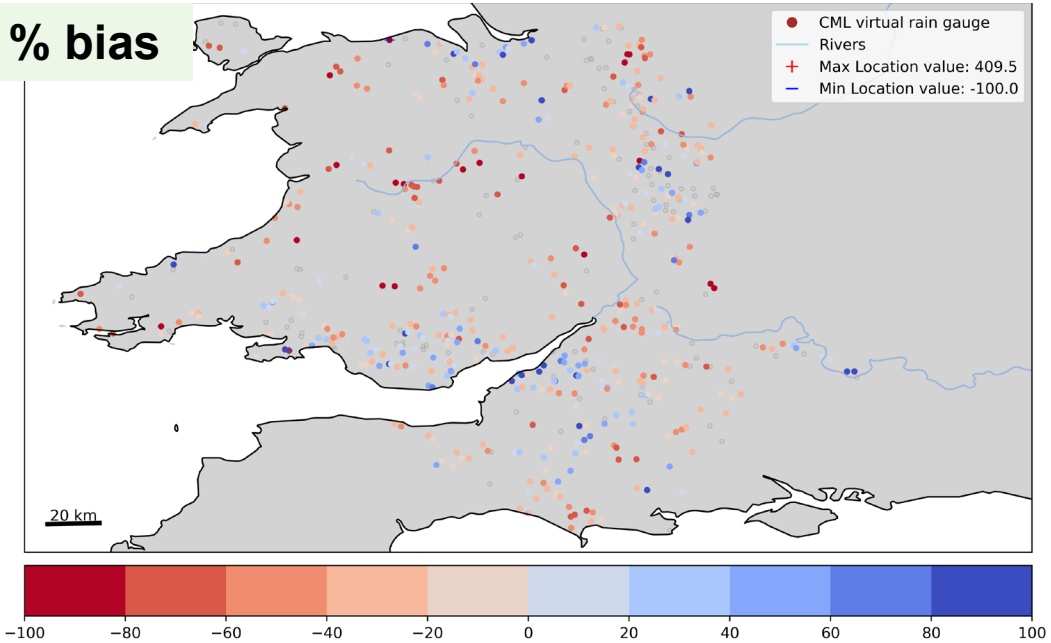
Correlation Coefficient (**r**)



R^2

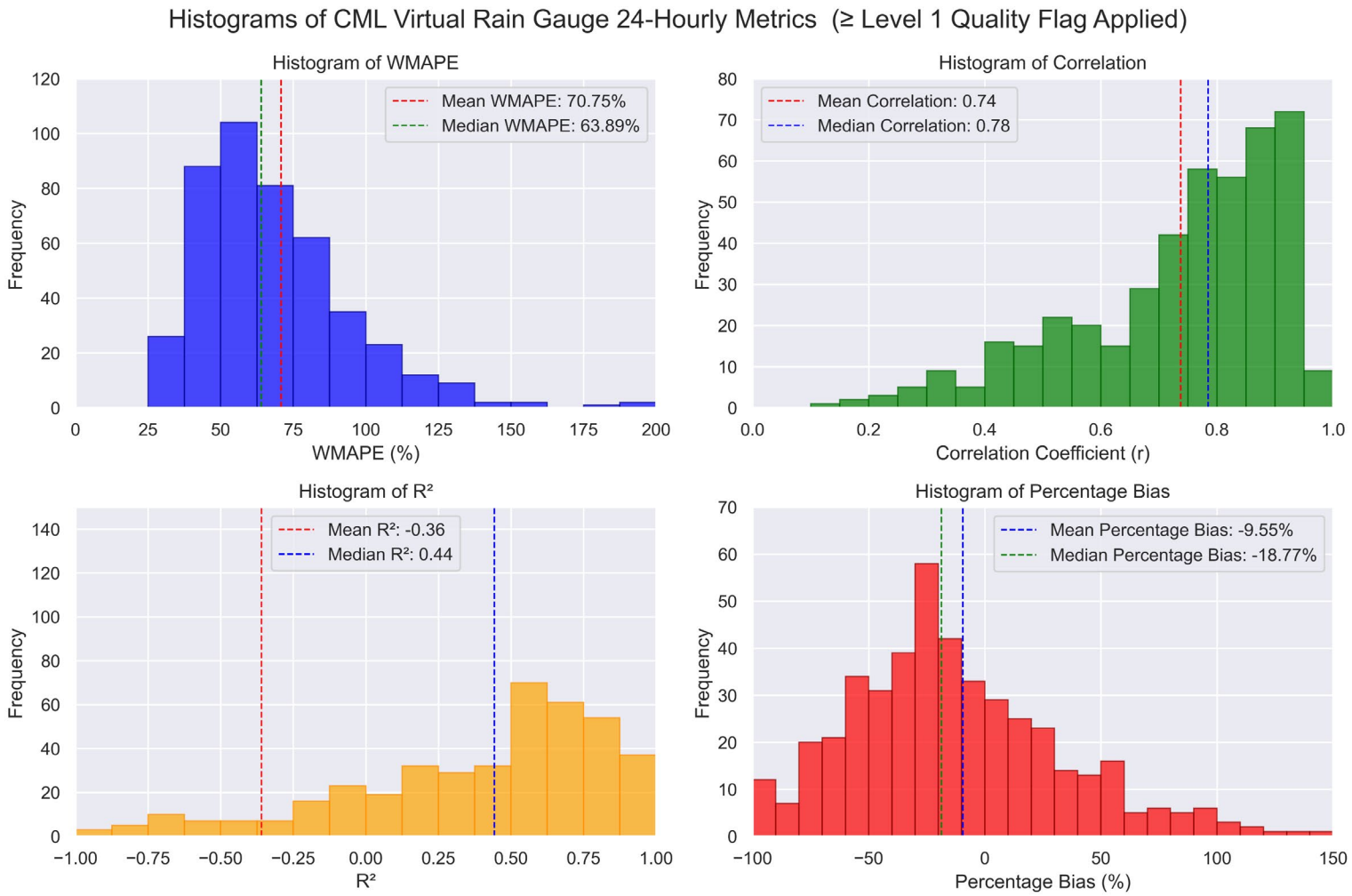


% bias



- **High spread of performance** for CML VRGs, but some particularly in the south are performing better

Evaluation of CML VRG performance with RG accumulation timesteps: 15min, 1hr, 24hrs



15 min

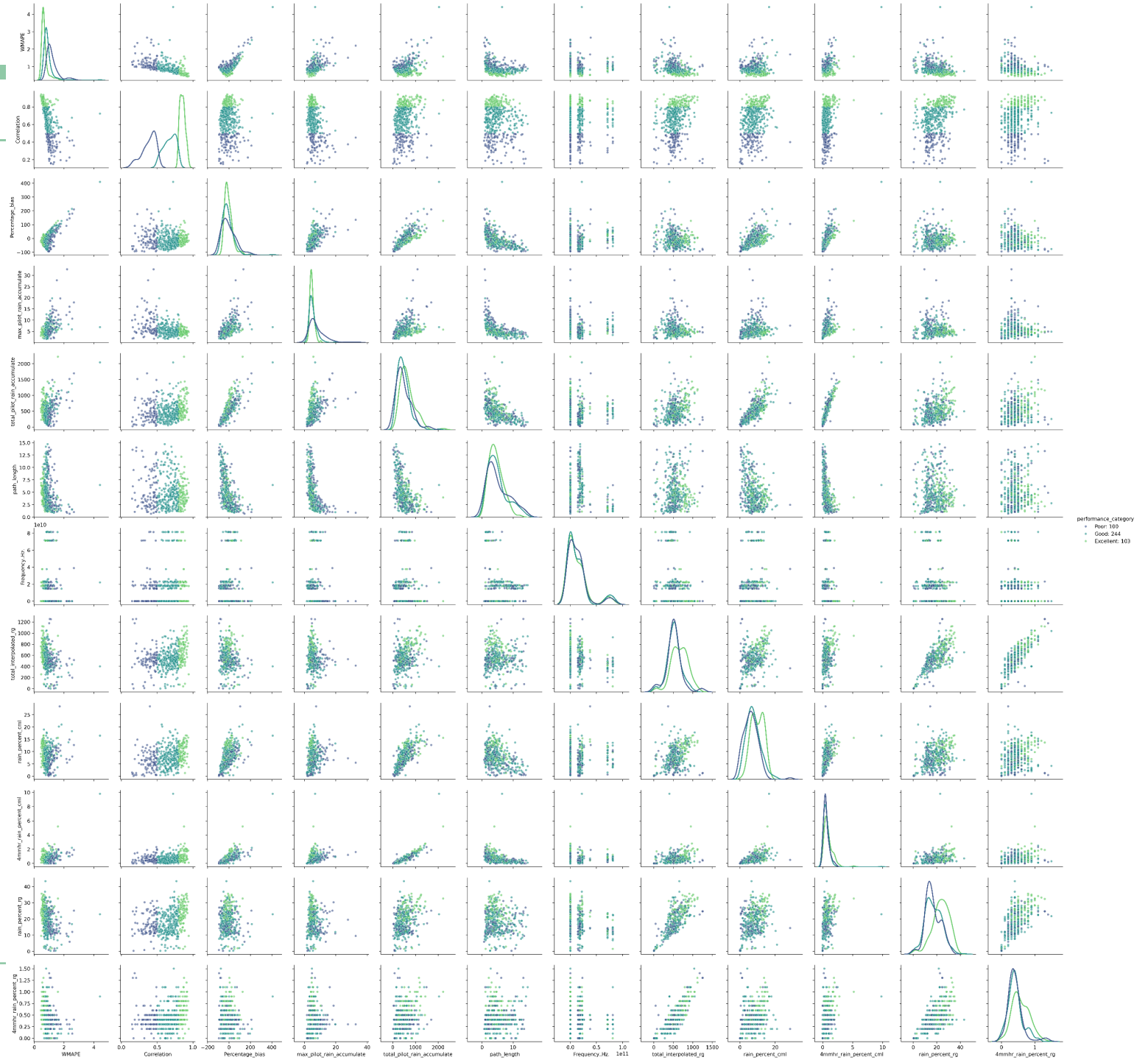
- Metrics **improve as the accumulation time increases** except % bias
- **Spread of correlation** coefficient **decreases** as the accumulation time increases

24 hr

Correlation of variables

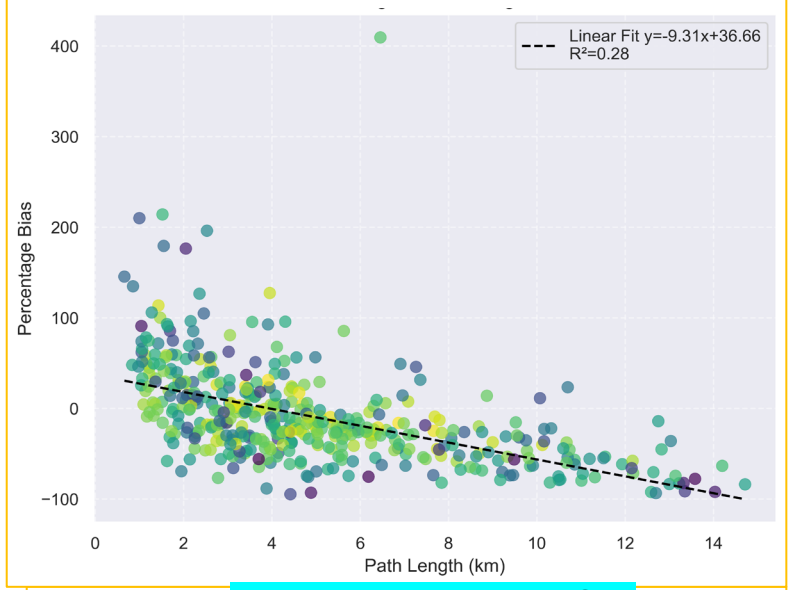
Some tested variables:

- Path lengths
- total pilot period accumulated rainfall
- max pilot period accumulated rainfall
- max 15min accumulated rainfall
- frequency of links
- percentage of time steps with rain (CML/ RG)

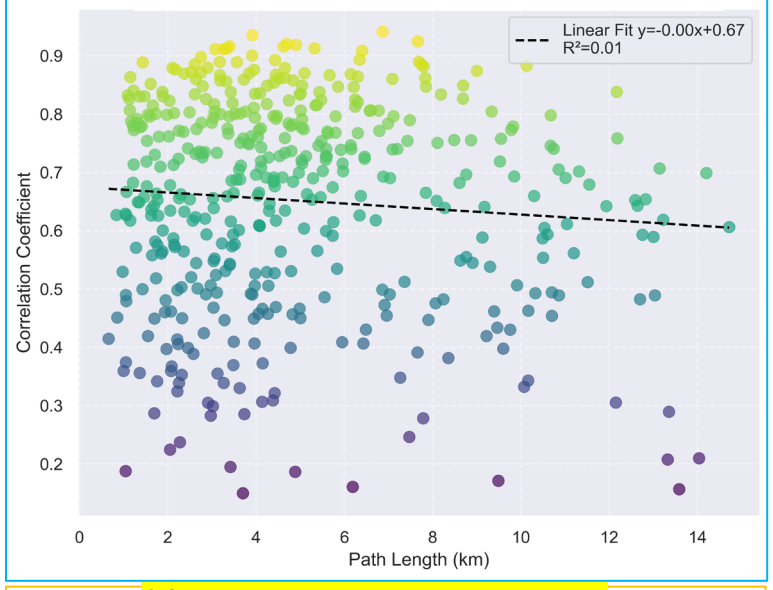


Selected Variables vs Path Length

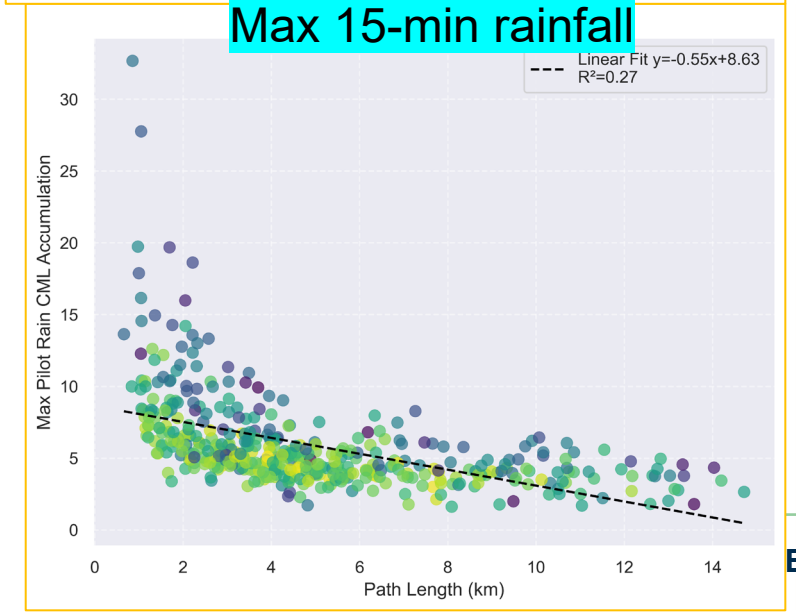
Percentage bias of CMLs



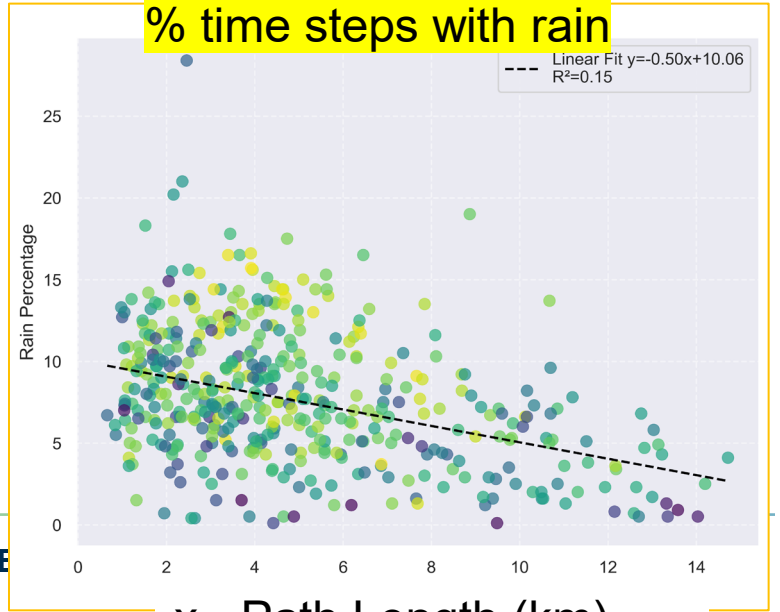
Correlation coef. of CMLs



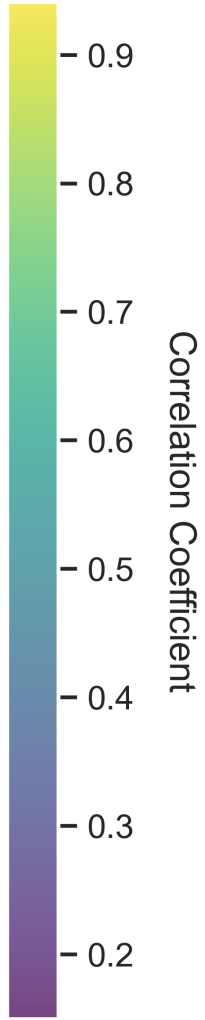
Max 15-min rainfall



% time steps with rain



x - Path Length (km)



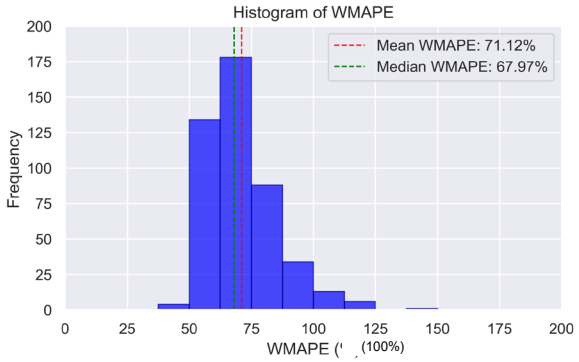
% bias negatively correlates with **path length** of CMLs

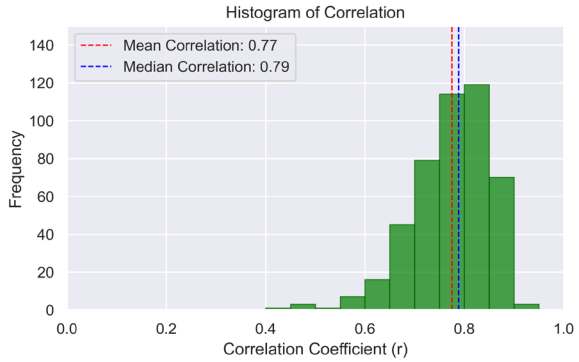
No correlation between correlation coefficient and CML path lengths

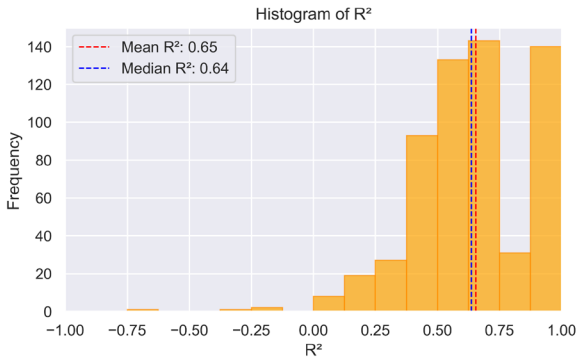
Shorter links appear to be detecting **larger quantities of rain** at each timestep and detecting rain for a **greater percentage of timesteps**.

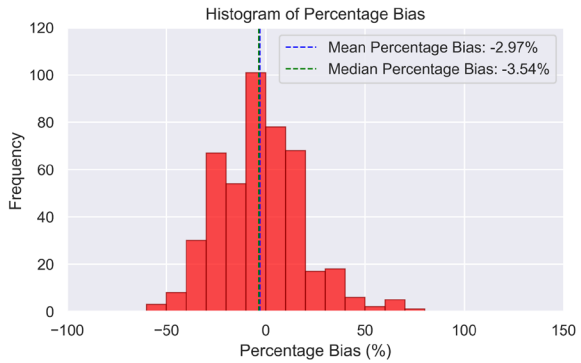
Evaluation of hourly Radar performance compared to RG at the CML locations

Option 1) Compare at mid point only

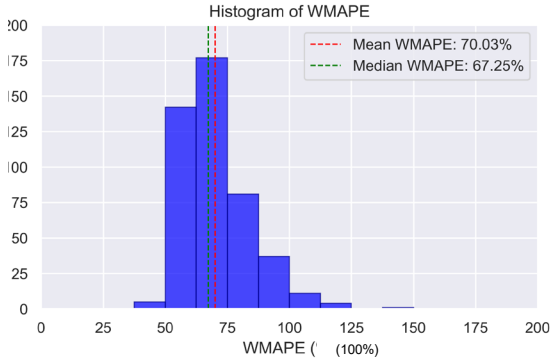


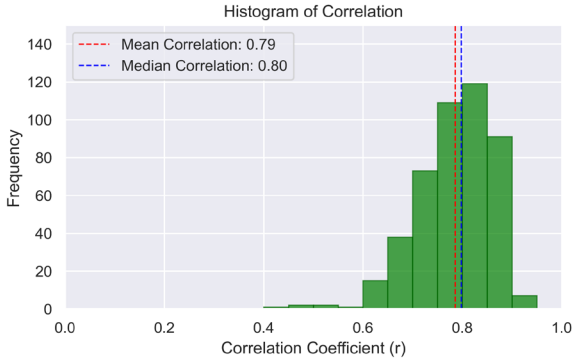


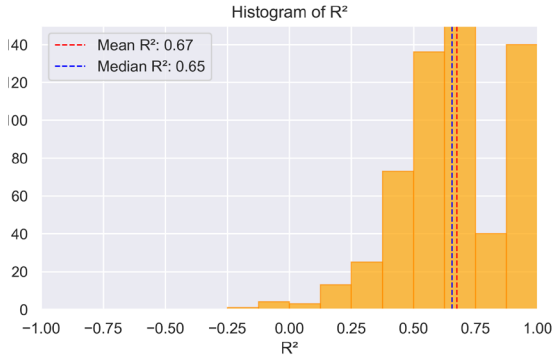


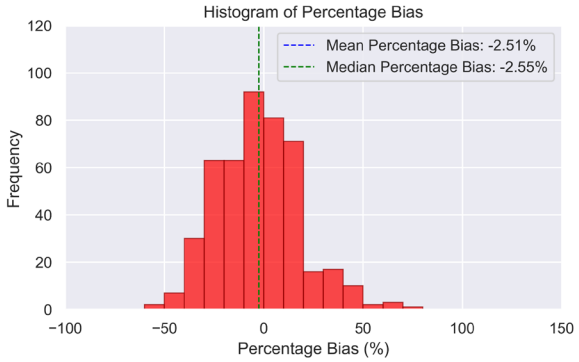


Option 2) Radar averaged over CML Path









- Only a **small part of the inaccuracy** in the CML precipitation is due to comparing a **path integrated precipitation** (CML) to a **point precipitation** (rain gauge).
- The **radar** precipitation estimate is **more accurate** for most, **but not all, locations studied**. The **spread** is still **large**.

Conclusions

The **accuracy** of CML precipitation estimates from the pilot study are **strongly location dependent**. For **most** but not all locations, their accuracy is **lower than radar**. There are persistent (systematic) **biases** at some locations that **could be corrected**. **Further data** provided by Vodafone processed using an alternative Machine Learning method appear to show noticeably **better performance for light rain scenario** than the pilot data.

Possible next steps:

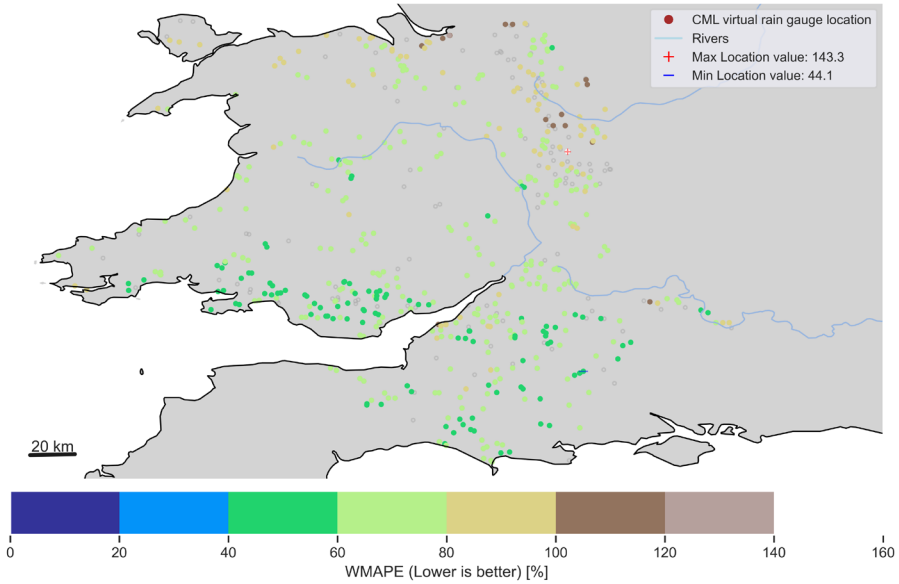
- Explore the best approach to **blend** the data
- **Triple Collocation** to establish simultaneously the inaccuracies in the CML, RG and radar data.
- Identification of **factors producing errors**: dew, wet antenna locations, snow, light rain, heavy rain…
- Full examination of the new study data supplied by Vodafone (12/8/25)

Q&A/ Further Information

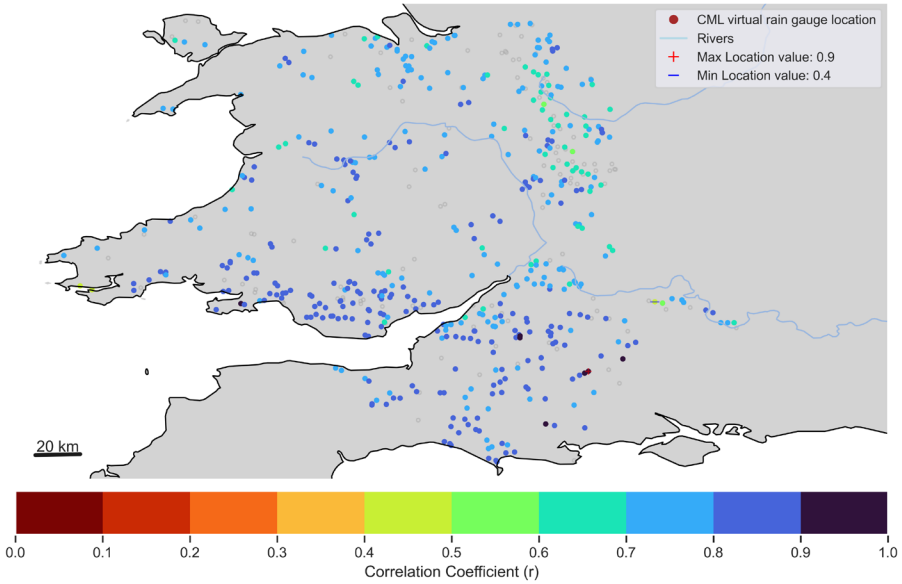
UK Centre for
Ecology & Hydrology

Evaluation of Radar performance with RG (Path Avg)

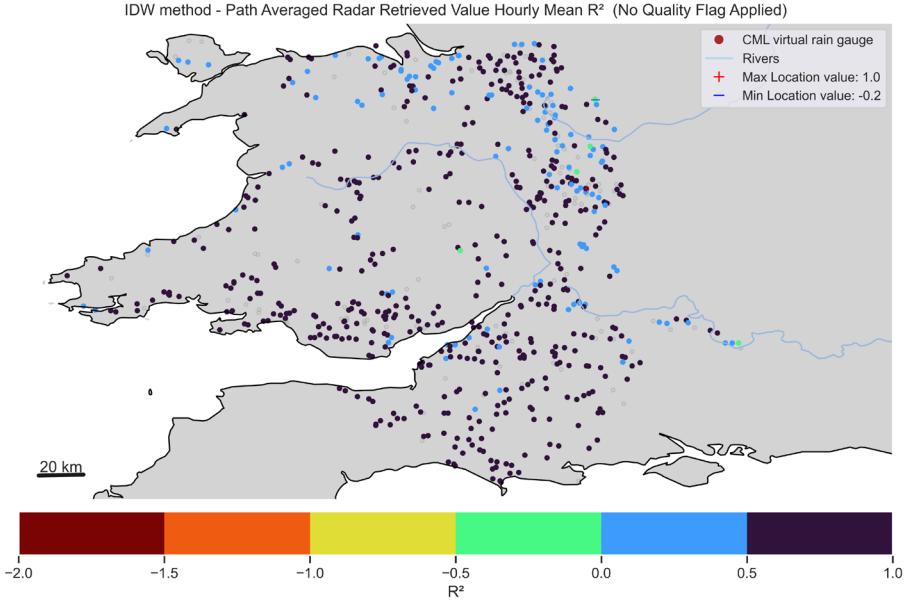
IDW method - Path Averaged Radar Retrieved Value Hourly Mean Weighted Mean Absolute Percentage Error (No Quality Flag Applied)



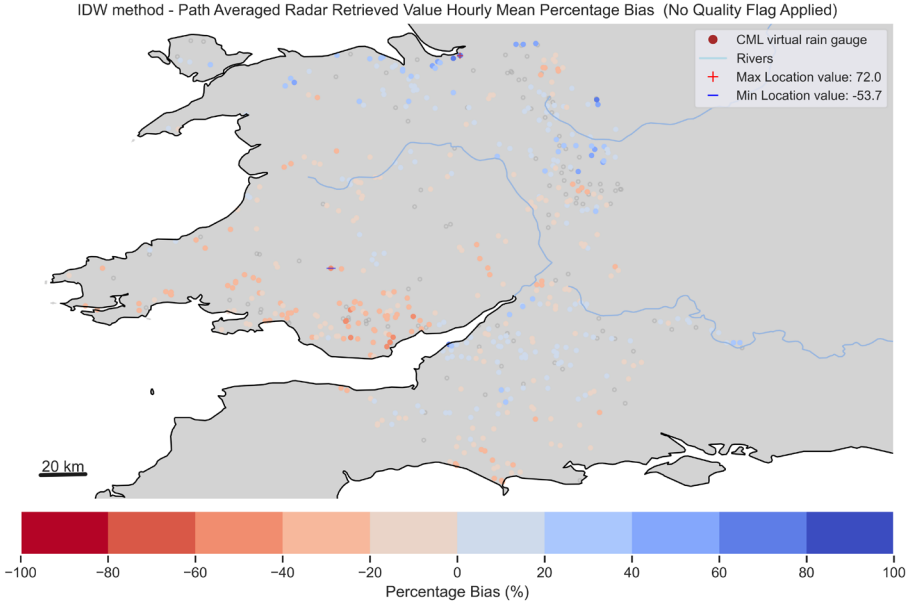
IDW method - Path Averaged Radar Retrieved Value Hourly Mean Correlation (No Quality Flag Applied)

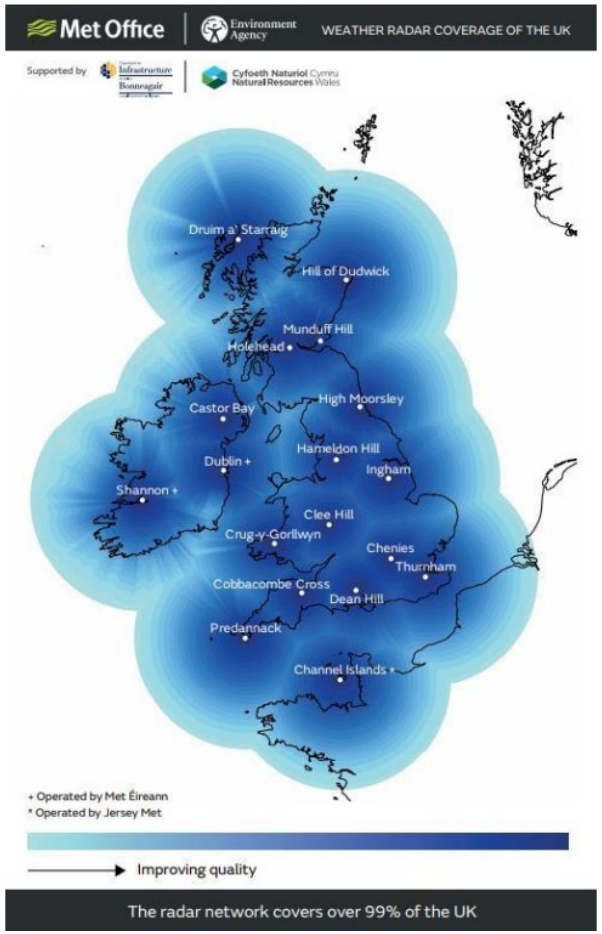


IDW method - Path Averaged Radar Retrieved Value Hourly Mean R^2 (No Quality Flag Applied)



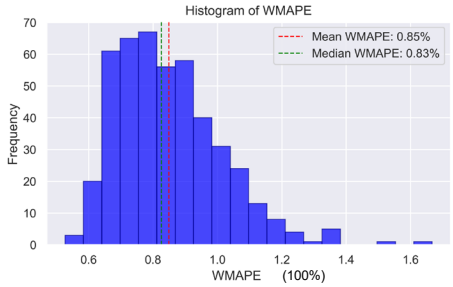
IDW method - Path Averaged Radar Retrieved Value Hourly Mean Percentage Bias (No Quality Flag Applied)

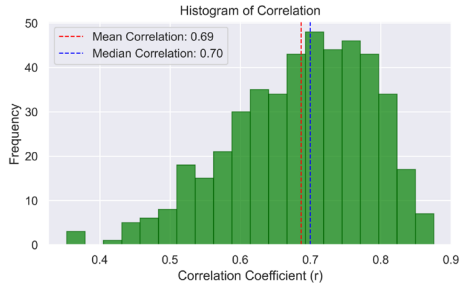


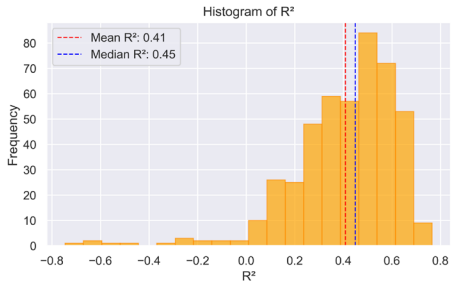


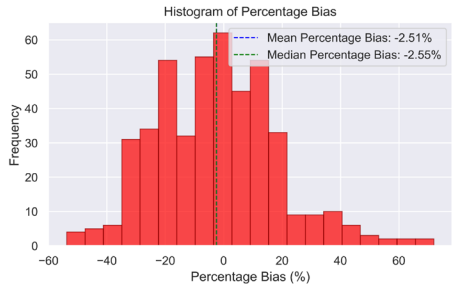
Evaluation of Radar performance with RG at various time steps:

Histograms of Path Averaged Radar Retrieved Value 15.0-Minute Metrics (No Quality Flag Applied)

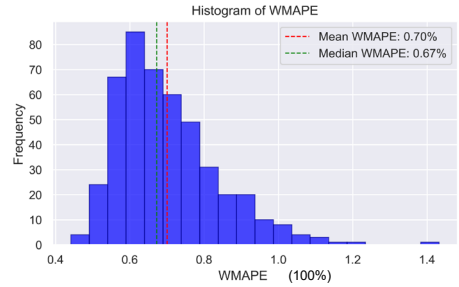


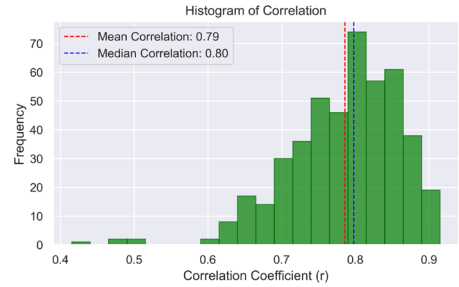


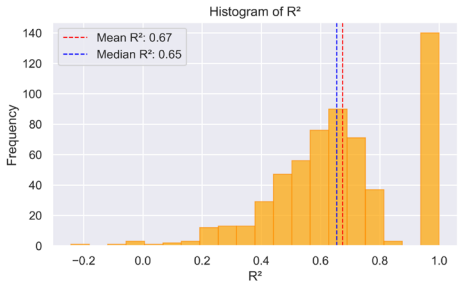


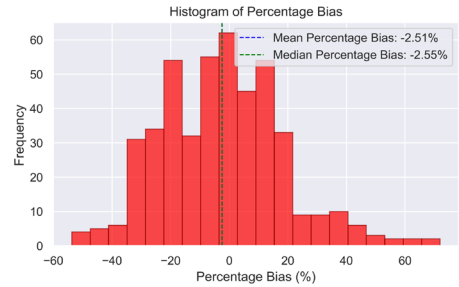


Histograms of Path Averaged Radar Retrieved Value Hourly Metrics (No Quality Flag Applied)

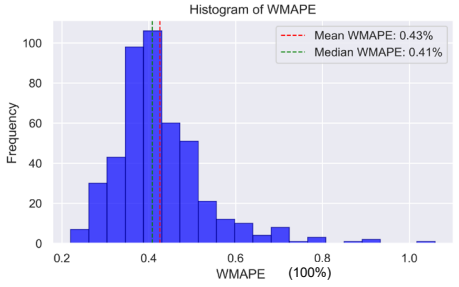


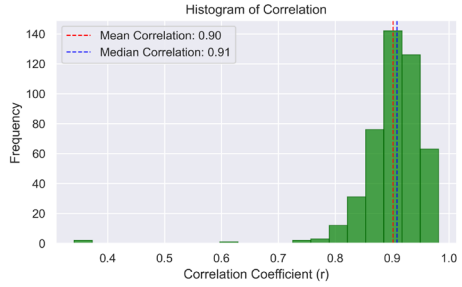


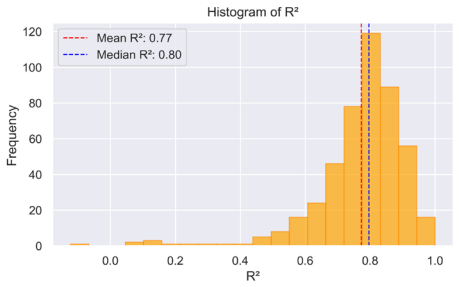


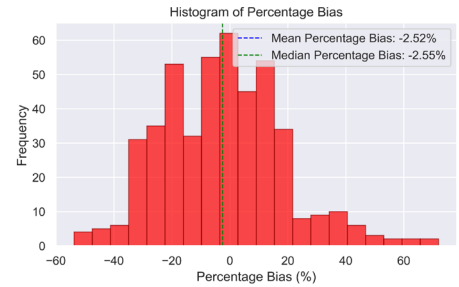


Histograms of Path Averaged Radar Retrieved Value 24-Hourly Metrics (No Quality Flag Applied)









Metrics improve as the accumulation time increases

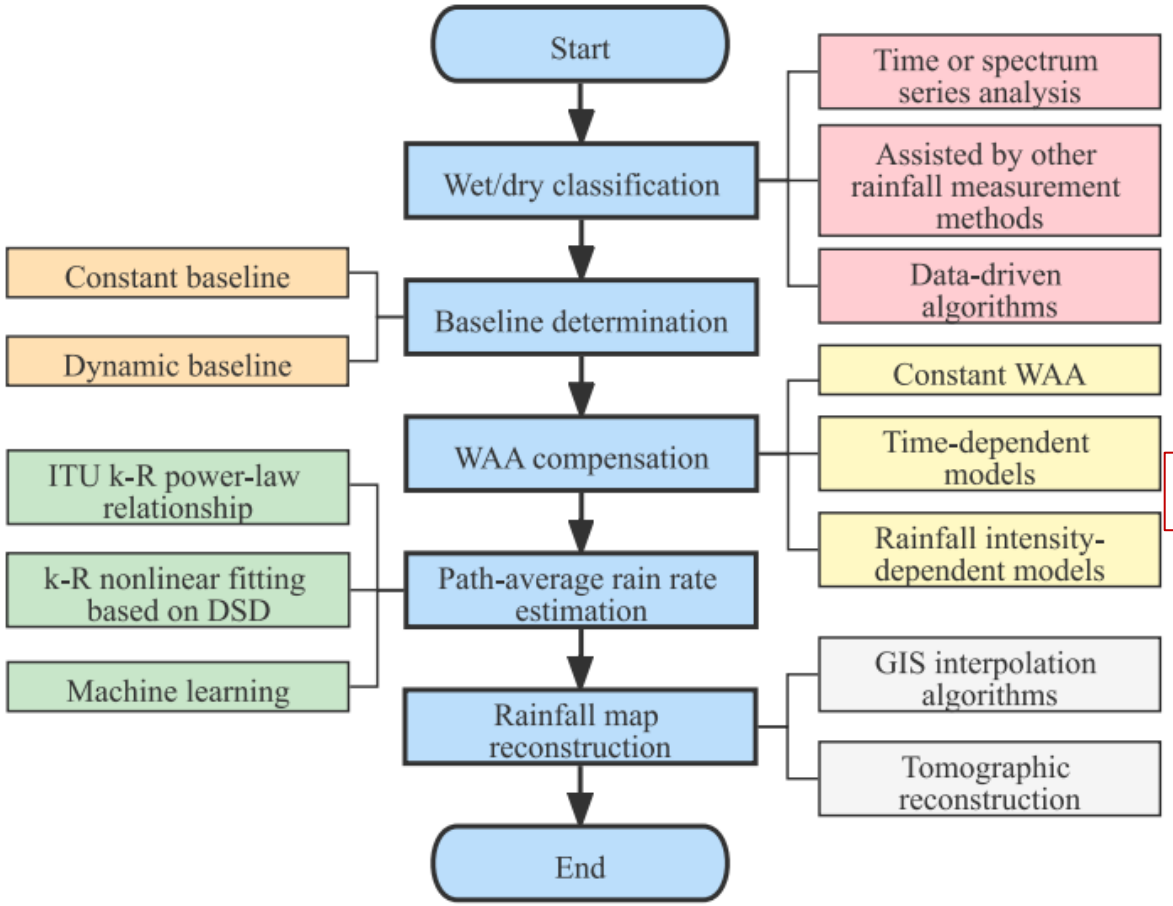
Median Correlation: 0.91
Median WMAPE: 41%
Median % Bias: -2.6%
Seen in Day accumulation

CML-based rainfall measurement Flow Chart

Approximately linear at microwave wavelengths

Note polarisation matters

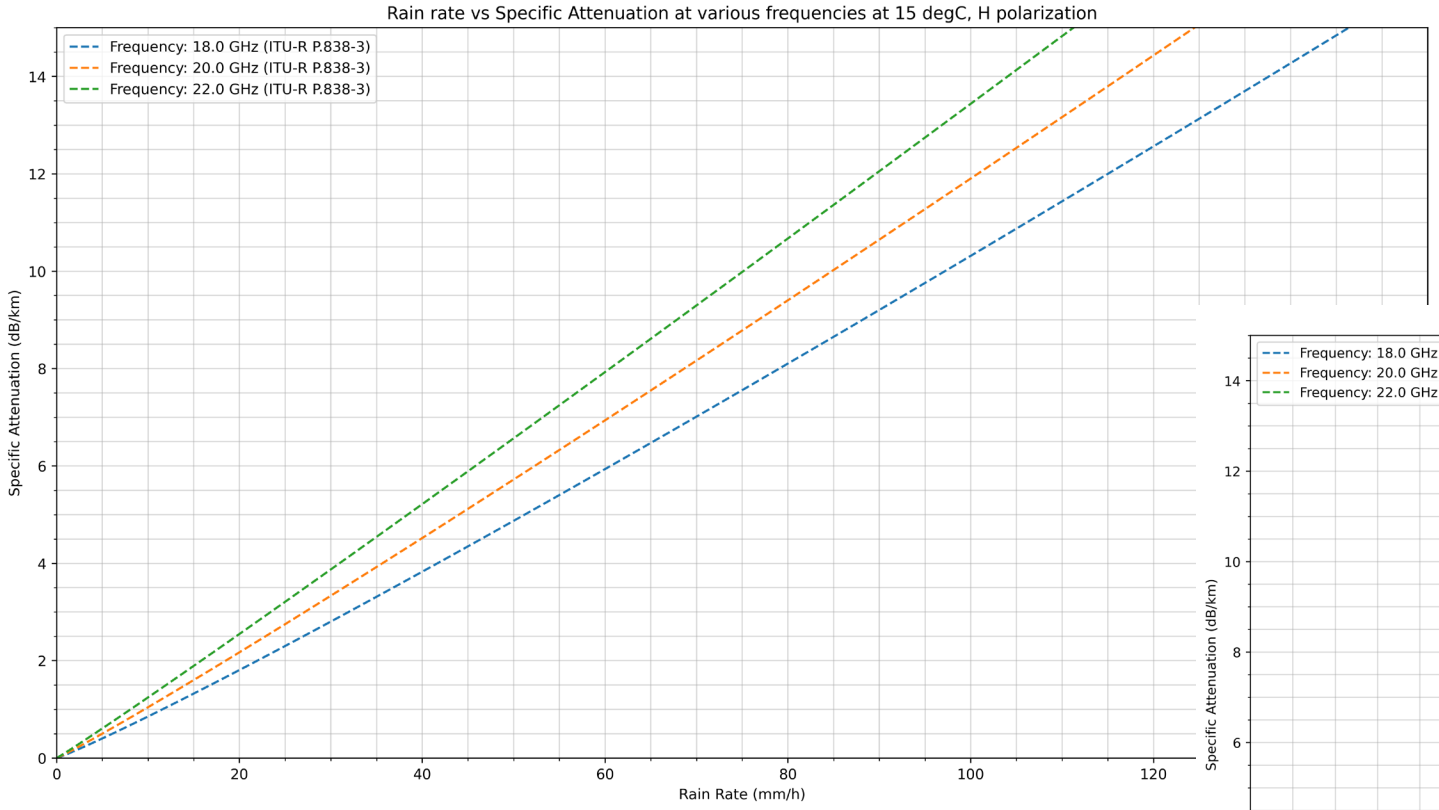
Location/ climate dependent



Wet antenna ~1-2 dB Error

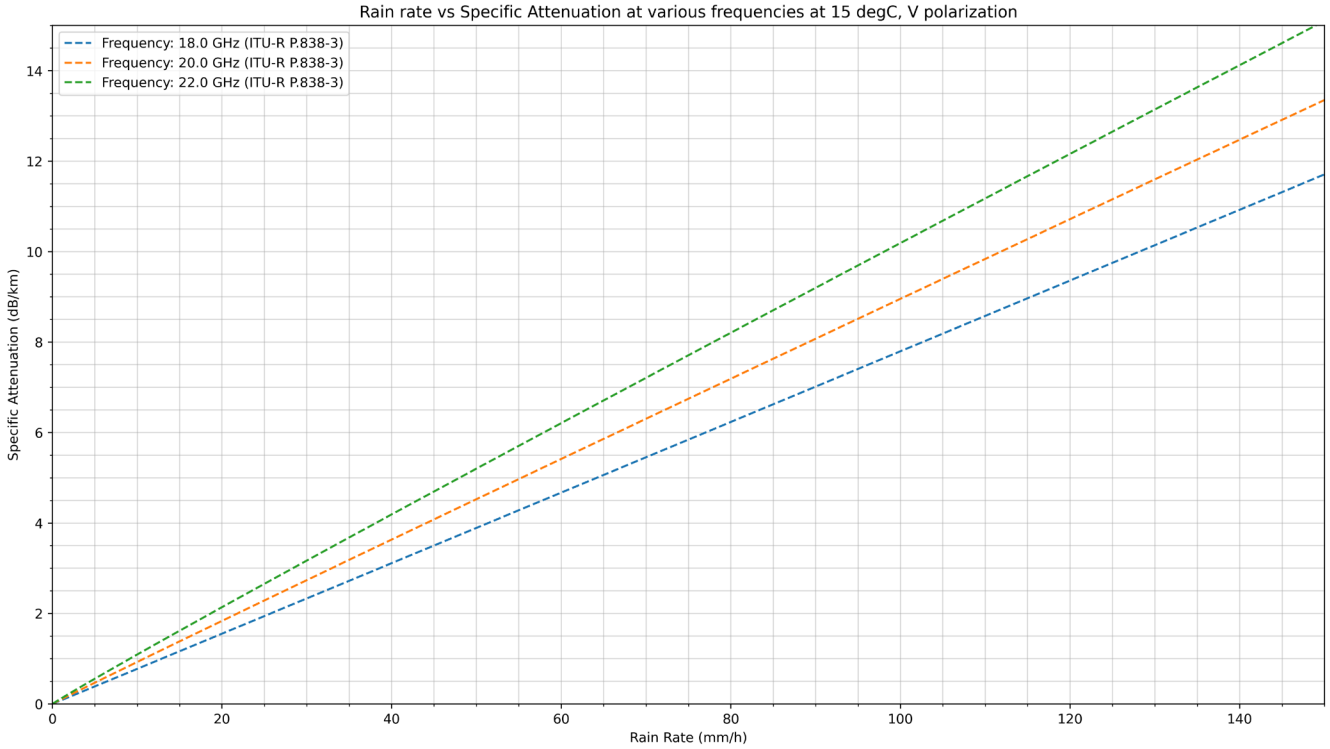
Dependence of attenuation on frequency, polarisation & temperature

15 °C
18-22 GHz



Horizontal Polarisation

Vertical Polarisation



Evaluation Metrics

A single metric cannot reflect the performance of stations holistically…

$$\mathrm{WMAPE} = \frac{\sum_{t=1}^{n} |A_t - F_t|}{\sum_{t=1}^{n} |A_t|}$$

Lower is better

Measures “closeness”.
More weights to larger values.
Make sense when rainfall only recorded intermittently.

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[\, n\Sigma x^2 - (\Sigma x)^2 \,][\, n\Sigma y^2 - (\Sigma y)^2 \,]}}$$

Strength and direction relationship
How good are the links at following an onset?
Dry wet detection

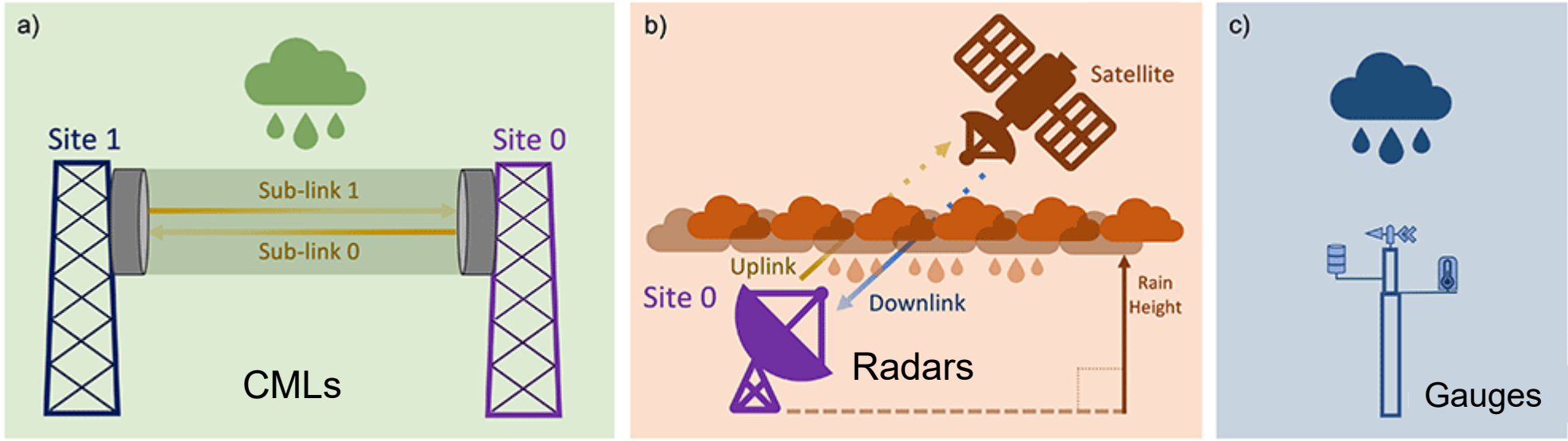
$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y}_i)^2}$$

Explains extent of one variance concerning the other variable

$$\%Bias = \frac{Mean - True\ value}{True\ value}$$

Average direction of forecast errors
Systematic error

Why Microwave links?



(Fencl M. et al., 2024)

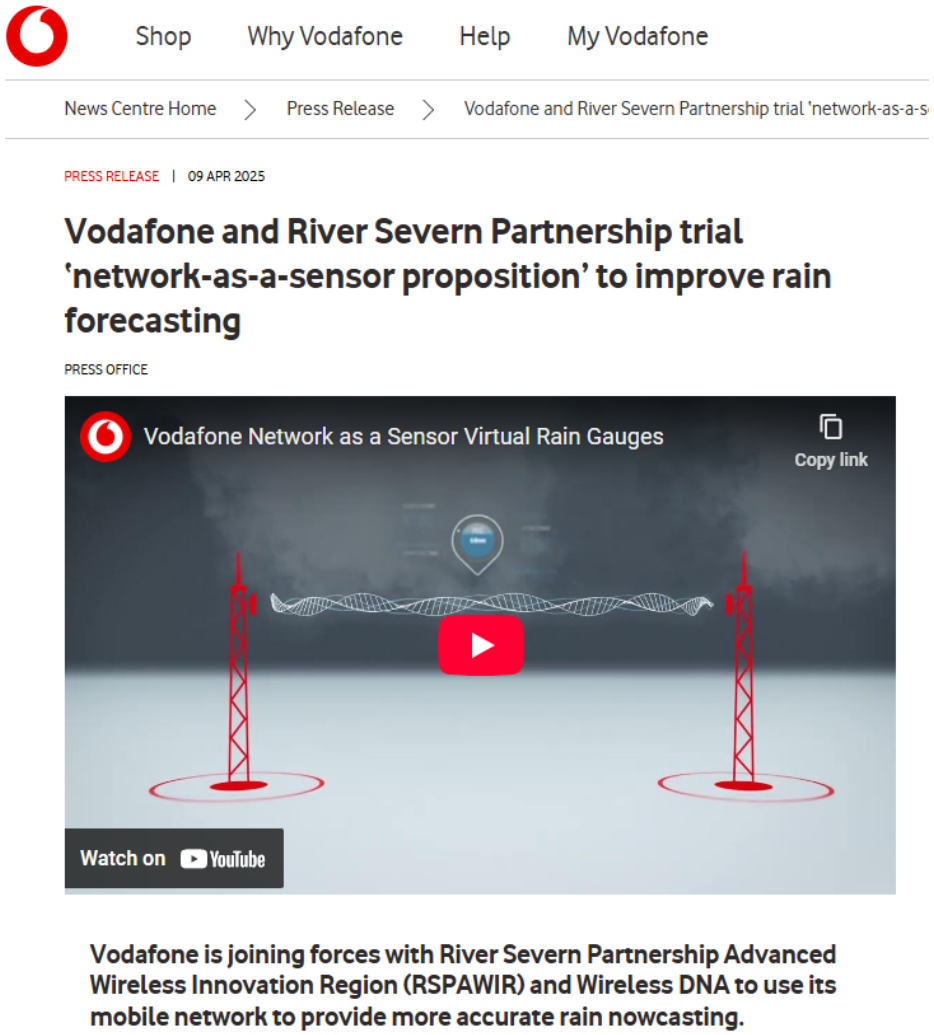
Pros

- No upfront costs
- Higher spatial density
- Close to surface
- Spatial-temporal scale relevant for hydrology & meteorology

Cons

- Calibration are location, precipitation type dependent
- Hard to obtain signal level data from telecommunication companies

HydroJULES project Overview: Understanding Precipitation Estimates from CMLs



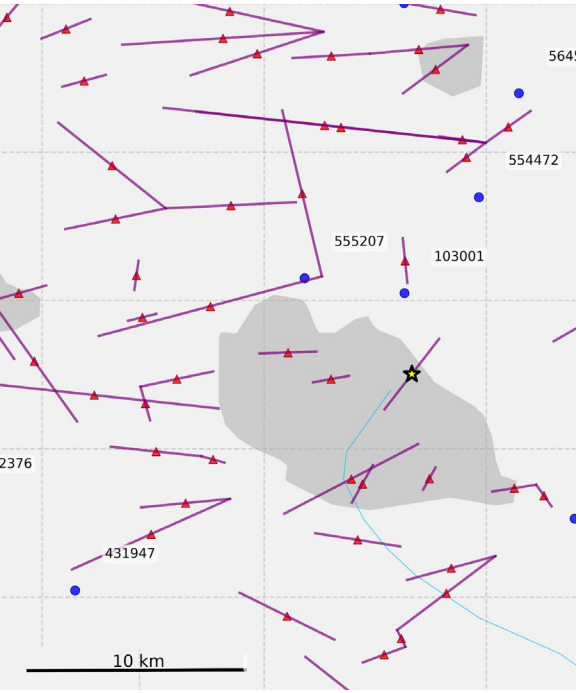
Study Period:
Vodafone has provided precipitation estimates for a 6-month period for Wales and South-West England.

Start date & time: 2024-07-14 23:00
End date & time: 2025-01-13 23:30

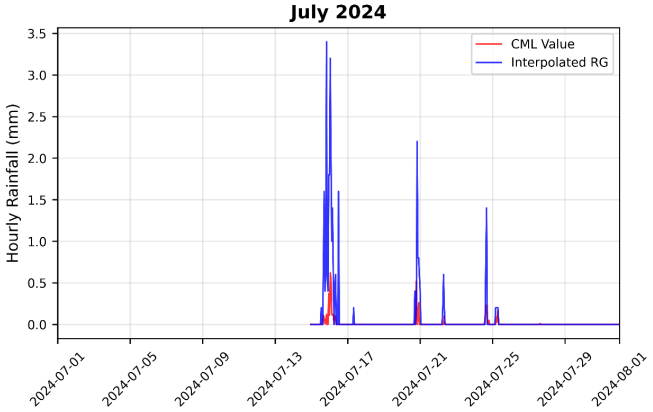
Time steps of 15 minutes

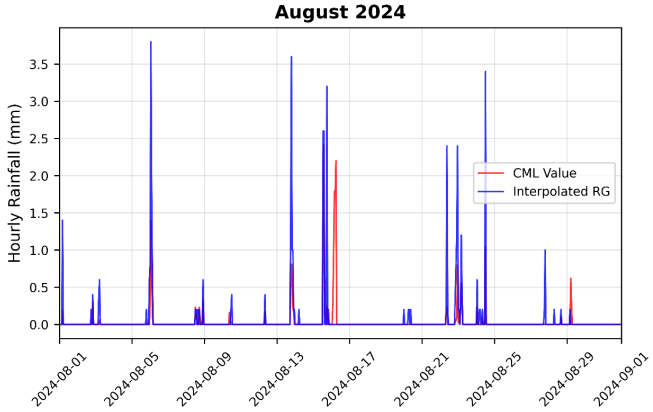
Metadata and 1-2 days of Raw data, available

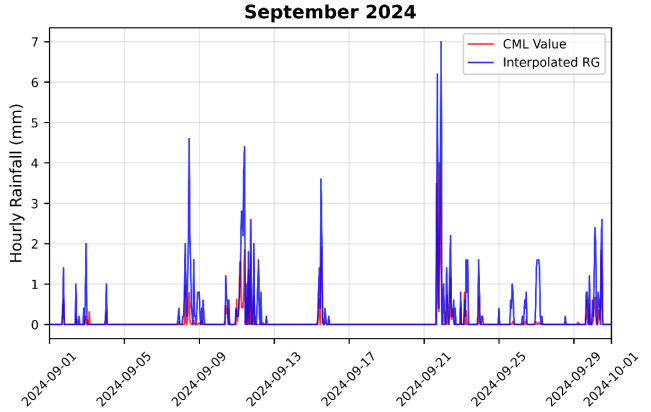
River Trent – Poor

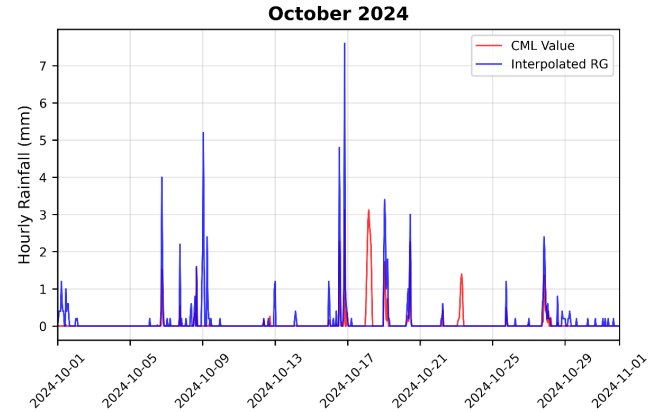


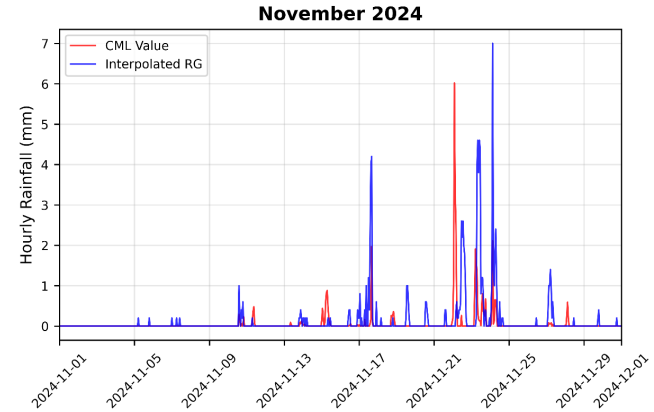
Hourly Rainfall Time Series - Station GBSTE_GBSTS_003476

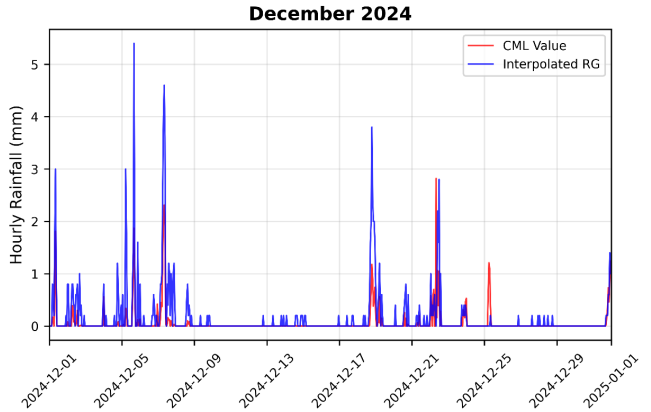


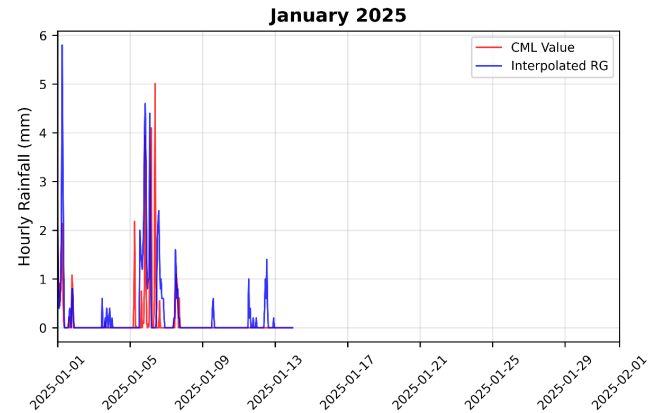


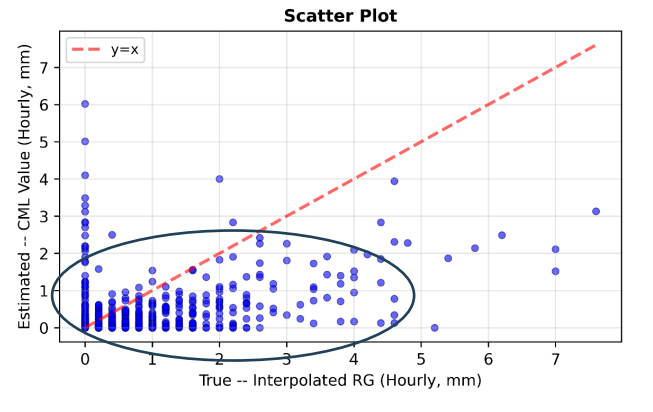










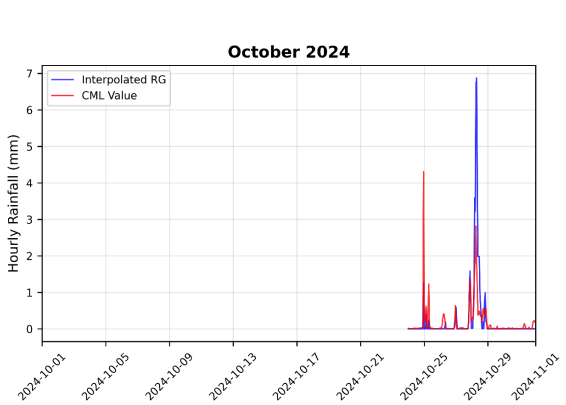


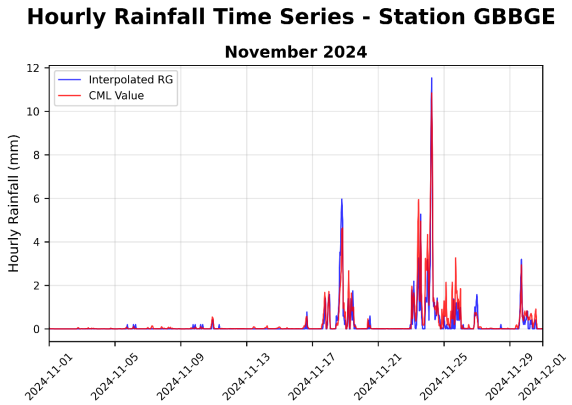
Station Metadata & Performance

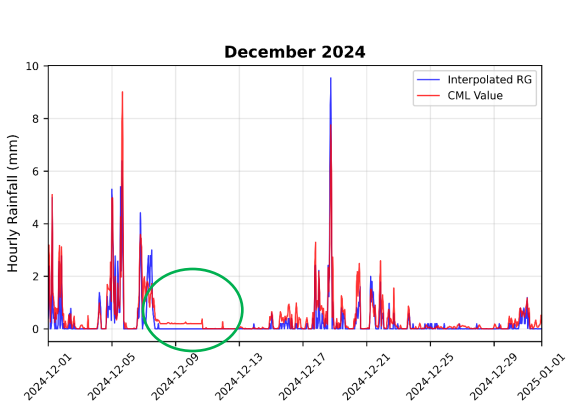
Metadata	Value	Statistics	Value
Middle Longitude	-2.1502	Min 15min Rain	0.01 mm
Middle Latitude	53.0504	Total Pilot Rain	257.13 mm
Start Latitude	53.0741	Completeness	99.3%
Start Longitude	-2.1318	WMAPE	0.88
End Latitude	53.0266	Correlation	0.53
End Longitude	-2.1686	R^2	0.26
Path Length	5.83 km	Bias (%)	-58.46
Max 15min Rain	2.71 mm		

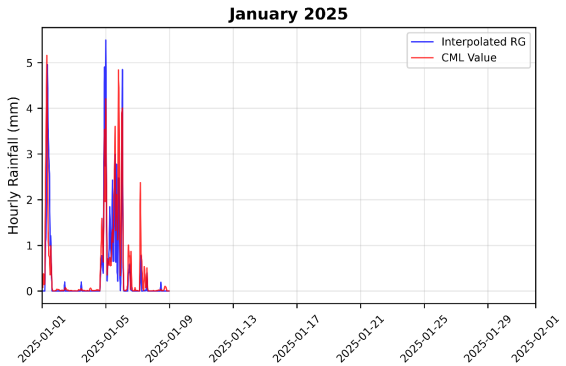
Re-evaluation data from Vodafone (11/08/25 in-house model)

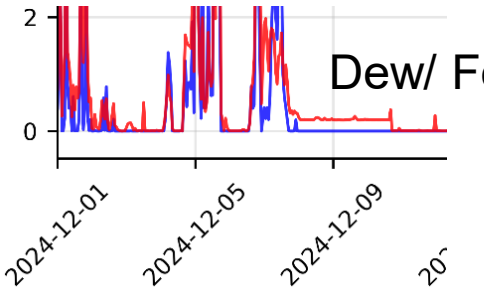
- Better estimate of light rainfall and dry-wet classification
- Note attenuation due to non-rainfall related hydrometers



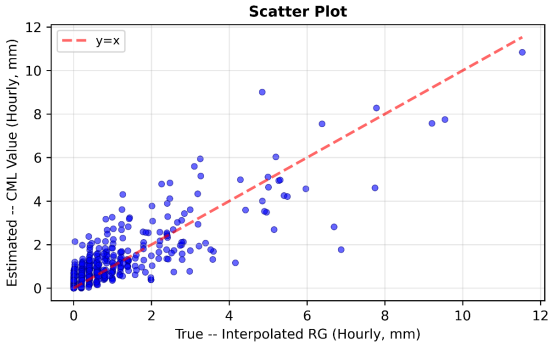








Dew/ Fog?



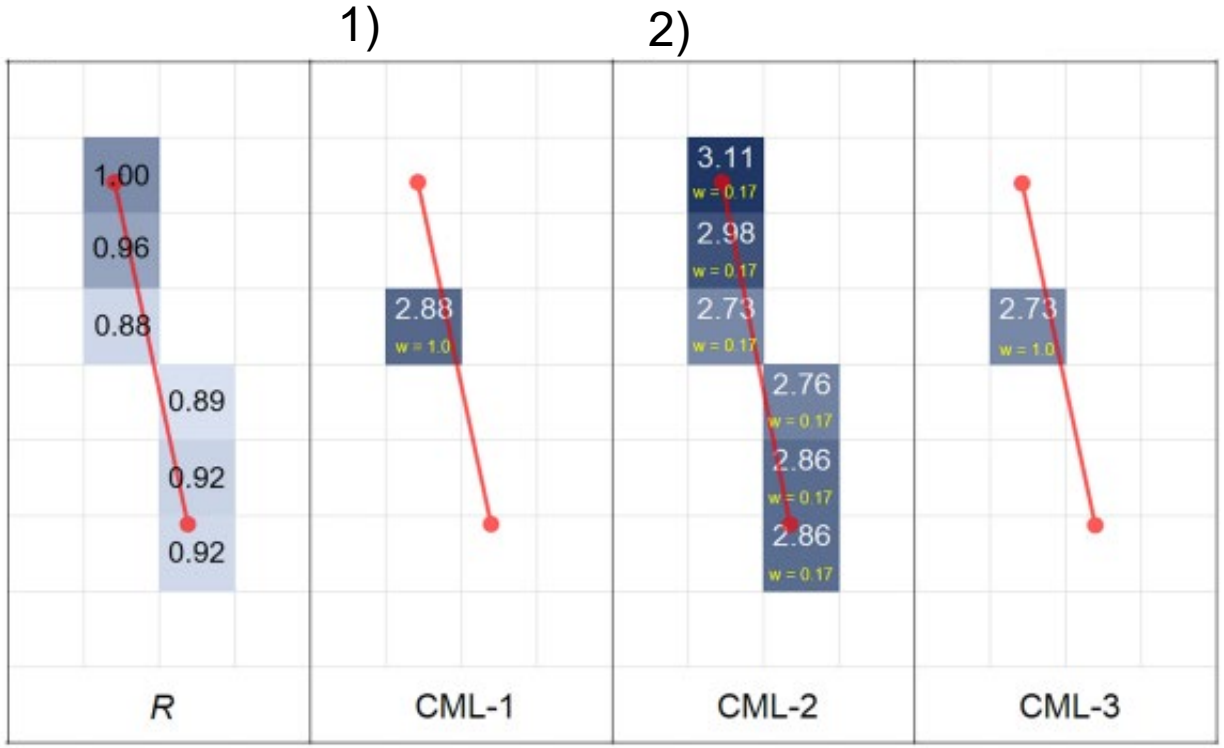
Station Metadata & Performance

Metadata	Value	Statistics	Value
WMAPE	0.70	R^2	0.74
Correlation	**0.88**	Bias (%)	31.63

Evaluation of Radar performance with RG

Central question: How does the **path-integrated nature** of the precipitation estimate (between base stations separated by several km) **affect the accuracy**?

Meteorological Applications
Science and Technology for Weather and Climate
RMetS
7 of 19



Ways of deriving of Radar rainfall value:

1) Pixel value at midpoint &

2) CML Path average pixel values